Ulrike Höppner

Drei Irrtümer über Anonymität im Netz – und was sie mit Privatsphäre zu tun haben

Im Internet ist nichts so, wie es uns erscheint. Oder doch zumindest ist einiges ganz anders. Man könnte auch sagen, ums Internet ranken sich besonders viele Mythen, die falsch sind – und doch wichtige Hinweise darauf geben, was Vernetzung mit uns macht und wie wir damit umgehen können. In meiner Forschung zu Privatheit, Internet und Governance sind mir drei häufige Mythen immer wieder begegnet, die eigentlich Irrtümer und gerade deswegen besonders erhellend sind. Darum greife ich sie hier auf und zeige, was man daraus für den Umgang mit Vernetzung lernen kann.

1 Erster Irrtum: Im Netz ist jeder anonym

In gewisser Weise ist dies der fundamentalste Irrtum, denn das Gegenteil ist der Fall: Im Netz ist niemand anonym. Um zu verstehen, warum das so ist, ist es nötig, sich mit der Technik zu befassen, auf der das Internet beruht. Technisch funktionieren alle Anwendungen des Internets nur, weil jedes Gerät eindeutig identifizierbar ist. Sonst wüsste kein Rechner, wohin er eine angeforderte Information übermitteln soll.[1] Erreicht wird das über das IP (Internet Protocol), durch das jedem Rechner eine eindeutige Nummer zugewiesen wird. Durch diese Nummer wissen Provider beispielsweise, für welchen Kunden ein Datenpaket gedacht ist. Es ist auch diese Nummer, die bei der Vorratsdatenspeicherung genutzt wird, um zu rekonstruieren, wer wann von wem Daten bekommen hat. Niemand, der am Netz ist, kann sie umgehen. Zuständig für die Entwicklung und Verwaltung dieses technischen Rückgrats des Netzes ist die „Internet Engineering Task Force" (IETF), ein Expertengremium, das dafür zuständig ist, Standards zu entwickeln, die einen reibungslosen technischen Datenaustausch zwischen allen Servern im Netz ermöglichen. (Allein dieses Gremium mit seinen informellen Entscheidungsmechanismen, seiner spannenden Genese und, für eine so wichtige Institution, ungewöhnlichen Zusammensetzung wäre einen eigenen Artikel wert.)[2]

---

1 Tatsächlich ist es in der Informatik eine unbeantwortete Frage, ob es überhaupt technisch möglich ist, ein Netzwerk zu realisieren, in dem alle Teilnehmer anonym sind. Was es allerdings gibt, sind technische Wege, die Anonymität deutlich zu erhöhen, beispielsweise das Tor-Netzwerk (https://www.eff.org/whatistor, Abruf der in diesem Beitrag angegebenen Internetseiten: 7.11.2019).

2 Genaueres findet man unter https://www.ietf.org.

Die einem Rechner zugeteilte Nummer wird nach einem mathematischen Prinzip generiert (welches, ist hier nicht wichtig), das eine bestimmte Anzahl an Adressen zur Verfügung stellen kann. Bis 2011 basierte das gesamte Netz auf IPv4 (Internet Protocol Version 4), einer in den 1980er Jahren entwickelten Version. Damals gab es eine überschaubare Anzahl an verbundenen Rechnern, doch mit der Kommerzialisierung des Netzes stieg die Zahl der verbundenen Rechner stark an. 2011 wurden die letzten Adressen dieses Protokolls verteilt. Das heißt, diese sind im Wesentlichen aufgebraucht, neue Geräte könnten nicht mehr ins Netz. Nun ist es zurzeit so, dass noch nicht alle Geräte immer am Netz sind und eine Strategie, mit den knappen Adressen umzugehen, ist die, dass Rechner nur dann eine Adresse erhalten, wenn sie ins Netz gehen, und zwar jedes Mal eine andere. Das heißt, die IP-Adresse ändert sich für das Gerät, wenn es sich neu einwählt, beispielsweise nach einem Neustart, Verbindungsabbruch oder Stromausfall. Wer wann welche IP hatte oder hat, wird beim Provider gespeichert, aber nicht jeder im Netz hat einfach Zugriff darauf. Tatsächlich kann es ausgesprochen schwer sein herauszufinden, welcher Rechner hinter einer bestimmten IP-Adresse steht, und noch schwerer, welche Person zu dem Gerät gehört (auf Ausnahmen komme ich später). Und das ist nun der Grund, warum auch Informatiker von Anonymität im Netz sprechen, denn Anonymität ist informatisch relativ definiert – je mehr Aufwand ein potenzieller Angreifer betreiben muss, um zu ermitteln, wer man ist, desto anonymer ist man. Das wird mit einem Faktor ermittelt, und ab einem bestimmten Wert wird von faktischer Anonymität ausgegangen, auch wenn die technische Erkennbarkeit gegeben ist.[3]
Aber was kümmert es mich, ob theoretisch ein technisch sehr versierter Angreifer ermitteln könnte, wer ich bin? Ist also alles nur eine hypothetische Möglichkeit? Nicht ganz. Zunächst handelt es sich hier um „persönliche Daten", deren Anfallen wir nicht vermeiden können: Um die Kommunikation überhaupt zu ermöglichen, ist es zwingend erforderlich zu wissen, wer wann mit wem kommuniziert. Diese Tatsache spielt in Debatten beispielsweise um die Vorratsdatenspeicherung eine erhebliche Rolle. Denn da geht es um die Frage, ob Provider diese Daten sofort löschen oder über einen längeren Zeitraum aufbewahren und gegebenenfalls herausgeben sollen. Nur weil wir eben nicht anonym sind, kann so etwas überhaupt erwogen werden. Darüber hinaus wird sich künftig auch technisch einiges ändern. Um dem Adressmangel zu begegnen, hat die IETF nämlich bereits vor vielen Jahren mit der Entwicklung (und zwischenzeitlich Einführung) von IPv6 begonnen, das ein Vielfaches der Adressen von IPv4 zur Verfügung stellt und es theoretisch ermöglicht, jedem Gerät auf absehbare Zeit eine permanente IP-Adresse zur Verfügung zu stellen. Damit wäre jeder Rechner immer unter derselben Nummer im Netz. Die Wiedererkennbarkeit der Rechner hätte durchaus Vorteile. Wenn alle Seiten und Anbieter durch die IP sicher wüssten, mit welchem Gerät sie kommunizieren,

[3] Vgl. Latanya Sweeney: k-anonymity. A model for protecting privacy, in: International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10/5 (2002), 557 – 570.

könnte man beispielsweise Authentifizierungsverfahren vereinfachen. Allerdings wird es dann auch schwerer, zumindest im Ansatz zu kontrollieren, wer welche Daten über mein Nutzungsverhalten im Internet sammelt.[4]
Die grundlegendste Erkenntnis, die ich aus der Beschäftigung mit dem Irrtum, im Netz sei jeder anonym, gewonnen habe, ist die, dass man, um die Wirkungen des Netzes auf die Privatsphäre wirklich abschätzen zu können, die technischen Grundlagen, auf denen die Kommunikation beruht, zumindest ansatzweise verstehen muss. Und zwar nicht nur als technischer Experte, sondern auch als Nutzer und erst recht als Beteiligter an den politischen Diskussionen darüber, wann, wie und von wem digitale Daten gesammelt, verwendet, gespeichert und ausgewertet werden dürfen. Wer nicht weiß, was möglich ist – also zum Beispiel denkt, er wäre im Netz anonym –, kann auch nicht informiert entscheiden, wie die eigene Privatsphäre zu schützen ist.

2 Zweiter Irrtum: Anonymität ist die Hauptursache für die Verbreitung von Hass im Netz

Wer die Kommentarspalten im Internet liest (es ist fast schon egal wo), muss einiges lesen, was diese Schlussfolgerung nahelegt. Da finden sich Hasstiraden, Unwahrheiten und allzu oft auch einfach „Argumente", die keine sind, sondern nur Meinungen. Da wird oft nicht differenziert diskutiert, sondern beleidigt, abgewertet und nicht aufeinander gehört, wie wir es im „echten" Leben nicht für möglich halten (wollen). Und es ist fast schon zum Allgemeinplatz geworden, dass wegen der (relativen) Anonymität Hemmungen abnehmen – wer seinen Namen nicht damit verbinden muss, sagt Dinge, die er sonst nicht sagen würde. So richtig dieser Eindruck ist, so wichtig ist es, sich der Ambivalenz dieses Satzes bewusst zu sein.
Zunächst die Schattenseite: Anonymität ermöglicht es Menschen, Hass auf eine Weise zu äußern, wie sie es aus berechtigter Furcht vor den sozialen Folgen sonst vielleicht nicht tun würden. Unter dem Mantel der Anonymität leugnet sich leicht der Holocaust, wird schneller zu Gewalt aufgerufen und auf Stereotype zurückgegriffen. Die Spontaneität und Unmittelbarkeit der Kommunikation im Netz verstärkt diesen Effekt noch – ein Kommentar ist innerhalb weniger Sekunden geschrieben oder ein Hasspost geliked,[5] ein Leserbrief ist im Vergleich viel aufwändiger. Liegt das alles nur an der Anonymität, derer sich die Hetzer bedienen?
Zwei Beobachtungen lassen sich dazu machen. Auch wenn es so scheint, als gäbe es unglaublich viele Menschen, die aktiv Hass verbreiten, ist dem nicht so, wie sich bei

[4] Bisher geschieht dies zu einem großen Teil durch Cookies, kleine, auf meinem Rechner abgelegte Dateien, die ich zum Beispiel löschen kann.
[5] Vgl. Alexander Brown: What is so special about online (as compared to offline) hate speech?, in: Ethnicities 18/3 (2018), 297 – 326, https://ueaeprints.uea.ac.uk/64133/1/Accepted_manuscript.pdf.

näherer Betrachtung herausstellt. Tatsächlich wird die überwiegende Zahl der Hasskommentare von einer verhältnismäßig kleinen Anzahl sehr aktiver Nutzer geschrieben.[6] Anonymität vereinfacht diesen Hetzern möglicherweise ihre Kampagnen, aber sie würden wohl nicht aufhören, wenn sie nicht anonym wären. Gegen den harten Kern der Hetzer helfen eher Strategien wie die des Vereins „ichbinhier e. V.", der darauf setzt, Hass nicht unkommentiert stehen zu lassen. Die Aktivisten kommentieren selbst aktiv und machen den Lesern der Kommentare deutlich, dass die Hetzer eine Minderheit sind. Die andere Beobachtung ist, dass zunehmend versucht wird, auch rechtlich gegen Hass im Netz vorzugehen. Viele Länder, auch Deutschland, haben in den vergangenen Jahren spezielle Gesetze gegen Hass im Netz verabschiedet, die die Verfolgung vereinfachen sollen.[7] Und wie wir oben gesehen haben, ist eine Identifizierung der Autoren technisch durchaus möglich, auch wenn sie sich anonym wähnen. Allerdings ist diese Strategie, wenn auch vielleicht wirkungsvoller als argumentatives Dagegenhalten, auch ein Einfallstor für Zensur – denn wer entscheidet letztlich, was Hass und was noch legitime Kritik ist? So eindeutig, wie es uns manchmal erscheint, ist das keineswegs. Unmittelbar einsichtig ist das, wenn wir statt auf rechtliche Regulierung auf die technische blicken. Wenn Algorithmen automatisiert die Unmengen an Kommentaren, Bildern und Filmen, die jede Minute in Soziale Netzwerke hochgeladen werden, sichten und Hass herausfiltern, kommt es regelmäßig vor, dass auch legitime Kritik unterbunden wird, weil alles, was potenziell anstößig sein könnte, herausgefiltert wird.[8] In aufgeheizten Debatten zwischen legitimer Meinungsäußerung und Hetze zu unterscheiden, ist weder technisch noch rechtlich einfach. Langer Rede kurzer Sinn: Weder ist Anonymität die Hauptursache für die Verbreitung von Hass im Netz, noch ist ihre Aufhebung die Lösung.

Oft vergessen wird die Sonnenseite der Anonymität: Sie ermöglicht es Menschen, Dinge zu sagen, die sie sonst nicht sagen könnten. Da sind diejenigen, die in der Anonymität des Netzes endlich ihre Regierungen kritisieren können. Da sind die Whistleblower, die anonym Missstände in Institutionen aufdecken. Da sind auch die, die aus dem einen oder anderen Grund nicht alles über sich offenlegen wollen, weil sie eine andere sexuelle Orientierung haben, ungewöhnliche Hobbys oder auch chronische Krankheiten und sich in aller Anonymität mit Gleichgesinnten austauschen wollen. Oder Eltern, die es ihren Kindern ermöglichen wollen, ohne Folgen für ihre Zukunft das Internet zu erkunden. Sie alle haben ein legitimes Interesse an Anonymität. Und

---

6 Vgl. Philip Kreißel / Julia Ebner / Alexander Urban / Jakob Guhl: Hass auf Knopfdruck. Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz, 2018, https://www.isdglobal.org/wp-content/uploads/2018/07/ISD_Ich_Bin_Hier_2.pdf.

7 Vgl. Iginio Gagliardone / Danit Gal / Thiago Alves / Gabriela Martinez: Countering online hate speech, 2015, https://unesdoc.unesco.org/ark:/48223/pf0000233231.

8 Vgl. Pew Research Center: The future of free speech, trolls, anonymity, and fake news online, 2017, https://www.pewinternet.org/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online.

sie tragen dazu bei, dass sich im Netz auch fairer Austausch, intensive Diskussionen, legitime Kritik von Stimmen, die sonst nicht gehört würden, und gegenseitiger Beistand verbreiten können. Nicht umsonst kämpfen Institutionen wie die Electronic Frontier Foundation für den Erhalt von Anonymität im Netz, denn diese hilft auch gerade den Schwachen.[9]

So zeigt sich: Anonymität ist ambivalent und eben nicht eindeutig gut oder schlecht. Insofern verweist auch dieser Irrtum über Anonymität im Netz auf ein grundsätzlicheres Problem: Bei den allermeisten Fragen rund um die Privatsphäre (und auch anderes) im Netz handelt es sich um Fragen, bei denen Vor- und Nachteile vorsichtig gegeneinander abgewogen werden müssen, aber eindeutige und für alle optimale Lösungen gibt es nicht. Natürlich wollen wir, dass Hass verfolgt werden kann. Wir wollen aber auch, dass sich Menschen ohne Angst vor Verfolgung äußern können. In den politischen Debatten geht im Eifer des Gefechts – je nachdem, um welchen Skandal es geht – entweder das eine oder das andere gern unter. Zu oft scheuen wir uns, die schwierigen Wertediskussionen zu führen und darüber zu streiten, was mehr zählen sollte. Aber wir sollten diese wichtigen Entscheidungen nicht den Technikern der großen Konzerne oder auch nur den gewählten Regierungen überlassen. Wir müssen uns immer wieder selbst beteiligen, denn täglich bringen neue Medien, Technologien und soziale Praktiken neue Herausforderungen hervor, die sich oft einer einfachen „One size fits all"-Lösung entziehen.

3 Dritter Irrtum: „Ich bin ja nicht bei Facebook" – und wahre meine Privatsphäre

Unter denjenigen, denen ihre Privatsphäre wichtig ist, gibt es einige, die versuchen, sich durch Rückzug zu schützen, und wenn die Diskussion auf private Daten im Netz kommt, sagen sie überzeugt: „Ich bin ja nicht bei Facebook" (Facebook kann beliebig auch durch Whatsapp, Paypal, Ebay, Google ersetzt werden). Oberflächlich betrachtet sollte es bei der Wahrung der Privatsphäre helfen, wenn man keine Daten preisgibt. Aber bei genauerer Betrachtung erweist sich das als Illusion, wie sich leicht an drei Beispielen illustrieren lässt.

Das erste Beispiel ist der Ebayer, der die Gebrauchsanweisung der zu kaufenden Kamera nicht googeln will, weil Google „böse" ist, aber anschließend über Paypal zahlt. Beide Dienste sind nicht das Gleiche, aber die Daten bezüglich Gewohnheiten, Vermögen und Vorlieben, die bei Paypal „anfallen", sind sicherlich genauso wertvoll wie die Kenntnis meines Surfverhaltens. Das soll heißen: Wer einen oder auch zwei Dienste vermeidet, mag dort keine Daten hinterlassen, überhaupt keinen „Datensammler" zu nutzen, ist aber kaum möglich, wenn jemand im Internet irgendetwas machen will.

[9] Vgl. Electronic Frontier Foundation: Anonymity, https://www.eff.org/issues/anonymity.

Ein zweites Beispiel ist besonders ausgeprägt bei Diensten wie Whatsapp. Nicht bei Whatsapp zu sein, scheint auf den ersten Blick eine gute Strategie, Datensammler zu vermeiden. Und doch nützt es wenig, wenn alle um einen herum ihr Adressbuch für Whatsapp freigegeben haben[10] – denn dann kennt der Dienst meine Nummer und kann mit hoher Wahrscheinlichkeit ein gutes Abbild meines sozialen Netzwerks errechnen, ohne dass ich selbst irgendetwas preisgebe. Eingeschränkt gilt das auch für andere Soziale Netzwerke – aus den Informationen der anderen lässt sich viel über mich erschließen, ohne dass ich den geringsten Einfluss darauf habe.

Ein letztes Beispiel sind die Spuren, die wir, teilweise sogar recht anonym, notwendigerweise im Netz hinterlassen. Wenn ich anonymisiert auf meinem Rechner bei Google nach Kinderschuhen suche, wie kann es dann sein, dass mir später im Café auf meinem Handy bei Facebook Werbung für ebensolche Schuhe angezeigt wird? Kurz gesagt: Facebook rät. Mein Rechner und mein Handy sind oft zusammen in einem Netz, und der Werbealgorithmus „rät" einfach, dass beide wohl zum selben Nutzer gehören. Und der interessiert sich gerade für Kinderschuhe einer bestimmten Marke. Sicher liegt Facebook dabei auch manchmal falsch, aber selbst wenn der Algorithmus nur in sechs von zehn Fällen mit seiner Annahme richtig liegt, ist das immer noch besser, als zufällig Werbung anzuzeigen. Datensparsamkeit hilft natürlich, diese Effekte zu reduzieren, aber ganz ausschalten kann allein die Sorgfalt des Nutzers sie nicht.

Die Beispiele haben gemeinsam, dass es hier um Daten geht, bei denen uns allen prinzipiell klar ist, dass es persönliche Daten sind. Es sind Informationen über uns, von denen wir zumindest im Ansatz ermessen können, was sich daraus ablesen lässt. Wir können darüber nachdenken, ob wir sie weitergeben, wem und wozu. Und auch wenn es keinen perfekten Schutz gibt, können wir durch Vorsicht einiges erreichen, um uns nicht erkennbar oder zumindest weniger durchschaubar zu machen. Alle diese Strategien stoßen jedoch auch an technologische Grenzen. Dazu muss man sich zwei Dinge klarmachen.

Erstens reicht es bei unternehmerischen Entscheidungen in vielen Fällen, wenn man sich auf Wahrscheinlichkeiten verlässt. Man benötigt gar keine absolute Sicherheit darüber, ob ein bestimmter Fakt über eine konkrete Person stimmt oder nicht. Bei Werbung ist das, in eher harmlosem Zusammenhang, schnell deutlich: Wenn ich jemandem, der sich wahrscheinlich für Kinderschuhe interessiert, Werbung für ebensolche zeige, ist das erheblich besser, als wenn ich einfach allen diese Werbung zeige. Diese Art der Personalisierung ist gang und gäbe und hat sogar für die Nutzerin bzw. den Nutzer Vorteile: Man bekommt eher Dinge angezeigt, die einen auch tatsächlich interessieren.

[10] Rechtlich betrachtet, bestätigt jeder Nutzer beim Hochladen seines Adressbuchs, dass er das schriftliche Einverständnis aller im Adressbuch verzeichneter Personen für das Hochladen hat. Ohne Freigabe des Adressbuches lässt sich der Dienst nur mit gravierenden Einschränkungen nutzen. Vgl. https://www.spiegel.de/netzwelt/netzpolitik/whatsapp-upload-von-kontaktdaten-ist-illegal-a-1154667.html.

Moderne Datenauswertung ermöglicht es aber auch, anhand weniger Daten mit einer gewissen Wahrscheinlichkeit vorauszusagen, ob jemand einen Kredit wirklich zurückzahlt, eine chronische Krankheit hat oder bekommen wird oder einen angebotenen Job ausfüllen kann. Da kann es passieren, dass jemand wegen seiner Postleitzahl keinen Kredit bekommt, wegen seiner kranken Tante keine Lebensversicherung oder die Anzeige für die Führungsstelle gar nicht erst zu Gesicht bekommt. Entschieden wird dabei nicht, ob das Individuum tatsächlich zahlungsunfähig, krank oder faul ist, sondern nur, dass dies mit einer hohen Wahrscheinlichkeit der Fall ist. Unternehmen minimieren also ihr Risiko, auch wenn sie dadurch einige Personen von Dienstleistungen oder Gelegenheiten ausschließen, obwohl sie an ihnen durchaus hätten verdienen können. Aufs Ganze gerechnet, lohnt sich diese Strategie für die Unternehmen, aber für den Einzelnen kann sie einen Verlust an Möglichkeiten bedeuten, nicht auf Basis über ihn gesammelter Daten, sondern auf Basis aggregierter Daten, die der Betroffene weder einsehen noch beeinflussen kann.

Zweitens gibt es eine ganze Industrie von Datenverarbeitern, deren Strategie darin besteht, unterschiedliche Datensätze, die einzeln nicht besonders aussagekräftig sind, so zusammenzuführen, dass sie einerseits sinnvolle Aussagen über mögliche Interessen, Verhaltensweisen und Einstellungen bestimmter Gruppen ermöglichen und andererseits Individuen identifizierbar machen, die zu diesen Gruppen gehören. Das bekannteste, wenn auch bei weitem nicht das einzige Beispiel dafür ist Cambridge Analytica. Die Praktiken dieser Firma gerieten in Verruf, nachdem gezielte Werbekampagnen, die auf einer derartigen Datenanalyse beruhten, den Ausgang der US-Wahl 2016 entscheidend beeinflusst haben. Dabei wurde nichts Illegales getan und auch nicht gegen geltende Datenschutzgesetze verstoßen. Cambridge Analytica nutzte unterschiedliche Datensätze, um möglichst genau zu ermitteln, wer wahrscheinlich bei der Wahl noch zur Stimmabgabe für Donald Trump motiviert werden könnte, und zwar nur in den „swing states", dort also, wo wenige Stimmen über den Wahlausgang entscheiden konnten. Genutzt wurden Daten zum sozioökonomischen Status, Bildungsgrad, Wohnort, Alter und zu Interessen – Daten, die Facebook in „anonymer" Form, also ohne Namen, seinen Werbekunden zur Verfügung stellt, eben damit diese sie nutzen können, um Gruppen gezielt für Werbung auszusuchen. Kombiniert mit anderen – öffentlich zugänglichen – Daten, ergibt sich ein erschreckend akkurates Vorhersagepotenzial. Auf Basis der Analyse von Cambridge Analytica wurde dafür gesorgt, dass nur speziell ausgewählte Nutzer negative Informationen zu Hillary Clinton und positive zu Donald Trump eher angezeigt bekamen als andere Nutzer. Mithilfe von Anzeigen wurde diesen Nutzern eine andere Wirklichkeit vorgegaukelt, sodass ihnen die Stimmabgabe für Trump plausibler erscheinen musste, als wenn sie objektive Informationen erhalten hätten. Auch wenn diese Nutzer also versuchten, sich ein objektives Bild zu machen, konnten sie das nicht ohne

Weiteres im Internet finden – dazu hätten sie um die Möglichkeiten der Manipulation wissen und sie gezielt umgehen müssen.[11]
Analysen großer anonymisierter Datenmengen können also genutzt werden, um mit sehr hoher Treffsicherheit nicht nur derzeitige Gewohnheiten, Vorlieben und Eigenschaften konkreter Einzelner zu bestimmen (über die man dann nur noch wenig wissen muss). Sie ermöglichen es auch, erstaunlich korrekte Vorhersagen über das Verhalten Einzelner zu treffen. Zum Beispiel lässt sich statistisch gut vorhersagen, mit welcher Wahrscheinlichkeit ein Arbeitnehmer einem Unternehmen langfristig erhalten bleibt. Geschaut wird dabei darauf, mit wie vielen Menschen im Unternehmen er kommuniziert, wie viele Angebote er nutzt (z. B. Betriebssport) und ob er mit Kollegen oder eher alleine zu Mittag isst. Im Idealfall nutzen Unternehmen diese Informationen, um Mitarbeiter zu halten, besser zu motivieren und gerechter zu entlohnen. Denkbar ist auch, dass bei Entlassungen eben diejenigen dran sind, bei denen der Algorithmus vorhersagt, dass sie ohnehin schon auf dem Absprung sind. Ob das im konkreten Einzelfall korrekt ist, ist nicht entscheidend.[12]
Auch Gesundheitsdaten bieten sich für Analysen an – „Wearables" können durch die gesammelten Daten über Bewegung, Schlaf, Herzfrequenz etc. schon bald dazu beitragen, dass man gewarnt wird, wenn sich die Indikatoren für eine ernsthafte Erkrankung verdichten.[13] Vorstellbar ist zum Beispiel, dass der Manager gewarnt wird, dass er auf einen Burnout zusteuert, wenn er über einen längeren Zeitraum zu schlecht schläft, sich seine Arbeitsgeschwindigkeit am Computer verringert und sich der Blutdruck erhöht. Allen diesen Anwendungsfällen ist die Ambivalenz gemeinsam, denn einerseits bieten sie Chancen, andererseits wird das Unerwartete, die Ausnahme nicht wahrgenommen und bekommt auch keine Chance mehr, sich zu entfalten.

Drei Schlussfolgerungen lassen sich ziehen:

- Anonymität, so die *erste*, ist weder einfach zu erreichen, noch löst sie jedes Problem – Identifizierung bleibt unter Umständen auch in anonymen Datensätzen grundsätzlich möglich und wenn nicht, dann reicht für viele Anwendungen auch eine mit hoher Wahrscheinlichkeit richtige Einschätzung, auch wenn sie zu Resultaten führt, die wir als ungerecht empfinden.

---

[11] Vgl. Paul Lewis / Paul Hilder: Leaked: Cambridge Analytica's blueprint for Trump victory, in: The Guardian, 23.3.2018, https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory.

[12] Vgl. Tim Adams: Job hunting is a matter of Big Data, not how you perform at an interview, in: The Guardian, 10.5.2014, https://www.theguardian.com/technology/2014/may/10/job-hunting-big-data-interview-algorithms-employees.

[13] Vgl. Xiao Li et al.: Digital health. Tracking physiomes and activity using wearable biosensors reveals useful health-related information, in: PLOS Biology 15/1 (2017), https://doi.org/10.1371/journal.pbio.2001402.

- Sich individuell zu schützen ist, so die *zweite* Schlussfolgerung, oftmals eben nicht möglich. Auch wenn man selbst nichts zu den Datensilos beiträgt, wirken sie sich darauf aus, wie man Informationen im Internet findet, was man sieht und welche Möglichkeiten man hat.
- Die *dritte* Schlussfolgerung ist etwas weniger offensichtlich, aber doch wesentlich. Der Schutz persönlicher Daten greift da zu kurz, wo große anonyme Datensätze verwendet, aus ihnen Wahrscheinlichkeiten generiert und diese auf das Nutzererlebnis angewendet werden. Datenschutzgesetze beziehen sich aber oft eben nicht auf solche Datensätze, die an sich nicht identifizierbar machen – und das ist auch sinnvoll, denn es ist heute noch überhaupt nicht wirklich alles bekannt, was sich aus diesen Daten auch an wünschenswerten Erkenntnissen gewinnen lässt, welche Datenanalysetechniken noch entwickelt werden und was wir davon haben könnten. Ein generelles Verbot könnte also leicht kontraproduktiv werden. Auch hier gilt es also wieder, in einen Prozess der kontinuierlichen Abwägung einzutreten, für den man technisch mit den Entwicklern auf Augenhöhe sprechen können muss.

4 Was kann man lernen? Was lässt sich tun?

Über den Detailblick auf das „Problem" der Anonymität lassen sich viele typische Probleme aufzeigen, nicht nur in Bezug auf Privatheit, sondern durchaus darüber hinaus. Anonymität ist ambivalent, sie ist an sich weder gut noch schlecht, aber sie hat sowohl wünschenswerte als auch weniger wünschenswerte Konsequenzen. Nur wenn beides sorgfältig abgewogen wird, können wir überhaupt zu guten Regelungen kommen. Anonymität ist – trotz aller Möglichkeiten, sich individuell zu anonymisieren – kein Schutz gegen einige Formen der Datenanalyse; Konsequenzen von Identifikation (auch partieller) können nicht immer auf der individuellen Ebene verhindert werden. Es braucht kollektive Lösungen. Anonymität ist auch eine Illusion, die wir haben, weil wir zu wenig von der Technik hinter dem Netzwerk oder den Daten verstehen. Dadurch sind diejenigen, die beides entwickeln und die Grenzen des Machbaren ausreizen, im Vorteil. Hier müssen wir lernen mitzuhalten und die Technik, die wir nutzen und regulieren wollen, zumindest im Ansatz verstehen. Ansonsten werden wir immer wieder Probleme haben, gute Entwicklungen zuzulassen und Fehlentwicklungen zu unterbinden. Am schwierigsten aber ist wohl, dass es keine Patentlösungen gibt, keine Möglichkeit, das Problem ein für alle Mal (oder wenigstens für viele Jahre) ad acta zu legen. Wir müssen einsehen, dass die Verregulierung des Bereiches der Privatheit ein nicht abzuschließender Prozess ist, in dem sich wandelnde Werte, neue Technologien und zunehmend unterschiedliche Kulturen zusammengebracht werden müssen.