Artificial Intelligence, Smart Orders, and the Problem of Legal and Moral Responsibility

Klaus Günther

The more precise the possible predictions of individual behavior and the more effective the possible behavior modifications through AI become, the more obvious it is to use them for the prevention of deviant behavior. The goal of using technical means to encourage the addressees of the law to behave in accordance with the law or to make unlawful behavior completely impossible without the threat or application of legal sanctions seems to have so much persuasive power that the question of what consequences this would have for the essential characteristics of normative orders, especially the law, is neglected. Above all, we must ask what would then become of a central prerequisite of law: The concept of a person who can be held responsible for their deviant behavior. Would this concept, which is based on the ability to decide individually for or against following the law, be rendered superfluous by smart orders that make deviant behavior impossible? And what consequences would this have for the related ability to question the law itself and its legitimacy?

A. How is it possible to limit instrumental power of AI?

It is not only since the accelerating development of artificial intelligence and its diverse applications that there has been growing concern about the power it can wield and the dangerous consequences for the freedom of the individual and society as a whole. This danger arises from the fact that the "data collected by the operators of such technologies during individual use is increasingly detached from its purpose of collection and transformed into independent information capital."[1] This information capital can be used to generate both economic and state power.

1 Spiros Simitis, in: Simitis, Bundesdatenschutzgesetz Kommentar (BDSG), Einleitung: Geschichte – Ziele – Prinzipien, 14. Aufl. Baden-Baden 2014. Rn. 111.

At the heart of this is a phenomenon characterized by Soshana Zuboff as "instrumental power".[2] It is made up of two components that constitute the opportunity, corresponding to Weber's classic definition of power, to impose one's will on third parties, even against their resistance.[3] This opportunity goes hand in hand with the increasing ability to predict future behavior on the basis of precise observations and to use this predictive knowledge to modify the future behavior of third parties in the pursuit of one's own interests and intentions. This is made possible by the fact that users are incentivized to produce and use as much data as possible, which is collected by the providers and analysed according to aspects determined by them at an increasing distance from the original purpose of collection. The more frequently and comprehensively this happens and the more precise and intensive the technical possibilities of data analysis become, the more effective (and efficient) the possibilities of predicting the future thoughts, feelings, intentions and actions of users on the basis of so-called profiles become. The more precise the predictive knowledge becomes, the greater the opportunities to influence or modify future behavior. Instrumental power is therefore made up of the potential for predicting and modifying behavior that comes with the use of digital media. Unlike the totalitarian power responsible for the human catastrophes of the 20th century, it does not aim to ideologically penetrate body and soul in order to motivate them to behave as desired from within, but rather to modify behavior externally through data-supported observation and prediction knowledge. The more accurate the personality profile distilled from this becomes, the more detailed the external incentives can be specified to which the users will react in the predicted way (so-called micro-targeting).

Instrumental power is driven by a "utopia of certainty" of human behavior, which should become as predictable and controllable as the functional processes of a machine.[4] The main aim is to anticipate possible deviations or errors that could impair the smooth functioning of the machine in accordance with the specified program in order to prevent them from occurring. If this utopia is transferred to human behavior, namely social interactions and communications, it can be applied above all to the practice

2 The concept of instrumental power has been introduced and explained by Shoshana Zuboff, *Das Zeitalter des Überwachungskapitalismus* (2018) 412.

3 Max Weber, Wirtschaft und Gesellschaft, 5. A., Tübingen 1921/1976, S. 29.

4 Zuboff, Überwachungskapitalismus, S. 461-480. See also: Klaus Günther, Die Zukunft der Freiheit in smarten Ordnungen, in: WestEnd – Neue Zeitschrift für Sozialforschung, 17/2 (2020), S. 165-175.

of following norms and rules by their addressees. If norms and rules are there to guide their addressees in their behavior so that they do what the norm requires or refrain from doing what it prohibits, then the avoidance of deviations is part of their meaning. However, since experience teaches us that deviations always occur in practice and that the mutual expectation of following the rules is repeatedly disappointed as a result, instrumental power could be used to perfect rule-following in such a way that the mutual expectation at least approximately reaches the level of certainty. The utopia of certainty would merge with the claim to compliance with normative orders in such a way that it would become the utopia of the certainty of norm compliance. Instrumental power would have to start at the source of deviant behavior in the norm addressee - at her/his subjectivity in relation to the norm's claim to be followed. Norms are addressed to people who must first decide to comply with the norm or at least have decided to do so at some point, even if they may currently do so out of habit or quasi-automatically in many cases.

If instrumentary power is aimed at drying up this source of constant uncertainty in rule-following, any attempt to tame instrumental power normatively, e.g. through legal regulation, could lead to a dilemma: To the extent that legal systems move towards using instrumental power to ensure general rule-following, legal regulation to limit instrumental power would in turn be dependent on this kind of power or, by shutting down the source of uncertainty in norm-following behavior, it would have lost its specific addressee, the legal person with its subjectivity. As I will show below, the reason for this is the freedom to engage in deviant behavior, which is a necessary precondition for normative orders in general and for legal orders in particular.

B. The internal connection between norm, responsibility and deviant behavior

Lon Fuller's 1964 monograph on the "Morality of Law" contains the following remark: "To embark on the enterprise of subjecting human conduct to the governance of rules involves of necessity a commitment to the view that man is, or can become, a responsible agent, capable of understanding and following rules, and answerable for his defaults. Every departure from

the principles of law's inner morality is an affront to man's dignity as a responsible agent."[5]

For Fuller, there is an internal, necessary connection ("involves of necessity the commitment") between rules and the conception ("view") of addressees of these rules as responsible persons. No rule without responsible addressees. Fuller only says in a subordinate clause what exactly this means: "capable of understanding and following rules, and answerable for his defaults." While the first part refers to the cognitive and motivational abilities required for rule compliance (ability to understand and follow rules), the second part refers to a specific position or status associated with the concept of a responsible person: To be able and supposed to give a response in case of non-compliance or violation of the rule. Fuller refers here to the literal sense of the term ver-answering (Latin respondeo). Anyone who does not follow a rule is able and obliged to respond. It is therefore the specific situation of breaking the rule that makes the necessary connection between rule and responsibility evident. However, there are two further aspects to this description:

(1) When it comes to answering, a communicative relationship is presupposed, i.e. in addition to the answering speaker, there is another speaker/listener to whom the answer is addressed. It is an answer to the other person's critical question as to why the person acting did not follow the rule although he should have done so. Following a rule is therefore not only an intersubjective practice because, as Wittgenstein has shown, no one can follow a rule only once and for themselves alone, i.e. because intersubjectivity is part of the meaning of "rule".[6] Furthermore, because there is a practice of reacting to the non-observance of a rule, in which critical questions are asked and answered. Fuller's opponent, H.L.A. Hart, described this practice in a similar way: Whether someone follows a rule can only be inferred from an internal point of view, which consists of a critical reflective attitude that manifests itself in an intersubjective practice of criticizing rule violations.

"What is necessary is that there should be a critical reflective attitude to certain patterns of behavior as a common standard, and that this should display itself in criticism (including self-criticism), demands for conformity, and in acknowledgements that such criticism and demands are justified, all

5 Fuller, The Morality of Law, New Haven 1964/69, 162.

6 Ludwig Wittgenstein, Philosophische Untersuchungen, Frankfurt/Main 1975, p. 127ff. (Nr. 199ff.).

of which find their characteristic expression in the normative terminology of 'ought', 'must', and 'should', 'right' and 'wrong'."[7]

(2) The second element contained in Fuller's "answerability" is only indirectly apparent. If responsibility refers to the case of non-compliance with the rule, this presupposes that a rule cannot actually be followed. This sounds trivial prima facie - but it is not. The fact that a rule should be followed is part of its meaning. It would be nonsensical to establish a rule without the requirement that it be followed. But rules and norms are not of such a nature that they would guarantee and ensure their own compliance. A rule always includes the possibility of its violation. A rule wants to be obeyed (which makes it susceptible to the utopia of the certainty of norm compliance), but only in a way that includes the possibility of non-compliance. To put it in paradoxical terms: Following a rule presupposes the possibility of not following it. This is presumably an unavoidable presupposition of the practice of following rules.

Of course, a distinction must be made with regard to the possible reasons for non-compliance. One reason may be the incorrect application of the rule. In this case, the addressee wants to follow the rule, but makes mistakes because he either does not understand its meaning correctly (cognitive error) or because he does not manage to control his behavior in such a way that he follows the rule despite understanding it correctly (motivational error). However, if he is cognitively and motivationally capable of following the rule, but does not actually follow it, then he finds himself in a position where he has to answer to third parties. It is only because this possibility - or freedom - exists that we can also be responsible for following the rule. If we were to follow a rule automatically, the concept of a responsible person would be superfluous - any more than a machine is responsible for functioning in accordance with the technical rules of its mechanical construction. The utopia of norm-following certainty would therefore, if realized, transform norm-following subjects into machines.

7 H.L.A. Hart, The Concept of Law, Oxford 1961, p. 55f.

C. The freedom to engage in deviant behavior

It therefore seems sensible to speak of norms instead of rules in general - which is presumably also the meaning that corresponds to the use of the word by Lon Fuller (and also H.L.A. Hart). It would then apply to every social practice that is constituted and regulated by norms that it presupposes the concept of a responsible person and thus includes the possibility that the addressees may or may not follow their norms. This applies to every kind of norm, including the most elementary of all norms, which is inherent in a promise. Anyone who promises something to another person creates a norm (or, in a specific case, updates the current norm that promises should be kept) for which they are responsible. He thereby commits himself to the future behavior of another person. However, he is only responsible for keeping the promise because he has the option of not keeping the promise. It is therefore solely up to him whether he will do what he has promised or not (apart from extreme changes in circumstances that make it impossible to keep the promise - clausula rebus sic stantibus). The ability to commit oneself to one's own behavior towards others in the future is anything but natural, and it includes the freedom to behave differently. According to Nietzsche, the greatest and at the same time paradoxical task for man is therefore: "To breed an animal that is allowed to promise."[8]

Following Lon Fuller's quote, people who embark on the project of regulating their coexistence through norms are therefore taking a certain risk if following a norm also means not being able to follow the norm. The risk lies in the norm addressee as a responsible person. Instead of creating mechanisms that ensure automatic and perfect compliance with norms (as with a machine), a practice is established that H.L.A. Hart has described as a critical reflective attitude towards norms. It is the practice of criticizing deviant behavior, possibly even the practice of criticism at all.

I. Techniques of risk minimization

However, communities with a norm-guided social practice do not completely refrain from minimizing the risk of deviant behavior in other ways. Which additional measures are taken depends, among other things, on the

8 Friedrich Nietzsche, Zur Genealogie der Moral, Dritte Abhandlung, in: Werke, hrsgg. v. Karl Schlechta, Band 2, Darmstadt 1994, S. 239 (Herv. F.N.).

type of norm in question. Essential are processes of socialization, education and cultivation - what Michel Foucault called "subjectivation" - which not only contribute to the development of those cognitive and motivational skills and dispositions to follow basic social norms, but also to a person learning what it means to be a responsible person. This begins with children learning to promise and to trust in a promise made to them - but then also experiencing that not every promise made to them is kept. Of course, they also learn that there is a social practice of criticizing the breaking of a promise and how to participate in this practice. It is possible that the experience of being able to violate norms and being criticized by others for doing so is part of the process of learning what it means to follow norms as a responsible person - as can be seen in adolescents in the adolescent phase, for example.[9] In addition, there are various social conditions that must be fulfilled, at least to a certain extent, in order to become a responsible person and to be criticized as such for one's own actions. To the extent that compliance with norms becomes unreasonable because the addressees are hardly in a position to do so given the social conditions, the criticism of an addressee for violating these norms becomes unfair and the practice of critical reflective attitude among responsible persons becomes pointless.[10]

I am concentrating here only on one particular type, the legal norm. Not the only, but a central way of making norm compliance more likely is to link the legal norm with the threat of coercion or even sanctions in the event of norm violation. Some authors, such as the proponents of the coercion thesis, even consider this coupling to be a necessary part of the concept of law.[11] Irrespective of this, however, it is easy to see that the addition of a threat of coercion (and, in the case of de facto deviation, execution) does not change the fact that the addressee of a legal norm is subject to the presupposition that he has the option of not complying with the norm. If legal norms are coupled with a threat of coercion, this only means (provided that it is credible and the addressee is aware of the threat)

9 See, e.g., Gertrud Nunner-Winkler´s research results, in: Rainer Döbert/Gertrud Nunner-Winkler, Adoleszenzkrise und Identitätsbildung, Frankfurt/Main 1975; Nummer-Winkler, Prozesse moralischen Lernens und Entlernens, in: Zeitschrift für Pädagogik 55 (2009), S. 528-548 (534ff.).

10 Klaus Günther, Zwischen Ermächtigung und Disziplinierung. Verantwortung im gegenwärtigen Kapitalismus, in: Axel Honneth (Hg.), Befreiung aus der Mündigkeit. Paradoxien des gegenwärtigen Kapitalismus, Frankfurt/M. u. New York 2002, 117 – 140.

11 See, recently: Himma, Coercion and the Nature of Law, Oxford 2020.

that the violation of the norm appears less preferable to the addressee than compliance due to his individual preferences. According to Robert Nozick's analysis, it merely provides an additional reason for deciding in favour of compliance and against non-compliance; according to Joseph Raz, it even provides only a subsidiary and partial auxiliary reason.[12] However, this does not eliminate either the possibility or the freedom of the addressee to decide against compliance with the norm and to accept the sanction with its disadvantages. The threat of coercion does not eliminate the responsibility of the norm addressee; on the contrary, it is even the justifying reason for imposing the sanction on him in the event of a violation of the norm. The alternatives would be a system of brutal terror or a system of manipulation and control that penetrates into the smallest capillaries of the psyche, of complete conditioning.[13] Here, responsibility would lie, if at all, with a centre that manages the lives and psyches of the norm addressees in order to ensure compliance with the norm. In contrast, compliance with legal norms by responsible actors involves a kind of decentralized ontology of individual subjects who each comply with norms on their own - or not.[14] With regard to the alternative between a conventional criminal law with a criminal sanction for deviant behavior and a prevention that makes this impossible from the outset through technical precautions, Bernhard Haffke has clearly marked the consequences that endanger freedom - at the same time as a warning against a superficial understanding of the ultima ratio principle in criminal law, which approves of any alternative regulation that does not require a criminal sanction. "While psychological prevention, albeit by means of reward and punishment, still chooses the path - the rocky but decent path - via the subject, in technical prevention the subject is no longer considered from the outset: Deviant behavior has become impossible."[15] With the responsible subject, its basis, the freedom to decide

12 Robert Nozick, Coercion, in: Morgenbesser/Suppes/White, M. (eds.): Philosophy, Science, and Method: Essays in Honor of Ernest Nagel, New York NY 1969, 440–472; Joseph Raz, Practical Reason and Norms, London 1975, 162f.

13 Like the *ludovico technique* in Antony Burgess (1962)/Stanley Kubricks (1971) dystopical Novel/Film „A Clockwork Orange", together wirh the warning of a priest(*sic!*) to the protagonist, before he participates voluntarily in the conditioning experiment: "If a man cannot choose he ceases to be man.".

14 Günter Jakobs, Das Schuldprinzip, Rheinisch-Westfälische Akademie der Wissenschaften, Vorträge G 319, 1993, p. 34, and the parable on p. 34 f.

15 Bernhard Haffke, Die Legitimation des staatlichen Strafrechts zwischen Effizienz, Freiheitsverbürgung und

for or against compliance with the norm, also disappears: "Classical liberal criminal law deliberately chooses the path via the offender as a moral personality, as a responsible subject and, by proceeding in this way, respects his freedom to deviate from the norm."[16]

II. Compliance with norms - certainty or trust?

Every norm-guided social practice, and in particular every legal system, is therefore dependent on the existence of institutions, procedures and practices for criticizing deviant behavior, coercion and other sanctions, as well as on mutual trust that the responsible person will behave in accordance with the norm. Every legal system is based not only on this mutual trust, but also on the fact that in the event of deviant behavior, the institutionalized procedures for criticizing deviant behavior (e.g. court proceedings) are activated and the previously threatened sanctions are also imposed (legal trust).

Despite all the measures mentioned to ensure average compliance with norms, the risk of deviant behavior remains, albeit certainly to a lesser extent than without them. Trust, too, is only necessary because we have reasons to rely on others, but no certainty. According to Georg Simmel's well-known formulation, trust is "a hypothesis of future behavior that is certain enough to base practical action on, (...) as a hypothesis a middle state between knowing and not knowing about people."[17] With the help of new digital technologies, especially AI, there now seems to be a possibility of eliminating this risk or at least minimizing it to such an extent that the probability of choosing this behavioural alternative is significantly reduced. This is the promise or vision of smart orders. They are designed to minimize or completely eliminate deviations from their norms through intelligent design and with the help of algorithmic operations.[18] The trust in the ability and willingness of norm addressees to comply with the norm, which is risky due to the ever-present possibility of deviation, could thus be transformed into a certainty of compliance with the norm.

Prävention, in: Bernd Schünemann, Hans Achenbach u.a. (Hrsg.), Festschrift für Claus Roxin zum 70. Geb., Berlin/New York 2001, Sp 955 – 975, p. 967.

16 Haffee, p. 967.

17 Georg Simmel, Soziologie, Berlin 1908/1983, 263 (trans. K.G.).

18 Günther, Von normativen zu smarten Ordnungen?, in: Forst/Günther (Hrsg.), Normative Ordnungen, Berlin 2021, S. 523-552; ders., Die Zukunft der Freiheit in smarten Ordnungen, in: WestEnd – Neue Zeitschrift für Sozialforschung, 17/2 (2020), 165-175.

The risk that a promise will not be kept can be eliminated in a smart contract, for example, by automating the execution of performance and consideration in a blockchain. The risk of criminal offences can be minimized through situation- and person-related predictive policing and algorithm-based prevention of future criminals. Projects such as anticipatory governance and smart cities are motivated by the prospect of defusing social conflicts preventively ("prevention rather than cure") and organizing the "confluence" of urban interactions without conflict. The extreme case is the social credit model practiced in some regions of China.[19] With the help of such technologies, whose effectiveness can be greatly enhanced by AI, a society can come even closer to the supposed ideal of perfect compliance with norms, without it even being a question of "the view that man is, or can become, a responsible agent, capable of understanding and following rules, and answerable for his defaults."

D. AI as a new technology of self-commitment?

Of course, attempts have always been made to develop technologies that make compliance with norms more likely - coercion and its threat are perhaps the most primitive form. However, this also includes the technologies of self-coercion (or self-discipline). At the latest since economics abandoned the rationally calculating *homo oeconomicus* as the standard model, rational strategies for dealing with imperfectly rational behavior have become the focus of attention, such as the liberal-paternalistic model of nudging.[20] One possible strategy is self-commitment. The example of Ulysses has become famous: he has his companions tie him to the mast of his ship so that he can pass the island of the sirens and listen to their beguiling song without surrendering to their power, which tempts him to commit suicide. Jon Elster used this example in his early studies on imperfect rationality to show that the knowledge of one's own imperfections and weaknesses does not have to lead to fatalism, but can instead be the reason for choosing a strategy of self-binding:

19 For further elaboration on tnese examples see: Klaus Günther, Von normativen zu smarten Ordnungen?

20 Richard H. Thaler u. Cass Sunstein, *Nudge. Improving Decisions About Health, Wealth and Happiness*, New Haven: Yale UP, 2008.

"Ulysses was not fully rational, for a rational creature would not have to resort to this device; nor was he simply the passive and irrational vehicle for his changing wants and desires, for he was capable of achieving by indirect means the same end as a rational person could have realized in a direct manner."[21]

Anyone who, like Ulysses, foresees at time t1 that he will make a wrong, i.e. at least self-interestedly irrational, choice at time t2 and wants to avoid this, is behaving rationally if he takes precautions at time t1 that prevent him from behaving irrationally at time t2. This is generally more rational than relying on having sufficient psychological resilience in the decisive situation against the strong tendency to make the wrong choice. The foreseeable deficit in rationality in t2 is compensated for by the rationality in t1, so that the person behaves just as rationally as if they had acted completely rationally in t2.

It is obvious to conceive of smart orders as a means by which behavior that deviates from the norm in t1 can be technically excluded in t2. They would then be nothing other than a technically optimized tool for *precommitting oneself*[22], which would also have the advantage that it would operate with much less coarse means than Odysseus' shackles, which cut painfully into the body and were lashed even tighter at the decisive moment in response to the command given in advance. The use of a large number of apps offered on smartphones has become a widespread everyday practice in order to encourage a healthier lifestyle with more physical exercise or a reduction in body weight, for example, as a quantified self with self-tracking and an exchange with others on corresponding platforms that serves the purpose of mutual observation and control. And why shouldn't providers, software companies, health insurance companies and medicine, which can make diagnosis and therapy more effective and efficient through personalization, collect and analyse the data produced in the process and generalize it into behavioural patterns in order to perfect technically optimized self-restraint for preventive healthcare? One of the possible interpretations is that this, too, is only a technical optimization of practices of self-observation and self-care for the sake of a virtuous good life, which date back to antiquity.[23]

21 Jon Elster, *Ulysses and the Sirens*, Cambridge UK 1979/2013, p. 36.

22 Elster, *Ulysses and the Sirens*, p. 37.

23 See, e.g.: Stefan Meißner, *Lifelogging. Selbstvermessung als Möglichkeit von Selbststeigerung, Selbsteffektivierung und Selbstbegrenzung*, Berlin 2016; against rush critism

Would this not also apply to smart orders, provided they are only produced through collective self-determination? Jon Elster did indeed oppose the simple transfer of individual to collective self-binding when he interpreted constitutions primarily as a strategy for binding future political majorities to fundamental norms and less as a self-binding of the constitution-making actors themselves.[24] However, at least a democratically established coercive law could ideally be understood as an order through which the co-legislators bind themselves in their future role as norm addressees by means of the threat and execution of coercion. In this way, they would also use a technical means, acting on the body, the psyche and the emotions, to ensure compliance with the norm even in the more frequent case that someone acts not out of insight into good normative reasons or out of respect for the law, but in order to avoid disadvantages. For example, theories of negative prevention rely on the psychological effect of the threat of punishment and penalties, which create so much fear in potential delinquents that they avoid norm-violating behavior. Although the threat of coercion does not have the effect of depriving the person concerned of all freedom to behave in a deviant manner (accepting the disadvantages), it does reduce the probability. However, there are at least three reasons to doubt the equation of smart orders with analogous techniques of self-binding.[25]

The *first* reason relates to the assumption that we are actually dealing with collective autonomy through self-binding or self-intervention. When Ulysses orders his companions to bind him, it is he himself who binds himself by influencing his own future behavior with technical means and the help of third parties. Only under this condition does he preserve his autonomy, even if he obliges his companions in t1 not to listen to his expected command in t2 to untie him now. The situation is different when companies or states apply such technologies to customers or citizens in a way that is not or only partially transparent to them and over which they have no or only limited decision-making and control. In this case, those who decide on the use of technical means to modify future behavior are different from those who are affected by it. In addition, despite all efforts to ensure transparency and information, the individual modalities, duration, mode of action and, not least, the techniques used to skim and use further

see also Kathrin Passig, Internetkolumne. Unsere Daten, unser Leben, in: *Merkur* (756) 2012, S. 420 - 427.

24 Jon Elster, *Ulysses Unbound*, Cambridge UK 2000, S. 88 - 118.

25 See Günther, Die Zukunft der Freiheit.

behavioural data will remain largely hidden from them, if this is not already covered by commercial or state secrecy. A publicly applicable legal norm is different from an algorithm. Even if it is in the best interests of those affected, i.e. if it would be rational for them to submit to the externally determined smart binding, there is still an asymmetrical relationship between the two. In this respect, it makes no difference whether we are talking about smart technologies or analogue ones. Liberal paternalism, which makes use of practices such as nudging, is not immune to this criticism either, even if the asymmetrical relationship remains largely transparent here. It is all the more likely to apply if those affected are simultaneously moving in digital filter bubbles that, with the help of patterns generated from their own behavioural data, preferably only provide them with information and impulses that dispose them to accept smart ties without question.[26]

Of course, it could be the case that those affected voluntarily and in full knowledge of all the circumstances agree to a technical influence on their practical attitudes to standards, be it contractually in relation to a company or by way of collective self-binding through democratic legislation. Would it then not be in accordance with its own will? Rousseau had conceived the social contract in a similar way under similar conditions, which contains the clause that the dissenter may be forced to be free. For Kant, too, it was clear that the mechanics of coercive law, which act solely on the body, are morally unsuspicious if they are coerced in the name of practical reason.[27] Autonomy is preserved by Kant with the republican form of self-legislation. In this respect, wouldn't smart technologies for enforcing norms only be an optimization of analogous ones? However, Rousseau and Kant also foreshadow the dualism characteristic of modern capitalist societies between the two worlds of the *citoyen* and the *bourgeois*, the virtuous republican co-legislator and the private individual driven by his selfish passions. If Kant considers the task of establishing a state to be solvable even for a nation of devils, he makes it clear that the naked self-interest of each individual is sufficient to recognize that a state order of coercive law is preferable to a state in which everyone must fear for their lives in a permanent civil war.

At the same time, however, both authors maintain that man is not lost in his diabolical, instinctive nature, which tends towards selfishness and

26 S. dazu Günther, Die Zukunft der Freiheit.

27 S. Marcus Willaschek, Recht ohne Ethik? Kant über die Gründe, das Recht nicht zu brechen, in: Volker Gehrhardt (Hrsg.), Kant im Streit der Fakultäten, Berlin/New York 2005, S. 188 – 204, S. 195.

is corrupted by passions, with its cognitive and motivational deficits, but always has general human reason at his disposal. The physical mechanics of coercive law can then only be justified in the event that rational insight alone is not sufficient to comply with a rationally justified norm - not the other way around. This is why they trust the subjects of general human reason as flesh-and-blood human beings to be co-legislators and to be able to follow the law out of rational motives. There is no other way to justify the everyday mutual trust in a general willingness to follow the law in the event that the law cannot directly compel. Without such trust, no legal system would be stable in the long term, because the law cannot be enforced always and everywhere. Ulysses knows when his passions will lead him to his doom and takes specifically tailored, temporary technical precautions, thus retaining the upper hand with his insight.[28]

Smart orders, however, and this is the *second* reason for scepticism about the suitability of AI as a tool for self-commitment to self-legislation or political autonomy, silence this often conflictual interplay between the two worlds. With their appearance, the negative side of the norm addressee's fallibility becomes absolute, as if deviant behavior were the constantly disruptive normal case that constantly endangers general security. With technology, the view of norm addressees is changing. To the extent that norm compliance with or without persons qua norm addressees can be technically (re-)produced and thus perfected, the enterprise of realizing normative orders via persons who are at the same time autonomous and fallible flesh-and-blood human beings appears risky, dysfunctional and prone to disruption in comparison to the more effective technologies. The distrust in the general willingness to follow norms may also increase in modern, globalized, pluralistic and fragmented societies, so that for this reason too, smart norm realization appears to many to be a better alternative. As in the analogue two-world concept, human nature is then considered to be in need of control and mastery, but no longer with coercion, repression and the subjectification of control functions and techniques of self-control, which in extreme cases produce an authoritarian character, but with smart technologies that are softer, more sustainable, more effective and more

28 For Max Horkheimer und Theodor W. Adorno, *Dialektik der Aufklärung*, Frankfurt a. M., 1944/1973, S. 42, Ulysses is nevertheless already "the archetype of the bourgeois individual, whose concept originates in that unified self-assertion whose pre-worldly pattern is provided by the driven individual.".

efficient. In the best-case scenario, even people with fallible behavior can be completely replaced by technologies.

Based on comprehensive personality profiles, which are already being used for commercial purposes by collecting data and analysing it with the help of AI, laws could be individually tailored for each citizen. These would no longer take the form of abstract and general norms, which would then have to be cognitively and motivationally concretised by the addressees for their respective situation, but would instead take the form of individualised behavioural directives. Casey and Niblett analysed what such norms could look like back in 2013: "Predictive technology will generate greater ex ante information that can be used by lawmakers to write highly specific, complex laws. And individuals will receive notice of these complex laws in a simple form thanks to technological advances in communication. This will be the death of rules and standards and the rise of microdirectives."[29] Such microdirectives could, for example, set individualised maximum speed limits for every car driver in every situation: "For example, a microdirective might provide a speed limit of 51.2 miles per hour for a particular driver with twelve years of experience on a rainy Tuesday at 3:27 p.m. The legislation remains constant, but the microdirective updates as quickly as conditions change."[30] Legislation would only set the political goal of determining driving speeds in road traffic in such a way as to enable safe and trouble-free mobility for everyone at the same time as everyone else. This goal would then be transformed into an individual behavioural directive for each driver with the help of AI-generated micro-directives, which could be adapted to changing conditions at any time. The problem that a general and abstract standard cannot foresee all future cases of application and that the addressees have different capabilities and capacities in different situations, e.g., do not have sufficient motivation to comply with the standard in each individual case or make mistakes when specifying the standard to their case with a specific context and complexity would be solved. As Omri Ben-Shahar and Ariel Porat who both defend such a vision of personalized law frame it: "Rather than blindfolded, let the law know everything that is relevant about people, apply the underlying legal principles to facts of each person, and thus tailor personalized legal regimes. If medicine, education, or parenting can treat, teach, or nurture better when personalized and

29 Anthony J. Casey and Anthony Niblett, The Death of Rules and Standards, in: Indiana Law Journal, 92, (2017), pp. 1407 – 1447 (1446).

30 Ibid., p. 1404.

adjusted to the subjective, why not law?"[31] This even more so if the (self-driving) cars were technically equipped in such a way that they implement the behavioural directive directly, so that a driver no longer has to act at all. What applies to road traffic can be applied mutatis mutandis to all other legally regulated areas.

Thirdly, this leads to the problem of the tendency to demoralize individual behavior, as described and critically explained by Roger Brownsword and Evgeny Morozov, among others.[32] To the extent that people become accustomed to smart orders, their willingness and ability to form a moral judgment in confrontation with their own freedom to behave differently and to actively use practical reason, which can be put into action, decreases. The automation of virtue could result in a "moral disability" in the medium term.[33] In a similar way to how the constant use of a calculator atrophies the ability to apply mathematical rules and perform more complex mathematical operations by hand, let alone calculate simpler tasks in one's head, smart systems can lead to a loss of moral deliberation by either allowing norm-following behavior to be performed automatically by the person or replacing it entirely. The "citizen does not have to weigh the reasonableness of her actions, nor does she have to search for the content of a law. She just obeys a simple directive".[34] The possibility of deviating from the norm is, as it were, the sting that constantly challenges the norm addressee to use his practical reason by having to ask himself about the reasons for his judgments and actions as well as their justification to himself and others. [35] This is not least in order to consider and weigh up the reasons that may speak for or against following this norm - or perhaps another, perhaps even contradictory norm - in the specific situation. In this way, people

31 Omri Ben-Shahar, What is Persoanlised Law?, Faculty of Law Blogs/University of Oxford, 27 June 2022; Omri Ben-Shahar and Ariel Porat, Personalized Law - Different Rules for Different People, Oxford 2021.

32 Roger Brownsword, Lost in Translation: Legality, Regulatory Margins, and Technological Management, in: Berkeley Technology Law Journal 26 (2011), pp. 1321-1366 (1356); Evgeny Morozev, Smarte neue Welt. Digitale Technik und die Freiheit des Menschen (engl. Orig.: To Save Everything, Click Here), München 2013, p. 326 u. 343.

33 Morovzev, p. 337f; referring to: Ian Kerr, Digital Locks and the Automation of Virtue, in: Michael Geist, ed., From "Radical Extremism" to "Balanced Copyright": Canadian Copyright and the Digital Agenda, Irwin Law, 2010, p. 247 – 303 (282).

34 Casey/Niblett (n. 29) 1402.

35 Deviant behavior is therefore often also a source of innovation - in both a bad (criminal) and a good (enabling moral learning) sense.

develop (especially during and after adolescence, *see IV above*) into morally capable of judgement and action, i.e. into responsible persons who possess the critical reflective attitude described above (IV), who are able to criticize each other for behavior that deviates from the norm and thus participate freely and actively in rule-guided intersubjective moral practice. In smart orders, on the other hand, they adopt an observational attitude towards norms and allow their behavior to be passively guided by technology.

As a result, they are not only gradually losing their capacity for moral judgment and action, but also their moral critical faculties towards their normative orders themselves, which are taking on a smart form. Their ability and need for justification is increasingly fading because the norm addressees, weaned from autonomous compliance with the norm, are neither able nor willing to take a critical stance on the claim to validity and question the democratic legitimacy or, in the case of morality, the justification in the interests of all. However, normative orders are dependent on the reasons for their validity being publicly demanded, discussed and criticized, and that their factual, currently practiced interpretations can also be criticized and changed. However, this also presupposes that the norm addressees see themselves as co-legislators and have the corresponding ability and make active use of it, i.e. that something like "active moral citizenship" exists and is actually practiced. [36]

E. The normative constitutionalization of smart orders as a way out?

Of course, there have been and still are areas of society in which the benefits to society as a whole of adhering to norms that can be ensured through technical control and technical innovations outweigh the disadvantages. This is the case, for example, wherever technologies with greater risk potential are used, such as in motorized road traffic. In view of the virulent dangers there, the deadly realization of which can be measured by the annual number of traffic fatalities, why should motor vehicles not be

36 Brownsword, Lost in Translation, p. 356. This would also be an argument against the objection raised by Luna Rösinger, Der Autonomiebegriff im Kontext künstlicher Intelligenz als Prüfstein für die Rechtsphilosophie, in: Zs f. Rechtsphilosophie (ZRph), 6-7 (2022/2023), pp. 209-226 (225), that the Impossibility Structures "under the aspect of making it impossible to break the law do not represent a new challenge for legal autonomy", but would "once again expose the old misconception of a law misunderstood as behavioral control" (trans. K.G.).

equipped with a chip that automatically reduces the speed of the vehicle without the driver's intervention if a traffic rule requires this (and perhaps even be able to switch itself off in emergency situations)?

Given the ambivalent consequences of smart regulations, the question of how a democratic constitutional state should react to efforts to implement them should therefore not be answered in the sense of an either/or. Rather, it is a question of determining the degree or extent to which smart orders should coexist with normative orders and be effectively legally bound to them. Only when smart orders begin to call the concept of a responsible person into question or make it completely dispensable would it be necessary to consider whether our lifeworld of normativity is not fundamentally changing and whether the ideal of perfect compliance with norms is not a false ideal.

This could only be achieved by ensuring that normative orders with responsible persons who participate in a practice of intersubjective criticism both in the setting of norms and in the application and observance of norms remain the medium in which the substitution of normative sub-area orders by smart orders (e.g. in motorized road traffic) as well as these themselves, i.e. their goals and purposes as well as their probable consequences, must be justified. The task would therefore be to constitutionalize smart sub-orders of society through a normative order with legally institutionalized forums and procedures for public criticism and justification. This constitution of smart orders should therefore not be designed and institutionalized as a smart order, but only and exclusively as a normative order.[37] Furthermore, as the logic of constitutionalization implies both the primacy and the reservation of the constitution in a normative hierarchical relationship to smart orders must be ensured and legally controllable.

Whether such a constitutionalization of smart orders is sufficient to prevent them from a successive and comprehensive colonization of normative orders, including their subjects, is, however, questionable. Ultimately, it is likely to depend on whether and to what extent lifeworld and institutional spaces are maintained in which citizens can cultivate and participate in intersubjective practices, in which they understand themselves as responsible persons and also want to understand themselves as such. [38] Above all,

37 See for a proposal pointing into this direction: Andreas Werkmeister, Erste Überlegungen zum Begriff der »politischen Datenwirtschaftsstraftat«, in: GA 2021, Jg. 168, 570-587.

38 S. Morozev, Smarte neue Welt, p. 337.

this includes the practice of mutual criticism. Only in this way can they acquire the ability and develop and cultivate the will to bear the risk of setting imperfect norms as well as individual deviant behavior in mutual trust in a general willingness to follow norms. The fact that this self-image is a constitutive element of their role as co-legislators in a public process of democratic legitimization has already been made clear by the census ruling of the German Federal Constitutional Court. It insisted that the unrestricted collection and processing of personal data would not only jeopardize "the individual's opportunities for development", but also the functioning of "a liberal democratic community based on the ability of its citizens to act and participate."[39]

Smart orders, on the other hand, appear attractive and can be justified by the fact that they are able to minimize this risk and offer more security against deviant behavior with an approximate certainty of compliance with the norm, assuming that there is a kind of natural consensus on the meaning and purpose of the order whose security would be guaranteed by AI. The opportunity to make road traffic safer with a smart traffic code then appears as a utopian or dystopian paradigm for a social order that could make coexistence safer overall, depending on perspective and attitude.

39 BVerfGE 65, 1, 43 (trans. K.G.); Simitis, BDSG, Rn. 30.