

Interview Bart Van Moerkerke mit dem niederländischen Datenexperten Mirko Tobias Schäfer

Daten haben auch eine politische und ethische Komponente

„Daten und Algorithmen bieten viele Möglichkeiten, sind aber nicht neutral. Was analysiert wird, welche Algorithmen entwickelt werden, welche Daten verwendet werden und wofür, das sind alles politische und ethische Entscheidungen. Die politischen Entscheidungsträger sind sich dessen nur unzureichend bewusst. Grundkenntnisse in Sachen Datenkompetenz sind auch für Kommunalpolitiker unerlässlich.“ Das sagt der niederländische Datenexperte Mirko Tobias Schäfer in diesem Interview mit Bart Van Moerkerke.

(<https://www.vvsg.be/kennisitem/vvsg/aan-data-zit-een-politieke-en-ethische-component>)

Mirko Tobias Schäfer ist Associate Professor für Neue Medien und Digitale Kultur an der Universität Utrecht und leitet dort die Utrecht Data School. Er untersucht die sozialen Auswirkungen von Datenpraktiken. Daten und Algorithmen werden oft als neutral angenommen, aber das ist nicht immer der Fall. Daten werden auch das ‚neue Öl‘ genannt, aber dieser Vergleich passt nicht. Daten sind keine Rohstoffe im Boden; jemand hat sich entschieden, bestimmte Daten zu sammeln und zu speichern, andere nicht. Daten sind nicht ‚gegeben‘, sondern eher gemacht; das hat eine politische und ethische Komponente. Die sogenannten Big Data und die neuen Praktiken diese zu analysieren und zu instrumentalisieren, beeinflussen unser Verständnis von Kultur, Bürgerschaft und Demokratie. Deshalb ist es so wichtig, dass auch die Geisteswissenschaften sich damit auseinandersetzen und bei der Umsetzung dieser Praktiken Orientierung geben und darüber nachdenken, was wünschenswert ist und was nicht. „Ich bin in erster Linie Medienwissenschaftler, kein Digitalexperte.“

Sind die Daten nicht per Definition neutral?

Nein. Es gibt Algorithmen, die keine direkte Auswirkungen auf Menschen haben, die beispielsweise sehr genau vorhersagen, welche Wassermühle wann gewartet werden muss oder wann sie ersetzt werden muss. Aber viele Daten sind nicht neutral. Es ist gut, dass diesen Fallstricken jetzt viel Aufmerksamkeit geschenkt wird. Autoren wie Cathy O'Neil oder Virginia Eubanks zeigen überzeugend, dass es Probleme mit Algorithmen und Daten gibt, dass sie rassistisch und voreingenommen sein können und dass sie unerwünschte soziale Ansichten verstärken können. Aber über negativen Meldungen sollten wir nicht vergessen, dass Algorithmen und Daten Dinge ermöglichen, die in der Vergangenheit so nicht möglich waren.

Sie ermöglichen es, auf neue Herausforderungen zu reagieren, sie eröffnen innovative Möglichkeiten der Problemlösung, sie führen zu neuen, multidisziplinären Formen der Zusammenarbeit. Wir müssen erkennen, dass sie politische Verantwortung erfordern. Wir müssen dazu Kompetenzen entwickeln, wie Datenpraktiken oder Algorithmen funktionieren, oder wie ein Datensatz zusammengesetzt wird.

Können Sie einige Beispiele für Bias in Daten nennen?

Das Lebensqualitätsbarometer der niederländischen Regierung ist ein Index für die Lebensqualität in jeden Wohnbezirk in den Niederlanden. Der Bezirk in Rotterdam, in dem ich viele Jahre lang gerne lebte, bekam eine niedrige Punktzahl. Offensichtlich war die Anzahl der Personen mit

Migrationshintergrund einer der Indikatoren, die der Algorithmus bei der Berechnung der Lebensqualität berücksichtigt.

Die niederländische Polizei nutzt ein System für ‚predictive policing‘. Das Ergebnis ist eine Karte, die die Wahrscheinlichkeit bestimmter Verbrechen in jedem Gebiet zeigt. In einem bestimmten Stadtteil nahm die Zahl der Fahrraddiebstähle stark ab. Waren die Polizeistreifen dort sehr erfolgreich? Nein, die investigative Journalismus Plattform Investico zeigte, dass zwei Polizeibüros in diesem Gebiet geschlossen wurden. Die betroffenen Bürger zeigten einfach keinen Fahrraddiebstahl mehr an. (<https://www.groene.nl/artikel/smileys-scores-platlullen-en-downgraden>)

In den Vereinigten Staaten wurde ein Algorithmus verwendet, um das Risiko auf Rückfallkriminalität abzuschätzen. AfroamerikanerInnen, die nur leichte Verbrechen begangen hatten, wurde ein viel höheres Risiko zugeschrieben als Weißen, die sich schwerer Verbrechen schuldig gemacht hatten. Als Journalisten nachträglich untersuchten, wer tatsächlich rückfällig wurde, entsprach das nicht den Vorhersagen. Die Grundlage des Algorithmus war zum Teil rassistisch. (<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>)

Die Arbeitslosenbehörde in Michigan (USA) hatte ein neues Computersystem zur Bekämpfung von Leistungsbetrug eingeführt. Es war so streng, dass selbst die falsche Schreibweise eines Namens oder die Eingabe eines falschen Geburtsdatums ausreichte, um des Betrugs beschuldigt zu werden. 20.000 Arbeitslose, bedürftige Menschen, die das Geld dringend benötigten, wurden zu Unrecht beschuldigt. (<https://www.theguardian.com/us-news/2016/feb/12/michigan-unemployment-insurance-benefit-automated-system-fraud-penalties>)

Daher ist es sehr wichtig zu wissen, welche Daten und Indikatoren für einen Algorithmus verwendet werden und auf welcher Grundlage dieser erstellt wurde. Um beim letzten Beispiel zu bleiben: Sie könnten auch einen Algorithmus entwickeln, der Personen identifiziert, die Anspruch auf Leistungen haben, diese aber noch nicht beantragt haben.

Den Algorithmus, der entwickelt wird und die Daten, die verwendet werden, sind daher eine politische und ethische Entscheidung?

Ja, die Nutzung von Algorithmen, um Leistungsbetrug zu ahnden, berührt beispielsweise grundlegende Fragen unseres Rechtssystems und unserer Gesellschaftsordnung. Sie machen alle Leistungsempfänger von vorneherein verdächtig. Dies passt nahtlos in eine lange Geschichte, in der Regierungen arme Menschen überwachen und kontrollieren. Wenn wir das jetzt in Algorithmen festschreiben, können wir es dann später wieder ändern? Wir müssen jetzt ein Bewusstsein dafür entwickeln, dass offene Standards und transparente Algorithmen für eine Demokratie notwendig sind. Wo steht Europa in dieser Sache?

Auf der einen Seite sind die USA mit einem libertären Marktverständnis, wo Unternehmen diese Funktionen bereits übernehmen und die Algorithmen Geschäftsgeheimnisse sein können. Auf der anderen Seite ist China, wo die Regierung die Handlungen von Unternehmen lenkt und direkten Zugang zu Daten erhält. Europa muss seine eigene Technologie entwickeln, die demokratische Werte und Bürgerrechte nicht untergräbt, sondern Kontrolle und Rechenschaftspflicht unterworfen sind.

Wenn wir diese Technologien lediglich von den USA oder China kaufen, erhalten wir letztendlich auch das jeweilige Wertesystem, das darin eingebettet ist.

Haben Sie den Eindruck, dass die politischen Entscheidungsträger sich dessen ausreichend bewusst sind?

Meine Sorge gilt zwei Sorten von Menschen: dem Ingenieur ohne Verständnis von der sozialen Welt und dem Technokraten ohne Verständnis von Technologie. Aber als Dozent bin ich natürlich ein unverbesserlicher Optimist und ich denke, dass Bildung hilft. Den politischen und ethischen Aspekten von Daten wird nun viel mehr Aufmerksamkeit geschenkt. Das Bewusstsein wächst. Immer mehr Gemeinden, Ministerien, Organisationen und Unternehmen bitten unsere Data School um Hilfe bei der Analyse der ethischen Fallstricke in den Datenprojekten, die sie starten wollen. Auch BürgermeisterInnen, StadträtInnen und GemeinderätInnen müssen in diesem Bereich wesentlich kompetenter werden. Es ist notwendig, dass sie über Grundkenntnisse in digitalen und ethischen Fähigkeiten verfügen. Ein Datenprojekt anzugehen ist eine politische Entscheidung: Was ist das Problem und wie kann uns ein Datenprojekt helfen, es zu lösen? Wollen wir das tun, oder lieber nicht? Wie machen wir das, und wie verhindern wir, dass unsere Normen und Werte dadurch angetastet oder untergraben werden? PolitikerInnen müssen auch die Ergebnisse interpretieren, und dazu müssen sie wissen, was hinter den Daten steckt. Sie erhalten einen Bericht, eine Infografik oder ein Dashboard und müssen auf dieser Grundlage eine verantwortungsvolle Entscheidung treffen. Sie können die gleichen Daten auf verschiedene Weise verwenden: Sie können sie verwenden, um eine Gruppe zu stigmatisieren oder eine Gruppe zu unterstützen.

Nehmen wir das Beispiel eines Algorithmus, der ziemlich gut vorhersagt, welche SchülerInnen die Schule vorzeitig verlassen werden. (<https://www.ynformed.nl/klantcases/overig/voorkomen-voortijdig-schoolverlaters/>) Es reicht nicht aus, wenn der Algorithmus angibt, welche SchülerInnen zur Risikogruppe gehören. Es geht vor allem darum, wie mit diesen Informationen umgegangen wird. Welche Kinder sind das? Was ist die Familiensituation? Gibt es einen Migrationshintergrund? Spielen Sprachbarrieren eine Rolle? Aus welcher Nachbarschaft kommen die SchülerInnen? Wer wird sich an die Kinder und ihre Eltern wenden: jemand, der mit der Erfahrungswelt der Jugendlichen verbunden ist, oder ein darin unerfahrener Beamter? Es handelt sich also auch um eine Vielzahl von Nicht-Datenaspekten, die jedes Mal eine Entscheidung erfordern. Es ist auch wichtig, dass eine Gemeinde offen darüber mit den Eltern, mit den Jugendlichen kommuniziert. Kann ich entscheiden, dass die Daten meines Sohnes nicht für diese Analyse verwendet werden? Gibt es eine Opt-out-Funktion? Wie sieht es mit der Datensicherheit aus? Angenommen, ein solcher Datensatz gelangt in die Hände eines Personalvermittlers. Stellen Sie sich vor, dieser hätte ein Modell, mit dem er nachweisen kann, dass jemand, der in der Risikogruppe der Schulabbrecher war, mit hoher Wahrscheinlichkeit schlecht im Job funktioniert. Er könnte dazu neigen, diese Kandidaten sofort abzulehnen. All diese Aspekte müssen auch von der Politik berücksichtigt werden.

Wie hilft die Utrecht Data School den Gemeinden?

Wir haben die Data Ethics Decision Aid (DEDA) entwickelt; das ist eine dialogische Folgenabschätzung für ethische Fallstricke in Datenprojekten (<https://dataschool.nl/deda/?lang=en>). DEDA ist ein Prozess, in dem politische EntscheidungsträgerInnen, Datenspezialisten, Datenschutzbeauftragte, Projektmanager und ExpertInnen gemeinsam am Tisch sitzen. Da werden verschiedene Aspekte des Projekts besprochen. Eine der ersten Fragen dieses Workshops ist: Wird ein Algorithmus verwendet? Die Daten-Spezialistin sagt: Natürlich. Die politischen

EntscheidungsträgerInnen sind dann oft überrascht. Die zweite Frage ist, ob sie den Algorithmus erklären können. Die Daten-SpezialistInnen können das, die EntscheidungsträgerInnen hingegen denken, sie müssten so etwas ja niemandem erklären. Das würde bedeuten, dass sie es weder im Gemeinderat diskutieren, noch den Bürgern oder Journalisten erklären können.

Durch diese Fragen wächst das Bewusstsein, dass alle MitarbeiterInnen in der Organisation tatsächlich über ein Basiswissen verfügen müssen. Die StadträtInnen müssen auch wissen, wie der Algorithmus funktioniert. Sie müssen nicht in der Lage sein, den Algorithmus zu programmieren, sondern die Logik verstehen, und das Entscheidungsmodell des Algorithmus erklären können. Während des Workshops werden alle Arten von Fragen und Fallstricken diskutiert. Dies hilft, gute Entscheidungen zu treffen, z.B. dass sie nicht alles wissen müssen, dass sie mit Daten vorsichtig sein müssen und nur die Daten sammeln, die sie wirklich brauchen, dass sie sie gut schützen und so weiter. Auf diese Weise entwickeln sie bessere Daten-Projekte, bei denen der Entscheidungsprozess hinsichtlich ethischer Fragen dokumentiert wird. Eine kritische Öffentlichkeit kann das in Frage stellen.

In den Niederlanden gibt es bereits mehrere Gemeinden, die für alle ihre Datenprojekte eine ethische Folgenabschätzung oder einen ethischen Schnellscan durchführen. Auch der Verband der niederländischen Gemeinden nutzt unser Tool, das jeder von unserer Website herunterladen kann.

Wie starten Kommunen eine Datenpolitik?

Pilotprojekte sind sehr wichtig, um neue Datenpraktiken zu erlernen. In kleinen Projekten werden so Fähigkeiten und Kompetenzen entwickelt. So arbeitete beispielsweise ein Team der Utrecht Data School mit einer niederländischen Gemeinde zusammen, um zu untersuchen wer welche Sozialhilfeleistungen erhält und um Indikatoren zu definieren, die angeben, warum manche Personengruppen zahlreiche Leistungen und andere nur wenige beziehen. Während der Untersuchung traten verschiedene Schwierigkeiten auf: Nicht alle Daten waren verfügbar, andere durften nicht verwendet werden, Regeln zur Erhebung und Nutzung von Daten waren unklar oder nicht ausreichend formuliert. Das führte dazu, dass auch Gemeindevorstand und Gemeinderat sich mit diesen Fragen befassen mussten. Sie beschlossen, das Projekt fortzusetzen, um daraus zu lernen. Der Datenschutzbeauftragten war involviert, eine ordnungsgemäße Verarbeitung der Daten wurde sichergestellt und ein Modus Operandi entwickelt, der für zukünftige Datenprojekte genutzt werden kann.

Oft sehen wir in Kommunen auch Initiativen, wie eine „Datenbrigade“ oder „Datenakademie“. Dabei handelt es sich um informelle Netzwerke von interessierten MitarbeiterInnen aus allen Verwaltungs-Ebenen, die gemeinsam diskutieren, wie sie besser mit Daten arbeiten können. Das ist vielversprechend. Aber es geht nicht ohne die Führung und Unterstützung durch das Top-Management und die politischen Entscheidungsträger.

Wissen die Gemeinden, welche Daten sie haben?

Nein, es liegen weitaus mehr Daten vor, als den meisten Kommunen bewusst ist. Ein Bericht hebt oft bestimmte Dinge hervor, aber der Datensatz kann auch Antworten auf andere Fragen liefern. Die erste Aufgabe besteht darin, andere Fragen zu den verfügbaren Datensätzen zu stellen. Zweitens muss die Qualität der Datensätze verbessert werden. Und dann ist da noch das Problem der Standards für den Datenaustausch. Jede Gemeinde sammelt Daten auf unterschiedliche Weise. Es

wäre sinnvoll, europäische offene Normen einzuführen, um den Austausch von Daten und Wissen zwischen den Kommunen zu erleichtern.

Sollten die Bewohner an Datenprojekten beteiligt werden?

Das hängt von der Art des Projekts ab. Es gab Versuche, die Bürger in Datenkompetenzprojekte einzubeziehen, aber diese waren nicht immer erfolgreich und wer daran teilnahm war nicht repräsentativ für die Bevölkerung. Ich habe ernsthafte Zweifel daran; ich bevorzuge einen Expertenansatz. Mir geht es mehr um offene Kommunikation und Rechenschaftspflicht, Rechenschaftspflicht gegenüber dem Gemeinderat und der vierten Macht. So wie eine Gemeinde über ihre Verkehrsplanung offen berichten muss, so muss sie auch transparent sein über die Daten, die sie hat und was sie mit Daten macht und was sie damit erreichen will. Dann können Gemeinderat und die Medien kritische Fragen stellen. Wenn eine Gemeinde einen Schulabbrecher-Algorithmus will, muss sie kommunizieren, von welchen Vorüberlegungen und welchen Daten man ausgehen will, wie der Algorithmus erstellt wurde, was mit den Daten passiert und so weiter. Dann gibt es vielleicht keinen Grund, um Bürger einzubeziehen. Aber in einem anderen Projekt kann das doch notwendig sein. In den Niederlanden gibt es einen Vorschlag, ‚intelligente‘ Straßenlaternen mit Kamera und Mikrofon einzusetzen, um abweichendes Verhalten auf der Straße zu signalisieren. Ein Algorithmus entscheidet, ob Auffälligkeiten vorliegen: Rennen, Schreien, Raufereien. Die AnwohnerInnen müssen sich dazu äußern können, da dies ihre Privatsphäre beeinträchtigt.

Was ist die größte Aufgabe für die Kommunen?

Sie sollten nicht in die Technokraten-Falle treten. Sie dürfen nicht vergessen, dass die Daten, die Analyseprozesse und die Algorithmen in unser demokratisches Wertesystem passen müssen, und dass Datenpraktiken rechenschaftspflichtig sind.

Utrecht Data School: www.dataschool.nl

2019