




# Automation of Decision-Making in Guardianship Authorities and the Ethics of Computational Management

Ahmarow A.V.,<sup>1</sup> <sup>a</sup>, Vakhrusheva E.Y.,<sup>2</sup> <sup>b</sup>, Mezinova G.N.<sup>2</sup> <sup>c</sup>

<sup>1</sup>*Candidate of Philosophical Sciences, Assistant Professor, Department of Theory and Technology of Social Work, Federal State Budgetary Educational Institution of Higher Education "A.A. Kadyrov Chechen State University", Grozny, Russian Federation.*

<sup>2</sup>*PhD in Philosophy, Associate Professor, Department of Social Work Don State Technical University, Rostov-on-Don, Russian Federation., ahmed.ahmarow@mail.ru, ktyf1555@yandex.ru, mezinoff-sergei@yandex.ru*


**Keywords:** automation of decision-making, guardianship authorities, computational management, algorithmic.


**Abstract:** This article analyzes the automation of decision-making in child welfare institutions as a form of computational governance that transforms regimes of power and care. The theoretical framework combines concepts of governmentality, risk society, quantification, the performativity of classifications, and an actor-network perspective to demonstrate how algorithmic scales of "risk" define new orders of visibility and intervention. Methodologically, the work draws on normative-ethical and comparative legal analyses of international and European standards, as well as on the critical literature on digital surveillance and administrative classification. It is shown that automation amplifies structural asymmetries: biased data and closed feedback loops sustain surveillance cycles; inevitable metric trade-offs distribute the burden of error across groups; "automated authority" narrows professional discretion; and functional displacement expands goals and data sharing beyond the original mandate, threatening privacy and trust. The regulatory framework is defined by the principles of prioritizing the best interests of the child, non-discrimination, participation, respect for private and family life, as well as a data protection regime and requirements for high-risk systems, including human oversight, documentation, and follow-up monitoring. A responsible implementation model consisting of seven modules is proposed: justification of objectives; double impact assessments; careful data management; phased pilot; effective human oversight; institutional transparency and participation; ongoing audit, adjustment, and child well-being metrics.


## 1 INTRODUCTION

In the digitalization of public administration, the field of guardianship and trusteeship occupies a special place, where decisions directly affect the well-being of the child and the limits of state intervention in family life. Sociological debates about "governance" demonstrate that modern regimes increasingly rely on computational practices of population management: the classification, regulation, and prevention of risks (from the "micro-powers" of institutions to macro-regulations). This perspective, rooted in Michel Foucault's analysis of power and governance, helps describe the transition from individual expert

expertise to sociotechnical constructs, where an algorithm defines the framework of acceptable decisions and the individual acts as the executor of the envisioned scenario. At the same time, Ulrich Beck's concept of "risk society" and developments in reflexive modernization emphasize the expansion of preventive rationality: governance is based on the anticipation of hypothetical threats and the redistribution of risks among population groups. For child protection agencies, this means proactive intervention based on statistical models of adversity, even before harm occurs—with inevitable questions about the fairness, proportionality, and accountability of such interventions.

<sup>a</sup>  <https://orcid.org/0000-0201-5305-1664>

<sup>b</sup>  <https://orcid.org/0009-0007-1493-2561>

<sup>c</sup>  <https://orcid.org/0000-0001-6314-9492>

In recent years, child protection agencies in a number of countries have been implementing predictive decision support systems: scoring models that assess the likelihood of repeat cases, child removal, or other adverse family outcomes. One of the most studied cases is the family risk assessment algorithm in Allegheny County (USA), which aggregates administrative data and generates recommendations for hotline dispatchers. Publications by the developers and independent methodological reports document that the model, calibrated against historical service data and a number of metrics, can improve the consistency of decisions between specialists, while requiring ongoing monitoring of fairness and Type II errors. Such systems illustrate a shift from manual case management to a sociotechnical infrastructure where data quality, threshold design, and review procedures are key elements of decision ethics.

At the same time, criticism of the computational governance of child welfare is mounting among sociologists of inequality and surveillance researchers. Virginia Eubanks has demonstrated the mechanism of "poverty profiling," in which the use of social benefits becomes a proxy for "risk" and leads to an overemphasis on low-income families; this logic reproduces structural inequalities in access to private services and the "invisibility" of wealthy households. Journalistic investigations and advocacy have led to legal scrutiny of some tools for discrimination, including on the basis of disability, while academic work on racial disparities in interventions highlights the dangers of automating existing biases. These narratives resonate with Shoshana Zuboff's broader critique of "surveillance capitalism," where everyday digital traces become a resource for behavioral prediction and management—a process that threatens to shift the focus of welfare states from assistance to control. Regulatory guidelines are provided by international and regional law: the "best interests of the child" doctrine and the protection of privacy in the Convention on the Rights of the Child define the boundaries of permissible processing of sensitive data and require the child's participation in decisions affecting them; in European jurisdictions, the right not to be subject to decisions based solely on automated processing (including profiling) and the right to an explanation and review apply. Specialized recommendations from international organizations on the use of artificial intelligence systems in relation to children emphasize the principles of necessity, data minimization, non-discrimination, and independent impact assessments. This framework establishes an ethical "safety contour" within which it is possible to

discuss the proportionality of algorithmic prevention and the distribution of responsibility between developers, agencies, and child protection professionals.

Looking sociologically at the care apparatus itself, computational governance redistributes forms of capital and autonomy within the bureaucratic field: digital rules and metrics consolidate the symbolic power of central actors (standard-setters, program providers), while simultaneously narrowing the discretion of frontline workers. In Pierre Bourdieu's terms, this is a shift in the boundaries of the field—from professional judgment to algorithmically determined orthodoxy. At the same time, the dynamics of Michael Lipsky's "street-level bureaucracy" are realized: frontline employees continue to make decisions under time pressure and resource constraints, but now their practices are "mediated" by scoring cues and digital reporting. An "audit culture" is emerging, in which procedural verifiability replaces discussion of care goals, and control is transformed into a ritual of confirming the correctness of numbers.

The key operation of computational governance is quantification. Historians of science demonstrate that "trust in numbers" is not simply an epistemic virtue, but a political technology of impersonal trust, allowing decisions to be justified by reference to standardized indicators and procedures. Numbers create the "appearance" of objects of governance and thereby define the scope of permissible interventions. In social policy, this manifests itself in the proliferation of indicators and scales that define who is considered vulnerable and what signals to respond to. Methodological guidelines for child statistics capture precisely this logic: the prioritization of indicators of violence, alternative care, and social vulnerability as the basis for planning and monitoring.

Classification here is not neutral: as Ian Hacking has shown, the types ascribed to people "turn" back on them, changing their behavioral trajectories and the trajectories of institutional interventions; categories live, interacting with those they describe. In child welfare, this means that labels of "risk" and "distress" can produce self-perpetuating cycles of surveillance and sanctions. In the terms of Bruno Latour's actor-network theory, algorithms, forms, checklists, and databases are fully-fledged actors in the sociotechnical network of care: they connect professionals, families, and normative expectations, stabilizing "regimes of visibility" and decision-making pathways. This is how a "reality of governance" is constructed, in which the computable

begins to determine what counts as care, intervention, and evidence.

The concept of a "risk society" provides a broader context: preventative management of uncertainty becomes the norm rather than the exception; prediction and prevention are prioritized over reaction. Within this framework, digital surveillance and "surveillance capitalism" transform the everyday traces of family activity into raw material for profiling and prediction, which risks shifting care toward control and deepening inequality for low-income households. Research shows that automated systems in social policy often reproduce and exacerbate the vulnerability of those already dependent on assistance. Hence the demand for public accountability and transparency of algorithms in the public sector as a condition for the legitimacy of interventions in such an ethically charged area as guardianship.

The ethical and legal foundations for automation in child care stem from fundamental norms of child protection and from an analysis of the consequences of "computational governance" for vulnerable groups. The normative core is based on the principles of the child's best interests, non-discrimination, respect for dignity, the right to be heard, and the right to private and family life. These principles are enshrined in international child rights instruments and are further elaborated in general comments, which emphasize that any administrative procedures affecting children must ensure their "procedural audibility" and effective legal remedies, including in the digital environment. For automated decision-making systems, this entails the obligation to define the conditions for the participation of the child and legal representative, transparency in the logic of the system's actions, and verifiability of its consequences for a specific case.

At the same time, the right to respect for private and family life defines the limits of permissible digital surveillance. In the logic of "governmentality," the excessive accumulation and linking of administrative data sets, transformed into predictive "risk" scores, transforms care into a hidden form of discipline, reinforcing power asymmetries between families and bureaucracy. Shoshana Zuboff's concept of "surveillance capitalism" helps us understand how the transformation of lived experience into behavioral data pushes institutions toward predictive and ranking governance, which in child protection settings threatens stigmatization and a reduction in family autonomy. Legally, this shift should be constrained by tests of necessity and proportionality of intervention,

stemming from the standard of respect for private and family life.

The legal framework for data protection in the European Union operationalizes these principles. The core principles of processing—lawfulness, purpose limitation, minimization, accuracy, retention time limitation, integrity and confidentiality, and controller accountability—are mandatory in the design and operation of any systems that touch children's personal data. Accountability means that the guardianship authority must not only adhere to the principles but also be able to demonstrate this—from a description of the purposes to the decision-making process and the rationale for the selection of data sources. Of particular importance are compliance with age-related consent requirements when providing online services, protection from purely automated decisions with legally significant consequences, and the right to access data and "meaningful information about the logic" of processing. All of this establishes an ethical framework for "explainable" and "contestable" governance in social policy.

A separate procedural barrier is the data protection impact assessment: for systematic, large-scale, and profiling processing that could have significant effects on rights and freedoms, such an assessment is mandatory before the system is launched. For child care institutions, this entails analyzing sources of error (selection, incompleteness, historical biases), verifying methodologies and the validity of indicators, and documenting risk mitigation measures and human oversight mechanisms. The ethical significance of the assessment is to identify potential "side effects" of the technology and establish the limits of its autonomy before intervening in a child's life.

From a global ethical perspective, guidelines are provided by the recommendation of the Intergovernmental Organization for Education, Science, and Culture, which enshrines the principles of human-centeredness, non-discrimination, fair distribution of benefits and burdens, appropriate human oversight, transparency, and accountability throughout the entire life cycle of systems. Regarding children, the obligations to prevent harm, support participation, and create safe conditions for digital socialization are emphasized. These principles define the language for ethical codes and procedural "ethics by design" in child welfare agencies: from the interpretability of models to appeal procedures and the suspension of implementation in the event of unforeseen consequences.

European technology regulation takes a risk-based approach. Acts on artificial intelligence systems prohibit practices incompatible with human dignity and rights, including the social assessment of citizens by public authorities. For high-risk systems (which include systems affecting access to key social services), requirements are established for risk management, data quality, documentation, registration, human oversight, logging, and post-market monitoring. The provisions are being introduced in stages after official publication and entry into force; this creates a window for institutional "learning" and adjustment of practices before the widespread implementation of predictive tools in the social sphere.

The regional human rights organization's standards on children's rights in the digital environment deserve special attention. They explicitly articulate states' obligations to ensure safety, information, access to justice, and redress mechanisms, as well as to involve children in the development of policies and services. For child protection authorities, this means that technical architectures should include feedback and participatory channels, and administrative regulations should guarantee the "voice of the child" in decision-making based on assessment scores.

Sociological analysis points to the risk of exacerbating structural inequalities when using "predictive" tools. Researchers demonstrate that digital systems tend to transform social problems into individual "risks" and historical data into self-perpetuating predictions, creating stigmatizing effects and "locking" families into categories of disadvantage. Using case studies of social assistance and child protection practices, a "digital care home" is described, where poverty and marginality are digitized and accompanied by intense surveillance. From an ethical perspective, this conflicts between the rationalization of record-keeping and the principle of fairness as equal respect for the agency of families, especially when "indicators" replace dialogue and social work.

This entails the requirement for procedural justice: explainability and verifiability of results, the right to human intervention and appeal, equal access to information about the reasons for decisions, and institutional guarantees against discriminatory profiling. In the legal realm, this means an obligation to provide data subjects with access to information that affects the outcome of a case and the ability to challenge decisions based solely on automated processing. These guarantees must be integrated into frontline work—application intake, initial screening,

case referral—and supported by independent oversight and external audit of models.

Children's right to participation requires additional institutional mechanisms: adapted communication, representation, support, and deliberative processes that take into account the child's experience, not just statistical correlations. International documents specifically emphasize that in any administrative procedure affecting a child, they must be heard in a manner appropriate to their age and maturity, and in the digital environment, protected from arbitrary interference with their privacy. These provisions form the "ethical infrastructure" for algorithmic services: without them, automation turns the right to participation into a formality.

The risky horizon of "computational management" in guardianship stems from the very logic of quantification: the translation of life situations into formalized "risk" scales and binary thresholds is not neutral. Classifications operate performatively: the types assigned to people "return" to them, changing their behavioral trajectories and institutional impacts. In Ian Hacking's terms, this is the "looping effect" of human categories; in the language of Bowker and Starr's standardization studies, any classification embodies someone's point of view, transforming the social world into manageable rubrics from which priorities for control and resource allocation directly follow. For guardianship, this means that labels of "distress" and "risk" can stabilize a regime of surveillance, where the very appearance of "problem" generates grounds for further intervention.

The first major node of vulnerability is data. Historical data sets document inequalities in access to services and selective oversight: poor families and groups already under supervision are more visible in administrative registries and therefore more likely to be included in training sets. Virginia Eubanks has shown how "poverty profiling" transforms assistance use into a proxy for "risk," reinforcing stigma and redistributing the burden of oversight onto the vulnerable. Practices in implementing predictive systems in child protective services in the United States, described by advocates, confirm the rapid proliferation of such tools amid low awareness among families themselves. This is not simply an epistemic problem of accuracy: it is a structural bias in visibility, triggering an "audit culture" and accustoming organizations to respond with metrics instead of caring.

Real-world cases demonstrate how incomplete and biased data translates into discriminatory effects. A pre-screening algorithm in Allegheny County in the

United States has come under scrutiny from the Department of Justice due to the risk of adversely affecting parents with disabilities; An independent analysis showed that the design of the features and the learning method could have led to systematic "re-flagging" of such families. Journalistic investigations and legal initiatives around these tools highlight the danger of perpetuating previous biases under the guise of "objective" accounting.

A similar lesson emerged from the European agenda: the Dutch SyRI system for identifying the "risk" of benefit fraud was found by a court to be disproportionate and opaque, violating the right to privacy; international courts have ruled that large-scale data aggregation for the purpose of "profiling" vulnerable areas is unlawful. This precedent highlighted "functional bias": the tendency to expand the purposes of data reuse beyond those initially stated once the infrastructure is in place. In the social assistance sector, this manifests itself in the temptation to push thresholds, add new sources and tasks, turning the assistance tool into a filtering mechanism for "risky" households.

Feedback loops reinforce these dynamics. When past system decisions ("confirmed" notifications, withdrawals, "processed" signals) become features in the learning process, the algorithm begins to predict and reproduce its own visibility contours. Researchers warn that such contours create "self-fulfilling" trajectories—increased attention to a category produces more "problem" detections, which retrospectively confirms the "accuracy" of the model and justifies increased surveillance. In child welfare, this means that districts and families once subject to intensive monitoring receive more contact with the system and, as a result, more "risk signals."

The next dilemma is the inevitable tradeoffs between metrics. Theorems about the "impossibility" of simultaneously satisfying several popular fairness criteria for risk scoring demonstrate that, with different baseline outcome frequencies across groups, it is impossible to achieve both calibrated outcomes and equal error rates (except in trivial cases). This is not an academic abstraction: in child welfare, the choice between "not missing a danger" and "not triggering a false alarm" inevitably impacts different groups and family resources asymmetrically. Attempts to "remove" contradictions by clarifying practices only highlight that the stakes are distributed among types of errors—and this is a political and ethical, rather than a purely technical, solution.

The behavioral effects of "automated authority" also cannot be underestimated. Studies of the interaction between people and algorithms in the

public sector document two persistent trends: excessive reliance on recommendations even in the presence of countersignals (automation bias) and selective acceptance of algorithmic advice when it aligns with stereotypes. Combined with the time constraints of street-level bureaucrats, this narrows the space for professional discretion, turning "decision support" into a de facto norm from which it is difficult to deviate. For families, this regime is experienced as "the rule of numbers": explanations are reduced to thresholds and scores, even when the context and alternatives are obvious to specialists.

Another vulnerability is lack of transparency and accountability. European practice emphasizes that public algorithms must be transparent to the public and oversight bodies: registers of the systems used, disclosure of purpose and logic, and independent assessments. Where such mechanisms are absent or merely formal, the risk of a "black box" increases, in which errors or biases can persist for years. A prime example is the massive false flags of fraud in the British benefit system, which led to thousands of unnecessary checks on households. For child welfare services, such failures not only entail administrative costs but also serious biographical consequences for children and parents.

Legal restrictions partially address these risks, but also create new dilemmas. The European legal order includes the right not to be subject to decisions based solely on automated processing, and stricter requirements are being imposed on high-risk systems accessing social services: risk management, quality of enrollments, record-keeping, human oversight, and post-market monitoring. At the same time, bans on social assessment and certain forms of profiling are being introduced. This strengthens safeguards, but raises questions of practical compatibility for child protection authorities: How can we ensure the "person in the loop" is protected without reducing it to a formal signature? How can we reconcile the requirements of explainability with the protection of a child's sensitive data? Where is the line drawn between necessary prevention and disproportionate intervention?

The context of international recommendations expands this ethical framework: organizations working on children's rights emphasize the need for children to participate in procedures, the right to information and redress, "transparency in the case," and impact assessments before systems are launched. But it is here that the "transparency versus confidentiality" dilemma manifests itself: how much information about indicators and weights is permissible to disclose so that a family can

understand and challenge a decision without further stigmatizing the child or opening the door to behavior "in line with the model"? In the context of digital inequality, this is especially acute: resources and competencies for appeals are unequally distributed. Finally, institutional settings often "shift" the original goals. "Functional displacement" occurs when the established data and scoring infrastructure begins to serve new purposes—from expanding screening to interagency exchanges—beyond the original logic of assistance. Such expansions, as reviews of social protection and identification show, threaten the principle of goal limitation and proportionality, especially if accompanied by higher thresholds and increased risk-based triage. In child welfare, this threatens to normalize "preventive suspicion," where families become the subject of intense monitoring without apparent justification.

Another vulnerability is lack of transparency and accountability. European practice emphasizes that public algorithms must be transparent to the public and oversight bodies: registers of the systems used, disclosure of purpose and logic, and independent assessments. Where such mechanisms are absent or merely formal, the risk of a "black box" increases, in which errors or biases can persist for years. A prime example is the massive false flags of fraud in the British benefit system, which led to thousands of unnecessary checks on households. For child welfare services, such failures not only entail administrative costs but also serious biographical consequences for children and parents.

Legal restrictions partially address these risks, but also create new dilemmas. The European legal order includes the right not to be subject to decisions based solely on automated processing, and stricter requirements are being imposed on high-risk systems accessing social services: risk management, quality of enrollments, record-keeping, human oversight, and post-market monitoring. At the same time, bans on social assessment and certain forms of profiling are being introduced. This strengthens safeguards, but raises questions of practical compatibility for child protection authorities: How can we ensure the "person in the loop" is protected without reducing it to a formal signature? How can we reconcile the requirements of explainability with the protection of a child's sensitive data? Where is the line drawn between necessary prevention and disproportionate intervention?

The context of international recommendations expands this ethical framework: organizations working on children's rights emphasize the need for children to participate in procedures, the right to

information and redress, "transparency in the case," and impact assessments before systems are launched. But it is here that the "transparency versus confidentiality" dilemma manifests itself: how much information about indicators and weights is permissible to disclose so that a family can understand and challenge a decision without further stigmatizing the child or opening the door to behavior "in line with the model"? In the context of digital inequality, this is especially acute: resources and competencies for appeals are unequally distributed. Finally, institutional settings often "shift" the original goals. "Functional displacement" occurs when the established data and scoring infrastructure begins to serve new purposes—from expanding screening to interagency exchanges—beyond the original logic of assistance. Such expansions, as reviews of social protection and identification show, threaten the principle of goal limitation and proportionality, especially if accompanied by higher thresholds and increased risk-based triage. In child welfare, this threatens to normalize "preventive suspicion," where families become the subject of intense monitoring without apparent justification.

1. Another vulnerability is lack of transparency and accountability. European practice emphasizes that public algorithms must be transparent to the public and oversight bodies: registers of the systems used, disclosure of purpose and logic, and independent assessments. Where such mechanisms are absent or merely formal, the risk of a "black box" increases, in which errors or biases can persist for years. A prime example is the massive false flags of fraud in the British benefit system, which led to thousands of unnecessary checks on households. For child welfare services, such failures not only entail administrative costs but also serious biographical consequences for children and parents.

2. Legal restrictions partially address these risks, but also create new dilemmas. The European legal order includes the right not to be subject to decisions based solely on automated processing, and stricter requirements are being imposed on high-risk systems accessing social services: risk management, quality of enrollments, record-keeping, human oversight, and post-market monitoring. At the same time, bans on social assessment and certain forms of profiling are being introduced. This strengthens safeguards, but raises questions of practical compatibility for child protection authorities: How can we ensure the "person in the loop" is protected without reducing it to a formal signature? How can we reconcile the requirements of explainability with the protection of a child's sensitive data? Where is the line drawn

between necessary prevention and disproportionate intervention?

3. The context of international recommendations expands this ethical framework: organizations working on children's rights emphasize the need for children to participate in procedures, the right to information and redress, "transparency in the case," and impact assessments before systems are launched. But it is here that the "transparency versus confidentiality" dilemma manifests itself: how much information about indicators and weights is permissible to disclose so that a family can understand and challenge a decision without further stigmatizing the child or opening the door to behavior "in line with the model"? In the context of digital inequality, this is especially acute: resources and competencies for appeals are unequally distributed. Finally, institutional settings often "shift" the original goals. "Functional displacement" occurs when the established data and scoring infrastructure begins to serve new purposes—from expanding screening to interagency exchanges—beyond the original logic of assistance. Such expansions, as reviews of social protection and identification show, threaten the principle of goal limitation and proportionality, especially if accompanied by higher thresholds and increased risk-based triage. In child welfare, this threatens to normalize "preventive suspicion," where families become the subject of intense monitoring without apparent justification.

4. Pilot and phased rollout. Before scaling up, a limited pilot in several locations will be conducted with independent ethical and legal review, feedback from families and children, and specialist observation diaries. Thresholds influencing case routing (e.g., priority of screening) will be calibrated in an open format, with the protocol and expected tradeoffs between the costs of Type I and Type II errors for different subgroups published. Research shows that it is impossible to simultaneously satisfy multiple fairness criteria with different initial event prevalence; therefore, the choice of thresholds is a publicly agreed-upon normative decision, not a purely technical adjustment. Such a choice should undergo a public consultation process and be enshrined in regulations, specifying the frequency of review.

5. Human oversight and professional competence. The right to "disobey" an automated decision, the obligation to review complex cases, and the genuine ability of a specialist to change it are essential elements. This isn't a formal "interface label," but rather material conditions: time for analysis, professional freedom protected from sanctions,

methodological guidelines for interpreting forecasts, and standards for recording justifications. European law enshrines human control over high-risk systems; data protection regulations recognize the individual's right not to be subject to decisions based solely on automated processing and that have significant consequences. For guardianship authorities, this means developing local regulations for "cancelling" automated prompts, as well as training programs on critically interpreting predictions and identifying situations where the model may be systematically erring.

6. Institutional transparency and participation. Transparency must be systemic: a public registry of the systems used, describing their purpose, data, operating logic, control procedures, and responsibilities; published model passports; threshold calibration protocols; accessible CRIA and DPIA reports; and feedback and notification channels for families. Cities and agencies across Europe are developing algorithm registries based on agreed standards, complemented by the ability to provide feedback and ask questions. For child protection agencies, such a registry enhances the legitimacy of practices, facilitates interagency knowledge sharing, and reduces the risk of "hidden automation." The introduction of the registry should be accompanied by staff training and regular publication of reports on implementation, including the frequency of interventions, the distribution of errors, and the social impact.

7. Continuous monitoring, audit, and adjustment. After deployment, a monitoring plan with metrics related to child well-being (access to services, safety, health, learning, protection from harm, material conditions) should be developed, regularly disaggregated by gender, age, migration status, disability, and other indicators of vulnerability. International platforms and dashboards offer proven sets of indicators that can be adapted at the agency level. External periodic audits are mandatory: data verification, replicability testing, bias analysis, and social impact assessment. Audit results are published, and corrective measures are budgeted in advance.

## 2 CONCLUSION

Automated decision-making in child protection agencies is not a neutral technological modernization, but a profound restructuring of the practices of power and care. It is shaping a new regime of governmentality, in which the quantification of uncertainty, the standardization of categories, and

preventive rationality transform the biographies of children and families into manageable indicators. The performative nature of classifications means that the scoring of "risk" does not simply describe the situation but reconfigures intervention trajectories, reinforcing a "regime of visibility" where suspicion and prevention become the norm. An actor-network perspective shows that forms, checklists, weighting coefficients, and databases act as if they were people: they connect specialists, families, and norms, stabilizing decision patterns and the distribution of services' attention.

An ethical and legal framework delineates the limits of what is permissible and defines the language of responsibility. The principle of prioritizing the best interests of the child, non-discrimination, respect for private and family life, and the right to be heard require that digital procedures ensure participation, explainability, and the possibility of revision. Data protection regimes and risk-based regulation of high-impact systems translate these principles into operational requirements: the necessity and proportionality of processing, dual impact assessments, documentation of objectives, human oversight, and subsequent monitoring. In this logic, algorithmic support does not replace professional judgment, but is limited to the role of an auxiliary tool whose influence is transparent and debatable.

The risks and dilemmas identified in the study are structural in nature. Incomplete and biased data create closed loops of reinforcement: past decisions become predictors of future ones, and the appearance of "accuracy" masks the reproduction of old biases. Tradeoffs between error types and criteria of group fairness are inevitable; the choice of thresholds inevitably distributes the burden between "not missing a threat" and "not generating unnecessary intervention." The effect of "automated authority," combined with the time pressures of street-level bureaucrats, narrows the space for autonomy and transfers responsibility to the "rule of numbers." Functional drift pushes infrastructures to expand their goals and share data beyond their original mandate, threatening privacy and undermining trust.

The proposed responsible implementation model demonstrates how to transform computational tools from a mechanism of stratification into a resource for equitable care. Its core consists of seven interconnected modules: justification of purpose and legal basis; child rights impact assessment, coupled with a data protection assessment before implementation; careful data management and the exclusion of variables that lead to indirect discrimination; phased deployment with open

threshold calibration and public recognition of inevitable trade-offs; meaningful human oversight with the right to disagree without sanctions; institutional transparency through a public registry and accessible explanations; ongoing monitoring, external audit, and adjustments linked to child well-being indicators. This architecture is complemented by capacity-building programs for specialists and sustainable channels for child and family participation, so that the "voice of the child" is embedded in every decision loop, rather than being replaced by statistical correlation. Thus, "computational management" can serve as a tool for care only if the lens shifts—from managing suspicion to managing trust and support. This presupposes a political and ethical agreement between developers, agencies, the professional community, and families: transparent rules of the game, a fair distribution of benefits and burdens, the priority of dignity, and a real possibility of challenge. The scientific agenda for the coming years includes the development of agreed-upon metrics of child well-being to assess the impact of interventions, research into the long-term social consequences of automated case routing, and sociotechnical trials combining quantitative and qualitative methods. Only in this framework can algorithms not replace care, but make it more equal, targeted, and accountable, preserving the final say and responsibility for decisions that truly serve the child's interests.

## REFERENCES

- Andreev, V.V., Arshinov, M.Yu., 2022. Koncentratsiya troposfernogo ozona nad territoriej Rossii v 2021 g. *Optika atmosfery i okeana*, 35 (7), 559–571.
- Bezrukova, N.A., Chernokulsky, A.V., 2023. Rossijskie issledovaniya oblakov i osadkov v 2019–2022 gg. *Izvestiya RAN. Fizika atmosfery i okeana*, 59 (7), 882–914.
- Karamov, D.N., 2018. Matematicheskoe modelirovanie solnechnoj radiatsii s ispolzovaniem mnogoletnih meteorologicheskikh ryadov. *Izvestiya Tomskogo politehnicheskogo universiteta. Inzhiniring georesursov*, 329 (1), 69–88.
- Mordynsky, A.V., 2024. Metody opredeleniya rasseyannogo solnechnogo izlucheniya s tselyu povysheniya tochnosti prognozirovaniya pochasovoj vyrabotki PV. *Problemy energetiki*, 48 (2), 73–82.
- Sakerin, S.M., Gorbarenko, E.V., Kabanov, D.M., 2008. Osobennosti mnogoletej izmenchivosti aerodol'noj opticheskoy tolshchiny atmosfery i otsenki vliyaniya razlichnykh faktorov. *Optika atmosfery i okeana*, 21 (7), 625–631.

- Salamov, O.M., Mamedova, L.G., Salmanova, F.A., Yusupov, I.M., 2021. Vybory modeley raschyota summarnoj solnečnoj radiatsii yasnogo neba v naklonnoj ploskosti v vysokih shirotah Rossii. Problemy energetiki, 44 (6), 58–69.
- Zavyalov, D.D., Solomakha, T.A., 2021. Parametrizatsiya pogloshcheniya solnečnoj radiatsii snezhno-ledyanym pokrovom v termodinamicheskoj modeli l'da Azovskogo morja. Morskoy gidrofizicheskiy zhurnal, 37 (5), 538–553.