

Big Data – Chancen und Herausforderungen

Große Chancen bieten Big-Data-Verfahren unter anderem in der Medizin, in der Forschung verschiedener wissenschaftlicher Disziplinen oder für den Straßenverkehr. Durch die Analysen von großen Datenmengen sind neue Erkenntnisse möglich. Zusammenhänge werden sichtbar, die man ohne Big Data nicht erkannt hätte. Dem stehen große Herausforderungen gegenüber. Sie sind verbunden mit der Frage: Wie sieht ein verantwortungsvoller Umgang mit den riesigen Datenmengen und den aus Analysen gewonnenen Erkenntnissen aus – insbesondere, wenn sie personenbezogene Informationen enthalten?

Chancen von Big Data sehen und fördern, ohne die Risiken auszublenden.

Die im März 2018 in Kraft gesetzte Datenschutz-Grundverordnung der Europäischen Union (DSGVO) enthält umfangreiche Regeln und Vorschriften, die den Umgang mit Big Data betreffen. Dazu gehören die datenschutzrechtlichen Prinzipien, grundsätzlich nicht ohne Einwilligung des Datensubjekts Daten zu sammeln, stets nur solche personenbezogenen Daten zu erheben, die für einen bestimmten, vorher festgelegten Zweck erforderlich sind, und sie auch nur so lange wie unbedingt notwendig zu speichern.

Hier entsteht ein Dilemma: Einerseits gilt es, die Bürgerinnen und Bürger vor Datenmissbrauch zu schützen und ausreichend zu informieren. Andererseits sollen die Grundsätze der freien Marktwirtschaft geschützt und die Innovationskraft einer Gesellschaft gestärkt werden. Gesetzliche Regulierungen und Vorschriften dürfen demnach auch die Aktivitäten von neuen, innovativen Unternehmen, die auf Onlinedienste und Big-Data-Verfahren spezialisiert sind, nicht unnötig behindern. Gesellschaft und Politik sind also gefordert, die Chancen von Big Data zu sehen und zu fördern. Durch die computergestützten Datenanalysen sind neue Erkenntnisse möglich. Beispielsweise stellte der Suchmaschinenbetreiber Google vor vielen Jahren beim Analysieren von Suchanfragen eine Korrelation zwischen der vermehrten Suche nach Begriffen wie „Halsweh“ und der Anzahl an Grippepatientinnen und -patienten fest. Das Unternehmen nutzte diese Erkenntnisse, um vorherzusagen, wo sich eine Grippe auszubreiten drohte und wann es sich für bestimmte Unternehmen lohnte, gezielt für Medikamente oder andere Maßnahmen gegen Grippe zu werben. Zusammenhänge solcher Art können nicht nur für Werbetreibende nützlich sein, sondern auch für Individuen oder das Gesundheitssystem. Wer davon profitiert, ist abhängig davon, wer Zugriff auf die Ergebnisse der Datenanalysen hat und wie sie interpretiert werden.

Bei dem Beispiel der Grippewelle zeigte sich, dass die Medienberichterstattung zur Grippeprognose das Suchverhalten der Internetnutzerinnen und -nutzer beeinflusste. Der von Google diagnostizierte Zusammenhang zwischen den Suchanfragen und Grippewellen wird also durch weitere Faktoren beeinflusst. Google änderte daraufhin den Umgang mit den Daten und den Zugang zu ihnen. Dennoch sind einige Big-Data-Auswertungen für das Leben der Menschen von unmittelbarem Nutzen, z. B. die täglichen und sehr detaillierten Wettervorhersagen für zahlreiche Orte, Frühwarnsysteme für Katastrophen wie Vulkanausbrüche oder Tsunamis, aber auch die vielen Erkenntnisse aus umfangreichen Datenanalysen in Bezug auf den Klimawandel. Sie alle beruhen darauf, unzählige Messdaten auszuwerten, um darin Zusammenhänge zu erkennen, die auch langfristige Prognosen ermöglichen.

Wissenschaft und Forschung mit Big Data: Unbekanntes aufdecken

Mitunter ergeben sich auch Zusammenhänge, nach denen die entsprechenden Expertinnen und Experten gar nicht explizit gesucht hatten und die ihnen ohne Big Data wahrscheinlich verborgen geblieben wären. Beispielsweise entdeckten Forscherinnen und Forscher durch entsprechende

Big-Data-Auswertungen bis dahin unbekannte Infektionssymptome bei Frühgeborenen. Diese Erkenntnis ermöglicht es, Infektionen früher zu diagnostizieren und zu behandeln. Generell sieht die Medizin große Chancen in breit angelegten Datenanalysen. Mit ihrer Hilfe ist es möglich, Zusammenhänge zu finden, beispielsweise zwischen Genen und Blutwerten einerseits sowie Krankheiten wie Alzheimer, Krebs oder Schizophrenie andererseits. Dies wiederum kann helfen, die Therapien dieser Krankheiten zu verbessern.

Zu den Voraussetzungen, um solche Chancen von Big Data nutzen zu können, gehört unter anderem, möglichst viele Daten zusammenzuführen. In Bezug auf das Beispiel der Frühgeborenen bedeutet das, dass die Daten aus zahlreichen unterschiedlichen Regionen stammen und über lange Zeiträume hinweg erhoben worden sein sollten. Diese Daten müssen darüber hinaus für die computergestützte Bearbeitung freigegeben sein. Das ist insofern eine Herausforderung, als große Datenbanken viel Programmieraufwand und damit Investitionen erfordern, sodass man ihre Nutzung deshalb schützen will. Zum anderen, weil die Daten viele persönliche Informationen über Individuen beinhalten, die in bestimmten Zusammenhängen zu ihrem Nachteil eingesetzt werden könnten. Zudem stehen vielfach auch die Inhalte der Datensammlungen unter rechtlichem Schutz. Das ist meist dann der Fall, wenn es sich nicht nur um Messwerte oder Symptome handelt, sondern beispielsweise um zusammenhängende Texte, die als geistiges Werk angesehen werden. Sie sind durch das Urheberrecht geschützt.

Wenn Wissenschaftlerinnen und Wissenschaftler etwa analysieren, wie Sprache verwendet wird, nutzen sie beispielsweise Songtexte von Künstlerinnen und Künstlern in einer Art und Weise, wie es nach dem Urheberrecht erlaubt ist. Wenn die Songtexte aber in einer wissenschaftlichen Studie zitiert werden und diese Studie veröffentlicht wird, liegt eine Urheberrechtsverletzung vor. Deshalb wurde und wird derzeit in Deutschland sowie der Europäischen Union an weiteren gesetzlichen Rahmenbedingungen zum Umgang mit Daten gearbeitet. Sie sollen dafür sorgen, dass sich die Chancen, die Big Data bietet, besser und kontrollierter nutzen lassen, ohne dabei Grundsätze zu verletzen, wie z. B. den Schutz geistiger Schöpfung und geistigen Eigentums oder etwa der Persönlichkeitsrechte. Zugleich finden in Medien, Politik und Fachkreisen Diskussionen darüber statt, ob diese Vorhaben ausreichend sind oder ob sie womöglich sogar andere grundlegende Rechte verletzen.

Herausforderungen für den Datenschutz

Eine große Herausforderung für alle, die mit Big Data zu tun haben, stellt sich bereits bei der Datenerfassung. Immer wieder entdecken Fachleute, dass ans Internet angeschlossene Geräte viel mehr Daten sammeln, als für ihr Funktionieren nötig ist, etwa wenn Taschenlampen-Apps auf dem Smartphone auf die vom GPS-Tracker gelieferten Ortskoordinaten oder auf das Adressbuch zugreifen. Ein anderes Beispiel sind interaktive Lautsprecher, die auf Zurufe reagieren, aber immer wieder auch dann Gesprochenes aufzeichnen, wenn sie eigentlich im Ruhemodus sind. Zwar lassen sich solche überflüssigen Zugriffe verhindern. Doch häufig machen die Hersteller nicht transparent, in welchem Umfang ihre Apps oder Geräte Daten erfassen, sodass man als Nutzerin oder Nutzer nichts davon mitbekommt.

Dies markiert die nächste große Herausforderung: verantwortungsbewusst mit den erfassten Daten umzugehen. Hersteller von Endgeräten wie Smartphones oder Wearables, die zahlreiche Daten erheben, können dazu beitragen, indem sie beispielsweise transparent und deutlich darüber aufklären, welche Daten sie von welchen Sensoren für welche Zwecke übernehmen. Zudem zählt zu einer verantwortungsbewussten Datenbehandlung, die Nutzerin oder den Nutzer verständlich anzuleiten, wo und wie man in Geräten oder Apps Einstellungen zum Schutz der eigenen Daten vornehmen kann. Zuletzt ist auch jede und jeder Einzelne verantwortlich dafür, sich über die Weitergabe der eigenen Daten zu informieren und Daten möglichst sparsam preiszugeben.

Einige Hersteller erteilen bereits solche Hinweise zur Datennutzung bzw. dem Datenschutz. Hier sind die Nutzerinnen und Nutzer gefordert: Sie müssen diese Hinweise lesen und nicht einfach schnell und vollständig abhaken, z. B. wenn die App eines städtischen Verkehrsbetriebs den Zugriff auf die Standortdaten erbittet, zugleich aber diese und darüber hinaus weitere Daten sammeln will, „um den Service zu verbessern“. Hier wird pauschal nach der Zustimmung gefragt. Bei einer Ablehnung ist häufig damit zu rechnen, dass der jeweilige Dienst nur eingeschränkt oder gar nicht genutzt werden kann. Abgestufte Zustimmungsverfahren, die eine differenziertere Entscheidung ermöglichen, existieren nicht. So bleibt es trotz bereits bestehender Hinweise eine Herausforderung für Nutzerinnen und Nutzer, sich darüber bewusst zu sein, wann sie welche Daten weitergeben. Die Verantwortung des oder der Einzelnen, sich über die Datenweitergabe zu informieren, ist demnach bislang häufig schwer zu erfüllen.

Auch für Staat und Gesellschaft stellt Big Data eine Herausforderung dar. Wie sich in den vergangenen Jahren an vielen Skandalen gezeigt hat, sind große Internet-Unternehmen wie Facebook, Google oder Amazon immer wieder darin verwickelt, dass mehr Daten als angegeben oder sehr persönliche Daten gesammelt und/oder Dritten zugänglich gemacht werden. Forscherinnen und Forscher der Universität Oxford untersuchten Ende 2018 eine Million Apps und stellten fest: Ungefähr 90 Prozent sind so programmiert, dass Daten mit Google geteilt werden könnten. Bei etwas weniger als 43 Prozent der Apps könnten die Daten an Facebook gehen.

Solche Entdeckungen oder skandalösen Fälle von Datenweitergaben im großen Stil bzw. programmierten Möglichkeiten hierfür verdeutlichen, wie wichtig es ist, in Bezug auf Big Data gesetzliche Rahmenbedingungen festzulegen und Grenzen zu setzen. Ob beim datenbasierten Wahlkampf oder bei personalisierter Werbung – in vielen Bereichen besteht die Gefahr, dass durch das Ansammeln und Kombinieren von Daten mehr über einzelne Personen ermittelbar ist, als es der Datenschutz erlaubt. Dies wiederum schürt die Gefahr und Angst der Bürgerinnen und Bürger, dadurch transparent oder „gläsern“ zu werden.

Um die Chancen von Big Data zu nutzen und den Herausforderungen zu begegnen, ist es notwendig, dass alle, die mit Big Data arbeiten, die gesetzlichen Rahmenbedingungen einhalten sowie neue ergänzende Regelungen gefunden werden. Denn mit der Digitalisierung und den damit einhergehenden Möglichkeiten im Bereich Big Data ergeben sich Veränderungen für das Individuum und die Gesellschaft, die rechtlich bislang nicht berücksichtigt sind. Zugleich gilt für jede und jeden Einzelne/n wie auch für Unternehmen, Organisationen und den Staat, grundsätzlich verantwortungsbewusst mit den umfangreichen Daten umzugehen. Darüber hinaus ist es notwendig, eine breite gesellschaftliche Debatte zu den Chancen und Herausforderungen zu führen, die nicht nur Expertinnen und Experten einbezieht. Denn dann verbreiten sich auch Wissen und eine kritische Auseinandersetzung mit Big Data. Das kann schließlich dazu beitragen, dass sich eine größere digitale Mündigkeit ausbildet, die zu bewussterem Handeln führt.