

# Mit Big Data Risiken effektiv identifizieren

---

Der Einsatz von Big-Data-Technologien bei der Risikoidentifikation entlastet Mitarbeiter, verbessert die Qualität von Prognosen und spart Kosten. Gleichwohl bleibt es wichtig, Daten als Entscheidungsgrundlage kritisch zu hinterfragen.

*Patrick Eckstein, Thomas Hartung, Nadine Rohatsch*

Globalisierung und Digitalisierung machen das Unternehmensumfeld dynamischer und weniger vorhersehbar, das hat nicht zuletzt die Corona-Pandemie deutlich gemacht. Umso essenzieller ist es für die Unternehmenssteuerung, mögliche beziehungsweise tatsächliche Risiken frühzeitig identifizieren zu können, um den nachhaltigen Erfolg und das Überleben des Unternehmens zu sichern. Big Data – große Datenmengen und ihre Analysen – erleichtert diese Aufgabe zunehmend, wie Beispiele aus dem Gesundheitswesen und dem Supply Chain Management zeigen.

Big Data spielt in der Unternehmenssteuerung eine immer wichtigere Rolle. Zwar ist der Begriff selbst nicht eindeutig definiert, grundsätzlich verbindet die Forschung damit aber die „Generierung großer Datenmengen“ (Schröder et al. 2019, S. 122) aus internen sowie externen Quellen. Der Begriff meint oftmals zugleich auch die Analyse dieser Daten, um neue Erkenntnisse zu gewinnen, Risiken zu minimieren und unternehmerische Entscheidungen zu optimieren (vergleiche McAfee/Brynjolfsson 2012, S. 4 ff.; Hertweck/Kinitzki 2015, S. 15 f.). Das Konzept von Big Data ist vor allem durch drei grundlegende Attribute, die sogenannten 3 V, gekennzeichnet, die auch für ein erfolgreiches Risiko-Management eine wesentliche Rolle spielen: Volume (Datenmenge), Velocity (Geschwindigkeit) und Variety (Vielfalt) (vergleiche Kudyba/Kwatinetz 2014, S. 1 f.).

Volume beschreibt dabei den Umfang an verfügbaren Daten: Unternehmen stehen vor der Herausforderung, eine Masse an Daten und Informationen für ihre Risikoanalysen zu berücksichtigen. Velocity ist damit eng verknüpft: Die Menge entscheidungsrelevanter Daten wächst schnell und kontinuierlich, zudem treten solche Datenmengen immer häufiger auf. Eine manuelle Analyse der Daten kann den immer häufiger eintretenden Veränderungen im Risikoumfeld nicht mehr gerecht werden. Variety bezieht sich auf die eingehenden Informationen: Diese sind oft sehr heterogen, eine einheitliche, händische Bearbeitung ist schwierig zu realisieren.

Diese drei Attribute (3 V) wurden mit der Zeit um zwei zusätzliche erweitert (vergleiche Malik/Singh 2019, S. 260 f.): Veracity (Wahrhaftigkeit) zeigt, dass dem Unternehmen Daten in unterschiedlichen Qualitätsgraden zur Verfügung stehen. Sie müssen deshalb immer überprüft und veredelt werden. Value (Wert) ist eng mit dem Attribut Veracity verknüpft: Nicht jedes zusätzliche Dataset schafft unternehmerischen Mehrwert. Daten ohne Relevanz können zum Beispiel Scheinkorrelationen erzeugen oder die Suche nach relevanten Zusammenhängen erschweren und damit Value reduzieren (vergleiche Kudyba/Kwatinetz 2014, S. 13).

Der Umgang mit diesen Attributen ist essenziell, um verlässliche Entscheidungsgrundlagen für Risiko-Manager zu

schaffen. Während die Attribute Volume, Velocity und Variety inzwischen mittels moderner Software und Rechenkapazitäten standardisiert bewältigt werden können, stellen Value und Veracity Unternehmen noch vor Herausforderungen. Sie brauchen dafür aufwendigere Lösungen, die Datenexpertise mit tiefgehendem Branchenfachwissen verbinden. Der Bewertungsalgorithmus beispielsweise muss überwacht und durch Experteneinschätzungen verbessert werden, um aus Fehlern zu lernen. Nur wenn Analysen anhand sinnvoll ausgewählter und verifizierter Daten verwendet werden, können Risiken effizient identifiziert werden.

### Risikoidentifikation mit Big Data

Bisher war Risikoidentifikation eine personalintensive Aufgabe, die sich stark auf Expertenwissen und Erfahrung gestützt hat. Inzwischen kann ein Großteil davon mittels digitaler Auswertung von Daten und Prozessen erledigt werden (vergleiche Schröder et al. 2019, S. 125). Das entlastet nicht nur personelle Ressourcen und spart damit Kosten ein, sondern reduziert auch die Anfälligkeit für Fehler im gesamten Risiko-Management-Prozess. Zudem können Risiken in Echtzeit identifiziert werden. Dies ermöglicht es Risiko-Controllern, schnell auf Veränderungen in der Risikolandschaft hinzuweisen und damit die Qualität ihres Berichtswesens zu steigern. Die Unternehmensführung kann auf deren Berichte schnell reagieren, erforderliche Maßnahmen ohne Verzögerung einleiten und somit das Risiko-Management effektiver machen.

Ein datenbasiertes Risiko-Management und insbesondere eine datenbasierte Risikoidentifikation tragen dazu bei, dass Ursache und Wirkungen der Risiken schneller und präziser erfasst und vorhergesagt werden können. Ein Ad-hoc-Monitoring kann kleinste Veränderungen in der Risikosituation kontinuierlich erfassen – beispielsweise mittels Sensortechnologie – und macht damit eine konsistente und zeitnahe Risikoidentifikation möglich. Wirkungsvolle Frühwarnsysteme können erstellt werden. Für das Risiko-Controlling bedeutet das eine zunehmende Verlagerung der Identifikation von Risiken aus dem zeitlich nachgelagerten, nur erkennenden in den prädiktiven Bereich.

Tatsächlich ist eine notwendige, kontinuierliche Risikoidentifikation heute überhaupt nur mit geeigneten Big-Data-Lösungen umsetzbar, weil datenbasierte Identifikationsprozesse weder vollständig noch effizient durch Mitarbeiter repliziert werden können (vergleiche Diederichs 2017, S. 134). Unternehmen brauchen deshalb eine Big-Data-Infrastruktur, die interne Systeme um externe Daten erweitert, Informatio-

### Zusammenfassung

- Die Verfügbarkeit großer Datenmengen und ihre Analyse (Big Data) liefert für das Risiko-Management erheblichen Nutzen.
- Eine datenbasierte Risikoidentifikation kann Ursache und Wirkung der Risiken genauer und mit höherer Aktualität erfassen.
- Die Transparenz wird erhöht und ermöglicht zielgerichtete risikopolitische Maßnahmen und Kosteneinsparungen.

nen bündelt und auswertet. Nur so können die Daten verfügbar gemacht und effizient analysiert werden. Für die automatisierte Auswertung ist eine lernfähige Künstliche Intelligenz hilfreich, die Risiken auf Basis der Daten identifizieren und Entwicklungen prognostizieren kann.

Wie Big Data bereits für die Risikoidentifikation genutzt wird, zeigen Beispiele aus dem Gesundheitssektor und dem Bereich der Supply Chain.

### Beispiel Gesundheitssektor

Zu den primären Zielen im Gesundheitssektor gehört es, Krankheiten und ihre Ursachen frühzeitig zu erkennen, Patienten individuell optimiert zu behandeln und somit vermeidbare Kosten im Gesundheitswesen zu sparen. Big-Data-Anwendungen können die Risikoidentifikation im Gesundheitsbereich deutlich verbessern.

Voraussetzung dafür ist, dass relevante Daten digital verfügbar sind. Mit der Einführung der elektronischen Patientenakte 2021 eröffnen sich neue Möglichkeiten. Der Abgleich patientenspezifischer Daten mit bestehenden anonymisierten Datenbanken ermöglicht tiefgehende automatisierte Auswertungen – etwa zur Verbreitung von Krankheiten, zu Einflussfaktoren und Gesundheitsmerkmalen. Ebenso können öffentlich zugängliche Daten aus dem Internet oder von Social-Media-Kanälen miteinbezogen werden (vergleiche Bates et al. 2014, S. 1124). Gleichzeitig eröffnen diese Möglichkeiten einen Qualitätssprung im Krankenhaus-Risiko-Management und -Controlling. Bei der Durchführung von Big-Data-Analysen mit Gesundheitsdaten von Patienten sind jedoch insbesondere seit der Einführung der Datenschutz-Grundverordnung komplizierte Regeln zu beachten.

In den folgenden Bereichen macht sich der Gesundheitssektor Big Data bereits zunutze:

**Risikopatienten:** Ein großer Anteil der Gesundheitsaufwendungen entfällt auf eine kleine Anzahl von Patienten. Eine frühzeitige Identifikation dieser Risikopatienten anhand eines Abgleichs individueller Messwerte mit großzahligen, aber auch granular aufgebauten Datenbanken ermöglicht, dass ihr Gesundheitszustand laufend beobachtet und prognostiziert werden kann. Eine bedarfsorientierte Betreuung sowie Präventionsmaßnahmen werden dadurch leichter möglich.

**Alarmzeichen/Signale:** Der Zustand gefährdeter Patienten wird durch eine Reihe unabhängiger Systeme überprüft, die jedoch oft falsche Alarme auslösen. Werden die in solchen Systemen gesammelten Daten verbunden und durch einen lernenden Algorithmus automatisiert und in Echtzeit ausgewertet, können Alarme zuverlässiger werden.

**Tracking-Apps:** Im Rahmen der COVID-19-Pandemie werden in verschiedenen Ländern smartphonegestützte Tracking-Apps eingesetzt, um mögliche Ansteckungsketten zu identifizieren. Neben individuellen Bewegungsdaten werden damit auch registrierte Kontakte aus einer Vielzahl von Bewegungsmustern ausgewertet. Ziel ist es, Verdachtsfälle frühzeitig zu identifizieren und gegebenenfalls zu isolieren.

**Prognosen zu Entlassungen:** Ein Kostentreiber ist die Wiedereinlieferung von behandelten Patienten, bei denen erneut eine Krankheit auftritt. Auf Basis historischer und aktueller Daten lässt sich der Zustand der Patienten analysieren und ihr Krankheitsverlauf besser abschätzen. Eine Prognose wird möglich, die die Entscheidungen der Ärzte über eine Entlassung des Patienten unterstützen kann.

**Prognosen zu Wiedereinlieferung:** Im Bereich der Wiedereinlieferung wird bereits ein Algorithmus genutzt, um Patienten mit Herzversagen besser einzuordnen (vergleiche Evans et al. 2016, S. 872 ff.). Basierend auf Vergangenheitsdaten berechnet der Algorithmus aus relevanten Parametern Indikatoren zur Sterbewahrscheinlichkeit aufgrund von Herzversagen und zur ursachenunabhängigen Wiedereinlieferungswahrscheinlichkeit des Patienten in den nächsten 30 Tagen. Datengrundlage sind die elektronischen Patientenakten sowie Berichte ärztlicher Untersuchungen, die mittels Methoden zur maschinellen Verarbeitung natürlicher Sprache, dem sogenannten Natural Language Processing, ausgelesen wurden. Die so berechneten Indikatoren wurden in Besprechungen genutzt, um die Entlassung von Patienten zu planen. Auch die Dauer ärztlicher Untersuchungen konnte drastisch gesenkt werden. Mit dem Algorithmus konnten zudem Hochrisikopatienten besser identifiziert und Kosten eingespart werden, indem Patienten mit geringer Gefährdung frühzeitiger entlassen werden konnten.

## Beispiel Supply Chain Risk Management

Die Wertschöpfungsketten befinden sich aufgrund des technologischen Wandels und stetig neuer Nachfrage-Trends wie der zunehmenden Individualisierung von Produkten im Umbruch. Umso wichtiger wird es für Unternehmen, Risiken in ihren Supply Chains zu identifizieren und frühzeitig gegensteuern zu können. Big Data schafft hier neue Möglichkeiten.

Eine der Hauptursachen für Risiken war bisher die Intransparenz von Supply Chains. Informationen sind oftmals nur für einzelne Teile der Supply Chain verfügbar, ein Gesamtüberblick für alle fehlt. Treten Komplikationen in einem Teil der Supply Chain auf, bleiben andere Teile unwissend und haben keine Möglichkeit, eventuelle Risiken für den eigenen Geschäftsbetrieb zu identifizieren. Die Digitalisierung von Prozessen ebenso wie neu gewonnene Daten aus maschinellen Prozessen wie beispielsweise Radio-Frequency-Identification-Sensorik, die kontaktlose Übertragung von Daten über Funk, erhöhen die Transparenz bestehender Wertschöpfungsketten.

Um der Intransparenz entgegenzusteuern, ist eine End-to-End-Transparenz beziehungsweise Visibility in der gesamten Supply Chain notwendig. Sie kann durch die Nutzung von Big Data erreicht werden. Dazu werden wesentliche interne prozess- und partnerrelevante Daten systematisch gesammelt, ausgewertet und durch externe Daten unter anderem aus zahlungspflichtigen oder öffentlich zugänglichen Quellen zum Beispiel aus Social-Media-Kanälen oder zum Wetter ergänzt. Diese Verknüpfung verschiedener interner und externer Datensätze ermöglicht eine zuverlässige Identifikation sowohl von Risiken, die endogen in der Supply Chain bestehen, als auch von exogenen Risiken, die von außen auf das Unternehmen einwirken, wie beispielsweise Wetterereignisse.

Die verfügbaren Daten erlauben eine effizientere Prognose zukünftiger Risiken, sie bilden die Basis für leistungsfähigere Frühwarnsysteme und ein aussagekräftiges Risikoberichtswesen.

Hierfür werden die jeweils gesammelten Daten ausgewertet und darauf basierend automatisiert Szenarien erstellt. Um eine Risikostrategie zu entwickeln, wird anschließend deren Einfluss auf die Supply Chain modelliert. Das Supply Chain Risk Management lässt sich also durch eine gezielte Nutzung von großzahligen Daten verbessern.

## Schlussbetrachtung

Die Nutzung von Big Data kann wesentlich dazu beitragen, das Risiko-Management und -Controlling zu verbessern. Durch die anwachsenden Mengen an Daten kann größere Transparenz in Unternehmen hergestellt werden, die die Identifikation

tifikation relevanter Risiken und den Umgang mit diesen Risiken vereinfacht. Risiko-Controller benötigen dafür die Unterstützung von Big-Data-Anwendungen, die die Daten verarbeiten, automatisiertes Monitoring ermöglichen, Risiken selbstständig identifizieren und Veränderungen anzeigen können. Selbstlernende Systeme können laufend aktualisierte Prognosen über entstehende Risiken erstellen und Hinweise auf in der Zukunft eintretende Risiken geben. Die Beispiele zeigen, dass eine automatisierte, datengetriebene Risikoidentifikation bereits möglich ist. Eine breite Datenbasis, über die Unternehmen inzwischen häufig verfügen, kann somit zu einer Professionalisierung der Risikoidentifikation beitragen und einen Teil der inhärenten Unsicherheit reduzieren.

Mit dem Einsatz von Big Data wird es für Risiko-Controller leichter, jederzeit die aktuelle Risikosituation des Unternehmens zu erkennen und Einschätzungen zur Relevanz erkannter Risiken zu treffen. Das wirkt sich positiv auf den gesamten Risiko-Management-Prozess aus.

Es wäre jedoch verfrüht anzunehmen, alle relevanten Risiken könnten bereits automatisiert vollumfänglich und zeitpunktgenau identifiziert werden. Es ist weiterhin notwendig, die Ergebnisse – ebenso wie die Daten – durch Experten validieren zu lassen. Big-Data-Lösungen können damit aktuell das klassische Risiko-Management unterstützen, sie können es jedoch nicht gänzlich ersetzen. Speziell im Bereich ungewisser Zukunftsszenarien ist es noch erforderlich, dass Experten auf Basis von Daten entscheiden und ihre Bedeutung realistisch beurteilen, sich jedoch nicht von ihnen einengen lassen. Der Einsatz von Big Data in der Risikoidentifikation ist dennoch notwendig. Risiko-Controller dürfen sich diesem Trend nicht verschließen, denn manuelle Risikoidentifikation stößt in zunehmend digitalisierten Unternehmen an ihre Grenzen. Mit der Einführung neuer Datenübertragungsstandards und der fortschreitenden Digitalisierung sowie dem technischen Fortschritt werden sich in der Risikoidentifikation neue Chancen bieten – und das Potenzial der Nutzung von Big Data wird weiter zunehmen.

### Literatur

Bates, D. W./Saria, S./Ohno-Machado, L./Shah, A./Escobar, G. (2014): Big data in health care: using analytics to identify and manage high-risk and high-cost patients, in: *Health Affairs*, 33 (7), S. 1123-1131.

Diederichs, M. (2017): *Risikomanagement und Risikocontrolling*, 4. Auflage, München.

Evans, R. S. et al. (2016): Automated identification and predictive tools to help identify high-risk heart failure patients: pilot evaluation, in: *Journal of the American Medical Informatics Association*, 23 (5), S. 872-878.

Fan, Y./Heilig, L./Voß, S. (2015): Supply chain risk management in the era of big data, in: Marcus, A. (Hrsg.): *Design, user experience, and usability: design discourse*, Cham, S. 283-294. [www.springerprofessional.de/link/2489748](http://www.springerprofessional.de/link/2489748)

Hertweck, D./Kinitzki, M. (2015): Datenorientierung statt Bauchentscheidung: Führungs- und Organisationskultur in der datenorientierten Unternehmung, in: Dorschel, J. (Hrsg.): *Praxishandbuch Big Data*, Wiesbaden, S. 15-32.

Kudyba, S./Kwatinetz, M. (2014): Introduction to the big data era, in: Kudyba, S. (Hrsg.): *Big Data, mining, and analytics: components of strategic decision making*, Boca Raton (USA), S. 1-15.

Malik, V./Singh, S. (2019): Cloud, big data & IoT: risk management, in: *2019 International conference on machine learning, big data, cloud and parallel computing, proceeding*, Faridabad, S. 258-262.

McAfee, A./Brynjolfsson, E. (2012): Big data: the management revolution, in: *Harvard Business Review*, 90 (10), S. 1-9.

Schröder, M./Findeis, K./Späth, H. (2019): Neue Anwendungsmöglichkeiten für das Risikomanagement durch den Einsatz von Big Data – Zwei Fallbeispiele, in: Schröder, M./Wegner, K. (Hrsg.): *Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains*, Wiesbaden, S. 121-136. [www.springerprofessional.de/link/16567546](http://www.springerprofessional.de/link/16567546)

### Angaben zu den Autoren



#### Patrick Eckstein

ist Masterabsolvent der Fakultät für Wirtschafts- und Organisationswissenschaften der Universität der Bundeswehr München, Deutschland.  
E-Mail: [patrick.eckstein@unibw.de](mailto:patrick.eckstein@unibw.de)



#### Prof. Dr. Thomas Hartung

ist Inhaber der Professur für Versicherungswirtschaft an der Universität der Bundeswehr München, Deutschland.  
E-Mail: [thomas.hartung@unibw.de](mailto:thomas.hartung@unibw.de)



#### Nadine Rohatsch

ist wissenschaftliche Mitarbeiterin an der Professur für Versicherungswirtschaft an der Universität der Bundeswehr München, Deutschland.  
E-Mail: [nadine.rohatsch@unibw.de](mailto:nadine.rohatsch@unibw.de)