

# VOR DIE LAGE KOMMEN: PREDICTIVE POLICING IN DEUTSCHLAND

Chancen und Gefahren datenanalytischer Prognosetechnik und Empfehlungen für den Einsatz in der Polizeiarbeit

Dr. Tobias Knobloch

# **INHALT**

<b>Vorwort</b>	<b>3</b>
<b>Executive Summary</b>	<b>5</b>
<b>Hintergrund und Aufbau der Untersuchung</b>	<b>8</b>
<b>Grundlagen des Predictive Policing</b>	<b>9</b>
Definition und Ziel	<b>9</b>
Das Für und Wider in Kürze	<b>10</b>
<b>Internationales Lagebild</b>	<b>11</b>
Situation in Deutschland	<b>13</b>
<b>Kriminologische Grundlagen und Modellierung</b>	<b>16</b>
<b>Datenbasis</b>	<b>21</b>
<b>Institutionelle Implementierung</b>	<b>25</b>
<b>Wirkungsmessung und Evaluierung</b>	<b>27</b>
<b>Soziale Einbettung</b>	<b>30</b>
<b>Diskussion und Ausblick</b>	<b>33</b>
<b>Empfehlungen</b>	<b>37</b>
Nutzungszwecke und -grenzen vorab klar festlegen	<b>37</b>
Interdisziplinäre Expertise aufbauen und transdisziplinär arbeiten	<b>37</b>
Bereits bestehende Datenerfassungs- und -verarbeitungssysteme hinterfragen und eventuell modifizieren	<b>37</b>
Empirische Wirkungsforschung fördern und Evaluierungsergebnisse zugänglich machen	<b>38</b>
Systemvielfalt als Vorteil nutzen und Standards erwägen	<b>38</b>
Interne Kommunikation ernst nehmen und passgenaue Mitarbeiterschulungen anbieten	<b>39</b>
Aktive externe Kommunikation betreiben	<b>39</b>
<b>Literatur</b>	<b>41</b>

## VORWORT

Das digitale Zeitalter ist das Zeitalter der Algorithmen. Algorithmisch gesteuerte Entscheidungsprozesse werden in immer mehr Lebensbereichen eingesetzt – und zunehmend kritisch reflektiert. Der Wunsch nach mehr Transparenz und Kontrolle, mitunter auch nach Verboten, ist auch in der Bevölkerung weit verbreitet. Tatsächlich zeigen etliche Beispiele, dass algorithmische Entscheidungsfindung zu unfairen Ergebnissen oder systematischen Benachteiligungen führen können. Deswegen ist es zweifelsohne wichtig, sich mit ihren Schwächen und Risiken intensiv zu befassen. Allerdings bergen Algorithmen auch große Potenziale für die Lösung sozialer Herausforderungen. Sie können menschliche Entscheidungen an vielen Stellen sinnvoll ergänzen. Denn auch diese sind natürlicherweise mit erheblichen Schwächen behaftet, etwa weil Menschen von subjektiven Einschätzungen geleitet werden, oder weil die Komplexität einer Sachlage schlicht ihre zeitlichen oder kognitiven Kapazitäten übersteigt.

Unser Kooperationsprojekt „Algorithmen fürs Gemeinwohl“ verfolgt daher einen zu den verbreiteten eher kritischen Ansätzen komplementären Weg, indem es die gesellschaftlichen Chancen algorithmischer Entscheidungsfindung ins Zentrum des Interesses stellt. Anhand einer Analyse von exemplarischen Anwendungsbereichen, die auf das Leben der Menschen besonders großen Einfluss haben, soll geklärt werden, wie algorithmische Systeme per Design gemeinwohlorientiert gestaltet werden können. Ziel ist es, möglichst allgemeingültige Prinzipien zu formulieren, damit Algorithmen ihr gesellschaftliches Verbesserungspotenzial voll entfalten können.

Die Diskussionen über algorithmische Entscheidungsprozesse finden bislang meist in abgeschlossenen Kreisen mit relativ wenig transdisziplinärer Zusammenarbeit statt. Wir wollen daher in drei Themenfeldern, von denen vorausschauende Polizeiarbeit (Predictive Policing) das erste ist,<sup>1</sup> über Sektorengrenzen hinweg die Problemanalyse schärfen und den unmittelbaren Austausch über gemeinwohlorientierte Gestaltungsprinzipien ermöglichen. Im April 2018 kamen deshalb Vertreter aus vier der sechs in diesem Bereich aktiven Bundesländer sowie Entwickler:innen, Wissenschaftler:innen und Vertreter:innen der Zivilgesellschaft zu einem ersten intersektoralen Workshop zum Thema Predictive Policing in Deutschland zusammen. Die in diesem Papier aufgezeichneten Erkenntnisse und Empfehlungen gehen einerseits auf die Workshop-Ergebnisse und andererseits auf Recherchen und weitere Gespräche mit Experten:innen zurück.<sup>2</sup>

Insgesamt ist positiv hervorzuheben, wie methodenbewusst und reflektiert Behörden in Deutschland die jeweiligen Instrumente getestet, implementiert und immer wieder kritisch geprüft und weiterentwickelt haben. Predictive-

---

<sup>1</sup> Es folgen noch in diesem Jahr Robo Recruiting (verstärkter Technologieeinsatz in der Personalauswahl) und eHealth (Algorithmen im Gesundheitsbereich). 2019 werden die Erkenntnisse in einer Publikation zusammengeführt und bei einer Abschlussveranstaltung mit einem erweiterten Expertenkreis diskutiert.

<sup>2</sup> Eine Dokumentation der Ergebnisse dieses Workshops findet sich unter [http://www.stiftung-nv.de/sites/default/files/ws-predpol\\_april2018\\_ergebnisse\\_0.pdf](http://www.stiftung-nv.de/sites/default/files/ws-predpol_april2018_ergebnisse_0.pdf)

Policing-Systeme gelten überall als Erweiterung des Instrumentenkastens polizeilicher Präventionsarbeit zur Unterstützung kriminalistischer Methoden. Es gehe letztlich darum, unter Ausnutzung des technisch Möglichen „vor die Lage zu kommen“, wie einer der Beteiligten sagte, also den Kriminellen gewissermaßen einen Schritt voraus zu sein. Das bestätigt auf der Ebene des professionellen polizeilichen Selbstverständnisses die These, dass Predictive-Policing-Systeme – jedenfalls bis auf Weiteres – keine Substitution und keinen Paradigmenwechsel, sondern eine Ergänzung polizeilicher Praxis bedeuten.

Das starke mediale und politische Interesse am Thema Predictive Policing scheint angesichts dessen beinahe überzogen zu sein. Das liegt sicher auch daran, dass die Nutzung von prädiktiven Algorithmen im Zuge der allgemein großen Aufmerksamkeit für alles, was mit Künstlicher Intelligenz (KI) zu tun hat, derzeit noch zu häufig vor dem Spiegel popkultureller Erzeugnisse diskutiert wird, statt nüchtern im Verhältnis zu den tatsächlichen Gegebenheiten behördlicher Praxis. Aber auch die allzu passive Außenkommunikation der Polizeibehörden mag dazu beitragen, dass häufig mehr dahinter vermutet wird als tatsächlich dahintersteckt.<sup>3</sup> Dieses Papier soll zu einer Versachlichung der Debatte beitragen und den Einsatz algorithmischer Systeme in der Polizeiarbeit ein Stück weit demystifizieren. Wir möchten politischen Entscheidungsträgern eine Orientierungshilfe bieten, bereits umsetzenden Behörden Verbesserungsvorschläge unterbreiten und in die Umsetzung neu einsteigenden Behörden ein Hilfsmittel für ihre Planungen an die Hand geben.



A handwritten signature in black ink that reads "Anna Wohlfahrt".

Anna Wohlfahrt  
Sprecherin des Vorstands  
Stiftung Neue Verantwortung e. V.



A handwritten signature in black ink that reads "R. Müller-Eiselt".

Ralph Müller-Eiselt  
Senior Expert  
Bertelsmann Stiftung

---

<sup>3</sup> Vgl. Knobloch 2018b und Abschnitt 7 unten.

## **EXECUTIVE SUMMARY**

Seit etwa drei Jahren wird Predictive Policing in Deutschland praktiziert. Sechs Bundesländer nutzen entsprechende Software meist, um städtische Gebiete, in denen am wahrscheinlichsten Wohnungseinbrüche, Gewerbeeinbrüche oder Kfz-Diebstähle stattfinden werden, zu identifizieren. Die Erfahrungen zeigen erstens, dass Predictive Policing, wie es derzeit in Deutschland praktiziert wird, keinen Eingriff in die Bürgerrechte darstellt. Zweitens wird deutlich, dass Predictive Policing abseits der Verhinderung von Verbrechen, was sich als Wirkung kaum nachweisen lässt, so viel positives Potential entfaltet, das es sich durchaus lohnt, diesen Ansatz mit Bedacht weiter zu verfolgen.

Die unterschiedlichen, in Deutschland im Einsatz befindlichen – lizenzierten und selbst entwickelten – Systeme verarbeiten ausschließlich ortsbezogene Informationen, keine personenbezogenen. Wie gut das Einkreisen von Kriminalitäts-Hotspots funktioniert, weiß man nicht, weil die Wirkung sehr schwer messbar ist. Andere mögliche Einflussfaktoren lassen sich nicht so isolieren, dass man Wirkungen unmittelbar auf Predictive Policing zurückführen kann. Obwohl von 2015 bis 2017 in Baden-Württemberg die weltweit erste wissenschaftliche Begleitevaluation des Testbetriebs eines Predictive-Policing-Systems stattgefunden hat, lässt sich kaum mehr als vermuten, dass solche Systeme einen positiven Einfluss auf die Entwicklung von Wohnungseinbruchszahlen haben, indem sie eine gezieltere polizeiliche Präsenz und Prävention ermöglichen. Das macht es zweitens nicht nur schwierig, die Ausgaben der öffentlichen Hand für entsprechende Systeme zu rechtfertigen. Es erschwert den Anwendern der Systeme auch, den Erfolg ihrer Arbeit zu sehen. Denn Erfolg in der Logik des Predictive Policing heißt ja, dass gerade kein Einbruch oder Diebstahl stattfindet. Für Beamt:innen des Streifendienstes kann das frustrierend sein.

Um diesem drohenden Frust entgegen zu wirken, muss man Predictive Policing als das beschreiben, was es ist, nämlich als ein technisches Hilfsmittel zur Unterstützung der polizeilichen Intuition und des kriminalistischen Gespürs. Beides wird nach wie vor gebraucht, um ein Lagebild und einen darauf abgestimmten Einsatzplan zu erstellen – womit wir bei den positiven Effekten von Predictive Policing sind. Die Verwendung von Algorithmen zwingt dazu, bisherige Erkenntnisquellen und Vorgehensweisen auf den Prüfstand zu stellen. Um Algorithmen einsetzen zu können, muss man die Ziele seines Handelns benennen können. Und man muss sagen können, woraus man seine Erkenntnisse bezieht und wie man normalerweise vorgeht, um seine Ziele zu erreichen. Mit anderen Worten: Alles, was intransparent ist, was un- oder nur halb bewusst geschieht, muss explizit gemacht werden.

Um Predictive Policing als ein wirksames Zusatzinstrument einsetzen zu können, müssen Behörden beispielsweise die Datenquellen, aus denen die Software gespeist wird, auf den Prüfstand stellen: Welche Güte haben Polizeidaten und wie vollständig sind sie? Was lässt sich an Vorgangsbearbeitungssystemen – durchaus auch im Hinblick auf den Bedienkomfort für die Mitarbeiter:innen – verändern, um Qualität und Vollständigkeit der Datenbasis zu verbessern? Welcher weiterer Informationsquellen (z. B. soziodemografische Daten und Geodaten) bedarf es, um zu den gewünschten Erkenntnissen zu gelangen und entsprechend robuste Modelle zu erstellen? Auch die Arten der internen Zusammenarbeit an Lagebild, Einsatzplanung und Bewertung derselben müssen offengelegt und gegebenenfalls hinterfragt werden, wenn Software zum Einsatz kommen soll. Denn Maschinen brauchen, wenn sie uns unterstützen sollen, in weiten Teilen nach wie vor explizite Anweisungen. Bei Behörden, deren Auftrag es ist, die Sicherheit der Bevölkerung und ihres Besitzes zu gewährleisten, ist das ein guter Ansatzpunkt zur Verbesserung ihrer Arbeit – und zwar ganz unabhängig davon, wie man zu Predictive Policing steht.

Darüber hinaus stellen sich auch solche Fragen, die das polizeiliche Selbstverständnis tangieren und gesellschaftlich sehr relevant sind: Wenn empirische Daten die Realität abbilden und daraus Prognosen erstellt werden, die für die Polizei handlungsleitend sind, dann wird die Zukunft immer ein Stück weit der Vergangenheit gleichen, egal wie die Realität der Zukunft tatsächlich aussieht. Ferner können sich aus der datenanalytischen Automatisierung weder intendierte noch für die Beteiligten unmittelbar ersichtliche Effekte ergeben: Eine stärkere Polizeipräsenz kann zu einem besseren oder zu einem schlechteren Sicherheitsempfinden beitragen. Und sie kann schlussendlich sogar dazu führen, dass Wohnquartiere als unsichere Gebiete mit hoher Polizeipräsenz ökonomisch abgewertet werden. Behörden, die Predictive Policing als Erweiterung ihres Werkzeugkastens in Betracht ziehen, müssen sowohl realistisch die Frage beantworten, was sie damit erreichen wollen, als auch sich Aufschluss darüber geben, was sie damit unfreiwillig bewirken könnten.

Aus eben diesem Grund ist eine der am Ende dieses Papiers ausgesprochenen Empfehlungen als eine Art Mindeststandard zu betrachten: dass an

der Entwicklung und Einführung von Predictive-Policing-Systemen schon frühzeitig interdisziplinäre Teams mitwirken. Hier braucht es nicht nur kriminalistischen, kriminologischen und informationstechnischen Sachverstand, sondern auch Datenschutz-Expertise, rechtliches, soziologisches, sozialpsychologisches und ethisches Wissen. Eine weitere wichtige Empfehlung lautet, dass die Behörden untereinander nicht nur ihre Erfahrungen und Erkenntnisse austauschen, sondern beispielsweise auch hochwertige Trainingsdaten zum Anlernen von Systemen. Während eine gewisse Systemvielfalt besonders in der Frühphase eines Trends durchaus zu begrüßen ist, wäre eine Harmonisierung oder sogar Standardisierung über Bundesländergrenzen hinweg besonders im Bereich der Datenerfassung wünschenswert. Andernfalls ist die Güte unterschiedlicher Systeme nämlich gar nicht vergleichbar.

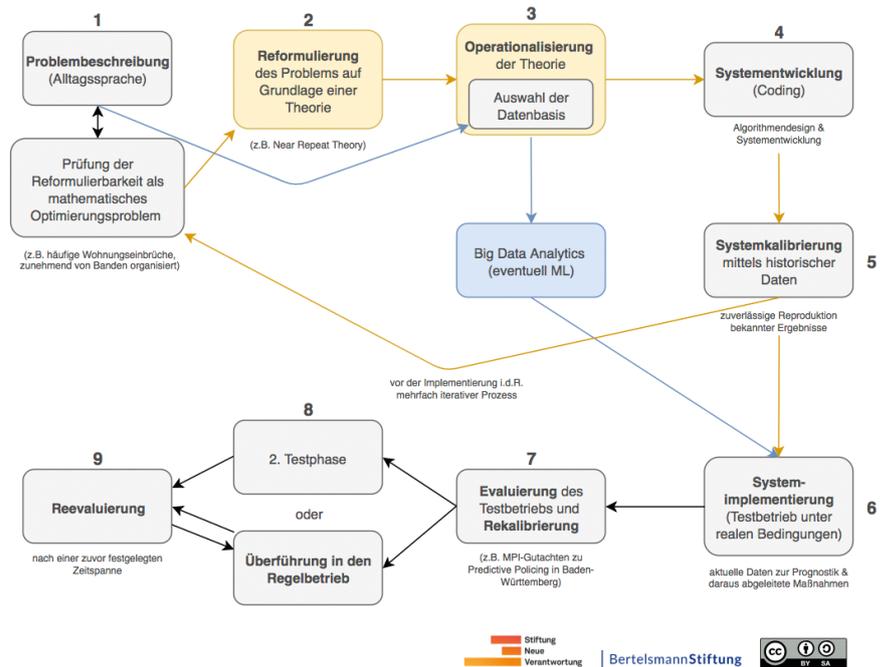
Durch den zunehmenden Einsatz Datenanalyse-gestützter staatlicher Überwachungstechniken wird sich die Polizeiarbeit insgesamt von der kriminalistischen Ermittlung hin zur präventiven Diagnostik entwickeln. Dadurch ist das rechtsstaatliche Prinzip der Unschuldsvermutung nicht unmittelbar in Gefahr. Potentiell gefährlich ist aber die Verschränkung von Techniken und Praktiken der Überwachung, wie wir sie heute bereits in den angelsächsischen und in einigen asiatischen Ländern sehen: Wenn überall Überwachungskameras mit Gesichtserkennung und Kennzeichenleser installiert sind, wenn jeder Polizeiwagen an eine lernende Prognosesoftware angeschlossen ist – dann ist es im Grunde gesellschaftlich nicht mehr hinnehmbar, dass überhaupt noch irgendwo ein Verbrechen geschehen kann. Es darf darüber aber nicht in Vergessenheit geraten, dass die Möglichkeit von Verbrechen die Kehrseite der Freiheit ist. Es lohnt sich daher, dass wir uns in Deutschland jetzt intensiv mit Predictive Policing auseinandersetzen und die eingesetzten Systeme kritisch beleuchten.

# HINTERGRUND UND AUFBAU DER UNTERSUCHUNG<sup>4</sup>

Der Begriff „vorausschauende Polizeiarbeit“ (oder „Predictive Policing“, nachfolgend „PP“ abgekürzt) bezeichnet verschiedene algorithmische Systeme, die gemein haben, dass sie zur Berechnung von Wahrscheinlichkeiten für das Auftreten von Verbrechen an bestimmten Orten zu bestimmten Zeiten genutzt werden, um mit entsprechenden polizeilichen Maßnahmen (präventiv oder repressiv) darauf reagieren zu können. Besondere Popularität haben solche Systeme seit Anfang der 2010er Jahre in den USA erlangt, seit einigen Jahren kommen sie auch in Deutschland zum Einsatz oder werden in Pilotprojekten getestet.

Bevor es um die spezifischen Chancen und Probleme von PP in Deutschland geht, benötigen wir zunächst eine begriffliche Eingrenzung, einen Überblick über den Ansatz und die Debatte darüber (Kapitel 2). Ferner wird als Ausgangsbasis für die anschließende Analyse eine Übersicht über den internationalen und nationalen diesbezüglichen Status quo (Kapitel 3) vorangestellt. Die weitere Gliederung der Kapitel 4 bis 8 orientiert sich an den Schwerpunkten des intersektoralen Workshops im April 2018.<sup>5</sup> Diese Schwerpunkte wiederum ergaben sich aus der Zerlegung des Entwicklungsprozesses algorithmischer Analyse- und Entscheidungssysteme im Allgemeinen und von Predictive-Policing-Software im Speziellen. Abbildung 1 stellt diesen Prozess dar:

**Abb. 1: Entwicklungsprozess algorithmischer Analyse- und Entscheidungssysteme.**  
Quelle: eigene Darstellung.



<sup>4</sup> Ein herzlicher Dank für viele kluge Kommentare und wertvolle Hinweise während der Entstehung dieses Papiers gilt meinen Projektkolleg:innen bei der Bertelsmann Stiftung, Carla Hustedt und Ralph Müller-Eiselt. Ein ebensolcher Dank gebührt auch Herrn Joachim Eschemann, der das Projekt SKALA in Nordrhein-Westfalen geleitet hat und inzwischen das Referat „Kriminalitätsangelegenheiten“ im Ministerium des Innern des Landes leitet.

<sup>5</sup> Den Teilnehmern dieses Workshops sei an dieser Stelle noch einmal herzlich für ihre Beteiligung und Unterstützung auch nach dem Termin gedankt!

Ausgangspunkt ist eine alltagssprachliche Problembeschreibung (1). Dann wird dieses Problem unter Zuhilfenahme theoretischer Annahmen als mathematisches Optimierungsproblem reformuliert; die eigentliche Modellierung findet im Bereich der gelben Kästchen statt (2 und 3). Wurde das Modell in Code gefasst und kalibriert (4 und 5), kann die Implementierung angegangen werden (6). Dabei finden immer wieder Funktionstests und Anpassungen statt, im Idealfall auch noch während des Regelbetriebs (7 bis 9). Der alternative Weg, der im Schaubild von der Auswahl der Datenbasis über den blauen Kasten („Big Data Analytics“) zur Systemimplementierung führt, deutet eine heute vielfach beschrittene Abkürzung des klassischen Softwareentwicklungsweges an: Hier werden eben nicht theoriegeleitet kausale Zusammenhänge hergestellt und automatisiert verknüpft, sondern Korrelationen in Daten gesucht, und diese Korrelationssuche wird sukzessive optimiert. Wir gehen hier, wenn nicht anders angemerkt, vom „Normalverfahren“ der Entwicklung genuiner Software und nicht von der Mustererkennung in großen, z. T. unstrukturierten Datenmengen mittels bestehender, mehr oder weniger standardisierter Instrumente aus.<sup>6</sup>

Der Übersichtlichkeit halber haben wir die eher technischen Prozessschritte für dieses Papier so modifiziert, dass sie dessen Zielsetzung gerecht werden: In Abschnitt 4 geht es um die Modellierung, in Abschnitt 5 um die Datenbasis, in Abschnitt 6 um die Verankerung in den Polizeibehörden, in Abschnitt 7 um die Evaluierung und in Abschnitt 8 um die soziale Einbettung von PP-Systemen. Es schließt sich eine kritische Diskussion der Befunde an, bevor am Ende einige Empfehlungen für den Einsatz von PP-Systemen folgen.

## GRUNDLAGEN DES PREDICTIVE POLICING

### DEFINITION UND ZIEL

Von „vorausschauender Polizeiarbeit“ (oder „Predictive Policing“) spricht man dann, wenn Polizeibehörden Software einsetzen, die auf der Grundlage einer Theorie aus historischen Kriminalitätsdaten (meist im Zusammenspiel mit aktuellen, kriminologisch relevanten Lagedaten) Prognosen zukünftiger Verbrechensareale oder zukünftiger Straftäter erstellen und diese für die Einsatzplanung<sup>7</sup> nutzen. PP-Systeme berechnen „die statistische Wahrscheinlichkeit ausgewählter Delikte auf Basis von raum- und zeitbezogenen Daten zur Prognose von Kriminalitätsrisiken“<sup>8</sup> (z. B. für Einbrüche oder andere Straftaten in bestimmten Quartieren). Dieser Aspekt ist äußerst wichtig im Hinblick auf die Erwartungen an die Effekte des Einsatzes solcher Systeme: Im Grunde geschieht damit im Wesentlichen nichts anderes, als

<sup>6</sup> Natürlich gibt es Überschneidungen, zumal die Mustererkennung bei neu zu entwickelnden Systemen und im Zusammenhang mit der Debatte um Künstliche Intelligenz eine immer größere Rolle spielt; vgl. Büнау 2018.

<sup>7</sup> Hierbei ist die Art der mit dieser Methode zu planenden Einsätze (präventiv oder repressiv) in diesem allgemeinen Zusammenhang zunächst zu vernachlässigen. Wichtig wird sie später, wenn es um die intendierten und tatsächlich erzielten Wirkungen des PP geht.

<sup>8</sup> LKA NRW 2018b: 19.

dass im Bereich geringer Wahrscheinlichkeiten signifikante Abweichungen von der Standardwahrscheinlichkeit ermittelt und auf einer Karte sichtbar gemacht werden.

Doch wenn das verlässlich funktioniert, kann damit bereits eine wesentliche Effizienzsteigerung der Polizeiarbeit im Bereich der Straftatenverhütung und -prävention erzielt werden. In der Leistungsbeschreibung des LKA Nordrhein-Westfalen für das Projekt *Predictive Policing* vom 9.3.2015 (seit 2016 *SKALA* genannt) wird das unmissverständlich so formuliert: „Zweck ist somit eine strategische und zielgerichtete Polizeiarbeit, die auf Basis bekannter, kriminalitätsrelevanter Faktoren aufkommende Brennpunkte frühzeitig identifiziert. Dadurch soll ein ressourceneffizienter Einsatz der Polizeikräfte und eine Senkung der Kriminalitätshäufigkeit erreicht werden.“<sup>9</sup> Dieses Motiv ist in den deutschen Bundesländern, die bereits PP-Systeme einsetzen, prävalent,<sup>10</sup> obwohl die Kriminalitätsrate in Deutschland insgesamt rückläufig ist.<sup>11</sup>

## DAS FÜR UND WIDER IN KÜRZE

Vom Einsatz von PP-Systemen verspricht man sich also vor allem einen effektiveren Einsatz von knappen Ressourcen, letztlich also ein „Mehr mit Weniger“. Wäre diese in PP-Systeme gesetzte Hoffnung berechtigt, und wären die Vermarktungsbehauptungen der Systemhersteller<sup>12</sup> zutreffend, so läge der gesellschaftliche Vorteil auf der Hand: mehr öffentliche Sicherheit mit weniger öffentlichen Mitteln durch den Einsatz von State-of-the-Art-Technologie. Die zugrundeliegende Motivation und das entsprechende Narrativ haben jedoch aus unterschiedlichen Richtungen Anlass zu grundsätzlicher Kritik gegeben. Neben Befürchtungen des Abbaus von Stellen bei bereits notorisch überlasteten Polizeibehörden und des Verschwindens beziehungsweise der drastischen Veränderung von Job-Profilen ist auch angeführt worden, dass der Einsatz prädiktiver Algorithmen dem vor allem von Technologieunternehmen und Politikern jeweils im Eigeninteresse befeuerten Trend zu sogenannten „Technofixes“ Vorschub leiste.<sup>13</sup> „Technofix“ ist eine pejorative Bezeichnung für den Einsatz von mehr oder weniger avancierter Technik an falschen Stellen beziehungsweise in überzogenem Maße. Meist wird von Kritikern ins Feld geführt, dass gesellschaftliche Problemlagen zu komplex seien, um sie mit einfachen technischen Mitteln auflösen zu können. So werde eine einfache Schuldnerberatungssoftware das Problem steigender Privatverschuldung allein nicht aus der Welt schaffen. Ebenso könne Prognosesoftware das gesellschaftliche Phänomen der Kriminalität in seiner Komplexität nicht so weit reduzieren, dass es fortan leicht zu beherrschen wäre.

<sup>9</sup> Zitiert nach LKA NRW 2018b: 4.

<sup>10</sup> Vgl. Egbert 2018, Abschnitte 3 und 4.

<sup>11</sup> 2017 hat es die niedrigste Zahl an verübten Straftaten seit 1992 gegeben; vgl. BMI 2018.

<sup>12</sup> Vgl. z.B. PredPol 2018.

<sup>13</sup> Für den politischen Umgang mit dem Thema Predictive Policing wird dies schon in Egbert 2018 gezeigt; für den gesamtgesellschaftlichen Bereich, insbesondere ausgehend von den großen Internetkonzernen des Silicon Valley, wird die Weltsicht der Technofixes (ohne die bei anderen Fundamentalkritikern wie Evgeny Morozov anzutreffende Polemik) in Hankey & Tuszinski 2017 dargestellt und kontextualisiert.

Untersuchungen von verschiedener Tiefe und Unabhängigkeit haben in der Tat wiederholt gezeigt, dass PP allenfalls eine geringe positive Wirkung auf die Effektivität polizeilicher Maßnahmen zur Verbrechensvorbeugung und -bekämpfung hat.<sup>14</sup> Dabei ist allerdings zu beachten, dass die Interferenz zahlreicher (möglicher) Einflussfaktoren es hier – wie übrigens bei der Evaluierung jeglicher Polizeitechnik<sup>15</sup> – unmöglich macht, die tatsächliche Effektivität von PP direkt zu messen. Vielmehr bedürfte es dazu aufwendiger quasi-experimenteller Langzeitstudien, die bislang noch nicht stattgefunden haben. Ausführlicher dazu im siebten Abschnitt unten.

## INTERNATIONALES LAGEBILD<sup>16</sup>

Vorausschauende Polizeiarbeit hat in den 1990er Jahren in New York unter dem Druck massiver, die öffentliche Sicherheit stark einschränkender Kriminalität und mit einer Verbesserung der Datenaufzeichnung und -auswertung durch die Einführung des Statistikprogramms *CompStat* begonnen. Es handelt sich also um ein noch junges Phänomen – zumal dann, wenn man in Rechnung stellt, dass das erste regelrechte PP-System im oben definierten Sinne erst im Jahr 2011 von der Foothill Division des Los Angeles Police Departments implementiert wurde.<sup>17</sup> Das in Los Angeles eingesetzte und bis heute international am weitesten verbreitete System *PredPol*<sup>18</sup> ist von dem Computerwissenschaftler George Mohler und dem Anthropologen Jeffrey Brantingham an der University of California auf der Basis einer Software zur Vorhersage von Erdbeben entwickelt worden. Seitdem sind erstens Ableger und Nachahmer von *PredPol*<sup>19</sup>, zweitens Erweiterungen<sup>20</sup>, drittens neue kommerzielle Produkte mit einem anderen theoretischen Ansatz<sup>21</sup> und viertens, zumal in Deutschland, Eigenentwicklungen von Polizeibehörden entstanden.

Mittlerweile sind algorithmische Systeme zur Unterstützung der polizeilichen Arbeit – meist ohne Wissen der Öffentlichkeit – weltweit bereits recht weit verbreitet. Die Debatten über PP ähneln sich in den einzelnen Ländern zwar, aber es sind unterschiedliche Schwerpunkte zu verzeichnen – vor allem, was die kritischen Aspekte betrifft. Dazu im Folgenden lediglich schlaglichtartig einige Beispiele.

14 Vgl. dazu Abschnitt 7 unten.

15 Vgl. Manning 2014.

16 Da sich nicht nur die Situation, sondern auch die rechtlichen Möglichkeiten in anderen Ländern z. T. drastisch von den hiesigen Verhältnissen und Möglichkeiten unterscheiden, wird der internationale Teil der Darstellung bewusst kurz gehalten. Einen detaillierten, wenn auch schon älteren Überblick über die Situation in den USA (sowie einen nach wie vor sehr guten Gesamtüberblick über das Thema) bietet das RAND-Forschungspapier Perry et al. 2013.

17 Natürlich versucht die Polizei aber schon seit dem 19. Jahrhundert anhand von Daten zu verstehen, wie sich Kriminalität räumlich ausbreitet.

18 Laut Fasman 2018 (Abschnitt „The promise and peril of big-data-justice“) wird es von über 50 Polizeidienststellen weltweit eingesetzt.

19 Dazu zählt beispielsweise das in Süddeutschland und der Schweiz eingesetzte deutsche System *Precobs*.

20 Hier kann etwa die Gefährderdatei „Heat List“ als Ergänzung zu *PredPol* in Chicago angeführt werden; vgl. Montag 2016: 3.

21 Vgl. etwa das in Abschnitt 4 kurz vorgestellte Produkt *HunchLab*.

In den **USA** geht es vor allem um das ethnische Diskriminierungspotenzial in der Datenbasis von PP-Systemen. Vermutet wird (und teilweise auch belegt werden konnte) ein sich fortwährend selbstverstärkender Effekt, der die ohnehin gegebene polizeiliche Diskriminierung der afro-amerikanischen und lateinamerikanischen Bevölkerungsanteile verstärkt. Einen solchen Effekt konnte die US-amerikanische Bürgerrechtsorganisation **ProPublica** beispielsweise bei einer Software nachweisen, die von Richtern in den USA dazu genutzt wird, das Rückfälligkeitsrisiko bei Bewährungsentscheidungen einzuschätzen.<sup>22</sup> Bei personenbezogenen PP-Systemen, wie sie in den Vereinigten Staaten genutzt werden, treten ähnliche Effekte schlicht dadurch auf, dass historische Kriminalitätsdaten die Basis für die zukünftige Polizeistrategie sind. Bei ortsbezogenen PP-Systemen besteht zwar kein Risiko der individuellen Diskriminierung, dafür gibt es die Befürchtung, dass ihr Einsatz zu einer Überrepräsentanz von Polizeikräften in ärmeren Gegenden mit einem hohen Anteil an ohnehin marginalisierten Bewohnern führen könnte. Dies ist selbst dann der Fall, wenn diese Systeme nicht auf Daten über Einkommen und Hautfarbe zugreifen, da diese Faktoren in ethnisch stark segregierten Gebieten oft mit der Postleitzahl korrelieren; vermeintlich ausgeschlossene Diskriminierung kommt so durch die Hintertür wieder herein.<sup>23</sup> Die Kritik in **Großbritannien** beschäftigt sich ebenfalls mit der möglichen Verstärkung bereits bestehender Diskriminierung durch PP. Ethnische Diskriminierung wird hier aber etwas weniger stark diskutiert als die allgemeineren sozial-ökonomischen Diskriminierungspotenziale.<sup>24</sup> In **Südafrika** bestehen Vorbehalte gegenüber PP, die davon ausgehen scheinen, dass die traditionelle Polizeiarbeit dadurch mehr oder weniger obsolet wird und ein Verlust von Arbeitsplätzen im großen Stil damit einhergeht.<sup>25</sup> In der **Schweiz** wurde zuletzt über die Häufigkeit von falschen Verdächtigungsfällen bei der Auswertung von Gefährderdatenbanken mittels des aus Deutschland stammenden Programms *DyRiAS* berichtet. „Polizei-Software verdächtigt zwei von drei Personen falsch“, lautet die Überschrift eines entsprechenden Beitrags des Schweizer Radios und Fernsehens (SRF).<sup>26</sup>

Bedenken aus Gründen des Schutzes der **Privatsphäre** (Privacy Concerns) nehmen vielerorts zu, gerade in den USA, wo in manchen Städten, beispielsweise Los Angeles oder Chicago, sehr viele unterschiedliche Datenquellen angelegt, miteinander verknüpft und ausgewertet werden. Kritisch hinterfragt wird hier, dass der „Datenhunger“ der Behörden sich verselbstständigt und, wenn überhaupt, nur noch auf dem Klageweg zu stoppen ist.<sup>27</sup> Generell bemängeln Vertreter:innen der Zivilgesellschaft die fehlende funktionale Transparenz der verwendeten Algorithmen (insbesondere bei kommerziellen PP-Systemen), ohne die eine unabhängige Überprüfung nicht möglich ist.<sup>28</sup> Dies sowie die Erzeugung von Fehlern und die Durchführung unangemessener Einsätze aufgrund von Verzerrungen in den Datengrundlagen

22 Vgl. Angwin et al. 2016.

23 Vgl. spezifisch Shapiro 2017, S.459; allgemein dazu Angwin & Larson 2016.

24 Vgl. Rowe 2018.

25 Vgl. Matinde 2015.

26 Stadler 2018

27 Vgl. Kartheuser 2018.

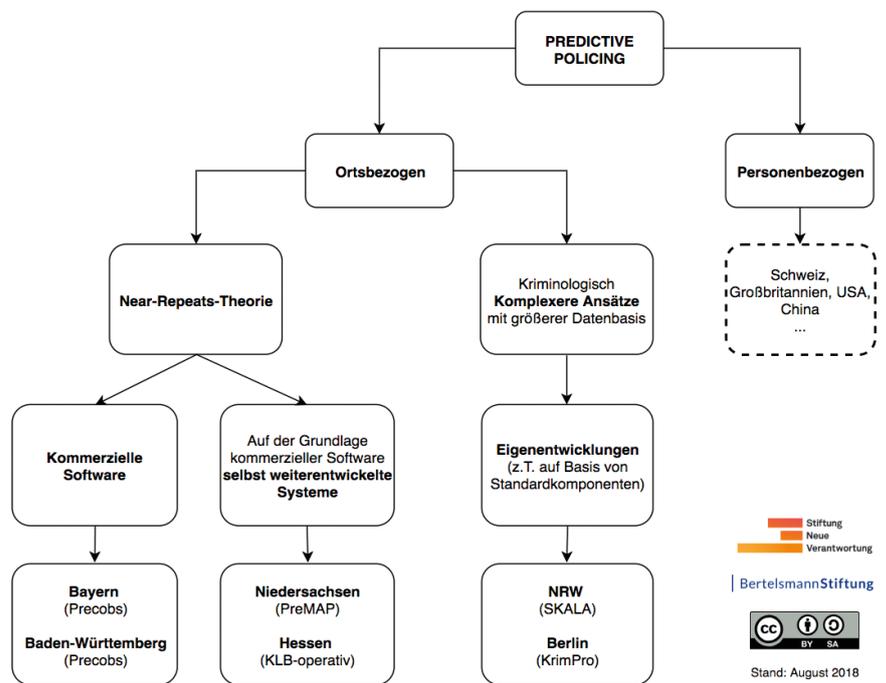
28 Vgl. Kartheuser & Jaume-Palasi 2018.

sind die Hauptansatzpunkte der journalistischen, politischen und zivilgesellschaftlichen Kritik an PP-Systemen in der westlichen Hemisphäre.<sup>29</sup>

## SITUATION IN DEUTSCHLAND

Sechs deutsche Bundesländer<sup>30</sup> setzen derzeit fünf verschiedene algorithmische Systeme zur ortsbezogenen vorausschauenden Polizeiarbeit ein, bislang vornehmlich zur Prävention von Wohnungseinbrüchen. Während die Bundesländer Bayern und Baden-Württemberg mit der Software *Precobs* (Pre Crime Observation System) des Oberhausener Unternehmens Institut für musterbasierte Prognosetechnik (IfmPt) GmbH arbeiten, nutzen Nordrhein-Westfalen und Berlin vollständig selbst entwickelte Systeme. Die Bundesländer Hessen und Niedersachsen haben wiederum Standard-Software für ihre Zwecke weiterentwickelt. Abbildung 2 bietet einen Überblick über den Stand der Dinge:

**Abb. 2: Situation des Predictive Policing in Deutschland.**  
Quelle: eigene Darstellung.



Die Unterscheidung zwischen ortsbezogenen und personenbezogenen Systemen ist grundlegend. Bei ortsbezogenen Systemen werden ausschließlich raum-zeitliche Variablen ohne Personenbezug berücksichtigt. Bei personenbezogenen Systemen werden zusätzlich personenbezogene Merkmale einbezogen. In Deutschland kommen ausschließlich ortsbezogene Systeme

29 In Asien ist das kulturelle Verhältnis zu staatlicher Überwachungstechnologie ein grundlegend anderes. Südkoreaner etwa bekommen, wenn sie ins Ausland reisen, automatisch eine Nachricht ihres Außenministeriums auf ihrem Smartphone mit Informationen und ggf. Warnungen zum Gastland angezeigt. Auch in Singapur und China werden informationstechnische Eingriffe in die Privatsphäre seitens der Bürger – nolens volens, sollte man vielleicht hinzufügen – eher hingenommen. Die kritische Distanz zum behördlichen Einsatz solcher Technologien ist hier im Vergleich zu Europa und Amerika unterausgeprägt.

30 Abweichend zur kartenbasierten Darstellung in Diehl & Kartheuser 2018, wird die Freie und Hansestadt Hamburg hier nicht berücksichtigt, weil in Hamburg lediglich der potentielle Nutzen von PP seitens der kriminologischen Forschungsstelle des LKA exploriert wird, sich jedoch kein PP-System in einer Testphase oder gar im Produktivbetrieb befindet. Der finale Bericht zum Hamburger Forschungsprojekt ist für Ende 2018 angekündigt, vgl. LKA Hamburg 2018.

zum Einsatz, in anderen Ländern gibt es sowohl orts- als auch personenbezogene Systeme, bisweilen auch in Mischformen. Die Unterscheidung nach kriminologischer Komplexität wird in Abschnitt 4 erläutert.

Das in Süddeutschland genutzte System *Precobs* ist dem US-amerikanischen Marktführer *PredPol* ähnlich. Es basiert auf dem Near-Repeats-Modell, welches besagt, dass bei einer Straftat in einem Gebiet die Wahrscheinlichkeit für Folgetaten steigt. Ähnlich wie *PredPol* nutzt auch *Precobs* für die Erstellung der Prognosen nur wenige Datenpunkte mit Bezug zur Tat selbst. Wie die Abschnitte 3 und 4 zeigen, bedeutet dieser überschaubare Mechanismus auf Grundlage einer übersichtlichen Datenbasis jedoch nicht automatisch, dass diese Systeme per se als völlig unproblematisch einzustufen sind – zumal sie stetig weiterentwickelt werden, (potenziell) mehr Informationsquellen integrieren und auf andere Deliktarten ausgeweitet werden. Die Bundesländer Niedersachsen (*PreMAP*) und Hessen (*KLB-operativ*) haben bewährte Statistikprogramme an ihre Zwecke und Bedürfnisse im Bereich des PP angepasst. Einen ähnlichen Weg über die Standardsoftwarekomponenten ist Nordrhein-Westfalen gegangen, hat dabei aber das eigenständige *System zur Kriminalitätsauswertung und Lageantizipation* (SKALA) mit Schnittstellen zu verschiedenen anderen Informationssystemen geschaffen. SKALA nutzt über die polizeiliche Datenerfassung hinausgehende Informationen, wie infrastrukturelle Gegebenheiten und demographische Daten, und das Deliktsspektrum umfasst nicht nur Wohnungseinbrüche<sup>31</sup>, sondern auch Gewerbe- und Autodiebstähle. Auch sonst geht die Eigenentwicklung SKALA in seinem Anwendungsbereich recht weit über die am weitesten verbreiteten Systeme hinaus.<sup>32</sup> Das Berliner System *Kriminalitätsprognose Wohnraumeinbruch* (KrimPro) ist ebenfalls eine Eigenentwicklung und setzt auf dem *Polizeilichen Landessystem zur Information, Kommunikation und Sachbearbeitung* (POLKIS) auf. In allen anderen Bundesländern befinden sich PP-Systeme weder im operativen Einsatz noch in Testverfahren, allenfalls wird deren Einsatz erwogen oder – wie in Hamburg<sup>33</sup> – wissenschaftlich erforscht.

Der derzeitige Stand des operativen Einsatzes von PP-Systemen in den sechs aktiven Bundesländern ist aktuell wie folgt:<sup>34</sup>

- Bayern setzt *Precobs* seit 2015/16 operativ ein.
- Baden-Württemberg hat nach der umfassenden wissenschaftlichen Evaluation eines Pilotbetriebs von *Precobs* im Jahr 2017<sup>35</sup> weitere

31 Mit dieser Deliktart wurde aber auch in Nordrhein-Westfalen die Arbeit am Predictive Policing begonnen.

32 Beispielsweise wurde während des Projekts das Karten-Visualisierungstool SKALA | MAP eigens entwickelt, vgl. LKA NRW 2018c und 2018d. Damit lassen sich nicht nur Kriminalitätsprognosen abbilden, sondern es erlaubt mittlerweile auch allen Kreispolizeibehörden des Landes die Darstellung des sehr viel komplexeren Gesamtlagebildes. Dieser Aspekt nicht-intendierter positiver Effekte, auf den unten zurückzukommen sein wird, ist insgesamt kennzeichnend für die Einführung und Nutzung von PP-Systemen. Im Unterschied dazu konnte die Erzielung unmittelbar intendierter Wirkungen bislang noch nicht wissenschaftlich eindeutig nachgewiesen werden; vgl. dazu Abschnitt 7 unten.

33 Vgl. LKA Hamburg 2018.

34 Stand Mitte August 2018.

35 Vgl. Gerstner 2017.

Testbetriebe durchgeführt; diese werden gerade ausgewertet,<sup>36</sup> die Ergebnisse werden bis Ende 2018 veröffentlicht werden.

- Hessen setzt die Eigenentwicklung *KLB-operativ* flächendeckend ein.
- In Nordrhein-Westfalen wird die Eigenentwicklung *SKALA* seit Anfang 2018 von sechs städtischen Dienststellen operativ genutzt, ab dem 1.10.2018 werden weitere zehn hinzukommen; die Ausweitung des Einsatzes auf den ländlichen Bereich ist für 2019 geplant.
- Niedersachsen bewertet derzeit den Testbetrieb der Eigenentwicklung *PreMAP*, was bis Ende des Jahres 2018 zu einem Ergebnis führen soll.
- Berlin führt seit längerer Zeit einen „Proberechtbetrieb“ der Eigenentwicklung *KrimPro* durch; die Überführung in den regulären Echtbetrieb ist noch für dieses Jahr geplant.

In der Antwort des Berliner Senats auf eine Schriftliche Anfrage des Abgeordneten Hanno Bachmann vom 16.3.2017 zum Thema PP heißt es über *KrimPro*: „Die Entwicklung des Analysemodells wurde in Teilen mit externer Unterstützung unter Ausschöpfung bereits vorhandener vertraglicher Beziehungen realisiert.“<sup>37</sup> Dies deutet darauf hin, dass PP-Systeme – unabhängig davon, welcher Ansatz letztlich dahintersteckt – in der Regel Weiterentwicklungen bestehender behördlicher Datenverarbeitungssysteme sind. Sie fallen nicht vom Himmel auf ein weißes Blatt, sondern müssen sich – schon aus praktischen, vergaberechtlichen und finanziellen Erwägungen heraus – in bestehende informationstechnische Strukturen einfügen. Genau hier – das wird eine der Hauptthesen dieser Untersuchung sein – liegt nach derzeitigem Stand der Technik auch die größte Chance des Einsatzes von PP-Systemen: weniger in der Etablierung eines genuin neuen Erkenntnismitels und Ermittlungswerkzeugs der Polizei als vielmehr darin, dass dadurch positive Entwicklungen beim Informations- und Wissensmanagement angestoßen werden.<sup>38</sup> Denn verlässliche, aktuelle und gut aufbereitete Informationen sind nun einmal die Grundlage für eine erfolgreiche Polizeiarbeit in allen Zuständigkeitsbereichen.

---

36 Und zwar erneut vom Max-Planck-Institut für ausländisches und internationales Strafrecht in Freiburg im Breisgau.

37 Drucksache 18/10732: 1

38 Diesen Aspekt hat Thomas Schweer, Geschäftsführer des Oberhausener Unternehmens, das die im deutschsprachigen Raum verbreitete PP-Software *Precobs* vertreibt, wiederholt betont; zitiert wird er damit z.B. in Montag 2016: 4 f. Stark betont hat dies auch die Kriminologische Forschungsstelle im Landeskriminalamt Hamburg, vgl. dazu Abschnitt 5 unten.

# KRIMINOLOGISCHE GRUNDLAGEN UND MODELLIERUNG

*„Die Vorhersagesoftwares stützen sich auf kriminologische Theorien, denen zufolge einschlägige Deliktsbegehungen bestimmte Muster aufweisen. Die spezielle Datenanalysesoftware soll anhand der Kerndaten vergangener Straftaten bestimmen, mit welcher Wahrscheinlichkeit es an bestimmten Orten und Zeiten zu Folgetaten kommt, und ein präventives Einschreiten der Polizei ermöglichen.“<sup>39</sup>*

Diese Beschreibung aus einer Kleinen Anfrage der FDP-Bundestagsfraktion von April 2018 veranschaulicht gut sowohl die theoretische Ausgangsbasis als auch das Dilemma von PP-Systemen: In Kriminalstatistiken sollen raumzeitliche Muster erkennbar sein, die Schlüsse von der Vergangenheit auf die Zukunft nahelegen. Wissenschaftstheoretisch sind solche Schlüsse abseits sehr gut gesicherter physikalischer Theorien an sich problematisch und seit Langem Gegenstand grundlegender Kontroversen.<sup>40</sup> Der Traum von der Modellierung der „sozialen Wetterlage“, wie es ein Beamter gegenüber *Spiegel Online* ausgedrückt haben soll,<sup>41</sup> ist für Personen, deren Aufgabe im weiteren Sinne in der Kontrolle des sozialen Raumes besteht, natürlich verlockend. Die Realität sieht bislang allerdings sehr viel nüchterner aus: Bei der Bekämpfung von Wohnungseinbruchsdiebstahl (WED), auf den PP-Systeme in Deutschland derzeit vor allem angewendet werden und wo sie auch international ihren Ursprung haben, geht man meistens von der sogenannten **Near-Repeats-Theorie** aus.<sup>42</sup> Der Algorithmus der US-amerikanischen Software PredPol wurde ursprünglich genutzt, um Erdbeben vorherzusagen. Hier gilt nämlich ein ähnliches Prinzip: Wenn ein Erdbeben stattgefunden hat, erhöht sich in der Nachbarschaft das Risiko, dass ein weiteres stattfindet. Täter, insbesondere bandenmäßig organisierte, begehen in der Nähe von Einbruchstatorten und in zeitlicher Nähe weitere Einbrüche, und zwar schlicht um der Maximierung ihres Ertrags bei gleichzeitiger Minimierung des Aufwands willen.<sup>43</sup> Dieser, sich aus dem **Rational-Choice-Kalkül** ergebenden Beweggrund ist die Grundlage dafür, dass die klassische Kartographierung von Kriminalität mittels Stecknadeln, manch einem aus Krimiserien bekannt, nicht (nur) retrospektiv, sondern mit Zuhilfenahme algorithmischer Technologie auch prospektiv angewendet werden kann. Die hierfür verwendeten Algorithmen können anhand der Daten vergangener Taten und unter Berücksichtigung von mehr oder weniger zahlreichen Randbedingungen (z. B. Bebauung, Wetter, Verkehr, demographische Struktur) die Wahrscheinlichkeit für zukünftige Taten für unterschiedliche Gebiete errechnen und auf einer Karte abbilden. Das ist nichts Magisches, sondern Mathematik, wenn auch nicht (nur) triviale.

39 Kleine Anfrage der FDP-Fraktion, Drucksache 19/1513, Vorbemerkung der Fragesteller.

40 Vgl. Poser 2012, Teil A, Kapitel 5.

41 Vgl. Diehl & Kartheuser 2018.

42 Dies trifft auf alle mit PP aktiven Bundesländer außer Nordrhein-Westfalen und Berlin zu; vgl. Abbildung 2 oben sowie Egbert 2018, Abschnitt 2.2. Zur Near-Repeats-Theorie im Praxiseinsatz vgl. Balogh 2016.

43 „They use areas that they know well and where they understand vulnerabilities to avoid risks.“ – Hustedt 2016: 7

Bei der Analyse von PP-Systemen ist eine Reihe an Unterscheidungen relevant. Die wichtigste ist die zwischen **ortsbezogenen** und **personenbezogenen** Systemen. Erstere verdichten Daten jüngst vergangener Straftaten und weitere raumbezogene Informationen (z. B. zu Infrastruktur und soziodemographischer Situation) zu einem räumlichen Lagebild, das helfen soll, knappe Polizeiresourcen möglichst effizient einzusetzen – sprich in den Arealen Streifen einzusetzen, wo die Verbrechenswahrscheinlichkeit am höchsten ist. Letztere errechnen für einzelne Personen anhand bestimmter Risikofaktoren Wahrscheinlichkeiten, mit welchen diese Personen Straftaten begehen oder in Kapitalverbrechen involviert werden, und geben diese Informationen gegebenenfalls auf so genannten „Heat Lists“ (Gefährderlisten) aus. Da in Deutschland derzeit keine personenbezogenen Analysesysteme im Einsatz sind, kommen hier allein ortsbezogene Systeme in Betracht.<sup>44</sup>

Ortsbezogenes PP eignet sich für manche **Deliktarten** (viel) besser als für andere. Weil eine geringe Dunkelziffer Voraussetzung ist und die Anwendbarkeit der Near-Repeats-Theorie die Modellierung stark vereinfacht, wird PP vor allem auf Wohnungseinbruchsdiebstahl angewendet; gerade in Deutschland ist dies aktuell der hauptsächliche Anwendungsfall. Doch Gespräche mit Verantwortlichen aus mehreren Bundesländern haben ergeben, dass überall an einer deliktischen Weiterentwicklung etwa auf Gewalttaten, etwa Kfz- und Taschendiebstähle sowie an der – wegen der geringen räumlichen Dichte der Fallzahlen schwierigen – Anwendbarkeit auch im ländlichen Raum gearbeitet wird.

Im Anwendungsbereich ortsbezogener Systeme sind weitere grundsätzliche Unterscheidungen maßgeblich. So unterscheiden sich die im Einsatz befindlichen Systeme im Hinblick auf ihren **Komplexitätsgrad**: Anwendungen wie das US-amerikanische *PredPol* oder das deutsche *Precobs* beruhen auf einer sehr einfachen theoretischen Annahme (nämlich der Near-Repeats-Theorie, s.o.) und auf einer Auswertung nur weniger Datenpunkte (nämlich Beutetyp, Datum/Uhrzeit, Ort und gegebenenfalls Einbruchsort). Systeme wie das US-amerikanische Produkt *HunchLab* oder die Eigenentwicklung des nordrhein-westfälischen Landeskriminalamtes *SKALA* hingegen beziehen neben den polizeilich aufgezeichneten Fällen auch infrastrukturelle Gegebenheiten und soziodemographische Informationen mit ein und sind von der Modellierung her anspruchsvoller.<sup>45</sup>

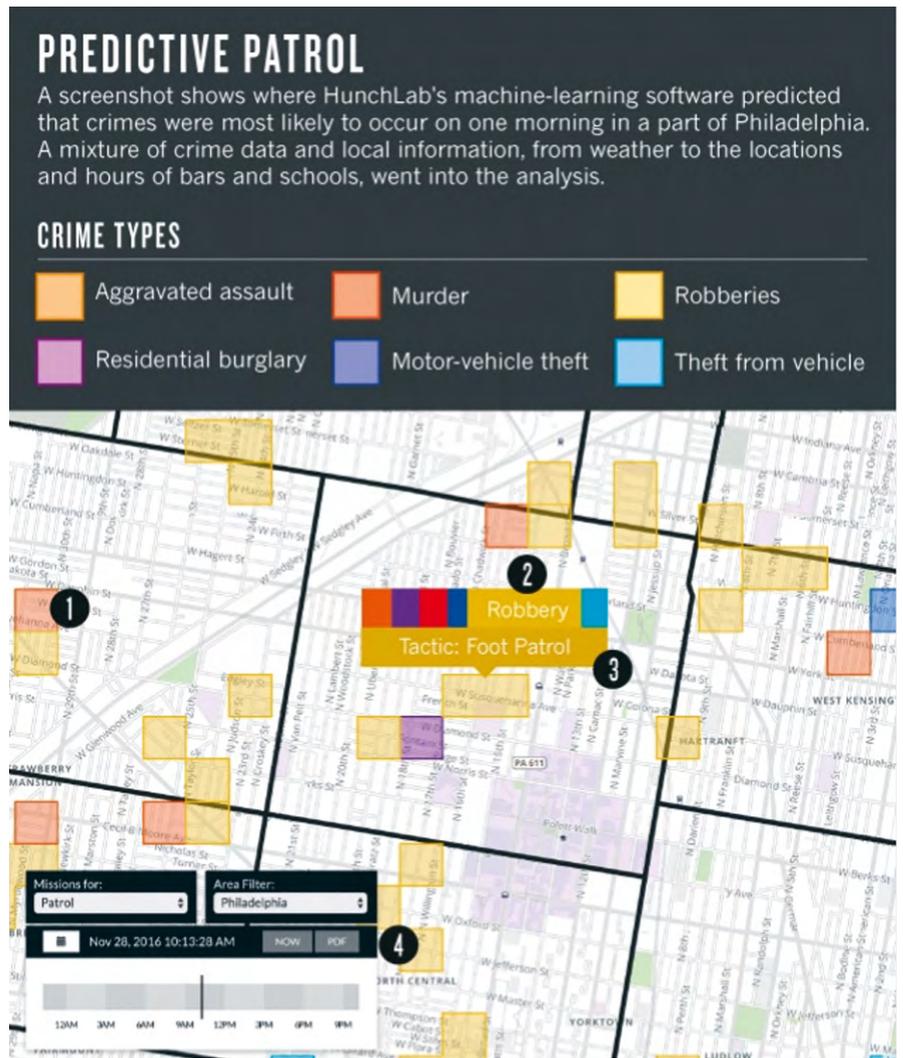
Dann gibt es Unterschiede hinsichtlich der **theoretischen Fundierung**. *HunchLab* ist beispielsweise theoretisch offen, es kommt völlig ohne kriminalistische Hypothesen aus und wendet maschinelles Lernen auf einen großen Datenpool an, um Vorhersagen zu generieren, die nicht durch theoretische Vorannahmen gestützt, aber eben auch nicht dadurch eingeschränkt sind. Auch bietet *HunchLab* Beamten die Möglichkeit, individuelle Voreinstellungen vorzunehmen (z. B. was die zur Verfügung stehende Anzahl an Streifen

<sup>44</sup> Lediglich im Rahmen der Darstellung internationaler Entwicklungen (Abschnitt 3 a.A.) sowie des Ausblicks auf zukünftige Entwicklungen (Abschnitt 9) wird am Rande auch personenbezogenes PP erörtert.

<sup>45</sup> Ausführlicher dazu in Abschnitt 5 unten.

**Abb. 3: Prospektives Kriminalitätslagebild in Philadelphia aus der Software HunchLab.**  
**Quelle: Nature, <https://www.nature.com/news/reform-predictive-policing-1.21338>; mit Bildmaterial von HunchLab, <https://hunchlab.com>**

angeht) und gibt taktische Empfehlungen, wie die Abbildung 3 verdeutlicht.<sup>46</sup>



Das LKA Nordrhein-Westfalen hat bei der Entwicklung des Systems *SKALA* dagegen bewusst einen theoriegeleiteten Ansatz gewählt. Demzufolge seien im Teilprojekt „Wissenschaftliche Begleitung und Analyse“ auf der Grundlage von Theorien und empirischen Befunden, die der wissenschaftlichen Literatur entnommen wurden, Hypothesen formuliert und operationalisiert worden; die Auswahl der Datenquellen habe sich aus der Operationalisierung ergeben.<sup>47</sup> Im Projektbericht wird dieses Vorgehen wie folgt (selbst) bewertet: „Das hypothesengestützte Verfahren gewährleistete, dass die Modell- und Prognoseerstellung auf belastbaren wissenschaftlichen Theorien und Forschungsbefunden basiert. Hierdurch grenzt sich die Vorgehensweise von vielen anderen Predictive-Policing-Verfahren ab, die häufig nur auf dem Near-Repeats-Ansatz beruhen.“<sup>48</sup> Was die Nachvollziehbarkeit und Transparenz des resultierenden Systems angeht, ist die nordrhein-westfälische Vorgehensweise gegenüber dem theorieleeren Machine-Learning-Ansatz grundsätzlich von Vorteil. Allerdings hat Azavea Inc., der Hersteller von *HunchLab* andere Maßnahmen, wie die Offenlegung von Quellcode und

<sup>46</sup> Vgl. Azavea 2015 und Shapiro 2017.

<sup>47</sup> Vgl. LKA NRW 2018a: 3.

<sup>48</sup> ebd.

den Einbezug von neutralen Beobachtern, ergriffen, um einen Blackbox-Charakter der Software zu vermeiden.<sup>49</sup> In Nordrhein-Westfalen hat man die Performanz der Entscheidungsbaumalgorithmen gegen ein neuronales Netz getestet, konnte aber keinen nennenswerten Vorzug feststellen und hat aus Gründen der besseren Nachvollziehbarkeit daher an der „klassischen“ Softwarearchitektur festgehalten.<sup>50</sup> Ob sich theorieaufgeladene Systeme wie *SKALA* in Sachen Effektivität gegenüber theoriearmen Systemen wie *Pred-Pol* oder *Precobs* als überlegen erweisen oder nicht, kann zum jetzigen Zeitpunkt schlicht deshalb nicht beantwortet werden, weil sich die Wirksamkeit sämtlicher im Einsatz befindlichen Systeme bislang noch kaum empirisch nachweisen ließ. Das hat einerseits mit der Schwierigkeit zu tun, Variablen zu kontrollieren, und andererseits mit dem Aufwand, der für adäquate Evaluationsdesigns betrieben werden muss.<sup>51</sup>

Zusätzlich zu den oben genannten Aspekten unterscheiden sich PP-Systeme nach der **Art ihrer Entstehung** und ihrem **Daseinszweck**: Erstens gibt es kommerzielle Software, die Firmen für die vorausschauende Polizeiarbeit mit Gewinnabsichten entwickelt haben (Nutzungslizenzmodell); zweitens verbreitete Statistikprogramme, die (mit oder ohne Dienstleisterhilfe) von den Behörden selbst an die Zwecke des PP angepasst wurden; drittens vollständige Eigenentwicklungen der Polizeibehörden, die mehr oder weniger weit über den Funktionsumfang von statistischen Standardwerkzeugen hinausgehen. Diese Vielfalt konzeptioneller Ansätze deutet auf einen wichtigen Umstand hin: PP-Systeme ersetzen ebenso wenig wie andere – mehr oder weniger automatisierte – algorithmische Analyse- und Entscheidungssysteme im öffentlichen Sektor die dort im Einsatz befindlichen IT-Systeme. Vielmehr treten sie in den meisten Fällen hinzu und sind daher auf die Interoperabilität mit den bestehenden Systemen angewiesen. So kommt es zu einer relativen Vielfalt an PP-Systemen, auch wenn bestimmte Lösungen zwischenzeitlich einen gewissen Marktanteil erzielen können. Aus Sicht der gesellschaftlichen Teilhabe ist eine solche Vielfalt (der Ziele und Betreiber algorithmischer Systeme) zu begrüßen, denn dadurch kann beispielsweise die Verbreitung von Schadenspotenzialen minimiert werden.<sup>52</sup>

Der **Prozess** der Entwicklung, Implementierung, des Einsatzes und der Evaluierung von PP-Systemen ist – wie die Entwicklung algorithmischer Systeme generell – ein komplexer Vorgang, in den zahlreiche Personen involviert sind. Die folgende Abbildung zum System *SKALA* in Nordrhein-Westfalen illustriert dies:

---

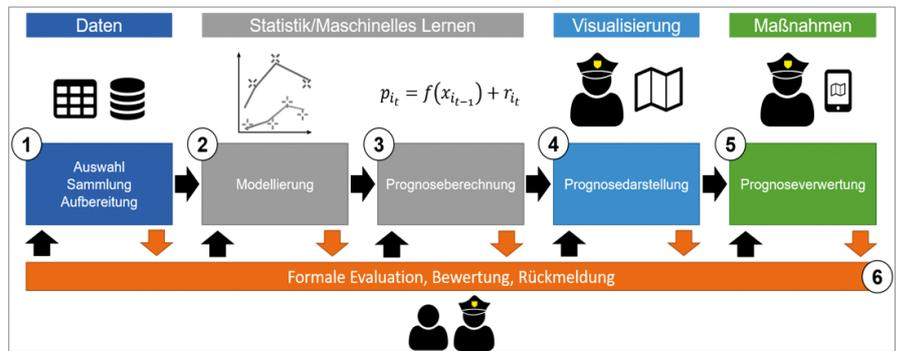
49 Vgl. Shapiro 2017 u. Azavea 2015.

50 Diese Information stammt vom ehemaligen Leiter des *SKALA*-Entwicklungsteams, der inzwischen das Referat „Kriminalitätsangelegenheiten“ im nordrhein-westfälischen Innenministerium leitet.

51 Ausführlicher dazu in Abschnitt 7 unten.

52 Vgl. Krüger & Lischka 2018: 49 ff.

**Abb. 4: Modellierung und Implementierung von SKALA.**  
 Quelle: Bode et al. 2017, S.2.



Hierbei muss sowohl bei der Entwicklung als auch bei der Implementierung von einer Fülle an möglichen Fehlerquellen ausgegangen werden. Katharina Zweig hat im Auftrag der Bertelsmann Stiftung solche Fehlerquellen systematisch erhoben und einige Lösungsansätze skizziert.<sup>53</sup> Es ist wichtig, dass sich Auftraggeber, Auftragnehmer und Anwender im Kontext von PP-Systemen der Fehleranfälligkeit bewusst sind und entsprechende Gegenmaßnahmen vorsehen.

Zuletzt ist in diesem Abschnitt ein **erkenntnistheoretisches Caveat** auszusprechen: Aus empirischen Befunden auf die Zukunft zu schließen impliziert die Annahme, dass die Dinge nach bestimmten Gesetzmäßigkeiten ablaufen und die Zukunft der Vergangenheit ähnelt.<sup>54</sup> Das trifft auf viele Erscheinungen zu, mit denen wir es zu haben. Beispielsweise sind wir bislang gut mit der Erwartung gefahren, dass die Sonne am Morgen des morgigen Tages erneut aufgehen und am Abend wieder untergehen wird. Auch im Bereich des sozialen Verhaltens ist es in beschränktem Umfang vernünftig, von Quasigesetzmäßigkeiten auszugehen. So sind etwa die meisten Menschen, die beleidigt wurden, nicht ohne Weiteres dazu bereit, demjenigen, von dem sie beleidigt wurden, einen Gefallen zu tun. Doch je komplexer die Sachverhalte, über die wir die Zukunft betreffende Aussagen machen wollen, sind, desto geringer ist normalerweise die Verlässlichkeit dieser Aussagen. Das trifft natürlich gerade auf ein so komplexes soziales Phänomen wie Kriminalität zu. Wir sollten endlich wenigstens den Versuch der Befreiung von dem Zauber, der uns seit dem Spielfilm *Minority Report* umfängt, unternehmen und uns ein für alle Mal klarmachen: Auch mit Erkenntnismitteln wie Big Data Analytics und Künstliche Intelligenz ist es nach wie vor nicht möglich, die Zukunft zu wissen. Wir können keine wirklich verlässlichen Aussagen über zukünftige Ereignisse machen. Allenfalls lassen sich Wahrscheinlichkeitsaussagen treffen, und nichts Anderes versucht man mit PP-Systemen zu tun.

<sup>53</sup> Vgl. Zweig et al. 2017.

<sup>54</sup> Vgl. Perry et al. 2013: 8.

## DATENBASIS

*„Weiterhin zeigte sich, dass die Modellgüte entscheidend von der Qualität und zeitlichen Verfügbarkeit der Daten abhängt.“<sup>55</sup>*

Entscheidende Faktoren für den Erfolg des – wie gerade gezeigt: anspruchsvollen – Prognoseanspruchs sind, neben der Modellgüte und der Angemessenheit der zugrundeliegenden theoretischen Annahmen, der Umfang, die Qualität, Aktualität und Verfügbarkeit der Datenbasis. Systemanbieter und Systembetreiber sind daher stets darum bemüht, die Datenbasis in den genannten Hinsichten zu verbessern. Der Aufwand, der dazu nötig ist, wird oft stark unterschätzt und kann daher bei der Projektplanung vernünftigerweise gar nicht hoch genug eingeschätzt werden. Im Zeitalter algorithmischer Steuerung steht und fällt alles mit Daten, denn Daten sind der Treibstoff algorithmischer Systeme. Nicht nur das: Unabhängig davon, ob es sich um die momentan gehypten selbstlernenden Systeme handelt oder nicht, prägen Daten auch die Gestalt und Funktionsweise der Modelle. Schließlich wird jedes Modell vor seiner Überführung in den Betrieb kalibriert, das heißt es muss vor der Inbetriebnahme in so genannten „Input-Output-Tests“ bereits als korrekt bekannte Resultate reproduzieren können. Hierbei kann es, wie an anderen Stellen des Softwareentwicklungsprozesses, zu Fehlern kommen, die sich im Realbetrieb später als verheerend herausstellen können.

Daten spielen schon viel früher, als im Zusammenhang mit algorithmischen Systemen gewöhnlich diskutiert wird, eine große Rolle, nämlich im Zuge ihrer Erfassung. Polizeidaten sind die wichtigste informationelle Stütze von PP-Systemen. Erfasst werden sie mittels so genannter „Vorgangsbearbeitungssysteme“. Polizeibeamt:innen tragen in Eingabemasken eines Computerprogramms alle für einen verwaltungsrechtlichen Vorgang relevanten Daten ein, bei einem Verkehrsunfall beispielsweise die Anschrift der beteiligten Personen, von Zeugen und vieles mehr. So ist es auch bei den Delikten, auf die derzeit PP-Systeme angewendet werden, vorzugsweise Wohnungseinbruch. Oben wurde bereits gesagt, dass eine möglichst vollständige datenmäßige Erfassung eines Deliktsfeldes die Voraussetzung dafür ist, dass prognostische Algorithmen überhaupt sinnvoll eingesetzt werden können. Das ist beim Wohnungseinbruch ganz überwiegend der Fall, doch das ist nicht das Ende des Themas Datenvollständigkeit und Datenqualität. Denn in der Praxis sind die polizeilichen Systeme zur Datenerfassung und Vorgangsbearbeitung natürlich nicht in allen Ländern gleich. Auch werden sie von den Anwendern nicht immer geschätzt, im Gegenteil gibt es vielerorts Klagen, was die Komplexität und Nutzerfreundlichkeit angeht. In einem der Bundesländer, die in unserem Workshop zu PP-Systemen im April 2018 vertreten waren, gibt stellt das Vorgangsbearbeitungssystem bei der Anlage eines neuen Wohnungseinbruchsfalles eine dreistellige Anzahl an Einbruchsmethoden zur Wahl.

---

55 LKA NRW 2018a: 6

Eine solch überfordernde Vielfalt führt der Tendenz nach zu einer unsauberen Datenlage. Auch müssen die Datenerzeugungs- und verarbeitungssysteme der Polizei dem Arbeitsalltag der Beamt:innen geschuldete Restriktionen spiegeln. Beispielsweise müssen zusätzlich gewonnene Informationen wie ein Tatwerkzeug auch nachträglich zu einem Fall hinzugefügt werden können. Da es diesbezüglich mannigfachen Verbesserungsbedarf gibt, wird in einem anderen Bundesland gerade der landesweite Wechsel auf ein neues Vorgangsbearbeitungssystem betrieben. Dabei handelt es sich um ein IT-Projekt mit einem hohen Millionenvolumen – weitaus größer und teurer als jedes PP-System, das sich weltweit auf dem Markt befindet. Es gehört zu den Eigentümlichkeiten des Diskurses über vorausschauende Polizeiarbeit, dass PP-Systemen dennoch eine große mediale Aufmerksamkeit zuteilwird, während die Systeme, mit denen die Polizei sämtliche Daten aufnimmt und verarbeitet, lediglich in Fachzeitschriften für Verwaltungs-IT eine Rolle spielen.

Da die Erstellung von Daten nur einen kleinen Teil des Arbeitsalltags von Polizeibeamt:innen ausmacht und die zurzeit dafür in Gebrauch befindlichen Systeme nicht besonders nutzerfreundlich sind, ist davon auszugehen, dass die Polizeidatenbasis weder vollständig noch fehlerfrei ist. Das spricht nicht per se gegen den Einsatz von PP-Systemen, aber es sensibilisiert erstens dafür, dass die mittels historischer Polizeidaten kalibrierten, beziehungsweise auf diesen Daten trainierten Modelle kritisch zu überprüfen sind. Lernende Algorithmen wie die von *PredPol* oder *Precobs* müssen mithilfe von Datenpools trainiert werden, bei *PredPol* sind das in der Regel zwei bis fünf Jahre an Einbruchdaten der Polizei, die hier Verwendung finden.<sup>56</sup> In einem deutschen Bundesland, das *Precobs* nutzt, musste das System in der Simulation für einen in der Vergangenheit liegenden Fünfjahreszeitraum bekannte Einbruchmuster mit einer Genauigkeit von 70 Prozent reproduzieren können, bevor es in den prognostischen Realbetrieb überführt wurde.

Wie stark der Einfluss von Trainingsdaten auf die spätere Funktion eines algorithmischen Systems ist, haben Forscher vom Massachusetts Institute of Technology an einem scheinbar so unverfänglichen Beispiel wie Bilderkennung und -interpretation gezeigt: Ihr Testprogramm *Norman* wurde mit Informationen der dunkelsten Regionen des Internets trainiert und interpretierte später das, worin „neutrale“ Bilderkennungsprogramme als Vögel oder Blumen ausmachten, ausschließlich als Gewaltszenen.<sup>57</sup> In einem anderen Beispiel ließen Microsoft-Forscher einen Chatbot namens *Tay* ausschließlich von realer Twitter-Kommunikation das Kommunizieren lernen. Das Ergebnis war, dass sich der Bot bereits nach kurzer Zeit rassistisch, und in anderer Weise politisch so radikal und unflätig gerierte, dass er abgeschaltet werden musste.<sup>58</sup> Beispiele wie diese zeigen, wie sehr die Interpretationsgabe und der Möglichkeitsraum für Funktionen maschineller Systeme von ihrem „Erfahrungshorizont“ abhängt. Das gilt auch für PP-Software.

---

56 Vgl. *PredPol* 2018.

57 Vgl. Wakefield 2018.

58 Vgl. Graff 2016.

Polizeidaten – also Daten, welche die Polizeidienststellen selbst erheben und speichern – sind nicht immer die einzigen Datenquellen, mit denen PP-Systeme anfänglich trainiert und später, im laufenden Betrieb „gefüttert“ werden. Komplexere Systeme wie das mehrfach erwähnte nordrhein-westfälische SKALA nutzen auch geographische, infrastrukturelle und soziodemographische Informationen.

**Abb. 5: Datenquellen von SKALA.**  
Quelle: LKA NRW 2018c: 22.



Auch diesbezüglich gilt natürlich, dass die verwendeten Daten qualitätsgesichert und aktuell sein müssen. Das kann, wie im Fall Nordrhein-Westfalen, dazu führen, dass die Behörden Geld in die Hand nehmen, sprich die entsprechenden Daten bei zuverlässigen Lieferant:innen einkaufen müssen.<sup>59</sup> Wünschenswerter wäre es indes, wenn die öffentliche Hand hier verlässlich auf Daten der öffentlichen Hand, also Open Government Data, zurückgreifen könnte. Doch infolge der viele Jahre währenden stiefmütterlichen Behandlung dieses Themas in Deutschland stellt sich die Situation uneinheitlich und insgesamt unbefriedigend dar. Hier sehen wir nach wie vor auch politischen Nachholbedarf, etwa im Zuge der von Bund und Ländern vereinbarten Open-Data-Gesetzgebungen der Bundesländer nach dem Vorbild der Bundesgesetzgebung in Gestalt der im Sommer 2017 erfolgten Novellierung des E-Government-Gesetzes (EGovG).

In unserem Policy Paper „Das Datenzeitalter gestalten“ von 2016 haben wir argumentiert, wie fundamental die intensive Selbstbefassung der Behörden mit ihren Praktiken und Systemen der Datenerfassung und -verarbeitung im Datenzeitalter ist.<sup>60</sup> Seinerzeit haben wir das Argument im Zusammenhang mit der Öffnung von Verwaltungsdaten vorgebracht, doch es lässt sich unter geänderten Vorzeichen nun auch auf das Thema vorausschauende Polizeiarbeit anwenden: Eine solche wird nur dann state-of-the-art möglich sein, wenn sich die Behörden kritisch mit ihrem eigenen Datenverhältnis auseinandersetzen, ihre Vorgangsbearbeitungssysteme auf den Prüfstand stellen, Datenformate vereinheitlichen, Datenkompetenz in der Belegschaft aufbauen und so weiter. Auf der anderen Seite muss man irgendwo anfangen, und so kann die Beschäftigung mit neuen technischen Verbrechensbekämpfungsinstrumenten wie PP-Software den Anfang für Reformen im Bereich des Wissens- und Informationsmanagements sein. Diesen Ansatz verfolgt die Kriminologische Forschungsstelle des Landeskriminalamts Hamburg mit einem aktuellen Forschungsprojekt zur schweren Einbruchskriminalität.

<sup>59</sup> Nämlich bei der Nexiga GmbH, vgl. LKA NRW 2018a: 3.

<sup>60</sup> Vgl. Knobloch & Manske 2016.

Im Projektzwischenbericht heißt es ganz richtig über den Zusammenhang zwischen dem Einsatz algorithmischer Systeme und einem adäquaten Datenmanagement:

*„Die Suche nach Zusammenhängen durch ‚wissende‘ Computerprogramme übernehmen zu lassen, ist verführerisch und lässt sich leicht und abstrakt fordern. Die Erhebung und Zusammenführung von Informationen – der unspektakuläre, aber essentielle Prozess der Datenverwaltung – wird hierbei fälschlicherweise als bereits voll funktional vorausgesetzt. Erfolgreiche Technisierung setzt jedoch eine Auseinandersetzung mit den vorhandenen Datengrundlagen zwingend voraus. (...) Aus den bisherigen Erkenntnissen des Forschungsprojekts resultiert das Erfordernis, digitales Informationsmanagement in der Ermittlungsarbeit zu optimieren und zukunftssicher zu gestalten.“<sup>61</sup>*

An letzter Stelle in diesem Abschnitt, aber keineswegs von der Bedeutung her zuletzt ist darauf zu verweisen, dass zweifelhafte Datengrundlagen nicht nur zu offensichtlich fehlerhaften oder unbrauchbaren Modellen, sondern auch mehr oder weniger unbemerkt zu anderen, nicht wünschenswerten Effekten führen können. Stadtbezirke, Bevölkerungsschichten und Personengruppen können, wenn Modelle mittels eines bloßen Datenabbildes der Wirklichkeit trainiert werden, systematisch diskriminiert werden. Es droht die Gefahr eines vitiösen Zirkels: In einem Bezirk, der häufig von Einbrüchen betroffen ist, geht die Polizei vermehrt auf Streife. Dadurch werden mehr Straftaten (auch anderer Art) entdeckt als an anderen Orten der Stadt. Das wiederum nimmt das System zum Anlass, eine weiterhin verschärfte Polizeipräsenz für den Problembezirk zu verordnen – und so weiter, bis man auf diese Weise so genannte „No-go-areas“ geschaffen hat, so dass alleine eine Meldeadresse dort dafür sorgt, dass man nicht beliefert wird, keine Verträge abschließen kann und keinen Job findet. Gegen das Eintreten solcher Effekte müssen Vorkehrungen getroffen werden. Das mehrfach erwähnte US-amerikanische PP-Programm *HunchLab* fügt, um dem entgegen zu wirken, beispielsweise Zufallsparameter ein, die dafür sorgen, dass Beamte immer auch mal in Quartieren anzutreffen sind, die von der Software ansonsten als unproblematisch ausgewiesen würden. Am Ende helfen gegen räumliche Segregationseffekte keine Algorithmen, sondern nur politisch motivierte sozioökonomische Maßnahmen, die aus der Stadtsoziologie hinreichend bekannt sind.<sup>62</sup>

Algorithmen können helfen, Muster zu erkennen, an denen wir etwas ändern wollen sollten. Sie können uns auch dabei unterstützen, die Entstehung so genannter „Problembezirke“ zu bekämpfen. Aber sie können, unreflektiert eingesetzt, eben auch den genau gegenteiligen Effekt hervorbringen, beziehungsweise verstärken. Dies muss jedem, der solche Systeme einsetzt beziehungsweise für deren Einsatz eintritt, bewusst sein. Der Einbezug von

<sup>61</sup> LKA Hamburg 2018: 2 f. – Der Projektabschlussbericht ist für Ende 2018 angekündigt. Die Situation in Hamburg ist insofern besonders, als es dort seit 2012 ein Transparenzgesetz gibt, welches die umfassende Veröffentlichung von Verwaltungsinformationen vorschreibt. Dies mag den dort überdurchschnittlich ausgeprägten Sinn für die Themen Daten- und Informationsmanagement erklären.

<sup>62</sup> Locus classicus der Entwicklung von Wohnquartieren in Abhängigkeit von der Technologie- und Kreativszene ist Richard Floridas „The Rise of the Creative Class“, vgl. Florida 2003, S.215 ff.

Datenexpertise, wie sie z. B. (aber nicht ausschließlich) bei Data Scientists vorliegt, ist hier unbedingt geboten. Als fruchtbar herausgestellt hat sich ferner eine Zusammenarbeit von Fach- und Datenexpert:innen, um sämtliche möglichen Fehler oder Verzerrungen (Biases) in den zugrunde gelegten Daten zu vermeiden.

## INSTITUTIONELLE IMPLEMENTIERUNG

*„PP-Systeme sind im Grunde eine Beleidigung des kriminalistischen Gespürs von Polizeibeamten.“<sup>63</sup>*

Diese (bewusst zugespitzte) Formulierung soll verdeutlichen, dass technische Systeme, die das Lagebild erweitern und Maßnahmenvorschläge machen, in hergebrachte Prozesse der polizeilichen Arbeit eingreifen und diese verändern. Es besteht das Risiko, dass solche Systeme – wie übrigens in anderen beruflichen und gesellschaftlichen Kontexten auch – pauschal als Bedrohung wahrgenommen und von den Anwendern nicht als das angenommen werden, was sie sein sollen: Hilfsmittel.

Im Evaluationsbericht zu SKALA, dem zwischen 2015 und 2017 getesteten und eingeführten PP-System Nordrhein-Westfalens, heißt es zum Thema Konkurrenz von PP-System und herkömmlicher Polizeiarbeit:

*„Das Lagebild hat für die Polizei eine große Bedeutung. Nach ihm werden Entscheidungen über Maßnahmen getroffen. Mit den Prognosen bezüglich der Wahrscheinlichkeit eines Auftritts von WED (Wohnungseinbruchdiebstahl, Anm.d.Verf.) für einzelne Gebiete tritt nun ein neues Element auf den Plan, das das vorhandene Lagebild nicht verdrängt, sondern mit zusätzlichen Informationen versieht, es umfassender gestaltet. (...) In jedem Fall werden Prognosen mit entsprechendem Wissen abgeglichen – im besten Fall wird eine Synthese daraus.“<sup>64</sup>*

In den Empfehlungen der Gutachter heißt es an späterer Stelle, man müsse klarstellen, „dass SKALA und die Prognosegebiete nicht in Konkurrenz zum Lagebild stehen – vielmehr soll das Lagebild um die Prognoseinformationen erweitert werden“<sup>65</sup>. Diese Einschätzung ist zentral: Auch die Intuitionen der Beamt:innen müssen nach wie vor ernst genommen werden und in die Polizeiarbeit einfließen können, sonst werden PP-Systeme innerhalb der Polizeibelegschaft keine Akzeptanz finden. PP-Systeme treten in der Regel nicht in Konkurrenz zu bestehenden Informations- und Einsatzplanungssystemen, sondern ergänzen diese. Die Akzeptanz innerhalb der Belegschaft und die Wirksamkeit des Einsatzes von PP-Systemen hängen auch davon ab, wie gut die Einbettung technisch und prozessual funktioniert. Aller-

63 Aussage eines Teilnehmers am PP-Workshop in der Stiftung Neue Verantwortung am 20.4.2018.

64 LKA NRW 2018b: 10 f.

65 a.a.O.: 19

dings verändert sich, wie die obigen Zitate andeuten, durch PP die Art und Weise, wie geplant und gearbeitet wird. Wichtig ist, dass Führungskräfte und politische Entscheidungsträger darauf hinweisen, dass dies – jedenfalls nach derzeitigem Stand der Dinge – kein paradigmatischer Wandel ist. Gleichwohl müssen die Kolleginnen und Kollegen aber darauf eingestellt und vorbereitet werden. Da die Akzeptanz proportional mit der Höhe des Dienststranges zunimmt, bestehen vor allem bei den unteren Dienststrängen Aufklärungs- und Schulungsdesiderata.<sup>66</sup>

Auf ein besonderes Problem, das beinahe erkenntnistheoretischer, mit Sicherheit aber psychologischer Natur ist, weist der Evaluierungsbericht zum Projekt SKALA in Nordrhein-Westfalen hin: Während PP-Systeme raumzeitliche Informationen verdichten und Gebiete zur Präventionsarbeit empfehlen, tendiert die Intuition der Beamt:innen in der Regel zur konkreten täter- und ereignisbezogenen Lageeinschätzung. Die Güte dieser Einschätzung wird normalerweise daran gemessen, ob ein prognostiziertes Ereignis eintritt oder nicht. Der Erfolg von PP-Systemen besteht aber gerade darin, dass Deliktbegehungen eben nicht oder vermindert eintreten. Menschen tendieren nachvollziehbarerweise dazu, Erfolg mit wahrnehmbaren Ereignissen und der menschlichen Manipulationsfähigkeit bezogen auf diese Ereignisse zu verknüpfen. Dabei geht es um die Sichtbarkeit der eigenen Arbeit und um die Demonstration von Erfolgen nach außen.<sup>67</sup> Unter diesen Bedingungen haben es PP-Systeme nicht leicht sich – insbesondere bei den Einsatzkräften im Außendienst – Sympathiepunkte zu erwerben. Zumindest implizit werde von vielen Kolleg:innen unterstellt, so das LKA Nordrhein-Westfalen, „dass das tradierte Vorgehen (Erfahrungswissen, retrogrades Lagebild) besser ‚funktioniert‘ als SKALA und SKALA nun den Beweis antreten muss, dass es mehr ‚bringt‘ als die bislang eingesetzten Vorgehensweisen“<sup>68</sup>.

Umso wichtiger ist es, mit falschen Erwartungen „aufzuräumen“ und realistisch mögliche Mehrwerte aufzuzeigen. Positiv ist beispielsweise zu vermerken, dass PP-Systeme dafür sorgen, dass man sich überhaupt mit dem Zustandekommen und der adäquaten Nutzung des Lagebildes kritisch auseinandersetzt, wozu auch die erhobenen Daten gehören. In Nordrhein-Westfalen haben sich im Zuge der Evaluierung von SKALA abseits der ursprünglichen Ziele positive Nebeneffekte gerade auch hinsichtlich der institutionellen Weiterentwicklung der Polizeibehörden gezeigt:

*„Von wenigen Ausnahmen abgesehen, verbessert SKALA offensichtlich die Kommunikation zwischen einzelnen Organisationseinheiten. Dies bezieht sich nicht nur auf den Austausch von Informationen, sondern auch auf die Verknüpfung von Maßnahmen unterschiedlicher Direktionen. Hier wird angemerkt, dass dabei vor allen Dingen auch die persönliche Ebene des Kontaktes gestärkt wurde.“<sup>69</sup>*

---

66 Vgl. LKA NRW 2018b: 10.

67 Vgl. a.a.O.: 12.

68 ebd.

69 a.a.O.: 11

Stärkung persönlicher Bindungen in einer Behörde durch den Einsatz einer Prognosesoftware – das ist schon ein Befund, der weitab von den gängigen Schreckensvisionen bezüglich der Entsozialisierung durch Technologie als bemerkenswert hervorzuheben ist. Ferner geht der Evaluierungsbericht auch darauf ein, dass das PP-System SKALA innerhalb der Polizeibelegschaft grundsätzlich als das anerkannt wurde, was es ist, nämlich als ein Instrument der effizienten Kräftesteuerung.<sup>70</sup>

Ohne an dieser Stelle zu sehr ins Detail gehen zu können, ist grundsätzlich zu konstatieren, dass der Erfolg von PP damit steht und fällt, wie es in den Dienststellen über Ränge, Funktionen und Laufbahnen hinweg eingeführt und vermittelt wird. Die wenigen Informationen, die dazu zugänglich sind, beziehungsweise von Beteiligten in Gesprächen offenbart wurden, deuten darauf hin, dass es hier noch deutliches Verbesserungspotenzial gibt. Der zumindest anfangsweise Pilotcharakter eigentlich aller Einführungen in deutschen Bundesländern hat von vornherein zu einer starken Selektivität geführt: Manchen wurde der Sinn und Zweck vorab erläutert, anderen nicht. Hier ist zu wünschen, dass datengesteuerte Prognose- und Ermittlungstechniken zum Gegenstand der Polizeiausbildung und -fortbildung werden. Ferner muss für das kommende Jahrzehnt des Übergangs ein Weg gefunden werden, auch älteren Kolleg:innen den Zweck sowie die Vorzüge und auch die Gefahren von PP-Systemen in leicht nachvollziehbarer, nicht-technischer Weise zu vermitteln. Ein Zugang zum System für alle und die Möglichkeit, es zumindest mit allen jüngst eingeführten Dienstgeräten nutzen zu können, sollte selbstverständlich sein. Die Nutzung einer kleinen polizeiinternen Elite vorzubehalten kann allenfalls im Zuge von Tests eine Kurzzeitlösung, keineswegs aber im Zuge des Ausrollens ein dauerhafter Zustand sein.

## WIRKUNGSMESSUNG UND EVALUIERUNG

*„Tatsächlich finden sich für bisherige Versuche auch nur wenige Veröffentlichungen oder Berichte, bei denen Unbeteiligte eine begleitende Evaluation durchgeführt haben. (...) Vor diesem Hintergrund ist das Pilotprojekt P4 der erste Test von Predictive Policing mit einer unabhängigen wissenschaftlichen Evaluation.“<sup>71</sup>*

Anknüpfend an die Skepsis vieler Polizeibeamt:innen, die mit dem Thema in Berührung kommen,<sup>72</sup> muss – schon weil PP-Systeme mit Steuergeld bezahlt werden – natürlich die Frage nach den tatsächlichen Effekten und Verbesserungen der Polizeiarbeit durch PP gestellt werden. Die ernüchternde Kurzantwort ist: Es ist bislang völlig unklar und insgesamt sehr schwer belegbar, ob der Einsatz von PP-Systemen in der intendierten Weise (also

---

<sup>70</sup> Vgl. ebd.

<sup>71</sup> Gerstner 2017: 1 über den Stellenwert der Evaluation des baden-württembergischen Einsatzes der Software *Precobs* durch das Max-Planck-Institut für ausländisches und internationales Strafrecht.

<sup>72</sup> Vgl. LKA NRW 2018d: 92.

einen Kriminalitätsrückgang mit gleichen Ressourcen zu erwirken) Effekte zeitigt. Die gute Nachricht ist allerdings, dass es unerwartete positive Nebeneffekte gibt, die den Einsatz von PP-Systemen unter Umständen dennoch rechtfertigen.

Eine Stellungnahme des Kriminologischen Forschungsinstituts Niedersachsen zum Maßnahmenpaket zur Bekämpfung des Wohnungseinbruchsdiebstahls in Nordrhein-Westfalen hat schon 2016 auf „kaum belastbare und zudem widersprüchliche Evaluationsbefunde“<sup>73</sup> verwiesen. Während die Entwickler von *Precobs* im Jahr 2015 eine höhere Präventivwirkung der Streifentätigkeit durch den Einsatz ihrer Software nachgewiesen zu haben glaubten,<sup>74</sup> hatte eine erste unabhängige Studie 2014 keinen wesentlichen Einfluss auf die Fallzahlen von Eigentumsdelikten finden können.<sup>75</sup> Ferner haben die niedersächsischen Forscher darauf verwiesen, dass die Wahl (oder Nichtwahl) bestimmter Präventions- oder sonstiger polizeilicher Maßnahmen im Gefolge eines PP-Systems einen wesentlichen Einfluss auf die Entwicklung der Fallzahlen hat, sodass eine Rückführung auf den Softwareeinsatz an sich fragwürdig sei.<sup>76</sup> Auch intervenierende, kaum zu systematisierende Effekte, wie die spontanen Taten von Drogenabhängigen, seien in Betracht zu ziehen.<sup>77</sup>

Die Wirkung von PP wissenschaftlich-empirisch nachzuweisen ist generell sehr anspruchsvoll, weil man sich beim Wohnungseinbruch, dem derzeit bevorzugten Anwendungsfeld, im Bereich geringer Wahrscheinlichkeiten befindet, und weil es schwierig ist, den Einfluss intervenierender Variablen zu isolieren. Einfach ausgedrückt: Wie können PP-Systeme als entscheidender Faktor bei einem Rückgang von Kriminalität (der in mehreren PP-Testanwendungsgebieten in Deutschland tatsächlich festgestellt wurde) nachgewiesen werden, wenn es ebenso gut am Wetter, am soziodemographischen Wandel, an anderen Polizeimaßnahmen oder an ganz anderen Faktoren, an die man spontan gar nicht denken würde, liegen könnte? Es ist klar, dass eine solche Isolierung von Einflussfaktoren in komplexen sozialen Zusammenhängen ein anspruchsvolles Testdesign voraussetzt. Neben experimentellen Kontrollgebieten, in denen PP-Maßnahmen ausbleiben, muss der Erhebungszeitraum lang genug ausfallen, um etwa saisonale Effekte auszuschließen. Zusammengenommen ergibt das eine recht hohe Hürde für eine State-of-the-Art-Evaluierung von PP-Systemen. Das erklärt, dass unabhängige Studien zur Wirksamkeit von PP lange Zeit fehlten.<sup>78</sup> Trotzdem sind entsprechende Versuche in Baden-Württemberg und Nordrhein-Westfalen mit der dortigen Einführung von PP-Systemen inzwischen unternommen worden.

Die ausführlich in Gerstner 2017 dargestellte Evaluation für Baden-Württemberg hat zusammengefasst – unter der Einschränkung der Aussagekraft aufgrund der Unsicherheiten des Testdesigns – Folgendes ergeben:

73 Kriminologisches Forschungsinstitut Niedersachsen 2016: 3; vgl. auch Gluba 2014 und 2016.

74 Vgl. Mohler et al. 2015.

75 Vgl. Hunt et al. 2014.

76 Kriminologisches Forschungsinstitut Niedersachsen 2016: 4

77 Vgl. ebd.

78 Vgl. Uchida 2014 und ebenso noch Gerstner 2017: 5 f.

*„Inwieweit Predictive Policing zu einer Verminderung von Wohnungseinbrüchen und zu einer Trendwende in der Fallentwicklung beitragen kann, ist auch nach dem Pilotprojekt trotz einiger positiver Hinweise schwer zu beurteilen. (...) Der wichtigste Schluss ist, dass kriminalitätsmindernde Effekte von Predictive Policing im Pilotprojekt P4 wahrscheinlich nur in einem moderaten Bereich liegen und allein durch dieses Instrument die Fallzahlen nicht deutlich reduziert werden können.“<sup>79</sup>*

Trotz dieses nicht eindeutigen Befundes ist es als ein großer Verdienst des beauftragenden Landes Baden-Württemberg, des durchführenden Max-Planck-Instituts für ausländisches und internationales Strafrecht in Freiburg im Breisgau und des Evaluationsleiters Dominik Gerstner anzusehen, dass diese bislang umfassendste wissenschaftliche Evaluation von PP in Deutschland stattfinden konnte. Folgenden Untersuchungen wird sie als wichtiger Referenzpunkt dienen.

In Nordrhein-Westfalen ist eine kooperative Evaluation der Zentralstelle Evaluation des LKA Nordrhein-Westfalen (ZEVA) und der Gesellschaft für innovative Sozialforschung und Sozialplanung Bremen (GISS) des NRW-Projekts *SKALA* 2018 zu folgendem Schluss gekommen: „Es liegen keine belastbaren statistischen Ergebnisse vor, die auf eine Wirkung von SKALA – im Sinne eines Zusammenhangs von Maßnahmen und den darauf folgenden Ereignissen (z. B. Festnahmen, Verhinderung von WED) – hindeuten.“<sup>80</sup> Gleich angefügt wird ein möglicher Grund für dieses Scheitern: „Allerdings ist die Zeit, hier entsprechende Befunde zu finden, zu kurz gewesen.“<sup>81</sup> Da man in Nordrhein-Westfalen großes Vertrauen in das selbst entwickelte Modell hat, geht man dort davon aus, dass „mit einer längeren Beobachtungszeit jedoch die Wahrscheinlichkeit steigt, auch im statistisch messbaren Bereich unter Hinzuziehung von adäquaten Indikatoren, belastbare Effekte zu ermitteln“<sup>82</sup>.

Für die Wirkungsevaluation reicht es nicht aus, Kriminalitätsstatistiken heranzuziehen. Auch das Sicherheitsgefühl der Bevölkerung muss hier in Betracht gezogen werden.<sup>83</sup> Dieses Gefühl kann sich auch dann durch Maßnahmen, die im Zuge der Verwendung von PP-Systemen durchgeführt werden, verbessern, wenn ein objektiver Effektivitätsnachweis nicht erbracht werden kann. An dieser Stelle kann ein natürlicher Anreiz für Polizeibehörden entstehen, transparent mit PP-Maßnahmen umzugehen und mit der Bevölkerung zusammenzuarbeiten. Während dieser Aspekt bereits auf die im folgenden Abschnitt 8 zu diskutierende soziale Einbettung von PP-Systemen verweist, ist, wie im vorigen Abschnitt 6 dargestellt, auch mit positiven Nebeneffekten im Bereich der institutionellen Einbettung zu rechnen. Insbesondere die unbedingt notwendige Verbesserung des Datenmanagements und die potenzielle Optimierung polizeilicher Lagebilderstellungs-

---

79 Gerstner 2017: 84 f.

80 LKA NRW 2018b: 16

81 ebd.

82 ebd.

83 Vgl. Hustedt 2016: 12.

Einsatzplanungs- und Ermittlungsprozesse sind hier zu nennen. Deshalb ist zu empfehlen, diese möglichen indirekten Gewinne zu weiteren Hauptzwecken der Einführung von PP-Systemen zu erklären, um Enttäuschungen vorzubeugen und eine für alle Beteiligten und Betroffenen motivierende Erfolgsgeschichte wahrscheinlicher zu machen.

Selbst bezüglich des Wenigen, was derzeitige PP-Systeme zu leisten versprechen,<sup>84</sup> lässt sich ihre Wirksamkeit nicht leicht nachweisen. Ganz abgesehen davon kann man im Grunde nicht davon sprechen, dass mit diesen Systemen Verbrechen regelrecht bekämpft werden. Erstens werden von den einfachen, auf dem Near-Repeats-Ansatz beruhenden Systemen nur kleine, zuvor als tauglich definierte Teile eines Gesamtgebietes überhaupt abgedeckt.<sup>85</sup> Zweitens tendieren sie dazu, als „technological fixes“ lediglich die Symptombekämpfung zu unterstützen, nicht aber die Ursachen von Kriminalität.<sup>86</sup> Es liegt die Gefahr nahe, dass hier das Gefühl eines Paradigmenwechsels erzeugt wird, während tatsächlich nach wie vor reaktiv gearbeitet wird. Wirkliche Präventionsarbeit müsste beispielsweise beim Abbau sozialer Ungleichheit, der Entstehung segregierter Wohnquartiere etc. ansetzen.

Automatisierung, durch die der Mensch in evolutionärer Betrachtung erst zum Herrscher der Welt geworden ist<sup>87</sup>, hat eine dezidiert positive Seite. Sie ergibt sich daraus, dass individuelle Schwächen in der Beurteilung von Einzelsituationen und im Schlussfolgern außen vor bleiben können. Aber schlecht umgesetzte Automatisierung kann eben auch dazu führen, dass sich gefärbte Vorannahmen und fälschlicherweise angenommene Zusammenhänge auf große Fallzahlen auswirken. Die Skaleneffekte automatisierter Analyse- und Entscheidungssysteme machen jede Form von Fehler oder Unzulänglichkeit folgenreicher. In diesem Spannungsfeld bewegt sich das gesamte Feld der Algorithmisierung in immer mehr Bereichen, angefangen bei der Schulplatzvergabe<sup>88</sup> bis hin zur medizinischen Diagnostik, und mit hin auch das PP.

## SOZIALE EINBETTUNG

*„Die aktuelle kritische Beschäftigung mit algorithmischen Bewertungs- und Entscheidungssystemen ist insofern bemerkenswert, als dass die algorithmische Revolution zu großen Teilen bereits hinter uns liegt.“<sup>89</sup>*

Großformatige negative gesellschaftliche Auswirkungen fehlerhafter

84 Nämlich gezielte Präventionsarbeit in für bestimmte Verbrechen besonders anfälligen Gebieten zu ermöglichen und dadurch die Fallzahlen zu verringern.

85 Vgl. Schweer 2015: 15.

86 Hierher gehört auch die seit langem gemachte Beobachtung des Verdrängungseffektes, d.h. einer bloßen Verlagerung quantitativ gleichbleibender Verbrechen in andere Gebiete; vgl. Bowes & Johnson 2003. In Bayern will man hingegen Erfolge des Einsatzes von *Precobs* ohne Auftreten des Verdrängungseffektes festgestellt haben; vgl. Heitmüller 2017.

87 Ausführlich dazu Knobloch 2018a.

88 Vgl. hierzu das Gemeinschaftsprojekt „Schuleinzugsbereicherechner“ von idalab GmbH, Stiftung Neue Verantwortung und der Technologiestiftung Berlin unter <https://idalab.de/intelligent-zoning-engine>

89 Mohabbat Kar & Parycek 2018: 8

algorithmischer Systeme lassen sich am Ende nur dadurch vermeiden, dass man kritische transdisziplinäre Diskurse<sup>90</sup> dazu ermöglicht. Das war das Ergebnis im Grunde aller bis hierher geführten intersektoralen Dialoge im Projekt „Algorithmen fürs Gemeinwohl“ bisher. An dieser Stelle haben PP-Projekte ebenso wie andere Automatisierungsvorhaben Nachholbedarf. Bei der Entwicklung und beim Einsatz von PP-Systemen gehen die Behörden in Deutschland nach unseren Informationen überwiegend umsichtig vor – nur sie kommunizieren dazu zu wenig, beziehungsweise lediglich auf Nachfrage.<sup>91</sup> Das hat durchaus nachvollziehbare Gründe: Sicherheitsbehörden haben Vertraulichkeit aus guten Gründen gewissermaßen „im Blut“. Ein Zuviel an öffentlich zugänglichen Informationen über die Funktionsweise der genutzten Vorhersagealgorithmen könnte dazu führen, dass die Systeme (etwa von Einbruchsbanden) ausgetrickst werden. Allerdings müssen Transparenz und Nachvollziehbarkeit bei algorithmischen Systemen, die für die breite Gesellschaft von großer Tragweite sind (hier geht es etwa um nicht weniger als eine faire Verteilung des Gutes Sicherheit), als notwendig und nicht als optionaler Zusatz angesehen werden. Untersuchungen zeigen, dass polizeiliche Arbeit durch den gezielten Einbezug von Nachbarschaften und lokalen Gruppen verbessert werden kann.<sup>92</sup>

Weil hier nicht allein der Einsatz interdisziplinärer Teams zum Ziel führt, sondern ein transdisziplinärer Rahmen auch den Einbezug von Praktiker:innen und Betroffenen vorgibt, wäre es gerade bei Projekten, die in einem engeren oder weiteren Sinne in die Kategorie Verwaltungsinnovation fallen, wünschenswert, dass man den Instrumentenkasten der transdisziplinären (Nachhaltigkeits-)Forschung<sup>93</sup> auf übertragbare Methoden, wie beispielsweise den Reallabor-Ansatz<sup>94</sup>, untersucht. Dadurch trüge man dem essentiell sozialen Charakter algorithmischer Systeme Rechnung und würde einen Teil der gesellschaftlichen Folgenabschätzung bereits in die Phase der Systementwicklung verlegen.

Der Einsatz von PP führt in der Regel dazu, dass die gleichen polizeilichen Ressourcen nun anders verteilt werden. An manchen Orten nimmt die Präsenz zu, an anderen nimmt sie ab. Manche Menschen fühlen sich durch verstärkte Polizeipräsenz sicherer, bei anderen steigt dadurch die Gefährdungswahrnehmung. Schon um solchen menschlichen Reaktionen Rechnung zu tragen, ist eine entsprechende Informationspolitik zu fordern. Kurzbezeichnungen wie „vorausschauende Polizeiarbeit“ oder „Predictive Policing“ sind dabei wenig hilfreich, stattdessen müsste so kommuniziert werden, dass sowohl zu hoch gegriffene Erwartungen als auch überzogene Ängste vermieden werden:

*„Wir ergänzen unsere Vorgangsdaten nun um weitere kriminalitätsrelevante Daten und werten sie mit maschineller Hilfe aus. Das ermöglicht uns die*

90 „Transdisziplinär“ bedeutet einerseits fachliche Interdisziplinarität und andererseits Praxisbezug sowie Praktikereinbezug.

91 Vgl. zu diesem Aspekt auch Knobloch 2018b.

92 Vgl. Hustedt 2016: 11 ff.

93 Eine aus konkreten Projektbeispielen abgeleitete Methodensammlung findet sich in Bergmann et al. 2010.

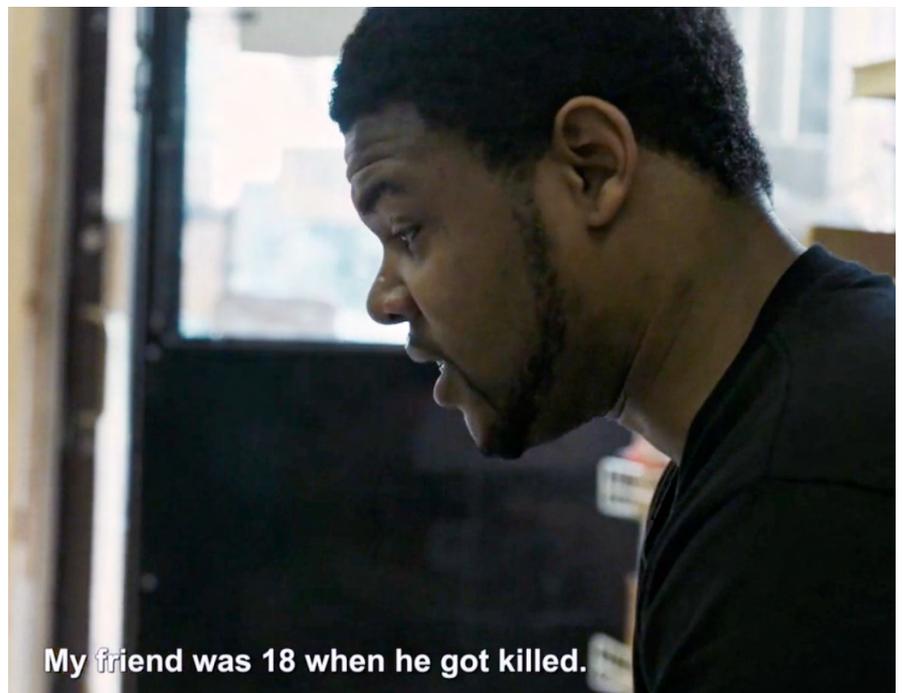
94 Vgl. Borner & Kraft 2018.

*Ermittlung statistischer Wahrscheinlichkeiten, mit denen an bestimmten Orten (eher als an anderen) zukünftig Verbrechen begangen werden. In solchen Gebieten werden wir unsere Präsenz zu Präventionszwecken in Zukunft kurzzeitig und variabel erhöhen.*<sup>95</sup>

Um mehr geht es ja im Grunde nicht, und es ist Sache der Polizeidienststellen und ihrer politischen Weisungsgeber, dies entsprechend korrekt und eingängig nach außen, das heißt in die breite Bevölkerung hinein zu vermitteln. Das ist auch deshalb geboten, weil die Auswirkungen viel weitreichender sein können als eine Änderung des Sicherheitsgefühls in der Bevölkerung. Ein nicht intendierter mittelfristiger Nebeneffekt von PP-Praxis kann beispielsweise die Veränderung von Immobilienpreisen und im Gefolge dessen langfristig auch eine Veränderung der soziodemographischen Struktur von Wohnquartieren sein. Auch solche möglichen Folgen müssen Behörden bedenken, wenn sie an PP-Strategien arbeiten.

Dieses Beispiel führt uns ein weiteres Mal zum schon mehrfach angesprochenen Thema Diskriminierungsgefahr. Die weitere Marginalisierung bereits benachteiligter Gegenden und Bevölkerungsschichten etwa wohnt der Nutzung von PP-Systemen aufgrund ihrer oben beschriebenen Design- und Datenverarbeitungsprinzipien gewissermaßen natürlicherweise inne.<sup>96</sup> Dagegen müssen auf der Ebene der Modellentwicklung, der Auswahl der Datenbasis, der Datenpflege sowie der institutionellen und sozialen Implementierung bewusst Vorkehrungen getroffen werden, sonst überwiegen am Ende die sozialen Kosten die möglichen Gewinne. Die US-Stadt New Orleans hat etwa aus eben diesem Grund die Zusammenarbeit mit dem Datenanalyseunternehmen *Palantir* beendet.<sup>97</sup>

**Abb. 6: Screenshot aus dem Dokumentarfilm „Pre-Crime“ von Monika Hielscher und Matthias Heeder, der einen der Protagonisten, Robert McDaniel aus Chicago, in einer Interview-Situation zeigt.**



95 Fiktive Verlautbarung einer ein PP-System einführenden Polizeibehörde.

96 Vgl. ACLU 2016.

97 Vgl. Winston 2018a und 2018b.

Die Absurdität und Tragik möglicher Folgen, wenn Menschen zum Gegenstand maschineller Prädiktion werden, zeigt der Fall Robert McDaniel aus Chicago. Dort geriet McDaniel auf die sogenannte „Heat List“, ein personenbezogenes PP-System, weil sein bester Freund ein Todesopfer von Gangkriminalität wurde. McDaniel selbst hatte seinerzeit lediglich kleinere Delikte begangen. Trotzdem wurde er als potentieller zukünftiger Schwerverbrecher auf die Liste genommen. Als Folge dessen wurde er regelmäßig von Polizeibeamt:innen aufgesucht, nach seiner Lebensführung befragt und davor gewarnt, sich irgendetwas zuschulden kommen zu lassen. Die Folgen für sein Sozialleben waren verheerend. Es folgte eine regelrechte soziale Ächtung, weil jemand, der von der Polizei aufgesucht und nicht verhaftet wird, als Spitzel gilt. McDaniel war aber natürlich nicht nur mit diesem, ihm zusätzlich durch die Heat List eingetragenen Problem konfrontiert, sondern auch mit den natürlicheren Folgen des Todes seines Freundes, beispielsweise der Übernahme der Vaterrolle gegenüber dessen Sohn. Dies ist nur ein Beispiel für eine nicht intendierte Folge des Einsatzes von Analyse- und Überwachungstechnologie, das vor Augen führt, wie wichtig es ist, die Nutzung von Technologie im gesellschaftlichen Gesamtkontext zu betrachten.

Es ist wichtig sich zum jetzigen Zeitpunkt hierzulande solche Fälle vor Augen zu führen, weil nicht ausgeschlossen werden kann, dass auch die deutsche Polizei (wie sie es beispielsweise bei der Fankriminalität längst tut) in Zukunft mit individuellen Personenprofilen arbeitet und Prognosemodelle darauf anwendet. Die kritischen Fragen müssen vor beziehungsweise bei der Entwicklung und Implementierung algorithmischer Systeme gestellt werden; „code first, fix later“ ist kein gangbarer Weg.

## DISKUSSION UND AUSBLICK

*„Whenever a reformer proposes to solve an algorithmic-sorting problem with ‚more data‘, that is an invitation to more surveillance, which can go to absurd lengths.“<sup>98</sup>*

Natürlich ist die Verlockung groß, angesichts eines womöglich hohen technischen und finanziellen Aufwandes zu suggerieren, es werde mit PP Rocket Science betrieben und sämtliche Kriminelle müssten sich fortan warm anziehen. Insbesondere auf Politiker übt dieses Narrativ Anziehungskraft aus. Doch nach allem, was wir bisher wissen, gibt es keine technische Wunderwaffe gegen Kriminalität, es existiert keine Glaskugel, die uns diesbezüglich alles verrät.<sup>99</sup> Jedenfalls nicht auf Grundlage des gesetzlich Abgesicherten, und wir sollten uns hüten, diese Grundlage zu verlassen oder infrage zu stellen.

PP stellt, insbesondere in der in Deutschland derzeit praktizierten Form, an sich keine Gefahr für Bürgerrechte dar, die Unschuldsvermutung als

---

98 Pasquale 2018

99 Vgl. Moraff 2016.

rechtstaatliches Prinzip wird durch *Precobs*, *SKALA* u.a. im Gebrauch befindliche Systemen nicht à la *Minority Report*<sup>100</sup> infrage gestellt. Das eigentlich Gefährliche sind Kombinationen von Überwachungstechniken und das Zusammenführen von Datenbanken und Informationsquellen (Pervasive Surveillance). In den USA lässt sich gut beobachten, welche Möglichkeiten und weiteren Begehrlichkeiten aufseiten der Polizei sich aus der Nutzung von PP-Techniken im Zusammenspiel mit verbreiteter Kameraüberwachung des öffentlichen Raumes, Kennzeichenlesern und Handyortung ergeben können.<sup>101</sup> Doch diese Entwicklung ist nicht auf das Ausland beschränkt,<sup>102</sup> auch hierzulande ist sie festzustellen. Die Ausweitung von Polizeibefugnissen (vgl. die Polizeigesetzesnovellen in Bayern und Nordrhein-Westfalen) in Kombination mit der immer effizienteren Überwachung des öffentlichen Raumes auch durch PP hat natürlich Einfluss auf den Schutzraum bürgerlicher Freiheiten. In der Phantasie mancher Ordnungshüter wird dieser Raum kleiner und enger. Allumfassendes Wissen ermöglicht demnach allumfassende Kontrolle.<sup>103</sup> Konkret bedeutet das den Einbezug von immer neuen Informationsquellen und das Zusammenführen von Datensätzen. Wenn etwa bekannt wird, dass Hessen, das die selbst entwickelte PP-Software *KLB-operativ* bereits flächendeckend einsetzt, zugleich in einen Vertrag mit dem US-Datenanalyseunternehmen *Palantir* eingetreten ist,<sup>104</sup> dann ist das keine beruhigende Nachricht für Bürger- und Menschenrechtler:innen. Dieser Vorgang ist umso erstaunlicher, da sich die ersten US-Kommunen aus eben jenem Grund der Wahrung von Bürgerrechten wieder aus der Zusammenarbeit mit dem in der Kritik stehenden Unternehmen zurückgezogen haben.<sup>105</sup>

Hinzu kommt ein weiterer Aspekt. Wie wir an anderer Stelle ausführlich gezeigt haben<sup>106</sup>, bestehen Gefahren für die Privatsphäre nicht nur dort, wo personenbezogene Daten verarbeitet werden. Kommerzielle Anbieter wie das US-amerikanische *PredPol* scheinen solche Group-Privacy-Effekte in ihrem Marketing verschleiern oder zumindest nicht nennen zu wollen. So heißt es auf der *PredPol*-Website: „No personally identifiable information is ever used. No demographic, ethnic or socio-economic information is ever used. This eliminates the possibility for privacy or civil rights violations seen with other intelligence-led policing models.“<sup>107</sup> Diese Aussage ist so nicht haltbar. Natürlich kann die Auswertung von nicht-personenbezogenen

100 Der Bezug zu diesem Hollywood-Film von 2002 taucht in der Debatte um PP bis heute immer wieder auf, obwohl er offensichtlich verfehlt ist. Im Film sagen zwei hellsehende Wesen zukünftige Verbrechen voraus, woraufhin die in der Zukunft mutmaßlich handeln werdenden Personen präemptiv verhaftet werden. Das hat der Sache nach nichts mit hierzulande aktuell eingesetzten PP-Systemen zu tun.

101 Vgl. Fasmann 2018.

102 Das 5 Millionen Euro schwere EU-Forschungsprojekt (HORIZON 2020) „CONNEXIONS“ (Projekt-ID 786731), an dem auch das Bayerische Staatsministerium des Innern sowie die Hochschule für den öffentlichen Dienst in Bayern beteiligt sind, ist ein guter Beleg für die Bemühungen, die auch in Europa in Richtung Datenintegration und ubiquitäre Überwachung zielen. In der Projektbeschreibung heißt es: „CONNEXIONS aims to develop and demonstrate next-generation detection, prediction, prevention, and investigation services. These services will be based on multidimensional integration and correlation of heterogeneous multimodal data, and delivery of pertinent information to various stakeholders in an interactive manner tailored to their needs, through augmented and virtual reality environments.“ – Quelle: [https://cordis.europa.eu/project/rcn/214843\\_en.html](https://cordis.europa.eu/project/rcn/214843_en.html)

103 Vgl. Diehl & Kartheuser 2018.

104 Vgl. Ulrich 2018.

105 Vgl. Winston 2018b.

106 Vgl. Manske & Knobloch 2017.

107 PredPol 2018

Daten negative Auswirkungen auf Gruppen und – wenn kombiniert mit anderen Datenquellen – sogar auf einzelne Personen haben. Bürgerrechte können durch einen schleichenden Prozess in Richtung umfassender staatlicher Überwachung sehr wohl ausgehebelt werden. Genau hier liegen auch die mittel- bis langfristigen Gefahren des PP in Deutschland, wo die Systeme tatsächlich noch keine Erweiterungen um Gefährderlisten oder Ähnliches erfahren haben.

Aber, so könnte man an dieser Stelle etwas ketzerisch fragen, was soll man von einem Unternehmen erwarten, das sich zwar um einen wissenschaftlich fundierten Anstrich bemüht, aber US-Gemeinden gleichzeitig Knebelverträge aufgezwungen und vor den Marketingkarren gespannt hat? US-Kunden von *PredPol* haben sich anscheinend vertraglich dazu verpflichtet, unabhängig von den eigenen Einsatzerfahrungen mit der Software (!) sich öffentlich stets nur positiv darüber zu äußern und anderen Städten und Gemeinden zur Anschaffung des Produkts zu raten.<sup>108</sup> Ein Vorgang, der in dieser Form hierzulande vielleicht aus kulturellen und rechtlichen Gründen weder möglich noch denkbar wäre. Allerdings zeigt die oben erwähnte Zusammenarbeit Hessens mit *Palantir* zum vermeintlichen Nulltarif für das Bundesland<sup>109</sup> auch, dass die Unterschiede so groß am Ende auch wieder nicht sind.

Die derzeit zu beobachtende Systemvielfalt in den Bundesländern stellt keinen Nachteil dar. Im Gegenteil kann darin sogar ein heuristischer Vorteil bei der Ermittlung der gemeinwohldienlichsten Vorgehensweise im Bereich einer noch jungen Technologie liegen. Mittelfristig ist aber aus bundespolitischer Sicht darauf zu achten, dass sich Standards nicht zu sehr auseinander entwickeln<sup>110</sup>, und dass die Verteilung des Gutes öffentliche Sicherheit als Folge nicht zu uneinheitlich ausfällt. Beim Thema Standards wäre es wünschenswert, dass sich folgende Prinzipien für die Auswahl und den Einsatz von PP-Systemen durchsetzen:<sup>111</sup>

- offenes System mit transparenter Funktionsweise (keine Blackbox)
- einfache Bedienbarkeit in den Behörden für alle Dienststränge
- Datenhoheit und zur Ergänzung von (nicht öffentlich verfügbaren) Polizeidaten vorzugsweise Nutzung von Open Government Data (z.B. Geodaten und soziodemographische Daten)

Als weiteres Qualitätssicherungsmerkmal könnten die Hersteller von PP-Systemen zu Fairness- und Nichtdiskriminierungstests und entsprechenden Dokumentationen verpflichtet werden, wie sie bei besonders sicherheitsrelevanten Systemen (z.B. Flugzeugautopiloten oder Bremssystemen) die Regel sind.<sup>112</sup>

<sup>108</sup> Vgl. Bond-Graham 2013.

<sup>109</sup> Was natürlich Fragen nach dem Nutzen für das Unternehmen (vgl. Borchers 2018), beziehungsweise nach den aus den USA bekannten Vermarktungsstrategien (vgl. Harris 2017) aufwirft.

<sup>110</sup> Dies trifft insbesondere auf die Daten zu. Bei einer massiven Zersplitterung der Datenhaltung und des Datenzugriffs steigt das Risiko von Fehlern und unentdeckter Zusammenhänge. Vgl. Kartheuser & Jaume-Palasi 2018 (am Ende des Interviews).

<sup>111</sup> Diese decken sich teilweise mit den Gestaltungsprinzipien des Projekts *SKALA* in Nordrhein-Westfalen, vgl. LKA NRW 2018c: 9.

<sup>112</sup> Diese Anregung stammt von Programmierern der Firma IBM, die sich damit auf die bei kritischen Systemen übliche *Supplier's Declaration of Conformity – SdoC* beziehen. Vgl. Robitzski 2018.

Es sind politische Entscheidungen, die darüber entscheiden, ob, wie und wozu genau PP-Systeme eingesetzt werden sollen. Die Prozesse, die zu diesen Entscheidungen führen, sollten maximal transparent sein, und dazu sollte zusätzlich ein gesellschaftlicher Deliberationsprozess – als dessen Teil dieses Papier zu verstehen ist – kommen. Denn Vertrauen ist ein elementar wichtiges Gut, wenn wir über den Erfolg eines für die öffentliche Sicherheit so wichtigen Instruments wie PP sprechen. Letztlich geht es dabei um die Transparenz sozio-technischer Prozesse an den für die Gesellschaft relevantesten Stellen, nämlich dort, wo die öffentliche Hand Systeme zur Wahrung oder Herstellung der öffentlichen Sicherheit entwickelt. Resa Mohabbat Kar und Peter Parycek bringen das in der Einleitung zu ihrem Sammelband über algorithmische Systeme im staatlichen Einsatz treffend so auf den Punkt:

*„Der gesamte soziotechnische Prozess, von der Problem- und Zieldefinition über die Entwicklung bis zur Implementierung kann demnach Interessenabwägungen und Werturteile beinhalten. Insbesondere beim Einsatz im öffentlichen Sektor wird dieser Umstand dort problematisch, wo diese Werturteile und Interessenabwägungen implizit, nicht offensichtlich oder schlicht unsichtbar sind, da sie, in Software eingebettet, aus dem Deliberationsprozess genommen wurden. Diese impliziten Wertsetzungen gilt es zu identifizieren und explizit zu machen, damit sie in den dafür vorgesehenen, legitimierten Prozessen entschieden werden können.“<sup>113</sup>*

Insgesamt ist hierzulande sowohl auf Softwareanbieter- als auch auf Behördenseite ein erfreulich reflektiertes Vorgehen bei der Einführung von PP-Systemen zu konstatieren. Gleichwohl bleiben die hier herausgearbeiteten aktuellen Defizite und Gefahren mit Blick in die Zukunft bestehen. Polizeiarbeit wird sich durch den Einsatz von PP auch strukturell wandeln. Der in der öffentlichen Verwaltung generell zu verzeichnende Trend zum Einsatz von Datenanalysewerkzeugen, um Ressourcen zu sparen, wird das Berufsbild und Selbstverständnis des Polizeibeamten verändern. Das hat durchaus auch seine guten Seiten, wie hier immer wieder dargestellt wurde. Beispielsweise werden dadurch tradierte (aber vielleicht wenig nützliche oder sogar schädliche) Praktiken der Lagebilderstellung und Einsatzplanung natürlicherweise erst einmal infrage gestellt. Die Attraktivität von PP-Systemen wird jedenfalls aller Wahrscheinlichkeit nach nicht abnehmen, weshalb es sich lohnt, abschließend einige allgemeine Empfehlungen für ihren Einsatz auszusprechen.

---

113 Mohabbat Kar & Parycek 2018: 33

# EMPFEHLUNGEN

## NUTZUNGSZWECKE UND -GRENZEN VORAB KLAR FESTLEGEN

- Politische Entscheidungen und gesellschaftlicher Dialog haben Vorrang vor behördeneigenen Überlegungen und Zwecksetzungen.
- Die intendierten Zwecke von PP-Systemen sind klar zu benennen (z.B. Rückgang von Wohnungseinbruchsfallzahlen); auf der anderen Seite ist explizit auszuschließen, in welche Richtung die Technologie (auch in Zukunft) nicht weiterentwickelt werden sollte (z.B. kein personenbezogenes PP).
- Auch indirekt erzielbare Ergebnisse (z.B. hinsichtlich einer Verbesserung interner Prozesse) können und sollten gegebenenfalls in die Liste der Zielsetzungen aufgenommen werden.
- Tatsächliche, mögliche und intendierte Verbindungen zu anderen Bereichen staatlicher Überwachung und datenanalysegestützten behördlichen Sicherheitsmaßnahmen sollten – wo keine klare Geheimschutzrelevanz vorliegt – benannt und diskutiert werden.

## INTERDISZIPLINÄRE EXPERTISE AUFBAUEN UND TRANSDISZIPLINÄR ARBEITEN

- Neben Kriminolog:innen, Informatiker:innen und Data Scientists müssen auch Datenschützer:innen, Menschenrechtler:innen, Ethiker:innen, Soziolog:innen und Sozialpsycholog:innen von Beginn an in PP-Projekte einbezogen werden. Für diesen Einbezug gibt es mannigfache Möglichkeiten: natürlich kann und muss entsprechende Expertise nicht überall in-house aufgebaut werden.
- Praktische Endanwender:innen von PP-Systemen sollten in der einen oder anderen Weise unbedingt von Anfang an in die Systementwicklung und -implementierung eingebunden sein.
- Diejenigen Personenkreise, die von PP in der Praxis mutmaßlich besonders betroffen sind (z.B. Anwohner:innen von sogenannten „Problembezirken“), sollten von Anfang an einbezogen werden.

## BEREITS BESTEHENDE DATENERFASSUNGS- UND -VERARBEITUNGSSYSTEME HINTERFRAGEN UND EVENTUELL MODIFIZIEREN

- PP-Systeme sind in der Regel Erweiterungen bestehender IT-Systeme; diesem Umstand sollte dadurch Rechnung getragen werden, dass PP-Systeme nicht isoliert betrachtet und eingeführt werden.

- Bestehende Strukturen, Prozesse und Systeme des Datenmanagements (z.B. die entsprechenden Vorgangsbearbeitungssysteme) sind besonders von der Einführung eines PP-Systems betroffen; sie sollten in diesem Zuge unbedingt kritisch hinterfragt und gegebenenfalls modifiziert werden.
- Diskriminierungsfreie Trainingsdatenpools sind für lernende Systeme essenziell; Polizeibehörden und ihre Dienstleister:innen, die welche angelegt oder von anderen angelegte genutzt haben, sollten sie anderen Behörden und Dienstleister:innen zugänglich machen (sofern sie die entsprechenden Rechte dazu besitzen).

### **EMPIRISCHE WIRKUNGSFORSCHUNG FÖRDERN UND EVALUIERUNGSERGEBNISSE ZUGÄNGLICH MACHEN**

- Effekte, die nicht-intendiert waren, aber dennoch eingetreten sind, müssen erfasst und in die entsprechenden Wirkungszusammenhänge eingeordnet werden.
- An gezielten Kontrollbereichsexperimenten mangelt es – vor allem in der Langzeitbetrachtung – nach wie vor.
- Systembedingte Verzerrungen (z.B. Überbestreifung von Quartieren) müssen aufgedeckt und entsprechende Gegenmaßnahmen (z.B. Zufallsparemeter) entwickelt werden.
- Mögliche soziodemographische und andere Veränderungen (z.B. hinsichtlich des Sicherheitsgefühls der Bevölkerung) in Bezirken, die nach den Empfehlungen von PP-Systemen bestreift werden, sollten – möglichst auch in der Langzeitbetrachtung – beobachtet und ausgewertet werden.
- Sämtliche Evaluierungsergebnisse sollten sich die Behörden – auch unterhalb der Ebene veröffentlichter Berichte und über Bundeslandgrenzen hinweg – untereinander zugänglich machen.

### **SYSTEMVIELFALT ALS VORTEIL NUTZEN UND STANDARDS ERWÄGEN**

- Versuche und Erfahrungen mit unterschiedlichen technischen und konzeptionellen Ansätzen in den verschiedenen Bundesländern sollten wechselseitig zugänglich gemacht und ausgetauscht werden, wobei der Bund seiner (wenn in diesem Bereich auch eingeschränkten) Zentralisierungsfunktion gerecht werden sollte.
- Die Formulierung von Standards auf System- und Datenebene sollte an möglichen und sinnvollen Stellen geprüft werden, insbesondere bei der

polizeilichen Datenerfassung und bei der Modellierung; bundeslandübergreifende Standards (z.B. hinsichtlich von Auswahlfeldern in den Vorgangsbearbeitungssystemen) sind wünschenswert, wobei eine vollständige Vereinheitlichung weder Selbstzweck sein kann noch in unserem föderalen System praktisch möglich ist.

- Entwicklungen auf internationaler Ebene sollten im Auge behalten, analysiert und gegebenenfalls für die eigenen Zwecke fruchtbar gemacht werden.

### **INTERNE KOMMUNIKATION ERNST NEHMEN UND PASSGENAUE MITARBEITERSCHULUNGEN ANBIETEN**

- Die Breite der Polizeibelegschaft auf allen Dienstrangebenen muss einbezogen werden, je früher desto besser. Dabei kommt es auf klare Kommunikation insbesondere im Hinblick auf die Ziele und Zwecke an (vgl. Empfehlung 10.1 oben).
- Der behördliche Erfahrungsaustausch auf Landesebene sollte intensiv sein; dies trifft insbesondere in Phasen zu, in denen Pilotbetriebe in einzelnen Direktionen und Dienststellen stattfinden. Der frühzeitige (wenn auch vielleicht nur informationelle) Einbezug von anfangs nicht betroffenen Dienststellen ist für die spätere landesweite Einführung zentral.
- Für ältere Kolleg:innen und andere, typischerweise eher skeptische Funktionsträger:innen (in der Regel untere Dienstränge) müssen spezifische Kommunikationsformate entwickelt und angewendet werden; hierbei empfiehlt sich wiederum ein intensiver bundeslandübergreifender Erfahrungsaustausch.
- Bei der internen Kommunikation kommt es vor allem darauf an, PP in die Reihe der polizeilichen Instrumente einzuordnen und dadurch realistische Erwartungen an die möglichen Erfolge zu wecken, während das Aufkommen überzogener Erwartungen zu vermeiden ist.

### **AKTIVE EXTERNE KOMMUNIKATION BETREIBEN**

- PP sollte als Systemerweiterung und weniger als revolutionärer Umbruch vermittelt werden; algorithmische Analyse- und Entscheidungssysteme kommen nicht isoliert in ein Vakuum hinein, sondern werden in eine bestehende Infrastruktur integriert.
- Es ist deutlich zu machen, dass PP-Systeme keine Autonomie besitzen, sondern nur die Aufgaben und Ziele ausführen, die Menschen ihnen (zuvor wohlüberlegt) vorgeben. Ferner handelt es sich um Assistenzsysteme, welche die polizeiliche Arbeit (bei der Erstellung des Lagebilds und bei der Einsatzplanung) unterstützen, sie aber nicht ersetzen.

- Polizeibehörden sollten sich die Chance bewusst machen, dass sie durch die Offenlegung der technischen und strukturellen Zusammenhänge von PP-Systemen mit bestehenden Systemen zur Unterstützung der polizeilichen Präventions- und Ermittlungsarbeit einen wichtigen Beitrag zur gesellschaftlichen Demystifizierung algorithmischer Systeme und Künstlicher Intelligenz leisten können.

Neben der Beachtung dieser sektorspezifischen Hinweise gibt es inzwischen eine Reihe an allgemeinen Designprinzipien für Künstliche Intelligenz und ebenso allgemeine ethisch-humanistische Gütekriterien für den Entwicklungsprozess algorithmischer Systeme.<sup>114</sup> Aus der Perspektive des Gemeinwohls ist es wünschenswert, dass solche Prinzipien und Kriterien bei der Entwicklung und Implementierung von PP-Systemen ebenfalls Berücksichtigung finden.<sup>115</sup>

---

114 Laura James hat bei Medium einige zusammengestellt: <https://medium.com/doteveryone/oaths-pledges-and-manifestos-a-master-list-of-ethical-tech-values-26e2672e161e> Die Liste stellt eine unvollständige Momentaufnahme dar, denn die Fülle an Ansätzen ist schon jetzt kaum noch zu übersehen, während beinahe wöchentlich neue hinzukommen. Der am MIT Media Lab entwickelte Ansatz *Society-in-the-Loop – SITL* beziehungsweise *Algorithmic Social Contract* (Rahwan 2017) ist zur Einordnung in den ideengeschichtlichen Kontext eine Betrachtung wert.

115 Im ersten Quartal 2019 wird die *Bertelsmann-Stiftung* einen mit der NGO *iRights* in einem partizipativen Stakeholderdialog entwickelten Gütekriterienkatalog für den Algorithmeinsatz vorstellen, vgl. Hustedt et al. 2018. Ein schöner Vergleich von drei bekannten Gütekriterienkatalogen findet sich in Rohde 2018.

# LITERATUR

- American Civil Liberty Union – ACLU (2016). Statement of Concern about Predictive Policing by ACLU and 16 Civil Rights Privacy, Racial Justice, and Technology Organizations. Online unter <https://www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice>.
- Angwin, J. & Larson, J. (2016): Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say. In: ProPublica vom 30.12.2016, online unter <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>.
- Angwin, J.; Larson, J.; Mattu, S.; Kirchner, S. (2016): Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks. In: ProPublica vom 23.5.2016, online unter <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Azavea Inc. (2015): HunchLab: Under the Hood. White Paper, online unter <https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf>.
- Balogh, D.A. (2016): Near Repeat-Prediction mit PRECOBS bei der Stadtpolizei Zürich. In: Kriminalistik 2016, S. 335-341.
- Bergmann, M.; Jahn, T.; Knobloch, T.; Krohn, W.; Pohl, C.; Schramm, E. (2010): Methoden transdisziplinärer Forschung. Ein Überblick mit Anwendungsbeispielen. Frankfurt am Main: Campus.
- Bode, F.; Stoffel, F.; Keim, D. (2017): Variabilität und Validität von Qualitätsmetriken im Bereich von Predictive Policing. Konstanzer Online-Publikations-System (KOPS), online unter <http://nbn-resolving.de/urn:nbn:de:bsz:352-0-402496>
- Bond-Graham, D. (2013): All Tomorrow's Crimes: The Future of Policing Looks a Lot Like Good Branding. In: San Francisco Weekly vom 30.10.2013, online unter <http://www.sfweekly.com/news/all-tomorrows-crimes-the-future-of-policing-looks-a-lot-like-good-branding/>.
- Borchers, D. (2018): Hessen will mit Palantir-Software islamistischen Terrorismus und organisierte Kriminalität bekämpfen. In: heise online vom 6.4.2018, online unter <https://www.heise.de/newsticker/meldung/Hessen-will-mit-Palantir-Software-islamistischen-Terrorismus-und-organisierte-Kriminalitaet-4012382.html>.
- Borner, J. & Kraft, A.H. (2018): Konzeptpapier zum Reallabor Ansatz. Arbeitspapier aus dem Kopernikusprojekt ENavi, gefördert durch das BMBF. Berlin: Kolleg für Management und Gestaltung nachhaltiger Entwicklung (KMGNE). Online unter: [https://www.researchgate.net/publication/323320642\\_Konzeptpapier\\_zur\\_Reallabor-Methode\\_im\\_ENavi-Projekt](https://www.researchgate.net/publication/323320642_Konzeptpapier_zur_Reallabor-Methode_im_ENavi-Projekt)
- Bowes, K.J. & Johnson, S.D. (2003): Measuring the Geographical Displacement and Diffusion of Benefit Effects of Crime Prevention Activity. In: Journal of Quantitative Criminology 19/3, S. 275-301.
- Bünau, P. (2018): Algorithmen und Transparenz. Unsere Sehnsucht nach Begründung. In: The Legal Revolutionary (LR) 2018, S.98-107. Online unter [https://www.legal-revolution.com/images/pdf/Algorithmen\\_und\\_Transparenz.pdf](https://www.legal-revolution.com/images/pdf/Algorithmen_und_Transparenz.pdf).
- Bundesministerium des Innern – BMI (2018): Niedrigste Zahl an verübten Straftaten seit 1992. Pressemitteilung vom 8.5.2018 zur Kriminalitätsstatistik 2017. Online unter [https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2018/05/pks-und-pmk-2017.html;jsessionid=E62COD97ABDF8CC5D12E1405BDC9FED1.2\\_cid364](https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2018/05/pks-und-pmk-2017.html;jsessionid=E62COD97ABDF8CC5D12E1405BDC9FED1.2_cid364).
- Diehl, J. & Kartheuser, B. (2018): Predictive Policing: Ich weiß, was du heute tun wirst. In: Spiegel Online vom 27.1.2018, online unter <http://www.spiegel.de/panorama/justiz/kriminalitaet-in-deutschland-polizei-setzt-auf-computer-vorhersagen-a-1188350.html>.
- Egbert, S. (2018): About Discursive Storylines and Techno-Fixes: The Political Framing of the Implementation of Predictive Policing in Germany. In: European Journal for Security Research 2018, S.1-20. Online unter <https://doi.org/10.1007/s41125-017-0027-3>.
- Fasman, J. (2018): I know what you'll do next summer. In: The Economist, Technology Quarterly vom 2.6.2018, online unter <https://www.economist.com/technology-quarterly/2018-05-02/justice>.
- Florida, R. (2003): The Rise of the Creative Class. And How It's Transforming Work, Leisure, Community and Everyday Life. New York: Basic Books.
- Gerstner, D. (2017): Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl. Evaluationsergebnisse zum Baden-Württembergischen Pilotprojekt P4. In: Max-Planck-Institut für ausländisches und internationales Strafrecht, forschung aktuell Nr.50, online unter [https://www.mpicc.de/files/pdf4/rib\\_50\\_gerstner\\_2017.pdf](https://www.mpicc.de/files/pdf4/rib_50_gerstner_2017.pdf).
- Gluba, A. (2014): Predictive Policing – eine Bestandsaufnahme. Historie, theoretische Grundlagen, Anwendungsgebiete und Wirkung. Landeskriminalamt Niedersachsen. Hannover.
- Gluba, A. (2016): Mehr offene Fragen als Antworten. Was für eine Bewertung des Nutzens von Predictive Policing noch zu klären ist. In: Die Polizei 107, S. 53-57.
- Graff, B. (2016): Rassistischer Chat-Roboter: Mit falschen Werten bombardiert. In: SZ Online vom 3.4.2016, online unter <https://www.sueddeutsche.de/digital/microsoft-programm-tay-rassistischer-chat-roboter-mit-falschen-werten-bombardiert-1.2928421>
- Hankey, S. & Tuszynski, M. (2017): Efficiency and Madness. Using Data and Technology to Solve Social, Environmental and Political Problems. In: Heinrich Böll Stiftung, online unter [https://www.boell.de/sites/default/files/efficiency\\_and\\_madness.pdf](https://www.boell.de/sites/default/files/efficiency_and_madness.pdf).

- Harris, M. (2017): How Peter Thiel's Secretive Data Company Pushed Into Policing. In: Wired vom 8.9.2017, online unter <https://www.wired.com/story/how-peter-thiels-secretive-data-company-pushed-into-policing/>.
- Heitmüller, U. (2017): Predictive Policing: Die deutsche Polizei zwischen Cyber-CSI und Minority Report. In: heise online vom 17.4.2017, online unter <https://www.heise.de/newsticker/meldung/Predictive-Policing-Die-deutsche-Polizei-zwischen-Cyber-CSI-und-Minority-Report-3685873.html>
- Hunt, P.; Saunders, J.; Hollywood, J.S. (2014): Evaluation of the Shreveport Predictive Policing Experiment. Santa Monica: RAND Corporation. Online unter <https://www.ncjrs.gov/pdffiles1/nij/grants/248883.pdf>.
- Hustedt, C. (2016): A Public Value Perspective on Predictive Policing in the US. Research Paper, online unter [https://www.researchgate.net/publication/327155597\\_A\\_Public\\_Value\\_Perspective\\_on\\_Predictive\\_Policing\\_in\\_the\\_US](https://www.researchgate.net/publication/327155597_A_Public_Value_Perspective_on_Predictive_Policing_in_the_US).
- Hustedt, C.; Rohde, N.; Puntschuh, M. (2018): Auf dem Weg zu Gütekriterien für den Algorithmeneinsatz. Blogbeitrag auf [algorithmenethik.de](http://algorithmenethik.de) vom 1.8.2018, online unter <https://algorithmenethik.de/2018/08/01/auf-dem-weg-zu-guetekriterien-fuer-den-algorithmeneinsatz/>.
- Kartheuser, B. (2018): Predictive Policing in den USA: Kontrolle ist gut, Überwachung ist besser. In: Spiegel Online vom 27.1.2018, online unter <http://www.spiegel.de/panorama/justiz/predictive-policing-in-los-angeles-kontrolle-ist-gut-ueberwachung-ist-besser-a-1188578.html>.
- Kartheuser, B. & Jaume-Palasi, L. (2018): Forscherin zu Predictive Policing: „Gefahr der Diskriminierung“. Interview vom 27.1.2018. In: Spiegel Online, online unter <http://www.spiegel.de/panorama/justiz/predictive-policing-wo-die-gefahren-in-der-arbeit-mit-der-neuen-software-liegen-a-1189340.html>.
- Knobloch, T. (2018a): Algorithmen steuern die Welt. Plädoyer für eine positive Interpretation. In: Beobachtungen zur Zeit, Nr. 19. Hamburg: M.M.Warburg & Co. Online unter <https://medium.com/@tobiasknobloch/algorithmen-steuern-die-welt-97b6353bf81a>.
- Knobloch, T. (2018b): Vorausschauende Polizeiarbeit: mit Algorithmen „vor die Lage kommen“. Blogbeitrag auf [algorithmenethik.de](http://algorithmenethik.de) vom 14.6.2018, online unter <https://algorithmenethik.de/2018/06/14/vorausschauende-polizeiarbeit-mit-algorithmen-vor-die-lage-kommen/>.
- Knobloch, T. & Manske, J. (2016): Das Datenzeitalter gestalten. Offene Daten sind der Schlüssel. Policy Brief der Stiftung Neue Verantwortung vom 11.7.2016, online unter [https://www.stiftung-nv.de/sites/default/files/snv\\_datenzeitalter-gestalten\\_7.7.2016.pdf](https://www.stiftung-nv.de/sites/default/files/snv_datenzeitalter-gestalten_7.7.2016.pdf)
- Kriminologisches Forschungsinstitut Niedersachsen e.V. (2016): Schriftliche Stellungnahme zum Innenausschuss des Landtags Nordrhein-Westfalen vom 7.10.2016. Maßnahmenpaket zur Bekämpfung des Wohnungseinbruchsdiebstahls (Antrag der Fraktion der CDU, Drucksache 16/12344). Online unter <https://www.landtag.nrw.de/Dokumentenservice/portal/WWW/dokumentenarchiv/Dokument/MMST16-4315.pdf>.
- Krüger, J. & Lischka, K. (2018): Damit Maschinen den Menschen dienen. Lösungsansätze, um algorithmische Prozesse in den Dienst der Gesellschaft zu stellen. Impuls Algorithmenethik Nr.6 der Bertelsmann Stiftung. Gütersloh. Online unter [https://algorithmenethik.de/wp-content/uploads/sites/10/2018/05/Algorithmenethik\\_Lösungsportfolio\\_final\\_online.pdf](https://algorithmenethik.de/wp-content/uploads/sites/10/2018/05/Algorithmenethik_Lösungsportfolio_final_online.pdf).
- Landeskriminalamt Hamburg (2018): Prädiktionspotenzial schwere Einbruchskriminalität. Projektstand Februar 2018. Online unter <https://www.polizei.hamburg/contentblob/10651848/ee061219712856c2dff60f29d205208/data/projektstand-02-2018-do.pdf>.
- Landeskriminalamt NRW (2018a): Abschlussbericht Projekt SKALA – Kurzfassung. Düsseldorf. Online unter [https://lka.polizei.nrw/sites/default/files/2018-06/180208\\_Abschlussbericht\\_SKALA\\_Kurzfassung.pdf](https://lka.polizei.nrw/sites/default/files/2018-06/180208_Abschlussbericht_SKALA_Kurzfassung.pdf).
- Landeskriminalamt NRW (2018b): Kooperative Evaluation des Projektes SKALA. Abschlussbericht der Zentralstelle Evaluation beim LKA NRW (ZEVA) und der Gesellschaft für innovative Sozialforschung und Sozialplanung e.V. Bremen (GISS). Kurzfassung des Endberichtes. Düsseldorf. Online unter [https://lka.polizei.nrw/sites/default/files/2018-06/160131\\_Evaluationsbericht\\_SKALA\\_Kurzfassung.pdf](https://lka.polizei.nrw/sites/default/files/2018-06/160131_Evaluationsbericht_SKALA_Kurzfassung.pdf).
- Landeskriminalamt NRW (2018c): Abschlussbericht Projekt SKALA. Düsseldorf. Online unter [https://lka.polizei.nrw/sites/default/files/2018-06/180208\\_Abschlussbericht\\_SKALA.pdf](https://lka.polizei.nrw/sites/default/files/2018-06/180208_Abschlussbericht_SKALA.pdf).
- Landeskriminalamt NRW (2018d): Kooperative Evaluation des Projektes SKALA. Abschlussbericht der Zentralstelle Evaluation beim LKA NRW (ZEVA) und der Gesellschaft für innovative Sozialforschung und Sozialplanung e.V. Bremen (GISS). Düsseldorf. Online unter [https://lka.polizei.nrw/sites/default/files/2018-06/160430\\_Evaluationsbericht\\_SKALA.pdf](https://lka.polizei.nrw/sites/default/files/2018-06/160430_Evaluationsbericht_SKALA.pdf).
- Manning, P. (2014): „Evaluation“. In: Encyclopedia of Criminology and Criminal Justice, hgg. v. Bruinsma, G. & Weisburd, D., New York: Springer, S. 2501-2513.
- Manske, J. & Knobloch, T. (2017): Datenpolitik jenseits von Datenschutz. Policy Brief der Stiftung Neue Verantwortung vom 2.11.2017, online unter [https://www.stiftung-nv.de/sites/default/files/datenpolitik\\_jenseits\\_von\\_datenschutz\\_0.pdf](https://www.stiftung-nv.de/sites/default/files/datenpolitik_jenseits_von_datenschutz_0.pdf).
- Matinde, V. (2015): Why We Need Predictive Policing in Africa. In: Hai Afrika, online unter <http://haiafrika.com/why-we-need-predictive-policing-in-africa/>.
- Mohabbat Kar, R. & Parycek, P. (2018): Berechnen, ermöglichen, verhindern: Algorithmen als Ordnungs- und Steuerungsinstrumente in der digitalen Gesellschaft. In: (Un)berechenbar? Algorithmen und

- Automatisierung in Staat und Gesellschaft, hgg. v. Mohabbat Kar, R.; Thapa, B.; Parycek, P., Berlin: Kompetenzzentrum Öffentliche IT, Fraunhofer-Institut für offene Kommunikationssysteme, S. 7-38. Online unter <https://www.oeffentliche-it.de/documents/10181/76866/1+Mohabbat+Kar+%26+Parycek+-+Editorial.pdf>.
- Mohler, G. O.; Short, M. B.; Malinowski, Sean; Johnson, Mark; Tita, G. E.; Bertozzi, Andrea L.; Brantingham, P. J. (2015): Randomized controlled field trials of predictive policing. In: Journal of the American Statistical Association, online unter [http://teamcore.usc.edu/projects/security/Muri\\_publications/Short\\_JASA\\_2015.pdf](http://teamcore.usc.edu/projects/security/Muri_publications/Short_JASA_2015.pdf).
- Montag, T. (2016): Der Algorithmus des Verbrechens. Potenzial und Grenzen von „Predictive Policing“. In: Konrad-Adenauer-Stiftung, Reihe Analysen & Argumente, Nr. 2015 (September 2016). Online unter [http://www.kas.de/wf/doc/kas\\_46279-544-1-30.pdf?160906110537](http://www.kas.de/wf/doc/kas_46279-544-1-30.pdf?160906110537).
- Moraff, C. (2016): Beware of “Big Data Hubris” When It Comes to Police Reform. In: Next City vom 7.3.2016, online unter <https://nextcity.org/daily/entry/big-data-police-reform-police-transparency-sites>.
- Pasquale, F. (2018): Odd Numbers. In: Real Life vom 20.8.2018, online unter <http://reallifemag.com/odd-numbers/>.
- Perry, W.L.; McInnis, B.; Price, C.C.; Smith, S.C.; Hollywood, J.S. (2013): Predictive Policing. The Role of Crime Forecasting in Law Enforcement Operations. In: RAND Safety and Justice Program, online unter [https://www.rand.org/content/dam/rand/pubs/research\\_reports/RR200/RR233/RAND\\_RR233.pdf](https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf).
- Poser, H. (2012): Wissenschaftstheorie. Eine philosophische Einführung; überarbeitete und erweiterte Auflage. Philipp Reclam.
- PredPol (2018): How Predictive Policing Works. Firmenwebsite, online unter <http://www.predpol.com/how-predictive-policing-works/>.
- Rahwan, I. (2017): Society-in-the-Loop: Programming the Algorithmic Social Contract. MIT Media Lab, Research Paper vom 20.7.2017, online unter <https://arxiv.org/pdf/1707.07232.pdf>.
- Robitzski, D. (2018): To Build Trust In Artificial Intelligence, IBM Wants Developers To Prove Their Algorithms Are Fair. In: Futurism vom 22.8.2018, online unter <https://futurism.com/trust-artificial-intelligence-ibm/>.
- Rohde, N. (2018): Gütekriterien für Algorithmen – Lehren aus bestehenden Forderungskatalogen. Impuls Algorithmenethik Nr. 8. Gütersloh: Bertelsmann Stiftung. Online unter <https://www.bertelsmann-stiftung.de/fileadmin/files/BSI/Publikationen/GrauePublikationen/GuetekriterienalgorithmischeProzesse29062018.pdf>.
- Rowe, M. (2018): AI profiling: the social and moral hazards of ‘predictive’ policing. In: The Conversation, online unter <http://theconversation.com/ai-profiling-the-social-and-moral-hazards-of-predictive-policing-92960>.
- Shapiro, A. (2017): Reform predictive policing. In: Nature 541, S. 458-460. Online unter <https://www.nature.com/news/reform-predictive-policing-1.21338>.
- Schweer, T. (2015): „Vor dem Täter am Tatort“ – Musterbasierte Tatortvorhersagen am Beispiel des Wohnungseinbruchs. In: Die Kriminalpolizei 1, S. 13-16.
- Stadler, K. (2018): Polizei-Software verdächtigt zwei von drei Personen falsch. In: SRF, online unter <https://www.srf.ch/news/schweiz/predictive-policing-polizei-software-verdaechtigt-zwei-von-drei-personen-falsch>.
- Uchida, C.D. (2014): „Predictive Policing“. In: Encyclopedia of Criminology and Criminal Justice, hgg. v. Bruinsma, G. & Weisburd, D., New York: Springer, S. 3871-3880.
- Ulrich, A. (2018): Schön billig. In: Der Spiegel 15/2018 vom 7.4.2018, S.41.
- Wakefield, J. (2018): Are you scared yet? Meet Norman, the psychopathic AI. In: BBC News Technology vom 2.6.2018, online unter <https://www.bbc.com/news/technology-44040008>.
- Winston, A. (2018a): Palantir has secretly been using New Orleans to test its predictive policing technology. In: The Verge vom 27.2.2018, online unter: <https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd>
- Winston, A. (2018b): New Orleans ends its Palantir predictive policing program. In: The Verge vom 15.3.2018, online unter <https://www.theverge.com/2018/3/15/17126174/new-orleans-palantir-predictive-policing-program-end>
- Zweig, K.A.; Fischer, S.; Lischka, K. (2017): Wo Maschinen irren können. Verantwortlichkeiten und Fehlerquellen in Prozessen algorithmischer Entscheidungsfindung. Impuls Algorithmenethik Nr.

## **STIFTUNG NEUE VERANTWORTUNG**

Neue Technologien verändern Gesellschaft. Dafür brauchen wir rechtzeitig politische Antworten. Die Stiftung Neue Verantwortung ist eine unabhängige Denkfabrik, in der konkrete Ideen für die aktuellen Herausforderungen des technologischen Wandels entstehen. Um Politik mit Vorschlägen zu unterstützen, führen unsere Expert:innen Wissen aus Wirtschaft, Wissenschaft, Verwaltung und Zivilgesellschaft zusammen und prüfen Ideen radikal.

## **BERTELSMANN STIFTUNG**

Die Bertelsmann Stiftung setzt sich für eine gerechte Teilhabe aller am gesellschaftlichen Leben ein. Sie engagiert sich in den Bereichen Bildung, Demokratie, Gesellschaft, Gesundheit, Kultur und Wirtschaft. Durch ihr Engagement will sie alle Bürgerinnen und Bürger ermutigen, sich für das Gemeinwohl einzusetzen. Die 1977 von Reinhard Mohn gegründete, gemeinnützige Einrichtung hält die Mehrheit der Kapitalanteile der Bertelsmann SE & Co. KGaA. Die Bertelsmann Stiftung arbeitet operativ und ist unabhängig vom Unternehmen sowie parteipolitisch neutral.

## **ÜBER DAS PROJEKT**

Das Kooperationsprojekt „Algorithmen fürs Gemeinwohl“ von Stiftung Neue Verantwortung und Bertelsmann Stiftung verfolgt eine zu den verbreiteten kritischen Ansätzen komplementäre Strategie, indem es das gesellschaftliche Potential von algorithmischen Systemen in einem Bottom-up-Ansatz ins Zentrum der Projektarbeit stellt. Anhand einer Analyse von Anwendungsbereichen, in denen das Potential algorithmischer Entscheidungsfindung besonders groß ist, soll geklärt werden, wie algorithmische Systeme per Design gemeinwohlorientiert gestaltet werden können. Ziel ist es, Prinzipien zu formulieren, über die bereits in der konzeptionellen und technischen Entwicklungsphase sichergestellt wird, dass Algorithmen ihr gesellschaftliches Verbesserungspotenzial voll entfalten können, ohne negative Effekte, wie die systematische Benachteiligung einzelner Personen oder Gruppen, über Skaleneffekte auf weite gesellschaftliche Teile auszudehnen.

So erreichen Sie den Autor:

Dr. Tobias Knobloch

[tknobloch@stiftung-nv.de](mailto:tknobloch@stiftung-nv.de)

T +49 (0)30 81 45 03 78 93

# IMPRESSUM

Stiftung Neue Verantwortung e. V.  
Beisheim Center  
Berliner Freiheit 2  
10785 Berlin  
[www.stiftung-nv.de](http://www.stiftung-nv.de)

Bertelsmann Stiftung  
Carl-Bertelsmann-Straße 256  
33311 Gütersloh  
[www.bertelsmann-stiftung.de](http://www.bertelsmann-stiftung.de)

Design: Jan Klöthe  
[www.jankloethe.de](http://www.jankloethe.de)

Kostenloser Download: [www.stiftung-nv.de](http://www.stiftung-nv.de)



Dieser Beitrag unterliegt einer CreativeCommons-Lizenz (CC BY-SA). Die Vervielfältigung, Verbreitung und Veröffentlichung, Veränderung oder Übersetzung von Inhalten der Stiftung Neue Verantwortung, die mit der Lizenz „CC BY-SA“ gekennzeichnet sind, sowie die Erstellung daraus abgeleiteter Produkte sind unter den Bedingungen „Namensnennung“ und „Weiterverwendung unter gleicher Lizenz“ gestattet. Ausführliche Informationen zu den Lizenzbedingungen finden Sie hier: <http://creativecommons.org/licenses/by-sa/4.0/>

