

Fachwissenschaftliche Projektarbeit
aus dem Gebiet IT in Produktion und Logistik
Prof. Dr.-Ing. Markus Rabe
Fakultät Maschinenbau
TU Dortmund

Systematische Untersuchung des Datenqualitätsbegriffs in Supply Chains

Ausgegeben am: 15.05.2013
Eingereicht am: 14.02.2014

Kandidat:	Daniel Piep
Matrikelnummer:	135988
Studienrichtung:	Wirtschaftsingenieurwesen B.Sc.
1. Prüfer:	Prof. Dr.-Ing. Markus Rabe
2. Prüfer:	Dipl.-Inf. Anne Antonia Scheidler

Inhaltsverzeichnis

1 Einleitung	3
2 Datenqualität	4
2.1 Grundlage: Daten und Informationen.....	4
2.2 Zweckeignung vs. statistisch belegte Quantifizierung	4
2.3 Ansatz der DGIQ.....	5
2.4 Datenqualitätsradar nach Würthele.....	8
2.5 Zwischenfazit	11
3 Datenqualitätsmanagement.....	12
3.1 Grundlage: DQM-Methoden	12
3.2 Lean Six Sigma	13
3.3 Risiko-Assessment	17
3.4 A3 Thinking.....	20
3.5 Zwischenfazit	23
4 Datenqualität im Kontext der Supply Chain	24
4.1 Grundlage: Supply Chain	24
4.2 Diskussion der Datenqualitätsthematik in Supply Chains	25
5 Zusammenfassung.....	26
6 Ausblick.....	27
7 Abbildungsverzeichnis	28
8 Literaturverzeichnis	29

1 Einleitung

Motivation

Zu Beginn des 21. Jahrhunderts sieht sich die weltweite Unternehmenslandschaft mit einer Vielzahl neuartiger Herausforderungen konfrontiert. So hat nicht nur die Digitale Revolution Einzug in sämtliche Wirtschaftssektoren gehalten, sondern ebenfalls die fortschreitende Globalisierung für einen starken Kostendruck durch erhöhte Marktkomplexität und -dynamik gesorgt. Um dieser Problematik entgegenzutreten, wurden Projekte wie bspw. Industrie 4.0 ins Leben gerufen, die auf eine Integration von modernen Informationstechnologien in den Produktionsalltag abzielen.

Obwohl diese Technologien zum jetzigen Zeitpunkt vielversprechend erscheinen, bergen sie doch ein großes Risiko: durch den rasanten Anstieg benötigter und verarbeiteter Datenmengen tragen sie wesentlich zu einer Erhöhung der Komplexität der Unternehmensprozesse bei. Verstärkt wird diese Problematik durch die Interdependenzen verschiedener Marktakteure, wie sie speziell in sogenannten Supply Chains, einer Art Verknüpfung oder Netzwerk verschiedener Unternehmen, auftreten. Besagte Unternehmen sind auf eine hohe Qualität ihrer Informationen und Daten als Basis für wichtige Entscheidungen angewiesen, um unternehmensübergreifende wie auch interne Prozesse effizient zu gestalten. Doch stellen sich an diesem Punkt die Fragen: was macht eine hohe Datenqualität überhaupt aus? Welche Voraussetzungen werden ihr zugrunde gelegt und wie kann man sie erreichen bzw. beibehalten?

Ziele dieser Arbeit und Vorgehensweise

Diese Arbeit verfolgt das Ziel, eine erste Definition des Datenqualitätsbegriffs im Kontext der Supply Chain aufzustellen. Weiterhin sollen verschiedene Methoden für Datenqualitätsmanagement (DQM) erfasst und gegenübergestellt werden. Zu diesem Zweck werden anhand von Fachliteratur zunächst unterschiedliche Datenqualitätsauslegungen gesammelt und verglichen, so dass eine angemessene Definition für die weitere Arbeit zur Verfügung steht. Weiterhin werden auf Basis dieser Definition verschiedene DQM-Methoden erfasst, gegenübergestellt und bewertet. Hierbei stellt nicht die Entwicklung neuartiger Lösungsansätze zum Thema DQM das Ziel dar, sondern vielmehr der Vergleich und die Charakterisierung bereits etablierter Methoden. Es wird somit angestrebt, einen Überblick über bisherige wissenschaftliche Arbeit zu verschaffen und eine kurze Einschätzung derselben zu geben. Im Anschluss daran wird eine gebündelte Definition von Supply Chains erarbeitet, da sich eine Vielzahl von Werken zwar ausführlich mit aktuellen Themen wie Supply Chain Management und Supply Chain Controlling befasst, allerdings keine grundlegende Präzisierung des Begriffs Supply Chain voraussetzt. Letztendlich werden die gewonnenen Erkenntnisse hinsichtlich der Datenqualitätsthematik im Kontext der Supply Chain diskutiert.

2 Datenqualität

2.1 Grundlage: Daten und Informationen

Bevor der Begriff der Datenqualität als Ganzes verstanden und untersucht werden kann, muss zunächst eine Betrachtung der Bestandteile des Wortes, nämlich Daten und Qualität, erfolgen. Betrachtet man in der Informatik gängige Definitionen, so bilden nach [Eng11, S.5] und [Tre06, S.26f.] einzelne Zeichen die Grundlage für Wissensbildung. Fügt man diese in einer bestimmten Syntax zusammen, so erhält man Daten. Diese wiederum werden zu Informationen, sobald man ihnen eine Bedeutung entnehmen kann. Erst wenn man Informationen vernetzt, in einen Kontext setzt und Erfahrungen ergänzt, dann werden diese Informationen zu Wissen. Es fällt also auf, dass die Begriffe Daten und Informationen nah beieinander liegen und nicht losgelöst voneinander betrachtet werden können [Eng11, S.11]. Tatsächlich werden beide in der Literatur oftmals synonym benutzt, ebenso Daten- und Informationsqualität. Trotz des aufgezeigten Unterschiedes zwischen Daten und Informationen ist in der Regel ein und derselbe Sachverhalt gemeint, wenn von Daten- respektive Informationsqualität die Rede ist [Lee01, S.2] [Geb11, S88]. Dies liegt darin begründet, dass die beiden Begriffe eng beieinander liegen und viele Autoren auf eine strikte Unterscheidung verzichten [Tre06, S.34f.]. In der Konsequenz wird festgelegt, dass auch in dieser Arbeit die Begriffe Daten- und Informationsqualität synonym verwendet werden. Andernfalls könnten viele Werke zu diesem Thema strenggenommen nicht beachtet werden, da ihren Ausführungen zur Daten- und Informationsqualität oftmals keine exakte Definition der Begriffe zugrunde liegt.

Was den Qualitätsbegriff angeht, so ist dieser vergleichsweise schnell und allgemeingültig zu erfassen, bspw. durch die vielzitierte Normenreihe ISO 9000 [Koc10, S.104ff]. Nach DIN EN ISO 9005:2005 beschreibt Qualität den „Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt“ [Har10, S.14f.] [Brü12, S.3]. Verknüpft man nun die Begriffe Daten und Qualität, so lässt sich nach [Wür03, S.21] folgende Definition für den Ausdruck Datenqualität aufstellen: „Mehrdimensionales Maß für die Eignung von Daten, den an ihre Erfassung/Generierung gebundenen Zweck zu erfüllen. Diese Eignung kann sich über die Zeit ändern, wenn sich die Bedürfnisse ändern.“. Unklar ist dabei jedoch, nach welchen Dimensionen Datenqualität bemessen wird bzw. welche Anforderungen ihr zugrunde liegen. Die eigentliche Schwierigkeit besteht also nicht im grundsätzlichen Erfassen des Begriffs Datenqualität, sondern vielmehr im Finden und Aufstellen geeigneter Ausprägungen. Die Herausarbeitung besagter Ausprägungen soll nun in diesem Kapitel erfolgen. Festzuhalten ist dabei, dass der Datenqualitätsbegriff nach [Ape09, S.18] stark subjektiv geprägt ist, also unterschiedliche Ausprägungen in der gängigen Literatur ausfindig gemacht werden können.

2.2 Zweckeignung vs. statistisch belegte Quantifizierung

Obwohl in der gängigen Literatur selten eine fundierte Definition für Datenqualität zugrunde gelegt wird, so sind doch im Allgemeinen ähnliche Vorstellungen von ihr vorhanden. So wird häufig von einer Zweckeignung [Mor11, S.16] oder „fitness for use“ gesprochen [Har10, S.18] [Tre06, S.33], korrelierend mit der im vorangegangenen Abschnitt aufgezeigten Definition von Würthele. Solch eine „Kurzdefinition“ lässt sich bspw. in [Cho13, S.501] finden. Hier werden Informationen - man beachte die Analogie zu Daten - als qualitativ hochwertig angesehen, wenn es die richtigen und korrekten Informationen sind, die in angemessener Zeitspanne zugänglich sind und ferner geteilt

werden können. Diese Definition oder vielmehr Beschreibung gibt bereits einen kurzen Einblick in den Datenqualitätsbegriff, allerdings lässt sie einen großen Interpretationsspielraum bzgl. der Auslegung einzelner Qualitätskriterien zu. Das Spektrum des Datenqualitätsbegriffs wird also nicht umfassend und gleichzeitig detailliert genug erfasst und dargestellt.

Einen wichtigen, weil innovativen und grundlegenden Schritt zur Erfassung des Datenqualitätsbegriffs lieferte Richard Y. Wang in den 1980er Jahren. Er gilt als der Begründer des Total Data Quality Management (TDQM) am renommierten Massachusetts Institute of Technology (MIT) [Har10, S.18], womit er eine der ersten Personen war, die sich mit dem Datenqualitätsbegriff im Digitalzeitalter befassten und somit die Bedeutung von Daten und Informationen in der Neuzeit erkannten. Ferner begann er, Daten und Informationen als Produkt anzusehen, was ihnen ebenfalls einen hohen Stellenwert einräumte [Wei11, S.69] [Lee06, S.125f., S.201]. Der Vollständigkeit halber dürfen auch die Namen Tom Redman und Larry English an dieser Stelle nicht unerwähnt bleiben, welche in den 1990er Jahren ebenfalls wichtige Arbeiten zur Thematik veröffentlichten. Allerdings reichen ihre Definitionen der Datenqualität nach Meinung des Autors in Sachen Umfang und Detaillierung nicht an das Werk Wangs heran, weswegen sie im Rahmen dieser Projektarbeit zwar erwähnt, jedoch nicht näher behandelt werden. Wangs Arbeit hingegen beruht auf statistischen Methoden und ist insofern sehr interessant für das Thema. Er führte 1996 eine Umfrage unter 137 IT-Nutzern durch, um herauszufinden, was ihrer Meinung nach Datenqualität auszeichnet [Kas11, S.25] [Tre06, S.19]. Herausgekommen sind 16 Dimensionen in 4 Kategorien, welche bspw. in [Lee01, S.5] und [Tre06, S.35f.] nachzulesen sind. Sie bilden die Grundlage für eine Vielzahl weiterer Arbeiten zum Thema, so auch für die beiden folgenden Werke der DGIQ (siehe Abschnitt 2.3) und von Würthele (Abschnitt 2.4). Allerdings werden sie nicht als Originalquelle in dieser Arbeit untersucht, sondern nur innerhalb der beiden aufbauenden Werke. Dies geschieht einerseits, um eine möglichst aktuelle Definition der Datenqualität zu erhalten, da annähernd 20 Jahre eine große Zeitspanne in der IT-Branche ausmacht. Weiterhin soll untersucht werden, inwiefern sich die originalen Dimensionen von Wang im Laufe der Zeit evtl. geändert haben.

2.3 Ansatz der DGIQ

Einen wichtigen, weil vergleichsweise weit verbreiteten Ansatz zur Erfassung und Quantifizierung des Datenqualitätsbegriffs liefert die Deutsche Gesellschaft für Informations- und Datenqualität e.V. (DGIQ). Die DGIQ hat es sich nach eigenen Angaben zum Ziel gesetzt, alle Aktivitäten zur Verbesserung der Informationsqualität in Gesellschaft, Wirtschaft, Wissenschaft und Verwaltung zu fördern. Essentiell ist dabei der gegenseitige Austausch; so fand im November 2013 bspw. die 11. von der DGIQ organisierte German Information Quality Management Conference (GIQMC) statt. Im 2007 abgeschlossenen Projekt „IQ-Dimensionen definieren“ (IQ für Informationsqualität, Anmerkung des Autors) hat eine Projektgruppe der DGIQ weiterhin 15 Dimensionen der Informations- bzw. Datenqualität definiert. Diese basieren auf der 1996 von Wang durchgeführten Umfrage und werden folgend vorgestellt.

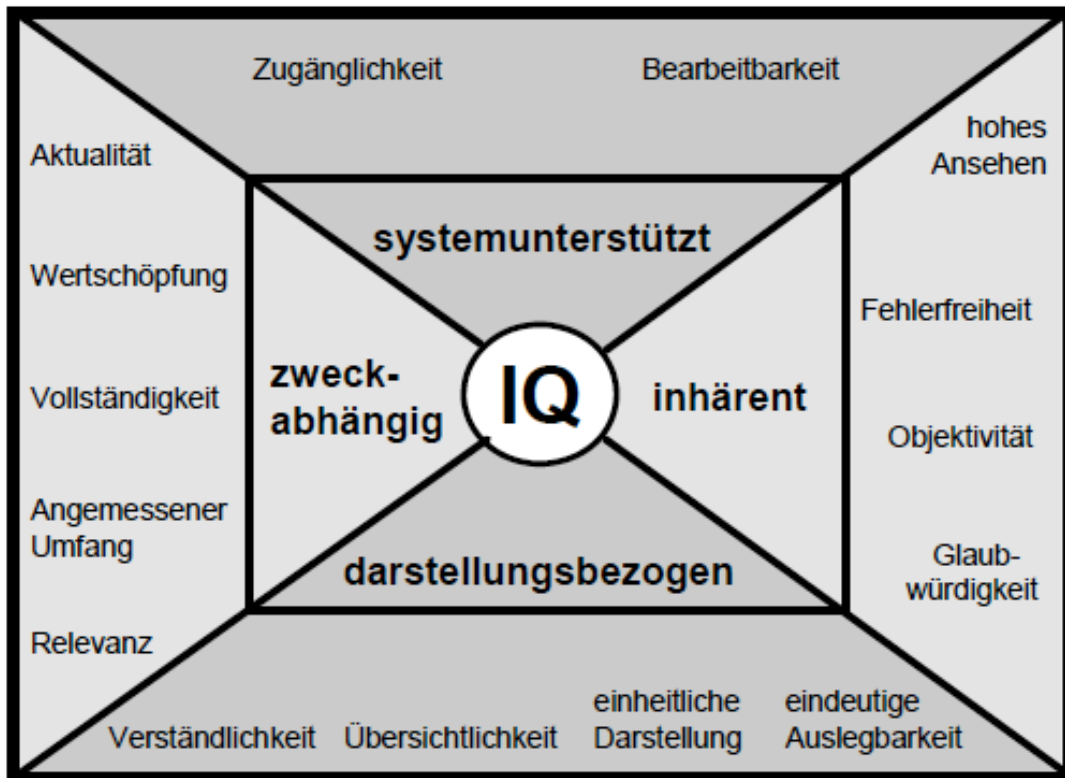


Abb. 1 Dimensionen der Informationsqualität nach der DGIQ [Kas11, S.29]

Die Gruppierung der Dimensionen in Kategorien (systemunterstützt, inhärent, darstellungsbezogen und zweckabhängig) und Untersuchungsgegenstände (System, Inhalt, Darstellung und Nutzung), wie sie in Abbildung 1 dargestellt ist, basiert auf einer Umfrage unter IT-Anwendern, die zu 80% dieselben Dimensionen in Gruppen zusammengefasst haben [Kas11, S.30]. Zu beachten ist weiterhin, dass einige der Begriffe umgangssprachlich einander sehr ähnlich sind und sich somit auf den ersten Blick überschneiden [Kas11, S.27]. Aus diesem Grund ist eine präzise Beschreibung der einzelnen Dimensionen nötig, wie sie in Abbildung 2 dargestellt ist. Außerdem ist anzumerken, dass die für den Anwender wichtigen Aspekte Dubletten und Redundanzen scheinbar nicht in den Dimensionen enthalten sind, nach [Kas11, S.43] jedoch durch die Dimension Wertschöpfung abgedeckt werden. Ebenfalls ist festzustellen, dass die Punkte Sicherheit und Datenschutz nicht enthalten sind. Während Sicherheit als technisches Hilfsmittel angesehen wird, dass zu den Dimensionen führt, wird Datenschutz ausdrücklich nicht als eigene Dimension betrachtet und spielt somit keine Rolle in dieser Darstellung [Kas11, S.43].

Dem Grundgedanken nach sind Informationen brauchbar bzw. qualitativ hochwertig, wenn sie den vom Nutzer beabsichtigten Zweck erfüllen. Auch hier gilt also das angesprochene „fitness for use“-Konzept. Dies ist allerdings erst erfüllt, wenn alle Dimensionen eine hohe oder wenigstens ausreichende Qualität aufweisen. Dabei nimmt keine Dimension einen höheren Stellenwert ein als eine andere. Allerdings steht es dem Nutzer von Informationen frei, die Qualitätsaspekte aus seiner Sicht zu priorisieren – wenn auch ausdrücklich erst, nachdem alle 15 Dimensionen bedacht und gegeneinander abgewogen wurden [Kas11, S.29f.]. Dieser Aspekt wurde bereits in der ursprünglichen Arbeit Wangs herausgestellt [Lee06, S.55].

- 1) **Zugänglichkeit (accessibility):** Informationen sind zugänglich, wenn sie anhand einfacher Verfahren und auf direktem Weg für den Anwender abrufbar sind.
- 2) **Angemessener Umfang (appropriate amount of data):** Informationen sind von angemessenem Umfang, wenn die Menge der verfügbaren Information den gestellten Anforderungen genügt.
- 3) **Glaubwürdigkeit (believability):** Informationen sind glaubwürdig, wenn Zertifikate einen hohen Qualitätsstandard ausweisen oder die Informationsgewinnung und -verbreitung mit hohem Aufwand betrieben werden.
- 4) **Vollständigkeit (completeness):** Informationen sind vollständig, wenn sie nicht fehlen und zu den festgelegten Zeitpunkten in den jeweiligen Prozess-Schritten zur Verfügung stehen.
- 5) **Übersichtlichkeit (concise representation):** Informationen sind übersichtlich, wenn genau die benötigten Informationen in einem passenden und leicht fassbaren Format dargestellt sind.
- 6) **Einheitliche Darstellung (consistent representation):** Informationen sind einheitlich dargestellt, wenn die Informationen fortlaufend auf dieselbe Art und Weise abgebildet werden.
- 7) **Bearbeitbarkeit (ease of manipulation):** Informationen sind leicht bearbeitbar, wenn sie leicht zu ändern und für unterschiedliche Zwecke zu verwenden sind.
- 8) **Fehlerfreiheit (free of error):** Informationen sind fehlerfrei, wenn sie mit der Realität übereinstimmen.
- 9) **Eindeutige Auslegbarkeit (interpretability):** Informationen sind eindeutig auslegbar, wenn sie in gleicher, fachlich korrekter Art und Weise begriffen werden.
- 10) **Objektivität (objectivity):** Informationen sind objektiv, wenn sie streng sachlich und wertfrei sind.
- 11) **Relevanz (relevancy):** Informationen sind relevant, wenn sie für den Anwender notwendige Informationen liefern.
- 12) **Hohes Ansehen (reputation):** Informationen sind hoch angesehen, wenn die Informationsquelle, das Transportmedium und das verarbeitende System im Ruf einer hohen Vertrauenswürdigkeit und Kompetenz stehen.
- 13) **Aktualität (timeliness):** Informationen sind aktuell, wenn sie die tatsächliche Eigenschaft des beschriebenen Objektes zeitnah abbilden.
- 14) **Verständlichkeit (understandability):** Informationen sind verständlich, wenn sie unmittelbar von den Anwendern verstanden und für deren Zwecke eingesetzt werden können.
- 15) **Wertschöpfung (value-added):** Informationen sind wertschöpfend, wenn ihre Nutzung zu einer quantifizierbaren Steigerung einer monetären Zielfunktion führen kann.

Abb. 2 Die 15 IQ-Dimensionen der DGIQ im Detail [Kas11, S.28]

Abschließend ist festzuhalten, dass die von der DGIQ erarbeitete Version der Datenqualitätsausprägungen in großem Ausmaß auf den originalen Dimensionen von Richard Wang basiert, nachzulesen bspw. in [Lee01, S.5]. Es wurden im Wesentlichen einige Formulierungen und Anordnungen der Begriffe auf Grundlage aktueller Untersuchungen angepasst. Auffällige Unterschiede bestehen darin, dass der besprochene Aspekt der Sicherheit (access security im Original) nach Ansicht der DGIQ kein Bestandteil der Dimensionen ist (siehe oben). Ebenfalls wurde der Aspekt der Bearbeitbarkeit („ease of operation“, später „ease of manipulation“) in die Übersicht

aufgenommen. Er wurde zwischenzeitlich von Wang als überflüssig erachtet, mittlerweile aber wieder von ihm aufgenommen, weswegen er auch von der DGIQ berücksichtigt wurde [Kas11, S.27]. Das Ausmaß, zu welchem die Arbeit Wangs von der DGIQ berücksichtigt wurde, zeigt dabei die Bedeutung des in Abschnitt 2.2 beschriebenen statistischen Ansatzes, welchen Wang mit seiner Umfrage 1986 verfolgte.

2.4 Datenqualitätsradar nach Würthele

Eine weitere bemerkenswerte Arbeit zur Aufstellung von Datenqualitätsausprägungen liefert Volker Würthele, von dem bereits eine allgemeine Definition des Datenqualitätsbegriffs in den vorangegangenen Abschnitten aufgezeigt wurde, mit seiner Abhandlung zur Erlangung eines Dokortitels an der ETH Zürich. Würthele fußt seine Arbeit dabei nicht einzig auf Umfragen, sondern auf grundsätzlichen Überlegungen. So unterscheidet er bspw. mehrere Datenarten, wobei dies nicht Bestandteil dieser Projektarbeit sein soll. Es sei jedoch vermerkt, dass sich Würthele in seinem Werk auf nichtflüchtige und gespeicherte Daten bezieht [Wür03, S.19]. Ebenfalls unterscheidet er zwischen Daten- und Informationsqualität, wobei er sich hier ausdrücklich mit Datenqualität beschäftigt und damit relevant für diese Projektarbeit ist.

Bevor nun aber die gewählten Datenqualitätsausprägungen von Würthele präsentiert werden können, müssen zwei seiner grundlegenden Vorüberlegungen behandelt werden: die Aspekte der Modell- und der Zeitabhängigkeit [Wür03, S.26ff]. Er geht davon aus, dass jede Person, also auch jeder Datennutzer ein eigenes Modell wählt, also einen eigenen Blick auf die Realität hat. Konkret bedeutet dies, wie er selbst im Wortlaut sagt: „Bezüglich Datenqualität ist festzuhalten, dass unterschiedliche Sichten auf die Realität auch zu unterschiedlichen Interpretationen von aus der Realität stammenden Daten führen.“ [Wür03, S.27]. Folgerichtig kann Datenqualität also in ihren Ausprägungen nicht ausschließlich auf die ihr zugrunde liegenden Daten reduziert werden, sondern bedarf umfassenderer Betrachtung. Gleiches gilt für den Punkt der Zeitabhängigkeit, da die angesprochene Realität ständigen Änderungen unterliegt. An dieser Stelle sei an den zweiten Teil Würtheles allgemeiner Definition des Datenqualitätsbegriffs erinnert: „Mehrdimensionales Maß für die Eignung von Daten, den an ihre Erfassung/Generierung gebundenen Zweck zu erfüllen. Diese Eignung kann sich über die Zeit ändern, wenn sich die Bedürfnisse ändern.“ [Wür03, S.21]. Des Weiteren versteht Würthele Datenqualität im Rahmen seiner Arbeit immer im Zusammenhang mit Informationsprozessen. Die Prozesssicht hat er gewählt, da die Qualität eines Endproduktes nur dann gewährleistet ist, wenn sie immer alle Teilprozesse hinweg sichergestellt wird. Er verweist hier, ähnlich wie Wang, auf die Analogie zu Produktionsprozessen [Wür03, S.23].

In einem weiteren Schritt hat Würthele die Ergebnisse einer umfassenden Literaturrecherche zusammengefasst, wobei er auch die Arbeit Wangs berücksichtigt. Daraus entstanden ist der Qualitätsbaum für Datenelemente, welcher bei [Wür03, S.24] nachgeschlagen werden kann. Er wird aber nicht eigens in dieser Arbeit präsentiert, da er nur einen Zwischenschritt zum Datenqualitätsradar und nicht die eigentliche Arbeit Würtheles darstellt. Der Qualitätsbaum besteht aus 18 Dimensionen, welche in fünf Kategorien eingeordnet werden. Er ähnelt damit dem Werk Wangs respektive der DGIQ. Allerdings stellt er die Zeitabhängigkeit der einzelnen Kriterien nicht dar und zeigt auch nicht auf, wie sie untereinander korrelieren und insgesamt zur Datenqualität beitragen.

Aus diesem Grund wurde das Datenqualitätsradar entwickelt [Wür03, S.25], welches nachfolgend in Abbildung 3, zusammen mit einer Erklärung der einzelnen Bestandteile in Abbildung 4 (analog zu 2.3), präsentiert wird.

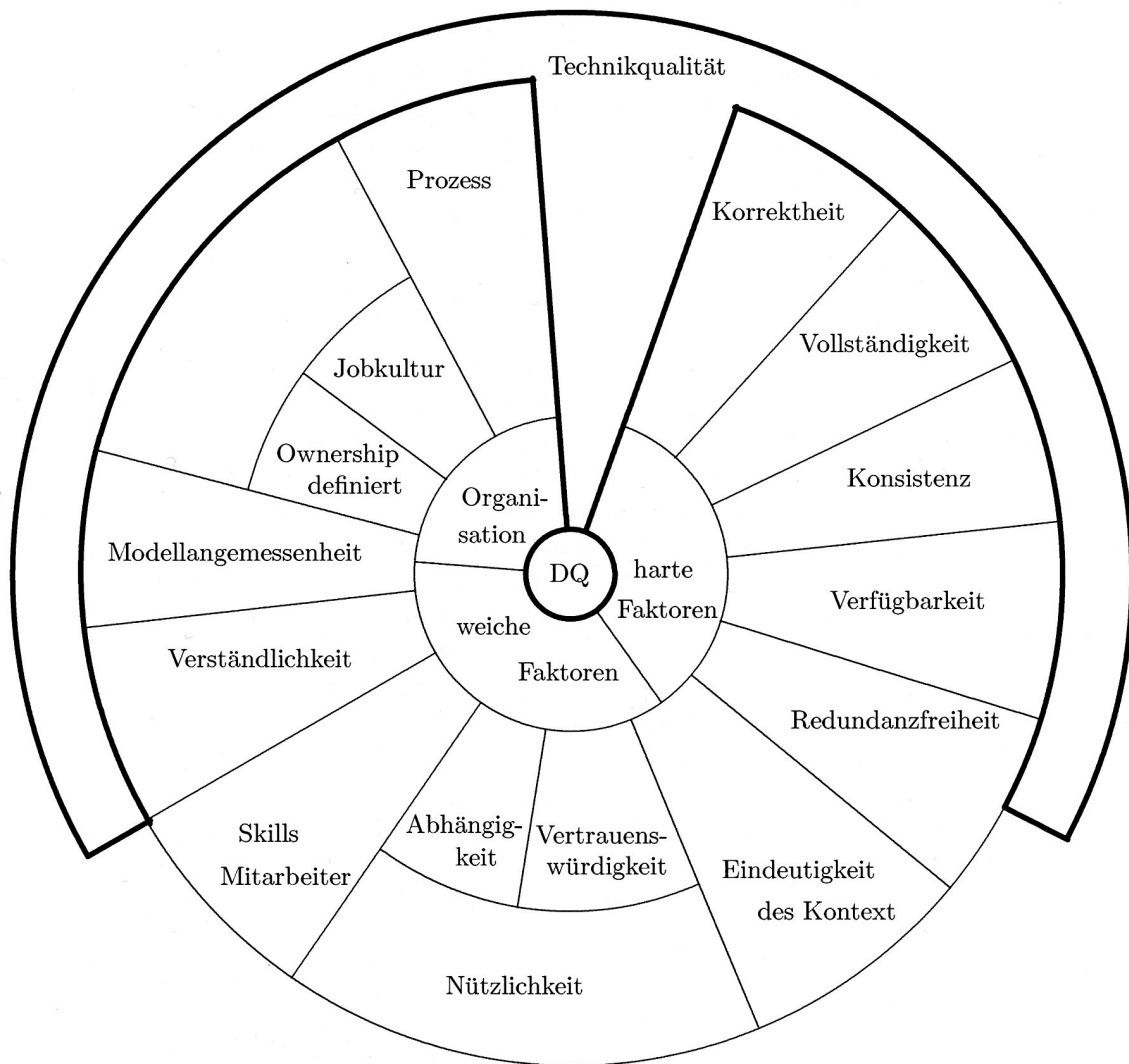


Abb. 3 Datenqualitätsradar nach Würthele [Wür03, S.31]

DQ	Datenqualität
harte Faktoren	Faktoren, die sich durch exakte Regeln prüfen und messen lassen.
weiche Faktoren	Faktoren, deren Bewertung nicht objektiv möglich ist.
Organisation	Faktoren, die stark von Aufbau- und Ablauforganisation sowie von der konkreten organisatorischen Ausgestaltung des Arbeitslebens im jeweiligen Betrieb abhängig sind.
Korrektheit	Gibt an, ob eine Wertausprägung richtig ist.
Vollständigkeit	Gibt an, ob zu <i>allen</i> relevanten Datenelementen vom Leerwert verschiedene Wertausprägungen vorhanden sind.
Konsistenz	Gibt an, ob mehrere für sich betrachtet brauchbare Datenelemente auch <i>im Zusammenhang</i> korrekt sind.
Verfügbarkeit	Gibt an, ob Daten innerhalb geforderter Zeit zur Verfügung stehen.
Redundanzfreiheit	Gibt an, in welchem Masse ein System frei von Redundanzen ist. Redundanzen sind eine Quelle für Inkonsistenz.
Eindeutigkeit des Kontexts	Korrektheit von Datenelementen alleine genügt nicht – sie müssen auch im richtigen Kontext und kontextgerecht angezeigt werden.
Vertrauenswürdigkeit	Subjektives Mass für die Glaubwürdigkeit eines (Informatik-)Systems oder einzelner oder mehrerer Datenelemente daraus.
Abhängigkeit	Mass für die Notwendigkeit, dass einzelne oder mehrere Datenelemente für eine Aufgabe / einen Prozess benötigt werden.
Nützlichkeit	Kombination von Abhängigkeit und Vertrauenswürdigkeit.
Skills Mitarbeiter	Eignung des Fach- und Methodenwissens der Mitarbeiter ³⁵ gemessen an den Anforderungen an einen <i>Standard-Benutzer</i> .
Verständlichkeit	Eignung der Datenpräsentation für Standard-Benutzer. (Missverständene Daten haben die selbe <i>Wirkung</i> wie falsche Daten)
Modellangemessenheit	Mass der Harmonie zwischen Geschäftsprozess-Modell, Modell des Benutzers und Modell des Informatiksystems.
Ownership definiert	Gibt an, ob zu jedem Datenelement ein eindeutiger <i>Owner</i> („Besitzer“) definiert und verantwortlich ist.
Jobkultur	Gibt an, ob alle Mitarbeiter für die Belange der Datenqualität sensibilisiert sind und die negativen Folgen einer schlechten Datenqualität in ihrem Umfeld kennen und ob sie „Qualität leben“.
Prozess	Gibt an, ob klar geregelte Prozesse eine geordnete Bewirtschaftung von Daten ermöglichen.
Technikqualität	Gibt an, wie gut die gewählte technische Lösung (etwa ein Informatiksystem) einschliesslich der definierten ablauforganisatorischen Prozesse in der Lage ist, eine gute Datenqualität zu gewährleisten (vgl. Abschnitt 3.4).

Abb. 4 Die Dimensionen des Datenqualitätsradars nach Würthele [Wür03, S.32]

Offensichtlich ist das Datenqualitätsradar nicht auf den ersten Blick so leicht zu lesen wie die Grafik der DGIQ. Folgende Erklärungen sind im Wortlaut [Wür03, S.30] zu entnehmen:

- jede weiße Fläche entspricht einem Qualitätsmerkmal
- die direkte Nachbarschaft zweier Flächen bedeutet einen engen Zusammenhang der entsprechenden Merkmale
- im Zentrum des Radars steht die „Datenqualität als Ganzes“ (DQ), die nach außen hin immer stärker aufgefächert wird
- außen steht die „Technikqualität“, welche zwei Drittel der Ringsektoren umfasst und deren Spitze bis zum Zentrum der Datenqualität reicht
- die Technikqualität berührt ausdrücklich nicht alle Qualitätsmerkmale
- einige Nachbarschaften sind bewusst durch ein „Loch“ ausgeschlossen
- das Radar liefert eine momentane Bewertung der Datenqualität zu einem bestimmten Zeitpunkt t_x

Zu beachten ist hier die Unterteilung der Datenqualität in verschiedene Faktoren. Die Anforderungen an Daten selbst sind nicht so detailliert aufgeschlüsselt wie bei Wang bzw. der DGIQ, dennoch sind durch die harten Faktoren die wichtigsten Dimensionen abgedeckt, welche andere, in Kapitel 2.3 aufgezeigte teilweise einschließen können. Auch [Hei11, S.49], eine DGIQ-Quelle, benennt diese Kriterien als die wichtigsten Faktoren. Zusätzlich sind durch die weichen Faktoren und die Punkte Organisation und Technikqualität viele wichtige Aspekte bedacht worden, die das Datenqualitätsradar zu einem ganzheitlichen Instrument machen. Besonders hervorzuheben ist die Technikqualität. So wird in [Wür03, S35] explizit auf ihre Bedeutung hingewiesen, indem der Vergleich zum Eisbergmodell gezogen wird. Weist ein Informatiksystem keinerlei Schutzmechanismen oder dergleichen auf, wird aber gut gepflegt, so ist dieser Pflegezustand in Form einer guten Datenqualität als Spitze des Eisbergs sichtbar. Unter der bildlich gesprochenen Wasseroberfläche liegt aber nicht sichtbar die Basis, nämlich die in diesem Fall schlechte Technikqualität. Sie birgt ein großes Risiko für den Fall, dass eine entsprechende Pflege des Systems nicht mehr erfolgen kann. Würthele zeigt an dieser Stelle also deutlich auf, dass Datenqualität nicht nur in Hinblick auf die ihr zugrunde liegenden Daten betrachtet und beurteilt werden darf, sondern weitergehender Überlegungen bedarf.

2.5 Zwischenfazit

Im Laufe dieses Kapitel wurde zunächst auf Daten und Informationen eingegangen und festgelegt, dass aufgrund von Überschneidungen in der Fachliteratur die Begriff Daten- und Informationsqualität zur Erfassung des Qualitätsbegriffs synonym verwendet werden können. Weiterhin wurde eine allgemeine Definition für Datenqualität nach Würthele präsentiert, aus der ersichtlich wurde, dass die Schwierigkeit bei einer angemessenen Definition des Datenqualitätsbegriffs im Wesentlichen die Bestimmung seiner Ausprägungen ist. Zu diesem Zweck wurden einige grobe Definitionsmöglichkeiten aufgezeigt und anschließend die beiden umfassenden Werke der DGIQ und von Volker Würthele zu dieser Thematik vorgestellt, welche beide zu einem, wenn auch unterschiedlich großem, Teil auf der Arbeit von Richard Wang basieren. Nun gilt es, sich auf eine der beiden Definitionen von Datenqualitätsausprägungen festzulegen, die als Grundlage für den weiteren Verlauf dieser Projektarbeit dienen soll.

Betrachtet man die Version der DGIQ, welche quasi das Werk Wangs präsentiert, so fällt sofort die starke Auffächerung des Datenqualitätsbegriffs auf. Es werden unterschiedliche Kategorien betrachtet, zu denen auch das System im weiteren Sinne zählt. Allerdings beziehen sich sämtliche Aspekte auf die Eigenschaften von Daten bzw. Informationen. Wie am Ende des Abschnitts 2.4 erklärt wird, zeigt Volker Würthele jedoch deutlich auf, dass der Begriff Datenqualität wesentlich komplexer ist, als dass er nur auf die Eigenschaften von Daten reduziert werden dürfe. Das Datenqualitätsradar von Würthele berücksichtigt diese Komplexität, indem es eine Vielzahl unterschiedlicher Einflussfaktoren auf Datenqualität darstellt und somit die Ganzheitlichkeit des Datenqualitätsbegriffs aufzeigt. Aus diesem Grund soll die Arbeit Würtheles als Basis für die weiteren Abschnitte dieser Abhandlung dienen.

Bevor nun dieses Kapitel abgeschlossen werden kann, muss explizit darauf hingewiesen werden, dass die Thematik der Datenqualität hier nicht ganzheitlich betrachtet wurde und wird. Ein elementarer Baustein zur kompletten Betrachtung sind

Datenqualitätsmetriken, welche eine Messung und damit Quantifizierung der Datenqualität erlauben [Hei11, S.49]. Ohne entsprechende Messungen lässt sich nach [Mor11, S.22] und [Tre11, S.52] keine adäquate Aussage über Datenqualität und ihre entsprechenden Auswirkungen treffen. Nach [Wür03, S.13] sollte es sich dabei genauer gesagt um ein Metriksystem handeln, welches einerseits harte und messbare, andererseits aber auch bloß schätzbare Kriterien berücksichtigt. In [Wür03] und [Hil11] sind solche Ansätze zur Erstellung entsprechender Metriken beschrieben. Auch die DGIQ arbeitet zum Zeitpunkt der Erstellung dieser Arbeit an ihrem Projekt „IQ-Dimensionen bewerten“, welches sich ebenfalls mit einer Messbarkeit der Datenqualitätsdimensionen beschäftigt. Hierdurch wird abermals aufgezeigt, dass Datenqualität eine komplexe Thematik darstellt, für die weiterhin umfassender Forschungsbedarf besteht.

3 Datenqualitätsmanagement

3.1 Grundlage: DQM-Methoden

Nachdem auf den vergangenen Seiten eine angemessene Definition des Datenqualitätsbegriffs gefunden wurde, soll es in diesem Kapitel nun darum gehen, geeignete Methoden für das Datenqualitätsmanagement zu betrachten. Dabei soll der aufgezeigte Datenqualitätsbegriff nach Würthele nicht als Basis für die beschriebenen Methoden dienen, sondern vom Leser lediglich im Hinterkopf behalten werden, um sich der Komplexität der Thematik bewusst zu sein.

Bevor nun aber entsprechende Methoden für das Datenqualitätsmanagement untersucht werden, muss zunächst einmal verdeutlicht werden, was darunter zu verstehen ist, angefangen mit Datenqualitätsmanagement. Anders als beim Begriff der Datenqualität soll hier keine exakte und aufgeschlüsselte Definition erfolgen, sondern der Ausdruck lediglich grob umrissen werden. Dies ist darin begründet, dass das Verständnis von Datenqualitätsmanagement in der Literatur nach Meinung des Autors starken Abweichungen unterliegt. Es mangelt hier also nicht an einheitlicher Definition, sondern vielmehr am allgemeinen Verständnis des Begriffs. So wird mitunter bereits die manuelle Bereinigung von Daten (Löschen von Dubletten etc.) als Datenqualitätsmanagement verstanden, wobei derartige Maßnahmen reaktiv, also nachträglich sind [Zwi11, S.102f.]. Hier soll es jedoch um ein ganzheitliches, proaktives Datenqualitätsmanagement gehen, welches nach [Zwi11, S.109] einzig als echtes Datenqualitätsmanagement bezeichnet werden kann und durch folgende Merkmale gekennzeichnet wird:

- Beseitigung der Fehlerquellen
- Fortlaufende Überwachung der Datenqualität
- Präventive Verhinderung neuer Datenfehler im laufenden Betrieb
- Regelmäßige Bereinigung neu entstandener Datenfehler

Auszeichnend für ein Datenqualitätsmanagement sind also steuernde und damit proaktive Aufgaben [Lüs11, S.232] [Koc11, S.24] [Brü12, S.122]. Offensichtlich stellt ein ganzheitliches Datenqualitätsmanagement also ein großes Unterfangen für ein Unternehmen dar, da durch präventive Maßnahmen im laufenden Betrieb auch Geschäftsprozesse tangiert werden. Für umfangreiche Veränderungen im Unternehmen empfiehlt sich nach [Wol11, S.257] die Initiierung durch Projekte. Ein solcher Projektablauf könnte nach [Wei11, S.71-85] folgendermaßen aussehen:

- 1) Initiierung des Datenqualitäts-Projekts
- 2) Definition der Datenqualitäts-Anforderungen
- 3) Messung der vorhandenen Datenqualität
- 4) Analyse der Fehlerursachen
- 5) Verbesserung der Datenqualität
- 6) Permanente Überwachung der Datenqualität

Diese Projektarbeit setzt an den Punkten 4-6 an, wobei insbesondere Punkt 5 interessant ist, welcher ebenfalls die Geschäftsprozessoptimierung enthalten kann. Ferner zeigt er auf, was alle Definitionen respektive Ansichten von Datenqualitätsmanagement gemeinsam zu haben scheinen: eine Verbesserung der Datenqualität.

Die in dieser Arbeit behandelten Methoden sind dabei nicht als eigenständiges und ganzheitliches Datenqualitätsmanagement zu verstehen, sondern unterstützen dieses wie bereits angedeutet. Ferner ist anzumerken, dass es sich um Methoden im weiteren Sinne handelt. Streng genommen werden Systeme wie das folgend behandelte Six Sigma gerne als eigenständige Methoden gesehen, wobei sie selbst eine Vielzahl verschiedener (bspw. Analyse-) Methoden enthalten und damit eigentlich eine Überordnung einnehmen. Für die folgenden Methoden wäre also eigentlich der Begriff Managementsystem oder teilweise auch –philosophie treffender. Sie stellen keine gesonderten Lösungen für die Datenqualitätsproblematik dar, sondern sind in der Regel vielseitig anwendbar. Dies basiert auf der Aussage von [Wei11, S.68f.], nach welcher die Behandlung von Daten und Informationen analog zu Produkten vorzunehmen ist. Es sei hier an Richard Wang erinnert, nach welchem die Nutzung bereits etablierter Methoden des Qualitätsmanagements aus der Fertigungsindustrie zu bevorzugen ist. Es werden im Rahmen dieser Arbeit also keine neuen, aber ggf. für die Problematik des Datenqualitätsmanagements neue respektive selten betrachtete Methoden aufgezeigt. Dies geschieht u.a. aus dem Grund, abermals die Komplexität der Thematik und daher die Notwendigkeit kreativer Herangehensweisen darzustellen.

3.2 Lean Six Sigma

Die erste der aufzuzeigenden Methoden stellt Lean Six Sigma dar. Dieses stellt eine Kombination der bekannten Konzepte Lean Management und Six Sigma dar, wobei letzteres explizit im Zusammenhang mit Datenqualitätsmanagement erwähnt wird [Wei11; S.69] und auch im Kontext des allgemeinen, unternehmensweiten Qualitätsmanagements häufig genannt wird [Koc11, S.183]. Außerdem fließen Elemente des Design for Six Sigma (DFSS) ein, wobei diese eher von geringerer Bedeutung für das Gesamtkonzept sind und daher nicht eigenständig behandelt werden. Das Ziel dieses Gesamtkonzepts ist nicht nur die Kombination der Vorteile besagter Konzepte [Gam13, S.273f.], sondern die Schaffung von Synergien [Töp09, S.8f.]. Im Kern handelt es sich dabei um eine Optimierung des betriebswirtschaftlichen Dreiecks Qualität – Zeit – Kosten [Töp09, S.3f.]. Es muss jedoch an dieser Stelle gesagt werden, dass Lean Management und Six Sigma jeweils umfassende und eigenständige Konzepte darstellen, die hier nicht in dem ihnen gebührenden Umfang behandelt werden können. Gleiches gilt für die folgenden Methoden, so dass dieses Kapitel als kurze Vorstellung inkl. Einschätzung besagter Konzepte zu verstehen ist.

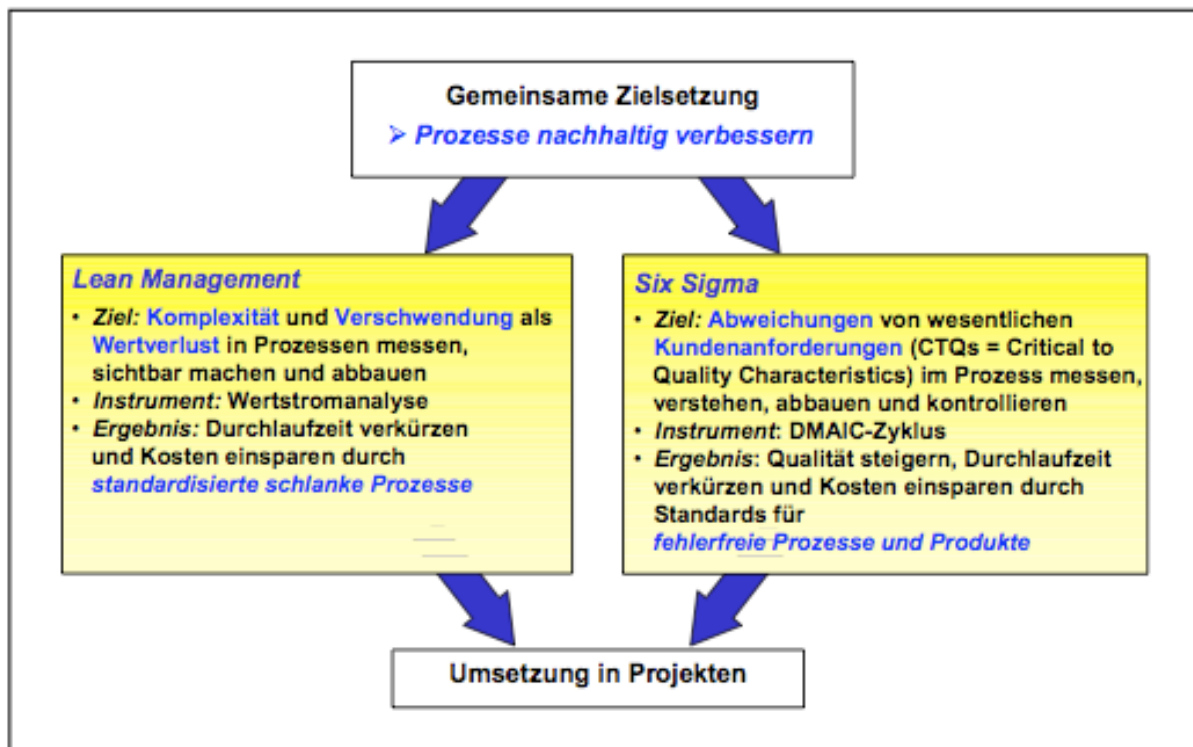


Abb. 5 Gleiches Ziel, unterschiedlicher Ansatz von Lean Management und Six Sigma [Töp09, S.8]

Abbildung 5 zeigt auf einen Blick, dass Lean Management und Six Sigma prinzipiell den gleichen Zielvorgaben folgen, dies jedoch auf unterschiedliche Art und Weise tun. Dadurch, dass beide die Prozesse nachhaltig verbessern, kann die Qualität, also auch Datenqualität, stark verbessert werden. Six Sigma ist direkt auf eine Qualitätssteigerung ausgerichtet, Lean Management bietet durch Verschlankeung der Prozesse das Potential für eine geringere Fehleranfälligkeit und damit indirekte Qualitätssteigerung [Töp09, S.5].

Lean Management

Das Lean-Konzept wurde in den 1950er Jahren in Japan entwickelt, wo es seitdem bei Toyota starke Anwendung findet. Lean Management, auch als Lean Production bezeichnet, beschreibt nach [Hum13, S.40] ein „Organisations- und Produktionsmodell, das das Management im Hinblick auf Ziele wie Qualität, Produktivität, Flexibilität und Mitarbeitermotivation unterstützt“. Es bildet nach Expertenmeinung die Grundlage für Toyotas Erfolg als Automobilhersteller. Kerngedanke des Konzepts ist die Schaffung schlanker Prozesse durch Vermeidung von Verschwendung, welche Toyota nach sieben Arten unterscheidet: Überproduktion, Wartezeit, Prozessübererfüllung, Transport, Nacharbeit, Bestand und Bewegung [Töp09, S.28] [Koc11, S.127ff]. Es wird also angestrebt, Verschwendungskosten respektive Blindleistungen zu vermeiden [Töp09, S.28f.]. Dies geschieht durch eine Fokussierung auf für den Kunden relevante, also wertschöpfende Prozesse und Beseitigung aller nicht-essentiellen Prozessschritte [Gor13, S.153ff].

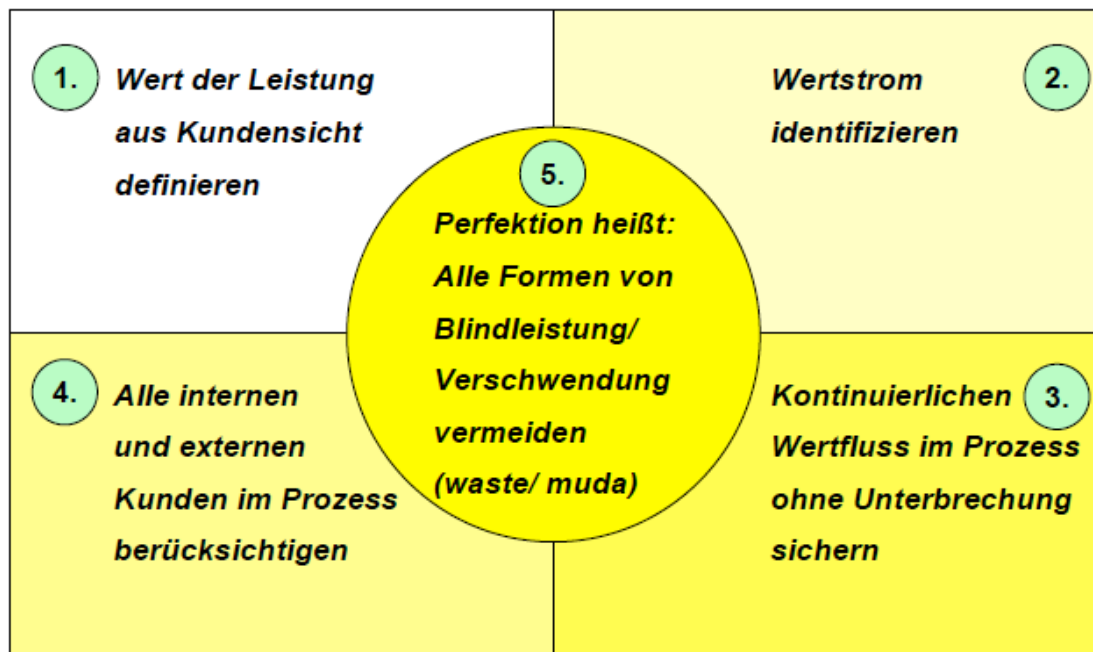


Abb. 6 Grundprinzipien des Lean Thinking [Töp09, S.31]

Lean Management ist, wie vorab bereits erwähnt, nicht als Maßnahme, sondern vielmehr als Konzept zu verstehen [Koc11, S.126]. Grafik 6 zeigt die fünf Prinzipien des Lean Thinking und verdeutlicht damit, dass Lean Management ganzheitlich ansetzt und sämtliche Unternehmensbereiche durchzieht [Töp09, S.42]. Ebenfalls wird ein elementarer Aspekt des Lean Managements beschrieben, nämlich die Betrachtung und Analyse des Wertstroms. Dieses Werkzeug trägt massiv zur Stärke von Lean Management bei: in jedem Prozess Verschwendung erkennen und beseitigen, das gesamte Unternehmen also schlanker und reaktionsschneller zu machen. Qualität wird folglich im Zuge der Prozessoptimierung erzeugt und gewährleistet. Treten allerdings schwerwiegende Abweichungen von im Rahmen des Lean Managements formulierten Standards auf, so ist das Konzept nicht in der Lage, diesen wirkungsvoll zu begegnen und Qualität weiterhin aufrecht zu erhalten. An dieser Stelle setzt Six Sigma an [Töp09, S.43].

Six Sigma

Anders als Lean Management ist Six Sigma ein selektiv, also gezielt ansetzendes Konzept zur Qualitätsverbesserung von wertschöpfenden Prozessen [Koc11, S.117]. Es zeichnet sich durch die Durchführung in klar definierten Projekten von üblicherweise 90 bis 180 Tagen Dauer und eine Zielstruktur aus, bei der finanzielle Ergebnisse, also die Bemessung von Veränderungen in monetären Maßen im Vordergrund steht [Töp09, S.43]. Namensgebend ist dabei der Sigma-Wert, welcher eine Messgröße für die tolerierte Abweichung auf Basis der Gauß'schen Normalverteilung und ihrer Standardabweichung darstellt [Koc10, S.106f.]. 6σ entspricht dabei einem Qualitätsniveau von 99,99966% und einer praktikablen Null-Fehler-Qualität (100% und damit absolute Perfektion sind per se auszuschließen), dem angestrebten Qualitätsziel bei Six Sigma [Koc11, S.148ff.]. Zum Vergleich: das durchschnittliche Qualitätsniveau in der deutschen Industrie wird mit $3,8\sigma$ angegeben, entsprechend 99% und einer Fehlerrate von 10.000 parts per million (ppm) [Töp09, S.10f.]. Six Sigma findet starke Anwendung im amerikanischen Raum, speziell bei General Electrics.

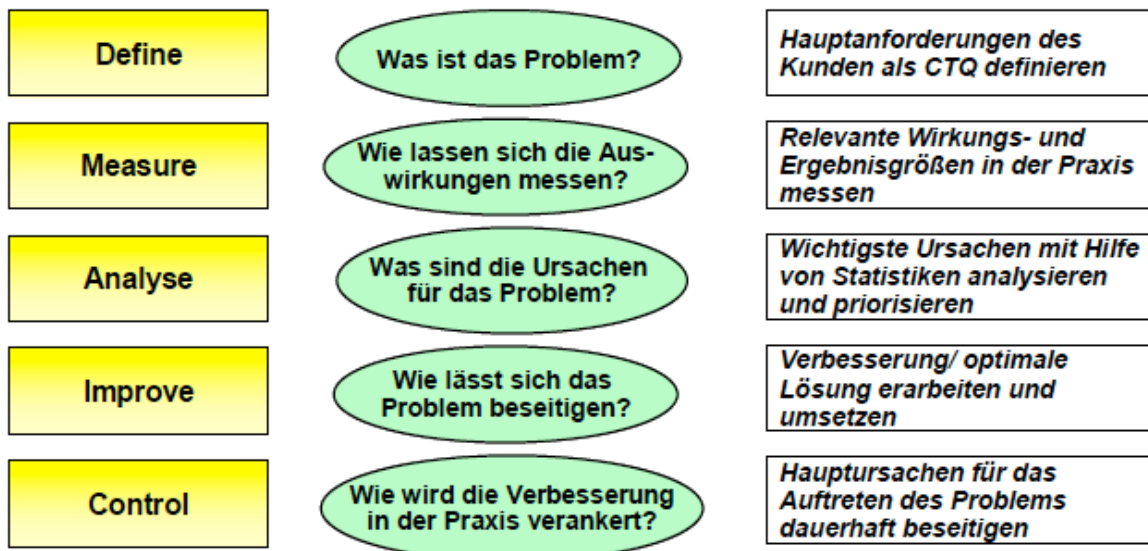


Abb. 7 Der DMAIC-Zyklus in Six Sigma [Töp09, S.45]

Die Six-Sigma-Philosophie besteht nach [Töp09, S.45] darin, „durch zielgerichtete Übersetzung der ‘Stimme des Kunden’ in die ‘Sprache der Prozesse’ Produkte und Dienstleistungen mit hoher Qualität zu erzeugen und so Wirtschaftlichkeit, also Effizienz, mit Kundenzufriedenheit, also Effektivität, zu verbinden.“. Der Grundgedanke ist dabei, dass qualitativ hochwertige Prozesse qualitativ hochwertige Produkte erzeugen [Koc11, S.131] [Kos13, S.137] [Gam13, S.249]. Zu diesem Zweck bedient sich Six Sigma einem elementaren und weit verbreiteten Werkzeug: dem DMAIC-Zyklus (siehe Abbildung 7). Analog zum PDCA-Zyklus (plan – do – check – act) steht der DMAIC-Zyklus (define – measure – analyse – improve – control) für wissenschaftliche und analytische Herangehensweisen [Koc11, S.118f.]. Aus diesem Grund eignet sich Six Sigma für die gezielte Bewältigung komplexer Probleme. Damit sind im Wesentlichen die Unterschiede zu Lean Management genannt: Six Sigma greift punktuell und konzentriert unter Einbeziehung weniger Mitarbeiter des Unternehmens dort an, wo Lean Management als unternehmensweiter Standard nicht wirken kann [Töp09, S.54f.]. Zur Ergänzung von Lean Management scheint es also durchaus geeignet zu sein.

Lean Six Sigma

Nach [Töp09, S.57] gibt es nun drei Möglichkeiten, Lean Management und Six Sigma zu kombinieren. Im Wortlaut:

1. Die Strategie, den aufwändigen Six-Sigma-Projektmanagementprozess mit dem Gedankengut und den Methoden des Lean Managements schlanker zu machen
2. Lean Management gezielt durch Six-Sigma-Projekte zu ergänzen
3. Beide Konzepte ganzheitlich in zweckmäßiger Bearbeitungsfolge und in leistungsfähigem Ausmaß inhaltlich untereinander zu vernetzen.

Hervorzuheben ist nach Meinung des Autors dabei die dritte Variante, da nur sie den Aussagen von [Töp09, S.58ff] folgend in der Lage ist, beide Konzepte sinnvoll und synergienschaffend zu verbinden. Abbildung 8 verdeutlicht dabei die Art und Weise der Verknüpfung.

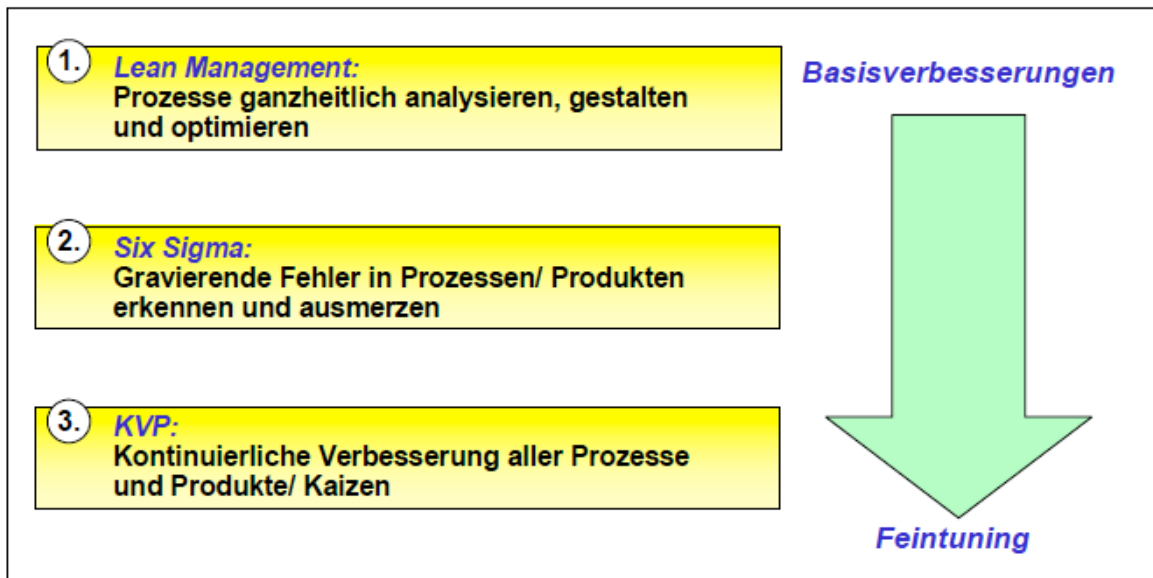


Abb. 8 Kombination von Lean Management und Six Sigma [Töp09, S.60]

Abschließend lässt sich also festhalten, dass Lean Six Sigma durch die sinnvolle Kombination von den sich ergänzenden Konzepten Lean Management und Six Sigma ein wirkungsvolles Gesamtkonzept zur Qualitätsverbesserung im Unternehmen darstellt. Da es nach [Töp09, S.5] auf alle Unternehmensprozesse anwenden lässt, ist die Aussage bzgl. des Verbesserungspotentials auch auf die Thematik der Datenqualität übertragbar.

3.3 Risiko-Assessment

Einen gänzlich anderen Ansatz als das besprochene Lean Six Sigma wählt Hakim Harrach in seiner Arbeit. Er sieht Datenqualitätsmanagement als ein Instrument der Unternehmensleitung, „um den Kosten-Nutzen-Faktor der Produktion guter Datenqualität zu optimieren.“ [Har10, S.20]. Dementsprechend merkt er an, dass die Kosten, welche bei mangelhafter Datenqualität, aber auch bei der Verbesserung der Datenqualität entstehen, in einem angemessenen Verhältnis zum erreichten Nutzen stehen müssen [Har10, S.20f.]. In der Konsequenz spricht er sich für eine Priorisierung von durch Daten induzierten Risiken aus, um einen guten Ausgangspunkt für weitergehende Maßnahmen zu haben [Har10, S.1].

Von einem Risiko ist nach [Har10, S.26] nicht bekannt, ob es eintreten wird und welchen Schaden es in diesem Fall verursachen wird. Es lässt sich aber mit einer Wahrscheinlichkeit beziffern, wobei das Risiko der Eintrittswahrscheinlichkeit multipliziert mit dem Schadensausmaß entspricht. Konkreter wird es bei dem Begriff IT-Risiko. Danach versteht man nach [Har10, S.27] die „Unfähigkeit, anforderungsgerechte IT-Leistungen effektiv und effizient erbringen zu können“. Im nächsten Schritt lässt sich auch ein IT-Risikomanagement definieren. Nach [Har10, S.27] umfasst dies „alle systematischen Maßnahmen zur rechtzeitigen Erkennung, Bewertung und Bewältigung von potentiellen Risiken.“. Harrach diskutiert in seiner Arbeit also die Bedeutung und den Ablauf eines Risikomanagements in der IT und zeigt anschließend seine Anwendungsmöglichkeit im Bereich des Datenqualitätsmanagements auf. Dieser Ansatz ist in der Fachliteratur möglicherweise ein einzigartiger oder zumindest seltener und soll hier daher vorgestellt, jedoch nicht in seiner genauen Methodik erläutert werden.

Ein Kernelement in der Arbeit Harrachs stellt der vorab bereits erwähnte PDCA-Zyklus dar [Har10, S.28]. Der PDCA-Zyklus, obwohl in den 1950er Jahren vom Amerikaner Deming (daher auch manchmal Deming-Zyklus genannt) entwickelt, findet weitläufige Anwendung in der japanischen Kaizen-Philosophie, dem kontinuierlichen Verbesserungsprozess [Har10, S.44ff]. Als solcher ist er nicht als geschlossenes System, sondern als offener Prozess zu verstehen, der immer wieder durchlaufen wird [Har10, S.28]. Er bildet die Grundlage für den von Harrach beschriebenen Risikomanagement-Prozess, welcher sich analog zum PDCA-Zyklus in die Phasen Risikoidentifikation (plan), -analyse (do), -handhabung (check) und -überwachung (act) aufteilt [Har10, S.28ff, S.62ff]. Damit nun jedoch die dem Risikomanagement gelieferten Daten qualitativ hochwertig sind, ist nach [Har10, S.34] die Einführung unternehmensweiter Qualitätsregelkreise notwendig. Dies wird weiterhin dadurch begründet, dass Datenqualität im Wesentlichen von den ihr zugrunde liegenden Prozessen bestimmt und daher auch am besten durch eine Prozessoptimierung verbessert wird [Har10, S.35].

Bevor eine solche Verbesserung der Datenqualitätsprozesse angestrebt werden kann, gilt es zunächst, die Voraussetzungen dafür zu schaffen. Hierfür bedient sich Harrach zunächst einmal sogenannter Reifegradmodelle, welche hier zum Ziel haben, die Qualität der relevanten Prozesse sichtbar zu machen. Da auf dem Markt verschiedene Reifegradmodelle existieren, muss je nach gewünschter Detaillierung und Darstellung ein für das jeweilige Unternehmen passende Modell gewählt werden [Har10, S.35ff]. Eine weitere Grundlage der Arbeit Harrachs ist ein ausgeprägtes Data-Ownership-Konzept, welches Verantwortlichkeiten für Unternehmensdaten und damit auch deren Qualität definiert. Im Wesentlichen sind dies Datenerfasser, -verarbeiter und -nutzer [Har10, S.41ff]. Sind diese Voraussetzungen geschaffen, kann auf Basis des PDCA-Zyklus nun ein Datenqualitätsregelkreis implementiert bzw. ein existierender effektiv genutzt werden. Die durchlaufenen Phasen sind dabei die Qualitätsplanung (plan), -lenkung (do), -sicherung (check) und -verbesserung (act) [Har10, S.45f., S.59f.].

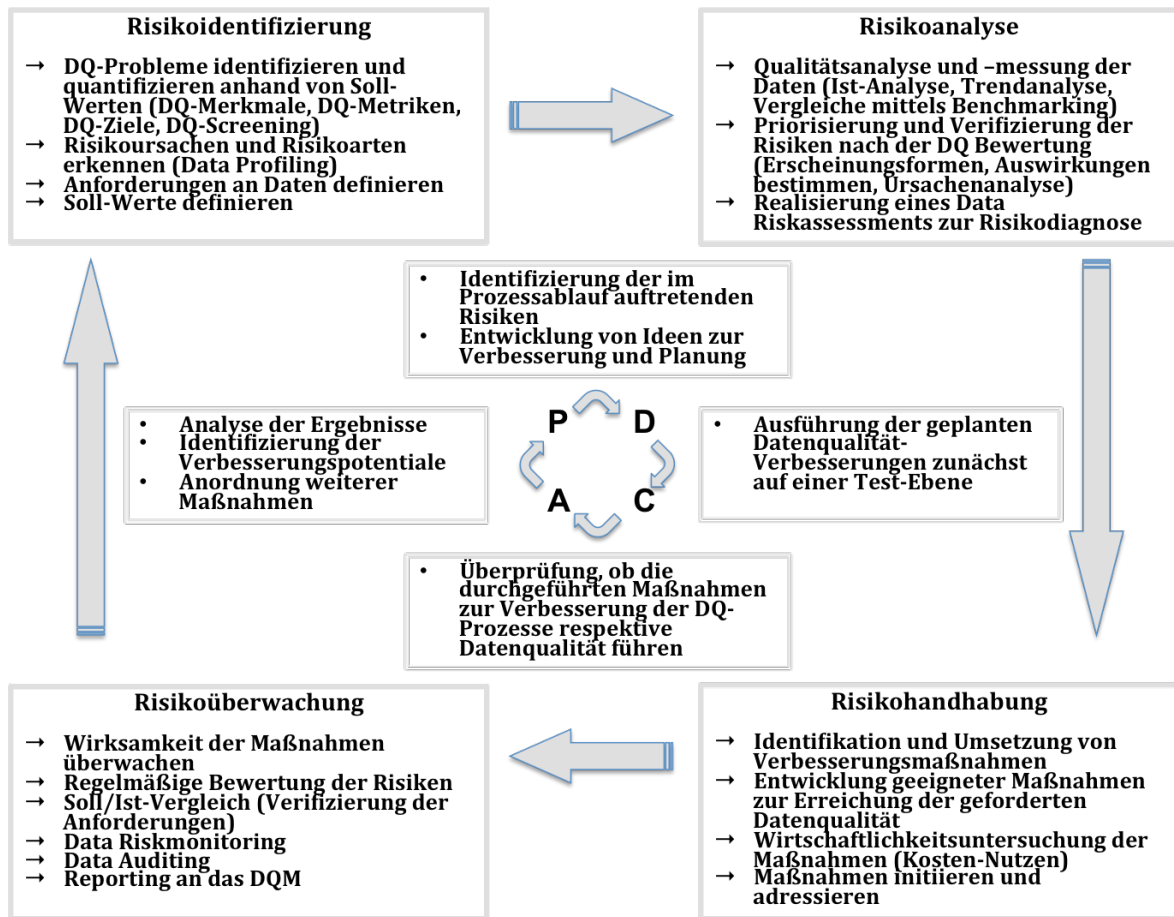


Abb. 9 Regelkreis des Risiko- und Qualitätsmanagements [Har10, S.61]

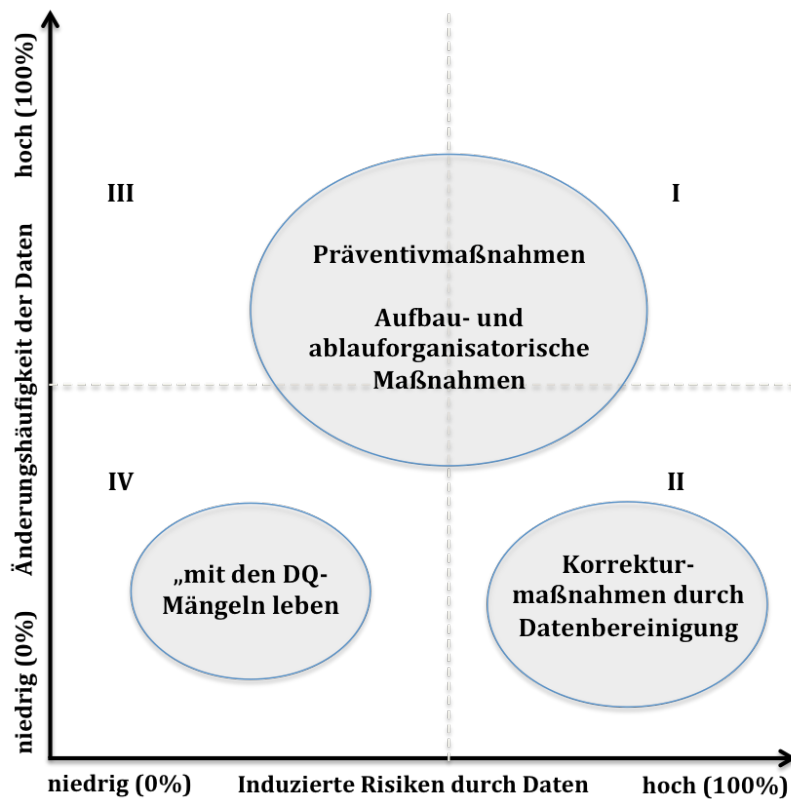


Abb. 10 Priorisierung der Risiken [Har10, S.66]

Abbildung 9 zeigt nun die Anordnung und Wirkungsweise eines Risikomanagement- im Verhältnis zu einem Datenqualitätsregelkreis. Abbildung 10 hingegen zeigt ein mögliches Ergebnis des Risiko-Assessments, nach welchem eine Priorisierung und damit effiziente Handhabung der identifizierten Risiken möglich ist. Beide Abbildungen verdeutlichen den Grundgedanken der Arbeit Harrachs, über die Einführung und Handhabung des Risikobegriffs, also einer Verminderung von Risiken, eine Steigerung der Datenqualität zu erreichen [Har10, S.68]. Diese Herangehensweise unterscheidet sein Werk grundlegend von den anderen in dieser Projektarbeit vorgestellten Konzepten und stellt einen interessanten Ansatz dar, ohne dabei spezifische Maßnahmen zur Steigerung der Datenqualität betrachten zu müssen. Allerdings ließen sich in dieser kurzen Abhandlung des Konzepts ebenfalls einige Gemeinsamkeiten feststellen, welche in Abschnitt 3.5 aufgegriffen werden.

3.4 A3 Thinking

Die dritte und letzte der in dieser Arbeit behandelten Methoden stellt nun A3 dar. Dabei handelt es sich mehr noch als die beiden zuvor behandelten Konzepte um eine Philosophie bzw. eine ganze Reihe von Prozessen und Gedankengerüsten [Sma08, S.8]. Es stellt laut [Sma08, S.2f.] einen der wesentlichen Faktoren für Toyotas Erfolg dar. Anders als das ebenfalls einen Eckpfeiler von Toyota bildende Lean Management verfolgt es dabei nicht die Verschlankeung von Prozessen und Qualitätssteigerung durch Vermeidung von Verschwendung, sondern vielmehr das tiefgehende Verständnis von im Unternehmen auftretenden Problemen und entsprechende Korrekturmaßnahmen [Sma08, S.5]. Kern bildet dabei auch hier die wissenschaftliche Arbeitsweise unter Zuhilfenahme des PDCA-Zyklus [Sma08, S.4]. Dabei wird A3 als das Fleisch auf dem PDCA-Skelett beschrieben [Sma08], was die Bedeutung des PDCA-Zyklus verdeutlicht. Namensgebend für das Konzept ist das A3-Papierformat, auf welchem kurze, knappe und übersichtliche Berichte zu behandelten Themen verfasst werden. Dabei ist zu beachten, dass diese Berichte ein Kernelement des Systems A3 ausmachen, aber nicht als das System selbst zu verstehen sind, welches sich wie gesagt durch umfassende Prozesse im Hintergrund auszeichnet [Sma08, S.8]. Da A3 nach [Sma08] als Werkzeug mit breitem Anwendungsgebiet angesehen wird, mit dem sich auch Qualitätsprobleme angehen lassen (siehe folgender Absatz), wird es hier ebenfalls im Kontext der Datenqualitätsthematik behandelt.

Um das Ausmaß des in A3 involvierten Denkens begreifen zu können, eignet sich folgende Anekdote hervorragend. In [Sma08, S.6f.] beschreibt einer der Autoren dabei seine Erfahrungen als Trainee in Japan. Er wurde mit der Aufgabe konfrontiert, Qualitätsprobleme an einer Präzisionsschleifmaschine zu lösen. Instinktiv schlug er eine Variierung verschiedener Parameter vor, um das Problem zu lösen. Sein Vorgesetzter hingegen lies ihn die Maschine mit ihren detaillierten Funktion erfassen, dokumentieren und eine Liste mit allen möglichen Defekten erstellen. Auf diese Weise ließ sich eine bestimmte Art von Defekt feststellen, welcher für die Qualitätsprobleme verantwortlich gemacht werden konnte. Im nächsten Schritt ordnete der Vorgesetzte an, sämtliche möglichen Ursachen für den Defekt aufzustellen. Da dies zur eindeutigen Ursachenbestimmung aber noch nicht ausreichend war, sollte der Autor für jede der möglichen Ursachen Korrekturmaßnahmen entwerfen und testen. Nach mehreren Tests konnte er das mit Bakterien verseuchte Kühlmittel als Fehlerquelle feststellen und durch einen Austausch des Kühlmittels Abhilfe schaffen. Er berichtete seinen Vorgesetzten und war überrascht, als diese nicht vollständig zufrieden mit seiner Arbeit

waren. Sie dankten ihm für seinen Aufwand und stellten ihm folgende Fragen. Wie konnte das Kühlmittel von Bakterien kontaminiert werden? Welche Tests können wir einsetzen, um das Kühlmittel zu überprüfen? Wer ist verantwortlich für den Prozess der Kühlmittelüberprüfung? Wie kann eine Kontamination des Kühlmittels in Zukunft vermieden werden?

Diese Geschichte verdeutlicht auf ausgezeichnete Art und Weise die Tragweite des A3 Gedankenkonstrukts. Die Behebung des Defekts an dieser einen Maschine war natürlich wichtig, wichtiger war es jedoch, die Fähigkeiten des Mitarbeiters zur eigenständigen Problemlösung zu entwickeln. Dies zeigt die langfristige Orientierung des Konzepts [Sma08, S.5, S.11]. Ferner zeigt es, wie A3 einen ewigen Kreislauf des Denkens und Handelns fördert [Sma08, S.12]. Dabei basiert die A3 Denkweise nach [Sma08, S.12ff] auf folgenden Elementen:

- 1) Logischer Denkprozess
- 2) Objektivität
- 3) Ergebnisse und Prozesse
- 4) Erstellung, Extraktion und Visualisierung (das Wichtigste kompakt)
- 5) Gegenseitige Abstimmung
- 6) Konsistenz innerhalb Problemlösung und gesamtem Unternehmen
- 7) Berücksichtigung des Gesamtsystems

Diese sieben Elemente stehen im Zusammenhang zueinander und verstärken sich gegenseitig [Sma08, S.18]. Man verfolgt hier also einen ganzheitlichen Ansatz und macht sich Synergieeffekte zunutze.

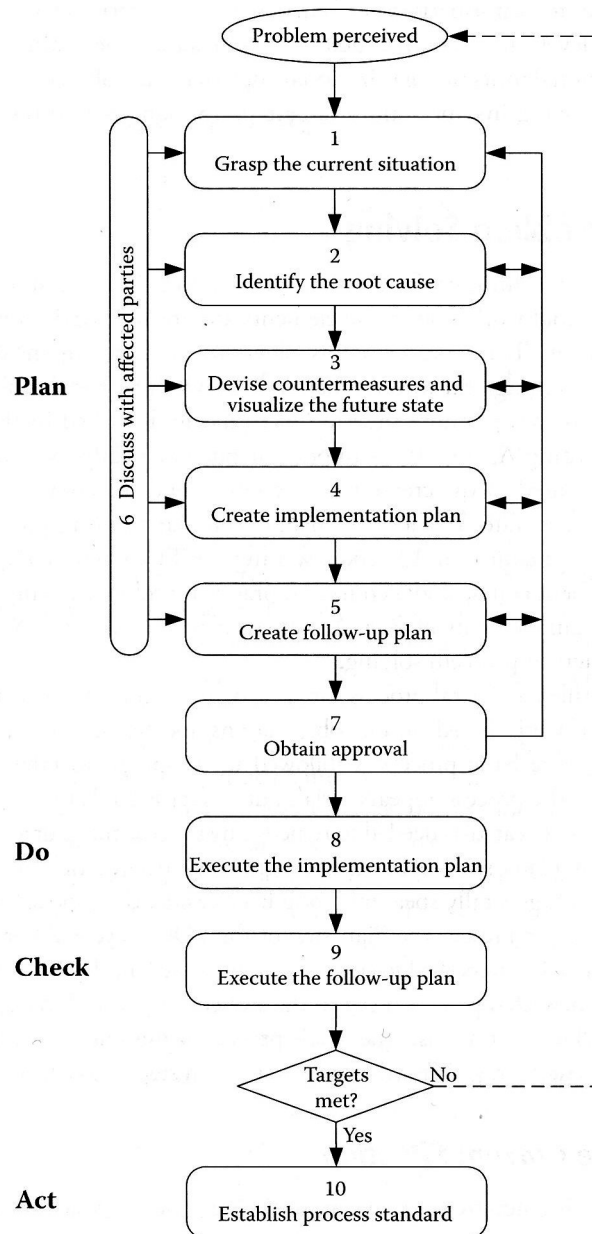


Abb. 11 Anwendung des PDCA-Zyklus in A3 [Sma08, S.20]

Abbildung 11 stellt nun kompakt und auf einen Blick – ganz im Sinne der A3 Berichte – die Herangehensweise an auftretende Probleme dar und demonstriert dabei die konsequente Anwendung des PDCA-Zyklus. Auch wenn A3 nicht in wenigen Seiten vollständig und angemessen erklärt ist, so zeigt die Abbildung doch viele der Kernelemente wie bspw. das geforderte tiefgehende Verständnis der Problematik, das konsequente Streben nach der Zielerreichung und die langfristige Orientierung. Diese Vorgehensweise, vielmehr aber die dahinterstehenden Gedankenkonstrukte und Philosophien machen A3 zu einem wirkungsvollen Allzweckwerkzeug [Sma08]. Seine Bedeutung hinsichtlich der Datenqualitätsthematik wird nun, zusammen mit den beiden anderen betrachteten Konzepten, im folgenden Abschnitt diskutiert und bewertet.

3.5 Zwischenfazit

Wie bereits zuvor herausgestellt wurde, wurde der Begriff der Methode des Datenqualitätsmanagements in dieser Arbeit sehr weitläufig betrachtet. Dies liegt u.a. daran, dass auch in der recherchierten Literatur zwar weitestgehend Übereinstimmung darin herrschte, was Datenqualitätsmanagement in seinen Grundzügen ausmacht, nämlich das Erhalten und potentiell auch das Verbessern der Datenqualität. Uneinig war man sich jedoch darüber, was Datenqualitätsmanagement und damit entsprechende Methoden ausmacht. Aus diesem Grund soll mithilfe von Abbildung 12 eine Betrachtung der Gemeinsamkeiten vorgenommen werden, da alle drei Methoden respektive Konzepte bekannte und erfolgreiche Werkzeuge für direktes oder indirektes Qualitätsmanagement darstellen. Die vom Autor gewählten Dimensionen und deren Ausprägungen sollen dabei einen Anhaltspunkt für eine Gegenüberstellung der Methoden geben.

	A3 Thinking	Risiko- management	Lean Management	Six Sigma	Lean Six Sigma
Verbreitung	★	★★	★★★	★★★	★★
Einbeziehung Mitarbeiter	★★★	★★	★★★	★	★★★
Akzeptanz	★★	★★	★★★	★★	★★
Komplexität Einführung	★★★	★★	★★★	★★	★★★
Lösung konkreter Probleme	★★	★★	★	★★★	★★★
Problemerkennung	★★★	★★★	★★★	★	★★★
Nachhaltigkeit	★★★	★★★	★★★	★	★★★
Ganzheitlichkeit	★★★	★★	★★★	★	★★★

★ wenig ausgeprägt ★★ ausgeprägt ★★★ stark ausgeprägt

Abb. 10 Methoden des Datenqualitätsmanagements im Vergleich

Betrachtet man die hier behandelten und in Abbildung 12 dargestellten Methoden, so fallen drei große Gemeinsamkeiten auf. Erstens handelt es sich um ganzheitliche Konzepte, die zahlreiche, wenn nicht sogar sämtliche Unternehmensprozesse in ihren Fokus einbeziehen. Dadurch eignen sie sich ebenfalls dazu, auftretende Probleme frühzeitig zu entdecken. Zweitens bedienen sich alle drei bewährter Regelkreise wie PDCA und DMAIC, überprüfen ihre Prozesse also permanent und passen ihre Zielvorgaben entsprechend an (siehe Dimension Nachhaltigkeit in Abbildung 12). Drittens bedarf es bei allen drei Methoden, schlussfolgernd aus den ersten beiden Punkten, eines hohen und regelmäßigen Aufwands zur Implementierung und Pflege, da grundlegende Strukturen eines Unternehmens einbezogen werden. Nicht zuletzt ist ein Grund hierfür in der Einbeziehung vieler oder sogar sämtlicher Mitarbeiter eines Unternehmens zu sehen, was sich allerdings wiederum positiv auf den Aspekt der Ganzheitlichkeit auswirkt.

Als Schlussfolgerung dieses Kapitels lässt sich also festhalten, dass Datenqualität und Datenqualitätsmanagement nicht als isolierte Thematiken betrachtet werden dürfen, die in einmaligen Maßnahmen behandelt werden können. Vielmehr muss es sich dabei um ganzheitliche Ansätze handeln, die unternehmensweit greifen und dabei einer regelmäßigen Prüfung unterliegen. Unter den aufgezeigten Konzepten ist dabei Lean Six Sigma hervorzuheben. Es bietet mit Lean Management genau den ganzheitlichen Ansatz, der unternehmensweit eine indirekte Qualitäts- und somit auch Datenqualitätssteigerung bewirkt, indem es für schlanke und funktionierende Prozesse sorgt. Ferner liefert es mit Six Sigma ein Werkzeug, explizit auftretende Probleme anzugehen und somit direkt Qualität zu steigern, womit es die Vorteile beider Konzepte vereint, wie anhand von Abbildung 12 ersichtlich wird. Unterstützt wird dies durch eine starke Verbreitung von Lean Management und Six Sigma, was in einem hohen Angebot an Arbeiten zu diesem Thema und Fachexpertise resultiert. Lean Six Sigma stellt somit einen komplettierten Ansatz zum Datenqualitätsmanagement dar.

4 Datenqualität im Kontext der Supply Chain

4.1 Grundlage: Supply Chain

Wie in der Einleitung bereits erwähnt, ist es unerlässlich für eine Betrachtung der Datenqualitätsthematik im Kontext der Supply Chain, zunächst eine geeignete Definition für den Begriff Supply Chain zu erarbeiten. Zwar findet sich in großen Mengen Fachliteratur zum Thema, doch wird eher selten eine umfangreiche Definition zugrunde gelegt. Aus diesem Grund soll nun ein Teil entsprechender Literatur untersucht werden, um aus der Querschnittsmenge eine treffende Definition des Begriffs Supply Chain zu gewinnen.

Um den Begriff Supply Chain einordnen zu können, ist es sicherlich hilfreich, zunächst einmal die deutsche Übersetzung zu betrachten: Versorgungs- oder Lieferkette. Damit wird fast sofort verständlich, was mit einer Supply Chain gemeint ist, nämlich ein Netzwerk aus Unternehmen, die direkt oder indirekt an der Erfüllung eines Kundenwunsches, bspw. der Herstellung eines Produkts, beteiligt sind [Cho13, S.13]. Obgleich Supply Chains prinzipiell nicht ausschließlich im produzierenden Gewerbe existieren, so eignet sich diese Einordnung dennoch gut dazu, die Weitläufigkeit des Begriffs zu verstehen. Auf das Beispiel des produzierenden Gewerbes bezogen, finden sich bspw. nicht nur Hersteller und Zulieferer innerhalb einer Supply Chain wieder, sondern ebenfalls Transportunternehmen, Lagerunternehmen, Verkäufer und auch die Kunden selbst [Cho13, S.13].

Ein ähnlicher Ansatz zur Definitionsfindung ist in [Cor08] zu finden. Hier wird von einer Logistikkette ausgegangen, welche den Realgüterprozess darstellt, erweitert um Beschaffungs-, Produktions-, Distributions- und Entsorgungslogistik [Cor08, S.94 ff]. Der entscheidende Unterschied zur Supply Chain ist jedoch, dass die Mitglieder einer logistischen Kette nach einzelwirtschaftlichen Interessen aus ihrer isolierten Sicht handeln, die Unternehmen innerhalb einer Supply Chain jedoch eine ganzheitliche, unternehmensübergreifende Zusammenarbeit und Verknüpfung der Kooperationspartner anstreben [Cor08, S.96 f.] [Haa08, S.18]. Eine Supply Chain lässt sich daher nicht als strikte Kette, sondern vielmehr als ein Kompetenznetzwerk verstehen [Bru10, S.11]. Von größter Bedeutung ist dabei die konsequente Kundenorientierung, welche als Ausgangspunkt für die gesamte Supply Chain gesehen

werden muss [Cor08, S.98]. Dies lässt sich u.a. damit begründen, dass der Kunde die einzige Quelle wirtschaftlichen Gewinns für die Supply Chain als Ganzes darstellt [Cho13, S.16]. Folgerichtig wird die Erfüllung von Kundenbedürfnissen als Primärziel einer Supply Chain angesehen und die Generierung von Profit sozusagen als Begleiterscheinung [Cho13, S.14], wobei der Erfolg einer Supply Chain wiederum am generierten Profit des gesamten Netzwerks als Kennzahl gemessen werden kann [Cho13, S.16].

Zusammenfassend lässt sich nun also eine erarbeitete Definition feststellen: eine Supply Chain ist ein Netzwerk aus verschiedenen Unternehmen, das Kompetenzen bündelt und dessen primäres Ziel die Erfüllung von Kundenbedürfnissen und im Zuge dessen Generierung von wirtschaftlichem Profit ist. Betrachtet man die Dynamik innerhalb einer Supply Chain, welche auf dem Fluss von Gütern, Finanzmitteln und Informationen in alle Richtungen beruht [Cho13, S.14] [Bru10, S.17], so wird die Relevanz einer hohen Datenqualität für alle Beteiligten deutlich. Als Beispiel sei hier der Bullwhip- oder Peitschenschlageffekt genannt, nach welchem sich Produktbestellungen mit jeder Stufe innerhalb einer Supply Chain als Reaktion auf die Bestellung der vorangegangenen Stufe aufschaukeln. Als Grund für dieses Phänomen ist sicherlich eine mangelhafte Datenqualität zu diskutieren, weshalb die bisherigen Ergebnisse dieser Arbeit nun im Kontext der Supply Chain diskutiert werden sollen

4.2 Diskussion der Datenqualitätsthematik in Supply Chains

Bevor nun eine spezifische Diskussion des Datenqualitätsbegriffs erfolgen kann, muss zunächst die Bedeutung von Datenqualität für Supply Chains aufgezeigt werden. Während Qualität im Allgemeinen als Erfolgsfaktor in vielen Branchen angesehen ist [Koc10, S.3], stellt Datenqualität noch einen relativ neuen Begriff dar [Wür03, S.12]. Tatsächlich wird das moderne Supply Chain Management nach [Hil01, S.41] aber maßgeblich von aktuellen Entwicklungen im Hard- und Softwarebereich beeinflusst. Dabei stellt es alleine aber keine geeignete Maßnahme dar, der Datenqualitätsproblematik zu begegnen, welche u.a. durch besagte technische Entwicklungen entsteht [Ape09, S.13]. Insofern verkörpert Datenqualität und Datenqualitätsmanagement eine höchst relevante Thematik für Supply Chains [Fan12, S.3], welche es näher zu betrachten gilt.

Betrachtet man nun die Datenqualitätsthematik im Kontext der Supply Chain, so ist zunächst keine gesonderte Definition des Datenqualitätsbegriffs aufzustellen. Dies liegt darin begründet, dass die hier betrachteten Definitionen allgemeiner Natur sind und für den individuellen Bedarf angepasst werden können, weswegen auch die Behandlung der Datenqualitätsthematik in dieser Arbeit vor der spezifischen Anwendung auf die Supply Chain erfolgt ist. Ebenfalls muss festgestellt werden, dass Datenqualität generell als sehr subjektiv zu erachten ist, da nicht nur jede Sicht auf Daten individuell ist (vgl. Begriff der Modellabhängigkeit in Abschnitt 2.4), sondern jede Person, jedes Unternehmen und jedes System potentiell unterschiedliche Daten benötigt [Eng11, S.18]. Insofern wurde bewusst darauf verzichtet, bereits zu Beginn dieser Arbeit einen speziell auf den Kontext der Supply Chains zugeschnittenen Datenqualitätsbegriff aufzuzeigen, welcher möglicherweise eher hinderlich für eine tiefergehende Betrachtung der Thematik gewesen wäre. Jedoch ist nach zunächst allgemeiner Betrachtung der Thematik hervorzuheben, dass die in Abschnitt 2.2 angesprochene Teilbarkeit von Informationen innerhalb der Supply Chain nach [Cho13, S.501] einen erhöhten Stellenwert einnimmt.

Als Grund hierfür lässt sich die stark erhöhte Störanfälligkeit von Supply Chains für bestimmte Risiken feststellen [Haa08, S.81ff]. Zu diesen Risiken werden IT-Risiken im weiteren Sinne, speziell aber ein mangelhafter Informationsfluss innerhalb der Supply Chain gezählt [Haa08, S.101f.]. Nach [Hil01, S.15] ist ebenfalls eine hohe Informationstransparenz essentiell für ein Wertschöpfungsnetzwerk, wie es eine Supply Chain darstellt. Es lässt sich somit festhalten, dass für die Betrachtung des Datenqualitätsbegriffs in Supply Chains eine allgemeine Definition, wie sie die in Abschnitt 2.4 aufgezeigte Arbeit Würtheles darstellt, hinzugezogen werden kann, welche jedoch um einige spezifische Aspekte ergänzt werden muss. Diese beinhalten die Verwendbarkeit der relevanten Informationen innerhalb der gesamten Supply Chain, also einheitliche technische Rahmenbedingungen und ein übereinstimmendes Verständnis der ausgetauschten Daten.

Nach [Haa08, S.35] stellt ein ganzheitliches Denken innerhalb der gesamten Supply Chain einen wesentlichen Erfolgsfaktor für selbige dar. Folgerichtig lässt sich feststellen, dass auch das in der Supply Chain verwendete Datenqualitätsmanagement und seine Methode ganzheitlicher Natur sein müssen, um eine hohe Datenqualität zu erzielen. Dies deckt sich mit der in Abschnitt 3.5 getroffenen Aussage, nach welcher Methoden des Datenqualitätsmanagements zwangsläufig einen ganzheitlichen Ansatz verfolgen müssen, um der Datenqualitätsthematik adäquat begegnen zu können. Nach [Bla10, S.215, S.224] werden speziell die Praktiken des Lean Managements für die Verwendung in Supply Chains empfohlen, da sie ein hohes Potential für Effizienz- und Produktivitätssteigerungen aufweisen. Auch [Haa08, S.81ff] beschreibt Lean Management im Zusammenhang mit Supply Chains, weist dabei jedoch explizit auf die Notwendigkeit eines Risikomanagements bei Lean Supply Chains hin. Beachtet man nun die in Abschnitt 3.2 beschriebenen Möglichkeiten Six Sigmas, punktuell auftretende Fehler und damit Risiken gezielt zu beseitigen, so ist das Konzept Lean Six Sigma auch im spezifischen Kontext der Supply Chain eine mögliche zu berücksichtigende Methode des Datenqualitätsmanagements.

5 Zusammenfassung

Zu Beginn dieser Arbeit wurden verschiedenen Definitionen zur Datenqualität untersucht, wobei sich die Arbeit Volker Würtheles in Umfang und Detaillierung hervortat. Nachfolgend wurden drei verschiedene Methoden bzw. Konzepte des Datenqualitätsmanagements betrachtet, wobei es sich um allgemeingültige und vielfältig einsetzbare Methoden handelte. Als Quintessenz des Kapitels und zu einem großen Teil auch der gesamten Arbeit konnte festgehalten werden, dass Datenqualität und Datenqualitätsmanagement nicht als isolierte Begriffe und Problematiken anzusehen sind. Vielmehr müssen sie im Gesamtkontext des jeweiligen Unternehmens gesehen werden. Einen ganzheitlichen Ansatz bietet dabei die Methode Lean Six Sigma, welche mit der Kombination aus unternehmensweiter Prozessoptimierung und gezielter Problemlösung ein hervorragendes Werkzeug zum nachhaltigen Datenqualitätsmanagement darstellt. Abschließend wurde eine geeignete Definition für eine Supply Chain hergeleitet, auf deren Basis die bisherigen Ergebnisse der Arbeit im spezifischen Kontext diskutiert wurden. Dabei wurde zunächst festgestellt, dass Datenqualität und Datenqualitätsmanagement in hohem Maße relevante Thematiken für die Organisationsform Supply Chain darstellen. Ferner wurde aufgezeigt, dass allgemeine Datenqualitätsdefinitionen wie das Werk Würtheles durch geringfügige Anpassungen auch für spezifische Anwendungsfälle eingesetzt werden können, wie sie

die Supply Chain darstellt. So ist hier insbesondere die Verwendbarkeit von Informationen innerhalb der gesamten Supply Chain zu nennen, auf die besonderes Augenmerk zu legen ist. Schlussendlich konnte festgestellt werden, dass unter den betrachteten Methoden des Datenqualitätsmanagements Lean Six Sigma ebenfalls einen verfolgenswerten Ansatz darstellt, um im Kontext der Supply Chain angewendet zu werden.

6 Ausblick

Die im Rahmen dieser Arbeit vorgestellten Datenqualitätsdefinitionen von Wang und Würthele stellen umfangreiche und statistisch belegte Werke dar. Metriken zur Quantifizierung der aufgestellten Datenqualitätsdimensionen sind bspw. in den Büchern von Würthele und Hildebrand zu finden, auch die DGIQ beschäftigt sich aktuell mit dem Thema. Anzumerken ist allerdings, dass die Arbeit Wangs und damit zu einem Teil auch Würtheles auf einer Umfrage unter 137 IT-Nutzern in den USA aus dem Jahre 1996 basiert. Wünschenswert wäre eine Überprüfung der ermittelten Dimensionen in Form einer aktuellen Umfrage, die ein breiteres geographisches Spektrum sowie weitere Datenqualitäts-Stakeholder in insgesamt größerer Anzahl einbezieht, da der Datenqualitätsbegriff im Laufe der vergangenen 18 Jahre stärker in den Fokus der Forschung gerückt ist und besagte Umfrage damit potentiell obsolet ist. Wahlweise werden dabei auch Fachkräfte aus dem Bereich des Supply Chain Managements hinzugezogen, um eine im Bedarfsfall möglichst spezifische Datenqualitätsdefinition aufstellen zu können. Weiterhin bedarf es nach Meinung des Autors Forschungsbedarf hinsichtlich der aufgestellten These, nachhaltiges Datenqualitätsmanagement nur im Rahmen von ganzheitlichen Konzepten betreiben zu können, wie sie hier vorgestellt wurden. Es sei an dieser Stelle darauf hingewiesen, dass keine der vorgestellten Methoden eine explizite Anwendung hinsichtlich der Datenqualitätsthematik vorsieht – mit Ausnahme von Six Sigma als isoliertem Werkzeug. Eine explizite Eignung im Rahmen des Datenqualitätsmanagements, eine Akzeptanzentwicklung bzgl. der vorwiegend fernöstlich geprägten Konzepte und die Vermittlung der Notwendigkeit einer ganzheitlichen Betrachtung des Datenqualitätsbegriffs bilden speziell im Bereich der Supply Chain somit weitere Bereiche, die es zu untersuchen und zu entwickeln gilt.

7 Abbildungsverzeichnis

1 Dimensionen der Informationsqualität nach der DGIQ	6
2 Die 15 IQ-Dimensionen der DGIQ im Detail	7
3 Datenqualitätsradar nach Würthele	9
4 Die Dimensionen des Datenqualitätsradars nach Würthele	10
5 Gleiches Ziel, unterschiedlicher Ansatz von Lean Management und Six Sigma	14
6 Grundprinzipien des Lean Thinking	15
7 Der DMAIC-Zyklus in Six Sigma	16
8 Kombination von Lean Management und Six Sigma	17
9 Regelkreis des Risiko- und Qualitätsmanagements	19
10 Priorisierung der Risiken	19
11 Anwendung des PDCA-Zyklus in A3	22
12 Methoden des Datenqualitätsmanagements im Vergleich	23

8 Literaturverzeichnis

- [Ape09] Apel, Detlef; Behme, Wolfgang; Eberlein, Rüdiger; Merighi, Christian: Datenqualität erfolgreich steuern, 1. Edition. München: Carl Hanser, 2009
- [Bla10] Blanchard, David: Supply Chain Management, 2. Edition. Hoboken: John Wiley & Sons, 2010
- [Bru10] Brumme, Hendrik; Schröter, Norbert; Schröter, Ingo: Supply Chain Management und Logistik, 1. Edition. Stuttgart: Kohlhammer, 2010
- [Brü12] Brüggemann, Holger; Bremer, Peik: Grundlagen Qualitätsmanagement, 1. Edition. Wiesbaden: Vieweg + Teubner, 2012
- [Cho13] Chopra, Sunil; Meindl, Peter: Supply Chain Management, 5. Edition. Boston: Pearson, 2013
- [Cor08] Corsten, Hans; Gössinger, Ralf: Einführung in das Supply Chain Management, 2. Edition. München: Oldenbourg, 2008
- [Eng11] Engelmann, Florian; Großmann, Christoph: Was wissen wir über Informationen? In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 3-24
- [Fan12] Fan, Wenfei; Geerts, Floris: Foundations of Data Quality Management, 1. Edition. San Rafael: Morgan & Claypool, 2012
- [Gam13] Gamweger, Jürgen; Jöbstl, Oliver: Six Sigma. In: Kamiske, Gerd F.: Handbuch QM-Methoden, 2. Edition. München: Carl Hanser, 2013, S. 247-291
- [Geb11] Gebauer, Marcus; Windheuser, Ulrich: Strukturierte Datenanalyse, Profiling und Geschäftsregeln. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 88-101
- [Gor13] Gorecki, Pawel; Pautsch, Peter: Lean Management. In: Kamiske, Gerd F.: Handbuch QM-Methoden, 2. Edition. München: Carl Hanser, 2013, S. 153-190
- [Haa08] von Haaren, Britta: Konzeption, Modellierung und Simulation eines Supply-Chain-Risikomanagements, 1. Edition. Dortmund, 2008

- [Har10] Harrach, Hakim: Risiko-Assessments für Datenqualität, 1. Edition. Wiesbaden: Vieweg + Teubner, 2010
- [Hei11] Heinrich, Bernd; Klier, Mathias: Datenqualitätsmetriken für ein ökonomisch orientiertes Qualitätsmanagement. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 49-67
- [Hil01] Hildebrand, Knut; Hillek, Thomas; Lawrenz, Oliver; Nenninger, Michael: Supply Chain Management, 2. Edition. Braunschweig: Friedrich Vieweg & Sohn, 2001
- [Hil11] Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011
- [Hum13] Hummel, Thomas; Malorny, Christian: Total Quality Management (TQM). In: Kamiske, Gerd F.: Handbuch QM-Methoden, 2. Edition. München: Carl Hanser, 2013, S. 1-48
- [Kas11] Kasten, Gerhard; Rohweder, Jan Philipp: Informationsqualität – Definitionen, Dimensionen und Begriffe. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S.25-45
- [Koc10] Koch, Jan: Qualitätsmanagement in Logistikunternehmen, 1. Edition. Lohmar: Josef Eul, 2010
- [Koc11] Koch, Susanne: Einführung in das Management von Geschäftsprozessen, 1. Edition. Berlin, Heidelberg: Springer-Verlag, 2011
- [Kos13] Kostka, Claudia; Kostka, Sebastian: Kontinuierlicher Verbesserungsprozess (KVP). In: Kamiske, Gerd F.: Handbuch QM-Methoden, 2. Edition. München, Carl Hanser, 2013, S. 131-151
- [Lee01] Lee, Yang W.; Wang, Richard Y.; Ziad, Mostapha: Data Quality, 1. Edition. New York: Springer Science+Business Media, 2001
- [Lee06] Lee, Yang W.; Funk, James D.; Pipino, Leo L.; Wang, Richard Y.: Journey to Data Quality, 1. Edition. Cambridge: MIT Press, 2006
- [Lüs11] Lüssem, Jens: Organisatorische Ansiedlung eines Datenqualitätsmanagements. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 229-240

[Mor11] Morbey, Guilherme: Datenqualität für Entscheider in Unternehmen, 1. Edition. Wiesbaden: Gabler Verlag / Springer Fachmedien, 2011

[Sma08] Smalley, Art; Sobek, Durward K.: Understanding A3 Thinking, 1. Edition. Boca Raton: Productivity Press, 2008

[Töp09] Töpfer, Armin: Lean Six Sigma, 1. Edition. Berlin, Heidelberg: Springer-Verlag, 2009

[Tre06] Treiblmaier, Horst: Datenqualität und individualisierte Kommunikation, 1. Edition. Wiesbaden: Deutscher Universitäts-Verlag / GWV Fachverlage GmbH, 2006

[Wei11] Weigel, Niels: Datenqualitätsmanagement – Steigerung der Datenqualität mit Methode. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 68-87

[Wol11] Wolf, Jürg: Organisatorische Maßnahmen für gute Datenqualität. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 241-259

[Wür03] Würthele, Volker: Datenqualitätsmetrik für Informationsprozesse, 1. Edition. Norderstedt: Books on Demand GmbH, 2003

[Zwi11] Zwirner, Marcus: Datenbereinigung zielgerichtet eingesetzt zur permanenten Datenqualitätssteigerung. In: Hildebrand, Knut; Gebauer, Marcus; Hinrichs, Holger; Mielke, Michael: Daten- und Informationsqualität, 2. Edition. Wiesbaden: Vieweg + Teubner, 2011, S. 102-122