

## **DECONSTRUCTING CONTROVERSIAL PREDICTIVE TECHNOLOGIES FOR CHILDREN IN LAW ENFORCEMENT TO IDENTIFY, UNDERSTAND, AND ADDRESS ETHICAL ISSUES**

**Pinelopi Troullinou, Francesca Trevisan, Elodie Makhoul, Xenia Ziouvelou, Lilian Mitrou, Paola Fratantoni, Dimitris Kyriazanos, Sofia Segkouli**

Trilateral Research (Ireland), Trilateral Research (Ireland), National Centre of Scientific Research “Demokritos” (Greece), University of the Aegean (Greece), Zanasi and Partners (Italy), National Centre of Scientific Research “Demokritos” (Greece), Centre for Research and Technology-Hellas (Greece)

Pinelopi.Troullinou@trilateralresearch.com; Francesca.Trevisan@trilateralresearch.com;  
Elodie.Makhoul@trilateralresearch.com; xeniaziouvelou@iit.demokritos.com;  
L.mitrou@aegean.gr; paola.fratantoni@zanasi-alessandro.eu; dkyri@iit.demokritos.gr;  
sofia@iti.gr

### **EXTENDED ABSTRACT**

There is an increasing employment of AI technologies in the civil security sector in the promise of improving efficiency mainly with regard to resource allocation and automatic data analysis. However, the widespread and intrusive uses of AI introduce new challenges, posing threats to fundamental rights and democratic principles (European Parliament, 2021). AI systems may escalate surveillance practices, amplify discriminatory practices and exacerbate pre-existing societal inequalities (e.g., O'neil 2017, Zuboff, 2019). Vulnerable populations, particularly children<sup>14</sup>, require special attention in this context (Charisi, 2022; Rahman & Keseru, 2021). To raise awareness on how AI can uphold or undermine children lives and rights, in 2021, UNICEF released a policy guidance on AI for children pinpointing how predictive analytics on children can limit their identities and experience of the world. As more decisions regarding children are being taken with the aid of predictive systems (Hall et al. 2023), it becomes important to understand how these technologies are developed, used, and how they might impact children's rights and lives.

This paper aims at identifying and addressing the ethical and societal impact of predictive technologies designed to identify youth at risk of committing crime. More specifically, the paper discusses how the use of these technologies by law enforcement can result in portraying children as security concerns and the potential negative consequences that may arise from such characterization. We shed light on the risks involved in such practices by analysing the Prokid (Wientjes et al., 2017) case and the controversies surrounding it. Prokid is an identification tool designed for the early detection of young individuals at risk of (re)offending, originally developed by the Gelderland-Midden police force in The Netherlands. Prokid started

---

<sup>14</sup> According to the United Nations Convention on the Rights of the Child (UNCRC) (1989), children are referred to as those below the age of 18, in addition adolescent and youth refers to those aged 10-19 (WHO, 2014) and 15-24 (United Nations Department of Economic and Social Affairs (UNDESA). For the purposes of this study, we adopt an inclusive definition of 'children' that encompasses all three definitions, considering those 24 years or younger.

to be introduced in 2009 in four pilot regions: Gelderland-Midden, Amsterdam-Amstelland, Brabant Zuidoost and Hollands Midden (Abraham et al., 2011). Over the course of the years, Prokid has gone through several iterations. The initial version, Prokid 12, was designed to assess the risk of criminality of children under 12 years old. A subsequent version of the system has shifted its focus to the age group of 12 to 18 years (Wientjes et al, 2017). The latest version, which is expected to differ substantially from the others, is still under development and will include individuals up to 23 years (Tweede Kamer der Staten-Generaal, 2022). Prokid relies upon existing police data such as reports of children who have come into contact with the police as suspect, victims or witnesses, their addresses, age, gender, the number and types of crimes committed, and additional information about their family and peers. Children data are sorted in a semi-automated way into four risk categories where “red” indicates critical danger, “orange” indicates a problematic situation in regard to the child or their address, “yellow” indicates that a potential risk is developing, and “white” indicates no risk. It was agreed that the police would take follow up actions with children categorized within the red, orange, and yellow categories.

In the analysis of Prokid, we use social controversy mapping as socio-technical tool to unpack the “black box”, understand the functioning of the technology, and evaluate its ethical and societal impact. Mapping and analyzing social controversies is a methodology that draws on the traditions of Science and Technology Studies and Surveillance Studies (Trevisan et al., forthcoming). It consists of deconstructing social controversies as reported in public discourses to identify and map the stakeholders involved in the technology lifecycle and gain a more nuanced understanding of the diverse perspectives, experiences, needs, values, interests, risks and expectations surrounding the technology development. This structured analysis is key to account for the larger social context and needs, uncover common grounds and areas of contentions and ultimately favour human centered approaches to tech development.

We also evaluate compliance with ethical principles on AI for children to inform policy, advocacy, and ethics scrutiny on these practices. By so doing, we flag the diverse factors that need to be considered in order to build systems that are ethical and socially sustainable promoting children’s safety and security minimising potential harms. Furthermore, we specifically evaluate the impact on children’s rights as the potential to interfere with human dignity, right to personality, privacy, and their ability to make decisions about their own lives (right to self-determination).

With this work, we want to emphasize the importance of conducting ethical, societal and fundamental rights impact assessments employing the social controversies deconstruction method to guide technology development and governance models towards promoting the well-being of children and upholding the no-harm principle. Our work makes two unique contributions. Firstly, it offers an evidence-based framework designed to unpack the black box of controversial technologies, support explainability and accountability and understand the dynamics of the diverse discourses and interests. This approach enables a comprehensive analysis of the technology's impact. Secondly, our approach recognizes that technology does not exist in a vacuum, but it interacts with, shapes, and is shaped by society. Therefore, it delves into the broader social understanding and ethical implications of the technology under examination.

**KEYWORDS:** Children, societal impact assessment, ethics, controversies, responsible AI.

## REFERENCES

- Assembly, United Nations General. (1989). Convention on the Rights of the Child. *United Nations, Treaty Series*, 1577(3), 1-23: Retrieved from: [http://wunrn.org/reference/pdf/Convention\\_Rights\\_Child.PDF](http://wunrn.org/reference/pdf/Convention_Rights_Child.PDF).
- Abraham M., Buysse W. Loef L., Bram van Dijk B. (2011). Pilots ProKid Signaleringsinstrument 12- geëvalueerd. WODC. <http://hdl.handle.net/20.500.12832/1832>
- Charisi, V., Chaudron, S., Di Gioia, R., Vuorikari, R., Escobar Planas, M., Sanchez Martin, J.I. and Gomez Gutierrez, E. (2022). Artificial Intelligence and the Rights of the Child: Towards an Integrated Agenda for Research and Policy, EUR 31048 EN, Publications Office of the European Union, Luxembourg, 2022, <http://hdl.handle.net/20.500.12832/1832>
- European Parliament, (2021). Artificial Intelligence in Criminal Law and its use by the Police and Judicial Authorities in Criminal Matters. Cain, K. (2012, June 29). European Parliament Resolution of 6 October 2021. Retrieved from: [https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html)
- O'neil, C. (2017). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.
- Hall, S. F., Sage, M., Scott, C. F., & Joseph, K. (2023). A Systematic Review of Sophisticated Predictive and Prescriptive Analytics in Child Welfare: Accuracy, Equity, and Bias. *Child and Adolescent Social Work Journal*, 1-17.
- Rahman, Z., and Keseru, J. (2021). *Predictive Analytics for Children: An Assessment of Ethical Considerations, Risks, and Benefits*. UNICEF Office of Research-Innocenti.
- Tweede Kamer der Staten-Generaal (2022). Aanhangsel van de Handelingen. Vragen gesteld door de leden der Kamer, met de daarop door de regering gegeven antwoorden 1177. <https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2022Z22088&did=2022D52317>
- Trevisan F., Troullinou P., Fisher E., Kyriazanos D., Bertelli V. (2023) Deconstructing Social Controversies for a Trusted AI Future. Under Review.
- UNDESA- United Nations Department of Economic and Social Affairs, "Definition of Youth", New York, Retrieved from: <http://un.org/esa/socdev/documents/youth/fact-sheets/youth-definition.pdf>
- UNICEF, (2021). Policy guidance on AI for Children. Retrieved from: <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>
- WHO (2014), 'Recognizing Adolescents', World Health Organization, Geneva, 2014.
- Wientjes, J., Delsing, M., Cillessen, A., Janssens, J., & Scholte, R. (2017). Identifying potential offenders on the basis of police records: development and validation of the ProKid risk assessment tool. *Journal of Criminological Research, Policy and Practice*, 3(4), 249-260.
- Zuboff, S., (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Public Affairs.