

Artificial Intelligence as the End of Criminal Law?

On the Algorithmic Transformation of Society

(https://doi.org/10.47907/livro2021_4c7)

*Christoph Burchard*¹

Abstract:

Does Artificial Intelligence (AI) imply the end of criminal law and justice as we know it? This article submits that AI is a transformative technology that seemingly assumes and optimizes the rationalities of criminal law (the effective prevention of crime; the objective, neutral and coherent application of the law etc.), namely by replacing the

¹ Prof. Dr., LL.M. (NYU), Chair for Criminal Law and Criminal Procedure, International and European Criminal Law, Comparative Law and Legal Theory at the Goethe University Frankfurt am Main.

This contribution was first published in German with the title «Künstliche Intelligenz als Ende des Strafrechts? Zur algorithmischen Transformation der Gesellschaft», in JOERDEN Jan C./SCHUHR Jan C., *Jahrbuch für Recht und Ethik. Zugleich Gedächtnisschrift für Joachim Hruschka*, Berlin: Duncker & Humblot, 2019, p. 527-555, and later in Italian «L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società», *Rivista Italiana di Diritto e Procedura Penale*, 4/62 (2019) p. 1909-1942. I would like to thank the editors and the publisher for letting me now publish an English translation in this volume.

The English translation has been curated by Mr. Johannes Abrell. To him goes my utmost gratitude.

This contribution was published as part of the research project “The Normative Order of Artificial Intelligence | NO:AI”, which I am conducting as Principal Investigator at the Research Centre “The Formation of Normative Orders”.

counterfactual guarantees of the law with the factual guarantees of technology. As a consequence, AI must not be trivialized by criminal law theory. Likewise, it is not enough to subversively criticize the current weaknesses of AI (e.g. vis-à-vis the “bias in, bias out” problem). Rather, criminal law theory should draw on the highflying promises of AI to reflect upon the foundational premises of criminal law. For a criminal law that is mostly a governance tool in the administrative and/or welfare state, AI applications promise the culmination of the law’s very objectives (like the effective inhibition and prevention of crime, e.g. by means of predictive policing; or the political determination of fuzzy sentencing rationales in sentencing algorithms that ensure equal sentences for comparable crimes). For a criminal law, however, that protects liberal freedoms and rests on inter-personal trust, AI may well lead to the passing of the law’s very ideals (e.g. of the presumption of innocence, which can no longer be upheld once everyone, ordinary citizens and judges alike, is deemed a possible risk). The question about “AI as the end of criminal law?” thus eventually raises the two-pronged question “Which criminal law for which society?”. Indeed, what is the status of freedom (especially in a surveillance society needed to power Big Data driven algorithms), trust (especially under the zero trust paradigm that underlies many risk assessment algorithms) and future (especially when algorithms make predictions based on past data) once AI enters into the administration of criminal justice? These are the questions, or so I respectfully submit, that criminal law theory needs to address today in order to come up with a criminal law that is both (for pragmatic reasons) open to technology as well as (for humane reasons) sensible. In all of this, we must take to heart Joachim Hruschka’s great legacy and remain intellectually honest.

Key-words: criminal law; artificial intelligence; predictive society; big data policing; predictive policing; risk assessment algorithms.

Question: “Can you imagine smart machines powered by AI being used to support judicial decision-making in the future?” – Answer: “This is already a reality. And it puts a very significant burden on the judiciary and how it functions.” – Question to and answer by U.S. Supreme Court Chief Justice John Glover Roberts Jr. in April 2017²

INTRODUCTION

Shortly before I received the honorable invitation to contribute to this important commemorative volume, I had just examined the essay by *Byrd/Hruschka* on Kant’s theory of crime in a seminar on “Criminal Law between Morality and Politics”.³ A truly groundbreaking essay, which removed Kant from the (for many unappealing) wing of absolute theories of crime. This marked the end of a journey for me. I first encountered *Joachim Hruschka’s* groundbreaking oeuvre while working on my doctoral thesis. I can still remember vividly how impressed I was – then⁴ and ever since – by his methodically and jurisprudentially reflected, equally logical-analytical and in the best sense enlightened approach to dogmatics. Therefore, it is a real honor for me, in memory of *Joachim Hruschka*, to be able to offer some initial thoughts today on recent developments, precisely on how so-called artificial intelligence “threatens” to change our society and our criminal law. This choice of topic may seem surprising at first, since there seem to be only a few direct points of reference to *Joachim Hruschka’s* work. However, as I would like to show, the use of artificial intelligence poses fundamental challenges to our social and criminal legal order. What is needed no less than a reflection on the normative foundations of our criminal law, is a criminal law suited to a society, which is to be explained,

² Reported by LIPTAK, Adam, «Sent to Prison by a Software Program’s secret Algorithms», *New York Times*, May 1, 2017, p. A22.

³ BYRD, Sharon; HRUSCHKA, Joachim, «Kant zu Strafrecht und Strafe im Rechtsstaat», *Juristen Zeitung*, 20/62 (2007) p. 957.

⁴ I used the following titles, inter alia, in my doctoral thesis: HRUSCHKA, Joachim, «Die Herbeiführung eines Erfolges durch einen von zwei Akten bei eindeutigen und bei mehrdeutigen Tatsachenfeststellungen», *Juristische Schulung* (1982) p. 317; – *Strafrecht nach logisch-analytischer Methode*, Berlin/New York: De Gruyter, 1988; – «Der Standard-Fall der aberratio ictus und verwandte Fallkonstellationen», *Juristen Zeitung*, 10/46 (1991) p. 488.

understood and evaluated anew. And in this regard, the virtual conversation with *Joachim Hruschka* needs to be continued in confrontation with his philosophical, not solely alleged,⁵ but always justified insights, in order to be able to critically question today's developments. Less courageously than *Joachim Hruschka*, who once demanded no less than a rethinking of criminal law,⁶ I advocate here to pass a reckoning on the society that brings forth or is to bring forth, our criminal law. To this end I must – as *Joachim Hruschka* recognized a long time ago⁷ open the criminal law again to the social and political theory.

“Artificial intelligence (AI) and criminal law” is *not* science fiction. In light of the “digital revolution” permeating all areas of life, neither law in general nor criminal law in particular can escape the influences of AI.⁸ As far as AI is already taken into consideration by criminal law scholars, it is traditionally (especially in Germany) considered as an area potentially in need of regulation⁹ and also capable of regulation.¹⁰

This contribution is advocating for a *supplementing of perspective*.¹¹ In other words, to focus on AI (or more precisely its social practice) as a medium of contemporary social and (criminal) legal

⁵ Decidedly against this HRUSCHKA, Joachim, «Kann und sollte die Strafrechtswissenschaft systematisch sein?», *Juristen Zeitung*, 1/40 (1985) p. 10.

⁶ HRUSCHKA, Joachim, «Das Strafrecht neu durchdenken!», *Goltdammer's Archiv* (1981) p. 237.

⁷ *Id.*, p. 249. Also relevant to social criminal law theory HRUSCHKA, Joachim, «Utilitarismus in der Variante von Peter Singer», *Juristen Zeitung*, 6/56 (2001) p. 261.

⁸ See for instance the provocative questioning of SCHWINTOWSKI, Hans-Peter, «Wird Recht durch Robotik und künstliche Intelligenz überflüssig?», *Neue Juristische Online-Zeitschrift*, 42 (2018) p. 1601.

⁹ See in general MEYER, Stephan, «Künstliche Intelligenz und die Rolle des Rechts für Innovation», *Zeitschrift für Rechtspolitik*, 8/51 (2018) p. 233 for the relevant considerations to be made in this regard, including the need to maintain technological, innovative strength in international (economic and scientific) competition.

¹⁰ The *agency* question is often prominent in this context, *i.e.*, who is responsible if damage is caused by the use of an AI system, especially when self-learning AI is used. On this topic in detail see HILGENDORF, Eric, «Autonome Systeme, künstliche Intelligenz und Roboter», in BARTON, Stephan et al., *Festschrift für Thomas Fischer*, Munich: C.H. Beck, 2018, p. 111 ff.

¹¹ In general on this topic see BALKIN, Jack B., «The Path of Robotics Law», *California Law Review Circuit*, 6 (2015) p. 45.

transformations.¹² The question should therefore be asked: How does AI already change our social and criminal legal system?¹³ Does AI even mean, in exaggerated terms, the end of criminal law?¹⁴ Be it either in the sense of a dying death of its fundamental principles (such as the presumption of innocence) or a crowning culmination of its fundamental goals (such as the protection of legal interests)?¹⁵

In order to answer these questions, it is necessary to examine the justification narratives more closely that are already being used today for the introduction of AI into criminal law; to analyze more closely the normative conceptions of order that lie behind these narratives and that are also concealed by them;¹⁶ and in doing so, to highlight the power-political and ideological openness of these conceptions of order. In the spirit of this program, the promises of AI will first be illuminated with a view to current developments (see I. and II. below),

¹² AI can therefore be flagged as a potentially “*transformative technology*” to establish a dialectical connection between technological, societal, and legal change. On this topic in general see FATEH-MOGHADAM, Bijan, «Selbstbestimmung im biotechnischen Zeitalter», *Basler Juristische Mitteilungen*, 5 (2018) p. 205 (in particular p. 209 ff.). – By focusing on the social practice of AI as a medium of social transformations, we on the one hand oppose an essentialization of AI (see I. below) and on the other hand clarify that AI itself does not have a normative effect, but nevertheless conveys programmed ideas of order.

¹³ The question insinuates a causality which, on closer inspection, can be resolved as a dialectical process in which social etc. developments promote the development and use of certain AI systems, which in turn reinforce the first-mentioned developments etc. AI hereby “naturally” builds on the general mechanization of real life. For example, the use of electronic ankle bracelets, especially against so-called endangerers, could be seen as a glimpse of a partial surveillance society, which always deprives certain groups of people (precisely these so-called endangerers) of the counterfactual trust in their lawfulness. This can then be reinforced by AI with a general *zero trust* paradigm. In this respect see II. below.

¹⁴ This question is meant to shake things up and not to express a *fin de siècle*. The point is: how are we to deal with the polyvalences of today’s developments? For a similar play on words see HILDEBRANDT, Mireille, *Smart Technologies and the End(s) of Law*, Cheltenham: Edward Elgar Publishing, 2015.

¹⁵ My questions are *not* to be understood as speculative science fiction, so that general dystopias are not considered here.

¹⁶ For an introduction to this conceptual apparatus coined at the Cluster of Excellence “Die Herausbildung normativer Ordnungen” see FORST, Rainer; GÜNTHER, Klaus, «Die Herausbildung normativer Ordnungen: zur Idee eines interdisziplinären Forschungsprogramms», in FORST, Rainer; GÜNTHER, Klaus, *Die Herausbildung normativer Ordnungen*, Frankfurt a.M.: Campus, 2011, p. 11 (in particular p. 15 f.).

in order to then reflect on how a criminal law theory, that opens up critical perspectives,¹⁷ should deal with them (see III. and IV. below).

Particular attention is paid in this respect to promises that supposedly smart or intelligent algorithms that regularly evaluate Big Data will provide more effective and efficient protection of legal interests and allow for a more neutral, objective, and coherent law enforcement than human decision makers. These promises correspond *prima facie* to those of criminal law. As the supposedly *ultima ratio* of the state, criminal law also promises a particularly thorough protection of legal interests. Moreover, criminal law demands for its impartial and unbiased as well as consistent application. The difference between the two is that AI promises technological facticity, while criminal law – like law in general – can only profess counterfactual guarantees. Whether AI means the end of criminal law is in consequence synonymous with the question, which criminal law is meant. A liberal criminal law of freedom, which is based on interpersonal trust and does not manage people solely as controllable and constantly assessable potential risks (in other words, as endangerers), is fundamentally called into question by AI; a liberal (criminal) law theory would therefore have to seek the forward defense of the counterfactual of (criminal) law against the factual of AI. For a welfare-state security criminal law that sees itself as an instrument of social control or governance of social interactions, AI on the other hand enables a crowning culmination of its rationality.

I. AI AS A NON-ESSENTIALIST CONSTRUCT

AI is not an unambiguous term. The exact definition of the term is being fought over on all sides.¹⁸ One often encounters what is

¹⁷ Criticism is understood here axiomatically as a practice of justified doubt. The normative program lies in the opening of possibilities for critical, doubting inquiries, not in the development of a normative program from which criticism can be exercised. The latter would be reserved for a critical theory of criminal law, which cannot be developed here. A return to a rational concept of critique is also demanded by HRUSCHKA, Joachim, *Strafrecht nach logisch-analytischer Methode*, Berlin/New York: De Gruyter, 1988, p. XI.

¹⁸ Of fundamental importance are debates about whether the name AI even accurately represents the current state of research and development (e.g., if and because

essentially a structural essentialism that seeks to fathom the proprietary properties of intelligence in general and artificial intelligence in particular.¹⁹ This may also explain why profound discussions about the criminal responsibility of intelligent and self-aware machines of the future²⁰ are already being held.²¹

Here, nevertheless, a non-essentialist understanding of AI is advocated in order to derive insights from the perspective of an observer and make these of use for a modern and critical theory of criminal law, that thinks in the categories of the participant perspective. Thus, it is not about what constitutes or should constitute the essence of

today's "AI" systems do not go beyond classical machine learning and time-honored pattern recognition); whether AI is really "intelligent" (e.g., if and because today's "AI" systems cannot provide transfer services); and whether the German qualification "künstlich" is correct (e.g., if and because the qualification "maschinelle Intelligenz" is meant to represent the properties of algorithms more correctly, or the "artificial" is pejorative in the romantic sense). On this topic see HERBERGER, Maximilian, „Künstliche Intelligenz“ und Recht», *Neue Juristische Wochenschrift*, 39/71 (2018) p. 2825. – Cf. further (the only at first seemingly outdated work of) WEIZENBAUM, Joseph, *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt a.M.: Suhrkamp, 1977, p. 268 ff.

¹⁹ This is how for instance ERTEL, Wolfgang, *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*, Wiesbaden: Springer VS, 2016, p. 1 ff. views the questions "What is intelligence?", "How can intelligence be measured?", or "How does our brain work?" as significant for the understanding of AI. It is further alleged that for computer scientists and in particular engineers, the question "about the intelligent machine that behaves like a human being, that shows intelligent behavior," is decisive.

²⁰ This is reflected in an essentialist and anthropocentric understanding of criminal law, which needs to be questioned in the course of the worldwide triumph of associational and corporate criminal law. It does not seem far-fetched (and nothing further is put up for discussion here) to want to justify the criminal liability of AI systems strictly functionally, *i.e.* independent of how intelligent an AI system now is and whether it is aware of itself and whether it can thus be ascribed original human characteristics. The following points, in analogy to corporate criminal liability, could be mentioned functionally in favor of an AI criminal liability: (1) Indirectly, the owners of a deficient AI system should be targeted (e.g., if the shutdown of the AI system is ordered) (2) The recourse to responsible persons who have developed a deficient AI system or brought it to the market shall be cut off (e.g. because and if there are innumerable persons "behind" an AI system, from programmers to company managers of any hierarchical level, so that an individualization of responsibility would not be feasible). (3) Or, by attributing criminal responsibility, feelings of resentment and indignation are to be expressed in an institutionalized form (I thank *Boris Burghardt* for pointing out this aspect).

²¹ For further detail see GAEDE, Karsten, *Recht und Strafen für Roboter?*, Baden-Baden: Nomos, 2019 with further references.

(artificial) intelligence. Rather, the starting point is to ask how AI (especially through corresponding justification narratives) is socially constructed, represented and received; how AI is embedded in certain social relations and changes them; and which relations of domination and power are expressed in AI and stabilized, mystified, transformed or produced by it.²² In other words, it is a matter of assigning meanings that originate in the social reality of life and have an effect on it. In this process, normative conceptions of order (including ideologies) which are to be determined in each case are promoted and specific conflicts are either directed to the center of attention or cast into the background. Seen in this light, AI is a normatively open and malleable, and simultaneously political construct. It is exactly the ambiguity and power-political openness of the term AI that allows interested actors to instrumentalize it for their own – political, economic, etc. – purposes. AI is a particularly powerful name, because it attests to the general cognitive ability of an IT system to “understand and learn well, and to form judgements and opinions based on reason” in purely linguistic terms – i.e., independently of its “real” intelligence.²³

In this context, AI should *not* be seen as a neutral technology or simply as an information technology innovation. Rather, AI is directly linked to the basic principles of human sociality (freedom, tolerance, law, etc.), traces back to them and transforms them. Therefore, it is important to take the promises, hopes and fears associated with the term AI at any given time seriously in order to be able to subject them to a critical reflection, as they are fueled and spread with a varying urgency by the most diverse actors in business, politics, science, etc.

As an example, let us refer to the conventional image of AI that has been cemented in our minds, thanks to *Hollywood*.²⁴ Here, AI

²² Similarly BALKIN, Jack B., «The Path of Robotics Law», *California Law Review Circuit*, 6 (2015) p. 59; WEIZENBAUM, Joseph, *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt a.M.: Suhrkamp, 1977, p. 268 ff. See also MAU, Steffen, *Das metrische Wir: Über die Quantifizierung des Sozialen*, Berlin: Suhrkamp, 2017.

²³ Intelligence is understood here in terms of the third definition as provided by the online Cambridge dictionary.

²⁴ For an overview on this topic see XANKE, Lisa; BÄRENZ, Elisabeth, «Künstliche Intelligenz in Literatur und Film – Fiktion oder Realität?», *Journal of New Frontiers in Spatial Concepts*, 4 (2012) p. 36; IRSIGLER, Ingo; ORTH, Dominik, «Zwischen Menschwerdung und Weltherrschaft: Künstliche Intelligenz im Film», *Aus Politik und Zeitgeschichte*, 6-8/68 (2018) p. 39.

stands for autonomous robots (one thinks dystopic of “Terminator” or more overtly “I, Robot”), intelligent androids (one thinks of Lieutenant Commander Data in “Star Trek: The Next Generation”), and self-aware supercomputers (think of HAL 9000 in “2001: A Space Odyssey” or Central in “Star Trek: Discovery”). In this respect, AI appears (which is dramaturgically understandable) almost exclusively as so-called *strong AI*, which strives for the same general intelligence as humans or has already attained it, if not surpassed it.²⁵ The main focus of the discussion is then on both the *conditio humana* (whether and how humans can integrate into a machine world)²⁶ and the *conditio automata* (whether and how intelligent machines can integrate into a human society).²⁷ Since none of the systems in existence today falls under the category of strong AI,²⁸ these discussions admittedly do not yet have any immediate practical legal significance. This may be one reason why “AI and criminal law” has been an orchid topic at best until recently, and why the inherent characteristics of strong artificial intelligence continue to be fought over in an essentialist manner.

However, we may not forget that we are currently becoming the witnesses of another AI revolution.²⁹ This revolution is based, in substance, on so-called *weak AI*, which is optimized for solving problems for specific applications, is based on known methods of mathematics and computer science, and does not acquire a deeper – or actual – understanding of problem solving.³⁰ And in which – in contrast to Hollywood – no robots or androids are used.

²⁵ For the differences between weak and strong AI see the “classic” SEARLE, John R., «Minds, Brains, and Programs», *Behavioral and Brain Sciences*, 3/3 (1980) p. 417.

²⁶ As a literary example see: MCEWAN, Ian, *Machines like Me*, London: Penguin, 2019.

²⁷ As a cinemactical example see: *Ex Machina*, 2015.

²⁸ If one follows the conceptual distinction between weak and strong AI as developed by SEARLE, John R., «Minds, Brains, and Programs», *Behavioral and Brain Sciences*, 3/3 (1980) p. 417, this distinction is in its present use still thoroughly inconsistent. The central vanishing point of a strong AI is seen in the equivalent to human abilities. On this see RAMGE, Thomas, *Mensch und Maschine: Wie künstliche Intelligenz und Roboter unser Leben verändern*, Stuttgart: Reclam, 2018, p. 19.

²⁹ Essentially as here, the popular science work by FRY, Hannah, *Hello World: Was Algorithmen können und wie sie unser Leben verändern*, München: C.H. Beck, 2018.

³⁰ In favor of this, despite all opposition, RAMGE, Thomas, *Mensch und Maschine: Wie künstliche Intelligenz und Roboter unser Leben verändern*, Stuttgart: Reclam, 2018, p. 19.

The fact that these systems are also listed here as AI is a consequence of the methodological approach we have just adopted. Due in part to the corresponding efforts of industry,³¹ we in politics and society are increasingly detaching ourselves from AI à la Hollywood and understand AI to mean information technology systems that are “sold” to us as solutions to real-life problems – and that we also “buy” for this purpose. Indicative of this is the recent concept paper of the Council of Europe’s “European Committee on Crime Problems (CDPC)” on “AI and Criminal Liability”. There, self-driving cars along with their self-learning algorithms are flagged as prime examples of AI as though this were self-evident.³² The Council of Europe’s “European Commission for the Efficiency of Justice (CEPEJ)” is playing the same tune by advocating the use of AI in the administration of justice in certain areas and under certain conditions.³³ All this is an expression of today’s AI hype. This hype is carried by the multilaterally reproduced narrative that smart or intelligent algorithms are capable of mastering real-life problems that exceed human capabilities by means of information technology to the benefit of all; and in fact do so better, faster and more cost-effectively than human decision-makers (beginning with the safe control of cars and continuing over to the evaluation of all medical publications in support of disease diagnoses and therapy concepts up to the automation of legal services, so-called *legal tech*³⁴).

This “new” meaning of AI is not neutral, but normative and also ideologically charged. One should, despite the ongoing struggle for interpretive dominance, be careful not to speak of “the” – all the while dominant – meaning of AI. Nonetheless, *Katz’s* finding is convincing

³¹ For an elaboration and critique see ZUBOFF, Shoshana, *Das Zeitalter des Überwachungskapitalismus*, Frankfurt a.M./New York: Campus, 2018.

³² European Committee on Crime Problems (CDPC), “Artificial Intelligence and its Impact on CDPC Work: The case of automated driving”, CDPC (2018) 14 – 14.09.2018, p. 6 *et seq.* [Date of consultation 14.12. 2021], Access: <https://rm.coe.int/cdpc-2018-14-artificial-intelligence-and-criminal-law-project-2018-202/16808d6d09>.

³³ CEPEJ, “European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment” v. 3 – 4.12.2018, in particular p. 64 *et seq.* [Date of consultation 14.12. 2021], Access: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

³⁴ For overviews on this topic see FRIES, Martin, «Automatische Rechtspflege», *Rechtswissenschaft*, 4/9 (2018) p. 414.

that “the “AI” label has been rebranded to promote a vision of world governance through big data.”³⁵ This vision is often combined with a metaphysically and theologically charged hope of redemption. AI promises a self-learning and self-improving entity that promises transparency and predictability of the other as well as the self, and thus hyper-rational controllability of social interactions.³⁶ This narrative has particular social and political traction and explosive power, since it is generalized and permeates all areas of life (which already sets AI apart quantitatively from earlier particularistic “criminological” currents à la Lombroso). The narrative, moreover, is so generically construed that it becomes neoliberally,³⁷ rationally and scientifically,³⁸ liberally³⁹ as well as authoritatively⁴⁰ appropriable.⁴¹ AI, or rather their fundamental normative postulates of order and narratives of justification, are in other words latently open to (power) politics. And in them, a new manifestation of the dialect of enlightenment is also revealed.

II. THE PROMISES OF AI FOR CRIMINAL JUSTICE

After having previously addressed the general promises made by AI, we will now address the specific promises made by AI for criminal law or criminal justice. Namely, the promise that crime can actually

³⁵ KATZ, Yarden, «Manufacturing an Artificial Intelligence Revolution», SSRN (2017) p. 1, [Date of consultation 14.12.2021], Access: <https://ssrn.com/abstract=3078224>.

³⁶ In this regard, the stance of NIDA-RÜMELIN, Julian; WEIDENFELD, Natalie, *Digitaler Humanismus*, München: Piper, 2018, p. 44 *et seq.* is accurate.

³⁷ Critical in this regard ZUBOFF, Shoshana, *Das Zeitalter des Überwachungskapitalismus*, Frankfurt a.M./New York: Campus, 2018.

³⁸ Critical in this regard NIDA-RÜMELIN, Julian; WEIDENFELD, Natalie, *Digitaler Humanismus*, München: Piper, 2018, Introduction.

³⁹ As for instance CHIAO, Vincent, «Predicting Proportionality: The Case for Algorithmic Sentencing», *Criminal Justice Ethics*, 3/37 (2018) p. 238 ff.

⁴⁰ For an overview on this – and in particular regarding China – see MAU, Steffen, *Das metrische Wir: Über die Quantifizierung des Sozialen*, Berlin: Suhrkamp, 2017, p. 9 ff.

⁴¹ Theoretically, it can then be said that the justification narratives of AI (the first analytical point of reference) refer to underlying normative orders (the second analytical point of reference), which are themselves open to power politics or ideology (the third analytical point of reference).

be made impossible, or can in any case be drastically reduced, through the use of intelligent information technology (see 1. below); and that decision making in criminal justice can be exempt from human subjectivity and bias and therefore can “finally” be truly objective, neutral, and coherent (see 2. below).

1. EFFICACY AND EFFICIENCY IN THE INHIBITION OF CRIME

First, AI promises to make crime impossible, both directly and indirectly. Or more precisely: advocates promote that certain forms of crime are *eo ipso* no longer committable through the use of AI, or that the commission of certain forms of crime is *de facto* significantly reducible through AI-supported (sovereign, privatized, or internalized) enforcement and surveillance structures. This is illustrated by so-called Smart Contracts⁴² and Predictive or Big Data Policing⁴³.

a) Smart Contracts

Smart contracts are qualified as “smart”, i.e. as clever, witty and sophisticated contracts, on the basis of specific information technology requirements. They appear with the promise of being able to handle contractual interactions more effectively and efficiently. Smart contracts are implemented through computer programs that aim to algorithmically enable, verify, and enforce contractual rights and obligations without relying on third parties. The idea is to automate contract drafting and execution to the greatest possible extent, which is ideally “self-executing” and thus minimizes the transaction costs of traditional contract law. Through this, the legal system (including its representatives such as notaries or judges) but also private service providers who insure against payment or delivery defaults are to be made obsolete.

⁴² Smart contracts are generally discussed in private rather than criminal law literature, so further references have been omitted here. Smart contracts are not necessarily to be subsumed under AI. However, since corresponding developments are becoming apparent, they are presented here.

⁴³ Big Data Policing describes a recent development, particularly in the USA, which has yet to be developed in Germany in terms of terminology. For this see FERGUSON, Andrew G., *The Rise of Big Data Policing. Surveillance, Race, and the Future of Law Enforcement*, New York: New York University Press, 2017. See also the contributions in the Ohio State Journal Criminal Law 2018, p. 473 ff. on a “Round Table on Big Data and Criminal Law”.

Free by the credo “code is law”, the law as well as the application of the law is replaced by a corresponding IT infrastructure.

Smart contracts are based on “distributed ledger technology” such as the so-called blockchain, i.e. a database networked with the reality of life, which is de-centrally stored, verified and continuously updated. And in which, for example, is stored who has which goods or financial resources. The more comprehensive and accurate this database is – that is, the more data records (Big Data) it has at its disposal – the more precisely and securely the smart contract can and will be executed – at least that is the claim.

This is supported by the vision of a trust-free society in which contract and/or interaction partners no longer have to trust each other.⁴⁴ This is because – according to the justification narrative – they are given better, namely absolute or unchallengeable informational assurances (e.g. that the seller has the offered goods at his disposal and will hand them over and transfer ownership; and that the buyer is sufficiently liquid and will actually pay for the goods). This is accompanied by a shift in trust, away from interpersonal trust and towards trust in information technology systems (the database and programming).⁴⁵ The contract and interaction partners are basically considered as a risk, since it – indeed – cannot be ruled out that they will breach their word or the contract. To cope with this risk, a hand is laid on interpersonal trust as a counterfactually postulated, as it is simply a socially necessary, mechanism for reducing social complexity.⁴⁶ “Smart” databases and algorithms are supposed to ensure that people interacting socially rely on each other, because they know about one another, that the possibility of database- and algorithm-unfriendly behavior is taken away from them by “smart” algorithms and databases. One person relinquishes the possibility that they might behave in disconformity to the databases and algorithms, because, if and so that the other person

⁴⁴ In general see PALKA, Silvia; WITTPAH, Volker, «Vertrauen und Transparenz – Blockchain Technologie als digitaler Vertrauenskatalysator», Working Paper of the Institute for Innovation and Technology, 39 (2018).

⁴⁵ For more on this and in general see WAGNER, Gerald, «Vertrauen in Technik», *Zeitschrift für Soziologie*, 2/23 (1994) p. 145.

⁴⁶ According to the classical definition of LUHMANN, Niklas, *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*, 5. Ed., Munich: UVK, 2014, p. 27 ff. (in particular p. 30).

is deprived of this possibility as well. To put it bluntly: *homo homini lupus est* is not only psychologically displaced (trust-based interaction), but also averted from the outset in terms of information technology (so-called *zero trust* or *in tech we trust* interaction).

In terms of criminal law, this is intended to make it impossible to commit fraud in the case of exchange contracts, when leaving sovereigns and the privatizing of crime prevention out of the equation.⁴⁷ This can be seen in the “grail scripture” of the original blockchain movement, with and in which the Bitcoin idea was invented and explained. Here it is stated that:

“Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for nonreversible services. With the possibility of reversal, the need for trust spreads. Merchants must be wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as unavoidable. These costs and payment uncertainties can be avoided in person by using physical currency, but no mechanism exists to make payments over a communications channel without a

⁴⁷ Further examples of such a privatization of crime prevention through AI are provided, for instance, by AI-supported “criminal compliance” systems (i.e., Digital Compliance tools), which promise e.g., full monitoring of internal company communications with the aim of “flagging” suspicious interactions and thus subjecting them to further scrutiny. For an overview on this see SCHEMMEL, Alexander; DIETZEN, Alexandra, «Effective Corporate Governance” by Legal Tech & Digital Compliance», in BREIDENBACH, Stephan; GLATZ, Florian, *Rechtshandbuch Legal Tech*, Munich: C.H. Beck, 2018, p. 137 – In Japan, as can be seen from press reports, there is also controversy about whether AI should be used to prevent shoplifting. For this purpose, customers can be monitored algorithmically in order to be able to make predictions from the analysis of their body language, as to whether they are planning to shoplift. On this see LEWIS, Nell, «Should AI be used to catch shoplifters?», *CNN Business* (2019) [Date of consultation 14.12.2021] Access: <https://edition.cnn.com/2019/04/18/business/ai-vaak-shoplifting/index.html>.

trusted party. What is needed is an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Transactions that are computationally impractical to reverse would protect sellers from fraud, and routine escrow mechanisms could easily be implemented to protect buyers.”⁴⁸

b) Predictive/Big Data Policing

Predictive policing and *big data policing*, which originated in the United States and are now also taking hold in other Western countries, including Germany,⁴⁹ are also based on corresponding narratives. Here, as there, it is about – according to the presentation by producers and distributors – intelligent algorithms that provide resource-scarce authorities with powerful instruments for the prospective prevention of crimes.⁵⁰ *Brantingham*, a scientific pioneer of predictive policing, defines it according to the following three-step process:

“(1) data of one more type are ingested; (2) algorithmic methods use ingested data to forecast the occurrence of crime in some domain of interest; and (3) police use forecasts to inform strategic and tactical decisions in the field. A primary goal of predictive policing is to reduce uncertainty so that police can approach the allocation of resources in an optimal manner. The theory is that an optimal allocation of police resources has a better chance at disrupting opportunities for crime before they happen.”⁵¹

Big Data Policing is mainly set up accordingly. *Big Data Policing* also revolves around algorithmic predictions that are intended to

⁴⁸ NAKAMOTO, Satoshi (a pseudonym!), «Bitcoin: A Peer-to-Peer Electronic Cash System» (2009) p. 1, [Date of consultation 14. 12. 2021], Access: <https://bitcoin.org/bitcoin.pdf>.

⁴⁹ For an overview on corresponding (pilot) projects in Germany see RADEMACHER, Timo, «Predictive Policing im deutschen Polizeirecht», *Archiv des öffentlichen Rechts*, 3/142 (2017) p. 369.

⁵⁰ For a similar description see the coalition agreement between CDU Hessen and BÜNDNIS90/DIE GRÜNEN Hessen for the 20th legislative term, line 2535 ff., where it can be read: “New tools such as specialized data processing systems that pool and analyze existing information from police databases can be of great benefit in addressing current policing challenges.”

⁵¹ BRANTINGHAM, Jeffrey, «The Logic of Data Bias and Its Impact on Place-Based Predictive Policing», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 473.

enable authorities to prevent crimes as effectively and efficiently as possible – in this logic, “ideally” and with regard to specific persons.⁵² The technological difference to Predictive Policing lies in the quantity and quality of data sets that go into Big Data Policing. As the name suggests, it processes large, seemingly disjointed data sets that traditional analysis tools (including the human brain) were no match for.

Behind the scenes, big data policing is likely to be more far reaching than simple predictive policing, since it is directed not only at public authorities, but also at citizens. As *Brennan-Marquez* has aptly noted, Big Data Policing generates “a social order – a surveillance society – in which people constantly monitor and curate the data-trails they leave behind in everyday life.”⁵³ In other words: the more (data) intensive predictions on behavior and crime inhibition turn out to be, the more the internalization of this crime inhibition project is promoted and the more the external (sovereign/public) and internal (private) law enforcement complement each other. In this respect, a return to *Foucault’s* panoptism seems of interest. Because, by committing himself to the project of technological crime inhibition through risk surveillance of others, the individual does so at the price of being surveilled themselves, accepts this and thus becomes the enforcer of the power structures that lie beneath this surveillance project.⁵⁴

The fact that the algorithms, on which predictive or big data policing is based – as described by the relevant nomenclature – make “decisions”⁵⁵ (especially because they establish correlations between seemingly unrelated data sets), is considered sufficient by quite a few voices, especially in the USA.⁵⁶ This is considered to be the case, even

⁵² BRENNAN-MARQUEZ, Kiel, «Big Data Policing and the Redistribution of Anxiety», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 487.

⁵³ BRENNAN-MARQUEZ, Kiel, «Big Data Policing and the Redistribution of Anxiety», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 487.

⁵⁴ Euphemistically, this is called self-documentation, the other side of which is *Foucault’s* panopticon of the self and a moment of the exercise of power. On this see MAU, Steffen, *Das metrische Wir: Über die Quantifizierung des Sozialen*, Berlin: Suhrkamp, 2017, p. 249 ff. – For more see HAN, Byung-Chul, *Psychopolitik: Neoliberalismus und die neuen Machttechniken*, Frankfurt a.M.: Fischer, 2015, p. 84.

⁵⁵ Anthropomorphisms dominate the social construction of AI today, even as they eclipse the technical and algorithmic idiosyncrasies of AI. It is at least misleading to think that AI systems make decisions, since they can actually only generate output.

⁵⁶ As here HENDERSON, Stephen E., «A Few Criminal Justice Big Data Rules», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 527.

if these decisions are not comprehensible or explainable.⁵⁷ Simply put, it should be possible to keep the algorithms – in contrast to traditional prediction models – deliberately “atheoretical”.⁵⁸

In other words, and with the corresponding *termini technici*: so-called “*opaque AI*” is considered acceptable and “*explainable AI*”⁵⁹ is considered negligible.⁶⁰ It seems reasonable to see in this an increased faith in technology and AI, which in turn is fueled by a loss of confidence in human analytical and decision-making capabilities. In any case, this perspective of things demonstrates the vision of an AI-supported organization and control of human sociality and interactivity. This vision is particularly powerful in the context of criminal law. After all, who would want to deny that effective and efficient crime prevention is desirable from the point of view of society as a whole, as well as from the individuals (e.g., the notorious victim’s) point of view?

2. OBJECTIVITY, NEUTRALITY AND COHERENCE IN THE APPLICATION OF CRIMINAL LAW

Given that we have previously spoken of the promise that AI is capable of effectively and efficiently minimizing the possibilities of committing crimes, we should now go into the further promise that the introduction of AI into (criminal) legal decision-making will make it possible to vouch for its objectivity, neutrality and coherence. The idea that decisions on the imposition of pre-trial detention due to the risk of repetition, on early release from prison due to a positive social prognosis or on sentencing could be made on the basis of *algorithmic risk assessment* may sound outrageous in the German Republic of Judges, due to the seemingly associated encroachment on judicial independence. In

⁵⁷ Although this is strongly – and in my opinion rightly – disputed by computer scientists or legal scholars. Critically for instance LIU, Han-Wei; LIN, Ching-Fu; CHEN, Yu-Jie, «Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability», *International Journal of Law and Information Technology*, 2/27 (2019).

⁵⁸ As literally stated by BERK, Richard; HYATT, Jordan, «Machine Learning Forecasts of Risk to Inform Sentencing Decisions», *Federal Sentencing Reporter*, 4/27 (2015) p. 223.

⁵⁹ Which can also be conceived and designed as “self-explanatory AI”.

⁶⁰ As stated for instance by HENDERSON, Stephen E., «A Few Criminal Justice Big Data Rules», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 527.

the USA, however, this is already common practice, approved by the highest courts.

This is exemplified by *State v. Loomis*, a decision of the Supreme Court of Wisconsin.⁶¹ The accused *Eric Loomis*, who had multiple prior criminal convictions, was suspected of being the driver of a “drive-by-shooting”. Subsequently, he pleaded guilty to eluding an officer and did not contest the charge of operating a vehicle without its owner’s consent. For this he was sentenced to six years in prison. This draconian sentencing, as the sentencing court openly acknowledged,⁶² was due in part to the fact that *Loomis* had been diagnosed as having an abysmal social record and a high recidivism rate. And namely by COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), a proprietary (vulgo secret) algorithm developed and distributed by Northpointe, Inc. COMPAS calculated the accused’s pre-trial and general risk of recidivism as well as his risk to the community based on a complex analysis of a 137-item questionnaire and the accused’s public criminal record. *Loomis* appealed the decision on the grounds that it violated his right to due process. In particular, he challenged that he could not review the algorithmic processes because they were protected as trade secrets; that there was no individual penalty assessment because COMPAS worked with generalizing group data; and that the algorithm also processed the gender of the person(s) being assessed and thus an inadmissible variable, as it was gender discriminatory. – The Supreme Court of Wisconsin rejected the appeal, and the U.S. Supreme Court also ultimately did not accept the case for decision, after previously asking the U.S. federal government to comment on it. Of significance was the decision, that individuals should not enjoy a “right to explanation” of an algorithmic risk prediction as long as they can oversee its input and are informed of its output. Access to the throughput, e.g. why individual data blocks are weighted and how, was thus in principle legally denied to *Loomis*. The use of general group data and the inclusion of gender in the algorithmic risk prediction were also not objected to because this improved its accuracy and did not have a discriminatory objective. Nevertheless, the Supreme Court

⁶¹ *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016) 754 (US).

⁶² For further details, in particular on the (in Wisconsin obviously procedurally admissible) use of a broader crime suspicion, which was added to the files, cf. the summary of the process in *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016) 754 (US).

of Wisconsin was anxious to clarify that a judge may only take into account a corresponding algorithmic risk forecast, but may not regard it as binding.

State v. Loomis is received controversially and mostly negatively in the academic literature – inside and outside the USA.⁶³ However, this should not obscure the fact that algorithm-assisted legal decision-making is now a practice sanctioned by the highest courts in the USA. The causes (but note: not the reasons) for this are numerous. If one detaches oneself from the jurisdiction-specific analysis that is actually indicated,⁶⁴ the recurring motives of today's AI justification narrative can be found quickly. According to this narrative, the use of intelligent algorithms should make criminal justice decision-making processes more effective and efficient. More accurate risk predictions should be made and freed up resources should be able to be used elsewhere, e.g. in rehabilitation programs.⁶⁵ And while detention and sentencing decisions are dismissed as gut decisions, and even seen as *black art*,⁶⁶ algorithmic-based decision making is praised for apparently being able to minimize the influence of biases, prejudices, and idiosyncrasies.⁶⁷

⁶³ Cf. for instance BERIAIN, Iñigo De Miguel, «Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling», *Law, Probability and Risk*, 1/17 (2018) p. 45; DESKUS, Cassie, «Fifth Amendment Limitations on Criminal Algorithmic Decision-Making», *NYU Journal of Legislation and Public Policy*, 1/21 (2018) p. 237; LIU, Han-Wei; LIN, Ching-Fu; CHEN, Yu-Jie, «Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability», *International Journal of Law and Information Technology*, 2/27 (2019) p. 122 ff.; HUQ, Aziz Z., «Racial Equity in Algorithmic Criminal Justice», *Duke Law Journal*, 6/68 (2019) p. 1081. – Cf. also OSTERMEIER, Lars, «Der Staat in der prognostischen Sicherheitsgesellschaft», in PUSCHKE, Jens; SINGELNSTEIN, Tobias, *Der Staat und die Sicherheitsgesellschaft*, Wiesbaden: Springer VS, 2017, p. 103.

⁶⁴ To give just a few key words: A deep-seated racism, which is still fueled today at the highest political level, finds expression not least in the phenomenon of so-called “mass incarceration” and causes a – sad but understandable – feeling of resignation that the U.S. criminal justice system can no longer be “saved” by conventional means.

⁶⁵ See for instance BOTNICK, Claire, «Evidence-based Practice and Sentencing in State Courts: A Critique of the Missouri System», *Washington University Journal of Law & Policy*, 1/49 (2015) p. 166.

⁶⁶ As strikingly stated by CHIAO, Vincent, «Predicting Proportionality: The Case for Algorithmic Sentencing», *Criminal Justice Ethics*, 3/37 (2018) p. 238.

⁶⁷ For a critical approach to this see MARTINI, Mario; NINK, David, «Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz», *Neue Zeitschrift für Verwaltungsrecht – Extra*, 10/36 (2017) p. 9.

Not the law (due to its postulated inherent properties) or the legal staff (due to its formation), but algorithms, in other words, are supposed to guarantee (and in the U.S.: save) the objectivity, neutrality and coherence of the application of law. Here we once again encounter the shift in trust already noted above, away from trust in people and toward trust in high technology. It is exactly this shift of trust that occurs for almost all humans, since, here the practitioners of the law are also perceived as a risk in the matter, namely as a risk to the objectivity, neutrality and coherence of the application of the law.⁶⁸

III. THE PROMISES OF AI FROM THE PERSPECTIVE OF CRIMINAL LAW THEORY

The foregoing illustrates that criminal law theory can no longer ignore the promises of AI. These have developed too much social traction and explosive power to be ignored, and are in the process of becoming effectively entrenched in the criminal justice system. A trivialization of AI (see 1. below) as well as a subversive doubting of its promises (see 2. below) therefore does not seem to be a viable way to face the current challenges.

1. CRIMINAL SOCIOLOGICAL TRIVIALIZATION?

A first, almost involuntary reaction to the promises described under I. and II. is to downplay them in their effects and potency and therefore to meet them with empathetic forbearance. Especially because many of the developments mentioned in the foregoing originate in the USA (or even in China) and therefore “cannot” develop any significance for Europe and Germany. Moreover, AI is not infrequently downgraded to a “normal” technological innovation that may be capable of transforming the criminal justice system to the usual extent and according to the usual pattern, but not of revolutionizing it (comparable, for example, to the introduction of modern passenger cars,

⁶⁸ Similar developments seem to be taking place in China, although there the center of discussion may be the control exercised by apparently dependent judges. Instructive on this matter: MENG, Yu; GUODONG, Du, «Why Are Chinese Courts Turning to AI?», *The Diplomat* (2019) [Date of consultation 14.12. 2021], Access: <https://thediplomat.com/2019/01/why-are-chinese-courts-turning-to-ai/>.

which made an actual criminal traffic law necessary). All that can then be expected (and, depending on one's point of view, feared) is that the familiar developments of modern criminal law will continue, such as the forward shifting of criminal liability or the protection of collective legal interests. As a consequence, a fundamental questioning of criminal law could not be envisaged in terms of criminal law theory.

It could seemingly (!) be said, for instance: Even if *arguendo* the promises of smart contracts were to be taken at face value, this would merely lead to shifts and adjustments in where, when and how criminal (fraudulent) energy comes to bear.⁶⁹ In particular, an increase in cybercrime could be expected. Instead of directly deceiving a buyer into believing that one can dispose of the item to be sold as the seller, the latter will probably manipulate the database ("blockchain") that verifies the power of disposal and thus carry out this deception indirectly. Furthermore, active attacks on the computer programs contouring and executing a smart contract as well as the exploitation of their vulnerabilities are conceivable.⁷⁰ If such developments create gaps in criminal liability, it would be up to the legislator to close them. In doing so, the legislator will rely on the use of preemptive offenses aimed at protecting collective rights (such as the integrity and correctness of decentrally organized databases). – Furthermore, with regard to predictive or big data policing, it could be argued that algorithmic crime predictions only work, if at all, for specific crime areas (such as burglary and drug-related crime, which can be prevented by locally targeted police patrols).⁷¹ The majority of the "fight" against crime would therefore have to be carried out by classical means.

⁶⁹ Parallels could be drawn here from DURKHEIM, Emile, *Die Regeln der soziologischen Methode*, Frankfurt a.M.: Suhrkamp, 1984, p. 86 and 156.

⁷⁰ An illustrative example of this is provided by the so-called DAO-Hack. On this see HECKMANN, Jörn, «DAO- Hack: smart contracts auf dem rechtlichen Prüfstand», *Computer und Recht*, 9 (2016) R99.

⁷¹ On the spatial dimension of predictive policing see for instance STRAUBE, Till; BELINA, Bernd, «Policing the Smart City: Eine Taxonomie polizeilicher Prognoseprogramme», in BAURIEDL, Sybille; STRÜVER, Anke, *Smart City – Kritische Perspektiven auf die Digitalisierung in Städten*, Bielefeld: transcript Verlag, 2018, p. 223. Predictive policing is not limited to spatial crime forecasting, however. The – probably highly unfruitful – analysis of airline passenger data also falls under the heading of predictive policing. On this see *Süddeutsche Zeitung*-Online 24.04. 2019, "Überwachung von Flugpassagieren liefert Fehler über Fehler".

In addition, predictive or big data policing could give rise to new forms of crime, which would then have to be countered with a correspondingly modern criminal justice system. One might think of so-called *oracle attacks*, which criminals use to gain knowledge about the predictions of the corresponding predictive or big data policing software in order to adapt their criminal behavior accordingly (e.g., by breaking in exactly where the algorithm does not suspect a break-in). – Finally, the influence of AI in criminal law decision-making can also be trivialized. For example, by arguing that sentencing is too complex for machines to handle; or that in *State v. Loomis*, the Supreme Court ruled that algorithmic risk assessments may only be used to support and prepare independent judicial decisions, but may in no way bind or prejudice the latter.⁷²

As engaging and comforting as criminal sociological trivializations of AI may sound, and as important as it is to keep in mind *de lege lata et ferenda* adaptation and displacement movements triggered by AI, criminal law theory must nevertheless deal more fundamentally with the promises of AI. Otherwise, it would involuntarily become their stirrup holder and miss the decisive “initial” phases of the upcoming developments. The more the weaknesses of AI are emphasized by criminal law theory, the more incentives are created for such weaknesses to be closed by technological progress. The trivialization of AI thus led, in substance, to a development spiral that creates social facts without being aware of their normative foundations.

For example: by correcting faulty code, *exploits* can be made increasingly impossible, *hacks* can be made more difficult by increasingly better firewalls, or *oracle attacks* can be anticipated (by predicting the abusive prediction of the regular predictions, with the consequence that patrols are carried out where they should not be). The importance of keeping such developments in mind from the outset in terms of criminal law theory is again demonstrated by *State v. Loomis*. For even if the decision-making process in criminal law is “only” to be prepared and supported by algorithms, this already generates – depending on

⁷² On this the referenced decision in Fn. 61 above and in addition see BERIAIN, Iñigo De Miguel, «Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling», *Law, Probability and Risk*, 1/17 (2018) p. 47; KATYAL, Sonia K., «Private Accountability in the Age of Artificial Intelligence», *UCLA Law Review*, 54/66 (2019) p. 86.

the layout – anchor effects⁷³ and a “compliance” pressure that should not be underestimated. If judges were allowed to overrule the algorithmic social prognosis with their own prognosis, this would lead to an actual prevalence of algorithmic methods (keyword: saving of working time; fear of negative reactions if one’s own prognosis turns out to be wrong and e.g. the person released contrary to the algorithmic “advice” immediately recidivates), so that sentencers are enabled to shift responsibility (“blame shifting”) to algorithms (according to the motto: “It’s not me, but the machine that is responsible!”). These points must be well considered, which cannot be achieved by trivializing AI.

2. UNDERMINING THROUGH INFORMATION TECHNOLOGY?

It is therefore more important to critically examine the promises of AI from the inside out and question their durability in terms of information technology, than to trivialize them from a perspective of criminal law theory. Admittedly, the challenges of AI cannot be met conclusively in this way, but only temporarily, if at all, in terms of criminal law theory. Here we encounter AI as the many-headed Hydra; as soon as one head is cut off, others grow back.

Nevertheless, it is necessary to express strong doubts about the promise of objectivity and neutrality of today’s AI systems. Unrestrictedly neutral algorithms are difficult to imagine. Moreover, data-processing forecasts have to struggle with the so-called “*bias in, bias*” problem.⁷⁴

a) Algorithmic Normativity

Data-processing algorithms work with self-learned or human determined criteria, e.g. when the age of an offender is given particular importance in calculating his risk of recidivism.⁷⁵ These algorithms become problematic, when their criteria are subject to conscious or unconscious normative targets – especially in the case of proprietary,

⁷³ On this and with regard to sentencing in general see STRENG, Franz et al., *Strafgesetzbuch*, Baden-Baden: Nomos, 5. Ed., 2017, § 46 m.n. 3; TRAUT, Marcus; NICKOLAUS, Christoph, «Der Ankereffekt: Schattenseiten im Strafprozess», *Strafverteidiger Forum*, 12 (2015) p. 485.

⁷⁴ Following MAYSON, Sandra G., «Bias in, Bias out», *Yale Law Journal*, 8/128 (2019) p. 2218.

⁷⁵ On this see BERK, Richard, «Algorithmic criminology», *Security Informatics*, 5/2 (2013) p. 4.

i.e. non-verifiable systems. As *Berk/Hyatt*, for instance, explain with regard to recidivism risk predictions – although linguistically somewhat dressed up, but in substance with thankful openness:

“Many criminal justice stakeholders will treat false negatives as more costly than false positives. When this policy preferences applies, the standard of statistical proof necessarily will be lower for forecasts of homicide. The intent is to not release an individual who will commit a homicide and, in trade, to accept a larger number of false positives.”⁷⁶

This means that such a normatively oriented algorithm “approvingly accepts” (or more precisely: those behind the algorithm and those using it must approve of) leaving persons who are not dangerous (so-called false positives) in custody in order to prevent persons who are dangerous (so-called false negatives) from being erroneously classified as not dangerous and subsequently released from custody. It is obvious that such programming of criminal justice decisions – which is unfortunately not so far removed from German criminal law either, as the debates about preventive detention, for example, show⁷⁷ – is by no means neutral and objective, but rather highly politically and normatively charged.

In the literature, however, this is immediately turned into a positive aspect with critical intent, namely by linking it to the promise that AI forces the disclosure of normatively open and therefore politically fixable objectives.⁷⁸ Transparency thus becomes not only a basic requirement, but a normative good of the use of AI in the criminal justice system.

An example of this is *Chiao’s* recent (theoretical) discussion of algorithmic penalty assessment.⁷⁹ According to *Chiao*, an algorithm

⁷⁶ HYATT, Jordan, «Machine Learning Forecasts of Risk to Inform Sentencing Decisions», *Federal Sentencing Reporter*, 4/27 (2015) p. 223.

⁷⁷ Cf. BOETTICHER, Axel et al., «Zum richtigen Umgang mit Prognoseinstrumenten durch psychiatrische und psychologische Sachverständige und Gerichte», *Neue Zeitschrift für Strafrecht*, 9/29 (2009) p. 478 (p. 479 with further references).

⁷⁸ As here for instance BERIAIN, Iñigo De Miguel, «Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the *Wisconsin v. Loomis* ruling», *Law, Probability and Risk*, 1/17 (2018) p. 48 with further references.

⁷⁹ See CHIAO, Vincent, «Predicting Proportionality: The Case for Algorithmic Sentencing», *Criminal Justice Ethics*, 3/37 (2018) p. 238 ff.

should evaluate the appropriateness of retrospective sentencing rather than prospective risk; that is, it should calculate not how dangerous a defendant is, but what sentence other judges in a given jurisdiction would impose in a comparable case. In this way, the judge who is actually called upon to assess the sentence is to be given a concrete guideline corridor. This proposal preempts traditional control mechanisms of sentencing (such as professional socialization processes, obligations of the courts to give reasons, and appellate reviews of appropriateness or arbitrariness). The idea is that sentencing is examined in individual cases for its systemic justice and correctness before it is pronounced in a legally binding manner. And since – as *Chiao* of course recognizes – the assessment of punishment is determined by the most diverse, partly antinomic goals and purposes, the algorithm to be applied would have to specify in a binding manner whether and which goals and purposes are to gain determining influence, and which weighting each is to have. With all this, *Chiao* opposes in a critical tradition the intransparency of the subjectively political side of the application of law. The transparency of an objective-politically designed algorithm is to take its place.

The algorithm thus becomes the *bouche de la loi*.⁸⁰ In other words, the belief of enlightenment in the ordering and pacifying power of rationality no longer connects with the human but with the machine-like *Subsumtionsautomat*,⁸¹ which applies the norms of the legislator objectively, neutrally, and coherently.

b) Bias in, bias out

Moreover, the neutrality and objectivity of today's AI systems must be called into doubt by the *bias in, bias out*. This directly relates to the operation of AI-based (crime, recidivism risk, or *Chiao's* sentencing) forecasts, which draw conclusions about the likelihood of future events by evaluating current data about past occurrences.⁸² If, of course,

⁸⁰ Whether and how closely this figure is to be connected with Montesquieu may be left aside here.

⁸¹ In general on the Figure of the *Subsumtionsautomat* see OGOREK, Regina, *Richterkönig oder Subsumtionsautomat? Zur Justiztheorie im 19. Jahrhundert*, Frankfurt a.M.: Vittorio Klostermann, 1986.

⁸² An instructive case study from the USA is presented by BERK, Richard, «An Impact Assessment of Machine Learning Risk Forecast on Parole Board Decisions and Recidivism», *Journal of Experimental Criminology*, 2/13 (2017) p. 193.

these past events are then prejudiced or reconstructed by prejudiced data sets, the prediction of the future reproduces the prejudices of the past in the present.⁸³ This is of immense relevance, especially in the U.S. discussion. The racial segregation that characterizes the U.S. criminal justice systems is represented by data sets that lead an algorithm to attribute a disproportionately high level of criminal energy to individual young African American males today that will be realized tomorrow, because this population group was disproportionately litigated (arrested, convicted, not released early from prison, etc.) through the criminal justice system yesterday. For which social (e.g. racist) reasons this took place “yesterday” remains algorithmically out of consideration from the beginning.⁸⁴

Such a *bias in, bias out* does not have to be based on an error in the system, but can theoretically also have a system.⁸⁵ The promise of objectivity and neutrality would then contribute to algorithmically whitewashing a criminal justice system that has been tarnished (e.g., through racism), mystifying it as free of domination, and thus legitimizing it in social perception. All this can and should be criticized by a critical theory of criminal law that takes AI into account.

A way out of the information-technologically securitized dominance petrification of a *bias in, bias out* is promised by neutralized training data sets that reject normatively undesirable input factors (such as those that are directly or indirectly related to the skin color of the persons to be evaluated), as well as algorithms that normatively compensate for undesirable biases in the input data, by subtracting them.⁸⁶ Interestingly, this does not sustainably challenge the promise that AI

⁸³ See also, instructively SINGELNSTEIN, Tobias, «Predictive Policing: Algorithmenbasierte Straftatprognosen zur vorrausschauenden Kriminalintervention», *Neue Zeitschrift für Strafrecht*, 1/38 (2018) p. 4.

⁸⁴ In general and critical on this HANNA-MOFFAT, Kelly; MONTFORD, Kelly Struthers, «Unpacking Sentencing Algorithms», in DE KEIJSER, Jan W.; ROBERTS, Julian; RYBERG, Jesper, *Predictive Sentencing*, Oxford: Hart Publishing, 2018, p. 186 ff. With regard to the USA, FERGUSON, Andrew G., «Illuminating Black Data Policing», *Ohio State Journal Criminal Law*, 2/15 (2018) p. 504 speaks of a “black data problem [since data is] racially encoded, colored by the history of real-world policing that disproportionality impacts communities of color.”

⁸⁵ This is brought into the discussion by MAYSON, Sandra G., «Bias in, Bias out», *Yale Law Journal*, 8/128 (2019) p. 2218.

⁸⁶ Also critical on this and with further references: MAYSON, Sandra G., «Bias in, Bias out», *Yale Law Journal*, 8/128 (2019) p. 2218.

could organize an objective, neutral as well as coherent criminal justice system more effectively and efficiently than human decision makers. On the contrary, it is reproduced at a higher level of order and issued as a development goal. Paradoxically, the criticism of the current use of AI in criminal justice outlined above stabilizes and legitimizes the future use of AI that is to be further developed.

The crucial doubts that the promises of AI are the proverbial “hot air” and that an AI-supported criminal justice system will virtually fail because of itself or because expectations are set too high, precisely because algorithms are programmed normatively and learn on the basis of biased data sets, are in the endless far-reaching than they might first appear. These doubts only concern the concrete implementation, but not the fundamental normative ideas of order of an algorithmic guarantee of an effective and efficient protection of legal interests as well as an objective, neutral and coherent application of criminal law. Rather, the demands for objective-political programming and for normative balancing of AI systems sanction and perpetuate these ideals. Ignoring, trivializing or undermining “AI and criminal law” is therefore of no help.

IV. REFLECTION ON FUNDAMENTAL NORMATIVE POSTULATES OF ORDER

In order to work through the transformative aspect of the connection between “AI and criminal law”, a reflection on fundamental normative postulates of order is necessary – since ignoring, trivializing, and undermining are not sufficient; namely, the postulates of order of a desired criminal law as well as those of the desired society that is supposed to produce this criminal law. Since this raises large and major questions, only a few cursory considerations can be made here, but no definitive answers can be promised, so that in the following we must switch to the subjunctive at the crucial points.

1. CRIMINAL LAW AS A LIBERAL PROTECTION OF FREEDOM OR AS A WELFARE-STATE SECURITY LAW?

The real challenge for criminal law theory is that AI *prima facie* adopts – and optimizes – the central promises of criminal law.

If our “conventional human criminal law” (of any theoretical provenance) can only guarantee the protection of legal interests⁸⁷ normatively and counterfactually, because breaches of norms remain the order of the day, then AI wants to achieve the factual prevention or at least the substantial minimization of violations of legal interests in the long term. And can our “conventional human law” only provide a normative and counterfactual guarantee of objectivity and neutrality for legal decision-making, because personal idiosyncrasies and errors of the users of the law remain the “human” rule, with which the legal system has also come to terms (especially for pragmatic considerations),⁸⁸ AI wants to categorically prevent subjectivity, bias, and unequal treatment in the application of the law. If criminal laws would aim at being taken seriously by primary and secondary norm addressees nationwide,⁸⁹ in order to guarantee serious protection of the protected legal interests, and if the law were to take its claim to want to exhaustively guarantee objective, neutral, and coherent decision making seriously, then criminal law theory could not reject the promises of AI on principle. The factuality of (the promises of) AI would, in other words, resolve the end of the counterfactuality (of the promises) of criminal law – and thus the end in terms of the crowning conclusion of criminal law as we know it.

⁸⁷ On this paradigm see HASSEMER, Winfried, *Theorie und Soziologie des Verbrechens*, Frankfurt a.M.: Athenäum, 1973, p. 27 ff.

⁸⁸ By way of example, it should be remembered that the legal protection guarantee in Article 19 section 4 of the German Basic Law only guarantees legal protection by, but not against, the judge, in accordance with the (admittedly controversial) so-called Dürig dogma.

⁸⁹ This qualification challenges us to deal more openly with the factual selectivity of criminal law – not only at the supra-national but also at the intra-national level – to think of the protection of legal interests under criminal law not only as fragmentary but as conceptually selective. After all, the degree of effectiveness that AI promises may quickly turn out to be too expensive for society in real life. For instance, in a Chinese newspaper the AI-based anti-corruption system there, which significantly bears the zero trust maxim in its name, was critically questioned as follows: “Is China’s corruption-busting AI system ‘Zero Trust’ being turned off for being too efficient?” The background is the concern that the Chinese public administration would not be able to cope with a comprehensive prosecution of all identified corruption offenses. Access: <https://www.scmp.com/news/china/science/article/2184857/chinas-corruption-busting-ai-system-zero-trust-being-turned-being> [Date of consultation 14.12.2021].

A fundamental criticism of the entry of AI into the criminal justice system would therefore have to take a critical look not only at its “foreign” objectives, but also at its “own” objectives, which are supposedly only taken over by AI. To defend and justify would be, in other words, nothing less than the “only” normative and counterfactual aspects of criminal justice. These are: the real possibility of criminal violations of legal rights as well as the real possibility of biased, prejudiced, idiosyncratic human law practitioners.⁹⁰ At this point, we encounter the difference between a liberal foundation of criminal law, which dedicates the latter to the protection of freedom, and welfare-state conceptions, which (can) use criminal law to protect security.⁹¹

In order to be able to counter the promises of AI in principle, the dogma of the “protection of legal rights through criminal law” would have to be supplemented from a *liberal perspective* – and since we cannot give definitive answers here, the subjunctive can and must be used in the following. This could be achieved – (partly very!) loosely according to *Haffke*,⁹² *Tiedemann*⁹³ and *Prittwitz*⁹⁴ –, by dedicating criminal law fundamentally to the protection of liberty, to which other (possibly legitimate, but then not in this sense criminal-legal) instruments for the prevention of violations of legal interests are to be seen in contrast to. What is meant by this is the direct protection of a use of freedom (such as the consumption of alcohol free from state supervision and control), which indirectly also enables the abuse of freedom, i.e. the freedom to commit criminal acts (such as negligent or intentional driving under the influence of alcohol).

⁹⁰ E.g. by placing law in the service of discursive, justification- and critique-driven orientation toward objectivity, neutrality and coherence.

⁹¹ This ideal-typical juxtaposition is also found in GÜNTHER, Klaus, «Bedrohte individuelle Freiheiten im aufgeklärten Strafrecht – Welche Freiheiten?», *Kritische Justiz*, 4/49 (2016) p. 520.

⁹² HAFFKE, Bernhard, «Die Legitimation des staatlichen Strafrechts zwischen Effizienz, Freiheitsverbürgung und Symbolik», in SCHÜNEMANN, Bernd et al., *Festschrift für Roxin zum 70. Geburtstag*, Berlin: De Gruyter, 2001, p. 965.

⁹³ Concisely summarized in TIEDEMANN, Klaus, *Wirtschaftsstrafrecht: Einführung und Allgemeiner Teil*, 5. Ed., Munich: Vahlen, 2017, m.n. 228.

⁹⁴ PRITTWITZ, Cornelius, «Strafrecht als propria ratio», in HEINRICH, Manfred et al., *Strafrecht als Scientia Universalis: Festschrift für Claus Roxin zum 80. Geburtstag*, Berlin/New York: De Gruyter, 2011, p. 23 ff.

Only in this way could the non-commitment of a crime continue to be evaluated as a free decision for and its commission as a free decision against the law from an *individual perspective*. In a (hypothetical) world in which AI makes crimes *eo ipso* or *de facto* impossible, there can no longer be any question of this freedom, the famous ability to act differently, even if only as a fiction necessary for a liberal community.

Moreover, from a *social point of view*, even up to now (i.e. without AI) ways and means were conceivable with which violations of legal rights could be made *eo ipso* impossible or at least *de facto* drastically (and draconically) minimized.⁹⁵ The potentials of AI culminate in this sense what was also conceivable up to this point in the context of a rigid welfare-state, namely administrative or “technical prevention” (Hassemer⁹⁶), in particular a kind of “technological paternalism” (Hilgendorf).

If the counterfactual of criminal law is to be preferred to the factual of more effective instruments of crime prevention, a liberal theory of criminal law would have to deal more openly with balancing liberties. The loss of freedom of the many, who have to submit to rigid non-criminal measures, although they can also be reached by normative (criminal) measures, would then have to weigh more heavily than the possible loss of freedom of the few, whose legal interests are violated by those who were not normatively (criminally) addressable. In terms of penal constitutional law: penal and prohibition norms would have to be theorized as *prima ratio* of the protection of liberty, because and if a more rigid (e.g., administratively supervisory, regulative, or technically intervening) protection of legal rights would, on balance, not be necessary or appropriate, since it would excessively burden individuals or the general public.

⁹⁵ In criminal economic law, for example, through rigid administrative supervision and regulation of economic operators. Or, in criminal traffic law, through technical intervention against drivers. One might think here of mandatory alcohol tests before driving, linked to an immobilizer. This is not (!) science fiction either, but is being considered concretely (!) at EU level. On this see <https://www.europarl.europa.eu/news/en/press-room/20190410IPR37528/parliament-approves-eu-rules-requiring-life-saving-technologies-in-vehicles> [Date of consultation 14.12. 2021] with further references.

⁹⁶ HASSEMER, Winfried, «Aktuelle Perspektiven der Kriminalpolitik», *Strafverteidiger*, 6 (1994), p. 333 ff. (p. 336).

Illustrative of this is the reciprocal freedom balance that is opened up, for example, in the case of high-risk technology (think, for example, of commercial nuclear power). Here, the possible loss of freedom of the many (which would have to be feared, for example, in the case of a nuclear total meltdown) outweighs the loss of freedom of the few, who are subjected to rigid non-criminal measures (such as close-meshed state control of the operators of nuclear power plants). The painful price of this reconstruction is that the victims of crime have to accept their real loss of freedom in order to secure for the others their virtual other preservation of freedom. To put it bluntly: the parents of a child who was run over because a driver negligently relied on the approved “autopilot” of his car must be told openly that the ban on such “autopilots” was socially too “expensive”, e.g. too anti-innovation, and also not in the sense of the many who use these “autopilots” in a traffic-friendly way to realize their freedom.

A liberal criminal law to be legitimized in this way could be positioned against the promise of an effective as well as efficient protection of legal rights by AI. Of course, this would require a great deal of courage in terms of criminal policy and penal theory. After all, the public and private third-party and self-monitoring that is necessary to enable predictive and big data policing would have to be classified as excessively invasive. And the virtual gain of freedom of the many, who are not exposed to surveillance, would have to be preferred to the real loss of freedom of the few, who become victims of criminal acts, which could (probably) have been prevented by the use of AI.

The necessary trade-offs and balancing of freedoms will have to be made on a sector-specific basis in the future as well. The penal law as a liberal law for the protection of liberty would therefore not have to be applied everywhere and in general. The liberal theory of criminal law, however, is required to weigh freedom and the protection of rights in an open process.

2. WHICH CRIMINAL LAW FOR WHICH SOCIETY?⁹⁷

The pressure to show one’s colors also arises as soon as one takes a look at AI’s underlying normative postulates of order, which are easily

⁹⁷ This formulation of my question originates from discussions in the circle of colleagues in Frankfurt and I ultimately owe it to *Klaus Günther*.

obscured by many superficial justificatory narratives (more effective and efficient; more objective, neutral, and coherent). In this respect, criminal legal theory must open up to social or political theory and take a stand on fundamental issues (in particular, the social status of trust and openness to the future).

As has been shown (above II.), the use of AI in criminal justice is and would be an expression of a fundamental loss of interpersonal trust. Any other person (including the user of the law) is no longer to be trusted; but rather, to be managed as a potential danger and as a risk to be monitored and whose future behavior is to be algorithmically anticipated. In its own logic, this amounts to a generalization of automated suspicion – or, in terms of criminal procedure, of automated and general initial suspicion.⁹⁸

In order to oppose the promises of AI in principle, the presumption of innocence would have to be held higher in social-criminal theory, namely more comprehensively as a counterfactual interpersonal presumption of trust and thus, e.g., would have to, quite enlightenment-like – as *Hruschka* has clearly worked out –, be reconstructed as “everyone’s dignity”.⁹⁹ This presupposes the closing of ranks with a social philosophy that does not regard trust as a mere mechanism for reducing social complexity (*Luhmann*), but as functionally valuable in order to prevent a slide into a trustless surveillance society with an always potentially authoritarian and oppressive character.¹⁰⁰

It is no less challenging to meet the ideal of an *end of history*.¹⁰¹ The functioning of AI proclaims, as we have seen (above II.), in the matter a kind of end of history, simultaneously a closure of the future. After all, conclusions about the future are drawn from the past, which

⁹⁸ BRENNAN-MARQUEZ, Kiel, «Big Data Policing and the Redistribution of Anxiety», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 488 therefore aptly diagnoses that under the impression of predictive and big data policing, the constitutional figure of initial suspicion, which legitimizes state encroachments on fundamental rights, is being worn away (beyond recognition).

⁹⁹ HRUSCHKA, Joachim, «Die Unschuldsvermutung in der Rechtsphilosophie der Aufklärung», *Zeitschrift für die gesamte Strafrechtswissenschaft*, 2/112 (2000) p. 285.

¹⁰⁰ Whether the concept of trust as developed by PARSONS, Talcott, *Politics and Social Structure*, New York: Free Press, 1969, which represents a political concept of order in response to the Hobbesian situation, can be used for this purpose does not need to be further explained here.

¹⁰¹ Downright classic: FUKUYAMA, Francis, *The End of History and the Last Man*, New York: Free Press, 1992.

– for better or worse – leads to a petrification of yesterday in today and inhibits the dynamic development of tomorrow. The promise is that an abnormal future can actually be prevented by means of AI.

At first glance, this is also the aim of traditional criminal law theory. The individual is relieved of the concern about future violations of legal rights because and by guaranteeing the future existence of these legal rights in the present by means of criminal law. The difference is once again that AI promises a factual and criminal law a counterfactual or normative *end of history*. This has consequential effects. In an ideal (utopian or dystopian) AI world, dissidence and resistance against the *status quo* reproduced as *status quo ante* are not only futile (because and if they are made *eo ipso* or *de facto* impossible), but ideally also inconceivable (especially because and if the individual makes himself the executor of his own subjection to a panoptic form of rule).

Whoever wanted to reject this closure of the future and advocate its opening could find the intrinsic and added value of the counterfactual of criminal law in the *de facto* admission of dissidence and resistance. Deviant behavior should then no longer be labeled solely as a violation of legal rights that needs to be prevented, but should at least also be recognized as potential (objective or subjective) criticism of the *status quo* (one need only think of the criminal prohibition of homosexual intercourse, which was brought down not least by continuous acts of resistance, namely by acts of norm-breaking). This meant that in the discussion of the violation of norms, the criminal reaction would have to be justified again and again with good reasons and that it could not, for example, be presented as “natural”. Criminal law would thus have to be conceived as an evolutionary and discursive practice of (human) justification and critique, and the temporal contingency of criminal law would have to be accepted in order to contrast the open future of criminal law with the closed future of AI.

OUTLOOK

This contribution advocates for viewing AI as a social construct whose practice is capable of transforming our fundamental social and (criminal) legal notions of order. In light of concrete criminal law applications, AI as a transformative technology should no longer be ignored

or trivialized in terms of criminal law theory. And since even internal information-technological doubts about the performance of AI only promote further spirals of development, it is necessary to critically question the normative postulates of order that are programmed into AI. In particular, it must be questioned what the price is of AI's promise, that it is able to guarantee a more effective and efficient protection of legal rights and a more neutral, objective and coherent application of criminal law (keyword: victory of the welfare-state security and *zero trust* paradigm in a Big Data-based surveillance society). Conversely, a non-ideal theory of criminal law would – as can only be stated here in conclusion – oversimplify its opposition to the entry of AI into the criminal justice system, by countering it with (normative) ideals (keyword: criminal law as freedom protection law; law as discursive practice of justification and criticism).¹⁰² For this would presuppose a particularly sophisticated criminal law that is (and should not be “merely”) free of all authoritarian and oppressive borrowings.¹⁰³ Anyone who wants to present this as a desirable goal, but one that is not very realistic in terms of critical intent¹⁰⁴, and who must assume that the digitization of society will in fact continue to advance, will ultimately have to think about how criminal law, which in principle must be liberal, democratic and based on the rule of law, can be supplemented by AI systems in such a way that AI does not corrupt criminal law in an authoritarian manner on the one hand, and on the other hand can free it from illiberal tendencies that are contrary to the rule of law and authoritarian.¹⁰⁵

¹⁰² To hint at the famous methodological debate between ideal and non-ideal theory formation for criminal law theory, as set off by *Rawls*.

¹⁰³ As such in general, without a specific reference to AI, NAUCKE, Wolfgang, *Negatives Strafrecht*, Berlin: Lit Verlag, 2015, p. 114 ff.

¹⁰⁴ Ideal theory formation is exposed to the criticism that it stabilizes the status quo. A non-ideal theory formation does not let itself be driven by current developments, but takes note of them and wants to understand them in order to make them the point of reference for critical reflections.

¹⁰⁵ E.g. by using AI for educational purposes in order to show judges their prejudices, etc. On this impressively SOMMER, Ulrich, «Psychologie der richterlichen Entscheidungsfindung», *Zeitschrift für Rechtspolitik*, 2/50 (2017) p. 60, who considers scientific research into judicial prejudices to be necessary because it is still “taboo” in this country. In terms of legal methodology and sentencing methodology, the work of HRUSCHKA, Joachim, «Rechtsanwendung als methodisches Problem», *Archiv für Rechts- und Sozialphilosophie*, 4/50 (1964), p. 485 (especially p. 498) continues to be extremely worth reading.

In this (non-ideal and critical) sense, AI would then not have to be designed as the end of criminal law (neither in the sense of the dying death of a liberal criminal law of liberty nor in the sense of the crowning conclusion of a welfare-state criminal law of security), but as a building block of a criminal law of the (near) future to be designed today, which at the same time shows itself to be open to technology¹⁰⁶ and humane and thus, continues to be reasonable in a modern sense. In all of this, it is important to heed *Joachim Hruschka's* great legacy – which is not always easy to implement, especially in view of AI – that criminal law must always remain *intellectually honest*, i.e. it must use rational argumentation and pedantically avoid incantations and magic formulas.¹⁰⁷

REFERENCES

- BALKIN, Jack B., «The Path of Robotics Law», *California Law Review Circuit*, 6 (2015) p. 45-60.
- BERIAIN, Iñigo De Miguel, «Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the *Wisconsin v. Loomis* ruling», *Law, Probability and Risk*, 1/17 (2018) p. 45-53.
- BERK, Richard; HYATT, Jordan, «Machine Learning Forecasts of Risk to Inform Sentencing Decisions», *Federal Sentencing Reporter*, 4/27 (2015) p. 222-228.

¹⁰⁶ Attempts to close criminal law to technological innovations in law can seem artificial at best. For example, French criminal law now includes the following provision: “Les données d’identité des magistrats et des membres du greffe ne peuvent faire l’objet d’une réutilisation ayant pour objet ou pour effet d’évaluer, d’analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées.” Access, and with reference to where this is to be implemented: <https://www.legifrance.gouv.fr/eli/loi/2019/3/23/JUST1806695L/jo/texte> [Date of consultation 14.12. 2021]. – This raises the concern of making judicial decision-makers algorithmically “transparent”. If this should be associated with the concern that it can be made clear algorithmically that these decision-makers do not apply the law objectively and neutrally, but rather make decisions colored by subjectivity, the question naturally arises as to whether the protection of a myth or mere ideal by criminal law (i.e., the objectivity and neutrality of the application of the law) or of an unquestionable trust in judicial decision-makers represents a legitimate and meaningful legal interest.

¹⁰⁷ As explicitly stated by HRUSCHKA, Joachim, *Strafrecht nach logisch-analytischer Methode*, Berlin/New York: De Gruyter, 1988, p. XVIII.

- BERK, Richard, «An Impact Assessment of Machine Learning Risk Forecast on Parole Board Decisions and Recidivism», *Journal of Experimental Criminology*, 2/13 (2017) p. 193-216.
- BERK, Richard, «Algorithmic criminology», *Security Informatics*, 5/2 (2013) p. 1-14.
- BOETTICHER, Axel et al., «Zum richtigen Umgang mit Prognoseinstrumenten durch psychiatrische und psychologische Sachverständige und Gerichte», *Neue Zeitschrift für Strafrecht*, 9/29 (2009) p. 478-481.
- BOTNICK, Claire, «Evidence-based Practice and Sentencing in State Courts: A Critique of the Missouri System», *Washington University Journal of Law & Policy*, 1/49 (2015) p. 159-180.
- BRANTINGHAM, Jeffrey, «The Logic of Data Bias and Its Impact on Place-Based Predictive Policing», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 473-486.
- BRENNAN-MARQUEZ, Kiel, «Big Data Policing and the Redistribution of Anxiety», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 487-493.
- BYRD, Sharon; HRUSCHKA, Joachim, «Kant zu Strafrecht und Strafe im Rechtsstaat», *Juristen Zeitung*, 20/62 (2007) p. 957-1008.
- CEPEJ, «European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment» v. 3 – 4.12.2018, in particular p. 64 ff. [Date of consultation 14.12. 2021], Access: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.
- CHIAO, Vincent, «Predicting Proportionality: The Case for Algorithmic Sentencing», *Criminal Justice Ethics*, 3/37 (2018) p. 238-61.
- DESKUS, Cassie, «Fifth Amendment Limitations on Criminal Algorithmic Decision-Making», *NYU Journal of Legislation and Public Policy*, 1/21 (2018) p. 237-286.
- DURKHEIM, Emile, *Die Regeln der soziologischen Methode*, Frankfurt a.M.: Suhrkamp, 1984.
- ERTEL, Wolfgang, *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*, Wiesbaden: Springer VS, 2016.
- EUROPEAN COMMITTEE ON CRIME PROBLEMS (CDPC), «Artificial Intelligence and its Impact on CDPC Work: The case of automated driving», CDPC (2018) 14 v. 14. 09. 2018, p. 6 f. [Date of consultation 14.12. 2021], Access: <https://rm.coe.int/cdpc-2018>

-14-artificial-intelligence-and-criminal-law-project-2018-202/16808d6d09.

- FATEH-MOGHADAM, Bijan, «Selbstbestimmung im biotechnischen Zeitalter», *Basler Juristische Mitteilungen*, 4 (2018) p. 205-232.
- FERGUSON, Andrew G., «Illuminating Black Data Policing», *Ohio State Journal Criminal Law*, 2/15 (2018) p. 503-525.
- FERGUSON, Andrew G., *The Rise of Big Data Policing. Surveillance, Race, and the Future of Law Enforcement*, New York: New York University Press, 2017.
- FORST, Rainer; GÜNTHER, Klaus, «Die Herausbildung normativer Ordnungen: zur Idee eines interdisziplinären Forschungsprogramms», in FORST, RAINER/GÜNTHER, KLAUS, *Die Herausbildung normativer Ordnungen*, Frankfurt a.M.: Campus, 2011, p. 11-30.
- FRIES, Martin, «Automatische Rechtspflege», *Rechtswissenschaft*, 4/9 (2018) p. 414-430.
- FRY, Hannah, *Hello World: Was Algorithmen können und wie sie unser Leben verändern*, München: C.H. Beck, 2018.
- FUKUYAMA, Francis, *The End of History and the Last Man*, New York: Free Press, 1992.
- GAEDE, Karsten, *Recht und Strafen für Roboter?*, Baden-Baden: Nomos, 2019.
- GÜNTHER, Klaus, «Bedrohte individuelle Freiheiten im aufgeklärten Strafrecht – Welche Freiheiten?», *Kritische Justiz*, 4/49 (2016) p. 520-534.
- HAFKKE, Bernhard, «Die Legitimation des staatlichen Strafrechts zwischen Effizienz, Freiheitsverbürgung und Symbolik», in SCHÜNE-MANN, Bernd et al., *Festschrift für Roxin zum 70. Geburtstag*, Berlin: De Gruyter, 2001, p. 955-977.
- HAN, Byung-Chul, *Psychopolitik: Neoliberalismus und die neuen Macht-techniken*, Frankfurt a.M.: Fischer, 2015.
- HANNA-MOFFAT, Kelly; MONTFORD, Kelly Struthers, «Unpacking Sentencing Algorithms», in DE KEIJSER, Jan W.; ROBERTS, Julian; RYBERG, Jesper, *Predictive Sentencing*, Oxford: Hart Publishing, 2018, p. 175-196.
- HASSEMER, Winfried, «Aktuelle Perspektiven der Kriminalpolitik», *Strafverteidiger*, 6 (1994) p. 333-336.
- HASSEMER, Winfried, *Theorie und Soziologie des Verbrechens*, Frankfurt a.M.: Athenäum, 1973.

- HECKMANN, Jörn, «DAO- Hack: smart contracts auf dem rechtlichen Prüfstand», *Computer und Recht*, 9 (2016) R99-R100.
- HENDERSON, Stephen E., «A Few Criminal Justice Big Data Rules», *Ohio State Journal of Criminal Law*, 2/15 (2018) p. 527-541.
- HERBERGER, Maximilian, «„Künstliche Intelligenz“ und Recht», *Neue Juristische Wochenschrift*, 39/71 (2018) p. 2825-2829.
- HILDEBRANDT, Mireille, *Smart Technologies and the End(s) of Law*, Cheltenham: Edward Elgar Publishing, 2015.
- HILGENDORF, Eric, «Autonome Systeme, künstliche Intelligenz und Roboter», in BARTON, Stephan et al., *Festschrift für Thomas Fischer*, Munich: C.H. Beck, 2018, p. 99-113.
- HRUSCHKA, Joachim, «Der Standard-Fall der aberratio ictus und verwandte Fallkonstellationen», *Juristen Zeitung*, 10/46 (1991) p. 488-492.
- HRUSCHKA, Joachim, «Die Herbeiführung eines Erfolges durch einen von zwei Akten bei eindeutigen und bei mehrdeutigen Tatsachefeststellungen», *Juristische Schulung* (1982) p. 317-325.
- HRUSCHKA, Joachim, «Kann und sollte die Strafrechtswissenschaft systematisch sein?», *Juristen Zeitung*, 1/40 (1985) p. 1-10.
- HRUSCHKA, Joachim, *Strafrecht nach logisch-analytischer Methode*, Berlin/New York: De Gruyter, 1988.
- HRUSCHKA, Joachim, «Das Strafrecht neu durchdenken!», *Goltdammer's Archiv* (1981) p. 237-250.
- HRUSCHKA, Joachim, «Die Unschuldsvermutung in der Rechtsphilosophie der Aufklärung», *Zeitschrift für die gesamte Strafrechtswissenschaft*, 2/112 (2000) p. 285-300.
- HRUSCHKA, Joachim, «Rechtsanwendung als methodisches Problem», *Archiv für Rechts- und Sozialphilosophie*, 4/50 (1964), p. 485-501.
- HRUSCHKA, Joachim, «Utilitarismus in der Variante von Peter Singer», *Juristen Zeitung*, 6/56 (2001) p. 261-271.
- HUQ, Aziz Z., «Racial Equity in Algorithmic Criminal Justice», *Duke Law Journal*, 6/68 (2019) p. 1043-1134.
- IRSIGLER, Ingo; ORTH, Dominik, «Zwischen Menschwerdung und Weltherrschaft: Künstliche Intelligenz im Film», *Aus Politik und Zeitgeschichte*, 6-8/68 (2018) p. 39.
- KATYAL, Sonia K., «Private Accountability in the Age of Artificial Intelligence», *UCLA Law Review*, 54/66 (2019) p. 56-141.

- KATZ, Yarden, «Manufacturing an Artificial Intelligence Revolution», SSRN (2017) p. 1, [Date of consultation 14.12. 2021], Access: <https://ssrn.com/abstract=3078224>.
- LEWIS, Nell, «Should AI be used to catch shoplifters?», *CNN Business* (2019) [Date of consultation 14.12.2021], Access: <https://edition.cnn.com/2019/04/18/business/ai-vaak-shoplifting/index.html>.
- LIPTAK, Adam, «Sent to Prison by a Software Program's secret Algorithms», *New York Times*, May 1, 2017.
- LIU, Han-Wei; LIN, Ching-Fu; CHEN, Yu-Jie, «Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability», *International Journal of Law and Information Technology*, 2/27 (2019) p. 122-141.
- LUHMANN, Niklas, *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*, 5. Ed., Munich: UVK, 2014.
- MARTINI, Mario; NINK, David, «Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz», *Neue Zeitschrift für Verwaltungsrecht – Extra*, 10/36 (2017) p. 1-14.
- MAU, Steffen, *Das metrische Wir: Über die Quantifizierung des Sozialen*, Berlin: Suhrkamp, 2017.
- MAYSON, Sandra G., «Bias in, Bias out», *Yale Law Journal*, 8/128 (2019) p. 2122-2473.
- MCEWAN, Ian, *Machines like Me*, London: Penguin, 2019.
- MENG, Yu; GUODONG, Du, «Why Are Chinese Courts Turning to AI?», *The Diplomat* (2019) [Date of consultation 14.12. 2021], Access: <https://thediplomat.com/2019/01/why-are-chinese-courts-turning-to-ai/>.
- MEYER, Stephan, «Künstliche Intelligenz und die Rolle des Rechts für Innovation», *Zeitschrift für Rechtspolitik*, 8/51 (2018) p. 233-238.
- NAKAMOTO, Satoshi, «Bitcoin: A Peer-to-Peer Electronic Cash System» (2009) p. 1-8 [Date of consultation 14. 12. 2021], Access: <https://bitcoin.org/bitcoin.pdf>.
- NAUCKE, Wolfgang, *Negatives Strafrecht*, Berlin: Lit Verlag, 2015.
- NIDA-RÜMELIN, Julia; WEIDENFELD, Nathalie, *Digitaler Humanismus*, Munich: Piper, 2018.
- OGOREK, Regina, *Richterkönig oder Subsumtionsautomat? Zur Justiztheorie im 19. Jahrhundert*, Frankfurt a.M.: Vittorio Klostermann, 1986.

- OSTERMEIER, Lars, «Der Staat in der prognostischen Sicherheitsgesellschaft», in PUSCHKE, Jens; SINGELNSTEIN, Tobias, *Der Staat und die Sicherheitsgesellschaft*, Wiesbaden: Springer VS, 2017, p. 101-121.
- PALKA, Silvia; WITTPAH, Volker, «Vertrauen und Transparenz – Blockchain Technologie als digitaler Vertrauenskatalysator», Working Paper of the Institute for Innovation and Technology, 39 (2018) p. 1-15.
- PARSONS, Talcott, *Politics and Social Structure*, New York: Free Press, 1969.
- PRITTWITZ, Cornelius, «Strafrecht als propria ratio», HEINRICH, Manfred et al., *Strafrecht als Scientia Universalis: Festschrift für Claus Roxin zum 80. Geburtstag*, Berlin/New York: De Gruyter, 2011, p. 23-39.
- RADEMACHER, Timo, «Predictive Policing im deutschen Polizeirecht», *Archiv des öffentlichen Rechts*, 3/142 (2017) p. 366-416.
- RAMGE, Thomas, *Mensch und Maschine: Wie künstliche Intelligenz und Roboter unser Leben verändern*, Stuttgart, Reclam, 2018.
- SCHEMMEL, Alexander; DIETZEN, Alexandra, «"Effective Corporate Governance" by Legal Tech & Digital Compliance», in BREIDENBACH, Stephan/GLATZ, Florian, *Rechtshandbuch Legal Tech*, Munich: Beck, 2018, p. 137-152.
- SCHWINTOWSKI, Hans-Peter, «Wird Recht durch Robotik und künstliche Intelligenz überflüssig?», *Neue Juristische Online-Zeitschrift*, 42 (2018) p. 1601-1609.
- SEARLE, John R., «Minds, Brains, and Programs», *Behavioral and Brain Sciences*, 3/3 (1980) p. 417-457.
- SINGELNSTEIN, Tobias, «Predictive Policing: Algorithmenbasierte Straftatprognosen zur vorrausschauenden Kriminalintervention», *Neue Zeitschrift für Strafrecht*, 1/38 (2018) p. 1-9.
- SOMMER, Ulrich, «Psychologie der richterlichen Entscheidungsfindung», *Zeitschrift für Rechtspolitik*, 2/50 (2017) p. 60-62.
- STRAUBE, Till; BELINA, Bernd, «Policing the Smart City: Eine Taxonomie polizeilicher Prognoseprogramme», in BAURIEDL, Sybille; STRÜVER, Anke, *Smart City – Kritische Perspektiven auf die Digitalisierung in Städten*, Bielefeld: transcript Verlag, 2018, p. 223-236.
- STRENG, Franz et al., *Strafgesetzbuch*, Baden-Baden: Nomos, 5. Ed., 2017.

- TIEDEMANN, Klaus, *Wirtschaftsstrafrecht: Einführung und Allgemeiner Teil*, 5. Ed., Munich: Vahlen, 2017.
- TRAUT, Marcus; NICKOLAUS, Christoph, «Der Ankereffekt: Schatten-seiten im Strafprozess», *Strafverteidiger Forum*, 12 (2015) p. 485-492.
- WAGNER, Gerald, «Vertrauen in Technik», *Zeitschrift für Soziologie* *Zeitschrift für Soziologie*, 2/23 (1994) p. 145-157.
- WEIZENBAUM, Joseph, *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt a.M.: Suhrkamp, 1977.
- XANKE, Lisa; BÄRENZ, Elisabeth, «Künstliche Intelligenz in Literatur und Film – Fiktion oder Realität?», *Journal of New Frontiers in Spatial Concepts*, 4 (2012) p. 36-43.
- ZUBOFF, Shoshana, *Das Zeitalter des Überwachungskapitalismus*, Frankfurt a.M./New York: Campus, 2018.