Implications and Limitations of the Use of AI in Criminal Justice in Germany

von Prof. Dr. Carsten Momsen[*]

Abstract
Auch in der deutschen Strafverfolgungspraxis haben Beweiserhebungen in Form von massenhaft erhobenen Daten erheblich an Bedeutung gewonnen. Gleichwohl spielen Big-Data-Analysen (noch) eine größere Rolle als der Einsatz spezifischer KI, da diese über die Sammlung und Analyse der verschiedenen Datenströme hinausgeht. Ein etwas anderes Bild ergibt sich, wenn man die präventive Gefahrenabwehr betrachtet. Hier wird KI zunehmend eingesetzt, vor allem im Bereich der vorausschauenden Polizeiarbeit (Predictive Policing), wo sie über die retrograde Analyse hinausgeht, um echte Vorhersagen zu treffen und damit menschliche Entscheidungen (zumindest) konkret zu prognostizieren oder gar vorwegzunehmen. In den letzten Jahren hat sich ein neues Phänomen herauskristallisiert: Dieselben Werkzeuge werden sowohl im Sicherheitsbereich als auch bei der Strafverfolgung eingesetzt. Diese Gemengelage ergibt sich vor allem aus der Überwachung von "Gefährdern" im Bereich der präventiven Terrorismusbekämpfung. Neu ist vor allem, dass die rechtlichen Voraussetzungen für den Einsatz in Deutschland nahezu identisch normiert sind. Ein weiteres Beispiel aus dem Bereich der allgemeinen Kriminalität ist der Einsatz der sogenannten "erweiterten DNA-Analyse" (Forensic DNA Phenotyping – FDP). Auch hier gelingt die Analyse der gespeicherten Daten in großem Umfang nur mit algorithmen-basierten Programmen, die mit bestimmten Suchkriterien ausgestattet und mit sog. „Lerndaten" gefüttert werden. Damit zeigt sich zugleich der menschliche Einfluss auf das Programm: Desgin des Algorithmus, Auswahl der Lerndaten, Auswahl der Entscheidungs- und Zuordnungskriterien sowie die eigentliche Bewertung des Ergebnisses mit Bezug auf die zutreffende Entscheidung (reichen 95% oder bedarf es 99,9% Übereinstimmung?) werden außerhalb des Datenverarbeitungsvorgangs von menschlichen Akteuren getroffen. Dabei bleiben der nachfolgenden Ebene in der Regel die Handlungs- und Zuordnungsparameter der vorherigen Ebene verborgen.

Damit ergibt sich die Frage, ob die individualschützenden, dem Grundsatz des Nachweises der individuellen Schuld verpflichteten, rechtlichen Standards der Beschuldigtenposition aufgeweicht oder doch zumindest denjenigen der Gefahrenabwehr angenähert werden. Dabei muss man berücksichtigen, dass die Entwicklungen weit fortgeschritten sind und möglicherweise faktisch notwendig, um mit neuen Kriminalitätsformen und Beweisformen Schritt halten zu können. Ebenso ist die Frage, ob die Standards bei herkömmlichen Beweismitteln nicht vergleichbare Defizite aufwiesen. Last not least ist es aber von zentraler Bedeutung, dass die Ursache des Problems in der Regel nicht das digitale Beweismittel ist. Vielmehr ist die Verlagerung des Beweiswürdigungsvorgangs der entscheidende Punkt. Diese häufig als ureigenste Domäne der Tatgerichte bezeichnete (Be-) Wertung der Beweislage könnte zunehmend durch Algorithmen oder von diesen gesteuerte KI ersetzt werden. Sollen diese Ergebnisse lediglich zur Gegenkontrolle menschlicher Entscheidungen dienen, oder wie die US Sentencing Guidelines zwingend bei der Entscheidungsfindung berücksichtigt werden, so entstehen für die Gerichte zwar erhöhte Begründungserfordernisse, diese können jedoch zur Vermeidung von Wahrnehmungs- und Wertungsfehlern hilfreich sein. Vorausgesetzt, die Gerichte sind sich der in dem Algorithmus selbst potentiell angelegten Fehlerquellen und Verzerrungspotentiale bewusst.

Evidence gathering in the shape of mass-collected data has also gained considerable importance in German law enforcement practice. Nevertheless, Big Data analyses (still) still is more important than the use of specific AI, as this goes beyond the collection and analysis of the various data streams. A somewhat different picture emerges when looking at preventive threat defense. Here, AI is increasingly being used, especially in predictive policing, where it goes beyond retrograde analysis to make real predictions and thus anticipate human decisions (at least). In recent years, a new phenomenon has emerged: The same tools are being used in both security and law enforcement. This mixed situation arises primarily from the surveillance of "potential offenders" in the field of preventive counter-terrorism. What is new above all is that the legal requirements for their use in Germany are standardized almost identically. Another example from the area of general crime is the use of so-called "extended DNA analysis" (Forensic DNA Phenotyping - FDP). Here, too, the analysis of stored data on a large scale is only possible with algorithm-based programs that are equipped with specific

* *Prof. Dr. Carsten Momsen* leitet den Arbeitsbereich „Vergleichendes Strafrecht, Strafverfahrensrecht, Wirtschafts- und Umweltstrafrecht" an der Freien Universität Berlin. Auf der Grundlage dieses Textes hat der Verfasser am 12.11.2022 einen Vortrag auf der Arbeitstagung des Kriminalpolitischen Kreises in Würzburg gehalten. Änderungen auf der Grundlage der dortigen Diskussion sind in den Text eingeflossen. Die ursprüngliche Version wurde als Online-Publikationen des Center for International Human Rights am John Jay College of Criminal Justice der City University New York (CUNY) veröffentlicht. Der Verfasser dankt für die Ermöglichung der weiteren Veröffentlichung.

search criteria and fed with so-called "learning data". This shows at the same time the human influence on the program: Desgin of the algorithm, selection of the learning data, selection of the decision and assignment criteria as well as the actual evaluation of the result with reference to the applicable decision (are 95% or only 99.9% agreement enough) are made outside of the data processing procedure by human actors. In this process, the action and assignment parameters of the previous level usually remain hidden from the subsequent level.

This raises the question of whether the individual-protective legal standards of the accused position, which are committed to the principle of proving individual guilt, are being softened or at least approximated to those of danger defense. It must be borne in mind that developments are far advanced and may be factually necessary to keep pace with new forms of crime and evidence. Likewise, there is the question of whether the standards for conventional forms of evidence did not have comparable deficits. Last but not least, however, it is crucial to note that the source of the problem is usually not the digital evidence. Rather, the shift in the process of evaluating evidence is the crucial point. This (evaluation) of the evidence, which is often described as the very own domain of the factual courts, could increasingly be replaced by algorithms or AI controlled by them. If these results are to serve merely as a cross-check of human decisions or, like the U.S. Sentencing Guidelines, are to be mandatorily taken into account in the decision-making process, this will create increased justification requirements for the courts, but these can be helpful in avoiding errors of perception and evaluation. Provided that the courts are aware of the potential sources of error and bias inherent in the algorithm itself.

I. Introduction

In law enforcement practice globally, Big Data analytics (still) play a greater role than the use of specific AI, insofar as this goes beyond the collection and analysis process of the various data streams. A somewhat different picture emerges when looking at threat prevention. Here, AI is increasingly being used, especially in the area of predictive policing, where it goes beyond retrograde analysis to make real predictions and thus (at least) predetermine human decisions. In recent years, a new phenomenon has emerged: the same tools are being used in both security and law enforcement. This mixed situation arises primarily from the surveillance of "threats" in the area of preventive counterterrorism. What is particularly novel is the fact that the legal requirements for use are standardized almost identically in Germany. Another example of this from the area of general crime is the use of so-called "extended DNA analysis" (Forensic DNA Phenotyping – FDP).[1]

1. Changes in Central Areas of German Criminal Procedure

In addition to these structural aspects, the new technologies harbor specific risks. Some central potential dangers of the use of AI can be described with the key words "lack of understanding of processes," "lack of transparency," "lack of individual fairness," "promotion and reinforcement of existing inequalities," "lack of evaluation level," and "problem of trust-based decisions."[2] In addition, there are documented instances of discrimination and racism in algorithmic bias and datasets, behind which lie real social problems, including in law enforcement. The use of data analytics seems to make these problems at least more salient, as tendentious analyses can often still be tracked.

Bias can arise outside the analysis process performed by the AI, particularly by those involved in the design, process, and evaluation of data processing. In this context, the origin of such an error is de facto in the human domain – such as incorrectly or carelessly selected learning data or a lack of ability to interpret it properly. However, within the processes automated with the help of an AI, certain factors, such as the design of the algorithm based on certain assumptions and classifications, can also amplify, modify or distort the error. AI can increase the impact of errors or make them more difficult to detect. Multiple errors can occur, for example, because both the selection of the input data and the interpretation of the result of an analysis as evidence are not carried out appropriately.[3] In the area of security and also law enforcement, they can lead to serious misinterpretations and misjudgments, such as the surveillance and prosecution of innocent people or the violation of elementary principles, e.g., the presumption of innocence.

2. Law Enforcement and Public Safety in Germany

In addition, there is a tendency to mix the tasks of the police – which, at least in Germany, are historically and constitutionally separate: preventive danger defense and reactive prosecution of criminal offenses. If the same tools and data sets are used in both areas, individuals and groups that were only classified as dangerous with a certain probability may automatically become suspects if corresponding crimes are committed. The overlap is clearly visible at the boundary between predictive policing and suspect investigation.

Central elements of criminal proceedings are being - one could call it - "policed". This applies, for example, to the concept and function of suspicion as well as to the concept, function and legal status of the accused.

---

1 *Momsen/Weichert*, freispruch 13/2018, 37 f.

2 *Momsen/Rennert*, KriPoZ 2020, 160 ff.

3 *Momsen*, in: FS Beulke, 2015, S. 871 f.; *ders.*, in: FS Heintschel-Heinegg, 2015, S. 313 f.; *ders.*, in: Beck/Meier/Momsen (Hrsg.), Cybercrime und Cyberinvestigations – Neue Herausforderungen der Digitalisierung für Strafrecht, Strafprozessrecht und Kriminologie, 2015, S. 67 ff.

3. Privatization and Internationalization of Criminal Proceedings

Another change is the increasing involvement of private institutions with their own (profit-oriented) interests in security and criminal justice that are not primarily committed to the common good. In the resulting relationship between private prosecutors and private suspects or defendants, fundamental rights apply only to a limited extent, even if one assumes the theoretical third-party effect. At the same time, this weakens the fundamental rights-related safeguards of individual rights, especially the position of the accused in criminal proceedings. Therefore, a new legal protection paradigm needs to be developed that is tailored to this changed situation. This is made more difficult by the fact that private actors are increasingly entering the playing field in positions that are central to investigations. For example, the companies that design the tools and program the algorithms, but also those that (have to) make the data they collect available for completely different purposes. These often have contractual relationships with data subjects, as do social media providers. Citizens are thus no longer confronted only with the state, but with an opaque mixture of state authority and private (de facto or contractual) power. In addition, due to the structure of private actors, but also due to the increasing exchange of data between national authorities, for example within the EU, many questions of individual rights protection are taking on an international component. In addition to national data protection law, international data protection regulations and agreements are therefore also gaining influence on law enforcement.[4]

4. Human Rights – a New Architecture of Procedural Rights?

If technologies are to be used responsibly, it is therefore necessary to design a newly coordinated set of institutions to safeguard individual legal positions in criminal proceedings.

Due to their international structure, human rights come into focus here. Human rights, as formulated in the classic form in the Universal Declaration of Human Rights in 1948, must be adapted to living conditions in a digitized environment. Human rights must also be addressed vis-à-vis private individuals (companies) when they become an inseparable part of the government's power structure or themselves act as a public authority vis-à-vis citizens who are in fact hierarchically subordinate.[5] In part, this leads to a moderate reshaping of core rights such as privacy. In some cases, however, rights need to be reshaped to ensure vital access to digital resources. In some cases, European legal systems are ahead of U.S. legal practice in this regard, particularly with respect to privacy and so-called fundamental IT rights. However, many human rights also need to be completely rethought to ensure that the ideas originally associated with them remain valid in the digital environment, as discussed in the in augural white paper of this series on the concept of digital citizenship.

The problem of the human rights approach is also well known and touches on this discussion in various ways, as part of an emerging paradigm of digital citizenship outlined in the inaugural white paper in the CIHR-John Jay College - series.[6] Assuming that human rights are recognized by most states, legal relationships between private actors, companies and users or otherwise affected parties, would also have to be included in the scope. Accordingly, a distinction must be made between those private actors who can invoke the protection afforded by human rights (legal entities) and those on whom a corresponding obligation to protect is to be imposed in parallel with state actors. The latter group in particular needs to be outlined. The manifold considerations and regulations on "Corporate Social Responsibility" (CSR) can be made useful here.[7]

II. Artificial Intelligence in Criminal Proceedings

1. Algorithm

An algorithm is generally described as "a finite sequence of well-defined, computer- implementable instructions, typically for solving a class of problems or performing a computation. Algorithms are always unique and are used as specifications for performing computation, data processing, automated reasoning, and other tasks."[8]

2. Selection of Incoming Data

Crucial to our analysis is that algorithms depend not only very much on proper design, but even more on the input data and the selection and selection criteria of that data. Secondly, it is important to keep in mind that the algorithm works with distinctive data/information and is very much oriented towards the largest number of funds (as the criteria match). As Richard Berk wrote, there have been a number of important recent developments that go far beyond the state of the art, even 5 years ago. Last but not least, context must be considered. For example, algorithm-based analytic tools will not meet the needs of decision makers if the context on the one hand and the consequences of prediction errors on the other are ignored. Particularly in the area of offender attribution, there is a great danger of falsification by biased selection criteria – often already due to "criminalistic experience" in the sense that

4 Excerpt *Klaas*, Internal Investigations and Information Sharing: The Coherence of Data Protection, Procedural Rights and Procedural Principles, 2020.

5 *Momsen/Willumat*, KriPoZ 2019, 323-337.

6 See the inaugural CIHR white paper on digital citizenship: "Towards a Concept of Digital Citizenship: AI and the Universal Declaration of Human Rights", https://jjccihr.medium.com/towards-a-concept-of-digital-citizenship-ai-and-the-universal-declaration-of-human-rights-e16f18492e2 (zuletzt abgerufen am 6.1.2023).

7 *Ambos/Momsen*, in: Ambos/Momsen (Hrsg.), Criminal Law Forum Special Edition: Human Rights Compliance and Corporate Criminal Liability, 2018; *Momsen/Schwarze*, in: Ambos/Momsen (Hrsg.), Criminal Law Forum Special Edition: Human Rights Compliance and Corporate Criminal Liability; *Momsen/Willumat*, KriPoZ 2019, 323-337; *Momsen/Rennert*, KriPoZ 2020, 160-172.

8 The Definitive Glossary of Higher Mathematical Jargon, online abrufbar unter: https://mathvault.ca/math-glossary/#algo (zuletzt abgerufen am 6.1.2023).

certain factors such as residential area, income, origin, for example, are placed in an objectively non-existent dependency relationship with the commission of crimes. Therefore, heuristic biases must be checked throughout the process.[9] Starting with the design of the algorithm, taking into account the learning and input data, and last but not least the interpreting humans at the input, output and decision level. The same applies to recent controversies about racial bias in criminal justice prediction tools.[10]

3. Prognostic Decisions and Parole Decisions

One of the most problematic but relevant areas for the use of algorithms is probation predictions. They are of particular interest because the risk numbers closely resemble predictive policing. Richard Berk (2019) showed that the accuracy of parole predictions is difficult to determine, despite their widespread use. Even less is known about the accuracy of similar predictions in other criminal justice decision-making situations.[11] According to Berk, the most obvious obstacle is that too few forecasting procedures have been empirically evaluated. Even when serious evaluations are reported, it appears that they are often poorly conducted. For example, the same data are used to create and test a forecasting procedure. Such "double-dipping" has long been known to make forecasts appear more accurate than they actually are. As a result, the accuracy of criminal justice forecasts is still considered to be largely unknown and inconsistent.[12]

But efficiency seems to have increased. Recent advances in statistics and computer science are setting new standards for predictive accuracy, at least in principle.[13] Apparently, the effort will be enhanced when these tools are combined with the increasing availability of very large data sets with hundreds of potential predictors. Regardless of how past criminal justice performance has risked predictions, it may now be possible to make much better predictions. This, however, again involves the relationship to privacy discussed above. This is all the more true when, outside of parole predictions, data are at stake under the protection of the presumption of innocence.[14]

4. Perception and Interpretation Errors

Another general problem, which seems likely to be exacerbated by the use of algorithms or AI, is the impact of only seemingly objective presumptions, biases. Like cognitive dissonance,[15] these confounding factors have always challenged routine criminal risk predictions because of a lack of transparency and fairness. Especially in light of the selection of learning and control data used to feed AI-based systems, the causes of bias in the criminal justice system need to be examined more closely. Against the backdrop of practices, particularly in the United States over the past decade, such as mass arrests, racial profiling, suspicionless stop and frisk, and, of course, the use of lethal force by police, the question arises as to how the use of AI may play out in this regard. Recent research shows that unanalyzed or inadequately analyzed risk predictions can make matters worse.[16] Obviously, the risk of misinterpretation increases the more the process is automated and the less it is understood. Structurally, related problems are also evident in law enforcement in Germany.

III. Specific Impact of „Big Data" Processed and Analyzed with Algorithms/AI

1. Added Value of the Analysis of "Big Data" in Investigative Contexts

The basic conceptual understanding of "Big Data" contains four points: It is large amounts of data (volume) that have a different format (variety), move quickly (velocity), and have a pattern through which value can be created from the data (value). This added value can take many different forms. Currently, it is even more evident in its use by companies, especially in the IT sector. When talking about private companies as developers, producers, sellers or service providers of databases, tools or infrastructure, a currently underestimated but crucial new player enters both the police and law enforcement context. Crucial, even in the law enforcement context, is the ability to link different data resources to add value. Insofar as linking preventive police and criminal justice data is concerned, this can promote the aforementioned merging of the two areas.

2. Problems of Participation by Private Actors in Algorithm-Based Investigative Work

To the extent that private actors are now processing the data and, if necessary, making appropriate linkages, the following aspects are relevant: (1) Private companies are neither designed nor intended to improve the common good. In a capitalist economic system, they are designed to increase the profit of their shareholders. (2) They are not traditional addressees of fundamental or civil rights, and they have no formal duty to protect these rights – at least from a traditional perspective of these rights as defensive rights against state intervention. (3) In companies, lawsuits are often caused by a multitude of individual subdecisions. Criminal imputability threatens to fade away within a structure of "organized irresponsibility." (4) Private companies own the architecture and design of algorithmic tools as intellectual property. The protection of

9 *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 6, 7.

10 *Courtland*, Nature 558 (2018), 357-360; *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 7.

11 *Skeem/Monahan*, Current Directions in Psychological Science (2011) 21(1), 38-42; *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 7.

12 *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 7.

13 *Berk*, in: Hagan/Schepple/Tyler (Hrsg.), Annual Review of Law and Social Science (2008) 4, 173-192.

14 *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 7; *Kahneman/Slovic/Tversky*, Judgement under Uncertainty: Heuristics and Biases, 1982, S. 3-22 (overview).

15 *Momsen/Washington*, in: FS Eisenberg, 2019, S. 453 ff; *Richardson/Goff,* Self-Defense and the Suspicion Heuristic, Iowa Law Review, Vol. 98 (2012), 293 ff.; *Kahnemann*, Thinking, Fast and Slow, 2012.

16 *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 7-8.

this property allows them, to a large extent, not to disclose it to the public or to users.

Most "predictive policing" tools are not primarily used[17] to predict specific crimes by their original algorithm-design, but to focus on specific groups of addressees, customers, or voters. For many companies, the specific advantage lies primarily in collecting and evaluating as much data as possible about consumer behavior in order to promote an advertising message or service tailored to the individual. The necessary data can be collected by the company itself by storing, evaluating and later analyzing the ordering behavior of individual users. On the other hand, the services of other companies that have easy access to large amounts of personal data, such as Google and Facebook, can also be used for this purpose. Since the latter companies offer most of the services to end consumers without any financial consideration, the business model is essentially based on the profitable marketing of the collected data.

Accordingly, the added value for companies is created not only by the mere collection of raw data, but also by the subsequent processing and sale of the knowledge gained. The processing is done by rule- or example-based algorithms, i.e., the behavior of consumers on the Internet is analyzed by an algorithm that is either based on certain rules defined by the creator of the algorithm or recognizes a certain buying behavior based on empirical values. These preconditions can, in the worst case, lead to various confounding variables if the analysis tool designed for a different purpose is used in the field of law enforcement, where a large number of economically interesting criteria from the social sphere are not likely to be included in the analysis due to the presumption of innocence, among other things.

3. Big Data in Criminal Investigations

Algorithm-supported data analyses are used outside the field of criminal law in the legal sector, for example to reduce the effort involved in due diligence processes as part of corporate acquisitions. For this purpose, algorithms are used to search in digital documents and so-called red flags are recognized on the basis of empirical values. The microanalysis of voter data to predict and influence their future voting behavior (as in the case of Cambridge Analytica)[18] presents possibly the best transitional use case to full “predictive policing.” Obviously, this technique has good theoretical applications in law enforcement as well. Dataset analysis is shaping contemporary policing: sample-based algorithms are admittedly currently used primarily to predict the likelihood of committing crimes according to certain spatial criteria. Add DNA analysis, especially the newly added "advanced DNA analysis" (so-called "forensic DNA phenotyping"), and a smooth transition to law enforcement becomes apparent. The predictive criteria are equally suitable for offender profiling. At the same time, the redundancies involved are obvious: if the same data were evaluated using the same methods and criteria, there would be a certain danger of generating suspects for future offenses for oneself and also automatically prosecuting them as potential offenders later on, a form of "self-fulfilling prophecy."[19]

4. Flawing the Presumption of Innocence – German Constitutional Questions

The AI-supported analysis of "Big Data" makes it possible, for example, to monitor people in real time and theoretically almost completely, for example by using video cameras and smartphone data with communication and location information. But human-generated data can provide information about more than just the present or the past. By correlating past behavior with statistical probabilities, Big Data can (presumably) predict future behavior or estimate the dangerousness of places, e.g., probability of home burglaries in certain areas. Thus, two processes become relevant for criminal law consideration: first, the collection of data and, second, the use of the collected data for criminal proceedings.

Police work can already come into conflict with the presumption of innocence when collecting data. A restriction of the presumption of innocence is initially present if proportionality is not maintained. This can occur in particular when data on unsuspected persons is accessed too extensively and too intensively in order to obtain a quantity of data capable of analysis. "Big Data" is, however, by definition only meaningful if as much data as possible is available. Therefore, measures taken on the basis of "Big Data" must be reviewed for their proportionality. This applies to a greater extent in the case of AI- supported data analysis, since data of suspected and unsuspected citizens are used indiscriminately.[20] To this extent, the presumption of innocence is indeed modified at this level, since data collection and analysis are already genuine investigative measures that address groups of people formed according to certain criteria – independent of a specific suspicion of a crime and thus independent of its basis, the principle of individual guilt.

In the case of measures taken in the context of criminal prosecution or averting danger, a conflict of fundamental rights with, for example, the secrecy of telecommunications standardized in Article 10 of the German Constitution or with the fundamental right to freedom, Article 2 of the Basic Law, is conceivable. According to the three-sphere theory developed by the Federal Constitutional Court, interference with the closest sphere of the person, the intimate sphere of private life, is not permissible. In the opinion of the Federal Constitutional Court, this also includes the prohibition of total surveillance, as this would

17 In Germany, too, the first steps in this direction have been taken, for example by companies such as Palantir.

18 *Wisser*, American Criminal Law Review Vol. 56, S. 1811 ff.; *Wylie,* Mindf*ck: Cambridge Analytica and the Plot to Break America: Inside Cambridge Analytica's Plot to Break the World, 2019.

19 *Završnik*, Big Data, Crime and Social Control (Routledge Frontiers of Criminal Justice), 2018.

20 *Završnik*, Big Data, Crime and Social Control (Routledge Frontiers of Criminal Justice), 2018.

per se constitute an encroachment on the core area of the right of personality.[21] On the other hand, the police have unrestricted access to a wide variety of areas, such as openly accessible social media activities. This creates an intermediate area in which core area information is comparatively easily accessible and may even be made available – albeit with a different purpose - by those authorized to access it, at least to a limited group of people. Within the intermediate area, the requirement of proportionality applies:

"The reason for the absolute protection of a core area of personality development lies in the fact that people are given the opportunity to deal with their own ego in a final space of retreat without having to fear that the state authorities will monitor this. Thoughts are basically free, because thinking is a condition of existence for human beings.[22] These thoughts lack in themselves the community reference, which lies outside the core area of personality development."[23] For those readers who are not familiar with German constitutional law, the so-called diary decision of the Federal Constitutional Court is of interest. The court had to decide whether and which records in a hidden diary belonged to the most intimate sphere and therefore remained closed to the police taking of evidence. As a result, a diary was generally considered to belong to the most intimate sphere. However, the court made an exception: if it contains thoughts or sufficiently concrete fantasies about the intention to commit serious crimes, these records belong to the intermediate sphere and are accessible to police investigations.[24]

It follows that no absolute protection applies to the police when investigating criminal acts. The same also applies to the planning of future criminal acts. The weighing of interests therefore takes place both in the area of police law and in criminal proceedings according to the criterion of proportionality.[25] The Federal Constitutional Court has developed additional fundamental rights for IT-based measures and the handling of data. For example, in the so-called census ruling, in which it derives from Art. 1 in conjunction with.

Art. 2 GWB, the right to informational self-determination is derived: "A social order and a legal order that make this possible would not be compatible with the right to informational self-determination, in which citizens can no longer know who knows what, when and on what occasion about them. Those who are uncertain whether deviant behavior will be noticed at any time and permanently stored, used or passed on as information will try not to be noticed by such behavior. This would not only impair the development opportunities of the individual, but also the common good, because self-determination is an elementary functional condition of a free democratic community based on the ability of its citizens to act and participate. Further protection follows from this.

Under the modern conditions of data processing, the free development of personality requires the protection of individuals against the unrestricted collection, storage, use and disclosure of their personal data.[26] 2006, the Federal Constitutional Court recently ruled with regard to the proportionality of investigative measures that, for example, a dragnet is only permissible if high-ranking legal interests are affected, because otherwise the encroachment on the right to informational self-determination of an indeterminate number of citizens cannot be justified.[27]

5. Changing Concept and Function of Suspicion, the Presumption of Innocence and the Shifted Position of the Defendant in German Criminal Proceedings

As discussed above, the application of the same criteria for individualizing probable dangers and threats as in the prosecution of criminals can lead to the presumption of innocence in criminal investigations being overridden by an attribution of presumptive suspicion to members of particular groups or individuals. In particular, this leads to a problem of proportionality between suspicion and investigation. Insofar as suspicion is replaced by a statistical parameter, it is defined merely in terms of abstract probability parameters. Within this realm, because completely unsuspicious persons are inevitably also covered by the data analysis, the presumption of innocence becomes a kind of presumption of guilt or, to be more precise, the prognosis that takes its place serves, under certain circumstances, to legitimize otherwise inadmissible encroachments on fundamental rights. Thorburn has stated in this regard that profiling schemes lead to the most serious normative challenges of all. These schematized assignments of an accused status or suspicion not only involve the collection of masses of data, much of it without the consent of the person being monitored, but also jeopardize another central concern of criminal law: the presumption of innocence.

Although the creation of probabilistically based categories of suspicion does not in itself formally abrogate the presumption of innocence in procedural law terms, it does threaten the normative foundations on which that presumption rests. The presumption of innocence is fundamentally a normative obligation of the criminal justice system to treat everyone as a free agent, even if they behave in ways that are wholly inconsistent with their prior behavior. Regardless of what actuarial tables tell us about a particular person, the criminal justice system is obligated to treat each person as if we knew almost nothing about his or her past-and to require the prosecution to prove beyond a reasonable doubt that he or she actually

---

21 BVerfGE 65, 1 ff., 41 ff. As well as clarification of the basic concept of the right of personality as a constitutional fundamental right.

22 Cf. the "dissenting votes" of *Mahrenholz, Böckenförde, Graßhof* and *Franßen*, who argue for a constitutional violation, in *BVerfG*, Urt. v. 14.9.1989 – 2 BvR 1062/87 = BVerfGE 80, 367 (381).

23 BVerfGE 80, 367. In the event of a tie vote, no violation of the Constitution was established.

24 *BGH*, Urt. v. 22.12.2011 – 2 StR 509/10.

25 *Frase/Momsen/Washington/O'Malley*, in: Ambos et al. (Hrsg.), Core Issues in Criminal Law and Criminal Justice, Volume 1, 2020, S. 213 ff.

26 BVerfGE 65, 1 ff., 43.

27 BVerfGE 115, 120 ff.

committed the specific acts of which he or she is accused.[28] Criteria such as association with dubious characters, membership in suspicious organizations, and even prior convictions are generally not taken into account in the assessment of suspicion - because in each case they tend to color our judgment of the defendant's guilt, even though, strictly speaking, they are not evidence of his guilt in terms of the specific crime being investigated.[29]

6. Towards a Probabilistic Concept of Suspicion

Conflicts with the presumption of innocence are mainly due to the measurement of the data in a statistical probability. Since, in the context of security law, the police intervene before the crime is committed, a "false positive detection", although a statistical exception, is practically a constant occurrence, especially in so-called predictive policing, e.g. in the context of "dangerous person detection". Innocent people can thus become the target of preventive police action, even if they never intended to commit a crime. According to recent studies, this can lead to certain population groups being viewed as suspicious more often than others, i.e. discriminated against.

The use of AI and algorithms in data analysis can even reinforce such discriminatory tendencies, whom the selection of learning and comparison data already suffers from corresponding biases. Thus, with respect to DNA analysis, Pfaffelhuber, Lipphardt et. al. demonstrated in an empirical study of the influence of ancestry-related marker selection that algorithmic computation can produce significant amplification effects with respect to these erroneous prior assumptions. They used feature selection theory from statistical learning to obtain AIMsets for BGA inference. Using simulations, they were able to show that this learning procedure works in several cases and outperforms ad hoc methods based on statistics such as FST or informativeness for selecting AIMs.

By applying their method to data from the 1000 Genomes Project, they identified an AIMset of 12 SNPs that, like other published AIMsets, yields a vanishingly small misclassification error at the continental level. In fact, cross-validation shows that there are a variety of sets with comparable performance to the optimal AIMset. On a subcontinental scale, we then find a set of 55 SNPs to distinguish the five European populations. The misclassification error is reduced by a factor of two compared to the published AIMsets, but is still *30%* and therefore too large to be useful for forensic applications.[30] Law enforcement actions are unlikely to be legitimate on this basis. The problem becomes obvious: the result depends very much on the purity of the learning data. Literally, the more unique the learning data, the better algorithms seem to work. Black is good, white too, but brown is not. Put simply and provocatively, since the algorithm seems to learn better with black and white, it is very likely to later work with criteria that previously matched white and (or more likely) black subjects. The attribution of propensity to danger or suspicion may then quickly prove to be merely self-referential and highly biased.

7. Merging with Security Law and Dismantling Defendants' Rights

As explained, the use of Big Data and AI seems to dissolve the traditional concepts and categories of criminal law of guilt. As such, this is not a priori an argument against the use of modern techniques in criminal prosecution. What remains decisive is that certain threshold values must be reached in criminal investigations in order to legitimize investigative measures that sometimes intensively interfere with the human and civil rights of defendants. In this respect, it makes sense to continue to use the differentiated and elaborate concepts of suspicion and the accused and to transfer their core content to modern investigative methods.

As a control consideration, one might imagine that at the end of the data analysis, all criteria can be matched for 10 individuals. What level of suspicion is required, and what level of evidence is required? Is it sufficient if it relates to a group of people, or must it be directed against a specific individual?[31] What is the effect if, although it is indisputable that only one person could have committed the act himself, after the analysis all of them are considered as perpetrators with equal probability? Does suspicion in the criminal law sense arise in this way and can a person become an accused on the basis of such analyses? Will they all be given the necessary teachings and allowed to exercise a full right to silence? Or would it not be much more likely that they would be seen only as suspects to be investigated as if they were perpetrators, but not as defendants with legal authority to actively weaken the case, not only by gathering exculpatory evidence? If so, the likelihood analysis would directly result in the actual persons being considered as defendants, being deprived of key rights.[32]

In order to answer the question of whether the concepts of the methods are so similar that the danger prognosis automatically leads, or at least there is a high probability, that the person brought into focus in the context of the danger prevention methods can also be regarded as a suspect of concrete criminal acts at the same time only on the basis of these prognosis methods, it must be compared to what extent the various methods from both areas of law are consistent and whether the methods can be applied independently of each other at all. Studies on this are still pending.

---

28 Vgl. dazu auch *Berk*, Machine Learning Risk Assessments in Criminal Justice Settings, 2019, S. 116 ff., 128 ff.

29 *Thorburn*, in: Sullivan/Dennis, Seeking Security. Pre-Empting the Commission of Criminal Harms, 2012, S. 32.

30 *Pfaffelhuber/Grundner-Culemann/Lipphardt/Baumdicker*, Forensic Science International: Genetics, Volume 46, May 2020, online abrufbar unter: https://doi.org/10.1016/j.fsigen.2020.102259 (zuletzt abgerufen am 5.1.2023); *Bown*, The criminal justice system as a problem in binary classification, 2018, S. 9, 10.

31 *Momsen/Rennert*, KriPoZ 2020, 160 ff.

32 *Momsen/Rennert*, KriPoZ 2020, 160 ff.

8. Algorithmic Fairness

Of central importance is whether AI in law enforcement can translate the category of fairness.[33] Elements of such "algorithmic fairness" would be the development of formal fairness criteria and accuracy measures, as well as standards for accuracy, transparency, and validity of analyses.[34] It must also be examined whether the use of these instruments leads to area-specific risks of discrimination and bias. For example, since in criminal cases neither the intermediary procedure nor the main trial structure ensure effective review of possible investigative errors, provided that in many cases there is an overconfidence in the impartiality of investigators,[35] the risks could theoretically be reduced when AI is used in the context of retrospectively analyzing criminal justice. If, on the other hand, the same databases and algorithms are used as for threat prediction, amplification effects could occur on the contrary.

Thus, discriminatory sampling of learning data in the predictive policing phase will introduce biases. These biases are adjusted and extrapolated (as self-fulfilling prophecies or "slippery slopes"[36]) in a police-led criminal investigation. The decisive factor may be whether the original algorithm-based evaluation procedures are (or can be) verified or whether it remains unclear who designed the learning data with which standards or who determined the evaluation criteria with which individual errors.

Since fairness itself cannot be defined in a binary structure, but in only ethical or philosophical ways,[37] developers of appropriate analytic tools need to think in a utilitarian mode that goes beyond abstract concepts of fairness and focuses on the clearly defined outcome goals of algorithms designed for specific contexts. This requires a case-based approach - identifying a set of test problems against which algorithmic outcomes can be evaluated - in context.[38] Russel, for example, emphasized that algorithms and AI systems are "authored texts"[39], "written by individuals and carrying with them the implicit values, biases, and ideologies of their authors."

9. Paradigm of AI in Law Enforcement

AI can be characterized as "the ability of a system to correctly interpret external data, learn from that data, and use that learning to achieve specific goals and tasks through flexible adaptation."[40] Partially defined somewhat more broadly, AI deal with intelligent behavior in artifacts. Intelligent behavior, in turn, is said to involve perception, reasoning, learning, communication, and action in complex environments.[41] Or they may be defined as "intelligent agents," i.e., any device that perceives its environment and performs actions that maximize its chance of successfully achieving its goals."[42] The term "rational agent" is also used.[43] The definitions of algorithms and AI thus differ in one key respect, the attribution as an agent.

In this function as a decision agent, an AI appears to make decisions instead of a human individual. This conceptually allows for an AI to make decisions over other humans or by interfering (assisting) in human decision-making processes. The "loop model" of decision making can be used to differentiate whether AIs integrate humans into the decision process, keep them out of the process, or even subordinate them to the *decision*.[44] It is clear that whoever is outside the process cannot be a subject of decision making. The objects of such decisions do not see through the process and can be relatively easily disinformed or misdirected by fake news and information.[45] Obviously, this opens up a wide space to discriminate against disadvantaged populations or individuals. Whereby the disadvantage does not necessarily have to be systemic, but can already arise from the opacity of the decision design of the AI.

IV. Black Box Effects: AI and Law Enforcement

How will AI be used? Once we reach the point where decisions are made by non-human decision makers (a.k.a. machines), we could focus on the design of the algorithm and the selection and compilation of learning data and baseline information. Given that we have not yet reached this point in criminal justice, human decision makers are

33 *Barabas,* Beyond Bias – Re-imagining the Terms of Ethical AI in Criminal Law, 2019, S. 19-20.

34 *Barabas,* Beyond Bias – Re-imagining the Terms of Ethical AI in Criminal Law, 2019, S. 1.

35 *Momsen/Washington*, in: FS Eisenberg, S. 453 f.

36 *Momsen/Weichert*, freispruch 13/2018, 37 f.

37 It is agreed, since there are some decisions to be made by human beings, which are at least the core of legal thinking. These choices cannot be expressed in binary terms because they are deliberative. See *Simoiu/Corbett-Davies/Goel,* The Annals of Applied Statistics, 11:3, 1193-1216, cited by *Russel/Akkiraju*, Put AI in the Human Loop, 12-2019, HICSS Workshop-AI-and-Bias, S. 7.

38 *Russell/Akkiraju*, Put AI in the Human Loop, 12-2019, HICSS Workshop-AI-and-Bias, S. 7.

39 *Caliskan-Islam/Harang/Liu/Narayanan/Voss/Yamaguchi/Greenstadt,* De-anonymizing programmers via code stylometry, 24th Usenix Security Symposium, USENIX, cited by *Russell/Akkiraju*, Put AI in the Human Loop, 12 -2019, HICSS-Workshop- AI-and-Bias, S. 7.

40 *Poole/Mackworth/Goebel,* Computational Intelligence und Knowledge, 1998, S. 1, online abrufbar unter: http://people.cs.ubc.ca/~poole/ci/ch1.pdf (zuletzt abgerufen am 5.1.2023).

41 *Malouf*, Artificial Intelligence: An Introduction, 2017, online abrufbar unter: http://people.cs.georgetown.edu/~maloof/cosc270.f17/cosc270-intro-handout.pdf (zuletzt abgerufen am 5.1.2023).

42 *Poole/Mackworth/Goebel,* Computational Intelligence und Knowledge, 1998, S. 1, online abrufbar unter: http://people.cs.ubc.ca/~poole/ci/ch1.pdf (zuletzt abgerufen am 5.1.2023).

43 *Russell/Norvig,* Artificial Intelligence: A Modern Approach (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, 2003, online abrufbar unter: http://aima.cs.berkeley.edu (zuletzt abgerufen am 5.1.2023).

44 Lecture by *Hin-Yan Liu* (Associate Professor in the Faculty of Law at the University of Copenhagen and director of the Center for International Law, Conflict and Crisis) at John Jay College of Criminal Justice (CUNY), Center for Criminal Justice Ethics and Center for International Human Rights, New York, v. 6.3.2020, in: Human Rights, Digital Society and the Law: A Research Companion, Chapter: 5, S. 75-86, online abrufbar unter: https://www.researchgate.net/publication/326991445_The_Digital_Disruption_of_Human_Rights_Foundations (zuletzt abgerufen am 6.1.2023).

45 *Wylie*, Mindf*ck: Inside Cambridge Analytica's Plot to Break the World, 2019, provides a political ethics engineering perspective on right-wing authoritarianism, disinformation, agency and self-determination, and challenges to democracy and elections.

an important part of decision making.[46] Therefore, the interaction between human AI preparation, non- human decision preparation (AI), and human decision making provides the framework for analysis;[47] there is therefore no precise differentiation between "artificial intelligence" and other data-driven decision-making regimes in criminal law. The discourse on artificial intelligence in, for example, the (U.S.) criminal justice system encompasses a hodgepodge of computational technologies ranging from decades-old practices to machine learning algorithms that were not possible before the era of "Big Data."[48] Broadly speaking, these technologies are a mix of new and old statistical methods that measure the strength of associations between a set of data points and an outcome. These techniques are correlational at their core - their results are typically in the form of probabilistic distributions that are read as forecasts or predictions of future events. In criminal justice, the data used to build these statistical models is typically administrative information collected by local police departments and court administrators and then interpreted using probabilistic computational methods.[49]

In many cases, it seems difficult for legal users to understand, calculate, or even reconstruct the operations performed by the AI.[50] In view of this, the AI is given some specific authority in decision making. As long as there is no user who would actually be able to monitor the decisions made by the AI, the decision-making process cannot be efficiently challenged. In legal terms: there is no viable argument and factual basis to challenge this decision. This problem becomes apparent, for example, when an attempt is made to file appropriate motions for evidence. Since these are only intended to create the basis for uncovering errors, no concrete evidentiary error can yet be named. The request can be treated as a mere request for evidence or, after the 2018 reform of Section 244 of the German Code of Criminal Procedure, almost certainly be rejected as a so-called "request for evidence in the blue" without further justification.

Another problem that is very common in legal classification by AI also occurs in our context: the fuzziness of causality or the resolution of attribution. Insofar as the evaluation process of AI cannot be traced, there is, as it were, a black box in the middle of the criminal attribution process and the criminal attribution process. The use of AI to investigate suspicion can sever the chain of causation or attribution of responsible human action. In the substantive area of criminal product liability law, it is well known that it is often impossible to hold a specific person on the manufacturer's or seller's side responsible for accidents involving automatic vehicles because it is not possible to trace why the decision- making process was flawed. In some cases, the decision that violates legal interests is not, strictly speaking, faulty, but rather logical for an AI based on programmed assumptions, such as the greatest benefit for the greatest number of people. The ethical problem addressed in the context of fairness. This has its counterpart in the attribution of suspicion: as long as the learning and basic data are not known and their selection criteria are not disclosed, as long as the analysis and decision-making process cannot be deciphered, the attribution of the accused position solely on the basis of an AI-based analysis process is a violation of the presumption of innocence.

Moreover, since doubtful decisions cannot be mapped in a binary structure, recidivism predictions are also hardly possible in a constitutional manner. Accordingly, for the time being, AI can only be used in the area of criminal prosecution for cross-checking human decision-making processes. The complete relocation of criminal law evaluation decisions or the replacement of human decisions by AI-based decision-making regimes would be most likely unconstitutional in this respect, in the German context. However, the first step has to become aware of the potential biases and risks of the algorithm in general and the tool used in the specific case in particular. Therefore, legal and technical tools have to be designed to enable the defense in particular to proof the tools used for the decision-making of the courts.

---

46 *Foucault*, Power/Knowledge: Selected Interviews and other Writings, 1972-1977 (1980); *Murakawa,* The First Civil Right: How Liberals Built Prison America, 2014; *Muhammad*, The Condemnation of Blackness, 2011; *Platt,* "Street" Crime – A View from the Left, Soc. Jistice J. Crime Confl. World Order 26 (1978); more: *Barabas,* Beyond Bias: Re-imagining the Terms of "Ethical AI" in Criminal Law, 2019*,* S. 4.

47 *Barabas,* Beyond Bias: Re-imagining the Terms of "Ethical AI" in Criminal Law, 2019*,* S. 4.

48 *Brayne*, "Big Data Surveillance: The Case of Policing," American Sociological Review 82, no. 5 (2017), 977–1008.

49 *Ferguson*, The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement, NYU Press, 2017.

50 *Coleman*, A Human Algorithm, 201, S. XVII-XXII; *Bostrom*, Are you living in a computer simulation? Philosophical Quarterly (2003) Vol. 53, No. 211, S. 243-255, online abrufbar unter: https://www.simulation-argument.com/simulation.pdf (zuletzt abgerufen am 5.1.2023); *Pfaffelhuber/Grundner-Culemann/Lipphardt/Baumdicker*, Forensic Science International: Genetics, Volume 46, May 2020, 102259, online abrufbar unter: https://doi.org/10.1016/j.fsigen.2020.102259 (zuletzt abgerufen am 5.1.2023).