

Voraussetzungen für den Einsatz datengetriebener Methoden in der Produktentwicklung

Jan Mehlstäubl^{1*}, Simon Nicklas¹, Benjamin Gerschütz², Nicolai Sprogies³, Benjamin Schleich², Thomas Lohner³, Sandro Wartzack², Karsten Stahl³, Kristin Paetzold¹

¹ Institut für Technische Produktentwicklung, Universität der Bundeswehr München

² Lehrstuhl für Konstruktionstechnik, Friedrich-Alexander-Universität Erlangen-Nürnberg

³ Forschungsstelle für Zahnräder und Getriebebau, Technische Universität München

* Korrespondierender Autor:

Jan Mehlstäubl

Institut für Technische Produktentwicklung, Universität der Bundeswehr München

Werner-Heisenberg-Weg 39

85579 Neubiberg

Telefon: +49 89 6004-2811

Mail: jan.mehlstaebul@unibw.de

Abstract

Data mining and machine learning are successfully applied in many business areas such as marketing or production. Due to the increasing complexity in data and information flows and the large amount of data, data-driven solutions have also great potential in product development. Nevertheless, many companies are not able to integrate data-driven methods into their product development process. In order to support the application of data-driven methods, this paper identifies necessary prerequisites for their integration in product development. Based on a data analysis process, the prerequisites are elaborated and, subsequently, levers for their design are derived.

Keywords

Data-driven Methods, Product Development, Data Mining, Machine Learning

1. Einleitung

Die Entwicklung eines am Markt erfolgreichen Produkts ist das Ergebnis einer Reihe von Entscheidungen mit einem hohen Maß an Komplexität und Unsicherheit [1]. Je mehr Wissen und Erfahrung Entwickler besitzen, desto größer ist die Wahrscheinlichkeit richtige Entscheidungen zu treffen [2]. Um Entwickler bei ihren Tätigkeiten und Entscheidungen zu unterstützen, können die stetig anwachsenden Datenmengen in Industrieunternehmen genutzt und mit geeigneten Algorithmen analysiert werden [3]. Daten können kontextuelle Aspekte enthalten, welche mit dem impliziten Erfahrungswissen von Entscheidungsträgern vergleichbar sind. Dieses Erfahrungswissen ist oft schwer in Regeln zu formalisieren, kann aber mit Hilfe von Lernmodellen abgebildet werden [4]. Generell können Unternehmen durch die Analyse von Daten ein besseres Verständnis für ihre Produkte, Kunden und Märkte gewinnen [5]. Trotz des großen Potentials datengetriebener Methoden werden diese in der Produktentwicklung bislang kaum eingesetzt [6], sodass die Produktentwicklung diesbezüglich anderen Geschäftsbereichen, wie der Produktion oder dem Marketing, hinterherhinkt [7].

Das weitestgehend ungenutzte Potential datengetriebener Methoden in der Produktentwicklung führt zu der folgenden Forschungsfrage:

- Welche Voraussetzungen sind im Hinblick auf die Integration von datengetriebenen Methoden in die Produktentwicklung zu berücksichtigen?

Ziel dieses Beitrags ist es, die notwendigen Bedingungen für den Einsatz datengetriebener Methoden in der Produktentwicklung zu identifizieren, deren Zusammenhänge zu analysieren und Stellhebel abzuleiten. Der Fokus liegt hierbei auf technischen und organisatorischen Gesichtspunkten. Dadurch können in einem weiteren Schritt Maßnahmen entwickelt werden, um die erforderlichen Voraussetzungen für die Einführung datengetriebener Methoden in der Produktentwicklung zu schaffen.

In dieser Arbeit wird in Kapitel 2 ein Überblick über datengetriebene Methoden in der Produktentwicklung gegeben. Dabei werden die Begriffe datengetriebene Methode, Data Mining und Machine Learning definiert sowie deren Potentiale in der Produktentwicklung anhand von bisherigen Einsatzmöglichkeiten verdeutlicht. Kapitel 3 beschreibt das Forschungsvorgehen, welches zur Erstellung dieses Beitrags verfolgt wird. In Kapitel 4 werden die Voraussetzungen für den Einsatz datengetriebener Methoden anhand eines Datenanalyseprozesses abgeleitet. Diese werden anschließend in Kapitel 5 gruppiert und Stellhebel als Ansatzpunkte für eine Optimierung ermittelt. Abschließend folgen in Kapitel 6 eine Zusammenfassung der Ergebnisse und ein Ausblick auf zukünftige Forschungstätigkeiten.

2. Datengetriebene Methoden in der Produktentwicklung

Datengetriebene Methoden sind Vorgehen, bei denen Daten mithilfe von Algorithmen analysiert werden, um Entscheidungen, Aktivitäten oder das Erreichen von Zielen zu unterstützen. Sie beinhalten Verfahren des Data Minings und Machine Learnings. Data Mining beschreibt die Anwendung spezifischer Algorithmen zur Extraktion von Mustern aus Daten [8]. Beim Machine Learning werden lernende Algorithmen verwendet, um Muster in vorhandenen Datensätzen zu erkennen und darauf aufbauend Vorhersagen zu treffen [9]. Die Anwendung von Data Mining und Machine Learning Verfahren stellt nur einen Schritt in einem Datenanalyseprozess dar [8]. Für die Entwicklung einer datengetriebenen Methode muss ein Datenanalyseprozess vollständig durchlaufen werden.

In der Literatur sind unter anderem in [10] und [11] Übersichten bisheriger Einsatzgebiete datengetriebener Methoden in der Produktentwicklung zu finden. Grundsätzlich bieten diese dort die Möglichkeit, die Entscheidungsfindung mithilfe von Daten zu optimieren [12]. Dies

kann durch den Erwerb von deskriptiven oder prädiktiven Erkenntnissen aus den Datensätzen geschehen [13]. Deskriptive Analysen konzentrieren sich darauf, was in der Vergangenheit passiert ist und welche zugrundeliegenden Ursachen die Ereignisse besitzen [12]. Die Ergebnisse solcher Analysen müssen manuell interpretiert werden, bevor sie zur Entscheidungsfindung genutzt werden können [14]. Prädiktive Datenanalysen beschäftigen sich mit der Fragestellung, was in Zukunft passieren wird. Durch Verwendung von Methoden des Data Mining und Machine Learning werden zukünftige Entwicklungen prognostiziert und die Wahrscheinlichkeit unbekannter Ereignisse geschätzt [15]. Prognosen dieser Art basieren auf historischen Daten, die auf Muster und Abhängigkeiten untersucht werden, um Ergebnisse in die Zukunft zu extrapolieren und damit die Entscheidungsfindung zu unterstützen [15]. In der Literatur werden Einsatzszenarien sowohl von deskriptiven Analysen als auch von prädiktiven Analysen in unterschiedlichsten Phasen der Produktentwicklung beschrieben. Im Folgenden werden zur Veranschaulichung einige Einsatzmöglichkeiten beispielhaft aufgeführt.

Deskriptive Analysen werden in frühen Entwicklungsphasen insbesondere eingesetzt, um große Textmengen zu analysieren [10]. XU et al. [16] analysieren zum Beispiel Kundenrezensionen hinsichtlich der adressierten Produktmerkmale. WANG and ZHANG [17] extrahieren Kundenbedürfnisse aus Produktrezensionen und ordnen diese zusätzlich vordefinierten Produktspezifikationen mittels einer Klassifikation zu. Einen Ansatz zur Vorverarbeitung und Strukturierung von Anforderungen mithilfe von Natural Language Processing stellen HORBER et al. [18] vor. Daneben wird in [19] Text-Mining eingesetzt, um Produktentwickelnden Simulationswissen aus Berechnungsberichten bereitzustellen und sie dadurch zu befähigen valide Simulationen aufzubauen und ihre Gestaltentwürfe abzusichern. Neben der Verarbeitung von Textmengen ist es in der Produktentwicklung beispielsweise auch möglich, mittels deskriptiven Analysemethoden während der Entwicklung generierte Daten aus numerischer oder analytischer Berechnung, Prüfstand-, Feld- oder Fahrversuchen zentral abzubilden und damit zielgerichtet, effizient und ganzheitlich zu analysieren oder Korrelationen aufzudecken [20].

Ein Beispiel für eine prädiktive Datenanalyse in der Produktentwicklung liefern HOORNAERT et al. [21]. Sie verwenden eine Klassifikation, um zu bestimmen, ob eine neue Produktidee vom Kunden angenommen oder abgelehnt wird. Auch in späteren Phasen der Produktentwicklung werden prädiktive Datenanalysen eingesetzt. Beispielsweise werden Beziehungen zwischen ergonomischen Faktoren des Fahrzeuginnenraums und den Kundenbewertungen modelliert, um ein Optimum bestimmen zu können [22], [23]. BERTONI et al. [1] bilden dagegen durch die Analyse von Nutzungsdaten Zusammenhänge diverser Leistungsparameter eines Radladers auf Basis seiner Betriebsvariablen mittels Regression ab. Auch zur Absicherung der Fertigbarkeit von Gestaltentwürfen im Sinne des fertigungsgerechten Konstruierens können maschinelle Lernverfahren zum Einsatz kommen [24]. Darüber hinaus bietet die Entwicklung von virtuellen Sensoren, die redundant, substituierend oder erweiternd für reale Sensoren eingesetzt werden [25], große Potentiale zur Reduzierung von Kosten und Applikationsaufwand realer Sensoren oder zur Informationserweiterung in frühen und späten experimentellen Absicherungsprozessen [26]. Daneben werden datengetriebene Methoden verwendet, um zum Beispiel Belastungskollektive zu generieren, die bereits in der Entwurfsphase im Getriebeauslegungsprozess berücksichtigt werden können [27]. Datengetriebene Methoden finden ebenso in der frühen experimentellen Absicherung ihre Anwendung. Zum Beispiel können anforderungsspezifische, reale Belastungskollektive mittels Analyse der Daten aus vorhandenen Feld- oder Fahrversuchen bestimmt und in Folgeentwicklungen frühzeitig in experimentellen Absicherungsversuchen abgebildet werden [28].

3. Forschungsvorgehen

Zur Beantwortung der Forschungsfrage wird die Design Research Methodology (DRM) nach BLESSING und CHAKRABARTI [29] Typ 3 verfolgt (Bild 1). Es wird zunächst die Motivation und das Forschungsziel anhand der Literatur abgeleitet. Anschließend wird in der deskriptiven

Studie I ein tieferes Verständnis für datengetriebene Methoden in der Produktentwicklung entwickelt. Dazu werden die Begriffe datengetriebene Methode, Data Mining und Machine Learning näher betrachtet, deren Potentiale für die Produktentwicklung abgeleitet und mit Einsatzszenarien aus der Literatur veranschaulicht. In der präskriptiven Studie werden die Voraussetzungen, welche für die Einführung datengetriebener Methoden in die Produktentwicklung unerlässlich sind, durch die Analyse des Cross Industry Standard Process for Data Mining (CRISP-DM) von WIRTH und HIPP [30] abgeleitet und Stellhebel für deren Beeinflussung identifiziert. Die Stellhebel werden anschließend näher untersucht und deren Zusammenhänge offengelegt.

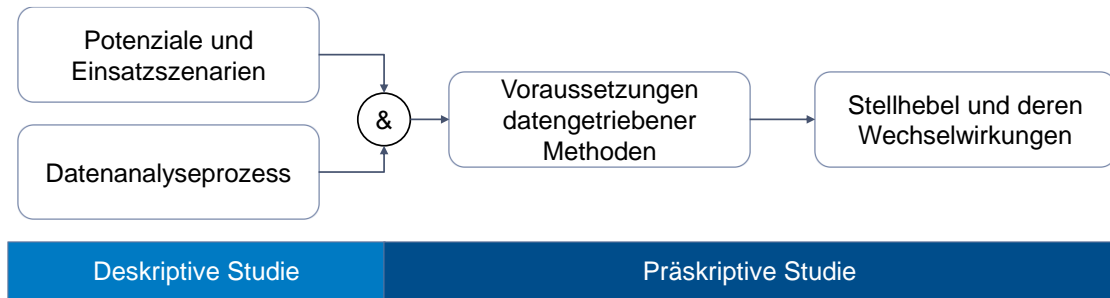


Bild 1: Angewandtes Forschungsvorgehen nach BLESSING und CHAKRABARTI [29]

4. Ableitung der Voraussetzungen für den Einsatz datengetriebener Methoden in der Produktentwicklung

Im Folgenden werden die Voraussetzungen, welche zur Einführung datengetriebener Methoden in die Produktentwicklung vorliegen müssen, unter Betrachtung eines Datenanalyseprozesses ermittelt. Hierfür wird der CRISP-DM von WIRTH und HIPP [30] genutzt. Der CRISP-DM legt den Schwerpunkt auf den industriellen Einsatz datengetriebener Methoden und beinhaltet eine domänenspezifische Betrachtung. Eine Übersicht des Prozesses ist in Bild 2 dargestellt.

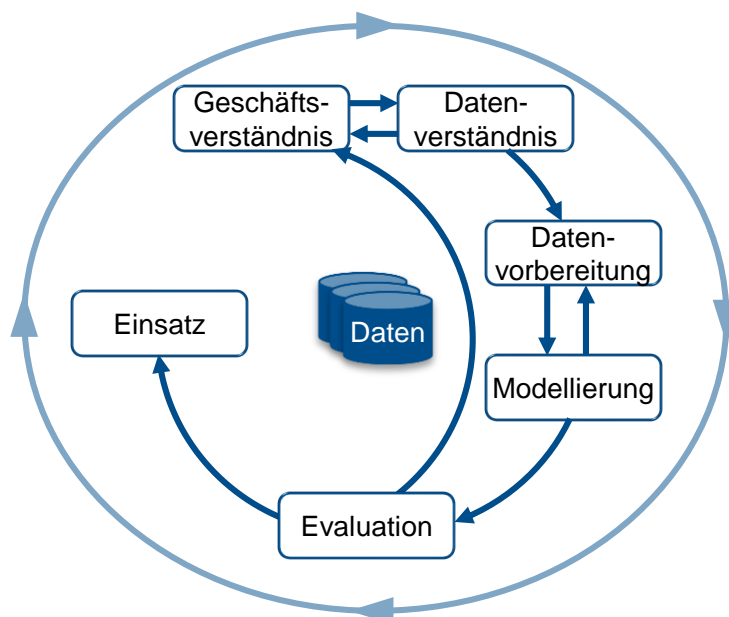


Bild 2: CRISP-DM Prozess [30]

Eine Übersicht der ermittelten Voraussetzungen und deren Zuordnung zu den einzelnen Phasen des Datenanalyseprozesses ist in Tabelle 1 dargestellt. Im Folgenden werden die einzelnen Voraussetzungen in Abhängigkeit der Phasen des CRISP-DM beschrieben.

Tabelle 1: Übersicht der ermittelten Voraussetzungen

Phase	Voraussetzungen	Phase	Voraussetzungen
Geschäftsverständnis	Managementsupport	Modellierung	Data Science und Data Analytics Know-how
	Qualifikationsanforderungen		Toolunterstützung
	Interdisziplinäre Zusammenarbeit		Rechenleistung
	Vertrauen in die Ergebnisse	Evaluation und Einsatz	Bedienfreundliches User Interface
	Relativer Vorteil des Anwendungsfalls		Kompatibilität mit IT-Landschaft
	Bewertungskriterien		Visualisierung
Datenverständnis und -vorbereitung	Datenerzeugung	Nutzungsbereitschaft	
	Datenspeicherung		
	Speicherkapazität		
	Datenmenge		
	Datenqualität		
	Datentransparenz		

4.1. Geschäftsverständnis

Die Phase des Geschäftsverständnisses rückt das Erfassen von Zielen des Datenanalyseprojekts und die Übersetzung dieser in einen konkreten Anwendungsfall für datengetriebene Methoden in den Mittelpunkt. Zu Beginn müssen erforderliche Kapazitäten freigesetzt und ein Datenanalyseprojekt gestartet werden, weshalb das Thema insbesondere von den Entscheidern im Unternehmen unterstützt werden muss. Die Ermittlung von Anwendungsfällen erfordert ein grundlegendes Verständnis über Data Mining und Machine Learning, sowohl im Management als auch in den jeweiligen Entwicklungsabteilungen [31], [32]. Das langfristige Ziel ist die Etablierung einer datengetriebenen Unternehmenskultur, welche die Bereitschaft ausdrückt, Entscheidungen wann immer möglich, auf Basis von Daten zu treffen und diesen zu vertrauen [33]. Darüber hinaus ist eine enge Zusammenarbeit zwischen Produktentwicklern beziehungsweise späteren Nutzern und Datenanalysten erforderlich. Die neue datengetriebene Methode muss einen relativen Vorteil gegenüber den momentan eingesetzten Methoden bieten. Der relative Vorteil beschreibt das Ausmaß, in dem eine Methode als besser wahrgenommen wird als diejenige, die von ihr abgelöst wird [34]. Im Kontext datengetriebener Methoden muss dabei sowohl die persönliche Ebene des Methodenanwenders als auch die betriebliche Ebene betrachtet werden. Somit muss für Methodenanwender der Nutzen dem Aufwand und für Unternehmen der Ertrag den Kosten überwiegen. Für die spätere Evaluation ist eine Spezifikation von Zielen und Parameter zur Überprüfung des Erfolgs datengetriebener Methoden erforderlich.

4.2. Datenverständnis und -vorbereitung

Zur Generierung eines Datenverständnisses werden Daten des zuvor ermittelten Anwendungsfalls gesammelt und näher analysiert, um Datenqualitätsprobleme, erste Erkenntnisse

und relevante Teilmengen zu identifizieren. Anschließend werden bei der Datenvorverarbeitung die gesammelten Rohdaten in den finalen Datensatz überführt. Zu diesen Aufgaben gehört die Datenbereinigung, die Bildung neuer Attribute und die Umformung von Daten für das Modellierungswerkzeug. In den verschiedenen Unternehmensbereichen werden Daten mit strukturierten, semi-strukturierten und unstrukturierten Informationen erzeugt und oft dezentral in unterschiedlichen Ablagesystemen gespeichert. Einzelne Abteilungen haben oft nur Kenntnisse über die bei ihnen erzeugten und gespeicherten Daten [35]. Für die systematische Datennutzung und -auswertung sind jedoch die Daten aus verschiedenen Bereichen erforderlich. Um zu wissen, in welchen Abteilungen und mit welchen IT-Tools Daten erzeugt und gespeichert werden, müssen Unternehmen für eine Transparenz in ihren Datenstrukturen sorgen. Daneben muss sichergestellt sein, dass die für eine Analyse erforderlichen Daten auch generiert und gespeichert werden [36]. Hierfür sind die erwarteten Arbeitsergebnisse zu definieren und gegebenenfalls entsprechende Sensoren an Produkten, Prüfständen oder Produktionsanlagen vorzusehen. Die erzeugte Datenmenge muss zudem ausreichend groß sein, um Muster erkennen und generalisierbare Aussagen treffen zu können. Dafür ist ausreichend Speicherkapazität vorzusehen. Eine weitere Voraussetzung ist die hinreichende Qualität der Daten. Datensätze in der Produktentwicklung enthalten Vorurteile, Unsicherheiten, Eindrücke, Unwahrheiten und fehlende Werte, welche einen starken Einfluss auf die Ergebnisqualität von datengetriebenen Methoden haben [37].

4.3. Modellierung

In der Modellierungs-Phase werden unterschiedliche Algorithmen ausgewählt, angewendet und verglichen, um die definierten Gütekriterien optimale Ergebnisse zu erfüllen. Für die Auswahl und Umsetzung der Algorithmen benötigen Industrieunternehmen Know-how im Bereich Data Analytics [38]. Aufgrund der Komplexität der Produktentwicklung sind Mitarbeiter erforderlich, die zusätzlich über technische und methodische Kenntnisse in der Produktentwicklung verfügen [33]. Eine weitere zentrale Anforderung ist eine ausreichende Rechenleistung, um die Daten im notwendigen Umfang und in angemessener Zeit verarbeiten zu können. Hierfür können Unternehmen entweder eigene Ressourcen einsetzen oder auch auf Cloud Services zurückgreifen [39].

4.4. Evaluation und Einsatz

In der Evaluation werden die zuvor gebildeten Modelle hinsichtlich der in der Geschäftsverständnis-Phase definierten Ziele bewertet und über deren Nutzung entschieden. Anschließend werden die Modelle in aktuelle Prozesse integriert. Der Einsatz einer Methode sollte für den Anwender so einfach wie möglich gestaltet werden und eine niedrige Komplexität aufweisen [34]. Die Analysemodelle datengetriebener Methoden müssen daher in einem bedienfreundlichen Tool implementiert werden. Ein solches Tool muss die Ergebnisse verständlich visualisieren und gegebenenfalls interpretieren [40]. Die Einführung neuer Methoden oder Werkzeuge in Unternehmen stößt dennoch bei den vorgesehenen Nutzern oft auf Widerstand [41]. Die Nutzung datengetriebener Methoden bedarf zunächst einer Bereitschaft zur Anpassung der Produktentwicklungsaktivitäten [35]. Die neue Methode muss konsistent mit gültigen Werten, bisherigen Erfahrungen und Bedürfnissen der potentiellen Anwender sein [34]. Für datengetriebene Methoden bedeutet dies vor allem, dass sie in bereits etablierte IT-Landschaften integriert werden und ein entsprechendes User Interface aufweisen müssen.

5. Stellhebel zur Gestaltung der Voraussetzungen

Aus den in Tabelle 1 dargestellten Voraussetzungen für die Einführung datengetriebener Methoden in der Produktentwicklung können fünf Stellhebel zu deren Beeinflussung abgeleitet

werden. Diese sind die datengetriebene Methode selbst, die Daten, die Methodenentwicklung, der Methodenanwender und das Management. Neue datengetriebene Methoden müssen zunächst einen relativen Vorteil gegenüber den aktuellen Verfahren aufweisen. Außerdem müssen sie bedienfreundlich und kompatibel zu der bisherigen IT-Landschaft sein. Der zweite Stellhebel sind die Daten, welche in ausreichender Menge und Qualität erzeugt und gespeichert werden müssen. Für die Entwicklung einer datengetriebenen Methode benötigen Unternehmen eine gewisse Datentransparenz, Qualifikationen im Bereich Data Science und Data Analytics, geeignete Toolunterstützung im Bereich der Modellierung, Datenerzeugung und -analyse sowie ausreichend Rechenleistung und Speicherkapazität. Die Entwickler müssen die grundlegenden Mechanismen von Data Mining und Machine Learning kennen sowie zur Zusammenarbeit mit Datenanalysten und der Nutzung der neuen Technologien bereit sein. Der letzte Stellhebel findet sich im Management. Dieses muss die Einführung datengetriebener Methoden in Form von Software- und Netzwerkressourcen und durch die Einstellung beziehungsweise Qualifizierung des geeigneten Personals unterstützen. Zudem müssen die Manager sich selbst mit den Technologien beschäftigen und Kenntnisse über deren Funktionsweise aufweisen. Zur Förderung einer datengetriebenen Unternehmenskultur müssen sie zudem den Ergebnissen der Methoden vertrauen und sich aktiv in die Erarbeitung von Anwendungsfällen einbringen.

Die aufgeführten Stellhebel sind dabei nicht unabhängig voneinander zu betrachten, sondern bilden vielmehr ein Netz von Abhängigkeiten. So ist ein möglicher Ansatzpunkt zur Verbesserung der Gesamtsituation schwierig zu identifizieren. Zweifelsohne wird die Einführung einer datengetriebenen Methode scheitern, wenn diese zwar auf Seiten der Anwender korrekt ein- und umgesetzt wird, deren Ergebnisse jedoch seitens des Managements kein Vertrauen und somit auch nicht den entsprechenden Rückhalt finden. Genauso wenig wird das Management in der Lage sein, datengetriebene Methoden erfolgreich in den Geschäftsprozess zu integrieren, wenn die je nach Position und Funktion notwendigen Fähigkeiten nicht im Repertoire der designierten Anwender zu finden sind. Daneben muss das Management auch die notwendigen Daten im Blick haben. Hier ist es erforderlich, die Methode und deren Datengrundlage rechtzeitig aufeinander abzustimmen. Selbes gilt für die Methodenentwicklung: die Methode muss letztlich nicht nur für sich gesehen, sondern besonders im Kontext der jeweiligen Entwicklungsumgebung funktionieren. Dazu gehört auch die weiterreichende Entscheidung, ob eine Methode an die vorherrschenden Bedingungen angepasst wird oder ob es notwendig ist, den Status Quo in Frage zu stellen, um die Weichen auch für künftige Prozesse zu stellen. Dies betrifft sowohl die Datensicht als auch die personelle Aufstellung der Entwicklungsabteilung. In beiden Fällen ist eine Änderung oder Anpassung meist ad hoc nicht möglich. Dementsprechend ist eine rechtzeitige und wohl überlegte Planung notwendig, um technische Voraussetzungen für ersteres zu schaffen und das Personal für zweiteres entsprechend zu verstärken oder weiterzubilden.

Das klassische Change-Management adressiert zwar einige der genannten Punkte, greift aber spätestens bei den spezifischen, technischen Herausforderungen hinsichtlich der notwendigen Aufbereitung der Daten im Entwicklungsprozess zu kurz. Insofern scheint es zielführend, die einzuführende datengetriebene Methode in den Mittelpunkt der Betrachtung zu rücken (siehe Bild 3) – nicht als gegebener Faktor, sondern vielmehr als zu diskutierende Variable. Ausgehend von der Methode können sodann die entscheidenden Fragen bezüglich ihrer selbst und der vier Rahmenbedingungen formuliert werden. In Anbetracht der komplexen Zusammenhänge bietet sich gerade auch für unerfahrene Unternehmen an dieser Stelle der Einsatz von Checklisten an, um zunächst den eigenen Digitalisierungsfortschritt zu messen, relevante Stellhebel effektiv zu identifizieren und zu adressieren und schlussendlich auch mittel- bis langfristig blinde Flecken zu vermeiden. Hierzu müssen zunächst die fünf Themengebiete samt ihren jeweiligen Einflussfaktoren in messbare Größen heruntergebrochen werden, sodass diese für die Beteiligten greif- und handhabbar werden.

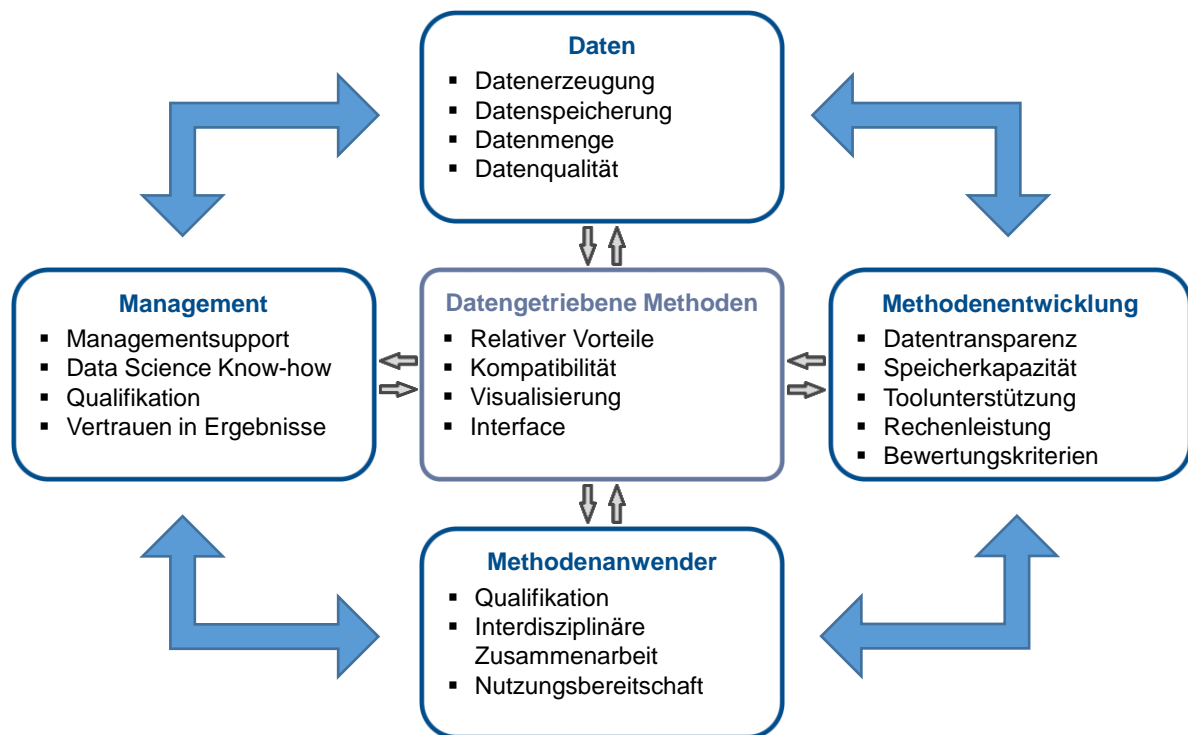


Bild 3: Stellhebel zur Schaffung der erforderlichen Voraussetzungen für den Einsatz datengetriebener Methoden in der Produktentwicklung

6. Zusammenfassung und Ausblick

In dem vorliegenden Beitrag wurden Voraussetzungen, die für die Einführung datengetriebener Methoden in der Produktentwicklung vorliegen müssen, ermittelt. Hierfür wurden zuerst datengetriebene Methoden sowie deren Potentiale und bisherige Anwendungsmöglichkeiten in der Produktentwicklung erläutert. Die Voraussetzungen wurden anhand des Datenanalyseprozesses CRISP-DM abgeleitet. Die identifizierten Voraussetzungen ermöglichen es Unternehmen ein Verständnis für diese zu erlangen und durch einen Abgleich mit den bei ihnen vorherrschenden Bedingungen einen Handlungsbedarf aufzuzeigen. Durch die Ermittlung von Stellhebeln und der Zuordnung der Voraussetzungen zu diesen wurden direkte Anknüpfungspunkte für eine Optimierung offengelegt. Eine Evaluation im industriellen Kontext und anschließende Konkretisierung ausgewählter Bedingungen steht noch aus. Dafür sollen Interviews mit Experten durchgeführt und die Einführung einer datengetriebenen Methode in der Produktentwicklung eines Industrieunternehmens begleitet werden. Daneben sind spezifische Maßnahmen zu entwickeln, um die erforderlichen Voraussetzungen in Unternehmen einzustellen.

Danksagung

Die vorgestellte Forschungsarbeit ist Teil des bayerischen Forschungsverbundes „FOR-CuDE@BEV - Customized Digital Engineering für Bayerische KMU“ und gefördert durch die Bayerische Forschungsförderung (BFS). Für den Inhalt dieser Publikation sind die Autoren verantwortlich. Die Autoren danken der Bayerischen Forschungsförderung (BFS) für die finanzielle Unterstützung.

Literaturverzeichnis

- [1] A. Bertoni, T. Larsson, J. Larsson, and J. Elfsberg, "Mining data to design value: A demonstrator in early design," *Proc. Int. Conf. Eng. Des. ICED*, vol. 7, no. DS87-7, pp. 21–29, 2017.
- [2] C. H. Dagli and H.-C. Lee, "Impacts of data mining technology on product design and process planning,"

- Comput. Appl. Prod. Eng.*, pp. 58–70, 1997, doi: 10.1007/978-0-387-35291-6_6.
- [3] B. Agard and A. Kusiak, "Standardization of components, products and processes with data mining," *Int. Conf. Prod. Res. Am.*, pp. 1–9, 2004.
- [4] A. Bertoni, "Role and Challenges of Data-Driven Design in the Product Innovation Process," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1107–1112, 2018, doi: 10.1016/j.ifacol.2018.08.455.
- [5] S. Jagtap and L. N. K. Duong, "Improving the new product development using big data: a case study of a food company," *Br. Food J.*, vol. 121, no. 11, pp. 2835–2848, 2019, doi: 10.1108/BFJ-02-2019-0097.
- [6] A. Kusiak, M. R. Smith, and Z. Song, "Planning product configurations based on sales data," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 37, no. 4, pp. 602–609, 2007, doi: 10.1109/TSMCC.2007.897503.
- [7] H. H. M. Kim, Y. Liu, C. C. L. Wang, and Y. Wang, "Data-driven design (D3)," *J. Mech. Des.*, vol. 139, no. 11, 2017.
- [8] T. Hendrickx, B. Cule, P. Meysman, S. Naulaerts, K. Laukens, and B. Goethals, "Mining association rules in graphs based on frequent cohesive itemsets," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9078, no. 3, pp. 637–648, 2015, doi: 10.1007/978-3-319-18032-8_50.
- [9] A. Testa and R. Langella, *Interharmonics from a probabilistic perspective*, vol. 3. MIT press, 2005.
- [10] A. Bertoni, "Data-Driven Design in Concept Development: Systematic Review and Missed Opportunities," *Proc. Des. Soc. Des. Conf.*, vol. 1, pp. 101–110, 2020, doi: 10.1017/dsd.2020.4.
- [11] S. S. Shabestari, M. Herzog, and B. Bender, "A survey on the applications of machine learning in the early phases of product development," *Proc. Int. Conf. Eng. Des. ICED*, vol. 2019-Augus, no. 1, pp. 2437–2446, 2019, doi: 10.1017/dsi.2019.250.
- [12] M.-H. Jank, *Produktportfoliosteuerung mittels präskriptiver Datenanalyseverfahren*. Aachen, GERMANY: Apprimus Wissenschaftsverlag, 2021.
- [13] S. S. Anand and A. G. Büchner, "Decision Support Using Data Mining," *Financial Times*. London, 1998.
- [14] P. T. Keenan, J. H. Owen, and K. Schumacher, "Introduction to analytics," *Inform. Anal. Body Knowl.*, pp. 1–29, 2018, doi: 10.1002/9781119505914.ch1.
- [15] C. McCue, *Data mining and predictive analysis: Intelligence gathering and crime analysis*. Butterworth-Heinemann, 2006.
- [16] B. Xu, T. J. Zhao, D. Q. Zheng, and S. Y. Wang, "Product features mining based on Conditional Random Fields model," *2010 Int. Conf. Mach. Learn. Cybern. ICMLC 2010*, vol. 6, no. July, pp. 3353–3357, 2010, doi: 10.1109/ICMLC.2010.5580679.
- [17] Y. Wang and J. Zhang, "Bridging the semantic gap in customer needs elicitation: A machine learning perspective," *Proc. Int. Conf. Eng. Des. ICED*, vol. 4, no. DS87-4, pp. 643–651, 2017.
- [18] D. Horber, B. Schleich, and S. Wartzack, "Conceptual Model for (Semi-) Automated Derivation of Evaluation Criteria in Requirements Modelling," *Proc. Des. Soc. Des. Conf.*, vol. 1, pp. 937–946, 2020, doi: 10.1017/dsd.2020.52.
- [19] P. Kestel, P. Kügler, C. Zirngibl, B. Schleich, and S. Wartzack, "Ontology-based approach for the provision of simulation knowledge acquired by Data and Text Mining processes," *Adv. Eng. Informatics*, vol. 39, pp. 292–305, 2019, doi: <https://doi.org/10.1016/j.aei.2019.02.001>.
- [20] P. Fietkau, S. Sanzenbacher, and B. Kistner, "Advantages of digital vehicle powertrain development in planning of reliability demonstration tests," *Forsch. im Ingenieurwesen/Engineering Res.*, vol. 85, no. 1, pp. 101–113, 2021, doi: 10.1007/s10010-020-00435-w.
- [21] S. Hoornaert, M. Ballings, E. C. Malthouse, and D. Van den Poel, "Identifying New Product Ideas: Waiting for the Wisdom of the Crowd or Screening Ideas in Real Time," *J. Prod. Innov. Manag.*, vol. 34, no. 5, pp. 580–597, 2017, doi: 10.1111/jpim.12396.
- [22] C. Hoyle, W. Chen, N. Wang, and G. Gomez-Levi, "Understanding heterogeneity of human preferences for engineering design," in *DS 58-5: Proceedings of ICED 09, the 17th International Conference on Engineering Design*, 2009, vol. 5, pp. 229–240.
- [23] K. Reed and D. Gillies, "Automatic derivation of design schemata and subsequent generation of designs," *Artif. Intell. Eng. Des. Anal. Manuf. AIEDAM*, vol. 30, no. 4, pp. 367–378, 2016, doi: 10.1017/S0890060416000354.
- [24] C. Sauer, B. Schleich, and S. Wartzack, "Simultaneous Development of a Self-learning Engineering Assistance System," in *Sheet Bulk Metal Forming*, 2021, pp. 127–146.
- [25] B. Kistner, S. Sanzenbacher, P. Fietkau, and J. Munier, *Die Herausforderungen an die digitale Antriebsentwicklung mit Systematik und neuen Möglichkeiten meistern*. Wiesbaden: Springer Vieweg, 2019.
- [26] H. Van Der Auweraer *et al.*, "Virtual Sensing Based on Design Engineering Simulation Models," *ICEDyn2017 Int. Conf. Struct. Eng. Dyn. Ericeira*, no. July, 2017.
- [27] M. Hein, C. Kunze, J.-K. Lee, T. Tobie, and K. Stahl, "Customer-focused modular test procedure for driveline components," *Drivetrain Veh. 2017*, pp. 41–42, 2017, doi: 10.51202/9783181023136-41.
- [28] A. Esser, M. Zeller, S. Foulard, and S. Rinderknecht, "Stochastic Synthesis of Representative and Multidimensional Driving Cycles," *SAE Int. J. Altern. Powertrains*, vol. 7, no. 3, pp. 263–272, 2018, doi: 10.4271/2018-01-0095.
- [29] L. T. M. Blessing and A. Chakrabarti, *DRM, a design research methodology*. Springer, 2009.
- [30] R. Wirth, "CRISP-DM: Towards a Standard Process Model for Data Mining," in *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, 2000, vol. 1, no. 24959, pp. 29–39.
- [31] N. Wirth and R. Wirth, "How to meet the four key challenges to turn Big Data into smart data-driven solutions," *Res. World*, vol. 2017, no. 64, pp. 31–36, 2017, doi: 10.1002/rwm3.20516.

-
- [32] C. Eckert, O. Isaksson, C. Eckert, M. Coeckelbergh, and M. H. Hagström, "Data Fairy in Engineering Land: The Magic of Data Analysis as a Sociotechnical Process in Engineering Companies," *J. Mech. Des.*, vol. 142, no. 12, 2020, doi: 10.1115/1.4047813.
- [33] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big data analytics: Applications, prospects and challenges," *Lect. Notes Data Eng. Commun. Technol.*, vol. 10, pp. 3–20, 2018, doi: 10.1007/978-3-319-67925-9_1.
- [34] J. Linton, *Diffusion of innovations*, vol. 9, no. 4. Simon and Schuster, 1998.
- [35] J. Wilberg, I. Triep, C. Hollauer, and M. Omer, "Big Data in Product Development: Need for a data strategy," in *PICMET 2017 - Portland International Conference on Management of Engineering and Technology: Technology Management for the Interconnected World, Proceedings*, 2017, vol. 2017-Janua, pp. 1–10, doi: 10.23919/PICMET.2017.8125460.
- [36] N. Ammu, M. Irfanuddin, and M. T. Students, "Big Data Challenges 1," vol. 2, no. 1, pp. 613–615, 2013.
- [37] R. V Zicari, "Big Data: Challenges and Opportunities," in *Big data computing*, 2014, pp. 103–128.
- [38] B. Gerschütz *et al.*, "Towards Customized Digital Engineering: Herausforderungen und Potentiale bei der Anpassung von Digital Engineering Methoden für den Produktentwicklungsprozess," 2021.
- [39] J. Al-Jaroodi and N. Mohamed, "Characteristics and requirements of big data analytics applications," *Proc. - 2016 IEEE 2nd Int. Conf. Collab. Internet Comput. IEEE CIC 2016*, pp. 426–432, 2017, doi: 10.1109/CIC.2016.062.
- [40] C. L. Philip Chen and C. Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Inf. Sci. (Ny)*, vol. 275, pp. 314–347, 2014, doi: 10.1016/j.ins.2014.01.015.
- [41] H. Birkhöfer, H. Kloberdanz, B. Berger, and T. Sauer, "Why Methods Don't Work and How To Get Them Work.," *Eng. Des. Integr. Prod. Dev.*, pp. 29–36, 2002.