

AN DEN RICHTIGEN STELLEN BESSERE ENTSCHEIDUNGEN TREFFEN

# DANK OPTIMIERTER DATENQUALITÄT

Um mit Hilfe von Daten die richtigen Entscheidungen zu treffen, ist die passende Datenqualität ausschlaggebend. Dabei muss jedoch eine Kosten-Nutzen-Abwägung getroffen werden, an welchen Stellen die Optimierung einen entscheidenden Mehrwert bietet und sich auszahlt. Wichtig ist das korrekte Zusammenspiel der fachlich getriebenen Prozesse durch eine passende IT-seitige Unterstützung – und genau an der Stelle beginnt häufig die Optimierung mit einer fundierten Analyse. Letzten Endes müssen die Daten in einer passenden Form zur Verfügung stehen, um sie als Grundlage für bessere Entscheidungen oder andere Zwecke nutzen zu können. Doch worauf muss geachtet werden?

Jan Schlechter und Henry Schnieders, Unternehmensberater im Bereich Performance Management haben sich mit Account Manager Eric Johanning ein erfolgreiches Kundenprojekt als Beispiel genommen und in einem Interviewgespräch das Vorgehen näher erläutert. So wird die Vorgehensweise nicht nur als blanke Theorie dargestellt, sondern direkt ganz konkret mit echter Praxis untermauert.

## EXPERTENINTERVIEW DATENQUALITÄT

**Interviewer** = Eric Johanning, Account Manager

**JS** = Jan Schlechter, Manager

**HS** = Henry Schnieders, Consultant Business Intelligence

Gefühlt werden heute zu jeder Zeit und an allen Ecken und Enden Daten generiert. Sowohl bei der Erfassung aber auch bei der späteren Weiterverarbeitung spielt die Datenqualität eine große Rolle. Warum das so ist, werden wir später sicher noch erfahren. Ich freue mich sehr, mich mit euch beiden Experten über dieses spannende und immer wichtiger werdende Thema zu unterhalten. Ihr habt als Unternehmensberater im Bereich Performance Management ja täglich mit dem bestmöglichen Kombinieren und Auswerten von Daten zu tun. Jan, kannst Du uns einen kurzen Einblick geben, was uns im Interview erwartet?

**JS:** Wir werfen heute einen ganzheitlichen Blick auf das Thema Datenqualität, insbesondere auf Ursachen und Lösungsansätze für Qualitätsengpässe. Ich werde vor allem die theoretisch-konzeptionelle Seite darstellen und Henry wird

anhand eines Projektbeispiels von der konkreten Umsetzung in einem Kundenprojekt berichten. Daran können wir sehr schön sehen und zeigen, wie die Wirkung einer schlechten und vor allem auch einer verbesserten Datenqualität ist. Denn was wir stets im Blick behalten sollten ist, dass wir alle IT-Systeme und Datenanalysen nicht zum Selbstzweck betreiben, sondern immer einen Mehrwert generieren wollen. Das kann sich beispielsweise in einer vereinfachten und verbesserten Entscheidungsfindung aufgrund einer soliden Datenbasis zeigen.

Du sprichst einen sehr wichtigen Punkt an: den Mehrwert. Dieser sinkt verständlicherweise mit einer schlechteren Datenqualität. Könnte man diesen geringeren Mehrwert dann nicht auch als die Kosten mangelnder Datenqualität bezeichnen?

**JS:** Richtig. Werden auf Basis fehlerhafter oder fehlender Daten schlechte Entscheidungen getroffen, ist das eine Auswirkung der unzureichenden Datenqualität. Die Kostenbetrachtung sollte im Umgang mit Daten allerdings deutlich früher einsetzen. Betrachten wir den Prozess der Datenerfassung bis zur Nutzung, steigen die Kosten mit jedem Prozessschritt deutlich an. Die Verhinderung von Engpässen in der Datenqualität unmittelbar bei der Datenerfassung ist häufig am günstigsten. Etwas teurer, da meistens mit mehr Aufwand verknüpft, ist die Datenkorrektur in den Systemen. Am teuersten jedoch ist häufig, wie schon angesprochen, das Zulassen von Fehlern. Das passiert allerdings oft un bemerkt oder ist leider schon der Fall, wenn wir zu Projekten gerufen werden. Da kann Henry bestimmt einen Einblick geben.

**HS:** Genau. Ich stelle in dem Zusammenhang gerne ein Projektbeispiel vor, bei dem wir es geschafft haben, die Datenqualität zu optimieren und vor allem die kostspieligen Konsequenzen der mangelnden Datenqualität zu verhindern. Es geht dabei um ein Projekt, bei dem Artikelstammdaten, insbesondere die Abmessungen der Artikel, eine wichtige Rolle spielen. Das ist oft so, doch in diesem speziellen Fall stellten die Abmessungen einen entscheidenden Faktor zur Berechnung des Rechnungsbetrags vom Auftragnehmer an den Kunden, einen Logistikdienstleister, dar.

Die Situation war hier bereits so weit eskaliert, dass der Endkunde aufgrund der für ihn unplausiblen Rechnungshöhe, mehrere Rechnungen abgewiesen hatte. Das Unternehmen, für das wir das Datenqualitätsprojekt durchgeführt haben, hätte bei andauerndem Zustand in einen Liquiditätsengpass rutschen können. Zusätzlich führten die Unstimmigkeiten zu einem hohen Aufwand und damit hohen Prozesskosten bei der Rechnungslegung, da immer wieder Rückfragen zu klären waren. In letzter Konsequenz führte die Situation zu einer hohen Unzufriedenheit des Endkunden und es drohte der Projektverlust für unseren Auftraggeber.

**Wahrscheinlich sind die Ursachen für solche Datenqualitätsprobleme vielschichtig. Du hast gerade schon das Thema Stammdaten angesprochen. Jan, kannst Du uns einen Überblick verschaffen, wie bzw. in welchen Bereichen Datenqualitätsprobleme häufig entstehen?**

**JS:** Zunächst einmal müssen wir sehen, welche Dimensionen bei der Qualität von Daten eine Rolle spielen, (vgl. Abbildung 1 auf Seite 3). Es geht also um vielschichtige Themen: von der Datenkonsistenz, über die Genauigkeit bis hin zur Vollständigkeit.

Die Qualität der Daten wird insbesondere, aber nicht ausschließlich, bei der Datenerfassung definiert. Die darauf aufbauenden Prozesse können zur Verbesserung der Datenqualität beitragen, diese aber auch gefährden. Generell betrachten wir vier Stufen.

**1. Datenerfassung:** Durch manuelle Dateneingabe, instabile Prozesse, unzureichend definierte Schnittstellen und weitere Datenerfassungsschritte entstehen fehlerhafte Datensätze.

**2. Stammdaten:** Dezentrale Verwaltung und Pflege von Stammdaten führt häufig zu Inkonsistenzen und trägt daher maßgeblich zu mangelnder Datenqualität bei.

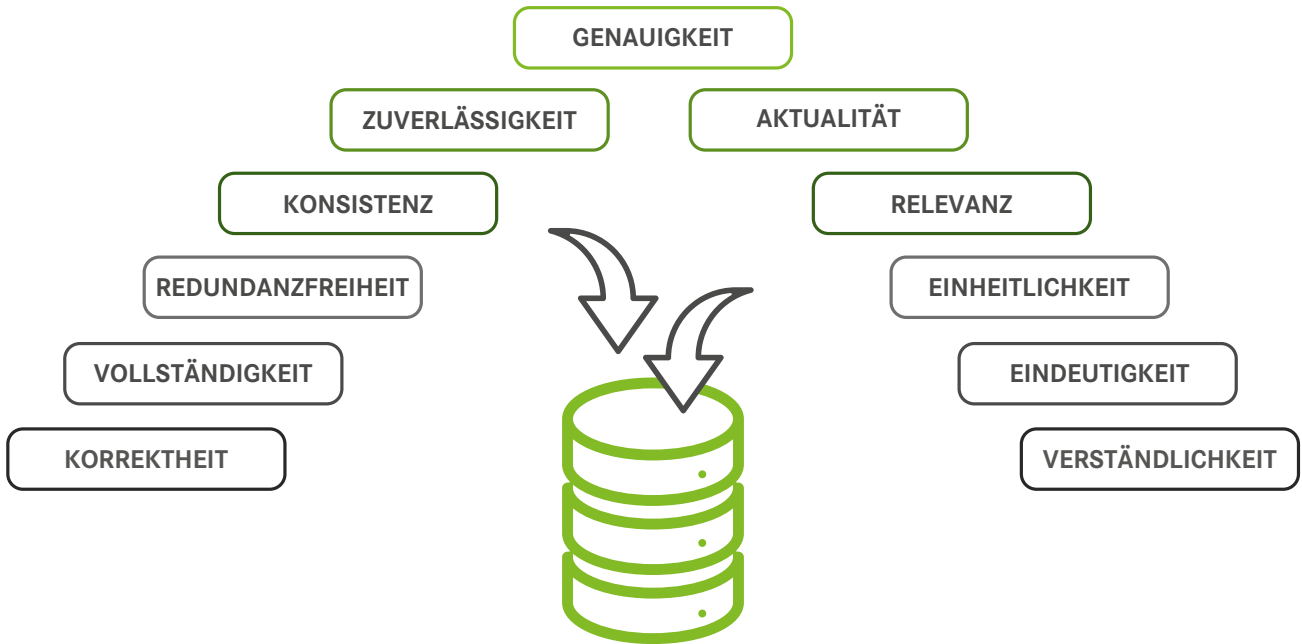
**3. Datenverarbeitung:** Bei der Weiterverarbeitung von Daten führt ein Nichtbeachten mangelnder Datenqualität und ein fehlendes Cleansing zu weiteren Datenqualitätsproblemen.

**4. Datenanalyse und -konsum:** Mangelnde Datenqualität führt nicht nur zu langsameren Prozessen, sondern insbesondere zu fehlerhaften Ergebnissen. Weiterhin bedeutet Daten-„Self-Service“ ohne jegliche Rahmenbedingungen weitere Fehler.

An genau diesen Stellen ist auch mit der Behebung der Fehlerursachen anzusetzen.

**Kannst Du uns das noch etwas genauer erläutern?**

**JS:** Wenn wir uns über die Lösungsansätze Gedanken machen, möchte ich zuerst einmal herausstellen, dass ein ganzheitlicher Blick gewählt werden sollte. Zwar kann mit Anpassungen an einzelnen Prozessen viel erreicht werden, für eine möglichst gesamthafte und umfängliche Lösung bedarf es jedoch einer Abstimmung zwischen Prozessen in den Fachbereichen und den IT-Prozessen.



Neben **externen Faktoren**, die ggf. die Notwendigkeit der **Datenqualität beeinflussen**, spielt der **Wert** der Daten für den **Geschäftserfolg** eine maßgebliche Rolle, wenn **Datenqualitätsstandards** festgelegt werden.

Abbildung 1: Dimensionen der Datenqualität

Auf Prozessebene können wir uns diese vier Aspekte dann genauer anschauen, um Lösungen zu implementieren:

**1. Datenerfassung:** Hier helfen die Automatisierung von manuellen Prozessen, Konsistenzprüfungen bei der Datenerfassung und Überarbeiten der Schnittstellen, um die Fehler bei der Datenerfassung deutlich zu reduzieren.

**2. Stammdatenpflege:** Mithilfe einer zentralen Stammdatenverwaltung und klaren Verantwortlichkeiten im Datensystem können Inkonsistenzen vermieden werden.

**3. Datenverarbeitung:** Sobald Qualitätsmängel in den Daten bekannt sind, können Korrekturalgorithmen in sogenannten Cleansinglayern in den Datenbankstrukturen die Datenqualität deutlich erhöhen.

**4. Datenanalyse und -konsum:** Hier hilft eine Governance in der Datenanalyse sowie die verbesserte Datenqualität aus den Vorsystemen zu schnelleren Prozessen und besseren

Entscheidungsgrundlagen.

Ich denke diese Themen werden einfacher verständlich, wenn wir uns das noch einmal in der Praxis anschauen.

**Henry, kannst Du uns einen Einblick geben?**

**HS:** Um Lösungen für die Datenqualitätsprobleme zu finden, ist eine genaue Analyse unumgänglich. Im konkreten Projekt stellten sich auf mehreren der von Jan dargestellten Ebenen Engpässe in den Prozessen dar. Beginnen wir mit der Datenerfassung: Hier war für uns der fachliche Prozess der Datenerhebung maßgebend. Generell fehlte eine belastbare Vereinbarung, welche Anforderungen bei der Ermittlung der Produktabmessungen überhaupt zu erfüllen waren. Diese ist wichtig, da sich die erforderliche Genauigkeit des Messmittels daraus ableitet. Allerdings wies das zuvor vorgesehene Messmittel mehrere hundert Prozent Ungenauigkeit bei derart kleinen Abmessungen der Produkte auf – und war außer Frage ungeeignet für die Datenerfassung. Vorübergehend wurde der Prozess bereits auf eine manuelle

Datenerfassung umgestellt, die allerdings zu einer weiteren Fehlerquelle, der manuellen Datenübertragung führte. Neben dem Prozess der Datenerfassung waren Inkonsistenzen zwischen den zwei verwendeten IT-Systemen auffällig. Als Ursache stellte sich heraus, dass beide Systeme zur Stammdatenverwaltung verwendet wurden, jedoch das eine System eine genauere Bewertung der Produktchargen ermöglichte als das andere. Die bis dato verwendete Schnittstelle konnte diese Informationen nicht valide synchronisieren. Eine weitere Asynchronität zwischen beiden Systemen kam einerseits durch eine recht lange Zeitdauer zwischen den Synchronisierungsvorgängen zustande und war andererseits auf Fehler bei der initialen Datenbeladung zurückzuführen. Ein weiterer Mangel bzw. Inkonsistenz zwischen mehreren, im selben Bericht ausgegebenen, Informationen entstand unmittelbar bei der Auswertung der Daten durch einen kleinen Unterschied in der Logik für die Rundung der Werte.

**Das klingt schlüssig. Du hast gerade schon gesagt, dass eine gründliche Analyse erforderlich ist. Kannst Du uns zusätzlich einen genaueren Einblick in die Vorgehensweise zur Optimierung der Datenqualität geben?**

**HS:** Die Basis zur Feststellung der Datenqualität ist immer eine Messung. Es muss also Kriterien geben, nach denen der Status Quo und auch der Sollzustand möglichst quantitativ bewertet werden können. In unserem Projekt war dies bereits grundlegend geschehen. Hier ist beispielsweise die Genauigkeit von numerischen Messwerten zu nennen. Um diese zu prüfen, haben wir im angesprochenen Projekt teilweise eine stichprobenartige Messung durchgeführt und mit den im führenden IT-System SAP vorhandenen Daten verglichen. Beim Abgleich der IT-Systeme eignet sich eine 100%-Prüfung. Die hier definierten Kriterien können außerdem auch bei Projektabschluss zur Bewertung des Erfolgs bzw. der Wirkung der Maßnahmen genutzt werden.

Als zweiter Schritt folgt die Bewertung der vorliegenden Messung. Dabei kann auch die Auswirkung des jeweiligen Datenqualitätsproblems auf den Geschäftserfolg bewertet werden, um beispielsweise einzelne Engpässe gegenüber anderen in der Abstimmung bzw. Maßnahmendefinition zu priorisieren.

Im dritten Schritt erstellen wir eine Roadmap, welche Maß-

nahmen mit welcher Priorisierung ergriffen werden können, um die Qualitätsprobleme abzustellen. Die Bewertung und das Ziel spielen dabei eine entscheidende Rolle. Nicht immer müssen beispielsweise auch alle Fehlerursachen vollständig behoben werden, wenn die Auswirkungen gering sind.

Es folgt die Umsetzung entsprechender Maßnahmen. In einem letzten, ebenfalls sehr wichtigen, Schritt schließt sich der Kreis: Die Einführung eines Monitorings hilft, den Erfolg der Maßnahmen dauerhaft zu prüfen, zu messen und bei neuen Abweichungen zukünftig schneller eingreifen zu können.

Insgesamt kann hier gut eine Parallele zu dem vielen bekannten PDCA-Zyklus aus dem klassischen Qualitätsmanagement gezogen werden.

**Hast Du zu diesen Schritten noch ein paar konkrete Beispiele aus deinem Projekt? Welche Maßnahmen hast Du ergriffen?**

**HS:** Durch unsere gründliche Analyse konnten wir die Ursachen gut identifizieren und direkt an der Entstehung der Datenqualitätsprobleme eingreifen. Es wurde ein geeignetes Messmittel mit zu den Anforderungen passender Genauigkeit eingeführt. Zusätzlich wurden in der Software Logiken implementiert, die einen einfachen Plausibilitätscheck durchführen. Eine Abmessung darf beispielsweise nicht gleich null sein. Entsprechende Werte werden nun gar nicht mehr vom System angenommen und eine erneute Messung verlangt. Eine weitere Änderung wurde in der Schnittstelle zwischen den beiden IT-Systemen eingeführt und das Synchronisationsintervall verkürzt. Zur fortlaufenden Messung der Datenqualität hinsichtlich der bekannten Fehler, wurde ein Dashboard erstellt. Dieses stellt eine Übersicht über Materialien mit unplausiblen Stammdaten dar, bei denen es aufgrund des Aufwands nicht angemessen gewesen wäre die Fehler über die Schnittstellen oder automatisierte Logiken abzustellen. Ein Alert ermöglicht ein kurzfristiges Eingreifen bei einer Auffälligkeit. Hier zeigt sich erneut wie wichtig es ist, die Dinge ganzheitlich zu betrachten, sowohl die Möglichkeiten der IT als auch im fachlichen Prozess.

In beiden Bereichen können entscheidende Optimierungs-

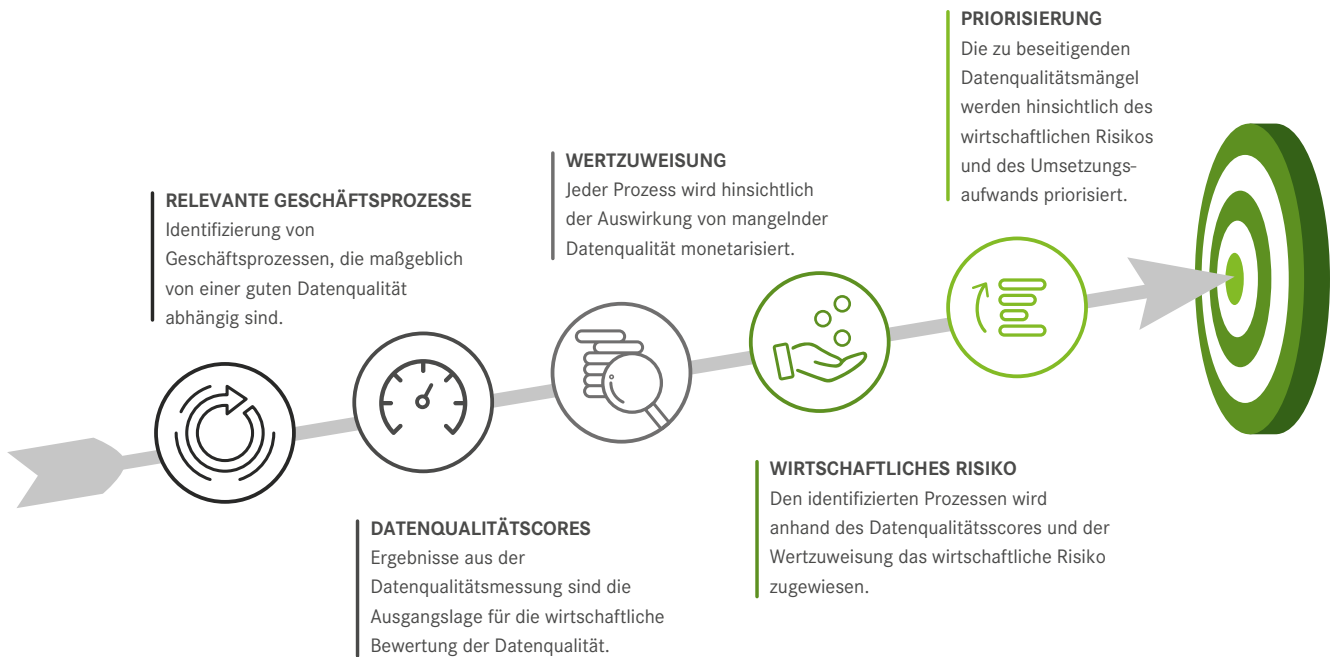


Abbildung 2: Bewertung der Datenqualität zur Priorisierung der Optimierung

potentiale liegen, die denselben Qualitätsengpass verbessern aber sehr unterschiedlich in ihren Kosten für die Umsetzung ausfallen.

**Danke für den Einblick. Ich finde mit einem Beispiel wird es immer gleich besser verständlich, was eigentlich hinter solchen Maßnahmen steckt. Beim Thema Bewertung der Datenqualität sprichst du von der Bewertung hinsichtlich der Auswirkungen auf den Geschäftserfolg. Könnt Ihr das für uns noch etwas konkretisieren?**

**JS:** Grundsätzlich geht es hier darum eine Priorisierung durchzuführen, welche Daten bzw. Informationen in einer hohen Qualität besonders wertvoll sind. Ein Überblick ist im Schaubild in Abbildung 2 zu sehen. Dazu sind zunächst Bewertungsmaßstäbe für die Datenqualität zu definieren und die relevanten Geschäftsprozesse zu identifizieren, die maßgeblich von der guten Datenqualität abhängig sind. Die Priorisierung entsteht schließlich, indem der jeweilige Prozess über den Wert einer mangelnden Datenqualität monetarisiert und das wirtschaftliche Risiko bewertet wird. Ergänzend wird der Umsetzungsaufwand für eine Maßnahme zur

Optimierung der Datenqualität betrachtet.

**Danke für die Erläuterungen. Wie ich es verstanden habe, wurde das Projekt erfolgreich abgeschlossen. Kannst Du uns etwas zum Ergebnis sagen, Henry?**

**HS:** Gerne, denn es ist eine wahre Erfolgsgeschichte, wenn wir auch die weiteren Folgen betrachten, die sich durch die verbesserten Prozesse und die verbesserte Datenqualität ergeben. Mit Abschluss des Projektes konnten einige Aufwände um ca. 80% reduziert werden, da keine aufwendigen Klärungsschleifen aufgrund unstimmiger Daten mehr erforderlich sind. Hinzu kommt, dass für unseren Kunden ein Reputationsschaden nicht nur abgewendet werden, sondern die Zusammenarbeit mit dessen Endkunde auch aufgrund des Projekterfolgs sogar erweitert werden konnte.

**Vielen Dank für die Einblicke und das interessante Gespräch.**

**JS:** Sehr gerne!

**SIE HABEN FRAGEN?**

# UNSERE EXPERTEN FREUEN SICH AUF DEN DIALOG MIT IHNEN!



**HENRY SCHNIEDERS**

Consultant Business Intelligence

**Tel.** +49 7031 6288-3226

**E-Mail** [henry.schnieders@star-cooperation.com](mailto:henry.schnieders@star-cooperation.com)



**JAN SCHLECHTER**

Manager Performance Management & BI

**Tel.** +49 7031 6288-3514

**E-Mail** [jan.schlechter@star-cooperation.com](mailto:jan.schlechter@star-cooperation.com)



**ERIC JOHANNING**

Account Manager

**Tel.** +49 7031 6288-3273

**E-Mail** [eric.johanning@star-cooperation.com](mailto:eric.johanning@star-cooperation.com)