Robert Rothmann*, Elisabeth Mayer

Künstliche Intelligenz im Strafvollzug: Zulässigkeit, Bedarf und Ethik multimodaler Überwachung im Kontext der österreichischen Justiz

Artificial intelligence in prison: Permissibility, demand and ethics of multimodal surveillance in the context of the Austrian justice system

https://doi.org/10.1515/mks-2024-0003

Zusammenfassung: Der vorliegende Beitrag widmet sich der Analyse multimodaler KI-Anwendungen im Strafvollzug zur Erkennung von selbst- und fremdgefährdendem Verhalten. Ausgangspunkt der Untersuchung ist ein entsprechendes Forschungsprojekt, das jüngst in österreichischen Justizanstalten durchgeführt wurde. In der Abhandlung der Thematik wird zunächst eine Beschreibung der Technologie vorgenommen und die damit verbundene Zielsetzung erläutert. Darauf aufbauend erfolgt eine Prüfung der rechtlichen Zulässigkeit, in welcher auf datenschutzrechtliche Aspekte ebenso Bezug genommen wird wie auf die spezifischen Vorgaben des österreichischen Strafvollzuggesetzes und den aktuellen Entwurf einer Verordnung zur Regulierung von KI auf europäischer Ebene. Darüber hinaus wird der praktische Bedarf des Technologieeinsatzes analysiert, wobei sich die Überlegungen auf empirische Einblicke und Gespräche in den Vollzugsanstalten stützen. Die Analyse mündet in einer interdisziplinären Diskussion der Verhältnismäßigkeit des Technologieeinsatzes sowie damit verbundenen ethischen Abwägungen in Bezug auf die Menschenwürde.

Schlagwörter: Künstliche Intelligenz, Videoüberwachung, Verhaltensanalyse, Biometrische Daten, Menschenwürde

Anmerkung: Für hilfreiche Hinweise, kritische Diskussion und Unterstützung in der Recherche danken wird insb Markus Kastelitz, Madeleine Müller, Philipp Poindl sowie David Schneeberger.

***Kontaktperson: Robert Rothmann MA, PhD**, Research Institute – Digital Human Rights Center, Amundsenstraße 9, 1170 Vienna, Austria, E-Mail: robert.rothmann@researchinstitute.at.
https://orcid.org/0000-0001-8049-9697
Elisabeth Mayer BA, BSc, MA, Research Institute – Digital Human Rights Center, Amundsenstraße 9, 1170 Vienna, Austria,
E-Mail: elisabeth.mayer@researchinstitute.ac.at

Abstract: This article is dedicated to the analysis of multimodal AI applications in prisons for the prevention of violence and suicide scenarios. The starting point of the analysis is a research project that was recently carried out in Austrian prisons. Therefore, we first provide a short description of the applied technology and the associated project objectives. Subsequently the legal admissibility of the technology in question is examined, referring to data protection law as well as to specific legal requirements of the Austrian penal system and the current version of the proposal for an Artificial Intelligence Act on the European level. In addition, the practical need for the use of said technology is analysed, whereby the considerations are based on empirical insights on everyday prison life. The analysis concludes with an interdisciplinary discussion of the proportionality of the envisioned technology and related ethical considerations regarding human dignity.

Keywords: Artificial Intelligence, Video Surveillance, Behavioural Analysis, Biometrics, Human Dignity

1 Einleitung

Studien zufolge sind Personen in Haft überdurchschnittlich häufig von Gewalt und selbstgefährdendem Verhalten betroffen. So zeigt beispielsweise eine von Hofinger & Fritsche (2021) in Österreich durchgeführte Dunkelfeldstudie, dass rund 72 % der befragten Insass*innen mindestens einmal während ihrer Inhaftierung psychische, körperliche oder sexuelle Gewalt erlebt haben. Auch Suizidversuche sind unter Insass*innen weit verbreitet und Selbsttötung stellt die häufigste Todesursache in Haft dar (Fazel et al., 2017; WHO, 2007).

Infolgedessen wird auch von staatlicher Seite vermehrt der Frage nachgegangen, wie KI-basierte Technologien ein-

gesetzt werden können, um die Sicherheit in Justizanstalten zu erhöhen und die Gesundheit der Insass*innen entsprechend zu gewährleisten (EUROPRIS, 2020). Der vorliegende Beitrag diskutiert diese Entwicklungen anhand des Forschungsprojekts »Künstliche Intelligenz im Strafvollzug« (KIIS), welches sich mit der multimodalen Detektion sicherheitskritischer Szenarien über spezielle bildgebende Sensoren und Smartwatches beschäftigt.[1] Ziel des Projekts war es, zu untersuchen, ob bzw. inwiefern die Technologie dazu geeignet ist selbst- und fremdgefährdendes Verhalten im Vollzugsalltag (frühzeitig) zu erkennen. Konkret sollte die automatisierte Erkennung spezifischer Verhaltensmuster den diensthabenden Beamt*innen ein schnelleres Eingreifen im Fall von gewalttätigen Auseinandersetzungen, Suizidversuchen und medizinischen Notfällen ermöglichen. Dies sollte mittelbar wiederum personelle Ressourcen für die persönliche Betreuung der Inhaftierten zu deren Rehabilitierung und Resozialisierung freimachen.

Das technische Vorhaben wird in weiterer Folge einer interdisziplinären Analyse unterzogen. Am Beginn steht eine kurze Beschreibung des Forschungsstandes und anknüpfend daran eine Erläuterung des Innovationsaspektes des KIIS-Projekts. Dabei werden die technischen Komponenten des multimodalen Systems erklärt sowie Details zur Datenerhebung und Implementierung der Test-Settings erörtert. Nach Darlegung des Sachverhalts werden rechtliche Aspekte zur möglichen Anwendung der entwickelten Technologie geprüft. Dabei wird im Kern die Frage nach einer geeigneten Rechtsgrundlage im Strafvollzug geklärt und eine Klassifizierung des KI-Systems über die aktuelle Version der KI-Verordnung der Europäischen Union vorgenommen. Die juristische Beurteilung wird durch eine Bedarfsanalyse ergänzt und mündet in einer interdisziplinären Prüfung der Verhältnismäßigkeit des mit dem Einsatz der Technologie verbundenen Grundrechtseingriffs. Der Beitrag schließt mit einer Zusammenfassung der zentralen Ergebnisse und einer ethischen Diskussion des Einsatzes KI-basierter Überwachungstechnologien im österreichischen Strafvollzug.

2 Stand der Forschung

2.1 Internationale Entwicklungen

Sucht man nach algorithmusbasierten Anwendungen im Bereich des Strafvollzugs, so stößt man zunächst vor allem auf Entwicklungen rund um den Begriff »Smart Prison« (Knight et al., 2023; Knight & Van De Steene, 2017; Puolakka & Van De Steene, 2021; Kaun & Stiernstedt, 2020) sowie auf sogenannte »Offender Management Systeme« (Baker et al., 2021; Berk, 2017; Hao, 2019; Kohn, 2021; Rizer & Watney, 2019; Završnik, 2021) und den Einsatz von KI bei Gericht (Aletras et al., 2016; Reiling, 2020). Während insb. Risiko- und Offender-Management Systeme zur Vorhersage bzw. Berechnung der Rückfallwahrscheinlichkeit oder der Empfehlung einer Bewährungsstrafe und erforderlicher Rehabilitationsprogramme im wissenschaftlichen Diskurs bereits wiederholt adressiert und (kritisch) behandelt wurden (Dressel & Farid, 2018; Thomas & Nunez, 2022; Završnik, 2021), lassen sich über KI-basierte Anwendungen zur (visuellen) Verhaltensanalyse und Gewaltprävention im Strafvollzug kaum wissenschaftlich belastbare Informationen finden. Unsere Recherchen zeigen stattdessen eine Reihe von Medienberichten und industrienahen Beiträgen, die in erster Linie das (künftige) Potential KI-basierter Technologien im Strafvollzug hervorheben und eher allgemein auf bereits bestehende Systeme verweisen (Chen, 2019; Houser, 2019; Yan, 2019; Redden et al., 2020; Rogers, 2021; Burns, 2022). Zur Verortung und räumlichen Verfolgung von Strafgefangenen werden im Schrifttum zudem vor allem RFID-basierte Ansätze beschrieben (Guercio, 2021; Mulholland et al., 2008; Purwito et al., 2018; Rogers, 2021; Xiao & Xiong, 2013).[2] Mit Blick auf den Bereich der »Computer Vision« kann weiters auf einen technischen Beitrag aus Kanada verwiesen werden (Bouachir et al., 2018) in dem die Entwicklung eines Systems zur Erkennung von Selbstmordversuchen beschrieben wird. Ähnliche Aktivitäten gibt es auch in Deutschland, wo sich z. B. das nordrhein-westfälische Justizministerium mit der Entwicklung künstlicher Intelligenz im Strafvollzug zur Suizidprävention befasst.[3] Zudem lassen sich einzelne rechtliche (bzw. ethische) Aufsätze finden (Esser & Reißmann 2019; Brobst,

1 Das Forschungsprojekt wurde gefördert vom Bundesministerium für Finanzen (BMF) und der Österreichischen Forschungsförderungsgesellschaft (FFG); KIRAS Sicherheitsforschung, Projekt Nr: 879744; das Konsortium besteht aus den folgenden Institutionen: Technische Universität Wien (Institut 193/1, Computer Vision Lab); Bundesministerium für Justiz (Bedarfsträger); CogVis Software und Consulting GmbH; PKE Holding AG; Research Institute AG & Co KG (Digital Human Rights Center); siehe auch https://www.kiras.at/gefoerderte-projekte/detail/kiis-kuenstliche-intelligenz-im-strafvollzug.

2 RFID steht für »Radio Frequency Identification«.

3 Siehe Landesregierung Nordrhein-Westfalen, »Einsatz künstlicher Intelligenz im Justizvollzug zur Suizidprävention« (2019.10.22.); https://www.land.nrw/pressemitteilung/einsatz-kuenstlicher-intelligenz-im-justizvollzug-zur-suizidpraevention; die Testimplementierung findet demnach in Düsseldorf statt und wird von der deutschen Firma Fusion Systems GmbH durchgeführt. Siehe auch NDR (17.05.2022). »Forschungsprojekt: Mehr Sicherheit in Gefängnissen durch KI?«; https://

2018). In der Regel wird in den genannten Beiträgen aber primär von technischen Entwicklungen und künftigen Anwendungen gesprochen. Nur wenige Beiträge behandeln auch den tatsächlichen Echtbetrieb in Strafvollzugsanstalten (EUROPRIS, 2020).

2.2 Einschlägige Vorstudien aus Österreich

Im Gegensatz dazu zeichnet sich das KIIS-Projekt durch eine Reihe nennenswerter Vorstudien aus.[4] Mit Bezug auf den Bereich der Justiz sind vor allem die Projekte SIMSTRA (Suizidprävention im österreichischen Strafvollzug, 2018)[5] und MaLeStV (Maschinelles Lernen von Bewegungsmustern im Strafvollzug, 2021) zu nennen.[6] Das Projekt SIMSTRA hatte die automatisierte Erkennung von Bewegungsabläufen zum Ziel, die im Fall eines Suizidversuchs durch Erhängen am Fenster typischerweise vorkommen könnten. Dazu zählen Szenarien wie z. B., eine Person, die sich (zu) lange am Fenster aufhält, das Fenster öffnet oder schließt, oder sich auf einen Stuhl stellt.[7] Im Projekt MaLeStV wurde der Fokus auf die allgemeine Erkennung von selbst- und fremdgefährdendem Verhalten erweitert.[8] Die Verhaltensanalyse wurde in beiden Studien anhand von 3D-Tiefendaten mit eigens entwickelten (visuellen) Sensoren, durchgeführt.

3 Beschreibung des Projekts

3.1 Technische Spezifikation und Zielsetzung

Die genannten Ansätze wurden im Fall des KIIS-Projekts weiterentwickelt. Die im Projekt zum Einsatz kommende Sensoreinheit wird als »Compact Tri-modal CAmera uniT« (CTCAT) bezeichnet und besteht grundsätzlich aus einer RGB-Bild-, einer Thermalbild- sowie einer 3D-Tiefenbild-Komponente (Strohmayer & Kampel, 2022). Aufgrund der Tiefen- und Thermaldaten eignet sich die Sensoreinheit für eine Reihe an Aufgaben, die von anderen Geräten, die lediglich RGB-Bilddaten erfassen, nicht effektiv erfüllt werden können. So dienen die Tiefendaten vor allem der robusteren Lokalisierung von Personen im Raum; die Thermaldaten erleichtern wiederrum die Detektion von Personen bei schwierigen Lichtverhältnissen.[9] Die CTCAT verfügt zudem über eine eigene Recheneinheit; diese ermöglicht die sensornahe Ausführung der Rechenprozesse für die Verhaltenserkennung (on-board). Auf diese Weise kann die Übertragung der RGB-Bilddaten in den Kontrollraum auf sicherheitskritische Szenarien reduziert und die Möglichkeit des (unbefugten) Zugriffs auf (sensible) Daten eingeschränkt werden.[10] Für die Verhaltensanalyse kommen Methoden des maschinellen Lernens auf der Grundlage von neuronalen Netzwerken (sog. Deep Learning) zum Einsatz; zudem handelt es sich um eine Form des überwachten Lernens (Supervised Learning).

Die KI-Anwendung der Sensoreinheit besteht aus zwei Hauptkomponenten: einer 3D-Objekt- bzw. Personenerkennung zur Extraktion semantisch wertvoller Merkmale und Muster aus den Bildern[11] und einer Mehrfachobjektverfolgung zur Erfassung von Objekten bzw. Personen und deren Bewegungslinien (Trajektorien) im Raum.[12] In Verbindung mit den bildgebenden Sensoren wird im Rahmen

www.ndr.de/nachrichten/niedersachsen/Forschungsprojekt-Mehr-Sicherheit-in-Gefaengnissen-durch-KI,ki160.html.

4 Im Kern gehen diese auf die Forschungsaktivitäten des *Computer Vision Lab (CVL)* der Technischen Universität (TU) Wien und das Spin-Off »Cogvis« zurück; siehe https://cvl.tuwien.ac.at/; siehe; https://cogvis.ai/.

5 Unveröffentlichter Projektbericht: SIMSTRA – Suizidprävention im österreichischen Strafvollzug – Ansätze zur Optimierung (2018); durchgeführt von CogVis und VICESSE in Kooperation mit dem BMJ.

6 Projekt MaLeStV – Maschinelles Lernen im Strafvollzug, 2021; KIRAS/FFG, Projekt-Nr.: 873495.

7 Im Rahmen des Projekts wurde über einen Zeitraum von 81 Tagen ein Sensor in einem Haftraum in einer niederösterreichischen Justizanstalt installiert, um alle Bewegungen als Lerngrundlage für das System aufzuzeichnen; Projektbericht SIMSTRA (2018).

8 Für die Entwicklung des Systems wurde zunächst ein eigener Datensatz mit gestellten Szenen erstellt. IPT (Identity Preserved Tracking) Datensatz; https://cvl.tuwien.ac.at/research/cvl-databases/ipt-dataset/. In weiterer Folge wurden Echtdaten einer Wiener Justizanstalt (Raucherzimmer, Werkhalle, 2er Zimmer; Aufnahmedauer: 2 Wochen) sowie einer niederösterreichischen Justizanstalt (Gemeinschaftsraum, Einzelzimmer; Aufnahmedauer: 1 Monat) erhoben.

9 So z. B. bei Rettungseinsätzen in Umgebungen mit schlechter Sicht oder Fußgängererkennung bei Nacht (Hayato et al. 2009; Zhilu & Xinming, 2019; Krotosky & Trivedi, 2007).

10 Durch dieses Edge-Computing und die Reduktion der visuellen Daten auf sicherheitskritische Szenarien wird im Projekt KIIS von technischer Seite die These vertreten, dass die CTCATs in ihrer Anwendung datenschutzfreundlicher sein könnten als herkömmliche Videoüberwachungssysteme.

11 Mit 3D-Personenerkennung ist hier nicht die Identifizierung einer natürlichen Person im juristischen Sinne, sondern das Erkennen (Detektieren) der Silhouette einer Person im Raum gemeint. Die 3D-Personenerkennung nutzt eine »MobileNet v2 Backbone-Architektur« (Sandler et al., 2018; Lin et al., 2017).

12 Das Multi Object Tracking (MOT) basiert wiederum auf der sogenannten Transformer-Architektur und deren Weiterentwicklungen (Meinhardt et al., 2022; Vaswani et al., 2017).

des Projekts zudem der Einsatz von Wearables (in Form einer Smartwatch)[13] getestet. Neben einer räumlich-örtlichen Positionsmessung werden über diese Geräte auch Puls, Temperatur- und Beschleunigungsdaten von Insassen erhoben.[14] Die gesammelten Informationen werden sodann mit den Bilddaten der CTCATs verknüpft (fusioniert) und zur Detektion und Lokalisierung von Personen im Raum, der Erkennung von Posen (stehend, sitzend, liegend), der Blickrichtung und der Bewegungsgeschwindigkeit sowie der Detektion von Interaktionen (zwischen zwei oder mehreren Personen) verwendet.[15] Auf dieser Basis soll das KI-System in weiterer Folge selbst- und fremdgefährdendes Verhalten (wie z. B. körperliche Kontakte mit Gewalthintergrund oder gesundheitsrelevante Ereignisse wie Suizidversuche) möglichst frühzeitig erkennen und melden. Ziel ist das Vollzugspersonal zu entlasten, indem das KI-basierte System sicherheitskritische Szenarien vorselektiert und die Aufmerksamkeit auf bestimmte Ereignisse lenkt. Auf diese Weise soll die Technologie zur Aufrechterhaltung der Sicherheit und Ordnung in Strafvollzugsanstalten beitragen.

3.2 Test-Settings und Datenerhebung

Im Zuge des Projekts wurden im Zeitraum zwischen April 2022 und März 2023 in zwei Strafvollzugsanstalten im Raum Wien-Niederösterreich lokale Test-Settings installiert. In Justizanstalt A, welche der Vollziehung von Untersuchungshaft, Strafhaft und Maßnahmenvollzug männlicher Jugendlicher und junger Erwachsener dient, wurde in zwei Freizeiträumen für den Zeitraum von 25 Tagen jeweils eine CTCAT zur Aufzeichnung von Wärme- und Tiefenbilddaten angebracht.[16] In Justizanstalt B, die männliche Strafgefangene mit kurzen bis mittellangen Freiheitsstrafen beherbergt, kamen neben zwei CTCATs auch Wearables (in Form von Smartwatches) zum Einsatz. Die Montage und Umsetzung erfolgte über 30 Tage im Unternehmerbetrieb (inkl. Raucherraum) der Justizanstalt.[17] Die Wearables wurden jeweils drei Insassen aus einer vorab definierten Gruppe von sechs Freiwilligen an Arbeitstagen bei Betreten des Unternehmerbetriebs vom diensthabenden Personal ausgehändigt und nach Beendigung der Arbeitszeit (von ca. sechs Stunden) an dasselbe retourniert.[18] Um eine ausreichende Menge von Trainingsdaten zu gewährleisten, wurden zusätzlich öffentlich verfügbare Datensätze[19] sowie gestellte Szenen aus dem Labor der TU Wien[20] und sogenannte synthetische Daten[21] in die Entwicklung der Algorithmen miteinbezogen. Das entstehende neuronale Netzwerk wurde anschließend auf der Sensoreinheit gespeichert und durch das dort laufende Programm ausgeführt.

13 Es wurde das Produkt »Bangle.js 2« verwendet; siehe: https://shop.espruino.com/banglejs2.

14 Während die Positionsdaten über eigens installierte Bluetooth-Messstationen (BLE location gateways) erfasst und auf einem (nur dem Projektteam zugänglichen) Server in der Justizanstalt gespeichert wurden, erfolgte die Speicherung der übrigen Daten direkt auf den tragbaren Geräten. Eine Verarbeitung von Luftdruck, Magnetometer und GPS-Daten wäre technisch ebenfalls möglich, wurde im Projekt aber nicht verfolgt.

15 Dies umfasst nicht nur die Entwicklung einer geeigneten Netzwerkarchitektur, sondern auch die Erstellung, Aufzeichnung und Annotation von Datensätzen.

16 Insgesamt kann von 27 betroffenen Insassen ausgegangen werden. Aus technischen Gründen wurden die Wearables in dieser Justizanstalt nicht getestet.

17 Im Zeitraum der Erhebung waren 88 Insassen im Unternehmerbetrieb zugeteilt und somit von der Datenverarbeitung der bildgebenden Sensoren betroffen.

18 Die gewählte Vorgehensweise ist als Maßnahme zur Gewährleistung der Pseudonymisierung zu verstehen. Im Schrifttum wird eine Gruppengröße von mehr als 5 Personen empfohlen. Siehe hierzu die Entscheidung der Datenschutzkommission (DSK, nunmehr DSB) vom 22.05.2013, GZ: K213.180/0021-DSK/2013; siehe außerdem § 79e Abs. 4 BDG (Beamten-Dienstrechtsgesetz 1979) im Hinblick auf Kontrollmaßnahmen gegenüber öffentlich Bediensteten.

19 Dabei wurden die Datensätze MIPT (2021) und TRISTAR (2023) verwendet. Siehe auch Heitzinger & Kampel (2021). Beide bestehend aus Indoor-Aufnahmen in den Räumlichkeiten der TU Wien sowie privaten Räumlichkeiten der technischen Entwickler (wie z. B. Büroräume, Flurszenen, Pausen- und Raucherräume, Küchen- und Wohnbereiche, Flur und Treppenhausbereiche, Konferenzräume, Eingangsbereiche). Zudem wurde der HARM (Human Actions Ranging from Mellow to Malevolent) Datensatz verwendet (Publikation in Arbeit); dieser bietet zusätzlich Informationen, um anzuzeigen, ob Handlungen freundlich oder aggressiv sind; siehe weiterführend auch https://cvl.tuwien.ac.at/category/research/cvl-databases/.

20 Insgesamt wurden etwa 35 »Gewaltszenen« gestellt und für das Machine Learning verwendet.

21 Im vorliegenden Fall wurde hierfür das Archive of Motion Capture As Surface Shapes (AMASS) verwendet; siehe https://amass.is.tue.mpg.de/. Synthetische Bild- bzw. Videodaten sind computergenerierte Daten, die reale Bilder/Videos simulieren. Im synthetischen Ansatz werden reale Personen mittels 3D-Scans erfasst (siehe hierzu https://alicevision.org/) und mit einer Wärmebildtextur sowie einem künstlichen Skelett (rigging) ausgestattet (siehe hierzu https://www.blender.org/). Anschließend können beliebige Animationen auf das entstandene Modell angewendet und gerendert werden. Im Gegensatz zu Aufnahmen aus echten Szenarien weisen diese Bilder kein – üblicherweise durch physikalische Effekte oder Sensorfehler entstehendes – Rauschen auf. Um eine realistischere Darstellung zu erzielen, wird das Fehlen von Rauschen künstlich durch die Anwendung eines Diffusionsmodells simuliert (Saharia et al, 2022).

4 Rechtliche Zulässigkeit

4.1 Anwendbares Recht

Die beschriebene KI-Anwendung wirft eine Reihe an Rechtsfragen auf. Im Folgenden geht es vor allem um die Klärung, ob bzw. inwiefern eine entsprechende Rechtsgrundlage für den künftigen Einsatz des in Rede stehenden Systems vorliegt. In der juristischen Beurteilung der KI-basierten Überwachung wird im vorliegenden Fall zunächst davon ausgegangen, dass es zu einer Verarbeitung personenbezogener Daten (iSd Art. 4 Nr. 1 und Nr. 5 DSGVO) und somit zu sogenannten »Informationseingriffen« (Berka, 1999) in die über Art. 8 der EMRK,[22] sowie Art 7 und Art 8 der GRC[23] geschützte Sphäre der Betroffenen kommt. Aufgrund der Thermal- und Pulsdaten sowie der Detektion gesundheitsrelevanter Szenarien liegt zudem eine Qualifikation als besondere Kategorie personenbezogener (sog sensibler) Daten vor (Art. 9 DSGVO). Die Verarbeitung personenbezogener Daten zum Zweck der Strafvollstreckung ist jedoch nicht vom sachlichen Anwendungsbereich der DSGVO umfasst (siehe Art. 2 Abs. 2 lit d DSGVO). Derartige Datenverarbeitungen werden stattdessen in der EU-Richtlinie für Justiz und Inneres (JI-RL) geregelt;[24] diese wurde in Österreich wiederum über das 3. Hauptstück (§ 36 ff) des Datenschutzgesetzes (DSG) umgesetzt.[25] Nach Maßgabe des anwendbaren Rechts sind Eingriffe staatlicher Behörden in die genannten Schutzbereiche laut Art. 8 Abs. 2 EMRK bzw. Art. 8 Abs. 2 GRC sowie Art. 8 Abs. 1 der JI-RL bzw. § 38 DSG nur statthaft, wenn diese auch gesetzlich vorgesehen und zur Erfüllung der Aufgabe der Strafvollstreckung notwendig bzw. erforderlich sind.[26] Die Verfassungsbestimmung des § 1 DSG konkretisiert zudem, dass »der Eingriff in das Grundrecht [auch im Falle zulässiger Beschränkungen] jeweils nur in der gelindesten, zum Ziel führenden Art vorgenommen werden [darf].«[27] In diesem Sinne wird in Folge die Frage nach einer geeigneten Rechtsgrundlage sowie auch die Verhältnismäßigkeit (Eignung, Erforderlichkeit und Notwendigkeit) des Grundrechtseingriffs behandelt.

4.2 Rechtsgrundlagen im StVG

Wie dargelegt ist als Ausgangspunkt für die Prüfung der Zulässigkeit die JI-RL bzw. deren Umsetzung über das DSG heranzuziehen. Die Bestimmungen des DSG werden in Österreich wiederum durch das Strafvollzugsgesetz (StVG)[28] als *lex specialis* konkretisiert (Drexler & Weger, 2018, § 15a). § 15a Abs. 1 StVG regelt den »Einsatz der Informationstechnik« im Strafvollzug, wobei sich die Bestimmung dem Wortlaut nach auf die allgemeine »Vollzugsverwaltung« bezieht. Die Vollzugsverwaltung kann sich demnach für Zwecke des Strafvollzuges der automationsunterstützten Datenverarbeitung bedienen. Die zuständigen Stellen dürfen dabei personenbezogene Daten über »strafbare Handlungen der Insassen« oder die »vollzugsrelevanten Lebensumstände in und außerhalb der Justizanstalt« verarbeiten. Dies gilt auch für die Verarbeitung sensibler (besondere Kategorien) personenbezogener Daten (i.S.d. § 39 DSG), sofern dies »unbedingt erforderlich« ist. Die Verarbeitung gilt jedenfalls nur dann als zulässig, wenn sie auch verhältnismäßig ist und »wirksame Maßnahmen zum Schutz der Rechte und Freiheiten der betroffenen Personen« getroffen werden.

Eine hinreichend bestimmte Ausgestaltung dieser Maßnahmen i.S.d. datenschutzrechtlichen Grundsätze gem. § 37 DSG (z. B. Zweckbindung, Speicherbegrenzung, Richtigkeit, Rechenschaftspflicht) fehlt im StVG jedoch. Vor allem der Umstand, dass weniger eingriffsintensive Datenverarbeitungen im StVG gesondert geregelt werden, lässt in systematischer Zusammenschau darauf schließen, dass der Einsatz des in Rede stehenden KI-Systems nicht ausschließlich über § 15a Abs. 1 StVG gerechtfertigt werden kann.[29] Es handelt sich bei § 15a Abs. 1 StVG um eine Norm zur Regelung administrativer Prozesse des Vollzugs, die andere – mitunter grundrechtsinvasivere Verarbeitungen – nicht per se umfasst.[30] Dabei kann exemplarisch auch auf die legislative Debatte rund um die StVG-Novelle 2019 verwiesen werden, in welcher eine Ausgestaltung des Gesetzes für den Einsatz von Bodycams (§ 102b Abs. 2a StVG) sowie die Einführung der Möglichkeit zur Videotelefonie (über § 24 Abs. 3 Z 1

22 Konvention zum Schutze der Menschenrechte und Grundfreiheiten (EMRK), StF: BGBl. Nr. 210/1958.

23 Charta der Grundrechte der Europäischen Union (GRC), (2012/C 326/02), 26.10.2012, ABl C 326/391. Zur spezifischen rechtstechnischen Inkorporation der GRC in Österreich siehe VfGH 14.03.2012, VfSlg 19.632/2012.

24 RL für Justiz und Inneres (JI-RL) (EU) 2016/680; ABl L 119/89.

25 Datenschutzgesetz (DSG); StF: BGBl. I Nr. 165/1999.

26 Siehe hierzu auch das in Art. 18 des österreichischen Bundes-Verfassungsgesetz (B-VG) verankerte Legalitätsprinzip (»Gesetzesvorbehalt«); StF: BGBl. Nr. 1/1930.

27 Zum spezifischen rechtstechnischen Aufbau des österreichischen Verfassungsrechts (fehlendes Inkorporationsgebot etc.), siehe Lachmayer & Rothmann (2020) 472 ff.

28 Strafvollzugsgesetz (StVG), StF: BGBl. Nr. 144/1969.

29 Siehe z. B. die in §§ 15a Abs. 2 ff StVG genannten (zulässigen) Fallgruppen; siehe auch § 102b StVG zur Videoüberwachung oder § 156b ff StVG zum elektronisch überwachten Hausarrest.

30 Briem (2022) stellt ähnliche Regelungsdefizite auch in der österreichischen Finanzverwaltung für automatisierte Einzelentscheidungen und Profiling im steuerlichen Abgabeverfahren fest.

StVG) vorgeschlagen wurde.[31] Dabei wurde ebenfalls keine Subsumtion unter die bestehende Regelung des § 15a StVG erwogen.

4.2.1 Bildgebende Sensoren als Form der Videoüberwachung

Von spezieller Bedeutung ist weiters vor allem § 102b StVG zur Regelung der »Videoüberwachung« in Strafvollzugsanstalten.[32] In § 102b Abs. 1 StVG werden »technische Mittel zur Bildübertragung« als Echtzeitüberwachung bezeichnet und von »technischen Mitteln zur Bildaufnahme« gem. § 102b Abs. 2 StVG unterschieden. Die Formulierung »technische Mittel« ist technologieneutral und umfasst in diesem Sinne Tiefen-, Thermal- und RGB-Bilddaten gleichermaßen, womit auch die bildgebende multimodale Sensoreinheiten (CTCAT) des in Rede stehenden KI-Systems unter diese Bestimmung fallen.[33] Weiters kann aus der Formulierung des § 102b Abs. 1 StVG abgeleitet werden, dass Formen der »Echtzeitüberwachung« zur »Sicherung der Abschließung der Strafgefangenen von der Außenwelt und zur Sicherung der Ordnung in der Anstalt«, grundsätzlich auf allen Anstaltsflächen (Räumlichkeiten und Höfe) zulässig sind.[34] § 102b Abs. 2 sieht in Abgrenzung dazu vor, dass »Videoüberwachung« (aufgrund des schwerwiegenderen Eingriffs in die Privatsphäre der Betroffenen durch die Speicherung) nur in bestimmten Bereichen eingesetzt werden darf;[35] im Wesentlichen sind dies jene Anstaltsflächen, die nicht dem höchstpersönlichen Lebensbereich dienen. § 102b Abs. 3 StVG enthält schließlich ein ausdrückliches Verbot; Videoüberwachung in »gewöhnlichen Hafträumen, gemeinschaftlichen Sanitärräumen und Räumen, die ausschließlich dem Aufenthalt von Vollzugsbediensteten vorbehalten sind« ist demnach unzulässig.[36] In den Erläuterungen wird allerdings auch ausgeführt, dass Echtzeitüberwachung in besonderen Hafträumen (z. B. besonders gesicherten Zellen nach § 103 Abs. 2 Z 4) zulässig sein kann. In diesen besonderen Hafträumen soll die Echtzeitüberwachung vor allem die Gefährdung des Lebens und der Gesundheit der Betroffenen verhindern.[37] Die mit den CTCATs verbundene Form der Datenverarbeitung (Thermaldaten, automatisierte Analyse), kann in ihrer Eingriffsqualität allerdings nicht mit herkömmlicher Echtzeitüberwachung gleichgesetzt werden. Auch dann, wenn das System so eingesetzt wird, dass die Analyse in Echtzeit ohne Speicherung erfolgt, kein Bildmaterial bereitgestellt wird und das Vollzugspersonal über sicherheitskritische Szenarien lediglich via Textmeldung (Alarmcode) informiert wird, stellt die KI-basierte Verhaltensanalyse eine technische Konkretisierung der Überwachungsprozesse dar. Anstelle einer sporadischen menschlichen Einsicht kommt es zu einer permanenten Analyse und Bewertung des Verhaltens (Rothmann & Vogtenhuber, 2013). Dabei werden bestimmte Verhaltensfiguren für das System als sicherheitskritisch bzw. verdächtig festgelegt und könnten über die automationsunterstütze Analyse eine »erhebliche Beeinträchtigung« für Betroffene im Sinne des § 41 DSG (Automatisierte Entscheidungsfindung im Einzelfall) entfalten. Folglich scheidet eine Klassifizierung des KI-System als reine Echtzeitüberwachung aus. Weder über § 15a noch über § 102b Abs. 1 StVG ließe sich der künftige Einsatz der KI-basierten Überwachung rechtfertigen.[38] Dies gilt insb. für persönliche Hafträume, besonders gesicherte Zellen und höchstpersönliche Bereiche wie Nassräume.

4.2.2 Einsatz der Wearables

Darüber hinaus ist für das in Rede stehende KI-System eine Kombination der CTCATs mit Wearables angedacht. Für den Einsatz der Smartwatches zur Position- und Pulsmessung gibt es im StVG derzeit keine normative Grundlage. Eine Rechtfertigung über § 15a ff StVG scheitert aufgrund der Unbestimmtheit der Norm. Um den Anforderungen des Legalitätsprinzips gem. Art. 18 B-VG sowie den Grundsätzen des Datenschutzrechts gem. § 37 DSG zu entsprechen, bedarf es einer hinreichend konkreten Ausgestaltung der Rechtsgrundlage (Berka, 2012). Eine solche sollte zunächst

31 Ministerialentwurf StVG-Novelle-2019; 166/ME XXVI. GP; siehe die Textgegenüberstellung unter https://www.parlament.gv.at/dokument/XXVI/ME/166/fname_764610.pdf.

32 Die Gründe für den Einsatz von Videoüberwachung im Vollzug liegen in der Sicherung der Abschließung (also va die Verhinderung von Fluchten) sowie allgemein der Gewähr der Sicherheit und Ordnung in den Anstalten (Drexler & Weger, 2018, § 102b); vgl. Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP; siehe Ausschussbericht (parlament.gv.at).

33 Auch Infrarot-Bilddaten würden darunterfallen; Tonaufnahmen sind jedoch nicht zulässig.

34 Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP.

35 Diese umfassen den Eingangsbereich, die Besucher*innen – und Vernehmungszonen, die Gänge im Gesperre, Örtlichkeiten, die der Beschäftigung und dem Aufenthalt von Strafgefangenen außerhalb der Hafträume dienen, und vergleichbare Bereiche sowie die Außengrenzen der Anstalt. Die Formulierung ist weit zu verstehen; vgl. Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP.

36 Soweit die gewöhnlichen Hafträume über sanitäre Einrichtungen verfügen, wird Videoüberwachung auch in diesem Bereich des Haftraumes als nicht zulässig erachtet; Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP.

37 Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP.

38 Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP.

vor allem angeben, unter welchen Umständen und Voraussetzungen der Einsatz der Technologie zulässig ist. Weiters sind wirksame Maßnahmen zum Schutz der Rechte und Freiheiten der Betroffenen vorzusehen (Auskunftsrechte, Protokollierungs- und Löschpflichten etc.). Letztere sind umso bedeutender, wenn es über das KI-System zu automatisierten Einzelentscheidungen kommen würde.

4.2.3 Automatisierter Entscheidungsfindung im Einzelfall

Die »automatisierte Entscheidungsfindung im Einzelfall« wird in Art. 11 JI-RL bzw. § 41 DSG geregelt. Automatisierte Entscheidungen im Einzelfall sind demnach zwar nicht per se verboten (Thiele & Wagner, 2020, § 41 Rz. 7), für Fälle, die für Betroffene »nachteilige Rechtsfolgen haben oder sie erheblich beeinträchtigen können«, braucht es allerdings eine hinreichend bestimmte Rechtfertigung im Gesetz.

Im Fall sensibler Daten sind zusätzliche Schutzmaßnahmen zu ergreifen. Aufgrund des höheren Risiko- und Gefährdungspotenzials unterliegen automatisierte Entscheidungen im Einzelfall auch strengeren Anforderungen an die Rechtmäßigkeit (Thiele & Wagner, 2020, § 41 Rz. 6 ff).[39]

Eine »erhebliche Beeinträchtigung« kann bereits vorliegen, wenn eine Person in ihrer »persönlichen Entfaltung nachhaltig gestört« wird (Buchner, 2018, Art. 22 Rz. 26).[40] Dies wäre z. B. der Fall, wenn vermehrte Alarmmeldungen des Systems der vorzeitigen Beendigung einer besonderen Sicherheitsmaßnahme entgegenstehen oder überhaupt erst zur Ergreifung besonderer Sicherheitsmaßnahmen führen. Darüber hinaus ist gem. § 41 Abs. 3 DSG die automatisierte Diskriminierung natürlicher Personen anhand besonderer Kategorien personenbezogener Daten ausdrücklich verboten. Führt die Analyse der Pulsfrequenz der Strafgefangenen zu nachteiligen Rechtsfolgen oder einer erheblichen Beeinträchtigung wäre der Einsatz des KI-System unzulässig. Es geht allerdings nur um »ausschließlich« automatisierte Entscheidungen, »ohne jegliches menschliche Eingreifen«.[41] Entscheidungen, die letztlich von natürlichen Personen getroffen werden, fallen nicht unter diese Regelung (Thiele & Wagner, 2020, § 41 Rz. 22). Die Beurteilung der Grenze kann im Einzelfall eine schwierige Aufgabe sein (Briem, 2022).[42] Die lediglich routinemäßige Einbeziehung eines Menschen oder symbolischer Gesten werden aber wohl noch nicht zur Ausnahme von der Regelung führen (Artikel-29-Datenschutzgruppe, 2018). Ob im Vollzugsalltag eine Einzelentscheidung im Sinne des Datenschutzrechts vorliegt, hängt letztlich auch von der künftigen (technischen und organisatorischen) Implementierung der Technologie (und etwaigen Schutzmaßnahmen) ab. Dabei ist auch das sozio-technische Phänomen des »automation bias« verstanden als »over-reliance on algorithmic advice« zu beachten (Alon-Barkat & Busuioc 2023; Logg et al. 2019); derartige Effekte wären in der datenschutzrechtlichen Beurteilung und Folgenabschätzung zu berücksichtigen. Dessen ungeachtet handelt es sich aber wohl jedenfalls um ein System im Sinne der KI-Verordnung, das »mit unterschiedlichem Grad an Autonomie operieren kann« und »Ergebnisse wie Vorhersagen, Empfehlungen oder Entscheidungen« hervorbringen soll, die das physische oder virtuelle Umfeld beeinflussen.[43]

4.3 Regulierung durch die KI-Verordnung

Als Reaktion auf die voranschreitende technische Entwicklung im Bereich der Künstlichen Intelligenz wurde in jüngster Zeit eine Reihe von Rechtsakten, Leitfäden, Richtlinien und anderen Dokumenten zur rechtlichen sowie ethischen Regulation veröffentlicht.[44] Die folgende Ausführungen konzentrieren sich auf die europäische KI-Verordnung (Gesetz über Künstliche Intelligenz). Da sich die Verordnung derzeit noch im Gesetzgebungsverfahren (gem. Art. 294 AEUV) befindet, stützt sich die Analyse auf die aktuelle Fassung vom 16.4.2024.[45]

39 Vgl. EuGH vom 07.12.2023, C634/21 (Schufa Holding), Rz. 57. Dies umfasst z. B. auch das Betroffenenrecht auf Darlegung des eigenen Standpunkts sowie zusätzliche Informationspflichten und Auskunftsrechte.

40 Vgl. EG 71 DSGVO; siehe auch EuGH vom 07.12.2023, Rs C634/21 (Schufa Holding), Rz. 48 ff.

41 Vgl. EG 71 DSGVO; EuGH vom 07.12.2023, Rs C634/21 (Schufa Holding), Rz. 45 ff.

42 Siehe hierzu auch die nicht-rechtskräftige Entscheidung des BVwG aus Dezember 2020 zum Arbeitsmarktchancen Assistenz-System (»AMAS«) des Arbeitsmarktservice (»AMS«), in der das BVwG zur Ansicht kommt, dass keine automatisierte Einzelentscheidung i.S.d. Art. 22 DSGVO vorliegt; BVwG 18.12.2020, W256 2235360-1/5E.

43 Vgl. Art. 3 Abs. 1 Nr. 1 KI-VO; siehe Abänderungen des Europäischen Parlaments vom 14. Juni 2023, (P9_TA(2023)0236).

44 Siehe z. B. die »European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment«, European Commission for the Efficiency of Justice (CEPEJ) (2018); siehe auch »Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Committee on Artificial Intelligence (CAI)«, Council of Europe (2023).

45 Siehe European Parliament, 16.4.2024, Corrigendum (cor01), https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_EN.pdf; siehe auch https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html, sowie https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-

Art. 2 der KI-VO regelt den Anwendungsbereich, welcher gem. Art. 2 Abs. 1 u. a. Anbieter bzw. Betreiber, die KI-Systeme in der Union in Verkehr bringen oder in Betrieb nehmen, umfasst. Während KI-Systeme, die ausschließlich für militärische Zwecke entwickelt oder verwendet werden, ausgenommen sind (Art. 2 Abs. 3 KI-VO), fallen Behörden des Strafvollzugs jedoch darunter (Art. 3 Nr. 46 KI-VO).[46] Weiters ist die Legaldefinition in Art. 3 Nr. 1 beachtlich; ein »System der künstlichen Intelligenz« (KI-System) ist im Vorschlag der Kommission als »eine Software« definiert »die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren«.[47] Die in Anhang I genannten Techniken und Konzepte umfassen auch Methoden des maschinellen Lernens wie sie im Fall des in Rede stehenden KI-Systems zur Anwendung kommen; der sachliche Anwendungsbereich der VO wäre damit erfüllt.

4.3.1 Biometrische Daten und Emotionserkennung

Für die Beurteilung des Systems anhand der KI-VO ist in Folge zu klären, ob das in Rede stehende KI-System sog »biometrische Daten« im Sinne von Art. 3 Nr. 34 KI-VO (bzw. § 36 Abs. 2 Z 13 DSG) verarbeitet.[48] Gemeint sind damit Daten, die eine »eindeutige Identifizierung [einer] natürlichen Person ermöglichen«, wobei exemplarisch auf »Gesichtsbilder oder daktyloskopische Daten« verwiesen wird. Eine derartige Verarbeitung soll im vorliegenden Fall zwar nicht stattfinden; das System zielt auf Basis der Puls- und Thermaldatenverarbeitung nicht auf eine Identifizierung der Personen; es reicht dem Wortlaut nach jedoch, dass die Daten eine solche »ermöglichen«.[49] Durch den Einsatz der in Rede stehenden Technologie kann es im Sinne dieser Bestimmung zu einer »spezifischen Verarbeitung von physischen, physiologischen sowie verhaltensbezogenen Signalen« einer natürlichen Person kommen.[50]

Findet über das in Rede stehende KI-System eine Verarbeitung biometrischer Daten statt, steht in weiterer Folge eine Klassifizierung als »Emotionserkennungssystem« i.S.v. Art. 3 Nr. 39 KI-VO zur Diskussion.[51] Als Emotionserkennungssystem gilt ein KI-System »das dem Zweck dient, Emotionen oder Absichten natürlicher Personen auf der Grundlage ihrer biometrischen und biometriegestützten Daten festzustellen oder daraus abzuleiten«. Als Emotion wird in der einschlägigen Fachliteratur »ein komplexes Muster von Veränderungen« verstanden, »das physiologische Erregung, Gefühle, kognitive Prozesse und Verhaltensweisen einschließt, die in Reaktion auf eine Situation auftreten, welche ein Individuum als persönlich bedeutsam wahrgenommen hat« (Kleinginna & Kleinginna, 1981). Zudem lassen sich verschiedene Modelle zur Beschreibung sog primärer Emotionen finden (Zimbardo, 1995).[52] Für die Detektion von selbst- und fremdgefährdendem Verhalten scheinen vor allem die Emotionen Wut, Aggression und Traurigkeit relevant. Soll das KI-System nun also gewaltvolle Interaktionen oder Suizidversuche erkennen, liegt auch eine Feststellung von Geisteszuständen oder Absichten i.S.v. Art. 3 Nr. 39 der KI-VO vor.

4.3.2 Risiko-Klassifizierung

Für die Risiko-Klassifizierung des vorliegenden KI-Systems ist schließlich Art. 5 KI-VO zur Regelung »verbotener Praktiken« relevant. Dabei springt zunächst Art. 5 Abs. 1 lit f der KI-VO ins Auge, wenn dort von einer »Ableitung von Emotionen einer natürlichen Person« die Rede ist. Die Bestimmung adressiert allerdings nur den Arbeitsplatz und Bildungseinrichtungen; der Strafvollzug ist nicht umfasst. Weiters ist Art. 5 Abs. 1 lit g beachtlich, wenn es dort um

meps-adopt-landmark-law und https://artificialintelligenceact.eu/de/dokumente/.

46 Der Begriff »Strafverfolgungsbehörde« umfasst auch Stellen zur Strafvollstreckung.

47 Vgl. hierzu auch den deutlich anders lautenden Vorschlag für eine Definition des EP; das in Rede stehende KI-System des KIIS-Projekts kann unter beide Legaldefinitionen subsumiert werden.

48 Siehe auch die Legaldefinition des Begriffs in Art. 4 Nr. 14 DSGVO.

49 Hier ist u. a. auf technische Entwicklungen hinzuweisen, die darauf abzielen Personen über die Frequenz ihres Herzschlags (das Elektrokardiogramm, EKG) zu identifizieren (Camara et al., 2023); siehe auch https://www.technologynetworks.com/informatics/news/algorithm-makes-it-possible-to-identify-people-by-their-heartbeat-360001.

50 Gang, Handbewegungen, Augenbewegungen, Gehirnaktivitäten sowie die Stimme werden auch als »verhaltensbiometrische Merkmale« bezeichnet (Hanisch et al., 2023).

51 Wenn keine biometrischen Daten verarbeitet werden, kommt es in der KI-VO zu einer anderen Risikoklassifizierung. Ohne biometrische Daten wäre die KI-Anwendung kein Hochrisiko-System. Dies wäre bspw. der Fall, wenn die Komponente der Wearables, bzw. die damit verbundene Verarbeitung der Pulsdaten, nicht zur Anwendung kommt.

52 Nach Plutchik (1980) gibt es acht grundlegende (angeborene) Emotionen, die aus vier Gegensatzpaaren bestehen: Freude und Traurigkeit, Furcht und Wut, Überraschung und Vorwarnung, sowie Akzeptanz und Ekel. Daraus ergeben sich (als Synthese) wiederum Liebe, Unterwerfung, Ehrfurcht, Enttäuschung, Treue, Verachtung, Aggressivität, und Optimismus.

Formen der biometrischen Kategorisierung geht. Die Bestimmung gilt jedoch nicht für die »Kategorisierung biometrischer Daten im Bereich der Strafverfolgung«.

Fällt das System nicht unter die verbotenen Praktiken gem. Art. 5 der KI-Verordnung, ist in einem nächsten Schritt Art. 6 zur Regelung von Hochrisikosystemen heranzuziehen. Bei der Spezifizierung von Hochrisikosystemen verweist Art. 6 (2) auf eine Liste in Anhang III der Verordnung. Anhang III 1 (b) regelt KI-Systeme die zur »biometrische[n] Kategorisierung nach sensitiven oder geschützten Attributen oder Merkmalen« verwendet werden sollen; Anhang III 1 (c) behandelt zudem Systeme, die »zur Erkennung von Emotionen« verwendet werden sollen. Wenngleich das in Rede stehende KI-System möglicherweise keine biometrische Kategorisierung im Sinne der Verordnung vornimmt, so kommt wohl eine Klassifizierung unter Anhang III 1 (c) in Frage. Die Erkennung von Aggression und Suizidszenarien auf Basis biometrischer und sensibler personenbezogener Daten stellt unserer Ansicht nach vor allem im Kontext des Strafvollzugs eine Hochrisiko-Praktik dar. Für derartige Systeme gelten gem. Art. 8 ff KI-VO wiederum erhöhte Sicherheitsanforderungen (Implementierung von Risikomanagementsysteme und Daten-Governance Prozessen, technische Dokumentations-, Aufzeichnungs- und Transparenzpflichten etc.).

Diese Risikoklassifizierung unter Anhang III 1 (c) ist insofern bemerkenswert, da sich das Europäische Parlament im Rahmen der legislativen Debatten zur KI-Verordnung dahingehend geäußert hat, dass »[i]m Hinblick auf die wissenschaftliche Grundlage von KI-Systemen zur Erkennung von Emotionen, physischen oder physiologischen Merkmalen wie Gesichtsausdrücken, Bewegungen, der Pulsfrequenz oder der Stimme [...] ernsthafte Bedenken [bestehen].« Zu den größten Schwachstellen solcher Technologien gehören demnach »die begrenzte Zuverlässigkeit [...], die mangelnde Spezifität (physische oder physiologische Ausdrücke stimmen nicht eins zu eins mit Emotionskategorien überein) und die begrenzte Verallgemeinerbarkeit (die Auswirkungen von Kontext und Kultur werden nicht ausreichend berücksichtigt).« So können z. B. heftiges Gestikulieren oder häufiger Körperkontakt im Gespräch je nach Sozialisation als aggressiv und bedrohlich wahrgenommen werden, oder aber als normaler Teil der Kommunikation. Dem Parlament nach hätten Systeme zur Feststellung des emotionalen Zustands einer Person im Bereich der Strafverfolgung (und -vollstreckung) daher verboten werden sollen. In der aktuellen Fassung der KI-VO gelten derartige KI-Systeme jedoch »nur mehr« als Hochrisiko-Systeme.

5 Verhältnismäßigkeit

5.1 Prüfkriterien

Zur Beurteilung des KI-Systems soll schließlich auch der unions- und verfassungsrechtliche Grundsatz der Verhältnismäßigkeit diskutiert werden. Die Prüfung der Verhältnismäßigkeit beginnt üblicherweise mit der Frage nach einem legitimen öffentlichen Interesse; danach wird die Eignung der Mittel, die Erforderlichkeit (bzw. Notwendigkeit) sowie die Adäquanz (als Güterabwägung und »Verhältnismäßigkeit im engeren Sinn«) untergliedert (Berka, 2005; Khakzadeh-Leiler, 2011; Öhlinger & Eberhard, 2014; Bezemek, 2016).[53] Im Kern geht es um eine Abwägung der Notwendigkeit grundrechtlicher Einschränkungen (Berka, 2005; Öhlinger & Eberhard, 2014). Mitunter ist auch vom Übermaßverbot die Rede; rechtskonform sind jedenfalls nur verhältnismäßige Eingriffe (Öhlinger & Eberhard, 2014).

Den Stellenwert der Verhältnismäßigkeitsprüfung im Datenschutzrecht hat jüngst auch der EuGH betont.[54] Unter Rückgriff auf die bestehende Spruchpraxis führt dieser aus, dass die Verarbeitung personenbezogener Daten auf das absolut Erforderliche einzuschränken ist, wobei, wenn mehrere geeignete Maßnahmen zur Erreichung der verfolgten legitimen Ziele zur Verfügung stehen, die am wenigsten belastende zu wählen ist. Die betreffende Regelung muss zudem präzise Anforderungen und Voraussetzungen für die Anwendung aufstellen.[55] Auch § 102b Abs. 4 StVG schreibt vor, darauf zu achten, dass »Eingriffe in die Privatsphäre der Betroffenen die Verhältnismäßigkeit zum Anlass wahren«. Der Einsatz von Videoüberwachung (und somit auch der Einsatz der bildgebenden Sensoreinheit) im Strafvollzug ist demnach nur verhältnismäßig, wenn kein anderes taugliches Mittel besteht, das bei gleicher Effizienz weniger eingriffsintensiv wäre.[56]

53 Siehe auch Art. 52 Abs. 1 GRC; siehe auch OGH, 30.1.1997, 6 Ob 2401/96y.

54 EuGH, Rs C37/20 und C601/20 (Luxembourg Business Registers) vom 22. November 2022, Rz. 63 ff.

55 EuGH, Rs C37/20 und C601/20 (Luxembourg Business Registers) vom 22. November 2022.

56 Weniger eingriffsintensiv wäre z. B. die Echtzeitüberwachung; diese erlaubt aber nur eine Reaktion bei Eintritt eines schädigenden Ereignisses (z. B. Körperverletzung), ohne Beweissicherung; Erläuterungen (Ausschussbericht NR), 2089 der Beilagen XXIV. GP. Zur Wahrung der Verhältnismäßigkeit können zudem technische Maßnahmen zur Datenminimierung und Speicherbegrenzung dienen. Der OGH geht in seiner Rsp allerdings davon aus, dass auch im Fall einer (nicht erkennbaren) Kameraattrappe eine »schwerwiegende Beeinträchtigung der Privatsphäre« vorliegen kann; siehe hierzu OGH 28.03.2007, 6 Ob 6/06k,

Im Folgenden werden in der Prüfung die Eignung (der Mittel) sowie die Erforderlichkeit (des Eingriffs) herausgestellt, da diese für die Beurteilung des vorliegenden Sachverhalt von besonderer Bedeutung sind und mit weiteren technischen und kriminologischen Argumenten verknüpft werden können.[57] Die Diskussion der Verhältnismäßigkeit mündet schließlich in einer interdisziplinäre Diskussion und ethischen Abwägung zentraler Rechtsgüter und Positionen.

5.2 Eignung der Technologie: Genauigkeit und Robustheit

Von besonderem Interesse für die Verhältnismäßigkeitsprüfung ist das Kriterium der Eignung, also die Frage, ob es sich um ein taugliches Mittel handelt. Damit bietet der Begriff der Eignung einen Anknüpfungspunkt für empirische Analysen und die wissenschaftliche Überprüfung der tatsächlichen Wirkung der Technologie in den verschiedenen Settings (Rothmann, 2012). Geht man davon aus, dass die Effektivität des Systems vorliegt, kann dessen Einsatz in spezifischen Situationen (insb. mit erhöhtem Risiko der Selbstverletzung bzw. -tötung) gegenüber der persönlichen Intervall-Kontrolle oder herkömmlichen Formen der Videoüberwachung besser geeignet und somit auch verhältnismäßig sein (Esser & Reißmann, 2019). Die darauf basierende Annahme der Verhältnismäßigkeit setzt jedoch einen objektiven Nachweis der tatsächlichen Eignung des Systems voraus. Um diese evidenzbasiert beurteilen zu können, bedarf es entsprechender Untersuchungen, die in unabhängiger sowie methodisch-systematischer Weise (empirische) Fakten für die Evaluationen der Eignung des KI-Systems für einen definierten Zweck liefern (Rothmann, 2012). So ist hinsichtlich der Gewaltprävention im Fall von körperlichen Auseinandersetzungen häufig von Affekthandlungen und impulsiven Reaktionen (auf Provokationen oder vermeintliche Ungerechtigkeiten) auszugehen (Allard et al., 2008). Die befragten Inhaftierten zweifeln aufgrund dessen an der abschreckenden Wirkung des KI-Systems. Darüber hinaus erklären die Befragten in Bezug auf Videoüberwachung, dass es möglich sei, in andere Bereiche auszuweichen: »[W]enn man wirklich raufen will, geht man wo anders hin«[58]. Die Betroffenen bestätigen damit das in der kriminologischen Literatur bekannte Phänomen der Verdrängung bzw. Verlagerung (Repetto, 1976; Welsh & Farrington, 2002; Rothmann, 2012). Zudem ist darauf hinzuweisen, dass die aktuelle technische Entwicklung des in Rede stehenden KI-Systems auf Technology Readiness Level (TRL) 4 (von 9) einzustufen ist.[59] Die technische Machbarkeit (im Sinne einer hinreichend exakten Detektion von selbst- und fremdgefährdendem Verhalten im Kontext des Strafvollzugs) ist derzeit (noch) nicht bewiesen. Die Ergebnisse der Untersuchungen belegen (noch) nicht, dass die neue Technologie die vorab aufgestellten Anforderungen zur KI-basierten Verhaltensanalyse erfüllen kann. So werden von technischer Seite hinsichtlich der Detektionsraten der CTCATs Werte um etwa 80 % ausgewiesen.[60] Bei technisch schwierigen oder seltenen Posen liegen die Raten entsprechend niedriger. Die Tracking-Performance liegt, je nach verwendeter Metrik, bei rund 60–80 % (Heitzinger & Kampel, 2023; Stippel et al., 2023).[61] Für die Beurteilung der Eignung sind zudem die in die Entwicklung der Algorithmen einbezogenen Datensätze mit gestellten Szenen und synthetischen Daten von Bedeutung (Mittelstadt 2021). Wenn diese ein unrealistisches Bild des Vollzugsalltags liefern, kann dies zu einer systematischen Verzerrung der Ergebnisse (Bias) und letztlich zu diskriminierenden Schlussfolgerungen führen.[62]

5.3 Erforderlichkeit des Eingriffs: Bedarf und gelindere Mittel

Für die Prüfung der Verhältnismäßigkeit des Grundrechtseingriffs ist weiters dessen Erforderlichkeit und die Frage nach dem gelindesten Mittel relevant (Öhlinger & Eberhard, 2014). Es geht somit auch um den eigentlichen Bedarf (die Notwendigkeit) des grundrechtlichen Eingriffs sowie mögliche Alternativen. Zur Beurteilung dieser Fragen war es im Zuge des Projekts möglich Einblicke in den Vollzugsalltag zu bekommen. Über Informationsveranstaltungen, Gruppendiskussionen und (informelle) Gespräche konnte auch

siehe auch OGH 26.06.2014, 8 Ob 47/14s; OGH 27.1.2010, 7 Ob 248/09k, sowie OGH 14.5.1997, 7 Ob 89/97g.

57 Wie sich zeigt, sind bereits die Kriterien der Eignung und Erforderlichkeit nicht ausreichend erfüllt, weshalb sich die Frage der Adäquanz im Folgenden gar nicht mehr stellt.

58 Insasse, Justizanstalt A (Gruppe C).

59 Siehe https://artes.esa.int/sites/default/files/TRL_Handbook.pdf

60 Die Detektionsrate wurde durch die Metriken Mean Average Precision (83 %) und Mean Average Heading Similarity (79 %) operationalisiert.

61 Die Tracking-Performance wurde durch die Metriken Multi-Object Tracking Accuracy (58 %) und Multi-Object Tracking Precision (78 %) operationalisiert.

62 Siehe auch Art. 10 KI-VO zur »Data Governance« für Trainings-, Validierungs- und Testdatensätze. Vgl. hierzu auch die Entschließung des Europäischen Parlaments vom 6. Oktober 2021 zum Thema, Künstliche Intelligenz im Strafrecht und ihre Verwendung durch die Polizei und Justizbehörden in Strafsache (2020/2016(INI)); P9_TA(2021)0405; https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_DE.pdf

die Perspektive der betroffenen Insassen und Bediensteten eingeholt werden.[63] Unter Hinzuziehung empirischer Daten aus der einschlägige Literatur wurde auf dieser Basis analysiert, ob der Einsatz des KI-Systems zur Prävention von selbst- und fremdgefährdendem Verhalten aus praktischer Sicht erforderlich ist. Dabei ist zunächst klarzustellen, dass sich der Begriff »Gewalt« als Form der Fremdgefährdung im Vollzugsalltag nicht auf körperliche Übergriffe beschränkt, sondern auch subtilere Formen wie psychische Gewalt umfasst. Für den österreichischen Vollzug zeigt die repräsentative Dunkelfeldstudie von Hofinger & Fritsche (2021), dass 69,9 % der befragten Insass*innen mindestens einmal von psychischer Gewalt betroffen gewesen sind, während die Werte für körperliche Gewalt bei 41,3 %, und sexueller Belästigung/Gewalt bei 9,6 % liegen. Auch im Zuge der von uns geführten Gespräche wird vom Vollzugspersonal attestiert: »Es passiert schon viel Gewalt, [es gibt] schon ein Gewaltproblem in der JA«.[64] Zudem wird darauf hingewiesen, dass es mehr Kameras als Bildschirme im Wachzimmer gibt[65] und somit durchaus Bedarf besteht, die Überwachung automationsunterstützt zu optimieren. Die Justizwache erläutert jedoch auch, dass »tatsächliche Gewalt in erster Linie verbal« in Form von Beschimpfungen und Mobbing stattfinden würde.[66] Zugleich wird vom Vollzugspersonal wiederholt darauf hingewiesen, dass es zur Beweissicherung das Klarbild braucht.[67] Das »herkömmliche Videobild« wird als »nützlich« bezeichnet und »sollte nicht durch den Sensor ersetzt werden«.[68] Für die Frage nach gelinderen Mitteln und alternativen Eingriffen, ist zudem die Ursache für gewalttätiges Verhalten in Haft zu beachten; dabei gelten sowohl biografische Merkmale als auch die institutionellen Rahmenbedingungen in den Justizanstalten als zentral (Sax, 2013; Baumeister, 2017; Hamedl & Monina, 2021; Hofinger & Fritsche, 2021). So verweisen auch die von uns befragten Insassen auf die beengten räumlichen Verhältnisse (»Stellen Sie sich vor, auf 10m2 mit einer völlig fremden Person!«[69]) und die Knappheit bestimmter Güter (»Die unnötigsten Gründe, weswegen sich draußen niemand schlagen würde. Zum Beispiel wegen Zigarettenpapier oder Tabak oder sowas«[70]). Folglich wären zur Gewaltprävention im Strafvollzug gelindere Mittel verfügbar, die die institutionellen Strukturen ins Zentrum der Überlegungen rücken. Zudem wäre, statt strenger Überwachung und technischer Kontrolle, die Handlungsfähigkeit und Selbstständigkeit der Betroffene zu fördern (Sax, 2013; Baumeister, 2017; Hofinger & Fritsche, 2021). Für die innere Sicherheit in den Justizanstalten und die Resozialisierung der Gefangenen gilt auch eine positive und vertrauensvolle Beziehung zu den Bediensteten als relevant (UNODC, 2016; Baumeister, 2017).

Doch aggressives Verhalten in Justizanstalten beschränkt sich nicht nur auf Fremdgefährdung und Gewalt gegenüber Dritten; auch selbstverletzende Verhaltensweisen und Suizid sind im Gefängniskontext von großer Bedeutung.[71] So handelt es sich bei Selbsttötung um die häufigste Todesursache in Haft; die Suizidraten in Gefängnissen sind deutlich höher als in der Allgemeinbevölkerung (WHO, 2007; Fazel et al., 2017). Als Hochrisikoperioden gelten insb. die ersten Stunden nach Haftantritt, Untersuchungshaft sowie die ersten Monate nach Einweisung in eine Justizanstalt und Tage vor oder nach relevanten Ereignissen (z. B. der Hauptverhandlung, negativer Rechtsbescheid etc.) (WHO, 2007; Esser & Reißmann, 2019). Als häufigste Methode wird das Erhängen mit an Verstrebungen im Haftraum geknüpftem Bettzeug angeführt und das meist, wenn der Personalschlüssel gering ist (z. B. in der Nacht oder am Wochenende) und die betroffene Person allein, also in Isolation oder Einzelhaft, ist (Volksanwaltschaft, 2022). Im Zusammenhang mit der Vulnerabilität der Betroffenen im Hinblick auf Selbstgefährdung und -tötung sind von staatlicher Seite im Strafvollzug auch erhöhte Schutz- und Gewährleistungspflichten zu beachten.[72] Praktischer Bedarf nach KI-Systemen bestünde demnach insb. in der Überwachung »besonders gesicherter Zellen« gem. § 103 Abs. 2 Z. 4 StVG. Ein technisch ausgereiftes System könnte hier zu einer effizienteren Erkennung sicherheitskritischer Handlungen führen (Esser & Reißmann, 2019); dies vor allem weil das Aufsichtspersonal praktisch nicht in der Lage ist, alle Monitore des Kontrollraums permanent über Echtzeitbeobachtung im Auge zu behalten. Esser & Reißmann (2019) erachten eine KI-basierte Überwachung nicht nur als effizienter, sondern dann, wenn die Einsicht in die Bilddaten lediglich im Alarm-

63 Insgesamt wurden auf diese Weise knapp 80 Insassen und 15 Bedienstete erreicht.

64 Justizwachebeamter, Justizanstalt B (Gespräch F).

65 Justizwachebeamter, Justizanstalt A (Gespräch E).

66 Justizwachebeamter, Justizanstalt A (Gespräch A). Zudem wird mit Bezug auf den Jugendstrafvollzug von »Spaßraufereien« gesprochen, die von ernsten Situationen zu unterscheiden seien.

67 Justizwachebeamter, Justizanstalt B (Gespräch F).

68 Psychologin, Justizanstalt A (Gespräch D).

69 Insasse, Justizanstalt B (Gespräch G).

70 Insasse, Justizanstalt B (Gruppe I).

71 Im Jahr 2021 ereigneten sich 34 Vorfälle in landesgerichtlichen Gefangenenhäusern, 10 in den Strafhäusern. Jeweils ein Todesfall war in zwei psychiatrischen Kliniken zu verzeichnen (Volksanwaltschaft, 2022).

72 Siehe EGMR, 10.02.2011, 44973/04, »Premininy vs Russia« im Zusammenhang mit einem Fall von »systematic abuse for at least a week at the hands of fellow inmates«.

fall erfolgt, im Vergleich zur konstante Beobachtung durch einen Menschen auch als weniger eingriffsintensiv. Dem ist aus datenschutzrechtlicher Sicht zu entgegnen, dass die KI-basierte Verhaltensanalyse die Effektivität zwar erhöhen kann, über eine damit verbundene Entscheidungsautomation jedoch auch als invasiver einzustufen wäre. Zudem ist auch hier auf gelindere Mittel abseits von Isolation und Überwachung hinzuweisen. Alternative Formen und Strategien der Suizidprävention, werden insb. in der psychosozialen Betreuung der Betroffenen gesehen (Baumeister, 2017). Somit kann grundsätzlich in beiden Use-Cases praktischer Bedarf attestiert werden; es besteht Erforderlichkeit, es gibt jedoch auch gelindere Mittel und nachhaltige Alternativen zur KI-basierten Überwachung.

6 Ethische Abwägung: Effizienz und Menschenwürde

In Ergänzung zur Verhältnismäßigkeitsprüfung wird für die Beurteilung des vorliegenden Sachverhalts zudem eine Abwägung zentraler ethischer Positionen vorgenommen. Diesbezüglich ist darauf hinzuweisen, dass der Begriff der Ethik im Zusammenhang mit KI jüngst eine Art Renaissance erfahren hat. In einschlägigen Leitfäden und Empfehlungen wird regelmäßig auf eine Reihe an Prinzipien und Grundsätzen, wie z. B. Transparency, Accuracy oder Accountability, verwiesen.[73] In der folgenden Abwägung geht es jedoch weniger um eine Diskussion derartige Begriffe als vielmehr um eine Gegenüberstellung und ethische Analyse utilitaristischer sowie deontologischer Positionen und deren Verbindung mit zentralen grundrechtlichen Gewährleistungen.

6.1 Utilitaristische Sicht

Der Einsatz des KI-Systems lässt sich in ethischer Hinsicht zunächst über den Ansatz des Utilitarismus diskutieren. Das moralische Prinzip des Utilitarismus liegt grundsätzlich darin, dass Personen so handeln sollten, dass sie den größtmöglichen Nettonutzen, d. h. den größten Nutzen oder die geringsten Kosten, für eine möglichst große Gemeinschaft fördern (Spinello, 2021). Die KI-basierte Überwachung kann in diesem Sinne als utilitaristische Maßnahme zur Erfüllung der Schutzpflichtendes Staates gegenüber den Gefangenen verstanden werden. Es handelt sich um einen rationalen Problemlösungsansatz auf Basis von Kosten-Nutzen Abwägungen; als Ziele für die Entwicklung des KI-Systems werden die bestmögliche Unterstützung und Entlastung des Personals, die frühzeitige Erkennung kritischer Verhaltensmuster sowie die Erhöhung der Sicherheit und Ordnung im Vollzugsalltag genannt. Für den Kontext des Strafvollzugs ist dabei auch das utilitaristische Konzept des Panoptismus von Bentham zu nennen (Foucault, 1975; Kammerer, 2008; Lyon, 2006). Das in Rede stehende KI-System kann in diesem Sinne als eine technische Weiterentwicklung der panoptischen Idee effizienter Überwachung gesehen werden. Die genannten Ziele des sozio-technischen Managements zur Verbesserung der Sicherheit und Ordnung im Strafvollzug sind demnach utilitaristisch begründet. Dabei lässt sich allgemein seit Jahrzehnten ein kriminalpolitischer Wandel beobachten, wobei technologiegestützte Strategien sozialer Kontrolle im Informationszeitalter in den Vordergrund treten; das wohlfahrtsstaatliche Inklusionsversprechen (auf Resozialisierung und Reintegration) erweist sich dabei als tendenzielles Auslaufmodell; der Besserungsdiskurs wird zunehmend von einem Überwachungsdiskurs und der Anwendung automationsunterstützter Systeme des Risikomanagements überlagert (Rothmann, 2010).

6.2 Deontologische Sicht

Den Ideen effizienter Überwachung stehen die Prinzipien der deontologischen Ethik gegenüber. Der Deontologie geht es grundsätzlich um universelle Pflichten und Maximen; der moralische Status einer Handlung wird nicht anhand ihrer Konsequenzen bestimmt, sondern auf Basis moralisch-ethischer Pflichten als intrinsisch (bzw. kategorisch) gut oder schlecht bezeichnet (Spinello, 2021). Relevant ist für den vorliegenden Zusammenhang insb. die Sittenlehre Kants; demnach soll nur nach derjenigen Maxime gehandelt werden, die zugleich ein allgemeines Gesetz werden kann. Zudem soll kein Mensch, für welche Zwecke auch immer, jemals als bloßes Mittel betrachtet und behandelt werden (Lutz-Bachmann, 2022; Spinello, 2021). Das damit umrissene Konzept schreibt jedem Menschen einen inhärenten Wert zu (Objektivierungsverbot); der Schutzbereich umfasst jede natürliche Person und stellt »nicht auf einen bestimmten Lebensbereich ab, sondern kommt jedem Menschen situationsunabhängig allein aufgrund seines Menschseins zu« (Fuchs & Segalla, 2019, Art. 1 Rz 30). Bielefeldt (2012) versteht

73 Siehe z. B. European Commission for the Efficiency of Justice (CEPEJ) (2018). European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment; Retrieved from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c; High-Level Expert Group on AI, European Commission (2019). »Ethics Guidelines for Trustworthy AI«; https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

die Menschenwürde auch als »unhintergehbare Prämisse normativer Verbindlichkeiten«.[74] Letztlich wird von einem dispositionsfesten Kern der Menschenwürde ausgegangen; dieser gilt mitunter auch als Kern der Privatsphäre (Moore, 1984; Rothmann, 2023). Die Relevanz der Achtung der Menschenwürde lässt nicht zuletzt auch auf Art. 1 der GRC sowie Art. 1 Abs. 4 des Bundesverfassungsgesetz über den Schutz der persönlichen Freiheit (PersFrBVG) stützen,[75] wobei letzterer vorsieht, dass jeder der »festgenommen oder angehalten wird, […] unter Achtung der Menschenwürde und mit möglichster Schonung der Person zu behandeln [ist]«.[76] Auch in § 22 Abs. 1 StVG ist normiert, dass die Strafgefangenen »[…] mit Ruhe, Ernst und Festigkeit, gerecht sowie unter Achtung ihres Ehrgefühls und der Menschenwürde zu behandeln« sind.

Der deontologische Ansatz zur Achtung der Menschenwürde führt in letzter Konsequenz aber auch zu positiven Schutz- bzw. Gewährleistungspflichten des Staates (Lukan, 2019, Art. 4 Rz. 23 ff). Nach Kant sollten Menschen bestimmte Handlungen nicht nur gegenüber anderen Menschen, sondern auch gegenüber sich selbst unterlassen. Moralisch ausgeschlossen ist daher nicht nur jede Form der freien Selbstunterwerfung (Versklavung, Entwürdigung), sondern auch eine vorsätzliche Selbstverstümmelung oder Suizid (Höffe, 1992; Lutz-Bachmann, 2022). In diesem Sinne kann ein staatliches Hinwirken auf die Herstellung bestimmter Zustände bzw. Lebensbedingungen zum Schutz der Menschenwürde auch geboten sein (Pavlidis, 2019, Art. 7 Rz. 11 ff). Setzt man die Effektivität des KI-Systems als gegeben voraus, könnte der Ansatz infolge dazu genutzt werden, diesen staatlichen Schutzpflichten nachzukommen. Auf der anderen Seite kann (mit Verweis auf Seneca, Nitsche u.A.) von einer »etablierten philosophischen Tradition« gesprochen werden, die argumentiert, dass gerade die »Möglichkeit zum Suizid […] einen Teil der Würde des Menschen ausmacht« (Stocker, 2019).[77] Somit stellt sich letztlich die Frage, ob Suizid als Ausfluss der Menschenwürde oder vielmehr als Verletzung derselben zu verstehen ist. Dies kann hier freilich nicht abschließend geklärt werden; das StVG geht in § 67 allerdings davon aus, dass die Vornahme ärztlicher Experimente auch dann unzulässig ist, wenn die Betroffenen dazu ihre Einwilligung erteilen. Dieses absolute Verbot bezieht sich zwar nicht auf Suizidszenarien, die Norm bringt jedoch eine deontologische (wie paternalistische) Werthaltung des Gesetzgebers zum Ausdruck, indem die Selbstbestimmung der Betroffenen zur Wahrung ihrer Würde ausgeschlossen wird.[78]

7 Fazit

Abschließend kann wie folgt zusammengefasst werden: Zur Diskussion steht die rechtliche Zulässigkeit eines KI-basierten Systems zur multimodalen Erkennung sicherheitskritischer Szenarien im Strafvollzug (wie z. B. Gewalt oder Suizidversuche). Das System soll das Personal künftig entlasten und die Sicherheit und Ordnung in Strafvollzugsanstalten erhöhen. Als technisch innovativ gilt dabei insb. der über die CTCATs verfolgte Ansatz zur »Human Segmentation and Action Recognition« (Heitzinger & Kampel, 2023; Stippel et al., 2023).

Für die Beurteilung des Systems sind zunächst die Regelungen des StVG beachtlich. Neben § 15a Abs. 1 StVG zur allgemeinen Vollzugsverwaltung bietet hier vor allem § 102b StVG zur Videoüberwachung einen normativen Anknüpfungspunkt. Die bildgebenden Sensoren können für sich zwar als eine Form der Videoüberwachung klassifiziert werden, in beiden Fällen fehlt jedoch eine hinreichende Rechtfertigung für den Einsatz des KI-Systems. Eine ausdrücklich im Gesetz vorgesehene Erlaubnis bräuchte es insb. dann, wenn über das KI-System eine automatisierte Entscheidung im Sinne des § 41 DSG stattfindet, da eine solche im Strafvollzug zu erheblichen Beeinträchtigungen der Betroffenen führen kann. Führt die Verarbeitung sensibler Daten (wie Pulsfrequenz oder Wärmebild) außerdem zu Formen der automatisierten Diskriminierung wäre der Einsatz des Systems gem. § 41 Abs. 3 DSG verboten. Somit braucht es für den datenschutzrechtlichen Eingriff mittels KI-basierter Technologie jedenfalls eine über § 15a StVG und § 102b StVG hinausgehende, auf den bestimmten Zweck des KI-Systems ausgerichtete, Rechtsgrundlage, die in ihrer Ausgestaltung nach Berka (2012) u. a. klären sollte,

- in welchen Bereichen der Vollzugsanstalten die Technologie eingesetzt werden darf,
- welche Daten konkret erfasst werden,
- wer wann Zugriff auf diese Daten hat,

74 Anders als in Deutschland (über Art. 1 GG) ist die Menschenwürde in Österreich nicht explizit in der Verfassung festgeschrieben; siehe VfGH 10.12.1993, VfSlg 13.635/1993, sowie VfGH 14.03.2012, VfSlg 19.632/2012.

75 Bundesverfassungsgesetz vom 29. November 1988 über den Schutz der persönlichen Freiheit; StF: BGBl. Nr. 684/1988.

76 Vgl. insb. VfGH 07.03.1994, VfSlg 13.708/1994.

77 Siehe hierzu die Debatte zum Thema Sterbehilfe oder Tötung auf Verlangen.

78 Siehe hierzu weiterführend auch das Erkenntnis des VfGH G 139/2019-71 mit dem das Verbot der »Hilfeleistung zum Selbstmord« (gem. § 78 StGB) aufgehoben wurde.

– ob bzw. inwiefern die erhobenen Daten weitergegeben oder mit anderen Informationen oder Aufgabenbereichen verknüpft werden und
– ob bzw. wie lange die erhobenen Daten gespeichert werden dürfen.

Darüber hinaus gilt es wirksame Maßnahmen gegen eine missbräuchliche Verwendung vorzusehen und die Umsetzung der Betroffenenrechte (auf Auskunft, Richtigstellung und Löschung, etc.) zu gewährleisten.

Im Hinblick auf die Beurteilung anhand der europäischen KI-Verordnung ist darauf hinzuweisen, dass sich diese derzeit noch im Gesetzgebungsverfahren befindet. Auf Basis der uns vorliegenden Vorschläge für die VO wäre das in Rede stehende KI-System zumindest als Hoch-Risiko-Anwendung einzustufen. So könnte das KI-System gem. Anhang III 1 (b) und (c) der KI-VO künftig, insb. aufgrund der Verarbeitung der Pulsfrequenz, als System zur »biometrische[n] Kategorisierung nach sensitiven [...] Attributen« oder als System »zur Erkennung von Emotionen« gelten. In welcher Version die KI-VO letztlich auch zu verbindlichem Sekundärrecht wird bleibt abzuwarten. Die Beurteilung der Zulässigkeit des Systems hängt aber auch wesentlich von der konkreten Ausgestaltung und künftigen Implementierung ab. Um die Verhältnismäßigkeit des damit verbundenen Grundrechtseingriffs zu wahren, ist insb. auch die Eignung und Erforderlichkeit dieses technischen Mittels beachtlich. Dabei ist es zunächst vor allem die aktuelle Stufe der Entwicklung, die einem Echtbetrieb im Vollzug noch entgegensteht; zur Beurteilung der Eignung der Technologie braucht es künftig einen evidenzbasierten Nachweis der Wirksamkeit und Zweckdienlichkeit. Im Fall der Erforderlichkeit ist wiederum festzuhalten, dass im Strafvollzug aus empirischer Sicht grundsätzlich der Bedarf zur Gewaltprävention erkennbar ist; darüber hinaus zeigt sich Bedarf zur frühzeitigen Erkennung autoaggressiver Verhaltensweisen; dies vor allem im Fall von Isolations- und Einzelhaft. In derartigen Situationen sind von staatlicher Seite auch erhöhte Schutzpflichten zu beachten. Die Justizwache sieht darin eine mögliche Unterstützung in ihren Aufgaben im Vollzugsalltag. Der Grundrechtseingriff über das System wäre dennoch nur dann gerechtfertigt, wenn kein gelinderes Mittel zur Zweckerreichung verfügbar ist; dabei kann im Zusammenhang mit Formen der Selbstgefährdung vor allem auch auf psychosoziale Maßnahmen verwiesen werden. Für die Wahrung der Rechte und Freiheiten der betroffenen Personen sollte der Einsatz des KI-Systems im Strafvollzug aber nicht zuletzt auch ausdrücklich im Gesetz erlaubt sein.

Referenzen

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective, *PeerJ Computer Science* 2:e93. Retrieved from https://doi.org/10.7717/peerj-cs.93.

Allard, T., Wortley, R., & Stewart, A. (2008). The Effect of CCTV on Prisoner Misbehavior. The Prison Journal 88(3), 404–422.

Alon-Barkat, S., & Busuioc, M. (2023). Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice, *Journal of Public Administration Research and Theory, 33*, 1,153–169, https://doi.org/10.1093/jopart/muac007.

Artikel-29-Datenschutzgruppe (2018). *Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679*, WP251rev01.

Baker, J., Hobart, L., & Mittelsteadt, M. (2021). AI for Judges. A Framework; Center for Security and Emerging Technology (CSET), Policy Brief; Georgetown University.

Baumeister, B. (2017). *Gewalt im Jugendstrafvollzug.* Nomos, Baden-Baden.

Berka, W. (1999). *Die Grundrechte. Grundfreiheiten und Menschenrechte in Österreich*. Springer, Wien, New York.

Berka, W. (2005). *Lehrbuch Verfassungsrecht*, Grundzüge des österreichischen Verfassungsrechts für das juristische Studium. Springer, Wien, New York.

Berka, W. (2012). *Das Grundrecht auf Datenschutz im Spannungsfeld zwischen Freiheit und Sicherheit*, 18. ÖJT, I/I; Manz, Wien.

Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism, *Journal of Experimental Criminology* 13, 193–216.

Bezemek, C. (2016). *Grundrechte in der Rechtsprechung der Höchstgerichte*. Facultas, Wien.

Bielefeldt, H. (2012) Universalität und Gleichheit, in: Pollmann, A. & Lohmann, G. (Hrsg), *Menschenrechte. Ein interdisziplinäres Handbuch*, Stuttgart/Weimar: J. B. Metzler Verlag.

Bouachir, W., Gouiaa, R., Li, B., & Noumeir, R. (2018). Intelligent video surveillance for real-time detection of suicide attempts, *Pattern Recognition Letters, 110*, 1–7.

Burns, G. (2022). The use and future of artificial intelligence monitoring in prisons, *Reason Foundation.* Retrieved from https://reason.org/commentary/the-use-and-future-of-artificial-intelligence-monitoring-in-prisons/.

Briem, C. (2022). Profiling, Scoring und Big Data im österreichischen Abgabenverfahren? *ELSA Austria Law Review, VII*, 38–45.

Brobst, J. A. (2018). The Metal Eye: Ethical Regulation of the State's Use of Surveillance Technology and Artificial Intelligence to Observe Humans in Confinement, *California Western Law Review, 55 (1)2.* Retrieved from https://ssrn.com/abstract=3297341.

Buchner, B. (2018). Art. 22, in: Kühling, J.; Buchner, B. (Hrsg) Datenschutz-Grundverordnung/BDSG, Kommentar, 2. Auflage; C. H. Beck, München; 508–520.

Camara, C.; Peris-Lopez, P.; Safkhani, M.; Bagheri, N. (2023). ECG Identification Based on the Gramian Angular Field and Tested with Individuals in Resting and Activity States. *Sensors*, *23*, 937; https://doi.org/10.3390/s23020937.

Chen, S. (2019). No escape? Chinese VIP jail puts AI monitors in every cell »to make prison breaks impossible«, *South China Morning*

Post. Retrieved from https://www.scmp.com/news/china/science/article/3003903/no-escape-chinese-vip-jail-puts-ai-monitors-every-cell-make.

Council of Europe (2023). *Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, Committee on Artificial Intelligence (CAI); Retrieved from https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f.

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism, *Science Advances, 4(1), eaao5580.* Retrieved from DOI: 10.1126/sciadv.aao5580

Drexler, K., & Weger, T. (2018). Strafvollzugsgesetz (StVG), Kommentar; 4 Auflage, Manz, Wien.

Esser, R., & Reißmann, L. (2019). Einsatz künstlicher Intelligenz zur Suizidprävention im Justizvollzug, *Juristen Zeitung 74*, 975–982.

EUROPRIS (2020). Artificial Intelligence in Correctional Services, Knowledge Management System; Retrieved from https://www.europris.org/epis/kms/?detail=391.

European Commission for the Efficiency of Justice (CEPEJ) (2018). *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment;* Retrieved from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

Fazel, S., Ramesh, T., & Hawton, K. (2017). Suicide in prisons: an international study of prevalence and contributory factors, *The Lancet Psychiatry* 4(12), 946–952.

Foucault, M. (1975). Überwachen und Strafen. Die Geburt des Gefängnisses; Suhrkamp 1994, Frankfurt a. M.

Fuchs, C., & Segalla, P. (2019). Art. 1, In Holoubek, R. & Lienbacher, G. (Hrsg) *GRC-Kommentar*, Manz, Wien.

Guercio, M. J. (2021). The One-Two Punch of RFID and AI, *Corrections Forum*, 30(2), 22–25.

Hanisch, S., Todt, J., Volkamer, M., & Strufe, T., (2023). Zu Risiken und Anonymisierungen von Verhaltensbiometrie, in: M. Friedewald et al. (Hrsg) *Daten-Fairness in einer globalisierten Welt* (423–444). Nomos, Baden-Baden.

Hao, K. (2019). AI is sending people to jail – and getting it wrong, *MIT Technology Review;* Retrieved from https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/.

Hamedl, P., & Monina, G. (2021). *Monitoring Prison Violence – A Handbook for National Preventive Mechanisms*. Ludwig Boltzmann Institute of Fundamental and Human Rights, Wien.

Hayato, M., Kazuyuki, K., Noritaka, S., Masataka, I., Hisashi, M., Kiyohiro, G., ... Fumitoshi, M. (2009). Treaded control system f or rescue robots in indoor environment. In *Proceedings of 2008 IEEE International Conference on Robotics and Biomimetics*, 1836–1843.

High-Level Expert Group on AI, European Commission (2019). *»Ethics Guidelines for Trustworthy AI«;* Retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Hofinger, V., & Fritsche, A. (2021): Gewalt in Haft: Ergebnisse einer Dunkelfeldstudie in Österreichs Justizanstalten. Wien, LIT Verlag.

Houser, K. (2019). Hong Kong Has a Plan to Make All of Its Prisons »Smart«, Futurism. Retrieved from https://futurism.com/smart-prisons-hong-kong.

Heitzinger, T., & Kampel, M. (2021). *A Foundation for 3D Human Behavior Detection in Privacy-Sensitive Domains*, in *32nd British Machine Vision Conference (BMVC).*

Heitzinger, T. & Kampel, M. (2023). A Fast Unified System for 3D Object Detection and Tracking, In *Proceedings of International Conference on Computer Vision (ICCV), Paris, France.*

Höffe, Otfried (1992). Ethik und Politik – Grundmodelle und -probleme der praktischen Philosophie. 4. Auflage, Suhrkamp, Frankfurt a. M.

Kammerer, D. (2008). Bilder der Überwachung. Suhrkamp, Frankfurt a. M.

Kaun, A., & Stiernstedt, F. (2020). Doing time, the smart way? Temporalities of the smart prison, *New Media & Society*, 22(9), 1580–1599.

Khakzadeh-Leiler, L. (2011). *Die Grundrechte in der Judikatur des Obersten Gerichtshofs*. Springer, Wien, New York.

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions with a suggestion for a consensual definition, *Motivation and Emotion, 5*, 263–291.

Knight, V., & Van De Steene, S. (2017). The capacity and capability of digital innovation in prisons: Towards smart prisons, *Advancing Corrections Journal*, 4, 88–101.

Knight, V., Reisdorf, B., & Van De Steene, S. (2023). *Digital Maturity of Prisons: A Global Survey*; De Montfort University.

Kohn, B. (2021). *Künstliche Intelligenz und Strafzumessung*. Nomos, Baden-Baden.

Krotosky, S. J., & Trivedi, M. M. (2007). On color-, infrared-, and multimodal-stereo approaches to pedestrian detection, *IEEE-Transactions on Intelligent Transportation Systems 8(4)*, 619–629.

Lachmayer, K., & Rothmann, R. (2020). Grundrechtswissen in Österreich – Eine empirische Untersuchung zum 100-jährigen Bestehen des Bundes-Verfassungsgesetzes 1920, *Juridikum 4*, 472–483.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2117–2125.

Logg, J. M., Minson, J. A., Moore D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment, *Organizational Behavior and Human Decision Processes*, 151, 90–103.

Lukan, M. (2019). Art. 4, In Holoubek, R. & Lienbacher, G. (Hrsg.) GRC-Kommentar. Manz, Wien.

Lutz-Bachmann, M. (2022). Würde, I. Moralphilosophisch. Retrieved from https://www.staatslexikon-online.de/Lexikon/W%C3%BCrde.

Lyon, D. (2006). Theorizing Surveillance – The panopticon and beyond. Routledge, Taylor & Francis Group, London, New York.

Meinhardt, T., Kirillov, A., Leal-Taixe, L., & Feichtenhofer, C. (2022). Trackformer: Multi-object tracking with transformers, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 8844–8854.

MIPT – Multimodal Identity Preserved Tracking Dataset (2021). Version 1.0.1; Retrieved from https://zenodo.org/records/5592323.

Mittelstadt, B. (2021). Interpretability and Transparency in Artificial Intelligence, In Carissa Véliz (Eds.), *The Oxford Handbook of Digital Ethics*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4317938.

Moore, B. (1984). Privacy – Studies in Social and Cultural History; M. E. Sharpe, Inc., Armonk, New York.

Mulholland, C. L., Sterritt, R., O'Hagan, P., & Hanna, E. (2008). Tagging and Tracking System for Prisons and Correctional Facilities – A Design Roadmap, In *Proceedings of Fifth IEEE Workshop on Engineering of Autonomic and Autonomous Systems*, 143–153.

Öhlinger, T., & Eberhard, H. (2014). Verfassungsrecht, 10. überarbeitete Auflage; Facultas, Wien.

Pavlidis, L. (2019). Art. 7, In Holoubek, R. & Lienbacher, G. (Hrsg.) GRC-Kommentar, Manz, Wien.

Plutchik, R. (1980). Emotion. A psychoevulutionary synthesis. Harper & Row, New York.

Puolakka, P. & Van De Steene, S. (2021). Artificial Intelligence in Prison in 2030: An exploration on the future of AI in prisons, *Advancing Corrections Journal*, 11 128–138.

Purwito, A., Sitanggang, D. W. F., Suprayitno, S., & Marbandi, A. (2018). Design of Prison Security Information System using RFID, International Journal of ASRO, 9(1), 129–135.

Redden, J., Inkpen, C., & DeMichele, M. (2020). *Artificial Intelligence Applications in Corrections.* Criminal Justice Testing and Evaluation Consortium; U. S. Department of Justice, National Institute of Justice, Office of Justice Programs.

Reiling, A. D. (Dory) (2020). Courts and Artificial Intelligence, *International Journal for Court Administration 11(2).* Retrieved from https://doi.org/10.36745/ijca.343.

Repetto, T. A. (1976). Crime Prevention and the Displacement Phenomenon, *Crime and Delinquency 22*, 166–177.

Rizer, A.; Watney, C. (2019). Artificial Intelligence can make our jail system more efficient, equitable, and just, *Texas Review of Law & Politics*, *23*(1), 181–227.

Rogers, D. (2021). How Real is Artificial Intelligence in Prisons?, *Corrections Forum 30(5),* 4,6,8.

Rothmann, R. (2010). Sicherheitsgefühl durch Videoüberwachung? Argumentative Paradoxien und empirische Widersprüche in der Verbreitung einer sicherheitspolitischen Maßnahme, *Neue Kriminalpolitik, 22(3),* 103–107.

Rothmann, R. (2012). Zur Evaluation der sicherheitstechnischen Eignung von Videoüberwachung? Regionale Defizite, internationale Standards, methodische Herausforderungen, *Juridikum, (4),* 483–495.

Rothmann, R. (2023). Die Rechtswirklichkeit der datenschutzrechtlichen Einwilligung, Mohr Siebeck, Tübingen.

Rothmann, R., & Vogtenhuber, S. (2013). Verdächtiges Verhalten und automationsunterstützte soziale Kontrolle: Über »intelligente« Videoüberwachung zur Detektion von Kfz-Delikten, *Zeitschrift für soziale Probleme und soziale Kontrolle, 24(2),* 271–298.

Saharia, C., Chan, W., Chang, H., Lee, C. A., Ho, J., … Norouzi, M. N. (2022). Palette: Image-to-image diffusion models. In *Proceedings of ACM SIGGRAPH 2022 Conference*, 1–10.

Sandler, M., Howard, A, Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks, In *Proceedings of the IEEE conference on computer vision and pattern recognition,* 4510–4520.

Sax, H. (2013). *Ending Violence against Children in Custody, Jugendliche im Strafvollzug – Gewalterfahrungen und Möglichkeiten der Veränderungen aus Perspektive der Betroffenen*. Länderbericht Österreich, Ludwig-Boltzmann-Institut für Menschenrechte, Wien.

Spinello, R. A. (2021). Cyberethics: Morality and Law in Cyberspace; 7th Edition, Jones and Bartlett Publisher; Bostin, Toronto, London, Singapore.

Stippel, C., Heitzinger, T., & Kampel, M. (2023). A Trimodal Dataset: RGB, Thermal, and Depth for Human Segmentation and Action Recognition, In *Proceedings of the German Conference on Pattern Recognition (GCPR)*, Heidelberg, Germany.

Stocker, R. (2019). *Theorie und Praxis der Menschenwürde*, Mentis Verlag, Paderborn.

Strohmayer, J., & Kampel, M. (2022). A Compact Tri-modal Camera Unit for RGBDT Vision. In *Proceedings of the 2022 5th International Conference on Machine Vision and Applications (ICMVA), Singapore,* 34–42.

Thiele, C., & Wagner, J. (2020). Praxiskommentar zum Datenschutzgesetz, Jan Sramek, Wien.

Thomas, C., & Nunez, A. (2022). Automating Judicial Discretion: How Algorithmic Risk Assessments in Pretrial Adjudications Violate Equal Protection Rights on the Basis of Race, *Law & Inequality, 40(2), 371–407.*

TRISTAR – A Trimodal Dataset: RGB, Thermal, and Depth for Human Segmentation and Action Recognition (2023). Version v1. Retrieved from https://zenodo.org/record/7996570.

United Nations Office on Drugs and Crime (2016). Handbook on the Management of High-Risk Prisoners. Retrieved from https://www.unodc.org/documents/justice-and-prison-reform/HB_on_High_Risk_Prisoners_Ebook_appr.pdf.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J.; Jones, L.; Gomez, A. N.; … Polosukhin, l. (2017). Attention is all you need. Advances in neural information processing systems. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*; Retrieved from https://arxiv.org/abs/1706.03762.

Volksanwaltschaft (2022). *Kontrolle der öffentlichen Verwaltung*, Bericht der Volksanwaltschaft an den Nationalrat und an den Bundesrat 2021, Retrieved from https://volksanwaltschaft.gv.at/downloads/20ag/pb-45-nachpruefend_2021_bf-1.pdf.

Welsh, B. C., Farrington, D. P. (2002). *Crime prevention effects of closed-circuit television: a systematic review*. Home Office Research, Development and Statistic Directorate. London.

World Health Organisation (2007). Suizidprävention: Ein Leitfaden für Mitarbeiter des Justizvollzugsdienstes; WHO, Genf.

Xiao, H. C., & Xiong, J. (2013). A Prison RFID Network System Using Position Computing, In *IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013)*, 1–5.

Yan, S. (2019). Chinese high-security jail puts AI monitors in every cell "to make prison breaks impossible", *The Telegraph*, Retrieved from https://www.telegraph.co.uk/news/2019/04/01/chinese-prison-rolls-facial-recognition-sensors-track-inmates/.

Završnik, A. (2021). Algorithmic justice: Algorithms and big data in criminal justice settings, *European Journal of Criminology*, 18(5), 623–642.

Zhilu, C., Xinming, H. (2019). Pedestrian detection for autonomous vehicle using multi-spectral cameras, In *Proceedings of IEEE-Transactions on Intelligent Vehicles 4(2),* 211–219.

Zimbardo, P. (1995). *Psychologie*, Springer, Berlin, Heidelberg, New York.