

# NORMATIVE ORDERS

Cluster of Excellence at Goethe University Frankfurt/Main

*Normative Orders Working Paper*

*02/2019*

## **Künstliche Intelligenz als Ende des Strafrechts? Zur algorithmischen Transformation der Gesellschaft**

*Von Christoph Burchard*

Cluster of Excellence  
The Formation of Normative Orders  
[www.normativeorders.net](http://www.normativeorders.net)

Goethe-University Frankfurt/Main  
Max-Horkheimer-Str. 2, 60323 Frankfurt/Main

[burchard@jur.uni-frankfurt.de](mailto:burchard@jur.uni-frankfurt.de)



This work is licensed under the Creative Commons Attribution-Non-Commercial-No Derivative Works 3.0 Germany License. To view a copy of this license, visit [http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed.en\\_GB](http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed.en_GB).

## Table of Contents

|   |    |
|---|----|
| EINFÜHRUNG .....  | 3  |
| I. KI ALS NON-ESSENTIALISTISCHES KONSTRUKT.....                           | 6  |
| II. DIE VERHEIßUNGEN VON KI FÜR DIE STRAFRECHTSPFLEGE .....               | 10 |
| III. DER STRAFRECHTSTHEORETISCHE UMGANG MIT DEN VERHEIßUNGEN VON KI ..... | 17 |
| IV. REFLEXION FUNDAMENTALER NORMATIVER ORDNUNGSPOSTULATE .....            | 24 |
| AUSBLICK.....   | 29 |

# Künstliche Intelligenz als Ende des Strafrechts? Zur algorithmischen Transformation der Gesellschaft

Von Christoph Burchard (Goethe-Universität Frankfurt am Main)

*Frage: „Können Sie sich vorstellen, dass smarte und durch KI angetriebene Maschinen in Zukunft zur Unterstützung der justiziellen Entscheidungsfindung eingesetzt werden?“ – Antwort: „Das ist bereits Realität. Und es belastet die Justiz und wie sie verfährt ganz erheblich.“ – Frage an und Antwort durch US Supreme Court Chief Justice John Glover Roberts Jr. im April 2017<sup>1</sup>*

Der finale Beitrag wird in der Gedächtnisschrift für Joachim Hruschka im Jahrbuch für Recht und Ethik (JRE) 2019 erscheinen. Den Herausgebern, Jan C. Joerden und Jan Schuhr, sowie dem Verlag Duncker & Humblodt gilt mein herzlicher Dank, dass der Beitrag in einem ersten Aufschlag auch in dieser Working Paper Reihe vorab erscheinen darf. Der Beitrag wurde im Rahmen des Clusterprojekts „Die normative Ordnung Künstlicher Intelligenz“ verfasst.

## EINFÜHRUNG

Kurz bevor ich die ehrenvolle Einladung erhielt, an diesem wichtigen Gedächtnisband<sup>2</sup> mitzuwirken, hatte ich just den Aufsatz von *Byrd/Hruschka* zu Kants Straftheorie in einem Lektüreseminar „Strafrecht zwischen Moral und Politik“ behandelt.<sup>3</sup> Ein wahrhaft bahnbrechender Aufsatz, der Kant aus der (für Viele: Schmutz-) Ecke der absoluten Straftheorien herausholte. Damit schloss sich für mich ein Kreis. Denn dem wegweisenden Œuvre *Joachim Hruschkas* war ich erstmals im Rahmen meiner Doktorarbeit begegnet. Ich kann mich noch gut erinnern, wie sehr ich – weiland<sup>4</sup> und seitdem – von seinem stets methodisch und rechtsphilosophisch reflektierten, eben logisch-analytisch gehaltenen und im besten Sinne aufgeklärten Zugang zur Dogmatik beeindruckt war. Daher ist es mir eine wirkliche Ehre, in Gedenken an *Joachim Hruschka* heute einige erste Überlegungen zu hochaktuellen Entwicklungen anstellen zu dürfen, eben dazu, wie sog. Künstliche Intelligenz unsere Gesellschaft und unser Strafrecht zu verändern „droht“. Diese Themensetzung mag zunächst verwundern, da sich scheinbar nur wenige unmittelbare Bezugspunkte zum Werk *Joachim Hruschkas* ergeben. Wie ich jedoch zeigen möchte, stellt der Einsatz

---

<sup>1</sup> Berichtet von *Liptak*, „Sent to Prison by a Software Program’s secret Algorithms“, New York Times v. 1.5.2017, S. A22 (eigene Übersetzung).

<sup>2</sup> Der Beitrag wird veröffentlicht in: *GS Hruschka* JRE 27 (2019).

<sup>3</sup> *Byrd/Hruschka*, „Kant zu Strafrecht und Strafe im Rechtsstaat“, JZ 2007, S. 957.

<sup>4</sup> So hatte ich u.a. folgende Titel in meiner Doktorarbeit verarbeitet: *Hruschka*, „Die Herbeiführung eines Erfolges durch einen von zwei Akten bei eindeutigen und bei mehrdeutigen Tatsachenfeststellungen“, JuS 1982, S. 317; ders., Strafrecht nach logisch-analytischer Methode, Berlin/New York: De Gruyter, 1988; ders., „Der Standard-Fall der aberratio ictus und verwandte Fallkonstellationen“, JZ 1991, S. 488.

Künstlicher Intelligenz unsere Gesellschafts- und Strafrechtsordnung vor fundamentale Herausforderungen. Notwendig wird nicht weniger als eine Besinnung auf die normativen Grundlagen unseres Strafrechts, das ein Strafrecht einer Gesellschaft ist, die es neu zu erklären, verstehen und bewerten gilt. Und insofern ist das virtuelle Gespräch mit *Joachim Hruschka* unbedingt fortzusetzen, um in Auseinandersetzung mit seinen philosophischen, also nicht schlicht behauptenden,<sup>5</sup> sondern stets begründenden Einsichten die heutigen Entwicklungen kritisch befragen zu können. Weniger mutig als *Joachim Hruschka*, der einst nicht weniger als ein Neudurchdenken des Strafrechts gefordert hat,<sup>6</sup> rege ich hier an, sich Rechenschaft über die Gesellschaft abzulegen, die unser Strafrecht hervorbringt bzw. hervorbringen soll. Dazu muss sich – wie dies *Joachim Hruschka* schon vor langer Zeit erkannt hat<sup>7</sup> – die Strafrechts- wieder zur Sozial- und politischen Theorie öffnen.

\*\*\*

„Künstliche Intelligenz (KI) und Strafrecht“ ist *keine* Science Fiction. Angesichts der alle Lebensbereiche durchwirkenden „digitalen Revolution“ kann sich weder das Recht im Allgemeinen noch das Strafrecht im Besonderen vor den Einflüssen von KI feien.<sup>8</sup> Soweit KI strafrechtswissenschaftlich bereits in den Blick genommen wird, wird sie (gerade in Deutschland) herkömmlich als potentiell regulierungsbedürftiger<sup>9</sup> und auch regulierungsfähiger Objektbereich geführt.<sup>10</sup>

Dieser Beitrag wirbt für eine *Perspektivenergänzung*.<sup>11</sup> Also dafür, KI (bzw. genauer ihre soziale Praxis) als Medium heutiger gesellschaftlicher und (straf-)rechtlicher Transformationen in den Blick zu nehmen.<sup>12</sup> Es soll also gefragt werden: Wie verändert KI bereits heu-

---

<sup>5</sup> Dagegen dezidiert *Hruschka*, „Kann und sollte die Strafrechtswissenschaft systematisch sein?“, JZ 1985, S. 1 (S. 10).

<sup>6</sup> So *Hruschka*, „Das Strafrecht neu durchdenken!“, GA 1981, S. 237.

<sup>7</sup> Ebd., S. 249. Gesellschaftsstraftheoretisch wichtig ebenfalls *Hruschka*, „Utilitarismus in der Variante von Peter Singer“, JZ 2001, S. 261.

<sup>8</sup> Provokant die Fragestellung von *Schwintowski*, „Wird Recht durch Robotik und künstliche Intelligenz überflüssig?“, NJOZ 2018, S. 1601.

<sup>9</sup> Zu den dabei notwendigen Abwägungen, nicht zuletzt mit dem notwendigen Erhalt technologischer Innovationskraft im internationalen (Wirtschafts- und Wissenschafts-)Wettbewerb, vgl. etwa allg. *Meyer*, „Künstliche Intelligenz und die Rolle des Rechts für Innovation“, ZRP 2018, S. 233.

<sup>10</sup> Dabei steht häufig die *agency*-Frage im Vordergrund, wer also verantwortlich ist, wenn durch den Einsatz eines KI-Systems, insbesondere beim Einsatz selbstlernender KI, Schäden verursacht werden. Hierzu eindrücklich *Hilgendorf*, „Autonome Systeme, künstliche Intelligenz und Roboter“, in: Barton u.a. (Hrsg.), Festschrift für Thomas Fischer, München: C.H. Beck, 2018, S. 99 (111 ff.).

<sup>11</sup> Allg. hierzu auch *Balkin*, „The Path of Robotics Law“, California Law Review Circuit 2015, S. 45.

<sup>12</sup> KI kann daher als potentiell „transformative Technologie“ ausgeflaggt werden, um einen dialektischen Zusammenhang zwischen technologischem, gesellschaftlichem und rechtlichem Wandel herzustellen. Hierzu allg. *Fateh-Moghadam*, „Selbstbestimmung im biotechnischen Zeitalter“, Basler Juristische Mitteilungen 2018, S. 205 (insbes. S. 209 ff.). – Indem hier auf die soziale Praxis von KI als Medium gesellschaftlicher Transformationen abgestellt wird, wird zum einen gegen eine Essentialisierung von KI angetreten (hierzu

te unsere Gesellschafts- und Strafrechtsordnung?<sup>13</sup> Bedeutet KI zugespitzt gar das Ende des Strafrechts?<sup>14</sup> Nämlich entweder im Sinne eines siechenden Hinscheidens seiner fundamentalen Prinzipien (wie der Unschuldsvermutung) oder eines krönenden Abschlusses seiner grundlegenden Ziele (wie des Rechtsgüterschutzes)?<sup>15</sup>

Um diesen Fragen nachzugehen, gilt es die Rechtfertigungsnarrative, die bereits heute für die Einführung von KI ins Strafrecht ins Felde geführt werden, näher zu untersuchen; die hinter diesen Narrativen liegenden, durch sie auch verborgenen normativen Ordnungsvorstellungen genauer zu analysieren;<sup>16</sup> und dabei auch die machtpolitische und ideologische Offenheit dieser Ordnungsvorstellungen herauszustreichen. Im Sinne dieses Programms werden zunächst mit Blick auf aktuelle Entwicklungen die Verheißungen von KI beleuchtet (hierzu unten I. und II.), um sodann zu reflektieren, wie eine kritische Perspektiven eröffnende<sup>17</sup> Strafrechtstheorie damit umgehen sollte (hierzu unten III. und IV.).

Besonderes Augenmerk wird dabei den Verheißungen geschenkt, dass vermeintlich smarte oder intelligente, regelmäßig Big Data auswertende Algorithmen einen effektiveren und effizienteren Rechtsgüterschutz ermöglichen und eine neutralere, objektivere und kohärentere Rechtsdurchsetzung garantieren als menschliche Entscheidungsträger. Diese Versprechungen entsprechen *prima facie* jenen des Strafrechts. Auch dieses verspricht, als vermeintlich schärfstes Schwert des Gesetzgebers, einen besonders durchgreifenden Rechtsgüterschutz. Zudem verlangt das Strafrecht nach seiner unbedingt unparteiischen und unvoreingenommenen sowie konsistenten Anwendung. Der Unterschied liegt darin, dass KI technologische Faktizität verheißt, während das Strafrecht – wie das Recht allge-

---

auch unten I.) und zum anderen klargestellt, dass KI zwar nicht selbst normativ wirkt, gleichwohl aber einprogrammierte Ordnungsvorstellungen vermittelt.

<sup>13</sup> Die Frage insinuiert eine Kausalität, die bei näherem Hinsehen als dialektischer Prozess aufzulösen ist, in dem gesellschaftliche etc. Entwicklungen die Entwicklung und den Einsatz bestimmter KI-Systeme fördern, die ihrerseits die erstgenannten Entwicklungen verstärken etc. KI baut dabei „natürlich“ auf der allgemeinen Technisierung der Lebenswirklichkeit auf. So könnte etwa auch im Einsatz elektronischer Fußfesseln, gerade gegen sog. Gefährder, ein Vorschein einer partiellen Überwachungsgesellschaft erkannt werden, die allemal bestimmten Personengruppen (eben den sog. Gefährdern) das kontrafaktische Vertrauen in ihre Rechtstreue entzieht. Dies kann dann KI mit einem allg. *zero trust*-Paradigma verstärken. Hierzu unten II.

<sup>14</sup> Diese Frage soll aufrütteln und kein *fin de siècle* zum Ausdruck bringen. In der Sache geht es darum, wie mit den Polyvalenzen der heutigen Entwicklungen umzugehen ist. Ein ähnliches Wortspiel findet sich bei *Hildebrandt*, *Smart Technologies and the End(s) of Law*, Cheltenham: Edward Elgar Publishing, 2015.

<sup>15</sup> Meine Fragestellungen verstehen sich *nicht* als spekulative Science Fiction, so dass allfällige Dystopien hier außer Betracht bleiben.

<sup>16</sup> Zu diesem am Exzellenzcluster „Die Herausbildung normativer Ordnungen“ geprägten Begriffsapparat einfühend *Forst/Günther*, „Die Herausbildung normativer Ordnungen: zur Idee eines interdisziplinären Forschungsprogramms“, in: *Forst/Günther* (Hrsg.), *Die Herausbildung normativer Ordnungen*, Frankfurt a.M.: Campus, 2011, S. 11 (insbes. S. 15 f.).

<sup>17</sup> Kritik wird hier axiomatisch verstanden als eine Praxis des begründenden Zweifelns. Das normative Programm liegt in der Eröffnung von Möglichkeiten zu kritischen, zweifelnden Anfragen, nicht in der Entwicklung eines normativen Programms, aus dem heraus Kritik geübt werden kann. Letzteres wäre einer Kritischen Theorie des Strafrechts vorbehalten, die hier nicht entwickelt werden kann. Eine Rückkehr zu einem rationalen Kritikbegriff fordert auch *Hruschka*, *Strafrecht nach logisch-analytischer Methode* (Fn. 4), S. XI.

mein – nur kontrafaktische Garantien aussprechen kann. Ob KI das Ende des Strafrechts bedeutet, ist in der Folge gleichbedeutend mit der Frage, welches Strafrecht damit gemeint ist. Ein liberales Freiheitsstrafrecht, das auf interpersonalem Vertrauen gründet und die Menschen nicht allein als steuer- und konstant bewertbare potentielle Risiken (neudeutsch also als Gefährder) führt, wird durch KI fundamental in Zweifel gezogen; eine liberale (Straf )Rechtstheorie müsste daher die Vorwärtsverteidigung des Kontrafaktischen des (Straf )Rechts gegen das Faktische der KI suchen. Für ein wohlfahrtstaatliches Sicherheitsstrafrecht, das sich als Instrument der Sozialkontrolle bzw. der Governance sozialer Interaktionen versteht, ermöglicht KI hingegen einen krönenden Abschluss seiner Rationalität.

## I. KI ALS NON-ESSENTIALISTISCHES KONSTRUKT

KI ist kein eindeutiger Begriff. Um seine genaue Begriffsdefinition wird allseits gerungen.<sup>18</sup> Häufig trifft man auf einen im Grunde genommen strukturell angelegten Essentialismus, der die proprietären Eigenschaften von Intelligenz im Allgemeinen und Künstlicher Intelligenz im Besonderen zu ergründen sucht.<sup>19</sup> Das mag auch erklären, warum bereits heute tiefschürfend über die strafrechtliche Verantwortung dereinst intelligenter und selbstbewusster Maschinen<sup>20</sup> räsoniert wird.<sup>21</sup>

---

<sup>18</sup> Prägend sind Debatten darüber, ob der Name KI den aktuellen Stand von Forschung und Entwicklung überhaupt akkurat abbildet (z.B. wenn und weil heutige „KI“-Systeme nicht über das klassische Maschinenlernen und die althergebrachte Mustererkennung hinausgehen); ob KI wirklich „intelligent“ ist (z.B. wenn und weil heutige „KI“-Systeme keine Transferleistungen erbringen können); und ob die deutsche Qualifikation „künstlich“ richtig ist (z.B. wenn und weil die Qualifikation „maschinelle“ Intelligenz die Eigenschaften von Algorithmen korrekter abbilden soll oder das „Künstliche“ im Sinne der Romantik pejorativ besetzt ist). Hierzu etwa *Herberger*, „Künstliche Intelligenz‘ und Recht“, NJW 2018, S. 2825. – Vgl. weiterführend (das nur auf den ersten Blick veraltete Werk von) *Weizenbaum*, Die Macht der Computer und die Ohnmacht der Vernunft, Frankfurt a.M.: Suhrkamp, 1977, S. 268 ff.

<sup>19</sup> So erachtet etwa *Ertel*, Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, Wiesbaden: Springer VS, 2016, S. 1 ff. die Fragen „Was ist Intelligenz?“, „Wie kann man Intelligenz messen?“ oder „Wie funktioniert unser Gehirn?“ für bedeutsam für das Verständnis von KI. Ferner wird ausgeführt, dass für den Informatiker und insbesondere den Ingenieur die Frage „nach der intelligenten Maschine, die sich verhält wie ein Mensch, die intelligentes Verhalten zeigt,“ entscheidend sei.

<sup>20</sup> Daraus spricht ein essentialistisch und anthropozentrisch strukturiertes Strafrechtsverständnis, das es im Zuge des weltweiten Siegeszugs des Verbands- und Unternehmensstrafrechts zu hinterfragen gilt. Es erscheint nicht fernliegend (und mehr sei hier nicht zur Diskussion gestellt), die Strafbarkeit von KI-Systemen strikt funktional begründen zu wollen, also unabhängig davon, wie intelligent ein KI-System nun ist und ob es sich seiner selbst bewusst ist und ob ihm damit originär menschliche Eigenschaft zugesprochen werden kann. Für eine KI-Strafbarkeit könnte vielmehr, analog zur Unternehmensstrafbarkeit, etwa funktional ins Felde geführt werden: (1) Indirekt sollen die Eigner eines defizitären KI-Systems getroffen werden (z.B. wenn dessen Abschaltung angeordnet wird). (2) Der Regress auf Verantwortliche, die ein defizitäres KI-System entwickelt oder auf den Markt gebracht haben, soll abgeschnitten werden (z.B. weil und wenn „hinter“ einem KI-System unzählige Personen stehen, von Programmierern bis Unternehmensverantwortlichen jedweder Hierarchieebene, so dass eine Verantwortungsindividualisierung nicht tunlich wäre). (3) Oder

Hier wird gleichwohl für ein *non-essentialistisches Verständnis von KI* eingetreten, um Erkenntnisse aus der Beobachterperspektive für eine in den Kategorien der Teilnehmerperspektive denkende, aktuelle und kritische Strafrechtstheorie fruchtbar zu machen. Es geht also nicht darum, was die Essenz von (Künstlicher) Intelligenz ausmacht bzw. ausmachen sollte. Vielmehr wird im Ausgangspunkt gefragt, wie KI (insbesondere durch entsprechende Rechtfertigungsnarrative) sozial konstruiert, repräsentiert und rezipiert wird; wie KI in bestimmte soziale Verhältnisse eingebettet ist und diese verändert; und welche Herrschafts- und Machtverhältnisse in KI zum Ausdruck kommen und durch sie stabilisiert, mystifiziert, transformiert oder produziert werden.<sup>22</sup> Es geht mit anderen Worten um Bedeutungszuweisungen, die der sozialen Lebenswirklichkeit entstammen und auf sie zurückwirken. Dabei werden je zu bestimmende normative Ordnungsvorstellungen (einschließlich Ideologien) befördert und spezifische Problemlagen in den Vorder- oder Hintergrund gespielt. So gesehen ist KI ein normativ offenes und formbares, gleichsam ein politisches Konstrukt. Gerade die Uneindeutigkeit und machtpolitische Offenheit des Terminus KI erlaubt es interessierten Akteuren, ihn für eigene – politische, ökonomische etc. – Zwecke zu instrumentalisieren. KI ist dabei eine besonders wirkmächtige Namensgebung, weil sie einem IT-System rein sprachlogisch – also unabhängig von seiner „wirklichen“ Intelligenz – die allgemeine kognitive Leistungsfähigkeit attestiert, „abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten.“<sup>23</sup>

In diesem Zugriff ist KI *nicht* als neutrale Technologie oder schlicht als informationstechnologische Innovation aufzufassen. Vielmehr ist KI unmittelbar mit den Grundprinzipien menschlicher Sozialität (Freiheit, Toleranz, Recht etc.) verknüpft, geht auf sie zurück und transformiert sie. Daher gilt es die mit dem Terminus KI zu einem je bestimmten Zeitpunkt verbundenen Versprechungen, Hoffnungen und Ängste, wie sie durch die verschiedensten Akteure in Wirtschaft, Politik, Wissenschaft etc. mit unterschiedlicher Dringlichkeit befeuert und geschürt werden, ernst zu nehmen, um sie so einer kritischen Reflexion unterziehen zu können.

Als Beispiel sei auf das herkömmliche Bild von KI verwiesen, das nicht zuletzt von *Hollywood* in unseren Köpfen verankert ist.<sup>24</sup> Hier steht KI für autonome Roboter (man

---

durch die Zurechnung strafrechtlicher Verantwortlichkeit sollen in institutionalisierter Form Gefühle des Übelnehmens und der Empörung zum Ausdruck gebracht werden (ich danke *Boris Burghardt* für den Hinweis auf diesen Aspekt).

<sup>21</sup> Hierzu ausf. *Gaede*, *Recht und Strafen für Roboter?*, Baden-Baden: Nomos, 2019 m.w.N.

<sup>22</sup> Ähnlich *Balkin* (Fn. 11), S. 59; *Weizenbaum* (Fn. 18), S. 268 ff. Weiterführend auch *Mau*, *Das metrische Wir: Über die Quantifizierung des Sozialen*, Berlin: Suhrkamp, 2017.

<sup>23</sup> Duden, Stichwort: Intelligenz, erste Wortbedeutung.

<sup>24</sup> Hierzu im Überblick bereits *Xanke/Bärenz*, „Künstliche Intelligenz in Literatur und Film – Fiktion oder Reali-

denke dystopisch an „Terminator“ oder offener an „I, Robot“), intelligente Androiden (man denke an Lieutenant Commander Data in „Star Trek: The Next Generation“) und selbstbewusste Supercomputer (man denke an HAL 9000 in „2001: Odyssee im Weltraum“ oder an Central in „Star Trek: Discovery“). KI tritt insofern (aus dramaturgischen Gründen verständlicherweise) fast ausschließlich als sog. *starke KI* auf, welche die gleiche allgemeine Intelligenz wie Menschen anstrebt oder bereits erlangt, wenn nicht gar übertroffen hat.<sup>25</sup> Verhandelt werden sodann im Schwerpunkt sowohl die *conditio humana* (ob und wie sich Menschen in eine Maschinenwelt einfinden können)<sup>26</sup> wie auch die *conditio automata* (ob und wie sich intelligente Maschinen in eine menschliche Gesellschaft einfinden können)<sup>27</sup> Da keines der heute existierenden Systeme unter die Kategorie der starken KI fällt,<sup>28</sup> haben diese Verhandlungen freilich noch keine unmittelbare rechtspraktische Bedeutung. Das mag eine Ursache dafür sein, dass „KI und Strafrecht“ bis vor Kurzem bestenfalls ein Orchideenthema war und dass weiterhin in essentialistischer Manier um die inhärenten Charakteristika von starker Künstlicher Intelligenz gerungen wird.

Darüber darf aber nicht vergessen werden, dass wir momentan Zeuge einer anderen KI-Revolution werden.<sup>29</sup> Diese basiert der Sache nach auf sog. *schwacher KI*, die für die Lösung konkreter Anwendungsprobleme optimiert ist, auf bekannten Methoden der Mathematik und Informatik basiert und kein tieferes – oder eigentliches – Verständnis für die Problemlösung erlangt.<sup>30</sup> Und bei der auch – Hollywood zuwider – keine Roboter oder Androiden zum Einsatz kommen.

Dass auch diese Systeme hier als KI geführt werden, folgt aus dem soeben anheimgestellten methodischen Ansatz. Unter anderem bedingt durch entsprechende Anstrengungen der Wirtschaft,<sup>31</sup> lösen wir uns in Politik und Gesellschaft zusehends von KI à la Hollywood und verstehen unter KI informationstechnische Systeme, die uns als Lösung für lebenswirkliche Probleme „verkauft“ werden – und die wir dafür auch

---

tät?“, *Journal of New Frontiers in Spatial Concepts* 2012, S. 36; *Irsigler/Orth*, „Zwischen Menschwerdung und Weltherrschaft: Künstliche Intelligenz im Film“, *APuZ* 2018, S. 39.

<sup>25</sup> Zur Unterscheidung zwischen starker und schwacher KI „klassisch“ *Searle*, „Minds, Brains, and Programs“, *Behavioral and Brain Sciences* 1980, S. 417.

<sup>26</sup> Als literarisches Beispiel: *McEwan*, *Machines like Me*, London: Penguin, 2019.

<sup>27</sup> Als filmisches Beispiel: *Ex Machina*, 2015.

<sup>28</sup> Folgt man der bei *Searle* (Fn. 25) entwickelten begrifflichen Trennung zwischen schwacher und starker KI, so ist diese bis in die Gegenwart in ihrer Verwendung durchaus inkonsistent. Als zentraler Fluchtpunkt einer starken KI gilt das Äquivalent zu menschlichen Fähigkeiten. Siehe hierzu nur *Ramge*, *Mensch und Maschine: Wie künstliche Intelligenz und Roboter unser Leben verändern*, Stuttgart: Reclam, 2018, S. 19.

<sup>29</sup> So in der Sache auch das populärwissenschaftliche Werk von *Fry*, *Hello World: Was Algorithmen können und wie sie unser Leben verändern*, München: C.H. Beck, 2018.

<sup>30</sup> Hierzu, statt aller, *Ramge* (Fn. 29), S. 19.

<sup>31</sup> Dies herausarbeitend und kritisierend *Zuboff*, *Das Zeitalter des Überwachungskapitalismus*, Frankfurt a.M./New York: Campus, 2018.

„einkaufen“. Bezeichnend hierfür ist das jüngste Konzeptpapier des „European Committee on Crime Problems (CDPC)“ des Europarats zum Thema „KI und strafrechtliche Verantwortlichkeit“. Dort werden selbstfahrende Automobile nebst ihren selbst-lernenden Algorithmen wie selbstverständlich als Paradebeispiele für KI ausgeflaggt.<sup>32</sup> Ins gleiche Horn bläst die „European Commission for the Efficiency of Justice (CEPEJ)“ des Europarats, die sich in bestimmten Bereichen und unter bestimmten Bedingungen für die Anwendung von KI in der Rechtspflege ausspricht.<sup>33</sup> All das ist Ausdruck des heute zu verzeichnenden KI-Hypes. Dieser wird von dem allorts reproduzierten Narrativ getragen, dass smarte oder intelligente Algorithmen lebenswirkliche, aber menschliche Fähigkeiten überfordernde Probleme informationstechnisch zum Wohle aller zu bewältigen im Stande sind; und zwar besser, schnell und kostengünstiger als menschliche Entscheidungsträger (beginnend bei der sicheren Steuerung von PKWs über die Auswertung sämtlicher medizinischer Publikationen bei der Unterstützung von Krankheitsdiagnosen und Therapiekonzepten bis hin zur Automatisierung von rechtlichen Dienstleistungen, sog. *legal tech*<sup>34</sup>).

Diese „neue“ Bedeutung von KI ist nicht neutral, sondern normativ und auch ideologisch aufgeladen. Man sollte sich zwar davor hüten, trotz des andauernden Ringens um die Deutungsherrschaft von „dem“ – allemal dominierenden – Bedeutungsinhalt von KI zu sprechen. Gleichwohl überzeugt der Befund von Katz, dass „the ‚AI‘ label has been rebranded to promote a vision of world governance through big data.“<sup>35</sup> Diese Vision verbindet sich nicht selten mit einer metaphysisch und theologisch aufgeladenen Erlösungshoffnung. KI verheißt eine selbst lernende und sich selbst verbessernde Entität, die die Transparenz und Berechenbarkeit des Anderen wie auch des Selbst, und damit die hyperrationale Steuerbarkeit sozialer Interaktionen, verspricht.<sup>36</sup> Dieses Narrativ hat besondere soziale und politische Zug- und Sprengkraft, da es allgemein angelegt ist und alle Lebensbereiche durchwirkt (was KI von früheren partikularen „kriminaltechnologischen“ Strömungen à la Lombroso schon quantitativ deutlich abhebt). Das Narrativ ist

---

<sup>32</sup> European Committee on Crime Problems (CDPC), „Artificial Intelligence and its Impact on CDPC Work: The case of automated driving“, CDPC (2018)14 v. 14.09.2018, S. 6 f. Abrufbar unter: <http://rm.coe.int/cdpc-2018-14-artificial-intelligence-and-criminal-law-project-2018-202/16808d6d09> (zuletzt abgerufen am 20.06.2019).

<sup>33</sup> CEPEJ, „European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment“ v. 3-4.12.2018, insbes. S. 64 ff. Abrufbar unter: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c> (zuletzt abgerufen am 20.06.2019).

<sup>34</sup> Überblicke bei Fries, „Automatische Rechtspflege“, RW 2018, S. 414.

<sup>35</sup> Katz, „Manufacturing an Artificial Intelligence Revolution“, 2017, S. 1. Abrufbar unter SSRN: <https://ssrn.com/abstract=3078224> (zuletzt abgerufen am 20.06.2019).

<sup>36</sup> Insofern treffend Nida-Rümelin/Weidenfeld, Digitaler Humanismus, München: Piper, 2018, S. 44 ff.

überdies so generisch ausgelegt ist, dass es neoliberal,<sup>37</sup> rational und wissenschaftlich,<sup>38</sup> liberal<sup>39</sup> wie auch autoritär<sup>40</sup> vereinnehmbar wird.<sup>41</sup> KI bzw. ihre grundlegenden normativen Ordnungspostulate und Rechtfertigungsnarrative sind mit anderen Worten latent (macht)politisch offen. Und in ihnen zeigt sich auch eine neue Ausprägung der Dialektik der Aufklärung.

## II. DIE VERHEIßUNGEN VON KI FÜR DIE STRAFRECHTSPFLEGE

Ausgehend von der zuvor herausgestrichenen allgemeinen gilt es nun auf die besonderen Versprechungen einzugehen, die KI für das Strafrecht bzw. die Strafrechtspflege ausspricht. Namentlich die Versprechung, dass Kriminalität durch den Einsatz intelligenter Informationstechnologie tatsächlich verunmöglicht oder allemal drastisch reduziert werden kann (unten 1.); und dass die strafjuristische Entscheidungsfindung von menschlicher Subjektivität und Voreingenommenheit befreibar ist und daher „endlich“ tatsächlich objektiv, neutral und kohärent erfolgen kann (unten 2.).

### 1. Effektivität und Effizienz in der Kriminalitätsinhibition

Zunächst verspricht KI die direkte wie indirekte Verunmöglichtung von Kriminalität. Bzw. genauer: Verfechter werben dafür, dass bestimmte Kriminalitätsformen durch den Einsatz von KI *eo ipso* nicht mehr begehbar bzw. die Begehung bestimmter Kriminalitätsformen durch KI-geförderte (hoheitliche, privatisierte oder internalisierte) Durchsetzungs- und Überwachungsstrukturen *de facto* signifikant reduzierbar ist. Illustrativ hierfür sind sog. Smart Contracts<sup>42</sup> sowie das Predictive<sup>43</sup> bzw. Big Data Policing<sup>44</sup>.

---

<sup>37</sup> Hierzu kritisch *Zuboff* (Fn. 32).

<sup>38</sup> Hierzu kritisch *Nida-Rümelin/Weidenfeld* (Fn. 36), Einführung.

<sup>39</sup> So etwa *Chiao*, „Predicting Proportionality: The Case for Algorithmic Sentencing“, 37 *Criminal Justice Ethics* 2018, S. 238 ff.

<sup>40</sup> Hierzu in der Übersicht – und insbes. mit Blick auf China – *Mau* (Fn. 22), S. 9 ff.

<sup>41</sup> Theoretisch lässt sich also sagen, dass die Rechtfertigungsnarrative von KI (erster analytischer Bezugspunkt) auf dahinterstehende normative Ordnungen (zweiter analytischer Bezugspunkt) verweisen, die ihrerseits machtpolitisch bzw. ideologisch offen sind (dritter analytischer Bezugspunkt).

<sup>42</sup> Smart Contracts werden in der Regel im zivil- und nicht im strafrechtlichen Schrifttum behandelt, so dass hier von weiteren Nachweisen abgesehen wurde. Smart Contracts sind nicht zwingend KI zu subsumieren. Da jedoch entsprechende Entwicklungen offenbar werden, werden sie hier dargestellt.

<sup>43</sup> Die deutsche Übersetzung „vorhersagende Polizeiarbeit“ gibt die Implikationen und Facetten der englischen Originalbegriffe nur unzureichend wieder, so dass hier letztere Verwendung finden. – Zum Predictive Policing im deutschen Schrifttum etwa *Rademacher*, „Predictive Policing im deutschen Polizeirecht“, *AöR* 2017, S. 366; *Singelstein*, „Predictive Policing: Algorithmenbasierte Straftatprognosen zur vorrausschauenden Kriminalintervention“, *NStZ* 2018, S. 1; *Härtel*, „Digitalisierung im Lichte des Verfassungsrechts – Algorithmen, Predictive Policing, autonomes Fahren“, *LKV* 2019, S. 49 (insbes. S. 54).

<sup>44</sup> Big Data Policing umschreibt eine neuere Entwicklung gerade in den USA, die in Deutschland terminologisch noch erschlossen werden muss. Hierzu *Ferguson*, *The Rise of Big Data Policing. Surveillance, Race,*

### a) *Smart Contracts*

Smart Contracts qualifizieren sich aufgrund spezifischer informationstechnischer Ansprüche bezeichnender Weise als „smarte“, d.h. als schlaue, pfiffige und raffinierte Verträge. Sie treten mit dem Versprechen auf, vertragliche Interaktionen effektiver und effizienter abwickeln zu können. Smart Contracts werden durch Computerprogramme umgesetzt, die auf die algorithmische Ermöglichung, Überprüfung und Durchsetzung vertraglicher Rechte und Pflichten zielen, ohne dabei auf dritte Parteien angewiesen zu sein. Die Idee ist die weitestgehende Automatisierung der Vertragsgestaltung und -abwicklung, die idealiter „self-executing“ ist und so die Transaktionskosten des klassischen Vertragsrechts minimiert. So soll das Rechtssystem (einschließlich seiner Vertreter wie Notare oder Richter) aber auch private Dienstleister, die gegen Zahlungs- oder Lieferausfälle versichern, obsolet gemacht werden. Frei nach dem Credo „code is law“ wird das Recht sowie die Rechtsanwendung durch eine entsprechende IT-Infrastruktur ersetzt.

Smart Contracts beruhen auf der „Distributed Ledger Technology“ wie der sog. Blockchain, also einer mit der Lebenswirklichkeit vernetzten Datenbank, die dezentral gespeichert, verifiziert und kontinuierlich aktualisiert wird. Und in der z.B. gespeichert ist, wer über welche Waren oder Finanzmittel verfügt. Je umfassender und akkurater diese Datenbank ist – gleichsam über je mehr Datensätze (vulgo Big Data) sie verfügt –, desto genauer und sicherer kann und wird der Smart Contract abgewickelt werden – so zumindest die Anspruchshaltung.

Getragen wird dies von der Vision einer vertrauensfreien Gesellschaft, in der sich Vertrags- bzw. Interaktionspartner nicht länger vertrauen müssen.<sup>45</sup> Denn ihnen werden – so das Rechtfertigungsnarrativ – bessere, nämlich absolute bzw. unhintergehbare informationstechnische Versicherungen gegeben (z.B. dass der Verkäufer über die angebotene Ware verfügt und diese übergeben und übereignen wird; und dass der Käufer hinreichend liquide ist und die Ware auch tatsächlich bezahlen wird). Damit geht eine Vertrauensverschiebung einher, weg vom zwischenmenschlichen Vertrauen und hin zum Vertrauen in informationstechnische Systeme (die Datenbank und die Programmierung).<sup>46</sup> Die Vertrags- und Interaktionspartner werden dabei dem Grunde nach als Risiko aufgefasst, da – in der Tat – nicht ausgeschlossen werden kann, dass sich diese wort- bzw. vertragsbrü-

---

and the Future of Law Enforcement, New York: New York University Press, 2017. Vgl. auch die im Ohio State Journal Criminal Law 2018, S. 473 ff. abgedruckten Beiträge eines „Round Table on Big Data and Criminal Law“.

<sup>45</sup> Vgl. allg. *Palka/Wittpahl*, Vertrauen und Transparenz – Blockchain-Technologie als digitaler Vertrauenskatalysator, Working Paper of the Institute for Innovation and Technology Nr. 39, 2018.

<sup>46</sup> Weiterführend und allg. *Wagner*, „Vertrauen in Technik“, Zeitschrift für Soziologie 1994, S. 145.

chig verhalten werden. Um dieses Risiko zu bewältigen, wird Hand an das zwischenmenschliche Vertrauen als kontrafaktisch postulierter, da sozial schlicht notwendiger Mechanismus zur Reduktion sozialer Komplexität gelegt.<sup>47</sup> „Smarte“ Datenbanken und Algorithmen sollen fürderhin dafür sorgen, dass sich die sozial Interagierenden auf einander verlassen, und zwar, weil sie übereinander wissen, dass ihnen die Möglichkeit des datenbank- und algorithmenwidrigen Verhaltens durch „intelligente“ Algorithmen und Datenbanken genommen ist. Der Eine begibt sich dabei der Möglichkeit seines datenbank- und algorithmenwidrigen Verhaltens, weil, wenn und damit dem Anderen diese Möglichkeit ebenfalls genommen ist. Zugespitzt gesagt: Das *homo homini lupus est* wird nicht nur psychologisch verdrängt (vertrauensbasierte Interaktion), sondern von Anfang informationstechnisch abgewendet (sog. *zero trust* bzw. *in tech we trust*-Interaktion).

Bezogen auf das Strafrecht soll so, Hoheitsträger außen vorlassend und in dem Sinne die Straftatprävention privatisierend<sup>48</sup>, der Eingehungs- und Erfüllungsbetrug bei Austauschverträgen verunmöglicht werden. Das ist pointiert in der „Gralschrift“ der ursprünglichen Blockchain-Bewegung nachzulesen, mit und in dem die Bitcoin-Idee erfunden und erklärt wurde. Wörtlich heißt es dort:

„Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for nonreversible services. With the possibility of reversal, the need for trust spreads. Merchants must be wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as

---

<sup>47</sup> So die klassische Definition bei *Luhmann*, *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*, München: UVK, 2014, S. 27 ff. (insbes. S. 30).

<sup>48</sup> Weitere Beispiele für eine solche Privatisierung der Straftatprävention durch KI liefern z.B. KI gestützte „Criminal Compliance“-Systeme (also neudeutsch Digital Compliancetools), die etwa eine Vollüberwachung der unternehmensinternen Kommunikation mit dem Ziel versprechen, verdächtige Interaktionen „zu flaggen“ und damit einer weiteren Prüfung zuzuführen. Hierzu in der Übersicht *Schemmel/Dietzen*, „‘Effective Corporate Governance’ by Legal Tech & Digital Compliance“, in: Breidenbach/Glatz (Hrsg.), *Rechtshandbuch Legal Tech*, München: Beck, 2018, S. 109. – In Japan wird ferner, wie entsprechenden Presseberichten zu entnehmen ist, kontrovers diskutiert, ob KI zur Verhinderung von Ladendiebstählen Verwendung finden soll. Dazu können Kunden algorithmisch überwacht werden, um aus der Analyse der Körpersprache Vorhersagen treffen zu können, ob sie einen Ladendiebstahl planen. Hierzu *Lewis*, „Should AI be used to catch shoplifters?“, *CNN Business* v. 18.5.2019. Abrufbar unter: <https://edition.cnn.com/2019/04/18/business/ai-vaak-shoplifting/index.html> (zuletzt abgerufen am 20.06.2019).

unavoidable. These costs and payment uncertainties can be avoided in person by using physical currency, but no mechanism exists to make payments over a communications channel without a trusted party. What is needed is an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Transactions that are computationally impractical to reverse would protect sellers from fraud, and routine escrow mechanisms could easily be implemented to protect buyers.”<sup>49</sup>

## **b) Predictive bzw. Big Data Policing**

Entsprechende Narrative tragen auch das sog. *Predictive Policing* und das sog. *Big Data Policing*, wie es ausgehend von den USA mittlerweile auch in anderen westlichen Staaten einschließlich Deutschland<sup>50</sup> Platz greift. Hier wie dort geht es um – so die Darstellung von Produzenten und Abnehmern – intelligente Algorithmen, die ressourcenknappen Hoheitsträgern schlagkräftige Instrumentarien zur prospektiven Straftatprävention an die Hand geben.<sup>51</sup> *Brantingham*, ein wissenschaftlicher Wegbereiter von *Predictive Policing*, definiert entlang eines dreistufigen Prozesses wie folgt:

„(1) data of one more type are ingested; (2) algorithmic methods use ingested data to forecast the occurrence of crime in some domain of interest; and (3) police use forecasts to inform strategic and tactical decisions in the field. A primary goal of predictive policing is to reduce uncertainty so that police can approach the allocation of resources in an optimal manner. The theory is that an optimal allocation of police resources has a better chance at disrupting opportunities for crime before they happen.”<sup>52</sup>

Vordergründig ist *Big Data Policing* ganz entsprechend aufgestellt. Auch hier geht es der Sache nach um algorithmische Vorhersagen, die Hoheitsträgern eine möglichst effektive wie effiziente – in dieser Logik „idealiter“ auch die personengenaue<sup>53</sup> – Straftatverhinderung ermöglichen sollen. Der technologische Unterschied zu *Predictive Policing* liegt in der Menge und der Qualität an Datensätzen, die in das *Big Data Policing* eingehen. Wie der Name zum Ausdruck bringt, werden große, scheinbar unzusammenhängende Datensätze

---

<sup>49</sup> *Nakamoto* (ein Pseudonym!), „Bitcoin: A Peer-to-Peer Electronic Cash System“, 2009, S. 1. Abrufbar unter: <https://bitcoin.org/bitcoin.pdf> (zuletzt abgerufen am 20.6.2019).

<sup>50</sup> Zu entsprechenden (Pilot-)Projekten in Deutschland in der Übersicht *Rademacher* (Fn. 43), S. 369.

<sup>51</sup> Bezeichnend der Koalitionsvertrag zwischen CDU Hessen und BÜNDNIS90/DIE GRÜNEN Hessen für die 20. Legislaturperiode, Zeile 2535 ff., wo nachzulesen ist: „Neue Instrumente wie spezielle Datenverarbeitungssysteme, die bereits vorhandene Informationen aus polizeilichen Datenbanken bündeln und auswerten, können bei der Bewältigung aktueller polizeilicher Herausforderungen von großem Nutzen sein.“

<sup>52</sup> *Brantingham*, „The Logic of Data Bias and Its Impact on Place-Based Predictive Policing“, *Ohio State Journal of Criminal Law* 2018, S. 473 (S. 473).

<sup>53</sup> *Brennan-Marquez*, „Big Data Policing and the Redistribution of Anxiety“, *Ohio State Journal of Criminal Law* 2018, S. 487 (S. 487).

verarbeitet, denen herkömmliche Analysewerkzeuge (einschließlich des menschlichen Gehirns) nicht gewachsen waren.

Hintergründig dürfte das Big Data Policing in seiner Adressierung weitergehen als das einfache Predictive Policing, ist es doch nicht nur an Hoheitsträger, sondern auch an Bürger gerichtet. Wie *Brennan-Marquez* treffend festgestellt hat, generiert Big Data Policing „a social order – a surveillance society – in which people constantly monitor and curate the data-trails they leave behind in everyday life.“<sup>54</sup> Anders ausgedrückt: Je (daten )intensiver Verhaltensvorhersagen zur Kriminalitätsinhibition ausfallen, desto mehr wird der Internalisierung des Kriminalitätsinhibitionsprojekts Vorschub geleistet, gehen also die externe (hoheitliche/öffentliche) und die interne (private) Normdurchsetzung Hand in Hand. In dieser Hinsicht scheint eine Rückbesinnung auf *Foucaults* Panoptismus von Interesse. Denn indem sich der Einzelne dem Projekt der technologischen Kriminalitätsinhibition durch eine Risikoüberwachung der Anderen verschreibt, macht er dies um den Preis der Überwachung des Selbst, nimmt all dies als eigenes an und wird so zum Vollstrecker der diesem Überwachungsprojekt unterliegenden Herrschaftsstrukturen.<sup>55</sup>

Dass die Predictive bzw. Big Data Policing zugrundeliegenden Algorithmen – so die bezeichnende Nomenklatur – „Entscheidungen“<sup>56</sup> treffen (insbesondere weil sie Korrelationen zwischen scheinbar unzusammenhängenden Datensätzen herstellen), wird von nicht wenigen Stimmen gerade in den USA für ausreichend erachtet.<sup>57</sup> Und zwar selbst dann, wenn diese Entscheidungen *nicht* nachvollziehbar bzw. erklärbar sind.<sup>58</sup> Zugespitzt sollen die Algorithmen also – im Gegensatz zu traditionellen Vorhersagemodellen – bewusst „atheoretisch“<sup>59</sup> gehalten werden können. Anders und mit den entsprechenden termini technici gesagt: Sog. „opake KI“ (*opaque AI*) wird für akzeptabel und „erklärbare KI“ (*explainable AI*)<sup>60</sup> für negligible erachtet.<sup>61</sup> Es liegt nahe, hierin eine gesteigerte Technologie-

---

<sup>54</sup> *Brennan-Marquez* (Fn. 53), S. 487.

<sup>55</sup> Euphemistisch heißt dies Selbstdokumentation, deren Kehrseite *Foucault's* Panoptikum des Selbst und ein Moment der Ausübung von Macht ist. Hierzu *Mau* (Fn. 22), S. 249 ff. – Weiterführend und einordnend auch *Han*, *Psychopolitik: Neoliberalismus und die neuen Machttechniken*, Frankfurt a.M.: Fischer, 2015, S. 84.

<sup>56</sup> Anthropomorphismen dominieren heute die soziale Konstruktion von KI, selbst wenn sie die technischen und algorithmischen Eigenheiten von KI in den Hintergrund treten lassen. Es ist zumindest missverständlich, dass KI-Systeme Entscheidungen treffen, da sie eigentlich nur Output generieren können.

<sup>57</sup> So etwa von *Henderson*, „A Few Criminal Justice Big Data Rules“, *Ohio State Journal of Criminal Law* 2018, S. 527.

<sup>58</sup> Obwohl dies scharf – m.E. zu Recht – von computer- oder rechtswissenschaftlicher Seite bestritten wird. Krit. etwa *Liu/Lin/Chen*, „Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability“, zur Erscheinung angenommen in: *International Journal of Law and Information Technology* 2019. Bereits jetzt verfügbar unter <https://ssrn.com/abstract=3313916> (zuletzt abgerufen am 20.06.2019).

<sup>59</sup> So wörtlich *Berk/Hyatt*, „Machine Learning Forecasts of Risk to Inform Sentencing Decisions“, *Federal Sentencing Reporter* 2015, S. 222 (S. 223).

<sup>60</sup> Die auch als „sich selbsterklärende KI“ denk- und ausgestaltbar ist.

und KI-Gläubigkeit zu erkennen, der ihrerseits durch einen Vertrauensverlust in menschliche Analyse- und Entscheidungsfähigkeiten Vorschub geleistet wird. Allemal belegt diese Sicht der Dinge die Vision einer durch KI getragenen Organisation und Lenkung menschlicher Sozialität und Interaktivität. Diese ist im Kontext des Strafrechts sozial besonders wirkmächtig. Denn wer würde schon bestreiten wollen, dass eine effektive und effiziente Kriminalitätsverhinderung aus gesamtgesellschaftlicher wie auch individueller (z.B. der notorischen Opfer-)Sicht wünschenswert ist?

## 2. Objektivität, Neutralität und Kohärenz in der Strafrechtsanwendung

War zuvor von der Versprechung die Rede, dass KI die Begehungsmöglichkeiten von Straftaten effektiv wie effizient zu minimieren im Stande ist, sei nun auf die weitere Versprechung einzugehen, dass sich durch den Einzug von KI in die (straf-)juristische Entscheidungsfindung deren Objektivität, Neutralität und Kohärenz zu verbürgen lassen soll. Die Vorstellung, dass Entscheidungen über die Verhängung von Untersuchungshaft wegen Wiederholungsgefahr, über die vorzeitige Haftentlassung wegen positiver Sozialprognose oder über die Strafzumessung jeweils auf der Grundlage algorithmischer Risikoeinschätzungen (*algorithmic risk assessment*) ergehen könnten, mag in der Richterrepublik Deutschland ob des scheinbar damit verbundenen Eingriffs in die richterliche Unabhängigkeit unerhört klingen. In den USA ist dies aber bereits gängige, höchstrichterlich abgesegnete Praxis.

Hierfür steht *State v. Loomis*, eine Entscheidung des Supreme Court of Wisconsin.<sup>62</sup> Der mehrfach vorbestrafte Beschuldigte *Eric Loomis* war verdächtigt, der Fahrer eines „drive-by-shootings“ gewesen zu sein. In der Folge bekannte er sich der Flucht vor der Staatsgewalt (*eluding an officer*) schuldig und trat dem Anklagepunkt des Fahrens eines PKW ohne Einverständnis des Eigentümers (*operating a vehicle without its owner's consent*) nicht entgegen. Hierfür wurde er zu einer Haftstrafe von sechs (sic) Jahren verurteilt. Diese drakonische Strafzumessung ging, wie das verurteilende Gericht offen einräumte, u.a.<sup>63</sup> darauf zurück, dass *Loomis* eine miserable Sozial- bzw. eine hohe Rückfallprognose attestiert wurde. Und zwar durch COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), einen proprietären (vulgo geheimen) Algorithmus, der durch Northpointe, Inc. entwickelt und vertrieben wurde. COMPAS berechnete auf der Grundlage ei-

---

<sup>61</sup> So etwa von *Henderson* (Fn. 57), S. 527.

<sup>62</sup> *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016) 754 (US). Dort auch zum Folgenden.

<sup>63</sup> Zu weiteren Details, insbes. der (strafprozessual in Wisconsin offensichtlich zulässigen, sic) strafschärfenden Verwendung eines weitergehenden Tatverdachts, die zu den Akten gegeben wurde, vgl. die Zusammenfassung des Verfahrensgangs in *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016) 754 (US).

ner komplexen Analyse eines 137 Punkte umfassenden Fragebogens sowie des öffentlichen Strafregisters des Beschuldigten dessen unmittelbares wie auch allgemeines Rückfallrisiko wie auch sein Risiko für die Gemeinschaft (*pre-trial and general risk of recidivism; risk to the community*). Dagegen legte *Loomis* Rechtsmittel wegen der Verletzung seines Rechts auf ein rechtsstaatliches Verfahren (*right to due process*) ein. Er rügte insbesondere, dass er keine Überprüfung der algorithmischen Abläufe vornehmen konnte, da diese als Geschäftsgeheimnis geschützt waren; dass keine individuelle Strafzumessung erfolgte, weil COMPAS mit generalisierenden Gruppendaten arbeitete; und dass der Algorithmus auch das Geschlecht des/der zu Bewertenden und damit eine unzulässige, da geschlechterdiskriminierende Variable prozessierte. – Der Supreme Court of Wisconsin verwarf das Rechtsmittel und auch der US Supreme Court nahm den Fall letztendlich nicht zur Entscheidung an, nachdem er zuvor noch die US Bundesregierung zur Stellungnahme aufgefordert hatte. Tragend wurde, dass der Einzelne kein „Recht auf Erklärung“ einer algorithmischen Risikoprognose genießen soll, solange er deren Input überblicken kann und über ihren Output informiert wird. Der Zugriff auf den Throughput, also z.B. warum einzelne Datenblöcke wie gewichtet werden, blieb *Loomis* somit rechtsgrundsätzlich verwehrt. Auch die Verwendung allgemeiner Gruppendaten sowie die Einbeziehung des Geschlechts in die algorithmische Risikoprognose wurde nicht gerügt, weil dies deren Genauigkeit verbessere und keine diskriminierende Zielsetzung verfolge. Freilich beeilte sich der Supreme Court of Wisconsin, klarzustellen, dass ein Richter eine entsprechende algorithmische Risikoprognose nur berücksichtigen, nicht aber als verbindlich werten dürfe.

*State v. Loomis* wird im Schrifttum – innerhalb und außerhalb der USA – kontrovers und mehrheitlich negativ aufgenommen.<sup>64</sup> Das darf aber nicht darüber hinwegtäuschen, dass die Algorithmen unterstützte juristische Entscheidungsfindung in den USA heute höchst-richterlich sanktionierte Praxis ist. Die Ursachen (wohlgemerkt: nicht die Gründe) hierfür sind vielfältig. Löst man sich von der eigentlich angezeigten jurisdiktionsspezifischen Analyse,<sup>65</sup> finden sich schnell die wiederkehrenden Motive des heutigen KI-Rechtfertigungsnarrativs. Diesem zufolge soll der Einsatz intelligenter Algorithmen die

---

<sup>64</sup> Vgl. exemplarisch *Berain*, „Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling“, *Law, Probability and Risk* 2018, S. 45; *Deskus*, „Fifth Amendment Limitations on Criminal Algorithmic Decision-Making“, *NYU Journal of Legislation and Public Policy* 2018, S. 237; *Liu/Lin/Chen* (Fn. 58), S. 122 ff.; *Huq*, „Racial Equity in Algorithmic Criminal Justice“, *Duke Law Journal* 2019, S. 1043 (S. 1081). – Vgl. auch *Ostermeier*, „Der Staat in der prognostischen Sicherheitsgesellschaft“, in: *Puschke/Singelstein* (Hrsg.), *Der Staat und die Sicherheitsgesellschaft*, Wiesbaden: Springer VS, 2017, S. 101 (S. 103).

<sup>65</sup> Um nur wenige Stichwörter zu geben: Ein tiefsitzender, heute auf höchster politischer Ebene noch befeu-erter Rassismus findet nicht zuletzt im Phänomen des sog. „mass incarceration“ Niederschlag und bedingt ein – trauriges, aber verständliches – Gefühl der Resignation, dass das US-amerikanische Strafrechtssystem nicht mehr mit herkömmlichen Mitteln zu „retten“ ist.

strafjuristischen Entscheidungsprozesse effektiver und effizienter machen. Es sollen genauere Risikoprognosen erstellt und frei werdende Ressourcen anderweitig, z.B. im Rahmen von Rehabilitationsprogrammen, eingesetzt werden können.<sup>66</sup> Und während Haft- und Strafzumessungs- überdies als Bauchentscheidungen, gleichsam als *black art* abgetan werden,<sup>67</sup> wird die algorithmisch fundierte Entscheidungsfindung dafür hochgehalten, dass sich der Einfluss von Vorurteilen, Voreingenommenheiten und Idiosynkrasien minimieren lasse.<sup>68</sup> Nicht das Recht (aufgrund seiner postulierten inhärenten Eigenschaften) oder der Rechtstab (aufgrund seiner Formation), sondern Algorithmen sollen mit anderen Worten die Objektivität, Neutralität und Kohärenz der Rechtsanwendung garantieren (und in den USA: retten). Hier begegnet uns einmal mehr die bereits oben konstatierte Vertrauensverschiebung, weg vom Vertrauen in Menschen und hin zum Vertrauen in Hochtechnologie. Gerade der Vertrauensfall in den Menschen erfolgt dabei auf ganzer Linie, werden hier doch auch die Rechtsanwender in der Sache als Risiko wahrgenommen, nämlich als Risiko für die Objektivität, Neutralität und Kohärenz der Rechtsanwendung.<sup>69</sup>

### III. DER STRAFRECHTSTHEORETISCHE UMGANG MIT DEN VERHEIßUNGEN VON KI

Die vorstehenden Ausführungen verdeutlichen, dass die Strafrechtstheorie die Verheißungen von KI nicht länger ignorieren darf. Dafür haben diese eine zu große soziale Zug- und Sprengkraft entwickelt und sind dabei, sich wirkmächtig in der Strafrechtspflege festzusetzen. Eine Bagatellisierung von KI (unten 1.) wie auch eine subversive Anzweifelung ihrer Verheißungen (unten 2.) scheint daher kein gangbarer Weg, um den aktuellen Herausforderungen begegnen zu können.

---

<sup>66</sup> Hierzu etwa *Botnick*, „Evidence-based Practice and Sentencing in State Courts: A Critique of the Missouri System“, *Washington University Journal of Law & Policy* 2015, S. 159 ff. (S. 166).

<sup>67</sup> So plakativ *Chiao* (Fn. 39), S. 238.

<sup>68</sup> Hierzu krit. *Martini/Nink*, „Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz“, *NVwZ-Extra* 2017, S. 1 ff. (S. 9).

<sup>69</sup> Ähnliche Entwicklungen scheinen auch in China Platz zu greifen, wobei dort die Kontrolle anscheinend abhängiger Richter im Vordergrund stehen dürfe. Instruktiv hierzu *Meng Yu/Guodong Du*, „Why Are Chinese Courts Turning to AI?“, *The Diplomat* v. 19.01.2019. Abrufbar unter: <https://thediplomat.com/2019/01/why-are-chinese-courts-turning-to-ai/> (zuletzt abgerufen am 20.06.2019).

## 1. Kriminalsoziologische Bagatellisierung?

Eine erste, fast schon unwillkürliche Reaktion auf die unter I. und II. geschilderten Verheißungen ist es, diese in ihren Auswirkungen und ihrer Potenz herunterzuspielen und ihnen daher mit tröstlichem Langmut zu begegnen. Insbesondere, weil viele der im Vorstehenden genannten Entwicklungen in den USA (oder gar in China) ihren Ursprung haben und daher für Europa und Deutschland keine Bedeutung entfalten „können“. Überdies wird KI nicht selten zur „normalen“ technologischen Innovation heruntergestuft, die die Strafrechtsordnung im üblichen Ausmaß und nach üblichem Muster zwar zu transformieren, nicht aber zu revolutionieren im Stande ist (vergleichbar etwa der Einführung moderner PKW, die ein eigentliches Verkehrsstrafrecht notwendig machte). Zu erwarten (und je nach Standpunkt: zu befürchten) wäre dann lediglich, dass sich die bekannten Entwicklungen des modernen Strafrechts fortsetzen werden, wie Vorverlagerungen der Strafbarkeit oder der Schutz von Kollektivrechtsgütern. Eine grundstürzende Infragestellung des Strafrechts wäre in der Folge strafrechtstheoretisch nicht zu vergegenwärtigen.

Beispielhaft ließe sich scheinbar (!) gut hören: Selbst wenn man *arguendo* die Verheißungen von Smart Contracts für bare Münze nehmen wollte, würde dies lediglich zu Verschiebungen und Anpassungen führen, wo, wann und wie kriminelle (betrügerische) Energie zum Tragen kommt.<sup>70</sup> Insbesondere wäre ein Ansteigen von Cyber-Kriminalität zu erwarten. Statt etwa einen Käufer direkt darüber zu täuschen, dass man über den zu verkaufenden Gegenstand als Verkäufer verfügt, wird Letzterer wohl die die Verfügungsgewalt verifizierende Datenbank („Blockchain“) manipulieren und damit diese Täuschung indirekt vornehmen. Weiterhin sind unproblematisch aktive Angriffe auf die einen Smart Contract konturierenden und ausführenden Computerprogramme wie auch die Ausnutzung ihrer Schwachstellen denkbar.<sup>71</sup> Soweit durch solche Entwicklungen Strafbarkeitslücken entstehen, wäre es am Gesetzgeber, sie zu schließen. Dabei wird dieser auf Vorverlagerungstatbestände setzen, die den Schutz von Kollektivrechtsgütern bezwecken (etwa die Integrität und Richtigkeit dezentral organisierter Datenbanken). – Ferner ließe sich mit Blick auf Predictive bzw. Big Data Policing einwenden, dass algorithmische Straftatvorhersagen wenn überhaupt nur für spezifische Kriminalitätsbereiche (wie Einbruchs- und Drogenkriminalität, der man durch örtlich

---

<sup>70</sup> Anleihen ließen sich hier ziehen bei *Durkheim*, Die Regeln der soziologischen Methode, Frankfurt a.M.: Suhrkamp, 1984, S. 86 und 156.

<sup>71</sup> Ein illustratives Beispiel hierfür liefert der sog. DAO-Hack. Hierzu *Heckmann*, „DAO-Hack: smart contracts auf dem rechtlichen Prüfstand“, CR 2016, R99.

ausgerichtete Polizeipatrouillen vorbeugen kann) funktionieren.<sup>72</sup> Der Großteil der Kriminalitäts„bekämpfung“ müsste daher weiterhin mit klassischen Mitteln betrieben werden. Zudem könnten Predictive bzw. Big Data Policing neue Kriminalitäts(erscheinungs)formen bedingen, denen dann wieder mit einer entsprechend modern aufgestellten Strafrechtspflege zu begegnen wäre. Man denke etwa an sog. *oracle attacks*, mit denen sich Kriminelle Kenntnisse über die Vorhersagen der entsprechenden Predictive bzw. Big Data Policing-Software verschaffen, um ihr kriminelles Verhalten entsprechend anzupassen (z.B. indem man genau dort einbricht, wo der Algorithmus keinen Einbruch vermutet). – Bagatellisieren lässt sich schließlich auch der Einfluss von KI in der strafjuristischen Entscheidungsfindung. Z.B. damit, dass die Strafzumessung zu komplex ist, als dass sie von Maschinen bewältigbar ist; oder damit, dass in *State v. Loomis* höchstrichterlich verfügt wurde, dass algorithmische Risikoeinschätzungen nur zur Unterstützung und Vorbereitung von unabhängigen justiziellen Entscheidungen herangezogen werden, erst letztere aber keinesfalls binden oder verbindlich präjudizieren dürfen.<sup>73</sup>

So eingängig und tröstlich kriminalsoziologische Bagatellisierungen von KI auch klingen mögen, und so wichtig es auch ist, durch KI ausgelöste Anpassungs- und Verdrängungsbewegungen *de lege lata et ferenda* im Blick zu behalten: die Strafrechtstheorie muss sich gleichwohl grundsätzlicher mit den Verheißungen von KI auseinandersetzen. Ansonsten würde sie sich unwillkürlich zu deren Steigbügelhalter machen und die entscheidenden „Anfangs“phasen der anstehenden Entwicklungen verpassen. Denn je mehr man strafrechtstheoretisch die Blößen von KI betont, desto mehr Anreize setzt man, dass solche Schwachstellen durch technologischen Fortschritt geschlossen werden. Die Bagatellisierung von KI führte also der Sache nach zu einer Entwicklungsspirale, die soziale Fakten schafft, ohne sich ihrer normativen Grundlagen bewusst zu sein.

Beispielhaft: Durch die Korrektur fehlerhaften Codes können *exploits* zusehends verunmöglicht, durch immer bessere Firewalls *hacks* erschwert oder aber *oracle attacks* vorweggenommen werden (indem die missbräuchliche Vorhersage der regulären

---

<sup>72</sup> Zur räumlichen Dimension von Predictive Policing etwa *Straube/Belina* „Policing the Smart City: Eine Taxonomie polizeilicher Prognoseprogramme“, in: Bauriedl/Strüver, *Smart City – Kritische Perspektiven auf die Digitalisierung in Städten*, Bielefeld: transcript Verlag, 2018, S. 223. Predictive Policing ist aber nicht auf räumliche Straftatprognosen beschränkt. Auch die – wohl höchst defizitäre – Auswertung von Fluggastdaten läuft unter der Rubrik Predictive Policing. Hierzu instruktiv SZ-Online v. 24.04.2019, „Überwachung von Flugpassagieren liefert Fehler über Fehler“.

<sup>73</sup> Hierzu die oben unter Fn. 62 nachgewiesene Entscheidung sowie ergänzend *Berain* (Fn. 64), S. 47; *Katyal*, „Private Accountability in the Age of Artificial Intelligence“, *UCLA Law Review* 2019, S. 54 (S. 86).

Vorhersagen vorhergesagt wird, mit der Folge, dass dort patrouilliert wird, wo eigentlich nicht patrouilliert werden sollte). Wie wichtig es ist, solche Entwicklungen von Beginn an strafrechtstheoretisch im Blick zu haben, zeigt abermals *State v. Loomis*. Denn selbst wenn die strafjuristische Entscheidungsfindung „nur“ durch Algorithmen unterstützt und vorbereitet werden sollte, erzeugt bereits dies – je nach Gestaltung – Ankereffekte<sup>74</sup> und einen nicht zu unterschätzenden „Compliance“-druck. Sollte es etwa Richtern freigestellt werden, mit einer eigenen die algorithmische Sozialprognose zu überstimmen, würde dies auf eine tatsächliche Prävalenz algorithmischer Methoden hinauslaufen (Stichwort: Arbeitszeiterparnis; Furcht vor negativen Reaktionen, wenn sich die eigene Prognose als falsch herausstellt und z.B. der dem algorithmischen „Rat“ zuwider Entlassene sofort rückfällig wird), so dass Strafzumessenden eine Verantwortungsverlagerung („blame shifting“) auf Algorithmen ermöglicht wird (frei nach dem Motto: „Nicht ich, sondern die Maschine ist verantwortlich!“). Diese Punkte müssen wohl bedacht sein, was sich durch eine Bagatellisierung von KI nicht erreichen lässt.

## **2. Informationstechnologische Unterminierung?**

Wichtiger als ihre strafrechtstheoretische Bagatellisierung ist es daher, die Verheißungen von KI von innen heraus kritisch zu beleuchten, sie gleichsam auf ihre informationstechnologische Haltbarkeit hin zu befragen. Freilich lässt sich auch so den Herausforderungen von KI nicht abschließend, sondern wenn überhaupt nur vorübergehend strafrechtstheoretisch beikommen. KI begegnet uns hier nämlich als die vielköpfige Hydra; sobald ein Kopf abgeschlagen wird, wachsen andere nach.

Gleichwohl gilt es nachdrückliche Zweifel an dem Objektivitäts- und Neutralitätsversprechen heutiger KI-Systeme anzumelden. So sind uneingeschränkt neutrale Algorithmen an sich nur schwerlich vorstellbar. Überdies haben datenverarbeitende Prognosen mit dem hier sog. „bias in, bias“-Problem<sup>75</sup> zu kämpfen.

### **a) Algorithmische Normativität**

Datenverarbeitende Algorithmen arbeiten mit – durch Menschen fixierten oder selbst erlernten – Gewichtungen, z.B. wenn dem Alter eines Straffälligen eine besondere Bedeutung bei der Kalkulation seines Rückfallrisikos beigemessen wird.<sup>76</sup> Kritisch wird es, wenn

---

<sup>74</sup> Hierzu mit Blick auf die Strafzumessung allgemein *Streng*, in: Kindhäuser/Neumann/Paeffgen (Hrsg.), *Strafgesetzbuch*, Baden-Baden: Nomos, 5. Aufl. 2017, § 46 Rn. 3; *Traut/Nickolaus*, „Der Ankereffekt: Schattenseiten im Strafprozess“, *StraFo* 2015, S. 485.

<sup>75</sup> In Anschluss an *Mayson*, „Bias in, Bias out“, *Yale Law Journal* 128 (2019), S. 2218.

<sup>76</sup> Hierzu etwa *Berk*, „Algorithmic criminology“, *2 Security Informatics* 2013, article 5, S. 4.

diesen Gewichtungen – gerade bei proprietären, also nicht überprüfbar Systemen – bewusste oder unbewusste normative Zielvorgaben unterliegen. Wie etwa *Berk/Hyatt* – sprachlich zwar etwas verbrämt, in der Sache aber in dankenswerter Offenheit – mit Blick auf Rückfallrisikoprognosen ausführen:

„Many criminal justice stakeholders will treat false negatives as more costly than false positives. When this policy preferences applies, the standard of statistical proof necessarily will be lower for forecasts of homicide. The intent is to not release an individual who will commit a homicide and, in trade, to accept a larger number of false positives.“<sup>77</sup>

Das bedeutet nichts anderes, als dass ein dergestalt normativ ausgerichteter Algorithmus „billigend in Kauf nimmt“ (bzw. genauer: die hinter dem Algorithmus Stehenden und diesen Verwendenden billigend in Kauf nehmen müssen), der Sache nach wohl ungefährliche Personen (sog. falsch Positive) in Haft zu lassen, um zu verhindern, dass der Sache nach wohl gefährliche Personen irrig als ungefährlich eingeordnet (sog. falsch Negative) und in der Folge aus der Haft entlassen werden. Es drängt sich auf, dass eine solche Programmierung strafjustizieller Entscheidungen – die auch dem deutschen Strafrecht leider nicht so fern ist, wie uns etwa die Debatten über die Sicherungsverwahrung vor Augen führen<sup>78</sup> – keinesfalls neutral und objektiv, sondern vielmehr hochgradig politisch und normativ aufgeladen ist.

Das wird freilich im Schrifttum in kritischer Absicht sogleich wieder ins Positive gewendet, nämlich indem es mit dem Versprechen verbunden wird, dass KI zur Offenlegung normativ offener und daher politisch zu fixierender Zielsetzungen zwingt.<sup>79</sup> Transparenz wird somit nicht nur zur Grundvoraussetzung, sondern zum normativen Gut des Einsatzes von KI in der Strafrechtspflege.

Bespielhaft hierfür steht die von *Chiao* jüngst zur (theoretischen) Diskussion gestellte Spielart algorithmischer Strafzumessung.<sup>80</sup> *Chiao* zufolge sollte ein Algorithmus nicht prospektive Risiken, sondern die Angemessenheit der retrospektiven Strafzumessung evaluieren; er soll also nicht kalkulieren, wie gefährlich ein Beschuldigter ist, sondern welche Strafe andere Richter in einer bestimmten Jurisdiktion in einem vergleichbaren Fall verhängen würden. So soll dem tatsächlich zur Strafzumessung berufenen Richter ein konkreter Richtkorridor an die Hand gegeben werden. Dieser Vorschlag greift traditionellen

---

<sup>77</sup> *Berk/Hyatt* (Fn. 59), S. 223.

<sup>78</sup> Vgl. nur *Boetticher u.a.*, „Zum richtigen Umgang mit Prognoseinstrumenten durch psychiatrische und psychologische Sachverständige und Gerichte“, NStZ 2009, S. 478 (S. 479 m.w.N.).

<sup>79</sup> So etwa *Berlain* (Fn. 64), S. 48 m.w.N.

<sup>80</sup> Hierzu *Chiao* (Fn. 39).

Kontrollmechanismen der Strafzumessung vor (wie beruflichen Sozialisierungsprozessen, instanzgerichtlichen Begründungspflichten und rechtsmittelrechtlichen Angemessenheits- bzw. Willkürprüfungen). Die Strafzumessung wird, so die Idee, im Einzelfall auf ihre Systemgerechtigkeit und -richtigkeit hin überprüft, bevor sie rechtsverbindlich ausgesprochen wird. Und da – wie *Chiao* natürlich erkennt – die Strafzumessung von den unterschiedlichsten, teils antinomischen Zielen und Zwecken bestimmt wird, müsste im zur Anwendung kommenden Algorithmus verbindlich festgelegt werden, ob und welche Ziele und Zwecke mit je welcher Gewichtung bestimmenden Einfluss gewinnen sollen. Mit all dem stellt sich *Chiao* in kritischer Tradition gegen die Intransparenz des subjektiv Politischen der Rechtsanwendung. An deren Stelle soll die Transparenz eines objektiv-politisch gestalteten Algorithmus treten.

Der Algorithmus wird damit gleichsam zur *bouche de la loi*.<sup>81</sup> Der aufklärerische Glaube an die ordnende und befriedende Kraft der Rationalität verbindet sich mit anderen Worten nicht länger mit dem menschlichen, sondern mit dem maschinellen Subsumtionsautomaten,<sup>82</sup> der die Normen des Gesetzgebers objektiv, neutral und kohärent zur Anwendung bringt.

### **b) *Bias in, bias out***

Die Neutralität und Objektivität heutiger KI-Systeme muss überdies durch das *bias in, bias out* in Zweifel gezogen werden. Dies betrifft unmittelbar die Funktionsweise von KI gestützten (Kriminalitäts-, Rückfallrisiko- oder *Chiao*'schen Strafzumessungs-)Prognosen, die durch die Auswertung aktueller Daten über vergangene Begebenheiten Schlüsse auf die Wahrscheinlichkeit zukünftiger Ereignisse ziehen.<sup>83</sup> Wenn diese vergangenen Begebenheiten dann freilich vorurteilsbehaftet sind bzw. durch vorurteilsbehaftete Datensätze rekonstruiert werden, reproduziert die Zukunftsvorhersage das Vorurteilsbehaftete der Vergangenheit in der Gegenwart.<sup>84</sup> Dies ist gerade in der US-Diskussion von ungemeiner Relevanz. Die die US-amerikanischen Kriminaljustizsysteme prägende Rassensegregation wird durch Datensätze repräsentiert, die einen Algorithmus wie von selbst dazu verleiten, einzelnen männlichen jungen Afroamerikanern heute eine sich morgen realisierende überproportional hohe kriminelle Energie zu attestieren, weil diese Bevölkerungsgruppe ges-

---

<sup>81</sup> Ob und wie eng diese Figur mit Montesquieu zu verbinden ist, kann hier dahinstehen.

<sup>82</sup> Allg. zur Figur des Subsumtionsautomaten *Ogorek*, Richterkönig oder Subsumtionsautomat? Zur Justiztheorie im 19. Jahrhundert, Frankfurt a.M.: Vittorio Klostermann, 1986.

<sup>83</sup> Eine instruktive Fallstudie aus den USA wird vorgestellt von *Berk*, „An Impact Assessment of Machine Learning Risk Forecast on Parole Board Decisions and Recidivism“, *Journal of Experimental Criminology* 2017, S. 193.

<sup>84</sup> Hierzu instruktiv auch *Singelstein* (Fn. 43), S.4.

tern überproportional häufig durch das Strafjustizsystem prozessiert (festgenommen, verurteilt, nicht vorzeitig aus der Haft entlassen etc.) wurde. Aus welchen sozialen (z.B. rassistischen) Ursachen dies „gestern“ erfolgte, bleibt im Ausgangspunkt algorithmisch außer Betracht.<sup>85</sup>

Ein solches *bias in, bias out* muss dabei nicht auf einem Fehler im System beruhen, sondern kann theoretisch auch System haben.<sup>86</sup> Das Objektivitäts- und Neutralitätsversprechen würde dann dazu beitragen, ein (z.B. rassistisch) bemakeltes Kriminaljustizsystem algorithmisch reinzuwaschen, als herrschaftsfrei zu mystifizieren und so in der sozialen Wahrnehmung zu legitimieren. All das kann und müsste eine kritische, KI in den Blick nehmende Strafrechtstheorie rügen.

Auswege aus der informationstechnologisch verbrieften Herrschaftspetrifizierung eines *bias in, bias out* versprechen neutralisierte Trainingsdatensätze, die normativ unerwünschte Inputfaktoren verwerfen (etwa solche, die direkt oder indirekt mit der Hautfarbe der zu Bewertenden zusammenhängen), sowie Algorithmen, die unerwünschte Verzerrungen in den Inputdaten normativ ausgleichen, gleichsam herausrechnen.<sup>87</sup> Interessanterweise wird dadurch die Verheißung, dass KI eine objektive, neutrale sowie kohärente Strafrechtspflege effektiver und effizienter als menschliche Entscheidungsträger organisieren könne, nicht nachhaltig in Frage gestellt. Im Gegenteil, sie wird auf einer höheren Ordnungsstufe reproduziert und als Entwicklungsziel ausgegeben. Paradoxerweise stabilisiert und legitimiert die im Vorstehenden skizzierte Kritik am jetzigen Einsatz von KI in der Strafrechtspflege den zukünftigen Einsatz einer weiterzuentwickelnden KI.

Die entscheidenden Zweifel daran, dass es sich bei den Versprechungen von KI um die sprichwörtliche „heiße Luft“ handelt und eine KI gestützte Strafrechtspflege quasi an sich selbst bzw. den zu hoch gesteckten Erwartungen scheitert, eben weil Algorithmen normativ programmiert sind und anhand voreingenommener Datensätze lernen, sind im Ergebnis weniger durchgreifend als sie zunächst erscheinen mögen. Denn diese Zweifel betreffen nur die konkrete Umsetzung, nicht aber die fundamentalen normativen Ordnungsvorstellungen einer algorithmischen Verbürgung effektiven wie effizienten Rechtsgüterschutzes sowie objektiver, neutraler und kohärenter Strafrechtsanwendung. Die Forderungen nach

---

<sup>85</sup> Hierzu allg. und krit. *Moffat/Montford*, „Unpacking Sentencing Algorithms“, in: Keijser/Roberts/Ryberg (Hrsg.), *Predictive Sentencing*, Oxford: Hart Publishing, 2018, S. 186 ff. Mit Blick auf die USA spricht *Ferguson*, „Illuminating Black Data Policing“, *Ohio State Journal Criminal Law* 2018, S. 503 (S. 504) von einem „black data problem [since data is] racially encoded, colored by the history of real-world policing that disproportionality impacts communities of color.“

<sup>86</sup> Dies stellt *Mayson* (Fn. 75) in den Raum.

<sup>87</sup> Hierzu ebenfalls krit. und m.w.N. *Mayson* (Fn. 75).

einer objektiv-politischen Programmierung und nach einer normativen Ausbalancierung von KI-Systemen sanktionieren und perpetuieren diese Ideale vielmehr. Ein Ignorieren, Bagatellisieren oder Unterminieren von „KI und Strafrecht“ hilft daher nicht weiter.

#### **IV. REFLEXION FUNDAMENTALER NORMATIVER ORDNUNGSPOSTULATE**

Um das Transformative der Verbindung von „KI und Strafrecht“ aufzuarbeiten, wird – da ein Ignorieren, Bagatellisieren und Unterminieren nicht ausreicht – eine Reflexion fundamentaler normativer Ordnungspostulate vonnöten; nämlich der Ordnungspostulate eines gewünschten Strafrechts wie auch jener der gewünschten Gesellschaft, die dieses Strafrecht hervorbringen soll. Da damit große und größte Fragen angesprochen sind, können hier nur einige kursorische Erwägungen angebracht, aber keine endgültigen Antworten versprochen werden, so dass im Folgenden an den entscheidenden Stellen in den Konjunktiv gewechselt werden muss.

##### **1. Strafrecht als liberales Freiheitsschutz- oder als wohlfahrstaatliches Sicherheitsrecht?**

Die eigentliche Herausforderung für die Strafrechtstheorie liegt darin, dass sich *KI prima facie* die zentralen Versprechungen des Strafrechts zu eigen macht – und sie optimiert. Kann unser „herkömmliches menschliches Strafrecht“ (jedweder theoretischer Provenienz) Rechtsgüterschutz<sup>88</sup> lediglich normativ und kontrafaktisch garantieren, weil Normbrüche an der Tagesordnung bleiben, will KI langfristig die faktische Verunmöglichung oder zumindest die substantielle Minimierung von Rechtsgutsverletzungen erreichen. Und kann unser „herkömmliches menschliches Recht“ nur eine normative und kontrafaktische Objektivitäts- und Neutralitätsgarantie für die juristische Entscheidungsfindung abgeben, weil persönliche Eigenheiten und Fehler der Rechtsanwender die „menschliche“ Regel bleiben, mit denen sich das Rechtssystem auch (insbes. aus pragmatischen Erwägungen) abgefunden hat,<sup>89</sup> will KI der Subjektivität, Voreingenommenheit und Ungleichbehandlung der Rechtsanwendung kategorisch vorbeugen. Ginge es strafrechtlichen Ge- und Verbotsnormen also darum, von den primären wie sekundären Normadressaten

---

<sup>88</sup> Zu diesem Paradigma nur *Hassemer*, Theorie und Soziologie des Verbrechens, Frankfurt a.M.: Athenäum, 1973, S. 27 ff.

<sup>89</sup> Beispielhaft sei an die Endlichkeit des Instanzenzugs und daran erinnert, dass die Rechtsschutzgarantie des Art. 19 Abs. 4 GG nach Maßgabe des (freilich umstrittenen) sog. Dürig'schen Dogmas nur Rechtsschutz durch, nicht aber gegen den Richter verbürgt.

flächendeckend<sup>90</sup> ernstgenommen zu werden, um so einen ernstlichen Schutz der geschützten Rechtsgüter zu garantieren, und nähme das Recht seinen Anspruch ernst, objektive, neutrale und kohärente Entscheidungen rechtstatsächlich flächendeckend garantieren zu wollen, so könnte die Strafrechtstheorie nicht prinzipiell gegen die Verheißungen von KI mobil machen. Das Faktische der (Verheißungen von) KI beschlösse mit anderen Worten das Ende des Kontrafaktischen (der Versprechungen) des Strafrechts – und damit das Ende im Sinne des krönenden Abschlusses des Strafrechts wie wir es kennen.

Eine grundsätzliche strafrechtstheoretische Kritik am Einzug von KI in die Strafrechtspflege müsste also nicht nur deren „fremde“, sondern vielmehr die „eigenen“, von KI ja vermeintlich nur übernommenen Zielsetzungen kritisch in den Blick nehmen. Zu Verteidigen und Rechtfertigen wäre mit anderen Worten nicht weniger als das „nurmehr“ Normative und Kontrafaktische der Strafrechtspflege. Will heißen: die reale Möglichkeit krimineller Rechtsgutsverletzungen wie auch die reale Möglichkeit voreingenommener, vorurteilsbehafteter, idiosynkratischer menschlicher Rechtsanwender.<sup>91</sup> An dieser Stelle begegnet uns die Differenz zwischen einer liberalen Fundierung des Strafrechts, die letzteres dem Freiheitsschutz verschreibt, und wohlfahrtstaatlichen Konzeptionen, die das Strafrecht zum Sicherheitsschutz einsetzen (können).<sup>92</sup>

Um den Verheißungen von KI prinzipiell entgegentreten zu können, müsste also – und da es hier nicht darum gehen kann, endgültige Antworten zu geben, kann und muss im Folgenden der Konjunktiv verwendet werden – in *liberaler Perspektive* das Dogma vom „Rechtsgüterschutz durch Strafrecht“ ergänzt werden. Bewerbstelligen ließe sich dies – (teilweise sehr!) frei nach *Haffke*,<sup>93</sup> *Tiedemann*<sup>94</sup> und *Prittwitz*<sup>95</sup> –, indem das Strafrecht

---

<sup>90</sup> Diese Qualifikation fordert dazu heraus, offener mit der faktischen Selektivität des Strafrechts – nicht nur auf supra-, sondern auch auf inner-nationaler Ebene – umzugehen, gleichsam den strafrechtlichen Rechtsgüterschutz nicht nur als fragmentarisch, sondern als konzeptionell selektiv zu denken. Denn der Grad an Effektivität, den KI verheißt, mag sich in der Lebenswirklichkeit schnell als gesellschaftlich zu teuer herausstellen. So wurde etwa in einer chinesischen Zeitung das dortige KI basierte Antikorruptionssystem, das bezeichnenderweise die *zero trust*-Maxime im Namen führt, wie folgt kritisch hinterfragt: „Is China’s corruption-busting AI system ‘Zero Trust’ being turned off for being too efficient?“ Hintergrund ist die Sorge, dass die chinesische öffentliche Verwaltung eine umfassende Verfolgung aller erkannten Korruptionsstraftaten nicht verkraften würde. Abrufbar unter: <https://www.scmp.com/news/china/science/article/2184857/chinas-corruption-busting-ai-system-zero-trust-being-turned-off-being-too-efficient> (zuletzt abgerufen am 20.06.2019).

<sup>91</sup> Z.B. indem das Recht in den Dienst der diskursiven, durch Rechtfertigung und Kritik angetriebenen Ausrichtung auf Objektivität, Neutralität und Kohärenz gestellt wird.

<sup>92</sup> Diese idealtypische Gegenüberstellung findet sich auch bei *Günther*, „Bedrohte individuelle Freiheiten im aufgeklärten Strafrecht – Welche Freiheiten?“, KJ 2016, S. 520.

<sup>93</sup> *Haffke*, „Die Legitimation des staatlichen Strafrechts zwischen Effizienz, Freiheitsverbürgung und Symbolik“, in: Schönemann u.a. (Hrsg.), Festschrift für Roxin zum 70. Geburtstag, Berlin: De Gruyter, 2001, S. 965.

<sup>94</sup> Prägnant zusammengefasst bei *Tiedemann*, Wirtschaftsstrafrecht: Einführung und Allgemeiner Teil, München: Vahlen, 5. Aufl. 2017, Rn. 228.

fundamental dem Freiheitsschutz verschrieben wird, dem andere (ggf. legitime, aber dann eben nicht in diesem Sinne strafrechtliche) Instrumente zur Prävention von Rechtsgutsverletzungen entgegenzustellen sind. Gemeint ist damit der unmittelbare Schutz eines Freiheitsgebrauchs (wie dem Alkoholkonsum frei von staatlicher Aufsicht und Kontrolle), der mittelbar auch den Freiheitsmissbrauch ermöglicht, also die Freiheit, Straftaten begehen zu können (wie fahrlässige oder vorsätzliche Alkoholfahrten im Straßenverkehr).

Nur so könnte in *individueller Hinsicht* weiterhin die Nichtbegehung einer Straftat als freiheitliche Entscheidung für und deren Begehung als freiheitliche Entscheidung gegen das Recht gewertet werden. In einer (hypothetischen) Welt, in der KI Straftaten *eo ipso* oder *de facto* unmöglich macht, kann von dieser Freiheit, dem berühmten Anders-Handeln-Können, und sei es auch nur als einer für ein liberales Gemeinschaftswesen notwendigen Fiktion, keine Rede mehr sein.

In *sozialer Hinsicht* waren überdies auch bis dato (d.h. ohne KI) Mittel und Wege denkbar, mit denen sich Rechtsgutsverletzungen *eo ipso* verunmöglichen oder zumindest *de facto* drastisch (und drakonisch) minimieren lassen.<sup>96</sup> Die Potentiale der KI spitzen in dem Sinne zu, was auch bis hierher im Rahmen einer rigiden wohlfahrtstaatlichen, nämlich administrativen oder „technischen Prävention“ (*Hassemer*<sup>97</sup>) denkbar war, nämlich eine Art „technologischen Paternalismus“ (*Hilgendorf*).

Wenn das Kontrafaktische des Strafrechts dem Faktischen effektiverer Instrumente der Straftatprävention vorgezogen werden soll, so müsste sich eine liberale Strafrechtstheorie offener mit saldierenden Freiheitsbilanzen auseinandersetzen. Der Freiheitsverlust der Vielen, die sich rigiden nicht-strafrechtlichen Maßnahmen zu unterwerfen haben, obwohl sie auch durch normative (strafrechtliche) Maßnahmen erreichbar sind, müsste dann schwerer als der mögliche Freiheitsverlust der Wenigen wiegen, deren Rechtsgüter durch jene verletzt werden, die normativ (strafrechtlich) nicht ansprechbar waren. Strafverfassungsrechtlich gedeutet: Strafrechtliche Ge- und Verbotsnormen müssten als *prima ratio* des Freiheitsschutzes theoretisiert werden, weil und wenn ein rigiderer (z.B.

---

<sup>95</sup> *Prittwitz*, „Strafrecht als propria ratio“, in: Heinrich u.a. (Hrsg.), *Strafrecht als Scientia Universalis*, Festschrift für Claus Roxin zum 80. Geburtstag, Band 1, Berlin/New York: De Gruyter, 2011, S. 23 ff.

<sup>96</sup> Im Wirtschaftsstrafrecht beispielhaft durch eine rigide administrative Supervision und Regulation der Wirtschaftsteilnehmer. Oder im Verkehrsstrafrecht durch eine technische Intervention gegenüber den Verkehrsteilnehmern. Hierzu denke man an verbindliche, mit einer Wegfahrsperre verbundene Alkoholtests vor Fahrtantritt. Auch das ist keine (!) Science Fiction, sondern wird konkret (!) auf EU-Ebene überlegt. Hierzu <https://www.europarl.europa.eu/news/en/press-room/20190410IPR37528/parliament-approves-eu-rules-requiring-life-saving-technologies-in-vehicles> (zuletzt abgerufen am 20.06.2018) m.w.N.

<sup>97</sup> *Hassemer*, „Aktuelle Perspektiven der Kriminalpolitik“, *StV* 1994, S. 333 ff. (S. 336).

administrativ superviser, regulativer oder technisch intervenierender) Rechtsgüterschutz *per saldo* nicht erforderlich oder nicht angemessen wäre, da er Einzelne oder die Allgemeinheit übermäßig belasten würde.

Illustrativ hierfür ist die reziproke Freiheitsbilanz, die etwa bei Hochrisikotechnologie (man denke etwa an die kommerzielle Atomkraft) aufgemacht wird. Hier überwiegen die möglichen Freiheitseinbußen der Vielen (die etwa bei einem atomaren Super-GAU zu befürchten wären) den Freiheitsverlust der Wenigen, die rigiden nicht-strafrechtlichen Maßnahmen unterworfen werden (etwa einer engmaschigen staatlichen Kontrolle der Betreiber von Atomkraftwerken). Der schmerzhafteste Preis dieser Rekonstruktion ist es, dass die Opfer von Straftaten ihre realen Freiheitseinbußen hinnehmen müssen, um den Anderen ihren virtuellen sonstigen Freiheitserhalt zu sichern. Plakativ: Den Eltern eines Kindes, das überfahren wurde, weil ein Autofahrer fahrlässig auf den zugelassenen „Autopiloten“ seines PKW vertraute, muss offen gesagt werden, dass das Verbot solcher „Autopiloten“ gesellschaftlich zu „teuer“, z.B. zu innovationsfeindlich, und auch nicht im Sinne der Vielen war, die diese „Autopiloten“ verkehrsgerecht zur Realisierung ihrer Freiheit einsetzen.

Ein so zu legitimierendes liberales Strafrecht könnte gegen die Verheißung eines effektiven wie effizienten Rechtsgüterschutzes durch KI in Stellung gebracht werden. Das setzte freilich gehörigen kriminalpolitischen und straftheoretischen Mut voraus. Immerhin müsste beispielhaft die öffentliche wie private Fremd- und Eigenüberwachung, die notwendig ist, um Predictive bzw. Big Data Policing zu ermöglichen, als übermäßig freiheitsinvasiv klassifiziert werden. Und der virtuelle Freiheitsgewinn der Vielen, die keiner Überwachung ausgesetzt werden, müsste dem realen Freiheitsverlust der Wenigen vorgezogen werden, die Opfer von Straftaten werden, die sich durch den Einsatz von KI (wohl) hätten verhindern lassen.

Die notwendig werdenden Abwägungen und Freiheitsbilanzierungen werden auch in Zukunft bereichsspezifisch ausfallen müssen. Das Straf- als liberales Freiheitsschutzrecht müsste also nicht überall und generell zur Geltung gebracht werden. Die liberale Strafrechtstheorie ist damit aber gehalten, Freiheit und Rechtsgüterschutz in einen offenen Abwägungsprozess einzustellen.

## 2. Welches Strafrecht für welche Gesellschaft?<sup>98</sup>

Der Druck, Farbe zu bekennen, stellt sich ebenfalls ein, sobald man die unterschwellig normativen Ordnungspostulate von KI in den Blick nimmt, die durch viele oberflächliche Rechtfertigungsnarrative (effektiver und effizienter; objektiver, neutraler und kohärenter) leicht verdeckt werden. Insofern muss sich die Strafrechts- zur Gesellschafts- oder politischen Theorie öffnen und zu Grundlegendem (insbesondere zum gesellschaftlichen Rang von Vertrauen und Zukunftsoffenheit) Stellung beziehen.

Wie gezeigt wurde (oben II.), ist und wäre der Einsatz von KI in der Strafrechtspflege Ausdruck eines fundamentalen Verlusts zwischenmenschlichen Vertrauens. Dem Anderen (einschließlich des Rechtsanwenders) soll nicht mehr zu vertrauen sein; er soll vielmehr als potentieller Gefährder und als Risiko geführt werden, den es zu überwachen und dessen zukünftiges Verhalten es algorithmisch zu antizipieren gilt. Das läuft in der Eigenlogik im Ergebnis auf eine Generalisierung automatisierten Misstrauens – oder strafprozessual gesprochen: eines automatisierten und generellen Anfangsverdachts – hinaus.<sup>99</sup>

Um den Verheißungen von KI prinzipiell entgegenzutreten, müsste die Unschuldsvermutung gesellschaftsstraftheoretisch höher aufgehängt, nämlich umfassender als kontrafaktische zwischenmenschliche Vertrauensvermutung und damit z.B. ganz aufklärerisch – wie *Hrschuka* dies glasklar herausgearbeitet hat – als „Jedermanns Würde“ rekonstruiert werden.<sup>100</sup> Dies setzte den Schulterschluss mit einer Sozialphilosophie voraus, die Vertrauen nicht als bloßen Mechanismus zur Reduktion sozialer Komplexität (*Luhmann*), sondern als funktional wertvoll führt, um ein Abgleiten in eine vertrauenslose Überwachungsgesellschaft mit allemal potentiell autoritärem und oppressivem Charakter zu verhindern.<sup>101</sup>

Nicht minder anspruchsvoll ist es, dem Ideal eines *end of history*<sup>102</sup> zu begegnen. Die Funktionsweise von KI proklamiert, wie wir gesehen haben (oben II.), in der Sache eine Art von Ende der Geschichte, gleichsam eine Schließung der Zukunft. Immerhin werden aus der Vergangenheit Rückschlüsse auf die Zukunft gezogen, was – im Guten wie im Schlechten – zu einer Petrifizierung des Gestern im Heute führt und die dynamische Ent-

---

<sup>98</sup> Diese Pointierung meiner Fragestellung entstammt Diskussionen im Kreis der Frankfurter Kolleginnen und Kollegen und verdanke ich letztlich *Klaus Günther*.

<sup>99</sup> *Brennan-Marquez* (Fn. 53), S. 488 diagnostiziert daher treffend, dass unter dem Eindruck von Predictive und Big Data Policing die verfassungsrechtliche Figur des Anfangsverdachts, die staatliche Grundrechtseingriffe legitimiert, (bis zur Unkenntlichkeit) geschleift wird.

<sup>100</sup> *Hrschuka*, „Die Unschuldsvermutung in der Rechtsphilosophie der Aufklärung“, *ZStW* 112 (2000), S. 285.

<sup>101</sup> Ob hierfür der von *Parsons*, *Politics and Social Structure*, New York: Free Press, 1969 geprägte Vertrauensbegriff, der als Antwort auf die hobbesche Situation formuliert einen politischen Ordnungsbegriff darstellt, in Ansatz gebracht werden kann, muss hier nicht weiter expliziert werden.

<sup>102</sup> Nachgerade klassisch *Fukuyama*, *The End of History and the Last Man*, New York: Free Press, 1992.

wicklung des Morgen hemmt. Eine normwidrige Zukunft wird, so das Versprechen, kraft KI tatsächlich verhinderbar.

Darauf zielt auf den ersten Blick auch die traditionelle Strafrechtstheorie. Dem Einzelnen wird die Sorge vor zukünftigen Rechtsgutsverletzungen genommen, weil und indem der zukünftige Bestand dieser Rechtsgüter im Jetzt strafrechtlich garantiert wird. Der Unterschied liegt einmal mehr darin, dass KI ein faktisches und Strafrecht ein kontrafaktisches bzw. normatives *end of history* verspricht. Das hat Folgewirkungen. In einer idealen (utopischen oder dystopischen) KI-Welt sind Dissidenz und Widerstand gegen den als *status quo* reproduzierten *status quo ante* nicht nur zwecklos (weil und wenn sie *eo ipso* oder *de facto* unmöglich gemacht werden), sondern idealiter auch unvorstellbar (insbesondere weil und wenn der Einzelne sich zum Vollstrecker seiner eigenen Unterwerfung unter eine panoptische Herrschaftsform macht).

Wer diese Schließung der Zukunft ablehnen und für deren Öffnung eintreten wollte, könnte den Eigen- und Mehrwert des Kontrafaktischen des Strafrechts in der faktischen Zulassung von Dissidenz und Widerstand finden. Deviantes Verhalten dürfte dann nicht länger allein als verhinderungsbedürftige Rechtsgutsverletzung abgestempelt, sondern zumindest auch als potentielle (objektive oder subjektive) Kritik am *status quo* anerkannt werden (man denke nur an das strafrechtliche Verbot homosexuellen Geschlechtsverkehrs, das nicht zuletzt durch fortwährende Akte des Widerstands, nämlich durch Akte des Normbruchs, zu Fall gebracht wurde). Das bedeutete, dass in der Auseinandersetzung mit dem Normbruch die strafrechtliche Reaktion immer wieder mit guten Gründen gerechtfertigt werden müsste und diese nicht z.B. als „natürlich“ hingestellt werden dürfte. Strafrecht müsste also als evolutive und diskursive Praxis von (menschlicher) Rechtfertigung und Kritik konzipiert sowie die zeitliche Kontingenz des Strafrechts akzeptiert werden, um die offene Zukunft des Strafrechts der geschlossenen Zukunft der KI entgegensetzen zu können.

## **AUSBLICK**

Dieser Beitrag wirbt darum, KI als soziales Konstrukt in den Blick zu nehmen, dessen Praxis unsere grundlegenden gesellschaftlichen und (straf-)rechtlichen Ordnungsvorstellungen zu transformieren im Stande ist. Angesichts konkreter strafrechtlicher Applikationen sollte KI als transformative Technologie nicht länger strafrechtstheoretisch ignoriert oder bagatellisiert werden. Und da auch interne informationstechnologische Zweifel an der Leis-

tungsfähigkeit von KI nur weitere Entwicklungsspiralen befördern, gilt es die normativen Ordnungspostulate, die in KI einprogrammiert sind, kritisch zu hinterfragen. Insbesondere ist zu hinterfragen, was der Preis der Verheißungen von KI ist, einen effektiveren und effizienteren Rechtsgüterschutz verbürgen und eine neutralere, objektivere und kohärenter Strafrechtsanwendung garantieren zu können (Stichwort: Obsiegen des wohlfahrtstaatlichen Sicherheits- und *zero trust*-Paradigmas in einer Big Data basierten Überwachungsgesellschaft). Umgekehrt würde es sich – wie hier nur abschließend in den Raum gestellt werden kann – eine non-ideale Strafrechtstheorie zu einfach machen, dem Einzug von KI in die Strafrechtspflege schlicht (normative) ideale Setzungen entgegensetzen zu wollen (Stichwort: Straf- als Freiheitschutzrecht; Recht als diskursive Praxis von Rechtfertigung und Kritik).<sup>103</sup> Denn das setzte ein besonders anspruchsvolles Strafrecht voraus, das von allen autoritären und oppressiven Anleihen befreit ist (und nicht „nur“ sein sollte).<sup>104</sup> Wer dies zwar als wünschenswertes, aber in kritischer Intention<sup>105</sup> wenig realistisches Ziel ausgeben will, und wer von einer faktisch immer weiter voranschreitenden Digitalisierung der Gesellschaft ausgehen muss, der wird im Endeffekt darüber nachzudenken haben, wie ein im Grundsatz liberal, rechtsstaatlich und demokratisch auszurichtendes Strafrecht so um KI-Systeme ergänzt werden kann, dass KI das Strafrecht einerseits nicht autoritär korrumpieren und es vielmehr andererseits von illiberalen, rechtsstaatswidrigen und autoritären Anwendungen befreien kann.<sup>106</sup> In diesem (non-idealen und kritischen) Sinne müsste KI dann nicht als Ende des Strafrechts (weder im Sinne des siechenden Hinscheidens eines liberalen Freiheitsstrafrechts noch im Sinne des krönenden Abschlusses eines wohlfahrtstaatlichen Sicherheitsstrafrechts), sondern als Baustein eines heute zu entwerfenden Strafrechts der (nahen) Zukunft entworfen werden, das sich zugleich technologieoffen<sup>107</sup>

---

<sup>103</sup> Um damit die berühmte, von *Rawls* losgetretene methodische Debatte zwischen idealer und non-idealer Theoriebildung für die Strafrechtstheorie anzudeuten.

<sup>104</sup> So allg., d.h. ohne spezifischen KI Bezug, *Naucke*, „Abhandlung über das Strafrecht als Machtbegrenzung“, in: *Negatives Strafrecht*, Berlin: Lit Verlag, 2015, S. 69 ff. (114 ff.).

<sup>105</sup> Ideale Theoriebildung sieht sich der Kritik ausgesetzt, den status quo zu stabilisieren. Eine non-ideale Theoriebildung lässt sich keinesfalls von aktuellen Entwicklungen vor sich hertreiben, nimmt diese aber zur Kenntnis und will sie verstehen, um sie so zum Bezugspunkt kritischer Reflexionen zu machen.

<sup>106</sup> Z.B. indem KI aufklärerisch zum Einsatz gebracht wird, um Richtern ihre Vorurteile etc. vor Augen zu führen. Hierzu eindrücklich *Sommer*, „Psychologie der richterlichen Entscheidungsfindung“, ZRP 2017, S. 60, der eine wissenschaftliche Erforschung richterlicher Vorurteile für geboten erachtet, weil sie hierzulande noch „tabu“ sei. – Rechts- und Urteilsmethodisch ist insofern weiterhin überaus lesenswert *Hruschka*, „Rechtsanwendung als methodisches Problem“, ARSP 50 (1964), S. 485 (insbes. S. 498).

<sup>107</sup> Der Versuch, das Strafrecht vor technologischen Innovationen im Recht zu schließen, kann bestenfalls künstlich wirken. So sieht das französische Strafrecht nun folgende Bestimmung vor: „Les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées.“ Abrufbar unter, dort auch mit dem Hinweis, wo diese Vorschrift eingefügt werden soll: <https://www.legifrance.gouv.fr/eli/loi/2019/3/23/JUST1806695L/jo/texte> (zuletzt abgerufen am 20.06.2019). – Daraus spricht die Sorge, justizielle Entscheidungsträger algorithmisch „durchleuchtbar“ zu machen. Wenn damit die Sorge verbunden sein sollte, dass algorithmisch deutlich gemacht werden kann, dass diese Ent-

*und* menschengerecht und damit in einem modernen Sinne weiterhin vernünftig zeigt. Bei alledem gilt es das – gerade mit Blick auf KI nicht immer leicht umsetzbare – große Vermächtnis von *Joachim Hruschkas* zu beherzigen, dass die Strafrechtswissenschaft stets *intellektuell redlich* bleiben muss, sich also rationaler Argumentation zu bedienen und Beschwörungswörter sowie Zauberformeln pedantisch zu vermeiden hat.<sup>108</sup>

---

scheidungsträger nicht objektiv und neutral das Recht anwenden, sondern durch Subjektives gefärbt entscheiden, stellt sich freilich die Frage, ob der strafrechtliche Schutz eines Mythos oder bloßen Ideals (also die Objektivität und Neutralität der Rechtsanwendung) oder eines nicht hinterfragbaren Vertrauens in justizielle Entscheidungsträger ein legitimes und sinnvolles Rechtsgut darstellen.

<sup>108</sup> So ausdrücklich *Hruschka*, Strafrecht nach logisch-analytischer Methode (Fn. 4), S. XVIII.