Big Data – das große Versprechen der neuen digitalisierten Welt

Im Informationszeitalter ist Big Data ein Schlüsselbegriff. In ihm spiegeln sich die Chancen und Risiken einer digitalen Gesellschaft wie im Brennglas. Dass die Nutzung und Auswertung von Massendaten nahezu alle Bereiche des Lebens durchdringt, steht längst außer Frage. Auch das Gesundheitswesen steht vor einer umfassenden Digitalisierung.

Rund 100 Millionen Einträge spuckt die Suchmaschine in Sekundenbruchteilen aus, wenn man den Begriff Big Data eingibt. Wofür steht dieser Begriff? „Big Data" hat meist zwei Bedeutungen. Der Begriff benennt ein Problem, nämlich die stetig anwachsende Flut von Daten, die längst jede menschliche Vorstellung übersteigt und herkömmliche Analyseverfahren überfordert (Klein). Der Begriff steht aber auch gleichzeitig für die Lösung dieses Problems. Denn wer Big Data sagt, meint auch diejenigen Techniken und analytischen Methoden, welche diese Datenflut erfassen, auswerten und speichern können.

Schillernde Ambivalenz

Informationelle Selbstbestimmung reicht im Big-Data-Zeitalter nicht mehr. Der Ethikrat fordert stattdessen Datensouveränität. Diese soll es ermöglichen, die Persönlichkeit interaktiv zu entfalten unter Wahrung von Privatheit in der vernetzten Welt. Viele Institutionen und staatliche Akteure sieht das Gremium in der Pflicht, den (außer- und) rechtlichen sowie technischen Rahmen zu setzen.
© Deutscher Ethikrat

Der Begriff selbst ist relativ jung. 2008 erobert er den Titel des Wissenschaftsmagazins „Nature". Denn er verweist auf ein drängendes Problem der Forschung: In vielen Disziplinen sprengen die in Versuchen und Simulationen erzeugten Datenmengen alle zuvor gekannten und beherrschbaren Größenordnungen. Zwei Jahre später schon thematisiert das Wirtschaftsblatt „The Economist" die Datenflut in ihren wirtschaftlichen Auswirkungen und beschreibt deren Chancen und Herausforderungen.

In seiner Stellungnahme "Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung" verwendet der Deutsche Ethikrat eine „Arbeitsdefinition". Das Gremium zieht damit aus der internationalen Diskussion und der Erkenntnis, dass es keine einheitliche Begriffsbestimmung gibt (vgl. de Mauro), nicht nur einen eigenen Schluss. Sondern es schreibt Big Data eine weitere Eigenschaft zu, die auch und gerade im Bereich der Gesundheit wichtig ist: die der Möglichkeit zur neuen Verknüpfung:

„Big Data ist der Umgang mit großen Datenmengen, der darauf abzielt, Muster zu erkennen und daraus neue Einsichten zu gewinnen, und der hierzu angesichts der Fülle und Vielfalt der Daten sowie der Geschwindigkeit, mit der sie erfasst, analysiert und neu verknüpft werden, innovative, kontinuierlich weiterentwickelte informationstechnologische Ansätze nutzt." (Deutscher Ethikrat, 2017, S. 36)

Plötzlich können alle Daten gesundheitsrelevant werden

Immer mehr, immer schneller und immer disparater. Big-Data-Nutzung braucht nicht nur künstliche Intelligenz, sondern – immer noch – den Menschen. Data Science, darin sind sich Fachleute einig, wird eine Schlüsseldisziplin im Digitalzeitalter.
© Europäische Kommission

Gut drei Jahre ist es her, da behauptete die deutsche Hochschulmedizin (Positionspapier IT-Lösungen, S. 2), dass auf den Gesundheitsbereich künftig ein „Drittel der weltweit erhobenen und austauschbaren Daten" entfallen. Jetzt setzt der Ethikrat noch eines drauf: Zur Logik von Big Data gehöre es, personenbezogene Daten aus einer immer größeren Zahl von Quellen ständig neu zu kombinieren, sodass nahezu alles gesundheitsrelevant werden kann, auch Daten sogenannter Wearables, Sensoren oder Äußerungen in sozialen Medien (Deutscher Ethikrat, 2017, S. 4ff., 54).

Big Data speisen sich im Gesundheitsbereich aus einer Vielzahl von Quellen: klinischen oder (Kohorten-) Studien, elektronischen Patientenakten (die in Deutschland immer noch in Entwicklung sind, aber bereits von Privatfirmen angeboten werden), Registern, Biobanken sowie den durch Hochdurchsatztechnologien erzeugten ‚Panomics'. Wohin diese Entwicklung führt, lässt sich schon heute an der UK Biobank absehen. Dort werden 2020 die Genome einer halben Million Briten öffentlich gemacht. Bereits jetzt zahlen Pharma-Firmen für die Nutzung dieser Big Data und versprechen sich davon einen Schub in der Arzneimittel-Entwicklung und für die biologische Forschung. (Technology Review, 08.01.2018).

IT-Größen nehmen Gesundheitsgeschäft ins Visier

Noch allerdings fehlen die infrastrukturellen Voraussetzungen in vielen Bereichen des Gesundheitswesens für eine systematische Nutzung von Big Data, um gesundheitsrelevante Daten wie die von EKG, Blutwerten, Bildern, Abrechnungen oder -omics-Daten von allen Patientinnen und Patienten einer Klinik oder Krankenhauskette zusammenzufassen, zu analysieren und auszuwerten.

eHealth heißt das Ziel, von dem das deutsche Gesundheitswesen noch ein Stück weit entfernt ist. Damit Big Data ihre mögliche Wirkung entfalten können, braucht es eine digitale Datenbasis für individuelle Patientenmodelle sowie Informations- und Ratgebersysteme. Diese Infografik beschränkt sich auf den ersten Gesundheitsmarkt und bildet den Datenfluss aus der Forschung nicht ab.
© acatech

Hinzu kommen im Zeitalter einer nahezu alle Lebensbereiche durchdringenden Digitalisierung Daten, die von Internet- und Mobilfunk-Unternehmen gewonnen werden, Daten, die in wissenschaftlichen Publikationen veröffentlicht wurden, sowie des Weiteren Daten, die in einer ärztlichen Umgebung oder mit ärztlicher Begleitung und mit Medizingeräten aufgenommen wurden, und schließlich Daten von ‚Wearables' und ähnlichen Geräten zur Selbstvermessung. Dass sich IT-Unternehmen wie Alphabet, der Mutterkonzern von Google, vom Geschäft mit Gesundheitsdaten viel versprechen, zeigt das dichte Geflecht von Beteiligungen, Investitionen, Forschungsprojekten und Kooperationen (Gigerenzer, 2016, S. 34f.) im Gesundheitsbereich.

Die intelligente Verknüpfung dieser Daten wird, so die Hoffnung, das medizinische Wissen in Zukunft deutlich erhöhen und absichern. Mit Methoden der Wissensmodellierung, des maschinellen Lernens und Cognitive Computing können Rechner selbstständig neues Wissen aus diesen Daten generieren (vgl. "Mit Klassifikatoren und multimodaler Datenfusion Big Data Sinnvolles entlocken"). Aus deren Integration und Analyse erhoffen sich die Akteure des Gesundheitswesens, bislang unerkannt gebliebene neue Muster zu erkennen und daraus neue Einsichten respektive Entscheidungshilfen zu gewinnen. Damit ergeben sich Chancen für verbesserte Diagnostik (z. B. Panomics-getriebene Identifizierung von Biomarkern), Therapie, Vorsorge und Kostensenkung.

Onkologie ist Vorreiter der Big-Data-Nutzung

Unter dem Stichwort Präzisionsmedizin, individualisierte resp. stratifizierte Medizin werden bereits heute, vor allem in der Onkologie, Big Data aus der Forschung mit einer Vielzahl klinischer Daten zusammengeführt (acatech, 48f.). Davon erhofft man sich genauere Diagnosen und auf das molekulare Patientenprofil zugeschnittene personalisierte Therapien (vgl. "Mit Big Data genauer therapieren"). Dafür jedoch müssen in Uniklinika erst die Voraussetzungen geschaffen werden, sowohl für Patienten als Datengeber wie auch für (nicht-)medizinische Forscher. Mit den „Eckpunkten für eine Heidelberger Praxis der Genomsequenzierung" (Marsilius-Kolleg der Universität Heidelberg und EURAT-Gruppe, 2013) haben sich die Heidelberger Forscher von Universität und DKFZ einer Art Selbstregulierung im Big-Data-Zeitalter unterworfen, die das Patientenwohl, die Forschungsfreiheit und den klinischen Fortschritt auszutarieren versucht.

Die Nutzung von Big Data im Gesundheitswesen erstreckt sich auf eine Vielzahl von Akteuren mit teilweise sich widersprechenden Interessen (Deutscher Ethikrat, 2017, S. 11). Die biomedizinische Forschung erhofft sich dadurch ein besseres Verständnis wissenschaftlich relevanter Zusammenhänge und Prozesse durch bildgebende und molekularbiologische Verfahren. Die Gesundheitsversorgung (sog. erster Markt) erwartet sich vom Big-Data-Einsatz stärker personalisierte Behandlungen und bessere Effektivität und Effizienz, Versicherer und Arbeitgeber präzise Profile einzelner Personen und Personengruppen. Global agierende IT- und Internetfirmen (sog. zweiter Markt) erstreben eine wachsende kommerzielle Verwertung durch die Verknüpfung und Analyse gesundheitsrelevanter Daten mit zahlreichen anderen Informationen. Betroffene bzw. Patienten schließlich erwarten durch die wachsende Selbstvermessung relevante Informationen zu einem gesundheitsbewussten Lebensstil und für ihr persönliches Wohlergehen (nach Deutscher Ethikrat, 2017, S. 11ff.)

Droht ein Angriff auf die evidenzbasierte Medizin?

All diesen Chancen bei der Nutzung von Big Data stehen im selben Maße Risiken (Dual Use) gegenüber. Im hochsensiblen Gesundheitsbereich sehen Experten das traditionelle Datenschutzrecht, gerade im Blick auf den Patienten, für den wachsenden Umgang und die Nutzung von Big Data überfordert. Nicht wenige Stimmen (vgl. "Mehr Daten bedeutet nicht automatisch mehr Information") warnen vor der Gefahr, Korrelationen und Datenmuster an die Stelle von Theorien und

Gesetzen zu setzen, sich vom Goldstandard evidenzbasierter Medizin klammheimlich zu verabschieden und die Korrelation über die Kausalität zu stellen (z. B. Wissenschaftstheoretiker wie Prof. Klaus Mainzer, 2015). „Big Data bleibt (...) ergänzungs- und kontrollbedürftig" (Deutscher Ethikrat, 2017, Vorsitzender Prof. Steffen Augsberg, PK Ethikrat, 30.11.2017)

„Der mögliche Nutzen von Big Data steht und fällt mit der Expertise und Integrität der Personen oder Institutionen, die Daten generieren, auswählen, verknüpfen und interpretieren." (Deutscher Ethikrat, 2017, S. 45). Angesichts des von verschiedenen Interessen befeuerten Hypes um Big Data sollte diese Feststellung stets Gehör finden: „Es wäre also ein Missverständnis zu glauben, dass mehr Daten auch automatisch zu mehr Wissen über kausale Effekte führen." (Deutscher Ethikrat, 2017, S. 47)

Literatur:

Acatech (Hg.): Individualisierte Medizin durch Medizintechnik, München 2017

Auffray, C. et al. : Making Sense of big data in health research: Towards an EU action plan, in: Genome Medicine (2016) 8:71, DOI: 10.1186/s13073-016-0323-y

BIO Deutschland (Hg.): Positionspapier der Bio Deutschland zum Thema Big Data im Gesundheitswesen, Berlin, 9.12.2016. https://www.biodeutschland.org/de/positionspapiere/positionspapier-zur-anwendung-von-big-data-im-gesundheitswesen.html

Bitkom (HG.) Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte, Berlin 2012, S. 7, S. 19ff.

Bundesverband der Pharmazeutischen Industrie (Hg.): Whitepaper. Digitalisierung in der Pharmaindustrie. Die Revolution ist Realität, Berlin Juni 2017.

De Mauro, A. et al.: What is big data? A consensual definition and a review of key research topics, AIP Conference Proceedings 1644, 97 (2015); doi: 10.1063/1.4907823, https://nowherevilleblog.files.wordpress.com/2015/05/what-is-big-data.pdf

Deutsche Hochschulmedizin: Positionspapier Medizinischer Fortschritt braucht leistungsstarke IT-Lösungen, Berlin, 1. Juli 2014 https://www.uniklinika.de/fileadmin/user_upload/pdf/2014-07-09_Deutsche_Hochschulmedizin_e.V._IT-Positionspapier_Final.pdf

Deutscher Ethikrat (Hg.): Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung. Stellungnahme, Berlin, 30.11.2017. http://www.ethikrat.org/dateien/pdf/stellungnahme-big-data-und-gesundheit.pdf

Deutscher Ethikrat (Hg.): Die Vermessung des Menschen – Big Data und Gesundheit, Jahrestagung, 21.5.2015, Simultanmitschrift. http://www.ethikrat.org/dateien/pdf/jt-21-05-2015-simultanmitschrift.pdf

Europäisches Parlament: Bericht über die Folgen von Massendaten für die Grundrechte: Privatsphäre, Datenschutz, Nichtdiskriminierung, Sicherheit und Rechtsdurchsetzung (2016/2225(INI)), 20.02.2017 http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A8-2017-0044+0+DOC+XML+V0//DE

European Commission: Study on Big Data in Public Health, Telemedicine and Healthcare, December, 2016, DOI:10.2875/734795

Floridi, L.: Big Data and Their Epistemological Challenge, Philos. Technol. (2012) 25: 435. https://doi.org/10.1007/s13347-012-0093-4

Fraunhofer-Institut für Intelligente Analyse-und Informationssysteme IAIS (Hg.): Big Data – Vorsprung durch Wissen. Innovationspotentialanalyse, Sankt-Augustin 2012. https://www.iais.fraunhofer.de/content/dam/iais/gf/bda/Downloads/Innovationspotenzialanalyse_Big-Data_FraunhoferIAIS_2012.pdf

Gigerenzer, G. et al.: Digitale Welt und Gesundheit. eHealth und mHealth – Chancen und Risiken der Digitalisierung im Gesundheitsbereich, in: Sachverständigenrat für Verbraucherfragen, Berlin, Januar 2016. http://www.telemedallianz.de/pdf/studien/01192016_Digitale_Welt_und_Gesundheit.pdf

Klein, D. et al.: Big Data, in: Informatik-Spektrum, June 2013, , DOI 10.1007/s00287-013-0702-3.

Krumholz,H. Big Data and New Knowledge in Medicine: The Thinking, Training, and Tools Needed for a Learning Health System, in: Health Affairs, 33: 7, 2014, S. 1163–1170. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5459394/

Leimbach, T./Bachlechner, D.: Big Data in der Cloud, in: Büro für Technikfolgenabschätzung beim Deutschen Bundestag (Hg.): Big Data in der Cloud (2014), S. 39ff., Berlin 2014 (Hintergrundpapier Nr. 19).

Mainzer, K. Die Berechnung des Menschen. http://www.ethikrat.org/dateien/pdf/jt-21-05-2015-mainzer.pdf

Marquardt, W.: Was ist Big Data? Versuch einer Positionsbestimmung, Jahrestagung des Deutschen Ethikrates, 2.6.2015. http://www.ethikrat.org/dateien/pdf/jt-21-05-2015-marquardt.pdf

Marsilius-Kolleg der Universität Heidelberg und EURAT-Gruppe. Eckpunkte für eine Heidelberger Praxis der Genomsequenzierung. https://www.uni-heidelberg.de/md/totalsequenzierung/informationen/mk_eurat_stellungnahme_2013.pdf

Nature: Science in the petabyte era http://www.nature.com/nature/journal/v455/n7209/covers/

PWC: Weiterentwicklung der eHealth-Strategie. Studie im Auftrag des Bundesministeriums für Gesundheit, Berlin, Oktober 2016. https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/E/eHealth/BMG-Weiterentwicklung_der_eHealth-Strategie-Abschlussfassung.pdf

Ramge, T.: Wie groß ist Big Data, in: Brand Eins 07/2016, https://www.brandeins.de/archiv/2016/digitalisierung/wie-gross-ist-big-data/

Regalado, Antonio, 8.1.2018, https://www.technologyreview.com/s/609897/500000-britons-genomes-will-be-public-by-2020-transforming-drug-research/

Rüping, S./Graf, N.: Big Data in der Medizin, Deutsches Ärzteblatt, Jg. 110, Heft 41, 11. Oktober 2013.

Springer Gabler Verlag (Herausgeber), Gabler Wirtschaftslexikon, Stichwort: Big Data, online im Internet: http://wirtschaftslexikon.gabler.de/Archiv/-2046774198/big-data-v5.html

Vfa, 18. Juli 2017 , "Panomics" in der Personalisierten Medizin.

https://www.vfa-bio.de/vb-de/aktuelle-themen/branche/panomics-in-der-personalisierten-medizin.html

Zusammenführen verschiedener Datenformate im Nationalen Centrum für Tumorerkrankungen, Uni Heidelberg, mit SAP
https://www.nct-heidelberg.de/das-nct/vorstellung/innovationen/nct-datatherehouse.html

Dossier

06.03.2018
Walter Pytlik
© BIOPRO Baden-Württemberg GmbH

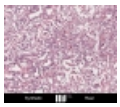
Weitere Artikel in diesem Dossier



22.09.2026
Der Zwilling aus Daten – Zukunftsmedizin für Neugeborene



04.05.2026
Frakturen im Fokus: Vorhersagen statt Abwarten



25.09.2025
Leistungsfähige KI-Systeme durch die Nutzung von synthetischen Trainingsdaten