# Artificial Intelligence (AI) in Prisons


- **Pia Puolakka**

  Project Manager

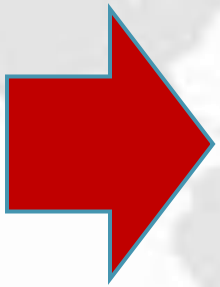
  Criminal Sanctions Agency

  Finland

# Introduction

Artificial Intelligence in Prisons in 2030.

An exploration on the future of Artificial Intelligence in Prisons.*



Literature review

Survey – 20 jurisdictions worldwide

Puolakka, P., & Van De Steene, S. (2021). Artificial Intelligence in Prisons in 2030. An exploration on the future of Artificial Intelligence in Prisons. *Advancing Corrections Journal*, Edition # 11, ICPA.

# Artificial Intelligence?

"Systems that display intelligent behaviour by analysing their environment and taking action — with some degree of autonomy — to achieve specific goals" (European Commission, 2018).

AI is currently not able to do more than the simulation of human intelligence processes like visual perception, speech recognition, decision-making and translation between languages (Tucci, Burns & Laskowski, 2020).

The development of AI is more about a revolution in computational statistics than a revolution in intelligence (Fry, 2018).

The **capabilities of AI are still extremely limited** relative to human intelligence (Andrew, 2017).

# AI Triad… Stages of AI

**ANI**

Artificial Narrow Intelligence

Better than humans
in one specific task

**AGI**

Artificial General Intelligence

Capable like humans
in every task

**ASI**

Artificial Super Intelligence

Better than humans
in every task

# Machine Learning

- Automated statistical technique that works by identifying patterns (relations) in available data and then applying this knowledge to new data. Relations in the data are turned into a set of rules = algorithms.
  - Algorithms learn by themselves from new data
  - More data → more accurate
- Examples: predicting phenomena, planning, optimization and pattern recognition (computer vision), speech recognition and natural language processing, robotics, optimizing the function of large systems (e.g., energy consumption) and improving expert systems

# Artificial Intelligence in Prisons

## Methodology

1) Literature review about the use of AI in prisons and ethical questions regarding its use
2) Digital survey about the use of AI in correctional services
   - EuroPris: member countries -> 13 / 32 respondents
   - The Corrections Technology Association: five North American jurisdictions
   - On authors' request: one Australian and two Asian jurisdictions

# The Survey

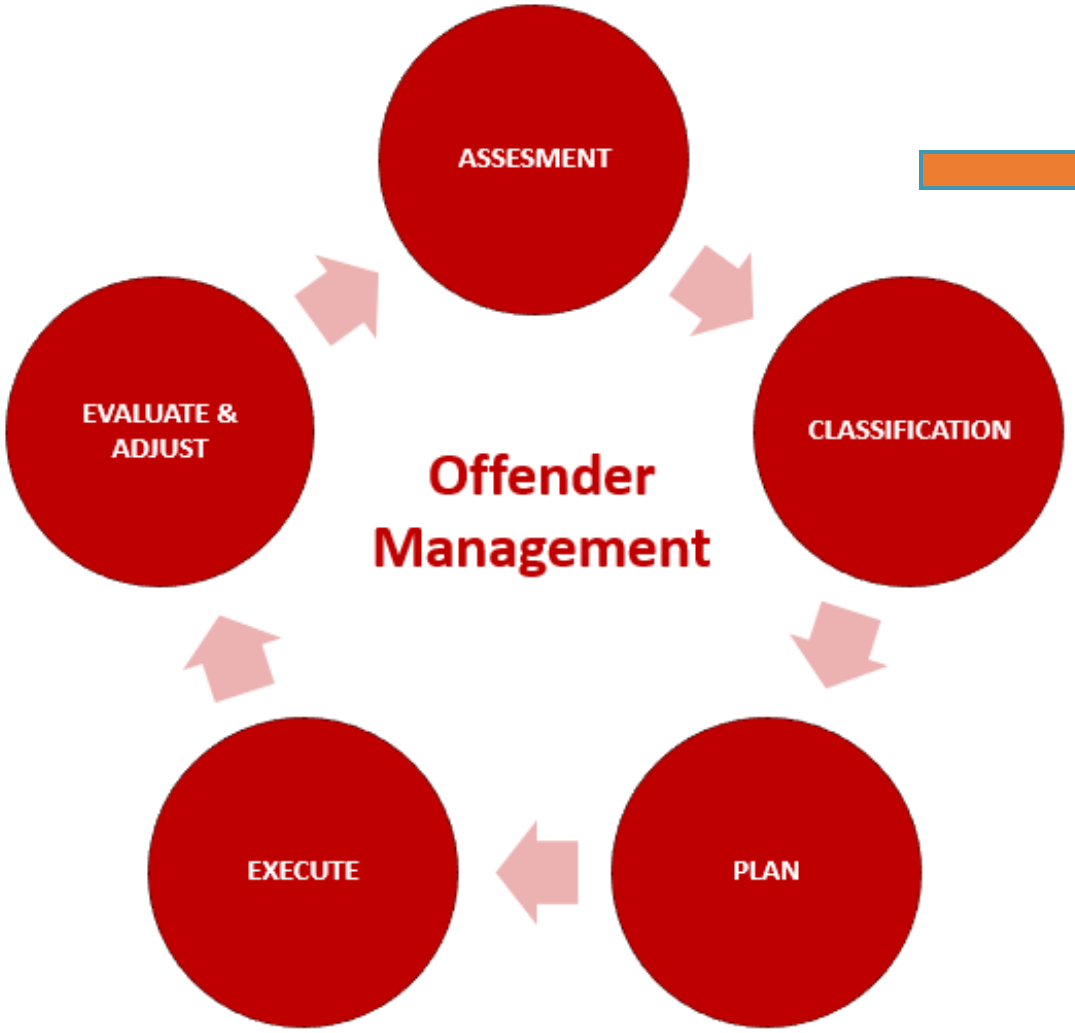
| |
|---|
| 1. Do you currently use AI in your correctional system? How? |
| 2. Are you planning to use AI? How? |
| 3. Are you planning / discussing the use of AI with your agency's management? |
| 4. Do you have concerns regarding risks or ethical questions connected to the use of AI? |
| 5. Have there been any new policies or legislation developed to support the current and the future use of AI in your organization? |

# Results: Literature Review

# Security Techniques

- The majority of AI applications in the prison context are in the security technique for "suspicious behaviour recognition" / surveillance
- **China:** smart surveillance in cells, facial identification and movement analysis (Yan, 2019)
- **Hong Kong:** smart video surveillance, tracking wristbands and robotic arms (Houser, 2019)
- **South Korea:** robot surveillance for violent and suicidal behaviour (BBC News, 2011)
- **UK:** security cameras monitored by AI to stop smuggling of drugs, phones and weapons into (McGoogan, 2016)
- **USA:** mass-monitoring system of inmate phone calls to detect suspicious behaviour (Cassens-Weiss, 2019)

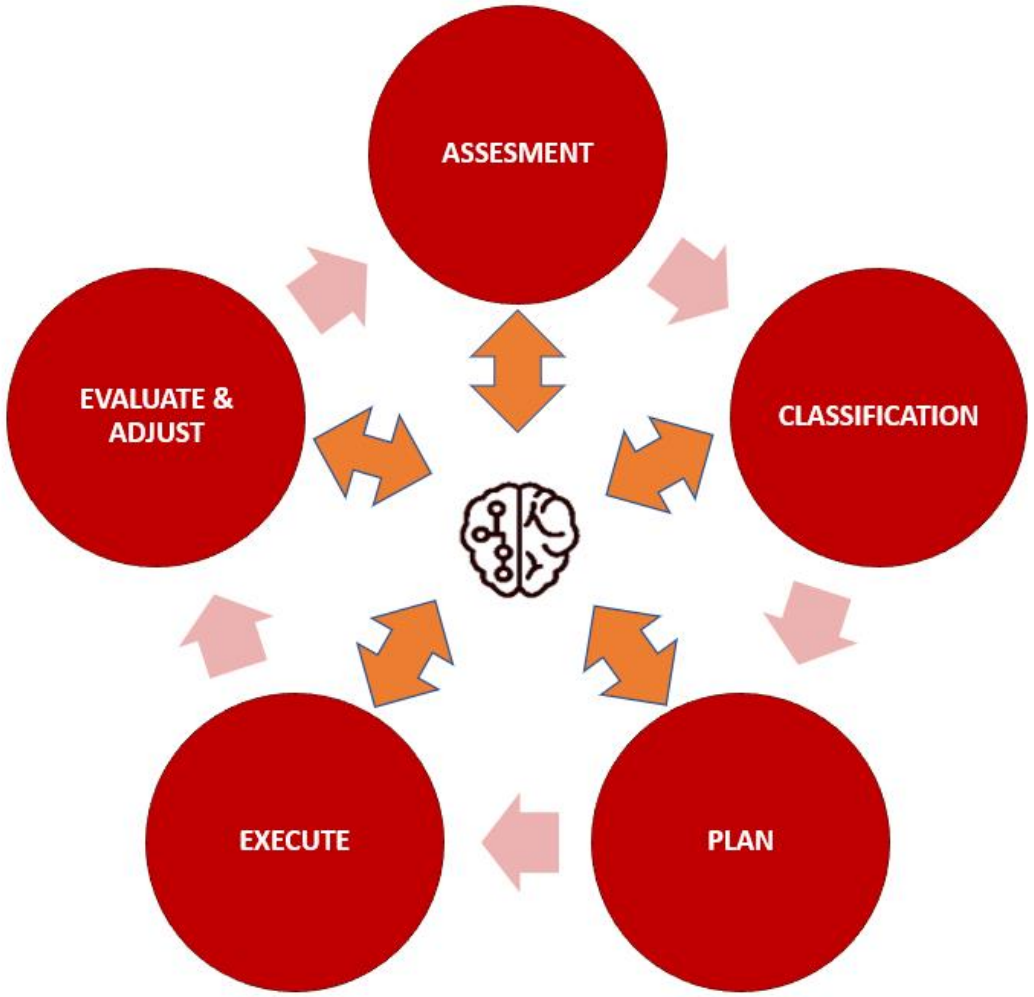
# AI in Offender Management




**Risk Assessment**

**Decision Making**

# AI in Offender Management




**Supporting decision making during the entire offender management cycle**

To enhance the efficiency of penal operations and even the effectiveness of rehabilitation programmes (Houser, 2019)

**Rise AI Recommender System (see Finnish example)**

# AI in Offender Management - Reflections

- Algorithms are no more accurate than laypeople's estimates of recidivism (Dressel & Farid, 2018)
- Nearly identical results, with humans and algorithms performing comparably (Lin et al., 2020). … but nuanced:
  - Algorithms can outperform human predictions of recidivism in "ecologically valid settings" (Lin et al., 2020.)

…a more **complex** set of parameters (an enriched set of risk factors)

# AI in Offender Management - Reflections

- Algorithms can overcome the harmful effects of cognitive biases (Sunstein, 2018)
- Algorithms can also easily be biased and start to repeat the same mistakes humans make
- Designing an algorithm for use in the prison context requires thinking deliberately about what it is that we exactly want to achieve and a solid understanding of the human failings they're supposed to be replacing (Fry, 2018)

# AI in Rehabilitation – Better than Human?

And what about the robots?

- Smart Assistants
- Buddy Robots
- Security Controls…

…A good virtual companion is perhaps a better solution than a bad physical one?...

# Ethical Questions

- Increase monitoring & control

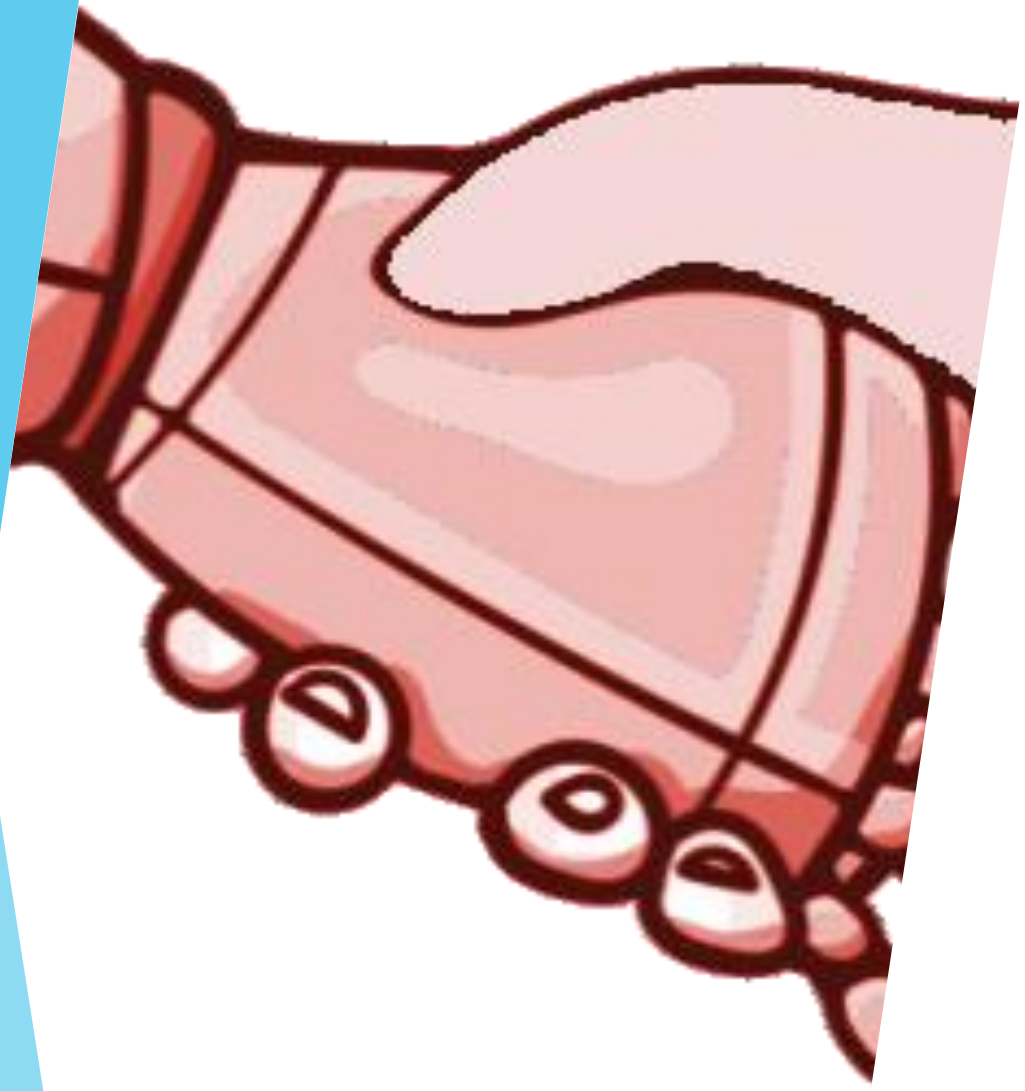
  → even more punitive interventions & negative side effects

  → bias and 'self fulfilling prophecy'

- Apply an **'ethics of technology'** (Nellis, 2019)

Make decisions taking in count the **intrusive character of the monitoring sensors: Where? When? What? How?**

# Ethical Questions



- Algorithms trained on what we do today?
- **Let's stay modest** and use this technology to help us improve our knowledge and understanding first…

→ AI in **advisory role only**

→ **Don't replace** face-to-face contact

→ **Ethical Design Principles** (Knight & Van De Steene, 2020)

# Ethical Charter EU*

- **Principle of respect for fundamental rights**
- **Principle of non-discrimination**
- **Principle of quality and security**
- **Principle of transparency, impartiality and fairness**
- **Principle “under user control”**

*** European Union Agency for Fundamental Rights, 2020**

# Results: Survey to Jurisdictions

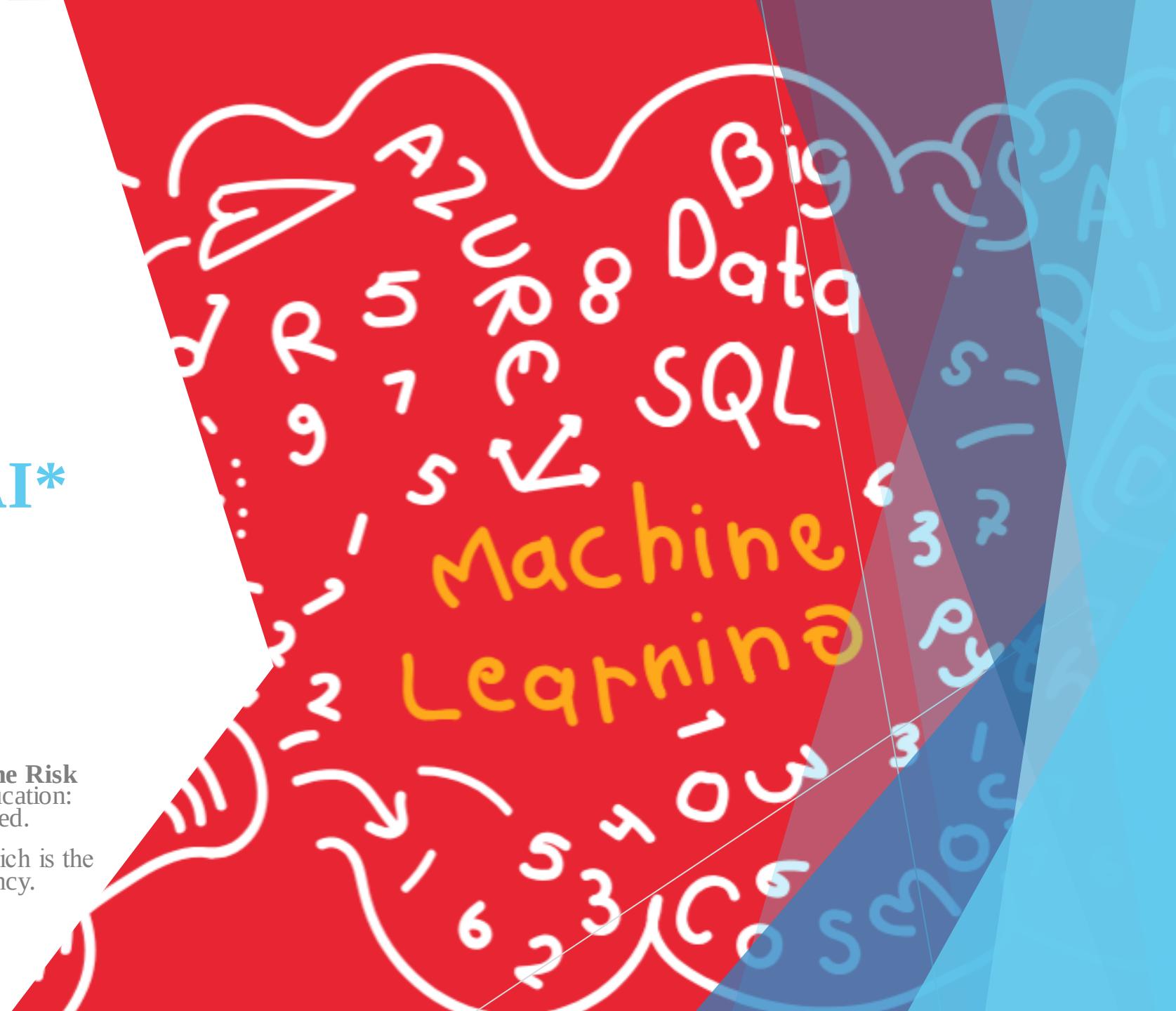
# AI in Prisons Today

- **Q1:** Most jurisdictions don't use AI yet
- **Q2:** The majority of the respondents indicated having plans to use AI both in security technique and offender management including decision making related to assessments and sentence planning, but they gave no real planned examples of possible AI projects
- **Q3:** The need for and interest in AI solutions in the future was acknowledged by all respondents, also considering the management level
- **Q4:** Ethical concerns included biased data-sets leading to wrong or inaccurate decisions, fairness, transparency, privacy, and GDPR questions. There was a need for ethical guidance
- **Q5:** There was also a need for new policies and legislations

# Jurisdictions Using AI

- **UK:** Natural language processing on case notes and free text is used to identify national and local themes / issues and some actuarial risk scores as part of the risk management processes. Machine learning is used to estimate the rate of violent incidents in prisons. Ethical guidance is built upon the Office of AI and Government Digital Service which was under development.
- **Singapore & Hong Kong:** AI is used in both security technique and offender management (see literature review). AI strategy is driven from broader national digital strategies on data-driven government (Singapore) or Smart Cities (Hong Kong).
- **Finland: RISE AI** project for the new offender management system

# The Finnish Example: RISE AI*




*Puolakka, P. (2020). **RISE AI: Reducing the Risk of Recidivism with AI.** Aalto Executive Education: Diploma in Artificial Intelligence. Unpublished.

'RISE' is short for Rikosseuraamuslaitos, which is the name of the Finnish Criminal Sanctions Agency.

# Sentence Planning Process

- Based on the mainly Anglo-American evidence-based research on offending and criminology since the 1990s' and the "What Works" debate emerging from it (McGuire, 1999)
  - Risk-Need-Responsivity (RNR) model (Andrews, Bonta, & Hoge, 1990) and assessment of strengths and level of motivation (Good Lives Model GLM by Ward, 2002)
- Database in Offender Management System (OMS): data from official documents and expert interview

**1) Assessment:** Risks and strengths

**2) Planning:** Objectives of the sentence

**3) Service counselling:** Rehabilitative activities

**1) Assessment:** Actuarial risk assessment instruments (ARAIs -> ARAT, SAQ etc.)

**2) Planning + 3) Service counselling:** Expert-view-based

# RISE AI Recommender System

- An algorithm that is based on relevant background factors (criminogenic factors) obtained from OMS and expert interviews and produces a list of **(1)** risk and strength factors and recommendations for **(2)** sentence objectives and **(3)** rehabilitative activities

- AI-based expert and recommender system using machine learning (learning from new data and sentence plans in OMS)
- An assessment and recommendation tool for senior coordinators responsible for the sentence planning process

→ Not automated decision making!

# Sentence Planning Process with RISE AI

- Based on the Risk-Need-Responsivity (RNR) model and strengths / level of motivation
- Database in Offender Management System (OMS): data from official documents and expert interview:

**1) Input data:** risk, strength and other relevant background factors (criminogenic factors)

**RISE AI ALGORITHM**

**2) Output data:** List of relevant risk and strength factors (1)

**3) Output data:** Recommendations for objectives of the sentence (2) and rehabilitative activities

**4) Expert analysis:** output data recommendations + expert-based view

# Effectiveness of / Training the Algorithm?

- Number of offenders having complied with the recommendations made and their progress assessed by the staff
- Number of participants in various activities in general
- Number of refusals or interruptions of activities
- Subjective but structured feedback from offenders
- Efficiency of the use of staff resources in sentence planning and possible reallocation of these resources
- Risk assessment before and after the sentence
- Recidivism records

# Benefits and Challenges of RISE AI

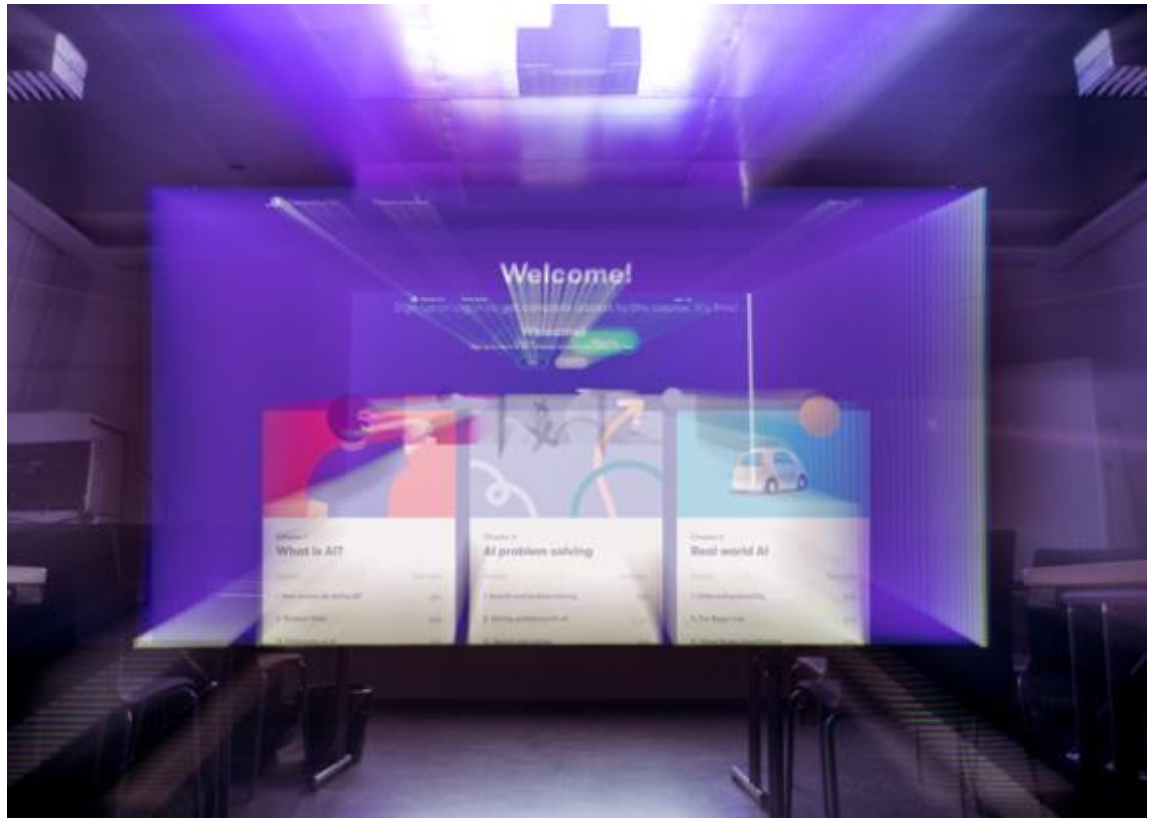
- **Benefits:** Offenders are directed to services and activities that better meet their risk levels and criminogenic needs, the compatibility of the offenders' needs and their rehabilitation and its impact on the risk of recidivism would thereby improve -> better use of sentence time and reduce risk of recidivism

- **Challenges:** The technical process starts to steer the process too much, to the extent that it excludes the offenders' and experts' analysis; RISE AI is affected by algorithmic discrimination and bias risk similar to other AI-systems: if the prejudices are 'built in' the data, the algorithm will repeat the typical fallacies of human reasoning…

# Aurora AI

- **National artificial intelligence (AI) programme**: a recommender system for public services to help citizens and companies find necessary and relevant services in a timely and ethically sustainable manner (Ministry of Finance)
- Criminal Sanctions Agency is part of the Aurora AI pilot (2022) and Aurora AI network
  - Public services providers whose services can be accessed from the joint network of smart services and applications
  - Target group: young offenders

# Elements of AI

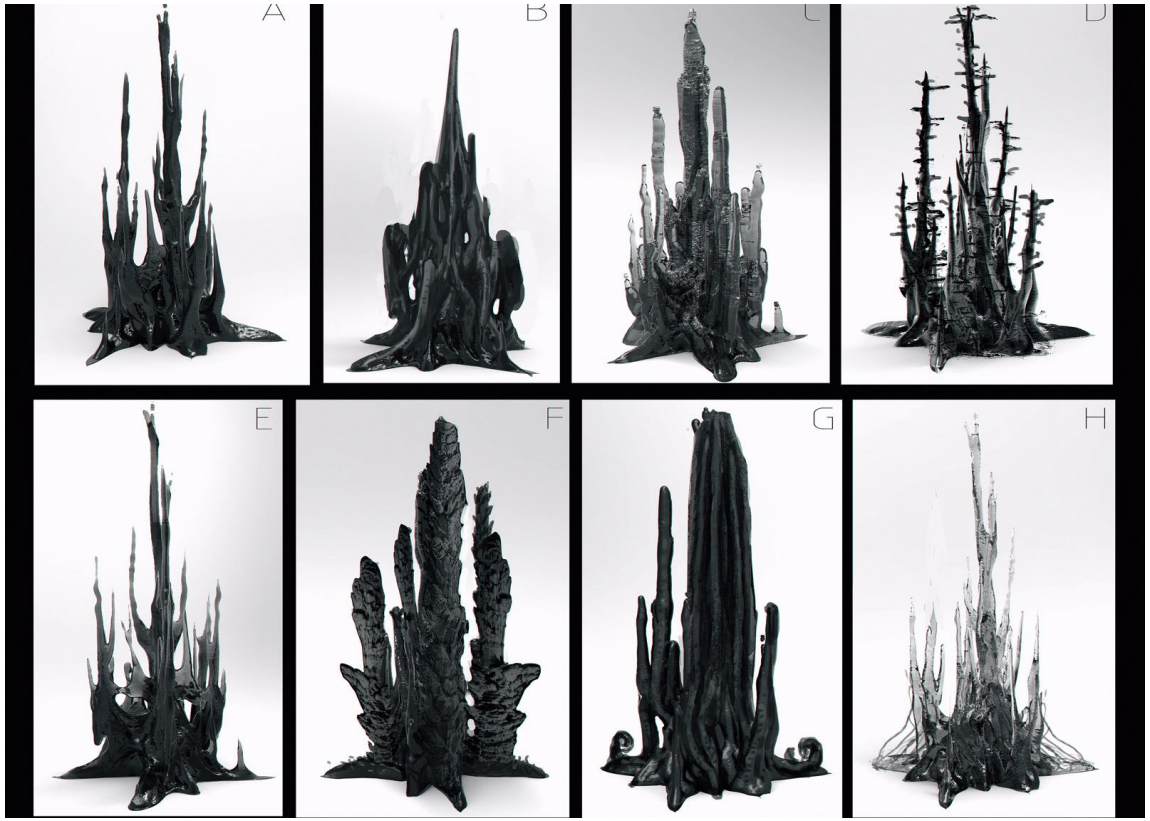



- Online course on the basics of AI by University of Helsinki and Reaktor Inc.
  - Available online for offenders in all prisons and probation offices in Finland
- *Elements of AI* has already trained more than 1% of the Finnish population on the basics of AI
  - Next challenge: 1 % of EU citizens, in 21 languages
  - 170 countries, 700 000 registered users

# My Vision of AI

- Combing advanced technology with heightened ethics and expert knowledge into algorithms
- Simplicity of algorithms: the fast way to optimal choices
- Clean data vs. biased distorted data -> Whose knowledege are we turning into algorithms? -> Involving experts by experience?
- Correlations have no ethics
  - Why there are correlations and what to do with the found correlations is up to humans to analyze

# To Know What To Do With It!



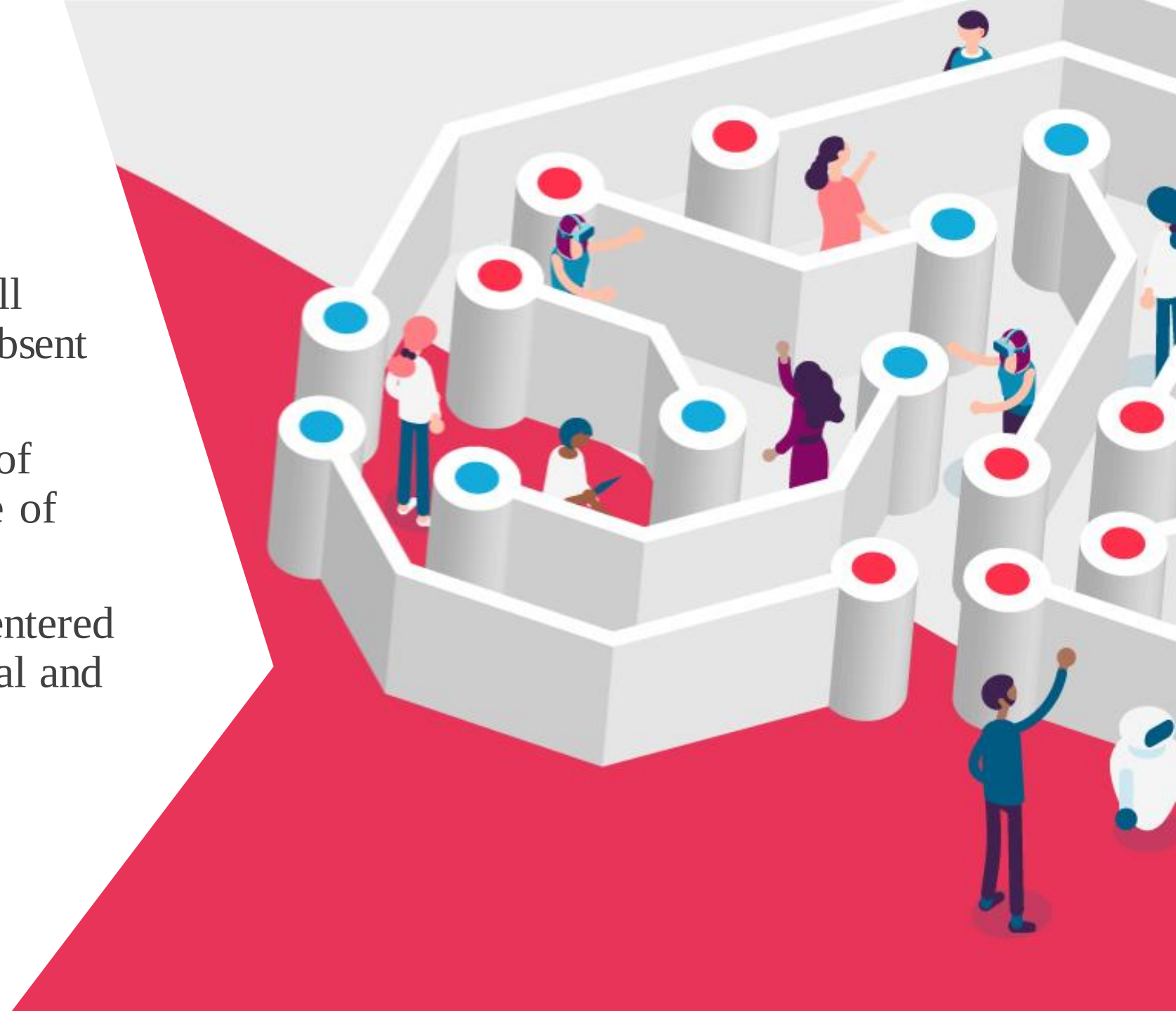

- Passing on our knowledge and wisdom to the future generations in the most delicate and simple way
- Not the hard and long way but the simplest possible way, the path leading to best / optimal decisions

# Discussion

- Use of AI in prisons is still rather limited or almost absent

  -> Are we behind the development of the rest of the society regarding use of AI?

  -> Ethical and human- centered corrections require ethical and human-centered AI!




Puolakka P., & Van De Steene, S. (2021). Artificial Intelligence in Prisons in 2030. An Exploration on the Future of Artificial Intelligence in Prisons. *Advancing Corrections Journal*, Edition # 11, ICPA.

https://icpa.org/17417-2/

# Thank you!

Pia Puolakka

pia.puolakka@om.fi

@Piapuolakkaomf1

/pia-puolakka-8218a9b3/

RISE

# References


- Andrew, N.G. (2017). *Artificial Intelligence is the New Electricity*. Presentation for The Stanford Graduate School of Business, 02/02/2017. https://www.youtube.com/watch?v=21EiKfQYZXc
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52.
- BBC News. (2011, November 25). *Robotic prison wardens to patrol South Korean prison.* BBC News. https://www.bbc.com/news/technology-15893772
- Cassens Weiss, D. (2019, October 25). Prisons and jails use artificial intelligence to monitor inmate phone calls. *Abajournal*. https://www.abajournal.com/news/article/prisons-and-jails-use-artificial-intelligence-to-monitor-inmate-phone-calls
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism, *Science Advances*, 4 (1), eaao5580. https://advances.sciencemag.org/content/4/1/eaao5580
- European Commission (2018). *Coordinated Plan on Artificial Intelligence, Communication from the commission to the European Parliament, The European Council, The Council, The European Economic and social committee and the Committee of the Regions.* EC, COM (2018) 795 final.
- European Union Agency for Fundamental Rights, FRA. (2020). *Getting the Future Right, Artificial Intelligence and Fundamental Rights Report.* FRA.

- Fry, H. (2018). *Hello World – Being human in the Age of Algorithms*. W.W. Norton & Company Ltd.
- Houser K. (2019, February 4). *China is Installing 'AI Guards' in Prison Cells. They'll make escape impossible – but the trade-off might be inmates' mental health*. Futurism. https://futurism.com/chinese-prison-ai-guards-cells
- Knight V., & Van De Steene S. (2020). The digital prison: Towards an ethics of technology. In Birch, P., & Sicard, L. A. (Eds.). *Prisons and Community Corrections. Critical issues and emerging controversies.* Routledge.
- Lin, Z., Jung, J., Goel, S. & Skeem, J. (2020). The Limits of Human Predictions on Recidivism, *Science Advances*, 6 (14), eaaz0652. https://advances.sciencemag.org/content/6/7/eaaz0652
- McGoogan, C. (2016, December 6). Liverpool prison is using AI to stop smuggling drugs and weapons. *The Telegraph.* https://www.telegraph.co.uk/technology/2016/12/06/liverpool-prison-using-ai-stop-drugs-weapons-smuggling/
- McGuire, J. (1999). *What Works: Reducing Reoffending: Guidelines from Research and Practice*. Wiley Series in Offender Rehabilitation. Wiley.
- Nellis, M. (2019, January 29). *"Better than Human"? Smartphones, Artificial Intelligence and Ultra-Punitive Electronic Monitoring*. Challenging Incarceration. https://www.challengingecarceration.org/wp-content/uploads/2019/01/TI-and-Smart-EM-Final-.pdf

- Puolakka P. (2020). *RISE AI: Reducing the Risk of Recidivism with AI*. Aalto Executive Education: Diploma in Artificial Intelligence. Unpublished.
- Puolakka, P., & Van De Steene, S. (2021). Artificial Intelligence in Prisons in 2030. An exploration on the future of Artificial Intelligence in Prisons. *Advancing Corrections Journal*, Edition # 11, ICPA.
- Sunstein, C.R. (2019, January 23). Algorithms, Correcting Biases. *Oxford Business Law Blog*. https://www.law.ox.ac.uk/business-law-blog/blog/2019/01/algorithms-correcting-biases
- Tucci, L., Burns, E. & Laskowski, N. (2020). Artificial Intelligence – definition. *Techtarget*. https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence
- Ward, T. (2002). The management of risk and the design of good lives. *Australian Psychologist*, 37, 172-179.
- Yan, S. (2019, April 1). Chinese high-security jail puts AI monitors in every cell 'to make prison breaks impossible'. *The Telegraph*. http://www.telegraph.co.uk/news/2019/04/01/chinese-prison-rolls-facial-recognition-sensors-track-inmates/