Vermittlung von Datensensibilität

Um eine zufriedenstellende Antwort auf die Kontroverse geben zu können, ob eine Zerstörung der *Membran von Welt und Privatheit* durch Big Data und Mustererkennung sich zwangsläufig ergeben muss, werden wir drei Perspektiven auf das Thema „Sichtbarmachen des bislang Unsichtbaren“ werfen.

Mustererkennung mit personenbezogenen Daten

Diese Strategie lässt sich in letzter Konsequenz so beschreiben: Über Mustererkennung werden die privaten Verhaltensdaten, die im Netz als Such-, Bewegungs-, Text-, Sprechdaten und vieles mehr anfallen, in ihre wertvollen Substanzen zerlegt. Es ist die Jagd auf das besondere Profil jedes Einzelnen, mit dem Ziel ein Profiling der Nutzer zu erreichen. Sind die „Bestandteile“ transparent, ist die Person, mit Unterstützung von Künstlicher Intelligenz, als „Besonderheit“ mit seinen persönlichen Neigungen, seiner Gemütsverfassung oder seiner individuellen ökonomischen Situation vorhersagbar, die als Datenspuren für die Profilbildung zu nutzen sind.

Das Ziel dieser Predictive Analytics ist: Mit Unterstützung von Deep-Learning-Methoden soll die Steuerung der Einzigartigkeit (Singularität) jedes Einzelnen, wie von sozialem Leben möglich werden. Es ist die Produktion vorhersagbarer Individuen durch individuelle Verhaltensüberwachung und -steuerung.

Das Netz „dringt nun von außen in die Privatsphäre ein – wo es nichts zu suchen hat, aber viel zu finden gibt“, so der Sozialwissenschaftlers Armin Nassehi. Die Mustererkennung rücke jetzt die Zerstörung der unsichtbaren Membran von Welt und Privatheit, in der wir geschützt leben wollen, in den Fokus. Big Data kombiniere Daten, die nicht für andere bestimmt sind und erzeuge dadurch für Dritte einen Mehrwert (Nassehi 2019, S. 302).

Noch schärfer drückt das Soshana Zuboff aus: Die Verwendung von privaten Verhaltensdaten sei ein historischer Wendepunkt, da diese als „Verhaltensüberschuss“ entdeckt und in ein marktfähiges Produkt verwandelt werden. „Die Nutzer sind die Quelle eines kostenlosen Rohstoffs für einen neuartigen Produktionsprozess“ (Zuboff 2018/ 1 & 2). „Unsere intimste alltägliche Realität ist im Überwachungskapitalismus wiedergeboren als Verhalten, das es zu überwachen und zu verändern, zu kaufen und zu verkaufen gilt“ (Zuboff 2019, (3).

Mustererkennung mit Forschungs- und Produktionsdaten

Ein Großteil der anfallenden Daten habe keinen Personenbezug, so eine verbreitete Meinung. Das gilt vor allem für Forschungsdaten, etwa bei der Klimamodellierung, die lange Laufzeiten haben. Sie laufen auf Hochsleistungsrechnern und erzeugen Datenmengen im Bereich Tera- und Petabyte für eine weltweite Nutzung, vorwiegend für wissenschaftliche und politische Debatten, aber u.a. auch für die

Versicherungswirtschaft (Ludwig/ Thiemann 2020).

Die Nutzung und Auswertung von Produktions- und Logistikdaten, insbesondere im Bereich Industrie 4.0, sind für Bestehen und Wachstum der Unternehmen im globalen Wettbewerb überlebenswichtig. Sie sind auf den ersten Blick unkritisch. Das wird in vielen Fällen zutreffen. Selbst wenn der überwiegende Teil als „reine Daten“ mit Sach- oder Materialbezug einzustufen ist, so ist in den Produktions- und Logistikketten in der Regel ein Anteil enthalten ist, der einen Arbeitsplatz oder menschliche Arbeitshandlungen einbezieht und extrahiert werden kann. Auch auf den ersten Blick unkritische Anwendungen, wie das Internet der Dinge oder digitalisierte Autobahnen, liefern keinen Datenmüll, sondern auch Daten zur fortschreitenden Kommerzialisierung und Überwachung (Hesse 2020).

Mustererkennung mit anonymisierten Daten

Bei der Mustererkennung mit anonymisierten Daten sind ursprünglich personenbezogene Daten so verändert, dass eine Zuordnung zu einer Person nur noch mit großem Aufwand realisierbar, aber eben immer noch möglich ist. Die Möglichkeit zur Identifizierung steigt logischerweise mit der Zahl der gespeicherten Datenmerkmale. Laut einer Studie in „Nature“, können mit nur 15 Merkmalen in jedem Datensatz, wie Alter, Wohnort oder Nationalität, 99,98 Prozent der US-Amerikaner identifiziert werden (https://netzpolitik.org/2019/).

Auch in dieser, auf den ersten Blick unkritischen Kategorie, ist keine Sicherheit gegeben, nicht in die von Zuboff und Nassehi genannte Kategorie zu geraten. Es steht im Befinden und im Wohlwollen der Datenbesitzer, allerdings wird der Aufwand beträchtlich sein.

Individuelle Interessen können darüber hinaus durch Verwertung nicht personenbezogener Daten beeinträchtigt werden, indem Daten kombiniert werden, die dafür ursprünglich nicht gedacht waren, was etwa bei anonymisierten Gesundheitsdaten oder Bewegungsdaten von Individuen der Fall ist. Die Verfügungsmacht über diese Daten, der sich Internetnutzer weder verweigern können noch von der sie oft etwas ahnen, generiert vor allem Macht bei den Plattformen über ökonomische und gesellschaftliche Prozesse und Entwicklungen, was tagtäglich an der Expansion von Google, Facebook & Co. zu erkennen ist.

Hier hat sich mittlerweile eine lohnende kommerzielle Beratungsszene aufgetan. Für Unternehmen und Regierungen bietet beispielsweise das Silicon Valley-Start-up Palantir seine Software zur Analyse großer Mengen heterogener Daten an. Investoren sind u.a. der US-Geheimdienst CIA. Regierungen, das Militär und Behörden sorgen für rund die Hälfte des Umsatzes von 743 Millionen Dollar, was die Vermutung aufkommen lässt, dass staatliche Institutionen sich zum Zweck der Verschleierung ihrer Aktivitäten oder der Gründung eines profitträchtigen Geschäftsmodells hier am Werke sind. Kunden

in Deutschland sind der Pharmakonzern Merck, der Palantir-Software und –Beratung für molekularbiologische und pharmazeutische Forschungen einsetzt, das Medienhaus Springer, Polizeibehörden sowie Airbus, das darauf Wert legte, dass Palantir die Speicherung der Daten auf Servern in Europa zusichert (O.V. Handelsblatt Online 2020.

Die Schlussfolgerun: Aus der Mustererkennung können sich positive wie negative Eingriffe in bestehende gesellschaftliche Strukturen ergeben. Negative, beispielsweise beim Schutz der Privatheit, wo es um Produktion vorhersagbarer Individuen durch Verhaltensüberwachung und –steuerung geht. Andererseits können komplexe Strukturen in Ökonomie, Ökologie und Forschung transparent gemacht werden, die anders nicht offenzulegen sind. Die Mustererkennung in Verbindung mit Algorithmen der KI ist ein effektives Datenanalysewerkzeug, das sowohl zur Forcierung ökonomischer Produktivität, von Innovationssteigerungen und Machterhalt genutzt werden kann, genauso wie zur Zerstörung der Demokratie oder zur Vertiefung und Absicherung sozial-ökologischer Narrative.

Essay Arno Rolf

Literatur

https://doi.org/10.1007/s11609-020-00403-9

O.V. Handelsblatt Online (2020): Überraschend mühselig und europäisch: Einblicke in das Geschäft von Palantir. Ein Insiderbericht, 28.09.2020

Hesse, Wolfgang (2020): Das Zerstörungspotenzial von Big Data und Künstlicher Intelligenz für die Demokratie, In: Informatik Spektrum(2020) 43:339-346 S.

Ludwig/ Thiemann (2020): Datenkompetenz – Data Literacy. Informatik Spektrum 45/2020

Nassehi Armin (2019): „Muster. Theorie der Gesellschaft“, C. H. Beck, München 2019.

Zuboff, Soshana (2019) (3) http://www.faz.net/aktuell/feuilleton/debatten/die-digital-debatte/shoshana-zuboff-googles- ueberwachungskapitalismus-14101816.html?printPagedArticle=true#pageIndex_2

https://netzpolitik.org/2019/weitere-studie-belegt-luege-anonymer-daten/#vorschaltbanner, abgerufen 14.9.20).

Fragen:
(1) Haltet ihr die Betonung von Datensensibilität erforderlich?
(2) Beschreibt bitte die drei unterschiedlichen gesellschaftlichen Datentypen.