

► MANAGEMENT

Data-Driven-Company.

Das technisch noch so perfekte Sammeln und Speichern von Daten allein garantiert Unternehmen keine besseren Informationen. Eine Information Supply Chain kann den Weg zu einem echten Nutzwert ebnen.

► Von Christoph Hein und Frank Hendricks

Die Menge elektronisch verfügbarer Daten steigt jährlich exponentiell. Der Bedarf an Speichermedien somit auch. Und die Speicherdichte von Festplatten oder SSDs ist am Rande des physikalisch Möglichen. Weiteren Bedarf erzeugt das Internet of Things, weil es mit seinen in jeden Winkel des Lebens vordringenden Sensoren die Flut an Daten zusätzlich kräftig steigen lässt.

Aber beinhaltet jeder Messwert auch eine Information? Das trifft insoweit zu, als dieser einen Status zu einem bestimmten Zeitpunkt wiedergibt. Aber ist diese Information wirklich wichtig und damit auch speicherwürdig?

Nehmen wir als Beispiel einen Verschleißsensor für einen Bremsbelag. Kontinuierlich meldet er die Stärke des Be-

lags. Diese Daten sind für einen Ingenieur, der den Verschleiß und damit den voraussichtlichen Wechselzeitpunkt berechnen möchte, sicherlich sinnvoll. Für den Autofahrer sind sie es nur bedingt. Ihm reicht in der Regel die Information, wann der Belag zu wechseln ist.

Erwartungshaltung bestimmt den Preis. Der Zweck also entscheidet über die Sinnhaftigkeit gespeicherter Daten und ihre Transformation zur Information. Hinzu kommt, dass ein Großteil der heute vorgehaltenen Daten redundant ist.

Ein Video etwa, das ich kaufe und auf meinen Laptop herunterlade, verdoppelt den Platzbedarf, weil eine Kopie auch auf dem Server gespeichert ist; sie vermehrt aber nicht die Information. Die Menge an

Informationen ist also in jedem Fall kleiner als das entsprechende Datenvolumen.

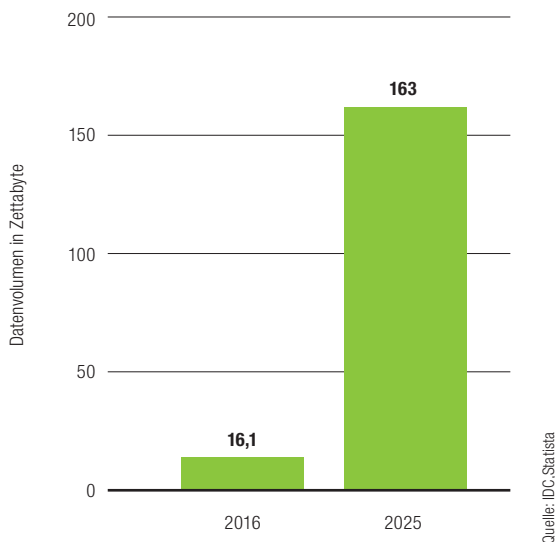
Und was bestimmt darüber hinaus den Wert einer Information? Antwort: allein ihr Verwendungszweck. Marktwirtschaftlich ausgedrückt könnte man sagen, dass die Erwartungshaltung hinsichtlich des Werts einer Information und die damit verknüpfte Nachfrage den tatsächlichen Preis bestimmen.

Und dieser in den Daten verborgene Schatz ist Gold wert. Eine raubkopierte Version der siebten Staffel von «Game of Thrones» beispielsweise hätte HBO in den Ruin treiben können. Somit haben sogar redundante Informationen einen potentiellen Wert. Doch wie sichert man ihn?

Wie findet man in einem Haufen Flusssand die wertvollen Nuggets? Die klare ►

Das Informationsangebot wächst exponentiell.

Volumen der jährlich generierten digitalen Datenmenge weltweit 2016 und 2025 (Prognose) in Zettabyte.



Die Präferenzen der Unternehmen bei der Datennutzung.

Obwohl das Datenmanagement als sehr wichtig gilt, wird es für Entscheidungsfindungen als weniger bedeutend angesehen.

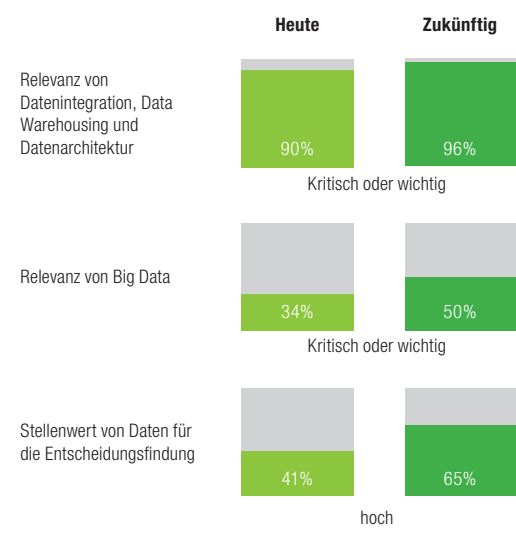
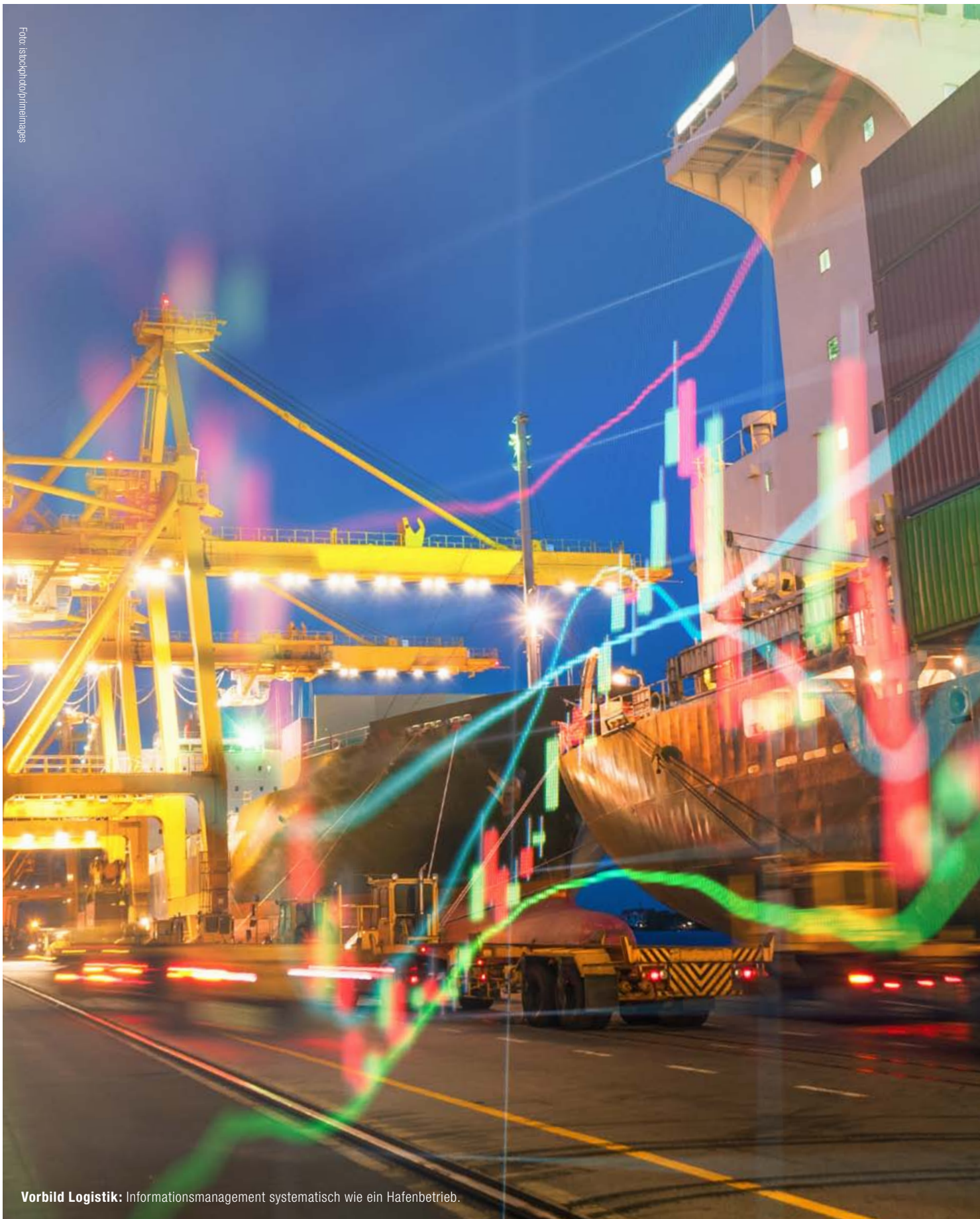


Foto: iStockphoto.com/merimages



Vorbild Logistik: Informationsmanagement systematisch wie ein Hafenbetrieb.

Antwort lautet: gar nicht. Genauso könnte man Michelangelo unterstellen, er hätte aus dem Block Marmor alles wegehauen müssen, was nicht nach der Statue von David aussah.

Daher ist die Informationsgewinnung, in volkswirtschaftlicher Hinsicht, eben kein Primärsektor, sondern ein produzierendes Gewerbe. Die Wertschöpfung liegt in der systematischen Erhebung, der Montage zu Informationen und der Lieferung an den Adressaten.

Aus dieser anderen Perspektive wird deutlich, warum der für die Informationsbeschaffung übliche Begriff des «Data Mining» eigentlich der falsche Terminus ist und man hier besser von «Information Production» sprechen sollte. Die Gewinnung von Information aus bestehenden Daten ist im besten Sinne ein kreativer Prozess, der aus vorhandenen Rohstoffen etwas Neues schafft.

Wie können nun Unternehmen eine solche Information Value Chain aufsetzen? Jedes Unternehmen muss seine eigenen wie auch externe Daten entlang von Geschäftsprozessketten konsistent sammeln, analysieren und mit zusätzlichen Informationen anreichern. Darüber hinaus besteht die Aufgabe der Unternehmen darin, dafür zu sorgen, dass die Daten sicher und nachhaltig gespeichert sowie in der geforderten Qualität und zum richtigen Zeitpunkt anderen Prozessen oder An-

wendern mit dem optimalen Nutzwert zur Verfügung gestellt werden können.

Datenmanagement reicht nicht.

Jedoch ist es mit diesem Datenmanagement allein nicht getan. Laut einer Umfrage des BARC-Instituts etwa verfügt zwar ein Großteil der Unternehmen über ein ausreichendes Datenmanagement, zieht aber die so gewonnenen Informationen nur geringfügig zur Entscheidungsfindung heran.

Das heißt nichts weiter, als dass die Unternehmen zahlreiche Daten sammeln und aufbereiten, ihnen aber nicht ausreichend vertrauen, um daraus fundierte Entscheidungen zu treffen. Sie sind noch ein gutes Stück entfernt davon, eine wirkliche «data-driven company» zu werden.

Erschwerend kommt hinzu, dass sie sich auf dem Weg dorthin auch mit gänzlich neuen Analyseverfahren auseinandersetzen müssen. Sei es nun Methoden wie «Data Discovery», «Data Visualizati-on», «Data Mining», «Predictive Analytics» oder eine der vielen anderen Optionen in diesem Umfeld. Von den Verheißungen der künstlichen Intelligenz mit den Methoden des «Machine Learning» und «Deep Learning» einmal ganz abgesehen.

Versorgungskette entscheidet.

Gutes Management der Daten allein führt jedoch nicht zu einem Erkenntnisgewinn.

Die umliegenden Prozesse müssen mitbe-dacht werden. Die besten Daten helfen nichts, wenn daraus keine Informationen gewonnen werden. Oder wie es der Zu-kunftsforscher John Naisbitt einmal tref-fend formulierte: «Wir ertrinken in Daten und dürsten nach Informationen.»

Aber wie gelingt der Ausweg aus die-sem Dilemma? Die Antwort lautet sehr pragmatisch: Analog zur Lieferkette im Logistikbereich ermöglicht es die Infor-mation Supply Chain (ISC), eine ganzheit-liche Betrachtung der Informationsprozes-se vorzunehmen. Dazu müssen sowohl die internen als auch die externen Informati-onsflüsse zwischen verschiedenen Berei-chen oder sogar unterschiedlichen Unter-nehmen zunächst dargestellt und später konsequent gesteuert werden.

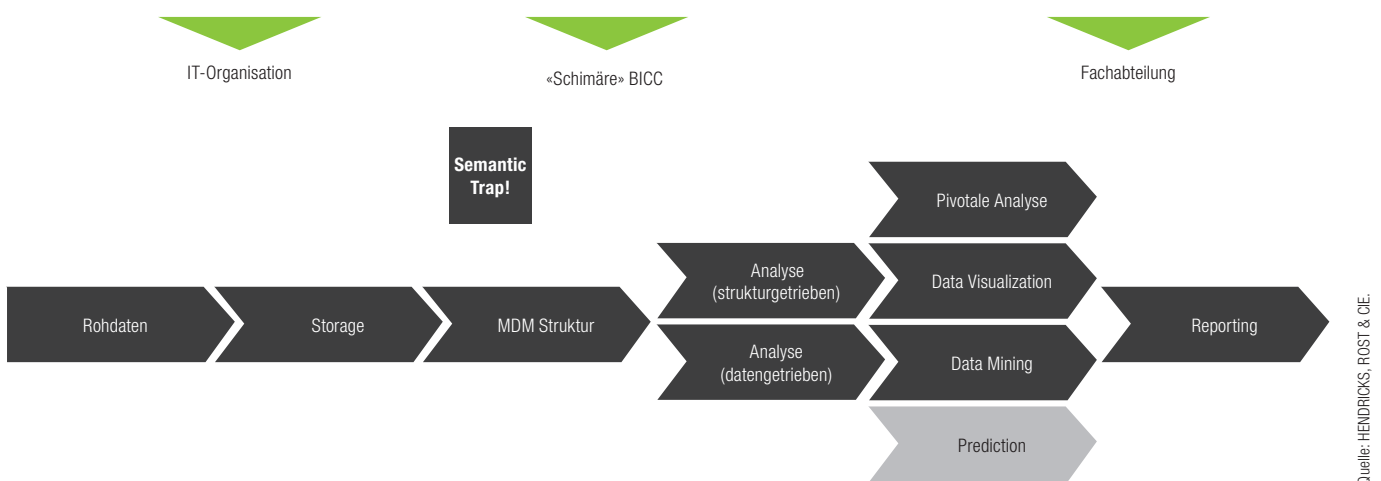
Der Fokus liegt hierbei auf der Gesamt-betrachtung, die von der Datenerstellung bis hin zur Entscheidung reicht. Letztere wird auf der Basis der gewonnenen Infor-mationen getroffen - und umfasst Aufga-ben aus verschiedenen Unternehmensbe-reichen, von der IT bis zum Management.

Ebenso bringt dieses Vorgehen einen Querschnitt durch verschiedene Metho-den und Technologien mit sich, ganz gleich, ob es sich um das etablierte Berichtswesen oder exploratives Data Mining handelt.

Wichtig ist in diesem Prozess, dass nicht die Methode im Mittelpunkt steht, sondern stets die Maximierung des Nut-

Die Wertschöpfungskette für die Transformation von Daten in praktisch verwertbare Informationen in den Unternehmen.

Die Informationsversorgungskette muss präzise definiert und kontrolliert werden. Entscheidend ist die Definition des Master Data Managements (MDM). Klare Zuordnungen vermeiden Schäden durch die «Semantische Falle». Die Existenz eines BI-Competence Center (BICC) reicht nicht.



zens für den Informationsempfänger: eine Zielsetzung, die die ISC mit der klassischen Supply Chain in der Logistik teilt.

Strukturierte Daten als Basis.

Die Informationsproduktion beginnt mit der Vorgabe einer Struktur, in deren Rahmen die Daten erhoben werden sollen. Ähnlich einer Stückliste wird beschrieben, wie Daten beschaffen sein müssen, um einer späteren Verwendung zugeführt werden zu können.

Diese Stückliste kann zum Beispiel ein Formular oder eine Maske für die Kontierung in der Finanzbuchhaltung sein. Nehmen wir die Erfassung eines Spesenbelegs: Die für die Transaktion notwendigen Einträge wie etwa Datum, Betrag, Währung, Konto, Kostenstelle oder Kostenträger sind Werte, mit denen eine Transaktion für spätere Analysen vorbereitet wird.

Solche Analysen sind struktur- oder besser hypothesengetrieben. Die Definition der Struktur basiert auf einer Hypothese darüber, was das Management mit den Daten und Informationen vornehmen möchte. In den meisten Fällen sind es Vergleiche, Zeitreihen, Budgetabweichungen, Deckungsbeiträge und ähnliches. Bei der Datenanalyse wird recherchiert, was Abweichungen von der Hypothese verursacht. Dies kann mit den vorliegenden Daten untersucht werden. Alle modernen OLAP-Systeme funktionieren auf diese Art der Ad-hoc-Analyse.

Die «Semantische Falle».

Allerdings gibt es bei der hypothesengetriebene Analyse mit solcherart strukturierten Daten ein Problem: den eingeschränkten Erkenntnisgewinn. Denn die Daten liefern nur die Information, die in der Struktur angelegt ist. Wir nennen dies die «Semantische Falle» («semantic trap»).

Am bereits genannten Beispiel der Kontierung eines Belegs wird diese Gefahr schnell deutlich: Einige Informationen werden nicht erfasst, weil sie für einen bestimmten Zweck nicht benötigt wurden. Die Bedeutung des Datensatzes ist starr festgelegt, eine andere Deutung damit technisch nicht mehr möglich.

Eine derartige Begrenzung kann fatal sein. Denn vielleicht hatte der Beleg noch viel mehr Informationen zu bieten, wie zum Beispiel eine Uhrzeit, einzelne Positionen mit ihren Preisen, den Ort und viele

weitere Daten, die analyserelevant sein könnten: Passt die Anzahl der Gerichte mit der Anzahl der bewirteten Personen zusammen? Sind gar Kindermenüs darunter? Fiel das Datum auf ein Wochenende? Dies alles sind Informationen, die etwa auf einen Spesenbetrug hinweisen könnten, der schon manchen Mitarbeiter den Job gekostet hat.

Die in diesem Fall beispielhaft angeführten Compliance-relevanten Informationen werden von den Unternehmen in der Regel nicht systematisch erfasst. Aber auch anderweitige Informationen ließen sich gegebenenfalls gewinnbringend einsetzen: Wenn etwa viele Mitarbeiter dasselbe Lokal regelmäßig besuchen, könnte das Unternehmen einen Firmenrabatt aushandeln und somit Kosten sparen.

Wissenschaftler statt Alchimisten.

Solche wertschöpfenden Analysen setzen lediglich eine komplette Erfassung des Belegs voraus und stellen keine allzu hohen Anforderungen an die Datenanalyse. Die heutigen Daten-Alchimisten, für die sich der Berufsbegriff «Data Scientist» etabliert hat, können mithilfe mathematisch-statistischer Werkzeuge sowie Methoden der künstlichen Intelligenz mehr Informationen aus den Daten extrahieren, als ein einzelner Datensatz auf den ersten Blick zu liefern vermag.

Diese Verfahren und Tools bedürfen allerdings eines fundierten Wissens, um Einsatzmöglichkeiten, Grenzen und Güte-

verfahren zur Qualitätssicherung der Ergebnisse einordnen zu können. Künstliche Intelligenz (KI) – englisch «Artificial Intelligence» (AI) – wird vielerorts als der Stein der Weisen der Neuzeit angepriesen.

Doch was bringt dieses neue Feld in der betrieblichen Praxis? Betrachten wir die Software Watson von IBM oder die Algorithmen hinter Google und Facebook, so können aus den Daten durch Aggregation und Mustererkennung Profile abgeleitet werden, die einen Mehrwert für die Wirksamkeit von Maßnahmen versprechen. In der Fertigung etwa ist es die Optimierung von Maschinen; im Falle von Google ist es die individualisierte Werbeansprache.

Fazit: Mehrwert aus Informationen.

Für die Unternehmen, wie beispielsweise wir sie beraten und für die wir Lösungen aufsetzen, ergibt sich auf dieser Grundlage in allen Branchen eine Vielzahl von Einsatzoptionen, mehr Profit aus den vorhandenen Daten zu generieren: Unter anderem können sie das Kaufverhalten viel besser prognostizieren, Abwanderungen von Kunden verhindern, Lieferantenanalysen und -bewertungen aufstellen, High Potentials in den Reihen ihrer Mitarbeiter entdecken oder auch Betrugsfälle aufspüren («Fraud Detection»).

Die Möglichkeiten sind vielfältig; entscheidend ist die Kreativität im Umgang mit den Daten. Eine gute Information Supply Chain bildet die Grundlage für eine verlässliche Wertschöpfung. ■



› Christoph Hein ist Berater und «Data Nerd» bei dem Düsseldorfer Consultingunternehmen HENDRICKS, ROST & CIE. Er beschäftigt sich seit vielen Jahren intensiv mit der Zukunft des Datenmanagements in Unternehmen verschiedener Branchen – beispielsweise mit Themen wie Industrie 4.0 oder Big Data, Predictive oder Prescriptive Analytics.



› Frank Hendricks arbeitet seit 1993 als Unternehmensberater und ist geschäftsführender Gesellschafter von HENDRICKS, ROST & CIE. Seine Schwerpunkte sind die Prozess- und Organisationsberatung, Controlling und Kostenrechnung, BI-Systeme sowie Data Warehouse. Vor seiner jetzigen Aufgabe war er unter anderem für Ernst & Young/Cap Gemini tätig.