

Algorithmen und Big Data als Politikum

Klaus Mainzer

Algorithmen und Big Data sind längst nicht mehr nur das Thema von Nerds. Sie verändern Wissenschaft, Technik und Gesellschaft in dramatischer Weise. Algorithmen und Big Data werden zum Politikum! Wir brauchen mehr Governance der Algorithmenwelt, damit uns Big Data nicht aus dem Ruder läuft!

Die erste digitale Revolution war das Internet der Personen, bei dem ein weltweites Computernetz zur Kommunikation zwischen Personen benutzt wurde. Im „Internet der Dinge“ werden Gegenstände und Objekte mit Kommunikationsschnittstellen (RFID-Chips, Sensoren etc.) versehen, um miteinander zu kommunizieren. Was treibt diese Entwicklung an?

Exponentielle Wachstumsgesetze als Treiber der IT-Welt

Seit den 1960er Jahren gibt das Mooresche Gesetz die exponentielle Entwicklung der Rechenkapazität vor: Alle 18 Monate verdoppelt sich im Schnitt die Rechenkapazität bei gleichzeitiger Miniaturisierung und Verbilligung der Geräte. Wir sind auf dieser exponentiellen Kurve im Zeitalter der Petaflops (10^{15} Rechenschritte pro Sekunde) für Superrechner angelangt. Nach dem Mooreschen Gesetz wird diese Rechenleistung wegen der vorausgesagten Miniaturisierung und Verbilligung Ende der 2020er Jahre von kleinen Rechengenäten wie Laptops und Smartphones realisiert.

Hinzu kommt eine exponentiell sich entwickelnde Sensortechnologie, die immer billigere und kleinere Sensoren entwickelt – wie in lebenden Organismen, von Sensoren in organischer Größe (z.B. Kameras) über zelluläre Größen (wie z.B. Bakterien in der Natur) bis zu molekularer Größe im Nanobereich. Wir reden derzeit von einem Markt von 14 Milliarden weltweit vernetzter Geräte, davon alleine ein Drittel in den USA. 2020 werden 28 Milliarden vernetzter Geräte erwartet.

Die dadurch produzierten Datenmassen führen zu Big Data. Auch hier sind wir im Peta-Zeitalter angelangt. Datenkonzerne wie Google setzen heutzutage täglich 24 Petabytes um, d.h. 6000mal der Dateninhalt der US Library of Congress. Die Datenmassen sind amorph, nicht nur strukturierte Nachrichten wie E-Mails, sondern Sensordaten von GPS und MobilPhones.

Diese Datenmassen können von herkömmlichen (relationalen) Datenbanken nicht bewältigt werden. Dazu bedarf es neuartiger Algorithmen wie die Google-Suchmaschine. Vereinfacht gesagt teilen Big Data Algorithmen eine Datenmasse in Teilaufgaben auf, um sie parallel nach Datenkorrelationen und Datenmustern zu durchsuchen.

Das Ziel sind Prognosen von Trends und Profilen von z.B. zukünftigen Produkten oder Kunden. Neu bei Big Data ist: Prognosen werden nicht statistisch aufgrund von

repräsentativen Stichproben hochgerechnet, sondern alle Daten und Signale werden durchforstet, um Korrelationen und Muster zu erkennen. Anschaulich kann man sagen: Um die Nadel zu finden, benötigt man einen möglichst großen Heuhaufen, der total durchforstet wird.

Die Medizin ist ein anschauliches Beispiel, das zeigt, wie Big Data Wissenschaft unsere Lebenswelt verändert. Das medizinische Wissen wird unübersehbar: So gibt es heute bereits z.B. ca. 400.000 Fachartikel über Diabetes, die ein Arzt in einem Menschenleben nicht lesen kann. Dazu bedarf es intelligenter Suchmaschinen, um für den jeweiligen Patienten die passenden Schlüsselinformationen zu finden.

Das führt in Richtung einer personalisierten Medizin, da Krankheiten im komplexen Organismus der Menschen sehr unterschiedlich ablaufen können. Beispiel ist der Krebs: Da Tumore individuell unterschiedlich mutieren, müssten laufend Daten erhoben werden, um eine passende Therapie individuell ständig anpassen zu können.

Das nächste zentrale Beispiel ist die Wirtschaft: In der Wirtschaft können durch Big Data Mining blitzschnell Kunden- und Produktprofile vorausgesagt werden. Datenmassen, so sagt man, sind das Rohöl der Zukunft! Daten werden aber durch Gebrauch nicht aufgezehrt, sondern immer wieder in neuen Kontexten verwendbar: Beispiel sind Firmen, die Datenmassen zur Verkehrssteuerung verwenden, dann an andere verkaufen, die sie für Informationen über den Arbeitsmarkt (Arbeitslosigkeit) aufgrund von zeitlich unterschiedlichen Verkehrsbelastungen auswerten und schließlich wieder andere, die diese Daten mit ABS-Daten korrelieren, um aus dem Bremsverhalten die sichersten und kürzesten Straßenverbindungen zu ermitteln.

Intelligente Infrastrukturen breiten sich aus!

Mit dem „Internet der Dinge“ zeichnet sich die zweite digitale Revolution ab, die Deutschland und Europa nicht nutzlos verstreichen lassen sollten. Mittlerweile wachsen nämlich die IT-Netzwerke mit den physischen Infrastrukturen unserer Gesellschaft zusammen. Grundlage ist neben der Rechner- vor allem die Sensortechnologie. Milliarden von Sensoren machen die Kommunikation und Selbststeuerung der Dinge erst möglich. Wir sprechen dann von Cyberphysical Systems, soziotechnischen Systemen oder „intelligenten Infrastrukturen“.

Erstes Beispiel sind Automobile, die immer stärker mit Software-Modulen ausgestattet und dadurch autonomer werden – zunächst der Airbag mit seinen Sensoren, dann Fahrerassistenzsysteme z.B. für das Einparken. Heute sind Automobile „Computer auf Rädern“ – schließlich vollständig autonom, technisch bereits möglich, nur noch eine Frage rechtlich-gesellschaftlicher Rahmenbedingungen.

Weiteres Beispiel sind die Energiesysteme nach der Energiewende, die zu „Smart Grids“ werden: Wir konnten uns bereits vor Jahren im Internet informieren, wann wir unsere stromfressenden Geräte zu welcher Zeit und zu welchem Tarif am besten einschalten sollten. Aber das war und ist vielen viel zu umständlich und zeitaufwendig. Daher wird in Zukunft intelligente Software diese Aufgabe übernehmen. Zudem werden sich die Netze dezentral organisieren können. Aufgrund der verbesserten Isolation der Gebäude und eigener Stromerzeugung mit z.B. Fotovoltaik, Biogas und Brennstoffzellen wird nicht nur Energie für

den eigenen Stromverbrauch erzeugt, sondern Energie ans Netz gegeben. Die Daten dieser Gesamtenergie werden dann im Netz verwaltet.

Der nächste Schritt sind „Smart Cities“, in denen Energiesysteme nur noch eine Teildomäne neben z.B. Verkehrssystemen, Krankenversorgung, Verwaltung, Logistik sind. Diese verschiedenen Domänen müssen in intelligenter Software integriert werden. Eine gewaltige Herausforderung an das Software-Engineering! Das Software-Engineering befindet sich in einem Umbruch zum Systems-Engineering. Zukünftige Ingenieurinnen und Informatiker werden es nicht nur mit einzelnen technischen Geräten zu tun haben, sondern mit soziotechnischen Systemen mit integrierten Geräten, IT-Netzen und Menschen.

Warum benötigen wir „intelligente Infrastrukturen“?

.....

Die zunehmende Komplexität der Infrastrukturaufgaben ist heute so gewaltig, dass wir ohne intelligente Infrastrukturen nicht mehr auskommen.

.....

Wir brauchen ihre Selbstorganisation und Automatisierung, um die Aufgaben der Logistik, Versorgungssysteme, Gesundheitssysteme, Verkehrssysteme überhaupt im Griff behalten zu können. Andererseits wachsen auch die Risiken empfindlicher Störungen solcher hochkomplexer Systeme – lokale Störungen, die sich kaskadenhaft aufschaukeln können („Schmetterlingseffekt“).

Auf dem Hintergrund der Cyberphysical Systems bzw. soziotechnischen Systeme verändert sich die Wirtschafts- und Industrielwelt grundlegend. Die Rede ist von Industrie 4.0:

Industrie 1.0 war die Dampfmaschine, Industrie 2.0 Henry Fords Fließband, Industrie 3.0 fixierte Industrieroboter am Fließband (z.B. Autoindustrie), Schließlich Industrie 4.0 das „Internet der Dinge“ in der Industrielwelt. Internet der Dinge meint, dass Gebrauchsgegenstände mit Sensoren, RFID-Chips und Softwarefunktionen ausgestattet sind und sich selber wahrnehmen und miteinander kommunizieren können. In der Arbeitswelt kommuniziert also das Werkstück mit Design, Werkbank, Transport, Logistik, Vertrieb und Versand, um die eigene Produktion zu organisieren.

Damit wird eine neue Kundenorientierte Produktion möglich: On Demand oder Tailored („maßgeschneiderte“) Production. Früher konnten sich nur wenige Reiche maßgeschneiderte Anzüge, individuell für den eigenen Bedarf, leisten. In Industrie 4.0 wird on-demand produziert, nach dem individuellen und personalisierten Kundendesign. Der individuelle und personalisierte Produktionsprozess kann sich selber organisieren. Auch in den Energiesystemen beobachten wir den Trend zu dezentraler und individueller Versorgung. Auf der ganzen Linie erleben wir also eine Abkehr von der Massen- und Standardproduktion à la Henry Ford – von der Industrie über die Ernährung bis zur personalisierten Medizin.

Im industriellen Internet („Industrie 4.0“) läge nun ein zentraler Standortvorteil Deutschlands und Europas. Deutschland ist ein klassisches Industrieland seit dem 19. Jahrhundert. Mit mehr als 22% hängt die deutsche Wirtschaft von der Industrie ab – ganz im

Gegensatz zum Mutterland der Industrialisierung Großbritannien, das heute maßgeblich von seinen Finanzgeschäften abhängt.

.....

Wenn allerdings Deutschland die Effizienzsteigerung mit IT nicht gelingt, verpufft der industrielle Standortvorteil und wir sind am Ende nur noch Zulieferungsindustrie für die IT-Marktführer in USA.

.....

Wie verändert sich die Arbeitswelt?

Die IT-Firmen brechen überall in andere Domänen ein. Umgekehrt werden sich aber auch IT-Firmen umstellen müssen. Ein IT-Gigant wie Microsoft hat bislang Software im Stil von Industrie 2.0 als Massen- und Standardfabrikation produziert. In der Welt von Industrie 4.0 werden Software-Häuser auf individuelle Wünsche und den Bedarf einzelner Firmenkunden zugehen müssen. Konzerne werden nicht länger Massenstandards vorgeben können, sondern sich in Consultingfirmen verwandeln, die individuelle Tools und passgenaue IT-Infrastrukturen mit dem Kunden entwickeln müssen. Das gilt auch für Energiekonzerne: Firmen wie Eon richten sich auf einen dezentralen Markt ein und setzen zunehmend auf individuelle Beratung, um die passgenaue Lösung zu finden. Das führt zu neuen Geschäftsmodellen wie „Buy and Build“. Nur durch Individualisierung und Personalisierung kann Vertrauen aufgebaut werden.

Vertrauen ist entscheidend für den Erfolg des industriellen Internets. Im Mittelstand als Rückgrad von Industriekultur meldet sich nämlich auch Skepsis. Der Grund ist die Cloud-Technologie: Wenn ein Mittelständler gutes Geld mit seinem Geschäftsmodell verdient, dann wird er sich hüten, die entsprechenden Daten in die Cloud zu stellen – vor allem nach der NSA-Debatte und Industriespionage. Die bisherige Sicherheitstechnik ist die Achillesferse von Industrie 4.0. Daher werden auch in diesem Fall individuelle Lösungen gefunden werden müssen. Man wird genau überlegen müssen, welche Daten in die Cloud gesetzt werden sollen, um einen effektiven Zugriff auf Daten durch Mitarbeiter und Kunden zu ermöglichen. Besonders sensible Firmendaten gehören eben nicht in die Cloud. Zudem sind auch 3.1 oder 3.3 eine individuell gute Lösungen, die vom jeweiligen Firmenprofil abhängen.

Aus Arbeitnehmersicht stellt sich ebenfalls die Frage der Datensicherheit: Die Automatisierung ist nur möglich, weil viele Sensoren, Kameras, Lichtschranken etc. dauernd massenhaft Daten aufnehmen. Wer hat Zugriff auf diese Daten, wo werden sie wie lange für wen gespeichert?

Schließlich geht es um den Arbeitsmarkt selber. Die Entwicklung wird ohne Zweifel die Effektivität der Industrie stärken. Wird die Automatisierung der Industrie aber nicht zu Arbeitslosigkeit führen? Deutschland mit bereits hoher Automatisierung hat deutlich geringere Arbeitslosigkeit als andere europäische Länder. Arbeitslosigkeit hat in diesen Ländern andere Gründe und hängt z.B. mit versäumten Reformen des Arbeitsmarkts zusammen. Ich teile auch nicht die Horrorvorstellung, dass wir am Ende nur noch hochqualifizierte Ingenieure mit Universitätsabschlüssen brauchen und den Rest machen Maschinen. Wir werden das Know-how der Menschen weiterhin auf allen Gebieten

benötigen. Allerdings werden sich die Anforderungen ändern. Das haben auch die Gewerkschaften verstanden.

.....

Die Innovationszyklen sind schon jetzt in vielen Bereichen schneller als unsere Ausbildungszyklen. Wir müssen uns also künftig genau überlegen, wozu wir die Menschen eigentlich ausbilden.

.....

Wenn wir jemandem heute in der Lehre ein bestimmtes Computerprogramm beibringen, ist das schon überholt, wenn er in den Betrieb kommt. Deswegen müssen wir die Fähigkeit des Menschen ausbilden, sich in neue Arbeitsprozesse einzuarbeiten und sich auf neue Situationen einzustellen (Motto: „Lebenslanges Lernen“). Ich denke, es wird in Zukunft absolut zur Normalität gehören, dass ein Teil der Mitarbeiter immer in Lehrgängen und Fortbildungen sein wird, um sich auf neue Abläufe vorzubereiten.

Fazit: Daten ohne Theorie und Governance sind blind!

Zusammengefasst: Wir müssen also wachsam sein, um die Entwicklung nicht zu verpassen! Am Ende dürfen wir uns aber von Big Data nicht blenden lassen. Datenmuster und Korrelationen helfen wenig, wenn wir die zugrunde liegenden Ursachen, Theorien und Modelle nicht kennen. Einige Propheten der Big Data-Welt sprechen ja bereits von „Big Data – the End of Theory“ (Chris Anderson): So wie in der Wirtschaft derjenige am Markt gewinnt, der die schnellsten Algorithmen für Produkt- und Kundenprofile hat, so auch in der Wissenschaft: Die Algorithmen entdecken neue Zusammenhänge, bevor Beweise und Begründungen nachgeliefert werden.

Diese Tendenz kann äußerst gefährlich werden:

- In der Medizin helfen Medikamente wenig, wenn wir nur auf kurzfristige Dateneffekte setzen, ohne die Ursachen (z.B. Krebs) verstanden zu haben.
- In der Wirtschaft laufen wir in die Irre, wenn wir uns nur auf unverstandene Formeln und Eckdaten verlassen: Das hat die Wirtschaftskrise von 2008 gezeigt.
- Hochfrequenzhandel ist zwar schnell, kann aber unterhalb der Wahrnehmungsfähigkeit von uns Menschen Krisen auslösen, weil die Koordination und Governance fehlt.
- Die Vorausberechnung von Kriminalität, Terror und Kriegseinsätzen hilft wenig, wenn wir die zugrunde liegenden sozialen Ursachen und Wirkungszusammenhänge nicht begreifen.

In der Wissenschaft gilt: Theorien ohne Daten sind zwar leer, aber Daten ohne Theorien sind blind! Daraus leite ich für die Gesellschaft ab:

.....

Governance ohne Daten ist leer, aber Daten ohne Governance sind blind!

.....

Am Ende zielt dieser Beitrag auf eine Stärkung unserer Urteilskraft, d.h. die Fähigkeit, Zusammenhänge zu erkennen, aber auch Maß zu behalten, damit eine immer komplexer werdende und von Automatisierung beherrschte Welt uns nicht aus dem Ruder läuft.

LITERATUR

Mainzer, Klaus 2014: Die Berechnung der Welt. Von der Weltformel zu Big Data, C.H. Beck: München

.....

Mainzer, Klaus 2014: Die Berechnung der Welt. Können Big Data-Ergebnisse Theorie und Beweis ersetzen? In: Glanzlichter der Wissenschaft (Hrsg. Deutscher Hochschulverband), Lucius Verlag, S. 117-121

.....

Mainzer, Klaus / Chua, Leon 2013: Local Activity Principle. The Cause of Complexity, Imperial College Press: London

.....

"Algorithms are powerful and useful. But by themselves they are blind." Interview mit Klaus Mainzer, in: Faszination Forschung (das Wissenschaftsmagazin der Technischen Universität München) 15/14, S. 64-75

Prof. Dr. Klaus Mainzer hat den Lehrstuhl für Philosophie und Wissenschaftstheorie und war bis 2014 Gründungsdirektor des Munich Center for Technology in Society (MCTS) an der Technischen Universität München.

