

Zeitschrift: Kriminologie / Schweizerische Arbeitsgruppe für Kriminologie SAK = Criminologie / Groupe Suisse de Criminologie GSC = Criminologia / Gruppo Svizzero di Criminologia GSC

Herausgeber: Schweizerische Arbeitsgruppe für Kriminologie

Band: 38 (2021)

Artikel: Algorithmen : zwischen blindem Vertrauen und panischer Angst

Autor: Arnold, Jörg

DOI: <https://doi.org/10.5169/seals-1051609>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 02.02.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Algorithmen: Zwischen blindem Vertrauen und panischer Angst

Jörg Arnold*

Inhaltsverzeichnis

Zusammenfassung	119
Résumé	120
1. Ausgangslage	120
1.1 Rechtliche Grundlagen	121
1.1.1 Beweismittel	121
1.1.2 Aufgaben der Polizei	122
1.2 Methodik und Logik	122
1.3 Das Problem der Zeitumkehr	123
2. Algorithmen	123
2.1 Befundbewertung und Bayesian Approach	123
2.2 Weitere Beispiele von Algorithmen	125
2.2.1 Der Aussergewöhnliche Todesfall (AgT)	126
2.2.2 Unfallrekonstruktion	127
2.3 Zwischenfazit: Umgang mit Algorithmen	129
2.4 Lernfähige Algorithmen und prädiktive Algorithmen	130
2.4.1 Einfache Algorithmen: beschränkte Aussagekraft	131
2.4.2 Komplexe Algorithmen: beschränkte Nachvollziehbarkeit	132
2.4.3 Prädiktive Algorithmen: beschränkte Zuverlässigkeit	133
3. Fazit	134
3.1 Forderungen im Umgang mit Algorithmen	134
3.2 Beispiel hochautomatisierte Fahrzeugsysteme	135
3.3 Schlussfolgerungen und Konsequenzen	135

Zusammenfassung

Wer kennt ihn nicht, den Spruch aus den Digesten: «*iudex non calculat*»? Das Verhältnis von JuristInnen zu *Big Data*, *Artificial Intelligence*, *Cybercrime* und Algorithmen schwankt zwischen sklavischem Gefolge und kategorischer Ablehnung. Im Beitrag werden Möglichkeiten und Grenzen von verschiedenen Algorithmen aufgezeigt. Für JuristInnen müssen nicht alle

* Stellvertretender Chef des Forensischen Instituts Zürich (FOR).

Details eines Algorithmus bekannt sein, aber sie sollten im Minimum verstehen, welche Parameter den Algorithmus wie stark beeinflussen, welche Randbedingungen beachtet werden müssen und welche Toleranzen zu berücksichtigen resp. zu akzeptieren sind. Einzelne Beispiele erlauben es, den Einsatz von Algorithmen in der Strafverfolgung zu illustrieren.

Résumé

Qui ne connaît pas le dicton du Digest: «*iudex non calculat*»? La relation entre les juristes et le *big data*, l'intelligence artificielle, la cybercriminalité et les algorithmes oscille entre suivisme servile et rejet catégorique. La contribution montre les possibilités et les limites de divers algorithmes. Les juristes n'ont pas besoin de connaître tous les détails d'un algorithme, mais ils doivent comprendre au minimum quels paramètres influencent l'algorithme et dans quelle mesure cela a lieu. Il faut qu'ils connaissent les conditions cadres qui doivent être respectées et le degré de tolérance qui doit être pris en compte ou doit être accepté. Des exemples permettent d'illustrer l'utilisation d'algorithmes dans le domaine de la poursuite pénale.

1. Ausgangslage

Im vorliegenden Beitrag¹ wird versucht, einige Grundbegriffe zum Thema «künstliche Intelligenz und prädiktive Algorithmen» zu klären. Es werden grundsätzliche Fragen gestellt und soweit möglich beantwortet. Um den Begriff des Algorithmus etwas greifbarer zu machen, werden zuerst Beispiele von klassischen Algorithmen besprochen, die *in der Vergangenheit liegende Ereignisse* beschreiben resp. verständlich machen. In einem zweiten Schritt werden in die Zukunft gerichtete, *prädiktive Algorithmen* thematisiert.

Im digitalisierten Umfeld besteht ein sehr grosses Bedürfnis, dass Begriffe gemeinsam definiert und Konzepte aus der Informationstechnologie mit den JuristInnen ausgetauscht und interdisziplinär verstanden werden. Der Gesetzgeber ist hier stark gefordert: Es muss wieder ein Weg gefunden werden, die *grundlegenden Prinzipien* gesetzlich zu regeln.²

1 Ich bedanke mich bei Thomas Ottiker, Chef FOR, und Marc Neracher, FOR, für das kritische Gegenlesen des Artikels und die wertvollen inhaltlichen Anregungen.

2 Als gutes Beispiel sei der Zweckartikel des Bundesgesetzes betreffend die elektrischen Schwach- und Starkstromanlagen (Elektrizitätsgesetz, EleG, SR 734.0, <https://www.admin.ch>) aus dem Jahr 1902 genannt:

Art. 3

¹ Der Bundesrat erlässt *Vorschriften zur Vermeidung von Gefahren und Schäden*, welche durch Stark- und Schwachstromanlagen entstehen. [...].

Die Tendenz, die sich insbesondere im Datenschutz und bei vielen Themen im hochtechnisierten Umfeld von Technik und Wissenschaft in der jüngsten Vergangenheit durchgesetzt hat, nämlich alles auf dem aktuellen Stand der Technik – dafür aber mangels grundlegenden Verständnisses – bis ins letzte Detail regeln zu wollen, ist fatal.

In der heutigen Gesellschaft gilt: Verfügbare Technologien werden gnadenlos genutzt – und allenfalls auch missbraucht. Umso wichtiger ist es, die entsprechenden Gesetze so zu schreiben, dass sie nicht laufend von der technischen Entwicklung und neuen Realitäten überholt werden.

Den Kopf in den Sand zu stecken und zu hoffen, dass der Kelch des technologischen Fortschritts an uns vorbeigeht, ist in jedem Fall falsch.

1.1 Rechtliche Grundlagen

1.1.1 Beweismittel

Zur Erinnerung – Art. 139 StPO: Das Ziel jeder Strafuntersuchung ist die Wahrheitsfindung.³ Als Grundsatz gilt, dass zur Wahrheitsfindung alle nach dem Stand von Wissenschaft und Erfahrung geeigneten Beweismittel einzusetzen sind, die rechtlich zulässig sind.

Die rechtliche Zulässigkeit bezieht sich dabei primär auf alle Arten von Zwangsmassnahmen – hier sind die Grundprinzipien des «*fair trial*» und der Verhältnismässigkeit selbstverständlich konsequent zu beachten.

Allerdings ist die rechtliche Grundlage für die Sachverhaltsfeststellung und insbesondere für die Spurensicherung und -auswertung als *Aufgabe der Polizei*⁴ in einer Muss-Formulierung in der StPO festgehalten.

Der Begriff des *Beweismittels* oder der *Spur* wird aber noch immer sehr stark als *physisch greifbar* verstanden.⁵ Sobald es sich um digitale Spuren handelt, die nur im virtuellen Raum existieren, werden Lücken in der Gesetzgebung und fehlende Begriffsdefinitionen sichtbar.

3 Schweizerische Strafprozessordnung (StPO, SR 312.0, <<https://www.admin.ch>> StPO)
Art. 139 Grundsätze (Beweiserhebung und Beweisverwertbarkeit)

¹ Die Strafbehörden setzen zur Wahrheitsfindung *alle nach dem Stand von Wissenschaft und Erfahrung geeigneten Beweismittel* ein, die rechtlich zulässig sind.

4 Art. 306 StPO Aufgaben der Polizei

¹ Die Polizei stellt ... den für eine Straftat *relevanten Sachverhalt* fest.

5 Siehe z.B. das Zürcher Polizeigesetz (PolG ZH, LS 550.1), § 35 bis § 40 in Bezug auf Durchsuchung, Sicherstellung etc. von Personen, Tieren und Gegenständen.

1.1.2 Aufgaben der Polizei

Art. 306 StPO mit dem Titel *Aufgaben der Polizei* wird in der polizeilichen Ausbildung leider oft vergessen – man kann in diesem Zusammenhang aktiv keine Fehler machen. Allerdings verpasst man dann *die* gesetzliche Grundlage für alles polizeiliche Handeln, *nachdem* sich ein strafrechtlich allenfalls relevantes Ereignis ereignet hat.

Auch muss immer kritisch hinterfragt werden, *wann* die Verhältnismässigkeit ein praktikabler Massstab ist. Ganz zu Beginn eines Strafverfahrens – im polizeilichen Ermittlungs-verfahren – taucht der Begriff des Sachverhalts resp. der Sachverhaltsfeststellung auf. Dies ist einer der ganz wenigen Artikel in der StPO, in dem der Polizei nichts verboten wird. Sie wird ohne Wenn und Aber beauftragt, etwas zu tun – nämlich den *Sachverhalt festzustellen*. Weiter sind in Art. 306 Abs. 2 StPO drei verschiedene Tätigkeiten der Polizei explizit genannt:

- a. die Spurensicherung und das Sichern und Auswerten von Beweisen;
- b. die polizeilichen Ermittlungen und Befragungen;
- c. die polizeilichen Zwangsmassnahmen wie Anhaltung, Festnahme und Fahndung. Hier sind selbstverständlich die Gesetzmässigkeit und die Verhältnismässigkeit zu beachten.

Und es stellt sich sofort die Frage: «In welcher Beziehung steht der Sachverhalt zur Wahrheit?». Ein zentrales, praktisches Problem ist, dass die Polizei in den meisten Fällen erst involviert wird, wenn «es» bereits passiert ist: Das Resultat eines Ereignisses ist also bereits konkret bzw. steht dann faktisch fest.

1.2 Methodik und Logik

Das führt zu einer ersten Erkenntnis: Der Sachverhalt ist sicher (viel) mehr, als das – im Rahmen der polizeilichen Tatbestandsaufnahme festgestellte – *Resultat eines Ereignisses*. Und nebenbei bemerkt: Über das eigentliche Resultat eines Ereignisses wird in den allerwenigsten Fällen gestritten.

Die zweite Erkenntnis ist: Zeitlich sind die Strafuntersuchungsbehörden grundsätzlich immer «zu spät» dran – dies aber nicht im Sinn eines Vorwurfs, sondern als wertneutrale, pragmatische Feststellung. Ganz praktisch gesprochen beginnt die Untersuchung – in Bezug auf die Zeitachse – immer erst, wenn sich das Ereignis bereits ereignet hat resp. das Resultat des – allenfalls strittigen – Ereignisses im Grundsatz feststeht.

Die Strafuntersuchungsbehörden müssen sich also bereits bei der Sachverhaltsfeststellung – und insbesondere bei der Wahrheitsfindung – bewusst sein, dass der *Ausgangspunkt der Untersuchung* das Resultat ist und dass das «Ergebnis der Untersuchung» zeitlich *vor* dem Ereignis liegt.

«Sachverhalt» und «Wahrheit» befinden sich auf der Zeitachse also nicht am gleichen Ort und müssen konsequent voneinander unterschieden werden. Insbesondere ist tunlichst darauf zu achten, dass Ursache und Wirkung nicht verwechselt oder vermischt werden.

1.3 Das Problem der Zeitumkehr

Die Frage nach dem Huhn oder dem Ei oder doch dem Huhn⁶ ist sowohl philosophischer Natur als auch ein grosses Thema in den Naturwissenschaften wie z.B. in der klassischen Mechanik von Newton, und Keppler. In die Umgangssprache übersetzt geht es um Ursache oder Wirkung resp. um ein Grundprinzip der Physik: *Actio = Reactio*.

Und dies führt zu einer weiteren wichtigen Erkenntnis für die Untersuchungsbehörden. Ausgehend vom Resultat – *rückwärts in der Zeit* – herauszufinden, wie dieses Resultat zustande gekommen ist, ergibt fast immer *mehrere Möglichkeiten*: Viele Wege führen nach Rom!

Hintergrund dieser Redewendung, die eigentlich lautet «Alle Wege führen nach Rom», ist eine vergoldete Säule, die Kaiser Augustus im Jahre 20 v. Chr. auf dem Forum Romanum aufstellen liess, auf der die Namen aller Hauptstädte der Provinzen des Römischen Reiches mit ihrer Entfernung zu Rom aufgeführt waren. Das Forum Romanum war damals auch der öffentliche Gerichtsplatz und gab der Forensik ihren Namen. Ein Betrachter der Säule musste also den Eindruck gehabt haben, alle Wege führten nach Rom.

2. Algorithmen

2.1 Befundbewertung und Bayesian Approach

Nun zu den Algorithmen mit einem ersten Beispiel: Der 4. Satz von Bayes ist ein mathematischer Satz aus der Wahrscheinlichkeitstheorie, der die Berechnung bedingter Wahrscheinlichkeiten beschreibt. Er ist nach dem englischen Mathematiker Thomas Bayes benannt und wird auch Bayes' Theorem genannt. Was aber heisst «bedingte Wahrscheinlichkeit» und was hat das mit einer Strafuntersuchung oder allenfalls mit Forensik zu tun? «Bedingt» heisst, dass man sich die Frage stellt, mit welcher Wahrscheinlichkeit etwas zutrifft, unter der Bedingung, dass etwas anderes zugrundegelegt wird. Darin ist einerseits eine zeitliche Abhängigkeit enthalten und andererseits entspricht das einem einfachen «*if ... then ... else ...*»-Algorithmus.

6 Siehe dazu auch J. Arnold, Wahrheitsfindung – Fiktion oder Realität?, in: J. Vuille/N. Oberholzer/M. Graf (Hrsg.), Schweizerische Arbeitsgruppe für Kriminologie (SAK), Band 33: Wahrheit, Täuschung und Lüge, Vérité, tromperie et mensonge, 1 ff., Bern 2016.

In der Forensik wird häufig diese Formel für die Berechnung des Beweiswertes resp. des Likelihood Ratios⁷ – zu Deutsch des Plausibilitätsverhältnisses – verwendet:

$$[1] \quad \frac{P(H_0|E)}{P(H_1|E)} = \frac{P(H_0)}{P(H_1)} \cdot \frac{P(E|H_0)}{P(E|H_1)}$$

Dabei geht es – vor dem Hintergrund von mindestens zwei gegensätzlichen Hypothesen H_0 resp. H_1 – um die Abschätzung, welches der vorgefundenen Spurenbilder – resp. welche Befunde – unter der einen oder der anderen Hypothese besser erklärbar ist resp. sind.⁸ Die beiden Hypothesen sind in vielen Fällen einerseits die der Anklage – beim Beschuldigten handelt es sich um den *richtigen* Verdächtigen – und andererseits die Hypothese der Verteidigung – es war jemand anders, ein Unbekannter, die festgestellten spurenkundlichen Übereinstimmungen sind *rein zufällig*. In der Formel resp. im Algorithmus sind es jeweils einfache Verhältnisse, also einfache Brüche, die es zu verstehen gilt. Die vom Richter aufgrund seiner Fallkenntnisse erlangten (a priori) Anfangswahrscheinlichkeiten $P(H_0)$ resp. $P(H_1)$ werden durch das forensische Gutachten des Sachverständigen beeinflusst. Das Gutachten vermittelt im Rahmen der Befundbewertung ein Beweiswertverhältnis $P(E|H_0)/P(E|H_1)$, also die Richtung und die Stärke der Beeinflussung zugunsten der einen oder der anderen Hypothese.

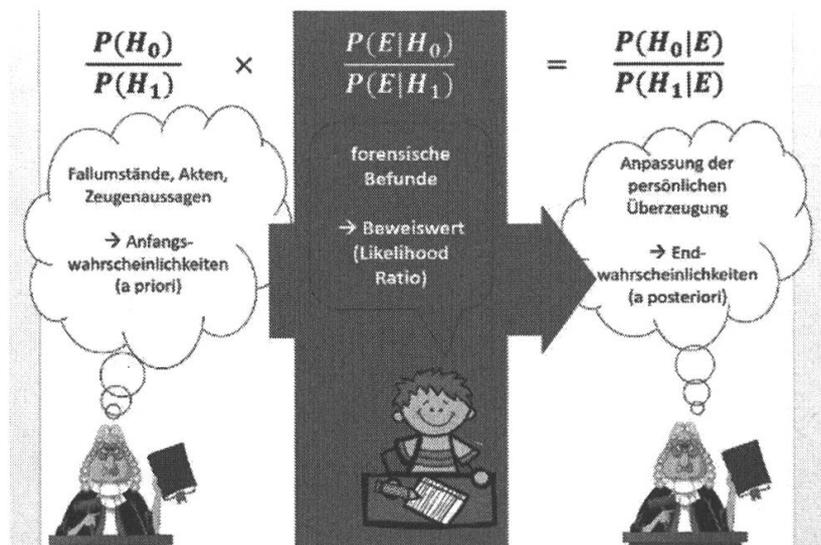


Abbildung 1: Vom «a priori»-Plausibilitätsverhältnis zum «a posteriori»-Plausibilitätsverhältnis

7 Siehe dazu u.a. C. Howson/P. Urbach, *Scientific Reasoning, The Bayesian Approach*, La Salle, Illinois 1989.

8 Ausführlich dazu siehe die *ENFSI guideline for evaluative reporting in forensic science*, öffentlich zugänglich auf der ENFSI-Homepage unter http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (zuletzt besucht am 20.9.2020).

Beides zusammen ergibt die Endwahrscheinlichkeiten (a posteriori) des Richters $P(H_0|E)$ resp. $P(H_1|E)$.

Dabei handelt es sich um ein logisches Denkmodell,⁹ das aufzeigt, wie das «aktualisieren» einer Meinung durch Berücksichtigung neuer Informationen funktioniert.

Zudem ist in [1] ersichtlich, dass es sich immer um ein Verhältnis von zwei Wahrscheinlichkeiten handelt, die jede für sich mit allenfalls grossen Unsicherheiten behaftet sind und einzeln nur sehr beschränkt Sinn machen.¹⁰

Im Gegensatz zu den ersten drei Sätzen aus der Wahrscheinlichkeitsrechnung, die sich darauf beschränken, statistische Aussagen zu machen, mit welcher Wahrscheinlichkeit eines von mehreren Ereignissen aus einer (beschränkten) Auswahl eintreffen wird, beschreiben bedingte Wahrscheinlichkeiten, wie wahrscheinlich etwas unter der Annahme (oder Voraussetzung) von etwas anderem eintritt.

Die Blickrichtung und damit indirekt der zeitliche Faktor ist in den beiden Termen des 4. Satzes von Bayes unterschiedlich. Dies entspricht auch den unterschiedlichen Blickrichtungen der Forensiker¹¹ $P(E|H_0)/P(E|H_1)$, die die *Plausibilität der Spurenbefunde unter den Hypothesen H_0 resp. H_1 bewerten*, resp. des Richters $P(H_0|E)/P(H_1|E)$, der die *Plausibilität der Hypothesen H_0 resp. H_1 unter Berücksichtigung der spurenkundlichen Befunde würdigt*.

2.2 Weitere Beispiele von Algorithmen

Im Folgenden sollen weitere, bekannte Algorithmen, wobei es vielleicht nicht offensichtlich ist, dass es sich hier um (allerdings relativ einfache) Algorithmen handelt. In der klassischen Physik sind es z.B. die Bewegungsgleichungen, die Ort und Zeit sowie allenfalls Geschwindigkeit und Beschleunigung miteinander verknüpfen. In der Unfallrekonstruktion werden diese Algorithmen in vielfältiger Weise angewandt, um den räumlich-zeitlichen Ablauf zu rekonstruieren resp. zu beschreiben, wie sich die beteiligten Objekte bewegt haben. In der Informatik arbeitet man häufig mit Entscheidungsbäumen die «if ... then ... else ...»-Verzweigungen enthalten.

9 Ausführlich dazu N. Ruckstuhl/V. Dittmann/J. Arnold, Lehrbuch Strafprozessrecht – unter Einschluss der psychiatrischen Forensik und Rechtsmedizin sowie des kriminal-technischen und naturwissenschaftlichen Gutachtens, Zürich 2011 (zit. Ruckstuhl/Dittmann/Arnold, Lehrbuch Strafprozessrecht).

10 Die Grundidee dabei ist, dass die fehlenden Informationen a priori für alle Szenarien etwa gleich wichtig sind und sich beim Bilden der «Wahrscheinlichkeitsverhältnisse» näherungsweise wegkürzen, siehe ausführlich dazu Ruckstuhl/Dittmann/Arnold., Lehrbuch Strafprozessrecht, 459 ff., N 1438 ff.

11 Ausführlich J. Arnold/T. Ottiker, Kriminalistisches Denken – systematisch, methodisch, logisch, Kriminalistik 1/2019, 50 ff.

Dies ist praktisch umsetzbar, solange die abzubildenden Prozesse nicht allzu komplex sind. Die Herausforderung z.B. bei Systemsteuerungen liegt darin, möglichst alle möglichen Betriebszustände – und allfällige Störungen derselben – vorausszusehen und die entsprechenden «richtigen» Reaktionen des Systems über solche «*if ... then ... else ...*»-Verzweigungen zu programmieren.

In der Justiz tauchen Algorithmen auf, wenn z.B. für Massendelikte sogenannte Strafmasstabellen definiert werden. Ein anderes Beispiel ist das Ablaufschema auf der Todesbescheinigung resp. sind die möglichen Entscheide, die bei einem **A**ussergewöhnlichen **T**odesfall (AgT)¹² gefällt werden müssen. Dieser Entscheidungsbaum ist auf den Formularen zur Todesbescheinigung dargestellt. In der Medizin verstecken sich hinter den verschiedenen Diagnose-tools, die von Ärzten angewendet werden, ebenfalls Algorithmen, die häufig mit Entscheidungsbäumen dargestellt werden.

Diese werden allenfalls ergänzt mit einer «*second opinion*» eines zweiten Arztes (4-Augen-Prinzip).

2.2.1 Der Aussergewöhnliche Todesfall (AgT)

Der Entscheidungsbaum auf der Todesbescheinigung ist im Grund genommen sehr einfach – trotzdem gibt es in der Praxis immer wieder Diskussionen bezüglich der zu fällenden Entscheidungen. Dabei kann es um Kompetenzen oder Verantwortlichkeiten gehen, die nicht allen Beteiligten klar sind: Entscheidet sich der (normale) Arzt für «natürlicher Todesfall», ist dieser Entscheid im Grundsatz endgültig.

Zutreffendes ankreuzen:

- Natürlicher Todesfall (*Erdbestattung oder Kremation zulässig*)
- Nicht-natürlicher Todesfall (*Unfall, Suizid, Delikt, inkl. Spätfolgen davon*)**
- Unklarer Todesfall (*nicht-natürlicher Tod möglich, unerwarteter und plötzlicher Tod ohne bekannte medizinische Vorgeschichte*) **
- Meldung an die Polizei / Institut für Rechtsmedizin erfolgt **

Abbildung 2: Entscheidungsbaum auf der Todesbescheinigung, aussergewöhnlicher Todesfall

Sobald ein Todesfall als (möglicherweise) nicht-natürlich oder unklar beurteilt wird, müssen zwingend die Polizei und durch diese ein rechtsmedizinisch ausgebildeter, sachverständiger Arzt sowie die Staatsanwaltschaft (Art. 253 StPO)¹² orientiert werden. Diese führen mit den kriminaltechnischen Spezialisten der Polizei die Legalinspektion durch und haben neben der rechtsgenü-

¹² Art. 253 Aussergewöhnliche Todesfälle.

¹ Bestehen bei einem Todesfall Anzeichen für einen unnatürlichen Tod, insbesondere für eine Straftat, oder ist die Identität des Leichnams unbekannt, so ordnet die Staatsanwaltschaft zur

genden Identitätsfeststellung auch das weitere Vorgehen wie Obduktion und toxikologische oder molekulargenetische Untersuchungen (DNA-Profilierung) anzuordnen. Der Entscheidungsbaum der Identitätsfeststellung ist auf der Todesbescheinigung vereinfacht dargestellt und verläuft parallel zu den medizinischen Abklärungen resp. Entscheidungen.

2.2.2 Unfallrekonstruktion

Ein weiterer Algorithmus ist die räumlich-zeitliche Gliederung eines (Unfall-) Ereignisses in die drei Phasen *Pre-Crash – Crash – Post-Crash*.¹³ Dies macht aus einer grossen und allenfalls sehr komplexen Aufgabenstellung drei kleinere, besser überschaubare Problemstellungen und erlaubt meist eine bessere Einschätzung des Ereignisverlaufs.¹⁴ Damit kann die Problematik der Zeitumkehr verdeutlicht werden – die Rekonstruktion startet beim Resultat. Weiter können die drei Phasen auch anders beschriftet werden – nämlich in Bezug auf objektive Anknüpfungstatsachen – und zu klärende Fragen – den Befundstatsachen. In den beiden Phasen Crash und Post-Crash werden einerseits viele Spuren verursacht – klassische Spuren, Schäden und Verletzungen (Crash). Andererseits lassen sich aus dem nachkollisionären Verhalten der beteiligten Objekte (Post-Crash) viele Informationen über die zum Unfallzeitpunkt herrschenden Bedingungen am Unfallort herleiten. Die für die rechtliche Würdigung relevanten Fragen stellen sich *nahezu ausschliesslich in der vorkollisionären («Pre-Crash»)-Phase*, wo häufig keine Spuren vorhanden sind.

Dies ist eine erste methodische Erklärung zur Relevanz dieses Schemas – um nicht zu sagen dieses Algorithmus'. Ein weiterer Grund, weshalb die Kenntnis dieses Algorithmus' sinnvoll ist: Abhängig von der Position auf der Zeitachse lässt sich abschätzen, wie gross die Unsicherheit bezüglich des realen Ablaufs ist resp. welche Toleranzen an verschiedenen Positionen auf dem Zeitstrahl zu berücksichtigen sind.

Dabei zeigt sich, dass die Unsicherheit massiv zunimmt, sobald keine objektiven Anknüpfungstatsachen resp. physischen Spuren mehr für die Ereignisrekonstruktion zur Verfügung stehen. Eine Extrapolation ist zeitlich nur wenige Sekunden vor die ersten Spuren überhaupt sinnvoll resp. mit vernünftigen Toleranzen möglich.

Klärung der Todesart oder zur Identifizierung des Leichnams eine Legalinspektion durch eine sachverständige Ärztin oder einen sachverständigen Arzt an.

² Bestehen nach der Legalinspektion keine Hinweise auf eine Straftat und steht die Identität fest, so gibt die Staatsanwaltschaft die Leiche zur Bestattung frei. [...].

13 Siehe ausführlich dazu Ruckstuhl/Dittmann/Arnold, Lehrbuch Strafprozessrecht, 483 ff., N 1518 ff.

14 Siehe ausführlich dazu J. Arnold, Unfall-Gutachten, Methodik – Möglichkeiten – Grenzen – Fallstricke – StPO, Strassenverkehr/Circulation Routière 1/2016, 41 ff.

Steht ein Event Data Recorder (EDR) zur Verfügung, der allenfalls digitale Spuren aus der «Pre-Crash»-Phase enthält, sind wir zurück beim Problem von virtuellen digitalen Spuren¹⁵ und Beweisen.

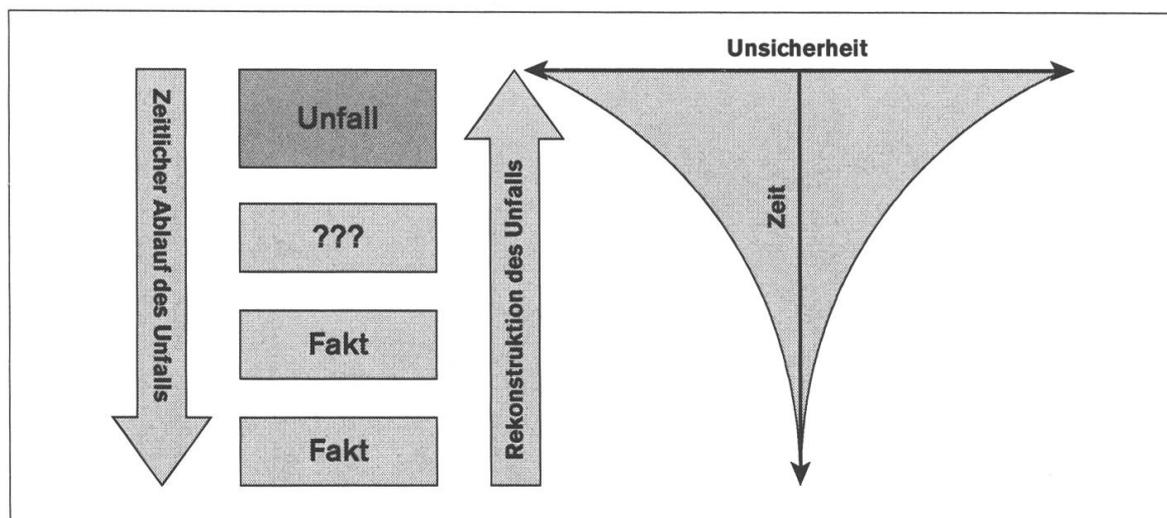


Abbildung 3: Zunahme der Unsicherheit vor den ersten Spuren – Extrapolationen müssen kritisch hinterfragt werden

Am Beispiel einer Kollision eines mit einem UDS (Unfalldatenspeicher) ausgerüsteten Personenwagens mit einem Fussgänger lässt sich dieser Algorithmus gut visualisieren und in einem zweiten Schritt ein weiterer Algorithmus – nämlich die unfallanalytische Klärung der Vermeidbarkeit des Unfalls – erläutern. Die Rahmenbedingungen, unter denen ein UDS eingesetzt werden kann oder muss, sowie die Sicherung und Auswertung dieser digitalen Spuren ist in der Verordnung über die technischen Anforderungen an Strassenfahrzeuge in Art. 102 VTS geregelt. Im Beispiel war der Personenwagen zu schnell unterwegs – nämlich mit 64 km/h statt der signalisierten 50 km/h. Die Kollision mit dem Fussgänger erfolgte während der eingeleiteten Vollbremsung mit einer Geschwindigkeit von ca. 50 km/h. Dank des UDS ist somit der gesamte räumlich-zeitliche Verlauf des Unfalls aus Sicht des Personenwagens im Detail dokumentiert. Der Fussgänger überquerte die Strasse aus Sicht des Lenkers von links nach rechts, was über die erlittenen Verletzungen (Schäden = Spuren) und die klassischen Anprallspuren (Mikrospuren, Fasern, Deformationen) belegt werden konnte. Aus dem Kollisionsraum und der Unfallendlage des Fussgängers ergab sich eine Längswurfweite von 13 m bis 17 m, was mit der Kollisionsgeschwindigkeit von ca. 50 km/h korrespondiert. Aus unfallanalytischer Sicht stellt sich die Frage nach der räumlichen Vermeidbarkeit der Kollision bezüglich der Eingangsgeschwindigkeit des Personenwagens. Damit

15 Siehe ausführlich dazu J. Arnold, Digitale Spuren im Strassenverkehr – die Zukunft hat begonnen!, in: R. Schaffhauser (Hrsg.), Jahrbuch zum Strassenverkehrsrecht 2017, Zürich/St. Gallen, 369 ff.

kann aus rechtlicher Sicht beurteilt werden, mit welchen Vorwürfen der Lenker konfrontiert wird.

Der Algorithmus für die Vermeidbarkeit der Kollision muss zwingend vom im konkreten Unfallgeschehen rekonstruierten Reaktionspunkt ausgehen – nur dann bleibt die Kausalität gewahrt. In den daraus zu berechnenden hypothetischen Varianten darf immer nur exakt ein Parameter – also z.B. die Eingangsgeschwindigkeit – variiert werden. Da die Geschwindigkeitsabhängigkeit der Vermeidbarkeit hochgradig nicht-linear ist, lässt sich die entsprechende Grenzggeschwindigkeit nicht abschätzen, sondern nur mit dem zugrundeliegenden physikalischen Algorithmus berechnen. Stellt sich die Frage nach der genügenden Aufmerksamkeit des Lenkers, darf nur der Zeitpunkt der Reaktion variiert werden, nicht jedoch die Eingangsgeschwindigkeit. Der richtige Algorithmus resp. das methodisch korrekte Vorgehen ist von zentraler Bedeutung, wenn rechtlich relevante Varianten ohne Bruch der Kausalität resultieren sollen.

Im Beispiel resultiert bei einer Eingangsgeschwindigkeit von ca. 64 km/h gemäss Aufzeichnung des UDS eine Kollisionsgeschwindigkeit von ca. 50 km/h. Versuchen Sie abzuschätzen, bei welcher Eingangsgeschwindigkeit, aber bei sonst unveränderten Parametern, die Kollision gerade nicht erfolgt wäre – also die räumliche Vermeidbarkeit der Kollision erstmals gegeben wäre.

Resultat: Im vorliegenden Fall wäre – bei sonst unveränderten Parametern – die Kollision bei der signalisierten Geschwindigkeit von 50 km/h räumlich gerade vermeidbar gewesen.

Nur dieses Vorgehen ergibt korrekte Varianten, die anschliessend von den JuristInnen rechtlich gewürdigt werden können (angepasste Geschwindigkeit, genügende Aufmerksamkeit, allenfalls adäquate Beherrschung des Fahrzeuges z.B. bei einer Vollbremsung). Ein Beispiel, wie solche Algorithmen programmiert und für interessierte Laien verständlich visualisiert werden können, ist das Tool «anhalteweg.exe» der Präventionsabteilung der Polizei Baden-Württemberg (siehe <https://www.gib-acht-im-Verkehr.de>). Es eignet sich sehr gut für die Darstellung verschiedener Ereignisvarianten, zum Beispiel im Rahmen einer Schlusseilvernehmung.

2.3 Zwischenfazit: Umgang mit Algorithmen

Das Zwischenfazit beim Blick zurück ergibt einige grundsätzliche Forderungen an die Entwickler resp. Programmierer von Algorithmen und insbesondere an deren Anwender, die sehr ähnlich sind, wie die grundsätzlichen Anforderungen, die ein Gericht an ein forensisches Gutachten stellen muss, damit es ein brauchbares Beweismittel wird:

– *Transparenz:*

Die umfassende Dokumentation der verwendeten Algorithmen und der eingesetzten Parameter ist unabdingbare Grundlage für deren Akzeptanz. Hier kann sich durchaus ein Spannungsfeld mit allenfalls geltend gemachten Geheimhaltungsansprüchen i. S. v. urheberrechtlichen Ansprüchen oder Produktionsgeheimnissen ergeben.

– *Nachvollziehbarkeit:*

Die wissenschaftlichen Grundlagen der Algorithmen müssen dargelegt werden und allgemein anerkannt sein, d.h. es müssen bei neuen Methoden entsprechende Publikationen in seriösen Fachzeitschriften oder Lehrbüchern existieren.

– *Überprüfbarkeit:*

Die Verifizierung der Algorithmen mit Testdaten ist erfolgt und ebenfalls publiziert – allenfalls muss in einem konkreten Fall ein zweiter Experte beauftragt werden, die verwendeten Parameter und die erhaltenen Resultate unabhängig zu überprüfen (Vier-Augen-Prinzip).

– *Anwendbarkeit:*

Die Rahmenbedingungen sind definiert, unter denen ein Algorithmus verwendet werden darf, damit sichergestellt ist, dass er innerhalb des Gültigkeitsbereichs der zugrundeliegenden Theorie – z.B. von anerkannten Naturgesetzen – eingesetzt wird und valide Resultate mit definierten Toleranzen liefert.¹⁶

2.4 Lernfähige Algorithmen und prädiktive Algorithmen

Nun zu den ungleich anspruchsvolleren, lernfähigen Algorithmen – also der künstlichen Intelligenz – oder den mit Blick in die Zukunft gerichteten prädiktiven Algorithmen. Wer hätte in den späten 60er-Jahren, als z.B. «*2001: a space odyssey*» gefilmt wurde, erahnen können, was im Jahr 2020 für uns alle selbstverständlich geworden ist. Einige Vorhersagen waren verblüffend gut, andere hingegen waren schlicht falsch! Und vergessen wir nicht «*Back to the future*» aus dem Jahr 1985 ... «*Never set it to 2020! ...*».

Es folgen einige Beispiele von Algorithmen, die entweder bereits im grossen Stil genutzt werden, oder bei denen man sich gut vorstellen kann, bei welchen Tätigkeiten sie Menschen unterstützen oder im Extremfall ersetzen sollen.

16 Ausführlich J. Arnold/B. Isenring/M. Lory/S. Walder, Vom Umgang mit Unsicherheiten und Fehlern, in: M. A. Niggli/L. Marty (Hrsg.), Risiken der Sicherheitsgesellschaft – Sicherheit, Risiko und Kriminalpolitik, Neue Kriminologische Schriftenreihe Band 115, Mönchengladbach 2014, 49 ff., abrufbar unter: <http://www.krimg.de/drupal/files/9783942865333.pdf> (zuletzt abgerufen am 5.6.2020).

All diese Fragestellungen haben eine Gemeinsamkeit: Was genau die nächste Aufgabe sein wird, die es zu lösen gilt, ist für den konkreten Fall weitgehend unbekannt.

In der Realität sind es häufig Probleme, bei denen eine grössere Anzahl von Parametern zu berücksichtigen sind, die sich allenfalls gegenseitig beeinflussen.

Und hier wird sich sehr rasch die Frage stellen, inwiefern aus «alten» Mustern erfolgreiche Lösungsansätze für «neue» Probleme abgeleitet werden können.¹⁷ Entscheidende Faktoren hierbei sind die zu respektierenden *Randbedingungen resp. Grenzen*, innerhalb denen solche Systeme zum Einsatz kommen können, und die Frage nach der *Verantwortung für die Systemreaktion resp. die vom System gefällten Entscheide*.

Für die folgenden Ausführungen wird auf die Homepage von <https://www.aisoma.de> verwiesen, um einige Begriffe in einer Übersicht einzuführen.

2.4.1 Einfache Algorithmen: beschränkte Aussagekraft

Zuerst drei mehr oder weniger gut bekannte Algorithmen, die einerseits verwendet werden können, um bestehende Daten zu interpolieren, die aber anschliessend auch eingesetzt werden können, um in die Zukunft zu extrapolieren: Die drei Algorithmen «lineare Regression», «logistische Regression» und «Entscheidungsbaum» sind zwar relativ einfach zu verstehen, so dass auch deren Ergebnisse mit wenig Aufwand zuverlässig interpretiert werden können. Sie sind aber allesamt wenig geeignet, um komplexe Fragestellungen darzustellen und gaukeln insbesondere bei der Extrapolation eine zu hohe Zuverlässigkeit vor, da meist nur eine relativ begrenzte Anzahl Parameter berücksichtigt werden können. Bereits hier müssen sowohl die Eignung der Algorithmen bezüglich der vorhandenen Daten für die Interpolation als auch – und ganz besonders – die Zulässigkeitsspanne allfälliger Extrapolationen kritisch beurteilt werden.

Mit diesen drei Typen von Algorithmen lassen sich zwei wichtige Problemfelder, die intelligente – sprich lernfähige – Systeme entschärfen sollen, nicht realisieren:

- Man kann vielfach in der Realität nicht alle zukünftigen Problemstellungen vorhersehen resp. programmieren. Dadurch wird es immer wieder Situationen geben, in denen solche – auf relativ einfachen Algorithmen basierende – Systeme überfordert sein werden.

¹⁷ Ausführlich dazu mit Blick auf Digitale Spuren: A. Biedermann/K. Kotsoglou, Digital evidence exceptionalism? A review and discussion of conceptual hurdles in digital evidence transformation, *Forensic Science International: Synergy* 2/2020, 262 ff.

- Durch die statische Programmierung ist es kaum möglich, die bekannten «menschlichen» Unzulänglichkeiten zu kompensieren, die in unvorhergesehenen Situationen zu Überforderungen führen.

2.4.2 Komplexe Algorithmen: beschränkte Nachvollziehbarkeit

Die rasant fortschreitende Digitalisierung erlaubt immer weitergehende Big Data Anwendungen, die es ermöglichen, ganz neue Fragestellungen zu bearbeiten und die riesigen Datenmengen, die laufend im Internet generiert werden, intelligent zu nutzen (machine learning, deep learning). Auch hier stellt sich die Frage, inwiefern man aus grossen Datenmengen für den konkreten Einzelfall zuverlässige Vorhersagen ableiten kann: Menschliches Verhalten ist nicht immer rational ...

Damit sind wir bei einigen Algorithmen angelangt, die zwar ausserordentlich leistungsfähig sind, deren Verständnis aber hohe theoretische Anforderungen an den Anwender stellt und deren Resultate schwierig zu interpretieren sind.

- *Random Forest* ist ein sehr leistungsfähiges Klassifikationsverfahren, das z.B. Zusammengehörigkeiten (Klassen) oder Zusammenhänge zwischen Klassen erkennen kann. Im Forensischen Institut Zürich wurde es schon für vertiefte Textanalysen in umfangreichen WhatsApp-Chats verwendet.
- *Gradient Boosting* ist ein Algorithmus, der u.a. für automatisches Ranking angewendet werden kann. Als Beispiel verwenden die Suchmaschinen Yahoo und Yandex angepasste Varianten von Gradient Boosting, um das Ranking von Homepages zu berechnen.
- *Neuronale Netze* entstanden mit der Entwicklung der ersten leistungsfähigen Computer und werden heute v.a. im Bereich der Mustererkennung eingesetzt (z.B. beim Handschriftenvergleich oder zur Erkennung von krankhaften Veränderungen auf CT- oder MRI-Bildern).

Die Genauigkeit der Vorhersagen, die Komplexität des Algorithmus sowie die Komplexität des beschreibbaren «Problems» und die Interpretierbarkeit der Resultate der Algorithmen resp. die Nachvollziehbarkeit stehen in einem direkten Spannungsfeld.

Diese grundsätzliche Problematik ist auf der folgenden Grafik auf <https://www.aisoma.de> anschaulich dargestellt:

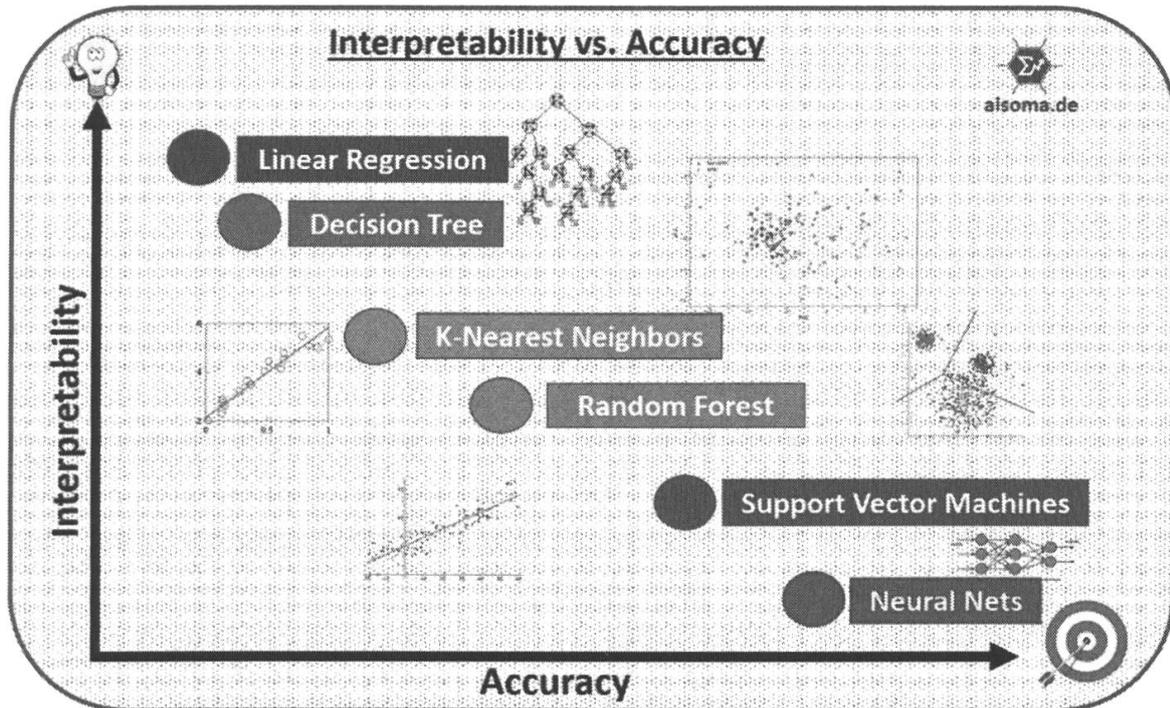


Abbildung 4: Genauigkeit und Aussagekraft vs Interpretierbarkeit und Nachvollziehbarkeit

2.4.3 Prädiktive Algorithmen: beschränkte Zuverlässigkeit

Bei den prädiktiven Algorithmen kommt ein weiteres grundsätzliches Problem hinzu, dass mit grosser Vorsicht zu beurteilen ist: Die Problematik einer Aussage in der (näheren oder fernerer) Zukunft. Hierzu sind zwei Aussagen zu machen, die jede für sich bei der Bewertung der Zuverlässigkeit von extrapolierten Aussagen in der Zukunft zu berücksichtigen sind: Egal, um welche Fragestellung es sich handelt – Aussagen in der Zukunft können nur sinnvoll gemacht und deren Unsicherheit abschätzbar sein, wenn klar deklariert wird, unter welchen Rahmenbedingungen (*boundary conditions*) die Extrapolation erfolgt. Sämtliche Ausgangswerte der Berechnungen müssen offengelegt werden.

Weiter muss bei prädiktiven Algorithmen transparent offengelegt werden, welche Parameter durch den Algorithmus variiert werden können und nach welchen Gesetzmässigkeiten.

Zuletzt muss nachvollziehbar begründet und abgeschätzt werden, wie sich die Unsicherheit der Ergebnisse des prädiktiven Algorithmus verändert in Abhängigkeit vom Zeitraum, der in der Extrapolation betrachtet wird.

Damit sollte – soweit wie möglich – offengelegt sein, wie sicher (oder eben unsicher) die vom prädiktiven Algorithmus in der Zukunft gemachten Prognosen überhaupt sein können.

In diesem Zusammenhang sei allen Experten ans Herz gelegt, hier die nötige Bescheidenheit an den Tag zu legen und diese Methoden mit der gebührenden Vorsicht anzuwenden.

3. Fazit

3.1 Forderungen im Umgang mit Algorithmen

Mit Blick auf die vier in Abs. 2.3 bei den «rückwärts» gerichteten Algorithmen diskutierten Kriterien, lassen sich Schlussfolgerungen ziehen, die sowohl für die damit befassten technischen Spezialisten von Bedeutung sind als auch für Anwender oder die mit den Resultaten konfrontierten Nutzer wie z.B. JurisInnen oder betroffene Personen (z.B. im Rahmen einer Gefährlichkeitsabschätzung):

– *Transparenz fordern!*

Die umfassende Dokumentation der Algorithmen und der verwendeten, relevanten Daten wird je länger desto wichtiger. Der «Datenschutz» muss sich in diesem Umfeld grundsätzlich wandeln – hier ist vor allem der Gesetzgeber stark gefordert. Es wird nur möglich sein, herauszufinden, was nicht «richtig» funktioniert hat, wenn bei einem «Versagen» solcher Systeme im praktischen Einsatz der Zugriff auf die relevanten Daten möglich ist und die notwendigen Werkzeuge für die Interpretation der Daten verfügbar sind. Es muss vom Gesetzgeber in Absprache mit den technischen Spezialisten rasch und detailliert geklärt werden, *was an solchen Daten zur Verfügung stehen muss, wer für die Sicherung dieser Daten verantwortlich ist, wie der berechtigte Zugriff auf diese Daten gewährleistet wird und wer über die nötigen Fachkompetenzen verfügt, um solche Daten korrekt zu interpretieren.* Weiter muss geklärt werden, unter welchen rechtlichen Aspekten man von digitalen Spuren sprechen kann – also weder von Personendaten noch von Zwangsmassnahmen – d.h. was genau unter den (strafprozessualen) Begriff der Aufzeichnungen fallen soll.

Dasselbe gilt für Daten, die bei weitem nicht immer Personendaten sind und den zu ziehenden Grenzen für den Datenschutz.

– *Nachvollziehbarkeit fordern!*

Die wissenschaftliche Grundlage der Algorithmen muss von unabhängigen Experten überprüft werden können. Dies bedingt auch, dass eine umfassende *Pflicht zur Offenlegung* der entsprechenden Grundlagen gefordert werden muss. Hier können sich heute die Systementwickler resp. -hersteller viel zu einfach hinter dem Argument von eigenem Knowhow und Produktionsgeheimnis verstecken. Der Umweg über die Produktesicherheit ist ebenfalls

heikel, da es sehr rasch zu einer Beweislastumkehr kommen kann und der betroffene Anwender dem Anbieter beweisen muss, dass etwas nicht richtig funktioniert hat.

– *Überprüfbarkeit fordern!*

Die *Verifizierung der Algorithmen* mit Testdaten muss Pflicht werden. Die Anbieter müssen für die Resultate ihrer Systeme die Verantwortung übernehmen – das wird umso entscheidender, je weniger die ein System bedienenden Personen noch wesentlich in die Funktionalität oder die Entscheidungen des Systems eingreifen können. Dies hat einen direkten Zusammenhang mit der Frage, wie detailliert die Betriebsdaten solcher Systeme aufgezeichnet werden müssen. Es muss sich z.B. eindeutig aus den Aufzeichnungen ableiten lassen, ob das System den Nutzer aufforderte, einzugreifen oder allenfalls die Kontrolle (wieder) zu übernehmen. Und genauso muss sich sagen lassen, ob, wann und wie der aufgeforderte Nutzer reagierte.

– *Anwendbarkeit eingrenzen!*

Es wird entscheidend sein, dass die Rahmenbedingungen genau definiert sind, unter denen ein Algorithmus sicher funktioniert und wie er verwendet werden darf. Da die Anbieter für Fehlfunktionen haften müssen, werden sie selber ein grosses Interesse daran haben, aufzeigen zu können, dass ihr System einwandfrei funktionierte oder dass es ausserhalb der zulässigen Rahmenbedingungen betrieben wurde.

3.2 Beispiel hochautomatisierte Fahrzeugsysteme

Die in Abs. 2.5 formulierten Forderungen werden u. a. dazu führen, dass sich unser Bild von hochautomatisierten Fahrzeugsystemen wandeln wird. Als Beispiel sei das quasi autonom verkehrende Postauto in Sion genannt, das sich mit einer Höchstgeschwindigkeit von 15 km/h auf festen Routen bewegt.

Das Bild der automobilen Freiheit wird sich mit autonomen (Privat-) Fahrzeugen stark relativieren – diese Fahrzeuge werden sich zuerst und immer an den geltenden Regeln z.B. des Strassenverkehrsrechts orientieren und diese konsequent einhalten, da sonst bei einem Zwischenfall der Hersteller zur Rechenschaft gezogen werden könnte. Will also ein Lenker das System übersteuern, wird das aufgezeichnet werden und er übernimmt damit nachweisbar die volle Verantwortung für sein weiteres Handeln.

3.3 Schlussfolgerungen und Konsequenzen

Für die Justiz wird es in diesem Kontext kein blindes Vertrauen in solche Systeme oder deren Entwickler, Hersteller oder Verkäufer geben können.

So unbequem es auch sein mag – das *grundsätzliche Verständnis* für die Funktionsweise solcher Systeme und insbesondere ihrer Grenzen werden sich auch die JuristInnen aneignen müssen. Dabei wird es nötig sein, genau hinzuschauen, was und vor allem wer sich zu diesen Grenzen äussert: Der verantwortliche Entwickler wird vermutlich deutlich sorgfältiger und zurückhaltender argumentieren als z.B. der provisionsgesteuerte Verkäufer.

Als Beispiel sei hier der «Autopilot» von Tesla genannt, der nur in den Verkaufsprospekten so genannt und angepriesen wird. Im Betriebshandbuch mutet die entsprechende Zurückhaltung – resp. die gebetsmühlenartig wiederholte Zuweisung der Verantwortung an den Lenker – mit Blick auf die Produktesicherheit zum Teil grotesk an.

Aber auch panische Angst ist hier fehl am Platz: Interessiertes Mitdenken, kritisches Hinterfragen von angepriesenen Fähigkeiten eines Systems und gesunder Menschenverstand, wenn es um die grundsätzlichen Grenzen des technisch Möglichen geht, sind hilfreiche und durchaus effiziente Mittel, um sich ein realistisches Bild machen und gute Fragen stellen zu können.

Hier gilt das Gleiche wie bei allen anderen Experten, mit denen die Strafjustiz allenfalls konfrontiert ist: Eine Frage ist erst dann beantwortet, wenn die Antwort für einen interessierten Laien verständlich – also nachvollziehbar – begründet, widerspruchsfrei und vollständig ist. Allerdings ist Mitdenken anstrengend und der erforderliche Aufwand an Zeit und allenfalls finanziellen Mitteln darf nicht unterschätzt werden.

Der Gesetzgeber ist bei diesen Fragen stark gefordert: Es muss ein Weg gefunden werden, um wieder die *grundlegenden Prinzipien* gesetzlich zu regeln. Die Tendenz, die sich insbesondere im Datenschutz und bei vielen Themen im hochtechnisierten Umfeld von Technik und Wissenschaft in der jüngsten Vergangenheit durchgesetzt hat, alles auf dem aktuellen Stand der Technik – dafür aber mangels grundlegendem Verständnis – bis ins letzte Detail regeln zu wollen, ist fatal. Gesetze sind so zu schreiben, dass sie nicht laufend von der technischen Entwicklung und neuen Realitäten überholt werden und uns lähmen.