

Brauchbare Prognosen setzen gute Datenqualität voraus

Der Wunsch, in die Zukunft blicken zu können, treibt den Menschen um, seit er auf der Welt ist. Allerhand Orakel, Wahrsager, Kartenleger und sonstige „Hellseher“ wurden und werden dazu befragt, was kommt. Die Vorhersage der wahrscheinlichen Zukunft und Trends spielt naturgemäß auch für Unternehmen eine große Rolle. Deshalb unternehmen sie erhebliche Anstrengungen, um Zusammenhänge, Muster und Korrelationen in den ihnen zur Verfügung stehenden Daten zu erkennen. Die Methode, mit deren Hilfe Unternehmen anhand von Daten aus der Vergangenheit und Gegenwart auf zu erwartende und mögliche Ereignisse und Entwicklungen schließen können, heißt Predictive Analytics. Sie ist nicht mehr ganz neu, aber immer mehr im Kommen.

Von Monika Pürsing*

Die Vorhersage zukünftiger Entwicklungen anhand historischer Daten sei wie im letzten Jahr ein wichtiges Thema für CIOs, seine Bedeutung sei sogar noch einmal leicht gestiegen, so Capgemini in den „IT-Trends 2017“. Unter den Top-IT-Themen des laufenden Jahres stehen Predictive Analytics auf Platz fünf. Ihre Bedeutung wird auf einer Skala von eins (sehr wichtig) bis sechs (völlig unwichtig) mit 2,6 bewertet. Angesichts der Digitalisierung und der zunehmenden Anzahl von Sensoren und Daten seien Predictive Analytics eine Methode, deren Nutzung in den kommenden Jahren stark ausgebaut werde. Denn derzeit nutzten sie zwar erst knapp sieben Prozent der Umfrageteilnehmer, aber mehr als 25 Prozent implementierten derzeit und 38 Prozent steckten in der Planung.

Predictive Analytics: Definition, Einsatzbereiche, Mehrwert

Was genau sind Predictive Analytics, wo werden sie eingesetzt und was bringen sie? Etwas sperrig, aber präzise definiert Wolfgang Martin: Bei Predictive Analytics gehe es um die „Anwendung von Analytik zum Berechnen von Wahrschein-

lichkeiten des Eintretens von Ereignissen“. Etwas eingängiger schreibt SAP, Predictive Analytics seien eine Methode, die dazu dient, wiederkehrende Muster in Daten aufzudecken und dann mithilfe bestimmter Algorithmen deren weitere Entwicklung zu prognostizieren. Vereinfacht ausgedrückt: Predictive Analytics beschäftigen sich mit zu erwartenden und möglichen Ereignissen und wollen herausfinden, was passieren kann und warum.

Predictive Analytics werden heute in vielen Bereichen eingesetzt, so beispielsweise für die Qualitätssicherung in internen Produktionsabläufen, bei der Steuerung von Produktionsanlagen, bei Instandhaltung/Wartung, Produktions- und Bedarfsplanung, bei der Lieferkettensteuerung und der externen Qualitätssicherung von Produkten/Anlagen. Der Mehrwert von Predictive Analytics liegt – branchen-

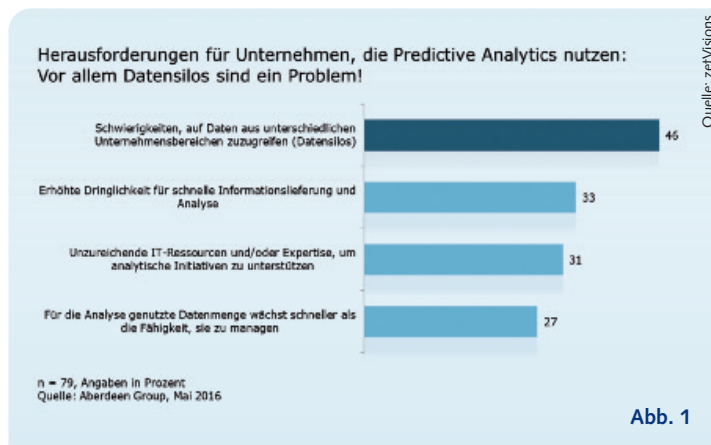


Abb. 1

übergreifend – beispielsweise in der Steigerung der Termintreue und Lieferfähigkeit, in der Ursachenforschung für Fehlerquellen, Ausfälle und Störungen, der besseren Planbarkeit von Bedarfen, der Verbesserung der Qualität in der Produktion und in der besseren Vorhersagbarkeit und Planung von Wartungseinsätzen.

Gute Datenqualität

ist ein kritischer Erfolgsfaktor

Es fällt nicht schwer, sich vorzustellen, dass sich, gleich welcher Vorteil, er sich auch immer nur dann einstellen wird, wenn zunächst eine Grundvoraussetzung erfüllt ist: Die Datenqualität muss stimmen. Zwar liegen zahlreiche relevante Informationen für Predictive Analytics in Form von unstrukturierten Daten vor allem in Textform vor, gleichwohl spielen Stammdaten eine nicht weniger gewichtige Rolle. „Der erste Schritt jeder Datenauswertung besteht im Sammeln und Konsolidieren der Daten, die beispielsweise Anlagen und Infrastrukturen erzeugen. Zu diesen Informationen gehören Stammdaten wie Anlagenidentifikationsnummern (Asset ID) sowie Messwerte oder Statuscodes mit Zeitstempeln. Diese Grunddaten zu sammeln und in ein geeignetes Werkzeug zur Speicherung und Visualisierung zu überführen, kann schon den ersten wesentlichen Mehrwert darstellen: Daten werden zentral gespeichert und einsehbar, Datenverläufe transparent und vergleichbar, zeitliche Zusammenhänge im Auftreten bestimmter Stadien zum ersten Mal beobachtbar“, so Michael ▶

*Monika Pürsing ist CEO der auf die Softwareentwicklung für Stammdaten- und Beteiligungsmanagement spezialisierten zetVisions AG in Heidelberg. Sie hat zahlreiche internationale Projekte zu Stammdatenmanagement-Lösungen geleitet und Kunden bei der Prozessgestaltung beraten.

Matzer in dem E-Book „Predictive Analytics“. Das Problem: Nicht wenige Unternehmen haben mit der Qualität ihrer Daten erhebliche Probleme. Das haben zahlreiche Untersuchungen in den letzten Jahren immer wieder gezeigt, zuletzt die Lünenonk-Studie „Revival der Stammdaten“ im Dezember 2016. Danach hat sich die Datenqualität in den letzten fünf Jahren

zwar deutlich verbessert, zufrieden sind die Unternehmen aller Größenklassen aber nicht. 40 Prozent sagen, ihre Datenqualität sei aktuell sehr gut (16 Prozent) oder gut (24 Prozent). Aber immerhin 60 Prozent betonen hier nur Mittelmaß. Mittelmaß reicht aber keinesfalls. Die Aberdeen Group hat deutlich gemacht, dass entgegen der Annahme, Predictive Analytics habe vor allem mit besonders fortschrittlichen Analysemethoden- und -aktivitäten zu tun, ihr Erfolg oder Misserfolg von einem außerordentlich simplen Mantra abhängt: „garbage in, garbage out“. „Companies still depend on clean, timely, and relevant data to generate insights – predictive or not“, so Aberdeen. Leider haben Unternehmen, die Predictive Analytics nutzen, mit einigen Problemen in ihrem Datenumfeld zu kämpfen (s. Abb. 1).

Problem Nummer eins ist demnach die Fragmentierung und Streuung von Daten über zahlreiche Datensilos hinweg. Da mehr und mehr die Geschwindigkeit der Auswertung den Mehrwert der Daten erzeugt, erweist sich der reduzierte oder zumindest zeitverzögerte Zugriff auf verstreute Daten als ernsthaftes Hindernis. Beispielsweise sind Kundendaten in der Regel in verschiedenen Datensilos gespeichert: CRM, Systeme für Vertrieb und Marketing, für Automation, Tickets und Services, ERP, Anwendungen für Buchhaltung und Controlling.

Multidomain-MDM schafft „Golden Record“

Predictive Analytics können in diesem Umfeld nur funktionieren, wenn es ein professionelles Stammdatenmanagement gibt. Nach einer Untersuchung des Fraunhofer IAO setzt sich für das Management

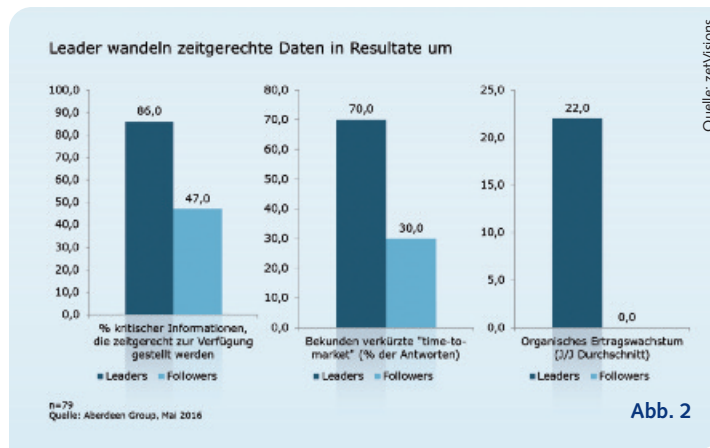


Abb. 2

von Stammdaten der Trend zu Multidomänen-Lösungen (Multidomain-MDM) weiter fort. Unternehmen stehen vor der Aufgabe, einen integrierten Überblick über alle Kunden-, Produkt-, Lieferanten- und weiteren Stammdaten (etwa Referenzdaten) herzustellen. Unternehmen mit nur einer Version der „Wahrheit“ und einem Master Data Record für jede Schlüsseldomäne verbringen weniger Zeit damit, Informationen zu suchen oder die Verlässlichkeit vorhandener Daten zu bestätigen und haben mehr Zeit für relevante Analysen.

Multidomain-MDM zentralisiert das gesamte Stammdatenmanagement. Alle relevanten Daten vom Einkauf bis zum Verkauf laufen in einem zentralen System zusammen. So lassen sich ein „Golden Record“ für zum Beispiel Kunden, Produkte, Lieferanten bilden sowie Verbindungen und Korrelationen zwischen diesen Domänen erkennen. Man erhält einen Rundumblick auf die Stammdaten über alle Domänen hinweg. Ein Multidomain-MDM schafft somit die „eine Wahrheit“ für unterschiedliche Stammdatendomänen über den gesamten Geschäftsprozess. Unternehmensweite Zusammenhänge und Wechselwirkungen werden sichtbar – und infolgedessen einerseits nicht selten ein beträchtliches Einsparpotenzial bei Zeit und Kosten und andererseits erfolgskritische Informationen für unternehmerische Entscheidungen.

Entscheidungseffizienz ist ein wichtiger Vorteil des Multidomain-Ansatzes im Stammdatenmanagement.

Nutzen zeitgerechter Informationsbereitstellung

Wer mit Predictive Analytics – oder jeder anderen Form von Business Intelligence – spürbare geschäftliche Performance erreichen will, braucht erst einmal effiziente Entscheidungsprozesse. „Users need access to clean, relevant, and timely data in order to make effective decisions“, weiß die Aberdeen Group. Stammdatenmanagement auf Basis eines Multidomain-MDM-Systems ist der Schlüssel dazu. Unternehmen, die so ihre Daten „im Griff“ haben und in der Lage sind, den Informationsfluss innerhalb des Unternehmens zu beschleunigen, verkürzen ihre Produkteinführungszeit (time to market) und generieren Möglichkeiten für zusätzliches Ertragswachstum, wie eine Untersuchung von Aberdeen zeigt (s. Abb. 2).

Entscheidungseffizienz, schnellere Produkteinführungszeiten und Ertragswachstum entständen nicht von allein. Die in Predictive Analytics führenden Unternehmen zeigten einen höheren Grad der Nutzerzufriedenheit bei einigen kritischen Aspekten ihres Daten- und Entscheidungsumfelds (s. Abb. 3). Einmal mehr zeichne die Top-Performer die Fähigkeit aus, Informationen zu teilen, sie schneller bereitzustellen und Nutzer mit den Informationen, die sie benötigen, zum richtigen Zeitpunkt zu versorgen. „However, it's also worth noting the things that can impede the flow of information and prevent timely insights. One of the biggest barriers is data quality. When information is corrupt, fragmented, or simply absent, time is wasted trying to correct and validate the data, severely slowing down the analytical process.“ Unternehmen, die mit Predictive Analytics Zusammenhänge, Muster und Korrelationen in Daten herausarbeiten wollen, um mögliche Ereignisse und Entwicklungen zu prognostizieren, müssen das im Auge behalten. (ap) ©

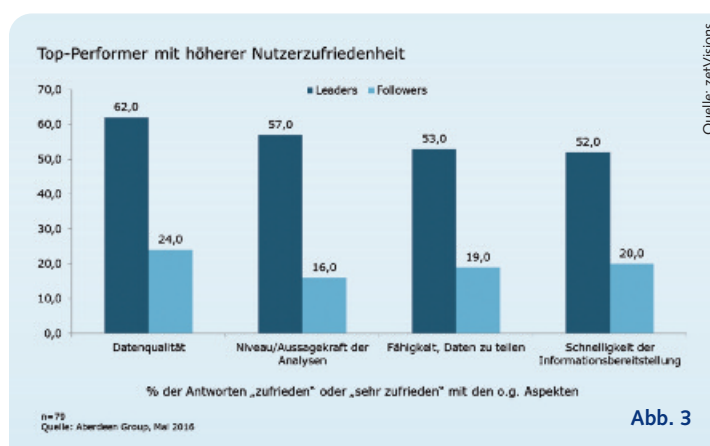


Abb. 3