

Trendkategorie: Digitalisierung

Big Data

Unter *Big Data* versteht man große und komplexe Datensätze, vor allem von neuen Datenquellen, die in großer Vielfalt, in großer Menge und mit hoher Geschwindigkeit anfallen¹. Diese Datenmengen sind so groß, dass sie mit herkömmlicher Hard- und Software nicht mehr verarbeitet werden können. *Big Data* löst diese Probleme durch spezielle Hard- und Software, die verteilt, d. h. in einem Verbund von vielen Rechnern (Clustern), arbeitet². Eine neuere Definition charakterisiert *Big Data* durch die sogenannten 6V: „Volume“ (Datenumfang), „Velocity“ (Geschwindigkeit, mit der Datenmengen generiert und transferiert werden), „Variety“ (Bandbreite der Datentypen und -quellen), „Veracity“ (Datenechtheit), „Value“ (unternehmerischer Wert der Daten) und „Validity“ (Datenkonsistenz und -qualität)³.



Um neue Erkenntnisse ableiten zu können, müssen die Daten durch leistungsfähige Analysemethoden aufbereitet werden. Big-Data-Analysen erfolgen mithilfe von Tools und Technologien wie statistischen Analysen, Data Mining, Künstlicher Intelligenz (KI), Predictive Analytics oder Machine Learning – als Unterkategorie von KI⁴. Machine-Learning-Tools verwenden datengesteuerte Algorithmen und statistische Modelle, um Datensätze zu analysieren und dann aus den identifizierten Mustern Rückschlüsse zu ziehen oder Vorhersagen zu treffen. Dabei lernen die Algorithmen aus den Daten, während sie diese durchlaufen⁵. Abhängig von der Art der Daten und dem gewünschten Ergebnis kann eines von vier Lernmodellen genutzt werden: überwacht, unüberwacht, teilüberwacht oder bestärkend⁶.

Big Data und Cybersicherheit sind eng miteinander verknüpft. Unternehmen, die große Datenmengen sammeln, sind einem hohen Risiko von Cyberangriffen ausgesetzt. Umgekehrt ist die Kombination von Big Data und maschinellem Lernen (KI) ein wichtiges Instrument im Kampf gegen Cyberkriminalität, indem man ungewöhnliche Aktivitäten schnell erkennen und entsprechende Gegenmaßnahmen einleiten kann⁷.



Was beschleunigt, was bremst den Trend?

Die Menge an Daten wächst rasant: Von 2012 bis 2022 hat sich das weltweit generierte Datenvolumen verzehnfacht⁹. 2022 betrug die jährlich generierte bzw. replizierte digitale Datenmenge weltweit etwa 104 Zettabyte, für das Jahr 2027 sind 284 Zettabyte prognostiziert¹⁰.

Die Digitalisierung bedingt den enormen Zuwachs. Zu den Treibern gehören das Internet und soziale Medien, aber auch die Integration digitaler Mess-, Steuer- und Regelsysteme in eine wachsende Zahl von Alltagsgegenständen¹¹. So fungieren Smart Homes, vernetzte Fahrzeuge, Wearables, Smartwatches, Smartphones, Kundenkarten und viele weitere vernetzte Geräte und Plattformen als Datenquelle¹². Auch neue Technologien wie Edge Computing (schnelle und effiziente Datenverarbeitung direkt oder nahe bei der Datenquelle), moderne Telekommunikationsinfrastrukturen (5G/6G), das Internet of Things (IoT), Augmented und Virtual Reality oder Kryptowährungen spielen bei der Generierung von Daten eine wichtige Rolle¹³.



Big Data hat Einfluss auf fast alle Aspekte des alltäglichen Lebens, Wirtschaft, Wissenschaft und Gesellschaft und gilt als einer der großen Innovationstreiber: In den kommenden Jahrzehnten werden sich auch in der Arbeitswelt durch Big Data zahlreiche Berufe stark verändern und neue Berufsbilder entstehen⁸.

Für Industrieunternehmen stellt die Verfügbarkeit (hochwertiger) Daten einen zentralen Wettbewerbsfaktor für Wachstum und Innovationen dar. Datenbasierte Geschäftsmodelle und Plattformen gewinnen an Bedeutung. Dies führt zu einem steigenden Einsatz solcher Anwendungen in der Industrie, insbesondere beim Einsatz von KI⁹. Auch sind die mit Big Data gesammelten Informationen zunehmend wichtig, um die Cybersicherheit von Unternehmen oder öffentlichen IT-Infrastrukturen zu stärken, um sich gegen Cyberkriminelle zu wappnen, die ihrerseits große Datenmengen zu ihrem Vorteil nutzen.

Auch die Wissenschaft verspricht sich Vorteile durch die Analyse von Big Data bei Umwelt- und Klimaschutz¹⁴, bei medizinischer Forschung und Diagnostik¹⁵, bei der Vorhersage von Erdbeben und Epidemien und der Analyse von Migrationen oder Verkehrsstaus¹². Den Behörden können Big Data-Technologien bei der Erfüllung vieler öffentlicher Aufgaben helfen, wie etwa Daseinsvorsorge, Umwelt-Monitoring, Stadtentwicklung, Smart-City-Konzepte und Wirtschaftsförderung. Möglich sind auch Anwendungen



zur Abschätzung von Gesetzesfolgen oder zur Meinungsanalyse. Weitere Einsatzgebiete sind die öffentliche Sicherheit, die Sicherheit kritischer Infrastrukturen¹⁶ und die Kriminalistik¹².

Im Jahr 2022 hatten laut einer Studie 95 % der Firmen weltweit Probleme, aus ihren (komplexen) Daten Erkenntnisse zu ziehen. Ein Großteil setzt immer noch auf zeitaufwendige Tabellenkalkulation bei der Datenverarbeitung und berichtet von mangelnder Automatisierung und Qualifikationsdefiziten, besonders bei fortgeschrittenen Fähigkeiten wie dem Machine Learning¹⁷.

Im Jahr 2024 sehen zwar fast alle deutschen Unternehmen (98 %) eine große Bedeutung von Datenanalysen für die Wettbewerbsfähigkeit, aber nur 37 % der Unternehmen nutzen Big Data bereits. Immerhin 48 % diskutieren darüber oder planen den Einsatz. Die Hürden bei der Digitalisierung sind begründet durch einen Mangel an Fachkräften, fehlende Zeit, erhebliche Anforderungen an die technische Sicherheit und einen zumindest bei der Implementierung hohen Kostenaufwand^{18; 19}.

„Big Data as a Service“-Plattformen (BDaaS) können jedoch interessierten Unternehmen hohe Einstiegsinvestitionen für die nötige Technik vor Ort ersparen, indem sie alle Ressourcen in der Cloud bereitstellen. Zudem kümmern sich die BDaaS-Anbieter um Wartung und Updates der Plattformen und die Einhaltung von Compliance- und Datenschutzvorgaben. So ermöglichen technische Vorteile und Kostenersparnisse, vor allem bei nutzungsbasierten Abrechnungsmodellen, dass auch kleinere Unternehmen in die Analyse großer Datenmengen einsteigen können²⁰. BDaaS ist ein stark wachsendes Modell: Bis zum Jahr 2032 erwartet man eine durchschnittlichen jährliche Wachstumsrate von 25,3 % auf fast 80 Milliarden US-Dollar²¹.

Vor allem hohe Anforderungen an den Datenschutz¹⁸ können die schnelle Verbreitung von Big Data einschränken. Die Studie „Datenwirtschaft in Deutschland“ des Instituts der deutschen Wirtschaft verdeutlicht, dass die Datenschutz-Grundverordnung (DSGVO) auch nach einigen Jahren Anwendungspraxis Rechtsunsicherheiten verursacht. 85 % der befragten Unternehmen gaben „datenschutzrechtliche Grauzonen“ als Hemmnis für eine stärkere wirtschaftliche Datennutzung an. Die fehlende Rechtssicherheit bei der Anonymisierung von Daten wird als konkretes Beispiel von 73 % der Betriebe genannt²².

Unternehmen stehen beispielsweise vor der Herausforderung, dass keine einheitlichen und rechtssicheren Standards für die Anonymisierung von personenbezogenen Daten existieren. Die DSGVO enthält keine positive Definition des Begriffs der Anonymisierung⁹. Wenn die Fallzahlen von nachträglich gebildeten Gruppen zu klein werden, lassen sich unter Umständen wieder reale Personen zuordnen, sodass die Daten nach den geltenden rechtlichen Anforderungen allenfalls als pseudonymisiert, aber nicht als anonymisiert gelten¹⁶.

Auch die Zweckbindung im Bundesdatenschutzgesetz kann die Datennutzung hemmen: Bei personenbezogenen Daten fehlt (bisher) oft die Erlaubnis für eine nachträgliche Verwendung für andere Zwecke als die ausdrücklich bei der Erfassung abgefragten. So können bestehende Datenbestände in vielen Fällen nur anonymisiert oder mit einer neuen Einwilligung ausgewertet werden. Beides erfordert einen hohen Aufwand und ist in der Praxis oft kaum zu leisten¹⁶.

KI zur Nutzung von Big Data kann einerseits helfen, den Energieverbrauch zu senken, die Energieeffizienz zu verbessern und dem Klimawandel entgegenzuwirken, andererseits hat die Technologie selbst einen erheblichen CO₂-Fußabdruck. Vor allem das Training von KI-Modellen mit großen Datenmengen ist sehr energieintensiv. Man hat ermittelt, dass der Stromverbrauch von Datacentern und Datenübertragungsnetzwerken jeweils 1 bis 1,5 % des weltweiten Gesamtverbrauchs beträgt. Der damit verbundene Treibhausgasausstoß beläuft sich auf 1 % der energiebezogenen weltweiten Emissionen²³. Der Stromverbrauch aller Rechenzentren der Welt zusammen entsprach mit 500 bis 650 TWh im Jahr 2021 in etwa dem Strombedarf von ganz Deutschland²⁴ – und die Tendenz ist stark steigend. Teilweise sind rechtliche Vorgaben zum Energieverbrauch von KI-Systemen im europäischen AI Act verankert, der im Jahr 2026 in Kraft tritt²⁵. Offen bleibt, welche konkreten Auswirkungen die Vorgaben zur Ressourceneffizienz auf die Ausbreitung von Big Data haben werden.



Wer ist betroffen?

Vorreiter beim Einsatz von Big-Data-Analysen waren die Automobilindustrie, Versicherungen, die Chemie- und Pharmaindustrie und die Energiewirtschaft²⁶. Auch Tech-Startups nutzen Big Data und Data Analytics besonders

intensiv²⁷. Die Bedeutung von Big Data wächst aber stetig, auch für kleine und mittlere Unternehmen²⁸.



Beispiele

Beispiel 1

☑ Data Stewards sind Fachleute in Unternehmen, die sich um die Datenverwaltung kümmern und Verantwortung für ethische Aspekte, Datenqualität und -schutz tragen

Beispiel 3

☑ Kostenlose Online-Kurse zu Daten- und Algorithmenethik

Beispiel 2

☑ Die Berufsgenossenschaft der Bauwirtschaft (BG BAU) setzt auf KI zur Vermeidung von Arbeitsunfällen



Welche Veränderungen ergeben sich für die Sicherheit und Gesundheit der Beschäftigten?

Big Data kann in verschiedenen Branchen Prozesse optimieren und Arbeitsabläufe erleichtern, beispielsweise bei der Analyse von Maschinendaten und Produktionsmengen in der Produktion oder bei der Analyse von Energiekosten, Laufzeiten, Takteinhaltung und Revisionsarbeiten. In der Logistik lassen sich Prozessabläufe verbessern sowie die Personaleinsatz- und Kapazitätsplanung steuern. Im Personalmanagement kann man datenbasiert überprüfen, ob die Beschäftigten entsprechend ihrer Qualifikation eingesetzt sind oder wo es Möglichkeiten zur Optimierung und Flexibilisierung von Arbeitszeiten gibt²⁹, um Beschäftigte zu entlasten, ohne dass Betriebsabläufe beeinträchtigt werden. In der Öl- und Gasindustrie sind Unfälle oft die Folge von Bränden oder Explosionen oder auf eine Belastung mit schädlichen Substanzen zurückzuführen. Hier kann eine intelligente Gaserkennung mithilfe von Sensoren, Big Data und smarter Software eine entscheidende Rolle bei der Früherkennung gefährlicher Situationen spielen³⁰.

Big Data und lernende Systeme können Arbeitsschutzinstitutionen bei Aufsichtstätigkeiten unterstützen, indem sie etwa in kurzer Zeit Unternehmen mit besonders hohen Risiken ermitteln. Durch die höhere Effizienz steht mehr Zeit für die Beratungs- und Überwachungsaufgaben zur Verfügung. Diese können auf Basis der breiteren Informationsbasis auch zielgenauer erfolgen. Big Data kann auch genutzt werden, um Muster in Unfallstatistiken zu erkennen und so Risiken besser einzuschätzen. So können Unternehmen gezielt in die Prävention investieren und die Unfallraten und Kosten senken³¹.

Eine weltweite Studie (2023) unter Beschäftigten – je zur Hälfte Führungskräfte und Beschäftigte ohne Führungsverantwortung – hat ergeben, dass enorme Datenmengen zu Überforderung führen können. So sagen 91 % der Führungskräfte, dass die Datenflut den Erfolg ihres Geschäfts beeinträchtigt, 72 % gaben sogar an, gar keine Entscheidungen anhand von Daten mehr

zu treffen³². Die mögliche Belastung durch Big Data ist besonders für kleinere Betriebe wahrscheinlicher: Laut der Studie „Big Data – Big Chance für kleine und mittlere Unternehmen“, einer Befragung von kleinen und mittelgroßen Unternehmen (KMU) und Einrichtungen, die sich mit Daten befassen, fehlt es vielen KMU an Know-how, Zeit und den richtigen Technologien, um das Potenzial ihrer Daten auszuschöpfen²⁹.

Ein Ressourcenproblem stellen auch extrem große und unstrukturierte Datenmengen (Dark Data) dar – oft mangelt es an Expertise, um aus den schwer zu verarbeitenden Daten Erkenntnisse zu ziehen. Obwohl Dark Data ohne weitere Verarbeitung wertlos sind, beanspruchen sie Speicherplatz und entsprechende (Energie-)Ressourcen. Schätzungen gehen davon aus, dass es sich bei bis zu 90 % der Unternehmensdaten um solche Daten handelt³³. Eine Vielzahl unstrukturierter Daten ist vor allem für die Industrie herausfordernd (43 %), denn eine Auswertung ist nur dann sinnvoll, wenn die Daten von Relevanz sind und in hoher Qualität vorliegen, ansonsten kann es zu Fehlinterpretationen kommen – mit möglichen Folgen für den Unternehmenserfolg und die Arbeitsplatzsicherheit der Beschäftigten²⁹.

Sogar physische Gefahren oder Unfälle sind durch eine falsche Auslegung von Daten denkbar, sofern die Funktionsfähigkeit von Anlagen oder Maschinen betroffen ist. Bei der datengetriebenen Produktion werden Fertigungsprozesse optimal gestaltet; durch den Zugriff auf Echtzeitdaten sind potenzielle Probleme und Ausfälle frühzeitig erkennbar. Auch Wartungs- und Instandhaltungsarbeiten lassen sich durchführen, bevor es zu Ausfällen kommt (Predictive Maintenance)^{34;35}. Umgekehrt können Daten zu (sicherheitsrelevanten) Funktionen Risiken bergen, wenn sie irrtümlich zu einem Eingriff in laufende Prozesse Anlass geben. Möglich wäre dies etwa in der chemischen Produktion, wenn Kühlmitteldruck, -temperatur und -menge falsch gewählt würden und Reaktionen außer Kontrolle geraten³⁶.

Der Zugriff auf Big Data ermöglicht oft Rückschlüsse auf konkrete Personen, sodass sich Fragen zum Datenschutz und zum Schutz der Privatsphäre stellen. Big-Data-Analysen bergen die Gefahr, anonymisierte oder pseudonymisierte Daten zu deanonymisieren. Besonders bei umfangreichen und detaillierten Daten besteht eine hohe Wahrscheinlichkeit, dass durch einen Abgleich der Merkmalsdaten eine Re-Identifizierung gelingt³⁷.

Tatsächlich können Big-Data-Technologien Entscheidungen von Arbeitgebern beeinflussen und sich so auf die Beschäftigungs- und Aufstiegsaussichten von Angestellten auswirken. Mithilfe umfangreicher Datenbestände lassen sich Algorithmen trainieren, um Verhaltensanalysen von

Beschäftigten durchzuführen und Vorhersagen zu treffen. Daten über Arbeitnehmende können aus verschiedenen Quellen innerhalb und außerhalb des Arbeitsplatzes gesammelt werden, z. B. Anzahl der Tastaturklicks, Informationen aus sozialen Medien, Anzahl und Inhalt von Telefonaten, besuchte Websites, physische Präsenz, Bewegungen im Büro, Inhalt von E-Mails und sogar Tonfall und Körperbewegungen³⁸.

Allerdings sind Analysen von Verhaltensdaten mit den im Bundesdatenschutzgesetz (BDSG) fixierten Grundprinzipien der Datenvermeidung und der Datensparsamkeit nur schwer vereinbar. Zudem scheidet eine wirksame Einwilligung zur Datenverwendung für Big-Data-Analysen in der Praxis meist aus. Und selbst die Erlaubnistatbestände nach dem BDSG sind aufgrund der fehlenden Verhältnismäßigkeit und der Persönlichkeitsrechte der Beschäftigten meist nicht gerechtfertigt³⁹. Insgesamt scheint die Situation auch für Betriebe schwierig: Eine Studie des Fraunhofer-Instituts für Sichere Informationstechnologie (2023) kritisiert, dass der aktuelle Rechtsrahmen für eine sichere Verarbeitung von Big Data und eine Anonymisierung personenbezogener Daten nicht ausreichend und für Rechtsunsicherheit bei Unternehmen sorgt⁴⁰.

Big Data benötigt nicht nur einen angemessenen Datenschutz bei Beschäftigten, sondern stellt gleichzeitig hohe Anforderungen an den Umgang mit großen Datenmengen seitens der verantwortlichen Beschäftigten. So sollte etwa ein Schutz der Daten über ihren gesamten Lebenszyklus im Sinne eines „Data Life-Cycle Management“ (DLCM) gewährleistet sein. Dies gilt naturgemäß in erster Linie für sensible Daten, die vom Zeitpunkt ihrer Erstellung gesichert sein müssen⁴¹. Bei Zeitdruck und knappen Personalressourcen kann es zu Arbeitsverdichtung und Überforderung kommen.

Eine zusätzliche Arbeitsbelastung kann darin bestehen, dass viele Daten in Unternehmen noch in veralteten Systemen oder analogen Formaten wie Papier oder Film gespeichert sind. Um alle Daten automatisiert analysieren zu können, müssen diese Informationen digitalisiert und zentral erfasst werden. Hier ist eine Digitalstrategie erforderlich, die die Dokumenten- und Informationssicherheit und auch die sichere Entsorgung von Papierdokumenten berücksichtigt⁴². Bei Behörden und öffentlichen Institutionen müssen Daten überdies über Jahrzehnte hinweg nutzbar bleiben, unabhängig vom verwendeten Speichersystem. Die Europäische Kommission will aber die interoperable Archivierung und Datenverwaltung fördern und unterstützt durch ihre „eArchiving Initiative“ Organisationen bei der langfristigen Bewahrung von Informationen⁴³.



Was sind Erkenntnisse und Perspektiven für den Arbeitsschutz?

- ❖ Die Datenmengen der digitalen Gesellschaft wachsen rasant. Daher ist der richtige Umgang mit Big Data branchenübergreifend von hoher Relevanz.
- ❖ Big Data ist untrennbar mit KI verwoben; zusammen können die Technologien wirtschaftlichen und wissenschaftlichen Fortschritt und Innovationen schaffen sowie Prozesse und Tätigkeiten verändern und effizienter machen.
- ❖ Die Nutzung von Big Data, z. B. zur Produktionssteuerung oder Absicherung gefährlicher Prozesse, birgt Potenzial zur Entlastung und zum Schutz von Beschäftigten. Dem Arbeitsschutz kommt die Aufgabe zu, sicherzustellen, dass diese Chancen genutzt und nicht durch Sorgen vor Rationalisierung, Überwachung und Überforderung auf Seiten der Belegschaften konterkariert werden. Nicht zuletzt wirft der Umgang mit Big Data am Arbeitsplatz Fragen hinsichtlich des Datenschutzes und dem Schutz der Persönlichkeitsrechte am Arbeitsplatz auf.
- ❖ Für eine menschengerechte Nutzung von KI und Big Data im Sinne des Arbeitsschutzes braucht es eine stärkere Vernetzung der Fachrichtungen Informatik, Statistik, Rechts- und Sozialwissenschaften.
- ❖ Auch der gesetzlichen Unfallversicherung bietet Big Data neue Möglichkeiten, interne Prozesse zu optimieren und Präventionsangebote zu schärfen. Eine abgestimmte Nutzungsstrategie und Kooperation können helfen, gemeinsame Handlungsfelder zu identifizieren und Synergien zu nutzen.
- ❖ Die Unfallversicherungsträger (UVT) und ihre Aufsichtspersonen als direkte Ansprechpartner für die Betriebe müssen für eine sachkundige betriebliche Beratung zum arbeitsschutzkonformen Einsatz von Big Data (und KI) befähigt werden. Besondere Unterstützung benötigen kleine und mittlere Unternehmen, für die der Umgang mit Big Data angesichts fehlender (Personal-)Kapazitäten meist eine große Herausforderung darstellt.
- ❖ Das "Kompetenzzentrum Künstliche Intelligenz und Big Data" am Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung (IFA) unterstützt die UVT bei der Planung und Durchführung konkreter Vorhaben. Zudem ist es Anlaufstelle gegenüber Politik, Forschung und Gesellschaft.

Herausgegeben von:

Deutsche Gesetzliche Unfallversicherung e.V.
(DGUV)
Glinkastraße 40 · 10117 Berlin
Telefon: 030 13001-0 (Zentrale)
E-Mail: info@dguv.de
Internet: www.dguv.de

Institut für Arbeitsschutz der Deutschen
Gesetzlichen Unfallversicherung (IFA),
Risikoobservatorium der DGUV

Verfasst von: Dr. Ruth Klüser

Ausgabe:

November 2024

Satz & Layout:

Atelier Hauer + Dörfler, Berlin

Copyright:

Diese Publikation ist urheberrechtlich geschützt. Die Vervielfältigung, auch auszugsweise, ist nur mit ausdrücklicher Genehmigung gestattet.

Bezug: www.dguv.de/publikationen

Die **Literaturliste** ist in der Online-Fassung der Trendbeschreibung verfügbar.

❖ www.dguv.de/ifa
🔍 risikoobservatorium

