UNIO - *EU Law Journal*. Vol. 10, No. 1, May 2024, pp. 43-54.

School of Law – University of Minho

The role of Artificial Intelligence (AI) in rehabilitation and in the reduction of the use of imprisonment

Anabela Miranda Rodrigues[*]
Sónia Fidalgo[**]

ABSTRACT: Over the past few decades, the criminal justice system has experienced a shift towards risk management (actuarial justice). The New Public Management ideology has also entered the field of criminal policy and currently there is a growing recognition of the importance of adopting an Offender Management System. This new way of managing the criminal justice is increasingly supported by Artificial Intelligence and related digital technologies. The use of these technologies has enormous potential from a rehabilitation perspective. However, the current trend is to use them mainly for security and control purposes. This analysis examines whether the goal of rehabilitation can still be accommodated in a criminal justice system with the aforementioned characteristics. Criminal policy has the responsibility to ensure that the control it imposes on citizens can be balanced by an ethical use of digital technologies, which effectively contributes to rehabilitation and to the reduction in the resort to imprisonment.

KEYWORDS: Artificial Intelligence – digital technologies – imprisonment – non-custodial sentences – rehabilitation

[*] Full Professor at the Faculty of Law of the University of Coimbra, President of the Portuguese Group of the International Association of Penal Law and Vice-President of the International Society for Social Defence.
[**] Assistant Professor at the Faculty of Law of the University of Coimbra, Research Associate at the Legal Institute of the Faculty of Law of Coimbra and Member of the Portuguese Group of the International Association of Penal Law.

1. Introduction

Over the past few decades, the criminal justice system has experienced a shift towards risk management (actuarial justice). The costs associated with a system that seeks to socialise or rehabilitate prisoners, particularly in times of an increasing prison population, have contributed to the economic crisis of the welfare culture within the system. It has been recognised that public resources are not inexhaustible and that better management of these public resources can also lead to improvements in the criminal justice system. The New Public Management model has entered the field of criminal policy and there is a growing recognition of the importance of adopting an Offender Management System. Currently, this new way of managing the criminal justice system is increasingly supported by Artificial Intelligence (AI) and related digital technologies.

In this context, the central question to be asked is whether there is still a place for rehabilitation as a dimension that shapes the system in general and the execution of sentences in particular. In this article, we emphasise that the meaning of the prisoner's right to rehabilitation needs to be reconsidered and redefined. We highlight that it is politics (criminal policy) that has the responsibility of ensuring that the (ethical) use of AI and related digital technologies can be operated in a way that is more favourable to rehabilitation practices and effectively contributes to reducing the use of imprisonment.

Section 2 focuses on the crisis of the Welfare State and the emergence of actuarial justice, which is based on risk management and is founded on the concepts of "security" and "incapacitation". This section highlights the fact that although policymakers have shown renewed interest in non-custodial sentences, the "new penology" has resulted in a widespread expansion of the criminal justice system that has favoured incarceration.

In section 3 we focus on the impact of the New Public Management model on the criminal justice system. We emphasise that this new administrative management model relegates the achievement of external objectives to second place in favour of productivity and performance standards. We highlight the fact that criminal policy is based on people – the society and the offender – and our view that efficiency must not take precedence over effectiveness.

Section 4 focuses on the effects of the use of AI and related digital technologies by prison and probation services. We refer to two areas where AI and related digital technologies can be used in criminal justice organisation: the use for safety, security and good order; and the use for offender management purposes (risk assessment, rehabilitation and reintegration). AI and related digital technologies are said to contribute to a rehabilitative environment within prisons, even when used primarily for security purposes. What we can see, however, is that the use of these digital security techniques managed by AI carries the risk of turning the prisons into institutions that seek to control rather than rehabilitate prisoners. Even in non-custodial sanctions, the use of AI and related digital technologies serves a primarily security-focused purpose, prioritising control over rehabilitation. On the other hand, we also emphasise that the use of AI in risk assessment tools, even with an essentially rehabilitative intention and the aim of introducing greater objectivity and efficiency into the decision-making process, thus eliminating the discretionary power of legal professionals, does not actually succeed in erasing bias and only has the effect of shifting discretion to different parts of the organisation.

In the last section, we present our conclusions concerning the important role of criminal justice policy in the context of the use of AI and related digital technologies in a way that is more favourable to rehabilitation practices and actually contributes to reducing the use of imprisonment.

2. The crisis of the welfare culture and the penal rationality of security and control

The "fiscal crisis" of the welfare state in the United States[1] led to the questioning of the "criminal model" that had developed within the framework of this state model and which lasted for about a century, until the mid-1970s. This system, based on the ideology of rehabilitation, could only work if the State was able to keep and maintain not only numerous "justice agencies" to support it – police, courts and prisons – but also a vast number of "justice agents". The costs that this system entailed, at a time when the prison population was already increasing, led to the economic crisis of the welfare culture in the criminal justice system, with a clear shortfall in budgets to feed the machine.[2]

With the expansion of neoliberal ideology – in which the market is presented as the ideal means of framing human action, not only in terms of the economy, but also labour, health and justice – there has been a significant reconversion of the welfare state model.[3] In the contractual relationship between state and individual, the individual is responsible for themselves, for their direction in life, for their success or failure. This means a shift to the work state in the social sphere. This social transformation of the welfare state underscores the idea of *less state*. This *less state* paradigm is concerned with the economic domain and seeks to promote deregulation and to reduce the supervising role of the state. Notwithstanding this, the *less state* welfare concept is counterbalanced by the *more state* paradigm, in which the state is needed to control the *losers* (*i.e.* those who do not win in the game of the markets).

In this context of a welfare state in crisis, it is important to problematise the issue of the execution of sentences. The central question to be asked is whether there is still room for rehabilitation as a dimension that shapes the criminal justice system in general and the execution of sentences in particular.[4]

The first aspect that needs to be considered is the role played either by prison sentences or by non-custodial sentences in the criminal system. In this regard, it should be remembered that the 1960s were marked in Europe by a "deinstitutionalisation" movement, characterised by the slogan that it was necessary to "empty the prisons". At the same time, there was the emergence of non-custodial sentences. But the truth is that this trend has been reversed and the widespread

[1] Iñaki Rivera Beiras, "Prologo", in *La cárcel. La experiencia histórica bajo las perspetivas criminológicas*, ed. Gilberto Giacóia and Denise Hammerschmidt (Curitiba: Juruá Editorial, 2012), 10.

[2] Anabela Rodrigues, "A política criminal no Estado de Direito do Século XXI – os desafios da segurança", *Brasília - Revista Brasileira de Ciências Policiais*, 11(1) (2020): 28.

[3] Paul De Hert, Serge Gutwirth, Sonja Snacken and Els Dumortier, "La montée de l'État penal: que peuvent les droits de l'homme?", in *Les Droits de l'Homme, bouclier ou épée de droit penal?*, ed. Yves Cartuyvels, Hugues Dumont, François Ost, *et al.* (Brussels: Facultés Universitaires Saint-Louis, 2007), 235 and following. Pierre Rosanvallon, *La nueva cuestión social. Repensar el Estado Providencia* (Buenos Aires: Manantial, 1995), 75 and following.

[4] Anabela Rodrigues, "Execução penal socializadora e o novo capitalismo – uma relação impossível?", *Revista Brasileira de Ciências Criminais*, 23 (2015): 22.

use of prison sentences has to be addressed in contemporary political-criminal discourse, and this – paradoxically, at least at first glance(!) – without abandoning the reference to penalties other than imprisonment.

To understand this evolution, we need to understand the changes – and the meaning of the changes – that have taken place in the ways in which individuals are controlled.

In the criminal justice system, prison sentences (and long-term prison sentences) are extensively imposed. This is a widespread phenomenon (affecting European countries as well as South American countries such as the United States of America), that contributes to prison overcrowding – a factor largely responsible for the failure of a rehabilitation policy in the penal system.[5]

This trend was initially associated with a purely retributive movement (*just deserts*). Later, it became associated with a "new penology" – an ideology that originated in the United States but quickly spread to Europe – based on risk analysis, where the key concepts were security and dangerousness. Despite being labelled "new",[6] this trend pursued a long-standing desire to eradicate crime, incorporating (economic) cost-benefit considerations into the discussion about the purposes of punishment. This "new penology" made the concept of security absolute and reinterpreted the concept of dangerousness in the light of economic rationality. It thus resurrected the old concept of "being innocuous", which was renamed "incapacitation". But this was a "selective incapacitation", which had significant economic advantages over indiscriminate incapacitation. Selective incapacitation means that a certain number of offenders, selected on the basis of their dangerousness, are kept in prison for as long as possible. These "risk offenders" are identified using a "risk assessment" method with the aim of "managing risk". This whole method relies on the use of "indicators", which are quantified to make prognostic judgments about the dangerousness of certain groups or classes of individuals. The sentence and its duration are not based on the nature of the crime or the personality of the offender, but rather on the assessment of the risk profile. The duration of control is determined by the level of dangerousness, with more prolonged control for higher levels of risk. This is a punitive policy affiliated with "administrative criminology", also known as "actuarial justice", which has been taken to its ultimate consequences in North American and Anglo-Saxon penology.

The change is significant. Whereas the "old" penology was based on the individual and was concerned with the causes of crime in order to "correct" them, the "new" penology is interested in the risk group in which the individual is placed, in order to incapacitate, monitor and control him or her.[7]

Security now relates to the dangerousness of the offender and the suspicion of being an offender, rather than the offender's guilt in committing the offence

[5] Anabela Rodrigues, "Superpopulação carcerária. Controlo da execução e alternativas", *Revista Eletrônica de Direito Penal*, 1 (2013): 13 and following. Jean-Paul Céré, "La surpopulation carcérale entre contraintes européennes et réalité française", *Revista Eletrônica de Direito Penal*, 1(1) (2013): 183 and following. See also Penal Reform International, "Global Prison Trends", 2023, 10.

[6] Anabela Rodrigues, "Política criminal – novos desafios, velhos rumos", in *Liber Discipulorum para Jorge de Figueiredo Dias*, ed. Manuel da Costa Andrade *et al.* (Coimbra: Coimbra Editora, 2003), 216 and following.

[7] Anabela Rodrigues, "Novo olhar sobre a questão punitiva", in *Educar o Outro. As Questões do Género, dos Direitos Humanos e da Educação nas Prisões Portuguesas* (Humana Global. Publicações Humanas, 2007), 117 and following.

and the proof of his or her involvement. In this context, the criminal justice system is now defined by the use of imprisonment, which is being employed for longer periods and purely as a mechanism for confining individuals. In light of this security logic, the idea that "prison works" is promoted and an institutional strategy of punishment is fostered. After all, if imprisonment can do nothing else, it can at least delay the resumption of criminal activity by dangerous individuals. So, this ideology reduces the impact of crime on society, albeit without changing either the offender or society.

However, the defenders of actuarial justice often forget that risk is not a one-dimensional concept. Risk is a plural and heterogeneous concept, which can lead to very different management models.[8] Furthermore, it is worth emphasising that risk cannot be understood as a neutral concept, nor can its management be understood as a purely objective technology. In fact, it is always necessary to make "normative decisions" regarding, for example, what should be valued as a risk, which risks should be prioritised or what should be considered a permissible risk.[9]

Alongside that new way of understanding prison sentences, there has also been – as already mentioned – a renewed interest in non-custodial sentences among political decision-makers. What now needs to be considered is the meaning of the sudden interest in these sanctions, which were originally rooted in a fight against (short-term and medium-term) prison sentences, and with the purpose of socialisation. We need, on the one hand, to question these "sanctions" in terms of their political and criminal significance in the new security context and, on the other hand, to prevent the risk of their rationale being subverted.

The truth is that, in many cases, these sanctions are in line with the principles of "new behaviourism" and do not aim to reform, reintegrate or rehabilitate. These sanctions are seen merely as managerialist techniques aimed at limiting freedom of movement, reducing their effects to intensifying control over the individual (*i.e.*, they do not combine treatment with control). And yet again, do not change anything about the offender or society.

Furthermore, an analysis of the movement of offenders over a certain period of time shows a continual flow of individuals between the community (complying with non-custodial sentences) and prison. This is because there is less tolerance in the case of violations of the conditions that make it possible to implement sanctions in the community and which lead to the serving of a prison sentence (by revoking the non-custodial sanction). The aspect now in question is that of *transincarceration*, resulting from the autopoietic system created by the proliferation of mutually reinforcing sanctions, which allows for what has already been called the "recycling" of the individual, facilitating their circulation through different instances of control.[10]

In this way, it is once again the increased use of prison that is at stake, now boosted by a generalised expansion of the criminal justice system that favours incarceration.

---

[8] José Ángel Brandariz García, "La difusión de lógicas actuariales y gerenciales en las políticas punitivas", *InDret*, 2 (2014): 17.

[9] José Ángel Brandariz García, "La difusión de lógicas...", 8.

[10] Anabela Rodrigues, "Novo olhar sobre a questão punitiva...", 124. José Ángel Brandariz García, "La difusión de lógicas...", 16. Anabela Rodrigues, "Controlar e punir – o Direito Penal em mudança?", *FIDES* 8 (2017): 165.

3. The impact of the New Public Management on the criminal justice system

The spread of this actuarial trend in criminal justice policy has also been due to another, more general, development in the management of public policy: *New Public Management*. It is now recognised that criminal policy is another public policy that needs to be carried out rationally, using the most modern management techniques. This New Public Management has entered the field of criminal policy and there is every indication that its influence will not be merely short-lived. It has been recognised that public resources are not inexhaustible, and that better management of these public resources can also lead to improvements in the criminal justice system.[11] In this context, there is a growing recognition of the importance of adopting an Offender Management System, *i.e.*, an information system that includes information related to all aspects of the life of offenders during their journey and which is needed in order to take decisions on their cases while they are serving their sentences in prison or community settings.[12]

Nevertheless, only an open and unprejudiced discussion of the characteristics of this New Public Management will make it possible to ascertain its rightful place in the management of the criminal justice system. The fulfilment of *performance* expectations, typical of (administrative) audit and evaluation policies, embodied in indicators such as OBTJ (Offenses Brought to Justice), KPI (Key Performance Indicators) or KPT (Key Performance Targets) can measure the *internal* functioning of the punishment system. But they totally neglect criteria for evaluating punishment centred on the social effects *external* to the system. New Public Management is concerned with *efficiency*, *i.e.* how the functions of government are carried out, to the detriment of *effectiveness*, *i.e.* the purpose of those functions. This new administrative management relegates the achievement of external objectives to second place in favour of productivity and performance standards.[13] If, as in the case of criminal policy, it is a social policy that focuses on people – society and the offender –, it must be recognised that a purely performative model for evaluating its success overlooks the social effects and benefits that it is supposed to produce. With this change in the management of the criminal system, the external goals of crime prevention and reduction lose their prominence. In addition, values that are essential to the resolution of criminal conflicts, such as the guarantees and fundamental rights of citizens, are relegated to the background.[14]

4. The effects of the use of AI and related digital technologies by prison and probation services

Nowadays, this new model of prison management is increasingly being supported by the use of AI and related digital technologies.

Historically – from the 18th century onwards –, prisons in the West were designed and built to optimise the observation of prisoners. The immediate benefit

[11] Anabela Rodrigues, "A política criminal no Estado de Direito...", 26.
[12] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030: an exploration on the future of AI in prisons", *Advancing Correction Journal*, 11 (2021): 132. Pedro das Neves, "Towards an intelligent offender manage system", *Justice Trends*, 10 (2023): 33.
[13] José Ángel Brandariz García, "El new public management y las políticas penales", *Revista Nuevo Foro Penal*, 87 (2016): 185 and following.
[14] Anabela Rodrigues, "A política criminal no Estado de Direito...", 26. José Ángel Brandariz García, "El new public management...", 186.

of this constant surveillance was compliant behaviour from prisoners. What is more, this form of surveillance also had the advantage of generating information about the prisoners that prison authorities used to classify them as particular criminal types and to predict recidivism.[15] In the late 20th century, new remote surveillance technologies obviated the need for a physical prison architecture that maximised such direct observation of inmates.[16] In the 21st century, newer surveillance technologies, sometimes managed by AI, have increased the reach and intensity of this scrutiny in unprecedent ways.[17]

Some prisons around the world are becoming *smart prisons*, *i.e.*, prisons with digital capabilities.[18] The most advanced AI applications in the prison context are currently seen in Asian countries. In Singapore there is a fully automated prison, that does not have guards; AI and robots do everything.[19] In China, a smart surveillance system has been designed to monitor prisons at every moment, including when they are in their cells.[20] In Hong Kong, a smart video surveillance system includes cameras with analytic monitoring that can detect unusual behaviour of inmates and alert officials.[21] In South Korea, robots are used to monitor violence and inmates' suicide risk.[22] But, for instance, in United States prisons, in New York and Alabama, there is also a mass-monitoring system of inmates' phone calls that uses speech-recognition technology, semantic analytics and machine learning software to build databases of searchable words that can flag suspicious calls.[23] And a prison in Liverpool also uses security cameras monitored by AI to detect suspicious behaviour from the prisoners.[24]

---

[15] Mike Nellis, "Ethical, Strategic and Operational Guidance on the Use of Artificial Intelligence in Prison and Probation Services and the Private Companies acting on their Behalf", written for the PC-CP, Strasbourg, 10 September 2021, 16.

[16] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16.

[17] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16.

[18] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 129 and following. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 15. Pia Puolakka, "Smart prisons and artificial intelligent systems expand in Finland", *Justice Trends*, 10 (2023): 91.

[19] Jan Kleijssen, "Justice and beyond. Council of Europe working on setting global benchmarks on artificial intelligence. Interview", *Justice Trends*, 7 (2021): 27 and following.

[20] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16. Sophia Yan, "Chinese high-security jail puts AI monitors in every cell 'to make prison breaks impossible'", *The Telegraph*, 1 April 2019, https://www.telegraph.co.uk/news/2019/04/01/chinese-prison-rolls-facial-recognition-sensors-track-inmates/. Aleš Završnik, "Criminal justice, artificial intelligence systems, and human rights", *ERA Forum*, 20 (2020): 572.

[21] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16. Kristin Houser, «China is installing "AI guards" in prison cells. They'll make escape impossible – but the trade-off might be inmates' mental health», *Futurism*, 4 February 2019, https://futurism.com/chinese-prison-ai-guards-cells.

[22] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. "Robotic prisons wardens to patrol South Korean prison", *BBC News*, 25 November 2011, https://www.bbc.com/news/technology-15893772.

[23] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16 and following. Debra Weiss, "Prisons and jails use artificial intelligence to monitor inmate phone calls", *Abajournal*, 25 October 2019, https://www.abajournal.com/news/article/prisons-and-jails-use-artificial-intelligence-to-monitor-inmate-phone-calls.

[24] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16. Cara McGoogan, "Liverpool prison using AI to stop drugs and weapons smuggling", *The Telegraph*, 6 December 2016, https://www.telegraph.co.uk/technology/2016/12/06/liverpool-prison-using-ai-stop-drugs-weapons-smuggling.

In fact, there are three key areas where AI and related digital technologies can be used in criminal justice organisation: (i) use for safety, security and good order; (ii) use for offender management purposes (risk assessment, rehabilitation and reintegration); and (iii) finally, use for staff selection, management, training and development.[25]

4.1. The use of AI and related digital technologies for safety, security and good order

The majority of AI applications in the prison context are in the realm of security technique.[26] Image recognition, audio recognition and movement analysis are some AI technologies that can be used to gather intelligence about offenders (and, in some cases, about their interlocutors). As we have just seen, these different techniques can also be used in combination to create surveillance systems of an extremely intrusive nature.[27]

The use of AI and related digital technologies in prisons should allow for better risk and crisis management. These new technologies, even when used for security purposes, should also help staff – who are relieved of their habitual repetitive tasks, such as opening and closing doors, monitoring movements and behaviour, etc. – develop and maintain positive human relations thus enhancing rehabilitation and social inclusion of offenders.[28] However, the truth is that the use of these digital security techniques managed by AI carries the risk of turning the prisons into institutions of control rather than rehabilitation.

There has also been an increase in the use of various technologies in non-custodial sentences. The use of AI could lead to profound changes in the way new technologies are used in probation and community correction (including parole) settings. Electronic monitoring, telephone check-ins and biometric check-in *kiosks* for automating probation reporting conditions (which had their roots in the eighties and nineties) could now expand and be managed by AI systems. Smartphones (and all the data they store...) are already being used in the framework of probation, both for use as tracking devices (a mainly United States development) and as a basis of "probation with apps" (developing both in Europe and United States). The use of all these technologies has an enormous potential from a rehabilitation perspective. However, the current trend is to use them mainly for security and control purposes.[29]

---

[25] Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services. Document prepared by Håkan Klarin, Pia Puolakka and Fernando Miró Llinares. Council of Europe – Council for Penological Co-operation. Strasbourg, 8 February 2023.

[26] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 131. Mike Nellis, "Ethical, Strategic and Operational Guidance...", 16. Penal Reform International, "Global Prison Trends", 2021, 47.

[27] Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 14.

[28] Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 13.

[29] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 18 and following. Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 14.

4.2. The use of AI and related digital technologies for offender management purposes

The use of AI and related digital technologies in prisons offers the chance to alleviate staff of the routine tasks that they are responsible for on an everyday basis (e.g. assisting offenders in their contacts with lawyers, psychologists, and social workers as well as with their families).[30] AI applications are being used in offender management systems to support decision-making throughout the entire offender management cycle, including the assessment and classification of offenders, planning, executing, evaluating, and adjusting services for offenders.[31] During and after imprisonment, algorithms may be used to determine security levels, eligibility for parole, and conditions of community supervision.[32] With the use of AI in the context of offender management, the punitive system aims to "improve decision-making related to finding the best trajectory for the offenders regarding their needs and minimizing their risks".[33]

The first (and still more common) use of AI applications in prisons was in the context of risk assessment tools.[34] AI applications have significantly enhanced the capabilities of existing risk assessment tools. In fact, data-driven decision-making is far from new in the criminal justice context. As mentioned before, over the past few decades, there has been a shift in the criminal justice system towards actuarial justice, in which the punitive system is based on risk management. However, current technologies differ significantly from their predecessors in several important ways. On the one hand, the amount and variety of available data has significantly increased. On the other hand, computational techniques have become much more advanced and refined.[35] The first wave of risk-assessment tools relied only on "static", unalterable factors (e.g. history of substance abuse, age at first offence, etc.). More recent instruments, based on a large amount of data, use sophisticated methods and "dynamic" risk factors (e.g. variables about employment status, criminal friends, etc.) that can be adjusted over time.[36] Nowadays, risk assessment tools have been described as specifically "algorithmic" in the sense that "they rely on omnivorous data collection and machine learning models to identify patterns, becoming opaque and 'black boxed' in the process".[37]

Advocates for algorithmic technologies emphasise the benefits of using "smart statistics" in order to reduce crime and improve a dysfunctional punitive system

[30] Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 16.
[31] Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 17.
[32] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts", *Social Problems* 68(3) (2021): 613.
[33] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 132.
[34] Pia Puolakka and Steven Steene, "Artificial Intelligence in prisons in 2030...", 132.
[35] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts", *Social Problems* 68(3) (2021): 612 and following.
[36] Angèle Christin, "Predictive Algorithms and Criminal Sentencing", in *The Decisionist Imagination,* ed. Nicolas Guilhot and Daniel Bessner (Berghahn Books, 2018), 277 and following.
[37] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction...", 611. Rodrigues AM. "O(s) tipo(s) de medida da pena ou a necessidade de um 'algoritmo a ser seguido'", in *I Congresso – Inteligência Artificial e Direito*, ed. Anabela Rodrigues e Susana Aires de Sousa (Coimbra: Almedina, 2023), 300 and following.

characterised by racial discrimination and mass incarceration.[38] On the other side, critics point out that algorithms tend to reinforce social and racial inequalities instead of reducing them and emphasise that these new risk assessment tools make part of a "culture of control", based on surveillance of "risky" groups.[39]

The use of predictive technologies is often justified by citing accountability pressures and new budgetary constraints.[40] Algorithms are usually described as a rationalising force; they help judges, prosecutors and probation officers make more informed decisions (e.g. about bail, sentencing and parole) by providing them with reliable information; algorithms would be better than humans at making decisions because they are value-neutral (decision-making will be free from subjective and emotional elements).[41] Hence, a dual justification emerged for using predictive algorithms in the criminal justice system: firstly, an *efficiency* argument, which described predictive algorithms "as a cost-cutting device at a time of funding and budgetary constraints"; and secondly, an *objectivity* argument, which presented algorithms "as a means to increase accountability and mitigate bias".[42] The use of algorithms is therefore, an attempt to remove discretionary power in order to diminish bias.

However, as we have already mentioned, in a social system such as the criminal justice system, it is important to prioritise effectiveness over efficiency. On the other hand, as is well known, most of the debate regarding AI technologies has focused on the internal bias of algorithm instruments (since the algorithms learn by being fed historical data, inequalities from the past will be projected into the future).[43] In this way, it is hard to see how the use of algorithms really helps to reduce bias in criminal decision-making. Questions also arise regarding whether predictive algorithms can erase discretion in the criminal decision-making process. In fact, the adoption of these digital technologies makes professionals feel that their autonomy and experiential knowledge are under threat. As a consequence, they tend to manipulate the data in order to regain the autonomy they feel is being threatened. Furthermore, the implementation of these new technologies results in the involvement of additional actors in the criminal justice system (data analysts, data entry specialists, technology teams, etc.). This, in turn, leads to the emergence of novel forms of discretionary power within these institutions. All of this leads us to the conclusion, with Brayne and Christin, that "far from eliminating discretionary power in its various forms, the adoption of predictive algorithms in fact displaces discretion to less visible parts of the organization".[44]

From the strict point of view of the rehabilitation and reintegration of offenders, AI can be used in education and training platforms, in various treatment procedures and in preparation for release. AI can assist rehabilitative processes, and programmes or individual therapeutic work can include AI based methods such as virtual reality. Virtual reality makes it possible to meet the special needs of different offender populations and overcome language barriers in an inexpensive

[38] Angèle Christin, "Predictive Algorithms and Criminal Sentencing", 277 and following.
[39] Angèle Christin, "Predictive Algorithms and Criminal Sentencing", 273.
[40] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction...", 614.
[41] Angèle Christin, "Predictive Algorithms and Criminal Sentencing", 275.
[42] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction...", 615.
[43] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction...", 620.
[44] Sarah Brayne and Angèle Christin, "Technologies of Crime Prediction...", 620.

way.[45] What must not be forgotten is the immense amount of data generated by each user while immersed in virtual reality.[46]

In any case, and as we have already seen, despite the great potential that AI and related technologies have from the point of view of rehabilitation, these new technologies are currently used mainly for safety and control purposes.

The use of robotic systems for rehabilitative tasks besides security tasks is another possibility being discussed. In the United States, some scholars are discussing the possibility of using AI to address the solitary confinement crisis, by employing smart assistants as a form of 'confinement companion' for prisoners. It has, however, been emphasised that even though these 'companions' might alleviate some of the physical stress of some prisoners, this apparent 'solution' might actually contribute to the legitimisation of a policy of solitary confinement rather than questioning it.[47] In Europe, a similar discussion might take place regarding 'special units' – prisons within prisons – for the most dangerous offenders. The use of digital technologies in such units may have some advantages (e.g. in cell communication with prison managers, on-line learning facilities or video contact with family), but possibly at the expense of a further decrease in human contact. We agree with Nellis when he emphasises that it is not inconceivable that the (established) concept of "meaningful human contact" will need to be debated in the context of highly digitised prisons.[48]

Conclusion

The use of AI and related technologies in imprisonment and non-custodial sentences is part of a broader movement that indicates a paradigm shift in the state's approach to crime control in a security sense.[49] In terms of legitimisation and effectiveness, the classic repressive model established until now, based on prevention through the threat of punishment (indirect prevention model), is at stake. Through the "technological shaping" of behaviour, constant and widespread surveillance of people is permitted, with the aim of preventing criminal activity. Progressively, the repressive model of indirect prevention is being replaced by a model of direct prevention, which seeks total control of behaviour and the full efficiency of penal regulation.[50]

---

[45] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 20. Draft Committee of Ministers Recommendation CM/REC(2023)XX. Ethical and organisational aspects of the use of artificial intelligence and related digital technologies..., 18. Ana Rita Pires, *et al.*, "The potential of virtual reality for education and training in prisons", January 2021, 33 and following, http://dx.doi.org/10.13140/RG.2.2.29588.63366. Melissa Teng and Eric Gordon, "Therapeutic virtual reality in prison: Participatory design with incarcerated women" *New Media & Society* 23 (2021): 2210-2229, https://doi.org/10.1177/1461444821993131.

[46] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 20.

[47] Aleš Završnik, "Criminal justice, artificial intelligence systems, and human rights", 573. On the situation in Japan, see Penal Reform International, "Global Prison Trends", 2023, 46.

[48] Mike Nellis, "Ethical, Strategic and Operational Guidance...", 18.

[49] Christoph Burchard, "Artificial Intelligence as the end of criminal law? On the algorithmic transformation of society", in *Artificial Intelligence in the Economic Sector. Prevention and Responsibility*, ed. Maria João Antunes and Susana Aires de Sousa (Coimbra: Instituto Jurídico da Universidade de Coimbra, 2022), 191 and following. Anabela Rodrigues, "Compliance Criminal Digital – O Oráculo da Era da Inteligência Artificial", *Bulletin of the Faculty of Law – FLL Bulletin*, Macau. 53 (in the process of being published).

[50] Anabela Rodrigues, "Compliance Criminal Digital - O Oráculo da Era da Inteligência Artificial".

This security strategy to control crime has been gradually implemented in our societies of fear. There are numerous reasons and causes for this change. We live in a time when "governing" no longer means guaranteeing safety, but managing dangers and risks. This process – as well as what Shoshana Zuboff describes as a global behaviour modification project that threatens to transform human nature in the 21st century –[51] cannot be seen solely through the lens of scholars and professionals who apply the criminal law. In the era of AI, criminal policy faces complex challenges. It is tempting to slide into increased and mechanised control. The rehabilitative function of criminal sanctions is juxtaposed with a model of surveillance in which the image is that of an "omnipresent digital architect".[52] The meaning of the prisoner's right to rehabilitation needs to be reconsidered and redefined.[53] New tools should not undermine the human-centred approach. Technology should be a form of complementing, not reducing or replacing human contact, which is a fundamental dimension of the social reintegration process.[54]

It is politics – not management or technology –, and in this case, it is criminal policy, that is responsible for creating a justice criminal system focused on control and security. Transforming this criminal justice system is a political task.[55] With political guidance and power, AI and related digital technologies can be used in a way that is more favourable to rehabilitation practices and can actually contribute to reducing the use of imprisonment. In this way, politics will ensure that the control it imposes on citizens can be balanced by an ethical use of AI and related technologies, with a dimension of social responsibility which should shape collective action.

[51] Shoshana Zuboff, *A era do capitalismo da vigilância. a disputa por um futuro humano na nova fronteira do poder* (Relógio D'Água, 2019).

[52] Shoshana Zuboff, *A era do capitalismo da vigilância. a disputa por um futuro humano na nova fronteira do poder*.

[53] Anabela Rodrigues, "Execução penal socializadora e o novo capitalismo - uma relação impossível?", 26 and following.

[54] Anabela Miranda Rodrigues, Maria João Antunes, Sónia Fidalgo, Inês Horta Pinto and Karla Tayumi Ishiy, "Non-custodial sanctions and measures in the Member States of the European Union", *Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra,* November 2022, 106, https://doi.org/10.47907/livro/2022/Custodial-sanctions-measures.

[55] Anabela Rodrigues, "A política criminal no Estado de Direito do Século XXI - os desafios da segurança", 37. José Ángel Brandariz García, "El new public management y las políticas penales", 209. Anabela Rodrigues, "Compliance Criminal Digital - O Oráculo da Era da Inteligência Artificial".