

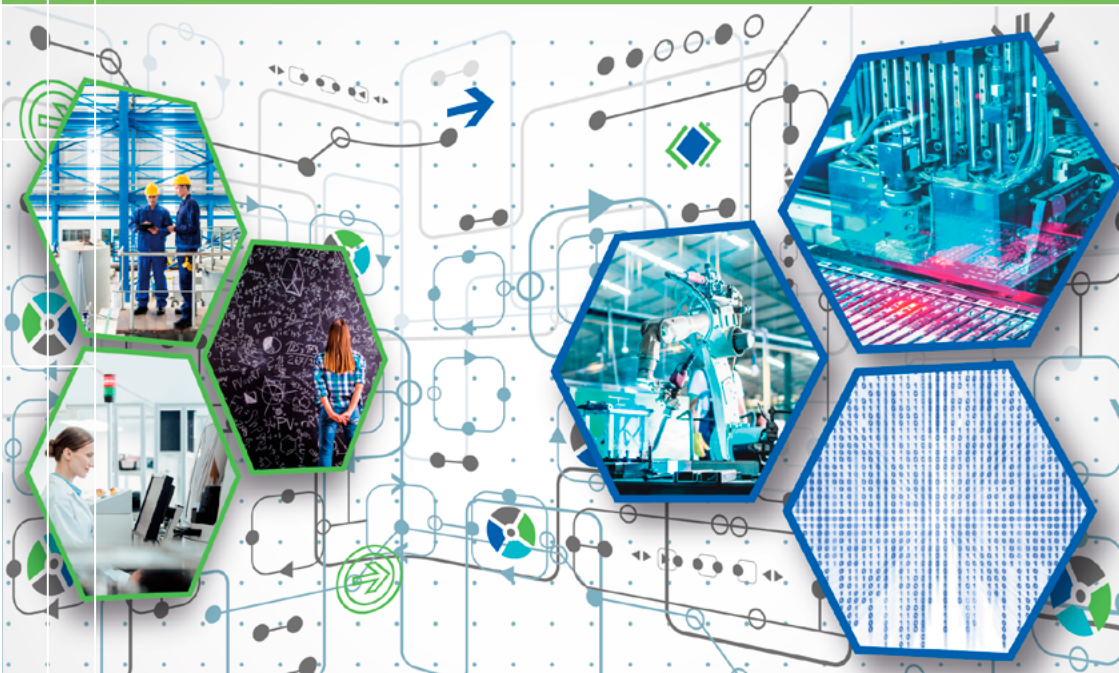
Dieses Kompendium kann für eine Kostenerstattung in Höhe von Euro 20,- bestellt werden. Hierzu senden Sie bitte eine schriftliche Bestellung mit vollständigen Kontaktangaben und der benötigten Anzahl an: wincheringer@hs-koblenz.de

Wissenschaftliche Schriften

**Data Science in der Produktion –
Grundlagen, Beispiele und Erkenntnisse**

Marc Wirtgen, M. Sc.
Prof. Dr.-Ing. Walter Wincheringer

Fachbereich Ingenieurwesen Nr. 01/2022



Wissenschaftliche Schriften des Digitalen Produktionslabors

Fachbereich Ingenieurwesen

Hochschule Koblenz - University of Applied Sciences

**Data Science in der Produktion –
Grundlagen, Beispiele und Erkenntnisse**

von

Marc Wirtgen, M.Sc.

Prof. Dr.-Ing. Walter Wincheringer

Aus Gründen der leichteren Lesbarkeit wird in der vorliegenden wissenschaftlichen Schrift die männliche Sprachform bei personenbezogenen Substantiven und Pronomen verwendet. Dies impliziert jedoch keine Benachteiligung des weiblichen Geschlechtes, sondern ist im Sinne der sprachlichen Vereinfachung als geschlechtsneutral zu verstehen.

DE – 56075 Koblenz, Juli 2022

Digitales Produktionslabor

Alle Rechte vorbehalten.

© Digitales Produktionslabor Hochschule Koblenz – University of Applied Sciences. Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechts ist ohne Zustimmung von Prof. Dr.-Ing. Walter Wincheringer unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Tabellenverzeichnis	III
1 Vorwort	1
2 Einführung	2
3 Data Science - Grundlagen.....	5
3.1 Big Data	7
3.2 Data Analytics	9
3.3 Data Mining und Maschinelles Lernen	11
3.3.1 Lernverfahren	12
3.3.2 Deep Learning und Künstliche Neuronale Netze	13
3.3.3 Automatisiertes Maschinelles Lernen	14
4 Data Science Projekte in der Produktion	16
4.1 CRISP DM als Standardvorgehensweise	16
4.2 Anwendungsfelder und Methoden	18
5 Analyse von Engpässen zur Durchsatzoptimierung.....	21
5.1 Lokalisierung von Engpässen	21
5.1.1 Anwendung eines statistischen Ansatzes	21
5.1.2 Anwendung des hierarchischen Clusters.....	23
5.2 Schlussfolgerungen und weitere Rechercheergebnisse	28
6 Vorhersage der Durchlaufzeit	31
6.1 Vorausschauende Identifikation von Störungswirkungen	32
6.1.1 Anwendung Künstlich Neuronaler Netze	33
6.1.2 Anwendung der Support Vektor Maschinen	38
6.2 Schlussfolgerungen	41
7 Zusammenfassung	43
8 Handlungsempfehlungen	45
9 Literaturverzeichnis.....	46

Abbildungsverzeichnis

Abbildung 1: Historische Einordnung des Digitalen Zeitalters [2, S. 25]	2
Abbildung 2: Detaillierungsgrad i.A.a. [3, S. 1033].....	3
Abbildung 3: Wissenspyramide i.A.a. [10].....	5
Abbildung 4: Bestandteile von Data Science i.A.a. [11]	6
Abbildung 5: Eigenschaften von Big Data: 5 V's [15, S. 49].....	8
Abbildung 6: Klassifikation von Big Data Technologien i.A.a. [16, S. 23].....	8
Abbildung 7: Ausprägungen von Data Analytics [18].....	9
Abbildung 8: Daten zur Beschreibung einer Blackbox [19].....	10
Abbildung 9: Qualitative Unterscheidung i.A.a. [5].....	11
Abbildung 10: Data Science und Maschinelles Lernen i.A.a. [20, S. 16]	12
Abbildung 11: Künstlich Neuronales Netz (KNN) [21].....	14
Abbildung 12: Cross Industry Standard Process for Data Mining i.A.a. [28].....	17
Abbildung 13: Anwendungscluster i.A.a. [30, S. 40]	18
Abbildung 14: Ableitung der Verteilungsfunktion i.A.a. [35, S. 161]	22
Abbildung 15: Statistische Engpassanalysen i.A.a. [54]	23
Abbildung 16: Layout der Produktionslinie [36, S. 148]	23
Abbildung 17: Zeitreihendiagramm für jede Maschine [36, S. 150].....	25
Abbildung 18: Dendrogramm mit Clustern [36, S. 153].....	27
Abbildung 19: Repräsentative Zeitreihen je Cluster [36, S. 154].....	28
Abbildung 20: Zeitschere i.A.a. [37, S. 1].....	31
Abbildung 21: Identifikation von Störungswirkungen [i.A.a. 49, S. 3].....	32
Abbildung 22: Exemplarische Darstellung eines Dashboards [38, S. 7]	33
Abbildung 23: Simuliertes Produktionsszenario [37, S. 6]	34
Abbildung 24: KNN-basiertes Frühwarnsystems [37, S. 3].....	35
Abbildung 25: Konzeptioneller Aufbau des KNN [37, S. 6]	36
Abbildung 26: Beispielhaftes Produktionsszenario [38, S. 6].....	38
Abbildung 27: SVM-basiertes Frühwarnsystems [38, S. 3].....	39
Abbildung 28: Abweichung Prognosewert und Ist-DLZ [38, S. 7]	40

Tabellenverzeichnis

Tabelle 1: Ereignisse und deren jeweilige Klassifikation [55, S. 148]	24
Tabelle 2: $T_{30 \times 13}$ Matrix [55, S. 149]	25
Tabelle 3: Distanzmatrix nach DTW aller Maschinenpaare [55, S. 151]	26
Tabelle 4: Maschinencluster	27
Tabelle 5: Verwendete Kennzahlen [i.A.a. 49, S. 4].....	33
Tabelle 6: Verfügbarkeiten der Arbeitsstationen [56, S. 7].....	34
Tabelle 7: Prozessdaten nach 100 Minuten [56, S. 7]	37
Tabelle 8: Ergebnis der Prognosegenauigkeit der KNN [56, S. 8]	37
Tabelle 9: Prognosefehler [57, S. 7]	41

1 Vorwort

Die Herausforderungen für Produktionsunternehmen in Deutschland steigen seit Jahren an. Dies ist nicht nur in der zunehmenden Individualisierung der Kundenwünsche und der damit verbundenen Abnahme der Losgröße begründet, sondern auch in der gestiegenen Kundenerwartung bzgl. einer kurzfristigen Lieferfähigkeit der Produkte. Weitere Herausforderungen sind in dem dramatisch abnehmenden technologischen Abstand zwischen Deutschland und anderen Industriestandorten, insbesondere aus Fernost, begründet. Damit nicht genug, sind in jüngster Vergangenheit die Herausforderungen durch die instabilen Lieferketten und extrem gestiegenen Energiekosten unverhältnismäßig angestiegen.

Vor diesem Hintergrund müssen die Potentiale der Industrie 4.0 und damit die Digitalisierung der Wertschöpfungsprozesse – auch als Smart Factory bezeichnet – forciert werden. Eine Smart Factory setzt Smart Operation voraus und damit rücken die Daten der Produktion in den Fokus der Betrachtung. Big-Data, Data-Analytics sind neben der KI die aktuellen Schlagworte. Hierbei ist die analytische Verarbeitung von Daten Gegenstand der Datenwissenschaft und eine wesentliche Voraussetzung für die digital vernetzte Produktion.

Das Digitale Produktionslabor (DPL) der Hochschule Koblenz beschäftigt sich mit der Planung, der Simulation und Optimierung von Produktionsabläufen und der Intralogistik. Ziel der vorliegenden Schrift ist es, Anwendungen zur Optimierung von Produktionsabläufen zu identifizieren, Methoden und Ergebnisse zu untersuchen sowie praxisrelevante Erkenntnisse abzuleiten. Dabei steht der Aufbau von Wissen und Fähigkeiten im Fokus.

Insbesondere für KMUs ist die Erschließung und Nutzung dieser Fähigkeiten eine Herausforderung, da es unter anderem an Expertise und digitaler Infrastruktur fehlt. Der Bedarf an Fachleuten sowie am notwendigen Technologietransfer scheint groß. Es gilt Chancen zu nutzen und Risiken zu vermeiden.

Daher geben wir Ihnen mit dieser Schrift einen kompakten Überblick über die Grundlagen, das Standardvorgehen und Methoden der Datenanalyse sowie die bisher erschlossenen Anwendungsfelder in der Produktion. Es werden anhand von Use Cases die Engpassanalyse und die Vorhersage der Durchlaufzeit fokussiert betrachtet.

Wir wünschen Ihnen beim intensiven Studium dieser praxisnahen, wissenschaftlichen Schrift viele Erkenntnisgewinne und den Mut, diese Erkenntnisse in der Praxis erfolgreich umzusetzen.

Prof. Dr.-Ing. Walter Wincheringer

Marc Wirtgen, M.Sc.

9 Literaturverzeichnis

- [1] A. Mockenhaupt, *Digitalisierung und Künstliche Intelligenz in der Produktion: Grundlagen und Anwendung*. Wiesbaden: Springer Fachmedien Wiesbaden, 2021. [Online]. Verfügbar unter: <http://swbplus.bsz-bw.de/bsz1756955328cov.htm>
- [2] „WGP-Standpunkt Industrie 4.0“. [Online]. Verfügbar unter: https://www.ipa.fraunhofer.de/content/dam/ipa/de/documents/Presse/Presseinformationen/2016/Juni/WGP_Standpunkt_Industrie_40.pdf
- [3] D. Rippel, M. Lütjen und B. Scholz-Reiter, „A framework for the quality-oriented design of micro manufacturing process chains“, *Journal of Manufacturing Technology Management*, Jg. 25, Nr. 7, S. 1028–1048, 2014, doi: 10.1108/JMTM-11-2012-0110.
- [4] „potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland“. [Online]. Verfügbar unter: https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.pdf?__blob=publicationFile&v=8
- [5] Benjamin Maschler, Dustin White und Michael Weyrich, „Anwendungsfälle und Methoden der künstlichen Intelligenz in der anwendungsorientierten Forschung im Kontext von Industrie 4.0“. [Online]. Verfügbar unter: https://www.ias.uni-stuttgart.de/dokumente/publikationen/2020-Anwendungsaefaelle_und-Methoden_der_kuenstlichen_Intelligenz_in_der_anwendungsorientierten_Forschung_im_Kontext_von_Industrie_4.0.pdf
- [6] N. Vojdani und B. Erichsen, „Anwendungspotenziale von maschinellem Lernen in der Produktion und Logistik“, 2020.
- [7] „WGP-Standpunkt_KI-final_20190906-2“. [Online]. Verfügbar unter: https://wgp.de/wp-content/uploads/WGP-Standpunkt_KI-final_20190906-2.pdf
- [8] The Economist, *The world's most valuable resource is no longer oil, but data*. [Online]. Verfügbar unter: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> (Zugriff am: 13. Dezember 2021).

- [9] Prof. Dr. Hans-Werner Wohltmann, „Definition: Daten“, *Springer Fachmedien Wiesbaden GmbH*, 19. Feb. 2018, 2018. [Online]. Verfügbar unter: <https://wirtschaftslexikon.gabler.de/definition/daten-30636>. Zugriff am: 13. Dezember 2021.
- [10] Raffael, „Wissenspyramide“, *Der WINF*, 12. Sep. 2012, 2012. [Online]. Verfügbar unter: https://derwirtschaftsinformatiker.de/2012/09/12/it-management/wissenspyramide-wiki/#google_vignette. Zugriff am: 13. Dezember 2021.
- [11] Michael Freitag, Mirko Kück, Abderrahim Ait Alla, Michael Lütjen, (PDF) *Potenziale von Data Science in Produktion und Logistik: Teil 1 - Eine Einführung in aktuelle Ansätze der Data Science*. [Online]. Verfügbar unter: https://www.researchgate.net/publication/284158759_Potenziale_von_Data_Science_in_Produktion_und_Logistik_Teil_1_-_Eine_Einfuehrung_in_aktuelle_Ansatze_der_Data_Science/related (Zugriff am: 14. Dezember 2021).
- [12] R. Brooks, „Die sieben Todsünden der KI-Vorhersagen“, *heise online*, 8. Nov. 2017, 2017. [Online]. Verfügbar unter: <https://www.heise.de/hintergrund/01-3876558.html>. Zugriff am: 14. Dezember 2021.
- [13] V. Brühl, „Big Data, Data Mining, Machine Learning und Predictive Analytics: Ein konzeptioneller Überblick“, Frankfurt a. M.: Goethe University Frankfurt, Center for Financial Studies (CFS), CFS Working Paper Series 617, 2019. [Online]. Verfügbar unter: <https://www.econstor.eu/handle/10419/191736>.
- [14] D. Frick *et al.*, Hg., *Data Science: Konzepte, Erfahrungen, Fallstudien und Praxis*, 1. Aufl. Wiesbaden: Springer Fachmedien Wiesbaden, 2021. [Online]. Verfügbar unter: <http://swbplus.bsz-bw.de/bsz1777508894cov.htm>
- [15] P. Gölzer, „Big Data in Industrie 4.0 - Eine strukturierte Aufarbeitung von Anforderungen, Anwendungsfällen und deren Umsetzung: Doctoralthesis“, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). [Online]. Verfügbar unter: <https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/8106>
- [16] „140228-Big-Data-Technologien-Wissen-fuer-Entscheider“. [Online]. Verfügbar unter: <https://www.bitkom.org/sites/default/files/file/import/140228-Big-Data-Technologien-Wissen-fuer-Entscheider.pdf>

- [17] M. Schreiber, P. Bausch, J. Best und J. Metternich, „Datenanalyse in Produktionsprozessen“, *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, Jg. 115, Nr. 5, S. 309–313, 2020, doi: 10.3139/104.112285.
- [18] Simon Schuller und Sebastian Werler, „Festo Automation Experience - Höhere Produktivität mit Künstlicher Intelligenz“.
- [19] L. Breiman, „Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)“ (en), ss, Jg. 16, Nr. 3, S. 199–231, 2001, doi: 10.1214/ss/1009213726.
- [20] B. Pokorni, M. Braun und C. Knecht, *Menschzentrierte KI-Anwendungen in der Produktion*. Stuttgart: Fraunhofer IAO.
- [21] Wikipedia, *Künstliches neuronales Netz*. [Online]. Verfügbar unter: https://de.wikipedia.org/w/index.php?title=Künstliches_neuronales_Netz&oldid=20294159 (Zugriff am: 26. Februar 2022).
- [22] cartacioS, *Was ist automatisiertes maschinelles Lernen? AutoML - Azure Machine Learning*. [Online]. Verfügbar unter: <https://docs.microsoft.com/de-de/azure/machine-learning/concept-automated-ml> (Zugriff am: 19. Dezember 2021).
- [23] Fraunhofer IAO | Digital Business, *Cloudbasierte KI-Plattformen*. [Online]. Verfügbar unter: <https://www.digital.iao.fraunhofer.de/de/publikationen/CloudbasierteKI-Plattformen.html> (Zugriff am: 19. Dezember 2021).
- [24] Systems, MIT Laboratory for Information and Decision, *ML 2.0: Machine learning for many*. [Online]. Verfügbar unter: <https://news.mit.edu/2018/ml-20-machine-learning-many-data-science-0306> (Zugriff am: 18. Dezember 2021).
- [25] <https://www.hannovermesse.de>, *KI & Maschinelles Lernen: Deutsch-Schwedische Machine Learning-Plattform*. [Online]. Verfügbar unter: <https://www.hannovermesse.de/de/news/news-fachartikel/deutsch-schwedische-machine-learning-plattform> (Zugriff am: 18. Dezember 2021).
- [26] S. Kipar, „Seeq erleichtert Integration von Open-Source-Algorithmen“, *Industry of Things*, 27. Okt. 2021, 2021. [Online]. Verfügbar unter: <https://www.industry-of-things.de/seeq-erleichtert-integration-von-open-source-algorithmen-a-1067860/?cmp=nl-345&uuid=>. Zugriff am: 18. Dezember 2021.

- [27] biljanagabric, „Microsoft Word - !VDMA Quick Guide_Machine Learning_Final.docx“. [Online]. Verfügbar unter:
[https://www.vdma.org/documents/34570/0/Quick_Guide_Machine_Learning%20\(3\)_1557231153950.pdf/d7ff7370-1d30-762d-cf7c-ff9aafd581c9](https://www.vdma.org/documents/34570/0/Quick_Guide_Machine_Learning%20(3)_1557231153950.pdf/d7ff7370-1d30-762d-cf7c-ff9aafd581c9)
- [28] Michael Freitag, Mirko Kück, Abderrahim Ait Alla, Michael Lütjen, (3) (PDF) *Potenziale von Data Science in Produktion und Logistik: Teil 2 - Vorgehensweise zur Datenanalyse und Anwendungsbeispiele*. [Online]. Verfügbar unter:
https://www.researchgate.net/publication/285537085_Potenziale_von_Data_Science_in_Produktion_und_Logistik_Teil_2_-_Vorgehensweise_zur_Datenanalyse_und_Anwendungsbeispiele (Zugriff am: 15. Dezember 2021).
- [29] W. Brenner, B. van Giffen, J. Koehler, T. Fahse und A. Sagodi, *Bausteine eines Managements Künstlicher Intelligenz*. Wiesbaden: Springer Fachmedien Wiesbaden, 2021.
- [30] J. Krauß, J. Dorißen, H. Mende, M. Frye und R. H. Schmitt, „Maschinelles Lernen in der Produktion - Anwendungsgebiete und frei verfügbare Datensätze“, *I40M*, Jg. 2019, Nr. 4, S. 39–42, 2019, doi: 10.30844/I40M_19-4_S39-42.
- [31] „Studie-Einsatzfelder-KI-im-Produktionsumfeld“. [Online]. Verfügbar unter:
<https://www.i40-bw.de/wp-content/uploads/2020/09/Studie-Einsatzfelder-KI-im-Produktionsumfeld.pdf>
- [32] Erich Schubert, *ML L2b: Validation and Model Selection - YouTube*. [Online]. Verfügbar unter:
https://www.youtube.com/watch?v=Vu3hAsDwvdU&list=PLElvkFQko9bc3LvEZJnROvjC_fl5Xt1a1&index=21 (Zugriff am: 25. Mai 2022).
- [33] P. Roh, „Methoden zur datengetriebenen Lokalisierung des Verbesserungspotentials in Produktionsabläufen“, ETH Zurich, 2019.
- [34] M. Subramaniyan, A. Skoogh, J. Bokrantz, M. A. Sheikh, M. Thüerer und Q. Chang, „Artificial intelligence for throughput bottleneck analysis – State-of-the-art and future directions“, *Journal of Manufacturing Systems*, Jg. 60, S. 734–751, 2021, doi: 10.1016/j.jmsy.2021.07.021.

- [35] P. Roh, A. Kunz und T. Netland, „Data-driven detection of moving bottlenecks in multi-variant production lines“, *IFAC-PapersOnLine*, Jg. 51, Nr. 11, S. 158–163, 2018, doi: 10.1016/j.ifacol.2018.08.251.
- [36] M. Subramaniyan, A. Skoogh, A. S. Muhammad, J. Bokrantz, B. Johansson und C. Roser, „A generic hierarchical clustering approach for detecting bottlenecks in manufacturing“, *Journal of Manufacturing Systems*, Jg. 55, S. 143–158, 2020, doi: 10.1016/j.jmsy.2020.02.011.
- [37] N. Vojdani und B. Erichsen, „Ein Ansatz für ein Predictive-Monitoring-System zur Identifikation von Störungswirkungen in der Produktionslogistik mittels künstlichen neuronalen Netzen“, 2018.
- [38] N. Vojdani und B. Erichsen, „Anwendung von Support Vector Regression zur vorausschauenden Identifikation von Störungswirkungen in der Produktionslogistik“, 2019.
- [39] „SmoSch04“. [Online]. Verfügbar unter:
<https://alex.smola.org/papers/2004/SmoSch04.pdf>



Hochschule Koblenz
Fachbereich Ingenieurwesen
Konrad-Zuse-Str. 1
56075 Koblenz

Hochschule Koblenz
z.H. Prof. Dr.-Ing. Walter Wincheringer
Konrad-Zuse-Str. 1
56075 Koblenz

Wir freuen uns, dass Sie sich für unser Schriftenreihe interessieren. Mit diesem Formular können Sie unsere Kompendien für eine Kostenerstattung (Betrag siehe unten) bestellen. Schicken Sie uns einfach das ausgefüllte Formular per mail an wincheringer@hs-koblenz.de zurück. Gerne nehmen wir Ihre Anfrage auch postalisch entgegen.

Firma _____

Vor- und Nachname, Titel _____

Straße, Hausnummer _____

Postleitzahl, Ort _____

Mail _____

Telefon oder E-Mail falls Rückfragen _____

Folgende Veröffentlichung/en möchte ich bestellen:

Titel:	Bestellen	Anzahl
Kompendium: Wertstromoptimierung ISSN: 2625-5782, 45 Seiten, Kosten: 15,--€	<input type="checkbox"/>	
Smart Factory – Eine Orientierungshilfe für die Praxis ISSN: 2625-5782, 72 Seiten, Kosten: 15,--€	<input type="checkbox"/>	
Ganzheitliche Produktionssysteme in Theorie und Praxis ISSN: 2625-5782, 176 Seiten, Kosten: 22,--€	<input type="checkbox"/>	
VR-Einsatz bei der Fabrik- und Arbeitsstättenplanung ISSN: 2625-5782, 48 Seiten, Kosten: 15,--€	<input type="checkbox"/>	
Routenzugsysteme in der Praxis ISSN: 2625-5782, 51 Seiten, Kosten: 15,--€	<input type="checkbox"/>	
Data-Science in der Produktion ISSN: 2625-5782, 51 Seiten, Kosten: 20,--€	<input type="checkbox"/>	

Datum

Unterschrift