

Jochen Deuse · Ralf Klinkenberg ·
Nikolai West *Hrsg.*

Industrielle Datenanalyse

Entwicklung einer Datenanalyse-
Plattform für die wertschaffende,
kompetenzorientierte Kollaboration in
dynamischen Wertschöpfungsnetzwerken

OPEN ACCESS

 Springer Vieweg

Industrielle Datenanalyse

Jochen Deuse · Ralf Klinkenberg · Nikolai West
(Hrsg.)

Industrielle Datenanalyse

Entwicklung einer Datenanalyse-
Plattform für die wertschaffende,
kompetenzorientierte Kollaboration in
dynamischen Wertschöpfungsnetzwerken

Hrsg.

Jochen Deuse
Institut für Produktionssysteme
Technische Universität Dortmund
Dortmund, Deutschland

Ralf Klinkenberg
RapidMiner GmbH
Dortmund, Nordrhein-Westfalen, Deutschland

Nikolai West
Institut für Produktionssysteme
Technische Universität Dortmund
Dortmund, Deutschland



ISBN 978-3-658-42778-8 ISBN 978-3-658-42779-5 (eBook)
<https://doi.org/10.1007/978-3-658-42779-5>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Technische Universität Dortmund 2024. Dieses Buch ist eine Open-Access-Publikation.

Open Access Dieses Buch wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Buch enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Petra Steinmueller

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Das Papier dieses Produkts ist recyclebar.

Vorwort der Herausgeber

Die integrierte und vernetzte Anwendung von industriellen Datenanalysen in dynamischen Wertschöpfungsnetzwerken ist von entscheidender Bedeutung für die Wettbewerbsfähigkeit produzierender Unternehmen. Die rasant fortschreitende Entwicklung der **Digitalisierung**, neue Technologien durch **Industrie 4.0** und die nie dagewesenen Potenziale von **Künstlicher Intelligenz** verändern das produzierende Gewerbe nachhaltig. Um langfristig bestehen zu können, sind ein tiefgreifendes Verständnis für die neuen Möglichkeiten und der Mut zur Anwendung von industriellen Datenanalysen erforderlich. Nur so können die neuen Möglichkeiten zur Wertschöpfung und für innovative Geschäftsmodelle erschlossen werden. Das vorliegende Werk widmet sich diesem hochaktuellen und zukunftsweisenden Thema, das schon längst im Fokus der Forschung, aber auch der Praxis, steht.

Dieser Sammelband verfolgt das Ziel, sowohl theoretische als auch praxisnahe Perspektiven auf die integrierte und vernetzte Anwendung von industriellen Datenanalysen in dynamischen Wertschöpfungsnetzwerken zu beleuchten. Die Leser erhalten so wertvolle Einblicke in aktuelle Forschungsansätze, innovative Anwendungen und gelungene Erfolgsgeschichten. Diese Kombination aus **Theorie** und **Praxis** soll Anwendern und Forschern als Vorlage für weitere Umsetzungen dienen, um die Potenziale von industriellen Datenanalysen im produzierenden Gewerbe zu erschließen und zukünftige Entwicklungen erfolgreich zu gestalten.

Der Sammelband umfasst insgesamt zwanzig Beiträge von über fünfzig Autorinnen und Autoren aus Wissenschaft und Praxis. Die Beiträge decken ein breites Spektrum von Themenfeldern ab und behandeln beispielsweise Methoden der industriellen Datenanalyse, die integrierte Vernetzung von Produktionsprozessen, innovative Kollaborationsmöglichkeiten und Geschäftsmodelle sowie die Rolle von Kompetenzaufbau und -sicherung für Datenanalysen in der Industrie. Jedes Kapitel bietet eine einzigartige und wertvolle Perspektive zum Gesamtverständnis des Themenkomplexes.

Die Arbeiten in diesem Sammelband sind das Ergebnis eines gemeinsamen Forschungsprojekts, an dem insgesamt vierzehn Konsortialpartner über einen Zeitraum von vier Jahren beteiligt waren. Das Projekt trägt den Titel *„Vernetzte und integrierte Anwendung industrieller Datenanalyse für die wertschaffende, kompetenzorientierte*

Kollaboration in dynamischen Wertschöpfungsnetzwerken“. Das Forschungsprojekt wurde unter der Konsortialführung der RapidMiner GmbH und der Projektkoordination des Instituts für Produktionssysteme der Technischen Universität Dortmund umgesetzt. Das Vorhaben trägt das Förderkennzeichen 02P17D210 und wurde im Rahmen des Förderprogramms „*Industrie 4.0 – Kollaborationen in dynamischen Wertschöpfungsnetzwerken (InKoWe)*“ vom Bundesministerium für Bildung und Forschung (BMBF) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Es wurde im Programm „*Innovationen für die Produktion, Dienstleistung und Arbeit von morgen*“ kofinanziert. Die Zusammenarbeit ermöglichte einen interdisziplinären und praxisnahen Ansatz, der das Verständnis der integrierten und vernetzten Anwendung von industriellen Datenanalysen in dynamischen Wertschöpfungsnetzwerken entscheidend prägt.

Die Umsetzung des Forschungsprojekts und die Erstellung des Sammelbands fanden unter den besonderen Herausforderungen der COVID-19-Pandemie statt. Die Projektbearbeitung erfolgte größtenteils ohne persönliche Treffen vor Ort, was die Zusammenarbeit erschwerte. Durch kontinuierliche und produktive Online-Treffen gelang es den Konsortialpartnerinnen und -partnern, trotz dieser Schwierigkeiten, vielversprechende Ergebnisse zu erzielen. Dieser Sammelband bündelt die wissenschaftlichen Beiträge zu besagten Ergebnissen und war nur durch ein hohes Maß an Kooperation und Kollaboration im Konsortium umsetzbar.

Der Sammelband trägt dazu bei, die Bedeutung von Datenanalysen, Industrie 4.0 und Digitalisierung für produzierende Unternehmen im Hinblick auf gegenwärtige und zukünftige Entwicklungen zu unterstreichen. Er zeigt auf, welche Chancen und Herausforderungen sich aus diesen Entwicklungen ergeben und wie Unternehmen ihre Wettbewerbsfähigkeit durch die gezielte Anwendung von Datenanalysen und vernetzten Wertschöpfungsnetzwerken stärken können. Ein besonderer Fokus der Arbeiten liegt auf den entwickelten Bausteinen des Referenzbaukastens, durch den Lösungskomponenten zielgerichtet ausgewählt und bedarfsgerecht genutzt werden können. Die Beiträge in diesem Sammelband liefern so zum einen Impulse für weitere Umsetzungen und zeigen zum anderen konkret auf, wie Unternehmen sich den Herausforderungen der Zukunft stellen und ihr volles Potenzial ausschöpfen können.

An dieser Stelle möchten wir uns ausdrücklich bei allen beteiligten Konsortialpartnerinnen und -partnern sowie bei allen weiteren unterstützenden Personen und Institutionen für ihr Engagement und ihre Beiträge bedanken. Ohne die vielfältige Expertise, den außergewöhnlichen Einsatz und die stets gute Zusammenarbeit wäre dieser Sammelband nicht in dieser Form zustande gekommen. Unser besonderer Dank gilt außerdem allen Gutachterinnen und Gutachtern aus dem Konsortium, die sich an dem zweistufigen Peer-Review-Prozess beteiligt haben und so dazu beigetragen haben, die Qualität der Beiträge zu sichern und den hohen wissenschaftlichen Anspruch des Sammelbands zu gewährleisten. Abschließend möchten wir uns nochmal herzlich beim BMBF und dem PTKA bedanken, die maßgeblich für die Förderung und für die organisatorische

Betreuung dieses Forschungsvorhabens waren und ohne die diese Ergebnisse nicht möglich gewesen wären.

Schlussendlich möchten wir uns nun an alle Leser dieses Werkes richten: Um den größtmöglichen Nutzen aus diesem Sammelband zu ziehen, empfehlen wir Ihnen, sich zunächst einen Überblick über die verschiedenen Beiträge und ihre Themen zu verschaffen. Es ist nicht zwingend notwendig, die Beiträge in der vorgegebenen Reihenfolge zu lesen. Vielmehr können sie je nach Interesse und Bedarf ausgewählt werden. Eine Orientierung bieten die vier Gruppen, in die die Beiträge dieses Sammelbands unterteilt sind.

- Beiträge in Teil I behandeln primär die projektspezifischen Aspekte und behandeln u. a. den entwickelten Referenzbaukasten.
- Teil II umfasst die wissenschaftlichen Grundlagen und konzeptionellen Arbeiten der als Leistungsbereiche bezeichneten Forschungsrichtungen.
- Teil III präsentiert insgesamt acht Beispiele für erfolgreiche Anwendungen industrieller Datenanalysen, die daher auch als Erfolgsgeschichten bezeichnet werden.
- Teil IV umfasst eine Reihe von weiteren exemplarischen Anwendungen und stellt verschiedene Technologie-Demonstratoren des Vorhabens vor.

Der Sammelband eignet sich sowohl für Fachleute aus der Industrie als auch für Wissenschaftler und Studierende, die sich mit dem Thema der integrierten und vernetzten Anwendung von industriellen Datenanalysen in dynamischen Wertschöpfungsnetzwerken befassen. Er trägt zum weiteren Diskurs und zur Forschung in diesem Bereich bei und soll dabei helfen, die Potenziale von Datenanalysen und vernetzten Wertschöpfungsnetzwerken für produzierende Unternehmen zu erschließen. Wir wünschen den Lesern viele spannende Erkenntnisse und hoffen, dass dieser Sammelband dazu beitragen kann, die **Zukunft der industriellen Datenanalysen** erfolgreich zu gestalten.

Dortmund
im April 2023

Jochen Deuse
Ralf Klinkenberg
Nikolai West

Inhaltsverzeichnis

Teil I Einleitung

1 Einleitung	3
Nikolai West, Ralf Klinkenberg und Jochen Deuse	
2 Industrielle Datenanalysen im Mittelstand	11
Nikolai West, Thomas Huber und Jochen Deuse	
3 Ausgangssituation und technische Grundlagen	21
Nikolai West, Marius Syberg, Jörn Schwenken, Thomas Eickhoff, Jens C. Göbel, Michael Frey, Andreas Emrich, Christopher A. Klupak, Felix Walker und Jochen Deuse	
4 Prozesskette der industriellen Datenanalyse	41
Nikolai West, Marius Syberg, Jörn Schwenken, Edin Klapic, Ralf Klinkenberg und Jochen Deuse	

Teil II Leistungsbereiche im Referenzbaukasten

5 Konzeption und Entwicklung eines integrierten Datenbackends für die industrielle Datenanalyse	51
Thomas Eickhoff, Jens C. Göbel, Christo Apostolov und Hardy Krappe	
6 Abstraktion und Übertragbarkeit von Datenanalysen im industriellen Kontext	65
Marius Syberg, Nikolai West, Edin Klapic, Jörn Schwenken, Ralf Klinkenberg und Jochen Deuse	
7 Kompetenzentwicklung und -sicherung in der industriellen Datenanalyse von Wertschöpfungsnetzwerken	77
Christopher Klupak, Felix Walker, Volker Zimmermann und Rebekka Adams	

8	Entwicklung neuer Kollaborationsmöglichkeiten und Geschäftsmodelle	87
	Michael Frey, Andreas Emrich, Rebekka Adams und Volker Zimmermann	
Teil III Erfolgsgeschichten aus der Anwendung		
9	Datengetriebene Arbeitsplangestaltung in der Automobilfertigung	103
	Christine Rese, Sven Krzoska, Edin Klapic, Nikolai West, Philipp Schlunder, Ralf Klinkenberg, Mathias Gebler und Jochen Deuse	
10	Kollaborative, modulare Datenanalyse als Werkzeug im Qualitätsmanagement	113
	Martina Ringeln, Sibylle Legner, Sebastian Pähler, Katharina John, Andreas Cibis, Ralf Gryga, Lennart Krüger, Corinna Osterbrink, Felix Reinhart, Edin Klapic, Ralf Klinkenberg, Marius Syberg, Hermann Ferstl, Tanja Sindram, Michael Doukas und Jochen Deuse	
11	Integrierte Datenanalyse zur Kollaboration in der Auftragsplanung	123
	Andreas Bohlmann, Claudia Scholten, Rachana Desai und Edin Klapic	
12	Potenzialanalyse industrieller Datenanalysen in der Produktion am Beispiel des Kunststoffschweißens	133
	Tobias Beiß, Edin Klapic, Tina Hardt, Jörn Schwenken und Jochen Deuse	
13	Kompetenzentwicklung und -sicherung mit einem digitalen Wissensdienst	145
	Volker Zimmermann, Rebekka Adams und Christopher Klupak	
14	Retrofitting und Datenakquisition mittels Industrie 4.0-Technologien	155
	Tina Hardt, Hans-Joachim Mieves, Christoph Marx und Christian Meisberger	
15	Integration der Kundenmeinung ins Geschäftsmodell	163
	Michael Frey	
16	Vorgehensweisen für menschenzentrierte Datenanalyseprojekte	173
	Adrienne Bühler, Andrea Rösch, Anette Schötz, Birko Ruzicka, Céline Nordmann, Daniaal Dar, Hermann Ferstl, Julian Joachimbauer, Michael Doukas, Tanja Sindram und Thomas Huber	
Teil IV Verwendung der AKKORD-Plattform		
17	Umsetzung des Backend im Autorenn-Demonstrator	187
	Thomas Eickhoff und Jens C. Göbel	
18	Umsetzung der Analysebausteine im Autorenn-Demonstrator	193
	Jörn Schwenken, Nikolai West, Marius Syberg und Jochen Deuse	

19 Realisierung eines I4.0-Lerndemonstrators im Rahmen des AKKORD-Projekts	201
Christopher Klupak und Tina Hardt	
20 Industrielle Datenanalyse im Einklang mit Mensch, Technik und Organisation	209
Jochen Deuse, Ralf Klinkenberg, Nikolai West und Marius Syberg	

Abbildungsverzeichnis

Abb. 1.1	Darstellung der beteiligten Konsortialpartner im Forschungsprojekt AKKORD unter Berücksichtigung der Unterscheidung von Anwendungs-, Entwicklungs- und Forschungspartnern.	6
Abb. 1.2	Darstellung der vier Leistungsbereiche aus dem Forschungsprojekt AKKORD im Spannungsfeld von Mensch, Technik und Organisation.	8
Abb. 2.1	Vorläufige Studienergebnisse zur Grundhaltung zu Datenwissenschaft und Künstlicher Intelligenz im produzierenden Gewerbe	15
Abb. 2.2	Anwenderbedarfe zur Nutzung von industriellen Datenanalysen	16
Abb. 3.1	Darstellung des Open System Interconnection Modells als Beispiel für ein Referenzmodell der Software-Entwicklung (i. A. a. ISO/IEC 7498-1:1994)	23
Abb. 3.2	Darstellung des Referenzarchitekturmodell Industrie 4.0 als Beispiel für eine Referenzarchitektur in der Software-Entwicklung (i. A. a. DIN SPEC 91345:2016-04).	24
Abb. 3.3	Darstellung des Cross-Industry Standard Process for Data Mining (CRISP-DM) im Kontext von Daten, Informationen und Wissen (i. A. a. Chapman et al., 2000, S. 13; West et al., 2021a, S. 133).	30
Abb. 4.1	Darstellung der Prozesskette der industriellen Datenanalyse.	42
Abb. 4.2	Darstellung des entwickelten AKKORD-Referenzbaukastens, der die Prozesskette der industriellen Datenanalyse abbildet und aus sechs Bausteinen besteht.	44
Abb. 5.1	Vorgehen beim Aufbau des Datenbackends	53
Abb. 5.2	Architektur der SP ² IDER-Plattform.	55
Abb. 5.3	Benutzeroberfläche des Data Model Canvas	55
Abb. 5.4	Architektur und Module des konsolidierten Datenbackends	58
Abb. 5.5	Datenmodell des konsolidierten Datenbackends	59
Abb. 5.6	Architektur des Data Warehouses im Referenzbaukasten von AKKORD	61
Abb. 5.7	Instanziierung des Data Warehouse für AKKORD	61

Abb. 6.1	Anforderungen aus dem Bereich der industriellen Datenanalyse an eine Kollaborations-plattform (Syberg et al., 2023, S. 66)	67
Abb. 6.2	Schematische Darstellung der Elemente eines Datenanalysemoduls . . .	70
Abb. 6.3	Exemplarische Pipeline für einen Analyseprozesses in RapidMiner Studio am Beispiel einer univariaten Zeitreihenprognose mit mehreren Vorverarbeitungsprozessen	71
Abb. 7.1	Grundkonzept der Work&Learn-Plattform	81
Abb. 7.2	Kompetenzanforderungen und Identifikation von Kompetenzrollen . . .	82
Abb. 7.3	Inhalte aus dem digitalen Wissensdienst	83
Abb. 7.4	Technisch-funktionale Validierungsmaßnahme mittels digitalem Fragebogen	84
Abb. 8.1	Schematische Darstellung der Bewertung	94
Abb. 8.2	Allgemeine Ontologie zum XES Mapping auf Geschäftsmodelle	95
Abb. 8.3	Architektur Geschäftsmodellentwicklung	96
Abb. 8.4	Abbildung von Inhalten im Text auf Geschäftsmodellaspekte	97
Abb. 8.5	Geschäftsmodellanalyse-Baustein in RapidMiner Studio	98
Abb. 9.1	Beispielhafte Struktur eines Arbeitsplans	106
Abb. 9.2	Vorgehensmodell zur datenbasierten Optimierung und Gestaltung von Montagelinien. (Nach Rese et al., 2023, S. 11 f.)	107
Abb. 9.3	Prozessbeschreibungen vor und nach der Komplexitätsreduktion	108
Abb. 9.4	Synthetisch generierte Prozessbeschreibung in kontrollierter Sprache	108
Abb. 9.5	Power-BI-Bericht zur datenbasierten Entscheidungsunterstützung für die Gestaltung und -optimierung von Montagelinien	110
Abb. 10.1	Allgemeiner Problemlöseprozess, Analysefragestellungen und beispielhafte Darstellung von Ergebnissen im Anwendungsfall	119
Abb. 10.2	Dashboard zu Übersicht bzw. Alarmsystem der Ersatzteilverkäufe für das Qualitätsmanagement (Ausschnitt)	119
Abb. 11.1	Darstellung eines Zeitreihenfensters mit Vorhersagehorizont	127
Abb. 11.2	Ergebnisdarstellung im Vergleich der der Prognosemethoden	128
Abb. 11.3	Darstellung von Prognoseergebnis und Abweichung	128
Abb. 12.1	Webapplikation zum Prozessstart mit Gesamtübersicht	138
Abb. 12.2	Zeitabhängiger Prozessverlauf beim Kunststoffschweißen	139
Abb. 12.3	Aufgezeichnete Daten eines ersten Schweißprozesses mit fehlerhafter Funktion der Maschine	140
Abb. 12.4	Fügeweg zu gering eingestellt	140
Abb. 12.5	Mittelwertkurven der Servo-Kraft in der Fügephase mit Abweichungskorridor	141
Abb. 12.6	Analysemodul zur Selektion und Darstellung von Einzelprozessen . . .	142
Abb. 13.1	Einblick in den Lern-, den Magazin- und den Communitybereich der Work&Learn-Plattform	148
Abb. 14.1	Visualisierung erfasster Werte im Autorenn-Demonstrator	159

Abb. 14.2	Schematischer Aufbau der Lernstation	160
Abb. 15.1	NLP-Pipeline zur Sentiment-Analyse auf Kundenmeinungen	166
Abb. 15.2	Schrittweise Vorbereitung der Ausgangstexte durch Tokenisierung, Normalisierung und Filterung	167
Abb. 15.3	Abbildung von Textsegmenten auf GM-Aspekte	168
Abb. 15.4	Die Benutzeroberfläche des Demonstrators zeigt die Bewertung der einzelnen Geschäftsmodellaspekte	169
Abb. 15.5	Empfehlung von Geschäftsmodellmustern	169
Abb. 16.1	Rahmenbedingungen der Studie zum „Erfolgsfaktor Mensch in Data-Science-Projekten“	175
Abb. 16.2	Das Reifegradmodell mit den Erfolgsfaktoren 01–21 und mit beispielhaften Ausprägungen in Grau	175
Abb. 16.3	Die vier Phasen der KMU-Interviewstudie	176
Abb. 16.4	Verwendung von DS/KI in KMUs	177
Abb. 16.5	Das Industrial Data Science-Vorgehensmodell	179
Abb. 16.6	Integriertes Vorgehensmodell Data Science – UX	181
Abb. 16.7	Der zu erzielende Nutzen durch den Einsatz menschenzentrierter Vorgehensweisen	181
Abb. 16.8	Change Cycles für die Change-Management Vorgehensweise in Data-Science-Projekten	182
Abb. 16.9	Potenzielle Change Aktivitäten entlang des DS-Phasenmodells	182
Abb. 17.1	Bausteine des Demonstrators	188
Abb. 17.2	Darstellung einer Zeitreihe in Contact Elements for IoT	190
Abb. 17.3	Anbindung von Datenquellen im Autorenn-Demonstrator	191
Abb. 18.1	Systemarchitektur AKKORD-Autorenn-Demonstrator	194
Abb. 18.2	Architektur des Datenzugriffs des AKKORD-Autorenn- Demonstrators	196
Abb. 18.3	Prozess des Analysemoduls Berechnung von Rennsieger:innen	197
Abb. 18.4	AI-Toolbox Nutzungsmodul Auswertung Rennsieger:in	198
Abb. 19.1	Darstellung der AKKORD-Lernstation	202
Abb. 19.2	Auszug aus der Zielgruppenanalyse des Ausbildungsberufs Anlagenmechaniker:in	204
Abb. 19.3	Exemplarische Darstellung einer Lernaufgabe zur Auswahl eines geeigneten Geräts	205

Teil I
Einleitung



Einleitung

1

Einführung ins Forschungsvorhaben AKKORD zur industriellen Datenanalyse in dynamischen Wertschöpfungsnetzwerken

Nikolai West , Ralf Klinkenberg  und Jochen Deuse 

1.1 Ausgangssituation

Die Anwendung industrieller Datenanalysen eröffnet produzierenden Unternehmen innovative Möglichkeiten zur nachhaltigen Optimierung von **Produkten** und **Prozessen**. Gleichzeitig wird die Initiierung neuer **Geschäftsmodelle** und **Kollaborationen** in Wertschöpfungsnetzwerken ermöglicht. Aus dieser Ausgangsage entstammt das BMBF-Forschungsvorhaben. Das auch mit der Kurzbezeichnung „AKKORD“ benannte Forschungsvorhaben, dessen Ergebnisse zentraler Bestandteil dieses Werks sind, trägt den folgenden Titel:

Vernetzte und integrierte Anwendung industrieller Datenanalyse für die wertschaffende, kompetenzorientierte Kollaboration in dynamischen Wertschöpfungsnetzwerken

In ingenieurwissenschaftlich geprägten Unternehmen werden **betriebliche Entscheidungen** seit jeher anhand von **faktenbasierten Auswertungen** getroffen (Deuse et al., 2022, S. 2 f.). Die Anwendung von industriellen Datenanalysen stellt somit eine logische Weiterentwicklung dar, die sich die digitalen Möglichkeiten für Auswertung immer größer werdender Datenmengen zu Nutze macht (West et al., 2021, S. 131). Zu Beginn des Vorhabens im Frühjahr 2019 zeichnete es sich ab, dass insbesondere

N. West (✉) · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: nikolai.west@ips.tu-dortmund.de

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

R. Klinkenberg
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: rklinkenberg@rapidminer.com, rklinkenberg@altair.com

Kleine und Mittlere Unternehmen (KMU) einen sinnvollen und zielgerichteten Einsatz moderner Analysetechnologien aus eigener Kraft nicht leisten können (Mazarov et al., 2019, S. 874 f.). Es mangelte zum einen an den erforderlichen Kompetenzen und Strategien innerhalb der Unternehmen und zum anderen an einem auf KMU ausgerichteten Dienstleistungs- und Technikangebot zur nachhaltigen Nutzung der umfangreichen Potenziale. In Kap. 2 werden die Hintergründe der Forschungsförderung weiterführend beschrieben. Neben einem Überblick der Projekthistorie wird dort insbesondere das Potenzial der Forschungsergebnisse für KMU beschrieben.

Zusammenfassend lässt sich sagen, dass die bestehenden Hindernisse bei der technischen, aber vor allem auch bei der organisatorischen Integration von Anwendern, Systemanbietern und Dienstleistern die Realisierung initialer Leuchtturmprojekte sowie die langfristige Durchdringung weiterer Industriebereiche mit Datenanalysetechnologien erschweren. Es fehlten sowohl ein integriertes Konzept zur Überwindung der existierenden Barrieren und Hemmnisse als auch konkrete Lösungen (Nolte et al., 2020, S. 734).

1.2 Verbundprojektziel

Zur erfolgreichen Anwendung von industriellen Datenanalysen ist eine ganzheitliche Betrachtung betrieblicher Abläufe erforderlich. Da die betrachteten Betriebsabläufe häufig komplex sind, involvieren industrielle Datenanalysen viele unterschiedliche Prozesse, Systeme, Datenquellen, Akteure und Zusammenhänge. Erst durch den Dreiklang von **Mensch**, **Technik** und **Organisation** (MTO) kann die dafür notwendige Harmonie entstehen. Kap. 20 nimmt einen weiterführenden Ausblick auf die künftige Entwicklung von MTO vor. Als Kurzbezeichnung für das Forschungsvorhaben wurde deshalb der bereits erwähnte Begriff „AKKORD“ gewählt, der in der Musik den Zusammenklang vieler Stimmen mit verschiedenen Funktionen bezeichnet. Auch zur industriellen Umsetzung von Datenanalysen ist ein derartiges Zusammenspiel erforderlich. An die Stelle von Melodie, Bass oder Begleitstimme treten in diesem Falle sowohl **menschliche Akteurinnen und Akteure** aus unterschiedlichen Fachbereichen als auch die **technischen und organisatorischen Einflussfaktoren**, unter denen die Datenanalysen stattfinden. Damit produzierende Unternehmen, und insbesondere KMU, diesen Akkord erklingen lassen können benötigen sie entsprechende Instrumente.

Projektsteckbrief AKKORD

Laufzeit: 48 Monate

Zeitraum: 04.2019-03.2023

Volumen: 7,7 Mio. €

Konsortium: 14 Partner

Webseite: www.akkord-projekt.de

Der im Rahmen des Verbundprojekts entwickelte **Referenzbaukasten** zur Anwendung industrieller Datenanalysen stellt in dieser Analogie einen **Werkzeugkoffer** zur erfolgreichen Umsetzung von integrierten und vernetzten Anwendungen von industriellen Datenanalysen dar. Typischerweise werden Datenanalyseprojekte von interdisziplinären Teams umgesetzt, die bei der Bearbeitung unterschiedliche Rollen einnehmen und entsprechende Werkzeuge in der Bearbeitung benötigen (Schwenken et al., 2023, S. 83). Der Baukasten beinhaltet ein umfangreiches Angebot von Werkzeugen, die bedarfsgerecht und modular zusammengestellt und genutzt werden können. Mithilfe dieser Werkzeuge wird eine kompetenzorientierte Kollaboration in dynamischen Wertschöpfungsnetzwerken anwenderfreundlich ermöglicht (Syberg et al., 2023, S. 65, Kap. 3). In Anlehnung an den Gedanken eines Baukastens werden einzelne Module im Weiteren als **Bausteine** bezeichnet. Um eine einsteigerfreundliche Verwendung des Referenzbaukastens und seiner Bausteine zu ermöglichen, wird der ganzheitliche Ansatz in einer kollaborativen Online-Service-Plattform vereint. Eine ausführliche Beschreibung des Referenzbaukastens und seiner Bausteine findet sich weiterführend in Kap. 4.

1.3 Forschungspartner

Im Forschungsprojekt AKKORD arbeiteten von 2019 bis 2023 vierzehn Konsortialpartner aus unterschiedlichen Branchen zusammen. Die Partner können anhand ihrer Ausrichtung in drei Gruppen unterteilt werden. Zu den **Anwendungspartnern** zählen vier produzierende Unternehmen, die die Nutzung von industriellen Datenanalysen anstreben. Die **Entwicklungspartner** stellen sechs Unternehmen dar, die entweder Software- oder Hardware-Lösungen anbieten und so die Grundlage zur Umsetzung sowie Nutzung von Datenanalysen liefern. Abgerundet wird das Konsortium durch vier **Forschungspartner**, die im Projekt eine Brücke zwischen den technischen Lösungen und der praxisorientierten Nutzung schlagen. Abb. 1.1 zeigt eine Übersicht der vierzehn Partner innerhalb dieser drei Gruppen.

Ein detaillierter Einblick in die Arbeiten der beteiligten Forschungspartner findet sich auf der Projekt-Webseite (www.akkord-projekt.de/partner). Neben einer ausführlichen Partnervorstellung wird dort die individuelle Rolle im Projekt erläutert und die individuell erzielten Ergebnisse vorgestellt. In diesem Kapitel werden die in Abb. 1.1 gezeigten Forschungspartner aus dem Vorhaben AKKORD stichpunktartig vorgestellt.

- **Anwendungspartner:**

- ERCO GmbH (*Elektroindustrie*): Architekturbeleuchtung mit besonderem Fokus auf präziser und effektiver LED-Technologie
- Miele & Cie. KG (*Weißes Ware*): Premium Haushalts- und Gewerbegeräte mit Hauptbeteiligung durch das zentrale Qualitätsmanagement und Serienpflege der Haushaltswascher



Abb. 1.1 Darstellung der beteiligten Konsortialpartner im Forschungsprojekt AKKORD unter Berücksichtigung der Unterscheidung von Anwendungs-, Entwicklungs- und Forschungspartnern

- PolyMerge GmbH (*Sondermaschinenbau*): Entwicklung, Konstruktion und Fertigung von Maschinen für die Automobilzulieferer und Kunststoffverarbeitendes Gewerbe
- Volkswagen AG (*Automobilindustrie*): Fahrzeuge mit Hauptbeteiligung durch die zentrale Industrial Engineering Abteilung
- **Entwicklungspartner:**
 - Arendar IT-Security GmbH (*Industrie 4.0 Kommunikation*): Industrieller Datensammler mit Schwerpunkt auf Cyber-Security
 - CONTACT Software (*Softwareentwicklung*): Lösungen für Produktdatenmanagement, Product Lifecycle Management, Projektmanagement und Internet of Things
 - mosaiic GmbH (*Unternehmensberatung*): Beratung, Konzeption und Einführung von IT für produzierende Unternehmen
 - NEOCOSMO GmbH (*Softwareentwicklung*): Lösungen für Kommunikations- und Lernplattformen für Informations- und Wissensvermittlung
 - PDTec AG (*Softwareentwicklung*): Lösungen für Daten- und Informationsmanagement innerhalb des Produktlebenszyklus
 - RapidMiner GmbH (*Softwareentwicklung*): Lösungen für Predictive Analytics, Data Mining und Text Mining [Konsortialführer]
- **Forschungspartner:**
 - Lehrstuhl für Berufspädagogik für technische Fächer (*Universität Hamburg*): Technologiebasierte Erfassung und Entwicklung von Fachkompetenzen
 - Deutsches Forschungszentrum für Künstliche Intelligenz (*Institut für Wirtschaftsinformatik, Saarbrücken*): Angewandte Informatikforschung

- Institut für Produktionssysteme (*Technische Universität Dortmund*): Datengetriebene Problemlösung mittels industriellen Datenanalysen für diverse Fragestellungen aus dem produzierenden Gewerbe [Projektleitung]
- Lehrstuhl für Virtuelle Produktentwicklung (*Technische Universität Kaiserslautern*): Produkt- und Service Lifecycle Management sowie Datenintegration und -vernetzung

Während der Bearbeitung des Forschungsprojekts AKKORD veröffentlichten die aufgezählten Projektpartner regelmäßig Arbeitsberichte zu Fortschritten und Ergebnissen auf der Projekt-Webseite. Diese Beiträge finden sich ebenfalls weiterführend auf der Projekt-Webseite (www.akkord-projekt.de/news).

1.4 Vorgehensweise

Um die Umsetzung des Referenzbaukastens als übergeordnetes Gesamtziel zu erreichen, werden im Forschungsprojekt mehrere wissenschaftliche und technische Teilziele verfolgt. Anhand dieser Teilziele strukturieren sich die Aufgaben der vierzehn Projektpartner.

1. **Übergreifende und vernetzte Datenbasis:** Durch ein leistungsfähiges PLM-basiertes Daten-Backend-System werden Erfassung, Integration und Verknüpfung unterschiedlicher Datenquellen entlang des Produktlebenszyklus umgesetzt und die Grundlage für anschließende Datenanalysen geschaffen (siehe Kap. 5).
2. **Integrierte und vernetzte Datenanalysen:** Durch einen benutzerorientierten Konfigurationsassistenten mit einheitlichen Analysemodulen und vorkonfigurierbaren Dashboards werden automatische und individualisierbare Analyseprozesse umgesetzt (siehe Kap. 6).
3. **Kompetenzentwicklung und -sicherung:** Durch einen digitalen Wissensdienst werden situierte Lernmodule zielgerichtet und anwendergerecht bereitgestellt, um die Kompetenzprofile von Mitarbeiterinnen und Mitarbeitern bedarfsgerecht zu erweitern (siehe Kap. 7).
4. **Kollaborationsmöglichkeiten und Geschäftsmodelle:** Durch einen Empfehlungsassistenten werden Vorschläge und Handlungsempfehlungen für innovative und wirtschaftliche Geschäftsmodelle zur kollaborativen Wertschöpfung abgegeben, die auf den Ergebnissen der vernetzten und integrierten Datenanalysen beruhen (siehe Kap. 8).

In Abb. 1.2 werden diese vier Teilziele im Kontext von Mensch, Technik und Organisation gezeigt. Aufgrund der vielen Akteure, die zum Erfolg von industriellen Datenanalyseprojekten beitragen, erfordert ein ganzheitlicher Ansatz auf stets die Berücksichtigung aller Einflussfaktoren. Im Forschungsvorhaben werden diese Teilziele

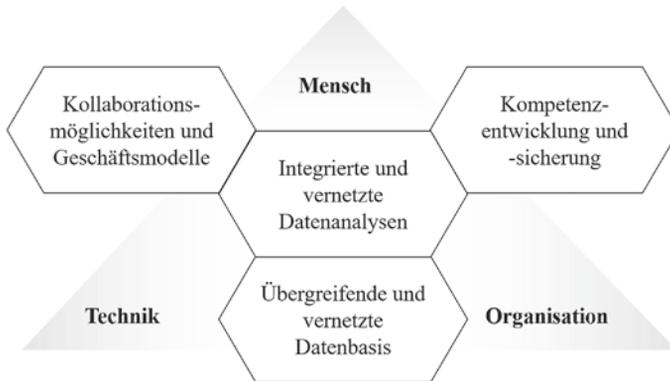


Abb. 1.2 Darstellung der vier Leistungsbereiche aus dem Forschungsprojekt AKKORD im Spannungsfeld von Mensch, Technik und Organisation

in den folgenden Kapiteln als Leistungsbereiche bezeichnet. Die wissenschaftlich-konzeptionellen Arbeiten orientierten sich an dieser Unterteilung und wurden in einem übergeordneten **Referenzbaukasten** zusammengeführt.

Neben der Umsetzung dieser Zielsetzungen in einem integrierten Referenzbaukasten erfolgt in AKKORD außerdem die exemplarische Erarbeitung von industriellen Anwendungsszenarien sowie der Aufbau von Demonstratoren. Durch die Umsetzung von Anwendungsszenarien in Form von Software- und Hardware-basierten Demonstrations- bzw. Pilotanwendungen werden sowohl der Anwendernutzen als auch die Akzeptanz für Datenanalysetechnologien gesteigert.

An der Struktur orientiert sich auch der weitere Aufbau des vorliegenden Sammelbands. In den verbleibenden Beiträgen von **Teil I** werden zunächst der unmittelbare Nutzen für mittelständige Anwender, die grundsätzliche Problemstellung zu Beginn des Vorhabens sowie das übergeordnete Konzept des Referenzbaukastens vorgestellt. Innerhalb von **Teil II** werden die als Leistungsbereiche bezeichneten Teillösungen des Referenzbaukastens beschrieben. Die Leistungsbereiche orientieren sich an den Teilzielen 1 bis 4 und wurden federführend von je einem der Forschungspartner und mit mehreren Anwendungs- und Entwicklungspartnern erarbeitet. Die angesprochenen Anwendungsszenarien werden als acht Erfolgsgeschichten innerhalb von **Teil III** vorgestellt. Der abschließende **Teil IV** fasst dann die Anwendung der erarbeiteten Service-Plattform zusammen und greift dafür u. a. die entwickelten Demonstratoren auf.

Literatur

Deuse, J., West, N., & Syberg, M. (2022). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering (IJPME)*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.

- Mazarov, J., Wolf, P., Schallow, J., Nöhring, F., Deuse, J., & Richter, R. (2019). Industrial Data Science in Wertschöpfungsnetzwerken. *Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF)*, 114(12), 874–877. <https://doi.org/10.3139/104.112205>.
- Nolte, V., Sindram, T., Mazarov, J., & Deuse, J. (2020). Industrial Data Science erfolgreich implementieren. Interviewstudie zu Erfolgsfaktoren und Hemmnissen. *Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF)*, 115(10), 734–737. <https://doi.org/10.1515/zwf-2020-1151020>.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Industrielle Datenanalysen im Mittelstand

2

Hintergrund der Forschungsförderung und Nutzen der Forschungsergebnisse für mittelständische Unternehmen

Nikolai West , Thomas Huber  und Jochen Deuse 

2.1 Rahmen der Forschungsförderung

Eine Vielzahl unterschiedlicher Aktivitäten zu Industrie 4.0 wurden seit dem Jahr 2011 in Deutschland durchgeführt. Die Forschungsarbeiten haben den Weg zu einer intelligenten **Vernetzung der Produktion** durch den Einsatz neuartiger cyberphysischer Produktionssysteme geebnet. Diese ermöglichen u. a. eine Vernetzung von Maschinen und Objekten sowie einen zentralen Zugriff auf Prozessdaten (BMW, 2015, S. 8 f.). Darüber hinaus wurden Lösungen für unterschiedlichste Anwendungsfälle entwickelt und auf den betrieblichen Hallenboden gebracht. Auf Bundesebene werden eine Vielzahl von Aktivitäten auf der *Plattform Industrie 4.0* zusammengeführt und koordiniert (BMW, 2020, S. 2). Im Jahr 2014 benannte die Bundesregierung dann mit der neuen **Hightech-Strategie** „Innovationen für Deutschland“ zentrale Aktionsfelder der Zukunftsaufgabe „Digitale Wirtschaft und Gesellschaft“. Dazu zählen beispielsweise die Anwendung von Industrie 4.0, Smart Services, Smart Data und die digitale Vernetzung (BMBF, 2021, S. 3). In diese Strategie fällt auch das Programm „Innovationen für die Produktion, Dienstleistung und Arbeit von morgen“ des *Bundesministeriums für Bildung und Forschung* (BMBF), das auf eine Stärkung der Technologiekompetenz, eine Vertiefung von wissenschaftlichem Austausch sowie auf die Unterstützung von Forschung

N. West (✉) · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: nikolai.west@ips.tu-dortmund.de

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

T. Huber
mosaic GmbH, München, Deutschland
E-Mail: thomas.huber@mosaic.de

und Lehre ausgerichtet ist. Die Forschungsergebnisse, aus denen dieses Buch entstanden ist, sind ein Teil der vom BMBF geförderten Maßnahmen für „Industrie 4.0 – Kollaborationen in dynamischen Wertschöpfungsnetzwerken“, kurz **InKoWe** (BMBF, 2017, S. 1 ff.).

Die Förderrichtlinie InKoWe zielte insbesondere auf die Entwicklung und Einführung innovativer Lösungen zu **Industrie 4.0** ab. Im Fokus der Arbeiten stand neben einer Stärkung der Zusammenarbeit von Unternehmen mit ihren Kunden und Lieferanten außerdem die Betrachtung des Unternehmens als sozio-technisches System. Dafür sollten unternehmensinterne sowie -übergreifende Prozesse entwickelt werden, um eine Steigerung der Flexibilität in der Verbundproduktion zu erreichen. Darüber hinaus sollte der Einsatz geeigneter Methoden zum Schutz von Daten das unternehmensspezifische Fachwissen langfristig sichern. Da insbesondere Kleine und Mittlere Unternehmen (KMU) häufig in Wertschöpfungsnetzwerken angesiedelt sind, lag ein besonderer Fokus der Richtlinie auf der Nutzbarkeit der Ergebnisse für mittelständische Unternehmen. Die Aufbereitung von Ergebnissen, wie Methoden, Werkzeuge, Prozesse, Leitfäden oder Konzepte, sollte derart erfolgt sein, dass ein breiter und branchenübergreifender Erfahrungsaustausch und Wissenszuwachs angeregt wird (BMBF, 2017, S. 4). Allgemein stellt dieser gesamte Sammelband einen Teil der Maßnahmen zur Verbreitung der Forschungsergebnisse und zum Erfahrungs- und Wissensaustausch dar. Insbesondere in diesem Kapitel soll daher nochmals der Nutzen der Forschungsergebnisse für KMU zusammenfassend erklärt werden, bevor die nachfolgenden Beiträge spezifisch auf konkrete Arbeiten und Forschungsergebnisse eingehen.

Bereits im vorangegangenen Kapitel wurde das Verbundprojektziel des Forschungsvorhabens zur vernetzten und integrierten Anwendung industrieller Datenanalyse für die wertschaffende, kompetenzorientierte Kollaboration in dynamischen Wertschöpfungsnetzwerken vorgestellt (siehe Kap. 1). Das im Weiteren stets als **AKKORD** bezeichnete Projekt ist Teil der Forschungsförderung durch InKoWe und behandelt die technischen, organisatorischen und menschlichen Aspekte zur **Anwendung industrieller Datenanalysen** in dynamischen Wertschöpfungsnetzwerken. Ein wichtiger Forschungsansatz in InKoWe ist die Verallgemeinerung unterschiedlicher interdisziplinärer Anwendungsfälle, aus denen neue und datengetriebene Innovationen in der Produktion entstehen. **AKKORD** folgt dieser Abstraktion von konkreten Anwendungsfällen, die im Projektkontext als **Erfolgsgeschichten** bezeichnet werden und exemplarische Umsetzungen der industriellen Datenanalyse sowie ihrer Peripherieprozesse darstellen. Die praktischen Anwendungen teilen sich in die beiden Bereiche **fachlich** und **strategisch** auf. Unter den Unternehmen dieser Erfolgsgeschichten finden sich also sowohl Großunternehmen wie auch mittelständische und kleine Unternehmen. Nachfolgend werden die Erfolgsgeschichten aus den vier **fachlichen** Anwendungsbereichen aufgezählt:

- **Vernetztes Industrial Engineering (Volkswagen AG):** Wissensbasierte Entscheidungsfindung zur Arbeitssystemgestaltung und Kollaboration im internen Unternehmensnetzwerk (Konzern/Marke/Werke; siehe Kap. 9)

- **Integriertes Qualitätsmanagement** (*Miele & Cie. KG*): Übergreifendes und vernetztes Analysesystem zur Qualitätssteuerung und Kollaboration entlang des Produktlebenszyklus (Produzent/Servicepartner/Kunde, siehe Kap. 10)
- **Prospektive Ressourcenplanung** (*ERCO GmbH*): Integrierte Auftragsprognose zur Kapazitätsplanung in der Produktion und Kollaboration im Wertschöpfungsnetzwerk (Produzent/Vertrieb/Lieferant, siehe Kap. 11)
- **Optimierung automatisierter Herstellungsprozesse zur Kunststoffverarbeitung** (*PolyMerge GmbH*): Überwachung, Analyse und Vorhersage von datenbasierten Mustern am Beispiel des Kunststoffschweißens und der Anomalieerkennung (siehe Kap. 12)

Außerdem wurden im Forschungsprojekt vier weitere Erfolgsgeschichten zu **strategischen** Handlungsfeldern umgesetzt. Diese wurden federführend von Partnern aus der Hardware- und Softwareentwicklung, der Beratung oder Forschung umgesetzt.

- **KMU-gerechte Kompetenzentwicklung für industrielle Datenanalyse** (*NEOCOSMO GmbH*): Kompetenzentwicklung im Industrial Engineering sowie werks- und funktionsübergreifende Kompetenzvernetzung (siehe Kap. 13)
- **Datenakquisition über I4.0-Technologien** (*Arendar IT-Security GmbH*): Synthese von Potenzialen zur Funktionserweiterung von IoT-Technologien für Retrofitting und Maschinendatenerfassung in der Produktion (siehe Kap. 14)
- **Adaptive Geschäftsmodellentwicklung** (*Deutsches Forschungszentrum für Künstliche Intelligenz*): Prozessbasierte Geschäftsmodellempfehlungen im Produktlebenszyklus und Co-kreative Leistungserbringung zur Datenerfassung und -vernetzung (siehe Kap. 15)
- **Handlungsstrategien zur industriellen Datenanalyse** (*mosaiic GmbH*): Einführungskonzept zur Kollaboration durch Datenanalyse und Vorgehen zur Adaptierung anwendungsspezifischer Analyseumgebungen (siehe Kap. 16)

Während die fachlichen Anwendungsbereiche konkrete Umsetzungen industrieller Datenanalysen bei produzierenden Unternehmen beschreiben, behandeln die strategischen Bereiche wichtige Funktionen und Prozesse, die für eine integrierte und vernetzte Anwendung erforderlich sind.

Neben der Umsetzung der Erfolgsgeschichten wurden im Projekt AKKORD auch **Technologie-Demonstratoren** erarbeitet, die (Teil-)Lösungen des Referenzbaukastens anschaulich darstellen. Die Demonstratoren wurden mit dem Ziel einer haptischen Veranschaulichung von hauptsächlich softwareseitigen Lösungen umgesetzt. Genutzt werden die Demonstratoren beispielsweise bei Messebesuchen oder sonstigen Praxisvorträgen, da sie es ermöglichen, die größtenteils softwarebasierten Entwicklungen mit einem physischen Ergebnis zu veranschaulichen. Insbesondere die Forschungspartner waren federführend an der Umsetzung der nachfolgenden Demonstratoren beteiligt:

- **Autorennbahn mit durchgängigem Datenbackend** (*Lehrstuhl für virtuelle Produktentwicklung*): Datenerfassung, -integration und -vernetzung mithilfe der modularen Lösungen des AKKORD-Referenzbaukastens (siehe Kap. 17)
- **Autorennbahn mit standardisierten Analysemodulen** (*Institut für Produktionssysteme*): Echtzeitfähige Aufnahme und Auswertung von Daten einer Autorennbahn von zwei bis vier Personen mithilfe des AKKORD-Referenzbaukastens (siehe Kap. 18)
- **Mobile Lernstation mit Industrie 4.0-Hardware** (*Arendar IT-Security GmbH, Lehrstuhl Berufspädagogik in den technischen Fächern*): Kompakte Station mit einem hybriden Lehrkonzept zur Veranschaulichung von industrieller Datenerfassung und den dazugehörigen Sicherheitskonzepten (siehe Kap. 19)

Weiterführende Informationen mit einer ausführlichen Darstellung der Arbeiten finden sich in den jeweils referenzierten Beiträgen aus diesem Sammelband.

2.2 Ausgangssituation und Status Quo im Mittelstand

Seit Jahrzehnten gewinnt die Verwendung von **industriellen Datenanalysen** im Mittelstand an Bedeutung, da Unternehmen immer mehr Daten erheben und analysieren, um ihre Geschäftsprozesse zu optimieren. Grundsätzlich handelt es sich bei dieser trendartig wirkenden Entwicklung jedoch nicht um ein neues Phänomen. Bereits seit dem Aufkommen der industriellen Herstellung von Gütern Anfang des 20. Jahrhunderts stellen **zahlen- und faktenbasierte Auswertungen** die wichtigste Grundlage für betriebliche Entscheidungsprozesse dar (Deuse et al., 2022, S. 2 f.). Die nachfolgende Definition basiert auf den Arbeiten von West et al. (2021, S. 131) und dient als allgemeine Grundlage für das Begriffsverständnis der Arbeiten in allen nachfolgenden Kapiteln.

Industrielle Datenanalysen bezeichnen den Prozess von zahlen- und faktenbasierten Auswertungen in einem industriellen Anwendungsfall.

Die Mehrheit der Unternehmen im produzierenden Gewerbe kennt zumindest den Trend bzw. das Technologiefeld der industriellen Datenanalysen. Ihr Einsatz stellt für den deutschen Mittelstand jedoch eine Herausforderung dar. Eine wichtige Frage ist beispielsweise, wie die Vielzahl an Datenquellen und -formaten effektiv und effizient miteinander verknüpft werden kann. Zudem müssen mittelständige Unternehmen geeignete Fachkräfte finden, die über die notwendigen Fähigkeiten und Kenntnisse verfügen, um die Datenanalysen durchzuführen und zu interpretieren. Bei der flächendeckenden Nutzung innerhalb des eigenen Unternehmens sowie bei der Integration der Potenziale in ein neues, digitales Geschäftsmodell zögern bisher viele Akteur:innen.

Die im Nachfolgenden vorgestellten Umfrageergebnisse entstammen einer im Rahmen des Forschungsprojekts AKKORD durchgeführten Studie. Die Studienergebnisse sind

bisher noch unveröffentlicht. An einer Publikation der Ergebnisse wird derzeit gearbeitet und die hier genannten Ergebnisse verstehen sich als Vorabinformationen. An der Studie nahmen 197 Personen teil, die in kleinen Unternehmen mit weniger als 50 Mitarbeitenden (35,7 %), mittleren Unternehmen mit 50 bis 500 Mitarbeitenden (21,11 %) und großen Unternehmen mit mehr als 500 Mitarbeitenden (42,21 %) tätig sind. Die Befragung wurde im Jahr 2022 online durchgeführt. Antworten erhielt die Studie aus Deutschland, Österreich und der Schweiz.

Zunächst wurde die grundsätzliche Haltung zur Verwendung von datenwissenschaftlichen Verfahren sowie zur Künstlichen Intelligenz erfragt. Die Ergebnisse dieser Frage sind in Abb. 2.1 zusammengefasst. Bei der Befragung wurde zum einen eine Unterscheidung anhand der Unternehmensgröße (klein, mittel oder groß) sowie zum anderen anhand der Perspektive im Unternehmen (allgemeine Einstellung im Unternehmen oder Einstellung der Geschäftsführung) vorgenommen. Die Aufteilung in Größen folgt den zuvor genannten Mitarbeiterzahlen der Unternehmen. Die Perspektive der beiden Rollen wurde durch Abfrage der generell vorherrschenden Einstellung im Unternehmen sowie der Einstellung der Geschäftsführung gewonnen. Zunächst fällt der überwiegende positive Anteil der Antworten auf. Die Mehrheit der Befragten bescheinigt beiden Perspektiven eine positive oder sehr positive Haltung. Auffällig ist außerdem, dass die Perspektive der Geschäftsführung stets einen höheren Anteil der sehr positiven Einstellungen aufweist, im Vergleich zur generellen Einstellung (zwischen 7,2 % und 14,1 %). Weiterhin bemerkenswert ist der erhöhte Anteil der negativen bzw. sehr negativen Antworten zur generellen Einstellung bei großen Unternehmen. Mit 21,4 %

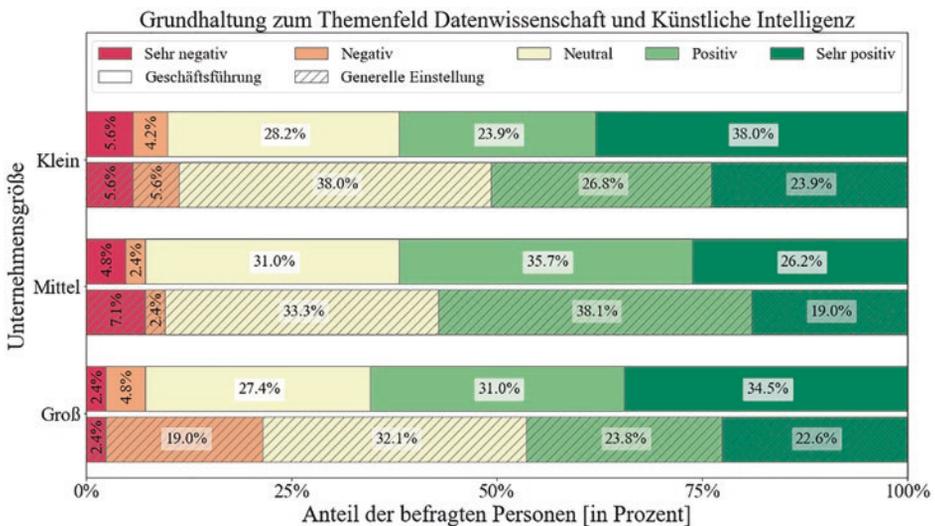


Abb. 2.1 Vorläufige Studienergebnisse zur Grundhaltung zu Datenwissenschaft und Künstlicher Intelligenz im produzierenden Gewerbe

liegt der Wert erheblich höher, als dies sowohl für die eigene Geschäftsführung als auch für mittlere und kleine Unternehmen der Fall ist.

Neben der Einstellung zum Themenfeld wurden in der Mittelstandsstudie auch die Wünsche bzw. Bedarfe der Anwenderunternehmen erfragt. Abb. 2.2 zeigt die Ergebnisse der Befragung auf. Dort sind sieben vorgegebene Anwenderbedarfe und die absolute Anzahl der Stimmen der befragten Personen dargestellt. Um die Lesbarkeit zu gewährleisten, wurde in dieser Darstellung auf eine Unterscheidung anhand der Unternehmensgröße verzichtet. Für ein eindeutiges Stimmungsbild wurden hier Antworten in lediglich vier Ausprägungsformen zugelassen. Die zur Gesamtanzahl von 197 Personen fehlenden Stimmen sind entsprechende Enthaltungen. Augenscheinlich ist, dass alle aufgeführten Wünsche einen hohen Grad an Zustimmung erhalten haben. Eine pragmatische Übersicht funktionierender Anwendungsfälle, eine entschlossene Geschäftsführung sowie ein interner Wissensaufbau belegen die ersten drei Plätze der Befragung. Lediglich eine Vernetzungs- und Austauschplattform mit Gleichgesinnten sowie der Wunsch nach schnellerer und einfacherer Förderung erscheint weniger wünschenswert. Dennoch liegt auch hier die Anzahl der Zustimmungen („stimme zu“ und „stimme eher zu“) deutlich über den Ablehnungen („stimme nicht zu“ und „stimme eher nicht zu“).

Diese beiden Umfrageergebnisse sollen zunächst ein grundsätzliches Stimmungsbild aufzeigen. Sowohl die tendenziell positive bis sehr positive Grundhaltung als auch die deutlichen Wünsche und Bedarfe der Anwender zeigen, dass industrielle Datenanalysen ihren Weg in den Alltag produzierender Unternehmen gefunden haben. Im nachfolgenden Abschnitt werden wir die konkreten Potenziale für Unternehmen benennen, die sich durch die Nutzung des entwickelten Referenzbaukastens ergeben.

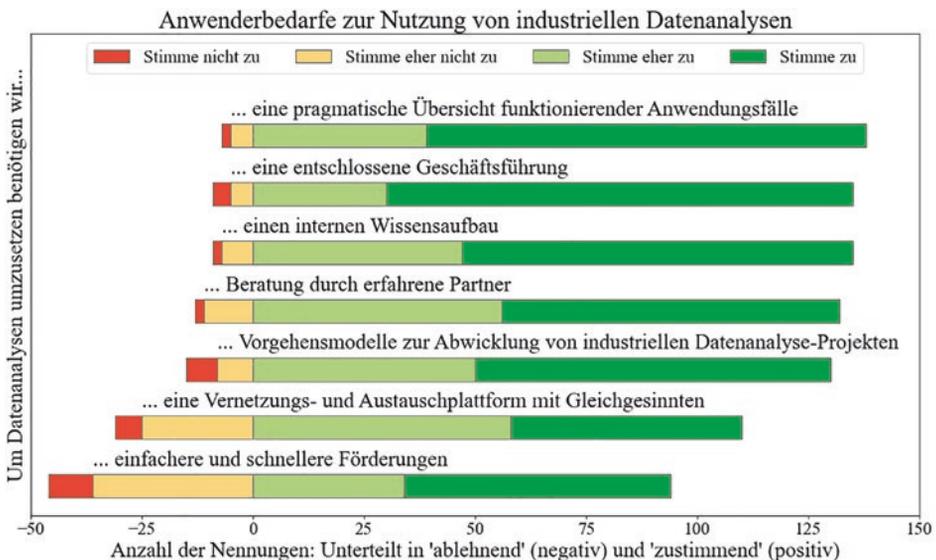


Abb. 2.2 Anwenderbedarfe zur Nutzung von industriellen Datenanalysen

2.3 Nutzenpotenziale für Kleine und Mittlere Unternehmen

Die im Forschungsprojekt AKKORD entwickelten Lösungen verfolgen sechs konkrete Projektziele, die bereits in Kap. 1 vorgestellt wurden. Durch die Umsetzung dieser Ziele ergeben sich wiederum sechs Nutzenpotenziale, die vor allem auf die Bedarfe von KMU zugeschnitten sind und eine Ausschöpfung der Möglichkeiten industrieller Datenanalysen umfassen.

1. Umfangreicher **Werkzeugkoffer** mit Verfahren, Methoden und Vorgehensweisen als Referenzbaukasten
2. Durchgängige **Datenanbindung und -speicherung** durch ein integriertes und vernetztes Datenbackend-System zur Erfassung, Verknüpfung und Speicherung entlang des Produktlebenszyklus
3. Niederschwellige **Datenanalyse und -anwendung** durch ein System zur Konfiguration und Wiederverwendung von umgesetzten Datenanalysen basierend auf einem Katalog aus modularen Bausteinen
4. Langfristiger **Kompetenzaufbau und -sicherung** durch ein umfangreiches Schulungsangebot für selbstgesteuertes Lernen entlang vorgegebener Lernpfade und durch ein integriertes Wissensmanagement mithilfe der implementierten Kollaborationsplattform
5. Neue **Kollaborationsformen und Geschäftsmodelle** durch den Zusammenschluss mehrerer Abteilungen oder Werke bzw. unterschiedlicher Unternehmen sowie durch die Nutzung der technologischen Potenziale industrieller Datenanalysen
6. Anschauliche **Beispiele und Vorlagen** durch die Umsetzung der vier fachlichen Anwendungsbereiche und durch die drei Technologie-Demonstratoren

Integrierter, modularer Referenzbaukasten. Um die volle Leistungsfähigkeit und breitenwirksame Übertragbarkeit der aus den Teilzielen hervorgehenden Lösungen gewährleisten zu können, sind diese in einem ganzheitlichen, vernetzten Ansatz vereint und anforderungsgerecht bereitgestellt. Vor diesem Hintergrund wurde im Forschungsvorhaben AKKORD ein modularer und datengetriebener Referenzbaukasten entwickelt, der Unternehmen bei der Ausgestaltung und Integration dynamischer Kollaborationen zur industriellen Datenanalyse unterstützen soll. Eine Vorstellung des Referenzbaukastens befindet sich weiterführend in Kap. 4. Prinzipiell erleichtert das Zusammenwirken der entwickelten Lösungen die notwendigen Maßnahmen zur Vorbereitung, Durchführung und Nutzung von industriellen Datenanalysen. Alle Ergebnisse basieren auf Anwenderanforderungen, die zu Beginn des Forschungsvorhabens erfasst wurden und sich in den in Abschn. 2.2 dargestellten Studienergebnissen widerspiegeln. Der Referenzbaukasten wurde als kollaborative Service-Plattform umgesetzt, wodurch eine breitenwirksame Nutzung und Erweiterung des Baukastens gefördert werden sollen.

Schaffung einer übergreifenden und vernetzten Datenbasis. Die Grundlage der Datenanalyse ist die Vernetzung von Daten eines dynamischen Wertschöpfungsnetzwerks durch Erfassung, Integration und Verknüpfung unterschiedlicher Datenquellen entlang des Produktlebenszyklus. Zur Sammlung, Integration, Vernetzung und Speicherung von Daten, Strukturen, Semantik und Referenzen zu Daten in Quellsystemen wurden daher im Rahmen des Vorhabens innovative Datenakquisitionslösungen, Datenformate, Schnittstellen und Kommunikationsprotokolle sowie anwendungsspezifische Integrations- und Vernetzungsmodule festgelegt und entwickelt. Diese ermöglichen den Aufbau eines leistungsfähigen PLM-basierten Daten-Backend-Systems (inkl. Rechtemanagement, siehe Kap. 5).

Integrierte und vernetzte Analyse industrieller Daten. Um in Wertschöpfungsnetzwerken auch KMU den Einsatz von Datenanalyse mit vertretbarem Aufwand zu ermöglichen, gilt es Lösungen bereitzustellen, welche eine automatische, individualisierte Analyseprozess-erstellung realisieren. Ein wesentliches Ziel des Vorhabens war daher die Entwicklung eines benutzerorientierten Konfigurationsassistenten zur anwendungsspezifischen Datenanalyse. Hierzu wurden Bausteine wie einheitliche Analysemodule und -Dashboards erarbeitet, die je nach Anwendungsfall vorkonfiguriert zusammengestellt werden können. Neben der Gestaltung der Analysemodule und -Dashboards wurden die Schnittstellen und Regeln definiert und mit der sogenannten AI-Toolbox in einem anwendungsgerechten Werkzeug zur teilautomatisierten Konfiguration umgesetzt (siehe Kap. 6).

Kompetenzentwicklung und -sicherung in Wertschöpfungsnetzwerken. Mit den tiefgreifenden Veränderungen der Arbeitsprozesse durch Industrie 4.0 wird eine Weiterentwicklung der Kompetenzprofile von Mitarbeitern auf unterschiedlichen Ebenen erforderlich. Auch der effiziente und effektive Einsatz von Datenanalysetechnologien stellt besondere Anforderung an sowohl personenbezogene Kompetenzen als auch Fachkompetenzen. Vor diesem Hintergrund wurden im Rahmen des Forschungsvorhabens einerseits situierte Lernmodule zur anwendungsbezogenen Kompetenzentwicklung erarbeitet. Andererseits wurden Strukturen und praxistaugliche Handlungsempfehlungen entwickelt, die es ermöglichen, Kompetenzen zu erfassen, übergreifend zu vernetzen und nachhaltig zu sichern. Hierzu wurde mit der sogenannten Work&Learn-Plattform ein digitaler Wissensdienst für die KMU-gerechte Kompetenzentwicklung und -vernetzung im Themenfeld „Industrielle Datenanalyse“ umgesetzt (siehe Kap. 7).

Neue Kollaborationsmöglichkeiten und Geschäftsmodelle. Datenanalysen ermöglichen den Zusammenschluss mehrerer Abteilungen bzw. Werke im Unternehmen sowie die Kooperation unterschiedlicher Akteure im Wertschöpfungsnetz zur Schaffung neuer, gemeinsamer Geschäftsmodelle. Daher wurde ein Empfehlungsassistent entwickelt, der ausgehend von den Ergebnissen einer vernetzten und integrierten Datenanalyse Vorschläge und Handlungsempfehlungen für innovative und wirtschaftliche Geschäftsmodelle zur kollaborativen Wertschöpfung erzeugt (siehe Kap. 8).

Umsetzung von Anwendungen und Aufbau von Demonstratoren. Um den Referenzbaukasten und die entwickelten Lösungen zugänglich für produzierende Unternehmen zu machen, wurden drei Demonstratoren entwickelt. Sie sollen unterschiedliche Aspekte der Ergebnisse anhand von konkreten Umsetzungen aufzeigen. Die Demonstratoren wurden bereits in Kapitel 2.1 aufgelistet und werden weiterführend in den jeweiligen Beiträgen erläutert. An dieser Stelle sei erneut auf Abb. 2.2 verwiesen: Eine pragmatische Übersicht von funktionierenden Anwendungsfälle wurde in der Umfrage als wichtigste Forderung genannt. Die realisierten Demonstratoren (siehe Kap. 17, 18 und 19), ebenso wie die entwickelten Anwendungsfälle in Form von AKKORD-Erfolgsgeschichten, verstehen sich als konkrete Umsetzungen der erarbeiteten Konzepte. Mit ihrer Hilfe sollen die Forschungsergebnisse verständlich und zugänglich gemacht werden.

2.4 Ausblick

In diesem Kapitel wurde ein Einblick in die Entstehung des Forschungsprojekts AKKORD gegeben. Anhand von vorläufigen Studienergebnissen wurde ein Stimmungsbild im produzierenden Gewerbe zur Nutzung industrieller Datenanalysen aufgezeigt. Stichpunktartig wurden die konkreten Potenziale des Projekts benannt und mit den durchgeführten Arbeiten verknüpft. Das nächste Kapitel behandelt weiterführend die technischen Grundlagen und den Status Quo des Vorhabens (siehe Kap. 3). Darüber hinaus stellt das Kap. 20 einen umfassenden Ausblick auf die künftige Anwendung Industrieller Datenanalysen im Dreiklang aus Mensch, Technik und Organisation vor.

Literatur

- BMBF. (2017). *Bekanntmachung im Rahmen des Forschungsprogramms „Innovationen für die Produktion, Dienstleistung und Arbeit von morgen“*. BAnz AT 10.05.2017 B3. Bundesministerium der Justiz und für Verbraucherschutz.
- BMBF. (2021). *Bericht der Bundesregierung zur Hightech-Strategie 2025. Erfolgsmodell Hightech-Strategie für ein starkes Innovationsland Deutschland*. Bundesministerium für Bildung und Forschung (BMBF).
- BMWi. (2015). *Industrie 4.0 und Digitale Wirtschaft. Impulse für Wachstum, Beschäftigung und Innovation*. Bundesministerium für Wirtschaft und Energie (BMWi).
- BMWi. (2020). *Von der Vision in die Praxis. Industrie 4.0-Umsetzungsprojekte*. Bundesministerium für Wirtschaft und Energie (BMWi).
- Deuse, J., West, N., & Syberg, M. (2022). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering (IJPME)*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Ausgangssituation und technische Grundlagen

3

Technische und organisatorische Rahmenbedingungen zur Anwendung industrieller Datenanalysen in dynamischen Wertschöpfungsnetzwerken

Nikolai West , Marius Syberg , Jörn Schwenken , Thomas Eickhoff ,
Jens C. Göbel , Michael Frey , Andreas Emrich ,
Christopher A. Klupak , Felix Walker  und Jochen Deuse 

N. West (✉) · M. Syberg · J. Schwenken · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: nikolai.west@ips.tu-dortmund.de

M. Syberg
E-Mail: marius.syberg@ips.tu-dortmund.de

J. Schwenken
E-Mail: joern.schwenken@ips.tu-dortmund.de

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

T. Eickhoff · J. C. Göbel
Technische Universität Kaiserslautern, Kaiserslautern, Deutschland
E-Mail: eickhoff@mv.uni-kl.de

J. C. Göbel
E-Mail: goebel@mv.uni-kl.de

M. Frey · A. Emrich
Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, Deutschland
E-Mail: michael.frey@dfki.de

A. Emrich
E-Mail: andreas.emrich@dfki.de

C. A. Klupak · F. Walker
Universität Hamburg, Hamburg, Deutschland
E-Mail: christopher.andre.klupak@uni-hamburg.de

F. Walker
E-Mail: felix.walker@uni-hamburg.de

3.1 Einleitung

In den beiden vorangegangenen Kapiteln wurde zum einen die grundsätzliche Zielsetzung des Forschungsvorhabens AKKORD vorgestellt (siehe Kap. 1) und zum anderen wurde der Anwendernutzen der anvisierten Zielgruppe der Entwicklungsarbeiten aufgezeigt (siehe Kap. 2). Im vorliegenden Kapitel werden weiterführend die **Ausgangssituation** zu Beginn des Forschungsvorhabens betrachtet und die relevanten, **technischen Grundlagen** eingeführt.

Da im Rahmen von AKKORD ein Referenzbaukasten zur Umsetzung von wiederverwendbaren, industriellen Datenanalysen erarbeitet wird, führt dieses Kapitel nachfolgend sowohl Referenzmodelle und -architekturen als auch Plattformkonzepte und Kollaborationsplattformen ein. Abschließend werden die wissenschaftlich-technischen Grundlagen der vier im Forschungsvorhaben vertretenden Leistungsbereiche erläutert. Ausgehend von den Voraussetzungen einer integrierten Datenvernetzung wird die Anwendung industrieller Datenanalysen sowie die Prinzipien zu Kompetenzaufbau und -sicherung in Wertschöpfungsnetzwerken erläutert. Abschließend werden die Notwendigkeit von Kollaborationsmöglichkeiten und Potenziale zur Entwicklung von Geschäftsmodellen beschrieben.

3.2 Technische Grundlagen

Das zentrale Ziel im Forschungsprojekt AKKORD stellt die Entwicklung eines modularen **Referenzbaukastens** dar, durch den industrielle Datenanalysen in Wertschöpfungsnetzwerken zielgerichtet und wiederverwendbar eingesetzt werden können (Mazarov et al., 2019, S. 875). Im Folgenden werden daher die technischen Grundlagen zu Referenzmodellen und -architekturen, sowie von Plattformkonzepten und Kollaborationsplattformen vorgestellt.

3.2.1 Referenzmodelle und -architekturen

Bei einem **Referenzmodell** handelt es sich um eine abstrakte und ideale Darstellung eines allgemeinen Systems, das als Grundlage für die Entwicklung konkreter Anwendungsszenarien dient (Schwegmann, 1999, S. 53 ff.). Im Kontext vom Forschungsprojekt AKKORD sind damit insbesondere Software-Systeme gemeint. Ein Referenzmodell beschreibt dabei grundsätzlich verschiedene Elemente eines Systems, wie allgemeine Konzepte, Prozesse und Funktionen, die für eine Modellierung der jeweils spezifizierten Ausprägungen benötigt werden. Kennzeichnend ist, dass diese Elemente in wiederkehrenden Anwendungsfällen erneut genutzt werden können. Da Referenzmodelle **allgemeingültig formuliert** sind, müssen sie insbesondere von

Kleinen und Mittleren Unternehmen (KMU) für die jeweils gearteten Markt- und Wettbewerbssituationen angepasst werden (Becker et al., 2007, S. 2). Für eine wirtschaftliche Nutzung müssen die höheren Adaptionen- und Pflegeaufwände von den potenziellen Einsparungen durch das Referenzmodell kompensiert werden. Diese Herausforderung wird auch als das Referenzmodellierungsdilemma bezeichnet (Becker et al., 2002, S. 26).

Demgegenüber steht der Begriff der **Referenzarchitektur**. Der Begriff ist in wissenschaftlichen und praxisorientierten Arbeiten nicht unumstritten. Für eine umfassende Diskussion zum Verständnis von Referenzarchitekturen in den vergangenen Jahrzehnten verweisen wir weiterführend auf Reidt (2019, S. 18 ff.). Im Kontext der weiteren Ausführungen bezeichnet der Begriff eine spezifische Architekturbeschreibung, die als Anleitung fungiert und eine Reihe von Beschränkungen auf der Grundlage von verwandten Systemen beinhaltet (Lin et al., 2022, S. 8). Eine Referenzarchitektur kann ihrerseits auf einem Referenzmodell basieren. Damit handelt es sich um eine Musterumsetzung für eine konkrete Problem- oder Aufgabenstellung, die als Vorlage für die anschließende Entwicklung von realer Systemen dient. Eine Referenzarchitektur ist eine **konkrete Implementierung** des abstrakten Referenzmodells.

Der wesentliche Unterschied zwischen einem Referenzmodell und einer Referenzarchitektur liegt also im Grad der Spezifität. Während das Referenzmodell abstrakt und allgemein strukturiert wird, stellt eine Referenzarchitektur eine konkrete Implementierung des Referenzmodells dar und umfasst auch problemspezifische Details, die für die Erfüllung einer bestimmten Problemstellung relevant sind. Um die Unterschiede, aber auch die Gemeinsamkeiten von Referenzmodellen und -architekturen aufzuzeigen, werden im nachfolgenden zwei prominente Beispiele vorgestellt und verglichen.

Als eines der bekanntesten Referenzmodelle in der Software-Entwicklung gilt das in Abb. 3.1 gezeigte *Open Systems Interconnection Reference Model*, kurz **OSI-Referenzmodell** (ISO/IEC 7498-1:1994). Es wurde von der Internationalen Organisation für Normung entwickelt und bereits im Jahr 1984 standardisiert. OSI beschreibt eine abstrakte Architektur für Netzwerkkommunikation, die aus insgesamt sieben Schichten besteht. Von jeder Schicht werden spezifische Funktionen erfüllt und Dienste angeboten,

Abb. 3.1 Darstellung des Open System Interconnection Modells als Beispiel für ein Referenzmodell der Software-Entwicklung (i. A. a. ISO/IEC 7498-1:1994)



die von darüber liegend angeordneten Schichten genutzt werden können. Das OSI-Referenzmodell dient als Grundlage für die Entwicklung von konkreten Netzwerkarchitekturen, die in diesem Kontext auch als Netzwerkprotokolle bezeichnet werden (Meinel & Sack, 2012, S. 40). Es besitzt also den allgemein formulierten Charakter eines Referenzmodells und wurde in vielen Anwendungsfällen genutzt.

Insbesondere im deutschsprachigen Raum ist das *Referenzarchitekturmodell Industrie 4.0*, kurz **RAMI4.0** genannt, weitläufig bekannt. Abb. 3.2 zeigt eine leicht vereinfachte Darstellung des Modells in ein deutscher Sprache. Das Modell beschreibt eine umfassende Architektur zur Entwicklung und Implementierung von vernetzten, cyberphysischen Systemen zur Anwendung in der Industrie 4.0 (DIN SPEC 91345:2016-04). RAMI4.0 liefert einheitliche Rahmenbedingungen für konkrete Umsetzungen von Industrie 4.0 Lösungen fest, in dem es u. a. interne Maschinendaten, externe Sensoren und angrenzende Systemumgebungen vereinigt (Wöstmann et al., 2019, S. 96). Das Modell wurde vom Verband Deutscher Maschinen- und Anlagenbauer und vom Zentralverband Elektrotechnik- und Elektroindustrie gemeinsam entwickelt. Es besitzt ebenfalls einen schichtenbasierten Aufbau und legt sowohl die horizontale als auch die vertikale Integration fest, um eine Interoperabilität von industriell genutzten Systemen und Anwendungen zu gewährleisten. RAMI4.0 gilt als wichtiges Werkzeug für produzierende Unternehmen, das sie bei der Planung und Umsetzung von Industrie 4.0 Lösungen unterstützt, indem es die konkreten Implementierungen abstrakter Architekturzusammenhänge vorbereitet.

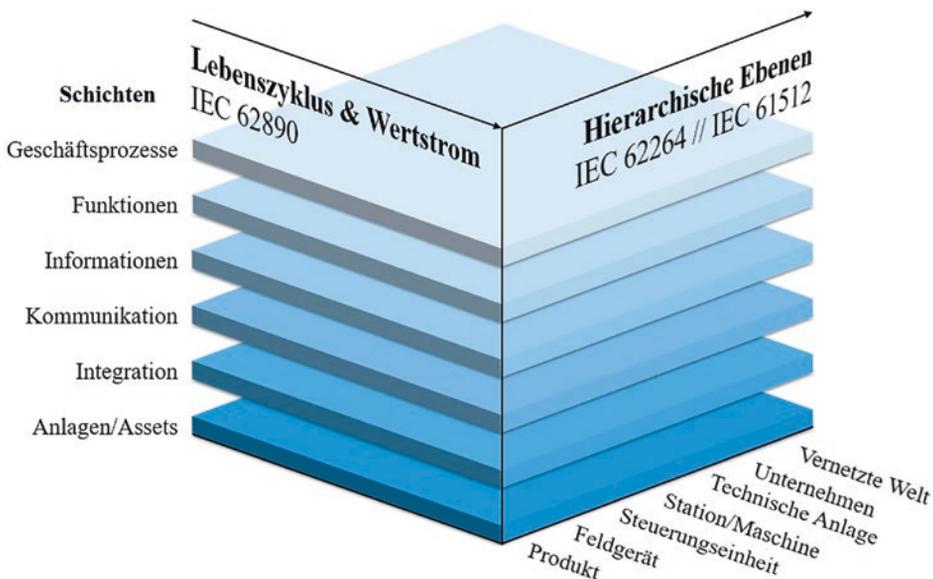


Abb. 3.2 Darstellung des Referenzarchitekturmodell Industrie 4.0 als Beispiel für eine Referenzarchitektur in der Software-Entwicklung (i. A. a. DIN SPEC 91345:2016-04)

Während das OSI-Referenzmodell also eine abstrakte und herstellerunabhängige Beschreibung der **Funktionsweise von Netzwerken** darstellt und die Grundlage für Netzwerkprotokolle bildet, dient RAMI4.0 als Werkzeug für Unternehmen und definiert die **Schnittstellen und Beziehungen** der vier Schichten in einer Wertschöpfungskette. Damit ist RAMI4.0 durch die konkrete Ausrichtung auf Anforderungen und Charakteristika der Industrie 4.0 deutlich spezifischer. OSI hingegen ist allgemeiner und beschreibt die Grundprinzipien der Netzwerkkommunikation, die in vielen Anwendungsbereichen eingesetzt werden können.

3.2.2 Plattformkonzepte und Kollaborationsplattformen

Der Begriff **Plattformkonzept** bezeichnet ein allgemeines Konzept der Informations- bzw. Organisationstechnik, das auf einer zentralen Infrastruktur zur Entwicklung und Bereitstellung von Diensten und Anwendungen basiert. Diese Infrastruktur wird als Plattform bezeichnet, da sie eine einheitliche Schnittstelle zur Erstellung, Integration und Nutzung von sowohl von Hardware- als auch von Software-Komponenten anbietet (Engelhardt et al., 2017, S. 11, 2017, S. 22). Plattformkonzepte zielen darauf ab, eine Vereinfachung in der Entwicklung von Anwendungen zu erreichen, indem verschiedene Entwicklungsressourcen bereitgestellt und die Zusammenarbeit zwischen Diensten und Entwicklern unterstützt wird. Gleichzeitig werden Skalierbarkeit und Flexibilität gewährleistet, durch die schnelle Bereitstellung und Aktualisierung von Diensten und Anwendungen. Der Siegeszug von Plattformen basiert vor allem auf den Fortschritten im Bereich des Cloud-Computing, das als wichtige Grundlage für die Bereitstellung von Anwendungen und Diensten gilt (Pfeiffer et al., 2019, S. 2). Während sich Plattformkonzepte auf spezifische Technologien oder Branchen beziehen können, wird nachfolgend insbesondere die Anwendung industrieller Datenanalysen behandelt.

Als **Kollaborationsplattformen** werden spezifische Software-Anwendungen bezeichnet, die einen Austausch oder die Zusammenarbeit zwischen verschiedenen Personen, Gruppen oder Organisationen ermöglichen, um gemeinsame Ziele zu erreichen. Dabei handelt es sich meist um eine digitale Plattform, die menschlichen Nutzern den Austausch von Informationen und Wissen ermöglicht, die Kommunikation unterstützt und die gemeinsame Arbeit an Projekten erleichtert. Kollaborationsplattformen stellen dafür zielgerichtete Funktionen bereit, wie beispielsweise Nachrichtenaustausch, Videotelefonie, Dateifreigabe, Projektmanagement-Werkzeuge oder gemeinsame Dokumentenbearbeitung (Weißmann & Hardwig, 2020, S. 7 ff.). Einige Kollaborationsplattformen integrieren zusätzlich Werkzeuge oder Anwendungen von Drittanwendern, beispielsweise E-Mail-Programme, Projektmanagement-Software oder Kundenbeziehungsmanagement. In den vergangenen Jahren haben Kollaborationsplattformen insbesondere in dezentralisierten oder globalisierten Arbeitsumgebungen Einzug gehalten, da sie Kooperationen über geographische und organisatorische Grenzen hinweg ermöglichen (Syberg et al., 2023, S. 64 f.).

Während eine Kollaborationsplattform also primär Unterstützung zur Zusammenarbeit und Kommunikation unterschiedlicher Akteure bereitstellt, bezieht sich ein Plattformkonzept hingegen auf eine umfangreiche Palette von Funktionen und Technologien, die zum Aufbau und Betrieb der Plattform notwendig sind. Zusammenfassend beschreibt ein Plattformkonzept also die strategische Ebene einer Plattform, während eine Kollaborationsplattform vor allem die operative Ebene umfasst, die Nutzern eine Zusammenarbeit ermöglicht.

Wie zuvor werden nachfolgend erneut zwei Beispiele betrachtet, um die Zusammenhänge und Unterschiede von Plattformkonzepten und Kollaborationsplattformen zu verdeutlichen.

Ein bekanntes Beispiel für ein Plattformkonzept stellen die *Amazon Web Services* dar, kurz **AWS**-Plattform genannt. Als Plattformkonzept basiert AWS auf einem serviceorientierten Ansatz, der Kunden Cloud-basierte Dienste und Infrastrukturen bereitstellt. Dazu zählen beispielsweise Ansätze wie Infrastruktur-, Plattform- und Software-as-a-Service (Amazon Web Services, 2021, S. 5 f.). Diese Leistungen gelten als skalierbar, robust und sicher und schließen datenwissenschaftliche Funktionen wie Verarbeitung, Speicherung und Analyse großer Datenmengen mit ein. AWS kann über eine einzige Schnittstelle und eine benutzerorientierte Konsole verwaltet werden, um eine durchgängige Verwendung sicherzustellen. Aufgrund der breiten Auswahl von Leistungen die AWS bereitstellt, gilt das Plattformkonzept als eines der erfolgreichsten Modelle für Cloud-Computing (Mufti et al., 2020, S. 4).

Ein ebenso verbreitetes Beispiel für eine Kollaborationsplattform ist *Microsoft Teams*, kurz auch oft nur **Teams** genannt, die als Teil von *Microsoft Office 365* angeboten wird. Als Anwendung bietet Teams Funktionen für Kommunikation, Dateiaustausch, Meetings und Zusammenarbeit an gemeinsamen Projekten an. So werden unterschiedliche Formen der Zusammenarbeit unterstützt, die in Echtzeit oder asynchron ablaufen können. Durch Teams können Menschen miteinander über Chat, Audio- und Videoanrufe, virtuelle Meetings oder durch den Austausch von Dateien kommunizieren (Buchal & Songsore, 2019, S. 2 ff.). Als Kollaborationsplattform ermöglicht Teams auch die Erstellung von Kanälen, um Gruppen von Teammitgliedern zusammenzufassen, um die Zusammenarbeit an spezifischen Projekten oder Themen zu erleichtern. Zusammenfassend gilt Teams als umfassende Kollaborationsplattform, da sie speziell für Unternehmensbedürfnisse entwickelt wurde und eine effiziente sowie koordinierte Zusammenarbeit operativer Tätigkeiten ermöglicht (Ilag & Sabale, 2022, S. 2 ff.).

3.3 Status Quo der vier Leistungsbereiche

Die vernetzte und integrierte Nutzung von industriellen Datenanalysen in dynamischen Wertschöpfungsnetzwerken erfordert das Zusammenwirken von sehr unterschiedlichen Themen- und Fachgebieten. Wie bereits in Kap. 1 erläutert werden dazu in AKKORD vier als Leistungsbereiche bezeichnete Teilbereiche betrachtet. Dazu zählen Werkzeuge

zur integrierten **Datenvernetzung**, zur Anwendung der industriellen **Datenanalyse**, zum Aufbau und zur Sicherung von **Kompetenzen** sowie zur Ausschöpfung von **Kollaborationsmöglichkeiten** und neuen Geschäftsmodellen. Die wissenschaftlich-technischen Grundlagen der vier Leistungsbereiche werden in den nachfolgenden vier Abschnitten sukzessiv erläutert.

3.3.1 Integrierte Datenvernetzung

Ein großes Problem für die industrielle Datenanalyse in KMU ist der Umstand, dass die Daten häufig redundant, auf verschiedenen Systemen und in unterschiedlichen Formaten verteilt sind. Dies erschwert die Arbeit mit den Daten, da unklar ist, ob auf alle relevanten Daten zugegriffen werden kann. Als zentrale IT-Lösung des Produktentwicklungsprozesses bietet sich das Produktdatenmanagement und das darauf aufbauende *Product Lifecycle Management* (PLM) an (Eigner et al., 2014, S. 267 ff.). Hier werden alle Dokumente, die von der ersten Idee bis zur finalen Produktstruktur anfallen, verwaltet. Im weiteren Verlauf der Produktentstehung werden die produktions- und logistikbezogenen Informationen in *Produktionsplanungs- und -steuerungssystemen* (PPS) verwaltet (Eigner et al., 2014, S. 301 ff.). Eine große Herausforderung stellt für viele Unternehmen das Aftersales-PLM dar. Der Umgang mit den Daten aus Ersatzteilmanagement, Wartung, Reparatur, Überholung und Dienstleistungen sowie das Zurückspielen von Felddaten, beispielsweise über IoT-Plattformen, an die Produktentwickler verlangt, dies alles auf einem gemeinsamen Produkt- und Prozessmodell umzusetzen.

Durch den technischen Fortschritt ändern sich sowohl die **Funktionen** als auch die entstehenden **Daten**. Davon sind auch bereits bestehenden Systemen betroffen. So ändert sich beispielsweise durch IoT-Plattformen die Möglichkeit des Datenzugriffs. Angebundene IoT-Assets können in Echtzeit Informationen an die Plattform liefern, was unmittelbare Analysen ermöglicht. Moderne Vorgehensmethoden des Model Based System Engineerings ermöglichen eine durchgängige Vernetzung von Produktdaten, erfordern im Gegenzug aber auch ein kohärentes Gesamtkonzept zur Datenverwaltung. Des Weiteren werden auch gänzlich neue Felder der Datengenerierung erschlossen, beispielsweise durch die Entwicklung und Anwendung von *Digitale-Zwillinge-Konzepten*. Der digitale Zwilling ist hierbei mehr als die Repräsentation eines realen Systems oder Produkts, sondern bietet auch eine Datenverarbeitung oder Interaktionen und Eingriffsmöglichkeiten für den laufenden Betrieb an (Göbel & Eickhoff, 2020, S. 74 ff.). Durch diese erweiterten Objekte kann die Vision eines entlang des Produktlebenszyklus durchgängigen und rückverfolgbaren Produktmodells realisiert werden (Gilz & Eigner, 2013, S. 293 ff.). In diesem Kontext wird auch von System Lifecycle Management gesprochen (Eigner et al., 2014, S. 374).

Einen **Anwendungsfall** für integrierte Datenvernetzung stellen Savarino und Dickopf (2022, S. 257 ff.) vor. Cyber-physische Systeme, die um Smart Services erweitert werden, haben einen hohen Anteil an interdisziplinären Komponenten. Durch die

Analyse der Stücklisten und Arbeitspläne in Kombination mit den Daten aus dem Echtzeit-Feedback der Nutzungsphase, kann ein ganzheitlicher Ansatz für intelligente Produktstücklisten und Arbeitsplänen entwickelt werden.

Für die zur industriellen Datenanalyse nötigen Integration der Daten in ein Backendsystem ergeben sich gleich mehrere Herausforderungen, nämlich die Verteilung, Autonomie und die Heterogenität der Daten (Leser & Naumann, 2007, S. 49 ff.). Bei der Verteilung besteht zum einen das physische Problem, das Daten sich auf unterschiedlichen Speicherorten befinden. Zum anderen kann es auch ein logisches Problem werden, wenn die Daten auf unterschiedlichen Systemen und teilweise redundant gespeichert werden. Vor allem wenn die Redundanz unkontrolliert geschieht, kann dies durch den fehlenden Abgleich zu Problemen führen. Die Autonomie der Datenquellsysteme führt bei jedem System zu eigenen Regeln und Strukturen für die generierten Daten, Schnittstellen und Zugriffsrechte. Durch die Verteilung und der Autonomie entsteht eine große Heterogenität an Datenstrukturen, Schnittstellen, Austauschformaten sowie Kommunikationsprotokollen (Leser & Naumann, 2007, S. 58 ff.).

Zum Umgang mit diesen Herausforderungen wurden bereits erste Ansätze entwickelt. Ihnen ist gemein, Daten in einem schlanken Backbone in neutralen, transparenten und archivierbaren Formaten zu verwalten (siehe beispielsweise Mogo Nem, 2011; Ernst, 2016). Der Sicherstellung der Datenqualität fällt in dieser Thematik eine Schlüsselrolle zu, um Multi Source Probleme wie beispielsweise widersprüchliche Werte oder unterschiedliche Einheiten aus dem Weg zu gehen (Helmis & Hollmann, 2009).

Ein Ansatz, um mit diesen Problemen umzugehen, ist die Erweiterung eines Data Warehouse um einen **Data Lake**. In einem Data Warehouse ist die Datenstruktur im Voraus definiert und für schnelle Abfragen optimiert. Die Daten sind bereinigt, umgewandelt und können als sogenannte *Single-Source-of-Truth* betrachtet werden. Ein Data Lake kann nicht nur wie ein Data Warehouse strukturierte Daten, sondern auch unstrukturierte Daten speichern. Diese unstrukturierten Daten können entweder durch den Verzicht der Strukturierung generiert werden oder aus Quellen wie mobilen Apps, IoT-Geräten oder sozialen Medien gewonnen werden. Der Vorteil der Kombination dieser Speicherkonzepte ist die Möglichkeit der Nutzung neuer Abfrage- und Analysemodellen. Die Umsetzung kann hierbei auf eigener Hardware oder unter Einsatz von Cloudlösungen geschehen. Die Verwendung standardisierter Softwareplattformen ermöglicht hierbei auch nachträglich ein hohes Maß an Flexibilität (vgl. Amazon Web Services, 2021).

3.3.2 Industrielle Datenanalyse

Für die Umsetzung von industriellen Datenanalysen müssen verschiedene **Voraussetzungen** erfüllt werden. Die industrielle Produktion ist ein komplexer Prozess, der aufgrund seiner zahlreichen Variablen und dynamischen Eigenschaften eine systematische Überwachung und Analyse erfordert. In diesem Zusammenhang sind innovative

Methoden der Datenanalyse von entscheidender Bedeutung, um eine effiziente Kontrolle und Steuerung der Produktionsprozesse zu ermöglichen (Deuse et al., 2022, S. 1 ff.). Eine erweiterbare Methode zur Verteilung und Konfiguration der Sensorik im gesamten Fertigungsprozess und die Speicherung der Daten in beschreibenden Prozessen und Modellen sind grundlegende Voraussetzungen, um einen umfassenden Überblick über den Fertigungsprozess zu gewährleisten. Weiterhin sind Methoden zur Erkennung und Fehlerbehebung von Ereignissen in der Produktion sowie Methoden zur Speicherung von Echtzeitdaten und Durchführung von Korrelationsanalysen für große Datensätze und Datenströme nach bestimmten Dimensionen unabdingbar, um mögliche Fehler und Abweichungen im Produktionsprozess frühzeitig zu erkennen. Darüber hinaus erfordert eine erfolgreiche Datenanalyse in der industriellen Produktion eine kontinuierliche Anpassung komplexer Ereignisvorhersagemodelle, um einen reibungslosen Produktionsablauf unter Berücksichtigung der auftretenden Dynamik zu gewährleisten. Eine Methode zur Erzeugung von Alarmmeldungen als Reaktion auf erwartete Abweichungen vom geplanten Produktions- bzw. Herstellungsziel auf der Basis einer rechnerischen Analyse ist dabei ein wichtiger Faktor (Krumeich et al., 2014, S. 16 ff.; Babiceanu & Seker, 2016, S. 132 f.).

Um Produktionsfehler zu reduzieren und die Effizienz des Produktionsprozesses zu erhöhen, ist es auch notwendig, Vorschläge und automatisierte Entscheidungsmethoden zur Verfügung zu stellen. Schließlich ist eine Methode zur aktiven Anpassung des Produktionsprozesses auf der Grundlage von Berechnungen und Analysen ein zentraler Aspekt, um eine reibungslose Produktion zu gewährleisten und die Produktionsqualität zu erhöhen. Zur automatisierten Auswertung großer Datenmengen stehen in der modernen Datenanalyse einige Verfahren zur Verfügung, darunter auch das Data Mining (Hastie et al., S. 1 ff.; Fayyad et al., 1996, S. 27 ff.). Durch die Anwendung des Data Mining können bisher unbekannte Strukturen und Zusammenhänge in großen Datenmengen entdeckt und wertschöpfend genutzt werden (Witten et al., 2011, S. 3). Die möglichen Lösungsansätze sind dabei ebenso vielfältig wie die jeweiligen Anwendungsfälle: In der Instandhaltung können beispielsweise durch Vorhersagen Aussagen über den bevorstehenden Ausfall von Bauteilen oder Anlagen getroffen werden (Wöstmann et al., 2019, S. 94 ff.). In der Produktionsplanung können beispielsweise durch selbstlernende Agenten Vorschläge für optimale Produktionspläne gemacht werden (West et al., 2021b, S. 347 ff.). In der Produktionssteuerung können beispielsweise durch gezielte Vorhersagen Engpässe antizipiert und proaktiv vermieden werden (West et al., 2022a, S. 612 ff., b, S. 3 ff.). Im Qualitätsmanagement können mithilfe von Anomalie-Erkennung bisher unentdeckte Fehler identifiziert und verhindert werden (Schlegl et al., 2022, S. 1 ff.; West et al., 2021c, S. 762 ff.).

Zur Informationsgewinnung durch das Data Mining stehen verschiedene Prozessmodellen zur Auswahl, wie beispielsweise *Knowledge Discovery in Database* (Fayyad et al., 1996, S. 27 ff.) oder *Cross-Industry Standard Process for Data Mining*, abgekürzt auch CRISP-DM genannt (Chapman et al., 2000, S. 1 ff.). Insb. das CRISP-DM besitzt eine gute Eignung für industrielle Datenanalysen, da es mit seiner klar definierten

Projektstruktur als intuitiv gilt und so entwickelt wurde, dass es branchenunabhängig genutzt werden kann. Das Vorgehensmodell ist in Abb. 3.3 dargestellt und zeigt den iterativen Charakter des Modells auf.

Zur Aufdeckung der möglich nutzbaren Potenziale muss zunächst das Ziel des Projekts definiert werden. Außerdem müssen alle nötigen Datensätze zur Verfügung stehen. Der für die Modellierung und somit für das Ergebnis zentralste Schritt ist die Vorverarbeitung der Datensätze. In diesem zeitaufwendigen Schritt werden Datenbestände selektiert oder zusammengeführt, fehlende Werte ersetzt, Ausreißer eliminiert und ggf. Daten transformiert. Mit den bereinigten Datensätzen als Ergebnis der Datenvorverarbeitung können verschiedene Modelle trainiert werden. Mittels überwachter oder unüberwachter Lernverfahren werden so neue Muster entdeckt und Zusammenhänge verifiziert (Witten et al., 2011).

Unüberwachte Lernverfahren, wie das Clustering oder Assoziationsregeln, dienen dem Auffinden neuer Zusammenhänge und Muster in historischen Datenbeständen, ohne dass eine Zielgröße bekannt ist. Bei den **überwachten Lernverfahren**, beispielsweise Klassifikations- oder Regressionsverfahren, ist dagegen die Zielgröße bekannt und kann im Sinne einer Prognose prospektiv abgeschätzt werden. Darüber hinaus können zur Bewältigung hoher Komplexität in gewissen Problemstellungen Agentensysteme eingesetzt werden. Im Bereich der Informatik agieren mehrere Agenten innerhalb eines solchen Systems, die jeweils mit einem Verhalten einen gewissen Zielzustand kontrollieren können. Ein Ansatz zur Kombination solcher Agentensysteme mit Künstlicher Intelligenz stellt das Reinforcement Learning, das **bestärkende Lernen**, dar. Dabei lernt ein Agent, selbstständig zu agieren und durch ein gewisses Verhalten

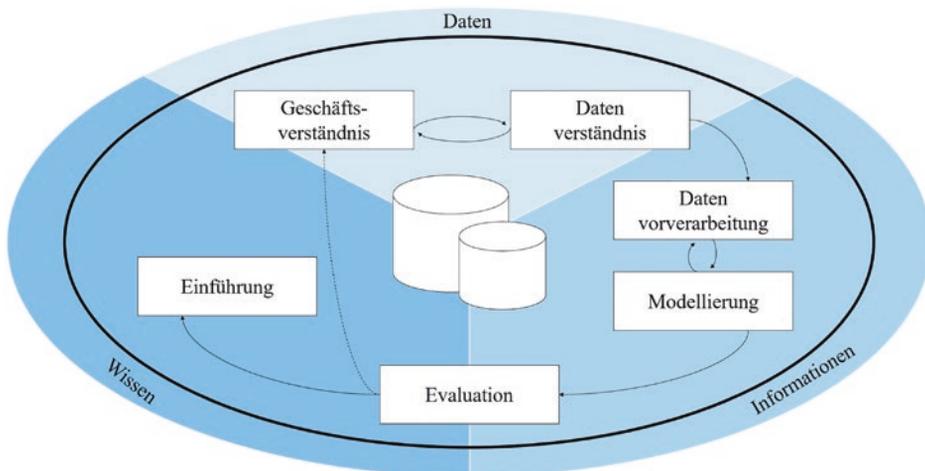


Abb. 3.3 Darstellung des Cross-Industry Standard Process for Data Mining (CRISP-DM) im Kontext von Daten, Informationen und Wissen (i. A. a. Chapman et al., 2000, S. 13; West et al., 2021a, S. 133)

langfristig seine Belohnung zu optimieren. Er trifft dadurch bessere Entscheidungen, was in unterschiedlichen Problemstellungen vorteilhaft ist (Hastie et al., 2009, S. 5 f.).

Die Erkenntnisse aus dem Data Mining und das praktische Erfahrungswissen der Mitarbeiter bilden in Kombination eine wesentliche Grundlage zu Unterstützung für zukünftige Entscheidungen in Unternehmen (Deuse et al., 2014, S. 5 f.). Pioniere in der Nutzung des Data Mining sind die Banken und Versicherungsgesellschaften (Prasad & Madhavi, 2012, S. 96). Bereits seit längerer Zeit prognostizieren sie Entwicklungen und schätzen deren Risiko ab, indem sie große Mengen historischer Daten auswerten und mit ihnen echtzeitnahe Daten generieren. Industrieunternehmen setzen in den letzten Jahren jedoch auch vermehrt auf die Methoden des Data Mining, vor allem in den Einsatzfeldern des Qualitätsmanagements, der Prozessoptimierung und der Instandhaltung (Dogan & Birant, 2021). Eine breite Adaption dieser Ansätze auf den Industriestandort Deutschland, in dem der Großteil der Unternehmen KMU-ähnliche Strukturen aufweist, ist noch nicht erfolgt. Hemmnisse, die einer breiten Anwendung von Data Mining Verfahren entgegenstehen, wurden im Rahmen einer Studie des Fraunhofer IPA identifiziert (siehe Weskamp et al., 2014). Neben dem Vorhandensein einer ausreichenden Menge an Datensätzen in der für die Aufgabe erforderlichen Qualität sind für die Anwendung von Data Mining neben Domänenwissen spezifische Kompetenzen und Fachwissen erforderlich, um Modelle trainieren und evaluieren zu können. Weiterhin bleiben den befragten Unternehmen die erforderlichen Voraussetzungen sowie der erzielbare Zusatznutzen unklar.

Trotz der anwenderfreundlichen Gestaltung der bestehenden Werkzeuge zur Datenanalyse, beispielsweise durch modulbasierte oder *Drag-and-Drop*-Gestaltung der Software, wirkt dies bisher kaum dem **fehlenden Fachwissen** der Anwender entgegen. Da noch immer die geeigneten Parameter und Verfahren vom Anwender auszuwählen sind, kann von keiner Anwendungsunterstützung durch die Software gesprochen werden. Um entsprechende Verfahren der Datenanalyse zielführend anzuwenden, ist es im industriellen Kontext entscheidend, die entsprechenden Gegebenheiten in diesem Umfeld zu berücksichtigen. Dies gilt in der **Kompetenzentwicklung** wie auch in der **Gestaltung geeigneter Werkzeuge** zur Anwendung. Dabei müssen u. a. sowohl die Individualität der Anwendungsfälle, als auch Sicherheitsaspekte oder technische Voraussetzungen beachtet werden (Syberg et al., 2023, S. 64 ff.). Eine ausführliche Betrachtung der umgesetzten Werkzeuge zur Unterstützung von industriellen Datenanalysen findet sich weiterführend in Kap. 6.

3.3.3 Kompetenzaufbau und -sicherung

Die vorhandenen Kompetenzen der Mitarbeitenden sind maßgeblich für den Umsetzungserfolg von industriellen Datenanalysen bei produzierenden Unternehmen entscheidend. Sie werden nicht nur für das erfolgreiche Anwenden von Methoden

benötigt, sondern auch um die richtige Methode für das passende Problem auszuwählen und die Ergebnisse zu interpretieren und richtige Schlussfolgerungen zu ziehen.

Die benötigten Kompetenzen der Mitarbeiter lassen sich in zwei Bereiche aufteilen: **Fachkompetenz**, die zur Bewältigung der sachlichen Aspekte der Berufsanforderung dient, und **methodenbezogene Kompetenz**, die universelle Problemlösefähigkeiten umfasst (Stefănică et al., 2017, S. 843 ff.). Die Kompetenzen und ihre Ausprägungen, die jeder Mitarbeiter benötigt, sind stark vom jeweiligen Aufgabengebiet abhängig (Becker et al., 2011, S. 1 ff.). Waren in der Vergangenheit die benötigten Kompetenzen bekannt, verschieben sich durch den Einzug der Industrie 4.0 diese, da sich auch die Arbeitsprozesse neu definieren. Dies nötigt die Unternehmen, die Kompetenzprofile ihrer Mitarbeiter weiterzuentwickeln (Arnold et al., 2016, S. 1 ff.).

In einer von Acatech durchgeführten Studie gibt es erste Hinweise, welche Kompetenzen zukünftig für die Mitarbeiter eines erfolgreichen Unternehmens relevant werden (acatech, 2016, S. 9 ff.). Im Fokus der Untersuchung stand die Erfassung, Förderung und Sicherung technologie-, daten-, prozess- und kundenorientierter Kompetenzen, sowie Kompetenzen zur Infrastruktur und Organisation. Demzufolge sind Fachkompetenzen im Bereich der Datenanalyse und des Verständnisses für Geschäftsprozesse ebenso relevant wie die personenbezogenen Kompetenzen zur Abstimmung, Kollaboration und Information aller Beteiligten. Ein Problem hierbei ist, dass es nur wenige Ansätze zur Erfassung des Standes der Kompetenzen der jeweiligen Mitarbeiter gibt.

Ein großes Problem bei der Ermittlung der Kompetenzen ist, dass zwar für die Fachkompetenzen objektive Leistungstests hinzugezogen werden können, für die personenbezogenen Kompetenzen allerdings oftmals nur über eine Selbsteinschätzung der Mitarbeiter Daten generiert werden können (Nickolaus & Walker, 2016, S. 7 ff.). Aus diesem Grund wird auf die Ergebnisse von Arbeiten zur Förderung von fachlichen Kompetenzen mittels digitaler Medien im beruflichen Bereich (Walker et al., 2016, S. 1 ff.) als auch im Bereich Industrie 4.0 (Faath-Becker & Walker, 2020, S. 109 ff.) zurückgegriffen. Die Kompetenzförderung wird besonders durch den Einsatz digitaler Technologien und den Prinzipien des situierten Lernens unterstützt (Stegmann et al., 2018, S. 967 ff.). Es lassen sich so vier Gestaltungsprinzipien für die technologieunterstützte, situierte Lernumgebung beschreiben (Rosen & Salomon, 2007, S. 1 ff.):

1. Authentizität und Anwendungsbezug mittels realer Probleme und Situationen.
2. Multiple Kontexte und Perspektiven zur Entwicklung der Fachkompetenzen.
3. Kooperatives Lernen und Problemlösen für den Umgang mit komplexen Problemen und die Entwicklung sozialer Kompetenzen.
4. Instruktionale Anleitung und Unterstützung, beispielsweise mittels Feedbacks.

Vor allem der letzte Punkt kommt in den situierten Lernprozessen eine besondere Rolle zu, da es hilft diesen zu regulieren. Allerdings hängt die Wirksamkeit des Feedbacks von verschiedenen Faktoren ab (Narciss, 2006, S. 1 ff.). So ist adaptives Feedback

beispielsweise wirksamer als nichtadaptives Feedback (Marschner, 2011, S. 1 ff.). Eine Studie von Birdman et al., (2022, S. 1139 ff.) gibt einen Hinweis, wie eine situierte Lernumgebung gestaltet werden kann. So wurden Studenten in einem Modul angehalten sowohl Vorlesungen als auch Gruppen- und Einzelarbeiten zu absolvieren. Der Fortschritt der Kompetenzbildung der Studenten wurde in regelmäßigen Interviews evaluiert. Durch den regelmäßigen Austausch der Studenten untereinander und mit den Lehrenden werden die vermittelten Kompetenzen nachhaltig verinnerlicht. Im Rahmen des Projektvorhabens wird erstmals dem oben formulierten Defizit begegnet, dass bislang auf Unternehmensebene große Unklarheit herrscht, welche Kompetenzen aufseiten der Mitarbeiter in Bezug auf industrielle Datenanalyse in Wertschöpfungsnetzwerken benötigt werden. Hierfür werden gezielte Recherchen, Analysen und Befragungen auf Unternehmensebene durchgeführt, aufbereitet und in entsprechende Instrumente zur Kompetenzmessung überführt.

Um dem Umstand, dass bezüglich der Kompetenzförderung im Bereich der industriellen Datenanalyse keine Ansätze bzw. Konzepte bestehen, zu begegnen, werden technologiebasierte situierte Lernmodule in Anhängigkeit der ermittelten Kompetenzen entwickelt. Die Basis der Lernmodule stellt der empirisch fundierte kognitionspsychologische Ansatz des situierten Lernens dar, welcher in besonderem Maße zum Kompetenzaufbau in Unternehmen geeignet ist. In die Lernmodule werden individuelle, adaptive Feedbackelemente implementiert, wodurch die Kompetenzförderung zusätzlich verstärkt werden soll.

Durch Integration der Instrumente und der Lernmodule innerhalb des als webbasierte, digitale Service-Plattform ausgeprägten Referenzbaukastens, wird ein digitaler Wissensdienst realisiert, welcher die Anwendung der Datenanalyse im Rahmen der Kompetenzentwicklung ermöglicht.

3.3.4 Kollaborationsmöglichkeiten und Geschäftsmodelle

Für die Forschung ist das Feld der Geschäftsmodelle seit dem Internetboom am Ende der 1990er Jahre zunehmend in den Fokus gerückt. Burkhart et al. (2011, S. 109 ff.) arbeiten verschiedene existierende Definitionen des Begriffs „Geschäftsmodell“ auf und versuchen auf der Basis den Begriff folgendermaßen zu definieren:

Ein Geschäftsmodell beschreibt die logische Funktionsweise eines Unternehmens und insbesondere die spezifische Art und Weise, mit der es Gewinne erwirtschaftet. Grundsätzlich soll die Beschreibung von Geschäftsmodellen helfen, die Schlüsselfaktoren des Unternehmenserfolges oder Misserfolges zu verstehen, zu analysieren und zu kommunizieren.

Aufgrund der Globalisierung und der zunehmenden Nachfrage nach kundenspezifischen Kleinserienprodukten werden die Fertigungsaktivitäten immer weiter verstreut und digital verwaltet. Gerade in Branchen mit geringer Fertigungstiefe sind häufig auch

viele kleine Akteure insb. KMUs einbezogen. Entsprechend müssen sich diese den neuen Integrationsbedürfnissen anpassen und ihre Geschäftsmodelle entsprechend ändern.

Im Rahmen der Digitalisierung sind neue Geschäftsmodelle zumeist datengetrieben und bieten ein Angebot an Services an, das ergänzend zum etablierten Ansatz des Produktverkaufs passt. Eine Vernetzung der Wertschöpfungsnetzwerke ist dabei essentiell für die Nutzung serviceorientierter Geschäftsmodelle, da ein Unternehmen in diesen Fällen mit einer Vielzahl von anderen Akteuren für die Leistungserbringung zusammenarbeiten muss. Die relevanten Netzwerke sind in diesem Fall das Entwicklungs- und Produktionsnetzwerk, OEM und das Servicenetzwerk. Diese Vernetzung wird erst dann möglich, wenn sich Unternehmen unterschiedlicher Kompetenzebenen im Wertschöpfungsnetzwerk einigen, wie Sensoren, Geräte, Infrastruktur, Plattformen und Services entlang der Wertschöpfungskette genutzt werden (BMW, 2015, S. 1 ff.). Der Einsatz einer zentralen Service-Plattform erscheint in diesem Zusammenhang äußerst sinnvoll, um die verschiedenen Akteure einzubinden und Kollaborationsmöglichkeiten zu diesen Fragestellungen anzubieten. Dadurch besteht die Möglichkeit zur Vernetzung der Akteure und zur Optimierung der individuellen oder gemeinsamen Geschäftsmodelle.

Gerade bei Produkten und Dienstleistungen mit einer komplexen Wertschöpfungskette wird es immer relevanter, sowohl positives als auch negatives Feedback aus der Kundenbasis möglichst genau einzelnen Geschäftsmodellkomponenten und Akteuren zuzuordnen. Für die Geschäftsmodelle selbst erlaubt dies ein selbstadaptives Geschäftsmodell, das sich entsprechend dem Feedback weiterentwickeln kann. Die Zuordnung zu Akteuren andererseits lässt die Lokalisierung von Problemen in der Zulieferkette einfacher realisieren. Somit können kollaborative Geschäftsmodelle kontinuierlich verbessert und angepasst werden, indem die jeweiligen Akteure entsprechend auf das Feedback eingehen und ihren Beitrag zum Geschäftsmodell und dem dazugehörigen Wertangebot verbessern.

Gerade für ein solches selbstadaptives Geschäftsmodell eignen sich Methoden der künstlichen Intelligenz, indem sie immer bessere Analysen kurzfristiger und langfristiger Kundenpräferenzen und der Nutzung von Produkten ermöglichen. Mittels natürlicher Sprachverarbeitung können hier Kundenaussagen passgenau automatisiert Komponenten eines Geschäftsmodells zugeordnet werden. Diese Methoden erlauben somit eine passive Integration des Kunden in die ständige Evaluation und Entwicklung des Geschäftsmodells.

Die **automatisierte Auswertung von Nutzungsdaten** von Produkten und Dienstleistungen liefert eine umfassende Datenbasis um die Qualität von Angeboten zu bewerten und den subjektiven Eindrücken der Nutzer im Sinne von Kundenfeedback eine objektive Erhebungsbasis gegenüber zu stellen. So lassen sich Korrelationen zwischen subjektivem und objektivem Feedback generieren, Inkonsistenzen und Prioritäten aufdecken sowie Zusammenhänge zwischen Kundenbewertungen und Produktdefekten identifizieren. Dies wiederum kann der Produktentwicklung wichtige Anhaltspunkte für die Entwicklung neuer Wertangebote und die Qualitätskontrolle bieten.

Aber nicht nur für bereits im Einsatz befindliches Geschäftsmodelle bieten diese Ansätze, sondern auch für die Entwicklung neuer Geschäftsmodelle im Sinne von Unternehmenskooperation. Nach dem Prinzip von Open Business Models müssen solche Wertschöpfungsnetzwerke ihr gemeinsames Geschäftsmodell möglichst offen und adaptiv gestalten. Die Konzeption und Weiterentwicklung von Geschäftsmodellen erfordert entsprechende Werkzeuge für die kollaborative Definition der Geschäftsmodelle (Emrich et al., 2018, S. 1309 ff.), aber auch kontinuierlichen Austausch und Kommunikation zwischen den Akteuren. Mitunter ist auch die Partnerauswahl in einem solchen Netzwerk von Interesse. Hier müssen oft Problemstellungen und Lösungsansätze verknüpft werden, Technologien für die Realisierung bestimmter Business Needs identifiziert werden. Auch hierfür gibt es bereits Ansätze (Emrich et al., 2018, S. 1309 ff.).

Die beschriebenen Technologie- und Managementtrends sind bisher nur in Einzelansätzen in der betrieblichen Realität angekommen. Eine durchgängige Anwendung dieser Ansätze würde eine konsequente Kollaboration zwischen den Akteuren in einer zentralen, digitalen Plattform erfordern. Neben dem kollaborativen Management der Geschäftsmodelle selbst sollen so auch Kundenfeedback- und Produktnutzungsdaten als Feedback in das Management der Geschäftsmodelle einfließen und Akteure können somit individuelle Maßnahmen für sich ableiten. So kann es beispielsweise im Zuge der Geschäftsmodellweiterentwicklung erforderlich sein, neue Kompetenzen im Wertangebot zu ergänzen. Dies kann entweder durch die Hinzunahme neuer Partner in das Netzwerk realisiert werden oder durch die Weiterbildung der bestehenden Partner durch einen entsprechenden Kompetenzaufbau (beispielsweise durch das neue Angebot eines Webshops).

Deshalb müssen **Kollaborationsmöglichkeiten** in allen Phasen des Geschäftsmodellmanagements zur Verfügung gestellt werden, nämlich Analysemodule zur gemeinsamen Untersuchung relevanter Datenphänomene, die Zuordnung zu Akteuren und Geschäftsmodellbereichen, sowie die kollaborative Bearbeitung der zugrunde liegenden Geschäftsmodelle.

3.4 Ausblick

In diesem Kapitel wurde mit der Darstellung der Ausgangssituation und der technischen Grundlagen die Basis für die folgenden konzeptionellen und technischen Entwicklungen gelegt. Während das nächste Kapitel die entwickelte Prozesskette zur Anwendung von industriellen Datenanalysen beschreibt (siehe Kap. 4), werden daraufhin die Ergebnisse der Leistungsbereiche in **Teil II** insb. die Erkenntnisse aus Abschn. 3.3 aufgreifen und weiter ausführen. Bereits in diesem Kapitel haben dazu die Vertreter:innen die entsprechende Perspektive ihres Leistungsbereichs eingenommen und in den Abschnitten ausgeführt. Abschließend, sei für einen tiefergehenden Ausblick auf Kap. 20 verwiesen, das die weiterführende Entwicklung der industriellen Datenanalyse unter Berücksichtigung von Mensch, Technik und Organisation diskutiert.

Literatur

- acatech. (2016). *Kompetenzen für Industrie 4.0. Qualifizierungsbedarfe und Lösungsansätze*. Herbert Utz.
- Amazon Web Services. (2021). *Eine Übersicht über das AWS Cloud Adoption Framework. Whitepaper zu AWS* (3. Aufl.). AWS Inc.
- Arnold, D., Butschek, S., Steffes, S., & Müller, D. (2016). Forschungsbericht 468 – Digitalisierung am Arbeitsplatz.
- Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, *81*, 128–137. <https://doi.org/10.1016/j.compind.2016.02.004>.
- Becker, J., Delfmann, P., Knackstedt, R., & Kuroepka, D. (2002). Konfigurative Referenzmodellierung. In J. Becker & R. Knackstedt (Hrsg.), *Wissensmanagement mit Referenzmodellen. Konzepte für die Anwendungssystem- und Organisationsgestaltung* (S. 25–144). Physica.
- Becker, J., Delfmann, P., & Rieke, T. (2007). Referenzmodellierung. Perspektiven für die effiziente Gestaltung von Softwaresystemen. In J. Becker, P. Delfmann, & T. Rieke (Hrsg.), *Effiziente Softwareentwicklung mit Referenzmodellen* (S. 1–10). Physica.
- Becker, M., Spöttl, G., & Fischer, M. (Hrsg.). (2011). *Von der Arbeitsanalyse zur Diagnose beruflicher Kompetenzen. Methoden und methodologische Beiträge aus der Berufsbildungsforschung*. Peter Lang GmbH Internationaler Verlag der Wissenschaften.
- Birdman, J., Wiek, A., & Lang, D. J. (2022). Developing key competencies in sustainability through project-based learning in graduate sustainability programs. *International Journal of Sustainability in Higher Education*, *23*(5), 1139–1157. <https://doi.org/10.1108/IJSHE-12-2020-0506>.
- BMWi. (2015). *Industrie 4.0 und Digitale Wirtschaft. Impulse für Wachstum, Beschäftigung und Innovation* (S. 7–12). Bundesministerium für Wirtschaft und Energie.
- Buchal, R., & Songsore, E. (2019). Using Microsoft teams to support collaborative knowledge building in the context of sustainability assessment. *Proceedings of the Canadian Engineering Education Association*, *10*(1), 1–8. <https://doi.org/10.24908/pceea.vi0.13882>.
- Burkhart, T., Krumeich, J., Werth, D., & Loos, P. (2011). Analyzing the business model concept. A comprehensive classification of literature. *Proceedings of the International Conference on Information Systems*, 109–130.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-step data mining guide, CRISP-DM consortium.
- Deuse, J., Erohin, O., & Lieber, D. (2014). Wissensentdeckung in vernetzten, industriellen Datenbeständen. *Industrie*, *4*, 373–395.
- Deuse, J., West, N., & Syberg, M. (2022). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering (IJPME)*, *10*(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- DIN SPEC 91345:2016-04, 04.2016: Referenzarchitekturmodell Industrie 4.0 (RAMI 4.0).
- Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, *166*, 114060. <https://doi.org/10.1016/j.eswa.2020.114060>.
- Eigner, M., Roubanov, D., & Zafirov, R. (2014). *Modellbasierte virtuelle Produktentwicklung*. Springer.
- Emrich, A., Klein, S., Frey, M., Fettke, P., & Loos, P. (2018). A platform for data-driven self-consulting to enable business transformation and technology innovation. In P. Drews, B. Funk,

- P. Niemeyer, & L. Xie (Hrsg.), *Data driven x – turning data into value (MKWI)* (S. 1309–1315). Leuphana Universität Lüneburg. <https://doi.org/10.22028/D291-33441>.
- Engelhardt, S. V., Wangler, L., & Wischmann, S. (2017). *Eigenschaften und Erfolgsfaktoren digitaler Plattformen*. VDI/VDE Innovation + Technik GmbH.
- Ernst, J. (2016). *Systemübergreifendes Änderungsmanagement: Graphbasierte Identifikation und Visualisierung betroffener Konfigurationselemente aus PLM und ERP*. Technische Universität Kaiserslautern.
- Faath-Becker, A., & Walker, F. (2020). Development of a video-based test instrument for the assessment of professional competence in the vocational teacher training course. In E. Wuttke, J. Seifried, & H. Niegemann (Hrsg.), *Vocational education and training in the age of digitization: Challenges and opportunities* (S. 109–130). Verlag Barbara Budrich. <https://doi.org/10.2307/j.ctv18dvv1c.9>.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34. <https://doi.org/10.1145/240455.240464>.
- Gilz, T., & Eigner, M. (2013). Ansatz zur integrierten Verwendung von SysML Modellen in PLM zur Beschreibung der funktionalen Produktarchitektur. In M. Maurer & S.-O. Schulze (Hrsg.), *Tag des Systems Engineering* (S. 293–302). Carl Hanser Verlag GmbH & Co. KG.
- Göbel, J. C., & Eickhoff, T. (2020). Konzeption von Digitalen Zwillingen smarterer Produkte. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 115(s1), 74–77. <https://doi.org/10.3139/104.112301>.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer Series in Statistics.
- Helmis, S., & Hollmann, R. (2009). *Webbasierte Datenintegration*. Vieweg+Teubner.
- Ilag, B. N., & Sabale, A. M. (2022). *Troubleshooting Microsoft teams. Enlisting the right approach and tools in teams for mapping and troubleshooting issues*. Apress Media LLC.
- ISO/IEC 7498-1:1994, 15.11.1994: Information technology – Open Systems Interconnection.
- Krumeich, J., Werth, D., Loos, P., Schimmelpennig, J., & Jacobi, S. (2014). Advanced planning and control of manufacturing processes in steel industry through big data analytics: Case study and architecture proposal. *IEEE International Conference on Big Data (Big Data)*, 16–24.
- Leser, U., & Naumann, F. (2007). *Informationsintegration. Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen*. Dpunkt.
- Lin, S.-W., Simmon, E., Young, D., Miller, B., Durand, J., Bleakley, G., et al. (2022). The industrial internet reference architecture. Version 1.10. An industrial IoT consortium foundational document: Industry IoT consortium.
- Marschner, J. (2011). *Adaptives Feedback zur Unterstützung des selbstregulierten Lernens durch Experimentieren*. Dissertation, Universität Duisburg-Essen, Duisburg, Essen.
- Mazarov, J., Wolf, P., Schallow, J., Nöhring, F., Deuse, J., & Richter, R. (2019). Industrial Data Science in Wertschöpfungsnetzwerken. Konzept einer Service-Plattform zur Datenintegration und -analyse, Kompetenzentwicklung und Initiierung neuer Geschäftsmodelle. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 114(12), 874–877. <https://doi.org/10.3139/104.112205>.
- Meinel, C., & Sack, H. (2012). *Internetworking. Technische Grundlagen und Anwendungen*. Springer.
- Mogo Nem, F. (2011). *Engineering networks: Holistic approach for multidisciplinary product lifecycle management*. Technische Universität Kaiserslautern.
- Mufti, T., Mittal, P., & Gupta, B. (2020). A review on Amazon web service (AWS), Microsoft azure & Google cloud platform (GCP) Services. *Proceedings of th International Conference on ICT for Digital, Smart and Sustainable Development*, 2(1), 1–9. <https://doi.org/10.4108/eai.27-2-2020.2303255>.

- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Waxmann.
- Nickolaus, R., & Walker, F. (2016). Kompetenzorientierung in der beruflichen Bildung. In A. Dietzen, R. Nickolaus, B. Rammstedt, & R. Weiß (Hrsg.), *Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen* (S. 7–27). W. Bertelsmann Verlag GmbH & Co. KG.
- Pfeiffer, M., Rossberg, M., Buttgerit, S., & Schaefer, G. (2019). Strong tenant separation in cloud computing platforms. *Proceedings of the International Conference on Availability, Reliability and Security*, 14(1), 1–10. <https://doi.org/10.1145/3339252.3339262>.
- Prasad, U. D., & Madhavi, S. (2012). Prediction of churn behavior of bank customers using data mining tools. *Business Intelligence Journal*, 5(1), 96–101.
- Reidt, A. (2019). Referenzarchitektur eines integrierten Informationssystems zur Unterstützung der Instandhaltung. Dissertation, Technische Universität München, München.
- Rosen, Y., & Salomon, G. (2007). The differential learning achievements of constructivist technology-intensive learning environments as compared with traditional ones: A Meta-analysis. *Journal of Educational Computing Research*, 36(1), 1–14.
- Savarino, P., & Dickopf, T. (2022). Enhancing product test management by using digital twins for device under test analysis. In O. C. Junior, F. Noël, L. Rivest, & A. Bouras (Hrsg.), *Product lifecycle management. Green and blue technologies to support smart and sustainable organizations* (Bd. 640, S. 257–271). Springer International Publishing (IFIP Advances in Information and Communication Technology).
- Schlegl, T., Schlegl, S., Tomaselli, D., West, N., & Deuse, J. (2022). Adaptive similarity search for the retrieval of rare events from large time series databases. *Advanced Engineering Informatics*, 52, 101629. <https://doi.org/10.1016/j.aei.2022.101629>.
- Schwegmann, A. (1999). *Objektorientierte Referenzmodellierung. Theoretische Grundlagen und praktische Anwendung*. Betriebswirtschaftlicher Verlag Dr. Th. Gabler GmbH; Deutscher Universitäts-Verlag GmbH.
- Stefănică, F., Abele, S., Walker, F., & Nickolaus, R. (2017). Modeling, measurement, and development of professional competence in industrial-technical professions. In M. Mulder (Hrsg.), *Competence-based vocational and professional education. Bridging the worlds of work and education* (Bd. 23, S. 843–861). Springer International Publishing.
- Stegmann, K., Wecker, C., Mandl, H., & Fischer, F. (2018). Lehren und Lernen mit digitalen Medien. Ansätze und Befunde der empirischen Bildungsforschung. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Handbuch Bildungsforschung* (4. Aufl., S. 967–988). Springer VS.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- Walker, F. (2016). Digitalisierung in Theorie und Praxis: Kommt die Ausbildung 4.0 in Technikberufen? In BMWi (Hrsg.), *Digitale Bildung. Themenheft Mittelstand-Digital* (S. 14–16).
- Weißmann, M., & Hardwig, T. (2020). *Arbeit mit Kollaborationsplattformen. Gestaltungsempfehlungen*. Soziologisches Forschungsinstitut Göttingen e. V.
- Weskamp, M., Tamas, A., Wochinger, T., & Schatz, A. (2014). *Einsatz und Nutzenpotenziale von Data Mining in Produktionsunternehmen*. Fraunhofer IPA.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining. Practical machine learning tools and techniques* (3. Aufl.), Morgan Kaufmann.

- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021a). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- West, N., Hoffmann, F., Schulte, L., Moreno, V. H., & Deuse, J. (2021b). An approach for deep reinforcement learning for production planning in value streams. In J. Franke & P. Schuderer (Hrsg.), *Simulation in Produktion und Logistik* (S. 347–356). Cuvillier.
- West, N., Schlegl, T., & Deuse, J. (2021c). Feature extraction for time series classification using univariate descriptive statistics and dynamic time warping in a manufacturing environment, *IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering*, 762–768. <https://doi.org/10.1109/ICBAIE52039.2021.9389954>.
- West, N., Schwenken, J., & Deuse, J. (2022a). Comparative study of methods for the real-time detection of dynamic bottlenecks in serial production lines. In H. Fujita, P. Fournier-Viger, M. Ali, & Y. Wang (Hrsg.), *Advances and trends in artificial intelligence. Theory and practices in artificial intelligence* (S. 3–14). Springer. https://doi.org/10.1007/978-3-031-08530-7_1.
- West, N., Syberg, M., & Deuse, J. (2022b). A holistic methodology for successive bottleneck analysis in dynamic value streams of manufacturing companies. *Proceedings of the 8th Changeable, Agile, Reconfigurable and Virtual Production Conference*, 3–14. https://doi.org/10.1007/978-3-030-90700-6_69.
- Wöstmann, R., Barthelme, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. In T. Schüppstuhl, K. Tracht, & J. Roßmann (Hrsg.), *Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Prozesskette der industriellen Datenanalyse

4

Durchführung von industriellen Datenanalyseprojekten entlang einer vierstufigen, tätigkeitsorientierten Prozesskette

Nikolai West , Marius Syberg , Jörn Schwenken , Edin Klasic ,
Ralf Klinkenberg  und Jochen Deuse 

4.1 Motivation und Hintergrund

Im Zuge der Industrie 4.0 und der damit einhergehenden Digitalisierung treten die Potenziale einer intensiven Vernetzung und intelligenten Analyse von Daten in der industriellen Produktion in den Vordergrund (siehe Kap. 3). Ebenso alt wie das Versprechen der zu erwartenden Potenziale ist jedoch auch die Erkenntnis, dass die Anzahl erfolgreich umgesetzter, ganzheitlicher Konzepte der Digitalisierung sowie darauf aufbauender Data Science-Anwendungen in der Produktion deutlich hinter den Erwartungen zurückbleiben (Deuse et al., 2022, S. 2–3).

Das AKKORD-Projekt knüpft an diese Erkenntnis an und stellt ein Konzept zur **Prozesskette der industriellen Datenanalyse** vor. Basierend auf einem Referenzbaukasten und zusammengesetzt aus sechs verschiedenen Bausteinen soll dieses Konzept

N. West (✉) · M. Syberg · J. Schwenken · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: nikolai.west@ips.tu-dortmund.de

M. Syberg
E-Mail: marius.syberg@ips.tu-dortmund.de

J. Schwenken
E-Mail: joern.schwenken@ips.tu-dortmund.de

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

E. Klasic · R. Klinkenberg
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: eklasic@rapidminer.com eklasic@altair.com

R. Klinkenberg
E-Mail: rklinkenberg@rapidminer.com; rklinkenberg@altair.com

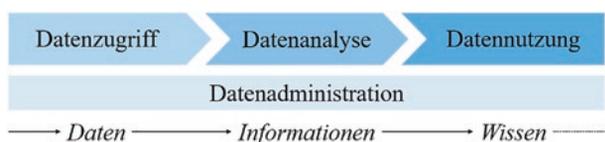
den Umgang mit Data Science vereinfachen, sodass die Erwartungshaltung an die Potenziale von IDS erfüllt werden kann (siehe Kap. 1). Auch die Zugangsschwelle zu IDS kann so reduziert werden, sodass auch unerfahrene Nutzer von diesem Trend profitieren können.

4.2 Prozesskette der industriellen Datenanalyse im Referenzbaukasten

Im Rahmen des Forschungsprojekts AKKORD wurde ein Referenzbaukasten mit Werkzeugen für die Anwendung von industriellen Datenanalysen entwickelt. Der Referenzbaukasten besitzt einen integrierten und datengetriebenen Charakter zur Durchführung von industriellen Datenanalysen. Er kombiniert die Prozesskette der Datenanalyse mit der kollaborativen Work&Learn-Plattform, exemplarischen Erfolgsgeschichten und einem umfangreichen Software- und Dienstleistungsangebot (Mazarov et al., 2019, S. 875). Sein Ziel ist die anwendungsorientierte Bereitstellung von modularen und einheitlichen Lösungsbausteinen, um die industrielle Datenanalyse einem breiteren Nutzerspektrum zugänglich zu machen. Zielgruppe sind vorzugsweise Kleine und Mittlere Unternehmen (KMU), die häufig Hemmnisse in der Durchführung von industriellen Datenanalysen haben. Zu den zu nennenden Hindernissen gehören beispielsweise fehlende Ressourcen und Kompetenzen für die Aufgaben der industriellen Datenanalyse. Der Referenzbaukasten adressiert die Hemmnisse indem er Datenanalysen unternehmensübergreifend wiederverwendbar macht und somit Ressourcen schont. Gleichzeitig reduziert die Wiederverwendbarkeit die Komplexität von Datenanalysen, sodass weniger Kompetenzen bei der Durchführung notwendig sind. An dieser Stelle wird der Anwender ebenso durch die Kollaborations- und Lernplattform unterstützt.

Als Basis des AKKORD-Referenzbaukastens dient die Prozesskette der Datenanalyse (West et al., 2021, S. 131). Diese ist in Abb. 4.1 dargestellt und besteht aus den vier Prozessschritten *Datenzugriff*, *-analyse*, *-nutzung* und *-administration* (engl. *Access, Analyze, Apply, Administrate*). Der Prozessschritt der Datenadministration wird dabei im Referenzbaukasten nicht explizit aufgeführt, sondern ist in den weiteren Bausteinen inbegriffen. Innerhalb der Bausteine der Prozesskette, welche sich den Prozessschritten des *Cross Industry Standard Process for Data Mining*, (CRISP-DM, Chapman et al., 2000, S. 13) zuordnen lassen, werden praxisorientierte Module zusammengefasst (siehe Kap. 3). Zunächst werden im *Zugriffsbaustein* alle notwendigen Datenquellen erschlossen. Dazu gehören Module für die Identifizierung aller relevanten

Abb. 4.1 Darstellung der Prozesskette der industriellen Datenanalyse



Datenquellen für die Analyse, für die Umcodierung fehlender Daten mit geeigneten Erhebungsmethoden und die Bereitstellung der Daten für das Analysesystem (Eiden et al., 2020, S. 82). Dieser Baustein bezieht sich auf die CRISP-DM Phasen des *Business Understanding* und des *Data Understanding*. Nachdem ein durchgängiger Zugang zu den relevanten Daten sichergestellt wurde, beginnt die Zuständigkeit des *Analysebausteins*. Dieser teilt sich dabei in die Vorverarbeitungs- und Analysemodule. In Bezug auf den CRISP-DM stellt der Analysebaustein die Phasen des *Data Understanding*, der *Data Preparation*, des *Modeling* und der *Evaluation* dar. Zuletzt ergibt sich erst durch den *Nutzungsbaustein* der Prozesskette ein wirtschaftlicher Nutzen. Dieser Baustein umfasst die Durchführung von punktuellen Analysen für gezielte Fragestellungen als auch sowohl die Durchführung kontinuierlicher Auswertungen von Langzeitbeobachtungen. Im Einklang mit dem CRISP-DM entspricht der Nutzungsbaustein dem *Deployment* (West et al., 2021, S. 131). Zu den angesprochenen weiteren Bausteinen gehören die *Erfolgsgeschichten*. Sie umfassen beispielhafte Anwendungsfälle im AKKORD-Projekt, welche potenziellen Anwendern einen Einblick in die Potenziale, in den Nutzen und die möglichen Anwendungsfelder von IDS ermöglichen. Mit dem Baustein der *Lern- und Kollaborationsplattform* werden Anwender befähigt, IDS-Projekte durchzuführen. Hierzu stellt die sogenannte Work&Learn-Plattform ein breites Spektrum an Lernpfaden und -modulen zur Verfügung. Gleichzeitig ermöglicht die Work&Learn-Plattform eine unternehmensübergreifende Kollaboration. Zuletzt bietet AKKORD ein *Erweitertes Software- und Dienstleistungsangebot*, welches valide Werkzeuge für die Umsetzung von IDS-Projekten bietet (Schwenken et al., 2023, S. 2 f.).

4.3 Bausteine im Referenzbaukasten

Die Bausteine des AKKORD-Referenzbaukastens bilden die Grundlage, um IDS-Projekte einer breiten Nutzerbasis zugänglich zu machen und die Hürden für deren Durchführung zu senken. Bereits in Kap. 1 und 2 wurde als zentrales Projektziel die Entwicklung eines Referenzbaukastens erklärt. Der in diesem Forschungsprojekt erarbeitete Referenzbaukasten ist in Abb. 4.2 dargestellt. Er besteht aus insgesamt sechs Bausteinen, die im Nachfolgenden ausführlicher erläutert werden.

Zugriffsbaustein Mit diesem Baustein beginnt die Prozesskette der industriellen Datenanalyse. Seine Anforderungen leiten sich aus dem Analysebaustein ab. Dieser gibt den Datenbedarf und die Datenstruktur an den Zugriffsbaustein vor. Im Zugriff werden alle relevanten Datenquellen für die Analyse identifiziert. Hierfür wurden während der Laufzeit des AKKORD-Forschungsprojekts diverse software- und hardwaretechnische Module, die die notwendigen Daten aufnehmen, erstellt. Anschließend codieren Module fehlende Daten mit geeigneten Erhebungsmethoden um und Schnittstellenmodule stellen die Rohdaten für die Analyse bereit. Der Baustein startet den datengetriebenen Analyseprozess und muss eine gute Datenqualität gewährleisten (West et al., 2021, S. 132).

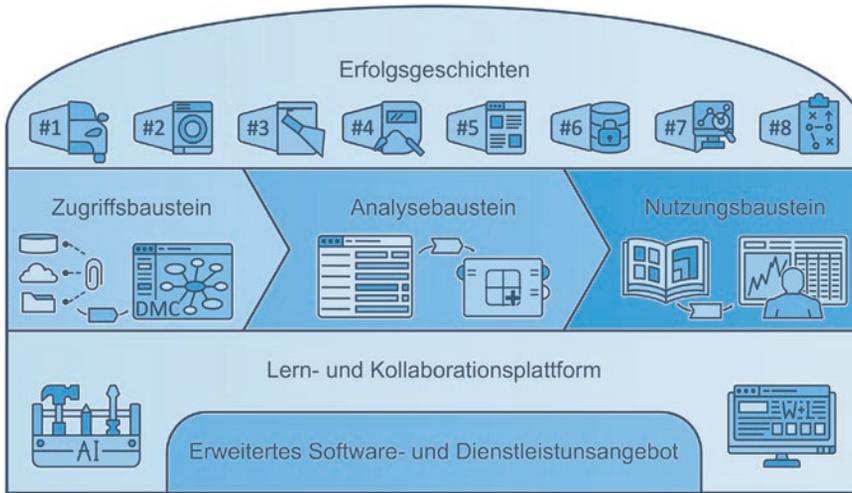


Abb. 4.2 Darstellung des entwickelten AKKORD-Referenzbaukastens, der die Prozesskette der industriellen Datenanalyse abbildet und aus sechs Bausteinen besteht

Qualitätsprobleme, wie fehlerhafte Werte, die während des Zugriffs auftreten, äußern sich später als Fehlentscheidungen im Nutzungsbaustein. Dadurch können wirtschaftliche Schäden entstehen, welche durch einen qualitativ hochwertigen Datenzugriff vermieden werden. Mit den AKKORD-Partnern Arendar IT-Security GmbH, Contact Software, PD Tec und dem Lehrstuhl für Virtuelle Produktentwicklung konnten im AKKORD-Forschungsprojekt vertrauenswürdige Module für den Datenzugriff erstellt werden (siehe Kap. 5). Sie werden auf der Work&Learn-Plattform bereitgestellt.

Analysebaustein Der zweite Schritt der Prozesskette ist der Analysebaustein. In diesem Baustein werden die Rohdaten vorverarbeitet und analysiert, um Informationen und Wissen zu generieren. Hierbei gibt der Nutzungsbaustein vor, welcher Anwendungsfall untersucht werden soll, indem er die Durchführung der Analyse veranlasst. Aus dem Anwendungsfall leitet der Analysebaustein den Datenbedarf und die Datenstruktur ab und gibt sie als Anforderung an den Zugriffsbaustein weiter. Dieser stellt die Rohdaten für die Analyse bereit. In diesem Schritt werden die bereitgestellten Rohdaten entsprechend vorverarbeitet und durch verschiedene Werkzeuge und Methoden analysiert. Sowohl jede Datenvorverarbeitung als auch jede Analyse entspricht hierbei einem wiederverwendbaren Analysemodul. Zu den potenziellen Werkzeugen und Methoden des Analysebausteins gehören z. B. zahlreiche Open-Source Pakete und Data-Science-Algorithmen. Weitere Schritte zur Informationsverarbeitung und -umwandlung können den Analyseschritt begleiten. Die Ergebnisse werden vom Analysebaustein dem Nutzungsbaustein bereitgestellt. Während der Laufzeit des AKKORD-Forschungsprojekts konnten die Partner RapidMiner und das Institut für Produktionssysteme

zahlreiche Module für die Datenvorverarbeitung und -analyse entwickeln (siehe Kap. 6). Sie werden generalisiert auf der Work&Learn-Plattform bereitgestellt.

Nutzungsbaustein Damit die gewonnenen Informationen aus dem Analysebaustein eingesetzt werden können, ist der Nutzungsbaustein notwendig. Er ist der finale Baustein der Prozesskette und stellt somit die Schnittstelle zum Nutzer der industriellen Datenanalyse dar. Die Aufgabe des Bausteins ist auf der einen Seite die Formulierung der Bedürfnisse des Nutzers und die daraus resultierende Veranlassung zur Durchführung der Datenanalyse an den Analysebaustein. Auf der anderen Seite ist es die Aufgabe des Nutzungsbausteins dem Nutzer die Informationen gebündelt und übersichtlich zu präsentieren, sodass dieser die Informationen für den Anwendungsfall interpretieren kann. Erst durch die Nutzung der aus der Analyse gewonnenen Informationen kann ein monetärer Nutzen erzielt werden (West et al., 2021, S. 131). Eine breite Menge an Modulen zur Nutzung von Datenanalysen haben die AKKORD-Partner RapidMiner und IPS, gemeinsam mit den Anwendungspartnern (siehe **Teil III**), bereitgestellt.

Erfolgsgeschichten Die Erfolgsgeschichten ermöglichen dem Nutzer einen Einblick in die Applikationsmöglichkeiten von Datenanalysen in der industriellen Praxis. Im Kontext des AKKORD-Forschungsprojekts bestehen verschiedene Erfolgsgeschichten, welche durch die verschiedenen Projektpartner ERCO GmbH, Miele & Cie. KG, Polymerge GmbH und Volkswagen AG in Form von konkreten Anwendungsfällen durchgeführt wurden. Diese Anwendungsfälle bestehen aus den folgenden verschiedenen IDS-relevanten Themenbereichen: Industrial Engineering, Qualitätsmanagement, integrierte Datenanalyse, Potenzialanalyse und Datenakquisition sowie Geschäftsmodellentwicklung. Innerhalb der Erfolgsgeschichten werden die Prozessschritte der Prozesskette der industriellen Datenanalyse exemplarisch durchlaufen. Die Erfolgsgeschichten können im Detail auf der Work&Learn-Plattform sowie auf der Projekt-Webseite (www.akkord-projekt.de/use-cases) nachvollzogen werden. Anwender erhalten durch die Ergebnisse der Use Cases die Möglichkeit, einen Einblick in die Potenziale industrieller Datenanalyse.

Lern- und Kollaborationsplattform Im Referenzbaukasten wird die Lern- und Kollaborationsplattform durch die eigenentwickelte Work&Learn-Plattform abgebildet. Die Plattform bündelt Wissen und Ressourcen wie Artikel, Kurse und Anwendungen rund um das Thema IDS. So schafft sie eine Umgebung, in der ihre Nutzer ihre Erfahrungen und Kenntnisse im Bereich von IDS entwickeln und verfeinern können. Anwender bekommen abhängig von ihrer Rolle in IDS-Projekten und ihrem IDS-Kennntnisstand Inhalte vorgeschlagen. Ebenso bildet sich in dieser Abhängigkeit ein Lernpfad für den individuellen Kompetenzaufbau. Technisch basiert die Work&Learn-Plattform auf der Plattform PIPE (Syberg et al., 2023, S. 4). Die fünf Bereiche der Work&Learn-Plattform, Kanal-, Service-, Lern-, Kollaborations- und Newsbereich,

wurden maßgeblich vom Lehrstuhl Berufspädagogik in den technischen Fächern und der Neocosmo GmbH umgesetzt (Schwenken et al., 2023, S. 6).

Software und Dienstleistungsangebot Den letzten Baustein des AKKORD-Referenzbaukastens bildet das ergänzende Software- und Dienstleistungsangebot. Über die Work&Learn-Plattform werden den Anwendern leicht zugängliche Beratungsangebote für die begleitende Durchführung von IDS-Projekten bereitgestellt. Des Weiteren wird über Open Source-Angebote hinaus auch ein erweitertes Angebot an Software- und Hardwarelösungen für die technische Umsetzung vorgeschlagen. Sowohl das Beratungsangebot als auch die Softwarelösungen unterstützen Anwender entlang der gesamten Prozesskette der industriellen Datenanalyse.

4.4 Ausblick

Abschließend lässt sich resümieren, dass der AKKORD-Referenzbaukasten mit seinen sechs Bausteinen die verschiedenen Bedürfnisse und Probleme von Nutzern im IDS-Umfeld adressiert. Er vereinfacht durch seinen modularen Charakter in verschiedenen Anwendungsfällen die Durchführung von IDS-Projekten und erhöht dadurch die Wiederverwendbarkeit von Datenanalysen. Die Nutzung der Bausteine erfolgt in der Praxis mithilfe der webbasierten **AI-Toolbox** (siehe Kap. 6).

Durch die **Work&Learn-Plattform** werden zudem die Kompetenzen von Mitarbeitern in den Rollen der Prozesskette der industriellen Datenanalyse verbessert. Gleichzeitig schafft die Plattform eine Möglichkeit, über Unternehmensgrenzen hinweg zu kollaborieren (siehe Kap. 7 und 13) In den Erfolgsgeschichten wird beispielhaft gezeigt, wie die Durchführung der Prozesskette aussehen kann. Mit dem Erweiterten Software- und Dienstleistungsangebot stehen die Konsortialpartner den Nutzern zur Verfügung, um individuell und zielgerichtet bei der Durchführung von Datenanalysen zu unterstützen. Ein weiterführender Ausblick für dieses Vorhaben findet sich in Kap. 20, in dem die weiterführende Entwicklung industrieller Datenanalysen im Einklang mit Mensch, Technik und Organisation diskutiert wird.

Literatur

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-step data mining guide, CRISP-DM consortium.
- Deuse, J., West, N., & Syberg, M. (2022). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- Eiden, A., Gries, J., Eickhoff, T., & Göbel, J. C. (2020). Anforderungen an ein Daten-Backend-System zur Unterstützung industrieller Datenanalyse-Anwendungen in digitalen Engineering-

- Prozessen dynamischer Wertschöpfungsnetzwerke. *Proceedings of the 31st Symposium Design for X*, 81–90.
- Mazarov, J., Wolf, P., Schallow, J., Nöhring, F., Deuse, J., & Richter, R. (2019). Industrial Data Science in Wertschöpfungsnetzwerken. Konzept einer Service-Plattform zur Datenintegration und -analyse, Kompetenzentwicklung und Initiierung neuer Geschäftsmodelle. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 114(12), 874–877. <https://doi.org/10.3139/104.112205>.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Teil II

Leistungsbereiche im Referenzbaukasten



Konzeption und Entwicklung eines integrierten Datenbackends für die industrielle Datenanalyse

Vorstellung entwickelter Konzepte, Lösungsbausteine sowie einer Entscheidungshilfe zur industriellen Implementierung

Thomas Eickhoff , Jens C. Göbel , Christo Apostolov 
und Hardy Krappe

5.1 Anforderungen an das Datenbackend

In der Softwareentwicklung bezeichnet ein (Daten-)Backend die Teile der Gesamtlösung, die sich mit der Speicherung und Verwaltung von Daten befassen. Das Backend, welches sich weiter vom Benutzer entfernt befindet als das Frontend, stellt Eingabemasken und andere Benutzeroberflächen und Interaktionsmöglichkeiten bereit. Grundlegende Ansätze zum Aufbau eines Backends, wie zum Beispiel Synchronisation oder Integration bestehender Datenquellen im Kontext der Produktentwicklung werden von Bergsjö et al. (2006, S. 1065 ff.) verglichen.

Die Entwicklung eines Datenbackends für die industrielle Datenanalyse setzt ein tiefgehendes Verständnis des adressierten Anwendungsszenarios voraus. Um im Kontext des integrierten Referenzbaukastens auf verschiedene grundsätzliche Ausprägungen eines möglichen Datenbackends eingehen zu können, wurden im Rahmen des Forschungsprojekts AKKORD (siehe Kap. 1) zunächst allgemeine funktionale und nicht-funktionale

T. Eickhoff (✉) · J. C. Göbel
Technische Universität Kaiserslautern, Kaiserslautern, Deutschland
E-Mail: eickhoff@mv.uni-kl.de

J. C. Göbel
E-Mail: goebel@mv.uni-kl.de

C. Apostolov
CONTACT Software GmbH, Bremen, Deutschland
E-Mail: christo.apostolov@contact-software.com

H. Krappe
PD Tec AG, Karlsruhe, Deutschland
E-Mail: krappe@pdtec.de

Anforderungen, an die im Projekt zu entwickelnde Backend-Lösung formuliert. Diese aufgenommenen Anforderungen lassen sich anhand einiger repräsentativer Beispiele illustrieren: Durch funktionale Anforderungen wurde der prinzipielle Funktionsumfang des Backends eingegrenzt. Das Datenbackend muss erforderliche Schnittstellen bereitstellen, um verschiedene Quellsysteme anzubinden. Das Backend soll darüber hinaus in der Lage sein, diese Daten so aufzubereiten, dass sie für eine Analyse verwendet werden können. Neben der Kernfunktionalität wurden auch weitere Aspekte, wie zum Beispiel ein flexibler Einsatz der entwickelten Lösung sowohl in der Cloud als auch auf eigener Hardware vorgesehen. Nichtfunktionale Anforderungen, beispielsweise zur verarbeiteten Datenmenge, wurden ebenfalls berücksichtigt. Um das im Projekt gewonnene Wissen sichern zu können, wurde durch weitere definierte Anforderungen die Erstellung diesbezüglicher Leitfäden berücksichtigt. Insgesamt bezogen sich 29 der aufgenommenen Anforderungen auf das Datenbackend. Alle Anforderungen wurden mit passenden Prüfkriterien verknüpft, die die Basis für die Validierung am Projektende liefern sollten.

5.2 Allgemeines Vorgehen zum Aufbau des Datenbackends

Basierend auf den Anforderungen wurde ein Vorgehensmodell für die Entwicklung des Datenbackends definiert. Der prinzipielle Aufbau orientiert sich an den Phasen des Cross- Industry Standard Process for Data Mining (CRISP-DM) nach (Chapman et al., 2000), der ein allgemeines Vorgehen für Data Mining-Projekte definiert und bereits in Kap. 3 vorgestellt wurde (West et al., 2021, S. 133). In der ersten Phase, die das *Geschäftsverständnis* umfasst, wird fachlich definiert, was mit den geplanten Analysen erreicht werden soll. Ein solches Verständnis ergab sich im Projektkontext aus der Bearbeitung der ersten Arbeitspakete, die sich mit der Anforderungsaufnahme und Konzeption des Referenzbaukastens befassten. Das hier beschriebene Vorgehen setzt daher ein fachliches Verständnis als gegeben voraus, welches jedoch bei der späteren Beschreibung der fachlichen Ebene noch nachgeschärft werden kann.

Der Aufbau des Backends, von der Konzeption bis hin zur technischen Umsetzung, wird in diesem Kontext als Teil der Datenverständnis-Phase aufgefasst, da es hier um die Verfügbarmachung der Daten für die Datenvorverarbeitungs-Phase geht. In der Prozesskette der Industriellen Datenanalyse behandelt das im weiteren vorgestellte Datenbackend also den Schritt des Datenzugriffs (siehe Kap. 4). Da die Vorverarbeitung der Daten im Projektkontext teilweise in den in Kap. 6 beschriebenen Analysebausteinen erfolgt, endet der Aufbau des Datenbackends zwischen der Datenverständnis- und Datenvorverarbeitungs-Phase. In der exemplarischen Umsetzung im Projektkontext werden unterschiedliche Abgrenzungen zwischen Backend-Aufbau und Datenvorverarbeitung aufgezeigt. Dies ist schematisch in der nun folgenden Abb. 5.1 dargestellt.

Innerhalb der Datenverständnis-Phase ergeben sich die folgenden vier Teilschritte: Zunächst wird auf der fachlichen Ebene das Geschäftsverständnis auf die im Unternehmen verfügbaren Datenquellen abgebildet. Zu diesem Zweck wird der sich aus dem

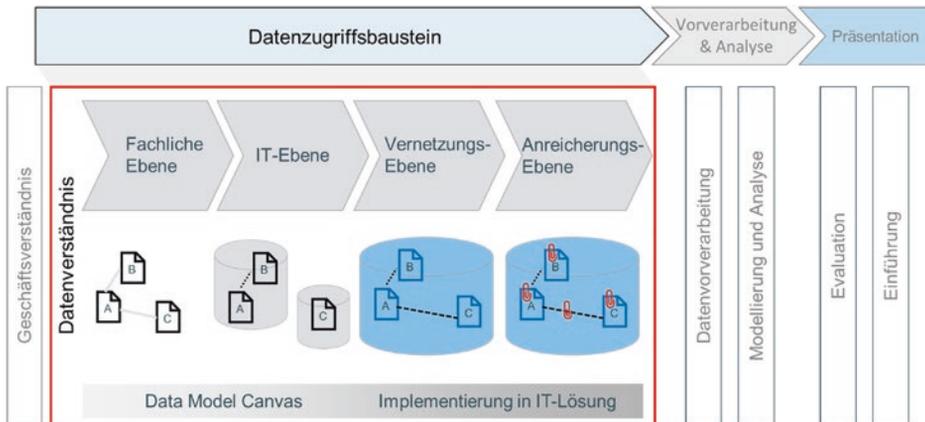


Abb. 5.1 Vorgehen beim Aufbau des Datenbackends

Geschäftsverständnis ergebende Datenbedarf in Form von benötigten Datenobjekten aufgebaut. Dieser wird dann den tatsächlich in den IT-Systemen des Unternehmens vorhandenen Datenobjekten gegenübergestellt. Hierbei besteht die Möglichkeit, in der Vernetzungsebene Querbeziehungen zwischen den einzelnen angebotenen Datenquellen zu ergänzen und diese in der Anreicherungsstufe mit weiteren Informationen, wie beispielsweise externen Links oder Metainformationen, anzureichern.

Abschließend können die zusammengeführten Daten entsprechend dem Datenbedarf an die für die Datenvorverarbeitung bzw. Analyse zuständigen Bausteine weitergegeben werden.

5.3 Lösungsbausteine

Das im vorherigen Abschnitt abstrakt beschriebene Vorgehen lässt grundsätzlich noch ein breites Spektrum an technischen Umsetzungen zu. Im Kontext des Forschungsprojekts wurden einige Lösungsbausteine entwickelt, die auf verschiedenen Umsetzungsstrategien basieren, die von einer zentralisierten Datenintegration über Datenduplizierung bis hin zu einer metadatenbasierten dezentralen Datenhaltung reichen. Die entwickelten Lösungsbausteine in Datenzugriffsbausteine, Backendbausteine sowie Schnittstellenbausteine untergliedern. Die Wahl einer passenden Umsetzung hängt vom Anwendungsszenario ab, jede Lösung bringt eigene Vor- und Nachteile mit sich (Bergsjö et al., 2006, S. 1065 ff.).

Die Datenzugriffsbausteine ermöglichen das Anbinden von Datenquellen. Dies kann sehr pragmatisch erfolgen. Beispielsweise wurde im Kontext des Autorenn-Demonstrators (siehe Kap. 17) ein Baustein entwickelt, der die Übertragung von Daten der Bahn (Rundenzeiten und simulierte Tankfüllstände) und weiteren Metainformationen (Fahrername, Fahrzeug, Alter) ermöglicht. Dieser Zugriffsbaustein liegt in der Form

eines einfachen Python-Skripts vor, das auf einem RaspberryPi ausgeführt wird, das über Bluetooth auf die Rennbahndaten zugreift. Ein etwas komplexeres Beispiel für Datenzugriffsbausteine sind die in Abschn. 5.3.2 beschriebenen SP²IDER-Konnektoren.

Die Schnittstellenbausteine ermöglichen das Übertragen von Daten zwischen den Datenzugriffsbausteinen und dem Backend, aber auch die Weitergabe von Daten aus dem Backend an die Analysebausteine. Exemplarisch wird nachfolgend der für das Projekt zentrale Lösungsbaustein „Data Model Canvas“ vorgestellt.

5.3.1 Dezentrale Datenvernetzung mit SP²IDER

Der dezentrale Ansatz zur Datenvernetzung basiert auf der Idee, die im Unternehmen verfügbaren Daten in ihren ursprünglichen Quellsystemen gespeichert zu lassen, und über ein gemeinsames Metadatenetz darauf zu verlinken. Zum Aufbau des Netzes und zum Abruf der Daten wird jedes Quellsystem über einen minimalistischen Konnektor mit einem zentralen Repository verbunden.

Das zentrale Repository verwaltet im Gegensatz zu monolithischen Datenintegrationsansätzen bewusst keine vollständige Kopie der angebotenen Daten. Stattdessen wird nur die Information „Datensatz A existiert in Quellsystem B“ und „Datensatz A steht in Beziehung zu Datensatz C“ verwaltet. Aus der Vielzahl so beschriebener Beziehungen entsteht ein Graph, der mit zusätzlichen Informationen angereichert werden kann, um beispielsweise systemübergreifende Querbeziehungen zwischen Datensätzen zu modellieren. Wird ein kompletter Datensatz benötigt, kann dieser im ursprünglichen System abgefragt werden.

Die technologische Basis für die dezentrale Datenintegration im Forschungsprojekt AKKORD bildet das am Lehrstuhl für Virtuelle Produktentwicklung der RPTU entwickelte SP²IDER-System. Die Architektur der SP²IDER-Plattform ist in Abb. 5.2 dargestellt. Das System wurde bereits in verschiedenen Publikationen beschrieben (Eiden et al., 2021, S. 690 ff.).

Um Daten aus verschiedenen Quellsystemen in der SP²IDER-Plattform verfügbar zu machen, werden bestehende IT-Systeme über eine einheitliche Schnittstelle mit der Plattform verbunden. Auch diese Konnektoren wurden im Kontext des Projekts in Python implementiert, können aber mithilfe beliebiger Programmiersprachen umgesetzt werden, da für die Kommunikation mit der Plattform ausschließlich sprachunabhängige REST-Schnittstellen zum Einsatz kommen. Um den Benutzer durch die in Abschn. 5.2 beschriebenen Phasen zu leiten, stellt die SP²IDER-Plattform mit dem Data Model Canvas eine intuitive Benutzeroberfläche zur Verfügung (siehe Abb. 5.3). Auf der fachlichen Ebene unterstützt der Data Model Canvas die Benutzer methodisch und informationstechnisch beim Aufbau einer fachlichen Datengrundlage zur Datenanalyse. Datenmodelle aus verschiedenen Quellen können analysiert und in ein gemeinsames Modell für die weitere Analyse integriert werden. Der Aufbau dieses fachlich orientierten Datenmodells soll hierbei vorrangig durch Domänen-Experten erfolgen.

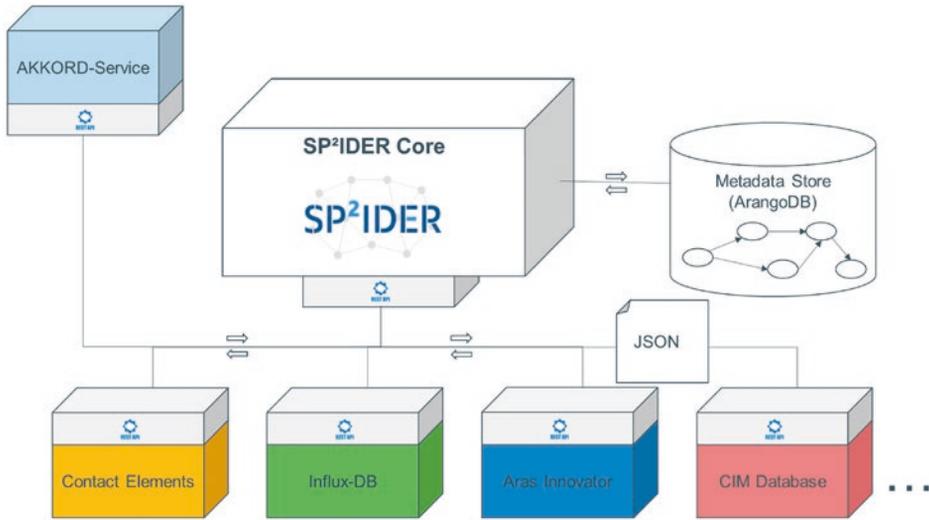


Abb. 5.2 Architektur der SP²IDER-Plattform

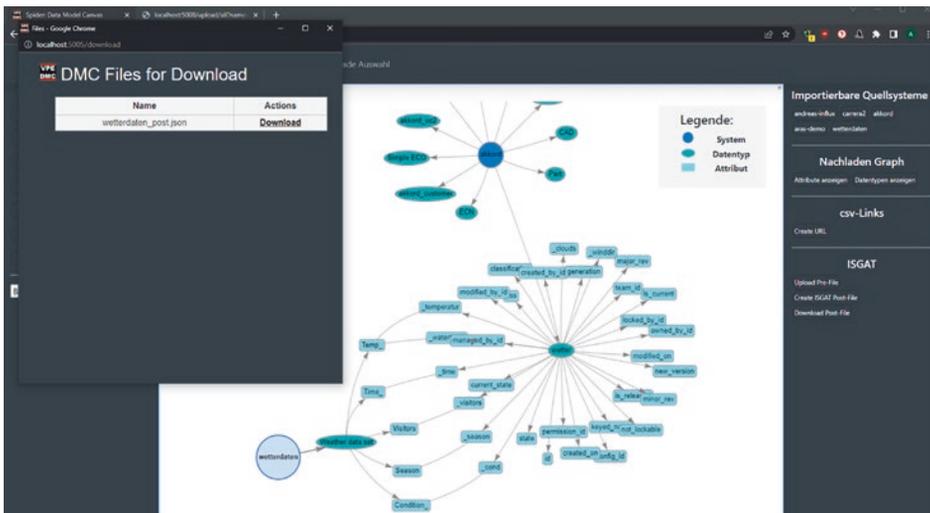


Abb. 5.3 Benutzeroberfläche des Data Model Canvas

Die prototypische Implementierung des Data Model Canvas wird als Web-App bereitgestellt und liefert die beschriebenen Möglichkeiten, Datenmodelle in verschiedenen Formaten sowie bereits bestehende Vorüberlegungen zu importieren. Die Daten werden in einem Graph mit Datentypen als Knoten, die durch als Kanten dargestellte Relation miteinander in Beziehung stehen, dargestellt.

Nach dem Import des ersten Datenmodells wird dieses auf die relevanten Datentypen reduziert. Der Aufbau des gemeinsamen Datenmodells folgt anschließend einem iterativen Prozess: Jedes weitere importierte Datenmodell wird ebenfalls auf die relevanten Datenobjekte reduziert. Querbeziehungen zwischen den Datenmodellen werden durch das Einfügen neuer Kanten abgebildet. Datentypen, die denselben Sachverhalt in unterschiedlichen Systemen abbilden, können zu einem gemeinsamen Knoten zusammengeführt werden.

Die zusammengeführten Informationen liefern ein initiales fachliches Verständnis über die vorhandenen Daten, und bilden somit die fachliche Ebene des Datenzugriffsprozesses ab. Gleichzeitig liefert das entwickelte Modell aber auch eine technische Abbildung der Verfügbarkeit dieser Daten in den angebenen Quellsystemen. Der Data Model Canvas ermöglicht somit die Verfeinerung des entwickelten Datenmodells hin zur technischen Ebene. Sind die benötigten Daten sowie ihre Beziehungen und ihr Speicherort bekannt, kann die eigentliche Vernetzung stattfinden. In diesem Schritt werden die im DMC entwickelten Modelle auf die reale Infrastruktur übertragen, damit die Daten in gebündelter Form für die Analyse zur Verfügung stehen. Dafür kommen verschiedene Lösungen infrage, wobei der DMC durch seinen Aufbau eine große Flexibilität in der technischen Umsetzung erlaubt. Eine Einschränkung auf konkret zu verwendende Softwaresysteme oder eine feste Struktur des endgültigen Datenmodells wird vermieden. Im Projektkontext ergibt sich die konkret gewählte Umsetzung aus den Anforderungen der betroffenen Use Cases.

In den Anwendungsfällen des Projekts kam der Data Model Canvas für die Gegenüberstellung eines Datenbedarfes gegen die in den mit SP²IDER verbundenen Quellsystemen verfügbaren Daten zum Einsatz. Der Datenbedarf wird in diesem Fall in Form einer JSON-Datei bereitgestellt. Benötigte Datenobjekte können mit verfügbaren Datenobjekten verbunden werden. Das daraus resultierende Datenmapping wird wiederum als JSON-Datei bereitgestellt, die u. a. URLs zum Abruf der ausgewählten Daten beinhaltet. Diese Datei wird anschließend an die in Rapidminer Studio implementierten Analysebausteine weitergegeben.

5.3.2 Aufbau eines zentralen Datenmodells mit der Contact Elements-Plattform

In diesem Ansatz werden alle Daten in einer zentralen Plattform gespeichert. Das Datenmodell kann hierbei unterschiedlichste Datenquellen umfassen und detaillierte Verknüpfungen abbilden. Die eingesetzte Datenbank kann bereits Werkzeuge zur weiteren Vernetzung, Bearbeitung oder Auswertung der Daten mitliefern. Im Kontext von Produktdaten oder Daten aus dem „Internet of Things“ (IoT) ist es sinnvoll, eine angemessene Plattform zu wählen, wie beispielsweise eine PLM-Softwarelösung für die Verwaltung von Produktlebenszyklusdaten.

Im Forschungsprojekt wurde ein zentrales Datenmodell in der Elements-Plattform des Herstellers Contact Software erstellt. Die Architektur der Lösung ist in Abb. 5.4 dargestellt. Im Zentrum stand hierbei der Aufbau eines durchgängigen Datenmodells, um eine möglichst breite Informationsbasis für Datenanalysen bereitstellen zu können.

Zur Anbindung von Datenquellen und zur Weitergabe der Daten an die Analysebausteine wird auch hier auf den im vorherigen Ansatz beschriebenen Data Model Canvas zurückgegriffen. Ein Einsatz der bestehenden Möglichkeiten zum Datenimport oder -export in der Contact Elements Plattform ist ebenfalls möglich.

Der Einsatz einer konsolidierten Backend-Lösung bietet vielfältige Möglichkeiten, Daten direkt miteinander zu verknüpfen und über den kompletten Produktlebenszyklus hinweg an einer Stelle zu verwalten. Konkret können beispielsweise Daten eines über die Zugriffsbausteine angebotenen IoT-Assets mit PLM-Stammdaten vernetzt werden. Das konsolidierte Backend liefert auch Möglichkeiten, erste Vorverarbeitungsschritte oder grundlegende Analysen direkt auszuführen und wichtige Performance-Indikatoren in Dashboards zusammenzufassen. Die für verschiedene Aufgabenbereiche relevanten Daten sind in Abb. 5.5 dargestellt.

Verschiedene Datenobjekte aus Modulen repräsentieren Daten aus heterogenen Quellen des Produktlebenszyklus in der Contact Elements Plattform: Das Modul Produktdatenmanagement besteht aus Objekten wie dem Produkt selbst, das auf generischer Ebene verschiedene Teile enthält. Das generische Produkt muss Spezifikationen erfüllen, die Teil von Anforderungen sind, die in Testläufen gemäß Akzeptanzkriterien getestet werden und Testergebnisse erzeugen. Jedes getestete Teil wird durch die Zuweisung einer Serie instanziiert, auf die in einem Testauftrag als Teil des Testausführungsmanagementmoduls Bezug genommen wird. Arbeits- und Prüfpläne bestehen aus einer oder mehreren Sequenzen (bzw. Prüfprozessen) und einer oder mehreren Operationen (bzw. Prüfschritten) und sind Teil des Moduls Arbeitsplanung bzw. Prüfplanung. Ein Arbeitsplatz stellt das Organisations- und Ressourcenmanagementmodul dar, das Teil einer Organisation oder eines Werkes ist, dem Personen, z. B. arbeitendes Personal, zugeordnet sind. Echtzeitdaten werden durch Objekte wie Assets repräsentiert, die den digitalen Zwilling einer Produktinstanz beschreiben. Der Produktionsauftrag schließlich beschreibt die Reihenfolge, in der Vorgänge ausgeführt werden. Als Teil verschiedener Module kommunizieren diese Objekte mit den Modulen zur Datenanalyse und der Datenmodellanalyse und -vorlage.

Wie oben beschrieben, bietet die CONTACT Elements Plattform ein umfangreiches Datenmodell, das es erlaubt, heterogene Daten aus verschiedenen Legacy-Quellen in eine State-of-the-Art-Plattform einzubringen, zu verknüpfen und diese Daten semantisch anzureichern. Diese Ansätze können mit den Bausteinen der anderen Ansätze kombiniert werden, um eine ganzheitliche Backendlösung zu realisieren. Das detaillierte Konzept wurde im Rahmen einer Veröffentlichung im Rahmen der DESIGN2022 Konferenz vorgestellt (Eiden et al., 2022, S. 693 ff.). Hier ist auch beschrieben, wie ein zentrales Backend mit dem SP²IDER-System aus der in Abschn. 5.3.1 beschriebenen dezentralen Lösung kombiniert werden kann, um gezielt Daten für eine Analyse zusammenzuführen und an die im Projekt entwickelten Analysebausteine weitergegeben werden.

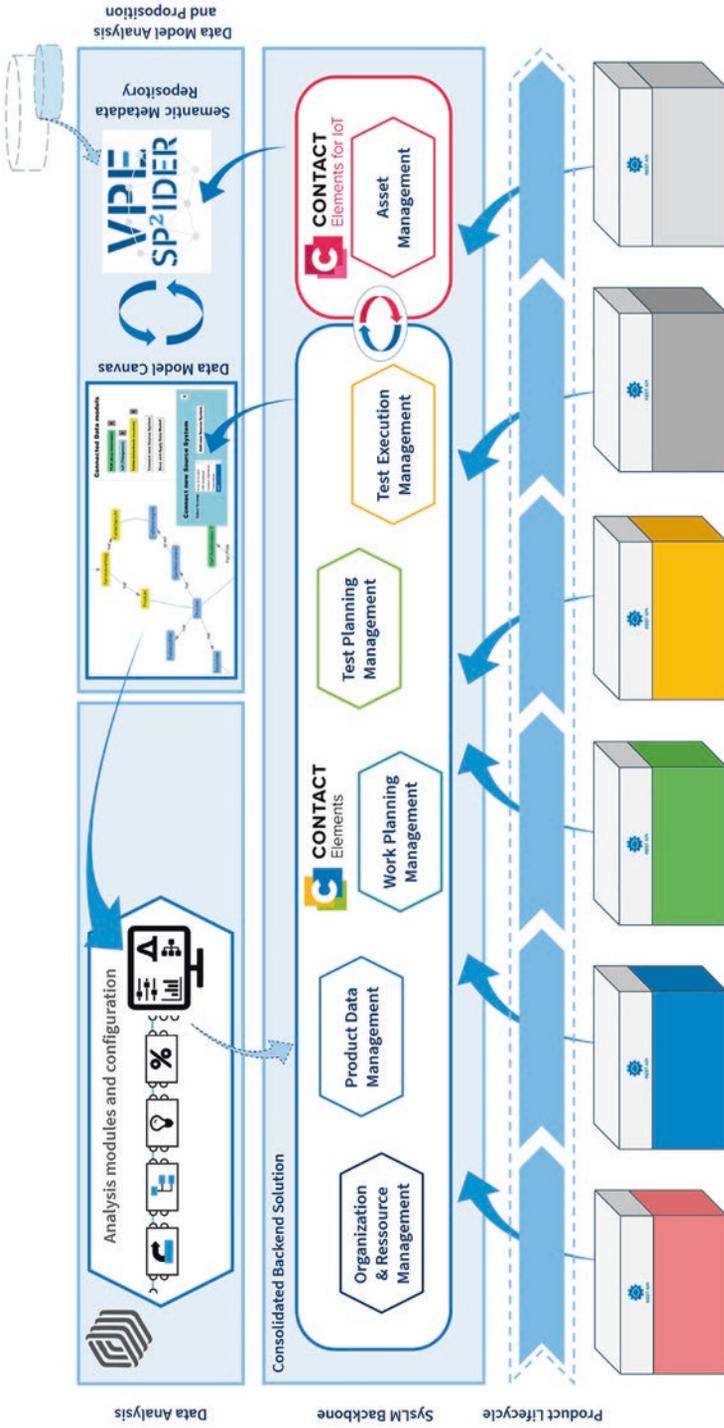


Abb. 5.4 Architektur und Module des konsolidierten Datenbackends

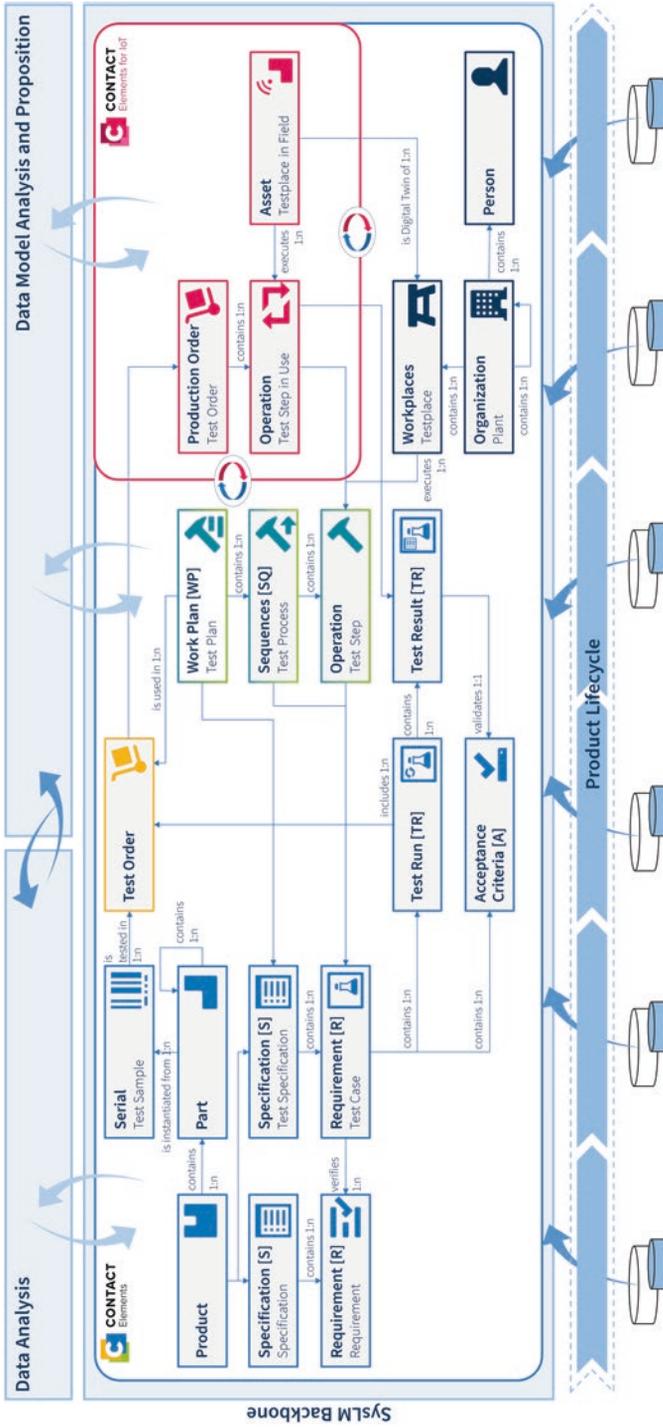


Abb. 5.5 Datenmodell des konsolidierten Datenbackends

5.3.3 Data Warehouse-Ansatz mit Power-BI

Ein Data Warehouse ist eine für Analysezwecke optimierte zentrale Datenbank, die Daten aus mehreren, in der Regel heterogenen Quellen zusammenführt (Devlin & Cote, 1996). Die Architektur eines Data Warehouse und von Microsoft Power BI umfasst in der Regel die Verwendung eines ETL-Tools (Extrahieren, Transformieren, Laden), um Daten z. B. aus SaaS-Anwendungen, aber auch aus strukturierten Dateien (Excel) oder unstrukturierten Formaten, sowie Datenbanken zu extrahieren und in diese über ein Staging/Publishing Mechanismus in das Data Warehouse zu laden. Dabei sind insbesondere das Verschieben von Business-Transformationslogik aus Power BI in die Datenbank und das Speichern von Dimensionsdaten im Data Warehouse für Berichts- und Analysezwecke zu beachten. Je nach Situation kann entweder ein Data Warehouse oder Power BI Dataflow die beste Option für eine bestimmte Lösung sein. Im Kontext AKKORD lag der Fokus auf der Implementierung des Data Warehouse. PowerBI wurde im Rahmen der Plausibilisierung und Validierung der Daten im Wesentlichen für die Visualisierung und das Reporting verwendet. Zu den verwendeten Verfahren für die Kopplung eines Data Warehouse mit Power BI gehörten neben der Verwendung eines Data Warehouse, das alle Daten an einem einzigen Ort für Analysen konsolidiert, auch die Erstellung eines dimensional Modells mithilfe von Datenflüssen, sowie die Möglichkeit zur Berichterstellung und die Implementierung von Self-Service-BI.

Im Rahmen von AKKORD zeigt das Abb. 5.6 die gewählte Architektur, die im Rahmen des Referenzbaukastens als Lösungsschablone dienen sollte. Ausgehend von verschiedenartigen Quellsystemen wird mittels verschiedener Mechanismen und Methoden, wie das Mapping von Modellinformationen (Klassen, Attribute, Relationen), Datenschemas und Datenstrukturen und deren Konsolidierung die Daten in ein Data Warehouse, bestehend aus Staging und Publishing System, hinterlegt und aufbereitet. Diese stehen im Anschluss zur Analyse mittels z. B. Visualisierung, Machine-Learning-Algorithmen oder anderen Verfahren u. a. für PowerBI oder anderen 3rd Party Solutions den Domänenexperten im Unternehmen zur Verfügung.

Für die Instanziierung der Architektur aus Abb. 5.6 wurde ein agiler Ansatz anstelle eines Big-Bang-Ansatzes gewählt. Es wurden sukzessive verschiedene Quellsysteme, je nach Verfügbarkeit, agil aufbereitet und in das Data Warehouse extrahiert. Ein großer Vorteil dieser Methode liegt darin, dass auf diese Weise sich die Erstellung abtrünniger „unabhängiger“ Data Marts vermeiden lässt und das Geschäftsmodell und die Architektur nur bei Bedarf und sobald die Data Marts einen echten Mehrwert liefern, berücksichtigt werden. Außerdem ist es wichtig, in der Testphase Sortiertests sowie Hierarchisierungstest durchzuführen und später optimierte Abläufe einzuführen. Dazu wurden entsprechende Läufe und Tests über eine speziell geschaffene Infrastruktur entwickelt und durchgeführt. Schließlich sind die Einhaltung von Best Practices der Branche, wie z. B. das Verständnis, warum man ein Data Warehouse braucht, und die Analyse der Quellsysteme, wichtige Konzepte, die bei der Implementierung berücksichtigt wurden.

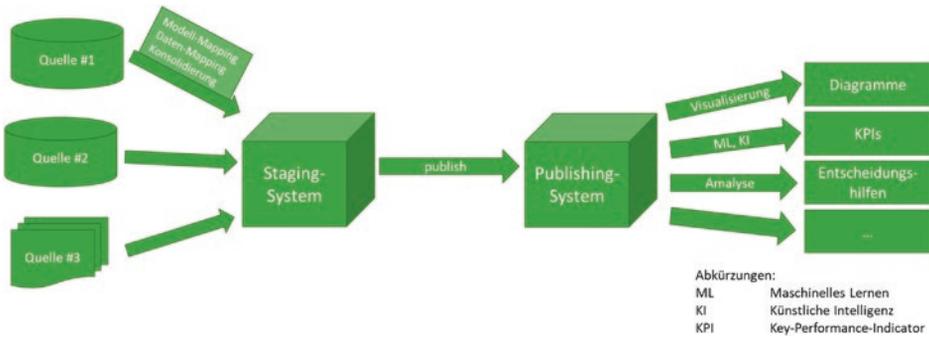


Abb. 5.6 Architektur des Data Warehouses im Referenzbaukasten von AKKORD

In Abb. 5.7 wird die Instanziierung der Architektur dargestellt. Dabei wird auf der linken Seite den Quellsystemen die zugehörige Instanz, sowie die Import- und Mappingmethode gezeigt. In der PD Tec eigenen ice.NET Plattform wird die Staging Datenbank gehostet. Ice.NET stellt mittels speziell im Projektkontext entwickelten Tools automatisierte Transformationsalgorithmen zur Verfügung, die im Wesentlichen eine Normalisierung der DB vornehmen, wobei der Fokus insbesondere auf Robustheit und Performance, z. B. bei verschachtelten Hierarchien, lagen. Im ice.NET Data Warehouse liegen letztendlich die normalisierten Daten bereit, um in verschiedenen 3rd Party Tools Anwendung zu finden. Beispielhaft wurde hierzu PowerBI, SQL View und SQL Queries herangezogen.

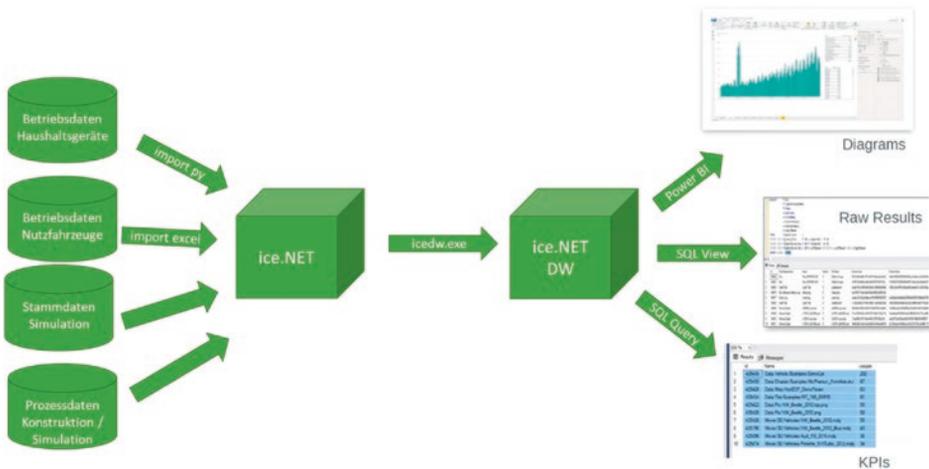


Abb. 5.7 Instanziierung des Data Warehouse für AKKORD

5.4 Diskussion und Entscheidungshilfe zur Auswahl der entwickelten Lösungsbausteine

Die Stärke des Data-Warehouse-Ansatzes liegt in der flexiblen Vorverarbeitung in der Staging-Datenbank, da hier eine uneingeschränkte Transformation der Daten vorgenommen werden kann. Wie bereits beschrieben können die anderen beiden Ansätze (d. h. dezentrale Datenvernetzung oder zentrales Datenmodell) eingebunden werden, indem die Daten durch ein geeignetes Austauschformat in die Staging-Datenbank übertragen werden. Der Ansatz ist also insbesondere geeignet, wenn ein exploratives Herangehen an die Daten und eine umfangreiche Datenvorverarbeitung bereits im Backend erwünscht sind. Zur Befüllung des Data Warehouses kann bei Bedarf auf Datenzugriffsbauusteine zugegriffen werden, die im Kontext der anderen Ansätze entwickelt wurden.

Der Aufbau eines zentralen Datenmodells ist sinnvoll, wenn im angedachten Einsatzszenario eine hohe Integrationstiefe und ein direkter Zugriff auf Daten aus dem Live-Betrieb erforderlich sind. Sofern eine vollständige Backend-Lösung ohne zuvor bestehende Datenbanken aufgebaut werden soll, liefert ein zentrales Backend direkt einen Ort, an dem Daten persistent gespeichert werden können. Für Datenanbindung und -weitergabe kann hier ebenfalls auf Bausteine aus dem dezentralen Ansatz zurückgegriffen werden, wie in Abschn. 5.3.1 aufgezeigt wird. Falls bereits andere datenverwaltende IT-Systeme im Einsatz sind, gilt es hier aber den „Single Source of Truth“-Grundsatz zu beachten, d. h. keine Datenbestände parallel weiter zu pflegen. Stattdessen wäre hier eine (potenziell aufwendige) Datenmigration angeraten.

Der dezentrale Ansatz kommt ohne die Einführung einer großen neuen Datenbank aus. Über offene Standards kann mit relativ geringem Aufwand auf neue Quellsysteme zugegriffen werden, weshalb der Ansatz im Kontext des Forschungsprojekts auch oft als „Datenlieferant“ für die beiden anderen Ansätze zum Einsatz kam. Dieser Daten-Minimalismus wird aber damit erkaufte, dass zunächst keine Datenbank für neue Daten ohne persistentes Quellsystem (z. B. ein neu angebundenes IoT-Asset) zur Verfügung steht.

5.5 Zusammenfassung

Insgesamt ergibt sich im Kontext des im Projekt realisierten Referenzbaukastens eine flexible, modulare Lösung, die auf spezifische Datenanalyse-Anforderungen des jeweiligen Unternehmens anpassbar ist. So stehen neben dem vorgestellten Prozess und IT-Tool zur Analyse vorhandener Datenmodelle Lösungen für die Sammlung und Verwaltung der relevanten Daten: Dieser kann durch den Aufbau einer Staging-Datenbank für Analysen in einem kontrollierten Umfeld erfolgen. Gleichermaßen stehen aber auch Konzepte für den dezentralen Zugriff auf Daten mittels Verwaltung eines gemeinsamen Metadatenmodells sowie skalierbare Konzepte für den Aufbau eines Datenbackends im Unternehmen bis hin zu einer vollständigen Lösung zur Datenverwaltung basierend auf

dem PLM-System Contact Elements zur Verfügung. Je nach gewähltem Lösungsansatz stehen die gesammelten Daten und/oder Analyseergebnisse zur weiteren Freigabe zur Verfügung. Die Anforderungen an das Rollen- und Rechte-Management variieren hierbei stark mit den anvisierten Anwendungsfällen. Besonders im Hinblick auf Kollaboration über Unternehmensgrenzen hinweg besteht ein großer Bedarf daran, den Zugriff auf freigegebene Daten des eigenen Unternehmens feingranular kontrollieren zu können. Im Projekt kommt das ausgereifte Rollenkonzept der Contact Plattform zum Einsatz, für Daten aus den Quellsystemen wird auf die sorgfältige Vergabe von Rechten in den jeweiligen Zielplattformen gesetzt, um den beteiligten Unternehmen maximale Sicherheit und Transparenz zu garantieren.

Im Rahmen der Use Cases im Projekt wurde ein breites Feld an unterschiedlichen Szenarien aufgezeigt, weshalb eine „One-Size-Fits-All“-Lösung für das Datenbackend nicht realistisch erscheint. Stattdessen wird im Rahmen des modularen Referenzbaukastens des AKKORD-Projekts die hier vorgestellte Bandbreite an Lösungsbausteinen bereitgestellt, um auf die individuellen Anforderungen von Unternehmen eingehen zu können.

Denkbare Erweiterungen der im Projekt entwickelten Lösung wären beispielsweise Konzepte zu entwickeln, wie die Datenfreigaben anonymisiert und gefiltert werden können. Zudem könnten zeitgebundene Freigaben zwischen Unternehmen betrachtet werden, was beispielsweise ein „Abo-Modell“ für Nutzungsdaten ermöglicht.

Die im Leistungsbereich entwickelten Lösungen werden im Rahmen der Projektergebnisse zur Verfügung gestellt. Zudem wird derzeit die Veröffentlichung der im akademischen Umfeld entstandenen Lösungskomponenten geprüft.

Literatur

- Bergsjö, D., Malmqvist, J., & Ström, M. (2006). Architectures for mechatronic product data integration in PLM systems. In *Proceedings of the, 9th international design conference – Design 2006*, S. 1065–1076.
- Devlin, B., & Cote, L. D. (1996). *Data warehouse: From architecture to implementation*. Addison-Wesley Longman Publishing Co., Inc.
- Eiden, A., Eickhoff, T., Göbel, J. C., Apostolov, C., Savarino, P., & Dickopf, T. (2022). Data networking for industrial data analysis based on a data backbone system. *Proceedings of the Design Society*, 2, 693–702. <https://doi.org/10.1017/pds.2022.71>.
- Eiden, A., Eickhoff, T., Gries, J., Göbel, J. C., & Psota, T. (2021). Supporting semantic PLM by using a lightweight engineering metadata mapping engine. *Procedia CIRP*, 100, 690–695. <https://doi.org/10.1016/j.procir.2021.05.146>.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step data mining guide, CRISP-DM consortium.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. In *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Abstraktion und Übertragbarkeit von Datenanalysen im industriellen Kontext

6

Entwicklung von modularisierten Analysebausteinen zur Nutzung in industriellen Anwendungsfällen

Marius Syberg , Nikolai West , Edin Klapic , Jörn Schwenken , Ralf Klinkenberg  und Jochen Deuse 

6.1 Einleitung

Aufgrund der zunehmenden Digitalisierung der industriellen Produktion rückt die **industrielle Datenanalyse** in den Fokus von Unternehmen, die im Zeitalter global konkurrierender Wertschöpfungsnetzwerke wettbewerbsfähig bleiben möchten. Das Verständnis industrieller Datenanalysen folgt dazu der Definition von Kap. 2. In diesem Kapitel wird außerdem der gebräuchliche, englischsprachige Ausdruck der *Industrial Data Science* (IDS) verwendet. Grundsätzlich folgt jedes IDS-Projekt im produzierenden Gewerbe einem ähnlichen Prozess, wobei bestimmte Prozesse und Abläufe in den Projekten wiederkehren (Schulz et al., 2021, S. 2 f.). Diese Erkenntnis ist

M. Syberg (✉) · N. West · J. Schwenken · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: marius.syberg@ips.tu-dortmund.de

N. West
E-Mail: nikolai.west@ips.tu-dortmund.de

J. Schwenken
E-Mail: joern.schwenken@ips.tu-dortmund.de

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

E. Klapic · R. Klinkenberg
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: eklapic@rapidminer.com; eklapic@altair.com

R. Klinkenberg
E-Mail: rklinkenberg@rapidminer.com; rklinkenberg@altair.com

den Nutzern der Datenanalyse seit der Entwicklung des *Knowledge Discovery in Databases* (KDD) spätestens bekannt (Fayyad et al., 1996, S. 13 ff.). Für eine Vorstellung von Vorgehensmodellen der industriellen Datenanalyse verweisen wir auf Kap. 3. Neben einem Prozessmodell sind menschliche Kompetenzen von entscheidender Bedeutung für den erfolgreichen Abschluss eines Datenanalyseprojekts. Hierfür definieren Mazarov et al. (2020, S. 63 ff.) vier **Kompetenzbereiche** und zugewiesene **Rollen** für die unterschiedlichen Verantwortlichkeiten in einem IDS-Projekt:

1. Management
2. Data Scientist
3. Domänenexperte
4. IT-Personal

In ähnlicher Weise können die Kompetenzbereiche für verschiedene Verantwortlichkeiten in vier Gruppen unterteilt werden. Daher empfehlen Deuse et al. (2022, S. 3) den Ablauf eines IDS-Projekts in vier separate Elemente zu teilen (siehe Kap. 4). Das Ergebnis ist eine **Prozesskette**, die diese Aufgaben in einer bestimmten Reihenfolge anordnet und Interaktionspunkte definiert. Die Segmentierung von IDS-Projekten entlang der Prozesskette ermöglicht es den Mitarbeitern der genannten Kompetenzbereiche, ihre jeweiligen Fähigkeiten gezielt einzubringen. Darüber hinaus erleichtert die Prozesskette die Entwicklung von Analyseergebnissen in einer wiederverwendbaren Form, indem sie einzelne Teilanalysen klar voneinander trennt. Um jedoch eine gute Wiederverwendbarkeit von Lösungen für Datenzugriff, Datenanalyse, Datennutzung und Datenmanagement zu erreichen, ist eine geeignete Plattform für die Zusammenarbeit mit einer angemessenen IT-Infrastruktur erforderlich. Um dies zu erreichen, muss ein Unternehmen diese Rollen besetzen, um Projekte erfolgreich durchzuführen.

In den meisten Fällen müssen die erforderlichen Kompetenzen zunächst mit einem gewissen Aufwand aufgebaut werden. Gleichzeitig bietet die iterative Natur solcher Projekte den Mehrwert, dass diese Kompetenzen mittel- und langfristig nützlich sein können, indem sie in vielen weiteren Projekten eingesetzt werden. Es muss daher möglich sein, die eigene Datenanalyseaufgabe in Elemente (Module) zu unterteilen, die wiederum wiederverwendet werden können. Dieses Kapitel beschreibt die Methodik und Umsetzung der Abstraktion und die Generalisierung von Datenanalysen im Rahmen des AKKORD-Projekts (siehe Kap. 1). Somit ist es Unternehmen, insbesondere KMU, möglich, niederschwellig Datenanalysen auszuführen, (weiter-) zu entwickeln und im Unternehmen mit anderen Nutzern oder außerhalb mit anderen Unternehmen zu teilen. Dazu werden zunächst die bereits erhobenen Anforderungen zusammengefasst und danach die Modularisierung industrieller Datenanalysen methodisch und technisch mitsamt den entwickelten Funktionalitäten beschrieben. Spezifika des Projektergebnisses folgen im Anschluss. Ein Ausblick auf die zukünftige Nutzung schließt dieses Kapitel ab.

6.2 Anforderungen an die Generalisierung von Datenanalysen

Für den erfolgreichen Einsatz von Datenanalysen im industriellen Umfeld, insbesondere für kleine und mittlere Unternehmen, gelten bestimmte Anforderungen. Syberg et al. (2023, S. 64 ff.) haben diese Anforderungen durch Experteninterviews mit Anwendern und Anbietern sowie auf Basis einer umfassenden Literaturrecherche erhoben.

Eine wichtige Grundlage für die Erhebung der Anforderungen stellt das FURPS-Modell von Grady und Caswell (1987) dar. Die Abkürzung steht für *Functionality, Usability, Reliability, Performance, Supportability* und wurde in Abb. 6.1 aufgegriffen. Dieses Modell ermöglicht eine strukturierte Erfassung und Kategorisierung der Anforderungen an die Qualität einer Softwarelösung und lässt sich auf die Entwicklung einer Kollaborationsplattform für die industrielle Datenanalyse übertragen. In dieser Adaption des bewährten FURPS-Modells werden explizit IDS-spezifische Anforderungen definiert. Dieses sichert eine stabile Datenanalyse-Pipeline.

Funktionalität. Im Bereich *Functionality* werden Anforderungen an die Fragmentierung des Analyseprozesses in Module gestellt, die wiederverwendbar und in verschiedenen Konfigurationen nutzbar sein müssen. Zudem müssen die Analyseergebnisse durch Schnittstellen oder Visualisierungsoptionen nutzbar gemacht werden können. Dies ist insbesondere bei KMU unabdinglich, da häufig keine eigenen Ressourcen zur Programmierung von Schnittstellen an eigene Systeme bestehen. Gleichzeitig sollten Visualisierungen von Datenanalyseergebnissen entsprechend zugänglich gewählt werden können, um zielgerichtete Auswertungen zu ermöglichen (Mazarov et al., 2020, S. 63 ff.).

Nutzbarkeit. Die Anforderungen im Bereich *Usability* beinhalten die Entwicklung von Modulen für alle Phasen eines IDS-Projekts, die auf die Bedürfnisse der

Anforderung	Beschreibung
Functionality (<i>Funktionalität</i>)	Fragmentierung des Analyseprozesses; Konfigurierbarkeit; Nutzung der Ergebnisse in anderen Umgebungen
Usability (<i>Nutzbarkeit</i>)	Erstellung von Modulen für alle Phasen eines IDS-Projekts; Auswahl und Erstellung von Modulen durch Benutzer
Reliability (<i>Zuverlässigkeit</i>)	Vorverarbeitung der Daten und Vorbereitung der Analyse zur Sicherstellung der Datenqualität
Performance (<i>Leistung</i>)	Leistungsstarke analytische Umgebung (Cloud-basiert); Verknüpfung mit Kompetenz-Entwicklung
Supportability (<i>Unterstützung</i>)	Strukturierte Datenbank, Up- und Download von Daten und Ergebnisse; Unterstützung bei technischen Problemen

Abb. 6.1 Anforderungen aus dem Bereich der industriellen Datenanalyse an eine Kollaborationsplattform (Syberg et al., 2023, S. 66)

industriellen Produktion abgestimmt sind. Diese Module müssen entsprechend der in Kap. 3 genannten Prozessmodelle beschrieben werden, um den erfolgreichen Verlauf des Projekts zu gewährleisten. Benutzer müssen in der Lage sein, Module auszuwählen oder bei Bedarf selbst zu erstellen und anderen wiederum zur Verfügung stellen zu können. Gleichzeitig muss das Verfahren entsprechend dokumentiert werden, um es für neue Benutzer zugänglich zu machen.

Zuverlässigkeit. Unter *Reliability* wird die Notwendigkeit betont, eine bestimmte Menge an Vorverarbeitung der Daten sowie Vorbereitung des Analyseprozesses technisch zu handhaben, um die erforderliche Datenqualität für Datenanalysen zu erreichen (West et al., 2021, S. 133 ff.).

Leistung. Die Anforderungen im Bereich *Performance* beziehen sich auf eine leistungsstarke Serverumgebung als Teil einer Cloud-Lösung, die eng mit der technischen Infrastruktur verbunden ist. Inhaltlich kann das Potenzial der Plattform in Bezug auf die Performance nur durch die Verknüpfung der praktischen Datenanalyse mit der Kompetenzentwicklung und dem kontinuierlichen Lernen in geeigneter Weise realisiert werden. Dennoch muss die Lösung methodisch und technisch ermöglichen, auch erweiterte Analyseverfahren aus dem Bereich des überwachten, semi-überwachten und unüberwachten Lernen zu nutzen.

Unterstützung. Die *Supportability* wird durch den kollaborativen Ansatz der Plattform sichergestellt. Dazu gehört auch die Umsetzung einer strukturierten Datenbank, die Möglichkeit zum Hoch- und Herunterladen von Daten und Ergebnissen sowie die Verfügbarkeit eines technischen Supports bei Problemen.

Diese Aspekte decken den Bereich der technischen Infrastruktur ab. Gleichzeitig erfordert die Heterogenität der Anwendungsfälle in den Unternehmen explizit eine Individualisierung. Dies bedeutet vor allem, dass die Analysemodule individuell angepasst werden können. So kann zum Beispiel eine Zeitreihenanalyse für ein Unternehmen im Vertriebsbereich sinnvoll sein (siehe Kap. 11), aber für einen anderen Benutzer im Bereich der Prozessüberwachung relevant sein (siehe Kap. 12). Entsprechend soll ein Modul wiederverwendet werden können, wenn es konfigurierbar und entsprechend auf andere Daten übertragbar ist. Gleichzeitig müssen auch Assistenzsysteme implementiert werden.

6.3 Modularisierung industrieller Datenanalysen

Die Abstrahierung und Modularisierung von bisher einmalig durchgeführten Datenanalysen ist das zentrale Objekt in der Befähigung von Unternehmen zur Anwendung industrieller Datenanalysen. Da die Entwicklung durch die Probleme der Anwender getrieben ist, ist eine Abstrahierung der Datenanalyse-Pipeline sinnvoll. Im Rahmen der Arbeiten im Leistungsbereich wurde das Grundkonzept von verallgemeinerten, wiederverwendbaren Modulen zur Problemlösung aufgegriffen und in ein Nutzungskonzept überführt. Verallgemeinerte Module sind in der Informatik weit verbreitet, sei es in der

Medizintechnik oder im Bereich des Internet of Things (Banijamali et al., 2020; Sievi-Korte et al., 2019). Dieser Aufbau bietet zahlreiche Vorteile für die Anwender der Systeme:

- Ein wichtiger Vorteil von verallgemeinerten Modulen ist ihre **definierte Schnittstelle**. Die Verwendung solcher definierten Schnittstellen erleichtert die Integration des Moduls in ein bestehendes System. Da verallgemeinerte Module eben diese bieten, müssen nur wenige Änderungen am System vorgenommen werden, um sie zu integrieren. Dies vereinfacht den Integrationsprozess erheblich und senkt die Integrationshürden.
- Ein weiterer Vorteil verallgemeinerter Module ist die **Bildung von Teilmodulen**. Durch die Aufteilung eines komplexen Problemlösungsprozesses in Teilmodule kann die Komplexität reduziert werden. Jedes Teilmodul kann dann separat entwickelt und getestet werden, bevor es in das Gesamtsystem integriert wird. Dadurch wird das Fehlerrisiko minimiert und die Qualität des Gesamtsystems verbessert.
- Die Zerlegung der Analyse in Teilmodule erleichtert auch die **Skalierbarkeit**. Wenn ein System erweitert werden muss, kann einfach ein neues Modul hinzugefügt werden. Da das Modul immer gleiche Schnittstellen zur Verfügung stellt, müssen nur wenige Änderungen am Gesamtsystem vorgenommen werden. Dies erleichtert die Anpassung an sich ändernde technische oder inhaltliche Anforderungen.
- Ein weiterer Vorteil verallgemeinerter Module ist das **Ersetzen von Teilaspekten durch ‚bessere‘ Module**. Wenn ein Teilaspekt des Systems verbessert werden muss, kann einfach ein neues Modul erstellt werden, das den alten Teilaspekt ersetzt. Dies erhöht die Flexibilität und Anpassungsfähigkeit des Systems. Hauptvorteil ist jedoch die *Wiederverwendbarkeit* inner- und außerhalb von Unternehmensgrenzen. Wenn ein Modul in einem Projekt entwickelt wurde, kann es einfach in anderen Projekten wiederverwendet werden, um Zeit und Entwicklungskosten zu sparen. Dies reduziert auch die Wahrscheinlichkeit von Fehlern und verbessert die Qualität des Systems.

Technisch wird dieser modulare Aufbau in den Arbeiten zu diesem Kapitel mithilfe der Datenanalysesoftware **RapidMiner** umgesetzt. RapidMiner ist eine Softwareplattform für Maschinelles Lernen und Datenanalyse, mit der Benutzer Daten visuell analysieren, Modelle entwickeln und bereitstellen können, ohne direkt Programmieren zu müssen. RapidMiner enthält verschiedene Komponenten, die für verschiedene Phasen des Datenanalyseprozesses verwendet werden können.

Die Software bietet eine Vielzahl von Werkzeugen, um Datenvorverarbeitungsprozesse durchzuführen und darauf aufbauend niederschwellig ML-Modelle zu erstellen und zu trainieren. Das Portfolio enthält zahlreiche Algorithmen, die für verschiedene Anwendungsfälle wie Klassifikation, Regression, Clustering und Assoziationsanalysen geeignet sind. Darüber hinaus bietet die Software eine breite Palette von Reporting- und Visualisierungstools, um Ergebnisse und Erkenntnisse zu kommunizieren.

6.3.1 Methodische Idee der Datenanalysemodule

Betrachtet man die nötigen Elemente eines Datenanalyseprozesses als abstraktes System, benötigt dieser einen **Input**, eine **Analyse** und einen **Output**. Diese drei Elemente gilt es für jeden Analyseprozess zu abstrahieren und zu modularisieren. Analog zu einem simplen Transformationsprozess, bilden diese Elemente ein Analysemodul, wie es bereits in Kap. 2 beschrieben wurde (Abb. 6.2).

Zunächst muss für die Funktionalität sichergestellt werden, dass ein **Input-Datensatz** für einen Analyseprozess besteht und immer gleich aufgebaut ist. Um die Eigenschaften von Input-Datensätzen abzubilden, werden unterschiedliche Abstraktionslevel genutzt. Das im Forschungsprojekt **DaPro** (www.dapro-projekt.de) von RapidMiner entwickelte *Data Schema* hilft dabei, die Anforderungen an die Input-Daten zu definieren. Dies umfasst Dateninhalte, Statistiken und allgemeine Beschreibungen einzelner, für die Analyse notwendiger und optionaler Attribute. Eine grafische Benutzeroberfläche unterstützt die Anwender dabei, eigene Inputdaten den bestehenden Anforderungen zuzuordnen. Das *Data Schema* beinhaltet zusätzliche Hilfestellungen, wie zum Beispiel eine automatische Konvertierung von Datentypen. Dies ermöglicht beispielsweise die Anwendung der Module mit unterschiedlichen Datumsformaten.

Unter Berücksichtigung des vollständigen Datenanalyseprozesses werden die verschiedenen Funktionalitäten hinsichtlich ihrer Generalisierbarkeit und Notwendigkeit betrachtet. Auch die für die Analyse erforderlichen Daten werden kritisch untersucht, um festzustellen, welche davon unerlässlich, welche optional und welche für das Ergebnis überflüssig sind. Eine umfassende Beschreibung der Attribute ist wichtig, um die gewünschten Ergebnisse aus den Analyseprozessen zu erzielen. Dabei können auch bestimmte statistische Kennzahlen wie Verteilungen oder Anzahl von Fehlwerten als Vorgabe dienen. Das Ziel ist es, eine gekapselte Prozess-Einheit zu schaffen, die sich leicht mit Variablen modifizieren lässt, um komplexe Sachverhalte über *Schleifen*, *if-else*-Anweisungen oder weitere logische Verknüpfungen abbilden zu können. Auch die Kapselung von ganzen datenwissenschaftlichen Methoden und Modellen ist so möglich.

Das zweite Element ist die abstrahierte **Pipeline** für die industrielle Datenanalyse. Abb. 6.3 zeigt ein exemplarisches Beispiel einer Pipeline in RapidMiner Studio. Jede Verarbeitung eines Datensatzes wird mit einer solchen Kette aus **Operatoren** dar-

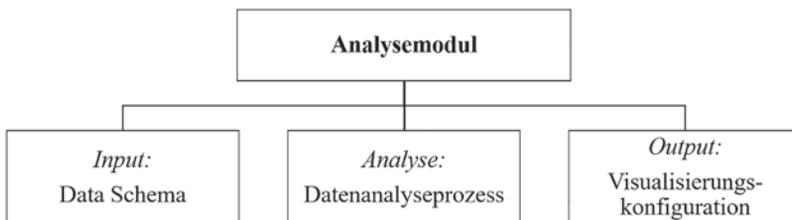


Abb. 6.2 Schematische Darstellung der Elemente eines Datenanalysemoduls

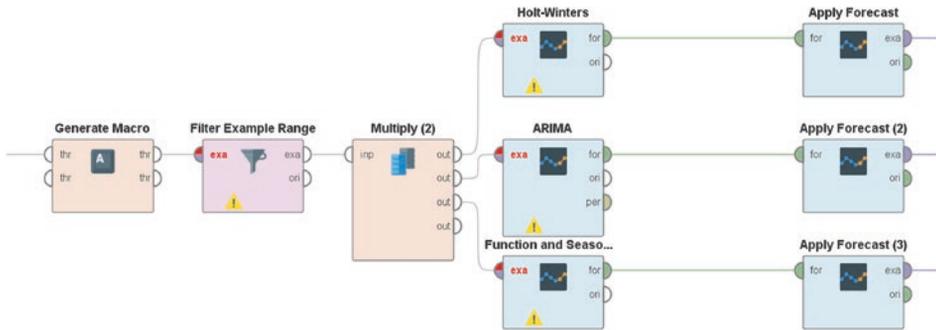


Abb. 6.3 Exemplarische Pipeline für einen Analyseprozess in RapidMiner Studio am Beispiel einer univariaten Zeitreihenprognose mit mehreren Vorverarbeitungsprozessen

gestellt Diese Kette aus verschiedenen Operatoren aus RapidMiner Studio wird einer Java-Archive-Datei zusammengeführt und zu einem eigenen spezifischen Operator zusammengeführt, der als *Custom Operator* bezeichnet wird. Dieser stellt das Hauptelement des Analysemoduls dar. Der Custom Operator enthält dabei außerdem auch allgemeinere textuelle Informationen, die der Ersteller eines Analysemoduls den späteren Nutzern kommunizieren möchte.

Das dritte Element ist die **Visualisierung** des Ergebnisses. Die vom Ersteller als am geeignetsten identifizierte Visualisierung kann von diesem gespeichert und ebenfalls als Teil des Analysemoduls bereitgestellt werden. Dabei ist diese Konfigurationsdatei austauschbar und so vom Nutzer entsprechend seiner Bedürfnisse auch im Nachgang noch anpassbar.

Diese Elemente des Analysemoduls werden schließlich zu einem Analysemodul zusammengeführt und als Module auf der *AI-Toolbox*, einer eigens entwickelten Datenanalyse-Plattform, zur Verfügung gestellt, die in Abschn. 7.4 ausführlicher beschrieben wird. Dort ist dieses Analysemodul auf den Daten eines ebenfalls hochgeladenen Datensatzes anwendbar.

6.3.2 Technischer Hintergrund der modularisierten Datenanalysen

Die Softwareplattform *RapidMiner Studio* basiert auf der Programmiersprache Java. Neben dem Kernprodukt wurden seitens RapidMiner einige Software-Erweiterungen entwickelt sowie öffentliche Schnittstellen bereitgestellt, welche auch externen Personen die Möglichkeit bieten, mithilfe von Erweiterungen eigenständige Lösungen zu programmieren. Beispielsweise wird ein eigenständiger Operator angeboten, mit dem es möglich ist, Python-Programmcode auszuführen und so eine noch höhere Individualisierung in der Datenanalyse zu erreichen. Im vorliegenden Anwendungsfall konnte eine bereits existierende Erweiterung zur Gruppierung von Analyseschritten

mittels umfangreicher Anpassungen für die Umsetzung im Kontext allgemeiner Datenanalysemodule verwendet werden.

Dabei werden die einzelnen Modulkomponenten verschiedentlich repräsentiert. Während die Prozesse im XML-Format existieren, ist die Visualisierungskonfiguration im JSON-Format und das Datenschema als HDF5-Datensatz verfügbar. Sobald die einzelnen Bestandteile die internen Qualitäts- und Funktionalitäts-Tests bestanden haben, erfolgt die Bündelung und Kapselung in Form des Custom Operators. Daran anschließend erfolgt die Transformation des Custom Operators in eine spezifische RapidMiner Erweiterung, die für die spätere Produktivnahme auf der im Abschn. 7.4 beschriebenen AI-Toolbox benötigt wird.

Im Zuge dessen werden eben jene in Real-Time Scoring Agents (RTSA) durchgeführt. Dabei handelt es sich um vordefinierte, allein lauffähige und für die operationalisierte Nutzung direkt einsetzbare Softwareprogramme, die als “Representational state transfer” (REST) Web Service Endpunkte genutzt werden. Als Besonderheit der RTSA ist zu nennen, dass diese auf jedem Gerät lauffähig sind, welches die Mindestvoraussetzungen erfüllt. Hierbei gilt insbesondere die Einschränkung, dass eine Ausführungs-umgebung für sogenannte Container, wie beispielsweise *Docker* verfügbar sein muss. Ein RTSA Deployment, und somit auch ein Datenanalysemodul, wird durch den Start des Containers aktiviert. Neben einem minimalistischen Webserver für die Verfügbarkeit als Webservice enthält das Deployment, wie im vorangegangenen Abschn. 7.3 beschrieben, das Datenschema, die Visualisierungskonfiguration und die Prozessdefinition.

Der RapidMiner RTSA ist für eine breite Palette an Einsatzmöglichkeiten vorgesehen. Als Input- und Output-Standard wird das Dateiformat *JavaScript Object Notation* (JSON) verwendet, welches vielfältige Varianten zur Datenrepräsentation bietet. Im Rahmen der programmatischen Umsetzung wurde zur besseren Nutzbarkeit für die Anwender das Datenformat CSV nach außen gegeben, welches intern in JSON umgewandelt wird. In der Anwendung wird beim Aufruf des RESTful Web Service der zu analysierende Datensatz als Payload des Web Service Aufrufs an das Datenanalysemodul übergeben, welches die vorgesehene Verarbeitung vornimmt und die Ergebnisse gemäß den bereitgestellten Informationen ausliefert.

6.4 AI-Toolbox als Enabler für das Nutzen und Teilen der Module

Die Abstraktion industrieller Datenanalysen in generalisierte Analysemodule ist der relevanteste Schritt, produzierenden Unternehmen den **Einstieg in industrielle Datenanalyseprojekte** zu ermöglichen. Für eine Anwendung ist aber die Bereitstellung der Module entscheidend. Die eigens dafür entwickelte Plattform AI-Toolbox ist dabei eine browserbasierte Softwarelösung, mit der die Analysemodule ausgeführt, angepasst und geteilt werden können. Das Kürzel AI steht dabei für *Artificial Intelligence*, die englische Bezeichnung für die **Künstliche Intelligenz**. Der Begriff *Toolbox* wurde gewählt, in

Anlehnung an die Analogie des **Werkzeugkoffers** für industrielle Datenanalysen (siehe Kap. 1 und 4). Technisch handelt es sich bei der AI-Toolbox um eine container-basierte *Spring-Boot-Applikation*, deren Frontend mittels des React-Frameworks realisiert worden ist. In der Anwendung wird diese an einen sogenannten *AI-Hub* von RapidMiner angedockt. So ist es möglich, in gängigen Internetbrowsern und auf mobilen Endgeräten Analysemodule zu verwenden. In der grundsätzlichen Anwendung erhält der Nutzer eine Übersicht über die Funktionen der verfügbaren Module, kann die gewünschten Daten im entsprechenden Format hochladen und das Modul auf der AI-Toolbox selbst ausführen. Als Ergebnis werden eine Visualisierung sowie ein Ergebnisdatsatz ausgegeben. Neben dem Nutzer als Hauptanwender der Analysemodule bestehen mit dem *Analysten* und dem *Admin* zwei weitere Rollen. Die Rollenverteilung ist hierbei an ein transdisziplinäres Rollenkonzept für Industrial Data-Science-Projekte (Schwenken et al., 2023, S. 82 f.) angelehnt. Dabei ist der Analyst neben der Nutzung der Module in der Lage, bestehende Module zu modifizieren und eigene Module hochzuladen. Der *Admin* ist darüber hinaus in der Lage, die Implementierung der Plattformlösung technisch zu überwachen und zu steuern und fungiert so als technischer Support.

Der Aufbau der AI-Toolbox ist an gängige Plattformen angelehnt, die im Consumer Bereich marktführend sind. Das **UX Design** ist dabei möglichst einfach gehalten, um eine leichte Bedienbarkeit zu gewährleisten. Im Bereich *Production Modules* findet sich eine Übersicht über alle vorhandenen Analysemodule. Diese sind über ihre Bezeichnung und Icons schnell den eigenen Anwendungsfällen und Problemstellungen zugeordnet. Gleichzeitig dienen die Demoprozesse und Beschreibungen jedes einzelnen Moduls dazu, den Nutzer bei der Auswahl des passenden Moduls zu unterstützen. Der Bereich *Dashboard* zeigt die Ergebnisse des ausgewählten Analysemoduls. Dabei ist die Visualisierung nicht statisch, sondern entsprechend der gewählten Konfiguration anpassbar. So lassen sich Bereiche der Abbildung vergrößern und Inhalte ein-/ausblenden. Auch dies begünstigt die Nutzererfahrung, da mit geringem Aufwand Übersichtlichkeit in einer Visualisierung hergestellt werden kann, um Wissen aus der Analyse zu generieren. Beispielhaft ist eine Anomalie in einer Zeitreihe näher analysierbar, auch im Vergleich verschiedener Prognosemodelle im Rahmen der Absatzprognose lassen sich durch gezieltes Ein- und Ausblenden entsprechender Verfahren schnelle Vergleiche auf Basis der Visualisierung ziehen. Außerdem lassen sich Analyseergebnisse als CSV-Dateien exportieren.

Neben der reinen Verwendung der Analysemodule ist die AI-Toolbox vielmehr **Plattformkonzept** als Softwarelösung (siehe Kap. 3). Hierbei wird neben der reinen Nutzung von Analysemodulen das Ziel angestrebt, die Entwicklung und Verbreitung von Modulen zu vereinfachen. Dies wird durch die Möglichkeit realisiert, nicht nur vorhandene Analysemodule anpassen und wieder bereitstellen zu können, sondern vollständig neue Analysen in RapidMiner zu erstellen und auf der Plattform Anwender zur Verfügung zu stellen. Dadurch wird Skalierbarkeit und Flexibilität gewährleistet. Technisch fungiert die AI-Toolbox so als service-orientierte Schnittstelle zwischen Nutzern ohne große Vorkenntnisse und Datenanalysen. Zudem erfüllt die Möglichkeit des Datenexports eine Schnittstellenfunktion in andere Systeme. Auf der operativen

Ebene wird die AI-Toolbox mit einer geschaffenen Verknüpfung zur Work&Learn-Plattform zu einer Kollaborationsplattform. Dieser Teil der Entwicklungen im AKKORD Projekt erfüllt Kommunikations- und Informationsfunktionen und ermöglicht außerdem eine Zusammenarbeit. Außerdem wurden durch textuelle und visuelle Anleitungen für einzelne datenanalytische Themenblöcke aber auch die Nutzung der AI-Toolbox selbst Voraussetzungen dafür geschaffen, das Personal in einem Unternehmen dazu zu befähigen, die Plattform zielführend einzusetzen.

6.5 Zusammenfassung

Die Abstraktion industrieller Datenanalysen und die Umsetzung in modularisierte Elemente ist den obigen Ausführungen folgend ein vielversprechender Ansatz zur Verbreitung datenanalytischer Methoden in KMU. Die Umsetzung in Form einer Plattform wie der AI-Toolbox ist ein Erfolgsfaktor für die Implementierung in der täglichen Arbeit. Nutzer können die Analysemodule in ihrem bekannten Internetbrowser nutzen und schnell auf der Grundlage von Visualisierungen Wissen aus ihren Daten generieren. Der skalierbare Ansatz auf Basis von RapidMiner ermöglicht außerdem spezielle Anwendungen und individuelle Analysen für komplexe Problemstellungen. RapidMiner ist dabei explizit für den Einsatz von Maschinellen Lernverfahren bzw. datenwissenschaftlichen Methoden entwickelt worden. Jeder in RapidMiner verfügbare Operator ist dabei in ein Analysemodul auf der AI-Toolbox überführbar. Dadurch ist die Komplexität der angewandten Analysen nur durch die Rechenleistung der Umgebung begrenzt, auf der der zu Grunde liegende AI-Hub gehostet wird. Selbst die Einbindung von nicht RapidMiner-basierten Datenanalyseprozessen ist möglich, da die Software die freie Programmierung in anderen Programmiersprachen wie Python oder C++ mit entsprechend dafür entwickelten Operatoren unterstützt. So stellt die Abstraktion der industriellen Datenanalyseprozesse in Kombination mit der Bereitstellung auf der AI-Toolbox einen niederschweligen Einstieg in die industrielle Datenanalyse dar, die gleichzeitig hoch spezialisiert und maximal individualisiert im Unternehmen eingesetzt werden kann. Kap. 18 stellt ein anschauliches Beispiel für den Einsatz von Analysebausteinen für den Anwendungsfall einer Spielzeug-Autorennbahn vor.

Die AI-Toolbox kann dabei eine **Schlüsseltechnologie** für Unternehmen in der datenbasierten **Entscheidungsfindung** und der Anwendung von Maschinellen Lernen sein. Entscheidend in der Roll-Out-Phase sind hierbei drei Aspekte. Erstens muss die technische Infrastruktur hergestellt werden. In der Entwicklung der Plattform hat sich gezeigt, dass das Hosting des zu Grunde liegenden AI-Hubs in die bestehende IT-Infrastruktur ein elementarer Punkt ist. Dabei müssen insbesondere auch Datenschutz- und Sicherheitsaspekte berücksichtigt werden. Cloud- und On-Premise-Lösungen sind beide im Einsatz. Die kontinuierliche Aufnahme von Daten innerhalb der Produktion ist als Grundlage für eine datengetriebene Optimierung Voraussetzung (Richter et al., 2022, S. 1). Zweitens ist der Roll-Out entsprechend zu organisieren. Zusammenfassend ist

anzuführen, dass die Hürde der umfassenden Mitarbeiterschulung, die mit einer Softwareimplementierung häufig einhergeht, durch die begleitend entwickelte Work&Learn-Plattform minimiert wird. Dennoch sind klare Prozesse zur Einführung zu etablieren. Auch müssen die entsprechenden Rollen und Zuständigkeiten definiert werden. Der letzte Aspekt betrifft die Managementunterstützung. Gerade in der datenbasierten Entscheidungsfindung mithilfe von erweiterten Verfahren des Maschinellen Lernens, wie in vielen Digitalisierungsmaßnahmen, ist es wichtig, durch eine entsprechende Förderung Vertrauen in die Technologie zu schaffen, um so die Nutzerakzeptanz zu erhöhen.

Literatur

- Banijamali, A., Pakanen, O.-P., Kuvaja, P., & Oivo, M. (2020). Software architectures of the convergence of cloud computing and the internet of things: A systematic literature review. *Information and Software Technology*, 122, 106271.
- Deuse, J., West, N., & Syberg, M. (2022). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17, 37–54.
- Grady, R. B., & Caswell, D. L. (1987). *Software metrics. Establishing a company-wide program*. Prentice-Hall.
- Mazarov, J., Schmitt, J., Deuse, J., Richter, R., Kühnast-Benedikt, R., & Biedermann, H. (2020). Visualisation in industrial data science projects (Translation). Visualisierung in industrial data-science-projekten (Original title). *Industrie 4.0 Management*, 36(6), 63–66.
- Richter, R., Syberg, M., Deuse, J., Willats, P., & Lenze, D. (2022). Creating lean value streams through proactive variability management. *International Journal of Production Research*, 1–12. <https://doi.org/10.1080/00207543.2022.2111614>.
- Schulz, M., Neuhaus, U., Kaufmann, J., Badura, D., Kerzel, U., Welter, F. et al. (2021). DASC-PM v1.0 – Ein Vorgehensmodell für Data-Science-Projekte. Unter Mitarbeit von Universitäts- und Landesbibliothek Sachsen-Anhalt und Martin-Luther Universität.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., Deuse, J. (2023). Development of a transdisciplinary role concept for the process chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- Sievi-Korte, O., Richardson, I., & Beecham, S. (2019). Software architecture design in global software development: An empirical study. *Journal of Systems and Software*, 158, 110400.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Kompetenzentwicklung und -sicherung in der industriellen Datenanalyse von Wertschöpfungsnetzwerken

7

Vorstellung einer modularen und praxisorientierten
Work&Learn-Plattform mit dem besonderen Fokus auf KMU

Christopher Klupak , Felix Walker , Volker Zimmermann und
Rebekka Adams

7.1 Einleitung und Motivation

Die zunehmende Digitalisierung und Automatisierung in der Industrie hat dazu geführt, dass immer mehr Daten erfasst werden. Diese Daten können besonders für kleine und mittelständische Unternehmen (KMU) eine wichtige Ressource darstellen, um den zahlreichen Herausforderungen, wie zum Beispiel den steigenden Kosten, dem erhöhten Wettbewerbsdruck und den schnellen Veränderungen im Marktumfeld, zu begegnen (West et al., 2021, S. 131 ff.; Haarmer, 2021, S. 5 ff.). Die Bedeutung kleiner und mittelständischer Unternehmen (KMU) in Deutschland ist dabei nicht zu unterschätzen, da sie das Rückgrat der deutschen Wirtschaft mit etwa 99 % aller Unternehmen in Deutschland bilden. Damit sind sie ein wichtiger Treiber für Innovation und Wachstum (Statistisches Bundesamt, 2019). Allerdings haben diese Unternehmen auch besondere Herausforderungen zu bewältigen. Sie verfügen oft über begrenzte Ressourcen und müssen sich gegenüber großen Konzernen behaupten. Dies erfordert ein hohes Maß an Flexibilität

C. Klupak (✉) · F. Walker
Universität Hamburg, Hamburg, Deutschland
E-Mail: christopher.andre.klupak@uni-hamburg.de

F. Walker
E-Mail: felix.walker@uni-hamburg.de

V. Zimmermann · R. Adams
NEOCOSMO GmbH, Saarbrücken, Deutschland
E-Mail: volker@neocosmo.de

R. Adams
E-Mail: rebekka@neocosmo.de

und Anpassungsfähigkeit, um in einem dynamischen Marktumfeld erfolgreich bestehen zu können (Seyda, 2012, S.1ff). Durch die Analyse von Daten können Unternehmen in Echtzeit auf Veränderungen im Marktumfeld reagieren und somit schnell und flexibel agieren. Der Einsatz von industrieller Datenanalyse bietet hierbei eine Möglichkeit, um Unternehmen dabei zu unterstützen, ihre Geschäftsprozesse und Produkte kontinuierlich zu verbessern und somit langfristig wettbewerbsfähig zu bleiben (Czernich, et al., 2019, S. 9 ff.). Für kleine und mittelständische Unternehmen (KMU) kann der Einstieg in die industrielle Datenanalyse allerdings eine Herausforderung darstellen, denn häufig verfügen sie nicht über die notwendige Infrastruktur und Expertise, um große Datenmengen effektiv zu verarbeiten und auszuwerten (Weskamp et al., 2014, S. 22 ff.). In diesem Kontext können Serviceplattformen für die industrielle Datenanalyse einen wichtigen Beitrag leisten, indem sie Unternehmen und ihre Mitarbeitenden dabei unterstützen, ihr Wissen und ihre Fähigkeiten im Bereich der datengestützten Analysen und Entscheidungen zu erweitern (siehe Kap. 1). Auf dem Markt gibt es bereits eine Vielzahl unterschiedlicher Aus- und Weiterbildungsangebote im Bereich der industriellen Datenanalyse. Dabei handelt es sich meist um sehr kosten- und zeitintensive Angebote, die gerade für KMU aufgrund der knappen Ressourcen nicht wahrnehmbar sind. Hier braucht es für die Unternehmen und Mitarbeitenden ein niederschwelliges, kostenloses und ganzheitliches Lösungskonzept, welches im Rahmen des Projektes AKKORD mit der Work&Learn-Plattform (W&L) umgesetzt wird. Dazu wird im ersten Abschnitt ein kurzer Überblick über die bereits am Markt verfügbaren Lernangeboten zur Einordnung gegeben. Im nächsten Abschnitt werden die Anforderung zur „Kompetenzerfassung – und Sicherung in der industriellen Datenanalyse von Wertschöpfungsnetzwerken“ mit dem besonderen Fokus auf KMU definiert. Daran anknüpfend wird die konzeptuelle Entwicklung des Lernbereiches dargestellt, in dem der inhaltliche Aufbau der Lernkurse als Mockup dargelegt wird. Im nächsten Abschnitt wird die Umsetzung der digitalen Lerninhalte als Mockup in die Serviceplattform (W&L) aufgezeigt. Im Anschluss erfolgt die Darstellung der Validierung sowie der Optimierungsansätze. Im letzten Abschnitt wird neben einer Zusammenfassung noch ein Ausblick auf die Zukunft gegeben.

7.2 Analyse und Anforderungsdefinition zum Themenfeld „Kompetenzaufbau und -sicherung in Wertschöpfungsnetzwerken“

Als Grundlage für die Konzeption und Entwicklung im Leistungsbereich ist es zunächst einmal notwendig, den Begriff der industriellen Datenanalyse sowie die notwendigen Kompetenzen näher zu spezifizieren (siehe Kap. 2). Im Bereich der allgemeinen **Datenwissenschaften** gibt es eine große Anzahl datengestützter Ansätze, die ausschließlich auf der Grundlage von Zahlenwerten Erkenntnisse liefern. Beim Einsatz im industriellen Produktionskontext (Engl. Industrial Data Science) scheitert dieser Lösungsansatz u. a. meistens aufgrund der komplexen Anlagenstruktur sowie der spezifischen Anforderungen

einer automatisierten Fertigung. Abhilfe können hier sogenannte Smart Data-Methoden schaffen, die Expertenwissen mit datengetriebenen Big-Data-Analysen vereinen (Trunzer et al., 2019, S. 90 ff.). Beim Einsatz industrieller Datenanalysen werden die gesammelten Daten häufig retrospektiv betrachtet, um anhand von historischen Produktionsdaten Modelle zum Systemverhalten der Fertigungsinfrastruktur zu erstellen. Durch diese Vorgehensweise kann beispielsweise das Wartungsmanagement optimiert werden, um eine höhere Anlagenverfügbarkeit zu erreichen. Die Anwendung von industriellen Datenanalysen gestaltet sich durch die Heterogenität der vorhandenen Daten als große Herausforderung, da der Zugriff auf die unterschiedlichen Systeme und Datenbanken sowie die sinnvolle Zusammenführung der Informationen ein hohes Maß an Fachwissen und Methodenkompetenz voraussetzt.

Um erfolgreich industrielle Datenanalysen durchführen zu können, benötigen gerade die Mitarbeiter in kleinen und mittelständischen Unternehmen (KMU) spezifische **Kompetenzen**. Ein solides Fachwissen über die Funktionsweise und Prozesse in der produzierenden Industrie ist dabei unerlässlich, um die gesammelten Daten richtig interpretieren und analysieren zu können (siehe Kap. 3). Zudem ist eine Methodenkompetenz erforderlich, um statistische und mathematische Methoden anwenden zu können und Datenanalyse-Tools und -Technologien effektiv zu nutzen (Lundborg & Gull, 2021, S. 2 ff). Die Kompetenz im Umgang mit Daten, Datenbanken und Datenmanagement ist ebenfalls von Bedeutung, um die Daten korrekt abzurufen, zu speichern und zu verarbeiten. Zudem benötigen Mitarbeiter eine ausgeprägte Fähigkeit, komplexe Probleme zu identifizieren und zu lösen, um die Herausforderungen bei der Analyse und Interpretation von Produktionsdaten zu bewältigen. Schließlich ist es von großer Bedeutung, dass Mitarbeiter in der Lage sind, die Ergebnisse der Datenanalyse in verständlicher Weise zu kommunizieren und Empfehlungen für Maßnahmen abzuleiten (Blumberg & Kauffeld, 2021, S. 204 ff.; Wangermann, 2020, S. 5 f.; Buchholz et al., 2017, S. 32 ff.).

Weitere Anforderungen an die Entwicklung einer Kollaborationsplattform für kompetenzbasierte Zusammenarbeit in industriellen Datenanalyseprojekten werden detaillierter durch Syberg et al. (2023, S. 64 ff.) beschrieben.

7.3 Konzeptuelle Entwicklung des Leistungsbereichs „Kompetenzen und Handlungsempfehlungen“

Um KMU bei der praxisorientierten Umsetzung von industrieller Datenanalyse effektiv zu unterstützen, ist es notwendig, die technischen und inhaltlichen Rahmenbedingungen festzulegen. Mit dem Projektpartner NEOCOSMO wurde die Entwicklung und Konzeptionierung einer E-Learning-Plattform geplant, um die Plattformnutzer optimal durch das Konzept des situierten und selbstgesteuerten Lernens bei der Aus- und Weiterbildung im Bereich der industriellen Datenanalyse zu fördern. Die Plattform muss dabei über eine intuitive Bedienoberfläche verfügen, die den Lernenden ermöglicht, sich

schnell zurechtzufinden. Im Referenzbaukasten von AKKORD ist diese Plattform im Baustein „Lern- und Kollaborationsplattform“ verankert (siehe Kap. 4). Die Plattform unterstützt insbesondere die peripheren Bausteine zur Durchführung von Industriellen Datenanalysen (siehe Kap. 5 und 6).

Der **Lernbereich** als zentraler Teil der Serviceplattform ist folgendermaßen aufgebaut: Beim ersten Login wird den Nutzenden ein Popup-Fenster angezeigt, in dem sie auf der Plattform begrüßt werden und der Link zur Eingangsbefragung hinterlegt ist. Durch die Absolvierung der Eingangsbefragung soll der IST-Zustand ihrer Datenanalysekompetenzen erfasst werden. Im weiteren Verlauf des Lernprozesses wird den Lernenden die Empfehlung von rollenbezogenen Lernpfaden ermöglicht, die an ihren individuellen Bedürfnissen orientiert sind.

Die einzelnen Kurse und Lektionen sind als Microkurse angelegt und bieten kompakte Einheiten von etwa 90min bis 300min, die im Gegensatz zu anderen Anbietern knappe und für die Nutzenden zwischendurch absolvierbare Kursbausteine zur Verfügung zu stellen. Jeder AKKORD-Kurs besteht aus mehreren Lektionen und Sublektionen, wobei der Schulungsinhalt durch verschiedene auditive (Tonspuren) und visuelle (Bilder, Videos) Unterlegungen unterstützt wird. Des Weiteren ist die Nutzung des tutoriellen Feedbacks innerhalb der Verständnisquizze in den Kurslektionen vorgesehen, um die Lernenden durch adaptive Hilfestellung zusätzlich zu unterstützen (Narciss, 2006, S. 19 ff.). Darüber hinaus soll der Lernfortschritt für die Lernenden ersichtlich sein. Zum Beginn jeder Lernlektion wird ein Einführungsvideo eingebettet, in dem die Lernenden begrüßt und die Lerninhalte dargestellt werden. Nach Abschluss eines rollenbasierten Kurses wird der Lernende gebeten, erneut an einem Fragebogen teilzunehmen, der dieselben Fragen enthält. Dies ermöglicht eine Einschätzung der vermittelten Kompetenzen durch die Lernenden vor und nach dem Kurs (Vorher-Nachher-Vergleich). Um abschließend sein erlerntes Wissen zu überprüfen, kann der Nutzer abschließend an einem rollengebundenen Abschlusstest teilnehmen, bei dessen Bestehen (>50%) ein AKKORD-Zertifikat ausgestellt wird. Zur Abbildung der Lerninhalte und der Abschlussquizze wird dazu die Plugin-Lösung H5P in die Plattform integriert.

Neben dem Lernbereich bietet das Konzept der Serviceplattform noch weitere Funktionen, wie beispielsweise einen **Kollaborationsbereich**, in dem sich die Lernenden untereinander, aber auch mit Experten, austauschen können. Weitere Bestandteile der Plattform sollen Best Practice-Vorgehensweisen der Akkord-Anwendungspartner sein, die als Leitfaden und Orientierung dienen sollen. Dies ist von besonderer Relevanz, da gerade in KMU Datenanalysevorhaben häufig an dem Start eines solchen Vorhabens scheitern. Im Themenbereich sollen Plattformnutzende zusätzlich über aktuell relevante Themen im Bereich Data Science (wie beispielsweise Big Data, Data Mining, etc.) informiert werden. In Abb. 7.1 wird das Grundkonzept der Work&Learn-Plattform noch einmal dargestellt. Im Kap. 13 werden dazu ausführlichere Informationen zur Konzeptionierung des Wissensdienstes und zu den einzelnen Bereichen zur Verfügung gestellt.

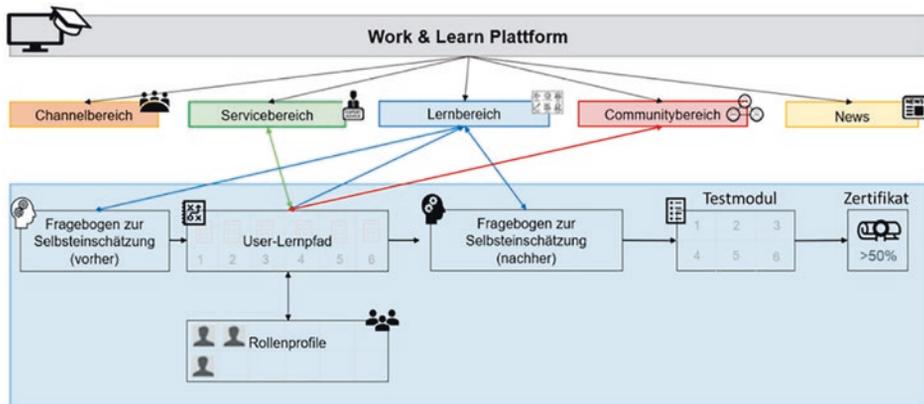


Abb. 7.1 Grundkonzept der Work&Learn-Plattform

7.4 Entwicklung und Implementierung im Bereich „Kompetenzen und Handlungsempfehlungen“

Um die Kompetenzen im Bereich der industriellen Datenanalyse im Rahmen der Work&Learn-Plattform zu erfassen, werden verschiedene Methoden angewendet. Hierzu wird zunächst ein digitaler Fragebogen entwickelt, der an unsere Projektpartner sowie externe Unternehmen gerichtet ist, um die vorhandenen Kompetenzen der Mitarbeitenden in der industriellen Datenanalyse präzise zu erfassen und Feedback von den Unternehmen einzuholen. Darüber hinaus soll der Erhebungsbogen auch die Identifizierung von notwendigem Wissen ermöglichen, das noch aufgebaut werden muss. Konkrete Kompetenzen, die im Fragebogen erfasst werden können, umfassen beispielsweise Kenntnisse in statistischer Datenanalyse, Maschinellem Lernen, Datenvisualisierung oder Datenbankmanagement. Dies ermöglicht eine fundierte Analyse der Kompetenzniveaus des Personals und dient als Grundlage für gezielte Schulungsmaßnahmen in der Work&Learn-Plattform.

Zusätzlich wird ein Screening relevanter Stellenanzeigen zum Thema Data Science durchgeführt, um daraus eine rangierte Anforderungsliste der notwendigen Kompetenzen zu erstellen. Dabei wurden verschiedene Kompetenzbereiche berücksichtigt, wie beispielsweise der Einsatz von Software (z. B. RapidMiner), Fachkompetenz (z. B. Prozessverständnis der Fertigung) sowie soziale Kompetenzen (z. B. Teamfähigkeit). Um die gewonnenen Erkenntnisse zusammenzuführen und daraus konkrete Kompetenzanforderungen abzuleiten, ist es ratsam, die Anforderungen mit den Vorgehensmodellen des *Knowledge Discovery in Databases* (KDD) nach *Fayyad et al. (1996, S. 40 ff.)* und des *Cross-Industry Standard Process for Data Mining* (CRISP-DM) nach *Chapman et al. (2000, S. 13 f.)* zu verknüpfen. Diese Vorgehensmodelle wurden bereits in *Kap. 3* näher erläutert. Auf diese Weise konnten Kompetenzprofile (Rollen) und entsprechende

Lerninhalte generiert werden, die auf die Bedürfnisse von KMU zugeschnitten sind (siehe Abb. 7.2). Durch die Anwendung der beschriebenen Methoden, wie dem Einsatz von Fragebögen oder dem Screening von Stellenanzeigen, können die Instrumente zur Erfassung der notwendigen Kompetenzen für Mitarbeiter und entsprechende Lerninhalte zur Aus- und Weiterbildung abgeleitet werden.

Anhand der erfassten Anforderungen an zu entwickelnde und zu erfassende Kompetenzen wird zunächst ein Basiskurs entwickelt, der auf dem KDD-Ansatz (Knowledge Discovery in Databases) aufbaut, um ein breites Grundlagenwissen zu vermitteln. Ein Zusatzkurs mit weiteren Lektionen ergänzt den Basiskurs und behandelt Themen wie KPIs (Key Performance Indicators), Geschäftsmodelle und Datensicherheit. Aufgrund seines industriellen Bezugs wird das CRISP-DM Modell als Grundlage für den Advancedkurs genutzt, der einen starken Anwendungsbezug durch die Vermittlung von Wissen zu unterschiedlichen Tools gewährleisten soll. Zusätzlich wird eine Best Practice-Lektion mit einem fiktiven Unternehmen und typischen datenanalytischen Fragestellungen konzipiert, um den Nutzenden die Möglichkeit zu geben, eigenständig zu trainieren und ihr Wissen zu vertiefen. Besonderes Augenmerk wurde dabei auf kleine und mittelständische Unternehmen (KMU) gelegt, für die fünf verschiedene Kompetenzrollen und Kurspfade definiert werden: Data Engineer, Data Analyst/ Scientist, Projektmanager, Domänenexperte und Geschäftsführung. Dabei orientiert sich das Rollenkonzept an der Prozesskette von industrieller Datenanalyse (Schwenken et al., 2023, S. 84 f.). Durch die Kombination von Basiskurs, Zusatzkurs, Advancedkurs und Best Practice-Übungen sowie unter Berücksichtigung der verschiedenen Kompetenzrollen und Kurspfade wird ein ganzheitliches und praxisnahes Schulungsprogramm im Bereich Datenanalyse und -management für KMU bereitgestellt (Abb. 7.3).

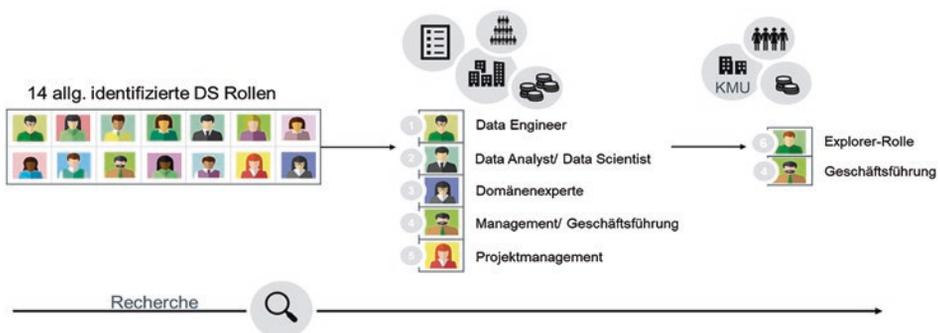


Abb. 7.2 Kompetenzerfordernisse und Identifikation von Kompetenzrollen

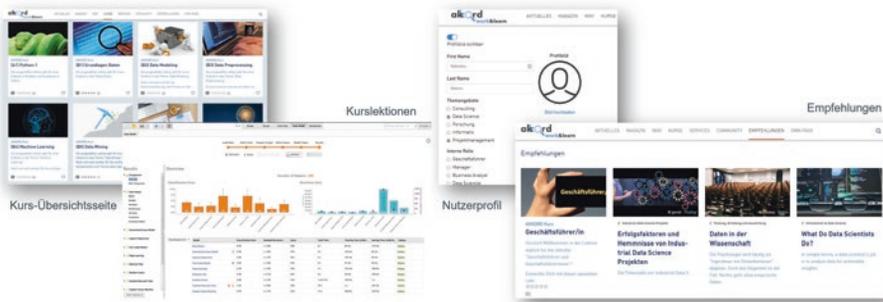
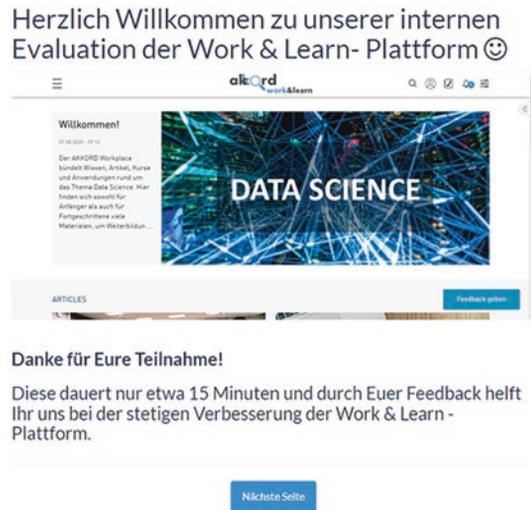


Abb. 7.3 Inhalte aus dem digitalen Wissensdienst

7.5 Validierung und Optimierung

Nach der Konzeptionierung, Entwicklung, Implementierung und Pilotierung des Leistungsbereiches „Kompetenzen und Handlungsempfehlungen“ erfolgt eine konzeptuell methodische sowie technisch funktionale Validierung. Hierfür werden die identifizierten Anforderungsbereiche und Kompetenzen in konkrete und messbare Kompetenzdefinitionen und -bereiche überführt, um entsprechende, technologiebasierte Lernmodule zur industriellen Datenanalyse zu erstellen und als Angebot auf einer modularen und erweiterbaren Work&Learn-Plattform für die Nutzenden bereitzustellen. Um sicherzustellen, dass die entwickelten Lösungen einen praxisorientierten Mehrwert für die Lernenden bieten, werden alle 40 Anforderungen (darunter 7 nicht funktionale und 33 funktionale) des Lasten- und Pflichtenheftes im Rahmen der Ergebnisauswertung iterativ validiert. Eine gängige konzeptionell-methodische Validierungsmaßnahme ist die Inhaltsvalidierung. Hierbei wird überprüft, ob z. B. der Inhalt einer Testaufgabe oder einer Messmethode tatsächlich das misst, was gemessen werden soll (Rammstedt, 2010, S. 239 ff). In unserem Fall erfolgt eine Überprüfung der Kurs- und Testaufgabeninhalte durch interne und externe Experten, um sicherzustellen, dass der relevante Lerninhalt angemessen abgedeckt wird. Zur Überprüfung der umgesetzten Serviceplattform (Wissensdienst) werden auch technisch-funktionale Beurteilungsmaßnahmen (Albers et al., 2016, S. 541 ff.) angewendet, bei der Testnutzer Accounts erhalten, um die einzelnen Funktionen der Plattform ausgiebig zu testen: Hierbei soll sichergestellt werden, dass die Plattform mindestens alle funktionalen Anforderungen nach dem Pflichtenheft erfüllt. Die Testprobanden erhalten während der Durchführung spezielle Anleitungen und Aufgaben, um alle möglichen Funktionen der Plattform zu testen, beispielsweise die Teilnahme an Online-Kursen, die Interaktion mit anderen Nutzern, das Stöbern in den Magazinbeiträgen, die Verwendung von Tools zur Bewertung und Überprüfung des Lernfortschritts. Durch einen verlinkten digitalen Fragebogen auf der Work&Learn-Plattform werden die Probanden nach der Testung aufgefordert, Rückmeldung zum gesamten Funktionsumfang (u. a. Benutzerfreundlichkeit, Lerninhalte, Magazinbei-

Abb. 7.4 Technisch-funktionale Validierungsmaßnahme mittels digitalem Fragebogen



träge, Kollaborationsmöglichkeit, usw.) zu geben. Im Anschluss kann das gesammelte Feedback automatisiert ausgewertet werden, um mögliche Schwachstellen und Verbesserungspotenziale zu identifizieren (Abb. 7.4).

7.6 Zusammenfassung

Viele kleine und mittlere Unternehmen in Deutschland sind derzeit nicht in der Lage, das Potenzial und die Möglichkeiten, die die industrielle Datenanalyse bietet, in ihrer Gänze auszuschöpfen. Dies führt zu einem Versäumnis der Wahrnehmung der Chancen, an innovativen Geschäftsmodellen und Kollaborationen in Wertschöpfungsnetzwerken und der datengestützten Optimierung von Produkten und Prozessen zu partizipieren. Den KMUs mangelt es besonders an den erforderlichen Kompetenzen und Einführungsstrategien. Um eine Möglichkeit zur Unterstützung und zur Deckung des Bedarfs für KMU zu schaffen, wird eine ganzheitliche und kostenlose Serviceplattform entwickelt, die bisher einzigartig ist. Dazu werden nach der Darlegung der Einführung und Motivation die Anforderungen an den Kompetenzaufbau und die -sicherung in Wertschöpfungsnetzwerken analysiert und definiert. Im Ergebnis werden drei Anforderungsbereiche: Allgemeine Anforderungen, Microkurse und Anwenderunterstützung fokussiert, die insgesamt 40 Kriterien beinhalten (z. B. Datenschutz, intuitive Bedienung, verschiedene Kurse, etc.) und in einem Lasten- und Pflichtenheft gebündelt. Im nächsten Abschnitt erfolgt die Konzeption des Bereichs „Kompetenzen und Handlungsempfehlungen“ im Referenzbaukasten, der technische und inhaltliche Perspektiven umfasst. Die technische Seite wird in Zusammenarbeit mit dem Projektpartner Neocosmo umgesetzt, während die inhaltliche Konzeption von Lernkursen,

Tests oder Themenbeiträgen vom Lehrstuhl für Berufspädagogik in den technischen Fächern der Uni Hamburg übernommen wird. Der nächste Abschnitt beschäftigt sich mit der Erfassung der notwendigen Kompetenzen der Lernenden und der Entwicklung der Kompetenzmessinstrumente in Form von Fragebögen und Testaufgaben durch den integrierten H5P-Funktionsbaustein. Auch die Entwicklung der einzelnen Kurse und Lektionen sowie die Umsetzung eines Wissensdienstes als Serviceplattform finden in diesem Kapitel statt. Die Inhalte der Serviceplattform sind dabei so gestaltet, dass sie von anderen Unternehmen generisch verwendet werden können. Schließlich werden die als Mockup entwickelten Kurse, Tests, Beiträge und Kompetenzmessinstrumente in die Umgebung der Work&Learn-Plattform integriert, dort getestet und in iterativen Prozessschritten angepasst. Nach der Konzipierung und Entwicklung des Bereiches „Kompetenzen und Handlungsempfehlungen“ erfolgt die Validierungsphase aus methodisch-inhaltlicher und technisch-funktionaler Perspektive, bei der gegenwärtige und zukünftige Optimierungspotenziale ermittelt werden. Hier ist es notwendig, verstärkt mit externen KMU zu pilotieren, um noch spezifischer deren Anforderungen erfassen und umsetzen zu können. Dazu könnten beispielsweise KI-gestützte Bots verwendet werden, die Handlungsempfehlungen im Kontext der individuellen Datenanalyse ermöglichen. Auch der Einsatz von stärkeren Gamification-Ansätzen spielt bei der zukünftigen Weiterentwicklung unter Berücksichtigung der technischen Umsetzbarkeit eine tragende Rolle, da dadurch die Motivation und Produktivität der Lernenden im Lern- und Arbeitsprozess entscheidend gefördert werden können.

Literatur

- Albers, A., Behrendt, M., Klingler, S., & Matros, K. (Hrsg.). (2016). *Handbuch Produktentwicklung. Verifikation und Validierung im Produktentstehungsprozess* (S. 541–570). Hanser.
- Blumberg, V., & Kauffeld, S. (2021). Kompetenzen und Wege der Kompetenzentwicklung in der Industrie 4.0. *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie*, 52(2), 203–225. <https://doi.org/10.1007/s11612-021-00579-5>.
- Buchholz, B., Ferdinand, J. P., Gieschen, J. H., & Seidel, U. (2017). Digitalisierung industrieller Wertschöpfung. Transformationsansätze für KMU. iit-Institut für Innovation und Technik in der VDI/VDE Innovation + Technik GmbH.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step data mining guide, CRISP-DM consortium.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazin*, 17(3), 37, 37–54.
- Czernich N., Fackler T., Falck O., Schüller S., Wichert S., Keveloh K., & Vijayakumar R. (2019). Digitale Kompetenzen – Ist die deutsche Industrie bereit für die Zukunft? Ifo institut.
- Haarmeier, M. (2021). *Künstliche Intelligenz für den Mittelstand*. Springer Fachmedien Wiesbaden.
- Lundborg, M., & Gull, I. (2021). Künstliche Intelligenz im Mittelstand. So wird KI für kleine und mittlere Unternehmen zum Game Changer. Begleitforschung Mittelstand-Digital. Bad Honnef.

- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Waxmann.
- Rammstedt, B. (2010). Reliabilität, Validität, Objektivität. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (Bd. 1, S. 239–258). VS Verlag.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process Chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and Systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- Seyda, S. (2012). Fachkräfte sichern. Hg. v. Bundesministerium für Wirtschaft und Technologie (BMWi).
- Statistisches Bundesamt. (2019). Kleine und mittlere Unternehmen. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Kleine-Unternehmen-Mittlere-Unternehmen/_inhalt.html. zuletzt aktualisiert am 12.08.2019, zuletzt geprüft am 12.03.2023.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- Trunzer, E., Weiß, I., Pötter, T., Vermum, C., Odenweller, M., Unland, S., et al. (2019). Big Data trifft Produktion. *atp*, 61(1–2), 90–98. <https://doi.org/10.17560/atp.v61i1-2.2394>.
- Wangermann, T. (2020). KI in KMU. Rahmenbedingungen für den Transfer von KI-Anwendungen in kleine und mittlere Unternehmen. Konrad-Adenauer-Stiftung (Analysen & Argumente, 381).
- Weskamp, M., Tamas, A., Wochinger, T., & Schatz, A. (2014). Einsatz und Nutzenpotenziale von Data Mining in Produktionsunternehmen. Fraunhofer IPA.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Entwicklung neuer Kollaborationsmöglichkeiten und Geschäftsmodelle

8

Vorstellung der entwickelten Konzepte und
Lösungsbausteine, sowie einer Entscheidungshilfe
zu ihrer Verwendung

Michael Frey , Andreas Emrich , Rebekka Adams und
Volker Zimmermann

8.1 Motivation

Data-Science-Anwendungen im betrieblichen Kontext benötigen für ihr Gelingen die Einbindung einer Vielzahl von Akteuren: Data Scientists, Domänenexperten und Entscheider im Unternehmen (Zhang et al. 2020 S. 4ff; Syberg et al., 2023, S. 64 ff.). Nicht zuletzt die resultierenden Data-Science-Systeme sind dabei ein nicht-menschlicher Akteur. Es geht darum, Verständnis zwischen diesen Akteuren zu generieren. Grundsätzlich benötigt jede der Phasen des CRISP-DM-Modells (Chapman et al., 2000, S. 1 ff) Unterstützung durch Kollaboration. Mit seinen Phasen des „Business Understanding“ und „Data Understanding“ wird deutlich, an welchen Stellen die Akteure interagieren müssen. Auch bei dem Schritt der „Evaluation“ sollte nicht nur nach rein algorithmischen und technischen Aspekten bewertet werden, sondern es sollten auch die initialen Unternehmensanforderungen

M. Frey (✉) · A. Emrich
Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, Deutschland
E-Mail: michael.frey@dfki.de

A. Emrich
E-Mail: andreas.emrich@dfki.de

R. Adams · V. Zimmermann
NEOCOSMO GmbH, Saarbrücken, Deutschland
E-Mail: rebekka@neocosmo.de

V. Zimmermann
E-Mail: volker@neocosmo.de

abgeglichen werden. Gerade kleine und mittlere Unternehmen stehen oft vor dem Problem, dass sie nicht die Kapazitäten haben, um die Durchführung von Datenanalyseprojekte eigenständig stemmen zu können. Gerade die Freistellung der benötigten Mitarbeiter für ein solches Projekt über die interne Wertschöpfungskette hinweg, kann bei kleineren Unternehmen schnell zu Verzögerungen in der Produktion und somit Verdienstausschlag führen. Eine unternehmensübergreifende Kooperation verteilt dabei die Belastung auf mehrere Unternehmen und erlaubt es auch KMU größere Datenanalyseprojekte durchzuführen.

Daten werden häufig als der Treibstoff der neuen Ökonomie gesehen und ermöglichen somit neue Geschäftsmodelle (Brownlow et al., 2015, S. 14). Sie können dabei gleichzeitig als Enabler für disruptive Geschäftsmodelle dienen, wie z. B. der Verkauf von Scheibenwischerdaten aus Autos für Wetterdienste, als auch die bestehenden Geschäftsmodelle im Sinne einer kontinuierlichen Geschäftsmodellentwicklung unterstützen (Bartos et al., 2019, S. 1 ff). Aggregiertes Feedback aus sozialen Medien kann Rückschlüsse auf gewünschte Produktfeatures oder Unzufriedenheit mit dem Service geben, während aus Nutzungsdaten von Produkten Rückschlüsse auf die Funktionstüchtigkeit des Produkts selbst abgeleitet werden können. Aus Sicht der Geschäftsmodellentwicklung kann so datengetrieben priorisiert werden, wo im Sinne einer Weiterentwicklung eines Geschäftsmodells die dringendsten Handlungsbedarfe bestehen. Da sie den Aufwand und die damit verbundenen Kosten für die kontinuierliche Weiterentwicklung bestehender Geschäftsmodelle reduzieren, eignen sich datengetriebene Ansätze besonders für kleine und mittlere Unternehmen. Letztere haben in der Regel weder die Kapazität noch das Fachwissen solche Projekte eigenständig durchzuführen und sind somit meist auf teure Beratungsdienstleistungen angewiesen.

Auch für die datengetriebene Geschäftsmodellentwicklung müssen die gleichen vorgenannten Akteure eingebunden werden und zusätzlich die Zulieferer und Vertriebspartner. Je nach bestehendem Geschäftsmodell kann es sogar relevant sein, die Endkunden passiv oder aktiv mit einzubeziehen. So können auftretende Datenphänomene besser analysiert und eingeordnet werden und Änderungen im Geschäftsmodell den jeweiligen Akteuren besser zugeordnet werden. Es ist also von entscheidender Bedeutung, den genannten Akteuren eine Plattform zu bieten, die einen konstruktiven Austausch zu den jeweiligen Analysen erlaubt.

Im Rahmen des Leistungsbereichs wurden im Projekt AKKORD (siehe Kap. 1 und 2) Konzepte und Werkzeuge zur Kollaboration für Akteure in den Bereichen Data Science und datengetriebener Geschäftsmodellentwicklung umgesetzt. Kollaborationswerkzeuge ermöglichen die gemeinsame Analyse und die punktuelle Schulung von Stakeholdern zu wichtigen Data-Science-Grundlagen. Die Werkzeuge zur kollaborativen Geschäftsmodellentwicklung erlauben die Analyse von Social-Media-Daten auf Basis von Natural Language Processing (NLP)-Verfahren. Dabei handelt es sich um eine Sammlung von Anwendungen zur computergesteuerten Verarbeitung von Sprache um spezifische Erkenntnisse zu gewinnen. Sie dienen dazu, positive und negative Bereiche der einzelnen Geschäftsmodellkomponenten zu visualisieren und Bearbeitern das Editieren des Ge-

schäftsmodells aufgrund entsprechender Empfehlungen zu ermöglichen. Folgende Anforderungen können daraus abgeleitet werden:

- Eine Lösung zur Unterstützung der Kollaboration für Data-Science-Projekte muss es Stakeholdern ermöglichen sich sowohl spezifisch zu einzelnen Projekten aber auch generell zu Data-Science-Methoden auszutauschen und relevante Kontakte zu identifizieren.
- Sie muss zudem eine zentrale Wissensquelle für die Durchführung von Data-Science-Projekten bieten, die allen Mitwirkenden eine gemeinschaftliche Basis zu Methoden und Terminologie bietet.
- Um Datenanalyseprojekte zur Entwicklung oder Erweiterung von Geschäftsmodellen in einem kooperativen Umfeld durchzuführen muss ein Konzept aufzeigen, wie in Data-Science-Projekten gewonnene Erkenntnisse auf Aspekte eines Geschäftsmodells abgebildet werden können, um so schließlich eine Anpassung des Geschäftsmodells durchzuführen.

8.2 Kollaboration für Data-Science-Projekte

Ein zentraler Aspekt ist das Konzept der Kollaboration in industriellen Datenanalyseprojekten und der Geschäftsmodellentwicklung. Wie eingehend erwähnt geht es dabei vor allem darum, den einzelnen Akteuren die Möglichkeit zu geben, ihr Wissen anzuwenden, es zu erweitern und mit anderen Akteuren zu teilen. Das in AKKORD entwickelte Konzept zum kollaborativen Arbeiten sieht folgende Punkte vor, die vom *Knowledge Management Cycle* abgeleitet wurden (Evans et al., 2014, S. 91 ff):

- **Wissensaustausch:** Den Akteuren muss ein Werkzeug bereitgestellt werden, mit dessen Hilfe sie sich zum Projekt austauschen können. Dieser Austausch kann auf technischer Ebene oder konzeptioneller Ebene nötig sein.
- **Wissensaneignung:** Den Akteuren muss Lehrmaterial bereitgestellt werden, mit dem sie ihr Wissen bei Bedarf so erweitern können, wie es für das zugrunde liegende Projekt relevant ist. Des Weiteren ist es hilfreich, wenn Ressourcen bereitgestellt werden, die bereits durchgeführte, ähnliche Projekte thematisieren, um somit die vorhandenen Potenziale hervorzuheben.
- **Wissensfindung:** Den Akteuren muss ein Werkzeug bereitgestellt werden, das es ihnen ermöglicht herauszufinden, wer potenziell über das nötige Wissen verfügt, und mit diesen Experten in Kontakt zu treten.
- **Wissensanwendung:** Den Akteuren muss ein Werkzeug bereitgestellt werden, dass es ihnen erlaubt, ihr Wissen zielführend einzubringen und so kollaborativ das Projekt durchzuführen.

Im Folgenden wird erläutert, wie das Konzept durch die in AKKORD entwickelte Work&Learn-Plattform umgesetzt werden konnte.

8.2.1 Nutzungskonzept

Um die Anforderungen, welche für die Kollaborationsmöglichkeiten in der AKKORD Work&Learn-Plattform bestehen, besser nachvollziehen zu können, wurden in Kap. 7 beispielhaft für potenzielle Nutzer der Plattform stehende Personas kreiert (Schwenken et al., 2023, S. 81 ff.), welche sich im Bereich Industrielle Datenanalyse und Data Science weiterbilden bzw. den Austausch mit weiteren Themenbereichs-Interessierten und Fachexperten suchen (West et al., 2021, S. 131). Dadurch konnte verdeutlicht werden, dass einige, aber nicht alle Nutzer der Plattform einen IT-Hintergrund haben werden. Außerdem ist davon auszugehen, dass eine ganze Reihe von Nutzern sich bisher noch nicht ausführlich mit dem Themenbereich „Industrielle Datenanalyse und Data Science“ beschäftigt haben.

Für die Kollaboration im Rahmen der entwickelten Online-Plattform konnte abgeleitet werden, dass schnelle und einfache Kommunikationswege von großer Wichtigkeit sind und von potenziellen Plattformnutzern und Stakeholdern bevorzugt werden. Über das umfangreiche Lernangebot hinaus besteht Bedarf an Kontaktmöglichkeiten zu Experten, die im Betrieb bzw. der Organisation des jeweiligen Nutzers unter Umständen nicht oder nur marginal gegeben sind. Aus diesem Grund sollten Fragen direkt auf der Plattform gestellt werden können und die Möglichkeit zur Kontaktaufnahme zu Fachexperten gegeben sein. Auf dieser Grundlage wurden passende Kollaborationsfeatures konzipiert, die im Abschn. 8.2.2 genauer dargestellt werden.

Die im Rahmen des Projekts konzipierte und entwickelte Work&Learn-Plattform für den Themenbereich Industrielle Datenanalyse und Data Science beinhaltet neben einem Lern-, einem Magazin-, einem Wiki- und einem News-bzw. Channelbereich den für die Kollaboration zentralen Community-Bereich sowie den Servicebereich, über den Nutzer Informationen zu den Best Practices, den Projektpartnern, den AKKORD Baustein-Inhalten und Success Story-Seiten der einzelnen Projektpartner erhalten. In den Beiträgen zu den AKKORD Baustein-Inhalten sind, die im Rahmen des Projekts entwickelten Bausteine verlinkt. Diese können von den Nutzern somit direkt verwendet werden, um spezifische Datenanalysen durchzuführen und somit ihr **Wissen anzuwenden**. Eine genauere Beschreibung des Aufbaus der Lernplattform ist in Kap. 7 zu finden.

Um darüber hinaus möglichst vielen Interessenten die Möglichkeit zu geben, das Lern- und Kollaborationsangebot in der digitalen AKKORD Work&Learn-Plattform zu nutzen, steht die Plattform als Open Source-System zur Verfügung. Eine Anmeldung wird per Azure-Account oder per Selbstregistrierung ermöglicht. Es gelangen somit nicht nur Themeninteressierte leichter auf die Plattform, sondern auch Fachexperten, durch deren Wissen und Expertise der Mehrwert der Kollaboration auf der Work&Learn-

Plattform steigt. Nutzer mit einem Account auf der Plattform können so beispielsweise Kommentare verfassen, ihr Profil anlegen und bearbeiten bzw. im Benutzerverzeichnis aufgelistet werden und sich mit anderen Nutzern der Plattform in Social Streams austauschen.

8.2.2 Kollaborations-Werkzeuge

Ein zentraler Aspekt, um kleinen und mittleren Unternehmen die Durchführung von Datenanalyseprojekten zu ermöglichen, ist die Verbesserung des unternehmens- und organisationsübergreifenden Austausches. Neben dem Lern- und Wissensbereich sind deshalb der Community-Bereich sowie weitere integrierte Kollaborationswerkzeuge ein weiterer zentraler Bereich der Work&Learn-Plattform. Die Kommunikation zwischen den Stakeholdern wird somit durch die Work&Learn-Plattform sichergestellt, auf der sich die Beteiligten austauschen und ihre Expertise teilen können. Beispielsweise können sich unter verschiedenen anderen Datenanalysten, Change-Manager und Domänenexperten vernetzen und austauschen. Im Folgenden wird ein genauerer Einblick in die Kollaborationswerkzeuge gegeben.

Ein wichtiges Kollaborationswerkzeug in der Work&Learn-Plattform ist die Kommentarfunktion. Jeder Nutzer der Plattform hat die Möglichkeit Kurslektionen und -sublektionen, Magazinartikel, Beiträge des Wikis und Service-Seiten zu kommentieren. Jeder Kommentar kann „geliked“ und erneut kommentiert werden, sodass ein geordneter Austausch zu den Plattform-Inhalten entstehen kann. Durch die Kommentarfunktion wird den Nutzern die Möglichkeit gegeben, sich zu allen Beiträgen der Plattform zu äußern oder Fragen mit direktem Inhaltsbezug zu stellen. Kommentare werden in der Marginalspalte des jeweiligen Beitrags angezeigt und sind in allen Bereichen der Plattform gleich aufgebaut, um ein konsistentes Aussehen und Bedienen für die Nutzer zu ermöglichen. Über den Aspekt hinaus, dass die Nutzer über die Kommentarfunktion in einem themenspezifischen Austausch treten können, können die Autoren einzelner Inhalte die Kommentare gleichzeitig als Feedback nutzen. Dadurch können eine fortlaufende Verbesserung und Optimierung der Inhalte der Work&Learn-Plattform ermöglicht werden.

Eine weitere Grundlage, um in der Work&Learn-Plattform kollaborieren zu können bzw. Kontakte zu knüpfen, bietet das Benutzerverzeichnis, das eine Übersicht der auf der Plattform registrierten Nutzer liefert. Es ergibt sich entsprechend aus deren Profilen. Jeder Nutzer kann in seinem Profil Angaben über sich als Person, Kontaktdaten und die eigene Expertise machen. In der Übersicht des Benutzerverzeichnisses werden diese Informationen erkenntlich, sodass sich zu verschiedenen Themen schnell und unkompliziert Ansprechpersonen bzw. Experten finden lassen. Haben Nutzer darüber hinaus Kontaktinformationen in ihrem Profil angegeben, können sie dem Verzeichnis entnommen und für die Kontaktherstellung genutzt werden. Das Benutzerverzeichnis ebnet dadurch den Weg für einen gezielten Austausch der Nutzer untereinander und für die

Gabe von fachspezifischen Hilfestellungen. Es adressiert somit den im Konzept eingeführten Aspekt der **Wissensfindung**. Im Verzeichnis lassen sich außerdem Favoriten markieren, welche an erster Stelle für den Nutzer angezeigt werden. Somit wird dem Nutzer die Möglichkeit gegeben, für ihn bzw. sie persönlich wichtige Kontakte hervorzuheben. Gerade dies ist ein Alleinstellungsmerkmal der entwickelten Lösung, das es Nutzern erlaubt, sich ein Portfolio von Expertenkontakten anzulegen und diese direkt auf bereits durchgeführte Projekte und den jeweiligen Beitrag zu beziehen.

Die Wiki-Wissensdatenbank (kurz: Wiki), die eine Kollektion von Beiträgen zu einem bestimmten Thema darstellt, ist eine weitere Rubrik in der Work&Learn-Plattform. Durch die Verortung in der Hauptnavigation der Plattform ist sie für Nutzer gut sichtbar platziert. Diese Rubrik bietet dem Nutzer die Möglichkeit, aktiv inhaltlich mitzuwirken und fachlichen Input zu bestimmten Themen einzubringen. So kann der Nutzer Einträge im Wiki zu verschiedenen Schlagworten rund um die Themen Industrielle Datenanalyse und Data Science erweitern, und Verknüpfungen zu weiteren Wiki-Beiträgen durch Verlinkungen herstellen. In der AKKORD Work&Learn-Plattform eignet sich das stets wachsende Netzwerk an Inhalten somit als Nachschlagewerk für Definitionen oder für Erklärungen rund um die zentralen Themenbereiche. Insgesamt fördert das Wiki ein kollaboratives Arbeiten und passt sich durch die offene Administration an die Bedürfnisse der Nutzer an. Eine alphabetische Sortierung der Beiträge fördert Übersichtlichkeit und Struktur auf der Wiki-Übersichtsseite. Zusätzlich zu den bereitgestellten, themenbezogenen Lerninhalten bietet das Wiki eine Möglichkeit der **Wissensaneignung**.

Ein gruppenbezogener Austausch wird im Community-Bereich ermöglicht, auf welchen Nutzer zentral über den Hauptnavigationspunkt "Community" zugreifen können. Im Community-Bereich wurden Themen- und Kompetenzrollenbezogene Social Streams angelegt, die einen Raum zum themenspezifischen Austausch bieten. So können Nutzer beispielsweise in der Gruppe "Changemanagement" themenspezifische Fragen klären und sich mit weiteren Stakeholdern vernetzen. Im Falle eines laufenden Projekts kann zudem eine projektspezifische Gruppe erzeugt werden, welcher dann den direkten Austausch der einzelnen Stakeholder ermöglicht. Darüber hinaus wurde ein Social Stream angelegt, in welchem sich zu generellen Fragen ausgetauscht werden kann. Nutzer können somit den für sie relevanten Gruppen beitreten und themenspezifische Fragestellungen und Anregungen über Social Streams diskutieren. Ihnen wird die Möglichkeit gegeben, Beiträge zu verfassen, die wiederum von anderen Nutzern „geliked“ werden können und untereinander aufgelistet werden. Auf diese Art und Weise können Fragen und Themen schnell und unkompliziert angesprochen bzw. beantwortet werden. Anders als Kommentare können Posts im Social Stream mit Links, Bildern oder Dokumenten versehen werden, um so einen besseren Austausch besonders in Datenanalyseprojekten zu ermöglichen. Der Community Bereich adressiert somit den Punkt des **Wissensaustauschs** des Konzepts zum kollaborativen Arbeiten. Im Gegensatz zu vorhandenen Angeboten ermöglicht dies sowohl einen auf ein spezifisches Projekt fokussierten als auch einen bereichs- oder unternehmensübergreifenden Austausch von Stakeholdern.

8.3 Datengetriebene Geschäftsmodellverbesserung

Es gibt in der Literatur eine Vielzahl von Definitionen für den Begriff Geschäftsmodell (Scheer et al., 2003, S. 7 ff). Im Kern geht es darum, welche Werte in einem Unternehmen aus welchen Ressourcen geschaffen werden und wie diese an wen vermarktet werden. Optimalerweise sollte ein Geschäftsmodell dabei flexibel genug gestaltet sein, dass es auf externe und interne Impulse reagieren und somit stetig weiterentwickelt werden kann (Simmert et al., 2019, S. 456 ff. Wie in Kap. 3 erwähnt wurde, ist ein Ziel des Leistungsbereichs die Entwicklung eines Konzeptes zur automatischen datengetriebenen Verbesserung von bestehenden Geschäftsmodellen. Gerade kleine und mittlere Unternehmen profitieren davon, diese oft Zeit- und Ressourcenintensiven Aktivitäten zumindest teilautomatisieren zu können. Im Folgenden wird dieses Konzept erläutert. Eine prototypische Umsetzung des Konzeptes erfolgte in einer der Erfolgsgeschichten von AKKORD (siehe Kap. 15).

8.3.1 Nutzungskonzept

Voraussetzung zur datengetriebenen Anpassung des Geschäftsmodells ist die Bewertung des bestehenden Modells anhand konkreter Gesichtspunkte und deren Wirkung. Der Business Model Canvas (Osterwalder and Pigneur 2010, S. 14 ff) beschreibt ein Modell aus Blöcken, die verschiedene Aspekte von Unternehmen und deren strategische Ausrichtungen logisch aufteilen und darstellen. Diese Blöcke beschreiben zum Beispiel die relevanten Partner, benötigten Ressourcen, Schlüsselaktivitäten, aber auch die kundengerichtete Sicht in Bezug auf Vertriebskanäle, Kundenbeziehungen und Kundensegmente. Eine solche Unterteilung vereinfacht die Bewertung einzelner Bereiche und ermöglicht es so, Bereiche des bestehenden Modells mit möglichem Verbesserungspotenzial zu identifizieren. Durch die unterschiedlichen Geschäftsmodelle der Partner in AKKORD konzentrieren sich die im Projekt erstellten Konzepte auf die Kundensicht des Geschäftsmodells, wodurch eine homogene Datenbasis in Form von Bewertungstexten genutzt werden kann. Die folgende Grafik (Abb. 8.2) zeigt den Business Model Canvas im Kontext der Sichten auf das Geschäftsmodell. Die linke Seite zeigt die interne Sicht des Unternehmens. Hier anfallende Daten beinhalten primär Informationen über Produktionsprozesse oder die Supply-Chain. Die interne Sicht hat direkte Auswirkungen auf die externe Sicht, da sich Änderungen dort in der Regel auf das Produkt und somit auf die Verkaufsstrategie oder direkt auf den Kunden auswirken. Jedoch können auch in der externen Sicht relevante Daten anfallen, welche sich schlussendlich auf das Geschäftsmodell auswirken können. Dazu zählt insbesondere das Feedback des Endkunden (Abb. 8.1).

Neben der Analyse der Kundenmeinung lassen sich jedoch besonders aus in Produktionsprozessen anfallenden Daten Rückschlüsse auf den Geschäftsmodellbereich Schlüsselaktivitäten ziehen. Bestehen Schwierigkeiten über mehrere Prozesse hinweg können einerseits die Prozesse angepasst werden, andererseits könnte es sich jedoch

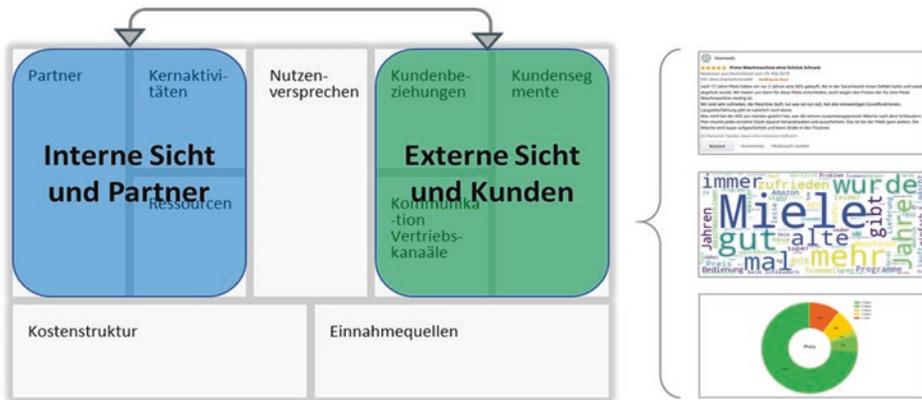


Abb. 8.1 Schematische Darstellung der Bewertung

auch um ein strategisches Problem handeln, das nicht auf Prozessebene adressiert werden sollte.

Um im Rahmen des Verbundprojekts AKKORD Daten aus Unternehmensprozessen hinsichtlich des bestehenden Geschäftsmodells bewerten zu können, wurde eine Ontologie erstellt, mit deren Hilfe eine Zuordnung von einzelnen Datenfeldern einer XES Datei, der de facto Standard für Process Mining Daten, hin zu einem Geschäftsmodell ermöglicht wird. Abb. 8.2 zeigt diese Ontologie. Die Ontologie wurde anhand eines Datensatzes definiert, durch die eine grobe Einteilung der Daten hin zu einem Geschäftsmodellbereich möglich wird. Dabei wurde im Besonderen die Art der Daten, also ob es sich um numerische oder textuelle Informationen handelt, berücksichtigt. In einem zweiten Schritt wurde die erste Version der Ontologie zu einer allgemeinen Version weiterentwickelt, die auf Basis von mehreren Datensätzen erstellt wurde. Hierbei wurden Fallunterscheidungen der ersten Version verallgemeinert.

8.3.2 Lösungsarchitektur

Es wurde eine generische Architektur zur datengetriebenen Geschäftsmodellentwicklung erstellt (siehe Abb. 8.3). Die *Datenebene (Data Layer)* umfasst dabei sowohl prozessnahe Daten, die direkt aus Fertigungsprozessen extrahiert werden können oder bereits über ein ERP-System aggregiert wurden. Als weitere Datenquellen inkludiert die entwickelte Architektur die Möglichkeit von web-basierten Schnittstellen oder Crawlern. Diese dienen in erster Linie dazu die letzte Meile des Wertschöpfungsnetzwerks abzugreifen. Ein spezieller Fall ist der in Kap. 15 porträtierte Feedbackmechanismus der Kundenbewertung, wie er z. B. häufig beim Onlineversandhandel eingesetzt wird. Um die vorhandenen Datenquellen semantisch miteinander und dem Geschäftsmodell zu verknüpfen, beinhaltet die vorgestellte Architektur zudem, wie oben erwähnt, eine oder

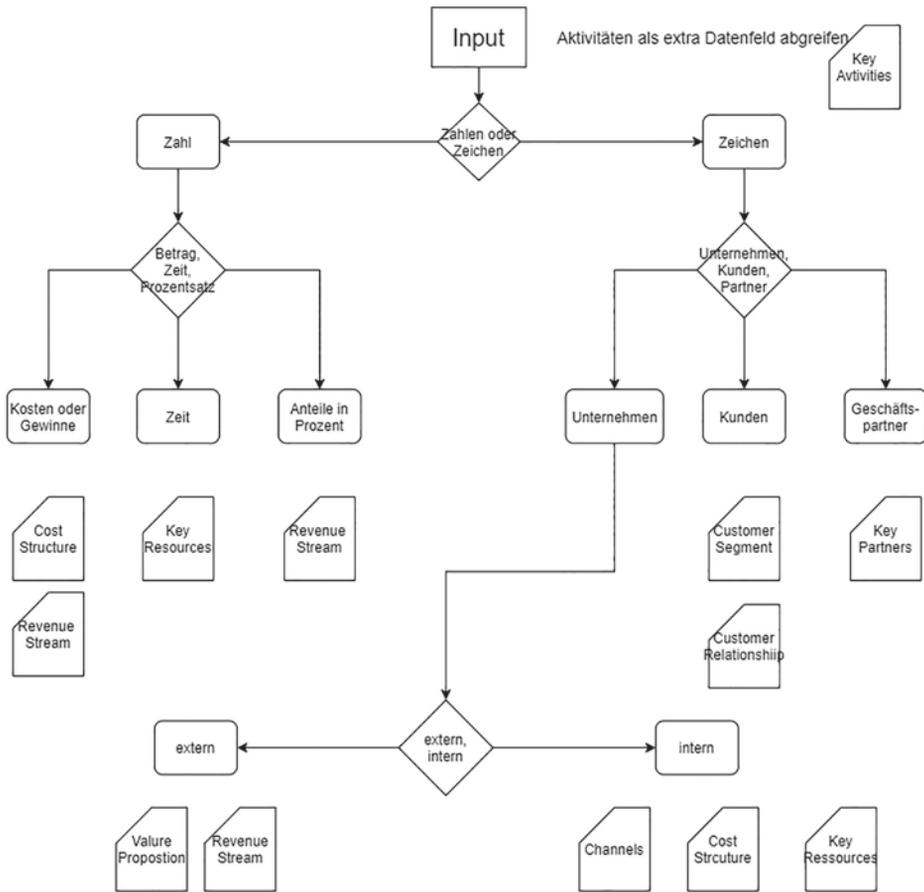


Abb. 8.2 Allgemeine Ontologie zum XES Mapping auf Geschäftsmodelle

mehrere Ontologien. Diese erlauben es logische Zusammenhänge zwischen Entitäten maschinenlesbar abzuspeichern, um so während der Datenakquise als auch bei der folgenden Analyse Beziehungen zwischen einzelnen Elementen zu berücksichtigen.

Hierzu wurden Verknüpfungen zwischen Geschäftsmodellen und Produktionsdaten identifiziert, um so die Integration im Hinblick auf die Analysemodule zu ermöglichen. Im Besonderen wurde die Prozessebene als optimale Schnittstelle identifiziert, da hier die benötigten Verbindungen zu Geschäftsmodellen vorhanden sind und sich Prozesse aus aggregierten Unternehmensdaten erzeugen lassen.

In der *Analyseebene (Analytics Layer)* werden auf den eingehenden Prozessdaten klassische Mining-Operationen des Process Mining durchgeführt. Hierbei werden aus zeitlich sortierten Log-Daten Schlussfolgerungen über den zugrunde liegenden Prozess gezogen und mit vorhanden Prozessmodellen abgeglichen. Die aus dem Web extrahierten Informationen zu Kunden- und Produktbewertungen werden zunächst durch diverse

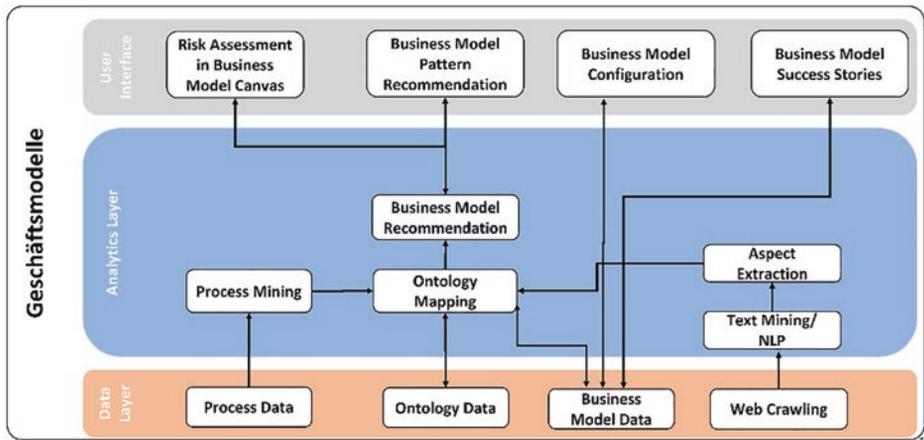


Abb. 8.3 Architektur Geschäftsmodellentwicklung

Methoden des Natural Language Processing (Text Mining) vorverarbeitet. Im Anschluss werden die in den Bewertungen identifizierten Aussagen über das jeweilige Produkt oder den Hersteller bzw. dessen Service extrahiert und zusammen mit den aus dem Prozess gewonnen Erkenntnissen unter Verwendung der semantischen Verknüpfungen der zugrunde liegenden Ontologien auf die jeweiligen Aspekte des Prozesses bzw. des Geschäftsmodells projiziert.

Die durch die Analysen erzeugten Erkenntnisse werden den Nutzern über diverse Nutzerschnittstellen (User Interface) präsentiert. Dazu zählen in erster Linie eine Risikoeinschätzung für das Geschäftsmodell, dargestellt über den Business Model Canvas, und Änderungsempfehlungen für das Geschäftsmodell unter Verwendung von etablierten Geschäftsmodellmustern. Um den Nutzern weiteres Hintergrund- und Referenzmaterial zu liefern, werden zudem Möglichkeiten der Geschäftsmodellkonfiguration und Erfolgsgeschichten zu diversen Geschäftsmodellen angeboten.

8.3.3 Geschäftsmodellanalyse-Bausteine

Der im Prototypen umgesetzte Geschäftsmodellanalyse-Baustein ermöglicht die Bewertung eines bestehenden Geschäftsmodells auf Basis von klassischen Kundenbewertungen wie man sie von gängigen Verkaufsplattformen kennt. Zu diesem Zweck wurde eine Taxonomie erstellt, welche es ermöglicht Aspekte von Kundenbewertungen, die mit Hilfe von NLP-Methoden (Natürliche Sprachverarbeitung) extrahiert wurden, auf Geschäftsmodellaspekte des Business Model Canvas abzubilden. Eine anschließende Sentimentanalyse bewertet den jeweiligen Aspekt dann mit der Stimmung des Verfassers.

Dies führt nach durchgeführter Analyse dazu, dass die abgebildeten Geschäftsmodellaspekte eine kumulative Bewertung auf Basis der Stimmung der Bewertenden erhalten. Dieses Wissen ermöglicht es Entscheidern, Anpassungen am bestehenden Geschäftsmodell durchzuführen, um negative Aspekte zu verbessern.

Die Datenakquise wird dabei wie in der obigen Architektur aufgeführt mithilfe von Crawlern durchgeführt. Da diese Art der Anbindung jedoch sehr unzuverlässig ist, können bereits kleinste Änderungen am Layout der Webseite dazu führen, dass keine Daten mehr erfasst werden. Der in RapidMiner Studio umgesetzte Baustein (siehe Abb. 8.5) verwendet daher einen Datei-Upload, um die benötigten Daten anzubinden. Die modularen Analysen stellen also einen eigenen Durchlauf der **Prozesskette der industriellen Datenanalyse** dar (siehe Kap. 4) und nutzen dabei die Ergebnisse von Kap. 6. Ebenso ist es möglich die verwendete Taxonomie über einen Datei-Upload auszutauschen, um die Analysen an die jeweilige Domäne des Vorhabens anzupassen. Das Analyseergebnis wird in Form einer Tabelle dargestellt, welche Geschäftsmodellaspekte ihrer kumulierten Stimmungsbewertung gegenüberstellt (Abb. 8.4).

Detaillierte Informationen über die prototypische Implementierung des Konzepts zur datengetriebenen Geschäftsmodellverbesserung finden sich in Kap. 15).

Das vorgestellte Konzept zur Erweiterung von Geschäftsmodellen ist stark abhängig von dem jeweiligen Anwendungsfall, da insbesondere die Taxonomie sehr abhängig von der Domäne ist, in welcher das System eingesetzt wird. Besonders für die Anwendung zur Analyse von Prozessdaten zur Weiterentwicklung von Geschäftsmodellen, sind größere Anpassungen nötig. Das in Kap. 15 vorgestellte System ist dabei als prototypische Validierung des Ansatzes anhand eines speziellen Anwendungsfalls zu sehen. Es wurde somit gezeigt, dass der Ansatz machbar ist und der daraus entstandene Prototyp kann als Ausgangspunkt für eine weitere Entwicklung verwendet werden.



Abb. 8.4 Abbildung von Inhalten im Text auf Geschäftsmodellaspekte

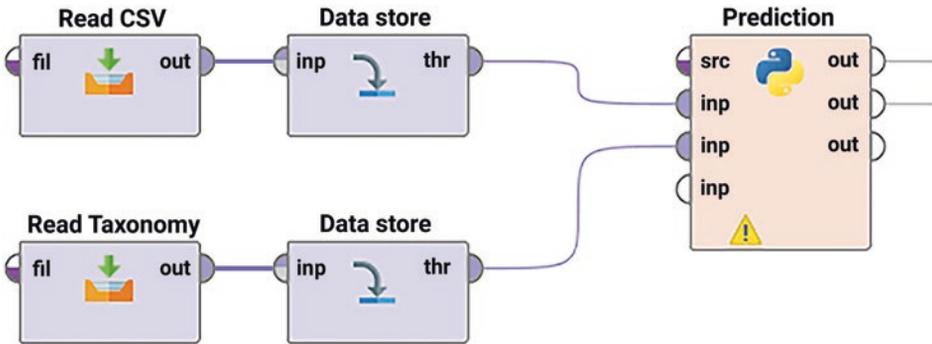


Abb. 8.5 Geschäftsmodellanalyse-Baustein in RapidMiner Studio

8.4 Zusammenfassung

Im Leistungsbereich wurde zum einen ein Konzept zur Durchführung von kollaborativen Data-Science-Projekten entwickelt, zum anderen wurde ein Konzept zur datengetriebenen Geschäftsmodellentwicklung erstellt.

Die im Konzept zur kollaborativen Data-Science aufgestellten Konzepte konnten in Zusammenarbeit mit dem Leistungsbereich „Kompetenzen und Handlungsempfehlungen“ erfolgreich in die entwickelte Work&Learn-Plattform integriert werden. Eine mögliche Weiterentwicklung der Integrationen könnte eine engere Anbindung an die durchgeführten Datenanalyseprojekte sein, indem die in der Work&Learn-Plattform vorhandenen Communities und Kommentarfunktionen direkt in Datenanalysetools integriert werden.

Das Konzept zur datengetriebenen Geschäftsmodellverbesserung demonstriert eine Architektur und Vorgehensweise, wie sich diverse Datenströme im Unternehmen auf Geschäftsmodellaspekte abbilden lassen, um somit eine Bewertung dieser zu erstellen. Dies bietet eine datengetriebene Entscheidungsgrundlage zur späteren Adaption des Geschäftsmodells.

Der im Leistungsbereich durchgeführte Use-Case zur datengetriebenen Geschäftsmodellentwicklung zeigt das Potenzial auf, das selbst bei der Analyse von relativ unstrukturierten Daten für die Geschäftsmodellentwicklung existiert. Eine mögliche Erweiterung der im Leistungsbereich entstandenen Ergebnisse wäre die Verwendung von Daten aus der Produktion oder Supply-Chain zur Bewertung weiterer Aspekte des Geschäftsmodells, so wie es das entwickelte Konzept vorsieht.

Die im Leistungsbereich entwickelten Lösungen werden im Rahmen der Projektergebnisse zur Verfügung gestellt. Eine Veröffentlichung der Ergebnisse im akademischen Umfeld ist bereits in Arbeit.

Literatur

- Bartos, M., Park, H., Zhou, T., et al. (2019). Windshield wipers on connected vehicles produce high-accuracy rainfall maps. *Science and Reports*, 9, 170.
- Brownlow, J., Zaki, M., Neely, A., & Urmetzer, F. (2015). Data and analytics-data-driven business models: A blueprint for innovation. *Cambridge Service Alliance*, 7, 1–17.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step data mining guide, CRISP-DM consortium.
- Evans, M., Dalkir, K., & Bidian, C. (2014). A holistic view of the knowledge life cycle: The Knowledge Management Cycle (KMC) model. *Electronic Journal of Knowledge Management*, 12, 85–97.
- Osterwalder A., Pigneur Y. (2010) *Business model generation: a handbook for visionaries, game changers, and challengers*. John Wiley & Sons.
- Scheer C., Deelmann T., Loos P. (2003) Geschäftsmodelle und internetbasierte Geschäftsmodelle- Begriffsbestimmung und Teilnehmermodell. Johannes Gutenberg Universität Main.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process Chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- Simmert, B., Ebel, P. A., Peters, C., et al. (2019). Conquering the challenge of continuous business model improvement: Design of a repeatable process. *Business & Information Systems Engineering*, 61, 451–468. <https://doi.org/10.1007/S12599-018-0556-Y>
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and data science in engineering management. CIO 2022. Lecture notes on data engineering and communications technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- Zhang AX., Muller M., Wang D. (2020) How do data science workers collaborate? roles, workflows, and tools. *Proc ACM Human-Computer Interact* 4:1–23.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Teil III

Erfolgsgeschichten aus der Anwendung



Datengetriebene Arbeitsplangestaltung in der Automobilfertigung

9

Entwicklung eines mehrstufigen Vorgehensmodells
zur datenbasierten Optimierung und Gestaltung von
Montagelinien

Christine Rese , Sven Krzoska , Edin Klasic , Nikolai West , Philipp Schlunder , Ralf Klinkenberg , Mathias Gebler  und Jochen Deuse 

9.1 Motivation

Im produzierenden Gewerbe werden betriebliche Entscheidungen traditionell anhand von zahlenbasierten Beobachtungen getroffen (Deuse et al., 2021, S. 1). In Zeiten stetig zunehmender Möglichkeiten zur Datenerfassung und -speicherung stellen daten-

C. Rese (✉) · S. Krzoska · M. Gebler
Volkswagen AG, Wolfsburg, Deutschland
E-Mail: christine.rese@volkswagen.de

S. Krzoska
E-Mail: sven.krzoska@volkswagen.de

M. Gebler
E-Mail: mathias.gebler@volkswagen.de

E. Klasic · R. Klinkenberg
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: eklasic@rapidminer.com eklasic@altair.com

R. Klinkenberg
E-Mail: rklinkenberg@rapidminer.com; rklinkenberg@altair.com

N. West · P. Schlunder · J. Deuse
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: nikolai.west@ips.tu-dortmund.de

P. Schlunder
E-Mail: philipp.schlunder@udo.edu

J. Deuse
E-Mail: jochen.deuse@ips.tu-dortmund.de

basierte Assistenzsysteme eine geeignete Weiterentwicklung dieser Entscheidungsunterstützung dar. Solche Systeme eignen sich, um Datenpotenziale in komplexen Unternehmensumfeldern zu erschließen und um daraus nutzbare Informationen und Wissen zu extrahieren (West et al., 2021, S. 133). Insbesondere die Mitarbeitenden benötigen digitale Werkzeuge zur Beherrschung der Datenvielfalt und -menge, um mit fachlicher Expertise und entsprechendem Domänenwissen schnell und zielgerichtet Entscheidungen treffen zu können.

Aus Sicht der Produktion und des Produktivitätsmanagements ist die Automobilindustrie durch eine mehrdimensionale Komplexität geprägt. Neben einer möglichst effizienten Abarbeitung kundenindividueller Fahrzeugbestellungen in standardisierten Fließfertigungen, gilt es in Mixlinien unterschiedliche Fahrzeugmodelle zu integrieren, sodass Verrichtungszeitdauer bzw. deren Spreizung, Wertschöpfungsanteil, Materialanstellung, Betriebsmittelauswahl, Verbaureihenfolge und weitere Planungsperspektiven berücksichtigt werden (Gebler, 2021, S. 32 ff.). Dabei ist wettbewerbsentscheidend, dass sowohl für die Mitarbeitenden in der Produktion als auch für die übergreifende Produktivität, ein optimaler Fertigungsablauf als Kombination der Komplexitätstreiber zu einem Gesamtoptimum ausgestaltet werden.

Die Herstellung eines Fahrzeugs bedarf der Ausführung einer Vielzahl von komplexen Montageschritten (Demlehner et al., 2021, S. 2), welche typischerweise anhand von Arbeitsplänen strukturiert und dokumentiert werden (Manns et al., 2015, S. 349; Rese et al., 2023, S. 11 f.; Wallis et al., 2014a, S. 181 f.). Ein Arbeitsplan basiert dabei auf Daten aus der Stückliste und weiterer Ressourcentübersichten und wird entlang des Produktentstehungsprozesses mit Daten angereichert, um die oben dargestellten Planungsperspektiven zu integrieren und abzubilden (Kropik, 2009, S. 55).

Insbesondere für Unternehmen mit verteilten Fertigungsstandorten zählt es in der Produktionsplanung zur gängigen Praxis, bei der Gestaltung und Optimierung von Montageprozessen auf bewährte Lösungsansätze aus Arbeitsplänen zurückzugreifen. Durch eine Verbreitung dieser etablierten Praktiken kann eine werks- und standortübergreifende Konsistenz der Prozessgestaltung unterstützt werden. Zusätzlich wird die Anpassbarkeit und Nutzung von bestehendem Prozesswissen über Produktlinien hinweg gewährleistet. Einer standardisierten Dokumentation dieser Fertigungsabläufe für unterschiedliche Produktkonfigurationen kommt somit vor dem Hintergrund einer zukünftigen Nutzung datengetriebener Assistenzsysteme eine wettbewerbsrelevante Bedeutung zu.

Dieses Kapitel zeigt am Beispiel der Fahrzeugmontage, wie industrielle Datenanalysen eingesetzt werden können, um diesen Herausforderungen zu begegnen. Im Speziellen wird die Entwicklung einer datenbasierten Analyse, die sich aus Modulen zum automatischen Auffinden, Wiederverwenden und Vergleichen relevanter Prozessplanungsdaten zusammensetzt, betrachtet. Dieser Anwendungsfall stellt eine der acht Erfolgsgeschichten im Referenzbaukasten des Forschungsprojekts AKKORD dar (siehe Kap. 4).

9.2 Datenbasierte Optimierung und Gestaltung von Montagelinien

Das durch Erfahrungen und inkrementelle Verbesserungen gewonnene Wissen zur Erstellung von adäquaten Arbeitsplänen ist in produzierenden Unternehmen häufig nicht an einem Ort gebündelt (Wallis et al., 2014b, S. 262). Zum einen findet es sich dezentral als Wissen bei den Mitarbeitenden der Planungsabteilungen und zum anderen ist es innerhalb der erprobten und etablierten Arbeitspläne selbst enthalten.

Sowohl entlang des Produktentstehungsprozesses als auch in der sich anschließenden Serienfertigung muss der Aspekt der Datendurchgängigkeit Berücksichtigung finden. Angesichts der Produktvarianz und der Komplexität globaler Produktionsnetzwerke sind damit einige Herausforderungen verbunden. Ein Arbeitsplan ist analog dem Fertigungsprozess strukturiert und umfasst sowohl produktbezogene, prozessuale als auch organisationale Informationen. Er beinhaltet alle Fertigungs- und Prüfschritte, die für die Fertigung eines Fahrzeugmodells erforderlich sind. Zudem dienen Arbeitspläne als Datengrundlage, die in IT-Systemen im Produktionsumfeld wie z. B. in Informationssystemen für Werker:innen weiterverwendet werden (Kropik, 2009, S. 55, S. 235 ff.).

Im Folgenden wird am Beispiel der Fahrzeugmontage die Charakteristik dieser Arbeitspläne dargestellt. Die Fließfertigung lässt sich in einzelne Arbeitssysteme zerlegen. Ein Arbeitssystem besteht aus mehreren Mikroprozessen, die beispielsweise durch einen Mitarbeitenden im Rahmen der Taktzeit bearbeitet werden. Ein einzelner Arbeitsplan beinhaltet mehrere zehntausend einzelne Mikroprozesse, die als Arbeitsvorgänge bezeichnet werden. Diese lassen sich nicht mehr sinnvoll aufteilen. Es können wertschöpfende, bauteilbezogene Arbeitsvorgänge, in denen z. B. ein einzelnes Bauteil dem Fahrzeug hinzugefügt wird, von prozessual erforderlichen Mikroprozessen, in denen z. B. vorbereitende oder prüfende Tätigkeiten erfolgen, unterschieden werden. Nicht alle Arbeitsvorgänge sind für jede kundenspezifische Produktkonfiguration relevant. Zwischen verschiedenen Fahrzeugmodellen, die auf einer Montagelinie gefertigt werden, kann es Unterschiede zwischen den jeweils gültigen Arbeitsvorgängen geben (Kropik, 2009, S. 55 ff.; Bracht et al., 2015, S. 291 ff.). Abb. 9.1 zeigt beispielhaft die Struktur eines Arbeitsplans.

In einem globalen Produktionsnetzwerk mit einer Vielzahl von Produktionsstandorten und verschiedenen Produkten sind im Rahmen der Produktentstehung und in der Serienphase verschiedene Fragestellungen zur Produktivitätsgestaltung relevant. Als Teil der Planungsarbeit müssen besonders bewährte Lösungen in den Produktionsprozessen identifiziert werden. Dazu stellen quantitative und qualitative Vergleiche einen wichtigen Ausgangspunkt für Verbesserungen von Arbeitsplänen dar. Eine manuelle Suche nach ähnlichen Mikroprozessen ggfs. in fremdsprachigen Arbeitsplänen ist jedoch vor dem Hintergrund der oben dargestellten Anzahl an Arbeitsvorgängen je Arbeitsplan eine komplexe und zeitaufwendige Aufgabenstellung. Zur Arbeitsprozessgestaltung werden häufig computerbasierte Planungssysteme eingesetzt. In diesen existieren umfangreiche

Prozessplanungssystem

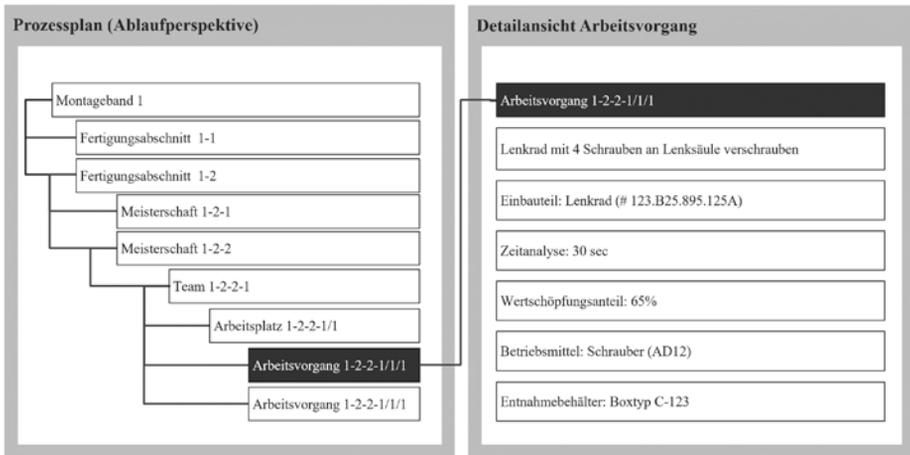


Abb. 9.1 Beispielhafte Struktur eines Arbeitsplans

Daten in Form von Arbeitsplänen, die verschiedene Szenarien der Fahrzeugfertigung innerhalb des globalen Produktionsnetzwerkes beschreiben. Werkzeuge, um dieses Planungswissen zugänglich zu machen und insbesondere spezifische Arbeitsvorgänge eindeutig zu identifizieren und diese Eindeutigkeit maschinenlesbar abzubilden und bereitzustellen, fehlen heutzutage.

Im Rahmen des Anwendungsfalls wurden datengetriebene Methoden und Werkzeuge entwickelt, die ein ganzheitliches Produktivitätsmanagement im Industrial Engineering unterstützen. Durch die Erweiterung des Kompetenzportfolios um Methoden und Werkzeuge der industriellen Datenanalyse und die damit verbundene Implementierung von datenbasierten Entscheidungsunterstützungen entlang des Produktentstehungsprozesses und in der Serienphase wird der Fokus stärker auf die Gestaltungsaufgaben im Industrial Engineering gelegt. Durch die Integration, Aufbereitung, Analyse und kontextbezogene Bereitstellung von vorhandenen Daten aus computerbasierten Planungssystemen, wird im Unternehmen bereits bestehendes Wissen zum Produktivitätsmanagement leichter zugänglich und teilweise überhaupt erst nutzbar. Heute sind Gestaltungs- und Optimierungslösungen für manuelle Arbeitsprozesse in der Fertigung stark von den beteiligten Personen und deren Erfahrungswissen abhängig. Zukünftig wird durch die automatische Analyse von Arbeitsplänen der manuelle Aufwand zur Identifikation und die Auswertung von relevanten Prozessplanungsdaten stark reduziert und mit dem beschriebenen Assistenzsystem wird ein Beitrag zur Verbesserung der Auffindbarkeit und Wiederverwendbarkeit von bestehendem Prozesswissen aus Arbeitsplänen geleistet. Mithilfe der zielgerichteten Entscheidungsunterstützungen und Planungsempfehlungen wird die Gestaltung und Optimierung von Montagelinien vereinfacht und verbessert.

9.3 Umsetzung der datenbasierten Entscheidungsunterstützung

Für den AKKORD-Referenzbaukasten wurden verschiedene Lösungsbausteine entwickelt (siehe Kap. 1 und 2). Die konzeptionelle und technische Umsetzung der datengetriebenen Assistenzfunktionen zur Gestaltung und Optimierung von Arbeitssystemen erfolgt im Rahmen eines mehrstufigen Vorgehensmodells (siehe Abb. 9.2) und fokussierte deutschsprachige Arbeitspläne der TÜVormontage (Rese et al., 2023, S. 11 f.).

Für die erfolgreiche Anwendung datengetriebener Assistenzfunktionen ist die Datenvorverarbeitung unerlässlich (García et al., 2016, S. 1). Die technische Realisierung dieser vorbereitenden Aktivitäten sowie des mehrstufigen Vorgehensmodells erfolgte in RapidMiner Studio, einer Software-Lösung für Data Mining und Maschinelles Lernen sowie in Python. Zu den entwickelten Lösungen für die Datenvorverarbeitung zählen beispielsweise die Visualisierung von fehlenden Werten in Prozessplänen, der Einsatz von Datenschemata oder die Vereinfachung von Prozessbeschreibungen zur Verringerung der sprachlichen Komplexität. Abb. 9.3 zeigt beispielhaft die Ergebnisse zur automatischen Reduktion der sprachlichen Komplexität der Prozessbeschreibungen. Diese erfolgt wie in Rese et al. (2023) ausführlich beschrieben (S. 5 f.). Reguläre Ausdrücke ersetzen im Vorfeld definierte Satzbausteine, z. B. bei der Angabe von Multiplikatoren oder Trennzeichen. Die Vereinfachung dient der besseren Identifikation von Mustern und Zusammenhängen in den Daten. Auf eine Rechtschreibkorrektur wird zugunsten der Datenintegrität verzichtet (Rese et al., 2023, S. 5 f.).

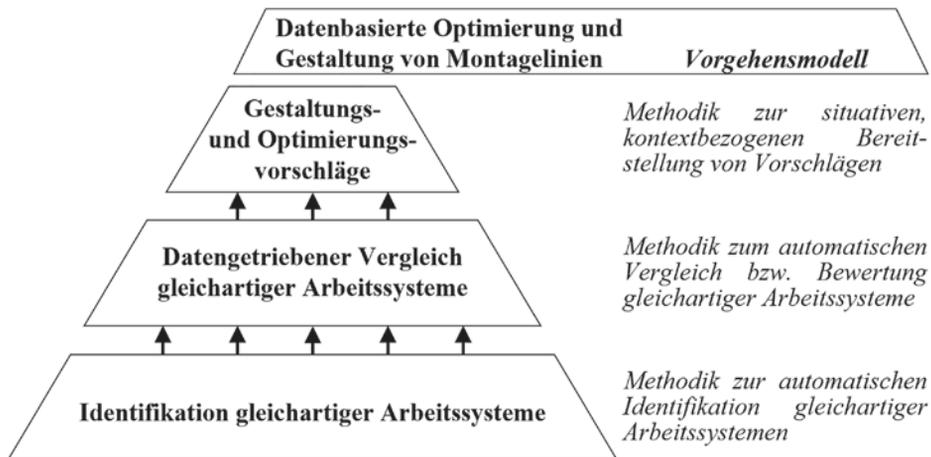


Abb. 9.2 Vorgehensmodell zur datenbasierten Optimierung und Gestaltung von Montagelinien. (Nach Rese et al., 2023, S. 11 f.)

Original Prozessbeschreibung	Bereinigte Prozessbeschreibung
Hebgestell Fensterheber vorn rechts 4x verschrauben	Hebgestell Fensterheber vorn rechts <MULT> verschrauben
(L0R) Schließzylinder montieren	<GUELT> Schließzylinder montieren
TIV: Türsteuergerät elektrisch stecken, links	TIV <SEP> Türsteuergerät elektrisch stecken <SEP> links

Abb. 9.3 Prozessbeschreibungen vor und nach der Komplexitätsreduktion

Zur automatischen **Identifizierung gleichartiger Arbeitssysteme** innerhalb verschiedener Prozesspläne bedarf es zudem der maschinellen Auslesbarkeit der Montageprozesse (Rese et al., 2023, S. 11 f.). Durch die Überführung der Textdaten in ein strukturiertes semantisches Modell, wie eine kontrollierte Sprache, kann der Analyseaufwand maßgeblich reduziert oder teils überhaupt erst möglich werden. In der Produktionsplanung ist die Verwendung einer kontrollierten Sprache jedoch nicht verbreitet. Gründe hierfür liegen u. a. im Zeitaufwand, die dessen Etablierung erfordert (Manns et al., 2015, S. 349). Der entwickelte Referenzbaukasten enthält Analyseschritte, die natürlichsprachliche Prozessbeschreibungen automatisch in eine kontrollierte Sprache überführen. Durch die Verwendung der synthetisch generierten, kontrollierten Sprache können unstrukturierte, textuelle Prozessbeschreibungen standardisiert und Fehler bei der Beschreibung vermieden werden. Damit trägt die maschinelle Auslesbarkeit der Montageprozesse zu einer höheren Qualität der Prozesspläne bei. Abb. 9.4 zeigt eine manuell erstellte Prozessbeschreibung und die entsprechende automatisch durch verschiedene Text-Mining-Verfahren (u. a. Part of Speech Tagging, Named Entity Recognition) generierte Beschreibung inklusive ihrer numerischen Repräsentationen. Für die Überführung der Montageprozessbeschreibungen in ein maschinenlesbares Format wurde die in Rese et al. (2023) vorgestellte Methodik verwendet und erweitert. Es erfolgte beispielsweise eine Erweiterung des Trainingsdatensatzes für das

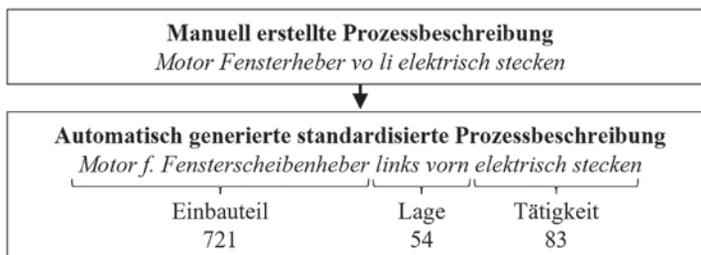


Abb. 9.4 Synthetisch generierte Prozessbeschreibung in kontrollierter Sprache

Named-Entity-Recognition-Modell von ursprünglich 134 auf mehr als 2000 Montagebeschreibungen).

Neben den textuellen Prozessbeschreibungen werden für die automatische Überführung in die kontrollierte Sprache die hinterlegten Teilenummern als nicht textuelle Informationen genutzt. Die Validierung dieses Teilschrittes erfolgte auf Grundlage von 2740 Montagebeschreibungen der Türenvormontage und ergab, dass durchschnittlich 87 % der gefundenen numerischen Repräsentationen korrekt zurückgeführt werden konnten. Eine Rückführung auf die kontrollierte Sprache war insbesondere nicht möglich, wenn die vorliegenden Daten nicht vollständig gepflegt waren (z. B. fehlende Informationen zu Tätigkeiten oder Wortneuschöpfungen) (Rese et al., 2023, S. 10). Durch die standardisierte, numerische Abbildung der Prozessbeschreibungen sind Montageschritte einem Einbauteil zuordenbar. Das identifizierte Einbauteil bildet die Grundlage eines bauteilbezogenen Montageprozesses. Ihre inhaltliche Ähnlichkeit bestimmt sich über die automatisch extrahierte Bauteilzugehörigkeit. Damit gelingt es Montageprozesse mit ähnlichen Bauteilen in verschiedenen Prozessplänen aufzufinden, um die in ihnen enthaltenen Informationen zu extrahieren und in einem späteren Schritt Vergleiche zu ermöglichen.

Der **automatische Vergleich bzw. die Bewertung der Montageprozesse** basiert auf der im Vorfeld durchgeführten automatischen Identifikation der bauteilbezogenen Montageprozesse. Derzeit sind Gestaltungs- und Optimierungslösungen für manuelle Prozesse in der Fertigung stark vom Domänenwissen und der individuellen Erfahrung abhängig. Mit einer steigenden Anzahl von Produkten und einem großen Produktionsnetzwerk verlieren das persönliche Netzwerk und frühere Erfahrungen an Bedeutung. Gleichzeitig ist der Aufbau von Querverbindungen und der Wissenstransfer in komplexen Unternehmensumfeldern eine anspruchsvolle Aufgabe. Die Ergebnisse der Extraktion reduzieren den manuellen Aufwand zur Identifikation relevanter Daten. Insbesondere ermöglicht die maschinelle Auslesbarkeit die Identifikation von besonders bewährten Montageprozessen. Die Bewertung und der Vergleich der bauteilbezogenen Montageprozesse basieren auf entwickelten Gestaltungskriterien. Grundlage dieser sind Kriterien, die zur Bewertung von arbeitsplatzbezogenen Taktungsvorschlägen entwickelt wurden (Gebler, 2021, S. 119 ff.). Die dort enthaltenen Bedingungen wurden an die Anforderungen einer montageprozessorientierten Betrachtungsweise angepasst und erweitert. Somit können Montageprozesse u. a. hinsichtlich ihrer Wertschöpfung, Fertigungszeit, Materialanstellung, Arbeitshöhe und Betriebsmittel bewertet und verglichen werden. Abb. 9.5 zeigt die Umsetzung der datenbasierten Entscheidungsunterstützung für die Gestaltung- und Optimierung von Montageprozessen in Power-BI. Der erstellte Bericht ermöglicht einen marken- und werksübergreifenden Vergleich, der automatisch identifizierten Montageprozesse hinsichtlich verschiedener Vergleichskriterien.

Auf Basis der ermittelten Unterschiede der betrachteten Montageprozesse werden Optimierungspotenziale aufgezeigt. Dies ermöglicht die automatische **Bereitstellung von Vorschlägen zur Prozessgestaltung und -optimierung**. Dem Arbeitsgestalter werden somit automatisch besonders bewährte Lösungen für Montageprozesse bereit-



Abb. 9.5 Power-BI-Bericht zur datenbasierten Entscheidungsunterstützung für die Gestaltung und -optimierung von Montagelinien

gestellt. Für die untersuchten Prozesspläne der Türenvormontage konnten mehr als 80 % der bauteilbezogenen Montageprozesse automatisch aus sieben Prozessplänen identifiziert und marken- und werksübergreifend verglichen werden. Hierfür bedurfte es keinen manuellen Anpassungen an den Eingangsdaten beispielsweise Korrekturen der Montagebeschreibungen.

9.4 Fazit

Am Beispiel der Automobilfertigung wurde gezeigt wie der Prozess zur Gestaltung und Optimierung von Montageprozessen datengetrieben unterstützt werden kann. Durch die Integration, Aufbereitung, Analyse und kontextbezogene Bereitstellung von bereits vorhandenen Prozessplanungsdaten, wird bestehendes Wissen zum Produktivitätsmanagement leichter zugänglich bzw. überhaupt erst nutzbar. Das entwickelte dreistufige Vorgehensmodell ermöglicht das automatische Auffinden, Wiederverwenden und Vergleichen der vor und nach Produktionsbeginn generierten Prozessplanungsdaten. Die Methodik zur Überführung von textuellen Prozessbeschreibungen in eine strukturierte semantische Darstellung liefert einen wertvollen Beitrag für die Extrahierung von Prozesswissen im Rahmen der Produktionsplanung und kann mitunter für weiterführende Aufgaben wie Textgenerierung, Wissensgenerierung oder Simulationserstellung genutzt werden. Durch die Implementierung von datenbasierten Entscheidungsunter-

stützungen entlang des Produktentstehungsprozesses und in der Serienphase wird der Fokus stärker auf die Gestaltungsaufgaben im Industrial Engineering gelegt.

Der Vergleich und die darauf aufbauende Identifikation der besonders bewährten Lösungen sind nicht nur für einzelne bauteilbezogene Montageprozesse möglich. Zukünftig können Fertigungsabschnitte bzw. komplette Arbeitspläne mit sämtlichen Bauteilen der Fahrzeugmontage gegenübergestellt werden. Hierfür bedarf es u. a. der Weiterentwicklung der Methodik zur Überführung der natürlich-sprachlichen Montagebeschreibungen in die kontrollierte Sprache. Verbesserungsmöglichkeiten bieten eine datenbasierte Erweiterung der Referenzdaten (z. B. der Synonymlisten) sowie der Einsatz von Wortvektoren. Für die Integration in den Planungsprozess eines globalen Produktionsnetzwerkes ist zudem die Berücksichtigung der Mehrsprachigkeit von Arbeitsplänen notwendig sowie die Ausdehnung auf weitere Gewerke wünschenswert.

Schlussendlich sei für eine ausführlichere Betrachtung der allgemeinen Weiterentwicklung industrieller Datenanalysen im Einklang mit Mensch, Technik und Organisation auf Kap. 20 verwiesen.

Literatur

- Bracht, U., Geckler, D., & Wenzel, S. (2015). *Digitale Fabrik. Methoden und Praxisbeispiele*. Springer. <https://doi.org/10.1007/978-3-540-88973-1>.
- Demlehner, Q., Schoemer, D., & Laumer, S. (2021). How can artificial intelligence enhance car manufacturing? A Delphi study-based identification and assessment of general use cases. *International Journal of Information Management*, 58. <https://doi.org/10.1016/j.ijinfomgt.2021.102317>.
- Deuse, J., West, N., & Syberg, M. (2021). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- García, S., Luengo, J., & Herrera, F. (2016). Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Knowledge-Based Systems*, 98, 1–29. <https://doi.org/10.1016/j.knosys.2015.12.006>.
- Gebler, M. (2021). *Industrialisierung von Optimierungsmethoden zur automatisierten Fließbandabstimmung in der Automobilindustrie*. Dissertation, Friedrich Schiller University, Jena.
- Kropik, M. (2009). *Produktionsleitsysteme in der Automobilfertigung*. Springer. <https://doi.org/10.1007/978-3-540-88991-5>.
- Manns, M., Wallis, R., & Deuse, J. (2015). Automatic proposal of assembly work plans with a controlled natural language. *Procedia CIRP*, 33(1), 345–350. <https://doi.org/10.1016/j.procir.2015.06.079>.
- Rese, C., West, N., Gebler, M., Krzoska, S., Schlunder, P., & Deuse, J. (2023). Pipeline for the automatic extraction of procedural knowledge from assembly instructions into controlled natural language. *Journal of Software*, 18(1), 1–14. <https://doi.org/10.17706/jsw.18.1.1-14>.
- Wallis, R., Erohin, O., Klinkenberg, R., Deuse, J., & Stromberger, F. (2014a). Data mining-supported generation of assembly process plans. *Procedia CIRP*, 23. <https://doi.org/10.1016/j.procir.2014.10.095>.

- Wallis, R., Stjepandic, J., Rulhoff, S., Stromberger, F., & Deuse, J. (2014b). Intelligent utilization of digital manufacturing data in modern product emergence processes. *Advances in Transdisciplinary Engineering*, 1, 261–270. <https://doi.org/10.3233/978-1-61499-440-4-261>.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Kollaborative, modulare Datenanalyse als Werkzeug im Qualitätsmanagement

10

Potenziale von Citizen Data Science im Qualitätsmanagement zur Erweiterung der Feldbeobachtung

Martina Ringeln , Sibylle Legner , Sebastian Pähler , Katharina John , Andreas Cibis , Ralf Gryga , Lennart Krüger , Corinna Osterbrink , Felix Reinhart , Edin Klavic , Ralf Klinkenberg , Marius Syberg , Hermann Ferstl , Tanja Sindram, Michael Doukas und Jochen Deuse 

M. Ringeln (✉) · S. Legner · S. Pähler · K. John · A. Cibis · R. Gryga · L. Krüger · C. Osterbrink · F. Reinhart
Miele & Cie. KG, Gütersloh, Deutschland
E-Mail: martina.ringeln@miele.com

S. Legner
E-Mail: sibylle.legner@miele.com

S. Pähler
E-Mail: sebastian.paehler@miele.com

K. John
E-Mail: katharina.john@miele.com

A. Cibis
E-Mail: andreas.cibis@miele.com

R. Gryga
E-Mail: ralf.gryga@miele.com

L. Krüger
E-Mail: lennart.krueger@miele.com

C. Osterbrink
E-Mail: corinna.osterbrink@miele.com

F. Reinhart
E-Mail: felix.reinhart@miele.com

E. Klavic · R. Klinkenberg
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: eklavic@rapidminer.com eklavic@altair.com

10.1 Motivation

Die Vision von industriellen Datenanalysen, Daten entlang der gesamten Wertschöpfungskette mit Daten aus dem Produktlebenszyklus zu vernetzen, beinhaltet großes Potenzial für ganzheitliche Qualitätsverbesserungen. Industrielle Datenanalyse ist ein Sammelbegriff für datenwissenschaftliche Werkzeuge zur faktenbasierten Entscheidungsfindung in industriellen Anwendungsfällen (West et al., 2021a, S. 131). Um dieses zunehmende Potenzial ausschöpfen zu können, bedarf es vor allem zielgerichteter Werkzeuge und Strategien für Datenwissenschaften sowie der Befähigung von Expertinnen und Experten im Qualitätsmanagement (siehe Kap. 2). Diese entwickeln sich somit in die Rolle von sogenannten *Citizen Data Scientists*, welche Werkzeuge der Datenwissenschaften gewinnbringend verwenden, weiterentwickeln oder bei der Weiterentwicklung mitarbeiten (Schwenken et al., 2023, S. 84). Gleichzeitig bedeutet dies auch, dass die bekannten und etablierten Methoden und Systeme des Qualitätsmanagements mit Methoden der Datenwissenschaften in Einklang gebracht werden müssen (siehe hierzu Schäfer et al., 2018, S. 190 ff.). In diesem Kapitel wird beschrieben, wie mithilfe von modularen Analysebausteinen eine Lösung zur Unterstützung der Feldbeobachtungen bei Miele entwickelt wurde. Die Zielsetzung war hierbei herauszufinden, wie das für konkrete Fragestellungen aus der Praxis mit Hilfe des für solche Aufgaben konzipierten Referenzbaukastens im Forschungsprojekt AKKORD konkret gelingen kann (siehe Kap. 4).

R. Klinkenberg

E-Mail: rklinkenberg@rapidminer.com; rklinkenberg@altair.com

M. Syberg

Technische Universität Dortmund, Dortmund, Deutschland

E-Mail: marius.syberg@ips.tu-dortmund.de

H. Ferstl · T. Sindram · M. Doukas · J. Deuse

mosaiic GmbH, München, Deutschland

E-Mail: hermann.ferstl@mosaiic.com

T. Sindram

E-Mail: tanja.sindram@mosaiic.com

M. Doukas

E-Mail: michael.doukas@mosaiic.com

J. Deuse

E-Mail: jochen.deuse@ips.tu-dortmund.de

10.2 Datenwissenschaft und Feldbeobachtung im Qualitätsmanagement

Der Anwendungsfall ist im Bereich der Feldbeobachtung im Qualitätsmanagement angesiedelt, dessen wichtigste Aspekte im Folgenden kurz erläutert werden.

10.2.1 Datenwissenschaft im Qualitätsmanagement

Seit jeher arbeiten im Unternehmen Miele, genauso wie in anderen Unternehmen, Expertinnen und Experten mit verschiedenen fachlichen Hintergründen gemeinsam an der Sicherung der Qualität von Produkten und Prozessen. Qualität entsteht dabei nicht vorwiegend im Qualitätsmanagement, sondern beispielsweise in der Entwicklung oder der Produktion. Wobei das Qualitätsmanagement als Führungsinstrument für die gesamte Organisation fungiert und dabei als technischer Controller dient. Dies lässt sich z. B. aus den grundlegenden Anforderungen der ISO 9001 für ein Qualitätsmanagementsystem ableiten (DIN EN ISO 9001: 2015-11). Hierin werden Unternehmen dazu aufgefordert, ihre gesamte Organisation und die Prozesse auf die Erfüllung der Kundenanforderungen auszurichten und damit die gewünschte Qualität zu erreichen.

Vorreiter beim Thema Data Science sind Entwicklung, Service und Produktion. Dort werden unter der Überschrift „Internet of Things“ (IoT) bzw. „Industrial Internet of Things“ (IIoT), neue Produkte und Features für Endanwender bzw. neue, effiziente Fertigungstechnologien ermöglicht. Aus der zentralen Sicht des Qualitätsmanagements entstehen durch die Verfügbarkeit besserer und umfangreicherer Daten und Methoden einzigartige Möglichkeiten, Qualität ganzheitlicher über den gesamten Produktlebenszyklus zu bewerten (Schmitt & Pfeifer, 2015, S. 36). Die Expertise der Mitarbeitenden im Qualitätsmanagement als Daten- und Informationsvernetzer zu Qualitätsthemen, mit Erfahrung in der Einschätzung und Bewertung von Daten aus unterschiedlichen Kontexten, ist bei der Umsetzung solcher Themen unerlässlich.

10.2.2 Feldbeobachtung

Die Daten aus der Nutzungsphase der Produkte, die Gegenstand der Feldbeobachtung sind, stammen klassischerweise aus der Auswertung von Kundendienst-Einsätzen (VDA QMC, 2009). Inhalt der Feldbeobachtung ist, neben der Erfüllung der Produktbeobachtungspflichten (§ 823 BGB und ProdHaftG), die Beobachtung der erlebten Produktqualität im Feld, mit dem Ziel, mögliche Qualitätsabweichungen schnell zu erkennen und langfristig abzustellen. Darüber hinaus lassen sich Erkenntnisse über die erlebte Qualität beim Kunden auch für Neuentwicklungen nutzen (Schmitt & Pfeifer, 2015, S. 249 ff.).

Da sich sehr unterschiedliche Einflüsse, z. B. von Produktausprägung, Produktionsmengen und Produktnutzung, in den Felddaten niederschlagen, erfordert die Analyse und Auswertung dieser Daten, oft in Form von Zeitreihen, Expertenwissen. Eine Besonderheit der Feldbeobachtung erwächst daraus, dass hier Daten mit potenziellem Kundenbezug betrachtet werden. Diese Daten und Informationen sind sehr wertvoll, da sie direkte Rückschlüsse auf die erlebte Qualität beim Kunden erlauben (Schmitt & Pfeifer, 2015, S. 97 ff.). Gleichzeitig schränkt die Berücksichtigung von Datenschutzanforderungen die Möglichkeiten für Auswertungen und Datenverknüpfungen stark ein. In diesem Kontext stellt die Feldbeobachtung beim Qualitätsmanagement hohe Anforderungen an eine durchgängige Qualität der zu analysierenden Daten (West et al., 2021, S. 131 ff.).

Für die Feldbeobachtung stehen zunehmend weitere Datenquellen zur Verfügung, insbesondere Daten aus vernetzten IoT-Geräten (Ewerszumrode et al., 2021). Außerdem ermöglichen es neue Werkzeuge, aber auch die wachsende Qualität und Verfügbarkeit von Daten aus Datenbanken, systematische Analysen auf weitere bekannte Datenquellen auszuweiten.

10.3 Zielsetzung im Forschungsvorhaben

„Wie können die Experten im Umfeld der Feldbeobachtung dazu befähigt werden, die Potenziale von neuen und bekannten Daten mittels Datenwissenschaften optimal zu heben?“ – so lässt sich die Fragestellung des Anwendungsfalls an AKKORD grob umreißen. Während der Anwendungsfall in Kap. 2 bereits kurz dargestellt wurde, wird nachfolgend die Unterstützung, die der Referenzbaukasten hinsichtlich der konkreten Problemsetzung bietet, detaillierter beschrieben.

10.3.1 Befähigung und gezielte Qualifikation von Mitarbeitern

Die Mitarbeitenden aus dem Qualitätsmanagement haben zwar eine hohe Expertise im Umgang mit Daten auf ihrem Gebiet sowie zu Qualitätsthemen und relevanten Prozessen, aber meist keine Erfahrung mit Datenwissenschaften. Das zeigte auch die qualitative Befragung, die zur Erstellung von sogenannten *Personas* im Rahmen des Anwendungsfalls durchgeführt wurde. Damit diese Mitarbeitenden in die Rolle als Citizen Data Scientists hineinwachsen können, war es ein Anliegen, Instrumente zur Qualifikation der Mitarbeitenden in der entwickelten *Work&Learn-Plattform* zur Verfügung zu haben. Eine ausführliche Beschreibung dieser Plattform findet sich weiterführend in Kap. 7 und 11. Dabei zeigte sich, dass eine Weiterbildung sowohl zu reinen Datenthemen als auch zu Vorgehensmodellen und Methoden sehr hilfreich ist. Das u. a. für die Verwendung des Referenzbaukastens entwickelte Vorgehensmodell wird darüber hinaus beschrieben in Kap. 16.

10.3.2 Niederschwelliger Einstieg in Datenanalysen für konkrete Anwendungsfälle

Wie in Abschn. 10.4.1 näher ausgeführt wird, ähneln sich viele Fragestellungen an Zeitreihendaten (siehe Schlegl et al., 2022, S. 52 ff.; West et al., 2021a, S. 762 ff.), die beispielsweise auch in der Feldbeobachtung relevant sind, sehr. Daher war ein Anliegen aus dem Anwendungsfall, Analysebausteine, die in bekannten Zusammenhängen erfolgreich erprobt wurden, zur Nutzung in anderen Kontexten zur Verfügung stellen zu können.

Die im Referenzbaukasten umgesetzte Lösung begegnet genau diesen Anforderungen (siehe Kap. 6). Ein für den hier beschriebenen Anwendungsfall sehr nützlicher Bestandteil ist die im Vorhaben entwickelte *AI-Toolbox*. Diese bietet die Möglichkeit, Analysen auf einer Web-Oberfläche bereitzustellen, sodass die Berechnungen auf anderen Datensätzen durchgeführt werden können, ohne den „Umweg“ über eine dezidierte Software oder Programmierung. Um die Nutzbarkeit auch im Umfeld von potenziell sensiblen Daten aus dem Unternehmen zu gewährleisten, ist es dabei wichtig, dass die Plattform auch in der geschlossenen IT-Umgebung innerhalb des Unternehmens nutzbar ist.

10.4 Einsatz Industrieller Datenanalysen in der Feldbeobachtung

Im Rahmen des Forschungsprojekts wurden Analysen zu verschiedenen Datenquellen erarbeitet, mit dem Ziel, die serienbegleitende Qualitätsüberwachung zu erweitern und zu vereinfachen. Für das Forschungsprojekt wurden u. a. Daten aus den Ersatzteilverkäufen und anonymisierte Daten aus vernetzten IoT-Geräten gewählt. Diese beiden Datensätze sind nicht direkt mit anderen Daten verknüpfbar und nicht personenbezogen, was in Hinblick auf den Datenschutz und die Vereinfachung von Fragestellungen in diesem Zusammenhang hilfreich ist.

Ein weiterer Anwendungsfall mit internationalen Kundendienst-Daten diene ebenfalls der Validierung.

Die wesentlichen Ergebnisse und Erkenntnisse werden im Folgenden vorgestellt.

10.4.1 Inhaltliche und Methodische Umsetzung

In datenwissenschaftlichen Projekten beruht die Erstellung von erfolgreichen Analysen und Modellen stets auf einem Verständnis der Daten und des Anliegens der Nutzer, auch Geschäftsverständnis genannt, sowie einer entsprechenden Evaluation der Ergebnisse (Chapman et al., 2000, S. 13). Die Personen, die im hier betrachteten Anwendungsfall geeignete Analysebausteine erstellten, brachten bereits sehr gute Vorkenntnisse in diesen Bereichen ein. Die vorgesehenen Akteurinnen und Akteure sind angelehnt an das erarbeitete Rollenkonzept des Forschungsvorhabens (Schwenken et al., 2023, S. 84).

Es zeigte sich, dass ein Vorgehen mit wenigen, passenden Methoden die Projektarbeit strukturiert, die Ergebnisse objektiviert sowie die Akzeptanz und Sichtbarkeit der Arbeit erhöht. Folgender Ansatz wurde in zwei Anwendungsfällen (Ersatzteilverkäufe und internationale Kundendienst-Daten) erfolgreich erprobt:

- **Datenverständnis:** Daten und deren Zusammenhänge wurden in Informationsmodellen bzw. in einem Business Glossar charakterisiert. Die entsprechenden Datenstrukturen wurden auch zum Abgleich als Anwendungsbeispiel für das Datenbackend in AKKORD zur Verfügung gestellt (siehe Kap. 5).
- **Geschäftsverständnis und Nutzerbezug:** Die wichtigsten Nutzergruppen von Ergebnissen der Felddatenanalyse wurden befragt und in Form von 7 Personas charakterisiert. Personen aus diesen Nutzergruppen wurden zusätzlich vor der Modellbildung über Fragebögen zu ihren Einschätzungen und Priorisierungen der Anwendungsfälle befragt.
- **Evaluation:** Die o. g. Nutzergruppen wurden in die Evaluation der Ergebnisse, wieder über Fragebögen, mit einbezogen. Zusätzlich wurde das Gesamtergebnis immer anhand historischer Befunde auf Plausibilität geprüft.

Inhaltlich bezieht der Anwendungsfall seine Analyse-Fragestellungen aus der serienbegleitenden Qualitätsüberwachung, in der Hinweise auf Qualitätsabweichungen aufgenommen und bewertet werden. Für vorhandene Qualitätsabweichungen werden in diesem Bereich auch Abstellmaßnahmen erarbeitet und umgesetzt. Dieser Prozess wird in einem interdisziplinären Team durchgeführt.

In einer vereinfachten Darstellung dieses Vorgehens, in Abb. 10.1 analog zum Six Sigma DMAIC Zyklus skizziert, ist gut erkennbar, dass viele wichtige Analysefragestellungen unabhängig von der betrachteten Datenquelle sehr ähnlich sind: Um mögliche Probleme zu erkennen, wird typischerweise eine Analyse von Abweichungen in Zeitreihendaten benötigt. Die so gefundenen Auffälligkeiten sollten dann näher charakterisiert und quantifiziert werden. Typische Fragestellungen in diesem Themenbereich sind die nach Prognose-Szenarien oder nach möglichen Clustern mit ähnlichen Attributen. Und, nachdem Maßnahmen umgesetzt wurden, werden Visualisierungen von Kenngrößen als Tool zur Überwachung der Qualität benötigt, die dann auch wieder bei der Erkennung von Abweichungen unterstützen können.

In diesen Kategorien sind in der Projektlaufzeit verschiedene generalisierte Analysebausteine für konkrete Anwendungsfälle bei Miele und durch das Projektteam von Miele mit Unterstützung von RapidMiner entstanden. Ergänzend widmete sich das IPS der Aufgabe, solche Bausteine speziell für den allgemeineren Gebrauch auf der AI-Toolbox zu erstellen, und nutzte dazu anonymisierte Daten aus vernetzten IoT-Geräten (siehe Kap. 6). So befinden sich heute unterschiedliche Bausteine zur Visualisierung von Daten als Liniendiagramme, zur Prognose mit unterschiedlichen Verfahren, zur vergleichenden

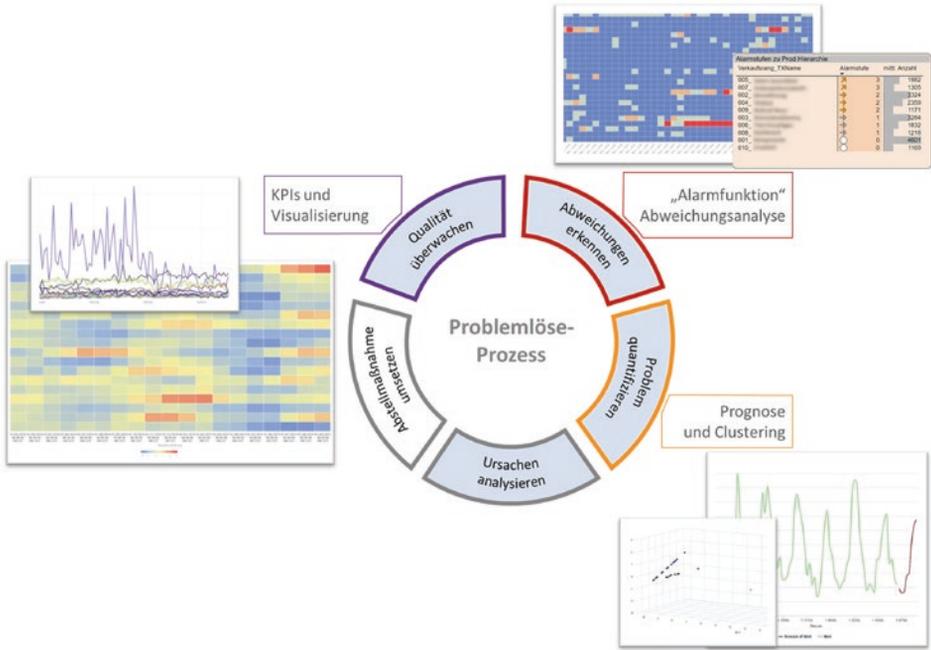


Abb. 10.1 Allgemeiner Problemlöseprozess, Analysefragestellungen und beispielhafte Darstellung von Ergebnissen im Anwendungsfall

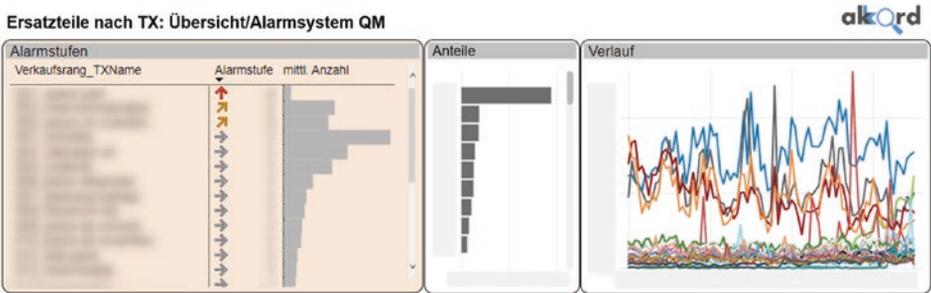


Abb. 10.2 Dashboard zu Übersicht bzw. Alarmsystem der Ersatzteilverkäufe für das Qualitätsmanagement (Ausschnitt)

Bewertung von Prognoseverfahren, zum Clustering und zur Aufbereitung von Attributen in einer sogenannten Heatmap auf der *AI-Toolbox*. Die Ergebnisse zur „Alarmfunktion“ sind für die AI Toolbox weniger geeignet. Sie wurden bei Miele als Bestandteil eines Dashboards in Power BI operationalisiert (siehe Abb. 10.2).

10.4.2 Technische und strategische Einbindung

Einige Rahmenbedingungen, um nicht nur weiterhin Citizen Data Science im Qualitätsmanagement zu betreiben, sondern diese Initiativen noch ausbauen zu können, wurden bereits in der Projektlaufzeit gesetzt: Auf der technischen Seite steht die Sicherstellung der Geheimhaltung und des Datenschutzes durch eine Version der *AI-Toolbox* in Form von Docker Containern. Diese wird innerhalb der IT-Landschaft der Firma betrieben und erhöht damit die Praxistauglichkeit der *AI-Toolbox* auch für andere Firmen. Die strategische Einbindung erfolgt über Change-Maßnahmen. Hier wurde im Projekt unter Federführung von mosaiaic ein Nutzenvideo erstellt, das die Motivation aus AKKORD für Miele und den Use Case illustriert und den Rollout begleiten soll. Näheres zu den Hintergründen und zum Change-Management in AKKORD ist in Kap. 16 zu finden.

10.5 Fazit

Auch nach Abschluss des Projekts bei Miele sollen alle Ergebnisse aus dem Forschungsprojekt weiterhin genutzt und weiterentwickelt werden. Mit dem Rollout und durch die strategische Ausrichtung im Qualitätsmanagement bei Miele ist zu erwarten, dass sich der Anwender- und Nutzerkreis erweitern wird. Die Konzepte und Bausteine, die im Anwendungsfall erarbeitet wurden, wurden auch immer in Hinblick auf Übertragbarkeit und Generalisierbarkeit hin entwickelt und sind daher in ähnlichen Kontexten allgemein nutzbar.

Durch die sehr gute Kombination und Vernetzung von Wissensaufbau, Datenwissenschaften, Mitarbeiterereinbindung und Change Aktivitäten konnte für den Themenkomplex „Industrial Data Science“ ein hohes Bewusstsein mit gleichzeitiger Akzeptanz geschaffen werden. Dieser Umstand wird uns als QM im vernetzten Miteinander in der Miele Organisation entlang der kompletten Wertschöpfungskette, von der Idee bis zur Nutzungsphase, in die Lage versetzen, zukünftig einfach und gemeinsam Data-Science-Projekte durchzuführen. Ein weiterer wichtiger Aspekt ist die Verfestigung des Wissens und der Akzeptanz an der Basis. Die gesamte Organisation wird somit in die Lage versetzt Ideen für Data Science Aktivitäten zu entwickeln.

Darüber hinaus wurden im interdisziplinären Team bei Miele, befeuert durch das Forschungsvorhaben AKKORD, Gruppen und Initiativen entwickelt, die von allen Funktionen genutzt werden können. In Kap. 20 wird dazu eine tiefere Analyse der bevorstehenden Weiterentwicklung industrieller Datenanalysen im Zusammenspiel von Mensch, Technik und Organisation dargestellt.

Literatur

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step data mining guide, CRISP-DM consortium.
- DIN EN ISO 9001:2015-11, Qualitätsmanagementsysteme_- Anforderungen (ISO_9001:2015); Deutsche und Englische Fassung EN_ISO_9001:2015.

- Ewerszumrode, J., Schöne, M., Godt, S., & Kohlhase, M. (2021). Assistenzsystem zur Qualitätssicherung von IoT-Geräten basierend auf AutoML und SHAP. In H. Schulte, F. Hoffmann, & R. Mikut (Hrsg.), *Proceedings – 31. Workshop computational intelligence*: Berlin, 25.–26. November 2021. KIT Scientific Publishing. (S 285–305).
- Schäfer, F., Zeiselmaier, C., Becker, J., & Otten, H. (2018). Synthesizing CRISP-DM and quality management: A data mining approach for production processes *IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*, Marrakech, Morocco. (S. 190–195). <https://doi.org/10.1109/ITMC.2018.8691266>.
- Schlegl, T., Tomaselli, D., Schlegl, S., West, N., & Deuse, J. (2022). Automated search of process control limits for fault detection in time series data. *Journal of Process Control*, 117, 52–64. <https://doi.org/10.1016/j.jprocont.2022.07.002>.
- Schmitt, R., & Pfeifer, T. (2015). *Qualitätsmanagement: Strategien, Methoden, Techniken* (5. Aufl.). Hanser.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process Chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- VDA QMC Verband der Automobilindustrie und Qualitätsmanagement-Center. (2009). *Schadteilanalyse Feld: Vermarktung und Kundenbetreuung* (1. Aufl.). VDA QMC.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021a). Towards integrated data analysis quality. Criteria for the application of industrial data science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- West, N., Schlegl, T., & Deuse, J. (2021b). Feature extraction for time series classification using univariate descriptive statistics and dynamic time warping in a manufacturing environment. In *IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering* (S. 762–768). <https://doi.org/10.1109/ICBAIE52039.2021.9389954>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Integrierte Datenanalyse zur Kollaboration in der Auftragsplanung

11

Implementierung einer prädiktiven Analyse zur optimalen Ressourcenallokation und Kollaboration der Geschäftsbereiche

Andreas Bohlmann , Claudia Scholten , Rachana Desai  und Edin Klasic 

11.1 Motivation

Wie bereits zusammenfassend in Kap. 2 beschrieben, entwickelt ERCO eine mengenbasierte Prognose des Auftragseingangs, um eine gezieltere Allokation von kapazitiven sowie von dispositiven Ressourcen zu erreichen. Derzeit werden Prognosen nur vergleichsweise ungenau und mit einer hohen Schwankungsbreite anhand der vergangener Absatzzahlen erstellt. Aufgrund der enorm hohen Anzahl von Produktvarianten wird die Vorhersage jedoch durch die sporadisch erfolgenden Abrufe erschwert. Gleichzeitig besitzt eine antizipative Planbarkeit der Aufträge jedoch eine hohe Bedeutung zur Stabilisierung der Fertigungs- und Montageprozesse. Das Ziel in diesem Kapitel ist daher die Implementierung von prädiktiver Analysen zur optimalen Ressourcenallokation und Zusammenarbeit zwischen den Geschäftsbereichen des Unternehmens.

Analyselätigkeiten werden häufig anhand von vier Aspekten unterschieden (West et al., 2022, S. 615):

A. Bohlmann (✉) · C. Scholten
ERCO GmbH, Lüdenscheid, Deutschland
E-Mail: a.bohlmann@erco.com

C. Scholten
E-Mail: c.scholten@erco.com

R. Desai · E. Klasic
RapidMiner GmbH, Dortmund, Deutschland
E-Mail: rdesai@rapidminer.com rdesai@altair.com

E. Klasic
E-Mail: eklasic@rapidminer.com; eklasic@altair.com

- **Deskriptiv:** Was ist geschehen?
- **Diagnostisch:** Warum ist etwas geschehen?
- **Prädiktiv:** Was wird geschehen?
- **Präskriptiv:** Was sollte getan werden?

Die in diesem Kapitel behandelte datenwissenschaftliche Problemstellung betrachtet mit der Vorhersage von Auftragsmengen einen prädiktiven Anwendungsfall. Vorhergesagt werden soll eine möglichst präziser Wert als Angabe der erwarteten Absatzmenge. Im Rahmen der unternehmensweiten Zusammenarbeit der Vertriebsgesellschaften sowie der Fertigung und Produktion bis hin zur Lieferantenintegration fließt die Prognose steuernd in das ERCO Wertschöpfungssystem ein. Dabei steht der Einsatz der KI-Software Rapid Miner im Vordergrund.

Darüber hinaus sollen neue Kompetenzen aufgebaut und die Entwicklung der Organisation zu einem höheren Reifegrad in Data-Science-Projekten vorangetrieben werden (siehe Kap. 1). Als Anwendungspartner bringt ERCO seine Kompetenzen in der Ausbildung, Weiterentwicklung und Qualifizierung von Mitarbeitern im Rahmen eines mittelständischen Unternehmens in das Projekt AKKORD ein. ERCO verfügt bereits über eine Datenbasis mit einem hohen Vernetzungsgrad und die grundsätzliche Fähigkeit, analytische Verfahren im Supply Chain Management einzusetzen.

In einem volatilen Marktumfeld bietet das Unternehmen Produkte und Dienstleistungen an, die mit einer hohen technologischen Fertigungstiefe produziert werden. Das eingesetzte System für *Enterprise Resource Planning* (ERP) bildet die einheitliche Datenbasis und bietet die Grundlage für die Zusammenarbeit im Wertschöpfungsnetzwerk.

Die Auftragsprognose bei ERCO basiert auf zwei Säulen von Datenstrukturen in der *SAP Business Suite*. Zum einen sind dies die Projektinformationen aus dem *Customer Relationship Management* (CRM) System, die dort auch als „Opportunities“ und zum anderen die Angebotsdaten aus dem ERP-System.

Eine neue Opportunity wird mit dem Meilenstein zum voraussichtlichen Datum des Auftragseingangs versehen. Dieses Datum kann sich während der Projektlaufzeit ändern und wird regelmäßig angepasst. In Verbindung mit den ERP-Angebotsdaten, bestehend aus einer Angabe Produkts sowie der entsprechenden Menge, die mit dem Projekt verknüpft sind, wird der voraussichtlicher Bedarf für einen Zeitraum von vier Monaten ermittelt.

Die Prognose des voraussichtlichen Auftragseingangs wird zur kapazitiven Auslegung der Produktion in der Bedarfsplanung und zur Anpassung der Lieferkette bei erkennbaren Nachfrageveränderungen genutzt. Die Verantwortung für die Koordination der Bedarfsplanung und -prognose ist wiederum in einer eigenständigen Abteilung konzentriert.

11.2 Absatzprognose in der Auftragsplanung

Aus wissenschaftlich-technischer Perspektive ermöglichen präzise Vorhersagen von Kundenaufträgen eine größere Unabhängigkeit der Planungstätigkeiten vom Kundenentkopplungspunkt. Ohne Vorliegen einer Auftragsbestätigung, also ohne dokumentierte Lieferverpflichtung des Unternehmens, kann die Auftragserfüllung veranlasst und Aufträge proaktiv produziert werden. Im nun folgenden Abschnitt werden die datenwissenschaftlichen Schritte zur dieser prädiktiven Absatzplanung vorgestellt.

11.2.1 Zielsetzung und Vorgehensweise

Am Anfang industrieller Datenanalyseprojekte steht typischerweise die Definition des Geschäftsprozesses und die Entwicklung einer quantifizierbaren Zielsetzung. Es ist zu berücksichtigen, dass zu Beginn eines datenwissenschaftlichen Projekts keine Garantie für die Erreichung der definierten Ziele besteht. Als Zielvorgabe wurde deshalb ein Korridor der Prognosegenauigkeit festgelegt. Um eine adäquate Planungsunterstützung zu ermöglichen, soll die prognostizierte Absatzmenge nicht mehr und nicht weniger als 20 % tatsächlich Absatzmenge betragen. Dadurch sollen Nachfrageschwankungen mithilfe von geeigneten Kapazitäts- und Dispositionsreserven abgefangen werden.

Die Vorgehensweise während der Datenanalyse ist angelehnt an das in der Praxis verbreitete Modell *Cross-Industry Standard Process for Data Mining* (CRISP-DM, Chapman et al., 2000). Der CRISP-DM beschreibt einen iterativen Prozess aus sechs Phasen: Geschäftsverständnis, Datenverständnis, Datenvorbereitung, Modellierung, Evaluation und Einführung. Als Vorgehensmodell ermöglicht CRISP-DM eine systematische und schrittweise Umsetzung datenanalytischer Projekte (West et al., 2021a, S. 131 ff.). Bereits in Kap. 3 wurde das Modell vorgestellt, wobei dort zusätzlich die Entwicklung von Daten über Informationen hin zu Wissen dargestellt ist. Im Forschungsprojekt wurde der Schwerpunkt zunächst auf die drei vorbereitenden Elemente des CRISP-DM gelegt:

- Data Understanding
- Data Preparation
- Modeling

Diese drei aufeinanderfolgenden Phasen, die auf der Datenbereinigung, der Vorverarbeitung und dem Aufbau eines *Data Dictionary* basieren, sind mit ständigen Iterationschleifen verbunden. Die hier investierte Zeit zahlt sich in den nachfolgenden Phasen der Modellanwendung und Validierung aus (Chapman et. al., 2000, S. 13 ff.). In den Prozessphasen *Data Understanding*, *Data Preparation* und *Modeling* wird iterativ und regelmäßig an der Evaluierung der Ergebnisse mit Abgleich zu den Projektzielen

gearbeitet. Auch in dieser Phase ist es immer wieder relevant, die Ergebnisse in Bezug auf den Geschäftsprozess zu erläutern.

11.2.2 Modellbildung und Validierung der Absatzprognose

Die mengen- und wertbasierte Prognose des Auftragseingangs erfolgt unter Zuhilfenahme von Zeitreihenanalysen. Abhängig von den Eingangsdaten wird meist zwischen uni- bzw. multivariaten Verfahren zur Prognose von Zeitreihen unterschieden (West et al., 2021b, S. 762 ff.), wobei Maschinelle Lernverfahren auch die Kombination qualitativer und quantitativer Inputdaten ermöglicht, beispielsweise durch sogenanntes *One-Hot-Encoding* (Cerdea & Varoquaux, 2020, S. 1 f.). Allgemein beschreibt eine Zeitreihe eine sequenziell erfasste Folge von Beobachtungen, wobei eine intrinsische Eigenschaft von Zeitreihen besagt, dass angrenzende Zeitreihenwerte in einem Zusammenhang stehen (Chatfield, 1982, S. 15). Diese implizierten Zusammenhänge bilden den Ausgangspunkt für Zeitreihenprognosen, die basierend auf den Abhängigkeiten Aussagen über das zukünftige, sequenzielle Verhalten einer Zeitreihe ermöglichen (Box et al., 2008, S. 1). Da das Kundenverhalten von deterministischen Faktoren beeinflusst wird, beispielsweise den abgesetzten Mengen vergangener Perioden oder von der aktuell vorherrschenden Wettbewerbssituation, werden diese Abhängigkeiten auch im weiteren Verlauf dieser Analyse impliziert.

Grundlage der industriellen Datenanalyse ist eine vorgegebene feste Datensequenz und die Betrachtung der Prognose als überwachtetes Lernmodell. In diesem Modell wird für eine definierte Menge von Eingangsdaten der Wert einer Zielvariablen vorhergesagt. Diese Zielvariable, auch als Label bezeichnet, wird durch die historischen Ausgabedaten des Vier-Monatsfensters bestimmt. Bei den Ausgabedaten handelt es sich um die tatsächlichen Auftragseingänge im Zeitraum des letzten 4-Monats-Fensters.

In Abb. 11.1 ist in blau die Größe der Zeitreihe für das Training des Modells dargestellt, der orange markierte Bereich beschreibt die Größe des Vorhersagezeitraums. Zur Bestimmung des Vorhersagewertes wurden drei Zeitreihen-Algorithmen in RapidMiner angewandt:

- **ARIMA** (Autoregressive Integrated Moving Average): ein statistisches Verfahren zur Zeitreihenanalyse, das sowohl Autoregression als auch bewegten Durchschnitt verwendet und häufig zur Vorhersage zukünftiger Werte verwendet wird
- **Holt-Winters**: eine Methode zur Vorhersage zukünftiger Werte von Zeitreihen, die saisonale Effekte sowie Trends berücksichtigt und durch die Verwendung von exponentieller Glättung berechnet wird
- **Functional and Seasonal Component**: ein Verfahren zur Zerlegung von Zeitreihendaten in funktionale und saisonale Komponenten, um Muster in der Datenreihe zu identifizieren und so Vorhersagen zukünftiger Werte zu erreichen

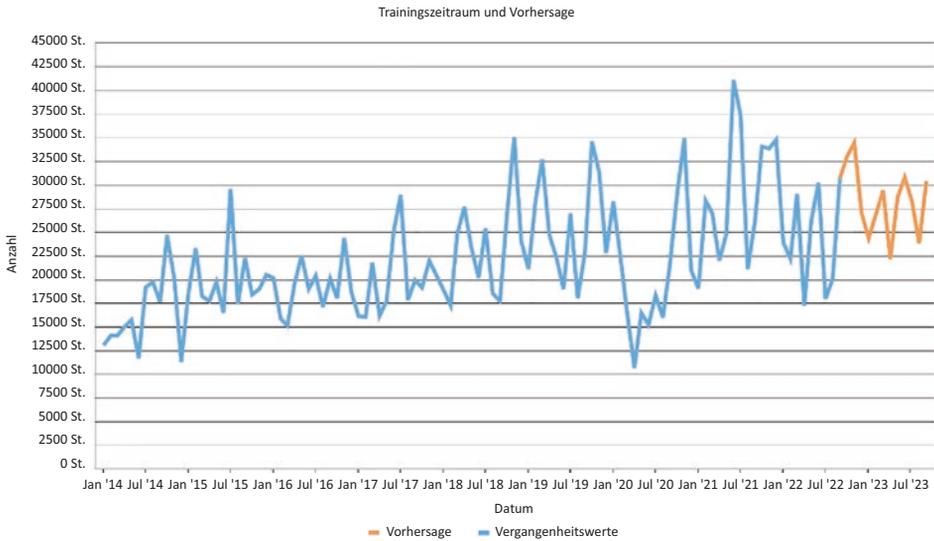


Abb. 11.1 Darstellung eines Zeitreihenfensters mit Vorhersagehorizont

Der Ergebnisvergleich der drei Methoden zeigte keine eindeutige Entscheidungsrichtung auf. Das wesentliche Kriterium für die Auswahl war die Abweichung von der tatsächlich gemessenen Auftragseingangskennzahl. Bei dieser Verifizierung zeigte Holt-Winters (blaue Linie) in Abb. 11.2 die geringste Abweichung vom vorgegebenen Ziel von 20 % Abweichung vom eingetroffenen Auftragseingang in Menge und ggfs. Wert. Somit wurde dieser Algorithmus in der Folge angewandt.

Für das Forschungsprojekt wurden die Daten aus dem ERP-System manuell monatlich übertragen und für die Prognoseprozesse bereitgestellt. Die Validierung der Ergebnisse ergab, dass eine Fokussierung in der Grundgesamtheit der Daten auf die Produktgruppe der „Spotlights“ eine sinnvolle und nachvollziehbare Einschränkung ergibt. Dieser Filter wurde innerhalb der Prognoseprozesse angewandt. In der weiteren Anwendung der Prognosemethode konnte eine visualisierte Ergebnisdarstellung für eine Gruppe von Produkten vorhergesagt werden.

Die Abb. 11.3 zeigt das Ergebnis der Prognose und die tolerierte Abweichung vom Zielkorridor 20 %. Wobei der blaue Graph die eingetretene reale Menge zeigt, der grüne Graph beschreibt die Prognose der Menge im jeweils aktuellen Ausführungsmonat und der orangene gepunktete Graph gibt die prozentuale Abweichung zum Zielkorridor an.

In vier Datenpunkten wurde der Zielkorridor bei der angewandten Prognosemethode deutlich unterschritten. Der Prognosewert war somit deutlich höher als der real eingetretene Wert. An dieser Stelle erfolgt außerhalb der Analyse eine Validierung der Datenqualität, um ggf. Extremwerte zu identifizieren, bzw. die Datenqualität zu verbessern. Nach der Intervention, die eine Verbesserung der Datenqualität bei der Bearbeitung von Projektinformationen zum Ziel hatte, konnte ab ca. August 2022

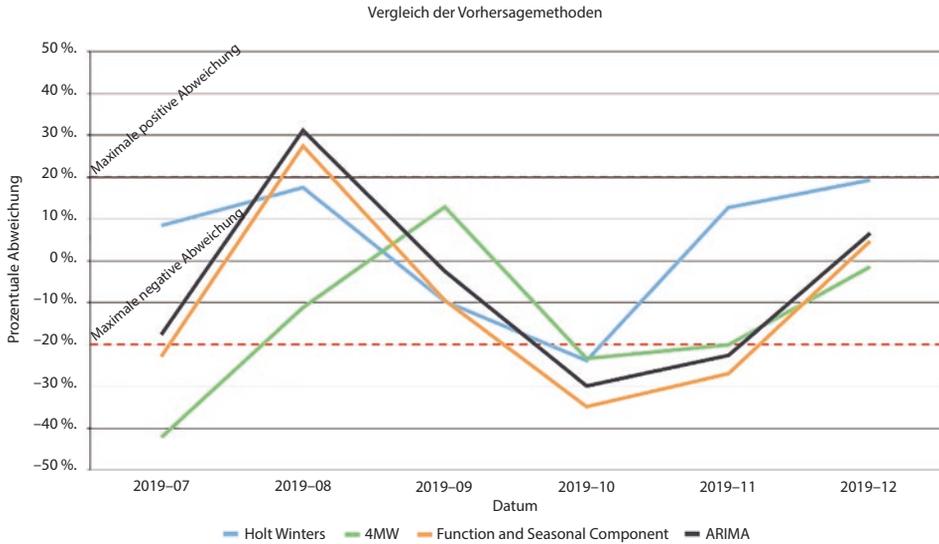


Abb. 11.2 Ergebnisdarstellung im Vergleich der der Prognosemethoden

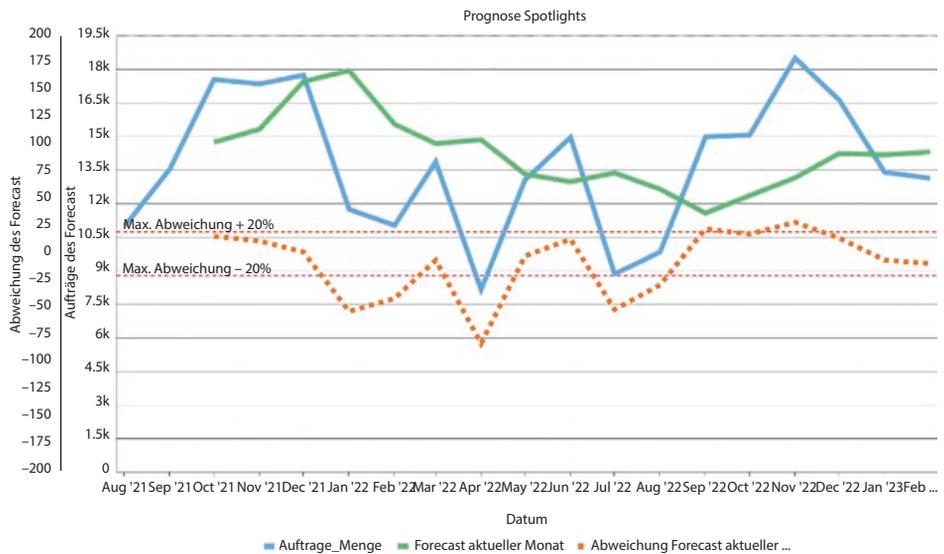


Abb. 11.3 Darstellung von Prognoseergebnis und Abweichung

eine Verringerung der Prognoseabweichung erkannt werden. Ein eindeutig kausaler Zusammenhang zwischen der Intervention und dem Prognoseergebnis ist jedoch noch nicht bewiesen.

11.2.3 Einführung der Prognose in die betrieblichen Entscheidungsprozesse

Die Einführung, häufig auch Englisch als *Deployment* bezeichnet, beschreibt das Vorgehen zum Ausrollen der industriellen Datenanalyse in die betriebliche Nutzung. Zur technischen Umsetzung gehört in diesem Fall die lokale Anwendung *RapidMiner Studio*, welche in der Verbindung mit der Web-Applikation RapidMiner AI-Hub einen Datenaustausch der Prognoseprozesse sowie deren Ergebnisse zwischen unterschiedlichen Anwendern sicherstellt. Eine Beschreibung dieser Werkzeuge findet sich in Kap. 6. Gemeinsam erarbeitete Ergebnisse können somit schnell verteilt und eingesetzt werden. Die Implementierung des Prozessablaufs für die Prognose wurde dafür in vier wesentliche Schritte aufgeteilt und in der Ausführungsreihenfolge automatisiert verknüpft. Die Reihenfolge ist:

1. Hinzufügen der Daten des abgeschlossenen Vormonats zur Analyse
2. Durchführung der Multi-Monats-Prognose für die Einheit Stück
3. Durchführung der Multi-Monats-Prognose für die Einheit Euro
4. Evaluierung der Prognose mit den tatsächlichen Auftragseingangswerten der Vergangenheit

Um für den Endanwender eine einfache Durchführung der Prognose zu ermöglichen, können die Schritte sequenziell automatisiert ausgeführt werden. Sie benötigen nur die Ablage der aktuellen Daten in einem dezidierten Format und Verzeichnis. Diese vier Schritte wurden zum Abschluss des Forschungsprojekts umgesetzt und können durch den Anwender autonom durchgeführt werden, die Ergebnisse dienen nun zur besseren Einschätzung der Anforderungen hinsichtlich zukünftiger Kapazitätsressourcen.

11.3 Ergebnisse Validierung des Geschäftsmodells

Im Rahmen des Forschungsprojekts bestand die Aufgabe, die vorhandenen Daten aus dem ERP System mit Informationen aus sozialen Netzwerken und oder der ERCO-Homepage zu verknüpfen. In Zusammenarbeit mit dem Konsortialpartner DFKI wurde das Geschäftsmodell des Unternehmens mit der Architektur des vom DFKI entworfenen Modells abgeglichen. Dies diente somit der Validierung des im Kap. 15 dargestellten Modells.

Das Unternehmen ERCO vertreibt seine Produkte und Dienstleistungen in der Regel an den Großhandel und Installateure, daher steht das Business-to-Business Geschäft im Vordergrund. Bei den vorhandenen Social-Media-Kanälen handelt es sich zumeist um geschlossene Benutzergruppen innerhalb der dezidierten Plattformen. Das Sammeln und Auswerten von Erfahrungsberichten oder Produktbewertungen über automatisierte Suchroutinen ist hier in der Regel nicht möglich. Die im Kap. 15 genannten Social-Streams konnten daher im vorliegenden Anwendungsfall nicht erhoben werden und wurden in der Projektbearbeitung nicht weiterverfolgt.

11.4 Data Science Reifegrad und Mittelstand

In einem mittelständischen Unternehmen werden Arbeitsprozesse häufig in einer Matrix der Aufbauorganisation zur Prozessorganisation umgesetzt bzw. implementiert. Die Mitarbeiter in der Aufbauorganisation arbeiten in der Regel an der Erstellung des von ihnen als relevant definierten Prozessergebnisses, dem sogenannten *Tagesgeschäft*. Analytische und prädiktive Aufgaben werden meistens bis zu einem gewissen Grad innerhalb des definierten Arbeitsvorrats zu erstellen und sind auch von der Kompetenz der Mitarbeiter abgedeckt.

Die Anwendung der Methoden im Data-Science-Umfeld und die Lösung der gestellten Problemaufgaben erfordern explizite Fachkompetenzen der ausführenden Person. Aber auch die Unternehmensorganisation muss Erfolgsfaktoren berücksichtigen, die erfolgreiche Data-Science-Projekte ausmachen. Die Beschreibung der Erfolgsfaktoren erfolgt im Kap. 16 des Forschungspartners *mosaiic*. Bei ERCO wurde die zugrunde liegende Reifegradanalyse auf Basis einer Umfrage mit 19 Teilnehmern aus unterschiedlichen Bereichen der Aufbauorganisation durchgeführt. Die Teilnehmer sind alle mit der Lösung analytischer Aufgaben beschäftigt und stellen einen Querschnitt der in diesem Aufgabenbereich tätigen Personen dar. Im Ergebnis zeigte sich ein hoher Reifegrad im Bereich des Willens, ein Data-Science-Projekt bzw. eine Aufgabenstellung zu lösen, aber auch ein eher niedriger Reifegrad im Bereich des „Könnens“, der Umsetzungskompetenz. Insgesamt konnte aus der Befragung ein eher „mittlerer Reifegrad“ abgeleitet und Handlungsempfehlungen für das Unternehmen gegeben werden, die den Erfolg von Data-Science-Projekten untermauern können.

11.5 Fazit

Im Forschungsprojekt AKKORD haben die Projektbeteiligten an der Lösung einer komplexen Prognoseaufgabe gearbeitet. Das angewandte CRISP-DM Modell bietet eine wichtige Richtschnur, um die Herangehensweise in Data-Science-Projekten zu unterstützen. Auch wenn die Prognose/das Projektergebnis nicht immer im definierten

Zielbereich lag, ist deutlich geworden, dass die Anwendung prädikativer Methoden unerlässlich ist, um einen andauernden und nachhaltigen Geschäftserfolg abzusichern.

Bedeutsam ist auch, dass der Kompetenzaufbau und die Etablierung von Data-Science-Projekten, mit dem Vorhandensein von freier Arbeitskapazität in der Aufbauorganisation zwingend einhergehen muss. Data-Science-Projekte benötigen Mitarbeitende die einen hohen Anteil Ihrer Ressource dem Projekt widmen müssen, die Vorkenntnisse als sogenannte *Citizen Data Scientist* sind hier essentiell, (siehe Kap. 10; Schwenken et al., 2023, S. 81 ff.). Sind diese Voraussetzungen geschaffen, wird es möglich sein, das Forschungsergebnis betrieblich anzuwenden sowie die Prognosemethoden und Algorithmen stetig zu verbessern und anzupassen. Eine weiterführende Untersuchung zur Entwicklung industrieller Datenanalysen, die Mensch, Technik und Organisation berücksichtigt, wird darüber hinaus in Kap. 20 vorgestellt.

Literatur

- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis. Forecasting and Control* (4. Aufl.). Wiley.
- Cerda, P., & Varoquaux, G. (2020). Encoding high-cardinality string categorical variables. *IEEE Transactions on Knowledge and Data Engineering*, 1–17. <https://doi.org/10.1109/TKDE.2020.2992529>.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step Data Mining Guide, CRISP-DM consortium.
- Chatfield, C. (1982). *Analyse von Zeitreihen. Eine Einführung*. Hanser.
- Schwenken, J., Klupak, C., Syberg, M., West, N., Walker, F., & Deuse, J. (2023). Development of a transdisciplinary role concept for the process chain of industrial data science. In A. Khanna, Z. Polkowski, & O. Castillo (Hrsg.), *Proceedings of data analytics and management. Lecture notes in networks and systems* (Bd. 572). Springer. https://doi.org/10.1007/978-981-19-7615-5_7.
- West, N., Syberg, M., Deuse, J. (2022). A Holistic Methodology for Successive Bottleneck Analysis in Dynamic Value Streams of Manufacturing Companies. In: *Proceedings Changeable, Agile, Reconfigurable and Virtual Production Conference*, 8(1), 612–619. https://doi.org/10.1007/978-3-030-90700-6_69.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021a). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- West, N., Schlegl, T., & Deuse, J. (2021b). Feature extraction for time series classification using univariate descriptive statistics and dynamic time warping in a manufacturing environment. *IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering*, 762–768. <https://doi.org/10.1109/ICBAIE52039.2021.9389954>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Potenzialanalyse industrieller Datenanalysen in der Produktion am Beispiel des Kunststoffschweißens

Umsetzungsszenario für Neuanlagen oder im Retrofit bei der automatisierten, industriellen Fertigung

Tobias Beiß , Edin Klapic , Tina Hardt , Jörn Schwenken  und Jochen Deuse 

12.1 Motivation

Beim Schweißen von Kunststoffen wird der Kunststoff im Schweißbereich der Bauteilhälften in einen schmelzförmigen Zustand gebracht, der dann durch Druck zusammengepresst und abgekühlt wird, um eine Schweißnaht zwischen den Bauteilhälften zu erzeugen. Die Anforderungen an Schweißverbindungen sind vielfältig und umfassen u. a. hohe Festigkeit sowie Dichtheit gegen Gase und Flüssigkeiten. In der Serienfertigung werden Kunststoffbauteile häufig nach dem Urformen im Spritzgießprozess durch Schweißen finalisiert. Dabei muss der Schweißprozess alle Toleranzen und Fehler

T. Beiß (✉)

PolyMerge GmbH, Geretsried, Deutschland

E-Mail: tobias.beiss@polymerge.de

E. Klapic

RapidMiner GmbH, Dortmund, Deutschland

E-Mail: eklapic@rapidminer.com eklapic@altair.com

T. Hardt

Real-Time-Systems GmbH, Wittlich, Deutschland

E-Mail: tina.hardt@real-time-systems.com

J. Schwenken · J. Deuse

Technische Universität Dortmund, Dortmund, Deutschland

E-Mail: joern.schwenken@ips.tu-dortmund.de

J. Deuse

E-Mail: jochen.deuse@ips.tu-dortmund.de

bewältigen, die aus dem Material, der Lagerung und den Vorverarbeitungsschritten resultieren, um ein spezifikationsgemäßes Bauteil zu erzeugen. Dafür steht ein gewisses Prozessfenster zur Verfügung, das während des Ablaufs eingehalten werden muss.

Der Erfolg des Schweißprozesses bzw. die Qualität der Schweißverbindung wird derzeit fast ausschließlich durch das Abfragen statischer Werte bestimmt. Dabei werden beispielsweise das Erreichen einer bestimmten Achsposition (Fügeposition) oder bestimmter Prozesskräfte abgefragt und in ein Toleranzfenster eingegeben. In Fällen komplexer Schweißvorgänge, z. B. mit vielen einzelnen Heizzonen bei großen Bauteilen, gehen jedoch viele Informationen aus dem Prozess verloren, und die erreichten Endwerte sind nur repräsentativ für das gesamte Bauteil. Es kann lokal zur Entstehung von Fehlern kommen, die anhand der detektierten Prozess-Endwerte nicht erkannt werden können. Die Idee des vorliegenden Anwendungsbeispiels besteht darin, mittels vollständiger Datenaufzeichnung und Inline-Auswertungen Anomalien zu erkennen, die durch die Kontrolle fixer Endkennwerte nicht erkannt werden können. Die Erkennung von Anomalien mittels Datenanalysen gilt insbesondere im Qualitätsmanagement als ein relevanter Anwendungsfall industrieller Datenanalysen (Schlegl et al., 2022, S. 1 ff.; West et al., 2021a, b, S. 762). Im Sinne der Prozesskette der industriellen Datenanalyse wird in diesem Kapitel also ein vollständiger Durchlauf der Schritte Datenzugriff, -analyse und -nutzung angestrebt (siehe Kap. 4).

12.2 Ablauf des Kunststoffschweißens

Das Schweißen von Kunststoffen erfolgt auf unterschiedliche Weise, abhängig von der Wärmebringung. Beispielsweise kann der Kunststoff mittels direkten Kontaktes mit einem heißen Werkzeug, einer Konvektionsströmung, Infrarotstrahlung oder durch Reibung aufgeschmolzen werden (Baur et al., 2022, S. 369; DIN 1910 Teil 3). Der Anwendungsfall entstammt dem Forschungsprojekt AKKORD und behandelt ein Beispiel industrieller Datenanalysen das sich mit dem Heizelementschweißen (HE) beschäftigt (siehe Kap. 1). Der in diesem Kapitel behandelte Anwendungsfall wurde bereits kurz in Kap. 2 eingeführt. Dabei wird der Schweißsteg des Kunststoffbauteils mit einem heißen, metallischen Heizelement in direkten Kontakt gebracht.

Der Schweißzyklus umfasst eine Angleichphase, eine Anwärmphase, eine Umstellphase, eine Fügephase und eine Abkühlphase (Ehrenstein, 2004, S. 179 ff.). Da Kunststoffbauteile meist uneben oder verzogen sind, wird zu Beginn des Schmelzvorgangs durch die bewegliche Achse der Bauteile ein kleiner Teil des Schweißsteges abgeschmolzen und verdrängt, sodass es zu einem vollflächigen Kontakt des Schweißsteges mit dem Heizelement kommt (Angleichphase). Anschließend verweilen die Bauteilhälften auf dem Heizelement, bis ein definiertes Schmelzpolster erreicht ist (Anwärmphase). Nachdem das Heizelement entfernt ist (Umstellphase), werden die Bauteilhälften zusammengepresst, wobei ein Teil der Schmelze in einen Schweißwulst

verdrängt wird und die Bauteilhälften miteinander den Stoffschluss durch Abkühlen ausbilden (Fügephase und Abkühlphase).

12.2.1 Prozessparameter und Datenquellen beim Schweißen

Um den Schweißprozess reibungsfrei durchzuführen, müssen verschiedene Randbedingungen und Parameter innerhalb eines für das Material und die Aufgabe geeigneten Prozessfensters liegen. Bei der *PolyMerge GmbH* wird Expertenwissen über den Prozess genutzt, um Erkenntnisse aus der Datensammlung für das Kunststoffschweißen zu gewinnen.

Für eine erfolgreiche Schweißung muss das Heizelement eine bestimmte, materialabhängige Temperatur aufweisen. Das Temperaturfenster variiert je nach dem zu verarbeitenden Kunststoff zwischen 120 und 480 °C. Im Rahmen des Anwendungsfalles werden die Temperaturen kontinuierlich aufgezeichnet, um sie für die Analyse zu nutzen.

Die Achsen der Maschine, die das Heizelement und die Bauteilhälften der Schweißteile in die jeweiligen Positionen bewegen und gegebenenfalls Druck ausüben können, werden durch mehrere Parameter überwacht. Die Zeit, die benötigt wird, um eine bestimmte Position zu erreichen, sowie das Auslesen des elektrischen Stroms (Servo-Antrieb), der für die Bewegung erforderlich ist, enthalten zusätzliche Informationen über die Reibungskräfte von Führungen und Antrieben. Für elektrische Antriebe ist es darüber hinaus möglich, Beschleunigungs- und Verzögerungswerte auszuwerten. Bei pneumatischen Antrieben könnten die Zeit für die Bewegung, der Druck und der Luftverbrauch gemessen werden. In aktuellen Maschinen wird dies oft nicht detailliert überwacht, was das Anwendungsbeispiel potenziell neue Erkenntnisse liefern lässt. Die Daten können für die einzelnen Prozessphasen folgende Aussagen liefern und auf Probleme im Prozess hinweisen

- **Angleichphase:** Eine kombinierte Weg-Zeit-Messung überwacht, ob das Erreichen der Angleichposition immer gleich lang dauert. Eine Verzögerung könnte auf Probleme beim Antrieb (schwergängig), beim Material (falscher Kunststoff) oder bei der Heizelementtemperatur (zu niedrig) hinweisen. Ein zu schnelles Erreichen der Position deutet auf eine eventuelle Materialverwechslung oder eine zu hohe Heizelementtemperatur hin. Die Überwachung der Kraft oder der benötigten Antriebsenergie (Strom oder Luftdruck) ermöglicht eine weitere Einschränkung potenzieller Probleme. Wenn die benötigte Kraft bei einer verkürzten Angleichzeit geringer ist als bei den Vergleichsdaten, könnte die Viskosität des Kunststoffes niedriger sein (falsches Material oder zu hohe Temperatur).
- **Anwärmphase:** Direkt im Anschluss an die Angleichphase wird die Kraft an den Antrieben reduziert, damit die Wärme den Kunststoff in der Tiefe aufschmelzen kann, ohne die Schmelze weiter zu verdrängen. Der Verlauf der Anpresskraft in dieser Phase kann Aussagen über das Relaxationsverhalten des Kunststoffes, während das Material

erweicht wird und schmilzt, liefern. Es ist jedoch aus dem bisherigen Erfahrungsschatz über das Kunststoffschweißen nicht gesichert, ob sich in dieser Phase Auffälligkeiten erkennen lassen, die eine Aussage über die Qualität der Schweißnaht erlauben.

- **Umstellphase:** Beim Umstellen werden die geschmolzenen Schweißstege der Bauteilhälften vom Heizelement abgelöst. Eine Kraftmessung ermöglicht die Messung und Überwachung der Ablösekraft. Durch eine Kombination mit der Wegmessung kann ein Kraft-Weg-Verlauf aufgezeichnet und analysiert werden. Es gibt jedoch keine Gewissheit darüber, ob diese Daten Aussagen über die Eigenschaften des Prozesses enthalten. Daher ist es im Rahmen des Projekts interessant, weitere spezifische Daten zu erfassen und mithilfe des vorhandenen Expertenwissens zu analysieren.
- **Fügephase:** In dieser Phase werden die geschmolzenen Schweißstege unter Druck zusammengepresst, um den Stoffschluss in der Schweißverbindung zu erzeugen. In dieser Phase werden ebenfalls Kraft-Weg- und Kraft-Zeit-Verläufe aufgezeichnet und analysiert. Hier können Einflüsse von Antrieben, Führungen, Temperatur der Schmelze und Werkstoffcharakteristika erkannt und charakterisiert werden.

12.3 Nachrüstung zur Datenerfassung mittels ARENDAR

Zur Erfassung der Daten aus dem Prozess wurde das IoT-Edge-Gateway ARENDAR eingesetzt. Dieses erfasst Werte aus dem Prozess und übergibt sie an Datenbanken. Zudem kann über eine Webapplikation die Datenaufzeichnung gestartet und aktuelle Werte überwacht werden (siehe Kap. 14). Diese Nachrüstung bestehender Anlagen wird auch als *Retrofitting* bezeichnet (Wöstmann et al., 2019, S. 94 f.) und weiterführend beschrieben. Die weiteren Schritte in dieser Arbeit sind an die Prozesskette der Industriellen Datenanalyse angelehnt (siehe Kap. 4) und betrachten im Folgenden die Erfassung, Analyse und Darstellung von Daten aus einem Anwendungsfall des Kunststoffschweißens.

12.3.1 Datenakquisition

Der ARENDAR ist in einer Testanlage von PolyMerge mit einer Kraftmesszelle (analog) und mit der Speicherprogrammierbaren Steuerung, kurz SPS, (digital, LAN) verbunden. Über das Protokoll S7 werden 15 Werte ausgelesen. Im Wesentlichen handelt es sich dabei um Boole'sche Werte, die bestimmte Prozessphasen der Schrittkette anzeigen. Zu Beginn wurden deutlich mehr Datenpunkte als nötig adressiert, um festzustellen, welche für die Triggerung der Aufzeichnung geeignet sind. Nach einigen Versuchen wurden

folgende Datenpunkte als relevant für die automatisierte Prozessdatenaufzeichnung identifiziert:

- Startsignal zur Triggerung der Aufzeichnung über den Startknopf der Maschine (Bool)
- Wegsignal des Servoencoders am Achsantrieb des Oberwerkzeugs (Digital, Integer)
- Stromaufnahme des Servoantriebes in der SPS intern umgerechnet in einen Kraftwert (Digital, Integer)
- Kraftsignal einer in das Oberwerkzeug eingebauten Kraftmesszelle (Analog, Integer)
- Zwei Temperatursignale aus den Heizelementen für die obere und die untere Bauteilhälfte (Digital, Integer)
- Signal für Prozessende „EndeFügen“ aus der SPS-Schrittkette (Bool)

Die Daten werden mit einer Abtastrate von 100 Millisekunden erfasst. Ein limitierender Faktor für eine schnellere Erfassung ist dabei hauptsächlich die CPU der SPS beziehungsweise die Anzahl der erfassten Werte. Der ARENDAR führt alle erfassten Werte in einem JSON-Objekt zusammen und gibt die Daten an eine Influx-Datenbank zur permanenten Speicherung weiter. Es stehen zwei Datenbanken zur Auswahl: Eine für Bauteile, die ohne erzeugte Bauteilfehler in den Prozess gegeben wurden und eine für Teile, welche mit Fehlern präpariert wurden. Durch die Triggerung wird die Aufzeichnung zwischen den Schweißungen pausiert, um die Datenmenge zu reduzieren.

12.3.2 Webapplikation

Zur nutzerfreundlichen Bedienung der Datenaufzeichnung, zur Überwachung aktueller Werte und zur Darstellung erfasster Werte in Line-Charts wurde eine Webapplikation erstellt. Diese besteht aus derzeit zwei Seiten: Eine für die Steuerung der Messung und die Echtzeitdarstellung der Werte und eine zur Darstellung einzelner Messungen in Form von Line-Charts, auf der gezielt Messungen anhand des Namens ausgewählt und dargestellt werden können.

Abb. 12.1 zeigt exemplarisch die erste Seite mit der Darstellung aktueller Werte (grau hinterlegter Bereich) samt der Benutzeroberfläche (weiß hinterlegter Bereich) mit der Möglichkeit dem Test einen Namen zu geben, eine Datenbank für die Testergebnisse auszuwählen sowie einen Namen und einen Kommentar zum Test in der Datenbank zu hinterlegen. Die Aufzeichnung der Daten wird über die Schaltfläche „Aufzeichnung starten“ initiiert. Dadurch wird gleichzeitig ein Zeitstempel für die Aufzeichnung in der Datenbank hinterlegt.

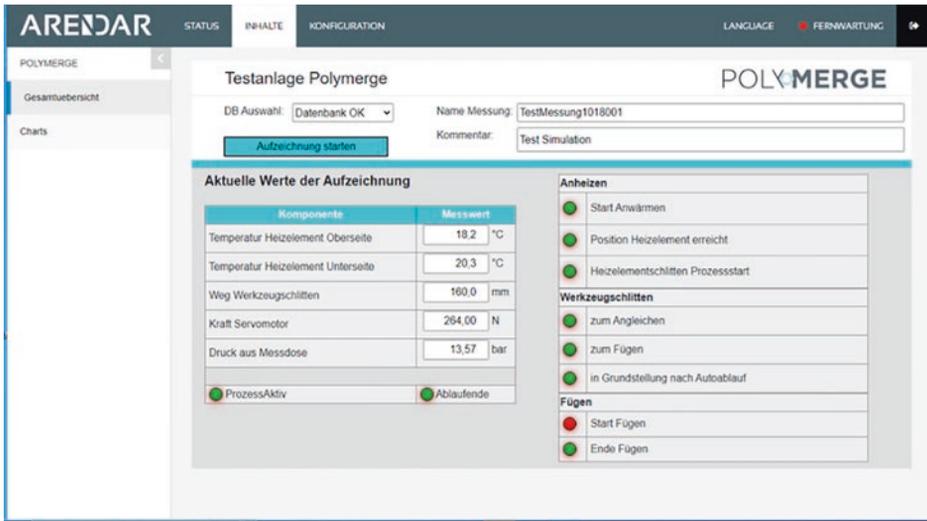


Abb. 12.1 Webapplikation zum Prozessstart mit Gesamtübersicht

12.4 Datenauswertung

Die Datenauswertung eines typischen Prozessablaufs, wie in Abschn. 12.2.1 beschrieben, ist in Abb. 12.2 dargestellt. Die einzelnen Prozessphasen sind gut voneinander abzugrenzen. Zum Prozessstart bekommt der Servoantrieb (grün) einen starken Impuls, um in die Angleich- und Anwärmposition am Heizelement zu fahren. Die Position der Schweißteile (schwarz) am Heizelement bei ca. 118 mm zeigt sich in einem Peak der Kraft in der Messzelle (hellblau). Die Oberfläche des Kunststoffteils wird zunächst mit erhöhter Kraft auf das Heizelement gedrückt, um Oberflächenunebenheiten abzuschmelzen und einen vollflächigen Kontakt herzustellen. Im Anschluss wird die Kraft reduziert und der Kunststoff mittels Wärmeleitung weiter aufgeschmolzen. Anschließend wird das Heizelement entfernt und die geschmolzenen Bauteiloberflächen werden unter Druck gefügt und abgekühlt. Der Regler des Servoantriebes wird dabei permanent angesteuert, um eine voreingestellte Fügeposition zu erreichen und die gewählte Fügekraft dort zu halten.

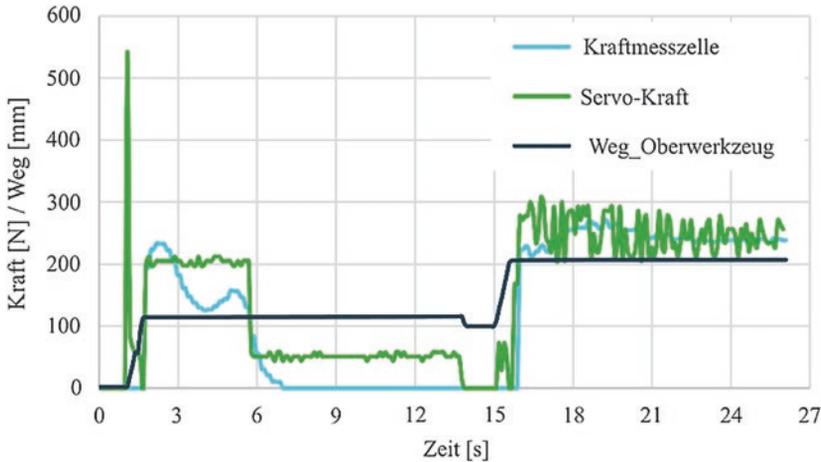


Abb. 12.2 Zeitabhängiger Prozessverlauf beim Kunststoffschweißen

12.5 Diskussion und Interpretation der Ergebnisse

Nach der Datenauswertung der Versuche konnte ein Programmierfehler in der Maschine identifiziert werden. Normalerweise läuft der Prozess automatisch ab, einschließlich des automatischen Entspannens des geschweißten Bauteils am Ende des Prozesses und dem Zurückkehren der Maschine in die Ausgangsposition. Allerdings erfordert das verwendete Werkzeug einen manuellen Schritt zum Lösen der Spannkraft. Die Maschine muss in der Fügeposition verbleiben, bis der Startknopf betätigt wird, bevor sie in die Ausgangsposition zurückkehrt. Durch die Messungen wurde festgestellt, dass der Antrieb nach Ablauf der Abkühlzeit nicht in der Lageregelung verbleibt, sondern die Kraft stark erhöht wird (siehe Abb. 12.3). Als Reaktion darauf wurde die Schrittkette angepasst und der Fehler behoben.

Durch die Analyse der aufgezeichneten Daten konnte festgestellt werden, dass die Einstellung der Fügeposition zu Beginn nicht korrekt war. Nach Abschluss des Prozesses erreicht die Achse des Oberwerkzeugs eine vorgegebene Position, woraufhin sich der Antrieb ausschaltet und die Kraft am Servoantrieb sofort reduziert wird. Aufgrund der Schwindung, die durch das Abkühlen und Erstarren des Kunststoffes entsteht, nimmt auch die Kraft in der Kraftmesszelle am Oberwerkzeug schnell ab (siehe Abb. 12.4). Dadurch erstarrt der Kunststoff ohne Rest-Fügedruck, was oft zu einer schlechten Schweißnahtqualität führt, da sich Fehlstellen in der Schweißnaht bilden können.

Trotz korrigierter und bereinigter Daten konnten in der automatisierten Datenanalyse nur geringfügige Unterschiede zwischen den untersuchten Proben festgestellt werden. Tendenzielle Abweichungen in der Angleich- und Anwärmphase sowie in der Fügephase sind erkennbar. Diese sind jedoch nicht signifikant genug, damit sie automatisch

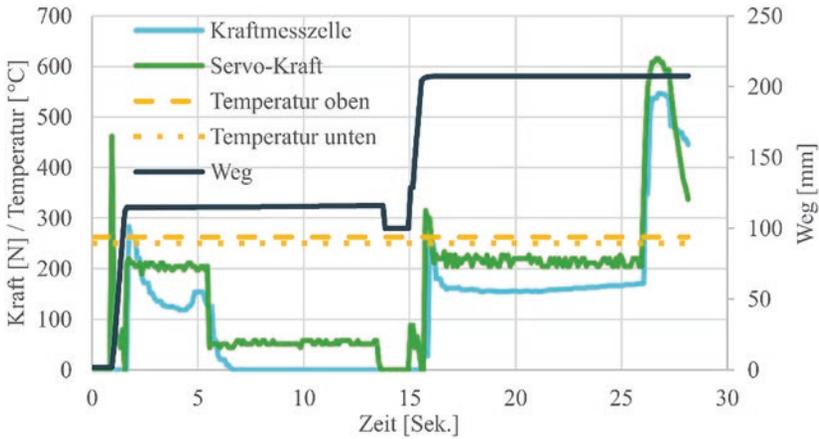


Abb. 12.3 Aufgezeichnete Daten eines ersten Schweißprozesses mit fehlerhafter Funktion der Maschine

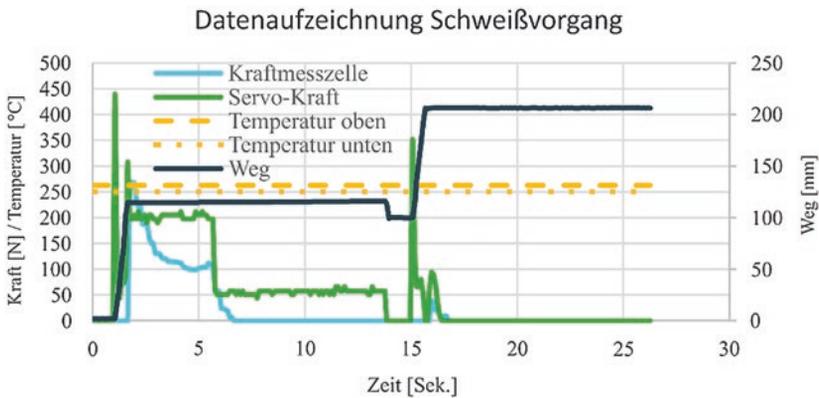


Abb. 12.4 Fügeweg zu gering eingestellt

von KI-Technologie erkannt werden. Eine Ausnahme stellen hier die Proben aus einem anderen Kunststoff-Material dar. Sie zeigen erkennbare Unterschiede bei der Schmelzebildung und dem Fügeweg. Daher ist eine tiefere Analyse der Prozessdaten mit hoher Aufzeichnungsfrequenz weiterhin lohnend. Ein Hinweis auf weitere Effekte ist bereits in der Analyse der Servo-Kraft entlang der Zeitachse zu erkennen (siehe Abb. 12.5). Vor allem in der Fügephase zeigt die Servo-Kraft unterschiedliche Verläufe abhängig von der Fehlerart. So zeigt sich deutlich in Grau, dass das Material 2 mit einem höheren Kraftbedarf gefügt werden muss als das erste Material. Während der Fügephase ist der Servo-Kraft-Verlauf von Gutteilen (blaue Kurve) insgesamt niedriger als bei Fehlteilen. Auch zwischen den Fehlerarten sind Unterschiede im Verlauf zu erkennen.

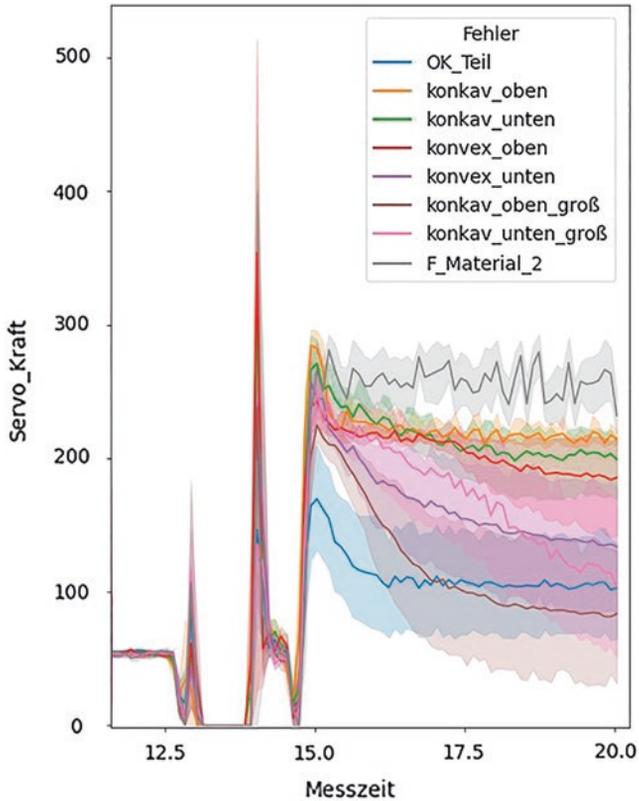


Abb. 12.5 Mittelwertkurven der Servo-Kraft in der Fügephase mit Abweichungskorridor

Zur Darstellung der Messungen wurde auf der AI-Toolbox ein Analysemodul hochgeladen (siehe Abb. 12.6). Dieses visualisiert den Kraft-Weg-Druck-Verlauf der einzelnen Messungen. Anwender:innen können sich über die Messungsnummer die zu analysierende Messung ausgeben. Das Modul soll dabei helfen, dass sich Anwender:innen einen schnellen Überblick über die Messung verschaffen können, damit sie ggf. erste Auffälligkeiten entdecken können.

Bei der beispielhaften Messung handelt es sich um eine fehlerhafte Schweißung. Die Messung zeigt nur im hinteren Verlauf Auffälligkeiten. Es ist zu sehen, dass die Servo-kraft (grün) und die Druckmessdose (blau) nach oben ausschlagen. Dies ist zurückzuführen darauf, dass der Weg des Schlittens scheinbar minimal zugenommen hat (orange). Diese geringen Abweichungen in der Steuerung können zu einem fehlerhaften Teil führen.

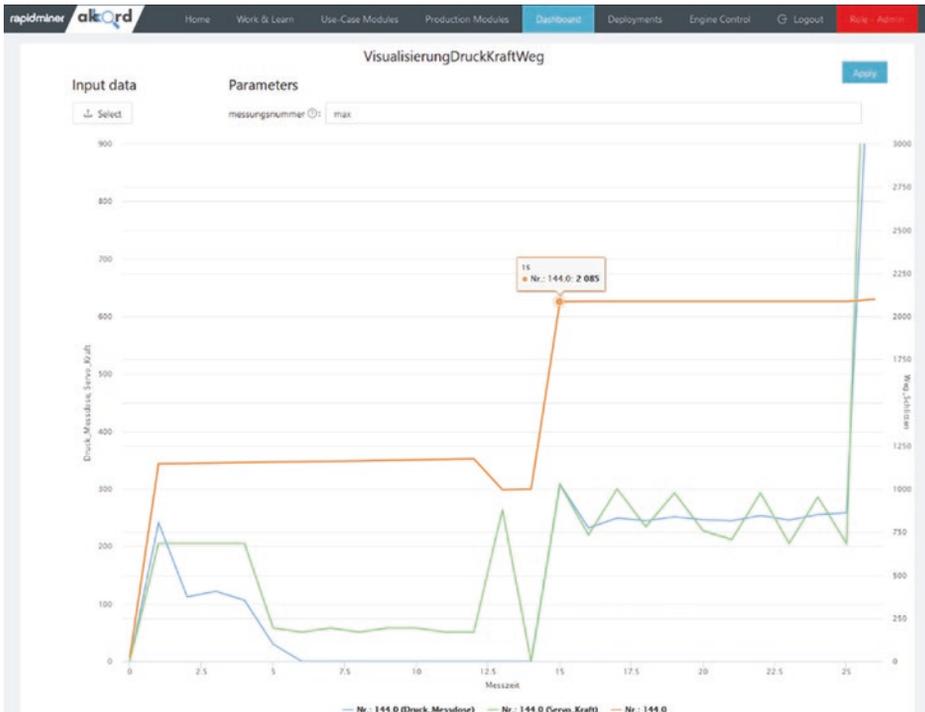


Abb. 12.6 Analysemodul zur Selektion und Darstellung von Einzelprozessen

12.6 Fazit

Die automatisierte Sammlung von Daten einer industriellen Produktionsmaschine hat schon bei der manuellen Auswertung zahlreiche Schwächen im Prozess aufgezeigt, die dadurch behoben werden konnten. Die automatisierte Erkennung von Anomalien oder Fehlern im Prozess funktionierte aufgrund des möglichen Versuchsumfangs und der Parametervariation im Labormaßstab nicht zuverlässig. Doch auch hierbei konnten bei entsprechender Darstellung und kombiniert mit Expertenwissen Unterschiede in den Prozessabläufen erkennbar gemacht werden. Es wird erwartet, dass insbesondere bei den Daten aus industrieller Serienproduktion durch die entsprechende Stückzahl deutlich mehr Effekte erkennbar sind. Eine weitere Möglichkeit stellt die erweiterte Betrachtung einer durchgängigen Datenanalysequalität im Prozess dar (West et al., 2021a, b, S. 131 ff.). Es ist geplant, mit dem Referenzkunden aus dem Projekt weiterhin Daten der Anlage zu erfassen und zu analysieren. Abschließend sei für eine ausführlichere Betrachtung der Weiterentwicklung industrieller Datenanalysen im Einklang mit Mensch, Technik und Organisation auf die zusammenfassenden Arbeiten in Kap. 20 hingewiesen.

Literatur

- Deutsches Institut für Normung e. V., (1977). *DIN 1910 Teil-3:1977-09 Schweißen; Schweißen von Kunststoffen, Verfahren*. Beuth.
- Gehde, M. (2022). Fügen In E. Baur, T. A. Osswald, & N. Rudolph (Hrsg.), *Saechtling Kunststoff Taschenbuch* (32. Aufl.). Hanser
- Gehde, M. (2004). Heizelementschiweißen. In G. W. Ehrenstein (Hrsg.), *Handbuch Kunststoffverbindungstechnik* (S. 1789–196). Hanser.
- Schlegl, T., Schlegl, S., Tomaselli, D., West, N., & Deuse, J. (2022). Adaptive similarity search for the retrieval of rare events from large time series databases. *Advanced Engineering Informatics*, 52, 101629. <https://doi.org/10.1016/j.aei.2022.101629>.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International Conference on Information Reuse and Integration for Data Science, (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- Wöstmann, R., Barthelmey, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. In T. Schüppstuhl, K. Tracht, & J. Roßmann (Hrsg.), *Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10.
- West, N., Schlegl, T., & Deuse, J. (2021). Feature extraction for time series classification using univariate descriptive statistics and dynamic time warping in a manufacturing environment. *IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering* (S. 762–768). <https://doi.org/10.1109/ICBAIE52039.2021.9389954>.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Kompetenzentwicklung und -sicherung mit einem digitalen Wissensdienst

13

Konzipierung und Entwicklung eines
auf Konzepten und Methoden zum
Kompetenzaufbau und zur Kompetenzsicherung
basierenden digitalen Wissensdienstes

Volker Zimmermann, Rebekka Adams und Christopher Klupak 

13.1 Motivation

Die Verbreitung moderner Informations- und Kommunikationstechnologien sowie die technologischen Möglichkeiten zur systematischen und umfassenden Erfassung von Daten ermöglichen den Aufbau dynamischer Informationsspeicher bisher unbekannter Größe und Güte (Eickelmann et al., 2015, S. 738 f.). Dies hat zur Folge, dass die Analyse und Interpretation ebendieser Daten sowie die effiziente Nutzung des generierten Wissens zunehmend in den Fokus von Unternehmen rücken (Deuse et al., 2014, S. 373 ff.; West et al., 2022a, S. 3 ff., b, S. 615 f.). Dem steht jedoch ein Mangel an erforderlichen Kompetenzen und Einführungsstrategien sowie an strategisch ausgerichteten, praktikablen Dienstleistungs- und Technologieangebote gegenüber, um die massiven Potenziale moderner Methoden der Datenvernetzung und -analyse nutzen zu können.

Die Schaffung einer vollständigen, vernetzten und integrierten Datenbasis, die Aufbereitung der Daten und die Verknüpfung komplexer Datenstrukturen erfordern darüber hinaus neben einem großen Zeit- und Ressourcenaufwand vor allem auch differenzierte

V. Zimmermann (✉) · R. Adams
NEOCOSMO GmbH, Saarbrücken, Deutschland
E-Mail: volker@neocosmo.de

R. Adams
E-Mail: rebekka@neocosmo.de

C. Klupak
Universität Hamburg, Hamburg, Deutschland
E-Mail: christopher.andre.klupak@uni-hamburg.de

und tiefgreifende Kompetenzen der Mitarbeitenden. Daher ist es von großer Wichtigkeit, Methoden und Plattformen zur effizienten Kompetenzvermittlung von Beginn an mitzugestalten und hier moderne Lösungen anzubieten, einzuführen und nachhaltig zu etablieren.

Die Konzipierung und Entwicklung eines digitalen Wissensdienstes für den Themenbereich Industrielle Datenanalyse, der jederzeit von Mitarbeitenden in KMU zum Kompetenzaufbau genutzt werden kann, wurde deshalb zu einem wichtigen Ziel im Rahmen des AKKORD-Projekts (siehe Kap. 1 und 2). Während der Leistungsbereich sich primär mit der inhaltlichen und fachlichen Konzeption beschäftigt (siehe Kap. 6), wird in diesem Anwendungsfall der Fokus auf die technische Umsetzung des digitalen Wissensdienstes gelegt. Im Kapitel wird die im Rahmen des Anwendungsfalls entwickelte und umgesetzte Work&Learn-Plattform zur Kompetenzentwicklung und -sicherung genauer dargestellt sowie auf deren Erprobung in der praktischen Anwendung und auf mögliche Transfermaßnahmen eingegangen.

13.2 Umsetzung des Anwendungsfalls mit den entwickelten Ergebnissen

Dass Industrie 4.0 es erfordert, die Kompetenzprofile von Mitarbeitenden auf unterschiedlichen Ebenen weiterzuentwickeln (Spöttl et al., 2016, S. 4 ff.; Arnold et al., 2016, S. 1), ist eine wichtige Erkenntnis, die für die Umsetzung des Anwendungsfalls „KMU-gerechte Kompetenzentwicklung für industrielle Datenanalyse“ relevant ist. Außerdem ist von Relevanz, dass der informelle Wissenserwerb zwischen 70–90 % der Kompetenzentwicklung in der betrieblichen Realität ausmacht (Cerasoli et al., 2018, S. 203 f.), gerade wenn es sich um neue und schnell entwickelnde Themen handelt. Um eine entsprechende digitale Plattform zur Verfügung stellen zu können, wurden zunächst messbare Kompetenzen definiert und Messinstrumente entwickelt. In der Folge konnten im Rahmen des Projekts eine digitale Serviceplattform entwickelt, technologische Lernmodule generiert und verschiedene Kollaborationsmöglichkeiten innerhalb der Plattform integriert werden. Abschließend wurde die Plattform durch Testläufe validiert und optimiert.

13.2.1 Definition und Ermittlung datenwissenschaftlicher Kompetenzen

Der effiziente und effektive Einsatz von Datenanalysetechnologien stellt besondere Anforderungen an die Kompetenzen in datenwissenschaftlichen Rollen. Entsprechend besteht einerseits die Notwendigkeit zur Bereitstellung situierter Lernmodule zur anwendungsbezogenen, individuellen Kompetenzentwicklung. Andererseits ist die Entwicklung von Strukturen und praxisorientierten Handlungsempfehlungen erforderlich, die die Erfassung von Kompetenzen und die anwenderübergreifende Vernetzung

ermöglichen. Dabei sind objektive Leistungstests einer generellen Selbsteinschätzung vorzuziehen (Nickolaus et al., 2016, S. 12 f.), weshalb im Rahmen des Projekts die Möglichkeit zur Absolvierung von objektiven Abschlusstests entwickelt und bereitgestellt wurde.

Generell ist davon auszugehen, dass bei weitem nicht alle Nutzer:innen der Plattform einen IT-Hintergrund haben werden und dass sich der Großteil der potenziellen zukünftigen Nutzer:innen bisher noch nicht ausführlich mit dem Themenbereich der Industriellen Datenanalyse beschäftigt haben. Deshalb besteht die Notwendigkeit, dass sowohl grundlegende Kurse für Anfänger:innen als auch für Fortgeschrittene auf der Plattform angeboten, eine übersichtliche Darstellung gewährleistet und eine Suchfunktion zur Unterstützung bereitgestellt werden sollten. Außerdem sind schnelle und einfache Kommunikationswege von Bedeutung: Fragen sollten direkt in der Plattform und im entsprechenden Inhalt gestellt werden können. Zudem wird ein Bedarf an Kontaktmöglichkeiten zu Expert:innen erforderlich sein, die im Betrieb des jeweiligen Nutzenden unter Umständen nicht oder nur marginal gegeben sind.

Auf Literaturbasis konnten Kompetenzanforderungen ermittelt und zentrale Kompetenzrollen in Form von sogenannten *Personas* (Data Engineer, Data Scientist, Domänenexpert:in, Projektmanager:in, Geschäftsführer:in) definiert werden, worauf in Kap. 7 genauer eingegangen wird. Neben dem Anfänger- (Basic-) und dem Fortgeschrittenen- (Advanced-)Kurs sowie Zusatzlektionen und Best-Practice-Trainings wurden entsprechend Kompetenzrollenbezogene Kurspfade zusammengestellt, die von Nutzer:innen der Plattform absolviert und mit einem Abschlusstest abgeschlossen werden können. Das Gesamtergebnis der einzelnen Nutzer:innen für einen solchen rollenbezogenen Abschlusstest kann von entsprechend berechtigten Administrator:innen im Backend eingesehen werden. Basierend auf dem Ergebnis wird den Absolvent:innen ein Zertifikat für die erfolgreiche Beendigung eines Kompetenzrollenbezogenen Kurses ausgestellt.

13.2.2 Entwicklung eines digitalen Wissensdienstes

Um Wissen, Kollaboration und Kompetenzentwicklung im Themenbereich Industrielle Datenanalyse für möglichst viele Nutzer:innen in Unternehmen einfach zugänglich zu machen, bietet sich die Entwicklung eines digitalen, plattformbasierten Ansatzes an. Durch die Entwicklung einer digitalen Plattform wird zudem der verstärkten Nachfrage nach Lösungen nachgegangen, die einen zeit- und ortsunabhängigen Kompetenzaufbau gestatten (Syberg et al., 2023, S. 64 ff.).

Der im Rahmen des Projekts konzipierte und entwickelte Wissensdienst für den Themenbereich Industrielle Datenanalyse hat sich im Projektverlauf durch die Integration verschiedener Funktionen und durch die Verlinkung der einzelnen Teillösungen zu einer Serviceplattform entwickelt, für die sich die Bezeichnung Work&Learn-Plattform etabliert hat. Diese beinhaltet einen Lern-, einen Community-, einen Service-, einen Magazin-, einen Wiki- und einen News-bzw. Channelbereich. Die



Abb. 13.1 Einblick in den Lern-, den Magazin- und den Communitybereich der Work&Learn-Plattform

Registrierung auf der Plattform wird durch die Möglichkeit zur Selbstregistrierung oder per Azure Login gegeben (Abb. 13.1).

Nach dem Login gelangen Nutzer:innen auf die Startseite, von welcher aus sie in alle Bereiche der Work&Learn-Plattform gelangen können. Es gibt verschiedene Bereiche, die Inhalte thematisch verorten und nachfolgend beschrieben werden.

- **Neuigkeiten:** Der Einstieg in den Neuigkeitenbereich ist am Ende der Startseite verortet, wo News zu verschiedene Themenkanälen angezeigt werden. Diese Themenkanäle können vom Nutzer:innen abonniert werden.
- **Kursübersicht:** In der Rubrik „Kurse“, die den zentralen Kern des Lernbereiches darstellt, kann auf die in der Plattform erfassten Kurse (Inhalte des Basic-Kurs, Advanced-Kurs, Kompetenzrollenbezogene Kurse) zugegriffen werden. AKKORD-Kurse wurden in Zusammenarbeit mit den Partnern des Konsortiums konzipiert und eingestellt, externe Kurse verweisen auf Weiterbildungsmöglichkeiten auf anderen Plattformen. Genauere Informationen zu den entwickelten Lernmodulen sind im Abschn. 13.2.3 zu finden. Auf die im Rahmen des AKKORD-Projekts entwickelten sowie erfassten AKKORD-Kurse wird im Kap. 7 detaillierter eingegangen.
- **Magazin:** Die Rubrik „Magazin“ bietet die Möglichkeit, den Nutzer:innen relevante und aktuelle Artikel rund um das Thema Industrielle Datenanalyse zur Verfügung zu stellen. Neben der Artikelübersicht in Kachelform werden eine Auflistung der meistgelesenen Magazinartikel sowie der verfügbaren Themen in Form von Tags angezeigt.

Über die Tags können sich Nutzer:innen die zu einem Data Science-Schlagwort im System erfassten Artikel anzeigen lassen.

- **Wissensdatenbank:** Die Wiki-Wissensdatenbank (kurz: Wiki) stellt eine weitere Rubrik in der Work&Learn-Plattform dar, die Nutzer:innen die Möglichkeit bieten, aktiv, inhaltlich mitzuwirken und fachlichen Input zu bestimmten Themen einzubringen. So können Nutzer:innen Wiki-Einträge zu zentralen Data Science-Schlagworten erweitern, sodass sich innerhalb Plattform das Netzwerk an Inhalten als Nachschlagewerk für Definitionen oder für Erklärungen zu den einzelnen Themen eignet. Insgesamt fördert das Wiki ein kollaboratives Arbeiten und passt sich durch die offene Administration an die Bedürfnisse der Nutzer:innen an.
- **Community:** Der „Community-Bereich“ erlaubt den Nutzer:innen einen Gruppen- bzw. Rollenbezogenen Austausch durch die Möglichkeit zum Verfassen von Beiträgen, die anschließend für alle Gruppenmitglieder einsehbar sind. Auf diese Art und Weise können themenspezifische Fragen in entsprechenden Gruppen diskutiert und beantwortet werden. Genauere Informationen zu dem Community-Bereich sind im Abschn. [13.2.4](#) und im Kap. [8](#) zu finden.

Darüber hinaus wurden verschiedene Personalisierungsoptionen fachlich konzipiert und entwickelt. Innerhalb der Einstellungen bietet sich dem Nutzenden die Möglichkeit, bestimmte Themen-Kanäle zu abonnieren, deren Inhalte entsprechend auf der Startseite angezeigt werden. Außerdem kann konfiguriert werden, über welche Aktionen im System (z. B. ein neuer Kommentar im abonnierten Inhalt) der Nutzende benachrichtigt werden möchte. In der Toolbar sind die von den Partnern entwickelten Services und Bausteine verlinkt, sodass sie als Absprungbrett für weitere AKKORD-Tools dient. Auch hierfür findet der Nutzende Konfigurationsoptionen in den Einstellungen, um die Toolbar personalisieren zu können. Im Profil lassen sich neben dem Hochladen eines Profilbildes auch Angaben zum Nutzenden machen wie beispielsweise über die Organisation/das Unternehmen, die eigene Position und Expertise sowie über Kontaktinformationen, um passende Ansprechpersonen auf der Plattform finden zu können. In der Rubrik „Meine Seite“ wird jedem Nutzenden die Möglichkeit gegeben, sich eine persönliche Übersicht der wichtigsten Inhalte aus dem Service-Bereich zu erstellen. Hier können also Best-Practice-Beschreibungen, Projektpartnerinformationen, AKKORD Baustein-Inhalte und Success Story-Seiten der einzelnen Projektpartner ausgewählt werden.

Um Nutzer:innen die Möglichkeit zu geben, ihre Kenntnisse in einzelnen Themenbereichen wie beispielsweise „Big Data“ oder „Data Mining“ zu vertiefen, aber auch um einen auf ihre Rolle im Unternehmen bezogenen AKKORD-Kurs vorgeschlagen zu bekommen, wurde eine Empfehlungsfunktion in die Work&Learn-Plattform integriert. Nach dem Abspeichern der individuellen Angaben im Nutzerprofil (Kompetenzrolle, Themengebiete) werden passende Content-Empfehlungen für den jeweiligen Nutzenden gemacht, welche unter der Rubrik „Empfehlungen“ in der Hauptnavigation eingesehen werden können. Durch diese Empfehlungsfunktion wird eine auf den Nutzenden und damit seine/ihre Interessen und Position im Unternehmen zugeschnittene Empfehlung

der auf der Plattform zur Verfügung stehenden Lerninhalte und -materialien ermöglicht. Die Anpassung auf die individuellen Bedürfnisse der Nutzer:innen wird damit gewährleistet. Der Nutzende hat jederzeit die Möglichkeit, seine Angaben anzupassen oder auch ohne die Verwendung der Empfehlungsfunktion die Work&Learn-Plattform zu nutzen.

13.2.3 Generierung technologiebasierter Lernmodule

Es wurden zahlreiche Lektionen zum Themenbereich der Industriellen Datenanalysen in die Kollaborationsplattform integriert, die sich dem Basic-Kurs, dem Advanced-Kurs bzw. den Kompetenzrollenbezogenen Kursen zuordnen lassen. Durch diese Einteilung wird Einsteiger:innen sowie Nutzer:innen mit ersten Kenntnissen in diesem Themenbereich die Möglichkeit gegeben, sich basierend auf dem eigenen Kenntnisstand weiterzubilden.

Die Kurs-Übersichtsseite der AKKORD Work&Learn-Plattform ist so strukturiert, dass ein bestmöglicher Überblick über das Kursangebot gewährt wird. Zunächst werden die dem Basic-Kurs zugeordneten AKKORD-Kurse unter einem entsprechenden Reiter angezeigt, darunter die dem Advanced-Kurs zugeordneten AKKORD-Kurse. Außerdem werden die Kompetenzrollenbezogenen AKKORD-Kurse gebündelt dargestellt sowie abschließend die externen Kurse, die auf Weiterbildungsmöglichkeiten auf anderen Plattformen verweisen.

Um den Nutzer:innen einen kurzen Einblick in die AKKORD-Kurse von der Kursübersichtsseite zu geben, wird in den Kurskacheln neben einem Vorschaubild, der Überschrift und einer kurzen Zusammenfassung die Bewertung des Kurses durch weitere Nutzer:innen (Sternenranking) ersichtlich. Den Nutzer:innen bietet sich außerdem die Möglichkeit einen AKKORD-Kurs zu favorisieren, sodass dieser für ihn oder sie zu Beginn der Übersicht angezeigt wird. Sollte der Nutzer bereits einen Kurs begonnen haben, wird ein Fortschrittsbalken sichtbar, sodass der Bearbeitungsstand immerwährend ersichtlich ist.

Wählt ein Nutzender einen der Kurse aus, so gelangt er oder sie auf die Kursseite, die immer einem ähnlichen Aufbau folgen, um den Nutzer:innen die bestmögliche Orientierung zu ermöglichen. Nach einer Einführung in den Kurs wird ein Überblick über die zugehörigen Sublektionen, den zeitlichen Umfang und die Lernziele gegeben. Darunter kann über die Kacheln auf die einzelnen Sublektionen zugegriffen werden. Die Kurse und Sublektionen bestehen aus einem Haupttext, welcher mit verschiedenen Edelsteinen („Gemstones“) und interaktiven Elementen ausgebaut ist. Zu diesen zählen u. a. Bilder, Youtube-Videos, Slide-Shows, Quizzes, Zitate, weiterführende Links und viele weitere. Nutzeraktivierende und interaktive Elemente sind ein wesentlicher Bestandteil, welche die Motivation der Nutzer:innen steigern und einen positiven Einfluss auf den Lernprozess haben (Hamari & Koivisto, 2015, S. 422).

13.2.4 Integration von Kollaborationsmöglichkeiten

Neben kreativen Lehr- und Lernmethoden werden Lernerlebnisse insbesondere durch Interaktion, beispielsweise mit weiteren Nutzer:innen und Themeninteressierten, erzeugt (Geschwill et al., 2019, S. 38). Um entsprechend schnelle und einfache Kommunikationswege sowie Kontaktmöglichkeiten zu Nutzer:innen und Fachexpert:innen zu ermöglichen, wurden verschiedene Kollaborationsfeatures in die Work&Learn-Plattform integriert, die im Folgenden aufgelistet und in Kap. 8 genauer dargestellt werden.

Die Kommentarfunktion ermöglicht den Nutzenden, sich zu Beiträgen zu äußern und Fragen direkt auf den Beitragsseiten zu stellen. Das Benutzerverzeichnis bietet darüber hinaus die Möglichkeit, Kontaktinformationen aus den Nutzerprofilen zu entnehmen und mit Expert:innen und weiteren Themeninteressierten in Kontakt zu treten. Über die Möglichkeit, Beiträge in den Community-Gruppen zu verfassen, können generelle und themenspezifische Fragen schnell und unkompliziert beantwortet und ein Themenbezogener Austausch ermöglicht werden.

13.2.5 Erprobung in der praktischen Anwendung

Die einzelnen Bereiche der Work&Learn-Plattform wurden im Rahmen des Projekts durch verschiedene Testphasen seitens der Projektpartner und einzelner externer Partner (beispielsweise von Messebesucher:innen verschiedener Fachmessen und Tagungen) getestet und validiert. Außerdem wurden im Rahmen einzelner Konsortialtreffen gemeinsame Testszenarien durchgespielt und evaluiert.

Im Projektverlauf wurde deutlich, dass die entwickelte Kollaborationsplattform weitere Möglichkeiten im Hinblick auf die inhaltliche Erweiterung bietet, was bereits im Rahmen des Projekts getestet wurde, jedoch insbesondere auch zukunftsprospektive Potenziale aufzeigt. Beispielsweise wurde die Work&Learn-Plattform um relevante Inhalte für die Zielgruppe „Berufsschule“ erweitert, eine entsprechende Rolle im Nutzerprofil hinzugefügt und ein zugehöriger AKKORD-Kurs speziell hierfür eingepflegt. Abschließend erfolgte die praktische Anwendung und Validierung im Rahmen eines Berufsschulprojekttages.

13.3 Transfermaßnahmen und Umsetzung im Referenzbaukasten

Die Work&Learn-Plattform dient als Bündelungsplattform für die einzelnen AKKORD-Services. Auf ihr sind die entwickelten Lernmodule sowie verschiedene Möglichkeiten zur Kollaboration der Nutzenden integriert. Zu den seitens der Projektpartner entwickelten Lösungsansätze wurden Informationen im Servicebereich zur Verfügung gestellt und entsprechende Verlinkungen zu diesen in die Toolbar sowie in der Rubrik „Services“ aufgenommen. Über diese Verlinkungen können die Nutzer:innen zu den

jeweiligen Lösungsansätzen (Datenbackend-Systeme, Analysebaustein) aus den einzelnen Leistungsbereichen gelangen.

Während die Testphasen im Rahmen der Projektlaufzeit überwiegend von Projektpartnern absolviert wurden, wäre das Ausrollen für externe Unternehmen zu Optimierungs- und Validierungszwecken eine mögliche Transfermaßnahme.

13.4 Fazit

Im Rahmen des Anwendungsfalls konnte die Konzipierung und Entwicklung eines digitalen Wissensdienstes für den Themenbereich Industrielle Datenanalyse erfolgen, der jederzeit und kostenlos von Mitarbeitenden in KMU zum Kompetenzaufbau genutzt werden kann – sowohl um formell Kompetenzen aufzubauen als auch um informell Wissen abzurufen. Nach der Definition und Ermittlung datenwissenschaftlicher Kompetenzen wurde im Rahmen des Projekts der digitale Wissensdienst entwickelt, technologiebasierte Lernmodule generiert und Kollaborationsmöglichkeiten sowie ein Servicebereich integriert. Durch Testläufe wurde die Plattform abschließend validiert und optimiert.

Mit der Work&Learn-Plattform konnte eine moderne Lösung zum Kompetenzaufbau und -sicherung geschaffen werden, auf der Lerninhalte für Einsteiger:innen und Fortgeschrittene bereitgestellt werden und die durch die Umsetzung als Online-Plattform mit der Möglichkeit zur Selbstregistrierung für Interessierte jederzeit zur Verfügung steht. Durch die Bündelung von Lern-, Community-, Service-, Magazin-, Wiki- und Neuigkeiten-Bereich in einer Online-Plattform wird den Nutzer:innen über Kompetenzaufbau und -sicherung hinaus die Aneignung von Wissen zum Themenbereich ermöglicht und die Chance zum Austausch und zur Vernetzung mit weiteren Themeninteressierten und Fachexpert:innen über die integrierten Kollaborationsfunktionen gegeben.

Wie bereits in Abschn. 13.2.5 hervorgehoben wurde, bietet die Work&Learn-Plattform Erweiterungs- und Ausbaupotenzial in verschiedener Hinsicht. Einerseits ist die Möglichkeit zum Ausbau und der Aktualisierung des bestehenden Lernangebots gegeben und andererseits können weitere Zielgruppen adressiert und Inhalte zu weiteren Themengebieten hinzugefügt werden. Darüber hinaus besteht die Möglichkeit, externe Kursangebote zu präsentieren und bereitzustellen, um das AKKORD-Kursangebot bestmöglich zu erweitern. Kap. 20 stellt einen abschließenden Ausblick auf die derzeit stattfindende Transformation Industrieller Datenanalysen im Einklang mit Mensch, Technik und Organisation vor.

Literatur

- Arnold, D., Butschek, S., Steffes, S., & Müller, D. (2016). *Forschungsbericht 468 -Digitalisierung am Arbeitsplatz*.
- Cerasoli, C. P., Alliger, G. M., Donsbach, J. S., Mathieu, J. E., Tannenbaum, S. I., & Orvis, K. A. (2018). Antecedents and outcomes of informal learning behaviours: A meta-analysis. *Journal*

- of *Business and Psychology*, 33(2), 203–230. Springer. <https://doi.org/10.1007/s10869-017-9492-y>.
- Deuse, J., Erohin, O., & Lieber, D. (2014). Wissensentdeckung in vernetzten, industriellen Datenbeständen. In H. Lödning (Hrsg.), *Wie intelligente Vernetzung und kognitive Systeme unsere Arbeit verändern* (S. 373–395). Gito (Schriftenreihe der Hochschulgruppe für Arbeits- und Betriebsorganisation e. V).
- Eickelmann, M., Wiegand, M., Konrad, B., & Deuse, J. (2015). Die Bedeutung von Data-Mining im Kontext von Industrie 4.0. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 110(11), 738–743.
- Geschwill, R., Nieswandt, M., & Zimmermann, V (2019). *EdTech in Unternehmen – Lernen als Schlüssel für Innovation und Wachstum in Zeiten der Digitalisierung*. Springer Gabler.
- Hamari, J., & Koivisto, J. (2015). Why do people use gamification services? *International Journal of Information Management*, 35(4), 419–431. <https://doi.org/10.1016/j.ijinfomgt.2015.04.006>.
- Nickolaus, R., & Walker, F. (2016). Kompetenzorientierung in der beruflichen Bildung. In A. Dietzen, R. Nickolaus, B. Rammstedt, & R. Weiß (Hrsg.), *Berichte zur beruflichen Bildung. Kompetenzorientierung: Berufliche Kompetenzen entwickeln, messen und anerkennen* (S. 7–28). W. Bertelsmann Verlag GmbH & Co. KG, Bielefeld.
- Spöttl, G., Gorldt, C., Windelbrandt, L., Grantz, T., & Richter, T. (2016). *Auswirkungen auf Aus- und Weiterbildung in der M+E Industrie*. Bayme vbm.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the development of a collaboration platform for competency-based collaboration in industrial data science projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and Data Science in Engineering Management. CIO 2022. Lecture Notes on Data Engineering and Communications Technologies* (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- West, N., Schwenken, J., & Deuse, J. (2022a). Comparative study of methods for the real-time detection of dynamic bottlenecks in serial production lines. In H. Fujita, P. Fournier-Viger, M. Ali, & Y. Wang (Hrsg.), *Advances and trends in artificial intelligence. Theory and practices in artificial intelligence* (S. 3–14). Springer. https://doi.org/10.1007/978-3-031-08530-7_1.
- West, N., Syberg, M., & Deuse, J. (2022b). A holistic methodology for successive bottleneck analysis in dynamic value streams of manufacturing companies. *Proceedings Changeable, Agile, Reconfigurable and Virtual Production Conference*, 8(1), 612–619. https://doi.org/10.1007/978-3-030-90700-6_69.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Retrofitting und Datenakquisition mittels Industrie 4.0-Technologien

14

Probleme und Lösungsansätze bei
Prozessdatenerfassung allgemein und
beispielhafte Umsetzung in AKKORD

Tina Hardt , Hans-Joachim Mieves , Christoph Marx  und Christian Meisberger

14.1 Motivation

Bei der Umsetzung von Projekten im Kontext von Industrie 4.0 (I4.0) ist meist die erste Aufgabe, Daten aus der Produktion (Operations Technology; OT) in dafür vorgesehene Verarbeitungsinfrastruktur (Information Technology; IT) zu transferieren. Dieser Schritt entspricht der ersten Phase der Prozesskette der Industriellen Datenanalyse (siehe Kap. 4). Diese Aufgabe bringt zu einer Herausforderung in Bezug auf die Sicherheit der OT mit (s. z. B. Hardt et al., 2021, S. 40 f.; Hengel & Hardt, 2020, S. 58 f.) und zu anderen weiteren verschiedenen Herausforderungen, die direkt mit der Datenerfassung zusammenhängen und die im nächsten Abschnitt aufgeführt werden.

14.1.1 Herausforderungen nachträglicher Datenakquisition

Ein Problem bei der Datenerfassung stellen heterogene Datenquellen dar. In modernen Anlagen werden viele relevante Daten in der *Speicherprogrammierbaren Steuerung*

T. Hardt (✉) · H.-J. Mieves · C. Marx · C. Meisberger
Real-Time-Systems GmbH, Wittlich, Deutschland
E-Mail: tina.hardt@real-time-systems.com

H.-J. Mieves
E-Mail: joachim.mieves@real-time-systems.com

C. Marx
E-Mail: christoph.marx@real-time-systems.com

C. Meisberger
E-Mail: christian.meisberger@real-time-systems.com

(SPS) der Anlage erfasst und können von dort ausgelesen werden. Es gibt mehrere Möglichkeiten, Daten aus einer SPS auszulesen. Eine besonders einfache Methode ist z. B. bei SPS-Systemen der Firma SIEMENS möglich. Hier kann über das S7 Protokoll direkt auf alle SPS-Variablen (z. B. Merker, Datenbausteine, Eingänge, Ausgänge, Zähler) lesend und schreibend zugreifen, sofern die jeweilige SPS nicht im sogenannten „optimierten Modus“ läuft. Eine potenzielle Folge, die dabei allerdings beachtet werden muss, ist eine eventuell resultierende Verlängerung der Zykluszeit. Generell hat jeder Hersteller für den Zugriff auf eine SPS sein eigenes proprietäres Protokoll. Darüber hinaus kann auch z. B. über Standards wie *modbus* oder *OPC UA* auf viele SPS-Systeme zugegriffen werden. Allerdings muss der Zugriff über *OPC UA* speziell eingerichtet werden und kann z. B. nach einem Update der SPS durch den Hersteller wieder überschrieben werden. Außerdem erhöht diese Art des Zugriffs die CPU-Last und kann die Echtzeitfähigkeit beeinträchtigen.

In älteren Anlagen oder wenn z. B. die Datenlage in der SPS nicht ausreichend ist, müssen Daten zusätzlich aus verbauter oder externer Sensorik/Aktorik oder aus Feldbussen ausgelesen werden. Beim Anbringen von externer Sensorik oder auch beim Zugriff auf die SPS ist übrigens stets darauf zu achten, dass keine Aktionen durchgeführt werden, die die Garantie des Herstellers verletzen.

Liegen verschiedene Datenquellen in einem System vor, bringt dies wiederum weitere Herausforderungen mit sich. Zum einen haben diese Datenquellen meist unterschiedliche physikalische Schnittstellen (z. B. LAN, seriell, Funk, ...) mit denen man sich physikalisch verbinden können muss und zum anderen herrschen für gewöhnlich in unterschiedlichen Produkten auch unterschiedliche Protokolle und Schnittstellen zur Datenkommunikation (z. B. *OPC UA*, *MQTT*, *modbus*, *CoAP*, *RS485*, *S.7*, *5G*, *LoRaWAN*, *LTE*, *WLAN*, ...), die man wiederum „sprechen“ können muss, um die Daten abzugreifen. Außerdem müssen die Daten unter Umständen bei nichtechtzeitfähigen Systemen oder Transportwegen zeitlich synchronisiert werden, um zu passenden Datensätzen zu einem Event zusammengestellt werden zu können. D. h. es sind entweder zeitsynchronisierte Datenquellen notwendig, die einen Erfassungszeitpunkt zum erfassten Wert mitliefern oder bei der Abtastung durch z. B. ein Gateway muss dieses einen Zeitstempel vergeben.

Ebenso kann es vorkommen, dass ein Prozess unterschiedliche Datensinken hat (Datenbanken, wie z. B. *mysql*, *SQLite*, *Influx-DB*, *CSV-Dateien*, *ERP-Systeme*, ...), auch mit diesen muss man kommunizieren können, um die Daten dort ablegen zu können. Eventuell ist auch eine Visualisierung der Daten oder eine Darstellung historischer Daten in Linien-Diagrammen notwendig, die umgesetzt werden muss.

Unter Umständen stellen zudem die Schnelligkeit bzw. Abtastrate sowie die Auflösung bzw. Bittiefe der benötigten Daten, ein Problem dar. Es muss sichergestellt werden, dass alle Komponenten, aus denen Daten erhoben werden, die Daten schnell genug erheben und auch kommunizieren können und die gewünschte Genauigkeit liefern. Ist dies nicht der Fall, müssen Strategien erarbeitet werden und gegebenen-

falls zusätzliche Komponenten recherchiert werden, mit denen das betrachtete System nachgerüstet werden kann, um die vorliegenden Anforderungen an Schnelligkeit und Genauigkeit zu erfüllen. Weiterhin müssen Daten zwischengespeichert werden können, falls es zu kurzen Unterbrechungen und Verzögerung beim Schreiben der Daten kommt.

Dieser Abschnitt hat zusammenfassend aufgezeigt, wie vielfältig die Probleme und Herausforderungen zur Bereitstellung der benötigten Datengrundlage für I4.0-Prozesse sind. Im folgenden Abschnitt werden praktische Tipps und Lösungsstrategien für die Erfassung von Daten vorgestellt.

14.1.2 Gateways und Retrofitting im produzierenden Gewerbe

Maschinen und Anlagen, die nicht über entsprechende Technologien zur Vernetzung verfügen, können bei Bedarf mithilfe vom sogenannten *Retrofitting* auch nachträglich in I4.0-Prozesse integriert werden (Wöstmann et al., 2019, S. 94 ff.). Zur Erfassung und Anbindung von Daten verschiedener Quellen sind typischerweise industrielle Netzwerknoten erforderlich, die Systeme mit unterschiedlichen Übertragungsprotokollen miteinander verbinden und auch als *Gateways* bezeichnet werden. Diese haben in der Regel mehrere physikalische Schnittstellen und beherrschen verschiedene Protokolle, um so Daten aus diversen Quellen auslesen zu können. Einige Gateways ermöglichen die Vorverarbeitung von Daten und können so erhobene Daten z. B. auf gewünschte Formate vor der Datenweitergabe transformieren oder diverse Kenngrößen wie beispielsweise Minima, Maxima, Durchschnitt oder Varianz berechnen. Durch letzteres kann die Übertragung von Rohdaten entfallen und der Datenverkehr mitunter erheblich reduziert werden. Aus der Kombination von unterschiedlichen Rohdaten können auch neue Informationen generiert werden.

Bietet ein betrachtetes industrielles System an sich nicht genügend oder nicht hinreichende Sensorik für gewünschte Analysen, müssen Komponenten nachgerüstet werden. Beim Einsatz von Gateways sowie beim Nachrüsten von Komponenten muss darauf geachtet werden, dass die Herstellergarantie nicht gefährdet wird. Außerdem ist es bei der Anschaffung zusätzlicher Sensorik u. U. hilfreich darauf zu achten, dass diese nicht nur benötigte Genauigkeit und Schnelligkeit erfüllen, sondern idealerweise auch Protokolle beherrschen, die schon im System vorherrschen. Weiterführend liefern Hardt et al., (2021, S. 31–34) nützliche Hinweise zur Auswahl von Sensorik. Integriert man zusätzliche Sensorik, ist es oft ratsam, die erhobenen Werte über Gateways zu erfassen und nicht in die SPS weiterzugeben, um dort nicht in die Konfiguration und Programmierung eingreifen zu müssen.

Kann man auf vorhandene Sensorwerte in einem System nicht einfach über die SPS zugreifen, bieten sich *Signalverdoppler* an. Diese können für gewöhnlich zwischen Sensor und SPS geklemmt werden und am Verdoppler kann dann das Signal z. B. mit einem Gateway abgegriffen werden. Ebenso sind Schütze ein einfaches Hilfsmittel, um

digitale Daten aus einem Schaltschrank ohne Eingriff in die Anlage abzugreifen. Nach der Prozessdatenerfassung schließt sich unter Umständen die Zusammenführung mit anderen Daten an, eventuell die Einspeisung in ein Datenbackend-System (siehe Kap. 5).

14.2 Datenerfassung in AKKORD

Nachdem der vorige Abschnitt die allgemeinen Probleme und Lösungsmöglichkeiten bei der Erfassung von Daten in industriellen Prozessen umrissen hat, wird in diesem Abschnitt speziell auf die Umsetzung der Datenerfassung in drei Szenarien aus dem Forschungsvorhaben AKKORD eingegangen (siehe Kap. 1). Dabei kam das industrielle Edge-Gateway ARENDAR des Projektpartners Real-Time-Systems GmbH (RTS) zum Einsatz.

14.2.1 Die Testanlage von PolyMerge

Im Anwendungsfall des Projektpartners PolyMerge (siehe Kap. 12) bestand in einer Pilotanlage das Problem, dass die Daten für die Auswertung in höherer Frequenz aus dem Prozess benötigt wurden, als die integrierte SPS sie bereitstellen kann. Hier wurde versucht, mit einem Switch und einem Profinet Gateway der Klasse 2 an den IO-Link-Master im Prozess anzuknüpfen. Das Edge-Gateway ARENDAR sollte dann über MQTT auf die Daten des Profinet Gateways zugreifen. Es stellte sich in einem Test-szenario heraus, dass dieser Weg zur schnelleren Datenerfassung mit den integrierten Komponenten so nicht umsetzbar ist (da der integrierte Master keine „shared devices“ unterstützt, d. h. es sind keine 2 Master im Szenario einsetzbar). Zwei Alternativen wurden recherchiert, leider mit nicht unerheblichem Programmieraufwand sowie hohen Kosten für andere Komponenten (z. B. schnellere SPS). Die Arbeiten resultierten also lediglich in zwei aufwendigeren Lösungsansätzen, eine kostengünstige schnelle Lösung war hier nicht umsetzbar.

In einer anderen Anlage ist die integrierte SPS leistungsfähig genug. Hier werden 15 Werte vom ARENDAR über das Protokoll S7 mit einer Abtastrate von 100 ms aus der SPS ausgelesen. Zudem ist ein zusätzlicher Drucksensor ins Szenario integriert, der direkt mit der IO-Karte des ARENDAR verbunden ist. Alle 15 Werte werden im ARENDAR in einem JSON-Datensatz zusammengefasst und in eine Influx-Datenbank geschrieben. Zudem werden die aktuellen Werte über den Webserver in einer Applikationswebseite angezeigt. Diese Seite ermöglicht auch den Start-Stop der Aufzeichnungen, die Eingabe von Kommentaren in die Datenbank, die Möglichkeit eine Datenbank für die Ablage auszuwählen und die Anzeige von historischen Werten in Line-Charts. Weitere Informationen zu den Daten und zur Auswertung finden sich in Kap. 12.

14.2.2 Der AKKORD-Autorenn-Demonstrator

Eine ausführliche Beschreibung des Demonstrators findet sich in Kap. 17 und Kap. 18. Während Kap. 18 das Datenbackend-System den Autorenn-Demonstrator vorstellt und Kap. 17 die modularen Datenanalysen behandelt, werden in diesem Kapitel die Arbeiten zur Erfassung von Daten als Retrofit-Szenario beschrieben.

Es existiert ein Python-Skript des Projektpartners VPE zur Datenerfassung im Demonstrator. Dieses wurde dahingehend erweitert, dass es sich parallel zu mehreren MQTT-Brokern verbinden kann. Somit kann sich der ARENDAR an die bereits in der Bahn implementierte Datenerfassung über MQTT anschließen. Zudem wurden Erweiterungen vorgenommen, die die Berechnung und Erfassung der Sektorzeiten umsetzen (Erzeugung von neuen Informationen aus den Rohdaten). Die Sektorzeiten werden alle 250 ms abgefragt. Limitierende Faktoren für eine schnellere Abtastung war hier ein Overhead des Kommunikationsprotokolls, der eine Kommunikation in Echtzeit an den Server nicht zuließ.

Ein Temperatur- und Umweltsensor zur Erfassung von Temperatur, Luftfeuchte und Luftdruck wurde zur Demonstration von Retrofitting nachgerüstet. Der Umweltsensor wurde mit einem separaten Raspberry Pi verbunden, der die Daten über MQTT an den ARENDAR weiterleitet. Hier wurde eine langsame Abtastrate gewählt, da nicht mit schnellen Veränderungen der erfassten Werte zu rechnen ist. Erfasste und berechnete Werte werden über eine Webseite angezeigt (siehe Abb. 14.1).

In diesem Szenario konnte also durch eine Skriptänderung über das leichtgewichtige Protokoll MQTT mit einem Gateway Daten erfasst und verarbeitet werden. Durch einen nachträglich angebrachten Sensor wurden zusätzliche Informationen im Szenario erzeugt, mit den übrigen Daten im Gateway zusammengeführt und über eine Webseite visualisiert.

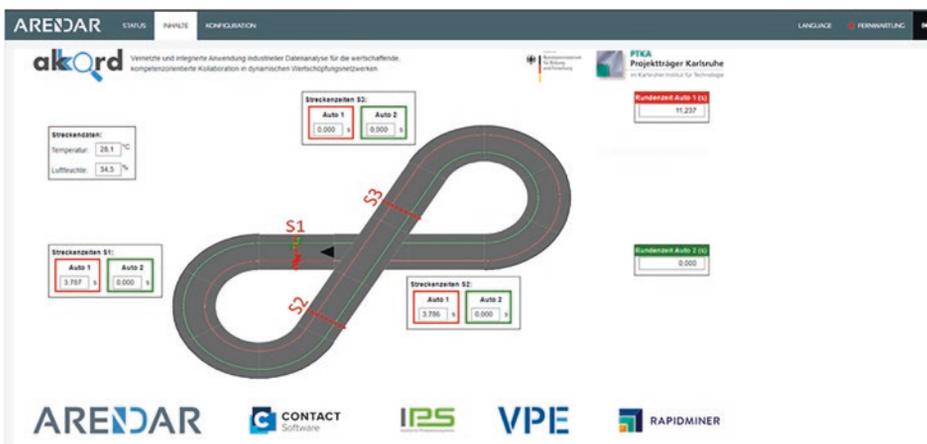


Abb. 14.1 Visualisierung erfasster Werte im Autorenn-Demonstrator

14.2.3 Die AKKORD Lernstation

Die AKKORD-Lernstation ist vor dem Hintergrund entstanden, Schulen und Ausbildungsbetrieben ein kompaktes, preiswertes Hilfsmittel zum Erlernen von I4.0-Szenarien an die Hand zu geben (Hardt & Klupak, 2022, S. 271 ff.). Der hardwaretechnische Aufbau und eine Abbildung der Station sind im Kap. 19) beschrieben.

Abb. 14.2 zeigt den schematischen Aufbau der Lernstation geordnet nach Szenarien und Kommunikation. Der obere linke Teil (grün hinterlegt), stellt eine alte Anlage ohne Steuerung und Vernetzung dar. Die zugehörige Sensorik ist direkt an die IO-Karte des ARENDAR angeschlossen und soll somit das Retrofitting einer Anlage abbilden. Sechs digitale Eingänge können über die drei Wechselschalter vorne links an der Station angesteuert werden. Die Potenziometer in der Mitte der Station sind über Analogeingänge an den ARENDAR angebunden, sie liefern eine Spannung im Bereich 0–10 V. Die IO-Karte des ARENDAR tastet die Werte mit der Rate 1 kHz ab. Der ARENDAR fragt in einem einstellbaren Zeitintervall die aktuellen Werte ab und kommuniziert diese von der Karte über den I2C-Bus in die Hauptplatine des ARENDAR zur weiteren Verarbeitung.

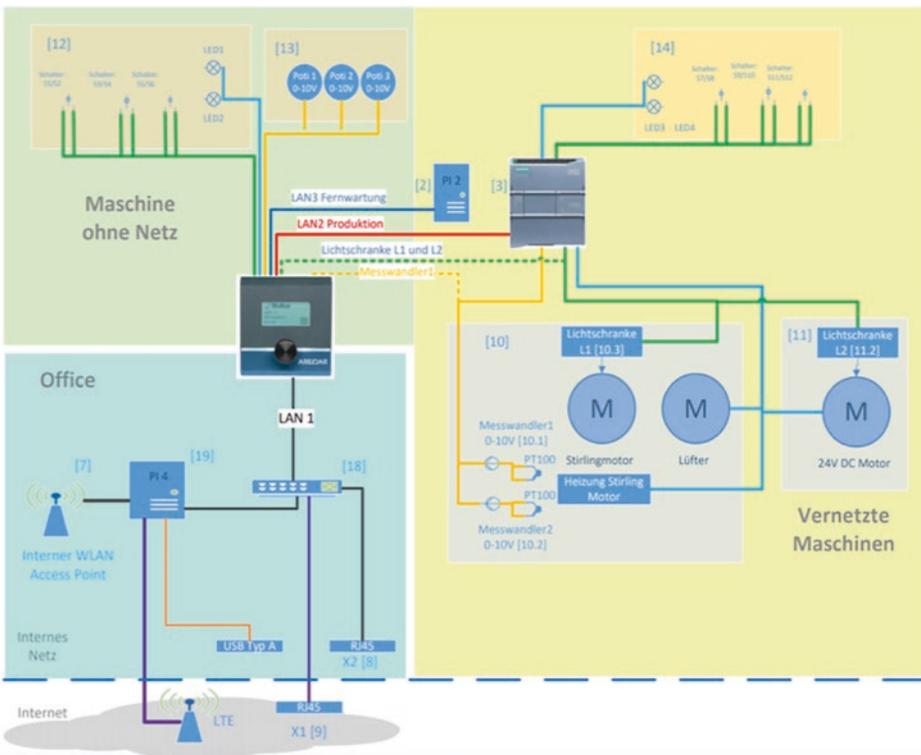


Abb. 14.2 Schematischer Aufbau der Lernstation

Der rechte Bereich der Abb. 14.2 (gelb hinterlegt) zeigt eine moderne Anlage mit Steuerung und zugehöriger Vernetzung. Auch hier stehen drei Schalter vorne rechts an der Station zur Verfügung. Diese sind an die Digitaleingänge der SPS angeschlossen. Die Lichtschranken der beiden im Produktionsbereich befindlichen Motoren sind ebenfalls mit Digitaleingängen der SPS und mit Digitaleingängen des ARENDAR verbunden. Zum Stirlingmotor gehören zwei Temperatursensoren, eine Lichtschranke, eine Heizplatte und ein Lüfter mit Windkanal zur Kühlung des Stirlingmotors. Bei den zwei Temperatursensoren handelt es sich um zwei PT100 Thermoelemente, diese sind über jeweils einen Messumwandler ($0-10\text{ V} = 0-100\text{ }^{\circ}\text{C}$) mit der SPS verbunden. Sie dienen zur Regelung des Stirlingmotors. Der Messumwandler des unteren PT100 ist zusätzlich noch auf einen Analogeingang des ARENDAR geführt. Alle in die SPS geführten Daten werden vom ARENDAR über die Schnittstelle LAN2 von der SPS über das Protokoll S7 eingelesen. Standardmäßig findet die Abfrage alle 100 Millisekunden statt, ist aber auf andere Werte konfigurierbar.

Über die digitalen Ausgänge der SPS lassen sich die Kontrollleuchten, zwei Relais und das Motorsteuerboard ansteuern. Mit Hilfe eines Relais wird eine Heizung geschaltet, die als Wärmequelle für den Stirlingmotor dient, mit einem weiteren Relais wird der Lüfter des Stirlingmotors geschaltet.

Der Bereich rechts unten in Abb. 14.2 (blauer Bereich) steht für die im Kofferboden integrierte Office-IT einer Firma. Diese ist über die LAN1-Schnittstelle mit dem ARENDAR verbunden und die Kommunikation ist z. B. über MQTT oder OPC-UA möglich. Die Office-IT ermöglicht Zugang zum Internet und damit verbundene Möglichkeiten wie Zugang zu Cloud-Services (AWS, Telekom-IoT, Azure IoT) und Darstellung von Werten über den integrierten Webserver. Durch die zwei getrennten LAN-Anschlüsse des ARENDARs ist eine Netzwerksegmentierung zwischen IT und OT gegeben, sodass die Komponenten der OT nicht den Gefahren des Internets ausgesetzt sind.

14.3 Fazit

Die Prozessdatenerfassung ist ein komplexes Thema. Technische Herausforderungen, Lösungsansätze und beispielhafte Umsetzungen wurden hier beschrieben. Daneben ist der Schutz der OT vor Angriffen von außen (Attacks, Spionage) ein weiteres großes Feld, das hier zu Anfang nur erwähnt wurde. Gateways stellen ein mächtiges Werkzeug bei der Datenerfassung dar. RTS konnte durch das Projekt AKKORD wichtige Erkenntnisse bezüglich der Weiterentwicklung des Gateways ARENDAR sowie auch Grenzen bei der Datenerfassung sammeln. Diese werden in die Weiterentwicklung des Produkts in den Bereichen Datenbankanbindung, Datenvisualisierung, Konnektivität und Nutzerfreundlichkeit einfließen.

Für einen umfassenderen Blick auf die derzeit stattfindende Entwicklung industrieller Datenanalysen im Zusammenhang mit Mensch, Technik und Organisation, verweisen wir abschließend auf Kap. 20.

Literatur

- Hardt, T., Meißner, T., Mohr, T., Riestler, B., Helou, K., Driemeyer Wilber, A., Klotz, M., Travner, S., Schmitt, L., Hoffmann, D., Tahiri, T., Frey, A., Kuhnle, A., Saier, T., Moschko, L., Zöllner, M., Lang, S., Racco, F., & Wojek, C. (2021). Abschlussbericht Forschungsprojekt ProData – Produktionsbezogene Dienstleistungssysteme auf Basis von Big-Data-Analysen.
- Hardt T., & Klupak, C. (2022) Lernstation für Datenerfassung & IoT-Applikationen. Unter Berücksichtigung der Sicherheit bei Datenerfassung und -kommunikation. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 5, 271–274. <https://doi.org/10.1515/zwf-2022-1066>.
- Hengel, S., & Hardt, T. (2020). Cybersicheres IoT-Gateway für die IT/OT-Kommunikation. *atp-magazin*, 4, 58–59.
- Wöstmann, R., Barthelmey, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. In T. Schüppstuhl, K. Tracht, & J. Roßmann (Hrsg.), *Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Integration der Kundenmeinung ins Geschäftsmodell

15

Einbindung von Kundenmeinungen über Sentiment Analysen und Anpassungen des Geschäftsmodells

Michael Frey

15.1 Motivation

Heute tauschen Menschen ihre Gedanken und Meinungen auf unterschiedlichen Social-Media-Plattformen, auf Online-Marktplätzen, Webshops und Foren aus. In Online-Shops und auf Social-Media-Seiten entstehen auf diese Weise täglich bewusst oder unbewusst Millionen neuer Produkt- und Dienstleistungsbewertungen. Wie in Kap. 3 eingehend erwähnt ist für Unternehmen die Kenntnis der Einstellung von Konsumenten zu einem Produkt oder der eigenen Marke wesentlich, um die Produktqualität oder die damit zusammenhängenden Dienstleistungen und Absatzkanäle und damit die Kaufentscheidung der Konsumenten zu beeinflussen (Hu & Liu, 2004; Kostyra et al., 2016). Trotzdem wissen Anbieter häufig nicht, wie Kunden ein Produkt einschätzen oder welche Änderungen sie sich wünschen. Diese online vorhandenen Daten könnten zur Verbesserung der Produkte und dem dahinterliegenden Geschäftsmodell genutzt werden.

Dieses Kapitel stammt aus dem Forschungsprojekt AKKORD (siehe Kap. 1 und 2). Er setzt sich zum Ziel Texte aus Onlinequellen, wie Webshops und Social-Media-Plattformen, systematisch zu erfassen und automatisiert auszuwerten. Dabei besteht aufgrund des Charakters der Daten, also der geschriebene Ausdruck einer persönlichen Meinung über das Produkt, natürlich ein Fokus auf die Kundensicht innerhalb des Geschäftsmodells.

M. Frey (✉)

Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, Deutschland

E-Mail: michael.frey@dfki.de

© Der/die Autor(en) 2024

J. Deuse et al. (Hrsg.), *Industrielle Datenanalyse*,

https://doi.org/10.1007/978-3-658-42779-5_15

163

15.2 Kundensicht im Geschäftsmodell

Die Wertschöpfungskette lässt sich mitsamt ihren Aktivitäten und Ressourcen im Modell des *Business Model Canvas* (BMC, Osterwalder & Pigneur, 2010), eine etablierte grafische Methode zur Bearbeitung und Überarbeitung von Geschäftsmodellen, einordnen. Dies erlaubt es eine simplifizierte Übersicht über das Geschäftsmodell eines Unternehmens zu erstellen, was eine vereinfachte Bewertung und subsequente Anpassungen von einzelnen Bereichen ermöglicht. Aufgrund des vorzeitigen Ausscheidens des Anwendungspartners Brabant & Lehnert konnte sich der Anwendungsfall nicht auf die Erfassung und Verarbeitung von Produktionsdaten erstrecken. Dadurch wurde der Anwendungsfall auf die Kundensicht beschränkt, da es hier möglich war ausreichende Datenmengen zu Bewertungen von Produkten oder Dienstleistungen zu erfassen. Durch die Reduzierung auf die Betrachtung der Kundensicht kann jedoch keine ganzheitliche Bewertung des Geschäftsmodells erfolgen. Mögliche Bereiche, in denen Kundenmeinungen eingeordnet werden können, finden sich im rechten Teil des BMC. Auf der obersten Ebene zählen dazu die folgenden Aspekte:

1. **Kundensegmente** – Adressiert das Geschäftsmodell die geplante Zielgruppe?
2. **Kundenbeziehung** – Werden Kundenerwartungen erfüllt?
3. **Verkaufskanäle** – Passen die im Geschäftsmodell gewählten Verkaufskanäle zu den Anforderungen der Kunden?
4. **Einnahmequellen** – Trifft das Geschäftsmodell die Preisvorstellung der Zielgruppe?
5. **Nutzenversprechen** – Werden die Kundenwünsche zu dem Produkt erfüllt?

Das im Kapitel vorgestellte Verfahren der automatischen Erfassung von Kundenmeinungen adressiert dabei den Punkt „Nutzenversprechen“.

15.3 Bewertung des Geschäftsmodells anhand von Kundenbewertungen

Der vorliegende Anwendungsfall thematisiert die Analyse von Online-Shop- und Social-Media-Daten durch den Einsatz von natürlicher Sprachverarbeitung, um fundierte Rückschlüsse auf das Geschäftsmodell zu ziehen. Der Fokus dieser Analyse liegt darin, potenzielle Risiken sowie Chancen für das Geschäftsmodell aufzudecken und entsprechende Maßnahmen zu ergreifen. Hierbei werden Datenquellen wie Kundenfeedback und Social-Media-Interaktionen analysiert, um das Geschäftsmodell zu optimieren und für zukünftige Marktbedingungen zu rüsten. Dabei werden im ersten Schritt passende Quellen mit Produkt- oder Dienstleistungsbewertungen identifiziert und mithilfe von Softwareschnittstellen oder Webcrawlern angebunden. In einem zweiten

Schritt wurde eine mehrstufige Geschäftsmodelltaxonomie auf Grundlage der aus dem BMC ausgewählten Aspekte erstellt. Diese erlaubt es Schlagwörter auf einen der oben genannten Geschäftsmodellaspekte abzubilden. Ähnlich wie in Kap. 9 werden mehrere Verfahren aus dem Bereich der natürlichen Sprachverarbeitung (Natural Language Processing, NLP) durchgeführt, um einzelne Textsegmente einem bestimmten Geschäftsmodellaspekt zuordnen zu können. Diese Abbildung von Textsegmenten auf Geschäftsmodellaspekte erlaubt es, die Ergebnisse der nachfolgenden Sentiment-Analyse direkt mit dem jeweiligen Geschäftsmodellaspekt zu verknüpfen. Das Ergebnis ist dann eine kumulative Bewertung des Kundenempfindens, welches sich mit den jeweiligen Geschäftsmodellaspekten verbinden lässt. Als letzten Schritt werden auf Basis der Bewertung des Geschäftsmodells aus Kundensicht, Empfehlungen generiert, wie das Geschäftsmodell verändert werden kann, um die negativen Aspekte zu negieren oder positive Aspekte zu verstärken. Die Empfehlungen werden dabei in Form von etablierten Geschäftsmodellmustern wie z. B. E-Commerce, Freemium oder Subscription gegeben.

15.3.1 Auswahl von Kundenportalen als Datenquellen

Grundlage für die im Kapitel durchgeführten Analysen ist eine bestehende Datenbasis an Kundenmeinungen zu Produkten oder Dienstleistungen. Aufgrund der großen Menge an vorhanden und öffentlich zugänglichen Daten in Bewertungsportalen wurde als Datenakquise das Web Crawling ebensolcher ausgewählt. Web Crawling, auch bekannt als Web-Scraping, bezieht sich auf den Prozess des automatisierten Durchsuchens des Internets, um Daten von Websites zu extrahieren. Ein Webcrawler nutzt dafür spezielle Algorithmen, um eine Website zu durchsuchen und alle relevanten Daten zu erfassen. Diese Daten werden dann in einer strukturierten Form gespeichert, um weiter analysiert zu werden. Für den vorliegenden Prototypen wurde das Python Open-Source-Framework Scrapy verwendet, um einschlägige Bewertungsportale zu crawlen.

Dabei ist zu beachten, dass sich diese Methode der Datenakquise nur für prototypische Implementierungen eignet, da bereits kleinste Änderungen an der Struktur von Webseiten bereits zu großen Fehlern führen können. Auch ist es aus ethischer Sicht angebracht bei einem produktiven Einsatz einer solchen Analysepipeline Verhandlungen mit dem Websitebetreiber aufzunehmen und eine Anbindung über eine dedizierte Schnittstelle anzustreben.

15.3.2 Sentiment-Analyse und Abbildung auf den Business Model Canvas

Die Auswertung der Bewertungstexte erfolgt in einem mehrstufigen Prozess, der verschiedene Methoden aus dem Bereich des Natural Language Processing umfasst. Dieser

Prozess beinhaltet die Anwendung von unterschiedlichen Verfahren, wie beispielsweise der Textvorverarbeitung, der Sentiment-Analyse und der Themenmodellierung, um die Texte gründlich zu analysieren. Da es häufig vorkommt, dass ein Kunde eine gemischte Bewertung abgibt, bei der selbst in einem einzigen Satz Aussagen zu unterschiedlichen Aspekten mit unterschiedlicher Bewertung getroffen werden, wird an erster Stelle eine aspektorientierte Sentiment-Analyse (ABSA, Federici & Dragoni, 2017, S. 1 ff.) durchgeführt. Im Kontext von Kundenbewertungen bedeutet dies, dass die Analyse sich auf die Bewertungen von Produkten oder Dienstleistungen konzentriert und untersucht, welche Aspekte des Produkts oder der Dienstleistung in den Bewertungen am häufigsten erwähnt werden. Diese Aspekte können zum Beispiel Qualität, Preis, Lieferung, Kundenservice usw. sein.

Abb. 15.1 zeigt die implementierte NLP-Pipeline der Sentiment-Analyse. Der nächste Abschnitt beschreibt die Vorgehensweise anhand des Bewertungstextes für eine Waschmaschine. In **Schritt 1** wird der Text für die weiteren Schritte vorbereitet. Dafür wird der Text einer Tokenisierung unterzogen. Dabei handelt es um einen Prozess, bei dem ein Text in kleinere Einheiten, sogenannte Token, aufgeteilt wird. Diese Tokens sind im Allgemeinen Wörter oder Satzzeichen, die dem Text eine Bedeutung geben und ihm eine Struktur verleihen. Der Text wird dazu zunächst anhand von Satzzeichen in einzelne Sätze unterteilt und anschließend anhand von Leerzeichen in einzelne Wörter bzw. Token transformiert. Eine anschließende Normalisierung entfernt Groß- und Kleinschreibung sowie Sonderzeichen. Der letzte Schritt der Textvorbereitung filtert die Liste an Tokens um irrelevante Tokens wie etwa Füllwörter oder Stoppwörter (und, oder, usw.) (Abb. 15.2).

Zu Beginn von **Schritt 3** fallen erneut Maßnahmen an, um die Liste an Tokens für die Sentiment-Analyse vorzubereiten:

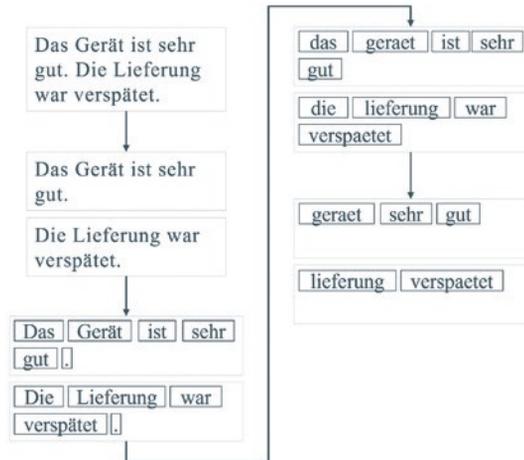
1. Spacy's „de_core_news_sm“ Modell wird eingesetzt um Abhängigkeiten zu modifizierenden Wörtern sowie Part-of-Speech-Tagging (POS) durchzuführen. Das Modell basiert auf einem Convolutional Neural Network (CNN) und einem Bidirectional Long Short-Term Memory Network (BiLSTM), die auf einer großen Menge an Textdaten trainiert wurden, wie etwa auf den TIGER und WikiNER Corpora.

NLP-Pipeline zur Sentiment-Analyse auf Kundenmeinungen

- Erfasste Bewertungen als Freitext
 1. **Schritt:** *In Sätze unterteilen, Tokenisieren, Sonderzeichen und Punctuation entfernen*
- Tokenisierte Sätze
 2. **Schritt:** Sentiment-Analyse mit SentiWS
- Bewertung Emotion
 3. **Schritt:** Klassifizierung mit RoBERTa
- Klassifizierung nach BMC
 4. **Schritt:** Bewertung der Emotion summieren
- Summierte Bewertung pro BMC Kategorie

Abb. 15.1 NLP-Pipeline zur Sentiment-Analyse auf Kundenmeinungen

Abb. 15.2 Schrittweise Vorbereitung der Ausgangstexte durch Tokenisierung, Normalisierung und Filterung



2. Mithilfe von Spacy's compound dependency werden zusammengesetzte Wörter identifiziert.

Mithilfe des German Sentiment Analysis Toolkit (Remus et al., 2010) werden in **Schritt 3** dann zu jedem identifizierten Aspekt die zugehörigen Meinungswörter identifiziert und Aspekt-Meinungs-Paare gebildet. Dafür werden die durch das POS identifizierten Beziehungen auf die folgenden drei Regeln hin überprüft:

- Wenn der Beziehungstyp als „adjectival modifier“ (amod) getaggt ist und das Hauptwort ein Adjektiv ist, dann liegt ein Meinungspar vor.
- Wenn der Beziehungstyp als „nominal subject“ (nsubj) getaggt ist und das Hauptwort verschiedene Rollen einnehmen kann (z. B. Nomen oder Verb) dann liegt ein Meinungspar vor.
- Wenn der Beziehungstyp als „direct object“ (dobj) getaggt ist und das Hauptwort ein Adjektiv ist, dann liegt ein Meinungspar vor.

Im Anschluss wird überprüft ob zu diesen Paaren einen Adjektiv-Modifikator (sehr, extrem, unglaublich) existiert oder sich eine Negation mit dem Paar verbinden lässt. Je nach Modifikator wird das Paar stärker oder schwächer gewichtet. Eine Negation kehrt das Ergebnis ins Negative um. Das Ergebnis der durchgeführten Analyse aller Kundenbewertungen ist eine Liste von Aspekten, welchen jeweils eine Liste von positiven und negativen Meinungswörtern zugeordnet ist, deren jeweilige Sentiment-Bewertung und die gesamte Sentiment-Bewertung.

Um die identifizierten Aspekte auf die einzelnen Kategorien und Subkategorien des Business Model Canvas abzubilden wurde das Sprachmodell RoBERTa (Liu et al., 2019)



Abb. 15.3 Abbildung von Textsegmenten auf GM-Aspekte

verwendet. RoBERTa ist ein Sprachmodell, das auf der Transformer-Architektur basiert und eine Variante von BERT (Bidirectional Encoder Representations from Transformers) ist. RoBERTa wurde mit einer größeren Menge an Trainingsdaten und einem längeren Training als BERT trainiert, was zu verbesserten Leistungen in verschiedenen Sprachaufgaben führt. Für Klassifizierung von Texten, wie sie im vorliegenden Kapitel benötigt wird, wird die Fähigkeit von RoBERTa genutzt, die semantischen Zusammenhänge in Texten zu verstehen und zu lernen, wie verschiedene Texte in bestimmte Kategorien oder Klassen eingeordnet werden können. Zu diesem Zweck wurde das Modell anhand der in Kap. 8 entwickelten Taxonomie trainiert. Eine auf dem trainierten Modell durchgeführte Mehrklassen-Klassifizierung ermöglicht es nun, die in den vorherigen Schritten identifizierten und emotional bewerteten Aspekte anhand des Business Model Canvas zu klassifizieren. Abb. 15.3 zeigt die Bewertung der einzelnen Textbausteine anhand der Taxonomie, sowie die Kategorisierung anhand der BMC.

In **Schritt 4** werden zuletzt sämtliche auf eine Kategorie des BMC abgebildeten Aspekte und deren emotionale Bewertungen summiert um, so eine Gesamtbewertung über eine Vielzahl von Kundenrezensionen zu erhalten.

15.3.3 Visualisierung der Analyse

Abb. 15.4 zeigt die für die Visualisierung entworfene grafische Benutzeroberfläche des Prototyps, welche die in Abschn. 15.2 aufgelisteten Geschäftsmodellaspekte der obersten Ebene sowie ihre direkten Unterkategorien darstellt. Positiv bewertete Aspekte werden in Grün hervorgehoben, negative in Rot. Zusätzlich können einzelne Aspekte mit dem Symbol einer Glühlampe gekennzeichnet sein. Dies bedeutet, dass zu diesem Aspekt eine Empfehlung existiert. Durch Anklicken der Schaltfläche öffnet sich ein weiteres



Abb. 15.4 Die Benutzeroberfläche des Demonstrators zeigt die Bewertung der einzelnen Geschäftsmodellaspekte

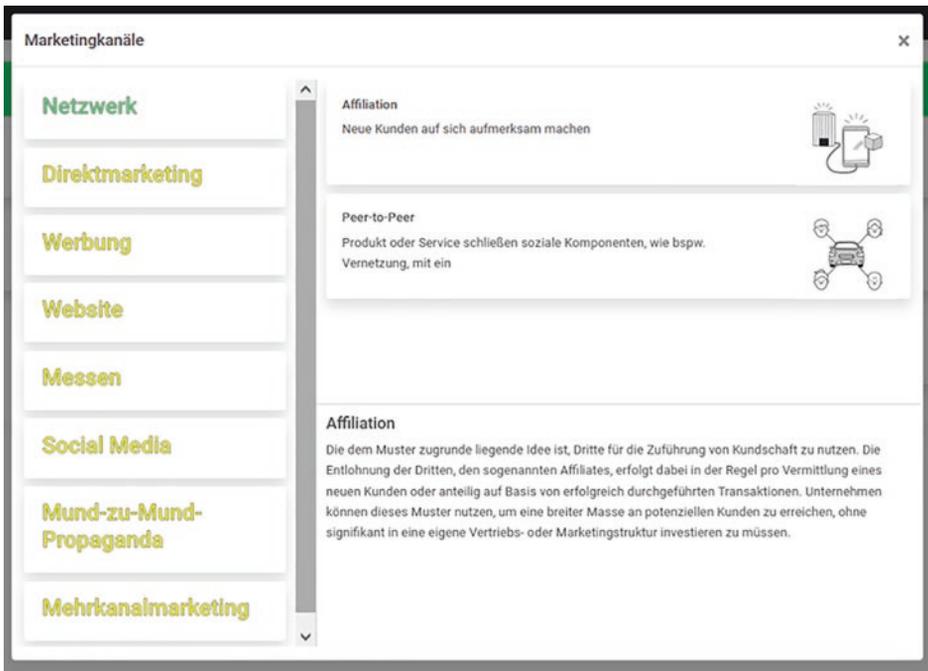


Abb. 15.5 Empfehlung von Geschäftsmodellmustern

Element (siehe Abb. 15.5), welches eine Liste von möglichen passenden Geschäftsmodellmuster zur Verbesserung des Geschäftsmodells auflistet. Die Liste besteht aus Mustern, welche die darauf abzielen die negativen Aspekte zu verbessern, ohne zu große negative Auswirkungen auf die positiv bewerteten Aspekte zu haben.

15.4 Fazit

Die im Kapitel durchgeführten Analysen haben gezeigt, dass mithilfe von Natural Language Processing aus, auf den ersten Blick unstrukturierten Daten, wie etwa im Freitext verfasste Produktbewertungen, sinnvolle Rückschlüsse auf Teile eines Geschäftsmodells gezogen werden können. Es wurden Möglichkeiten gezeigt, die Stimmung des Verfassers einer solchen Bewertung zu erfassen, selbst wenn diese sich innerhalb eines Satzes vom Positiven ins Negative wandelt und diese Stimmung in Form einer Bewertung einem Aspekt des Produktes zuzuordnen ist. Die hier gewonnenen Erkenntnisse bieten Unternehmen das Potenzial einen wichtigen Aspekt der Wertschöpfungskette, den Kunden, passiv in ihre Datenanalysen zu integrieren.

Der vorliegende Demonstrator zeigt exemplarisch, wie unstrukturierte Daten aus Produktbewertungen mit dem konkreten Geschäftsmodell eines Unternehmens verknüpft werden können. Weiterhin wird die Stimmung eines Verfassers erhoben, um so die Einstellung zum Produkt oder bestimmten Produkteigenschaften zu ermitteln. Zwar gibt es viele betriebswirtschaftliche Konzepte, die den einzelnen Baustein eines Geschäftsmodells zugrunde liegen, so müssen jedoch konkrete Produkteigenschaften, Absatzkanäle, Zulieferer, etc. und insbesondere Eigennamen individuell pro Unternehmen konfiguriert werden. Weiterhin empfiehlt es sich eine konkrete Auswahl der geeigneten Daten zu ermitteln, um so eine effizientere und zielgerichtete Suche zu ermöglichen. KMU erhalten somit ein einfaches, aber flexibles Werkzeug, um den Fit ihres Geschäftsmodells auf Basis der aktuellen Datenlage kontinuierlich weiter zu bewerten.

Im Sinne einer Weiterentwicklung des vorgestellten Prototyps in Richtung eines kommerziellen Produktes bedürfte es einer vollständigen Abbildung der vorgestellten Ontologie-basierten Taxonomie in die entsprechenden NLP-Konstrukte. Weiterhin muss der Aufwand für Anwender reduziert werden, in dem die Dateneingabe und -auswahl noch weiter vereinfacht wird und auf branchenspezifische Referenzmodelle zurückgreift.

Zum Abschluss dieses Kapitels sei auf den weiterführenden Ausblick von Kap. 20 verwiesen, die einen Blick auf die künftige Entwicklung Industrieller Datenanalysen vor dem Hintergrund menschlicher, technischer und organisatorischer Einflussfaktoren werfen.

Literatur

- Federici M., & Dragoni, M. (2017). A branching strategy for unsupervised aspect-based sentiment analysis. *EMSASW@ ESWC*, 1–16
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining* (S. 168–177)
- Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2016). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing*, 33, 11–26.

- Liu, Y., Ott, M., & Goyal, N., et al. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv Prepr arXiv190711692*, 1–13
- Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. Wiley.
- Remus, R., Quasthoff, U., & Heyer, G. (2010). SentiWS-A publicly available german-language resource for sentiment analysis. *LREC*, 1168–1171.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Vorgehensweisen für menschenzentrierte Datenanalyseprojekte

16

Eine Reise zum erfolgreichen Umsetzen von datengetriebenen Projekten

Adrienne Bühler, Andrea Rösch, Anette Schötz, Birko Ruzicka, Céline Nordmann, Daniaal Dar, Hermann Ferstl , Julian Joachimbauer, Michael Doukas, Tanja Sindram und Thomas Huber

A. Bühler · A. Rösch · A. Schötz · B. Ruzicka · C. Nordmann · D. Dar · H. Ferstl (✉) · J. Joachimbauer · M. Doukas · T. Sindram · T. Huber
mosaiic GmbH, München, Deutschland
E-Mail: hermann.ferstl@mosaiic.com

A. Bühler
E-Mail: adrienne.buehler@mosaiic.com

A. Rösch
E-Mail: andrea.roesch@mosaiic.de

A. Schötz
E-Mail: anette.schoetz@mosaiic.de

B. Ruzicka
E-Mail: birko.ruzicka@mosaiic.com

C. Nordmann
E-Mail: celine.nordmann@mosaiic.com

D. Dar
E-Mail: daniaal.dar@mosaiic.com

M. Doukas
E-Mail: michael.doukas@mosaiic.com

T. Sindram
E-Mail: tanja.sindram@mosaiic.com

T. Huber
E-Mail: thomas.huber@mosaiic.de

16.1 Motivation

Der Umgang mit *Industrial Data Science* (IDS) Projekten in einem Unternehmen ist ein facettenreiches Thema, das sowohl aus technischer als auch aus fachlicher, organisatorischer und menschlicher Perspektive betrachtet werden sollte. IDS, im deutschsprachigen Raum auch Industrielle Datenanalysen genannt, ist ein Sammelbegriff für datenwissenschaftliche Werkzeuge, die faktenbasierte Entscheidungen für industriellen Anwendungsfälle ermöglichen (siehe Kap. 4, West et al., 2021, S. 131).

Es gibt eine Reihe von iterativen Aktivitäten, die als Data Science Vorgehensmodell bezeichnet werden und dabei helfen können, IDS-Projekte methodisch zu unterstützen. Die Dimensionen Change-Management und User Experience (UX) sind allerdings bei diesen Modellen nicht ausreichend beleuchtet. Außerdem werden die unterschiedlichen Kundenkulturen in verschiedenen Branchen und Firmengrößen vernachlässigt. Das Ziel des Kapitels ist es diese Lücke zu schließen, indem ein modulares Vorgehensmodell definiert wird, welches die Einführung von IDS im Dreiklang aus Mensch, Technik und Organisation unterstützt. Mittels diesem Modell kann von der Idee bis hin zum produktiven Einsatz strukturiert vorgegangen werden. Das Vorgehensmodell liefert konkrete Handlungsempfehlungen für die technische Lösungsentwicklung und -umsetzung.

16.2 Interviewstudie zum Erfolgsfaktor Mensch und abgeleitetes Reifegradmodell

Im Rahmen des Forschungsprojekts AKKORD (siehe Kap. 2) wurde im Jahr 2020 eine branchen-, berufs- und altersübergreifende qualitative Interviewstudie zum Thema „Erfolgsfaktor Mensch in Data-Science-Projekten“ durchgeführt. Fokus der Studie war es zu verstehen, welche Erfahrungen zu Erfolgsfaktoren und Hemmnissen bei der Durchführung von IDS-Projekten von den Interviewpartner:innen gesammelt wurden, was die Mitarbeiter:innen (MA) im Bereich IDS und Künstliche Intelligenz (KI) antreibt, wie die MA diese Themen in ihrem Arbeitsumfeld erleben und welche Change-Management Maßnahmen bereits umgesetzt wurden und erfolgskritisch sind.

Es wurden mehr als 50 Personen unter Verwendung eines teilstrukturierten Interviewleitfadens befragt. Abb. 16.1 gibt eine Übersicht über die Verteilung und die Herkunft der Interviewpartner:innen. Die Ergebnisse der beschriebenen Studie zeigen, dass die Forschung rund um Data Science bei vielen Menschen Bedenken und Ängste auszulösen scheint. Die Gründe hierfür sind vielfältig. Die gesamten Ergebnisse aus der AKKORD Interviewstudie sind im Whitepaper zur Interviewstudie zusammengefasst (mosaic, 2021).

Ausgewählte Inhalte der Interviewstudie wurden in der Veröffentlichung „Industrial Data Science erfolgreich implementieren“ integriert (Nolte et al., 2020). Die Ergebnisse der Studie sind in die Vorgehensmodelle, die in den folgenden Kapiteln beschrieben werden, eingeflossen.

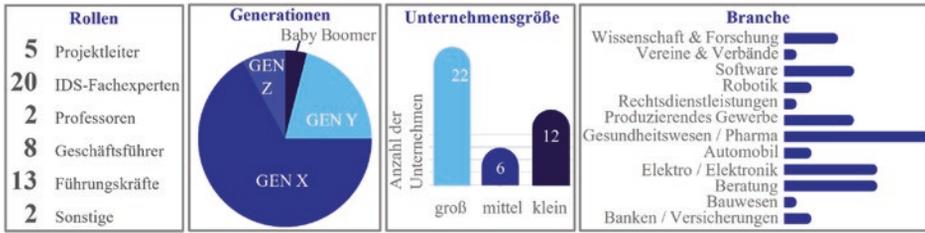


Abb. 16.1 Rahmenbedingungen der Studie zum „Erfolgsfaktor Mensch in Data-Science-Projekten“

In der erwähnten Interviewstudie wurden die Anforderungen und Erfolgsfaktoren für Data-Science-Projekte der beteiligten Unternehmen identifiziert und wie in Abb. 16.2 dargestellt in ein Reifegradmodell übertragen. Dabei wurden die Bereiche Organisation, Technik und Mensch zugrunde gelegt.

Das Reifegradmodell bietet mittelständischen Unternehmen die Möglichkeit, anhand eines Fragebogens den Reifegrad zu identifizieren. Das Ergebnis stellt eine Reifegradstufe zwischen 1 und 3 dar und zeigt die Stärke der Ausprägungen je Erfolgsfaktor. Damit ermöglicht das Reifegradmodell eine Identifikation potenzieller Entwicklungsfelder für das gesamte Unternehmen sowie ein systematisches Erarbeiten relevanter Maßnahmen (siehe Abb. 16.2).

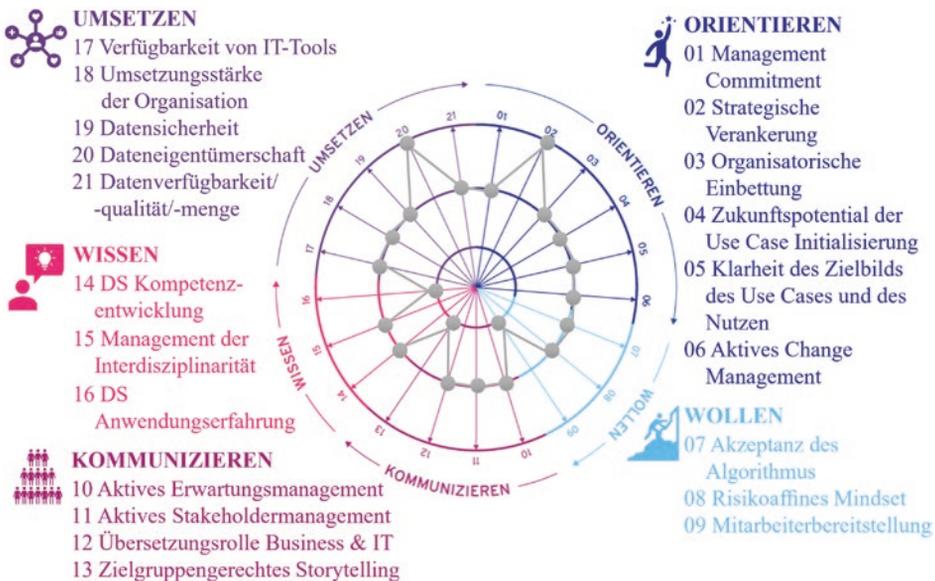


Abb. 16.2 Das Reifegradmodell mit den Erfolgsfaktoren 01–21 und mit beispielhaften Ausprägungen in Grau

Aus der Analyse der Ist-Situation ergibt sich eine Differenz zur optimalen Reifegradstufe 3. Diese Reifegradstufe indiziert, dass das Unternehmen gute Bedingungen für IDS-Projekte bietet. Das Reifegradmodell wurde im Jahr 2021 in zwei Durchläufen mit einem End-Anwender validiert.

16.3 Mittelstandsstudie zu Hemmnissen bei der Einführung von Data Science und Künstlicher Intelligenz

Die Digitalisierung und die Nutzung von IDS und KI bleiben im deutschsprachigen Raum hinter den Erwartungen zurück. Um besser zu verstehen, was Unternehmen derzeit daran hindert, sich mit *Data Science* und *Künstlicher Intelligenz* (DS/KI) zu beschäftigen und was aus Sicht der Unternehmen benötigt wird, wurde die Studie „Data Science und Künstliche Intelligenz für Unternehmen nutzbar machen“ durchgeführt.

Ziel der Studie war es, bestehende Hemmnisse, Vorgehensansätze sowie Bedürfnisse von KMUs bezüglich der Anwendung von Data Science zu verstehen und die notwendigen Voraussetzungen und Herangehensweisen zu ermitteln. Aus den gewonnenen Erkenntnissen sollten Vorgehensmodelle und Empfehlungen, insbesondere für KMUs abgeleitet und entwickelt werden. Für die Studie wurde eine mehrstufige Vorgehensweise gewählt (siehe Abb. 16.3).

Auf Basis von Recherchen und Analysen bereits existierender Studien wurde ein Onlinefragebogen erarbeitet, um die Erfahrungen, Voraussetzungen und Hemmnisse der Unternehmen erfassen bzw. verifizieren zu können. Aus dem Teilnehmerkreis der Online-Befragung wurden bei sieben Unternehmen zusätzlich Interviews durchgeführt, um die Erfahrungen und Hemmnisse noch besser zu verstehen und die gewählten Lösungsansätze und Vorgehensweisen aufzunehmen.

Die Ergebnisse der Online-Befragung als auch der Interviews wurden dann im Rahmen von Workshops ausgewertet und analysiert, um Vorgehensmodelle/Best Practices und Empfehlungen für Unternehmen abzuleiten (siehe Abb. 16.4).

An der Onlinebefragung haben 199 Personen aus Deutschland (49 %), Österreich (46 %), der Schweiz (4 %) und anderen Ländern (1%) teilgenommen. Bezüglich der



Abb. 16.3 Die vier Phasen der KMU-Interviewstudie

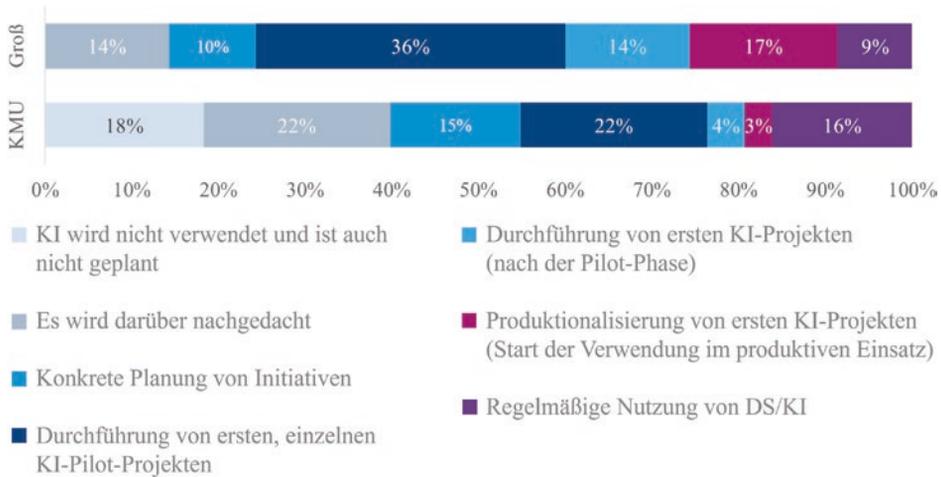


Abb. 16.4 Verwendung von DS/KI in KMUs

Unternehmensgröße waren 36 % Kleinunternehmen (<49 MA), 21 % mittlere Unternehmen (50–500 MA) und 43 % Großunternehmen (>500 MA). Zentrale Erkenntnisse der Online-Befragung sind:

- **Status Quo Digitalisierung:**
 - Ein Großteil der Unternehmen ist mittlerweile digitalisiert, oder hat mit der Digitalisierung begonnen. Es gibt aber weiterhin noch einen großen Digitalisierungsbedarf.
 - Kleine Unternehmen sind bezüglich der Digitalisierung schon weiter als Großunternehmen. Jedoch geht bei den kleinen Unternehmen die Schere weit auseinander. Die einen sind schon weit, die anderen stehen noch ganz am Anfang oder wollen nicht.
 - Angestellte sehen DS/KI am häufigsten als wichtiges Digitalisierungsthema. Im Management und Top Management werden hingegen digitale Geschäftsprozesse am häufigsten genannt.
- **Status Quo Data Science und KI:**
 - Generell gibt es ein überwiegend positives Mindset zum Thema DS/KI. Nur ein kleiner Teil (15 %) sehen in Ihrem Unternehmen ein eher negatives Mindset. In den überwiegenden Fällen steht die Geschäftsführung positiver gegenüber DS/KI als der Rest des Unternehmens.
 - Vielen Unternehmen (31 %) nutzen DS/KI bereits. Nur 10 % haben das gar nicht vor. Der überwiegende Teil der Unternehmen ist derzeit in der Planung bzw. führt bereits erste Pilotprojekte durch.
 - Descriptive und diagnostische Analyseverfahren sind am meisten verbreitet. Predictive und prescriptive sind derzeit noch sehr wenig verbreitet, werden aber bereits von sehr vielen diskutiert bzw. geplant.

Fast alle Teilnehmer geben an, dass die DS/KI Projekte in ihren Unternehmen erfolgreich bis sehr erfolgreich waren. Wie uns auch in den Interviews bestätigt wurde, stellt sich der Erfolg meist nach ca. zwei Jahren ein.

Hemmnisse: Als größte Hemmnisse werden fehlendes Fachwissen, fehlende Nutzenabschätzung und fehlendes Vertrauen in den möglichen Mehrwert angegeben. Fehlender Mut, Skepsis der Mitarbeiter:innen, Akzeptanz beim Kunden sowie Datenschutzgründe werden hingegen nicht als Hemmnisse gesehen. Tendenziell werden bei Großunternehmen mehr Hemmnisse gesehen als bei KMUs.

Voraussetzungen für den Start von DS/KI Projekten: Kleine Unternehmen (<50 MA) nennen deutlich die Finanzierung, die Nutzenbetrachtung, sinnvolle Anwendungsfälle und die Qualifizierung als wesentliche Voraussetzungen für die Initialisierung von DS/KI Projekten. Bei mittleren Unternehmen (50–500 MA) fallen die meisten Nennungen auf die Nutzenbetrachtung, die Ressourcenbereitstellung als auch die Datenverfügbarkeit. Bei den Großunternehmen liegt der Schwerpunkt auf der Nutzenbetrachtung, sinnvollen Use-Cases und der Datenverfügbarkeit als Voraussetzungen.

Wünsche der Unternehmen: KMUs wünschen sich vor allem schnellere und pragmatischere staatliche Förderungen und Budget, um DS/KI in ihrem Unternehmen schneller voranzubringen. Bei Großunternehmen besteht vor allem der Wunsch nach einem entsprechendem Mindset sowie nach mehr Fachwissen.

Die Ergebnisse wurden u. a. auch über die Akkord-Website (www.akkord-projekt.de) sowie über den Bundesverband Deutscher Mittelstand (BVMW, www.bm-mittelstand.com) verbreitet.

16.4 Vorgehensweisen in Data-Science-Projekten

Die Erkenntnisse aus den durchgeführten Interviewstudien fließen in das entwickelte Vorgehensmodell ein, welches eine Roadmap für die Initiierung und Durchführung von IDS-Projekten darstellt. Das Vorgehensmodell (Abb. 16.5) basiert auf dem Industriestandard Cross-Industry Standard Process for Data Mining (Chapman et al., 2000, S. 13 ff.) und erweitert den Prozess in Richtung Technologie, Mensch und Organisation. Mehr Details über das CRISP DM sind in Kap. 3 vorzufinden. Die einzelnen Module, sowie ihre Zwischeniterationen, lassen sich je nach Problemstellung unterschiedlich gewichten. Allerdings spielt jedes Modul eine entscheidende Rolle für den Erfolg des Gesamtprojekts. Die Rückkopplungsschleifen zwischen den Modulen deuten an, dass es keinen vordefinierten Endpunkt in dem Prozess gibt. Jeder iterative Durchlauf verbessert die Qualität.

Im Modell werden die unterschiedlichen Data Science Ansätze Descriptive, Diagnostic, Predictive und Prescriptive Analytics beleuchtet (vgl. West et al., 2022, S. 612 ff.) und alternative Vorgehensmodell, wie KDD (Knowledge Discovery in Databases) einbezogen. Es werden praxisnahe Priorisierungsverfahren für KI-Projekte und die fachlichen und organisatorischen Rollen in IDS-Projekten beschrieben. Sowohl Begriffe wie Datentypen, Skalenniveaus, Datenpipelines und Datenberichte als auch

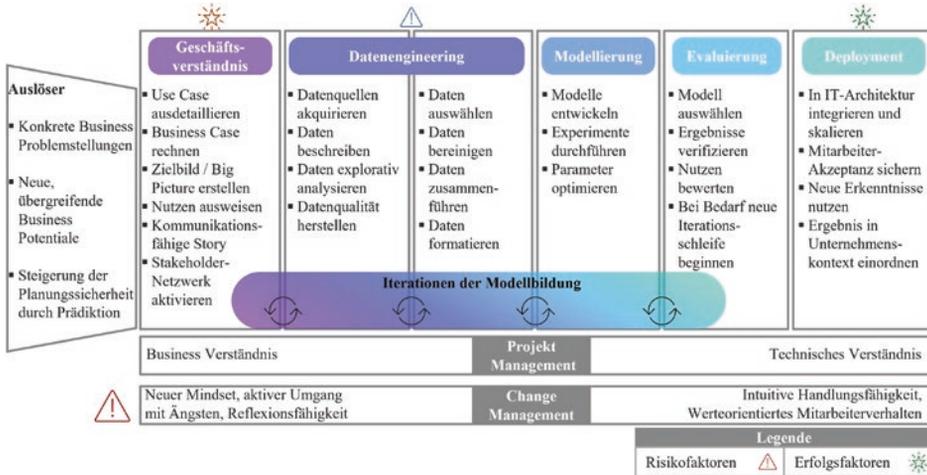


Abb. 16.5 Das Industrial Data Science-Vorgehensmodell

Key Performance Indicators, Measures und Metrics sind erläutert. Der Fokus liegt auf den Disziplinen Machine Learning und Deep Learning, sowie Learning-Typen, Problemstellungen und Modellierungsmethoden. Aus Unternehmenssicht sind wichtige Evaluierungsziele und bewährte Deployment-Strategien inkludiert.

16.4.1 Die Dimension menschenzentriertes Design im Vorgehensmodell

Die Ergebnisse der Studie bestätigen, dass IDS-Vorhaben vor allem durch technologische und wirtschaftliche Ziele motiviert sind. Strukturierte Vorgehensweisen zur Einbindung von Nutzern sind bisher wenig etabliert. Da der wertschöpfende Einsatz der IDS-Lösungen aber letztlich von der Akzeptanz und Nutzbarkeit durch die damit interagierenden Menschen abhängt, wird eine an den Nutzern orientierte Gestaltung zu einem entscheidenden Erfolgsfaktor.

Im bisherigen Vorgehensmodell für IDS-Projekte wird der Fokus hauptsächlich auf die Phasen des Daten-Engineerings und der Modellentwicklung gelegt. Zwar spielt das Geschäftsverständnis zu Beginn eines Vorhabens eine zentrale Rolle, jedoch bezieht dieses nicht explizit die Interessen und Bedürfnisse, der späteren Nutzer des Data-Science Produktes mit ein. Da die in den Interviews geschilderten Projekterfahrungen darauf hinweisen, dass die Nutzbarkeit und Nützlichkeit einer IDS-Lösung letztlich über deren Erfolg entscheiden, wurde eine Erweiterung des Modells um nutzerzentrierte Vorgehensweisen erarbeitet.

Ziel der Erweiterung des IDS-Vorgehensmodells ist es, relevante Aktivitäten aus dem „Menschenzentrierten Gestaltungsprozess“ (ISO 9241-210:2019 2019) zu integrieren.

Dabei handelt es sich um geeignete Methoden, die Erfordernisse der jeweiligen Nutzer systematisch zu erheben, in Konzeption und Entwicklung zu berücksichtigen und schließlich hinsichtlich der erreichten Nutzungsqualität zu evaluieren.

Die Recherchen zum aktuellen Forschungsstand für die Ergänzung des Vorgehensmodells fokussierten das Thema menschenzentrierte KI (Shneiderman, 2020; Xu, 2019; Xu et al., 2021; Yang et al., 2018). Dieses Themengebiet beschäftigt sich mit der nutzerzentrierten Gestaltung einer KI. Die wesentlichen Erkenntnisse aus der Forschungsliteratur wurden vor dem Hintergrund der praktischen Erfahrungen aus UX-Projekten der mosaici GmbH hinsichtlich ihrer Bedeutung und Wirkung auf UX-Aktivitäten bewertet. Dabei stand die Frage im Mittelpunkt, durch welche UX-spezifischen Phasen und Schritte das bestehende IDS-Vorgehensmodell erweitert werden kann, um heutige Defizite in der Nutzereinbindung und Nutzerakzeptanz zu beheben.

Für die Validierung des UX-spezifisch erweiterten IDS-Vorgehensmodell bildete der innerhalb des Akkord-Forschungsprojekts umgesetzte Anwendungsfall für ein „Übergreifendes, prädiktives Industrial Engineering“ den Rahmen (siehe Kap. 9). In Experteninterviews mit den zentralen Projektrollen wurde erfragt, welche Herausforderungen und Erfolge im Projekt verzeichnet wurden, die entweder auf einen Mangel an bzw. ein Vorhandensein von nutzerzentrierten Vorgehensweisen zurückzuführen sind.

Das resultierende, mit UX-Aktivitäten erweiterte Vorgehensmodell (Abb. 16.6) greift die Hauptphasen aus dem IDS-Modell – Geschäftsverständnis, Datenengineering, Modellierung, Evaluierung und Deployment auf.

Der Schwerpunkt der Erweiterung findet sich am Beginn des Vorgehensmodells mit der Aufsplittung der Phase „Geschäftsverständnis“ in die drei Schritte „Scoping“, „Understanding“ und „Definition“. Hier wird vor allem dem Ziel Rechnung getragen, neben den technologischen Aspekten die Bedarfe der beteiligten Menschen von Anfang in den Fokus zu stellen. Als neue Phase wurde „Conception“ ergänzt, die neben den bekannten konzeptionellen DS-Aktivitäten auch die Erarbeitung und prototypische Verprobung eines Nutzungskonzeptes beinhaltet.

Die in den einzelnen Phasen verorteten Aktivitäten wurden jeweils den Hauptakteuren zugeordnet. Berücksichtigt wurden in dieser integrierten Sicht die Rollen UX-Professional und Data Scientist. Bei einigen Aktivitäten sehen wir eine direkte Zusammenarbeit bzw. einen Austausch von Wissen zwischen den beiden Rollen als erforderlich (Rolle „Gemeinsam“). Durch das integrierte Vorgehensmodell kann die erfolgsrelevante Vernetzung der verschiedenen Disziplinen, hier IDS-Entwicklung und Human Centered Design (UX) erreicht werden (Abb. 16.7).

Die frühe Einbindung von Personen, die entweder als Datenlieferanten oder Nutzer des Outputs agieren, und die Integration nutzerzentrierter Methoden steigern sowohl die Effizienz des Projekts als auch die Qualität der Prognosen.

Durch das tiefe Verständnis des Business-Kontextes können schnelle und zielführende Modelländerungen vorgenommen werden. Ein nicht zu verachtender erfolgssteigernder Effekt ist die erhöhte Nutzungsbereitschaft, die durch die Einbindung der späteren Nutzer in den Entwicklungsprozess einhergeht.

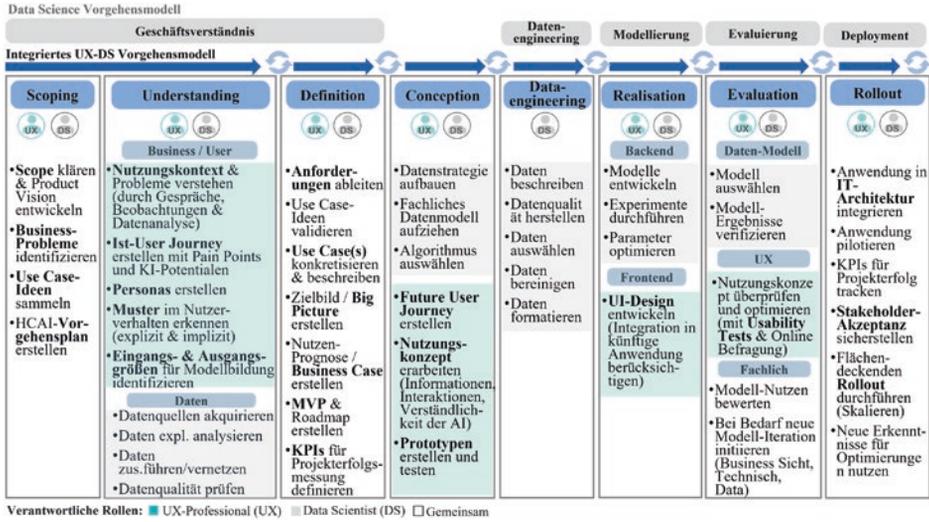


Abb. 16.6 Integriertes Vorgehensmodell Data Science – UX

Scoping	Understanding	Definition	Conception	Data-engineering	Realisation	Evaluation	Rollout
Klarheit über Ziele und Vorgehen	Tiefes Verständnis von Business, Nutzer und Datengrundlage	Use Case basiert auf abgeleiteter Nutzungsanforderungen & Datenverständnis	Verständliche und vom Nutzer steuerbare Nutzungskonzepte	Nutzerzentrierte Evaluierungsergebnisse für eine hohe Nutzerakzeptanz der DS-/KI-Anwendung			Erfolgreicher Rollout durch erhöhte Nutzungsbereitschaft

Abb. 16.7 Der zu erzielende Nutzen durch den Einsatz menschenzentrierter Vorgehensweisen

16.4.2 Die Dimension Change-Management im Vorgehensmodell

Das begleitende Change-Management ist ein wesentlicher Erfolgsfaktor für das Gelingen eines Data-Science-Projekts. Daher ist es sinnvoll im Rahmen eines Change-Vorgehensmodells jede Phase eines Data-Science-Projekts zu begleiten. In jeder Phase eines Data-Science-Projekts ist ein schlanker Change Cycle zu durchlaufen, der die aktuellen Anforderungen des Projekts validiert und das „Change System“ entsprechend flexibel und agil anpasst (Abb. 16.8).

Sinnvolle Change Aktivitäten entlang der Phasen des allgemeinen Vorgehensmodells sind in der folgenden Abbildung dargestellt. Sie stellen eine Zusammenfassung aller wichtigen Ergebnisse aus der Interviewstudie, Lessons Learned und Best Practices dar (Abb. 16.9).

Ansätze für konkrete und bewährte Change Maßnahmen finden sich im White Paper der Interviewstudie zum Erfolgsfaktor Mensch (mosaiic, 2021). Eine konkrete Change Maßnahme wurde mit einem End-Anwender entwickelt und verprobt.

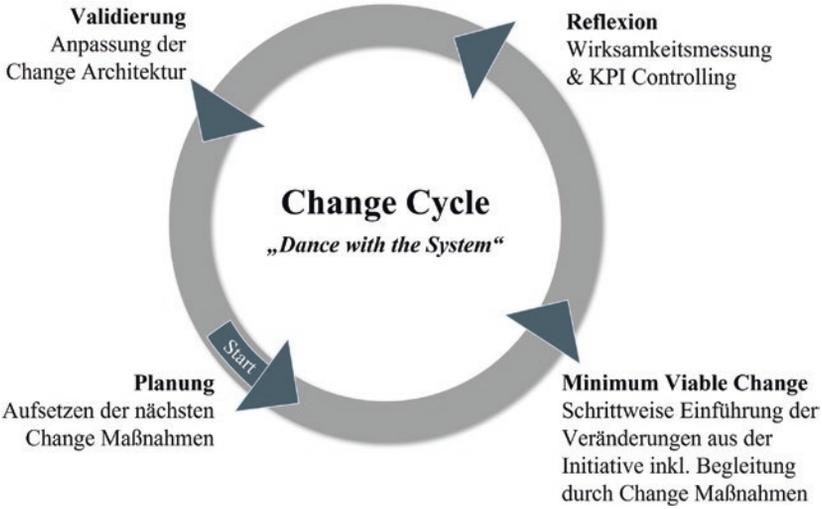


Abb. 16.8 Change Cycles für die Change-Management Vorgehensweise in Data-Science-Projekten

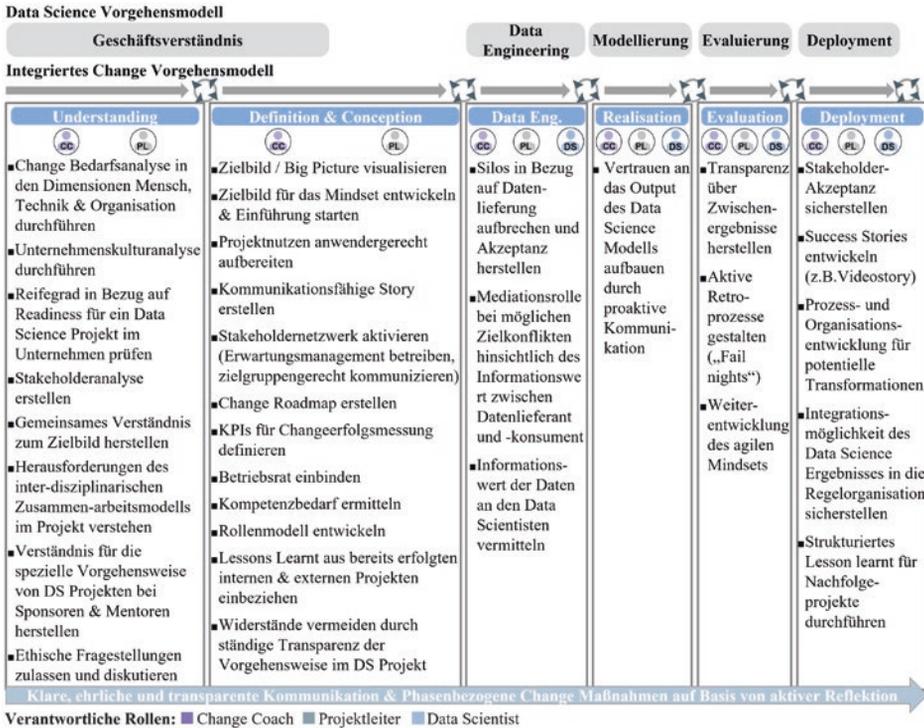


Abb. 16.9 Potenzielle Change Aktivitäten entlang des DS-Phasenmodells

16.5 Fazit

Zusammenfassend kann festgehalten werden, dass die Einführung von IDS-Projekten in Unternehmen ein komplexes Thema ist, welches eine ganzheitliche Betrachtung erfordert. Die in dieser Arbeit entwickelten Modelle und Maßnahmen können dabei helfen, die Einführung von IDS im Dreiklang aus Mensch, Technik und Organisation effektiver zu gestalten (siehe Kap. 20).

Die Analyse der bisherigen Vorgehensmodelle hat gezeigt, dass es bisher keine umfassenden Ansätze gibt, die sowohl technische, fachliche, organisatorische als auch menschliche Aspekte der Einführung von IDS abdecken. Die in diesem Kapitel entwickelten Modelle können dazu beitragen, diese Lücke zu schließen und Unternehmen bei der erfolgreichen Einführung von IDS zu unterstützen.

Das entwickelte Vorgehensmodell liefert konkrete Handlungsempfehlungen für eine technische und fachliche Lösungsentwicklung und -umsetzung von IDS-Projekten. Dabei wurde auch ein Reifegradmodell entwickelt und Change-Maßnahmen sind abgeleitet. Ein UX-Vorgehensmodell wurde ebenfalls erarbeitet, um das volle Potenzial der Benutzer-Interaktion mit dem zu entwickelnden System auszuschöpfen. Es wurden auch die Ergebnisse der umfassenden Studie zur Erfassung der aktuellen Verbreitung von IDS in KMUs präsentiert.

In Zukunft könnte die Anwendung des modularen Vorgehensmodells auf weitere Branchen und Unternehmen ausgeweitet werden. In diesem Zusammenhang sollte zudem die Wirkung der Change-Management-Maßnahmen und des UX-Vorgehensmodells auf den Erfolg von IDS-Projekten genauer zu untersucht bzw. gemessen werden. Darüber hinaus könnte die Studie zur Erfassung der Verbreitung von IDS in KMUs fortgeführt werden, um Trends und Entwicklungen in diesem Bereich besser zu verstehen.

Literatur

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0. Step-by-Step Data Mining Guide, CRISP-DM consortium.
- ISO 9241-210:2019: Ergonomie der Mensch-System-Interaktion – Teil 210: Menschenzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019).
- mosaiic GmbH. (2021). Whitepaper zur Interviewstudie zum Erfolgsfaktor Mensch in Data Science Projekten. https://www.mosaiic.de/wp-content/uploads/2020/11/Akkord_Interviewstudie-zum-Erfolgsfaktor-Mensch-in-Data-Science-Projekten.pdf. Zugegriffen: 24. Febr. 2023.
- Nolte, V., Sindram, T., Mazarov, J., & Deuse, J. (2020). Industrial Data Science erfolgreich implementieren: Interviewstudie zu Erfolgsfaktoren und Hemmnissen. *Zeitschrift für wirtschaftlichen Fabrikbetrieb.*, 115, 734–737. <https://doi.org/10.1515/zwf-2020-1151020>
- Shneiderman, B. (2020). Human-Centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human-Computer Interaction*, 12(3), 109–124. <https://doi.org/10.17705/1thci.00131>.

- Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Association for Computing Machinery, Interactions*, XXVI, 4, 42–46. <https://doi.org/10.1145/3328485>
- Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2021). From human-computer interaction to human-ai interaction: New challenges and opportunities for enabling human-centered AI, 1–73. <http://arxiv.org/abs/2105.05424>.
- Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating how experienced UX designers effectively work with machine learning, 585–596. <https://doi.org/10.1145/3196709.3196730>.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- West, N., Syberg, M., & Deuse, J. (2022). A holistic methodology for successive bottleneck analysis in dynamic value streams of manufacturing companies. In *Proceedings Changeable, Agile, Reconfigurable and Virtual Production Conference*, 8(1), 612–619. https://doi.org/10.1007/978-3-030-90700-6_69.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Teil IV

Verwendung der AKKORD-Plattform



Umsetzung des Backend im Autorenn-Demonstrator

17

Datenanbindung für eine Modellautorennbahn vom Modell Digital 132 als exemplarische Umsetzung des AKKORD-Backends

Thomas Eickhoff  und Jens C. Göbel 

17.1 Konzept des Demonstrators

Das zentrale Ziel des Forschungsvorhabens AKKORD ist die Entwicklung eines Referenzbaukastens für die industrielle Datenanalyse (siehe Kap. 1 und 2). Der Backend-Demonstrator soll ein greifbares und leicht verständliches Beispiel für die in diesem Vorhaben entwickelten Lösungskomponenten aus dem Bereich Datenbackend liefern. In Kap. 4 wurde der entwickelte Referenzbaukasten mit seinen Bausteinen vorgestellt. Konkret sollen die in den entsprechenden Beiträgen vorgestellten Bausteine des Referenzbaukastens demonstriert werden (siehe Kap. 5 und 6). In diesem Kontext dient eine Modellautorennbahn aus der *Digital 132 Serie* als Lieferant für Daten, auf die zunächst mit einem Datenzugriffsbaustein zugegriffen wird. Die einzelnen Modellautos wurden in diesem Kontext als *Internet of Things* (IoT)-Datenquellen interpretiert, die Daten in der Form von Zeitstempeln für Zwischen- und Rundenzeiten und simulierten Tankfüllständen liefern. Gerade in dieser Phase liegt ein besonderer Fokus auf der Gewährleistung einer durchgängigen Datenanalysequalität (West et al., 2021, S. 131 ff.). Die somit verfügbar gemachten Daten werden auf verschiedenen Wegen in einem Datenbackend abgespeichert. Anschließend werden die dort gespeicherten Daten für die Datenanalyse verfügbar gemacht. Der Aufbau des Demonstrators ist hierbei größtenteils mit dem in Kap. 18 beschriebenen Demonstrator für die Analysebausteine identisch, da zu Vergleichszwecken das gleiche Rennbahnmodell beschafft wurde. Eine detaillierte Beschreibung der verfolgten Anwendungsszenarien ist dort zu finden. In diesem Kapitel

T. Eickhoff (✉) · J. C. Göbel
Technische Universität Kaiserslautern, Kaiserslautern, Deutschland
E-Mail: eickhoff@mv.uni-kl.de

J. C. Göbel
E-Mail: goebel@mv.uni-kl.de

soll dementsprechend der Fokus auf dem technischen Aufbau, sowie den eingesetzten Protokollen und Technologien liegen. Der generelle Aufbau des Szenarios aus Backend-Sicht ist in Abb. 17.1 dargestellt.

Zu Testzwecken wurden an den Standorten der beteiligten Forschungspartner in Dortmund und Kaiserslautern jeweils eine Rennbahn in Betrieb genommen. Die Daten der Bahnen wurden über eine optional verfügbare Bluetooth-Schnittstelle ausgelesen und per Netzwerk an das Datenbackend gesendet. Dort können die Daten angereichert und mit weiteren Datenquellen verknüpft werden. Anschließend erfolgt die Weitergabe an die Analysebausteine. Diese konzeptionellen Schritte werden im Folgenden detaillierter beschrieben.

17.2 Auslesen und Übertragen der Daten

Die eingesetzte Produktreihe ermöglicht es, Daten über ein Bluetooth-Modul an mobile Endgeräte zu übertragen, um beispielsweise mit der offiziellen Smartphone-App des Herstellers die Rangliste des laufenden Rennens zu verfolgen. Die Daten stehen aber auch anderen Endgeräten zur Verfügung. So ermöglicht die Python-Bibliothek *carreralib* (github.com/tkem/carreralib) das Empfangen der Daten über standardkonforme Bluetooth-Hardware. Im Kontext des Backend-Demonstrators erfüllt ein RaspberryPi diese Rolle. Ein eigens entwickeltes Python-Skript liest die gesendeten Daten aus und bereitet sie für die Übertragung an das Datenbackend vor. Im Hinblick auf den Analysebaustein-Demonstrator wurde dieses Skript um die Eingabe von Fahrerdaten ergänzt. Diese Vorgehensweise ähnelt den typischen Ansätzen zum Retrofitting in industriellen

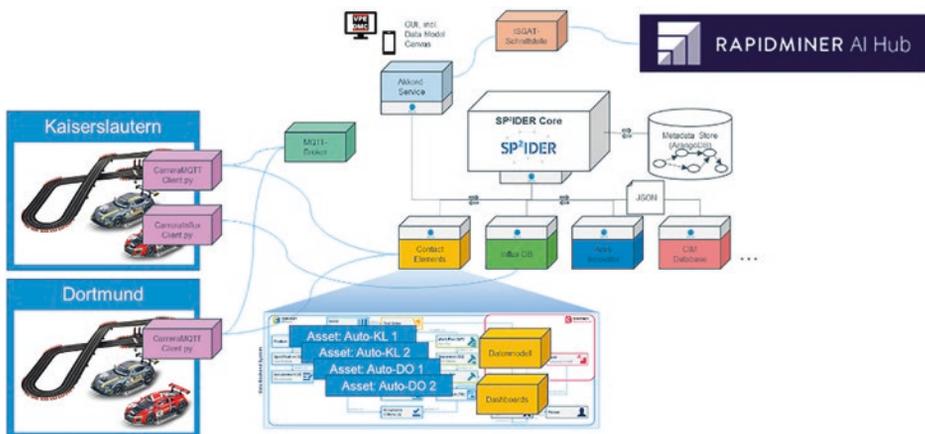


Abb. 17.1 Bausteine des Demonstrators

Anwendungsfällen, in denen häufig auch gezielt Low-Cost-Sensorik verwendet wird (Wöstmann et al., 2019, S. 94 ff.).

Alle verfügbaren Daten werden in einem JSON-Datensatz gesammelt und unter Verwendung des MQTT-Protokolls versendet. Die Verwendung einer solchen Standard-Schnittstelle ermöglicht unterschiedliche Endpunkte für diese Übertragung: So wurde zu Testzwecken zunächst ein bereits vorhandener MQTT-Broker angebunden, an dem die empfangenen Daten in Echtzeit ausgelesen werden konnten. Nach Sicherstellung der grundlegenden Funktionalität wurden zwei weitere MQTT-Broker hinzugefügt, die jeweils in die Software-Produkte Contact Elements for IoT und InfluxDB integriert laufen. Um maximale Autonomie des Demonstrators, beispielsweise bei Messebesuchen sicherzustellen, wurde schließlich auch eine in Kap. 18 beschriebene Anbindung an MariaDB realisiert. Diese etappenhafte Entwicklung liegt einerseits in verschiedenen praktischen Entscheidungen aus der Entwicklungsphase des Projekts begründet, spiegelt gleichzeitig aber auch den modularen Ansatz des entwickelten Referenzbaukastens wieder, da hier einzelne Lösungselemente passend für den anvisierten Einsatzzweck ausgewählt und gegeneinander ausgetauscht werden können.

17.3 Verwaltung der Daten im Backend

Die hauptsächliche Verwaltung der Daten erfolgt in der Software Contact Elements for IoT. Dort ist entsprechend des in Kap. 5 beschriebenen Ansatzes eines zentralisierten Datenbackends eine hochgradige Verknüpfung der Daten der Rennbahn mit beispielsweise PLM-Stammdaten zu den einzelnen Fahrzeugtypen vorgesehen. Für jedes Auto an einer der beiden Rennbahnen wurde ein sogenanntes IoT-Asset angelegt. In einem solchen Asset können Daten über den Standort und die Betriebsbereitschaft des Assets hinterlegt werden. Zusätzlich existiert für jedes IoT-Asset die Möglichkeit, Zeitreihendaten in einer angebundenen Zeitreihendatenbank abzuspeichern, was im Kontext des Demonstrators für die Speicherung der per MQTT empfangenen Daten genutzt wurde. Die Anzeige einer solchen Zeitreihe in der Contact-Oberfläche ist in Abb. 17.2 dargestellt.

Zusätzlich stehen die Vernetzungsmöglichkeiten der Contact Elements Plattform mit dem im Rahmen des Projekts ausdetaillierten Verknüpfungsmöglichkeiten zur Verfügung. Darüber hinaus können für verwaltete Assets interaktive Dashboards zusammengestellt werden, die über die aktuell empfangenen Daten informieren.

17.4 Weitergabe an die Analysemodule

Das bislang beschriebene Vorgehen zur Verwaltung orientiert sich am Ansatz für ein zentralisiertes Datenbackend (siehe Kap. 5). Um eine größere Bandbreite der entwickelten Lösungsbausteine darstellen zu können, wurde dieser für die Weitergabe

The screenshot shows the 'AKKORD - CONTACT Elements for IoT' interface. On the left is a red sidebar with navigation icons. The main content area is titled 'Auto 1 (AS-000003)' and shows a 'default' data table. The table has two columns: 'Datum' and 'Zeit'. The data rows are as follows:

Datum	Zeit
2022-01-14 11:20:54.588	
2022-01-14 11:19:31.486	
2022-01-14 11:19:22.430	
2022-01-14 11:01:36.187	

Abb. 17.2 Darstellung einer Zeitreihe in Contact Elements for IoT

der Daten an die Analysebausteine mit einem dezentralen Ansatz im Sinne von (siehe Kap. 5) verknüpft. Im Kontext dieses Ansatzes ist Contact Elements wiederum eine Datenquelle, welche über die an der Rheinland-Pfälzischen Technischen Universität entwickelte SP²IDER-Plattform angebunden ist.

Am Aufbau des Demonstrators lassen sich die in Abb. 17.3 dargestellten Schritte zum Aufbau der Datenanbindung nachverfolgen: Über einen entsprechenden Connector können die Daten als JSON oder CSV ausgelesen werden. Anschließend können sie im Data Model Canvas einem Datenbedarf gegenübergestellt werden. Die Definition des Datenbedarfs erfolgt hierbei durch eine JSON-Datei, die in der Gegenüberstellung um URLs für den Zugriff auf die ausgewählten Daten ergänzt wird. Diese Datei wird anschließend an die von RapidMiner entwickelte ISGAT-Schnittstelle übergeben, um die Weitergabe an die in RapidMiner AI Hub verwalteten Analysebausteine zu ermöglichen.

17.5 Fazit

Der Datenbackend-Demonstrator zeigt die praktische Umsetzung, der im Rahmen von AKKORD entwickelten Ansätze zum Aufbau einer Backend-Lösung. Die sich daraus ergebende Lösung wird in einem folgenden Schritt mit den Baustein „Datenanalyse“ und „Datennutzung“ integriert, um ein durchgängiges Anwendungsbeispiel für den

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Umsetzung der Analysebausteine im Autorenn-Demonstrator

18

Datenanalysen für eine Modellautorennbahn vom Modell
Digital 132 als exemplarische Umsetzung des AKKORD-
Bausteine

Jörn Schwenken , Nikolai West , Marius Syberg  und Jochen Deuse 

18.1 Konzeptionelle und technische Umsetzung des Demonstrators

Das zentrale Ziel des Forschungsvorhabens AKKORD ist die Entwicklung eines Referenzbaukastens für die industrielle Datenanalyse (siehe Kap. 1 und 2). Der in diesem Kapitel vorgestellte Autorenn-Demonstrator soll ein greifbares und leicht verständliches Beispiel für die entwickelten Lösungskomponenten aus dem Bereich *Datenanalyse* und *Datennutzung* liefern. Wie im vorherigen Kap. 17 ist dieser Anwendungsfall an industrielles Retrofitting bestehender Anlagen angelehnt (Wöstmann et al., 2019, S. 94 ff.). Bereits in Kap. 4 wurde der entwickelte Referenzbaukasten mit seinen Bausteinen vorgestellt. Die konzeptionelle Idee hinter dem Demonstrator ist es, ein anschauliches Anwendungsbeispiel zu erstellen, welches Anwender:innen den Nutzen von AKKORD spielerisch aufzeigt und ihre Akzeptanz für die Verwendung der Plattform erhöht. Voraussetzung hierbei ist, dass das Anwendungsbeispiel Daten erzeugt und sich somit Datenanalysen durchführen lassen. Ziel des Demonstrators ist es, dass er möglichst viele Module aus den Bausteinen der Prozesskette der industriellen Daten-

J. Schwenken (✉) · N. West · M. Syberg
Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: joern.schwenken@ips.tu-dortmund.de

N. West
E-Mail: nikolai.west@ips.tu-dortmund.de

M. Syberg
E-Mail: marius.syberg@ips.tu-dortmund.de

J. Deuse
Institut für Produktionssysteme, Technische Universität Dortmund, Dortmund, Deutschland
E-Mail: jochen.deuse@ips.tu-dortmund.de

analyse umsetzt und validiert. Dies soll demonstrieren, dass die AKKORD-Plattform modular, datengetrieben und für die Anwendung industrieller Datenanalysen geeignet ist. Alle genannten Anforderungen erfüllt der AKKORD-Autorenn-Demonstrator. Eine Zusammenfassung der technischen Systemarchitektur befindet sich in Abb. 18.1. Im Folgenden wird zunächst konzeptionell erläutert, wie der Demonstrator aufgebaut ist. Anschließend wird in den weiteren Unterkapiteln konkreter auf die erstellten Module der Bausteine eingegangen.

Die Arbeiten am Demonstrator basieren auf den vorgestellten Grundlagen von Kap. 3 und den konzeptionellen sowie technischen Ergebnissen von Kap. 6). Im Zentrum des Demonstrators steht eine digitale Modellrennbahn im Maßstab 1:32. Während eines Rennens können hier bis zu vier Rennfahrer:innen mit verschiedenen Autos gegeneinander antreten. Neben der Rennstrecke erweitert eine Boxengasse den technischen Umfang des Demonstrators. Durch die Verwendung von realen Tankfüllständen bekommt das Rennen eine weitere Dimension. Je länger ein Auto fährt, desto geringer wird dessen Tankfüllstand. Gleichzeitig wird das Auto virtuell leichter und kann dementsprechend höhere Geschwindigkeiten erreichen. Die digitalen Daten der Modellrennbahn werden über eine Bluetooth-Schnittstelle eines Raspberry Pis aufgenommen und verfügbar gemacht.

Zu den Datenquellen gehören *Rennevents*, die sich den jeweiligen Autos zuordnen. An drei Messstellen, sogenannten Sektoren, werden Zeitstempel und die Identifikationsnummern der Autos erfasst. Als weitere Datenquelle dienen Daten über den *Tankstand* der Autos. In regelmäßigen Abständen fragt der Raspberry Pi die Füllstände der Autos ab. Dabei erkennt die Abfrage ebenfalls, ob sich ein Auto zum Auftanken in der Boxengasse befindet. Zusätzlich geben die Rennfahrer:innen vor Rennstart *Fahrerinformationen* zu sich an. Zu den statischen Daten gehört ihr Name, ihr Alter und das Auto, das sie fahren. Während des Rennens zeichnen Rennleiter:innen manuell auf, wie viele *Renausfälle* die Rennfahrer:innen haben. Zu den Rennausfällen gehören u. a. das Verlassen der Strecke oder leere Tankfüllstände. Sinnvolle weitere Datenquellen, die

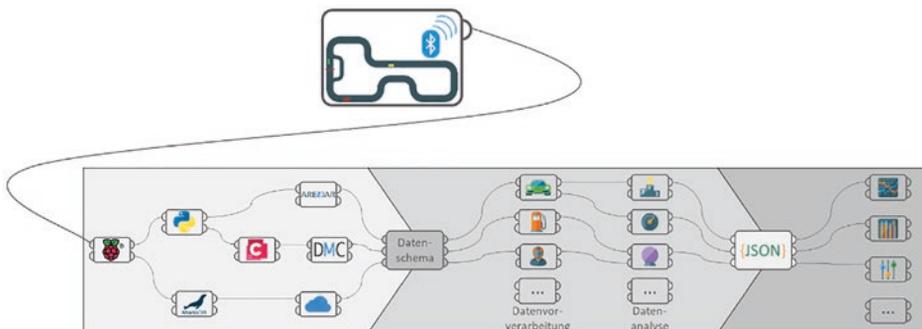


Abb. 18.1 Systemarchitektur AKKORD-Autorenn-Demonstrator

sich auf die Modellrennbahn beziehen, wären *Streckeninformationen* (z. B. Streckenlänge, Kurvenanzahl) oder *Fahrzeuginformationen* (z. B. Gewicht, Bremseinstellungen, Geschwindigkeitseinstellungen). In dieser Ausführung des Demonstrators werden diese jedoch konstant gehalten und dementsprechend nicht berücksichtigt.

Bevor auf Basis der aufgezeigten Daten diverse Analysen entstehen, müssen diese zunächst gesammelt, vorverarbeitet und miteinander verknüpft werden. Es entsteht ein Datensatz mit allen relevanten Attributen, welcher an den Analysebaustein übergeben wird. Hier können im Anschluss diverse Analysen erstellt werden. Zu den denkbaren Analysen gehören die Identifikation von Rennsieger:innen, die Berechnung der maximalen Geschwindigkeit (notwendig: Streckeninformationen), die Prognose zukünftiger Rennsieger:innen oder eine intelligente Ableitung von Tankstrategien.

Damit die Analysen den Anwender:innen einen Mehrwert bieten, sollen Dashboards mit geeigneten Visualisierungen die Analyseergebnisse präsentieren. Eine deskriptive Analyse wie die Identifikation von Rennsieger:innen wird von einem Dashboard mit einer diagnostischen Visualisierung des Rennverlaufs unterstützt. Hierdurch bekommen die Rennfahrer:innen gehaltvolle Informationen über ihr Fahrverhalten. Dashboards zu komplexeren Analysen wie die Prognose zukünftiger Rennsieger:innen sollen den Anwender:innen möglichst verständlich darlegen, warum das Modell diese oder jene Entscheidung getroffen hat.

Der Demonstrator soll den Eintritt für Anwender:innen in industrielle Datenanalysen möglichst niedrigschwellig zugänglich machen. Sie sollen die Analysen des Demonstrators auf ihre industriellen Anwendungsfälle projizieren können, sodass sie Potenziale für ihre individuellen Anwendungen der AKKORD-Plattform entdecken.

18.2 Zugriffsmodule des Demonstrators

Die Aufgabe der Zugriffsmodule ist die Realisierung des Datenzugriffs für die Analyse im Anwendungsfall. Eine ausführliche Definition zum Zugriffsbaustein und seiner Module findet sich in Kap. 4. Vor diesem Hintergrund ergibt sich für die Zugriffsmodule des Autorenn-Demonstrators die Aufgabe, die Daten aus den Datenquellen auszulesen, zu sammeln, zu verarbeiten, zu speichern und zur Verfügung zu stellen. Im Demonstrator wurden hierfür die folgenden Module erfolgreich getestet: MQTT-Modul, Contact Elements Datenbank-Modul, Data Model Canvas-Modul (DMC-Modul, siehe Kap. 5), ARENDAR-Modul, MariaDB Connector-Modul und MariaDB-Modul. Abb. 18.2 zeigt eine Zusammenfassung der Zugriffsmodule und wie diese miteinander verknüpft sind.

Auf dem Raspberry Pi wird ein Python-Skript mit der Bibliothek carreralib ausgeführt. Dieses schickt beispielsweise über MQTT die Daten an verschiedene Datenbanken. Carreralib arbeitet z. B. über eine Bluetooth-Schnittstelle, um die Daten der Modellrennbahn – *Tankstand* und *Rennevents* – zu empfangen. Gleichzeitig nimmt das Programm die statischen *Fahrerinformationen* über die Tastatureingabe auf und ver-

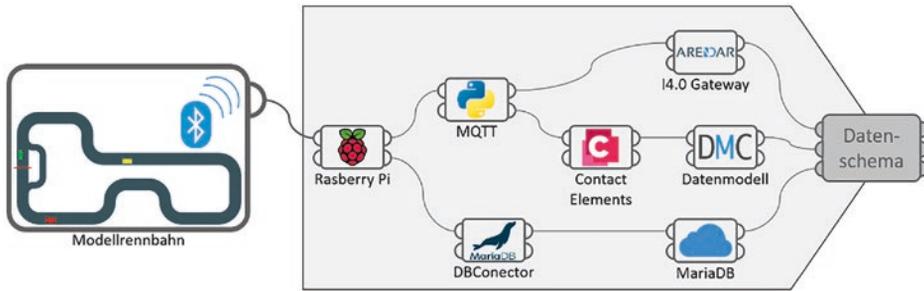


Abb. 18.2 Architektur des Datenzugriffs des AKKORD-Autorenn-Demonstrators

arbeitet sie gemeinsam mit den Daten der Modellbahn. Zuletzt sendet das Programm die Daten über eine MQTT-Schnittstelle an beliebige Empfänger.

Ein Empfänger des MQTT-Moduls ist das Contact Elements Datenbank-Modul. Das Modul sammelt und verarbeitet die gesendeten Daten und speichert sie in mehreren relationalen Datenbanken. Diese Datenbanken werden einzeln über eine Rest-API abgerufen. Das DMC-Modul unterstützt Anwender:innen anschließend methodisch und informationstechnisch beim Aufbau einer fachlichen Datengrundlage zur Datenanalyse. Es erstellt die erforderliche Datenvernetzung für das konkrete Analyseszenario. So erzeugt es aus den einzelnen Datenmodellen ein gebündeltes, gemeinsames Datenmodell für die weitere Analyse.

Ebenso wie das Contact Elements Datenbank-Modul empfängt auch das ARENDAR-Modul die Daten vom MQTT-Modul. Der ARENDAR ist ein sicheres Gateway für industrielle Anwendungen. Im Demonstrator verarbeitet, sammelt und speichert er die Daten. Neben seiner Sicherheit ist ein weiterer Vorteil des ARENDARs, dass er auch für schnelle Analysen eingesetzt werden kann, da er in prototypischen Dashboards die Daten bereits visualisieren kann.

Zum Auslesen der Daten aus der Modellrennbahn wurde ebenso ein zweites Programm geschrieben, welches auf dem Raspberry Pi die Daten über ein MariaDB Connector-Modul versendet. Der Connector ist ebenfalls eine Datenbankschnittstelle. Er basiert auf der Java Database Connectivity-Technologie. Speziell in der angewendeten Version schickt er die Daten an das MariaDB-Modul. Dieses ist eine Open Source-Server-Lösung für den Einsatz von relationalen Datenbanken. Auf dem Server werden die erzeugten Daten aus der Modellrennbahn gesammelt, verarbeitet, gespeichert und z. B. über eine CSV-Schnittstelle zur Verfügung gestellt.

18.3 Analysemodule des Demonstrators

Im Anschluss an die durchgängige Erstellung eines Datenzugriffs schließen sich die Analysemodule an. Über CSV- oder Rest-API werden die Daten den Modulen zugänglich gemacht. Zu den Aufgaben der Analysemodule gehört die Datenvorverarbeitung und die Datenanalyse (siehe Kap. 4). Zunächst beginnt jedoch jedes Modul mit einem Datenschema. Dieses gibt vor, welche Daten in welcher Form benötigt werden, damit das Modul benutzt werden kann. Gleichzeitig dient das Datenschema als Matching zwischen Zugriffs- und Analysebaustein. Während der Datenvorverarbeitung werden die Rohdaten auf die Analyse vorbereitet. Neben der Vorverarbeitung gehört hierzu die Bereinigung und Anreicherung der Daten, sodass die Datenanalyse ohne Hindernisse folgt. Im Anschluss werden drei Analysemodule des Demonstrators vorgestellt. Sie umfassen jeweils sowohl die Datenvorverarbeitung als auch die Datenanalyse.

Im Mittelpunkt des Demonstrators steht das Analysemodul zur Berechnung von Rennsieger:innen (siehe Abb. 18.3). Während des Rennens haben die Fahrer:innen in der Regel keinen genauen Überblick über den Rennstand. Damit am Ende Rennsieger:innen benannt werden, ist ein Modul entstanden, welches die Rohdaten des Demonstrators verarbeitet und analysiert. In der Datenvorverarbeitung besteht das Ziel darin, den finalen Datensatz für die Analyse zu modellieren. Hierzu werden die Daten aus den zur Verfügung stehenden Datenquellen ausgewählt. Für das Analysemodul ist nur der Datensatz *Rennevents* notwendig. Zur Verbesserung der Datenqualität führt das Modul zunächst eine Bereinigung der Daten durch (vgl. West et al., 2021, S. 131 ff.). Irrelevante Attribute, werden vom Modul gelöscht. Gleichzeitig ersetzt es fehlende Daten (Missing Values), die in diesem Beispiel nicht vorliegen, z. B. mit Mittelwerten. Zuletzt wird im Modul eine Datenformatierung durchgeführt. Im Vorverarbeitungsprozess ändert es z. B. im sogenannten Parsing die Datentypen der einzelnen Attribute. In der Analyse werden dann notwendige Kennzahlen berechnet und aggregiert, sodass ein/e Rennsieger:in bestimmt werden kann.

Weitere Analysemodule lauten Renngeschwindigkeit und Prognose der Rennsieger:innen. Das Modul für die Renngeschwindigkeit kombiniert *Streckeninformationen* mit *Rennevents*, wodurch auf einer kurzen Strecke die Renngeschwindigkeit der beiden Autos zwischen zwei Sektoren berechnet wird. Das Modul Prognose der Rennsieger:innen bestimmt anhand von historischen Daten die Siegwahrscheinlich-

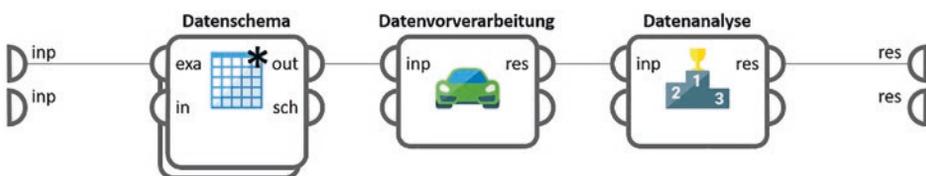


Abb. 18.3 Prozess des Analysemoduls Berechnung von Rennsieger:innen

keit der Fahrer:innen. Dieses Modul zieht vor allem die Datenquelle mit den *Fahrerinformationen* heran. Weitere Analysemodule, wie ein Modul zur Bestimmung der schnellsten Rundenzeit, sind umsetzbar.

18.4 Nutzungsmodule des Demonstrators

Damit die Analyseergebnisse von Anwender:innen interpretiert und verwendet werden können, ist der Nutzungsbaustein mit seinen Nutzungsmodulen notwendig. Die Nutzung umfasst die Durchführung von punktuellen Analysen für gezielte Fragestellungen als auch die Durchführung kontinuierlicher Auswertungen von Langzeitbeobachtungen (siehe Kap. 4). Auch im Demonstrator finden die Daten erst mit den Nutzungsmodulen eine Verwendung. Im Folgenden wird am Beispiel der zuvor dargestellten, punktuellen Analyse erläutert, wie der Nutzungsbaustein ausgestaltet werden kann (siehe Kap. 6).

Ziel des Nutzungsmoduls ist es, den/die Rennsieger:in möglichst selbsterklärend darzustellen. Hierzu sind als Input die Analyseergebnisse aus dem Analysemodul und eine Visualisierungsdatei im json-Format notwendig. Die Visualisierungsdatei wird in der Anwendung RapidMiner Studio erzeugt. Sie bildet eine visuelle Darstellung der Daten ab. Für dieses Nutzungsmodul eignet sich eine Darstellung des Rennverlaufs mittels eines Liniendiagramms. Auf der x-Achse ist das Attribut „zeitstempel_ms“ aufgetragen und auf der y-Achse das Attribut „runde“. Jede „startnummer“ erhält eine eigene Linie, sodass der/die führende/r Fahrer:in anhand der höheren Kurve identifiziert wird (siehe Abb. 18.4).

Das Anwendungsbeispiel des Demonstrators erklärt anschaulich, wie eine Datenanalyse entlang der Prozesskette der industriellen Datenanalyse (siehe Kap. 4) durchgeführt wird. Um Anwender:innen zu zeigen, welches Potenzial die AKKORD-Plattform

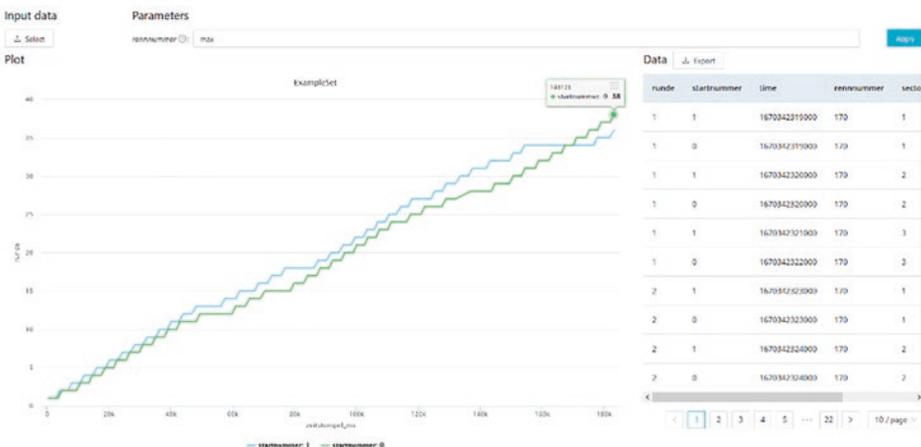


Abb. 18.4 AI-Toolbox Nutzungsmodul Auswertung Rennsieger:in

bietet, lässt sich das gezeigte Beispiel auf Anwendungen in der Produktion übertragen. Ein vergleichbares Beispiel wäre eine Analyse der Ausbringung mehrerer Maschinen. Die Maschinen erzeugen maschinenspezifische Zeitstempel, wenn sie Produkte („runde“) fertigstellen. Durch dieselbe Datenvorverarbeitung und -analyse lassen sich die Zeitstempel so auswerten, dass am Ende eine Visualisierung für die vergleichende Auswertung mehrerer Maschinen hinsichtlich ihrer Ausbringung entsteht. Das Liniendiagramm bildet dann den Produktionsverlauf der Maschinen ab. Maschinenplaner:innen sehen dann schnell, welche Maschine wann mehr Produkte produziert hat.

18.5 Fazit

Durch den AKKORD-Autorenn-Demonstrator wird anschaulich gezeigt, wie eine Datenanalyse aufgebaut wird. Entlang der Prozesskette der Datenanalyse zeigt der Demonstrator die Umsetzung diverser Module, sodass der modulare Charakter der AKKORD-Plattform deutlich wird. Dank der Modularität ermöglicht die Plattform ebenfalls eine Wiederverwendbarkeit der Datenanalysen. Dies zeigt auch die Übertragbarkeit der Ergebnisse auf weitere Anwendungsszenarien. Insgesamt repräsentiert der Demonstrator die AKKORD-Forschungsergebnisse und validiert diese hinsichtlich Anwendbarkeit, Modularität und Wiederverwendbarkeit.

Literatur

- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>
- Wöstmann, R., Barthelmey, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. In T. Schüppstuhl, K. Tracht, & J. Roßmann (Hrsg.), *Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Realisierung eines I4.0- Lerndemonstrators im Rahmen des AKKORD-Projekts

19

Ein hybrides Konzept für die berufsschulische und innerbetriebliche Ausbildung im Bereich Industrie 4.0

Christopher Klupak  und Tina Hardt 

19.1 Motivation

Die **Digitalisierung** von Produktionsprozessen nimmt zu, gleichzeitig existiert noch wenig didaktisches Material für die Aus- und Weiterbildung von Personal. Zudem sind derzeit erhältliche Demonstratoren eher sperrig, kostenintensiv und werden oft ohne didaktisches Konzept und Aufgaben für den Einsatz in der Ausbildung angeboten (Oks et al., 2021, S. 119 ff.). Deshalb haben sich die Real-Time-Systems GmbH, vormals Arendar IT-Security GmbH, und der Lehrstuhl für Berufspädagogik in den technischen Fächern (BtF) der Universität Hamburg innerhalb des Projekts AKKORD zusammengetan (Kap. 2), um prototypisch eine kompakte und erschwingliche Lernstation zur Veranschaulichung von industrieller Datenerfassung und zugehörigen Sicherheitskonzepten umzusetzen (siehe Kap. 14). Die Arbeiten stammen aus dem Umfeld des Konsortialprojekts AKKORD (siehe Kap. 1) und leisten einen Beitrag zur Realisierung der Prozesskette der industriellen Datenanalyse (siehe Kap. 4). Die abstrakten Industrie 4.0-Paradigmen werden durch den in diesem Kapitel beschriebenen Demonstrator als reale Lösungen erlern- und erlebbar. Der Demonstrator reiht sich in eine Gruppe von Demonstratoren im Forschungsprojekt ein, die u. a. auch das modulare Datenbackend (siehe Kap. 5 und 17) sowie die modularen Datenanalysen (siehe Kap. 6 und 18) beschreiben.

C. Klupak (✉)
Universität Hamburg, Hamburg, Deutschland
E-Mail: christopher.andre.klupak@uni-hamburg.de

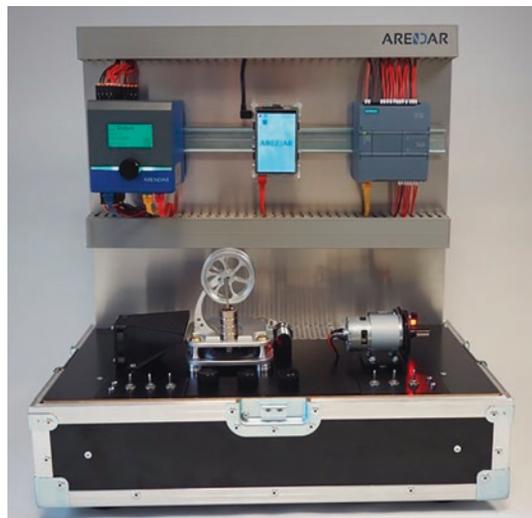
T. Hardt
Real-Time-Systems GmbH, Wittlich, Deutschland
E-Mail: tina.hardt@real-time-systems.com

19.2 Hardware-technischer Aufbau und Funktion der Elemente der AKKORD-Lernstation

Auf der Wand der Lernstation (siehe Abb. 19.1) sind von links nach rechts das cybersichere Edge-Gateway ARENDAR, ein Raspberry Pi 2 und eine SPS montiert. Das Gateway verfügt über 3 LAN-Schnittstellen und ist über LAN 1 mit der Office IT im Boden des Koffers, über LAN 2 mit der SPS und über LAN 3 mit dem Raspberry Pi verbunden. Die Office-IT kann mit dem Internet verbunden werden, sodass auch das Gateway aus dem Netz erreichbar ist. Der Pi ist nur zur Demonstration des sicheren Fernzugriffs über das LAN 3 des Gateways nach 2-Faktor-Authentifizierung vorhanden und hat darüber hinaus keine weitere Funktion.

Auf dem Kofferboden befindet sich mittig ein Stirlingmotor mit einer Kühlung links daneben. Dieser steht für die erste technische Revolution und wird für Regelszenarien in Zusammenhang mit der Kühlung und für Produktionssimulationen eingesetzt. Im Vordergrund der Kühlung ist ein Schaltboard mit drei Kippschaltern und 2 LED-Leuchten, die jeweils direktverdrahtet mit den digitalen Eingängen des Gateways sind und zur Anzeige von Prozesszuständen bzw. zum Anschalten von Demonstrator-komponenten dienen. Unmittelbar rechts vom Stirlingmotor sind eine Lichtschranke zur Drehzahlmessung sowie zwei Temperatursensoren verortet, die sowohl direkt mit dem Gateway als auch mit der SPS verbunden sind. Links befindet sich ein Elektromotor, ebenfalls mit einer rechts daneben angebrachten Lichtschranke. Davor befindet sich wieder ein Schaltboard mit gleichem Aufbau wie in der rechten Hälfte. Dessen Elemente sind mit der SPS verbunden. Zwischen den beiden Schaltboards vorne in der Mitte befindet sich das Potenziometer-Board mit 3 Potis, die mit den analogen Eingängen des

Abb. 19.1 Darstellung der AKKORD-Lernstation



Gateways verbunden sind. Diese dienen zur Simulation von Werten in den Prozessen. Unsichtbar in dieser Ansicht sind zwei USB-Schnittstellen des Servers aus dem Office-Bereich hinter dem Elektromotor. Durch die Direktverdrahtung von Elementen der Station mit dem Gateway werden Retrofitting-Szenarien simuliert (vgl. Wöstmann et al., 2019, S. 94 ff.). Die Erfassung von Daten der Elemente über die SPS dient zur Darstellung von moderneren Szenarien, bei denen auf eine SPS zugegriffen werden kann.

Im Boden der Lernstation sind verschiedene Netzteile für die verschiedenen Komponenten, ein gemanagter LAN-Switch (zwei VLANs), ein Raspberry Pi 4 mit WLAN, der als Server dient, und ein LTE Stick untergebracht.

Die Station bietet zwei Möglichkeiten der Visualisierung erfasster Daten: über eine Weboberfläche des im Gateway integrierten Webservers und durch eine externe Visualisierung über die Open-Source-Software Node-RED. Mit der Visualisierung lassen sich praxisnahe Betriebsfälle aus diversen industriellen Bereichen zu unterschiedlichen Lernfeldern und Lernniveaus konstruieren (siehe Hardt & Klupak, 2022, S. 271 ff.). Gleichzeitig ist es möglich, durch die Konfiguration des Gateways, Parameter anzupassen oder Informationen (z. B. Durchschnittswerte, Min-Max-Bestimmung, etc.) durch Datenvorverarbeitung im ARENDAR zu erzeugen (Edge-Computing).

19.3 Entwicklung des hybriden Lernkonzeptes

Im ersten Schritt wurden zunächst die möglichen technischen Einsatzfelder der vorhandenen ARENDAR-Lernplattform untersucht, um ein **hybrides Lernkonzept** daraus ableiten zu können. In diesem Zusammenhang bezeichnet *hybrides Lernen* eine Lehr- und Lernmethode, die sowohl traditionelle Präsenzveranstaltungen als auch digitale Elemente der Fernlehre kombiniert (Kerres, 2002, S. 1 ff.). Im weiteren Verlauf wurde eine Zielgruppenanalyse durchgeführt, bei der die duale technische Erstausbildung favorisiert wurde. Unter diesem Oberbegriff werden zahlreiche Ausbildungsberufe subsummiert, die einen inhaltlichen Bezug zum Schulungsinhalt aufweisen. Bei der dualen Berufsausbildung findet die Vermittlung von Lerninhalten an zwei Orten, im Ausbildungsbetrieb sowie in der berufsbildenden Schule, statt.

In diesem Zusammenhang wurden im nächsten Schritt aus den Rahmenlehrplänen der Kultusminister Konferenz zur dualen Berufsausbildung, welche auf der Internetseite der Kultusministerkonferenz zur Verfügung gestellt werden, vorab die technischen Ausbildungsberufe mit den passenden Lernfeldern erfasst und in einer Gesamtübersicht zusammengetragen (siehe KMK, 2023).

Unter den insgesamt 306 Rahmenlehrplänen wurden 26 technische Ausbildungsberufe mit möglichen Einsatzfeldern der Lernplattform identifiziert. Im Rahmen der Analyse wurden zu den 26 Ausbildungsberufen die spezifischen Lernfelder, das Ausbildungsjahr und deren Zeitrichtwerte erfasst. So hat beispielsweise der Ausbildungsberuf Anlagenmechaniker:in im adressierten Lernfeld 10 einen Zeitrichtwert von 80 h im dritten Ausbildungsjahr. Zusätzlich wurden unter dem Reiter ARENDAR die jeweiligen

Lernfeld-Inhalte, die durch die Lernplattform abgedeckt werden, aufgelistet. Dies wird in Abb. 19.2 exemplarisch dargestellt.

In der Gesamtübersicht kann somit die inhaltliche Passung zwischen den identifizierten Ausbildungsberufen und den möglichen Schulungsinhalten des Demonstrators abgeglichen werden. Als Leitfaden kann die Gesamtübersicht die Lehrkräfte später beim Einsatz des hybriden Lernplattform-Konzeptes im Unterricht unterstützen, und eine Orientierungshilfe für die konkrete inhaltliche Anwendung bieten.

Im Anschluss erfolgte die didaktisch-/methodische Vorüberlegung, die sich am Konzept „Vom Lernfeld zur Lernsituation“ orientiert. Dabei werden in der Lernsituation die Lerninhalte aus dem Lernfeld konkretisiert, in der die Auszubildenden in möglichst realitätsnahe Handlungssituationen hineinversetzt werden. Dazu werden die jeweiligen Lerninhalte in einem simulierten, didaktisch aufbereiteten, beruflichen und/oder lebensweltlichen Kontext vermittelt, was eine motivierende Wirkung bei den Lernenden mit sich bringt. Als Szenario wurde ein fiktives Sägewerk für die Lernenden gewählt, in dem sich alte und neue Produktionsanlagen befinden, die den einzelnen Stufen der industriellen Revolution zugeordnet werden können.

In diesem Zusammenhang wird für die Lernenden ein sinnstiftender Bezug zwischen dem fiktiven Unternehmen und der AKKORD-Lernstation geschaffen (siehe Kap. 7). Bei der späteren Durchführung als Industrie 4.0- Projekttag an den berufsbildenden Schulen schlüpfen die Auszubildenden in die fiktive Rolle eines/-r Experten/-in, die von ihrem Onkel als Inhaber des „Sägewerks Fichtenast“ um Hilfe gebeten werden. Er möchte ein weiteres Sägewerk mit dem Namen „Tannental“ aufkaufen und in seine bestehende Firmenstruktur integrieren. Das erworbene Sägewerk soll unter den Aspekten von Industrie 4.0 modernisiert werden, sodass eine standortübergreifende Koordination möglich ist. Durch die Aufgabenstellung werden die Lernenden so in eine berufliche und lebensweltliche Situation versetzt.

Daraus wurde ein zweiteiliges Lernkonzept abgeleitet, welches im ersten Abschnitt die Lernsituation und die ersten drei Lernaufgaben bildet. In einem für die Schüler

Anlagenmechaniker*in			
Link zum RLP: Anlagenmechaniker/Anlagenmechanikerin [pdf, 275 KB]			
Lernfeld	Inhalte	ARENDAR	Zeitrichtwert
Lernfeld 10	Einbinden von Komponenten der Steuerungs- und Regelungstechnik	<ul style="list-style-type: none"> ○ SPS-Programmierung ○ Steuern ○ Regeln ○ Messen 	80h

Abb. 19.2 Auszug aus der Zielgruppenanalyse des Ausbildungsberufs Anlagenmechaniker:in

greifbaren Setting werden sie zunächst in die Grundlagen eingearbeitet, lernen etwas über die Industrierevolutionen und über die Lernstation. Die einzelnen Lerninhalte wurden dabei als Präsentationsvorlage entwickelt und in einer digitalen Work&Learn-Plattform als rollenbezogener Kurs „Berufsschule“ hinterlegt (Abb. 19.3).

Den zweiten Teil bilden die Lernaufgaben vier bis sieben. Diese vier Aufgaben sind als umfangreiche, von den Schülern möglichst eigenständig zu erarbeitende kleine Projekte entworfen. Das Konzept ist angelehnt an Stationen, in denen ein spezielles Thema im Team vertieft bearbeitet wird. Jede der Stationen greift ein zentrales Thema der AKKORD-Lernstation auf. Dabei kann die Klasse der Lernenden in vier Gruppen unterteilt werden, in der jede Gruppe an einer Lernstation arbeitet. Abgeschlossen wird dieses Lernkonzept mit einer Präsentationsphase. Hier sollen die einzelnen Gruppen ihre Lernstation sowie deren Ergebnisse vorstellen. Zusätzlich erhalten die Schüler, die sich nicht mit der vorgestellten Station auseinandergesetzt haben, eine Idee, was mit dem Gateway ARENDAR alles realisierbar ist.

19.4 Validierung und Optimierung

Das fertige hybride Lernkonzept wurde an der *David Röntgen Berufsschule* mit zwei Mechatroniker Klassen als Projekttag getestet. Dabei zeigten sich bei der Durchführung noch kleine Schwachstellen (z. B. Zugangsprobleme auf die Work&Learn-Plattform, Multiuserfähigkeit AKKORD-Lernstation). Das hybride Lernkonzept kam in der Pilotierung in der Gesamtbetrachtung bei den Lernenden und Lehrkräften gut an und sollte durch nachgelagerte Optimierungsmaßnahmen weiter an verschiedenen Berufsschulen getestet und verfeinert werden.

The screenshot displays the user interface of the AKKORD work&learn platform. At the top, the navigation bar includes the logo 'akkord work&learn' and menu items: AKTUELLES, KURSE, MAGAZIN, WIKI, SERVICES, COMMUNITY, EMPFEHLUNGEN, and MEINE SEITE. A search icon and other utility icons are also present. The main content area is titled '[SL1.3.1] Lernaufgabe 3' and contains the text: 'Hier findest du die Lernaufgabe 3.' Below this, a task card titled 'Lernaufgabe 3 – Auswahl eines geeigneten Geräts' is shown. The task description reads: 'Nachdem du das geeignete Gerät ausgewählt hast erhältst du von deinem Onkel ein Infoblatt über das Gerät. Es handelt sich um ein Gateway, den Arendar, der Firma Arendar IT-Security GmbH.' The task instruction is: 'Aufgabe: Beschreibe die Anschluss- und Einsatzmöglichkeiten des Arendars anhand der Abbildung „Kommunikationsmöglichkeiten des Arendars“. Mache dir hierfür handschriftliche Notizen.' An image of the Arendar device is shown next to an information icon. A red hint box states: 'Hinweis: Das Infoblatt über den Arendar findest du unter „Infoblatt_Arendar“ auf der rechten Seite deines Desktop!' On the right side, a 'DOWNLOADS' box lists two PDF files: 'Lernaufgabe_2_0.pdf' and 'Infoblatt_Arendar.pdf'. At the bottom of the task card, a progress indicator shows '9 / 12'.

Abb. 19.3 Exemplarische Darstellung einer Lernaufgabe zur Auswahl eines geeigneten Geräts

19.5 Fazit

Die Digitalisierung von Produktionsprozessen nimmt zu, gleichzeitig existiert noch recht wenig didaktisches Material für die Aus- und Weiterbildung von Personal. Derzeit erhältliche Demonstratoren sind eher sperrig, kostenintensiv und werden oftmals ohne didaktisches Konzept und Aufgaben für den Einsatz in der Ausbildung angeboten. Aus dieser Notwendigkeit entstand die erarbeitete Lösung, bestehend aus der in Abschn. 19.2 vorgestellten Hardware und digitalen Lernmodulen. Das vorgestellte Schulungskonzept ist nur eines von vielen Möglichkeiten, zur Nutzung der AKKORD-Lernstation. Hier kann, geführt durch eine von den Schülern greifbare Lernsituation, das Konzept der Lernstation als Ganzes erlebt werden. Alle angeführten Inhalte wurden in eine digitalen Work & Learn- Lernplattform integriert, sodass es für die Schüler aber auch für die Lehrpersonen möglich ist, damit zu arbeiten. Die Ergebnisse der ersten Pilotierung waren bis auf kleinere Probleme durchweg positiv.

Aktuell wird daran gearbeitet, die Stationen mit einem Multi-User-Interface auszustatten, sodass mehrere Gruppen gleichzeitig an einer Station arbeiten können, ohne einander zu beeinflussen. Dazu werden Simulationen kreiert, die dann über die ARENDAR-Weboberfläche zur Verfügung stehen sollen.

Literatur

- Hardt T., & Klupak, C. (2022). Lernstation für Datenerfassung & IoT-Applikationen. Unter Berücksichtigung der Sicherheit bei Datenerfassung und -kommunikation. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 5, 271–274. <https://doi.org/10.1515/zwf-2022-1066>.
- Kerres, M. (2002). Online – und Präsenzelemente in hybriden Lernarrangements kombinieren. In A. Hohenstein & K. Wilbers (Hrsg.), *Handbuch E-Learning* (S. 1–15). Fachverlag Deutscher Wirtschaftsdienst.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, KMK. (2023). Übersicht der Rahmenlehrpläne für die duale Berufsausbildung, Online-Adresse: <https://www.kmk.org/themen/berufliche-schulen/duale-berufsausbildung/downloadbereich-rahmenlehrplaene.html>. Zugegriffen: 28. Febr. 2023.
- Oks, S. J., Jalowski, M., Zansinger, N., & Möslein, K. M. (2021). Die Rolle von Industrie 4.0-Demonstratoren in der digitalen Transformation. Eine Standpunktbestimmung am Portable Industrial Demonstrator for Cyber-Physical Systems. In K. Wilbers & L. Windelband (Hrsg.), *Lernfabriken an beruflichen Schulen. Gewerblich-technische und kaufmännische Perspektiven* (S. 119–158). Leibniz-Institut für Bildungsforschung und Bildungsinformation.
- Wöstmann, R., Barthelme, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. *Thorsten Schüppstuhl, Kirsten Tracht und Jürgen Roßmann (Hrsg.): Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Industrielle Datenanalyse im Einklang mit Mensch, Technik und Organisation

20

Ein Ausblick auf die Entwicklungen der Industriellen Datenanalysen

Jochen Deuse , Ralf Klinkenberg , Nikolai West  und Marius Syberg 

20.1 Ausgangssituation

Durch die zunehmende Verbreitung von Digitalisierungslösungen und Künstlicher Intelligenz sowie durch den steigenden Wert von Daten für Unternehmen ist der Einsatz industrieller Datenanalysen unumgänglich geworden, um im Markt zu bestehen (Han & Trimi, 2022). Die Anwendungsmöglichkeiten sind dabei ebenso divers wie funktionspezifisch: Beispielsweise können im Qualitätsmanagement durch eine Anomalieerkennung Produktionsfehler in großen Datenmengen identifiziert werden (Schlegl et al., 2022, S. 1 ff.; West et al., 2021c, S. 762 ff.). Mit prädiktiver Instandhaltung können Anlagenausfälle vorausgesagt werden, bevor sie auftreten (Wöstmann et al., 2019, S. 94 ff.). Durch selbstlernende Agentensysteme kann die Produktionsplanung durch zielgerichtete Empfehlungen unterstützt werden (West et al., 2021b, S. 347 ff.). Durch die Vorhersage von dynamischen Engpässen kann die Produktionssteuerung künftige Entwicklungen antizipieren und proaktiv darauf reagieren (West et al., 2022a, S. 612 ff., b, S. 3 ff.).

J. Deuse

Institut für Produktionssysteme, Technische Universität Dortmund, Dortmund, Deutschland

E-Mail: jochen.deuse@ips.tu-dortmund.de

R. Klinkenberg

Rapid Miner GmbH, Dortmund, Deutschland

E-Mail: rklinkenberg@rapidminer.com, rklinkenberg@altair.com

N. West (✉) · M. Syberg

Technische Universität Dortmund, Dortmund, Deutschland

E-Mail: nikolai.west@ips.tu-dortmund.de

M. Syberg

E-Mail: marius.syberg@ips.tu-dortmund.de

© Der/die Autor(en) 2024

J. Deuse et al. (Hrsg.), *Industrielle Datenanalyse*,

https://doi.org/10.1007/978-3-658-42779-5_20

209

Gerade kleine und mittlere Unternehmen (KMU) unterliegen in diesem hochgradig spezialisierten Umfeld einem erheblichen Innovationsdruck, um sich in globalisierten Märkten und gegen größere Wettbewerber behaupten zu können. Durch die fortschreitende digitale Transformation stehen viele Daten zur Verfügung, die bessere und schnellere Entscheidungen ermöglichen. Unternehmen erkennen, dass die Nutzung dieser Daten bedeutend für sie ist (Safar, 2023, S. 254). So sind sie in der Lage, Produkte und Prozesse nachhaltig zu optimieren.

Im Rahmen des AKKORD Projekts konnte ein Referenzbaukasten entwickelt werden, der spezielle Lösungen insbesondere für KMU bietet. Dabei wurden Möglichkeiten zur Kollaboration und Kompetenzentwicklung, für eine einfache, integrierte Datenanalyse sowie zur Schaffung einer vollständigen und vernetzten Datenbasis entwickelt. Zusätzlich konnten durch die breit gefächerten industriellen Anwendungsfälle Handlungsempfehlungen auf organisatorischer Ebene abgeleitet werden. So wurden Anforderungen menschlicher, technischer und organisatorischer Art beachtet (Syberg et al., 2023, S. 64 f.). Insbesondere für KMUs, die oft über begrenzte Ressourcen unterschiedlicher Ausprägung und begrenzte Erfahrung in der Datenanalyse verfügen, hat sich jedoch gezeigt, dass neben vielversprechenden Potenzialen bei der Implementierung neue Herausforderungen entstehen. Die Betrachtungsweise der Einflussfaktoren für den erfolgreichen, nachhaltigen Einsatz in Unternehmen verschiebt sich.

In diesem Kapitel diskutieren wir diese Herausforderungen und Hindernisse bei der Implementierung von industrieller Datenanalyse in KMUs und beleuchten die Potenziale für die Zukunft in diesem Bereich.

20.2 Schlussfolgerungen

Folgt man in der Implementierung den Gestaltungsprinzipien Mensch, Technik und Organisation (vgl. Pardo Escher, 1997), ist im Rahmen der Einführung industrieller Datenanalyse erkennbar, dass insbesondere die menschliche und technische Komponente in vielen Lösungen adressiert wird.

Mensch. Im menschlichen Bereich steht die Kompetenzentwicklung zur Qualifizierung der an der Implementierung beteiligten Menschen im Vordergrund, die nicht entsprechend ausgebildet sind. Dazu existieren zahlreiche Open-Source-Plattformen wie Udemy oder Coursera. In AKKORD wurde mit der Work&Learn-Plattform eine maßgeschneiderte Lösung für den industriellen Einsatz entwickelt, die sich in diese Reihe eingliedert. Ihr berufsbegleitender und integrierter Ansatz von Kompetenzentwicklung ist ein zielführender Ansatz für die Entwicklung beteiligter Mitarbeiter in die entscheidenden Rollen für den Einsatz industrieller Datenanalysen.

Technik. Technisch gibt es unterschiedliche Strömungen und Ansätze, da die Herausforderungen in KMU unterschiedlich sind. In den Anwendungsfällen des Forschungsprojekts galt es in einigen Anwendungsfällen, überhaupt erst entscheidende Daten zu generieren, in anderen Fällen ist die Verfügbarkeit entsprechender IT-Tools zur Analyse

nicht gegeben. Darüber hinaus müssen Datensicherheit und -eigentümerschaft geklärt und sichergestellt werden. Auch hier konnte in AKKORD ein Resultat entwickelt werden (siehe Kap. 5).

An die Schnittstelle **Mensch-Technik** ordnen wir technische Lösungen ein, die den Datenanalyseprozess selbst vereinfachen und so zugänglicher machen. Dies verringert die Einstiegshürde und erfordert damit eine weniger intensive Qualifizierung. Marktreife Lösungen wie RapidMiner, die Advanced Analytics mit einem Low Code Ansatz ermöglichen, sind verfügbar. Die entwickelte AI-Toolbox ist ein weiterer Ansatz an genau dieser Stelle, die aber durch die browserbasierte Umsetzung mit den auf industrielle Anwendungsfälle zugeschnittenen Analysemodulen speziell für KMUs den Einstieg in den Bereich schafft (siehe Kap. 6).

Organisation. Die beteiligten Unternehmen befinden sich nicht nur im Rahmen des Projekts in einer Phase der organisatorischen Transformation im Hinblick auf eine zukünftig digitale Ausrichtung. Die Implementierung des Referenzbaukastens und seiner Teillösungen zeigte in allen Unternehmen, dass organisatorische Maßnahmen immer begleitend einhergehen. Diese Maßnahmen sind facettenreich. Dabei sind eine klare Zielsetzung und Priorisierung sowie ein Verständnis der zu vergebenden Rollen. Als strukturgebendes Werkzeug ist ein ganzheitliches *Change Management* als vielversprechend anzusehen.

Die Gestaltung einer solchen Digitalisierungsimplementierung wie der Einführung von industrieller Datenanalyse ist mit der Betrachtung von **M-T-O** als Gestaltungsprinzipien nicht abgeschlossen. So erscheint durch den hohen Aufwand im Bereich der Datenakquise und -verarbeitung der Bereich *Daten* als eigener Betrachtungsgegenstand in der Gestaltung sinnvoll. In AKKORD stehen mit Plattformen zur Kompetenzentwicklung und Datenanalyse *Systeme* im Vordergrund, die nicht zuletzt durch die benötigten Schnittstellen zu bestehenden Systemen der Unternehmen ein eigener Betrachtungsgegenstand sein müssen. Ergänzend ist es der *Analyst* selbst mit seinen vorhandenen Fähigkeiten und der Rolle, die einzunehmen ist, entscheiden für das Ergebnis eines Prozesses und muss in der Lage sein, industriellen Kontext und analytische Komplexität miteinander zu verbinden.

20.2.1 Herausforderungen

Im Projekt haben sich anwenderübergreifend Herausforderungen und Hindernisse in verschiedenen Bereichen gezeigt, die für den nachhaltigen Einsatz von industrieller Datenanalyse anzugehen sind.

Organisation. Allein funktionierende Plattformen mit Datenanalyse als Kern und guten Ideen reichen nicht, um KMU für den flächendeckenden Einsatz von Datenanalysen zu befähigen. Der Roll Out von Digitalisierungslösungen – und die industrielle Datenanalyse im speziellen – erfordert einen strategischen Ansatz. Da es sich im Rahmen der Implementierung unabdinglich um komplexe Vorgänge handelt, erfordern

sie Personaleinsatz und Zeit. Im Projekt hat sich, nicht zuletzt im Rahmen einer Umfrage (siehe Kap. 2 und Kap. 16, gezeigt, dass darüber hinaus Managementunterstützung unabdinglich ist. All diese Punkte fallen in den Bereich der Organisation. Dieser Aspekt ist für die nachhaltige Implementierung der Datenanalyse in Prozesse entscheidend. Es lässt sich beobachten, dass viele Unternehmen bereits einzelne Tools einsetzen und im Rahmen der Implementierung Mitarbeiter insofern qualifiziert haben, als dass sie vereinzelte Pilotanwendungen umgesetzt haben. Es fehlt jedoch die angesprochene organisatorische Grundlage für eine flächendeckende Umsetzung. Aufbauend auf Managementverständnis und -unterstützung in der strategischen Herangehensweise ist ein organisatorischer Change-Prozess der Schlüssel.

Zielsetzung. Dabei ist auch auffällig, dass in vielen Unternehmen prädiktive und präskriptive Ansätze zwar langfristiges Ziel sind, viele akute Probleme aber durch eine klare, problembeschreibende und –erklärende Lösung erfolgreich vereinfacht werden können. Gleichzeitig helfen Lösungen dieser Art, beispielsweise Visualisierungen von Ist-Daten sowie Erklärungen auf den ersten Blick auf Basis von Korrelationsanalysen oder Klassifikationen in der Akzeptanz von Entscheidungen von Mitarbeitern durch Kunden oder Management. Die Transparenz von Prozessen wird entscheidend erhöht.

Return on Invest für technische Infrastruktur. Darüber hinaus ist auffällig, dass der Weg zu prädiktiven und präskriptiven Anwendungsfällen in vielen KMU weit ist. Die Digitalisierung an sich bietet große Potenziale zur Datenaufnahme und -verarbeitung, ein häufiges Hindernis ist es aber, diese nicht heben zu können. Ein häufiges Hindernis ist der Mangel an geeigneten Datenquellen mit entsprechender Datenqualität durch die Verwendung alter Anlagen oder von individuellen Systemen mit fehlenden Schnittstellen (West et al., 2021a, S. 131 ff.). Die in AKKORD erarbeitete Lösung bedient dieses Hindernis, es ist aber an den KMU, technisch für die Anbindung entscheidender Datenquellen zu sorgen. Hier schließt sich an, dass viele Unternehmen nicht zur Investition in die technische Verfügbarkeit bereit sind, da sie den Mehrwert nur langfristig oder gar nicht sehen. Eine gute Datenverfügbarkeit ist heute wertschöpfend. Dieser Gedanke muss im Rahmen der organisatorischen und kulturellen Veränderung in Unternehmen Einzug halten.

20.2.2 Potenziale

Über die genannten Hindernisse hinaus birgt die industrielle Datenanalyse zahlreiche Potenziale für KMU.

- **Datenbasierte Entscheidungsfindung als Standard.** Datenbasierte Entscheidungen sind elementar, wenn man im kompetitiven und zugleich volatilen Umfeld bestehen möchte. Das Risiko von Fehlentscheidungen durch den Rückgriff auf Erfahrungswerte sinkt. Der Einsatz von industriellen Datenanalysen und Maschinellern Lernen in diesen Prozessen ermöglicht es, weitaus mehr Informationen in solche

Entscheidungen mit einzubeziehen. Außerdem wirkt dies indirekt dem demographischen Wandel und komplexem Wissensmanagement entgegen, welches insbesondere in KMU zu Problemen führt (Tachkov & Mertens, 2016, S. 29 f.).

- **Schutz vor Wissensverlust.** Je mehr Entscheidungen datenbasiert getroffen werden können, desto weniger Gefahr besteht, aufgrund des Verlustes von Erfahrungswissens in einen Wettbewerbsnachteil zu gelangen. Hier setzt auch die Work&Learn-Plattform an.
- **Anwender als Akteure im Markt.** Durch gezielte Kompetenzentwicklung der Mitarbeiter im Bereich der industriellen Datenanalyse ist man in einem zukunfts-trächtigen Feld führend und auch in der Lage, gestaltend im Markt zu agieren. Neben dem verbesserten Überblick über das eigene Unternehmen hilft die Fähigkeit zum zielgerichteten Einsatz von Datenanalysewerkzeugen Trends und Verhaltensweisen im Markt, von Kunden und von Wettbewerbern früh erkennen zu können.
- **Geschäftsmodellentwicklung.** Eigene Geschäftsprozesse werden optimiert. Dabei spielt auch die Weiterentwicklung des eigenen Geschäftsmodells selbst eine Rolle. Smarte Produkte und Services können so in den Fokus rücken und auch für KMU eine näherliegende Option werden. Weiterentwicklungen in der industriellen Datenanalyse bringen Unternehmen so in die Lage, datenbasiert einen neuen Mehrwert für den Kunden zu generieren (Machchhar et al., 2022, S. 1).
- **Schneller Hochlauf.** Ein weiteres Potenzial besteht für KMU explizit im Einsatz der entwickelten modularen Datenanalyseplattform des Referenzbaukastens, der AI Toolbox. So ist eine Schlussfolgerung des Projekts, dass es vielen KMU primär insbesondere beim Einstieg nicht um das beste Analyseergebnis geht, sondern um eine gute Lösung, die aber schnell anwendbar und adaptierbar ist. Der modulare Ansatz zur Datenanalyse ist hier ein passender Einstieg. Anwenderunternehmen sind in der Lage, aufwandsarm Analysen durchzuführen und zu experimentieren. Gleichzeitig schafft die Work&Learn-Plattform ein Grundverständnis für datenanalytische Zusammenhänge. Wenn gute Ergebnisse schnell erzielt werden können, wird Entscheidern der Mehrwert deutlich. Aufbauend auf diesen Entwicklungen können Unternehmen dann spezifischere und komplexere Lösungen erarbeiten.

20.3 Fazit

Auch die Zukunft der industriellen Produktion wird von innovativen Technologien wie dem Internet der Dinge und Künstlicher Intelligenz geprägt sein. Unternehmen werden zukünftig weitreichende Möglichkeiten haben, Daten zu sammeln, sie auszuwerten und für sich nutzen, um bessere Produkte zu entwickeln, die sie effizienter produzieren. Die unterschiedlichen Anwendungsfälle im AKKORD-Projekt zeigen, dass die Anwenderunternehmen diese Entwicklung sehen und sich entsprechend rüsten wollen. Dabei spielen Sensor- und Produktionsdaten genauso eine Rolle wie Kundenbewertungen sowie Markt- und Vertriebsdaten, mit denen sich die Unternehmen beschäftigen. Sie

wollen eine Infrastruktur schaffen, mit deren Hilfe diese Pipeline von der Datenaufnahme mit der Transformation in Wissen bis hin zur datenbasierten Entscheidung Standard ist. Entscheidend sind hier immer Datenzugriff, Datenanalyse und -anwendung mitsamt entsprechender Kompetenzentwicklung in Kombination, die zu einem nachhaltigen Erfolg führen (Deuse et al., 2021, S. 3).

Industrielle Datenanalyse durchgängig als wichtiges Werkzeug im digitalisierten Unternehmen zu integrieren, ist unter dieser Prämisse vielmehr erreichbarer Zielzustand als Vision. Es ist außerdem erkennbar, dass viele Teillösungen oder Ansätze auf dem Markt verfügbar oder gar im industriellen Einsatz sind, es durch ihren Einsatz ohne Berücksichtigung der anderen Aspekte aber häufig bei Pilotumsetzungen bleibt. Wenn Unternehmen die Aspekte *Daten*, *Analyst*, *System* und *Organisation* entsprechend in der Gestaltung berücksichtigen, ist die nachhaltige Integration industrieller Datenanalyse in Unternehmen gesichert. AKKORD hat mit dem Referenzbaukasten eine Lösung geliefert, die den Einstieg flächendeckende und wertschöpfende Nutzung von industriellen Datenanalysen ermöglicht.

Literatur

- Deuse, J., West, N., & Syberg, M. (2021). Rediscovering scientific management. The evolution from industrial engineering to industrial data science. *International Journal of Production Management and Engineering*, 10(1), 1–12. <https://doi.org/10.4995/ijpme.2022.16617>.
- Han, H., & Trimi, S. (2022). Towards a data science platform for improving SME collaboration through Industry 4.0 technologies. *Technological Forecasting and Social Change*, 174, 121242. <https://doi.org/10.1016/j.techfore.2021.121242>.
- Machchhar, R. J., Toller, C. N. K., Bertoni, A., & Bertoni, M. (2022). Datadriven value creation in smart product-service system design: State-of-the-art and research directions. *Computers in Industry*, 137. <https://doi.org/10.1016/j.compind.2022.103606>.
- Pardo Escher, O. (1997). Unternehmen arbeitspsychologisch bewerten. Ein Mehr-Ebenen-Ansatz unter besonderer Berücksichtigung von Mensch, Technik, Organisation. Hrsg. Oliver Strohm. Vdf, Hochschulverl. an der ETH Zürich (Mensch, Technik, Organisation, 10).
- Safar, M. (2023). KI und IoT – Treiber der Digitalen Transformation. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 118(4), S. 254–257. <https://doi.org/10.1515/zwf-2023-1043>.
- Schlegl, T., Schlegl, S., Tomaselli, D., West, N., & Deuse, J. (2022). Adaptive similarity search for the retrieval of rare events from large time series databases. *Advanced Engineering Informatics*, 52, 101629. <https://doi.org/10.1016/j.aei.2022.101629>.
- Syberg, M., West, N., Schwenken, J., Adams, R., & Deuse, J. (2023). Requirements for the Development of a Collaboration Platform for Competency-Based Collaboration in Industrial Data Science Projects. In F. P. García Márquez, I. Segovia Ramírez, P. J. Bernalte Sánchez, & A. Muñoz del Río (Hrsg.), *IoT and Data Science in Engineering Management. CIO 2022. Lecture Notes on Data Engineering and Communications Technologies*, (Bd. 160, S. 64–69). Springer. https://doi.org/10.1007/978-3-031-27915-7_12.
- Tachkov, P., & Mertens, M. (2016). Implizites Mitarbeiterwissen in KMU: Den Schatz stiller Wissensressourcen heben. *Wissensmanagement*, 2.
- West, N., Gries, J., Brockmeier, C., Göbel, J. C., & Deuse, J. (2021a). Towards integrated Data Analysis Quality. Criteria for the application of Industrial Data Science. *IEEE International*

- Conference on Information Reuse and Integration for Data Science (IRI)*, 22(1), 131–138. <https://doi.org/10.1109/IRI51335.2021.00024>.
- West, N., Hoffmann, F., Schulte, L., Moreno, V. H., & Deuse, J. (2021b). An approach for deep reinforcement learning for production planning in value streams. In J. Franke & P. Schuderer (Hrsg.), *Simulation in Produktion und Logistik* (S. 347–356). Cuvillier.
- West, N., Schlegl, T., & Deuse, J. (2021c). Feature extraction for time series classification using univariate descriptive statistics and dynamic time warping in a manufacturing environment. *IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering* (S. 762–768). <https://doi.org/10.1109/ICBAIE52039.2021.9389954>.
- West, N., Schwenken, J., & Deuse, J. (2022a). Comparative study of methods for the real-time detection of dynamic bottlenecks in serial production lines. In H. Fujita, P. Fournier-Viger, Ali, M., Wang, Y. (Hrsg.), *Advances and trends in artificial intelligence. Theory and practices in artificial intelligence* (S. 3–14). Springer. https://doi.org/10.1007/978-3-031-08530-7_1.
- West, N., Syberg, M., & Deuse, J. (2022b). A holistic methodology for successive bottleneck analysis in dynamic value streams of manufacturing companies. *Proceedings of the 8th Changeable, Agile, Reconfigurable and Virtual Production Conference* (S. 3–14). https://doi.org/10.1007/978-3-030-90700-6_69.
- Wöstmann, R., Barthelmey, A., West, N., & Deuse, J. (2019). A retrofit approach for predictive maintenance. *Thorsten Schüppstuhl, Kirsten Tracht und Jürgen Roßmann (Hrsg.): Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter* (S. 94–106). Springer Vieweg. https://doi.org/10.1007/978-3-662-59317-2_10

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

