

Six Sigma und Data Mining wachsen zusammen

Wie steigende Datenmengen und Automatisierung Six Sigma verändern

Data Scientists und Data Engineers verfolgen das Ziel der Prozess- und Produktverbesserung ebenso wie Six Sigma-Belts. Deshalb liegt der Gedanke einer Zusammenarbeit nahe. Eine wichtige Voraussetzung dafür ist eine gemeinsame Sprache. Um dieses Ziel in der Praxis zu unterstützen, hat der ESSC-D-Arbeitskreis „Six Sigma weitergedacht“ in gemeinsamen firmen- und branchenübergreifenden Workshops die Ausbildungsrichtlinien erweitert.

Franziska Schäfer, Thomas Rittler, Erik Schwulera, Christian Glock, Björn Noreik, Roswitha Pietsch

Six Sigma ist branchenübergreifend als erfolgreiche Prozessverbesserungsmethode etabliert. Die European Six Sigma Club (ESSC-D) Community vereint eine breite Wissensbasis und pflegt regen Austausch über die Anwendung in der Praxis. Im geschäftlichen wie im privaten Umfeld wird der Ruf nach digitaler Transformation immer lauter. Mit dem stets steigenden Automatisierungsgrad von Prozessen und Produkten rücken neue Möglichkeiten als Erweiterung für die etablierte Methode in den Fokus. So besteht auf

der einen Seite oft die Möglichkeit, auf eine bereits vorhandene große Datenmenge für die Analyse zuzugreifen. Auf der anderen Seite ist es viel einfacher geworden, relevante Daten aufzunehmen und für eine Analyse vorzubereiten.

Klassisches Six Sigma wird herausgefordert

Die aktuellen Ausbildungsrichtlinien für klassische Six Sigma Green- und Black Belts haben sich längst in der Praxis bewährt und werden akzeptiert. Das Umfeld, in dem die-

se Qualifikationen angewandt werden, ist jedoch von einem starken Wandel geprägt. Die Beherrschung komplexer Interaktionen und Prozesse über Standorte und Firmen hinweg stellt dabei eine Herausforderung dar, die es zu meistern gilt. Hinter diesen End-to-End (E2E) -Prozessen verbergen sich zahlreiche Prozessketten mit einer großen Anzahl möglicher Einfluss- und Störgrößen. Diese müssen oft über mehrere Organisationseinheiten und Organisationen hinweg mit unterschiedlichen Zielen und Zielgrößen orchestriert werden. Dies >>>

erschwert die Beantwortung der Frage nach der Prozesshoheit und auf welcher Basis Entscheidungen getroffen werden. Um hier zahlen-, daten- und faktenbasiert Verbesserungen anstoßen und Entscheidungen treffen zu können, sind Unternehmen auf eine umfangreichere Datenlage und deren strukturierte und zielgerichtete Analyse angewiesen. Die gute Nachricht: Daten stehen meist schon zur Verfügung. Deren Aufbereitung, Analyse und kontextspezifische Verwendung ist allerdings nicht trivial.

Bei genauerer Betrachtung der Datenlage bekommen wir durch die Analyse von E2E-Prozessen nicht nur mehr Daten, sondern müssen auch viel mehr Daten für jeden Kunden, jeden einzelnen Vorgang, jeden einzelnen Prozessschritt etc. berücksichtigen. In der Vergangenheit wurden derartige Daten meist über einen Betrachtungszeitraum von einem Monat oder einer Woche kumuliert betrachtet – oft auch nur über Teilprozesse. Heute ist der Anspruch allerdings, alle Einzeldaten zu messen, vorzuarbeiten, ggf. bereits zu analysieren, direkt auf Dashboards bereitzustellen und den Prozess darauf basierend mit möglichst niedriger Reaktionszeit zu steuern.

Die Komplexität der Prozesse in Kombination mit der hohen Datenverfügbarkeit bringt klassische Methoden an ihre Grenzen. Um diese Daten also zielgerichtet mit Six Sigma nutzen zu können, drängen sich Data Mining-Werkzeuge geradezu auf.

Der Arbeitskreis des ESSC-D sieht daher die Stärken und Kompetenzen des zukunftsorientierten Belts in der Kombination von Six Sigma und Data Mining. Diese Zusatzqualifikationen sind in der ergänzenden Ausbildungsrichtlinie „Data Mining-Aufbaukurs“ beschrieben.

Die neue, ergänzende Ausbildungsrichtlinie des ESSC-D

Ausgangspunkt für die Entwicklung der ergänzenden Ausbildungsrichtlinie ist eine intensive Analyse und Bewertung der Werkzeuge und Methoden für die einzelnen Phasen im DMAIC hinsichtlich vorher festgelegter Szenarien. Diese orientieren sich an den Erfahrungswerten aus der Praxis und werden entsprechend der Anwendungsfälle betrachtet, in denen keine Daten oder auch viele Daten bereits vorhanden sind. Je nach Datenszenario hat sich basierend auf einer Expertenbewertung der

Bedarf nach Modifikationen klassischer oder Hinzufügen neuer Werkzeuge und Methoden ergeben. Diese Erkenntnisse fließen als alte und neue Inhaltsmodule entsprechend der Belt-Rollen in die ergänzende Ausbildungsrichtlinie ein (Bild 1).

Exemplarisch sind in Tabelle 1 die modifizierten und neuen Werkzeuge aufgelistet, die dann auch im Praxisbeispiel Anwendung finden.

„Six Sigma weitergedacht“ in der Praxis

Die neue ergänzende Ausbildungsrichtlinie dient als Basis für die praxisorientierte Anwendung in der sich verändernden Produktionswelt. Dies wird anhand eines generischen Prozessbeispiels, das stellvertre-

tend für ähnliche Prozesse unabhängig von Branche und Unternehmen steht, aufgezeigt. Im Beispiel wird ein End-to-End-Prozess mit zahlreichen, teilweise parallelen Prozessen, die ineinandergreifen, betrachtet (Bild 2). Der Fokus liegt auf einem klassischen Make-to-stock-Prozess (Lagerfertigung) mit der Eingrenzung auf die Liefertreue. Ziel ist die Optimierung der Ausbringung.

Für eine vollumfängliche Prozessverbesserung müssen hier verschiedene Interessen berücksichtigt werden. In der Folge hat die Stakeholderanalyse eine ganz andere Relevanz als in klassischen Six Sigma-Projekten: Im Vergleich zur Vergangenheit ist die Anzahl der Stakeholder stark angestiegen. Hinzu kommt, dass diese meist un-

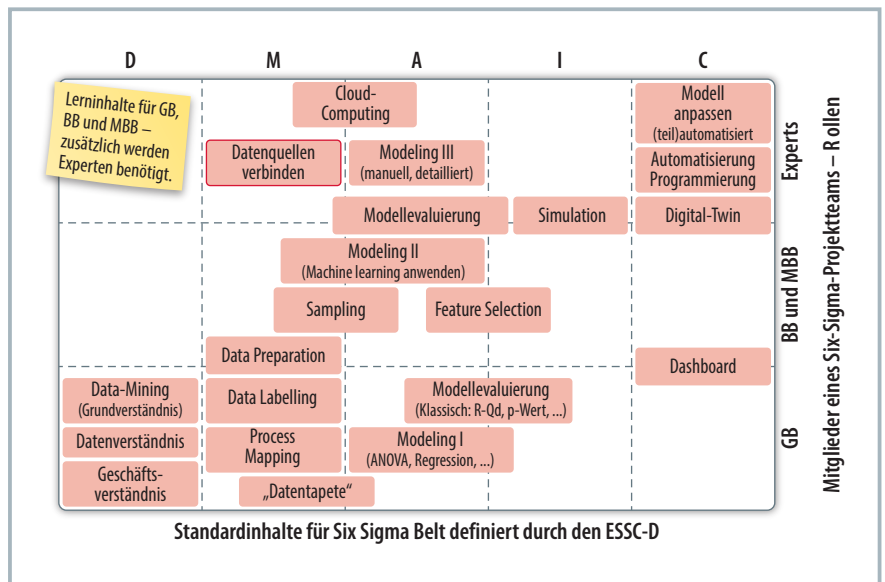


Bild 1. Erweiterung des DMAIC durch Big Data Tools. Quelle: ESSC-D © Hanser

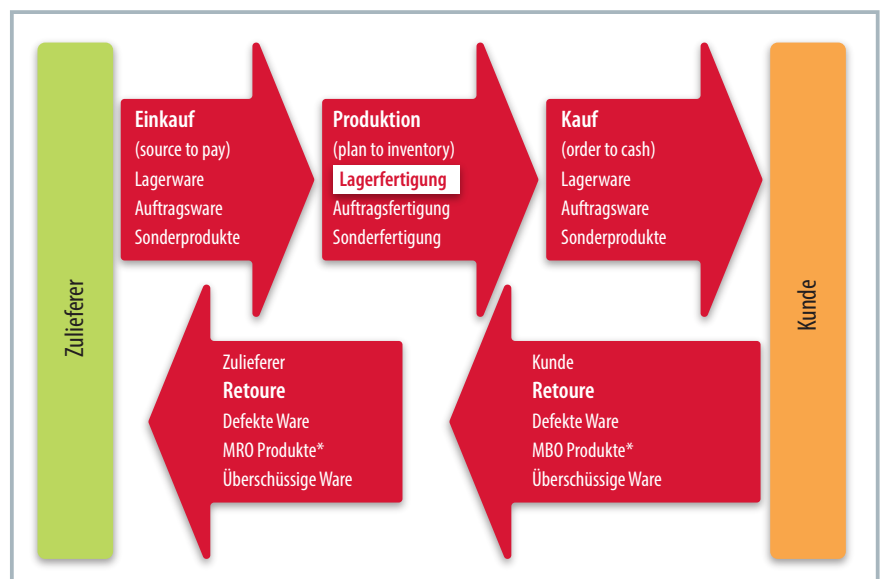


Bild 2. E2E-Prozessmodell aus dem „Supply Chain Operations Modell“. Quelle: ESSC-D © Hanser

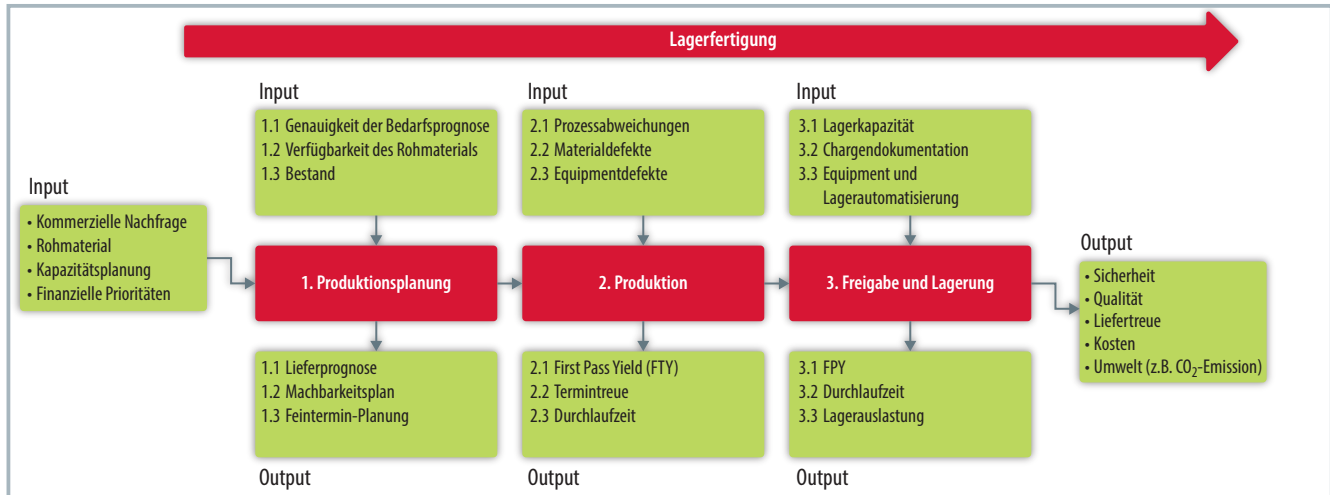


Bild 3. Make-To-Stock-Prozess (Lagerfertigung). Quelle: ESSC-D ©Hanser

verschiedene Ziele verfolgen, welche es in der Praxis abzustimmen und zu gewichten gilt. Für das jeweilige Unternehmen muss der optimale Arbeitspunkt gefunden werden und nicht eine Ansammlung lokaler Maxima.

Bei näherer Betrachtung der Inputs wird schnell klar, dass die Anzahl möglicher relevanter Inputs und Outputs sehr hoch ist (Bild 3). Dies gilt insbesondere, wenn man die Betrachtung auf jeden Einzelprozess erweitert. Um hier den Überblick nicht zu verlieren, ist ergänzend zur klassischen Process Map eine Process Data Map sinnvoll. Diese stellt zusätzlich je Prozessschritt die relevanten Daten, Datenquellen und deren Verfügbarkeit dar. Um daraus praktischen Nutzen ziehen zu können, ist es notwendig, sich mit Themen wie *Extract-Transform-Load (ETL)* oder *Extract-Load-Transform (ELT)*, also

der Speicherung, Orchestrierung und Verarbeitung der Daten zu beschäftigen. Wichtig ist dabei insbesondere die in E2E-Prozessen verorteten Daten mit den relevanten Metadaten zu versehen, um alle Daten in den richtigen, beispielsweise zeitlichen Bezug bringen zu können. Wird dieser Schritt weggelassen, kann das in der Praxis dazu führen, dass wichtige Beziehungen und Wechselwirkungen nicht erkannt werden.

Bei der Auswahl der potenziell relevanten Inputs bietet es sich an, auf Verfahren wie *Feature Selection* zurückzugreifen, um wenig relevante Inputs aus der Betrachtung zu entfernen. Insbesondere für die Entwicklung von Zusammenhangs- und Prognosemodellen ist dies unabdingbar. Wenn man sich auf Data Mining-Modelle verlassen möchte, ist es auf der anderen Seite wichtig,

die Ergebnisse der Modelle zu validieren, um keinen Scheinkorrelationen nachzueifern. An dieser Stelle bietet es sich an, eine bereits klassisch ermittelte *Cause-and-Effect-Matrix* um weitere Spalten zu ergänzen, in denen die von den Modellen ermittelten Gewichtungen eingetragen werden können.

Auf diese Weise werden die umfangreichen Daten und das Expertenwissen der Prozessbeteiligten synergetisch und sich gegenseitig bereichernd kombiniert. So können Widersprüche erkannt, Erfahrungen bestätigt und Ansatzpunkte für weitere Untersuchungen identifiziert werden. Dies bietet Chancen für weitere Verbesserungen bzw. für ein erweitertes Prozessverständnis. Ebenso unterstützt dies bei der Auswahl der relevanten Daten für ein zielgerichtetes Prozessmonitoring und auch zur Steuerung der Prozesse.

Bei klassischem Prozessmonitoring wird die Ausbringung oft über einen Zeitraum von einer Woche oder gar eines Monats kumuliert betrachtet. Bei dieser Betrachtung liegt man „im Mittel“ grob richtig. Anhand einer tagesgenauen Darstellung kann man in diesem Beispiel auf einen Blick erkennen, wie viele Lieferscheine erstellt wurden und ob die Lieferung zum vereinbarten Zeitpunkt erfolgte (Bild 4).

Für die Aufdeckung von konkreten Verbesserungsansätzen ist jedoch eine feingranulare Darstellung notwendig. Es ist folglich sinnvoll, die Darstellung in eine technisch relevante, zeitliche Auflösung zu bringen, damit man Verbesserungsansätze erkennen kann. In Bild 5 wurde für einen Teilprozess aus Bild 4 eine Auflösung »»

Werkzeug / Methode	Status	Bemerkung
Stakeholderanalyse	modifiziert	Höhere Priorität bei der Betrachtung von Wertschöpfungsketten. Eine Gewichtung der Interessen ist sinnvoll.
Process Data Map	neu	Mit der Process Data Map erhält man eine prozessschrittweise Übersicht über Daten, Datenquellen und deren Verfügbarkeit.
C&E-Matrix	modifiziert	Zusätzliche Spalten ermöglichen die Gegenüberstellung des Domänenwissens mit den Einschätzungen aus statistischen Modellen.
Grundlagen ETL / ELT	neu	Grundsätzliche Prinzipien zur Datenspeicherung und Weiterverarbeitung. Orchestrierung von Datenquellen.
Feature Selection	neu	Data-Mining und Machine Learning Ansätze können dabei unterstützen, relevante Größen, besonders aus komplexeren Datensituationen, zu identifizieren.
Prognosemodelle	modifiziert	Mustererkennung, Zusammenhangsanalysen, Entwickeln von Prognosemodellen mit Hilfe moderner Verfahren, wie z.B. CART, Naive Bayes, Random Forest, neuronale Netze.

Tabelle 1. Auswahl neuer und modifizierter Werkzeuge aus der ergänzenden Ausbildungsrichtlinie orientiert am Praxisbeispiels. Quelle: ESSC-D:

von 15 Minuten gewählt und in eine *Heatmap* (grafische Darstellung von Daten mit Farbcodierung) überführt. Je dunkler die Farbe, umso größer ist der *First Pass Yield* (Erstaussbeute). Betriebszeiten bzw. typische Pausenzeiten sind in dieser Darstellung sehr gut zu erkennen.

Auf Basis umfangreicherer Datensätze, die aus der laufenden Produktion kommen und direkt grafisch dargestellt bzw. bereits analysiert werden, können so weitere wichtige Informationen zur Identifikation von Optimierungsmöglichkeiten und zum Treffen von Entscheidungen zur Verbesserung der Ausbringung herangezogen werden. Zusätzlich können diese Daten zur Mustererkennung und Erstellung von Prognosemodellen verwendet werden. So werden Abweichungen schneller festgestellt, die

entsprechenden Stakeholder schneller eingebunden und regulierende Steuerungsaktivitäten schneller und vorbeugend eingeleitet.

An diesem Beispielprozess wird deutlich, dass durch die Einbindung umfangreicher Datenquellen, deren Vorverarbeitung und Nutzung in Kombination mit dem Prozesswissen die Informationslage zur Optimierung der Ausbringung erweitert wird.

Fazit: Das Produktionsumfeld inmitten der digitalen Transformation steht vielen Herausforderungen und täglichen Entscheidungen gegenüber. Es stellt allerdings auch wesentlich mehr Daten zur Verfügung, die in diesen Entscheidungsprozessen unterstützen können. Die Daten erweitern auch die Möglichkeiten der Prozessoptimierung im Rahmen von Six Sigma, indem Prozessabläufe detaillierter gemessen und teilweise auch automatisiert ausgewertet und prognostiziert werden können.

Um dieses Potenzial nutzen zu können, benötigen aktuelle Belts zusätzliches Ex-

pertenwissen und den intensiven Austausch mit den KollegInnen aus dem Bereich des Data Mining. Die ergänzende Ausbildungsrichtlinie „Data Mining-Aufbaukurs“ erweitert und adaptiert die klassische Ausbildung um Data Mining-Aspekte:

- bezüglich der notwendigen Infrastruktur zur Verarbeitung von großen Datenmengen und
- der Informationsgenerierung als Kombination von Six Sigma-Methoden, Data Mining-Werkzeugen und Expertenwissen.

Nun gilt es die adaptierten und neuen Werkzeuge zu etablieren, weiter in der Praxis zu testen und laufend zu verbessern. Denn Six Sigma ist nach wie vor aktuell, wird aber weitergedacht. ■

INFORMATION & SERVICE

QUELLEN

- Köhler, M.; Kurz, G.; David, S.; Rietdorf, T.: Smart Knowledge Picking – Vorsprung durch ausgesuchtes Wissen. QZ 10/2020
- Berg, S.; Köhler, M.; Kurz, G.; Schäfer, F.; Schwulera, E.: Fahrplan in die Zukunft – Six Sigma weitergedacht für Business Excellence; QZ 02/2020
- Glock, C.; Noreik, B.; Pietsch, R.; Rittler, T.: Six Sigma 4.0. Wie verändert sich Six Sigma durch Digitalisierung und KI? QZ 04/2020

AUTOREN

Franziska Schäfer ist cand. Dr.-Ing. im Bereich vorhersagebasierter Prozessverbesserung in der Produktion und Senior Agile Coach bei Siemens Digital Industries.

Thomas Rittler ist Six Sigma Master Black Belt, Reliability Backbelt und als Experte zuständig für statistische Versuchsmethoden und Zuverlässigkeit bei Festo SE & Co. KG.

Erik Schwulera ist Six Sigma Master Black Belt und leitet als Principal Key Expert das Six Sigma Programm bei Siemens Motion Control.

Christian Glock ist Lean Six Sigma Master Black Belt bei CSL-Behring.

Björn Noreik ist selbständiger Lean Six Sigma Master Black Belt und Datenanalyst, BNB-Qualitätsstatistik und Training

Roswitha Pietsch ist Six Sigma Master Black Belt bei G+D Mobile Security.

KONTAKT

Thomas Rittler
thomas.rittler@festo.com

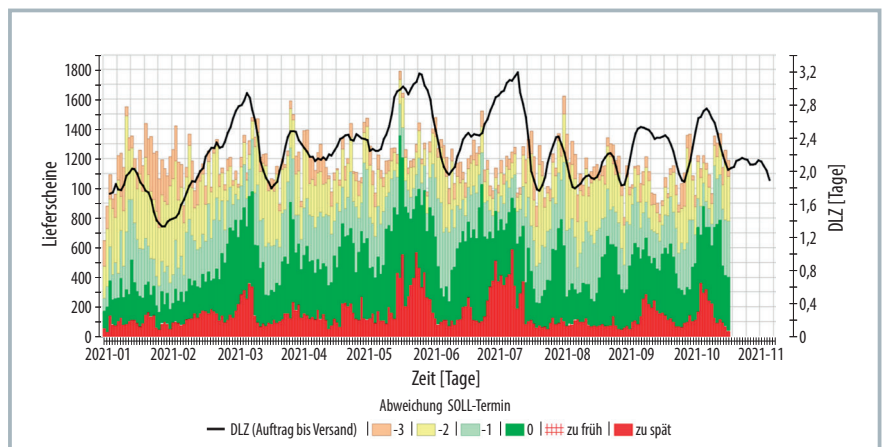


Bild 4. Liefergradtreue Versand. Quelle: ESSC-D © Hanser



Bild 5. Heatmap (Auszug) zur genaueren Darstellung der FPY eines Teilprozesses. Quelle: ESSC-D © Hanser