Autor:innen

Nikki Böhler und Nathalie Klauser stehen hinter dem sozialen Unternehmen **Intersections**, welches digitale Ungerechtigkeiten beleuchtet und neue Lösungsansätze für inklusive Innovation entwickelt. Sie hinterfragen bestehende Paradigmen, indem sie betroffene und marginalisierte Personengruppen von Anfang an bei Innovationsprozessen einbeziehen. Beide Gründerinnen haben langjährige Erfahrung aus unterschiedlichen Organisationen an der Schnittstelle zwischen Digitalisierung und Gesellschaft.
info@intersections.ch

Nikki Böhler studierte Volkswirtschaft an der Universität St. Gallen und gründete dort den Social Business Club, um das soziale Bewusstsein der Studierenden zu stärken. Als Vorstandsmitglied der SGG fördert sie soziale Innovation in der Schweiz und als Mitglied im Expertenkomitee des Digital Trust Labels setzt sie sich für mehr Transparenz und Vertrauen im Rahmen der Digitalisierung ein.

Nathalie Klauser studierte Publizistik, Politik und Wirtschaftspsychologie an der Universität Zürich. Sie ist Mitglied des Daten-Ethik-Beirats der Datengenossenschaft Midata, fördert als Co-CEO des nationalen Innovation Booster-Programms «Swiss Smart Cities» von Innosuisse und als Co-Präsidentin der SmartCity Alliance nachhaltige und smarte Ideen für Städte und Regionen. Weiter setzt sie sich für eine inklusive und bürgernahe Digitalisierungspolitik als Vorstandsmitglied von CH++ ein.

Abstract

Keywords: data gap, data justice, data bias, digital intersectionality, digital ethics, digital inclusion

Data divides our society, similar to raw materials
Data is considered as a raw material of the digital age. The comparison between data and conventional commodities is limp on various levels, but their impact on power and injustice unites them. Although data has become as central as commodities, sustainability and equity issues have not yet arrived in the data economy.

Keywords: Data Gap, Datengerechtigkeit, Data Bias, Digitale Intersektionalität, Digitale Ethik, Digitale Inklusion

Daten spalten unsere Gesellschaft, ähnlich wie Rohstoffe

Nikki Böhler und Nathalie Klauser

Daten gelten als Rohstoff des digitalen Zeitalters. Der Vergleich von Daten mit herkömmlichen Rohstoffen hinkt auf verschiedenen Ebenen, deren Auswirkungen auf Macht und Ungerechtigkeiten verbindet sie jedoch. Obwohl Daten mittlerweile eine ebenso zentrale Rolle wie Rohstoffe einnehmen, sind die Themen Nachhaltigkeit und Gerechtigkeit in der Datenökonomie noch nicht angekommen.

Der Rohstoff des digitalen Zeitalters?

Die Behauptung «Data is the new oil» wurde Anfang der Nullerjahre von einem britischen Mathematiker und Unternehmer auf einer Marketing-Konferenz geprägt (Palmer 2006). Die Unternehmen und Geschäftsmodelle, welche in den Folgejahren das stärkste Wachstum verbuchten, schienen von diesem «neuen» Rohstoff zu profitieren (Kershner 2021). Sind Daten also das neue Gold, das neue Öl, das neue Wasser in unserer digitalisierten Gesellschaft?

Unterstützt wird die häufig zitierte Analogie dadurch, dass Daten wie auch Rohstoffe gewonnen und verarbeitet werden und unsere digitalisierte Gesellschaft am Laufen halten. Wie Gold gefiltert, Öl raffiniert und Wasser geklärt, um eine wertigere Nutzbarkeit zu erlangen, werden Daten selektiert, aufbereitet und interpretiert. Ihr Wert ist wie bei Rohstoffen abhängig von ihrer Nutzbarkeit, welche allein durch ihr Vorhandensein noch nicht gegeben ist. Über 99 % des Wassers auf unserer Erde ist nicht trinkbar (National Geographic 2022) und auch bei Daten gehören bis zu 90 % zur Kategorie «Dark Data» (Johnson 2015). Das sind Daten, die

gesammelt und gespeichert, aber nie ausgewertet – teilweise nicht mal gesichtet – werden (Frenkel 2021). Die prozessierten Daten hingegen durchdringen im Zuge der sogenannten Datafizierung (Cukier/Mayer-Schoenberger 2013) mittlerweile fast jeden unserer Lebensaspekte: vom Konsum (bsp. automatisierte, zielgruppenspezifische Werbung und Minimierung der Lagerkapazitäten mittels Verkaufsprognosen) über die Mobilität (bsp. Planung des öffentlichen Verkehrs basierend auf Bewegungsdaten oder selbstfahrende Fahrzeuge) bis zur Kommunikation (bsp. algorithmisch bestimmte Inhaltsanzeige auf sozialen Medien und digitalen Nachrichten). Die breite Einsatzmöglichkeit und der hohe Wert dieser Daten können ähnlich wie bei der Rohstoffgewinnung zur Ausbeutung führen. Auch digitale Grosskonzerne nutzen Menschen im Globalen Süden aus, indem sie Personen zu niedrigsten Löhnen, ohne psychologische Unterstützung und unter stressigen Bedingungen verstörende Inhalte sichten lassen (Steiger 2021). Weltweit haben sich bei Digitalkonzernen missbräuchliche Praktiken etabliert, indem sie Berge an persönlichen Daten basierend auf unlesbaren AGBs (Belioniel 2019) ohne das (Ein-)Verständnis (Cakebread 2017) der Kundschaft sammeln und damit Milliardengewinne erzielen, welche sie mittels unfairer Steuerpraktiken nicht korrekt versteuern (Satyajeet 2022). Zudem werden Daten gemäss dem Prinzip von Angebot und Nachfrage in Märkten bepreist und gehandelt. Auch hier verzerren Quasi-Monopole analog zur Diamantenindustrie (Conklin 2005) die Marktposition in Preis und Innovations-Zugang: Die Big Five (Alphabet, Amazon, Apple, Meta Platforms und Microsoft) dominieren die Datenökonomie, wie die sich häufenden Gerichtsurteile[1] und die Bedenken von supranationalen Aufsichtsgremien[2] zeigen. Diese Monopolstellung verfestigt sich derzeit mit der Entwicklung der künstlichen Intelligenz, die nur den wenigen Unternehmen möglich ist, die enorme Datenmengen horten.

Ein Vergleich mit Limitationen

Doch der Vergleich von Daten mit analogen Rohstoffen hinkt. Der Wert von Daten entsteht im Gegensatz zu richtigen Rohstoffen hauptsächlich erst in der Verarbeitung und kontextbedingten Anwendung. Die grosse Menge an kostenlosen offen zugänglichen und unfiltrierten Datensätzen im Netz zeigt den tiefen Rohdatenwert an (Lundqvist 2015). Erst durch ihre Spezifizierung, Aufbereitung und Interpretation werden

1 Beispiele von Gerichtsurteilen: Google (2018): https://netzpolitik.org/2018/google-muss-milliardenstrafe-fuer-faule-taktiken-bei-android-zahlen/, Amazon (2020): https://www.tagesschau.de/wirtschaft/eu-kartellklage-gegen-amazon-101.html, Facebook (2020): https://www.faz.net/aktuell/wirtschaft/digitec/klage-gegen-facebook-nicht-nur-einen-klick-entfernt-17095966.html (28.10.2022).

2 UN General Assembly (2018): Promotion and protection of the right to freedom of opinion and expression, https://freedex.org/wp-content/blogs.dir/2015/files/2018/10/AI-and-FOE-GA.pdf (28.10.2022).

Daten zu einer nutzbaren Information und können einen Mehrwert schaffen. Auch sind Daten im Gegensatz zu analogen Rohstoffen (vgl. Pelosi 2022) rückführbar: Ihre Bewegung und Verarbeitung kann digital verfolgt und nachvollzogen werden. Und der wichtigste Unterschied: Daten sind nicht erschöpfbar. Sie sind nicht wie bspw. Öl in begrenzter Form vorhanden. Daten bergen also im Gegensatz zu materiellen Rohstoffen wie Gold das Potenzial, als universell anwendbare Informationsquelle allen zugutezukommen. Trotz der aufgeführten Widersprüche in der Analogie lässt sich konstatieren, dass Daten im 21. Jahrhundert wohl einen der wichtigsten Rohstoffe in Wissenschaft und Wirtschaft für den technologischen Fortschritt darstellen.

Mit Macht kommt die Verantwortung

In der Rohstoffindustrie ist ein Trend auszumachen, hin zu mehr Nachhaltigkeit und sozialer Verträglichkeit (World Economic Forum 2022). Diese Bedingungen werden als Voraussetzung gesehen, um langfristig zu wirtschaften. Die Datenindustrie hinkt da noch hinterher. So werden Daten-Centers in der ökologischen Nachhaltigkeitsbewegung noch vernachlässigt (ebd.), tun sich auch in den sozialen Bereichen wie den Arbeitsbedingungen schwer (Laeri 2019a), und auch gesellschaftlich relevante, selbstauferlegte ethische Kodizes der Big-Tech-Konzerne werden gebrochen (Harwell/Tiku 2020). Daten sind genauso wie Rohstoffe eine Form von gesellschaftlicher Macht geworden. Daten können genutzt werden, um Ungerechtigkeiten aufzudecken, die Gesundheitsversorgung zu verbessern oder die Klimakrise zu bekämpfen. Sie können aber auch dazu dienen, zu diskriminieren, zu überwachen und Privilegien auszuspielen. Dieses Potenzial zum positiven gesamt-gesellschaftlichen Fortschritt einerseits und zur missbräuchlichen Macht-Zementierung andererseits macht es unumgänglich zu fragen: Wer entscheidet, welche Daten gesammelt werden? Wie werden diese Daten gesammelt, bewertet und genutzt? Wer hat die Kontrolle über Daten und profitiert davon?

Programmierte Voreingenommenheiten

Voreingenommenheiten sind bewusst und unbewusst in uns allen verankert, wie die folgende Kurzgeschichte veranschaulicht:

> Vater und Sohn fahren im Auto. Sie haben einen schweren Unfall, bei dem der Vater sofort stirbt. Der Bub wird mit schweren Kopfverletzungen in ein Krankenhaus gebracht. Er wird in den OP-Saal gefahren, wo schon die diensthabenden Chirurgen warten. Als sie sich jedoch über den Jungen beugen, sagt jemand aus dem Chirurgen-Team mit erschrockener Stimme: «Ich kann nicht operieren – das ist mein Sohn!» (March 2016)

Das erstmalige Lesen der Geschichte an einer Vorlesung der Uni Zürich führt bei einem Grossteil von Studierenden zu Irritationen und Ratlosigkeit. Sogar weibliche Kommilitoninnen denken oft nicht an die Option der Mutter als Chirurgin (Scherer 2022). Dieses Beispiel illustriert eine mögliche Verzerrung in der Lesart von Daten.

Wie auch bei analogen Rohstoffen werden in allen Etappen der Datenverarbeitung Ungerechtigkeiten eingebettet: in der Sammlung, Auswertung und Anwendung. Beim Sammeln der Daten entstehen bereits Verzerrungen. Denn das Vorhandensein von Datensätzen selbst spiegelt ein verborgenes Ungleichgewicht in den institutionellen Infrastrukturen und den gesellschaftlichen Machtverhältnissen wider. Wikipedia zum Beispiel scheint eine reichhaltige und vielfältige Datenquelle zu sein. Aber weniger als 18 % der biografischen Einträge auf der Website sind über Frauen (Zou/Schiebinger 2018). Artikel über Frauen sind häufiger mit Artikeln über Männer verlinkt als umgekehrt, wodurch Männer für Suchmaschinen besser sichtbar sind. Je nachdem, wie Daten methodisch verarbeitet werden, können sie so ungewollt nicht nur geschlechtsspezifische, sondern auch ethnische und kulturelle Vorurteile enthalten und verstärken (ebd.).

Nicht nur Wikipedia, sondern die Datensammlung und -verarbeitung allgemein wird aktuell überwiegend von westlichen, weissen, männlichen Personen der Technikbranche geprägt (D'Ignazio/Klein 2020). Dadurch entstehen fehlende, unvollständige und verzerrte Datensätze. Die Verarbeitung solcher Daten wird branchentypisch als «garbage in, garbage out» bezeichnet. In der Auswertung wie auch Modellierung der Daten für automatisierte Entscheidungssysteme kommen zusätzlich persönliche Biases der Datenanalyst:innen zum Tragen und auch in diesem Prozessschritt werden betroffene Personengruppen selten eingebunden. Dies führt zu voreingenommenen und verzerrten, falschen Resultaten – sichtbar in Analysen, Produkten, Innovationen und Algorithmen (Perez 2019). In der Anwendung dieser Resultate verschärfen fehlendes Empowerment der betroffenen marginalisierten Personengruppen und Intransparenz im Prozess (Courtland 2018) und der Verarbeitung das Problem weiter.

Datenbasierte Ungerechtigkeit

Durch Datenlücken, -voreingenommenheiten und -fehlbewertungen werden Ungleichheiten verstärkt, Zugänge versperrt und Lösungen für gewisse Personengruppen verhindert. Die folgende «Smart City»-Umsetzung illustriert dies deutlich: Viele Städte haben das Problem von Schlaglöchern im Strassenbelag. Die Stadt Boston beschloss, dieses Problem mit einer technischen Applikation anzugehen. Diese nutzte den internen Beschleunigungsmesser des Mobiltelefons, um den Ruck eines Autos zu erkennen, das über ein Schlagloch fuhr, und übermittelte automatisch via GPS den Ort des Schlaglochs an die Stadtbehörden. So wussten Mitarbeitende der Strasseninstandhaltung genau, wo die Löcher zu reparieren wären. In der

Realität zeigte sich, dass vorwiegend Meldungen aus wohlhabenden Gegenden eintrafen, da der Besitz eines Autos und neueren Smartphones damit korrelierte. Streng nach diesem Datensatz zu reparieren, löst weder die generelle Schlaglochproblematik, sondern verschlimmert soziale Ungerechtigkeiten in der Stadt (Hand 2020). Dies ist nur eine exemplarische Veranschaulichung von zahlreichen Benachteiligungen, welche sich in der Datenökonomie des letzten Jahrzehnts grösstenteils unentdeckt auswirkten (Perez 2019). Während in der Rohstoffindustrie bereits ein Bewusstsein und längere Diskussionen über die Gefahren und Risiken im Gange sind, scheint in der Datenökonomie noch weitestgehend unklar: Wo werden keine Daten gesammelt und welchen Effekt haben diese Datenlücken auf Prozesse und die Gleichberechtigung aller Personengruppen? Welche Wirkung auf die Lebensrealität von Individuen und unsere Gesellschaft haben Daten, die sehr einseitig sind, nur bestimmte Personengruppen berücksichtigen und andere ausschliessen? Klar ist: Menschen werden basierend auf sich teilweise überkreuzenden Merkmalen (Geschlecht, Rasse, Status, Alter, Beeinträchtigung etc.) digital diskriminiert (Zheng/Walsham 2021).

Wer entscheidet, welche Daten gesammelt werden? Wie werden diese Daten gesammelt, bewertet und genutzt? Wer hat die Kontrolle über Daten und profitiert davon?

Die Beispiele dafür häufen sich: Google Maps sind grösstenteils nicht barrierefrei: Wenn Strecken mit einem Kinderwagen oder im Rollstuhl befahren werden, kommt Mensch plötzlich nicht weiter. Hunderte Arbeitgebende schalten ihre Werbeinserate auf Facebook und LinkedIn gezielt nur an Menschen unter 36 Jahre und schliessen ältere potenzielle Arbeitnehmende so von Anfang aus (Laeri 2019b). Kommerzielle Gesichtserkennungssysteme klassifizieren das Geschlecht viel häufiger falsch, wenn sie dunkelhäutige Frauen statt hellhäutige Männer vor sich haben, mit einer Fehlerquote von 35 % gegenüber 0,8 % (Buolamwini/Gebru 2018).

Lösungsansätze für mehr Datengerechtigkeit

Um auch in der Datenindustrie Gerechtigkeit voranzubringen, braucht es ein stärkeres Bewusstsein und eine neue Herangehensweise. Daten sind nicht neutral und emotionslos, sondern prägen unsere Lebensrealität auf voreingenommene Weise. Die Relevanz von Datenlücken und datengeprägter Ungerechtigkeiten muss mehr Aufmerksamkeit erhalten, und es gilt, diese Themen transparent aufzuarbeiten. Noch

sehr vereinzelt und isoliert sind erste Ansätze innerhalb unternehmerischer Ethikgrundlagen[3], der Community[4] und alternativer[5] Plattformen, verstärkter themenspezifischer wissenschaftlicher Publikationen und politischer Diskussionen erkennbar. Datengerechtigkeit bietet ein kollaboratives Feld für neue Lösungsansätze zwischen zivilgesellschaftlichen NGOs, Sozial-, Geistes- und technischen Akademien, Unternehmen und Behörden. Dabei müssen in allen Prozessschritten, wo Daten bzw. digitale Services und Produkte eingesetzt werden, betroffene und marginalisierte Personengruppen nicht nur mitgedacht, sondern konsequent involviert werden. Dafür braucht es neue Standards und Prozesse für die Datensammlung, -auswertung und -anwendung von Unternehmen und Behörden. So stellen wir sicher, dass die Digitalisierung langfristig allen zugutekommt.

3 https://digitaltrust-label.swiss/

4 https://data2x.org/

5 https://www.midata.coop/

Referenzen:

Belioniel, U. et al. (2019): *The Duty to Read the Unreadable*, in: Boston College Law Review, http://dx.doi.org/10.2139/ssrn.3313837 (28. Oktober 2022).

Buolamwini, J. und T. Gebru (2018): Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research, https://proceedings.mlr.press/v81/buolamwini18a.html (30. Oktober 2022).

Cakebread, C. (2017): You're not alone, no one reads terms of service agreements. Business Insider, https://www.businessinsider.com/deloitte-study-91-percent-agree-terms-of-service-without-reading-2017-11?r=US&IR=T (28.Oktober 2022).

Conklin D. W. et al. (2005): De Beers and the Global Diamond Industry. Ivey publishing, https://hbsp.harvard.edu/product/905M40-PDF-ENG (28. Oktober 2022).

Courtland, R. (2018): *Bias detectives: the researchers striving to make algorithms fair,* in: Nature 558(7710): 357–360, doi: https://doi.org/10.1038/d41586-018-05469-3

Cukier, K. N. und V. Mayer-Schoenberger (2013): The Rise of Big Data. Foreign Affairs, https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data (3. November 2015)

D'Ignazio C. und F. K. Klein (2020): *Data Feminism.* The MIT Press.

Frenkel, K. A. (2021): *«Is Data Collection a Waste of Time and Resources?»,* in: CIO Insight, https://www.cioinsight.com/it-strategy/is-data-collection-a-waste-of-time-and-resources/ (28. Oktober 2022).

Hand, David J. (2020): *Dark Data: Why What You Don't Know Matters.* Princeton University Press.

Harwell, D. und N. Tiku (2020): *Google's star AI ethics researcher, one of a few Black women in the field, says she was fired for a critical email,* in: The Washington Post, https://www.washingtonpost.com/technology/2020/12/03/timnit-gebru-google-fired/ (30. Oktober 2022).

Johnson, H. (2015): «Digging up dark data: What puts IBM at the forefront of insight economy». SiliconANGLE, http://siliconangle.com/blog/2015/10/30/ibm-is-at-the-forefront-of-insight-economy-ibminsight/ (3. November 2015).

Kershner, M. (2021): *Data is the New Oil,* in: Forbes, https://www.forbes.com/sites/theyec/2021/07/15/data-isnt-the-new-oil--time-is/?sh=550a4af935bb (28. Oktober 2022).

Laeri, P. (2019a): *Die neuen digitalen Tagelöhner,* in: Blick Online, https://www.blick.ch/meinung/kolumnen/aufbruch-mit-patrizia-laeri-die-neuen-digitalen-tageloehner-id15545161.html (30. Oktober 2022).

Laeri, P. (2019b): *Zeitphänomen Altershass,* in: Blick Online, https://www.blick.ch/meinung/kolumnen/aufbruch-mit-patrizia-laeri-zeitphaenomen-altershass-id15590346.html (30. Oktober 2022).

Lundqvist, B. et al. (2015): Open data and Competition Law: Some Issues Regarding Access and Pricing of Raw Data. HeinOnline, https://heinonline.org/HOL/LandingPage?handle=hein.journals/mujlt9&div=21&id=&page= (28. Oktober 2022).

March, E. (2016): I told a kid a riddle my dad told me when I was 7. His answer proves how far we've come. Upworthy, https://www.upworthy.com/i-told-a-kid-a-riddle-my-dad-told-me-when-i-was-7-his-answer-proves-how-far-weve-come (30. Oktober 2022).

National Geographic (2022): Earth's Freshwater. Education Guide, https://education.nationalgeographic.org/resource/earths-fresh-water (28. Oktober 2022).

Palmer, M. (2006): Data is the New Oil. ANA Marketing Maestros, https://ana.blogs.com/maestros/2006/11/data_is_the_new.html (28. Oktober 2022).

Pelosi, D. et al. (2022): Gold: Die schwierige Suche nach Transparenz. SRF Radio Trend, https://www.srf.ch/audio/trend/gold-die-schwierige-suche-nach-transparenz?id=12276022 (28. Oktober 2022).

Perez, C. C. (2019)*: Invisible Women: Exposing Data Bias in a World Designed for Men*. Abrams Press: New York.

Satyajeet, M. (2022): Western tech companies saving billions in taxes in the global south. netzpolitik.org, https://netzpolitik.org/2022/series-on-digital-colonialism-western-tech-companies-saving-billions-in-taxes-in-the-global-south/ (28. Oktober 2022).

Scherer, A. (2022): *Consumer Behavior in the Digital Age*. Lecture at University of Zurich.

Steiger M. et al. (2021): The Psychological Well-Being of Content Moderators. CHI Conference on Human Factors in Computing Systems, https://crowd.cs.vt.edu/wp-content/uploads/2021/02/CHI21_final__The_Psychological_Well_Being_of_Content_Moderators-2.pdf (3. November 2015).

Tost, M. et al. (2022): *Nachhaltiger Bergbau 2050,* in: BHM Berg- und Hüttenmännische Monatshefte 167: 151–157.

UN General Assembly (2018): Promotion and protection of the right to freedom of opinion
and expression, https://freedex.org/wp-content/blogs.dir/2015/files/2018/10/AI-and-FOE-GA.pdf (28. Oktober 2022).

World Economic Forum (2022): How data centres can join the circular economy to become greener, https://www.weforum.org/agenda/2022/07/datacentres-must-join-the-circular-economy-to-become-greener/ (30. Oktober 2022).

Zheng, Y. und G. Walsham (2021): *Inequality of what?* An intersectional approach to digital inequality under Covid-19, in: Information and Organization 31 (2021): 1–6, https://doi.org/10.1016/j.infoandorg.2021.100341

Zou, J. und L. Schiebinger (2018): *AI can be sexist and racist – it's time to make it fair,* in: Nature 559: 324–326, https://www.nature.com/articles/d41586-018-05707-8 (30. Oktober 2022).