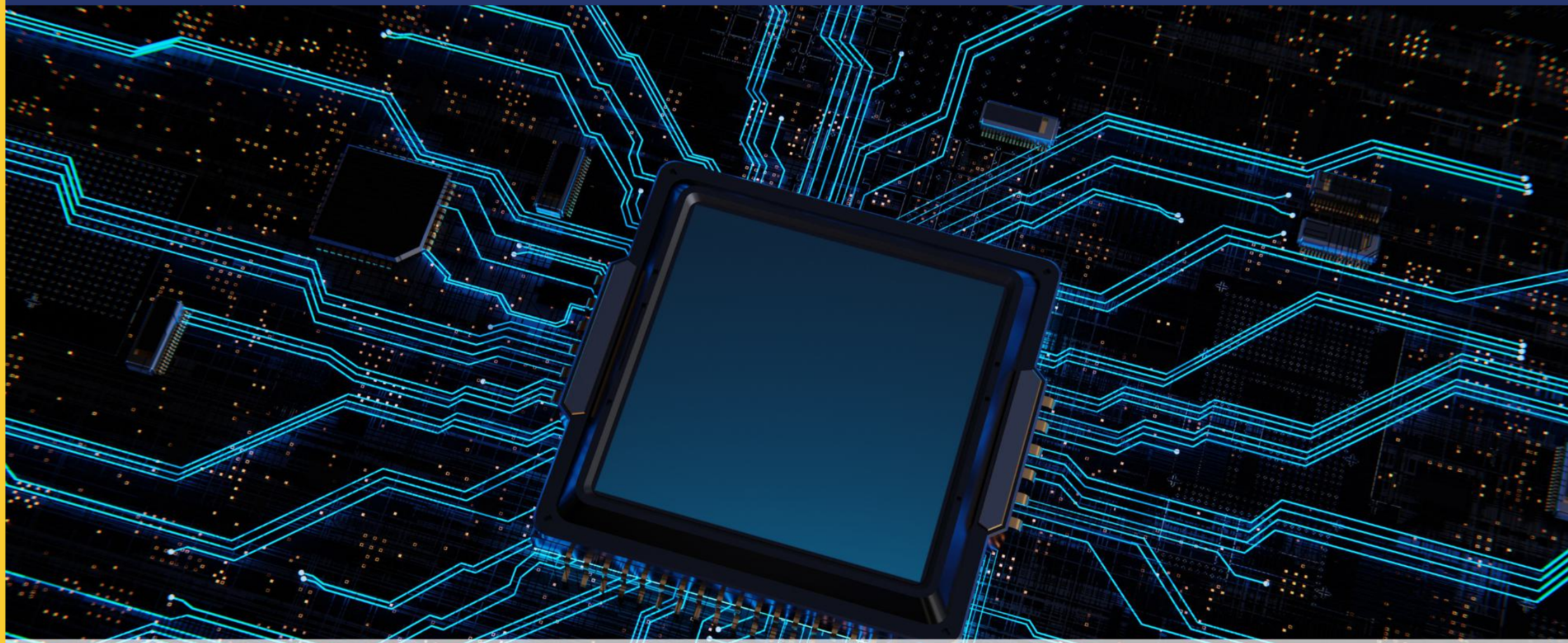
The Implementation of the Council of Europe Recommendations for the Use of Artificial Intelligence (AI) in Prison and Probation Services



EUROPRIS
PROMOTING PROFESSIONAL PRISON PRACTICE

About

The AI in Prisons Webinar brought together experts, practitioners, and policymakers to explore the transformative potential of artificial intelligence within the prison and correctional systems.

The event featured insightful presentations from leading figures in the field, followed by thought-provoking discussions among panellists, addressing opportunities, challenges, and ethical considerations surrounding AI implementation.

This report summarises the key points from the presentation. It also captures the critical discussions from the panel sessions, offering a well-rounded view of the topics addressed during the webinar.

Principles of the Recommendation

Starting with exploring the Council of Europe's Recommendation on Artificial Intelligence in Prisons and Probation, which was published in Oct 2024, there is an emphasis on the need for AI and related digital technologies to be used legitimately and proportionately within the prison systems. These principles are grounded in the overarching goals of contributing positively to:

- The rehabilitation and reintegration of offenders
- Supporting prison and probation staff
- Enhancing the effectiveness of the criminal justice system

These principles were further developed as a foundational framework for guiding the implementation and use of AI in prison systems. This development represents a step forward in ensuring the responsible and ethical integration of AI technologies in this context.

Principles and Categories

The RecommendationCM/Rec(2024)5 regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services has 30 principles which are grouped into six main categories: are grouped into six main categories:

1. **Basic principles**
2. **Data protection and privacy**
3. **Use for safety, security, and good order**
4. **Use for offender management, risk assessment, rehabilitation, and reintegration**
5. **Use for staff selection, management, training, and development**
6. **Research, development, evaluation, and regular revision**

These categories provide a comprehensive framework for addressing the diverse aspects of AI applications, ensuring that they align with operational needs and ethical considerations. The presentation underscored the importance of adhering to these principles to balance innovation with safeguards, ensuring AI's role as a supportive tool rather than replacing human judgment and interaction.

1. Basic Principles

The foundational principles for using AI in the prison sector emphasise respect for human rights and dignity. While many of these principles align with general AI guidelines, they have been tailored to the specific needs and sensitivities of prison systems.

The aim is to ensure that AI use results in the least negative impact on human rights while delivering meaningful benefits.

Data management also features prominently within these basic principles, with a significant focus on the protection of data to ensure ethical and lawful handling.

2. Data Protection and Privacy

A core pillar of the recommendations is that offenders retain their fundamental rights and freedoms, including their right to private life and data protection.

Key points include:

- All actors must comply with data protection laws, acting transparently and demonstrating adherence to these principles to those affected.
- Data should only be stored in a manner that allows personal identification for as long as is strictly necessary to achieve its original purpose.

These measures ensure the responsible handling of sensitive information and safeguard offenders' privacy.

3. Safety, Security, and Good Order

The use of AI for safety and security must be driven by consultation with prison services, ensuring that the specific needs are carefully identified and evaluated.

The recommendations emphasise a human-centred approach, reinforcing that AI should support rather than replace human judgement in maintaining security and order.

4. Offender Management, Risk Assessment, Rehabilitation, and Reintegration

AI can be a valuable tool in managing offender files, improving monitoring systems, and assisting decision-making processes. However, the recommendations stress that final responsibility must always rest with professionals.

Additionally, the use of AI should never replace the importance of face-to-face contact, recognising the critical role human interaction plays in rehabilitation and reintegration.

5. Staff Selection, Management, Training, and Development

AI has the potential to optimise human resources and management processes within the corrections sector. It can:

- Support the professional development of staff.
- Assist managers in predicting future organisational needs, such as capacity planning.
- Detect potential issues within staff resources, enabling proactive measures.

This principle highlights the use of AI as a tool to enhance, rather than diminish, the role and effectiveness of staff.

6. Research, Development, Evaluation, and Regular Revision

Continued research is vital to ensure that AI applications in this sector remain effective and ethical. Adequate funding and support are essential to drive innovation while maintaining safeguards.

Regular evaluation and revision of AI use are necessary to adapt to new challenges and uphold human rights and fundamental freedoms.

These principles call for an ongoing commitment to accountability, ensuring that AI remains a beneficial and responsible asset within the correctional system.

Key Insights from the Panel Discussions

The Q&A session during the AI in Prisons Webinar provided an engaging opportunity for participants to raise critical questions and explore practical applications of the recommendations.

The discussions delved into challenges, opportunities, and actionable steps for adopting AI within the corrections sector. Questions posed by the audience ranged from the drafting process of the recommendations to the readiness of European prisons to implement AI solutions.

It also captures the nuanced perspectives on AI's role in enhancing prison operations while preserving human-centred approaches.

Was there a consultation with civil society in drafting the recommendations?

The panellists acknowledged that the recommendations were primarily developed by experts in the field. While the process was largely expert-driven, they expressed interest in gathering public comments and feedback once the recommendations were released.

Revisiting the recommendations with insights from civil society could provide valuable perspectives and enhance their applicability in practice.

How should governments, services, or organisations use the recommendations? Is there a way to formally pledge support?

Panellists suggested that services can begin by drafting an AI policy aligned with the recommendations. This policy should then be tailored to meet the specific needs of the service.

Formal adoption might involve integrating these principles into existing structures and processes, enabling services to progressively implement AI solutions while remaining aligned with ethical standards.

How do you implement governance for AI without overwhelming the organisation?

The governance of AI is inherently challenging, particularly in prison settings where capacities vary. Panellists emphasised starting small, such as:

- Building AI literacy to familiarise staff with AI concepts.
- Using a data science approach to address foundational needs before directly adopting AI.
- Identifying risks and priorities early in the process to guide governance efforts effectively.

The path to governance requires balancing operational needs with manageable frameworks to ensure AI implementation does not overburden the organisation.

Is there a difference in how the recommendations approach narrow AI versus deep AI?

The recommendations distinguish between different levels of AI:

- **Narrow AI (e.g., Human Form Recognition cameras for analysing images or videos) often comes with lower risks and has more defined use cases.**
- **Deep AI (e.g., analysing documents, specifications, or conversations) involves more complex applications and inherently carries greater risks.**

This distinction underscores the need for tailored approaches to using AI, depending on its scope and potential impact.

How ready are European prisons to adopt AI solutions?

Readiness varies significantly across European countries, often correlating with their digital maturity. For instance:

- Western European countries and Nordic nations generally have higher levels of readiness due to advanced digital infrastructures.
- Other European in the East countries, such as Türkiye, have also made significant investments in digitalisation in the past years, improving their preparedness for AI adoption.

However, readiness is not solely about technology; it also depends on how prison services manage data and their broader legal and organisational frameworks.

What is the current level of AI awareness among prison authorities?

AI awareness among prison authorities is inconsistent across Europe. While some governments have actively supported AI adoption, others face challenges due to limited digitalisation and competing priorities, such as overcrowding and staff shortages.

Panellists highlighted that **raising AI awareness is crucial for unlocking its potential.** For example, using AI for practical applications like language translation demonstrates its benefits, though adapting such solutions for offline use may be necessary in prison environments.

How should prison services begin their AI journey?

Starting small was a recurring theme in the discussion. Panellists suggested:

- Focusing on AI literacy to help staff understand its potential and limitations.
- Using examples and case studies to demonstrate the value of AI in corrections.
- Addressing leadership by appointing someone to oversee AI initiatives and creating a governance plan.

AI adoption should also be framed around clear organisational priorities, using it as a tool to address specific challenges rather than as a blanket solution.

How can prison authorities convince decision-makers to adopt AI?

Convincing leadership, such as Directors General, requires a strategic approach:

- Establish a clear business case for AI, identifying how it can address specific organisational needs.
- Start with a policy and governance plan that outlines the areas AI can improve.
- Consider inter-agency collaboration or apply for EU funding to support AI implementation and share costs.

Panellists also emphasised that AI should assist rather than replace staff, with human interaction remaining essential to prison regimes.

Lastly, how can collaboration support AI adoption?

Collaboration across countries and services can play a pivotal role in enhancing AI adoption within the correctional system.

By working together, stakeholders can develop a comprehensive handbook based on shared recommendations, which would serve as a guiding framework for governance across different jurisdictions.

Furthermore, promoting inter-agency initiatives to share knowledge, resources, and best practices can foster a unified approach to implementing AI solutions.

Collaboration could not only accelerate the adoption of AI but also ensure that its deployment aligns with established principles and ethical guidelines, creating a more consistent and effective system for all involved.

EUROPRIS
PROMOTING PROFESSIONAL PRISON PRACTICE

Funded by
the European Union