VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7934

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

COLLABORATIVE DIGITAL SUPERVISION MODEL INVOLVING PARENTS, CHILDREN, AND AI IN PARENTAL CONTROL APPLICATIONS

Made Edy Listartha[1*], Gede Arna Jude Saskara[1], Bagus Gede Krishna Yudistira[1]

Universitas Pendidikan Ganesha, Singaraja, Indonesia[1]
www.undiksha.ac.id [1]
listartha@undiksha.ac.id[*]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License

Submitted: 16 December 2025, **Revised:** 01 September 2026, **Accepted:** 02 October 2026, **Published:** 13 November 2026.

Abstract— *Rapid advances in digital technology increase children's exposure to online risks, prompting widespread use of parental control applications. Traditional one-way control approaches raise concerns about privacy, autonomy, and family conflict. This study proposes TriadSafe, a collaborative supervision model developed through a User-Centered Design (UCD) approach. Unlike existing rigid tools, TriadSafe integrates three actors: parents, children, and artificial intelligence (AI) within a layered system architecture. The model positions AI as an active mediation engine that processes negotiation logic and adapts policies to family values and child feedback. Using a descriptive-exploratory qualitative method involving 12 purposively selected participants (parents, children, and IT experts), thematic analysis identified core themes: the shift from monitoring to mediation, the necessity of AI transparency, and the restoration of child autonomy. The result is a conceptual framework that balances technical oversight with digital ethics, offering a more sustainable alternative to traditional restrictive parental controls.*

Keywords: *AI facilitator, child privacy, collaborative parental control, family values, TriadSafe.*

INTRODUCTION

The swift progress in digital technologies has fundamentally changed the way youths and adolescents engage with information and their social surroundings, escalating their vulnerability to harmful material, online harassment, and risky interactions with unknown individuals. To mitigate these hazards, a variety of parental control tools have been designed to supervise, limit, and regulate children's internet navigation and device usage; yet, unilateral monitoring strategies frequently cause friction between parents and children, obstructing the growth of the child's digital independence [1], [2], [3], [4]. Furthermore, a significant number of commercial applications demand extensive data permissions, which sparks major privacy apprehensions and worries regarding the non-transparent distribution of information to external organizations [5], [6], [7].

Technically speaking, digital supervision applications have progressed from basic keyword-filtering techniques to incorporating machine learning models for adaptable content categorization and advanced user authentication systems [8], [9]. While artificial intelligence facilitates flexible identification of hazardous behaviors and permission-driven architectures strive to harmonize immediate access with youth independence, such technological breakthroughs have yet to completely address challenges related to mutual trust, the clarity of rules, and successful familial dialogue [10], [11]. Existing scholarly reviews highlight that the success of these applications relies on more than just complex algorithms; it also requires synchronization with household principles, active parental guidance, and the significant participation of youths as key actors in digital oversight [12], [13], [14].

Despite these advancements, most current parental control applications predominantly rely on "Black Box" machine learning algorithms that operate on rigid, binary logic: either allowing or blocking content without transparency [15]. These

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7934

models fail to support effective mediation because they lack interpretability and contextual awareness. As noted in recent technical evaluations, traditional algorithms often prioritize surveillance over educational dialogue, failing to provide the "supportive accountability" needed for developing healthy digital habits [16]. Consequently, these algorithms often produce "false positives" that restrict harmless educational content or "over-restriction" that fails to consider the child's growing need for digital exploration.

Existing gaps in both academic literature and practical application highlight the necessity of a novel framework—one that concurrently harmonizes protection, personal autonomy, and data privacy rather than simply intensifying technical restrictions. This technical rigidity creates a significant research gap; while existing literature extensively discusses the risks of online exposure, there is a lack of models that utilize AI as a social mediator to facilitate dialogue between parent and child rather than a mere surveillance tool. Current research suggests that for digital safety tools to be effective, they must move beyond "policing" and transition toward "joint family oversight," where parents and teens manage safety as equals [17]. The problem is not the lack of monitoring, but the absence of a collaborative framework that balances safety with autonomy.

To address this, a cooperative oversight model is suggested; it engages children as active decision-making participants and utilizes AI as an adaptable policy mediator aligned with family norms, rather than a mere restriction mechanism [12], [18], [19], [20]. Driven by these requirements, this research presents TriadSafe, a unified monitoring architecture uniting parents, children, and artificial intelligence to foster clear, engaging, and instructive supervision strategies while reducing domestic disputes and data privacy risks. To explicitly distinguish the proposed framework from existing paradigms and clear the ambiguity regarding its research novelty, a comparative state-of-the-art analysis is presented in Table 1.

Table 1. Comparative Analysis of State-of-the-Art Parental Control Paradigms

Proposed Model / Framework	Filtering & Control Paradigm	Algorithmic Approach	Child Engagement Role	Interaction Architecture
Kumar et al. [21]	Dynamic classification of risky sites	Black-box Machine Learning	Passive Object	Top-down, Automated binary blocking
Vishal [22]	Permission-based notification system	Identity verification (OTP via Ionic/Firebase)	Active Informant	Two-way permission exchange
Akter et al. [5]	Joint family oversight framework	General collaborative communication	Equal Stakeholder	Manual family negotiation
TriadSafe (Proposed Model)	Context-aware socio-technical mediation	Explainable AI Local-First Adaptation Loop	Active Partner with Negotiation UI	Circular Triadic Loop with Automated Scalability

Source: (Research Results, 2025)

MATERIALS AND METHODS

This study employs a descriptive-exploratory qualitative design integrated with a structured User-Centered Design (UCD) framework to bridge user experiential needs with technical system architecture development. The execution of the UCD framework was operationalized through four iterative phases: (1) Understand and Specify the Context of Use: Evaluating the limitations of current binary "Black Box" algorithms in existing parental control platforms through an extensive literature review to define the problem space. (2) Specify User Requirements: Empirical data collection was conducted via semi-structured interviews and Focus Group Discussions (FGD). A purposive sampling strategy was applied to recruit exactly 12 participants: four parents who have utilized digital control tools for at least six months, four adolescents (aged 10–17) acting as digital natives, and four experts (two serving as primary informants and two serving as independent validators). (3) Produce Design Solutions: Translating qualitative insights into system design specifications. This phase produced the three-layered system architecture and formalized the system's operational logic through a dynamic flowchart and pseudocode algorithm. (4) Evaluate Designs Against Requirements: Due to the conceptual nature of the current stage, the architecture and behavioral logic were evaluated using an expert-validated architectural walkthrough. Two independent experts (an AI ethics specialist and a child digital psychology analyst) reviewed the negotiation logic to verify structural feasibility. Thematic analysis was then performed following the Braun & Clarke framework to identify recurring patterns in the participants' responses. Initial codes were generated from the transcripts, such as "perceived surveillance," "negotiation triggers," and "AI as a neutral party." To ensure inter-coder reliability and

CC BY NC

trustworthiness, the coding process was cross-checked by all authors, achieving a consistency rate of over 85%, before the identified themes were mapped into the functional requirements of the TriadSafe model.

RESULTS AND DISCUSSION

As outlined in the methodology, data gathering involved semi-structured interviews with 4 parents and 4 children, alongside 2 focus group discussions (FGDs) featuring combined parent–child participants, and 2 supplementary expert interviews with specialists in AI and digital child psychology. Upon receiving participant consent, all sessions were audio-recorded, transcribed word-for-word, and enriched with field notes recording contextual nuances and nonverbal behaviors. The textual data underwent inductive thematic analysis, during which a codebook was iteratively crafted and validated via cross-coder consistency checks; initial open codes were synthesized into overarching themes using a process of continuous comparison and refinement. Key empirical observations revealed persistent parental anxieties regarding non-transparent data collection and boundary intrusion, children's desires for transparent explanations and dialogue avenues, and expert recommendations advocating for explainability alongside privacy-focused design—insights that laid the direct groundwork for the model components and interaction workflows detailed hereafter.

This qualitative inquiry yielded three primary thematic domains that underpin the TriadSafe model: (1) individual actor roles and expectations; (2) interaction dynamics and negotiation channels; and (3) architectural design requirements centered on user autonomy and transparency. Empirical data revealed a significant gap between parental expectations and children's digital reality. For instance, one parent (P3) highlighted the emotional strain of surveillance: "I use these tools because I'm worried, but I feel like a spy, and it creates distance between me and my son." This was echoed by a child participant (C2) who stated: "I don't mind being safe, but I want to be asked first. It feels unfair when an app just blocks my school research without any explanation." These findings are summarized in Table 2, which maps the qualitative themes to the functional requirements of the model.

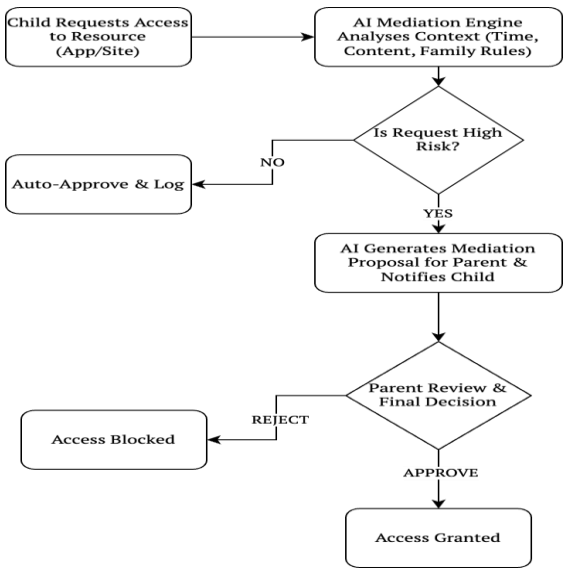
Table 2. Thematic Findings and Supporting Evidence

Theme	Sub-Theme	Representative Quote
Autonomy Restoration	Collaborative Decision Making	"I want to be able to request access and explain why I need it, not just be blocked." (C4)
AI Transparency	Interpretability of Logic	"The app should explain why something is considered risky based on our family values." (P1)
Conflict Reduction	Mediation & Negotiation Loops	"If the AI acts as a mediator, we avoid the direct arguments that usually happen." (P5)

Source: (Research Results, 2025)

Technical Framework: Layered Architecture and Negotiation Logic

To translate these qualitative findings into a technical solution, the TriadSafe model is conceptualized using a layered system architecture. This architecture consists of three primary layers: (1) The Presentation Layer, providing the interface for users; (2) The Logic Layer (AI Mediation Engine), where negotiation algorithms reside; and (3) The Data Layer, storing family policies and context rules. The core innovation is the Negotiation Loop Algorithm. Unlike traditional binary models (Allow/Block), this algorithm follows a four-step mediation process: (1) Request Capture, (2) Contextual Inference, (3) Mediation Proposal, and (4) Policy Adaptation. By visualizing this process in Figure 1, the model demonstrates how AI shifts from a surveillance tool to a technical mediator that promotes family dialogue.



Source: (Research Results, 2025)

Figure 1. TriadSafe Negotiation Flowchart

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7934

To provide a rigorous technical foundation for the negotiation behavior shown in the flowchart, the behavioral logic of the AI Mediation Engine is formalized in Algorithm 1.

Algorithm 1: TriadSafe Collaborative Negotiation Loop

```
Input: ChildRequest R (RequestedResource, UserJustification), FamilyValueMatrix M, ActivePolicy P
Output: AccessAction (GRANT, BLOCK, MEDIATE)
Initialize RequestContext C ← ExtractContext(R) // Extracts Time, ContentCategory, Prior Behavior
RiskScore L ← ContextualInference(C, M) // Evaluates risk against family value parameters
if RiskScore L == LOW then
    LogActivity(R, "Auto-Approved")
    return GRANT
else if RiskScore L == HIGH then
    MediationProposal Prop ← GenerateProposal(R, M, P) // Suggests parameters like time-extension
    SendNotificationToParent(Prop)
    SendNotificationToChild("Access restricted. Request routed to Parent with justification.")

    ParentDecision ← AwaitHumanResponse(Timeout = 24h)
    if ParentDecision == APPROVE then
        UpdateActivePolicy(P, R.RequestedResource, TemporaryGrant = True)
        UpdateAdaptationEngine(M, R.UserJustification) // Refines the value matrix dynamically
        return GRANT
    else
        LogActivity(R, "Parent Denied Request")
        return BLOCK
    end if
end if
```

Source: (Research Results, 2025)

Actor-level results

Parent model. Parents in the sample expressed a dual need: effective protection against clear online harms and tools that support constructive conversations rather than unilateral enforcement. They reported significant "monitoring fatigue" and a sense of guilt when using rigid surveillance tools. As stated by P3: "I use these tools because I'm worried, but I feel like a spy, and it creates distance between me and my son." Consequently, the parent model in TriadSafe shifts from being a "sole enforcer" to a "final decision-maker" who receives context-aware recommendations from the AI, allowing for more reasoned and less emotional interventions. Child model. Children reported greater acceptance of digital supervision when they were treated as partners and given clear, non-punitive explanations for restrictions. They emphasized that autonomy is not about total freedom but about the right to provide context for their activities. One child participant (C2) noted: "I don't mind being safe, but I want to be asked first. It feels unfair when an app just blocks my school research without any explanation." This feedback led to the inclusion of the "Negotiation Request" feature in TriadSafe, where children can submit justifications (e.g., for educational purposes) that the AI then processes for parental review.

AI facilitator model. Unlike commercial tools that act as silent filters, the AI in TriadSafe is designed as a transparent mediator. Expert interviews with IT specialists emphasized that the AI must move away from "Black Box" logic. According to the experts, the system should prioritize "explainability" so that both parents and children understand the reasoning behind a risk assessment. Within this model, the AI serves as a neutral buffer that reduces direct parent-child friction by providing data-driven suggestions based on the family's pre-defined value matrix. The integration of these actor requirements into a unified digital framework demonstrates that TriadSafe is not merely a social concept but a technically feasible model. By replacing rigid binary filters with an interpretable mediation engine, the system addresses the critical need for 'explainable AI' in domestic settings [23], ensuring that technology serves to strengthen, rather than disrupt, family dynamics.

Interaction patterns between actors

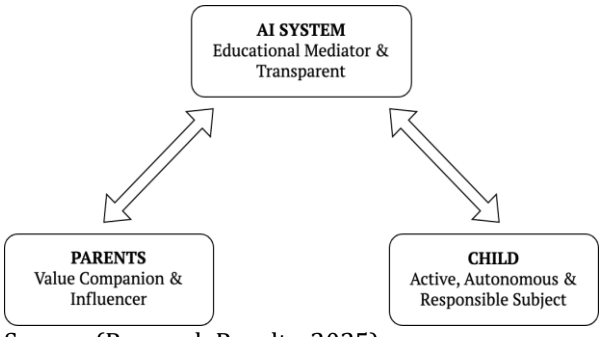
The interaction patterns in TriadSafe move beyond the traditional parent-to-child command structure. Instead, the model establishes a "triadic loop" where the AI serves as a dynamic buffer. When a child encounters a restricted digital boundary, the interaction is not immediately terminated. As shown in the Negotiation Flowchart (Figure 1), the system triggers a mediation sequence: the child provides a justification, the AI evaluates the contextual risk, and the parent receives a structured proposal. This pattern transforms digital supervision from a source of conflict into a pedagogical opportunity. By providing a "cooling-off" period through AI mediation, the model reduces impulsive parental restrictions and encourages

CC BY NC

children to reflect on their digital choices. Expert validation suggests that this layered interaction ensures that technical enforcement remains aligned with the fluid nature of family values.

Design considerations and trade-offs

The implementation of TriadSafe involves critical trade-offs between system security and user autonomy. One primary consideration is the latency in the decision-making process. Unlike traditional "Black Box" filters that block content in milliseconds, TriadSafe's negotiation loop introduces a delay while waiting for parental review or AI contextual analysis. To mitigate this, the model employs a "pre-approval" cache for educational domains, balancing immediate access with safety. Another significant trade-off is the privacy-accuracy balance. To provide highly accurate recommendations, the AI requires access to contextual data (such as the child's intent and current activity). However, to protect child privacy, TriadSafe adopts a "Local-First" data processing approach [24], where sensitive interactions are analyzed on the device rather than the cloud. While this enhances privacy, it may limit the AI's ability to learn from broader, cross-platform behavioral patterns. These design choices reflect a deliberate shift from maximize control to maximizing trust within the family ecosystem. As shown in Figure 2, the conceptual interaction model maps the three primary actors (parents, children, and AI) and their bidirectional sociological relationships within the TriadSafe ecosystem.



Source: (Research Results, 2025)

Figure 2. TriadSafe Model

Implementation requirements

Deploying the TriadSafe framework relies on three core technical foundations: an interpretable machine learning pipeline for context-aware classification, a collaborative mediation interface facilitating parent-child negotiation mechanisms, and robust privacy engineering aimed at limiting data exposure. The machine learning architecture should emphasize edge processing or local execution whenever feasible, leverage inherently transparent models or post-hoc interpretability methods, and deliver intuitive, human-readable justifications for automated actions. Additionally, the mediation user interface must present operational policies in straightforward language, permit minors to submit organized requests or input, and offer parents distilled, actionable insights rather than unparsed system logs. Lastly, privacy-by-design principles must mandate strict data minimization, local storage of confidential records, and highly specific, transparent consent protocols prior to off-device data transmission.

System architecture and integration

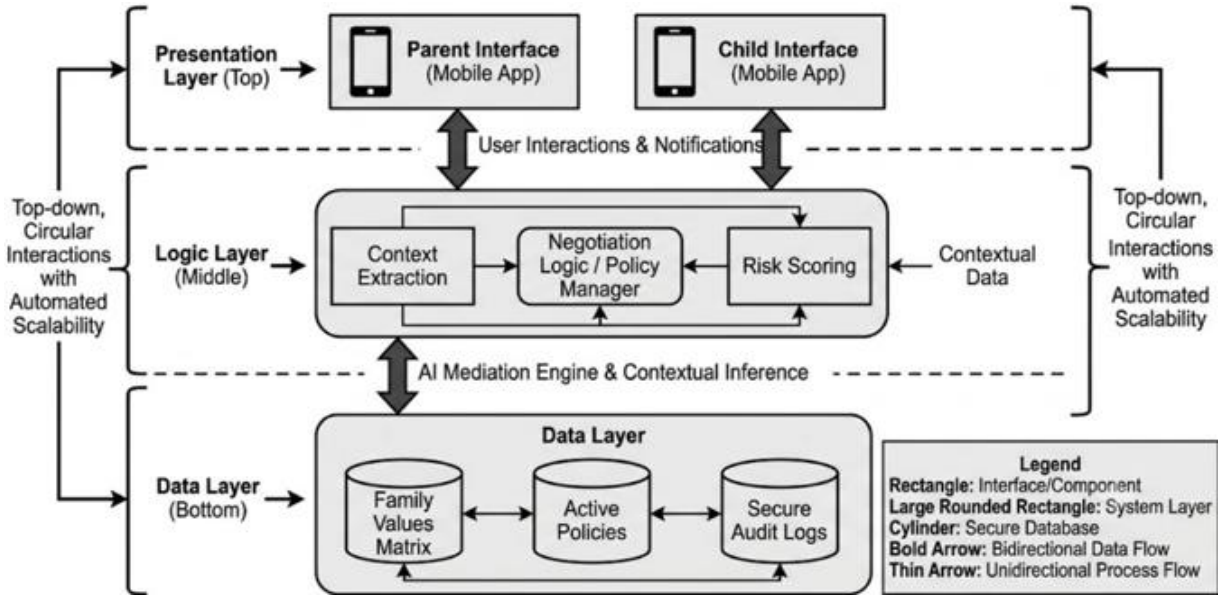
The TriadSafe model is conceptualized through a layered system architecture to ensure clear separation between user interaction and backend logic. Based on expert validation during the interview phase, the integration is designed across three functional layers:

1) Presentation Layer: This layer manages the user interfaces for parent and child devices. It is designed to be cross-platform, ensuring that the negotiation prompts are delivered in real-time. Experts suggested using asynchronous notification protocols (such as Firebase Cloud Messaging) to ensure the negotiation loop does not create significant system lag.
2) Logic Layer (AI Mediation Engine): This is the core "middleware" where the negotiation logic resides. The system integrates the child's access request with the Family Value Matrix. Instead of a simple "yes/no" database query, the integration here involves a contextual analysis service that evaluates the request's purpose against pre-negotiated rules.
3) Data Layer: This layer ensures data integrity and privacy. To address ethical concerns, the architecture employs a decentralized storage approach. Negotiation logs and personal justifications are stored locally on the family's primary device (e.g., the parent's phone) with encrypted cloud backups, ensuring that sensitive family dialogue is not accessible to service providers.

To provide a concrete technical visualization of these integrated layers, Figure 3 presents the TriadSafe Layered System Architecture Block Diagram. This technical blueprint explicitly illustrates the functional boundaries and data flow between the user-facing Presentation Layer, the core computational Logic Layer (AI Mediation Engine), and the underlying secure Data Layer.

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7934



Source: (Research Results, 2025)

Figure 3. TriadSafe Layered System Architecture Block Diagram

By defining these architectural layers, TriadSafe provides a robust blueprint for future technical development. This structured approach ensures that the model can be integrated into existing parental control infrastructures while maintaining high standards of data privacy and system interpretability.

Ethical and practical considerations

The deployment of TriadSafe raises critical ethical questions regarding data sovereignty and the power dynamics within the family. From an ethical perspective, the model prioritizes "Privacy by Design." Unlike traditional applications that often share children's behavioral data with third-party advertisers, TriadSafe's ethical framework dictates that all negotiation logs and personal justifications remain within the family's private encrypted silo. This ensures that the child's right to a "private digital space" is respected, even under supervision. From a practical standpoint, the success of this model depends on the digital literacy of the parents.

There is a risk that parents might become overly reliant on AI recommendations without critically evaluating the context. To mitigate this, TriadSafe is designed as a decision-support system, not a decision-maker. Practically, this means the system requires periodic "Value Alignment" sessions where parents and children manually review and update the AI's mediation rules, ensuring the technology evolves alongside the child's increasing maturity and responsibility.

Roadmap and next steps

In the short term, research efforts will focus on constructing a functional minimum viable product that incorporates essential negotiation pathways, localized contextual classification, and a foundational mediation interface, alongside pilot usability assessments across varied family demographics. Over the medium term, system models and user interfaces will be refined using empirical pilot feedback, robust data protection mechanisms (such as local differential privacy and encrypted backups) will be integrated, and system interoperability will be broadened. Finally, long-term objectives entail executing extensive field trials with impact evaluations, disseminating research findings, and collaborating with regulatory bodies and platform vendors to advance benchmarks for ethical, family-centered digital oversight mechanisms.

CONCLUSION

This study introduces TriadSafe, a collaborative digital supervision model designed to balance the needs for protection, autonomy, and privacy in parental control applications. By integrating a layered system architecture with a negotiation-based AI engine, the model shifts the digital parenting paradigm from unilateral surveillance to transparent mediation. The empirical findings, derived from qualitative thematic analysis, confirm that both parents and children seek more communicative oversight mechanisms that prioritize explainability over rigid,

CC BY NC

binary blocking. However, several inherent limitations must be rigorously acknowledged. As an expert-validated architectural concept, the TriadSafe model has not yet been deployed within a live production software environment, nor has it been subjected to empirical system latency, throughput testing, or usability testing under real-world parameters. Furthermore, the empirical patterns are derived from a localized purposive sample of 12 participants. Therefore, this conceptual design offers a highly potential architectural blueprint rather than a fully verified commercial system. Prototype software development and empirical usability testing are explicitly prioritized as the immediate next steps in future work to transform this architecture into a functional application.

Future work should focus on the technical development of the TriadSafe engine, specifically in implementing machine learning algorithms that can accurately interpret family values. Additionally, longitudinal studies are recommended to evaluate the long-term impact of AI-mediated negotiation on children's digital literacy and the quality of parent-child relationships. Such developments will be crucial in transforming this conceptual blueprint into a practical tool for the digital age.

DECLARATION OF AI USE

During the preparation of this work, the authors used Google Gemini Edu in order to improve language, readability, and grammar of the manuscript. After using this tool or service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

AUTHOR CONTRIBUTIONS

The individual contributions of the authors to the research and manuscript preparation are outlined as follows:

Author	Contribution Roles
Made Edy Listartha	Conceptualization, Methodology, Formal Analysis, Investigation, Writing – Original Draft.
Gede Arna Jude Saskara	Data Curation, Investigation, Validation, Visualization, Writing – Review & Editing.
Bagus Gede Krishna Yudistira	Project Administration, Funding Acquisition, Writing – Review & Editing.

ACKNOWLEDGMENTS

The authors extend their sincere appreciation to all participants (parents, adolescents, and experts) for their active involvement in the data collection process, as well as to the reviewers for their insightful suggestions that helped improve the quality of this manuscript.

FUNDING

This research was conducted under the independent research scheme of Universitas Pendidikan Ganesha and was self-funded by the authors. No external grant or financial support was received for the conduct of this study or the preparation of this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest regarding the publication of this manuscript. The research was conducted in the absence of any commercial, financial, or personal relationships that could be construed as a potential conflict of interest.

REFERENCE

[1] J. O. Richmond and C. Jemima N., "The Importance of Parental Monitoring and Its Impact on Children's Online Safety," in *Integrating Parental Consent and Child Engagement With Digital Protection Rules*, IGI Global, 2025, pp. 251–280. doi: 10.4018/979-8-3373-2716-7.ch010.

[2] S. Y. Jeyam, N. Mounika, S. Vignesh, and T. S. Devi, "Child Digital Monitoring and Controlling System," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, IEEE, May 2023, pp. 695–699. doi: 10.1109/ICAAIC56838.2023.10141525.

[3] D. Celestino, H. Medina, C. Lara, and N. Saboya, "Smart Mobile Apps for Responsible Child Management: A Systematic Literature Review," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 8, 2025, doi: 10.14569/IJACSA.2025.0160864.

[4] S. Kallur, Jambayya. V. Ganti, M. Kudari, A. S. Nandeppanavar, and P. N. Thotad, "ChildSafe: An Offline Android-Based Screen Time Management and App Restriction System for Children," in *2025 IEEE North Karnataka Subsection Flagship International Conference*

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7934

(NKCon), IEEE, Sep. 2025, pp. 1–6. doi: 10.1109/NKCon66957.2025.11345788.

[5] J. K. Park, M. Akter, P. Wisniewski, and K. Badillo-Urquiola, "It's Still Complicated: From Privacy-Invasive Parental Control to Teen-Centric Solutions for Digital Resilience," *IEEE Secur. Priv.*, vol. 22, no. 5, pp. 52–62, Sep. 2024, doi: 10.1109/MSEC.2024.3417804.

[6] W. Shafik, "Parental Consent in the Digital Age," in *Integrating Parental Consent and Child Engagement With Digital Protection Rules*, IGI Global, 2025, pp. 191–222. doi: 10.4018/979-8-3373-2716-7.ch008.

[7] M. Campbell, A. Barthwal, S. Joshi, A. Shouli, and A. K. Shrestha, "Investigation of the Privacy Concerns in AI Systems for Young Digital Citizens: A Comparative Stakeholder Analysis," in *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, Jan. 2025, pp. 00030–00037. doi: 10.1109/CCWC62904.2025.10903925.

[8] U. Aryanti, "Pemanfaatan Aplikasi Parental Control Sebagai Media untuk Mewujudkan Internet Sehat dan Aman," *Jurnal AbdiMU : Pengabdian Kepada Masyarakat*, vol. 3, no. 1, pp. 28–33, Mar. 2025, doi: 10.32627/abdimu.v3i1.726.

[9] P. Dumaru and M. N. Al-Ameen, "It's like educating us older people: Unveiling Needs and Expectations Regarding Educational Features within Parental Control Tools," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, May 2024, pp. 1–8. doi: 10.1145/3613905.3651113.

[10] S. T. Kaithathara and A. Jose, "Empowering Child Safety Using AI and ML to Address Online Risks," in *Child Protection Laws and Crime in the Digital Era*, IGI Global Scientific Publishing, 2025, pp. 361–398. doi: 10.4018/979-8-3373-5132-2.ch015.

[11] A. Rawat, A. Rajavat, P. Bhardwaj, and P. Muzumdar, "AI-Driven Consent Mechanisms," in *Integrating Parental Consent and Child Engagement With Digital Protection Rules*, IGI Global Scientific Publishing, 2025, pp. 417–440. doi: 10.4018/979-8-3373-2716-7.ch016.

[12] M. Stoilova, M. Bulger, and S. Livingstone, "Do parental control tools fulfil family expectations for child protection? A rapid evidence review of the contexts and outcomes of use," *J. Child. Media*, vol. 18, no. 1, pp. 29–49, Jan. 2024, doi: 10.1080/17482798.2023.2265512.

[13] S. Nichols, N. Selim, S. Nichols, and N. Selim, "Digitally Mediated Parenting: A Review of the Literature," *Societies 2022, Vol. 12,* vol. 12, no. 2, Mar. 2022, doi: 10.3390/SOC12020060.

[14] M.-H. Yu and R. A. Rose, "Algorithmic-Assisted Decision-Making Tools in Child Welfare Practice: A Systematic Review," *Res. Soc. Work Pract.*, vol. 36, no. 4, pp. 382–397, May 2026, doi: 10.1177/10497315251350933.

[15] J. Liu, Y. Cai, Z. Zhang, Y. Guo, and D. Li, "WiP: Beyond App Blocking: Fine-Grained Child Online Safety via MLLM-based GUI Provenance on the Edge," in *Proceedings of the 2nd International Workshop on Edge and Mobile Foundation Models*, New York, NY, USA: ACM, Nov. 2025, pp. 9–11. doi: 10.1145/3737902.3768353.

[16] O. Lindhiem *et al.*, "Novel Smartphone App and Supportive Accountability for the Treatment of Childhood Disruptive Behavior Problems: Protocol for a Randomized Controlled Trial (Preprint)," Oct. 11, 2024. doi: 10.2196/preprints.67051.

[17] M. Akter, A. J. Godfrey, J. Kropczynski, H. R. Lipford, and P. J. Wisniewski, "From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals?," *Proc. ACM Hum. Comput. Interact.*, vol. 6, no. CSCW1, Apr. 2022, doi: 10.1145/3512904.

[18] M. Kumar, M. Shuja, B. Ahmed, T. Gera, D. Thakur, and V. Bhardwaj, "Enhancing Digital Safety," in *Integrating Parental Consent and Child Engagement With Digital Protection Rules*, IGI Global, 2025, pp. 371–394. doi: 10.4018/979-8-3373-2716-7.ch014.

[19] S. Mathiyazhagan and K. La Fors, "Children's right to participation in AI: Exploring transnational co-creative approaches to foster child-inclusive AI policy and practice," *Information Polity*, vol. 28, no. 1, pp. 141–153, Feb. 2023, doi: 10.3233/IP-229014.

[20] S. Sharma *et al.*, "Age Against the Machine: A Call for Designing Ethical AI for and with Children," in *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference*, New York, NY, USA: ACM, Oct. 2022, pp. 1–6. doi: 10.1145/3547522.3547702.

[21] M. Kumar, M. Shuja, B. Ahmed, T. Gera, D. Thakur, and V. Bhardwaj, "Enhancing Digital Safety," in *Integrating Parental Consent and*

CC BY NC

Child Engagement With Digital Protection Rules, IGI Global, 2025, pp. 371–394. doi: 10.4018/979-8-3373-2716-7.ch014.

[22] V. Vishal, "Design and Development of Parental Notification Control System," *SSRN Electronic Journal*, May 2024, doi: 10.2139/ssrn.4817618.

[23] J. Singh, S. Rani, and G. Srilakshmi, "Towards Explainable AI: Interpretable Models for Complex Decision-making," in *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, IEEE, Apr. 2024, pp. 1–5. doi: 10.1109/ICKECS61492.2024.10616500.

[24] V. C. Andrade, R. D. Gomes, S. Reinehr, C. O. D. A. Freitas, and A. Malucelli, "Privacy by Design and Software Engineering," in *Proceedings of the XXI Brazilian Symposium on Software Quality*, New York, NY, USA: ACM, Nov. 2022, pp. 1–10. doi: 10.1145/3571473.3571480.

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.