

Oliver Jöbstl • Jürgen Lipp • Manfred Strohrmann

WORKBOOK – Digitale Transformation des QM

Data-Science-Innovationen
erfolgreich umsetzen

HANSER





Ihr Plus – digitale Zusatzinhalte!

Auf unserem Download-Portal finden Sie zu diesem Titel kostenloses Zusatzmaterial. Geben Sie dazu einfach diesen Code ein:

`plus-6dfps-ja9ud`

plus.hanser-fachbuch.de

Oliver Jöbstl / Jürgen Lipp / Manfred Strohrmann

Workbook – Digitale Transformation des Qualitätsmanagements

Data-Science-Innovationen erfolgreich umsetzen

HANSER

Print-ISBN 978-3-446-47762-9
E-Book-ISBN 978-3-446-48043-8

Alle in diesem Werk enthaltenen Informationen, Verfahren und Darstellungen wurden zum Zeitpunkt der Veröffentlichung nach bestem Wissen zusammengestellt. Dennoch sind Fehler nicht ganz auszuschließen. Aus diesem Grund sind die im vorliegenden Werk enthaltenen Informationen für Autor:innen, Herausgeber:innen und Verlag mit keiner Verpflichtung oder Garantie irgendeiner Art verbunden. Autor:innen, Herausgeber:innen und Verlag übernehmen infolgedessen keine Verantwortung und werden keine daraus folgende oder sonstige Haftung übernehmen, die auf irgendeine Weise aus der Benutzung dieser Informationen – oder Teilen davon – entsteht. Ebenso wenig übernehmen Autor:innen, Herausgeber:innen und Verlag die Gewähr dafür, dass die beschriebenen Verfahren usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt also auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die endgültige Entscheidung über die Eignung der Informationen für die vorgesehene Verwendung in einer bestimmten Anwendung liegt in der alleinigen Verantwortung des Nutzers.

Bibliografische Information der Deutschen Nationalbibliothek:

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

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Werkes, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Einwilligung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder einem anderen Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung – mit Ausnahme der in den §§ 53, 54 UrhG genannten Sonderfälle –, reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Wir behalten uns auch eine Nutzung des Werks für Zwecke des Text- und Data Mining nach § 44b UrhG ausdrücklich vor.

Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

© 2024 Carl Hanser Verlag München
www.hanser-fachbuch.de

Lektorat: Lisa Hoffmann-Bäumli
Herstellung: Carolin Benedix
Satz: Eberl & Koesel Studio, Kempten
Covergestaltung: Max Kostopoulos
Titelmotiv: © gettyimages.de/berya113
Druck und Bindung: Hubert & Co. GmbH und Co. KG BuchPartner,
Göttingen



Printed in Germany

Inhalt

1 Warum dieses Workbook?	X		
2 Einführung in Data Science	4		
2.1 Arten von Use Cases	5		
2.2 Machine Learning und künstliche Intelligenz	7		
2.3 Typische Anwendungsfälle von Machine Learning	9		
3 Vorgehensmodell	14		
4 Die richtigen Use Cases finden und verstehen	18		
4.1 Prozess identifizieren und abgrenzen	20		
4.2 Stakeholder-Anforderungen verstehen	22		
4.3 Prozessorientierte Use-Case-Ideen ableiten	24		
		4.4	Produkt- und serviceorientierte Use Cases finden
			26
		4.5	Use-Case-Ideen priorisieren
			28
		4.6	Use Cases aus Kundensicht detail- lierter beschreiben
			30
		4.7	Business Case beschreiben
			32
		4.8	Use Cases final bewerten und auswählen
			34
		5 Business-Ziele ableiten und Projekt planen	36
		5.1	Agile Projektorganisation aufsetzen
			38
		5.2	Use Case im Ablauf detailliert beschreiben
			40
		5.3	Wirkende Parameter erkennen ...
			42
		5.4	Geschäftsziele präzisieren und technische Ziele ableiten
			44
		5.5	Projektplan erstellen
			46

6 Daten erheben und verstehen	48	8 Lösungen ausarbeiten und bewerten	80
6.1 Relevante Features finden	50	8.1 Visualisierungslösungen designen	82
6.2 Datenerfassung planen	52	8.2 Dashboard-Lösungen ausarbeiten	84
6.3 Messsystemfähigkeit sicherstellen	54	8.3 Test- und Tuning-Strategie festlegen	86
6.4 Datenbedarf abschätzen	56	8.4 Regressionsmodelle trainieren, tunen und testen	88
6.5 IT-Infrastruktur für die Daten- speicherung entwickeln und realisieren	58	8.5 Klassifizierungsmodelle trainieren, tunen und testen	90
6.6 Data Ingestion Pipelines entwerfen	60	8.6 Neuronale Netze trainieren, tunen und testen	92
6.7 Datensätze durch Versuche erzeugen	62	8.7 Cluster-Verfahren trainieren und tunen	94
6.8 Eindimensionale Daten verstehen	64	8.8 Reinforcement-Verfahren trainieren, tunen und bewerten . .	96
6.9 Mehrdimensionale Daten verstehen	66	8.9 Large Language Models feintunen	98
7 Daten präparieren	68	8.10 Zeitreihen zerlegen und visualisieren	100
7.1 Daten bereinigen	70	8.11 Ausreißer in Zeitreihen identifizieren	102
7.2 Relevante Features auswählen und konstruieren	72	8.12 Bestes Modell auswählen	104
7.3 Features codieren	74		
7.4 Daten komprimieren (Dimension reduzieren)	76		
7.5 Zeitreihen vorverarbeiten	78		

9 Lösung validieren	106	11 Lösung warten und aktualisieren	128
9.1 Interpretierbarkeit von Machine-Learning-Modellen sicherstellen ..	108	11.1 Alterungsverhalten des Modells beschreiben	130
9.2 Lösung auf Plausibilität prüfen ...	110	11.2 Online-Überwachung für Modell-drift erarbeiten	132
9.3 Ergebnis aus der Business-Perspektive bewerten	112	11.3 Online-Überwachung für Daten-drift erarbeiten	134
9.4 Prozess reviewen	114	11.4 Aktualisierungsstrategie festlegen	136
10 Lösung einführen und industrialisieren	116	11.5 Training und Wartung von ML-Modellen automatisieren (MLOps)	138
10.1 Nicht-funktionale Anforderungen und Architekturtreiber identifizieren	118	12 Die Autoren	140
10.2 Entscheidung treffen bezüglich On-Premises oder Cloud-Service-modellen	120	13 Index	144
10.3 IT-Infrastruktur für die Produktivsetzung erarbeiten	122		
10.4 IT-Architekturvarianten bewerten und entscheiden	124		
10.5 Data-Science-Lösung einführen und Vertrauen aufbauen	126		

01

**Warum dieses
Workbook?**

Unternehmen können sich immer weniger auf den vergangenen Erfolgen ausruhen. Die Mitbewerber, die Technologien und der Markt müssen ständig beobachtet werden, um den Innovationsbedarf und die damit verbundenen Chancen rechtzeitig zu erkennen und zu nutzen. Innovationen systematisch umzusetzen, ist mehr denn je ein essenzieller Erfolgsfaktor, und ein wesentliches Innovationspotenzial liegt in der Digitalisierung und den damit generierten Daten, die viel zu oft in Unternehmen ungenutzt bleiben.

In dem Buch „Die digitale Transformation des Qualitätsmanagements“ zeigen die Autoren Möglichkeiten auf, die Digitalisierung dazu zu nutzen, das Qualitätsmanagement auf ein deutlich höheres Niveau zu heben, beispielsweise durch digitale QM-Systeme oder durch die Verbesserung der Produkt- und Prozessqualität mithilfe von Daten und digitalen Technologien. In dem Buch wird auch ein Vorgehensmodell zur systematischen Identifikation und Umsetzung von Use-Case-Ideen zur Prozessverbesserung vorgestellt. Aufgrund der häufig gestellten

Fragen zu dem Thema und der hohen Praxisrelevanz haben sich die Autoren entschlossen, diese Vorgehenslogik im Detail zu erläutern und in Form des vorliegenden Workbooks umsetzungsorientiert zur Verfügung zu stellen.

Dieser Praxisleitfaden fokussiert sich nicht ausschließlich auf das Thema Qualitätsmanagement, sondern bezieht sich auf alle Arten von datengetriebenen Innovationen mit Fokus auf das industri-

elle Umfeld, wobei das Spektrum von einfachen Visualisierungslösungen bis hin zu Prognosemodellen mithilfe von Machine Learning und künstlicher Intelligenz reicht. Diese Innovationen im Bereich Data Science können sich einerseits darauf beziehen, neue digitale Lösungen und Services mit Mehrwert für Kunden zu entwickeln, und andererseits darauf, die bestehenden Prozesse systematisch und kontinuierlich hinsichtlich Effektiv-



Bild 1.1 Die Wirkung von Innovationen im Bereich Data Science

tät, Effizienz und Flexibilität zu verbessern. Darüber hinaus können neue Business-Modelle generiert werden, die eine völlig neue Art der Leistungserbringung und -verrechnung mit sich bringen (Bild 1.1).

Die Möglichkeiten hierzu sind in den letzten Jahren enorm gestiegen – nicht nur, weil immer mehr Daten zur Verfügung stehen, sondern weil aktuell auch die Möglichkeiten zur Verarbeitung von großen Datenmengen vorhanden sind und entsprechende intelligente Algorithmen in rasender Geschwindigkeit weiterentwickelt wurden, wie am jüngsten Hype um ChatGPT zu sehen ist.

Die Vorgehensweise vom Finden der erfolgversprechendsten Use Cases hin zur gewinnbringenden und nachhaltigen Umsetzung muss jedoch von Unternehmen beherrscht werden. Dieser Prozess ist unternehmensspezifisch durch die Definition von entsprechenden Verantwortlichkeiten und Rollen zu organisieren. Es handelt sich hierbei um einen hochgradig interdisziplinären Prozess, weil die unterschiedlichsten Kompetenzen, beispielsweise Fachexperte, IT-Spezialist, Data Analyst und Software Designer eine gute Gesprächs- und Arbeitsbasis finden und ein einheitliches Bild für die Vorgehensweise erarbeiten müs-

sen. Dies geschieht in einer sehr komplexen Ausgangssituation, wo der Erfolg der Innovation keineswegs garantiert werden kann.

Erfolgsentscheidend ist eine Vorgehenslogik, die Orientierung und Klarheit schafft, für effektive Teamarbeit sorgt und je nach Aufgabenstellung geeignete Techniken vorschlägt, die in zielführender Art und Weise zu verwenden sind. Genau diese Vorgehenslogik wird in diesem Workbook vorgestellt.

02

Einführung in Data Science