

Big Data: Wissen ist Macht



Prof. Dr. med. Manfred Wildner

Der englische Philosoph und Lordkanzler seiner Majestät, Francis Bacon (1561–1626) prägte die Formulierung „ipsa scientia potestas est“, (dt.: „das Wissen selbst ist Macht“) vor über 400 Jahren. Dabei hatte er gar nicht so sehr die Macht seines Königs im Sinn, das philosophische Programm der Aufklärung oder sein „Novum organon scientiarum“, seine neue wissenschaftliche Methodenlehre. Der (Mit-)Begründer der modernen Naturwissenschaften dachte in seinen „Meditationes sacrae“ von 1597 vielmehr über die göttliche Macht nach und meinte, dass allein schon das göttliche Wissen einen Großteil der göttlichen Macht begründet. Eine ältere arabische Quelle aus dem 10. Jahrhundert, der schiitische Imam Ali im Buch Pfad der Eloquenz/Nadsch al-Balagha, Spruch 146, formulierte dem gegenüber sehr viel irdischer: „Wissen ist ein Herrscher und Reichtum ist ihm unterstellt“.

Wissen als Basis von Macht und Reichtum also: allein das gibt schon Grund zur Reflexion. Was hat es mit diesem seltsamen Konstrukt „Wissen“ auf sich? Wie entsteht es, welche Rohstoffe und Transformationsprozesse sind beteiligt, wie kann es sich als Macht auswirken? In fast schon ironischer Replik auf diese Frage könnte die Antwort lauten: Es sind genau diese Fragen selbst, die am Anfang des Wissens stehen. Fragen, die das um uns herum Vorgegebene (lat: das „Datum“, pl. „Data“) hinterfragen, nach Sinn suchen, diesen bisweilen mühsam konstruieren oder auch erst de-konstruieren. Wie richtig wir mit diesen Sinnsuchen dann im Ergebnis liegen, ist eine andere Frage. Dass eine zähe Folgerichtigkeit im Denken, exemplifiziert an den Naturwissenschaften, zu großer Macht führen kann, beweist kaum etwas so deutlich wie die Herrschaft über die Atome. Die aus physikalischem Wissen erwachsene Fähigkeit, Atome zu teilen oder auch zu fusionieren, ist in ihrer Mächtigkeit mehr als deutlich geworden: durch Atom- und Wasserstoffbombe. Das individuelle und gesellschaftliche Erschrecken vor aus solchem Wissen geborener Super-Macht gibt dieses Beispiel bewusst mit auf den Weg.

Nicht alle aus Wissen geborene Macht ist derartig erschreckend. Gerade im Bereich der Medizin – Individualmedizin wie Bevölkerungsmedizin – kann diese Macht auch heilend wirken, ihren fein dosierten und sorgfältig indizierten Einsatz vorausgesetzt. Die Wissensgrenzen werden in diesen Feldern in stetigem Bemühen immer weiter verschoben. Diese wissensuchende Dynamik drängt in der Medizin, misst man sie an den Meilensteinen, welche durch die Nobelpreise markiert sind, vor allem in den Mikrokosmos. Sie führt in die Welt der Zellen und Moleküle, und damit auch in die Welt der wiederum Wissen kodierenden Bausteine des Lebens: Genomik, Proteomik, Metabolomik und in ihrer thera-

peutischen Umsetzung Pharmakogenomik, Nutrimomik u.a.m.. Es ist eine Ironie des Lebens, dass diese Auflösung der Person im Mikrokosmos dann am Ende wieder eine Brücke in die sog. personalisierte Medizin und darüber hinaus auch in die überindividuelle Bevölkerungsmedizin schlägt.

Das Zauberwort für diesen technischen Brückenschlag? Es heißt „Big Data“. Was darunter zu verstehen ist? Vonseiten der Medizininformatik werden gerne 5 V's zur näheren Beschreibung von Big Data angeführt. Gemeint sind Daten, welche durch großes Volumen (Volume), große Geschwindigkeit der Datenverarbeitung (Velocity), große Heterogenität (Variety), große Aussagekraft (Validity) und großem potentielltem Informationsgewinn (Value) gekennzeichnet sind. Big Data bezieht sich auf verschiedenste Arten von Daten, wie sie in den unterschiedlichen Wirtschaftssektoren generiert werden: Verkaufsdaten, Daten aus Routinestatistiken, geografische Daten, Sozialdaten, sog. *mobile data* aus Smartphone-Anwendungen wie z.B. Bewegungsprofile und vieles andere mehr. Woher diese Daten eingespeist werden? Im Gesundheitsbereich lassen sich 3 große Datenquellen unterscheiden: die *Routinestatistiken* aus den gesetzlichen Regelungen z.B. der Sozialgesetzbücher, die in großen Studien gewonnenen *Forschungsdaten* einschließlich der darin zunehmend aus der Genomik gewonnenen Daten sowie die Daten der *personalisierten Gesundheitsanwendungen* des Internets und anderer, innovativer (medizin-)technologischer Angebote. Laborgestützte (Omic-)Analysen werden inzwischen nicht nur in Forschungsstudien durchgeführt, sondern werden für wenig Geld auch aus dem Bereich der personalisierten Gesundheitsangebote generiert. Technische Umsetzungen wie Gesundheits-Apps, *Wearable Computing* mit Sensoren zur mobilen Diagnostik und private *Health Cloud*-Angebote treiben die generierten Datenmengen weiter an. „Big“ bedeutet dabei in der Praxis häufig eine heterogene, unstrukturierte Datenzusammenführung, welche über eigene Algorithmen erst nutzbar gemacht werden muss.

Der Deutsche Ethikrat hat sich die Vermessung des Menschen, Big Data und Gesundheit zum Thema gesetzt [1]. Interessierende Aspekte sind die Auswirkungen der Datenvernetzung auf die Arzt-Patienten-Beziehung, die Folgen einer immer engmaschigeren Datensammlung auf die Freiheit des Individuums, die damit verbundenen Verantwortlichkeiten von Privatpersonen, Forschern und Firmen und auch die möglichen Beiträgen des Gesetzgebers. Von einem digitalen Universum der Möglichkeiten ist die Rede, Dessen Bitgehalt inzwischen die Zahl der Sterne im physikalischen Universum erreicht hat und dessen Größe sich alle 2 Jahre verdoppelt. Dessen Inhalte auf Analysen warten, aller-

Bibliografie

DOI <http://dx.doi.org/10.1055/s-0035-1559680>
Gesundheitswesen 2015;
77: 531–532
© Georg Thieme Verlag KG
Stuttgart · New York
ISSN 0941-3790

Korrespondenzadresse

Prof. Dr. med. Manfred Wildner
Bayerisches Landesamt für
Gesundheit und
Lebensmittelsicherheit
Veterinärstraße 2
85764 Oberschleißheim
manfred.wildner@lgl.bayern.de

dings in der Hälfte der gespeicherten Daten ohne den gebotenen Datenschutz [2]. Reichen herkömmliche Herangehensweisen an den Schutz von personenbezogenen Daten und auch von Daten im Zuständigkeitsbereich des Wettbewerbsrechts überhaupt noch aus? Das Recht auf informierte Aufklärung ist angesichts vielfacher Sekundärnutzungen rekombinanter und interoperabler Datenkörper kaum noch aufrecht zu erhalten, eine verfügte Datenlöschung aus genau diesem Grund kaum mehr durchzuführen. Auch die herkömmlichen Verfahren der Anonymisierung und Pseudonymisierung sind angesichts genetischer Informationen und einer kleinteiligen Datafizierung unseres sozialen und ökonomischen Verhaltens wie auch vieler physiologischer Funktionen, die buchstäblich bis unter die Haut reicht, nur mehr bedingt tragfähig [3].

Das Projekt der Vermessung der Welt (Daniel Kehlmann) ist nach einer Phase der *Digitalisierung* vorhandener Daten offenbar in eine neue Runde der *Datafizierung* – nämlich von bisher nicht erfassten Lebensbereichen – eingestiegen. Und dies auch mit neuen Auswertungsmethoden: So verdrängen informatikbasierte Suchstrategien nach größtmöglicher Ähnlichkeit und Präzedenzfällen in umfassenden Datenkörpern die wahrscheinlichkeitsbasierte Statistik anhand von Stichproben, einfache und schnell zu rechnende Korrelationen bekommen mehr Bedeutung als integrative Kausalitätsüberlegungen, sog. *reality mining* im Hier und Jetzt tritt vor die nachträgliche und nachhaltige Analyse von in der Vergangenheit gesammelten Daten. Zu befürchten ist, dass bei einem Wegfall theoretischer und kritisch-reflexiver Kausalitätsbetrachtungen eine scheinbar datengestützte Realitätsillusion erzeugt wird, die gleichzeitig ein vertieftes Verständnis von Gesundheits- und Krankheitsprozessen verhindert.

Die im Big Data-Kontext geäußerte Vermutung, dass „Menschen die Summe ihrer sozialen Beziehungen, online-Aktivitäten und Verbindungsaufnahmen mit im Internet angebotenen Inhalten“ sind ([2], S. 157), ist ethisch ebenso problematisch wie eine zunehmende *Datafizierung des Leibes* (Peter Dabrock) und ein *Internet der Dinge* mit seiner fortschreitenden elektronischen Markierung unserer Umwelt bis hinein in die alltäglichen Gebrauchs- und Konsumgüter. Auch wissenschaftstheoretische Bedenken sind anzumelden, wenn schnelle Algorithmen euphorisierend zu der vor-schnellen Aussage verleiten, dass das „Ende der Theorie“ (Chris Anderson) gekommen sei [4]. Digitale Selbstvermessung (*quantified self*) und Lebensprotokollierung (*life logging*), die Ableitung von Normen aus Deskription, die damit möglicherweise verbundene Schaffung eines „digitalen Über-Ichs“ (Thomas Heinemann) lassen die Gleichsetzung von schrankenlosen Big Data mit den totalitären Zügen eines Orwellschen Big Brothers als möglich erscheinen. Schnell können aus Utopien auch Dystopien werden. Eine wachsame und durchaus auch kritische Diskussion zum Umgang mit Daten nicht nur als Rohstoff einer neuen, globalisierten Ökonomie, sondern auch als diskursive Anknüpfungspunkte für den Schutz und den Erhalt der menschlichen Würde und ihrer Freiräume in dieser neuen Welt ist dringend angezeigt. Nicht zuletzt auch, um eine ausgewogene, menschenwürdige und heilsame Nutzung im Bereich von Medizin und Gesundheitswissenschaften zu ermöglichen bzw. zu erhalten. Gerade im Bereich der Bevölkerungsmedizin und der Gesundheitssystemforschung sind Daten unverzichtbare Informationsträger. Diese können im Dreischritt ihrer vernünftigen *Erhebung*, einsichtigen *Interpretation* und *Verdichtung zur Information* und einem daraus entstehenden und weiter diskutierten *Wissen* zu einer menschlichen, freien und zukunftsfähigen Gesellschaft beitragen. Einem *Empowerment* in einem solchen Rahmen kann mit freiheitlicher Zielsetzung und entsprechenden Kautelen – Wissen ist und bleibt Macht – zugestimmt werden.

Die Beiträge in dieser Ausgabe bieten vielfältige Anknüpfungspunkte zu diesen Fragestellungen: Mit einer Reflexion zu einem ethischen Framework für gute Versorgung sowie einem Bericht zur Dänischen Debatte zu Priorisierung und Posteriorisierung, einer Erhebung des Commonwealth-Funds über die Nutzung von Informationstechnologie durch Primärärzte in 11 Ländern, mit einem IT-Anwendungsbeispiel in der spezialisierten ambulanten Palliativversorgung, einem Qualitätsvergleich von kostenfreien Geokodierungsdiensten, einer Untersuchung zur Einheitlichkeit der G-DRG-Kodierungsprüfung durch den MDK sowie über die Inanspruchnahme begleitender Hilfen durch hörgeschädigte Arbeitnehmer/innen. Vorgestellt wird auch das Ergebnis einer Delphi-Studie des DNVF zu Begutachtungs-, Förder- und Evaluationskriterien für Projekte aus dem Innovationsfonds, ein Praxisbericht der DFG Nachwuchsakademie Versorgungsforschung, eine Stellungnahme zum Selbstverständnis von Ärztinnen und Ärzten in der sozialmedizinischen Begutachtung und Beratung sowie der Vorschlag eines Curriculums Klinische Sozialmedizin. Aus aktuellem Anlass befassen sich 2 Diskussionsbeiträge kritisch mit der Stellungnahme „*Public Health in Deutschland*“ der wissenschaftlichen Akademien Leopoldina, Akatech und Union der Deutschen Akademien der Wissenschaften. Und last but not least freuen wir uns sehr, wieder die Abstracts des diesjährigen wissenschaftlichen Kongresses von DGSM, DGMS, MDK Bayern und LGL in Regensburg präsentieren zu können: „Daten gewinnen, Wissen nutzen für die Praxis von Prävention und Versorgung“.

Wissen ist Macht – gilt dies bereits für die Daten von Big Data an sich? Es sind ja im Grunde nicht die Daten allein, sondern vor allem die Algorithmen ihrer Auswertung, die Daten zu Information und dann auch zu dem Wissen werden lassen, welches Macht gibt. Der Begriff *Algorithmus* bezeichnet eine eindeutige Handlungsvorgabe zur Problemlösung. Diese Handlungsvorgabe kann z. B. als von Computern übersetzbare Programmschritte niedergelegt werden. Die Bezeichnung ist ein Anklang an den persischen Mathematiker Al-Chwarizmi, der im 9. Jahrhundert die regelbasierten Überlegungen zur mathematischen Problemlösung des im 7. Jahrhundert lebenden indischen Mathematikers Brahmagupta aufgriff und als Buch veröffentlichte. Vielleicht könnte dieses Entwickeln, Anwenden und Einhalten von Regeln, die so typisch für die Arbeitsweise der Algorithmen sind, in einem übergeordneten ethisch-sozialen und wissenschaftstheoretischen Kontext auch auf Big Data und die Verwendung der Algorithmen selbst ausgeweitet werden: Im Sinne einer verantwortlichen Datenhaltung, eines Respekts vor Privatsphäre, einer Offenlegung der Rechenschritte und einer persönlichen Verantwortungsübernahme der Akteure im Falle von Missbrauch [2]. Hoffentlich auch im Sinne der Entwicklung einer über die rechtlichen Regelungen als ethischem Minimum hinausreichenden bereichsspezifischen Ethik, welche das Kleine, Unvorhergesehene und in einem Big Data Universum Unauffällige schützt: Dieses könnte – wer weiß – schon wieder den Samen für eine noch einmal andere Zukunft für uns Menschen in sich tragen.

Literatur

- 1 Deutscher Ethikrat. „Die Vermessung des Menschen – Big Data und Gesundheit“, Jahrestagung Berlin, 21.5.2015
- 2 Vernon T, Gantz JF, Reinsel D et al. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. IDC White Paper, Framing 2014
- 3 Mayer-Schönberger V, Cukier K. Big Data – A Revolution That Will Transform How We Live, work and Think. London: John Murray Publishers; 2013
- 4 Anderson C. End of Theory: The Data Deluge Makes the Scientific Method Obsolete. Wired Magazine: 16.07 vom 23.06.2008