

The
Software
Alliance

BSA

BIG DATA UND WAS ES DAMIT AUF SICH HAT



Vorwort

Softwareinnovationen bringen Tag für Tag neue Fortschritte, die unsere Welt verändern, dem Einzelnen neue Möglichkeiten eröffnen und unsere Wirtschaft beflügeln.

Das volle Potenzial dieses digitalen Wandels lässt sich jedoch nur dann realisieren, wenn wir das Potenzial der Daten nutzen, die diese Innovationen freisetzen. Wir erleben gerade eine Datenrevolution. Vorangetrieben wird diese Entwicklung nicht nur von der heutzutage nahezu endlosen Flut an Daten, sondern vor allem von den grundlegenden Technologien, die unsere Art der Erfassung, Speicherung, Auswertung und Umwandlung von Informationen verändern.

Heutzutage bezeichnen 90 Prozent der führenden Wirtschaftsvertreter Daten als wichtige Ressource und als ein wesentliches Differenzierungsmerkmal für Unternehmen. Sie sind damit ebenso bedeutsam wie grundlegende Ressourcen wie Land, Arbeitskraft und Kapital.

Es ist noch gar nicht so lange her, da man zur Erkennung von Niederschlagsmustern Wetterszenarien über Hunderte von Jahren hinweg beobachten und aufzeichnen musste. Wollte man ein Transportnetz planen, setzte man sich an den Straßenrand, um die Geschwindigkeit des vorbeiziehenden Verkehrs zu erfassen. Auf der Suche nach einem Heilmittel für Krankheiten arbeiteten sich Mediziner durch meterhohe Aktenberge voller handschriftlich festgehaltener Beobachtungen.

Heutzutage werden Daten durch Sensoren an Millionen von Geräten, Maschinen, Fahrzeugen und sogar Straßenlaternen generiert. Die Speicherung dieser riesigen Datenmenge war zunächst kostspielig und schwierig. Inzwischen sind die Speicherkapazitäten gewachsen und die Kosten rapide gesunken – die Datenspeicherung ist zu einer erneuerbaren Ressource geworden. Mit der Fähigkeit, Daten mehrfach und für verschiedene Zwecke zu verwerten, können wir sie immer wieder analysieren und transformieren. So erhalten wir wertvolle Erkenntnisse, die nicht nur Zeit und Geld sparen, sondern sogar Leben retten können.

Manche dieser erfassten Daten sind persönliche Informationen. Darum müssen aktuellste Modelle für Datenschutz und verantwortungsvolle Handhabung eingesetzt werden, um einen sicheren und korrekten Umgang mit diesen Informationen zu gewährleisten. Doch die Mehrheit der Daten entsteht hauptsächlich durch die Kommunikation zwischen Geräten und Maschinen untereinander und mit ihrem Betreiber. Vom Fließband in der Fabrik bis hin zum Passagierflugzeug in der Luft werden Millionen von Bytes an Daten generiert und analysiert. Dadurch lassen sich die Performance verbessern und die Produktivität steigern, und zwar in einem Maße, das früher nicht denkbar gewesen ist.

Obwohl Daten überall sind und ihre Allgegenwart und Nützlichkeit so viele Bereiche unseres Lebens verbessert, verstehen viele Menschen nicht, was genau diese Daten sind, woher sie kommen, wie man sie nutzt und welches Potenzial in ihnen steckt.

Diese Studie stellt anhand einiger konkreter Beispiele dar, wie die Dateninnovation bei einigen der größten Herausforderungen unserer Welt zu

außerordentlichen Fortschritten geführt hat. Sie zeigt auf, wie die wesentliche Veränderung bei der Erfassung, Speicherung, Analyse und Umwandlung von Daten uns an die Schwelle all der Möglichkeiten führt, die uns die digitale Wirtschaft des 21. Jahrhunderts bietet – und wie viel mehr noch möglich ist. Sie beschäftigt sich auch mit einigen der Mythen, die sich verbreiten, während die Menschen weiterhin versuchen, die wachsende Dateninnovations-Branche genauer zu verstehen. Abschließend bietet die Studie ein Glossar zur Terminologie der Dateninnovation als Leitfaden für alle, die sich neu mit dem Thema Datenwirtschaft befassen.

Die Möglichkeiten, welche die Dateninnovation mit sich bringt, sind wahrhaft einzigartig. Innovative Softwaretools revolutionieren unseren Alltag bereits in erstaunlicher Weise. Jetzt helfen diese Tools den Menschen, die Antworten zu finden, die sich in der wachsenden Fülle von Datenquellen verbergen. Diese transformativen neuen Tools übersetzen Daten in neue Produkte, neue Lösungen und neue Innovationen, die unser Leben verändern werden.

Wenn wir heute kluge Entscheidungen treffen, dann kann diese aufstrebende, „datenzentrierte“ Wirtschaft neue Arbeitsplätze und Branchen, neue Durchbrüche und neue Heilmethoden schaffen – und unser Wirtschaftswachstum auf Jahrzehnte hin antreiben.

WAS BEDEUTET „DATENINNOVATION“?

Es wurde viel über die „vier Vs“ der Dateninnovation geschrieben: „Volume“ (die Datenmenge); „Velocity“ (die Geschwindigkeit, in der Daten generiert werden), „Variety“ (die Form der Daten) und „Veracity“ (die Zuverlässigkeit). Aber es wurde deutlich weniger darüber diskutiert, wie wenig Wert eigentlich in Rohdaten steckt – und über die bahnbrechenden Möglichkeiten, die uns allen offen stehen, sie maximal zu nutzen. Wie diese Studie zeigt, müssen Daten erfasst, gespeichert, analysiert und umgewandelt werden, um uns praktische oder sogar lebensrettende Vorteile zu bieten. Diese Prozesse sind das eigentlich Wesen der Dateninnovation – die Schöpfung eines unermesslichen Mehrwerts aus einer Unmenge von ansonsten unproduktiven Informationen.

Inhaltsverzeichnis

- 5** EINLEITUNG
- 7** VON DATEN ZU ANTWORTEN: IN VIER SCHRITTEN
- 14** EINE DATENGETRIEBENE WIRTSCHAFT
- 17** MYTHOS ODER FAKT?
- 28** DIGITALER DISKURS: DIE DATENSPRACHE VERSTEHEN
- 34** ENDNOTEN
- 40** ÜBER BSA



6000 v. Chr.

15. Jhd.



1850er Jahre

21. Jhd.



MEILENSTEINE IN DER GESCHICHTE DER DATEN

Einleitung

Die Meilensteine der Menschheitsgeschichte, von den Anfängen bis hin zu unserer modernen Zivilisation, sind immer markiert durch die Verbesserung unserer Fähigkeit, Beobachtungen zu interpretieren und Informationen zu sammeln. Unsere Vorfahren entwickelten Instrumente, mit denen man Entfernungen, Gewicht, Menge, Temperatur oder Zeit messen und Standorte bestimmen konnte. Jedes dieser Werkzeuge wurde im Laufe der Zeit verbessert, und jedes war entscheidend für unsere Entwicklung von Jägern und Sammlern hin zu Bauern und schließlich zu Stadtbewohnern.

Schon 6000 v. Chr. nutzte man Daten über Ernteerträge und Brachjahre, um die Ernten zu steigern und mehr Menschen zu ernähren. Im 15. Jahrhundert nutzte man astronomische Informationen, um die Welt zu bereisen und die Weltmeere für den globalen Handel zu öffnen. In den 1850er Jahren wurden mithilfe von Daten Cholera-Ausbrüche mit verunreinigtem Wasser in Verbindung gebracht, wodurch Leben gerettet wurden.

In der gesamten Geschichte der Neuzeit haben uns sogar geringe Datenmengen wichtige Erkenntnisse für unerwartete Lösungen für einige unserer größten Herausforderungen geliefert. Ob auf Stelen, Papyrusrollen, in illuminierten Handschriften oder gedruckten Büchern: Daten und ihre zunehmende Verfügbarkeit und Bekanntheit haben die Entwicklung von Mensch und Wirtschaft maßgeblich vorangetrieben.

Im 21. Jahrhundert erfährt dieser Prozess nun eine rasante Beschleunigung. Die Menge an Informationen nimmt beständig zu, und die Kosten ihrer Speicherung sinken auf ein Minimum. Neue Technologien geben Datenwissenschaftlern modernste Werkzeuge an die Hand, mit denen sie wertvolle Erkenntnisse erschließen, die sich in der Masse an Daten verbergen. Je

Daten machen den Unterschied

- + **Barcelona** nutzt Daten, um die Stadt intelligenter zu machen. Mithilfe der erfassten Informationen werden die Reisemuster von Touristen studiert, wird ausgelotet, wo Mietstationen für öffentliche Fahrräder sinnvoll sind, und an welchen Stellen in der Stadt mehr Geldautomaten aufgestellt werden müssen.
- + In den **Vereinigten Arabischen Emiraten** wird mithilfe neuer Datentools das weltweit erste Plusenergie-Gebäude errichtet, das mehr Energie produziert als es verbraucht.
- + In **Kenia** werden mithilfe mobiler Daten Muster für Malariainfektionen erstellt und Infektionsherde identifiziert, um die Behörden bei der Ausrottung der Krankheiten zu unterstützen.
- + Landwirte von **Iowa** bis **Indien** nutzen Daten von Saatgut, Satelliten, Sensoren und Traktoren, um informierte Entscheidungen bezüglich der anzubauenden Pflanzen und des Zeitpunkts der Aussaat zu treffen und zu überlegen, wie man die Frische der Produkte vom Hof bis zum Verbraucher nachverfolgen kann und wie man sich am besten den veränderten klimatischen Bedingungen anpasst.

transformativer die Technologien zur Datenverarbeitung werden, desto nachhaltiger wird ihre Wirksamkeit und desto tiefgreifender werden die Möglichkeiten.

Wir bewegen uns auf eine Zukunft von einem nahezu endlosen Informationsangebot zu, die praktisch unbegrenzte Möglichkeiten bietet. Man bedenke, wie mithilfe von Daten Vorhersagen erstellt werden, die uns den Alltag erleichtern. Dank vorhersagender Daten wissen wir schon im Voraus, ob wir einen Schirm

SAMMELN

SPEICHERN

ANALYSIEREN

UMWANDELN
& UMSETZEN

mitnehmen oder lieber gleich mit dem Bus zur Arbeit fahren sollten. Mithilfe von Verkehrsdaten synchronisieren wir Ampeln, bestimmen die Ankunftszeiten von Zügen und finden den schnellsten Weg, damit unser Kind rechtzeitig zur Theaterprobe kommt. Tragbare Geräte ermitteln unsere persönliche Fitness, damit wir bessere Entscheidungen treffen können, um länger und gesünder zu leben. Wissenschaftler analysieren Terabytes an genetischen Informationen auf der Suche nach neuen Heilmethoden und effektiveren, individuellen Therapien.

Beim Autokauf richten wir uns nun nicht mehr nur nach dem Preisschild, sondern können dank Angaben zu Treibstoffverbrauch, Wartungs- und Versicherungskosten sowie zur Verkehrssicherheit des Modells eine informiertere Kaufentscheidung treffen. Auch das Auto selbst ist heutzutage im Prinzip ein Supercomputer auf Rädern. Es verfügt über einen Prozessor, der durch die Interaktion mit Sensoren die Leistung analysiert und den Fahrer rechtzeitig informiert, wann ein Ölwechsel nötig ist, ob er auf Elektromotor umstellen soll oder wenn ein Kind in der Einfahrt spielt, in die er gerade rückwärts einparken will.

Schon jetzt hilft uns diese wachsende Datenmenge, indem sie uns zahlreiche Informationen unmittelbar zur Verfügung stellt.

Aber was genau sind eigentlich „Daten“? Wer oder was generiert sie? Und welches Potenzial haben sie, unser Leben zu verbessern? Wie können wir den größtmöglichen Nutzen daraus ziehen? Und wie stellen wir sicher, dass sie entsprechend unserer Wertvorstellungen und unserem Sicherheitsbedürfnis verwendet werden?

Dies sind wichtige Fragen, da Daten, die ehemals nur schwer zu beschaffen waren, heutzutage zu einer zunehmend übermäßig verfügbaren, wertvollen und erneuerbaren Ressource werden und sich zu einer der wichtigsten Quellen für wirtschaftlichen und sozialen Nutzen entwickeln. Früher war der Zugang zu Land, Arbeitskraft und Kapital für den wirtschaftlichen Erfolg oder Misserfolg ausschlaggebend. Heute bezeichnen 90 Prozent der führenden Wirtschaftsvertreter Daten als wichtige Ressource und als ein wesentliches Differenzierungsmerkmal für Unternehmen,

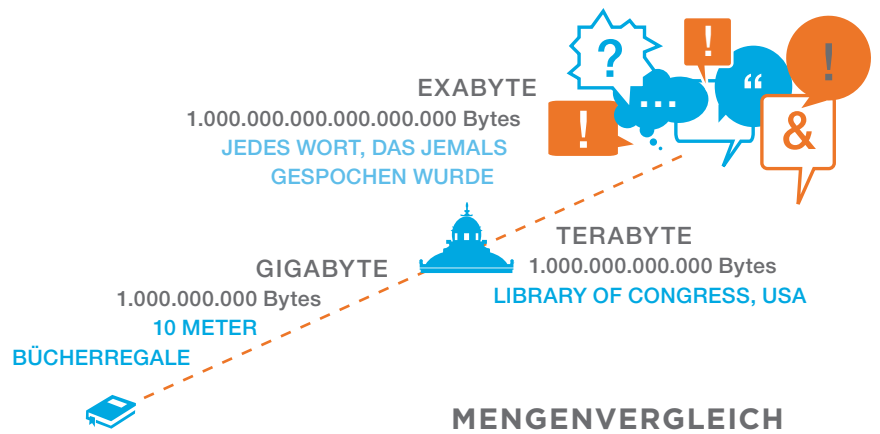
gleichbedeutend mit grundlegenden Ressourcen wie Land, Arbeitskraft und Kapital.¹

Ein Beispiel: Würden durch die effektivere Datennutzung kleine Gewinne realisiert, dank der die Branchen nur um ein Prozent effizienter würden, dann wäre dies gleichbedeutend mit einem Wachstum des globalen BIP um 15 Billionen US-Dollar bis 2030, so die vorsichtige Schätzung der Wissenschaftler.² Der nächste „große Durchbruch“ könnte von den Milliarden kleiner internetfähiger Geräte kommen, die uns noch genauere Daten über unsere Welt liefern, um uns noch effektivere datengetriebene Lösungen zu bieten.³ Schon jetzt finden wir Antworten auf Fragen, die wir noch nicht einmal gestellt haben.

Unsere Gesellschaft befindet sich inmitten eines großen Umbruchs. Fast alles, was wir tun, generiert Daten, und jeden Tag kommen komplett neue Datenströme hinzu. Tatsächlich wurden 90 Prozent der heute weltweit verfügbaren Daten in den letzten zwei Jahren generiert, und derzeit verdoppelt sich das Tempo der Datenproduktion alle zwei Jahre. Bei den meisten Daten handelt es sich jedoch nicht um persönliche Informationen. Das ist ein wichtiger Unterschied, denn obwohl der Schutz unserer Privatsphäre oberste Priorität hat, werden die meisten Daten, die unser Leben vereinfachen, durch einen Sensor an einer Maschine erstellt.

Unsere Herausforderung lautet, diese Daten für uns zu nutzen und mit unserem Einfallsreichtum die darin enthaltenen wertvollen Informationen sinnvoll umzusetzen. Diese Fähigkeit, Daten zu verarbeiten und Beobachtungen in Erkenntnisse umzuwandeln und aus Erkenntnissen wiederum Antworten zu generieren, ermöglicht es uns, die wesentlichen Herausforderungen unserer Zeit wirksam zu lösen.

Unternehmen und Behörden müssen jetzt aktiv daran arbeiten, den Innovationsmotor anzukurbeln



VON DATEN ZU ANTWORTEN IN VIER SCHRITTEN

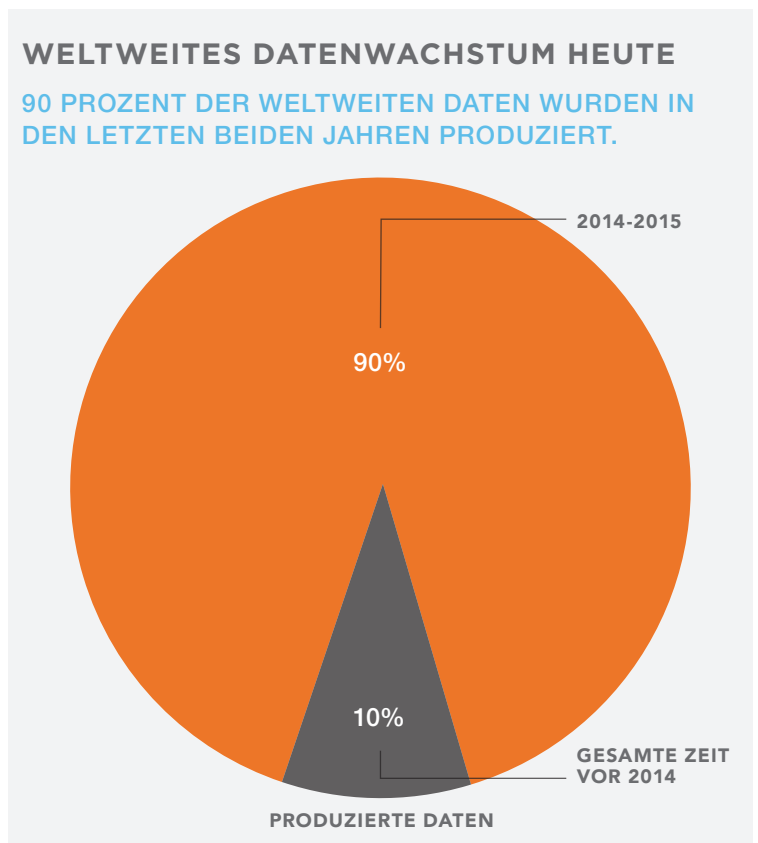
Die moderne Datenrevolution wird nicht nur durch die wachsende Menge an Informationen vorangetrieben. Auch grundlegende Technologien haben die Art, wie wir Daten sammeln, speichern, analysieren und umwandeln, geändert. Gemeinsam ermöglichen uns diese Treiber, neue, bahnbrechende Erkenntnisse aus den in der Datenflut verborgenen Informationen zu gewinnen und so neues Wissen zu erwerben, neue Verbindungen zu entdecken und neue Vorhersagen zu treffen.

1

DATEN SAMMELN

Seit jeher sammelt der Mensch Informationen und nutzt sie, um die Gesellschaft voranzubringen. Aber oft hatte er nicht genügend Daten zur Hand. Heute haben wir das Glück, mehr Daten über unsere Umwelt zu sammeln – aussagekräftigere Daten, die in nützlicherer Form erfasst werden und gezieltere Ergebnisse bringen. Die ehemals rare Informationsquelle ist zu einer zunehmend übermäßig verfügbaren, vitalen und erneuerbaren Ressource geworden.

Möglich wurde dies nicht durch unsere Fähigkeit, persönliche Informationen zusammenzufassen und zu nutzen, sondern durch unsere Fähigkeit, verschiedene Geräte und Sensoren mit dem Internet zu verbinden, die nun zu einer exponentiell wachsenden Flut an neuen Daten geführt hat. Daten werden nun überall generiert – durch Sensoren am Straßenrand, die den Verkehrsfluss messen, durch digitale Musik- und Filmaufnahmen, durch die Satelliten, die unseren Planeten umkreisen,



Source: IBM
<http://www.ibm.com/software/data/bigdata/what-is-big-data.html>

durch Sensoren und Systeme, die unsere Fabriken und Finanzmärkte kontrollieren und durch neue Werkzeuge, mit denen wir den nächsten neuen Trend digital entwerfen. Diese Informationen wachsen immer schneller, verbreiten sich immer weiter und nehmen an Bedeutung zu.

Heute werden schätzungsweise täglich etwa 2,5 Quintillionen Bytes an Daten generiert.⁴ Als Kindern des analogen Zeitalters fällt es uns schwer, die unglaubliche Menge dieses Datenvolumens voll zu erfassen. Zur Veranschaulichung: Würde man die allein im letzten Jahr weltweit gesammelte Menge an digitalen Daten auf DVDs brennen und stapeln, dann wäre der Stapel so hoch wie einmal von der Erde zum Mond und wieder zurück.⁵ Und die Geschwindigkeit, in der wir neue Daten generieren,

Die meisten Daten können nicht Personen zugeordnet werden. Schon bald werden uns vernetzte Geräte weltweit dabei helfen, unsere Umgebung besser zu verstehen und zu gestalten.

nimmt ebenfalls zu. Das Volumen der von allen Unternehmen weltweit erstellten geschäftlichen Daten verdoppelt sich derzeit alle 1,2 Jahre.⁶ Woher kommen diese Datenmengen? Nachstehend ein paar Beispiele:

- + Digitale Informationen in Krankenhäusern, größtenteils aus klinischer Bildgebung, werden Schätzungen zufolge bis 2015 auf 665 Terabytes pro Tag ansteigen – und helfen, Therapien zu entwickeln und Leben zu retten.⁷
- + Moderne Transkontinental-Flugzeuge verfügen über Unmengen an Sensoren auf den Triebwerken, Klappen und am Fahrwerk und produzieren bis zu einem halben Terabyte an Daten pro Flug, um die Flugleistung zu verbessern⁸, Turbulenzen auszuweichen, die Sicherheit zu verbessern und mögliche Triebwerksschäden etwa 2000 mal schneller als früher zu entdecken.⁹ Wenn man das auf mehr als 25.000 Flüge pro Tag hochrechnet, dann bekommt man eine ungefähre Vorstellung der Menge an nützlichen Daten, die heutzutage allein von Passagiermaschinen generiert werden.
- + Wettersatelliten, Wetterstationen, Radaranlagen und andere Sensoren erfassen 15 Mal pro Stunde über 2,25 Milliarden Wetterdatenpunkte. So werden täglich 20 Terabytes an Daten gesammelt, die weltweit genauere Wettervorhersagen ermöglichen.¹⁰
- + Die Börsen produzieren täglich etwa vier bis fünf Terabytes an Daten, die für Echtzeit-Analysen und die Entdeckung problematischer Handelsaktivitäten genutzt werden und gleichzeitig das Wachstum der Unternehmen und eine erfolgreichere Wirtschaft fördern.¹¹
- + Telematische Sensoren in zehntausenden Lieferfahrzeugen überwachen die Motorleistung, optimieren die Streckenführung und antizipieren Probleme. Dank der Daten der Fahrzeugsensoren und der Analyse der Streckenkarten können Unternehmen Millionen Liter Benzin sparen und so die Abgasemissionen in einem Maße reduzieren, das sich sonst nur erreichen ließe, würde man Tausende Fahrzeuge für ein Jahr von den Straßen nehmen.¹²
- + Der weltweit größte Teilchenbeschleuniger, der Large Hadron Collider am Europäischen Kernforschungsinstitut CERN, generiert während eines Experiments 40 Terabytes an Daten pro Sekunde und liefert so neue Kenntnisse, um die Geheimnisse des Universums zu entschlüsseln.¹³ Das Large Synoptic Survey Telescope in Chile erstellt jede Nacht 30 Terabytes an Informationen, wenn es die Tiefen des Universums abbildet.¹⁴
- + Die Sequenzierung eines einzelnen Genoms einer DNA kann bis zu 200 Gigabytes an Daten generieren. Da die DNA-Sequenzierung immer kostengünstiger wird, bauen die Wissenschaftler umfangreiche Datenbanken auf, gefüllt mit Hunderttausenden dieser Sequenzen, um Unterschiede oder Ähnlichkeiten zu identifizieren, die zusammen zu einem medizinischen Durchbruch führen und Leben retten können.¹⁵

Nicht nur die Menge der Daten wächst exponentiell, sondern auch die Art, wie sie erzeugt werden. Mit der ständig zunehmenden Anzahl an Geräten, die das Internet mit unserer Umgebung verbinden und so das „Internet der Dinge“ erschaffen, produzieren Massen von Sensoren täglich komplett neue Datenformate. Der nächste Trend könnte also auf zahlreichen kleinen Dingen basieren, denn Schätzungen zufolge werden bis 2020 bis zu 50 Milliarden Geräte mit leistungsstarken Sensoren mit dem Internet verbunden sein.¹⁶

Diese Geräte erfassen beispielsweise die Bodenfeuchtigkeit, die Motorleistung, die Effizienz der Energiesysteme und die Orte von Asthmaanfällen. Menschen nutzen nur fünf Sinne, um die Welt zu verstehen. Doch schon bald werden vernetzte Geräte weltweit die Eigenschaften unserer Umgebung in riesigen Mengen erfassen, damit wir sie besser begreifen und gestalten können. Dabei werden wir Exabytes an neuen, hilfreichen Daten erstellen.

Da Speicherplatz immer günstiger wird und die Datenmenge beständig wächst, wird die Nutzung von Daten zunehmend ausgebaut.

2

DATEN SPEICHERN

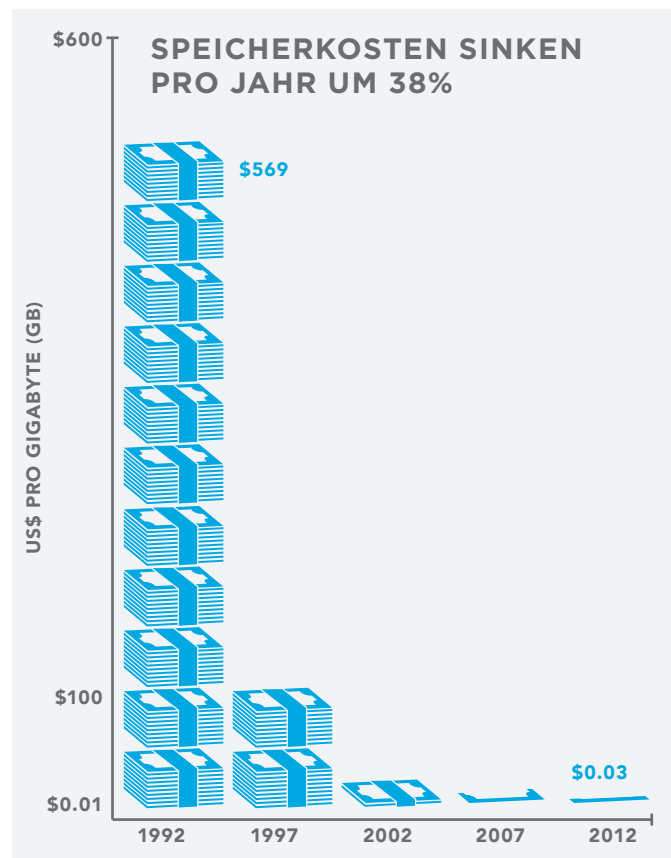
Datengestützte Innovation wird durch den Verfall der Kosten für die Datenspeicherung ermöglicht. 1980 war ein Gigabyte Speicherplatz nur begrenzt erhältlich und wenn, dann kostete er Tausende von Dollar und für deren Verwaltung musste eine Vollzeitkraft eingestellt werden.¹⁷ Heute kostet ein Gigabyte Speicherplatz nur wenige Cent, lässt sich problemlos verwalten und man kann jederzeit und von überall aus darauf zugreifen.¹⁸ Seit den 1980ern sind die Speicherplatzkosten um das 10-Millionenfache gesunken.¹⁹ Zur Veranschaulichung: Wären die Benzinpreise in gleichem Maße gesunken, dann könnte man heute für den Preis einer Tankfüllung Benzin aus dem Jahr 1980 etwa 10.000 km fahren.²⁰

Da die Kosten immer weiter sinken, können immer größere Datenmengen gespeichert werden. 1994 wurden nur drei Prozent der weltweiten Daten digital gespeichert.²¹ 2007 waren es 94 Prozent.²²

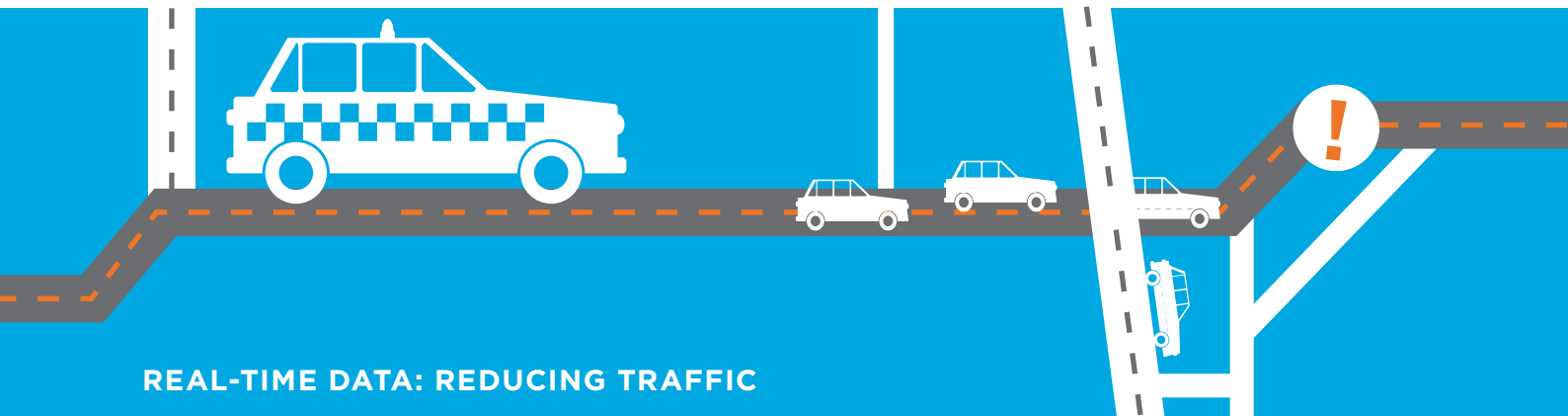
Da Speicherplatz immer günstiger wird und die Datenmenge beständig wächst, wird die Nutzung von Daten zunehmend ausgebaut. Cloud-Technologien – die Möglichkeit, die Daten remote zu speichern und auf diese über verschiedene Geräte zuzugreifen – hat zu einem weiteren deutlichen Preissturz für Speicherplatz geführt, sodass Daten nun nach der ersten Verwendung nicht mehr gelöscht werden müssen, um Platz für neue Daten zu schaffen.²³

Daher sind Daten anders als andere Ressourcen nach einmaligem Gebrauch nicht „aufgebraucht“. Daten sind eine erneuerbare Ressource, die mit anderen Datensätzen kombiniert und mehrfach verwendet werden können, um Antworten auf Fragen zu finden, die bei der ursprünglichen Generierung der Daten noch nicht im Raum standen. So werden Wetterdaten beispielsweise nicht nur zur Wettervorhersage verwendet, sondern zur Erstellung von Ertragsprognosen in der Landwirtschaft.

Cloud-Technologien geben Nutzern besseren, zuverlässigeren, erschwinglicheren und deutlich flexibleren Zugriff auf ihre Daten und entlasten gleichzeitig die Speicher der hausinternen IT-Systeme. Die Cloud-Technologien haben die Art, wie Speicherplatz gekauft, verkauft und bereitgestellt wird, grundlegend verändert. Zudem ermöglichen sie orts- und zeitunabhängigen Datenzugriff. Hierbei handelt es sich um eine der transformativsten Technologien des Jahrzehnts und einen der wichtigsten Wegbereiter für datengestützte Innovationen.



Source: Hagel III, John et al. *From Exponential Technologies to Exponential Innovation*. Deloitte University Press, 2013. Print. 2013 Shift Index Series.



REAL-TIME DATA: REDUCING TRAFFIC

3

DATEN ANALYSIEREN

Daten sind nur dann wertvoll, wenn sie verständlich sind. Sonst sind sie lediglich ein wirrer Haufen zufälliger Beobachtungen. Nur durch die Kombination menschlicher Erfindungsgabe und innovativer Software können diese Informationen sinnvoll gedeutet werden.

Trotz einer zunehmend autonomen Welt sind persönliche Neugier, menschliches Geschick und intensive Arbeit nach wie vor unerlässlich, um die in den Daten verborgenen Antworten zu entschlüsseln.

Im ersten Schritt müssen Rohdaten aufbereitet werden, ehe man damit arbeiten kann. Datenwissenschaftler verbringen geschätzt etwa 50 bis 80 Prozent ihrer Zeit damit, unsortierte Daten so vorzubereiten, dass man darin nach nützlichen Informationen suchen kann.²⁴

Dann ist menschliche Kreativität gefragt, um die richtigen Fragen zu stellen und bei der Auswertung der Daten die richtigen Antworten zu finden, schlechte Daten zu erkennen und die Ergebnisse sinnvoll zu interpretieren. Die Rolle des Datenwissenschaftlers wurde als eine Mischung aus Analyst, Künstler und Geschichtenerzähler definiert.²⁵ Jede Information ist wie ein Pixel auf einem Bildschirm. Für sich allein gibt es nur einen Teil des Gesamtbilds wider. Aber wenn man es mit genügend Pixeln in der richtigen Reihenfolge kombiniert, dann kann ein Datenwissenschaftler ein Bild erstellen, das mehr sagt als tausend Worte, und neue, manchmal auch unerwartete Informationen aus den Daten lesen.

Beim Durchforsten von Daten helfen Analysetools, den Berg an Informationen zu durchkämmen und neue Muster und Trends zu erkennen, unerwartete Erkenntnisse aus scheinbar nicht zueinander in Beziehung stehenden Daten zu gewinnen und automatisch statistisch interessante Verbindungen aufzudecken. Dank zunehmend umfangreicher Datenbanken und stetig verbesserter statistischer Algorithmen können mit Analysesoftware endlose

Datenmengen auf der Suche nach relevanten Informationen durchstöbert werden.

Mit den stärkeren Rechenleistungen der heutigen Computer und einfallsreicher Software haben Datenwissenschaftler modernste Werkzeuge zur Hand, um diese riesigen Massen an Informationen sinnvoll zu ordnen und die darin enthaltenen Erkenntnisse aufzudecken.

Obwohl die Netzwerke heutzutage bereits Beeindruckendes leisten, ist die Zusammenführung von Daten über Netzwerke hinweg an einen zentralen Ort, um sie dort gleichzeitig zu verarbeiten, finanziell und logistisch oft nicht machbar. Aber es gibt eine Lösung: riesige, parallel verteilte Cloud-Computing-Plattformen, die leistungsstärkere Analysemaschinen ermöglichen. Diese Plattformen stellen Nutzern erstklassige Datenanalysetools zur Verfügung, mit denen an verschiedenen Standorten gespeicherte Daten gleichzeitig analysiert werden können.

Welche Vorteile bietet uns diese Datenanalyse? Gezielte Zukunftsvorhersagen gehörten einst ins Reich der Märchen – heute sind sie Teil des Alltags. Dank der modernen Datenanalyse können fortlaufend verlässliche Vorhersagen getroffen werden. Wettervorhersagen sind deutlich zuverlässiger geworden und derzeit für bis zu zehn Tage im Voraus möglich. Fuhrparkmanager wissen, wann ein Motor reparaturbedürftig ist, noch bevor das Auto liegen bleibt. Bei diesen Prognosen hilft die Möglichkeit, aktuelle mit historischen Daten vergleichen zu können.

Wirtschaftswissenschaftler entwickeln immer präzisere Methoden, um die Entwicklung der Märkte, der Arbeitsmärkte und der Inflation zu prognostizieren. Viel zu lang haben die von den offiziellen Stellen veröffentlichten Wirtschaftsdaten Entscheidungsträger gezwungen, in den Rückspiegel zu schauen. Die von den Regierungen publizierten Statistiken, wie z.B. das BIP-Wachstum, haben sich immer an weit zurückliegenden Perioden orientiert, um mit langer Verzögerung über die vergangene Wirtschaftsentwicklung zu berichten und diese Zahlen als besten Maßstab für die Zukunft festzulegen. Heute ziehen Wirtschaftswissenschaftler eine Kombination verschiedener Echtzeitdaten wie neue



IBM und die Stadt Stockholm haben gemeinsam 1600 Taxis mit GPS-Systemen ausgestattet. Die Daten der GPS-Systeme werden mittels einer IBM-Streamingsoftware verarbeitet und geben Auskunft über Verkehrsflüsse, Reisezeiten und optimale Pendelrouten.

Reduzierung
der
Emissionen um
10%

Reduzierung
des
Verkehrsaufkommens
um
20%

Senkung der
durchschnittlichen
Reisezeit um
beinahe
50%

Steigerung
des Anteils
an grünen,
steuerfreien
Fahrzeugen
auf
9%

Source: Bertolucci, Jeff. 'Dublin Points Big Data Tech At Traffic Jams'. InformationWeek 2013. Web.
Nusca, Andrew. 'Stockholm Uses Real-Time GPS Data To Manage Traffic Congestion ZDNet, 2010. Web.

Stellenausschreibungen und Industrienaufträge heran und vergleichen sie mit den historischen Daten, um so ein genaueres Bild der heutigen Dynamik zu erstellen und bessere Strategien für eine gesunde Konjunktur zu formulieren.

Die Analyse von Echtzeitdaten ermöglicht auch eine autonome Entscheidungsfindung, sodass der Mensch oder die von ihm betriebenen Maschinen schneller und präziser Entscheidungen treffen können. Schon heute entwickeln führende amerikanische Automobilhersteller Fahrzeuge, ausgestattet mit Hunderten von Sensoren, Telematic und Echtzeit-Konnektivität, um fortschrittliche Funktionen wie selbständiges Parken anzubieten. Sie arbeiten außerdem an Echtzeitanalysetools, die selbständige Kollisionsvermeidung durch das Fahrzeug und selbstfahrende Autos ermöglichen. Solche Fortschritte könnten eines Tages Leben retten, da die Maschinen schneller und zuverlässiger reagieren können als der Mensch.

Dank der exponentiell wachsenden Menge an Echtzeitdaten über unsere Umgebung, können Experten, die diese Informationen ebenso schnell sinnvoll zuordnen können, wie sie sie erhalten, den Effekt der Datenanalysetools maximieren. Die Stärke der besten derzeit verfügbaren Tools ist ihre Fähigkeit, neue Korrelationen zu erstellen und unerwartete

Dank der stärkeren Rechenleistung der heutigen Computer, kombiniert mit einfallreicher Software, haben Datenwissenschaftler modernste Werkzeuge zur Hand, um diese riesigen Massen an Informationen sinnvoll zu ordnen und die darin enthaltenen Erkenntnisse aufzudecken.

Antworten in der Datenflut zu finden, für die wir noch nicht einmal die genaue Fragestellung kennen. Weltweit entdecken Analysetools wichtige Korrelationen und liefern überraschende Ergebnisse. Ein paar Beispiele:

- + Durch die Nachverfolgung und Korrelation von über 1000 Datenpunkten pro Sekunde schockierten kanadische Forscher die Mediziner, indem sie nachwiesen, dass ungewöhnlich stabile Vitalzeichen bei zu früh geborenen Säuglingen mit schwerem Fieber am nächsten Tag korrelierten. Dank dieser Erkenntnis konnten die Ärzte entsprechende Vorkehrungsmaßnahmen treffen.²⁶
- + Mithilfe von Zeitungsberichten der letzten 20 Jahre lassen sich Ort und Zeit von Cholera-Ausbrüchen in Ländern wie Angola prognostizieren.²⁷
- + Die Polizei von Los Angeles hat einen ursprünglich zur Vorhersage von Erdbeben entwickelten Algorithmus modifiziert und nutzt ihn nun, um mit einer Genauigkeit von etwa 150 Metern zu bestimmen, wo es wahrscheinlich zu Straftaten kommen wird. In den Gebieten, in denen diese Software eingesetzt wird, konnte die Zahl der Einbrüche um 33 Prozent und die Zahl der Gewaltverbrechen um 21 Prozent gesenkt werden.²⁸
- + Mithilfe von Datenanalyse und maritimen Sensoren, die Wellen, Strömungen und andere Daten überwachen, nutzen Forscher die Datenanalyse zur Prognose von Tsunamis und anderen Naturkatastrophen sowie deren Auswirkungen.²⁹
- + Daten von Arztbesuchen und Verschreibungsinformationen haben aufgezeigt, dass bei Patienten mit Autoimmunerkrankungen ein größeres Risiko einer Epilepsieerkrankung besteht.³⁰
- + Daten zur Kreditwürdigkeit wurden genutzt, um zu prognostizieren, welche Patienten „höflich“ daran erinnert werden sollten, ihre verschreibungspflichtigen Medikamente einzunehmen.³¹
- + Mithilfe historischer Flugdaten der letzten zehn Jahre korreliert mit Wettermustern können Fluggäste nun bestimmen, welche Flüge wahrscheinlich am pünktlichsten sind.³²

Wir sind umgeben von Daten – und somit auch von endlosen Möglichkeiten. Wenn die Innovatoren verantwortungsvoll und kreativ agieren, dann kann die Dateninnovation uns nicht nur den Alltag erleichtern, sondern auch zur Lösung schwerwiegender Herausforderungen unserer Welt beitragen.

4

DATEN UMWANDELN UND ÜBERSETZEN

Neue, leistungsstarke Softwaretools ermöglichen es uns, mithilfe von Daten bessere Entscheidungen zu treffen, die statt auf Bauchgefühl oder Intuition auf Fakten basieren.

Insbesondere ein neuer Satz von Werkzeugen unterstützt uns bei der Dateninterpretation, indem er die Informationen so transformiert, dass wir anhand der Daten neue Schlüsse ziehen, fokussieren, visualisieren, reflektieren, verfeinern, modellieren und Vorhersagen ableiten können. Hierzu gehört auch maschinelles Lernen, mit dem wir besser auf Daten reagieren können, Modellierungs- und Simulationstechnologien, die Szenarien testen und Daten in reale Lösungen umwandeln, sowie Werkzeuge, die Töne, Bilder oder Videos erkennen und in neue, aussagekräftigere Formate übertragen können.

Eine derartige Datenumwandlung führt zu besserer Planung, hochwertigeren Designs und klügeren Entscheidungen. Ein Beispiel: Ärzte werden heutzutage mit einer Fülle neuer Forschungsergebnisse bombardiert, sodass es nahezu unmöglich ist, bezüglich der aktuellen Entwicklungen auf dem neuesten Stand zu bleiben, ganz zu schweigen davon, die in Echtzeit erfassten Patientendaten zu interpretieren.³³ Folglich nutzen Krankenhäuser vermehrt Supportsysteme für klinische Entscheidungen. Hierbei handelt es sich im Wesentlichen um Softwaresysteme, die Daten aus unterschiedlichen Quellen analysieren und so eine schnelle und zuverlässigere Diagnose innerhalb einer komplexen Datenumgebung ermöglichen; dies hat sich in über 70 Prozent der Fälle als vorteilhaft erwiesen.³⁴

Andere Softwaretools übertragen Daten in aussagekräftigere Formate. Die Echtzeitverarbeitung von Audio-, Bild- und Videodaten führt zu lebensverändernden Durchbrüchen. Ein Beispiel: Da immer mehr Daten darüber erfasst werden, wie Menschen sprechen, konnte die Spracherfassungstechnologie kontinuierlich

verbessert werden. Dies hat neue Durchbrüche wie Echtzeitübersetzungen in zwei Sprachrichtungen bei kontinentübergreifenden Gesprächen ermöglicht und so neue Chancen für den globalen Handel eröffnet. Eine weitere sensationelle Entwicklung macht den weltweit über 360 Millionen Hörgeschädigten Hoffnung: Chinesische Forscher haben mithilfe von Mustererkennung und Echtzeitdatenverarbeitung eines 3D-Kinetik-Sensors ein System entwickelt, das die Gesten der Zeichensprache erkennt und sie in Echtzeit in gesprochene und geschriebene Sprache umwandeln kann – und umgekehrt.³⁵

Die schnelle Bildverarbeitung hat ebenfalls wesentliche Fortschritte auf den Gebieten der Krebserkennung, des Cognitive Computings, der Neurobiologie und der Robotik mit sich gebracht. Aufgrund der Unvorhersehbarkeit ihres Auftretens und ihrer Form sind Hirntumore beispielsweise mitunter nur schwer auf medizinischen Bildern zu erkennen. Mithilfe des Cloud Computings und fortschrittlicher Algorithmen zur Bildanalyse wetteifern die Wissenschaftler nun darum, die besten Softwarealgorithmen für eine noch genauere und schnellere Identifizierung von Hirntumoren zu finden.³⁶

Der Wechsel von der 2D- zur 3D-Mammografie hat die Erkennungsrate bei Brustkrebs deutlich verbessert. Bei der dreidimensionalen Mammografie werden mithilfe der Software aus verschiedenen Positionen aufgenommene Ultraschallbilder kombiniert, um so ein dreidimensionales Bild zu schaffen, das die Erkennungsrate von Brustkrebs steigert und die nervenzehrenden Fälle von Fehlverdacht reduziert.³⁷

Die Fähigkeit, Daten sowohl zur Visualisierung als auch zur Simulation verwenden zu können, macht es deutlich einfacher, sie zu verstehen und zu nutzen. Heutzutage modellieren und simulieren wir komplexe Systeme und testen Entwürfe schneller und mit größerer Genauigkeit, ohne sie tatsächlich zu bauen. In den 1980er Jahren testete Boeing 77 Prototypen seiner 767 mithilfe physikalischer Windkanäle. 2005 gab es nur elf physische Tests für die 787; die Prototypen wurden stattdessen mithilfe virtueller Windkanäle und Supercomputer getestet – so konnte Boeing nicht nur die Sicherheit für seine Fluggäste erhöhen, sondern auch Zeit und Geld sparen.³⁸



Virtuelle Windkanäle sind nur ein Beispiel für die Werkzeuge, die massive Datenvolumen verarbeiten, um einen einfacheren Einsatz und eine schnellere Umsetzung der dreidimensionalen numerischen Strömungsdynamik zu erreichen. Dank dieser Tools können die Strömung von Wärme, Flüssigkeiten oder Luft besser modelliert und so die Performance

optimiert werden. So kannherausgefunden werden, ob Schadstoffe im Grundwasser mitgeführt werden, wie man die Leistungskraft von Windturbinen steigert und wie man Gebäudekonstruktionen optimiert, damit sie den Naturgewalten standhalten.

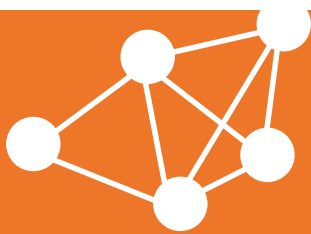
Kurz gesagt: Diese Werkzeuge verwandeln Daten in Lösungen

Vereint eröffnen uns die Fähigkeiten, Daten zu sammeln, zu speichern, zu analysieren und umzuwandeln neue Chancen für bessere Lösungen.

In der Praxis wird jeder dieser Treiber jedoch von verschiedenen Personen ausgeführt, die mit verschiedenen Datensätzen arbeiten, die wiederum an unterschiedlichen Standorten gespeichert sind. Das ist jedoch Teil dieser Datenrevolution. Verschiedene, bisher nicht zueinander in Beziehung stehende Datensätze können nun kombiniert und analysiert werden, auch wenn diese an verschiedenen Orten lagern, verbessert werden, wenn die Daten unstrukturiert sind, und auch dann genutzt werden, wenn die Parteien wichtige Antworten auf Fragen finden, die die Datenproduzenten bisher noch nicht einmal gestellt hatten. Man kann im Voraus nie genau einschätzen, welche bedeutenden Informationen vielleicht zu einem späteren Zeitpunkt einmal aus einem bestimmten Datensatz gewonnen werden können. Das liegt zum Teil daran, dass

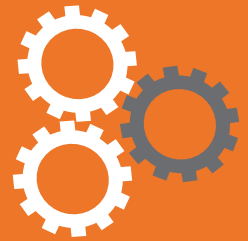
er später in Kombination mit anderen, scheinbar nicht in Zusammenhang stehenden Datensätzen wertvoll werden kann, da der Algorithmus, der die verborgenen Erkenntnisse aufdeckt, bisher noch nicht erfunden wurde.

Wir sind umgeben von Daten – und somit auch von endlosen Möglichkeiten. Wenn die Innovatoren verantwortungsvoll und kreativ agieren, dann kann die Dateninnovation uns nicht nur den Alltag erleichtern, sondern auch zur Lösung schwerwiegender Herausforderungen unserer Welt beitragen. Um die datengestützten Möglichkeiten so zu maximieren, dass wir Produktivität, Wirtschaftswachstum und die Schaffung individueller Vorteile beschleunigen können, müssen wir den Innovationsmotor noch einmal ankurbeln und den Weg für die bahnbrechenden Lösungen und Effekte bereiten, die er uns liefern kann.



Produktionsunternehmen, die sich ihre Daten vollständig zunutze machen, könnten innerhalb von vier Jahren

**371 Milliarden
US-Dollar einsparen.**



UNTERNEHMENS DATEN SENKEN KOSTEN



Eine datengetriebene Wirtschaft

Daten stellen mittlerweile einen der dynamischsten neuen Treiber von wirtschaftlichem Nutzen dar. Die ökonomischen Auswirkungen von Daten machen sich bereits in vielen Sektoren der Wirtschaft bemerkbar, darunter sowohl High-Tech- als auch Low-Tech-Branchen. Eine optimierte Datennutzung könnte allein in den nächsten vier Jahren zu einer weltweiten „Datendividende“ in Höhe von 1,6 Bio. US-Dollar führen.³⁹

Dateninnovation ist offensichtlich in der Lage, die Produktivität zu steigern. Unternehmen, die bereits auf eine datengestützte Entscheidungsfindung zurückgreifen, verbuchen einen Anstieg der Produktivität um fünf bis sechs Prozent.⁴⁰ Würden durch die effektivere Datennutzung kleine Gewinne über ein breites Spektrum an Branchen realisiert, durch die die Branchen nur um ein Prozent effizienter würden, könnte dies laut Schätzungen von Wirtschaftswissenschaftlern einem Wachstum des globalen BIP um 15 Bio. US-Dollar bis 2030 entsprechen. Dies würde dem Hinzufügen einer zweiten US-Wirtschaft gleichkommen. Ein Produktivitätsanstieg von einem Prozent mag gering erscheinen. Jeff Immelt, CEO bei General Electric, stellt dies allerdings wie folgt dar: „Sagen Sie einem Ölonternehmer, dass er Software nutzen kann, die ihm ein Prozent der Kosten spart, und Sie haben einen Freund für's Leben gefunden“.⁴¹

Dateninnovation schafft Arbeitsplätze.

Bei der Dateninnovation geht es nicht nur darum, das Wirtschaftswachstum anzukurbeln. Es geht ebenso darum, neue Arbeitsstellen zu schaffen und die Zahl der Arbeitskräfte zu multiplizieren. Dank der Dateninnovation gibt es bereits Tausende neuer, gut bezahlter Arbeitsplätze: von Datenanalysten über Softwareentwickler zu Personen, die riesige Datenbanken betreiben, die die Dateninnovation erst möglich machen. Des Weiteren geben 61 Prozent der leitenden Angestellten in den USA und 58 Prozent in Europa an, dass die Datenanalyse ein wesentlicher Bestandteil

Jeder
DATENBEZOGENE
Arbeitsplatz in den
USA schafft indirekt drei
weitere Arbeitsplätze.

61 %
der leitenden Angestellten in
den USA geben an, dass die
DATENANALYSE
ein wesentlicher Bestandteil der
Personalplanung
ihres Unternehmens ist.

der Unternehmenspläne zur MitarbeiterEinstellung darstellt.⁴² Für jeden neuen datenbezogenen IT-Job werden schätzungsweise drei weitere Arbeitsstellen außerhalb der IT-Branche geschaffen, wodurch in der gesamten Wirtschaft Millionen weitere Arbeitsplätze verfügbar werden.⁴³ Daraus ergeben sich breitgefächerte, neue wirtschaftliche Vorteile aus einem weiteren technologiebedingten Jobboom.

Daten machen Unternehmen flexibler, reaktionsschneller und wettbewerbsfähiger.

Die enormen IT-bedingten Produktivitätsanstiege, die zu einem Wirtschaftsaufschwung und einer Erhöhung



Unternehmen könnten allein durch intelligente Gebäudetechnik jährlich

25 Milliarden US-Dollar
an Energiekosten sparen.



Eine datenbedingte Produktivitätssteigerung um ein Prozent könnte in der Luftfahrt innerhalb von 1 5 Jahren weltweit

30 Milliarden US-Dollar
an Kraftstoff einsparen.



des Lebensstandards geführt haben, beschränkten sich üblicherweise auf einige wenige Sektoren wie den Technologiesektor selbst. Da jedoch immer mehr Sektoren immer mehr Arten von Daten generieren, sind dank der Daten nun Zuwächse in Sektoren möglich, die traditionell hinterherhinkten – von der Landwirtschaft bis zum Gesundheitswesen, vom Transportwesen bis zur Bildung und von Energie bis zum Finanzwesen. Tatsächlich gaben 79 Prozent der kürzlich befragten leitenden Angestellten in den USA und 80 Prozent in Europa an, dass die Datenanalyse eine wesentliche Rolle in den Plänen der Unternehmen zur besseren Befriedigung der Kundenbedürfnisse spielt.⁴⁴ 70 Prozent in den USA und 72 Prozent in Europa merkten außerdem an, dass die Datenanalyse bei der Entwicklung neuer Produkte und Dienstleistungen zum Tragen kommt.⁴⁵

GESUNDHEITSWESEN

Die Dateninnovation hilft Menschen anhand besserer Erkenntnisse, länger und gesünder zu leben.

Das Gesundheitswesen ist in der heutigen Zeit einer der datenreichsten Bereiche. Es generiert jeden Tag Hunderte Terabytes an Daten pro Krankenhaus.⁴⁶ Maximiert man die Verwendung dieser Daten, könnte dies das Rezept für eine bessere Versorgung und schnellere Heilmethoden sein. Daten bieten Ärzten heutzutage sehr gute Einsichten in den Gesundheitszustand der Patienten und ermöglichen eine fundiertere Entscheidungsfindung. Mithilfe von Analysesoftware können Ärzte die Ergebnisse früherer Patienten nutzen, um Behandlungsmethoden für künftige Patienten zu verbessern. Dennoch werden Gesundheitsdaten oftmals zu selten genutzt. Durch eine wirksamere Datennutzung im Gesundheitswesen zur Steigerung der Effizienz und Qualität könnte der Sektor Schätzungen zufolge jedes Jahr über 300 Milliarden US-Dollar einsparen, was die Ausgaben um stolze acht Prozent senken würde.⁴⁷ Doch die größten Auswirkungen zeigen sich nicht am gesparten Geld, sondern an den geretteten Leben. Forscher entwickelten beispielsweise eine Maschine, die Algorithmen erlernt und dadurch vier Stunden im Voraus und mit einer Genauigkeit von 66 Prozent Herzstillstände voraussagen kann, indem sie Echtzeitdaten mit der Krankengeschichte des Patienten abgleicht.⁴⁸

TRANSPORT

Die Dateninnovation spart Zeit, Geld und Kraftstoff und rettet Leben.

Daten weisen den Weg in eine intelligentere Zukunft im Transportwesen. Ob in der Luft, auf der Schiene oder auf der Straße: Echtzeit-Datenanalysen ermöglichen intelligentere Transportsysteme, durch die Güter und Menschen sicherer und effizienter befördert werden können. Neuwagen sind heute mit derart vielen Sensoren ausgestattet, dass sie pro Stunde bis zu 25 Gigabyte an Daten generieren können. Allein für die Verarbeitung dieser Daten sind mehr als zehn Millionen Zeilen an Code notwendig.⁴⁹ Diese Daten werden genutzt, um neue Systeme für mehr Sicherheit und die Vermeidung von Kollisionen zu betreiben, die ähnlich große Auswirkungen auf die Sicherheit haben könnten wie damals Sicherheitsgurte (Rückgang der Verletzungen und Todesfälle um 50 Prozent).⁵⁰ In einer Zeit, in der die Zahl der Fahrzeuge auf den Straßen stetig steigt, können Daten auf innovative Art genutzt werden, um die Zahl der Staus und das Verkehrsaufkommen zu verringern, wodurch Unmengen an verlorener Zeit, Tausende Liter an Benzin und Tonnen an Treibhausemissionen eingespart werden.

In der Luftfahrt startet die Dateninnovation ebenfalls durch. Heutzutage kann eine moderne Fluggesellschaft aus Sensoren, die im ganzen Flugzeug verteilt sind, bis zu einem halben Terabyte an Daten pro Flug generieren. Diese Daten werden genutzt, um die Flugleistung zu steigern, Turbulenzen zu meiden, die Sicherheit zu erhöhen und Triebwerkschäden 2.000 Mal schneller zu bemerken als zuvor.⁵¹ Luftfahrt-Daten dienen ebenso der effizienteren Flugroutenplanung und informieren die Crew über zu reparierende Teile, bevor diese ausfallen. Diese Vorteile machen sich bezahlt: Bereits eine datenbedingte Produktivitätssteigerung um ein Prozent könnte in der Luftfahrt weltweit 30 Milliarden US-Dollar an Kraftstoff einsparen.

ENERGIE UND UMWELT

Die Dateninnovation verringert den Energieverbrauch und trägt zur Verbesserung der Umwelt bei.

Daten treiben enorme Energieeinsparungen voran. Eine effizientere Nutzung von Daten kann in einer Reihe von Branchen zu Energieeinsparungen in

Im Gesundheitswesen könnten durch die intensivere Nutzung von Daten für mehr Qualität und Effizienz Schätzungen zufolge jedes Jahr mehr als 300 Milliarden US-Dollar eingespart und damit die Ausgaben um beachtliche acht Prozent gesenkt werden.

Milliardenhöhe führen und damit den Umweltschutz stärken. Datengetriebene Erkenntnisse ermöglichen eine intelligentere Energienutzung im Stromnetz, in Gebäuden, bei uns Zuhause, in Fabriken, in der Landwirtschaft und in der Umwelt. Durch die Verwendung von Daten-Design-Tools können beispielsweise energiesparende Gebäude entworfen werden, die mit intelligenten Systemen ausgestattet sind. Sie verwenden unzählige Sensoren, Analysen und Auslöser, um den Betrieb effizienter zu gestalten. Schätzungen zufolge könnten Unternehmen allein durch intelligente Gebäudetechnik 25 Milliarden US-Dollar jährlich an Energiekosten sparen.⁵²

VERARBEITENDES GEWERBE

Die Dateninnovation verbessert die Art und Weise, wie Produkte entwickelt, hergestellt und vertrieben werden.

Der Fertigungssektor speichert mehr Daten als jeder andere Sektor.⁵³ Daher profitieren Produzenten stark von einer besseren Datennutzung. Sie können die Effizienz und Qualität steigern und die Art und Weise verbessern, wie Produkte entwickelt, hergestellt und vertrieben werden. Eine vorteilhaftere Datennutzung im verarbeitenden Gewerbe kann die Produkt-Entwicklungszeit sowie die Herstellungskosten schätzungsweise um bis zu 50 Prozent senken.⁵⁴ So geht die IDC (International Data Corporation) davon aus, dass Produktionsunternehmen, die sich ihre Daten vollständig zunutze machen, innerhalb von vier Jahren eine Datendividende in Höhe von 371 Milliarden US-Dollar generieren könnten.⁵⁵ Durch die Verwendung von Echtzeitdaten können Firmen die globalen Lieferketten besser nachverfolgen und verwalten und zudem Produktionsfehler reduzieren.

Daten führen auch zu einer Verbesserung des Designs. Datengestützte digitale Designs verbessern den sich wiederholenden Design-Prozess und ermöglichen es Designern, unzählige Ideen zu erproben und diese in einer virtuellen Umgebung so lange zu verfeinern, bis schließlich das optimale Produktdesign gefunden wurde. Toyota, Fiat und Nissan haben die Entwicklungszeit von Neufahrzeugen durch die gemeinsame Nutzung von Daten- und Modellierungsverfahren um 30 bis 50 Prozent gesenkt.⁵⁶

FINANZWESEN

Die Dateninnovation steigert die Effizienz, erhöht die Compliance und verringert Betrugsfälle.

Die Verwendung der wachsenden Menge an Daten im Finanzdienstleistungssektor kann die Effizienz der Betriebsabläufe steigern, die Compliance erhöhen und Betrugsfälle aufdecken. 71 Prozent der Banken und Finanzmarktfirmen gaben in einer Studie an, dass die Nutzung von Informationen und Analysen ihren Unternehmen einen Wettbewerbsvorteil beschert.⁵⁷ Dateninvestitionen können außerdem hohe Dividenden einbringen. Die Datenanalyse verhalf einem Kreditkartenunternehmen dazu, einen vermeidbaren jährlichen Kreditkartenbetrug in Höhe von zwei Milliarden US-Dollar aufzudecken.⁵⁸

LANDWIRTSCHAFT

Durch Dateninnovation lassen sich mehr und bessere Nahrungsmittel bei geringerem Ressourcenverbrauch produzieren.

In der Landwirtschaft sprießen Daten überall aus dem Boden und fördern die Herstellung köstlicher und nahrhafter Lebensmittel für eine wachsende Anzahl an Menschen. Dank eines als Präzisionslandwirtschaft bekannten Konzepts ist es Landwirten weltweit möglich, Daten von Saatgut, Satelliten, Sensoren und Traktoren zu nutzen, um fundiertere Entscheidungen zu treffen, die die Erträge steigern, die Kosten senken und mehr Menschen ernähren. Da ein Viertel der Weltbevölkerung in der Landwirtschaft und der Lebensmittelproduktion tätig ist, können durch die intensivere Nutzung von Daten große Gewinne erzielt werden. Beispielsweise sind Landwirte durch eine Daten-Analysesoftware in der Lage, die Inputkosten sowie die Verwendung von Pestiziden und Chemikalien zu senken und gleichzeitig die Erträge um fünf bis zehn Bushel je Morgen Land zu steigern.⁵⁹ Ein kroatisches Start-up-Unternehmen hat eine Cloud-basierte Analysesoftware für Milchbauern entwickelt, die Echtzeitdaten darüber zur Verfügung stellt, wie sich die Futterqualität auf die Produktion und die Empfängnisrate bei jedem einzelnen Tier auswirkt. Das hat die Produktivität und Effizienz der landwirtschaftlichen Betriebe um 50 Prozent gesteigert.⁶⁰

Anhand von mehr als 1.000 Datenpunkten

pro Sekunde haben kanadische Wissenschaftler festgestellt, dass zu früh geborene Kinder mit ungewöhnlich stabilen Vitalzeichen ein erhöhtes Risiko für Fiebererkrankungen haben. Dadurch konnten Ärzte rechtzeitig Vorkehrungen treffen und Leben retten.

KANADA



Mythos oder Fakt?

Um die jüngsten Dateninnovationen und die Datenwirtschaft ranken sich zahlreiche Mythen. Zum Beispiel zu folgenden Themen:

- + Persönliche Daten und Datenschutz
- + Die wirtschaftliche Auswirkung der Datenwirtschaft
- + Die Zuverlässigkeit der Daten
- + Die Dateninnovationen des 21. Jahrhunderts
- + Globale Vorteile der Dateninnovation
- + Die Rolle der Regierungen bei der Datenregulierung

Persönliche Daten und Datenschutz

MYTHOS

Die einzigen, die von der Dateninnovation profitieren, sind die IT-Unternehmen – nicht der einzelne Mensch.

FAKT

Die Dateninnovation ermöglicht dem Verbraucher, informiertere Entscheidungen zu treffen, und dem Händler, seine Waren und Dienstleistungen besser auf die Bedürfnisse seiner Kunden abzustimmen. Durch personalisierte Fortbildung, individuell auf den Patienten abgestimmte Medikamente und personalisierte Unterhaltung können wesentliche Fortschritte erzielt werden. Sie kann uns dabei helfen, von einer durch Massenproduktion angetriebenen Wirtschaft auf eine Konjunktur umzusteigen, die auf Massenpersonalisierung beruht. Natürlich steht eine Gruppe von Unternehmen bei der Entwicklung und Bereitstellung der innovativen, neuen Technologien im Mittelpunkt, auf die diese Datenrevolution angewiesen ist. Wenn wir heute jedoch kluge Entscheidungen treffen, dann kann diese aufstrebende, datenzentrierte Wirtschaft neue Arbeitsplätze und Branchen hervorbringen und die Welt in die Hände der Verbraucher legen.

MYTHOS

Alle Daten sind persönliche Daten.

FAKT

Manche Daten sind personenbezogen (zum Beispiel Daten, die wir auf unseren Mobilgeräten oder durch die Nutzung sozialer Netzwerke generieren). Die meisten Daten sind es jedoch nicht.

Die riesigen Datenmengen, die täglich geschaffen werden, umfassen Informationen wie Wetterdaten

IN INDIEN GEBEN
INTERNET-KIOSKE ÜBER

4 Millionen Bauern

Zugang zu Erntepreisen, Wetter und
anderen Informationen in ihrer

regionalen Sprache.

INDIEN



von Satelliten, die Leistung von Jet-Triebwerken, computergenerierte Börsentransaktionen und Informationen von Sensoren, die nichts mit individuellen Personen zu tun haben. Auch wenn Daten personenbezogen sind, wird oft nicht von anderen Personen darauf zugegriffen. Und sie werden meist auch anonymisiert; das heißt, sie werden im Prinzip gespeichert und ohne die Informationen verwendet, die die Identität der betreffenden Person offenlegen.

MYTHOS

Unternehmen kümmern sich nicht um den Schutz persönlicher Daten.

FAKT

Wenn in den USA persönliche Daten generiert werden, sind sie entsprechend zu schützen. Um das Spektrum der Möglichkeiten zu erweitern, die Daten uns bieten, muss das öffentliche Vertrauen in Daten hoch sein. Unternehmen und Organisationen, die Daten nutzen, müssen diese Daten sorgfältig und sicher verwalten. Diese Praktiken könnten durch freiwillige Richtlinien für eine verantwortliche Nutzung der Daten standardisiert werden. Die Industrie selbst könnte das anregen. Viele auf diesem Gebiet führende Unternehmen legen ihren Kunden heutzutage bereits offen, wie sie bei der Datenerfassung vorgehen und ob sie diese Daten mit anderen teilen. Zahlreiche Firmen in den USA befolgen Best Practices, die sie dazu verpflichten, persönliche Informationen, sofern zweckmäßig, zu anonymisieren.

MYTHOS

Die Dateninnovation ist das Ende der Privatsphäre.

FAKT

Der Erfolg der Datenwirtschaft hängt vom Vertrauen der Verbraucher ab. Der Einzelne muss das Gefühl haben, dass seine persönlichen Daten sicher sind. Führende Softwareentwickler beziehen den Datenschutz bereits in die Konzeption ihrer Systeme ein. Diesen Ansatz nennt man „Privacy by Design“. Außerdem nutzen Entwickler häufig Tools zur Anonymisierung, De-Identifizierung und Verschlüsselung, um die Auswirkungen einer eventuellen Datenpanne so gering wie möglich zu

halten. Auch wenn Daten so erfasst werden, dass der einzelne Nutzer nicht persönlich identifiziert werden kann, können die Daten für gewöhnlich immer noch auf Verhaltensmuster untersucht werden, ohne das Vertrauen des Nutzers zu missbrauchen oder seine Privatsphäre zu verletzen. Des Weiteren können durchsetzbare Datenschutzregelungen den Kontext und die relativen Risiken einer Offenlegung oder eines Missbrauchs von Daten berücksichtigen. Die besonders vertraulichen Daten (wie Finanz- oder Gesundheitsdaten) genießen den höchsten Schutz. Das bedeutet, dass Wetterdaten oder Geschäftsanalysen, die keine persönlichen Informationen enthalten, nicht dasselbe Maß an Schutz benötigen wie patientenspezifische medizinische Daten.

MYTHOS

Man kann Daten nie vollständig de-identifizieren. Die De-Identifikation von Daten ist wirkungslos.

FAKT

Bei der De-Identifikation werden die Informationen so aufbereitet, dass man sie nicht mehr einer bestimmten Person zuordnen kann. Sind die Daten einmal de-identifiziert, können sie ohne Bezug zu ihrem Urheber analysiert werden. Experten haben Techniken entwickelt, mit denen man Daten so de-identifizieren kann, dass sowohl die Privatsphäre maximal geschützt als auch die Qualität der Daten maximiert werden kann.⁶¹ Wird die De-Identifikation sachgemäß durchgeführt, beträgt das Risiko einer Re-Identifikation der einzelnen Person anhand der anonymisierten Daten laut Experten in den meisten Fällen weniger als ein Prozent.⁶²

MYTHOS

Unternehmen, die Daten verwenden, kann man nicht vertrauen.

FAKT

Die Industrie hat die Sorgen der Verbraucher um den Schutz ihrer Daten gehört und darauf reagiert. Führende Unternehmen versuchen sich gegenseitig mit immer besseren Datenschutzregelungen zu übertrumpfen. So haben zum Beispiel die beiden Hersteller der Betriebssysteme, die zusammen auf

IN DEN USA ENTWICKELN FÜHRENDE
AUTOMOBILHERSTELER

neue Fahrzeuge, ausgestattet mit

hundertern von Sensoren

und Analysetools, die so fortschrittliche Funktionen wie selbstständiges Einparken und Kollisionsvermeidung ermöglichen.

USA



96,4 Prozent aller Smartphones weltweit laufen, vor kurzem Verbesserungen ihrer Datenschutzeinstellungen bekannt gegeben. Sie räumen Nutzern zusätzliche Kontrollmöglichkeiten ein und steigen zum Schutz der Privatsphäre auf verschlüsselte Daten um.⁶³ Führende Softwareunternehmen versuchen nun, die Verbraucher für sich zu gewinnen, indem sie ihr Engagement beim Datenschutz offensiv kommunizieren.⁶⁴

Die Unternehmen ändern ihre Dienstleistungen und Richtlinien ständig, um sie noch datenschutzfreundlicher zu gestalten. Zum Beispiel verzichten sie darauf, Kundenkommunikationen zu lesen, um gezielt Werbung zu platzieren.⁶⁵

MYTHOS

Der Einzelne hat keine Kontrolle über seine Daten.

FAKT

Obwohl es manchmal so scheinen mag, als hätten wir keine Kontrolle über unsere Daten, gibt es doch verschiedene Tools, mit deren Hilfe Verbraucher ihre Daten besser kontrollieren können. Die meisten von uns kennen diese Tools nur vielleicht noch nicht. So haben einige führende Softwareunternehmen in ihren Webbrowsern die Nicht-Verfolgen-Funktion (Do Not Track) als Grundeinstellung aktiviert. Das bedeutet, dass Websites, die man besucht, und deren Werbetreibende automatisch eine „Nicht-Verfolgen“-Anforderung erhalten. Auf diese Weise teilt man den Websites mit, dass man nicht möchte, dass das persönliche Surf-Verhalten aufgezeichnet wird. Zusätzlich haben einige Datenhändler Webseiten erstellt, auf denen die Kunden einsehen können, welche Art von Informationen über sie gesammelt wurden. Sie können dort eine weitere Erfassung von Daten untersagen und falsche Angaben korrigieren.⁶⁶ Gemeinsam ermöglichen es diese Tools den Verbrauchern, besser zu kontrollieren, wie Informationen über sie gesammelt und verwendet werden, oder der Nutzung zu bestimmten Marketingzwecken zu widersprechen.

Die wirtschaftlichen Auswirkungen der Datenwirtschaft

MYTHOS

Die Dateninnovation wird keine neuen Arbeitsplätze schaffen, sondern vielleicht sogar welche zerstören.

FAKT

Die Dateninnovation kann ein wesentlicher Treiber für das Wirtschaftswachstum sein. In den jüngsten Umfragen gaben 61 Prozent der befragten leitenden Angestellten in den USA und 58 Prozent in Europa an, dass die Datenanalyse für die Planung ihrer Unternehmen zur Einstellung neuer Mitarbeiter wichtig sei.⁶⁷ Aber bei der Dateninnovation geht es nicht nur um die Schaffung neuer Arbeitsplätze innerhalb der IT-Branche. Jede Aufgabe, die mit Daten zu tun hat, wird Schätzungen zufolge drei weitere Jobs außerhalb der IT-Branche schaffen und so noch mehr Arbeitsplätze in allen Bereichen der Wirtschaft kreieren.

Zwar werden sich einige Jobprofile ändern, da wir durch die Dateninnovation neue Wege finden, traditionelle Aufgaben zu erledigen. Diese Veränderung wird aber ähnlich sein wie die Ausbreitung des Internets, bei der für jeden verlorenen oder überflüssig gewordenen Arbeitsplatz 2,6 neue Stellen geschaffen wurden.⁶⁸ Einer der Bereiche mit dem größten Potenzial für Stellenwachstum ist die Datenanalyse. Oft können die in den Daten verborgenen Erkenntnisse nur durch menschlichen Scharfsinn sinnvoll gedeutet werden. Heute herrscht weltweit ein Mangel an fähigen Analysten und Datenmanagern, die in der Lage sind, die Daten zu deuten. McKinsey zufolge fehlen allein in den USA 140.000 bis 190.000 Fachkräfte mit dem entsprechenden analytischen Talent, das für die Erschließung des in den Daten verborgenen Potenzials erforderlich ist. Es fehlen 1,5 Millionen Manager und Analysten, die über die notwendigen Fähigkeiten verfügen, um die Analysen der Datenwirtschaft zu verstehen und darauf basierende Entscheidungen zu treffen.⁶⁹

IN BRASILIEN, investiert einer der größten

Sojaproduzenten

in Software und Datenanalyse, um die Effizienz der Verfahren zur Schadensbegrenzung zu steigern, die Kosten zu senken und die Anbaumethoden zu optimieren.

BRASILIEN



2014 lag das durchschnittliche Grundgehalt eines angestellten Datenwissenschaftlers bei 120.000 US-Dollar und das eines Managers bei 160.000 US-Dollar.⁷⁰ Es wurde bereits von vielen Seiten darauf hingewiesen, dass Unternehmen und Politiker weltweit Schritte einleiten müssen, um einem drohenden Mangel an Fachkräften vorzubeugen, wenn man das Potenzial der datengestützten Innovation in vollem Maße ausschöpfen will.

MYTHOS

Bei der Datenanalyse geht es darum, das menschliche Urteilsvermögen aus dem Prozess auszuschließen.

FAKT

Zwar können einige Fragen mittels Daten beantwortet werden (zum Beispiel: Nimmt die Anzahl der Einwohner in meiner Stadt zu oder ab?). Viele tiefer gehende Aussagen sind nicht so einfach. Man weiß nicht immer, in welcher Beziehung die verschiedenen Daten zueinander stehen. Und weil man manchmal am Anfang nicht weiß, wie die richtige Frage lautet, ist die Datenanalyse oft ein iterativer Prozess, bei dem immer wieder Fragen gestellt werden, bis man endlich die Antwort findet. Deshalb können wir niemals auf das menschliche Urteilsvermögen und den menschlichen Beitrag verzichten, um Differenzen auszugleichen und mögliche Widersprüche zu klären. Daten allein sind kein Allheilmittel, sie können keine Wunder bewirken. Tatsächlich sind Daten für sich meist nur von geringem Nutzen. Sie sind oft chaotisch, nicht von sich aus organisiert oder gar ordentlich strukturiert. Sie zu verstehen und die in ihnen verborgene Relevanz zu erkennen, ist keine leichte Aufgabe. Ob uns Daten bei der Lösung von Problemen helfen können, hängt von der effektiven Umsetzung einer cleveren Datenstrategie ab, die zu schnelleren und besseren Lösungen führen kann. Es hängt auch davon ab, dass man die richtigen Fragen stellt. Doch wenn wir die Daten auf korrekte Weise nutzen, können wir die Antworten auf einige der dringendsten Probleme unserer Gesellschaft finden, die Flamme der Innovation anfachen und die Schaffung zahlreicher IT-getriebener Arbeitsplätze sowie das Wirtschaftswachstum ankurbeln.

MYTHOS

Dateninnovation ist nur was für Großkonzerne – nicht für kleine Unternehmen.

FAKT

Daten werden zunehmend verfügbarer, die Speicherkosten sinken und die Analysetools werden immer effektiver und günstiger. Das bedeutet, dass sich heutzutage auch das kleinste Unternehmen die Datenanalyse zunutze machen und von Werkzeugen profitieren kann, die sich früher nur Großkonzerne leisten konnten. So ermöglicht zum Beispiel das Trends-Feature in QuickBooks Online von Intuit kleinen Unternehmen, das gesammelte Wissen anderer Intuit-User zu nutzen und ihre Einnahmen und Kosten mit denen anderer zu vergleichen, um mögliche Chancen zu erkennen. So können sie informierter und klügere Entscheidungen für das operative Geschäft fällen. Noch ist die Nutzung von Intelligence- und Analytics-Lösungen unter kleinen bis mittleren Unternehmen nicht stark verbreitet, aber man geht davon aus, dass sich das rasch ändern wird.⁷¹ Eine kürzlich durchgeführte Studie hat festgestellt, dass 60 Prozent der kleinen Unternehmen Datenanalysen für wichtig halten.⁷² Dazu gehören nach Aussage deren leitender Entscheidungsträger 57 Prozent der US-Unternehmen mit maximal 50 Mitarbeitern und 62 Prozent der entsprechenden Unternehmen in Europa. Bei mittelgroßen Unternehmen (51 bis 500 Beschäftigte) geben 87 Prozent der US-Geschäftsführer und 79 Prozent der europäischen Geschäftsführer an, dass Datenanalysen für ihr Unternehmen von Bedeutung sind.

MYTHOS

Daten nutzen nur der IT-Branche, nicht den anderen Wirtschaftssektoren.

FAKT

Daten stehen heutzutage im Mittelpunkt eines großen technologischen Wandels, der nahezu jeden Bereich der Wirtschaft verändern und verbessern wird. Viele Leute sind sogar der Ansicht, dass die Dateninnovation die Produktivität aller Wirtschaftssektoren steigern könnte. Obwohl Unternehmen, die datenbasierte Entscheidungen treffen, eine Produktivitätssteigerung

WISSENSCHAFTLER, die

Malariamuster untersuchen,

haben in Kenia kürzlich Mobilfunkdaten genutzt, um Hotspots zu identifizieren, an denen die Krankheit übertragen wird.

Damit unterstützen sie die Bemühungen der Regierung, die gezielte Ausrottung der Krankheit voranzutreiben.

KENIA



von 5 bis 6 Prozent melden, hätte bereits eine durch Dateninnovation erzielte Effizienzsteigerung um nur 1 Prozent branchenübergreifend enorme Auswirkungen – von Einsparungen bei Energie und Kraftstoff über eine bessere Gesundheit der Mitarbeiter zu geringeren Kosten und eine Steigerung der Leistung und Lebensdauer von Sachanlagen.⁷³

Da wir heutzutage in der Lage sind, aufgrund der gesammelten Echtzeitdaten die Effizienz der Triebwerke zu verbessern sowie Reisende auf optimaleren Routen an ihr Ziel zu bringen, prognostiziert GE, dass man allein in der kommerziellen Luftfahrt durch eine Senkung des Treibstoffverbrauchs um nur 1 Prozent in den kommenden 15 Jahren eine Kosteneinsparung von 30 Milliarden erzielen könnte.⁷⁴

MYTHOS

Korrelation impliziert auch immer Kausalität.

FAKT

Die Suche nach dem „unbekannten Unbekannten“ ist eine der wichtigsten Erkenntnisse, die uns die Daten bringen können. Manchmal scheinen Korrelationen jedoch keinen Sinn zu ergeben, und eine Korrelation ist nicht automatisch gleichbedeutend mit einer kausalen Beziehung. Ein Beispiel: Obwohl festgestellt wurde, dass die Anzahl an Gewaltverbrechen und Morden steigt, wenn der Absatz von Eiscreme steigt, ist es dennoch äußerst unwahrscheinlich, dass jeder, der sich ein Eis kauft, zum Mörder mutiert.⁷⁵ Trotzdem ist die Feststellung eines Kausalzusammenhangs eines der wichtigsten Ergebnisse, das wir durch die Auswertung von Daten erhalten können. Täglich werden immer mehr neue interessante Korrelationen und Zusammenhänge entdeckt, die eine zunehmend größere Bedeutung haben. Die Fähigkeit, zwischen einer einfachen Korrelation und einem tatsächlichen Kausalzusammenhang zu unterscheiden, ist einer der Gründe, warum wir auf qualifizierte Datenwissenschaftler angewiesen sind, die diesen Unterschied erkennen und unwahrscheinliche Ergebnisse eliminieren können.

MYTHOS

Dateninnovation erfordert ein großes Budget.

FAKT

Wer Daten zu seinem Vorteil nutzen möchte, braucht nicht zwingend ein großes Budget. Oft genügen schon ein paar kleinere Investitionen, um beeindruckende Gewinne zu erzielen. Prognosen zufolge werden Organisationen, die einen ganzheitlicheren Ansatz gegenüber ihren Daten verfolgen, einige der größten Erträge erwirtschaften, nämlich eine massive „Daten-Dividende“ in Form von etwa 1,6 Billionen USD an zusätzlichem Umsatz, niedrigeren Kosten und verbesserter Produktivität allein in den kommenden vier Jahren.⁷⁶ Ein guter Ausgangspunkt ist meist die Nutzung kleiner Datensätze oder bereits vorhandener Daten. Manchmal kann die Untersuchung auf Betrugsversuche bei den innerhalb einer Woche generierten Daten zu Finanztransaktionen Trends besser aufzeigen, als die Untersuchung der historischen Daten der letzten fünf Jahre. Eine Analyse ergab, dass Unternehmen, die Datenwirtschaftsanalysen effektiv einsetzten, 26 Prozent rentabler waren als ihre Branchenwettbewerber, durch ihre Mitarbeiter und Sachwerte 9 Prozent mehr Umsatz erwirtschafteten und einen 12 Prozent höheren Marktwert hatten.⁷⁷

Die Zuverlässigkeit von Daten

MYTHOS

Die aus den Daten gewonnenen Erkenntnisse sind von sich aus immer korrekt.

FAKT

Erkenntnisse sind nur dann korrekt, wenn die zugrundeliegenden Daten korrekt sind, intelligente Tools eingesetzt werden und ein kluger Datenwissenschaftler in die Entwicklung des Modells eingebunden wird, um verzerrte Ergebnisse zu vermeiden. So nutzt zum Beispiel die innovative „Street Bump“-App, die in Boston, USA, angeboten wird, Smartphones mit

In den Vereinigten Arabischen Emiraten wird mithilfe neuer Datentools das weltweit

erste energiepositive

Gebäude errichtet, das mehr Energie produziert als es verbraucht.

VEREINIGTE ARABISCHE EMIRATE



eingebauten Beschleunigungsmessern und die GPS-Position zahlreicher Nutzer, um die Position von Schlaglöchern in der Stadt zu bestimmen. Wenn diese Smartphonebesitzer in der Regel jedoch jung und eher wohlhabend sind, kann es sein, dass die von der App gesammelten Daten kein akkurates Bild des tatsächlichen Schlaglochvorkommens in allen demographischen Gebieten der Stadt wiedergeben – und die Reaktion der Stadt auf die Schlaglöcher könnte verzerrt werden. Darum ist es wichtig, Datenwissenschaftler hinzuzuziehen, um sicherzustellen, dass die statistische Relevanz stimmt und Verzerrungen der Ergebnisse vermieden werden.

MYTHOS

Entscheidungen aus dem Bauch heraus sind in der Regel richtig.

FAKT

Vom Tage unserer Geburt an lernen wir, bei der Entscheidungsfindung auf unsere Intuition zu vertrauen. Tatsächlich beschreiben sich 19 Prozent der Manager weltweit als „intuitive Entscheider“, die sich fast ausschließlich auf ihr Bauchgefühl und ihre Intuition verlassen.⁷⁸ Aber heutzutage können Daten uns dabei helfen, Beschlüsse auf einer informierteren Basis zu fällen – und somit exakter, schneller und wirkungsvoller zu entscheiden. Laut einer branchenübergreifenden Umfrage unter IT-Entscheidern gaben 59 der Befragten an, dass die Verbesserung der Entscheidungsfindung das Hauptziel der Investitionen in Datentechnologien sei.⁷⁹ Doch nicht nur am Arbeitsplatz ist diese Art der datengestützten Entscheidung von Nutzen, schließlich treffen wir auch im täglichen Leben ständig instinktive Entscheidungen. Und manchmal können wir dabei von Tools, die uns eine schnellere und bessere Entscheidungsfindung ermöglichen, deutlich profitieren. So sind zum Beispiel 93 Prozent aller Autounfälle auf menschliches Versagen zurückzuführen. Neue Fahrzeugsensoren, die riesige Datenmengen über ihre Umgebung erfassen, kombiniert mit Echtzeitanalyse und

automatisierter Entscheidungsfindung (ohne komplett autonom zu werden) könnten die Anzahl der durch Autounfälle verursachten Verletzungen oder Todesfälle um bis zu 50 Prozent reduzieren.⁸⁰

MYTHOS

Je größer die Datenmenge, desto besser.

FAKT

Große Datensätze sind nicht unbedingt gleich die besseren Datensätze. Die tatsächliche Größe ist eventuell sogar am wenigsten ausschlaggebend. Bei vielen Problemen und Fragestellungen können auch kleine Datenmengen wichtige Erkenntnisse liefern, vorausgesetzt sie werden mit den richtigen Werkzeugen analysiert. Das Herausfiltern der relevanten Informationen hängt meist von der Qualität der zugrundeliegenden Daten und der Qualität der Werkzeuge ab, mit denen sie bearbeitet werden. Dennoch wird bisweilen angenommen, dass ein Mehr an Daten auch ein Mehr an Wahrheit enthält und dass größere Datensätze zwangsläufig objektiver sind. Manchmal reichen jedoch einfache Daten aus, um die wichtigsten Informationen zu liefern: z. B. um wie viel Uhr der Laden im Ort schließt, wie viele Kilometer ein Fahrzeug seit dem letzten Ölwechsel zurückgelegt hat oder wie hoch die Telefonrechnung diesen Monat ausfällt. Je nach Fragestellung können auch kleine Datensätze, die auf Webseiten, in Excel-Tabellen oder CRM-Datenbanken zur Verfügung stehen, ebenso aussagekräftig sein wie große Datenmengen. In erster Linie kommt es darauf an, robuste Daten zu schaffen, sie sicher zu speichern, Zugang zu ihnen zu haben und in der Lage zu sein, sie zu bearbeiten, egal wie umfangreich sie sind. Dann kann man sie zu der Zeit und an dem Ort nutzen, wenn es ein Problem zu klären gilt.

Wetterdaten dienen

nicht nur dazu uns zu sagen, ob wir einen Schirm mitnehmen sollten oder nicht. Sie dienen auch zur Prognose von Ernteerträgen, der Wahrscheinlichkeit des Ausbruchs einer Grippewelle oder zur Einschätzung, wie viele Entsalzungsanlagen in Marokko gebaut werden müssen.

MAROKKO



MYTHOS

Unstrukturierte Daten sind unbrauchbar.

FAKT

Unstrukturierte Daten können durchaus nützlich sein. Viele der aussagekräftigsten Daten lassen sich nicht immer in ordentlich strukturierten Tabellen mit Spalten und Zeilen erfassen. Es können unstrukturierte Textdokumente sein oder Röntgenbilder. Der Schlüssel liegt darin, Daten in ihre nützlichste Form zu bringen. Wenn ein durchschnittliches Fortune-1.000-Unternehmen die Nutzbarkeit seiner Daten um nur zehn Prozent verbessern könnte, könnte es seine Erträge um über zwei Millionen USD steigern.⁸¹

MYTHOS

Daten sollten nur für den Zweck verwendet werden, zu dem sie ursprünglich gesammelt wurden.

FAKT

Die Cloud-Technologie hat die Kosten für die Datenspeicherung drastisch gesenkt, so dass Daten nach der ersten Verwendung nicht mehr gelöscht werden müssen, um Platz für neue Daten zu schaffen. Im Gegensatz zu anderen Ressourcen sind Daten nach dem ersten Gebrauch also nicht „verbraucht“. Daten sind eine erneuerbare Ressource. Man kann sie wiederverwenden, mit anderen Datensätzen kombinieren und sie mehrfach nutzen, um Antworten auf Fragen zu finden, die zum Zeitpunkt der ursprünglichen Erfassung der Daten noch nicht im Raum standen. So dienen Wetterdaten beispielsweise nicht nur dazu uns zu sagen, ob wir einen Schirm mitnehmen sollten oder nicht, sondern sie dienen auch zur Prognose von Ernteerträgen, der Wahrscheinlichkeit des Ausbruchs einer Grippewelle oder zur Einschätzung, wie viele Entsalzungsanlagen in Marokko gebaut werden müssen.

Dateninnovationen des 21. Jahrhunderts

MYTHOS

Die entscheidenden Vorteile der Daten werden erst in ferner Zukunft zum Tragen kommen.

FAKT

Obwohl die Dateninnovation auch in den kommenden Jahrzehnten Vorteile liefern wird, profitieren wir heute schon jeden Tag von den Ergebnissen der Datenrevolution. Täglich werden Terabytes von Daten verwendet, um uns noch genauere Wettervorhersagen zu ermöglichen. Anhand der Daten von Millionen von GPS-Signalen werden die Reisezeiten im Pendelverkehr berechnet, damit wir schneller ans Ziel kommen. Unzählige Gesundheitsdatenpunkte helfen uns, mögliche Krankheitsursachen zu entdecken. Terabytes an Finanzdaten werden eingesetzt, um Kreditkartenbetrug vorzubeugen.

Als die Economist Intelligence Unit die Teilnehmer einer Umfrage bat, zu beschreiben, wie sich Daten in den vergangenen fünf Jahre auf ihre Organisation ausgewirkt haben, gaben fast zehn Prozent an, dass sich ihre Art der Geschäftsführung dadurch komplett geändert hatte.⁸² 46 Prozent der Befragten sagten, dass Daten ein wichtiger Faktor geworden sind, vor allem wenn es um unternehmerische Entscheidungen geht. Neue Forschung deuten darauf hin, dass allein in den kommenden vier Jahren die bessere Nutzung von Daten zu einer weltweiten „Daten-Dividende“ in Höhe von 1,6 Billionen USD führen könnte, denn Daten helfen uns, neue Möglichkeiten und Lösungen zu entdecken.⁸³ In einer andere Studie, die sich auf das laufende Jahr bezog, führten 33 Prozent der leitenden Angestellten in den USA und 24 Prozent in Europa zehn Prozent oder mehr des Wachstums ihrer Unternehmen auf Datenanalyse zurück.⁸⁴ 58 Prozent dieser leitenden Angestellten in den USA und 43 Prozent in Europa rechneten für die kommenden fünf Jahren mit einem ähnlichen Ergebnis.

Nach dem verheerenden Tsunami 2004 in Südasiens wurden indonesische Fischer mit Mobiltelefonen ausgestattet. Ihre Einkommen stiegen um

30 Prozent,

da sie zum ersten Mal in ihrem Leben auf Daten über den aktuellen Marktpreis für Fisch zugreifen konnten.

INDONESIEN



MYTHOS

Dieser ganze Datenhype ist völlig überzogen.

FAKT

Die Verwendung von Daten zur Lösung der Probleme unserer Gesellschaft ist nichts Neues. Seit den Anfängen der modernen Zivilisation nutzen wir Daten, um bessere Entscheidungen zu treffen. Dies hat zur Brachwirtschaft geführt, dank der die Ernteerträge gesteigert und mehr Menschen versorgt werden konnten, zur Entwicklung von Navigationstechniken, die den globalen Handel förderten, und zu medizinischen Erkenntnissen, die Millionen von Menschen vor dem Cholera Tod bewahrten. In der Vergangenheit waren Daten jedoch eine seltene Ressource, deren Speicherung kostspielig und deren Bearbeitung schwierig war. Im Gegensatz dazu sind Daten heute im Überfluss vorhanden, die Speicherkosten sind rapide gesunken und die Werkzeuge zur Datenverarbeitung sind leistungsstärker und effektiver. Heute sehen wir uns mit neuen Herausforderungen konfrontiert, doch die neuen, fähigen Datenanalysetechnologien unterstützen uns bei der Auswertung der wachsenden Datenmengen, um wichtige Erkenntnisse und unerwartete Lösungen für einige der wichtigsten Probleme unserer Zeit zu entdecken.

MYTHOS

Die Ära des von der IT getriebenen Wirtschaftswachstums ist vorbei – Dateninnovation kann die Produktivität nicht ankurbeln.

FAKT

Die Fähigkeit von IT-Innovationen, Konjunkturen anzukurbeln, Arbeitsplätze zu schaffen und den Lebensstandard weltweit zu erhöhen, beruht auf ihrer nachweislichen Fähigkeit zur Produktivitätssteigerung. So wurde dank IT-Innovation die Produktivität in den 1990er Jahren um ein bis zwei Prozent verbessert.⁸⁵ Ein prominenter Wirtschaftswissenschaftler der Northwestern University argumentiert jedoch, dass die größten Errungenschaften der IT-Innovation bereits hinter uns liegen.⁸⁶ Aber die Produktivität ist nach wie vor in vollem Gange. Tatsächlich verbuchen Unternehmen, die bereits auf eine datengestützte Entscheidungsfindung zurückgreifen, einen Anstieg der Produktivität um fünf bis sechs Prozent.⁸⁷ Auch wenn die Nutzung dieser neuen Möglichkeiten nur die Produktivität in den USA um zum Beispiel 1,5 Prozent verbessern würde, so könnte man über einen Zeitraum von 20 Jahren genügend Mittel einsparen, um das durchschnittliche Einkommen der Amerikaner um bis zu 30 Prozent zu erhöhen.⁸⁸ Würden durch die effektivere Datennutzung kleine Gewinne über ein breites Spektrum an Branchen realisiert, durch die die Branchen nur um ein Prozent effizienter würden, könnte dies laut Schätzungen von Wirtschaftswissenschaftlern einem Wachstum des globalen BIP um 15 Bio. US-Dollar bis 2030 entsprechen. Dies würde dem Hinzufügen einer zweiten US-Wirtschaft gleichkommen.⁸⁹

Barcelona nutzt Daten, um die

Stadt intelligenter

zu machen, öffentliche Leistungen zu verbessern und nachhaltigere

Transportlösungen anzubieten.

SPANIEN



Die globalen Vorteile der Dateninnovation

MYTHOS

Heutzutage profitieren nur die USA von Daten.

FAKT

Überall in der Welt nutzt man heutzutage Daten, um wichtige Herausforderungen anzugehen. Ein paar Beispiele:

- + Durch die Verwendung von über 1.000 Datenpunkten pro Sekunde haben kanadische Wissenschaftler festgestellt, dass zu früh geborene Kinder mit ungewöhnlich stabilen Vitalzeichen ein erhöhtes Risiko für Fiebererkrankungen haben. Dadurch konnten Ärzte rechtzeitig Vorkehrungen treffen und so Leben retten.
- + In Brasilien investiert einer der größten Sojaproduzenten in Software und Datenanalyse, um die Effizienz der Verfahren zur Schadensbegrenzung zu steigern, Kosten zu senken und Ernteerträge zu maximieren.
- + Barcelona nutzt Daten, um die Stadt intelligenter zu machen, öffentliche Leistungen zu verbessern und nachhaltigere Transportlösungen anzubieten.⁹⁰
- + In den Vereinigten Arabischen Emiraten wird mithilfe neuer Datentools das weltweit erste energiepositive Gebäude errichtet, das mehr Energie produziert als es verbraucht.⁹¹
- + Landwirte in Indien nutzen Daten von Saatgut, Satelliten, Sensoren und Traktoren, um fundiertere Entscheidungen bezüglich der anzubauenden Pflanzen und der Art des Anbaus zu treffen und zu überlegen, wie man die Frische der Produkte vom Hof bis zum Verbraucher nachverfolgen kann und wie man sich am besten den veränderten klimatischen Bedingungen anpasst.⁹²

MYTHOS

Entwicklungsländer sind noch nicht in der Lage, die Vorteile der Datenanalyse zu nutzen.

FAKT

Die Datenrevolution und die Vorteile, die sie mit sich bringt, sind ein globales Phänomen. Der größte Nutzen und die größten Möglichkeiten dieses Phänomens bieten sich zum Teil in den Entwicklungsländern, die technologisch gesehen oft noch nicht auf dem neuesten Stand sind. Laut IDC werden die Schwellenländer die entwickelten Länder bis 2017 im digitalen Universum überholen und ihren Anteil in der wachsenden digitalen Welt zwischen 2012 und 2020 von 36 Prozent auf 62 Prozent erweitern.⁹³ In einer Umfrage unter Nichtregierungsorganisationen in den Entwicklungsländern waren 90 Prozent der Befragten der Ansicht, dass die Datenanalyse das wichtigste Werkzeug für das Erlangen von besseren Erkenntnissen für die Unterstützung ihrer Zielgruppe sei.⁹⁴

Beispiele für die Nutzung der Dateninnovation in Entwicklungsländern gibt es viele:

- + Wissenschaftler, die Muster von Malariainfektionen untersuchen, haben in Kenia kürzlich Mobilfunkdaten genutzt, um Hotspots zu identifizieren, an denen die Krankheit übertragen wird. Damit unterstützen sie die Bemühungen der Regierung, die gezielte Ausrottung der Krankheit voranzutreiben.⁹⁵
- + Nach dem verheerenden Tsunami 2004 in Südasiens wurden indonesische Fischer mit Mobiltelefonen ausgestattet. Ihre Einkommen stiegen um 30 Prozent, da sie zum ersten Mal in ihrem Leben auf Daten über den Marktpreis für Fisch zugreifen konnten.⁹⁶
- + In Peru sind historische Stätten von Erschließungsprojekten bedroht. Mithilfe von Luftaufnahmen und leistungsstarker Software, die das Bildmaterial zusammenfügt, hat Peru detaillierte dreidimensionale Datenpunkt-Clouds erstellt, um seine bedrohten Kulturschätze zu kartografieren, zu überwachen und zu schützen.⁹⁷

ZUM SCHUTZE SEINES KULTURELLEN ERBES setzt Vietnam 3D-Scanner ein, um in den kommenden fünf Jahren

40.000 historische Artefakte zu digitalisieren.

VIETNAM



- + Zum Schutz seines kulturellen Erbes setzt Vietnam 3D-Scanner ein und plant, in den kommenden fünf Jahren 40.000 historische Artefakte zu digitalisieren.⁹⁸
- + In Indien stellen Internet-Kioske über vier Millionen Bauern Zugang zu Verkaufspreisen für landwirtschaftliche Erzeugnisse, Wettervorhersagen und anderen Informationen in der lokalen Sprache zur Verfügung. Mittels fortschrittlicher Analyse- und Mobiltechnologien können sie die Daten der einzelnen Höfe verfolgen und den Bauern nach erfolgter Analyse Waren, zum Beispiel Düngemittel und Saatgut, je nach Bedarf anbieten.⁹⁹

Die Rolle der Regierungen bei der Datenregulierung

MYTHOS

Daten werden als Ausschlusskriterium genutzt – zum möglichen Nachteil Einkommensschwacher oder Benachteiligter

FAKT

Es ist schwer, Menschen die Sorge zu nehmen, dass einige Menschen oder gesellschaftliche Gruppen durch die Verwendung von Daten ungerechterweise benachteiligt werden. Darum sollten Industrie sowie Regierungen dieses Problem vorrangig angehen. Gleichzeitig sollte man jedoch nicht übersehen, dass die Nutzung von Daten auch Möglichkeiten bietet, unfaire Diskriminierung zu bekämpfen und Gemeinschaften zu stärken. Verantwortungsvoll verwendet sind Daten ein starkes neues Werkzeug, um vorhandene Diskriminierung aufzudecken, die den Zugang zu Arbeitsplätzen, Finanzmitteln, Bildung und Chancen verhindert.¹⁰⁰

MYTHOS

Die Regierungen spielen in dem ganzen Prozess keine Rolle.

FAKT

Das massive Wachstum von Remote-Computing, Speicherung, Analyse und Softwaredienstleistungen hat eine Menge politischer Fragen aufgeworfen. Die meisten Verträge, Gesetze und Vorschriften haben diese Möglichkeiten nicht vorhergesehen, als sie aufgesetzt wurden. Die daraus hervorgehende Ambiguität kann für Regierungen und das Wachstum der Datenwirtschaft problematisch werden. Seit langer Zeit bestehende politische Themen wie das Gleichgewicht zwischen dem Schutz der Privatsphäre und Sicherheitserfordernissen, der freie Handel, grundlegende Forschung und die Entwicklung von Arbeitsmarkt und Fachkräften müssen angesichts der rapiden Zunahmen dieser Technologien neu überdacht werden.

Der Politik bieten sich verschiedene konkrete Möglichkeiten, um das volle Potenzial der Datenrevolution auszuschöpfen und die neue Welle von Produktivität, Wirtschaftswachstum und Verbrauchervorteilen schneller auszulösen. Sie können in die Entwicklung von Datenlösungen investieren, mit denen sich die größten Herausforderungen unserer Gesellschaft angehen lassen, und somit eine Verbesserung in Bereichen wie Bildung, Lebensmittelproduktion, Alltagsleben, Energieeinsparung, Transport und Konjunktur erreichen. Sie können klare Regelungen für die internationale Zusammenarbeit bei der Strafverfolgung erlassen. Sie können diejenigen Marktfreiheiten fördern, die der Technologiebranchen helfen, ihr Potential der Investition, Innovation, des Handels und Wachstums auszuschöpfen. Sie können dabei helfen, den freien Datenfluss über physische Grenzen hinweg zu fördern. Sie können sicherstellen, dass ausreichend talentiertes Personal vorhanden ist, um dem drohenden Fachkräftemangel vorzubeugen. Und sie können den Innovationseifer anfeuern, um die Vorteile der Dateninnovation sowohl für Unternehmen als auch für

IN PERU SIND HISTORISCHE STÄTTEN VON
ERSCHLIESSUNGSPROJEKTEN BEDROHT.

Mithilfe von Luftaufnahmen und leistungsstarker Software,
die das Bildmaterial zusammenfügt,
hat Peru detaillierte dreidimensionale

Datenpunkt-Clouds erstellt,

um seine bedrohten Kulturschätze zu kartografieren, zu
überwachen und zu schützen.

PERU



Verbraucher schneller zu realisieren. Mit pragmatischen Gesetzen, die menschliches Talent und Beharrlichkeit nutzen, Innovation und Investition einsetzen und die Kapazitäten und Fähigkeiten erweitern, können Politiker die Vorteile maximieren, die uns diese datenzentrierte Wirtschaft bietet.

MYTHOS

Die Datenlokalisierung fördert den Schutz der Privatsphäre und sorgt für größere Sicherheit.

FAKT

Manche Leute sind der Meinung, dass die Datenlokalisierung, also die zwingende Speicherung von Daten innerhalb bestimmter Landesgrenzen, den Schutz der Privatsphäre und die Sicherheit fördert. Die modernen technologischen Vorteile werden heutzutage jedoch durch das globale Internet ermöglicht und basieren auf Daten, die zwischen verschiedensten Datenzentren weltweit hin- und hergeschickt werden. Der grenzüberschreitende Internetverkehr hat seit 2005 um über 50 Prozent zugenommen.¹⁰¹ Der ungehinderte grenzüberschreitende Datenaustausch ermöglicht es sogar den kleinsten Firmen und Unternehmern, zum Tante-Emma-Laden des gesamten Planeten zu werden, indem sie Produkte, Dienstleistungen und Ideen über Grenzen hinweg beziehen oder vertreiben. Dennoch erwägen Regierungen in aller Welt oft politische Regelungen, die den freien Datenfluss einschränken oder verlangen, dass Datenserver innerhalb ihrer Grenzen aufgestellt sind, um den Markt bedienen zu dürfen.¹⁰² Diese Einschränkungen untergraben die enormen Größen- und wirtschaftlichen Vorteile, die uns die Dateninnovation bieten kann, und beschränken die Möglichkeit, verschiedene Datensätze in verschiedenen Standorten zu kombinieren um neue, nützliche Erkenntnisse aus der wachsenden Datenmenge zu gewinnen. Sie können sich auch nachteilig auf die Sicherheit auswirken, da sie verhindern, dass an

mehreren Standorten Sicherungskopien wertvoller Daten erstellt werden, die ihrem Verlust im Fall einer Naturkatastrophe oder technischen Versagens vorbeugen. Natürlich müssen nicht alle Länder dieselben Gesetze erlassen, um die in den Daten schlummernden Vorteile optimal zu nutzen. Aber sie sollten zumindest miteinander kompatibel sein. Es ist einer der fundamentalen Grundsätze für diese datengetriebenen Vorteile, dass Daten ungehindert über Grenzen hinweg fließen können.

MYTHOS

Datenschutz kann nur gewährleistet werden, wenn die Regierungen einschreiten und verlangen, dass er gewährleistet wird.

FAKT

Bestehende staatliche Datenschutzgesetze können mit strengen, innovativen Weiterentwicklungen der Datenschutzfunktionen und freiwilligen Branchen-Best-Practices kombiniert werden, um die Sicherheit von Daten und die Vertraulichkeit persönlicher Informationen zu gewährleisten. Im Gegensatz dazu können staatliche Erlasse, die das Datenschutzproblem durch die zwingend vorgeschriebene Speicherung der Daten im Land lösen wollen, Innovationen verhindern und somit den Umfang der sozialen Vorteile, die aus der Dateninnovation hervorgehen können, einschränken.



DIGITALER DISKURS

Die Datensprache verstehen

Abundant data - Datenreichtum

Waren Daten einst nur begrenzt verfügbar, ist die Fülle an Daten heutzutage auf die zunehmende Fähigkeit zurückzuführen, aussagekräftige, digitale Daten auf völlig neue Art und Weise zu sammeln. Hinzu kommen die sinkenden Kosten für die Datenspeicherung und neue Wertschöpfungsmethoden.

Adaptive Intelligenz

Adaptive Intelligenz bezeichnet Computerintelligenz, die sich nicht nur auf die statistische Verarbeitung von Daten beschränkt, sondern diese mit Daten aus speziellen Bereichen vereint. Durch die Kombination von intelligenten Verhaltensmodellen mit Expertenwissen können Systeme mehr aus Beispielen erlernen und auf neue Situationen anwenden.

Algorithmus

Ein Algorithmus ist ein schrittweiser Vorgang oder eine Reihe von Computerbefehlen, die mathematische Erkenntnisse zur Datenanalyse nutzen, um Probleme zu lösen. Algorithmen kommen in beinahe jedem Softwareprogramm zum Einsatz.

Altsystem- Legacy system

Ein Altsystem ist ein Computer, eine Anwendung oder eine Technologie, die veraltet oder überholt ist, aber dennoch verwendet wird, da sie eine erforderliche Funktion angemessen ausführt.

Ambient intelligence - Umgebungsintelligenz

Der Begriff Umgebungsintelligenz steht für die Vorstellung, dass Technologie immer unsichtbarer wird und unbemerkt bleibt, da die allgegenwärtige und kostengünstige Informationstechnologie in wachsendem Ausmaß in unser Umfeld integriert wird. Die Umgebungsintelligenz stünde jederzeit bei Bedarf zur Verfügung, da alles vernetzt, intelligent und ansprechbar wird.

Analytik

Analytik ist die simultane Verwendung von Statistik und Software-basierten Algorithmen um aussagekräftige Einsichten, Muster und Zusammenhänge aus Daten zu erlangen.

Anomalie-Erkennung

Die Anomalie-Erkennung beschreibt die Ermittlung von Dateneinheiten in einem Datensatz, die nicht dem erwarteten Muster entsprechen. Anomalien werden ebenso als Sonderfälle, Ausnahmen oder Schadstellen in Daten bezeichnet und stellen oftmals wesentliche und nützliche Informationen zur Verfügung.

Anonymisierung

Die Anonymisierung von Daten umfasst das vollständige Entfernen personenbezogener Daten, die zur Identifizierung einer Person führen könnten.

Bad data - Ungültige Daten

Ungültige Daten sind Daten, die fehlen oder falsch sind. Dies kann so etwas Einfaches sein wie eine falsche Adresse. Dennoch verursachen ungültige Daten jedes Jahr Kosten in Milliardenhöhe bei den Fortune 1.000 Unternehmen.

Big Data

Big Data ist ein Überbegriff, der sich oftmals auf die Verwendung von Computeranalysen zur Auswertung großer Datenmengen bezieht, um neue Erkenntnisse zu erlangen und die Entscheidungsfindung zu verbessern. Er beschreibt oft Datensätze, die so groß und vielfältig und dabei so schnell sind, dass eine Verarbeitung mit herkömmlicher Software mitunter schwer fällt.

Brontobyte

Ein Brontobyte ist eine inoffizielle Maßeinheit für eine außergewöhnlich große Datenmenge. Ein Brontobyte entspricht für gewöhnlich 1.000 Yottabyte und wird durch eine 1 gefolgt von 27 Nullen oder 1×10^{27} dargestellt.



Business Intelligence (BI)

Der Begriff Business Intelligence bezieht sich auf Technologien und Anwendungen, die Rohdaten in operative Erkenntnisse umwandeln, welche die Unternehmensleistung und Entscheidungsfindung verbessern können.

Cloud

Cloud ist ein weitgreifender Begriff, der jede Anwendung, Dienstleistung oder Daten umfasst, die remote bereitgestellt werden. Allgemein wird dies durch große Gruppen an Remoteservern ermöglicht, die miteinander vernetzt sind, um überall auf Abruf Zugang zu Computer- oder Speicherkapazitäten zu gewähren.

Clusteranalyse

Bei der Clusteranalyse werden sich ähnelnde Datenobjekte identifiziert und gruppiert, um sowohl die Unterschiede als auch die Ähnlichkeiten zwischen den Daten besser zu verstehen.

Cognitive Computing

Cognitive Computing bezeichnet einen Prozess, der große Mengen an Informationen mit Verfahren des maschinellen Lernens, der Mustererkennung und zuweilen der natürlichen Sprachverarbeitung kombiniert, um die Funktionsweise des menschlichen Gehirns nachzuahmen. Diese Systeme sind oftmals in der Lage, Neues zu erlernen und mit Menschen zu interagieren, indem sie Informationsquellen, Kontext und Erkenntnisse miteinander verbinden.

Computer-generierte Daten

Computer-generierte Daten bezieht sich auf Daten, die automatisch von einem Computer ohne menschliches Eingreifen generiert werden, z.B. eine Protokolldatei, Telemetriedaten eines Satelliten oder Sensordaten einer Industriemaschine.

Dark Data

Dark Data besteht aus unstrukturierten und ungenutzten Daten, die gespeichert, allerdings noch nicht analysiert oder verarbeitet wurden und in irgendeiner Weise als vernachlässigt oder ungenutzt betrachtet werden.

Data-Mining

Data-Mining bezeichnet den Prozess, bei dem leistungsfähige Computeralgorithmen angewandt werden, um Muster oder Wissen aus großen Datenmengen zu extrahieren.

Data Science

Data Science ist eine Disziplin, die Statistik, Datenvisualisierung, Computerprogrammierung, Data-Mining, maschinelles Lernen und Datenbanktechnik verwendet, um aussagekräftige Erkenntnisse zur Lösung komplexer Probleme zu erlangen.

Data scientist – Datenforscher/-in

Ein/e Datenforscher/-in kombiniert menschliche Einsicht, mathematisches Fachwissen und technische Werkzeuge um den Sinn von Daten zu erschließen, etwa indem er/sie Computeralgorithmen entwickelt und einsetzt. Weltweit besteht ein Mangel an Datenforscher/-innen.

Daten

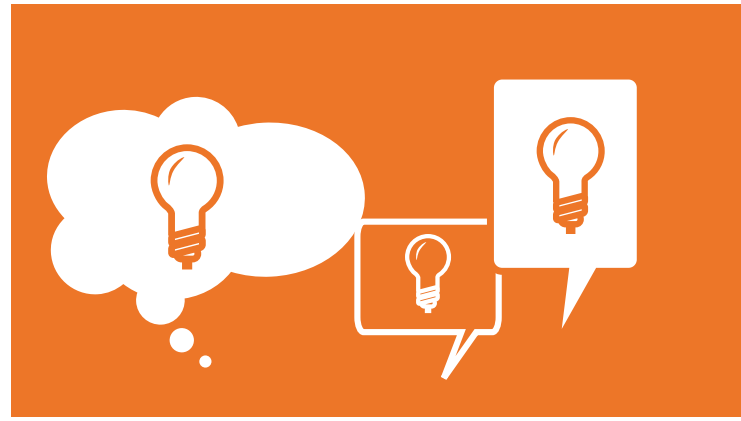
Daten sind Informationen in roher und ungeordneter Form, die digital verarbeitet werden können, um Bedingungen, Objekte oder Ideen darzustellen. Gängige Datentypen umfassen Verkaufszahlen, Marktforschungsergebnisse, Messwerte von Wettersensoren oder eine Liste von Städten und deren Bevölkerung. Heutzutage werden täglich schätzungsweise 2,5 Quintillionen Bytes an Daten generiert.

Datenanalyse

Die Datenanalyse ist die Verwendung von Software zur Umwandlung und Modellierung von Daten, um nützliche Informationen, Erkenntnisse oder Bedeutungen aus Daten abzuleiten. Sie wird oft genutzt, um versteckte Muster oder unbekannte Korrelationen aufzudecken und unterstützt die Entscheidungsfindung.

Datenanalyst

Ein Datenanalyst ist zuständig für die Vorbereitung, Bereinigung und Verarbeitung von Daten.



Datenarchitektur und -design

Datenarchitektur wird für gewöhnlich in der Planungsphase eines neuen Systems angewandt, um auszulegen und zu regeln, wie Daten verarbeitet, gespeichert, genutzt und aufgerufen werden. Indem zu Beginn festgelegt wird, wie bestimmte Daten zueinander in Beziehung gesetzt werden und der Datenfluss in Gang gebracht wird, kann dieser Datenfluss bestimmt und kontrolliert werden, um seinen Schutz im ganzen System zu gewährleisten.

Datenbank

Eine Datenbank ist ein großer strukturierter Satz geordneter digitaler Daten, der so konzipiert ist, dass die darin enthaltenen Daten schnell durchsucht, aufgerufen und aktualisiert werden können.

Datenbereinigung

Die Datenbereinigung beschreibt die Prüfung und Revision von Rohdaten, um Duplikate zu finden und zu löschen, Fehler zu korrigieren, fehlende Daten zu ergänzen, korrupte Daten zu entfernen und für mehr Konsistenz zu sorgen.

Datengestützte Entscheidungsfindung

Unternehmen, die auf eine datengestützte Entscheidungsfindung zurückgreifen, sammeln, verarbeiten und analysieren Daten, um auf deren Grundlage wichtige Entscheidungen zu treffen. Forschungen von Eric Brynjolfsson, Wirtschaftswissenschaftler der Sloan School of Management am Massachusetts Institute of Technology, belegen, dass Unternehmen, die sich einer datengestützten Entscheidungsfindung bedienen, von einem Anstieg der Produktivität um fünf bis sechs Prozent profitieren.

Datensatz

Ein Datensatz ist eine Sammlung zusammenhängender Informationsgruppen (üblicherweise einzelne Elemente) in tabellarischer Form, die als Einheit verarbeitet werden können.

Datensicherheit

Datensicherheit umfasst Verfahren zum Schutz der Daten vor Vernichtung, Missbrauch oder unbefugtem Zugriff. Angemessene Datensicherheitsmaßnahmen können Datenschutzverletzungen vorbeugen, die Datenintegrität sicherstellen und die Privatsphäre schützen. Oft liegt der Fokus dabei auf Menschen, Prozessen und Technologie.

Datenqualität

Datenqualität ist eine Maßeinheit zur Bestimmung des Nutzens von Daten für den Anwender. Sie bezieht sich dabei auf die Zuverlässigkeit, Effizienz und den Wert der Daten für die Entscheidungsfindung, Planung oder den Geschäftsbetrieb.

Datenquelle

Eine Datenquelle ist der primäre Ort, von dem Daten bezogen werden. Dies kann z.B. eine Datenbank, eine Tabelle oder ein Datenstrom sein.

Datenverdichtung

Die Datenverdichtung beschreibt das Verfahren zum Sammeln von Daten aus mehreren Quellen zum Zweck einer hochwertigeren Analyse.

Datenverdichtungstools

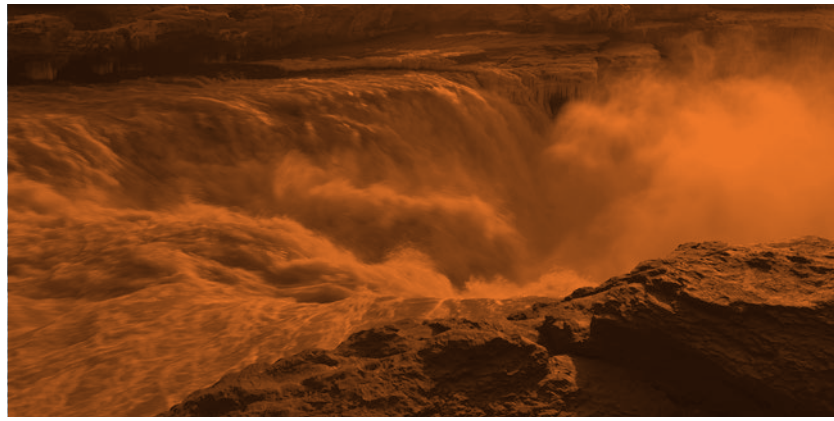
Datenverdichtungstools fassen verstreute Daten aus mehreren Quellen in einem einzelnen Datensatz zusammen.

Datenvirtualisierung

Datenvirtualisierung beschreibt das Abrufen und Verarbeiten verschiedener Datenquellen, ohne die genauen technischen Details zum Ort oder zur Formatierung zu kennen.

Datenvisualisierung

Die Datenvisualisierung umfasst die optische Darstellung von Daten mit dem Ziel, Informationen effizienter abzuleiten oder zu kommunizieren.



De-Identifizierung

Bei der De-Identifizierung von Daten werden Informationen entfernt, die bestimmte Informationsteile einer Person zuordnen.

Disruptive Verschiebungen

Disruptive Verschiebungen sind große, fundamentale Veränderungen in der Gesellschaft und in Unternehmen. Sie werden oftmals durch transformative neue Technologien ermöglicht, die einen völlig neuen Kontext für die Art und Weise schaffen, wie wir arbeiten, leben, spielen und Werte schaffen. Dateninnovation wird oft als Technologie beschrieben, die solche disruptiven Verschiebungen fördert.

Echtzeitdaten

Echtzeitdaten sind Daten, auf die bereits bei der Erstellung reagiert wird. Sie werden oft innerhalb von Millisekunden erstellt, verarbeitet, gespeichert und analysiert. Echtzeitdaten können von Börsenkursen bis zur Geschwindigkeit eines Laufradads (z.B. beim Antiblockiersystem eines Fahrzeugs) alles umfassen.

Empfehlungsdienst

Ein Empfehlungsdienst ist ein Computeralgorithmus, der basierend auf einer Vielzahl von Datenmustern (oftmals durch maschinelles Lernen abgeleitet) Empfehlungen und Vorschläge unterbreiten oder etwas personalisieren kann.

Erkennung von Sonderfällen

Ein Sonderfall ist ein Datenstück, das wesentlich vom allgemeinen Durchschnitt in einem größeren Datensatz abweicht. Es weicht zahlenmäßig vom Rest der Daten ab und weist daher auf eine Entwicklung hin, weshalb für gewöhnlich weitere Analysen erforderlich sind. (Siehe auch Anomalie-Erkennung.)

Exabyte

Ein Exabyte ist eine enorm große Datenspeichereinheit: eine 1 gefolgt von 18 Nullen. Um den Zusammenhang deutlich zu machen: Heutzutage generieren wir täglich ein Exabyte an neuen Informationen.

Fehler-Ursachen-Analyse - Root-cause analysis

Die Fehler-Ursachen-Analyse ist ein Problemlösungsverfahren, das sich auf den Zusammenhang zwischen Ursache und Wirkung konzentriert, um dadurch die Grundursache eines Fehlers oder Problems zu bestimmen. Als Grundursache ist diejenige Ursache zu sehen, deren Vermeidung die Wiederholung des unerwünschten Ereignisses verhindert.

Hadoop

Hadoop ist ein Open-Source-Softwareframework, das zur Verarbeitung und Speicherung großer Datenmengen über verteilte Dateisysteme entwickelt wurde.

Internet der Dinge - Internet of Things

Das Internet der Dinge beschreibt eine Welt, in der alltägliche Geräte intelligenter gestaltet und mit dem Internet verbunden sind, um die „Smart Revolution“ von den Geräten in unseren Händen auf unsere Umgebung auszuweiten. Denn alles, was vernetzt werden kann, wird auch vernetzt. Einige haben dieses Phänomen treffender als das „Internet von allem“ bezeichnet. Einer Schätzung zufolge haben wir bisher erst ein Prozent aller Dinge vernetzt, die vernetzt werden können. Bis 2020 werden schätzungsweise 50 Milliarden Geräte mit dem Internet verbunden sein.

Maschinelles Lernen

Maschinelles Lernen ist die Verwendung von Algorithmen, die es einem Computer erlauben, aus der Analyse von Daten die Handlungen zu „erlernen“, die bei einem bestimmten Muster oder Ereignis vorzunehmen sind.

Metadaten

Metadaten sind Daten über Daten selbst. Diese können grundlegende zusammenfassende Informationen über die Daten enthalten wie Autor, Erstellungsdatum, Dateigröße und Änderungsdatum.



Mustererkennung

Die Mustererkennung beschreibt die Suche und Ermittlung von Mustern innerhalb von Daten. Dies kann zum Beispiel die einfache Suche nach sich wiederholenden Sequenzen in einer DNA sein, oder die Suche nach einem Muster in der Interaktion von zwei Datensätzen, um herauszufinden, ob zwischen zwei Ereignissen ein Zusammenhang besteht. Oder es kann mithilfe von maschinellem Lernen die Suche nach komplexeren Mustern wie etwa das Auffinden numerischer Zeichen in einem Bild sein.

Petabyte

Ein Petabyte ist eine enorm große Maßeinheit für Speicherkapazität und wird durch eine 1 gefolgt von 15 Nullen dargestellt (entspricht einer Million Gigabyte). Ein Petabyte entspricht in etwa der vierfachen Menge der in der Library of Congress enthaltenen Daten.

Rechenzentrum

Ein Rechenzentrum ist eine physische Einrichtung, die eine Vielzahl an vernetzten Servern und Datenspeicher beherbergt und üblicherweise für die Remote-Speicherung und Verarbeitung großer Mengen an remote verfügbaren Daten genutzt wird. Weltweit gibt es schätzungsweise eine halbe Million Rechenzentren, von denen viele die Cloud bilden.

Regressionsanalyse

Die Regressionsanalyse ist ein statistischer Prozess, bei dem Daten genutzt werden, um den Zusammenhang zwischen zwei oder mehreren Variablen abzuschätzen.

Risikoanalyse

Bei der Risikoanalyse wird Daten-Analysesoftware verwendet, um das voraussichtliche Risiko eines Projekts, einer Handlung oder Entscheidung zu ermitteln. Neue Datentools können mögliche Risiken im Voraus identifizieren oder eine Reihe von Szenarien besser abbilden um das Risiko für Organisationen zu mindern, und Systeme überwachen, um festzustellen wenn etwas schief läuft.

Semistrukturierte Daten

Semistrukturierte Daten sind nicht wie ein formelles Datenmodell (z.B. Datenbankmodelle) aufgebaut, sondern stellen die Daten und Hierarchien in anderer Form dar. Semistrukturierte Daten verwenden oft Tags oder andere Datenmarkierungen. Dies wird mitunter als selbsterklärende Struktur bezeichnet.

Small Data

Bei Small Data geht es darum, selbst kleinste Datenmengen, wie die Ergebnisse einer Kundenumfrage, nutzbar zu machen, um umsetzbare Resultate zu erhalten. Der Begriff bezieht sich allgemein auf Datengrößen, die klein genug sind, dass ein Mensch diese verstehen und analysieren kann.

Strukturierte Daten

Strukturierte Daten sind sorgfältig organisiert und zumeist in Zeilen und Spalten gegliedert, was die Suche und Verarbeitung erleichtert.

Terabyte

Ein Terabyte ist eine Maßeinheit für Daten und wird durch eine 1 gefolgt von 12 Nullen dargestellt. Terabyte-Festplatten kommen heutzutage gemeinhin in Heim- und Arbeitsplatzrechnern zum Einsatz oder können über die Cloud genutzt werden. Zur Veranschaulichung: ein Terabyte kann rund 300 Stunden an HD-Videomaterial enthalten.

Textanalyse

Die Textanalyse beschreibt die Verwendung statistischer, sprachwissenschaftlicher und Verfahren des maschinellen Lernens für textbasierte Daten, um Informationen abzuleiten, Konzepte zu extrahieren oder Erkenntnisse zu gewinnen. Die Textanalyse wird für gewöhnlich auf Texte in natürlicher Sprache wie solchen in Dokumenten, Abschriften, Internetposts, Kommentaren oder Formularen angewandt. Sie kann bei der Zusammenfassung, Suche oder Klassifizierung von Inhalt sinnvoll sein.



Transaktionsdaten

Transaktionsdaten sind Daten aus bestimmten Ereignissen wie Finanzkäufe, Rechnungen, Zahlungen und Versanddaten. Sie enthalten für gewöhnlich einen Zeitstempel und unterstützen den täglichen Betrieb eines Unternehmens.

Unstrukturierte Daten

Unstrukturierte Daten haben keinen vordefinierten Aufbau. Als Beispiel zählen hier Notizen eines Meetings. Einigen Schätzungen zufolge könnten unstrukturierte Informationen über 70 bis 80 Prozent aller Daten eines Unternehmens ausmachen.

Variety

Variety, eines der vier Vs, welche die Dateninnovation definieren, stellt die verschiedenen Arten von Daten (oftmals aus unterschiedlichen Quellen) dar, die zusammengefasst und analysiert werden, um Erkenntnisse zu gewinnen. Die Vielfalt der Datentypen, die heutzutage durch Anwendungen verarbeitet werden, umfassen text-basierte Datenbanken, Transaktionsdaten, Streaming-Daten, Bilder, Audio- und Videomaterial.

Velocity

Velocity, eines der vier Vs, welche die Dateninnovation definieren, bezeichnet die Geschwindigkeit mit der Daten generiert, gespeichert, analysiert und visualisiert werden. Große Datenbanken können beispielsweise täglich mit Milliarden Zeilen an neuen Informationen gespeist werden. Zeitkritische Daten müssen bereits beim Streaming genutzt werden, um ihren Wert voll auszuschöpfen.

Veracity

Veracity, eines der vier Vs, welche die Dateninnovation definieren, beschreibt die Genauigkeit, Gewissheit und Exaktheit von Daten.

Volume

Volume, eines der vier Vs, welche die Dateninnovation definieren, bezieht sich auf die verarbeitete Datenmenge und reicht dabei von Megabytes bis Brontobytes.

Vorhersageanalyse - Predictive analytics

Die Vorhersageanalyse wendet unter anderem Softwarealgorithmen auf einen oder mehrere Datensätze an, um Tendenzen oder zukünftige Ereignisse vorherzusagen. Bei diesen Vorhersagen hilft uns die Möglichkeit, aktuelle mit historischen Daten vergleichen zu können.

Vorhersagemodellierung - Predictive modeling

Bei der Vorhersagemodellierung wird ein Modell entwickelt, das am ehesten eine Tendenz, zukünftiges Verhalten oder ein Ergebnis prognostiziert. Oftmals geschieht dies durch den Vergleich von vergangenen mit gegenwärtigen Ereignissen.

Yottabyte

Ein Yottabyte ist eine sehr große Maßeinheit für Datenspeicher und wird durch eine 1 gefolgt von 24 Nullen dargestellt. Zur Veranschaulichung: ein Yottabyte entspricht der Menge an Daten, die auf 250 Billionen DVDs gespeichert ist.

Zettabyte

Ein Zettabyte ist eine Speicher-Maßeinheit, die durch eine 1 gefolgt von 21 Nullen dargestellt wird. Die Größe des World Wide Web wird seit 2013 auf 4 Zettabyte geschätzt. Prognosen zufolge soll bis 2016 täglich mehr als ein Zettabyte an Daten in den weltweiten Netzwerken im Umlauf sein.

Endnoten

- 1 Economist Intelligence Unit: *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <https://www.capgemini.com/resources/the-deciding-factor-big-data-decision-making>
- 2 Ein Produktivitätsanstieg von einem Prozent mag gering erscheinen. Jeff Immelt, CEO bei General Electric, stellt dies allerdings wie folgt dar: „Sagen Sie einem Ölonternehmer, dass er Software nutzen kann, die ihm ein Prozent der Kosten spart, und Sie haben einen Freund fürs Leben gefunden“. Evans, Peter C., Annunziata, Marco: *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- 3 Bis 2020 werden schätzungsweise 50 Milliarden Geräte mit dem Internet verbunden sein und das „Internet von allem“ ermöglichen. Evans, Dave: *The Internet Of Things How The Next Evolution Of The Internet Is Changing Everything*. Cisco Internet Business Solutions Group (IBSG). 2011. Web. White Paper. http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf
- 4 „Bringing big data to the enterprise“. IBM. <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>
- 5 Golub, Ben: „Enough Data to Fill a Stack of DVDs to the Moon (and Back)“. *ComputerWorld*. 2011. Web. http://blogs.computerworld.com/18351/a_stack_of_dvds_to_the_moon_and_back
- 6 Corry, Will: „BIG Data / The Volume Of Business Data Worldwide, Across All Companies, Doubles Every 1.2 Years, According To Estimates“. *The Marketing Blog*, 2012. Web. <http://www.themarketingblog.co.uk/2012/10/big-data-the-volume-of-business-data-worldwide-across-all-companies-doubles-every-1-2-years-according-to-estimates/>
- 7 „Digital Imaging in the Cloud“. *There Magazine*, 2012:16. Web. http://www.agfahealthcare.com/he/global/en/binaries/THERE_12_tcm541-95647.pdf
- 8 Die 787 nutzt Datensensoren, um den Treibstoffverbrauch zu senken, die Systeme zu überwachen und verwendet Beschleunigungsmesser an der Flugzeugnase, um Turbulenzen entgegenzuwirken. Registrieren die Sensoren ein plötzliches Absinken, signalisieren sie den Flügelklappen sofort (innerhalb von Nanosekunden) eine Anpassung, wodurch das Absinken von 9 Fuß bei älteren Flugzeugen auf nur 3 Fuß bei der 787 reduziert werden kann, was einen reibungsloseren Flug garantiert. Humphries, Matthew: „The Boeing 787 produces over 500GB of Data during Every Flight“. *Geek.com*, 2013. Web. Gosling, Kevin: „E-Enabled Capabilities of the 787 Dreamliner“. *Aero Quarterly*. 2009: 22-24. http://www.boeing.com/commercial/aeromagazine/articles/qtr_01_09/pdfs/AERO_Q109_article05.pdf
- 9 Triebwerkshersteller GE gibt an, dass die Triebwerksdaten dem Unternehmen ermöglichen, beispielsweise mögliche Schäden 2.000 Mal schneller zu bemerken als zuvor. Hardy, Quentin: „What Cars Did for Today’s World, Data May Do for Tomorrow’s“. *New York Times*, 2014. B7. Print. <http://bits.blogs.nytimes.com/2014/08/10/g-e-creates-a-data-lake-for-new-industrial-ecosystem/>
- 10 Henschen, Doug: „Bad Winter Weather Meets Big Data Prediction“. *InformationWeek*, 2014. Web. <http://www.informationweek.com/big-data/software-platforms/bad-winter-weather-meets-big-data-prediction/d/d-id/1114192>
- 11 Groenfeldt, Tom: „At NYSE, The Data Deluge Overwhelms Traditional Databases“. *Forbes*, 2013. Web. <http://www.forbes.com/sites/tomgroenfeldt/2013/02/14/at-nyse-the-data-deluge-overwhelms-traditional-databases>
„NYSE Euronext Adapting To Market Changes with Near-Real-Time Insight into Information“. IBM Corporation, 2013. Print. NYSE Euronext. <http://www.ibmbigdatahub.com/sites/default/files/document/NYSE-Euronext-IMC14787USEN.PDF>
- 12 Mayyasi, Alex: „Why UPS Trucks Don’t Turn Left“. *Priceonomics.com*, 2014. Web. <http://priceonomics.com/why-ups-trucks-dont-turn-left/>
- 13 *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper*. Cisco Visual Networking Index, 2015. Print. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html
- 14 „LSST General Public FAQs“. LSST.org 2015. <http://www.lsst.org/lsst/faq>
- 15 Quazi, Ed: „Next Generation Sequencing and the Big Data Bottleneck“. *Arkivum*, 2014. Web. <http://arkivum.com/next-generation-sequencing-and-the-big-data-bottleneck/>
- 16 Valerio, Pablo: „Internet Of Things: 50 Billion Is Only The Beginning“. *EE Times*, 2014. Web. http://www.eetimes.com/document.asp?doc_id=1321229
- 17 1980 galt die Faustregel, dass für jedes Gigabyte an Speicherplatz ein Datenverwalter erforderlich war. Damals kostete ein Gigabyte an Speicherplatz

- rund eine Millionen Dollar. Daher war es sinnvoll, dass jemand diesen Speicherplatz optimierte und die Verwendung überwachte. Gray, Jim, Prashant, Shenoy: „Rules Of Thumb in Data Engineering“. Redmond, WA: Microsoft Research Advanced Technology Division. 2009. Print. Technical Report. http://research.microsoft.com/pubs/68636/ms_tr_99_100_rules_of_thumb_in_data_engineering.pdf
- ¹⁸ Wohlsen, Marcus: „Dropbox Slashes Its Price As The Cost Of A Gigabyte Nears Zero“. *Wired*, 2014. Web. <http://www.wired.com/2014/08/dropboxs-plan-to-stay-relevant/>
- ¹⁹ Von über 200.000 USD (sogar bis zu einer Million Dollar) für ein Gigabyte 1980 bis 0,02 USD pro Gigabyte 2013. Meeker, Mary: „Kleiner Perkins Internet Trends 2014“. 2014. Presentation. <http://cryptome.org/2014/05/internet-trends-2014.pdf>
- ²⁰ Auf Grundlage des durchschnittlichen Kraftstoffverbrauchs von Pkw 1980 (24,3 mpg). Entsprechend bedeuten 10 Millionen Mal mehr Kapazität zum selben Preis 10 Millionen Gallonen mehr Benzin, womit 243 Millionen Meilen gefahren werden können. Legt man den Umfang der Erde mit 24.091 Meilen zugrunde, könnte man die Erde damit 9.758 Mal umfahren. U.S. Department of Transportation.
- Table 4-23: Average Fuel Efficiency Of U.S. Light Duty Vehicles. Washington, DC: Bureau of Transportation Statistics, 2013. Print. http://www.rita.dot.gov/bts/sites/rita.dot.gov/bts/files/publications/national_transportation_statistics/html/table_04_23.html
- ²¹ Savitz, Eric: „Big Data: The Hidden Opportunity“. *Forbes*, 2012. Web. <http://www.forbes.com/sites/ciocentral/2012/05/01/big-data-the-hidden-opportunity/>
- ²² Die technologische Pro-Kopf-Speicherkapazität für Informationen hat sich laut Forschungen von Martin Holbert und Priscila López seit den 1980er Jahren weltweit alle 40 Monate in etwa verdoppelt. Hilbert, Martin, López, Priscila: „The World’s Technological Capacity to Store, Communicate, and Compute Information“. *Science*. 332.6025 (2011): 60-65. Web. <http://www.sciencemag.org/content/332/6025/60>
- ²³ Seit 1992 sind die Kosten für Speicherplatz von 569 USD pro Gigabyte auf nunmehr 0,02 USD pro Gigabyte (2013) gesunken. Dies entspricht einem jährlichen Rückgang von rund 38 Prozent. Meeker, Mary: „Kleiner Perkins Internet Trends 2014“. 2014. Presentation. <http://cryptome.org/2014/05/internet-trends-2014.pdf>
- ²⁴ Lohr, Steve: „For Big-Data Scientists, ‘Janitor Work’ Is Key Hurdle To Insights“. *New York Times*, 2014: B4. Print. http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html?_r=0
- ²⁵ „Data, data everywhere,“ *The Economist*, 25. Februar 2010. <http://www.economist.com/node/15557443>
- ²⁶ Crovitz, L. Gordon: „Why ‘Big Data’ Is A Big Deal“. *Wall Street Journal*, 2013. Print. <http://online.wsj.com/news/articles/SB10001424127887324077704578364632408717740>
- ²⁷ Berichte über Dürren in Angola 2006 lösten eine Warnung über einen möglichen Cholera-Ausbruch im Land aus, da vergangene Ereignisse dem System beigebracht hatten, dass die Wahrscheinlichkeit von Cholera-Ausbrüchen in den Jahren nach Dürren zunimmt. Die Warnungen des Systems trafen mit 70- bis 90-prozentiger Sicherheit zu. Simonite, Tom: „Software Predicts Tomorrow’s News by Analyzing Today’s And Yesterday’s.“ *MIT Technology Review*, 2013. Print. <http://www.technologyreview.com/news/510191/software-predictstomorrows-news-by-analyzing-todays-and-yesterdays/>
- ²⁸ „Ten Big Data Case Studies in a Nutshell“. *TechTarget*, 2013. Print. Essential Guide. <http://searchcio.techtarget.com/opinion/Ten-big-data-case-studies-in-a-nutshell>
- ²⁹ „Big Data to Predict Offshore Accidents, Tsunamis and Other Natural Disasters“. *Predictive Analytics Today*. 2013. Web. <http://www.predictiveanalyticstoday.com/big-data-predict-shore-accidents-tsunamis-natural-disasters/>
- ³⁰ „New Developments in Big Data Visualization“. *USTelecom Media*, 2014. Web. <http://www.ustelecom.org/blog/new-developments-big-data-visualization#sthash.HefD5H52.dpuf>
- ³¹ Quinn, Tom: „New and Unexpected Uses for Scoring Technology“. *Credit Score Blog*, 2011. Web. <http://blog.credit.com/2011/06/new-and-unexpected-uses-for-scoring-technology/>
- ³² Crovitz, L. Gordon: „Why ‘Big Data’ Is A Big Deal“. *Wall Street Journal*, 2013: Print. <http://online.wsj.com/news/articles/SB10001424127887324077704578364632408717740>
- ³³ Datenflut: Ein erfahrener Kliniker benötigt heutzutage knapp 2 Millionen Informationen, um als Arzt zu praktizieren. Mediziner abonnieren durchschnittlich sieben Zeitschriften, was über 2.500 neuen Artikeln jährlich entspricht. Dies macht es nahezu unmöglich, mit den neuesten Erkenntnissen der Diagnostik, Prognostik, Therapie und damit verbundenen Gesundheitsfragen Schritt zu halten. „Clinical Decisions Support Systems: The Time Has Come...“. Frost & Sullivan, 2009. Print. Market Insight. <http://www.frost.com/prod/servlet/cio/181298788>

- ³⁴ „Clinical Decisions Support Systems: The Time Has Come...“. Frost & Sullivan, 2009. Print. Market Insight. <http://www.frost.com/prod/servlet/cio/181298788>
- ³⁵ „Kinect Sign Language Translator Expands Communication Possibilities“. *Microsoft Research*, 2013. Web. <http://research.microsoft.com/en-us/collaboration/stories/kinect-sign-language-translator.aspx>
- ³⁶ „Brats 2012 - Multimodal Brain Tumor Segmentation Challenge“. *CodaLab*. 2012. Print. <https://www.codalab.org/competitions/191>
- ³⁷ Grady, Denise: „3-D Mammography Test Appears To Improve Breast Cancer Detection Rate“. *New York Times*. 2014: p. A1 Print. http://www.nytimes.com/2014/06/25/health/breast-cancer-3d-mammography-test-x-ray.html?emc=edit_th_20140625&nl=todaysheadlines&lid=435891&r=0
- ³⁸ „The game-changing technology that's transforming manufacturing“. *Manufacturing Weekly*, 31. Januar 2014. <http://web.archive.org/web/20140131233544/http://www.manufacturingweekly.com/supercomputers/>
- ³⁹ „The Return on the Data Asset in the Era of Big Data: Capturing the \$1.6 Trillion Data Dividend“. *Cloud Platform News Bytes Blog*, 2015. Web. <http://blogs.technet.com/b/stbnewsbytes/archive/2014/04/15/the-return-on-the-data-asset-in-the-era-of-big-data-capturing-the-1-6-trillion-data-dividend.aspx>
- ⁴⁰ Economist Intelligence Unit: *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁴¹ Ein Produktivitätsanstieg von einem Prozent mag gering erscheinen. Jeff Immelt, CEO bei General Electric, stellt dies allerdings wie folgt dar: „Sagen Sie einem Ölonternehmer, dass er Software nutzen kann, die ihm ein Prozent der Kosten spart, und Sie haben einen Freund fürs Leben gefunden“. Evans, Peter C., Annunziata, Marco: *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- ⁴² BSA/IPSOS Global Data Analytics Poll. November 2014. www.bsa.org/datasurvey
- ⁴³ Gartner: „Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally To Support Big Data By 2015“. 2012. Print. <http://www.gartner.com/newsroom/id/2207915>
- ⁴⁴ BSA/IPSOS Global Data Analytics Poll. November 2014. www.bsa.org/datasurvey
- ⁴⁵ BSA/IPSOS Global Data Analytics Poll. November 2014. www.bsa.org/datasurvey
- ⁴⁶ Laut ESG-Forschungen dürfte die verwaltete Datenmenge je Krankenhaus bis 2015 von 6 Terabyte (2010) auf 168 Terabyte steigen. „Digital Imaging in the Cloud“. *There Magazine*, 2012: 16. Print. http://www.agfahealthcare.com/he/global/en/binaries/THERE_12_tcm541-95647.pdf
- ⁴⁷ Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute. 2011. Print. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
- ⁴⁸ Forscher erarbeiteten einen Algorithmus für maschinelles Lernen aus Daten zu 133.000 Patienten. Das Modell muss noch optimiert werden, um Falschmeldungen zu verringern. Rutkin, Aviva: „Machine Predicts Heart Attacks 4 Hours Before Doctors – New Scientist“. *New Scientist*, 2014. Web. <http://www.newscientist.com/article/mg22329814.400-machine-predicts-heart-attacks-4-hours-before-doctors.html>
- ⁴⁹ Fords modernes Hybridmodell Fusion generiert bis zu 25 GB an Daten pro Stunde. Hemsoth, Nicole: „How Ford Is Putting Hadoop Pedal To The Metal“. *Datanami*, 2013. Web. http://www.datanami.com/2013/03/16/how_ford_is_putting_hadoop_pedal_to_the_metal/
- Der Chevy Volt enthält über 10 Millionen Zeilen an Software-Code, und Softwareentwickler ist eines der am schnellsten wachsenden technischen Berufsfelder in Südost-Michigan, einer Region die für ihre Leistungen bei Fertigung und Produktion bekannt ist. Trop, Jaclyn: „Detroit, Embracing New Auto Technologies, Seeks App Builders“. *New York Times*, 30. Juni 2013. <http://www.nytimes.com/2013/07/01/technology/detroit-embracing-new-auto-technologies-seeks-app-builders.html>
- ⁵⁰ Miller, Claire Cain: „If Robots Drove, How Much Safer Would Roads Be?“ *New York Times*, 2014: A3. Print. <http://www.nytimes.com/2014/06/10/upshot/if-robots-drove-how-much-safer-would-roads-be.html?ref=technology&r=0>
- ⁵¹ Die 787 nutzt Datensensoren, um den Treibstoffverbrauch zu senken, die Systeme zu überwachen und verwendet Beschleunigungsmesser an der Flugzeugnase, um Turbulenzen entgegenzuwirken. Registrieren die Sensoren ein plötzliches Absinken, signalisieren Sie den Flügelklappen sofort (innerhalb von Nanosekunden) eine Anpassung, wodurch das Absinken von 9 Fuß bei älteren Flugzeugen auf nur 3 Fuß bei der 787 reduziert werden kann, was einen reibungsloseren

- Flug garantiert. Gosling, Kevin: „E-Enabled Capabilities of the 787 Dreamliner“. *Aero Quarterly*. 2009: 22-24. http://www.boeing.com/commercial/aeromagazine/articles/qtr_01_09/pdfs/AERO_Q109_article05.pdf
- ⁵² Triebwerkshersteller GE gibt an, dass die Triebwerksdaten dem Unternehmen ermöglichen, beispielsweise mögliche Schäden 2.000 Mal schneller zu bemerken als zuvor. Hardy, Quentin: „What Cars Did for Today’s World, Data May Do for Tomorrow’s“. *New York Times*, 2014: B7. Print. <http://bits.blogs.nytimes.com/2014/08/10/g-e-creates-a-data-lake-for-new-industrial-ecosystem>
- ⁵³ Long, Jessica, Brindley, William: *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁵⁴ Der Fertigungssektor speicherte allein 2010 beinahe 2 Exabyte an neuen Daten. Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute. 2011. Print. http://www.mckinsey.com/~media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁵ Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁶ „Data Smart’ Strategies for Customers Are Yielding ‘Early But Impressive Returns‘“. *Microsoft Research the Fire Hose*. 2014. Web. <http://blogs.microsoft.com/firehose/2014/05/22/data-smart-strategies-for-customers-are-yielding-early-but-impressive-returns/>
- ⁵⁷ Somers, Dan: „Manufacturing 4.0 – From Industrialization to Data-Driven Product Lifecycle“. *Citizen.tekk*. 2013. Web. <http://citizentekk.com/2013/11/05/manufacturing-4-0-industrialisation-data-driven-product-lifecycle/>
- ⁵⁸ McKinsey berichtete, dass Toyota durch die Verwendung dieser datengestützten Designverfahren in der Lage war, 80 Prozent der Fehler vor dem Bau des ersten physischen Prototyps auszumerzen. Manyika, James et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁵⁹ Findings of the New Intelligent Enterprise Study. IBM. 2010 *New Intelligent Enterprise Global Executive Study*. 2010. Print.
- ⁶⁰ Geron, Tomio: „Cows in the Cloud: The Hot Startup Moving Farmers into the Cloud“. *Forbes*, 2012. Web. Helmer, Jodi: „Get Ready For Robot Farmers“. Yahoo, 2014. Web. <https://www.yahoo.com/tech/get-ready-for-robot-farmers-100613764059.html>
- ⁶¹ „Anonymisierung von persönlich zuordenbaren Daten“ (De-Identification of Personally Identifiable Information, DRAFT NISTIR 8053). National Institute of Science and Technology, April 2015 Web. http://csrc.nist.gov/publications/drafts/nistir-8053/nistir_8053_draft.pdf
- ⁶² Cavoukian, Ph.D., Ann, El Emam, Ph.D., Khaled: Dispelling the Myths Surrounding De-Identification: Anonymization Remains a Strong Tool for Protecting Privacy, Information and Privacy Commissioner of Ontario. (Juni 2011). Cavoukian, Ph.D., Ann, Castro, Daniel: *Big Data And Innovation, Setting The Record Straight: De-Identification Does Work*. ITIF, 2014. Print. <http://www2.itif.org/2014-big-data-deidentification.pdf>
- ⁶³ Siehe zum Beispiel der Werbeclip von Microsoft, in dem das Unternehmen seine Datenschutzfeatures bewirbt. https://www.youtube.com/watch?feature=player_embedded&v=bt51MWll1oY
- ⁶⁴ Timothy Cook, CEO von Apple, kündigte zur Unterstreichung des Datenschutzes und der Privatsphäre der Kunden an, dass das Unternehmen modernste Verschlüsselung in das iPhone-Betriebssystem integriert hat, so dass „persönliche Daten wie Fotos, Nachrichten (mit Anhängen), E Mails, Kontakte, Anruflisten, iTunes Inhalte, Notizen und Erinnerungen mit dem kundeneigenen Passwort geschützt“ sind.
- ⁶⁵ Hachman, Mark: „Microsoft’s updated privacy policy makes it clear it’s not selling ads against your words“. *PCWorld*, 11. Juni 2014. <http://www.pcworld.com/article/2362130/microsofts-updated-privacy-policy-makes-it-clear-its-not-selling-ads-against-your-words.html>. Timberg, Craig: „Newest Androids Will Join iPhone In Offering Default Encryption, Blocking Police.“ *Washington Post*, 2014: Print. <http://www.washingtonpost.com/blogs/the-switch/wp/2014/09/18/newest-androids-will-join-iphones-in-offering-default-encryption-blocking-police/>

- ⁶⁶ Data broker Acxiom opens consumer-facing data website, offers opt-out. <http://cir.ca/news/acxiom-gives-consumers-data-peek>
- ⁶⁷ BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁶⁸ McKinsey Global Institute. „Internet Matters: *The Net’s Sweeping Impact On Growth, Jobs, And Prosperity*“. McKinsey & Co, 2011. Print.
- ⁶⁹ Manyika, James et al: „Big Data: *The Next Frontier for Innovation, Competition, and Productivity*“. McKinsey Global Institute, 2011. Print. http://www.mckinsey.com/~media/McKinsey/dotcom/Insights20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx
- ⁷⁰ Laut der im April 2014 von Burtch Works veröffentlichten Studie „Salaries of Data Scientists“.
- ⁷¹ Bagley, Rebecca: „How The Cloud And Big Data Are Changing Small Business“. *Forbes*, 2014. Web. <http://www.forbes.com/sites/rebeccabagley/2014/07/15/how-the-cloud-and-big-data-are-changing-small-business/>
- ⁷² BSA/IPSOS Global Data Analytics Poll, November 2014, www.bsa.org/datasurvey
- ⁷³ Economist Intelligence Unit: *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁷⁴ Economist Intelligence Unit: *The Deciding Factor: Big Data & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁷⁵ Gerbis, Nicholas: „10 Correlations That Are Not Causations“. *HowStuffWorks*, 2015. Web. <http://science.howstuffworks.com/innovation/science-questions/10-correlations-that-are-not-causations.htm>
- ⁷⁶ Vesset, Dan, Morris, Henry D., Gantz, John F.: *Capturing the \$1.6 Trillion Data Dividend*. IDC, 2014. Print. IDC White Paper.
- ⁷⁷ Westerman, George, Bonnet, Didier, McAfee, Andrew: „The Advantages of Digital Maturity“. MIT Sloan, November 2012.
- ⁷⁸ „Meeting the Big Data Challenge: Don’t Be Objective“. *Forbes*, 2013. Web. <http://www.forbes.com/sites/darden/2013/02/01/meeting-the-big-data-challenge-dont-be-objective/>
- ⁷⁹ Big-Data-Studie 2014 von IDC Enterprise. IDG. *CEOs Call for Big Data and IT Continues To Lead Investment Decisions*. 2014. Print. <http://www.idgenterprise.com/press/ceos-call-for-big-data-and-it-continues-to-lead-investment-decisions>
- ⁸⁰ Miller, Claire Cain: „If Robots Drove, How Much Safer Would Roads Be?“. *New York Times*, 2014: A3. Print. http://www.nytimes.com/2014/06/10/upshot/if-robots-drove-how-much-safer-would-roads-be.html?ref=technology&_r=0
- ⁸¹ Clemens, Samuel: „7 Facts about Data Quality [Infographic]“. *InsightSquared*, 3. Januar 2012. Web. <http://www.insightsquared.com/2012/01/7-facts-about-data-quality-infographic/>
- ⁸² Economist Intelligence Unit: *Big Data Harnessing a Game-Changing Asset*. SAS, 2011. Web. http://www.sas.com/resources/asset/SAS_BigData_final.pdf
- ⁸³ „The Return on the Data Asset in the Era of Big Data: Capturing the \$1.6 Trillion Data Dividend“. *Cloud Platform News*. Bytes Blog, 2015. Web. <http://blogs.technet.com/b/stbnewsbytes/archive/2014/04/15/the-return-on-the-data-asset-in-the-era-of-big-data-capturing-the-1-6-trillion-data-dividend.aspx>
- ⁸⁴ BSA/IPSOS Global Data Analytics Poll. November 2014. www.bsa.org/datasurvey
- ⁸⁵ Während der starken Wachstumsphase in den 1990ern steigerten IT-Investitionen in der gesamten US-Wirtschaft, der Einzelhandel eingeschlossen, die jährliche Wachstumsrate der US-Produktivität um ein bis zwei Prozent. Farrell, Diana et al: *How IT Enables Productivity Growth*. San Francisco: McKinsey Global Institute High Tech Practice, 2002. Print. http://www.mckinsey.com/insights/business_technology/how_it_enables_productivity_growth
- ⁸⁶ National Bureau of Economic Research. NBER Working Paper No. 18315: *Is U.S. Economic Growth Over? Faltering Innovation Confronts The Six Headwinds*. 2012. Print. <http://www.nber.org/papers/w18315>
- ⁸⁷ Economist Intelligence Unit: *The Deciding Factor: Big Data 39 & Decision Making*. Cap Gemini, 2012. Web. Point Of View. <http://bigdata.pervasive.com/Solutions/Telecom-Analytics.aspx>
- ⁸⁸ Gertner, Joey: „GE for Making the ‘Internet of Things’ Real“. *Fast Company*, 2014. Web. <http://www.fastcompany.com/most-innovative-companies/2014/ge>
- ⁸⁹ Evans, Peter C., Annunziata, Marco: *Pushing the Boundaries of Minds and Machines*. GE, 2012. Web. <http://files.gereports.com/wp-content/uploads/2012/11/ge-industrial-internet-vision-paper.pdf>
- ⁹⁰ „City Of Barcelona Realizes Vision of Innovative City Governance with Cloud, Devices, and Apps“. Customers.microsoft.com, 2014. Web. <https://customers.microsoft.com/Pages/Home.aspx>

- ⁹¹ „Autodesk the Gallery Masdar Headquarters Positive Energy Building“. *Autodesk.com*, 2015. Web. <http://www.autodesk.com/gallery/exhibits/currently-on-display/adrian-smith-gordon-gill-architecture-masdar-headquarters>
- ⁹² Bunge, Jacob: „Big Data Comes To The Farm, Sowing Mistrust“. *Wall Street Journal*, 2014. Print. <http://online.wsj.com/news/articles/SB10001424052702304450904579369283869192124>
- „Supply Chain Management Solution for Fast Moving Consumer Goods & Food Industries - Farm to Fork Tech Mahindra“. *Techmahindra*, 2015. Web. http://www.techmahindra.com/en-US/www/solutions/Pages/Enterprises/retail_farm_fork.aspx
- ⁹³ Zwischen 2013 und 2020 wird sich die Aufteilung des digitalen Universums zwischen den entwickelten und aufstrebenden Märkten (z. B. China) umkehren – derzeit kommen 60 Prozent der Daten aus den entwickelten Märkten, doch dann werden die aufstrebenden Märkte 60 Prozent der Daten im digitalen Universum generieren. EMC Digital Universe. *Executive Summary Data Growth, Business Opportunities, and the IT Imperatives*. IDC, 2014. Print. <http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>
- ⁹⁴ Long, Jessica, Brindley, William: *The Role of Big Data and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁹⁵ Long, Jessica, Brindley, William: *The Role of BigData and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ⁹⁶ Levy, Stephen: „Bill Gates and President Bill Clinton on the NSA, Safe Sex, and American Exceptionalism“. *Wired*, 2013: Print. <http://www.wired.com/2013/11/bill-gates-bill-clinton-wired/2/>
- Chhachhar, Abdul Razaque, and Siti Zobidah Omar. „Use of Mobile Phone among Fishermen for Marketing and Weather Information“. *Archives Des Sciences* 65.8 (2012): 107-119. Print. http://www.academia.edu/4592505/Use_of_Mobile_Phone_among_Fishermen_for_Marketing_and_weather_information
- ⁹⁷ Neuman, William, Blumenthal, Ralph: „New to the Archaeologist’s Tool Kit: The Drone“. *New York Times*, 2014. Print. http://mobile.nytimes.com/2014/08/14/arts/design/drones-are-used-to-patrol-endangered-archaeological-sites.html?_r=1&referrer
- ⁹⁸ „Forty Thousand Relics to Be Digitized In Five Years“. *Thanhniem News*, 2010. Web. <http://www.thanhniemnews.com/entertainment/forty-thousand-relics-to-be-digitizedin-five-years-22816.html>
- ⁹⁹ Long, Jessica, Brindley, William: *The Role of BigData and Analytics in the Developing World*. Accenture, 2013. Print. Accenture Development Partnerships Insights into the Role of Technology in Addressing Development Challenges. https://www.accenture.com/us-en/~media/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Strategy_5/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf
- ¹⁰⁰ Future of Privacy Forum. *Big Data: A Tool for Fighting Discrimination and Empowering Groups*. Future of Privacy Forum and Anti-Defamation League, 2014. Print. <http://www.futureofprivacy.org/wp-content/uploads/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-Report1.pdf>
- ¹⁰¹ Wladawsky-Berger, Irving: „The Changing Nature of Globalization in Our Hyperconnected, Knowledge-Intensive Economy“. *Wall Street Journal*, 2014. Print. http://blogs.wsj.com/cio/2014/06/20/the-changing-nature-of-globalization-in-our-hyperconnected-knowledge-intensive-economy/?mod=wsj_ciohome_cioreport
- ¹⁰² Folgende Länder haben Regelungen vorgeschlagen oder erlassen, die Unternehmen die Übertragung persönlicher Daten in Gebiete außerhalb ihrer jeweiligen Rechtsprechung verbieten oder deutlich erschweren: Argentinien, Australien, Brasilien, Chile, China, Costa Rica, Griechenland, Hongkong, Indien, Indonesien, Kanada, Kolumbien, Korea, Mexiko, Peru, Russland, die Schweiz und Vietnam.

ÜBER BSA | THE SOFTWARE ALLIANCE

BSA | The Software Alliance (www.bsa.org) ist die globale Stimme der Software-Industrie. In der BSA sind weltweit führende Unternehmen versammelt, die neue Softwarelösungen erschaffen, welche die Wirtschaft antreiben und das moderne Leben von heute prägen.

In der Firmenzentrale in Washington, D.C., und in den Niederlassungen in über 60 Ländern weltweit leistet die BSA Pionierarbeit im Rahmen von Programmen zur Einhaltung und Durchsetzung von geistigen Eigentumsrechten und zur Einführung von Richtlinien, auf deren Basis technologische Innovationen und die digitale Wirtschaft gefördert werden.



www.bsa.org

BSA Worldwide Headquarters

20 F Street, NW
Suite 800
Washington, DC 20001

T: +1.202.872.5500
F: +1.202.872.5501

BSA Asia-Pacific

300 Beach Road
#25-08 The Concourse
Singapore 199555

T: +65.6292.2072
F: +65.6292.6369

BSA Europe, Middle East & Africa

2 Queen Anne's Gate Buildings
Dartmouth Street
London, SW1H 9BP
United Kingdom

T: +44.207.340.6080
F: +44.207.340.6090