

Predictive Analytics and Algorithmic Technologies in Youth Justice: Innovation at what cost?

Professor Claire Paterson-Young

Professor of Social Justice, University of Northampton

Contents

Grounding digital technologies within the UNCRC

Key Concerns

Labelling and stigma 2

Desistance 2

Transparency and Accountability 3

The NAYJ calls for:

Transparency 3

Child-specific impact 3

Social inclusion and participation 4

In conclusion 4

Acknowledgments 5

References 5

The National Association for Youth Justice (NAYJ) welcomes public debate on the growing use of predictive and probabilistic technologies within criminal justice systems. The NAYJ's position is grounded in longstanding legal and ethical principles, which place the child's welfare, rights and developmental needs at the centre of all youth justice policy and practice.

We recognise the need for early intervention to support children at the fringes of the youth justice system and for improved efficiencies, but have concerns that when predictive technologies are applied to children and young people, they risk undermining these principles that the youth justice system in England and Wales has adopted. Any proposal that risks prioritising prediction, surveillance or control over children's best interests must therefore be subject to the most rigorous scrutiny.

These developments sit alongside wider moves to expand the use of AI and predictive technologies within policing. This makes it increasingly important to consider how children may be affected by algorithmic systems across policing and youth justice.

The NAYJ is a registered charity (no. 1138177) and membership organisation campaigning for the rights of – and justice for – children in trouble with the law.

info@thenayj.org.uk

www.thenayj.org.uk

September 2026

Grounding digital technologies within the UNCRC

The growing use of artificial intelligence and data-driven technologies in justice systems has also prompted increasing concern within international children's rights scholarship.

Emerging policy work highlights that the development and deployment of digital technologies affecting children must be explicitly grounded in the standards set out within the UN Convention on the Rights of the Child (UNCRC), including protections relating to privacy, non-discrimination and children's participation in decisions affecting them (Lievens & Verdoodt, 2018; UN Committee on the Rights of the Child, 2021). More recent international guidance, including UNICEF's 2025 Guidance on AI and Children, reinforces the need for child-centred AI governance grounded in children's rights, participation and protection from harm.

Initiatives such as the i-ACCESS MyRights project (Kiss et al., 2024) further emphasise that AI tools used within justice contexts must be assessed against children's rights frameworks to ensure that technological innovation does not undermine procedural safeguards or children's developmental needs.

This debate is increasingly relevant to policy and practice in England and Wales. The Ministry of Justice's 2025 AI Action Plan for Justice sets out an ambition to embed AI across the justice system, including the use of predictive and risk-assessment models to support decision-making. This makes it particularly important that the development and deployment of such technologies within youth justice is grounded in children's rights, transparency and appropriate human oversight (Ministry of Justice, 2025).

Key Concerns

Key concerns of The NAYJ in relation to the growth in the adoption of artificial intelligence include the following:

Labelling and stigma

A central claim within predictive systems, in criminal justice contexts, is that algorithmic tools can identify those "at risk" of future offending or further justice system involvement. However, criminological research has long warned of the harms associated with labelling young people as "deviant" or "high risk" (see Becker, 1963; Creaney, 2012). Labelling theory (Becker, 1963) demonstrates that designation as "delinquent" can contribute to further marginalisation and secondary deviance.

Empirical studies show that early justice system contact increases the likelihood of continued system involvement (McAra & McVie, 2007). When predictive systems formalise such labels through algorithmic scores, they risk embedding this dynamic at scale, and without clear ethical, legal and responsible use, this may have significant and lasting impact on children's lives. Examples of the ethics and legal issues were raised in 2019 by the West Midlands Police and Crime Commissioner ethics committee in relation to a predictive re-offending tool (Guardian, 2019).

Desistance

Predictive systems typically focus on risk and deficit and may be unable to capture the

unmeasured (and unmeasurable) factors that may contribute to patterns of desistance or non-offending by children, such as the presence of a supportive adult or engagement in activities that may have a positive impact on a child.

Predictive systems are limited in their ability to accommodate the notion that children are capable of change as they grow and develop. These limitations reflect a broader critique within the literature that predictive analytics struggles to account for the fluid, developmental and relational nature of children's lives, meaning that algorithmic forecasts may fail to capture the dynamic processes through which young people desist from offending (Berk, 2017; Livingston & Risse, 2020).

This is particularly important in a youth justice system that has seen substantial reductions in offending and first-time entrants over recent decades, reinforcing the importance of approaches that support diversion and desistance rather than unnecessarily increasing children's contact with justice agencies (Ministry of Justice, 2026).

Transparency and Accountability

Risk assessment tools frequently operate as "black boxes," where the internal logic of decision-making is opaque to those affected (Pasquale, 2015; Burrell, 2016). Without full transparency, meaningful legal challenge becomes difficult, which undermines Article 6 ECHR rights to a fair hearing and Article 40 UNCRC protections regarding children in conflict with the law. Algorithmic outputs should therefore support, rather than replace, professional judgement. Decision-makers should remain responsible for interpreting and challenging algorithmic recommendations, rather than treating them as determinations of a child's future behaviour or risk.

The use of children (and their family's) personal data without meaningful information about how their data are collected, used and shared is also of concern (Kedell, 2023), as is data security and the risk of data breaches. Academic research further demonstrates that decision-makers may place undue trust in algorithmic outputs, a phenomenon described as "automation bias" (Alon-Barkat and Busuioc, 2023; Khera et al., 2023). The extensive collection and processing of children's personal data within such systems also raises significant privacy concerns.

The UN Committee on the Rights of the Child has emphasised that children's rights to privacy and data protection must be safeguarded within digital environments, particularly where automated systems may generate profiles that shape institutional responses to children's behaviour (UNCRC, 2021).

The NAYJ calls for...

Transparency regarding the use of predictive or algorithmic tools in the criminal justice system (including tools identifying children and young people).

Any predictive or algorithmic tool used must be subject to meaningful transparency about its design, purpose and operation, including disclosure of the data sources where appropriate and lawful.

Transparency must also include publication of independent, peer-reviewed validation studies demonstrating accuracy, error rates, and differential impacts across protected characteristics, using representative youth justice populations in England and Wales.

Ongoing monitoring, equality impact assessments and audit findings should be routinely published where appropriate, to ensure democratic oversight, legal accountability and public confidence. Where public bodies deploy algorithmic tools, information should also be made available through the Algorithmic Transparency Recording Standard or any successor framework, with sufficient detail to enable meaningful public and independent scrutiny.

Child-specific impact assessments must be undertaken on any predictive or algorithmic tool and should include a detailed equality impact assessment examining potential disproportionate effects across race, ethnicity, gender, disability, socio-economic status and care experience, alongside a structured human rights analysis assessing compatibility with the European Convention on Human Rights and the UN Convention on the Rights of the Child.

The assessment must explicitly consider children's developmental status, their heightened vulnerability within justice processes, and the long-term consequences of labelling or risk classification. It should consider how algorithmic outputs may influence professional decision-making, including the risk of automation bias and over-reliance on predicted risk. Such assessments should also draw on emerging approaches to Children's Rights Impact Assessment (CRIA), which provide structured methodologies for evaluating how policy or technological interventions may affect children's rights, welfare and developmental interests (Lundy, Byrne & McEvoy, 2021).

Social inclusion and participation of children and young people in governance discussions regarding digital technologies affecting their lives. Children must be meaningfully involved in the design, oversight, and evaluation of tools used in youth justice. This should include opportunities to influence decisions about implementation, data use, safeguards and how concerns or harms are identified and addressed.

Mechanisms for participation should be accessible, age-appropriate, and ongoing, enabling youth voices to shape policy, maintain accountability, and protect their rights throughout the life of any technology. Mechanisms for participation should be accessible, age-appropriate and ongoing, enabling children and young people to contribute to policy development and to inform decisions about technologies that may affect their rights and experiences. Evidence from participatory research with children and youth suggests that structured approaches such as Youth Advisory Boards can provide meaningful opportunities for them to influence programme direction, contribute to monitoring and evaluation, and shape knowledge and policy. However, meaningful participation requires more than creating a formal advisory structure, relying on clear roles, supportive relationships, opportunities for young people's views to influence decisions, and attention to practical barriers such as time, language and access to technology (Paterson-Young et al., 2024).

In conclusion

The aspiration to prevent harm and reduce risk is supported. However, greater reliance on predictive technologies risks undermining decades of progress toward a child-centred, welfare-based approach. Without robust safeguards, predictive technologies

may further embed youth justice practice into systems of surveillance and risk management rather than supporting children's development and desistance. Ensuring that emerging technologies remain firmly grounded in children's rights, transparency and democratic oversight is therefore essential if technological innovation is to support, rather than undermine, justice for children.

Acknowledgements

Thank you to Dr. Joe Allen-Fellows, Dr. Samantha Burns and Ross Little for comments on drafts of this briefing.

References

- Alon-Barkat, S. and 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), pp.153-169.
- Angwin, J., Larson, J., Mattu, S. and Kirchner, L. 2022. Machine bias. In *Ethics of data and analytics* (pp. 254-264). Auerbach Publications.
- Babuta, A. 2017. *Big Data and Policing: An Assessment of Law Enforcement Requirements, Expectations and Priorities*. Royal United Services Institute for Defence and Security Studies (RUSI).
- Babuta, A. and Oswald, M. 2021. ‘Machine learning predictive algorithms and the policing of future crimes’, in J.L.M. McDaniel and K.G. Pease (eds) *Predictive policing and artificial intelligence*. Abingdon, Oxford: Routledge.
- Becker, H. S. 1963. *Outsiders: Studies in the sociology of deviance*. New York: Free Press.
- Berk, R. 2017. *An impact assessment of machine learning risk forecasts on parole board decisions and recidivism*. *Journal of Experimental Criminology*, 13(2), 193–216.
- Burrell, J. 2016. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big data & society*, 3(1), p.2053951715622512.
- Creaney, S. 2012. Targeting, labelling and stigma: challenging the criminalisation of children and young people. Centre for Crime and Justice Studies
- Day, A-M, Allely, C., Robinson, L., Turner, K., Gerry, K.C & Forrester, A. 2024. The over-representation of neurodivergent children in Youth Justice Systems and The Youth Court. *Medicine, Science and the Law*. 64(4), 255-258
- Eubanks, V. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. Macmillan+ ORM.
- Fourcade, M. and Healy, K. 2017. Classification situations: Life-chances in the neoliberal era. *Accounting, Organizations and Society*, 63, 8–21.
- Keddell, E. 2023. The devil in the detail: Algorithmic risk prediction tools and their implications for ethics, justice and decision making. In: Taylor B. J., Keddell E., Graham J. C., et al. (eds) *The Sage Handbook of Decision Making, Assessment*

- and Risk in Social Work. Sage Publications Ltd.
- Khera, R., Simon, M.A. and Ross, J.S. 2023. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. *Jama*, 330(23), pp.2255-2257.
- Kiss, D., Foussard, C., Gil González, M., Lynch, N. & Perez-Cruz, M. 2024. Technology & Artificial Intelligence (AI) in Child Justice Systems: The Implications for Children's Human Rights. Terre des hommes Lausanne Foundation, Lausanne. <https://cora.ucc.ie/items/11298bca-c543-4fab-b8d5-25093307494f>. Accessed 8 September 2026.
- Lammy, D. 2017. *The Lammy Review: An independent review into the treatment of, and outcomes for, Black, Asian and Minority Ethnic individuals in the Criminal Justice System*.
- Lievens, E. and Verdoodt, V. 2018. Looking for needles in a haystack: Key issues affecting children's rights in the digital environment. *Computer Law & Security Review*, 34(2), 269–278.
- Lum, K. and Isaac, W., 2016. To predict and serve? *Significance*, 13(5), pp.14-19.
- Lundy, L., Byrne, B. and McEvoy, L. (2021). *Children's Rights Impact Assessment: A Guide for Policymakers*. Belfast: Queen's University Belfast
- McAra, L. and McVie, S., 2007. Youth justice? The impact of system contact on patterns of desistance from offending. *European journal of criminology*, 4(3), pp.315-345.
- Ministry of Justice (2025) AI action plan for justice. Ministry of Justice. Available at <https://www.gov.uk/government/publications/ai-action-plan-for-justice/ai-action-plan-for-justice>. Accessed 7 September 2026.
- Ministry of Justice (2026) A modern youth justice system: Foundation fit for the future. Ministry of Justice. Available at <https://www.gov.uk/government/publications/a-modern-youth-justice-service-foundations-fit-for-the-future>. Accessed 7 September 2026.
- Pasquale, F., 2015. *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Paterson-Young, C, Adhikari, J, Lee, L, Maher, M & Wright, LHV 2024, 'Creating ownership: Strengths and tensions in co-production with children, young people, and adults across contexts', *Childhood*, vol. 31, no. 3, pp. 329-348. <https://doi.org/10.1177/09075682241266814>
- Richardson, R., Schultz, J.M. and Crawford, K., 2019. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Rev. Online*, 94, p.15.
- Santos, R.B. 2017. *Crime analysis with crime mapping*. Fourth edition, Los Angeles: SAGE Publications Inc.
- Staines, J. 2016. *Risk, adverse influence and criminalisation understanding the over-representation of looked after children in the youth justice system*. London: Prison Reform Trust

The Guardian. 2019. Ethics committee raises alarm over 'predictive policing' tool. Available online at <https://www.theguardian.com/uk-news/2019/apr/20/predictive-policing-tool-could-entrench-bias-ethics-committee-warns>

UN Committee on the Rights of the Child. 2021. *General Comment No. 25 on children's rights in relation to the digital environment*.

UNICEF. 2025. Guidance on AI and Children. UNICEF. Available at <https://www.unicef.org/innocenti/media/11991/file/UNICEF-Innocenti-Guidance-on-AI-and-Children-3-2025.pdf>. Accessed 7 September 2026.