

HOW ARTIFICIAL INTELLIGENCE AND DATA-DRIVEN SYSTEMS CAN IMPROVE CHILD PROTECTION SERVICES

Mihail CIOBANU

Researcher, PhD Student

National Institute for Economic Research, AESM, MOLDOVA

E-mail: ciobanu.mihail@ase.md

ORCID: [0000-0003-1193-6018](https://orcid.org/0000-0003-1193-6018)

Abstract: AI and data-driven systems are transforming child protection as they enable early detection of risk factors and more effective intervention. AI-driven tools survey online conversations, behavior patterns, and real-time whereabouts to enhance child safety, while cybersecurity and blockchain technology are increasingly used to protect the identity and personal data of children. Virtual reality supports safety education by simulating dangerous scenarios in controlled environments, and big data analytics help predict and prevent abuse or neglect through early warning systems. These innovations are particularly promising in contexts with limited human resources, where automated systems can augment decision-making and optimize resource allocation. However, challenges such as data privacy, algorithmic bias, inconsistent data quality, and lack of transparency remain significant barriers to full-scale implementation. This paper presents a review of academic literature, institutional reports, and real-world case studies, including tools like the Allegheny Family Screening Tool and mobile health platforms, to evaluate the effectiveness and limitations of current AI applications in child welfare. It also explores the ethical implications of predictive modeling and decision support systems, especially their impact on marginalized communities. Based on policy analysis and best practices, this research highlights key recommendations to improve the design, governance, and accountability of AI in child protection services. The work was developed within the framework of Subprogram 030101 „Strengthening the resilience, competitiveness, and sustainability of the economy of the Republic of Moldova in the context of the accession process to the European Union”, institutional funding.

Keywords: Child, protection, data, tool, system, service.

UDC: 004.8:364.4-053.2(478)

Classification JEL: A13, C55, K36, M15, O33

1. Introduction

In an increasingly digital world, artificial intelligence and data-driven systems are emerging as transformative tools in the realm of child protection. These technologies offer unprecedented opportunities to enhance the efficiency, accuracy, and responsiveness of services aimed at safeguarding vulnerable children. From early identification of at-risk individuals to the optimization of social services and resource allocation, AI has the potential to revolutionize how governments, NGOs, and communities protect children from harm. However, alongside the promise of innovation comes the responsibility to ensure that these systems are ethical, transparent, and centered on the best interests of the child. This paper explores the ways in which AI and data-driven solutions can contribute to more effective child protection services, while also addressing the critical challenges of privacy, bias, and equitable access.

2. Literature Review

Artificial intelligence and data-driven systems have the potential to significantly enhance child protection services by improving risk assessment, streamlining decision-making processes, and facilitating the early identification of abuse and neglect cases. AI models, including artificial neural networks and natural language processing techniques,

have been developed to predict instances of child abuse and neglect. However, a systematic review highlighted that these implementations are still in nascent stages and often suffer from high risks of bias due to factors such as small sample sizes and overfitting [1]. The integration of machine learning-based risk models in child protection raises several ethical concerns, including data biases, lack of standardized documentation, and challenges in model evaluation. Addressing these issues is crucial to ensure the responsible deployment of AI in this sensitive field [2]. AI and mobile health (mHealth) approaches offer substantial potential in preventing and addressing violence against children on a large scale. These technologies can be particularly effective in low- and middle-income countries, provided they are implemented with caution to avoid exacerbating existing inequalities [3]. While AI and data-driven systems hold promise for improving child protection services, their current application is limited and fraught with challenges. Future research needs to focus on developing robust, unbiased models and addressing ethical considerations to fully harness the potential of AI in safeguarding children's welfare. In the scientific literature intelligence can be seen as "the ability to learn from experience (acquire and retain new knowledge), and to subsequently apply that new knowledge with flexibility to manipulate or adapt to a changing environment" or it can be seen as "the ability to create abstract thought, beyond instinct or responses to sensory input" [4]. In contrast, artificial intelligence is "the capacity of a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments" [5]. A data-driven system is a "collection of analytical systems that stores data from various processes and converts it into meaningful information needed for decision-making (often using third-party data sources)" [6].

3. Methodology

This research adopts a qualitative, exploratory approach to assess how artificial intelligence (AI) and data-driven systems enhance child protection services. The study is grounded in an integrative literature review of academic publications. Key sources include peer-reviewed journals, systematic reviews, and case studies from child welfare agencies and international organizations. Case examples, such as the Allegheny Family Screening Tool and predictive models developed by institutions like the University of Georgia and MIT, were analyzed to identify practical implementations, ethical concerns, and effectiveness outcomes. Additional insights were drawn from expert interviews and stakeholder commentaries featured in public reports and digital platforms. The analysis focused on four dimensions: technological capabilities of AI in risk prediction and intervention; ethical and human rights considerations; practical barriers to adoption; and policy and governance implications.

4. Results and Discussion

AI tools oriented towards children can be of various types, like educational apps (Duolingo, DreamBox), virtual assistants (Alexa), AI-powered games (Minecraft, Roblox), smart toys (Cozmo, Hello Barbie). With regard to social protection of children AI is frequently used in predictive risk assessment tools. A 4-year study by Daniel Gibbs (University of Georgia, US) examined AI algorithms that assign risk scores (1–20) to predict foster care placement likelihood within two years [7]. These tools aimed to reduce false negatives (missed risks) and false positives (unnecessary interventions) in child welfare investigations. However, their effectiveness remains debated due to concerns about data quality and algorithmic bias. AI tools don't make final decisions, only provide a risk score (1–20) based on public data. Data sources were public assistance, criminal records,

medical and mental health records, welfare, education records, etc. Social workers used the score as an additional piece of information to support their judgment. It needs to be added, that AI tool supplements, not replaces, human judgment in screening child maltreatment reports. D.Gibbs argued that “either we miss the needle in the haystack and a child or family experiences some sort of harm, or we grab too much hay and families get this intrusive intervention.” [8].

One of the problems of using AI in social protection of children is the insufficient, inconsistent data or record duplication may make AI tools unsuitable. Research in Canada identified record duplication as a major obstacle to AI implementation in child protection. Centralized case management systems require high-quality administrative data to train AI models, but inconsistent data practices undermine reliability. AI models for child abuse/neglect often use heterogeneous datasets and have high bias risks, emphasizing the need for larger, validated datasets to improve accuracy. The implementation of AI in the child abuse and neglect field lags compared to other medical fields. Tools like Save the Children’s “Ask Save the Children” use generative AI to provide real-time child protection guidance in emergencies, though their effectiveness depends on rigorous training data.

In using AI tools there are a number of ethical and human rights considerations. A chapter in the Handbook on Public Policy and Artificial Intelligence critiques AI’s role in child protection, arguing that predictive models risk exacerbating discrimination against marginalized groups [9]. It stresses balancing risk analysis with human rights, particularly in decisions involving family integrity and child safety.

Also, there are a number of benefits of AI. It could help reduce harmful errors, it may help identify high-risk cases more reliably, it can provide structure and objectivity to complex decisions. AI tools also have limitations. They can reflect only those data known to public systems and may leave out crucial context. Over-policing and poverty can lead to racial or socioeconomic discrimination. Workers sometimes distrust the AI risk prediction tools, especially when scores are ambiguous (for example, when score = 10). Some workers don’t understand how the algorithm works, reducing its effectiveness and adoption. Transparency is lacking - parents often don’t know that an algorithm was used in their case. It’s not about math - it’s about how the tool is used. Tools can be very useful at score extremes (for example, 1/20 or 20/20), but not helpful in middle-range cases. AI use is likely inevitable in child welfare due to limitations of human decision-making. Tools must be context-sensitive, explainable, and trusted to be adopted. AI can amplify both good and bad decisions - its impact depends on human use.

MIT Researchers [10] partnered with Colorado’s child welfare department to develop visual analytics tools that explain AI predictions for case screeners. Their work emphasized transparency in machine learning models, focusing on how factors like child age influence risk assessments. Screeners value factor-specific insights over technical explanations of the model. Simple models can still confuse users if features aren’t clearly explained in plain language. Screeners were more interested in validating or reconciling model predictions with their intuition. Four interfaces were built, including one comparing current cases to past similar ones. Case-specific details interface were found useful by 90%+ of screeners, increasing trust in predictions. Case comparison interface was not trusted - screeners worried it might lead to biased decisions based on past cases, rather than current evidence. Researchers modelled, designed and supported implementation of this world-first child welfare predictive analytics tool. The Allegheny Family Screening Tool (AFST) uses rich administrative data to generate a screening score for incoming calls alleging child maltreatment and neglect. The score is an additional piece of information that helps call screeners as they decide whether to open an investigation. Allegheny County

introduced this decision support tool with the aim of improving accuracy and consistency of call screening decisions. The evaluators' findings included that use of the tool improved the accurate identification of children in need of services and was associated with a modest reduction in racial disparities in case openings.

European Commission's Joint Research Centre (JRC) published a report analyzing AI's impact on children's rights, including risks from conversational agents and recommender systems. It underscores the need for transparency, accountability, and non-discrimination in AI design. It advocates for multistakeholder collaboration and ethical frameworks to balance innovation with child rights. The report data present a comparative overview of the types of questions raised by different stakeholder groups - policymakers, experts, and children/youth - on issues related to AI, children's rights, and digital technology [11, p.61]. Each group demonstrated distinct priorities and concerns, offering insights into how different perspectives shape the discourse on AI and child well-being. The input from policymakers (N=28), shows that highest percentage of questions were related to inclusion (25%) and education (25%), reflecting a strong emphasis on ensuring equitable access and learning opportunities. This was followed by cognition and development (21.4%), suggesting interest in how AI impacts children's growth and mental processes. Other concerns included exploitation (14.3%) and privacy (10.7%), with explainability (3.6%) being the least addressed topic, indicating limited focus on how understandable AI systems are. Findings from expert discussions (N=22) show that while education (22.9%) again ranked high, experts placed greater emphasis on privacy (17.1%) and evaluation/monitoring (17.1%), showing concern for ethical oversight and system accountability. They also raised issues related to cognition/development (11.4%), trust (8.6%), conflicting rights (8.6%), and employment (2.9%), suggesting a broader and more technical or ethical consideration of AI's impact. The topics raised by children and youth (N=35) demonstrate that their dominant concerns were cognition and development (27.3%) and inclusion (22.7%), followed by evaluation (13.6%) and explainability (13.6%). This group placed significantly more importance on understanding how AI systems work compared to policymakers and experts. Meanwhile, education (9.1%), conflicting rights (9.1%), and privacy (4.5%) were less frequently mentioned, possibly reflecting differences in awareness or perceived relevance. When comparing across the three groups we identify that inclusion was a top concern for both policymakers and children/youth, but not explicitly highlighted by experts; education was consistently important for policymakers and experts, though less so for children; cognition and development was highly valued by children and policymakers, with moderate attention from experts; explainability was scarcely addressed by policymakers and not by experts, but was relatively important to children, underscoring their desire to understand AI; privacy and evaluation were notably more emphasized by experts than by children or policymakers. These differences reveal how each group's perspective - policy-driven, technical/ethical, or experiential - shapes the framing of issues around AI and children's rights.

In order for children to safely engage with AI they need age-appropriate introduction, clear limits on their smart tech usage and prioritization of real-world social interactions, monitoring of app content and privacy settings, being taught about data privacy, encouragement of critical thinking skills, involvement of adults in their AI interactions, discussion about the (un)emotional side of AI, encouragement of a variety of learning and play. At the same time AI potential can benefit kids through building creativity, skill development, accessible learning to children with special needs, personalized learning, while there are risks related to privacy, over-reliance on technology,

exposure to inappropriate content, impact on emotional development, screen time and health concerns, bias in AI systems.

Current AI policy implementations in child protection lag behind other fields, with studies often plagued by small sample sizes and methodological flaws. The UNICEF and European Commission reports advocate for multistakeholder collaboration, children's participation, and policy frameworks that prioritize child rights over managerial efficiency. In this regard it is recommended to improve data quality and reduce duplication to support AI training, to ensure transparency, explainability, and accountability in AI systems, to develop guidelines that address conflicting rights (for example, privacy vs. protection) and to involve children in AI design.

The Alan Turing Institute from UK on 4th February 2025 organized Children's AI Summit: Putting children at the centre. The Manifesto of the AI Summit has set a series of demands for the world leaders formulated from the perspective of children. It highlights key areas of concern and outlines specific actions that children believe should be taken to ensure that AI is developed and used in ways that are ethical, inclusive, and safe. The demands emphasize the importance of listening to children's voices and taking their experiences into account when creating policies related to AI. A strong focus is placed on child safety, particularly in relation to social media and the broader digital environment. Another major point is the need for transparency in AI development. This includes tracking and removing biased or discriminatory training data, ensuring companies are open about how their AI systems are built and used, and requiring that AI development follows clear ethical standards through appropriate legislation. Environmental concerns are also addressed, with a call for AI technologies to rely on clean, sustainable energy sources. Additionally, the text underlines the necessity of improving public education about AI so that people, especially children and their caregivers, can understand and use it safely and effectively. Finally, the demands include ensuring cybersecurity protections for AI systems and promoting equitable access so that all children, regardless of background, can benefit from the opportunities AI may offer. Overall, the text reflects a thoughtful and multi-dimensional perspective on the responsible use of AI from the viewpoint of younger generations.

5. Conclusions

Artificial intelligence and data-driven systems present powerful opportunities to improve child protection services by enhancing risk detection, supporting decision-making, and optimizing resource allocation. These technologies can strengthen the ability of social workers and institutions to identify and respond to cases of abuse, neglect, and exploitation more quickly and accurately, especially in resource-constrained environments. Tools such as the Allegheny Family Screening Tool and AI-powered mHealth platforms demonstrate how predictive analytics and real-time data can contribute meaningfully to child welfare.

However, while the potential is significant, the current implementation of AI in child protection remains limited and fraught with challenges. These include inconsistent data quality, lack of transparency in algorithmic decision-making, biases that disproportionately affect marginalized communities, and limited user trust in AI tools. Furthermore, ethical concerns about the balance between efficiency and child rights, particularly regarding privacy, family integrity, and non-discrimination, remain unresolved.

Comparative analysis of stakeholder concerns - policymakers, experts, and children - reveals differing priorities that must be integrated into the development of AI tools. Policymakers focus on inclusion and education; experts prioritize privacy, monitoring, and ethical governance; children emphasize understanding, mental development, and inclusion.

This diversity highlights the need for multi-stakeholder involvement, including children's participation, in designing child-centered AI systems.

Moreover, the demands of children voiced during the AI Children Summit call attention to crucial policy gaps and offer a normative framework for action. These include the need for ethical legislation, transparent data practices, guidance for safe AI use, and education about AI's role and risks. The emphasis on environmental sustainability, cybersecurity, and equitable access further enriches the discourse on AI and children's rights.

In conclusion, realizing the full benefits of AI in child protection requires high-quality, interoperable data systems; transparent, explainable, and accountable AI models; ethical guidelines that prioritize children's rights over managerial efficiency; active participation of children and caregivers in AI governance; collaborative policymaking that bridges technical innovation with social justice. Only through such a comprehensive, rights-based, and inclusive approach can AI become a trustworthy partner in protecting and empowering children in the digital age.

Acknowledgements: *The work was developed within the framework of Subprogram 030101 „Strengthening the resilience, competitiveness, and sustainability of the economy of the Republic of Moldova in the context of the accession process to the European Union”, institutional funding.*

7. References

1. LUPARIELLO, F., SUSSETTO, L., DI TRANI, S., DI VELLA, G. *Artificial Intelligence and Child Abuse and Neglect: A Systematic Review [online]*. Children, 2023, 10(10), 1659. ISSN 2227-9067. [viewed 20 March 2025]. Available from: <<https://doi.org/10.3390/children10101659>>
2. LANDAU, A.Y., FERRARELLO, S., BLANCHARD, A., CATO, K., ATKINS, N., SALAZAR, S., PATTON, D.U., TOPAZ, M. *Developing machine learning-based models to help identify child abuse and neglect: key ethical challenges and recommended solutions [online]*. Journal of the American Medical Informatics Association, 2022, Volume 29, Issue 3, March 2022, pp. 576–580. ISSN 1067-5027 (print), ISSN 1527-974X (web). [viewed 20 March 2025]. Available from: <<https://doi.org/10.1093/jamia/ocab286>>
3. HUNT, X., TOMLINSON, M., SIKANDER, S., SKEEN, S., MARLOW, M., DU TOIT, S., EISNER, M. *Artificial intelligence, big data, and mHealth: The frontiers of the prevention of violence against children [online]*. Frontiers in Artificial Intelligence, 2020, 3, 543305. ISSN 2624-8212. [viewed 20 March 2025]. Available from: <<https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2020.543305/full>>
4. INTOTHEMINDS. Schwab, P.-N. *What is artificial intelligence and how it may save humanity [online]*. 28 June 2019, Latest update: 15 March 2022. [viewed 20 March 2025]. Available from: <<https://www.intotheminds.com/blog/en/what-is-artificial-intelligence-and-how-it-may-save-humanity/>>
5. U.S. DEPARTMENT OF STATE. *Artificial Intelligence (AI) [online]*. 2025. [viewed 20 March 2025]. Available from: <<http://web.archive.org/web/20250320121456/https://www.state.gov/artificial-intelligence/>>
6. LINKEDIN. VALURA. *Data-driven management systems [online]* 2025. [viewed 20 March 2025]. Available from: <<https://www.linkedin.com/pulse/data-driven-management-systems-valura-net-dsymf/>>
7. GIBBS, D. J., LOPER, A., FARLEY, A., AFKINICH, J. L., JOHNSON, I. C., METZ, A.J. *Implementing Algorithmic Decision-Making Tools in Child Welfare Systems: Practitioner Perspectives on Use and Usefulness [online]*. Journal of Technology in Human Services, 2024, 42(4), pp. 277–311. ISSN 1522-8835. [viewed 20 March 2025]. Available from: <<https://doi.org/10.1080/15228835.2024.2402982>>

8. THE IMPRINT. Suggs, B. *In a New Study, Georgia Professor Explores the Use of Artificial Intelligence in the Child Welfare Field [online]*. 11/26/2024. [viewed 20 March 2025]. Available from: <<https://imprintnews.org/top-stories/in-a-new-study-georgia-professor-explores-the-use-of-artificial-intelligence-in-the-child-welfare-field/256116>>
9. PAUL, R., CARMEL, E., COBBE, J. (Eds.). *Handbook on Public Policy and Artificial Intelligence*. Cambridge: Edward Elgar Publishing, 2024. ISBN 978-1-80392-216-4.
10. ZYTEK, A., LIU, D., VAITHIANATHAN, R., VEERAMACHANENI, K. *Sibyl: Understanding and addressing the usability challenges of machine learning in high-stakes decision making*. IEEE Transactions on Visualization and Computer Graphics, 2021, 28(1), pp.1161-1171. Electronic ISSN 1941-0506, Print ISSN 1077-2626
11. CHARISI, V., CHAUDRON, S., DI GIOIA, R., VUORIKARI, R., ESCOBAR-PLANAS, M., SANCHEZ, I., GOMEZ, E. *Artificial Intelligence and the Rights of the Child: Towards an Integrated Agenda for Research and Policy [online]*, EUR 31048 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-51837-2. [viewed 20 March 2025]. Available from: <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC127564/JRC127564_01.pdf>