

Privacy-Preserving Voice Interaction and Federated Learning for Intelligent Home IoT Systems

Umash Varmo

Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL,
USA.

verma2002@uab.edu

Andres Heirra

Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence,
KS, USA.

helloandres@ku.edu

Abstract

Intelligent home Internet of Things systems increasingly depend on voice interaction as a primary control interface, enabling users to manage lighting, climate, entertainment, security, and health-related devices through speech. However, continuous audio capture and cloud-based speech processing raise severe privacy concerns because spoken commands and incidental acoustic events can reveal identity, health status, interpersonal relationships, and behavioral routines. Federated learning offers a structural alternative to centralized voice data aggregation by allowing voice models to be trained across distributed home devices without transferring raw audio to central servers. This paper presents a system-level analysis of privacy-preserving voice interaction and federated learning for intelligent home IoT systems. It examines architectural choices, edge-cloud resource allocation, data governance, model aggregation, and security trade-offs. The discussion extends beyond algorithmic performance to consider infrastructure sustainability, deployment complexity, robustness against adversarial manipulation, fairness across households and demographic groups, and the regulatory implications of decentralized machine learning. By integrating perspectives from distributed systems, human-computer interaction, security engineering, and technology policy, the paper identifies structural tensions between personalization, privacy, and system utility. It argues that sustainable privacy-preserving voice interaction requires coordinated design across local voice processing, federated orchestration, governance mechanisms, and regulatory frameworks rather than isolated technical fixes. The analysis also highlights future directions including secure multiparty computation, on-device foundation models, and cross-domain learning from health and industrial IoT deployments.

Keywords

privacy-preserving voice interaction; federated learning; Internet of Things; smart home; edge computing; data governance; socio-technical systems.

1. Introduction

The expansion of voice-enabled devices in residential environments has transformed smart home systems from optional conveniences into integrated infrastructures for everyday life. Microphone-equipped speakers, thermostats, security panels, appliances, and mobile devices continuously listen for wake words and process commands that require access to personal schedules, media preferences, household composition, and sometimes health-related

information. The dominant deployment model sends voice recordings or acoustic features to cloud servers where large automatic speech recognition models perform transcription, intent recognition, and personalization. While cloud processing supports model accuracy and rapid iteration, it creates concentrated data repositories that are attractive targets for unauthorized access, commercial profiling, and state surveillance. The structural coupling between voice interaction and centralized data collection therefore conflicts with the privacy expectations of smart home occupants.

Federated learning has emerged as an influential paradigm for reconciling machine learning utility with data minimization. In federated learning, model updates are computed locally on participant devices and transmitted to a coordinator, while raw training data remain on local hardware [1]. This approach reduces the need to store voice recordings centrally and allows personalization to be learned from local usage patterns without exposing the underlying speech content. Federated systems nonetheless introduce significant complexity in communication scheduling, update aggregation, device heterogeneity, and security. Early work demonstrated that decentralized training can achieve competitive model quality under non-IID data distributions [2], but later analyses show that privacy gains are not automatic and must be actively engineered across the entire system stack [3]. Intelligent home environments add further constraints because residential devices vary in computational power, network stability, energy availability, and acoustic conditions.

This paper develops a system-level perspective on privacy-preserving voice interaction and federated learning for intelligent home IoT systems. Rather than focusing narrowly on model accuracy, the analysis addresses structural trade-offs among local processing, cloud coordination, privacy protection, robustness, fairness, sustainability, and policy compliance. The remainder of the paper proceeds as follows. Section 2 describes the architectural organization of voice interaction and edge-based learning. Section 3 examines federated learning integration and data governance. Section 4 discusses privacy, security, and structural trade-offs. Section 5 addresses infrastructure, deployment, and sustainability. Section 6 explores robustness, fairness, and socio-technical implications. Section 7 analyzes policy and regulatory dimensions. Section 8 offers a forward-looking comparison across domains, followed by the conclusion.

2. Voice Interaction Architecture and Edge Processing

An intelligent home voice interaction system typically comprises several functional layers that span edge devices, local gateways, and remote cloud services. The first layer includes microphones, digital signal processors, wake word detectors, and voice activity detection modules. These components must remain active for long periods while consuming minimal energy. Once a wake word is detected, subsequent speech is captured, preprocessed, and passed to an automatic speech recognition pipeline. The second layer performs feature extraction, acoustic modeling, language modeling, and natural language understanding. In centralized architectures this entire pipeline may reside in the cloud, but privacy-preserving designs increasingly move wake word detection, keyword spotting, and compact intent classifiers onto local hardware. A third layer manages device actuation, context matching, and dialogue state, translating recognized intents into commands for connected appliances.

The placement of processing functions across edge and cloud has profound implications for privacy, latency, reliability, and model capability. On-device voice models can operate without continuous streaming of raw audio, thereby reducing exposure to network interception and centralized retention. However, local models are constrained by memory,

processing speed, and battery capacity, which limits vocabulary size, contextual reasoning, and adaptation to novel commands. Federated learning architectures address this by distributing model updates rather than data. Local voice models are trained on speech interaction logs that remain on the device, and only encrypted or compressed weight updates are sent to a federated coordinator. Research on mobile keyboard prediction illustrates how personalized language models can be improved through federated averaging across many devices without transferring user text [4]. Analogously, federated keyword spotting has been explored to adapt wake word models to different voices and acoustic environments while avoiding the collection of private speech samples [5]. However, home devices exhibit high statistical heterogeneity because households differ in language use, accent, background noise, device placement, and interaction frequency. Federated optimization methods must therefore account for non-identical data distributions and asynchronous participation [6].

Voice interaction in multilingual households introduces additional architectural pressure. A single home may contain speakers who alternate between languages, use code-switching, or issue commands with accented speech. Centralized systems can draw on large multilingual datasets, but privacy constraints limit the feasibility of transferring those datasets to a cloud repository. The structural challenge is to maintain multilingual capability through decentralized model updates while ensuring that local utterances are not exposed. This requires careful design of vocabulary adaptation, pronunciation modeling, and intent parsing across distributed devices. Such considerations highlight that voice interaction cannot be treated as a monolithic application; it is a layered socio-technical system whose privacy properties depend on where data are generated, transformed, stored, and discarded.

3. Federated Learning Integration and Data Governance

Federated learning in home IoT systems operates through a coordination process in which a central server manages model initialization, participant selection, update aggregation, and distribution of new model versions. The local devices act as data-generating clients that perform stochastic gradient updates on their own interaction logs. The server combines these updates using weighted averaging or robust aggregation techniques, then redistributes the updated global model. This model-centric governance structure changes the nature of data protection because raw audio does not need to leave the home. However, the model updates themselves can encode sensitive information about speech patterns, vocabulary, and behavioral routines. Therefore, privacy preservation depends on the entire information lifecycle rather than a single data minimization decision. Federated machine learning has been framed as a general architectural pattern for collaborative learning across organizations and devices, with attention to security, incentive design, and regulatory constraints [7].

Data governance in federated voice interaction must address several dimensions. The first is consent and transparency. Household members may have different levels of awareness that a smart speaker is participating in federated training. Voice interfaces often blend into domestic life, making continuous consent difficult to obtain. The second dimension is purpose limitation. Audio data collected for wake word detection should not be reused for emotion recognition, identity inference, or marketing profiles without explicit authorization. The third dimension is retention. Local training logs can be stored for limited periods and sanitized after model updates. The fourth dimension is access control. Federated coordinators must not be able to reconstruct individual voice samples from aggregated updates. Privacy research on IoT environments has documented how seemingly innocuous sensor data can reveal sensitive behaviors and identities [8]. Security analyses of smart home applications further demonstrate

that fragmented access control and inconsistent firmware updates create vulnerabilities that undermine privacy assurances [9]. Protocol-level security mechanisms for constrained IoT devices must be integrated with federated orchestration to protect update integrity and authentication [10].

The governance of voice interaction also involves social expectations. People often respond to voice agents as social actors, attributing intentions and forming interaction habits that may reduce vigilance about data collection [11]. Federated learning does not automatically resolve these social dimensions. If a device informs users that data never leave the home, users may accept voice interactions more readily, but model updates may still reveal sensitive information under inference attacks. Therefore, governance frameworks should combine technical controls with clear communication, household-level consent interfaces, and audit mechanisms that verify how updates are generated and aggregated. Data minimization should be treated as an operational principle rather than a single architectural switch.

4. Privacy, Security, and Structural Trade-offs

Privacy-preserving voice interaction through federated learning introduces a set of structural trade-offs among accuracy, latency, energy consumption, communication overhead, and threat resistance. Moving computation to the edge reduces raw audio exposure but limits model size and adaptation speed. Secure aggregation and differential privacy can reduce the risk of extracting individual utterances from model updates, but they may degrade model performance and increase convergence time. Conversely, aggressive personalization improves user satisfaction but can strengthen the link between a model update and a specific household. The quest for privacy in IoT systems thus requires balancing multiple properties that are not fully compatible [12].

A key threat to voice systems is inference from speech. Even if raw audio is not transmitted, acoustic features, model gradients, and update trajectories can reveal demographic attributes, emotional states, or the content of commands. Large language models trained on broad web data may also reproduce sensitive associations, reinforcing the need for careful data curation and evaluation beyond accuracy metrics [13]. Differential privacy offers a formal mechanism to bound the influence of individual training samples on model outputs. When applied to federated voice models, differentially private noise can mask the contribution of a single utterance, but it may disproportionately harm accuracy for underrepresented accents or uncommon commands [14]. The choice of privacy parameters therefore has distributional consequences that must be assessed across different user groups.

Beyond privacy leaks, federated voice systems face active poisoning attacks. Malicious participants may submit manipulated updates that cause the global model to misrecognize certain wake words, trigger unintended commands, or exfiltrate private information through encoded gradients. Research on backdoor attacks in federated learning shows that a small number of adversarial clients can compromise the global model while remaining difficult to detect with simple averaging [15]. In home IoT settings, such attacks could cause locks to open, alarms to deactivate, or sensitive commands to be misinterpreted. Structural defenses include robust aggregation rules, outlier rejection, validation on trusted datasets, and isolated personalization layers that limit the influence of global updates. Multilingual natural language processing in smart home systems adds further complexity because adversarial perturbations may hide in minority language features that are difficult to monitor [16]. This underscores the need to view privacy and security as intertwined system properties rather than separate requirements.

Another structural trade-off concerns the relationship between local voice processing and cloud fallback. Some voice interactions require large knowledge bases, real-time information retrieval, or complex dialogue reasoning that cannot be executed locally. Hybrid architectures may process sensitive speech locally and send only anonymized intents to the cloud. But anonymization is often imperfect because intent sequences and contextual metadata can reveal identity. System designers must therefore decide which operations are strictly necessary in the cloud and which must remain under local control. These decisions implicate not only engineering performance but also organizational power, because cloud dependency shapes who can access, analyze, and monetize voice interaction data.

5. Infrastructure, Deployment, and Sustainability

Deploying federated voice interaction across residential IoT systems requires substantial attention to infrastructure heterogeneity. Smart home devices range from low-power sensors and microcontrollers to high-performance home gateways and dedicated edge accelerators. Federated training on battery-powered devices may be infeasible for large models because sustained computation and wireless communication shorten battery life. A practical architecture may use hierarchical federated learning in which local devices send updates to a home gateway, and the gateway aggregates updates before communicating with an external coordinator. This design reduces wide-area network traffic and enables localized adaptation, but it also creates a new aggregation point that must be hardened against intrusion.

Sustainability concerns extend beyond energy consumption. Continuous model training and frequent update distribution consume network bandwidth and increase the carbon footprint of smart home systems. If millions of devices retrain voice models every day, the cumulative energy cost can rival or exceed centralized training in certain configurations. However, federated learning can also reduce the need for massive cloud data centers dedicated to continuous speech collection and storage. The net sustainability impact depends on model size, update frequency, device efficiency, and the energy mix of local and remote infrastructure. System designers should therefore evaluate the environmental consequences of privacy-preserving architectures rather than assuming that decentralized computation is automatically sustainable.

Deployment reliability is another critical factor. Home networks vary widely in bandwidth, latency, and availability. Federated clients may drop out during training, submit stale updates, or operate under intermittent power. The coordinator must handle partial participation without introducing bias toward households with stable connectivity. If aggregation favors devices that are always online, the global model may learn primarily from affluent households or specific device configurations, reducing its utility for less connected populations. Scheduling mechanisms, asynchronous updates, and robust weighting can mitigate these effects, but they add complexity to the orchestrator and require transparent evaluation across diverse deployment contexts.

The lifecycle management of voice models also affects sustainability. Model updates must be tested for regression across multiple acoustic environments, languages, and device types before distribution. Continuous deployment pipelines for federated voice models need mechanisms for canary testing, rollback, and versioning. Without these mechanisms, a faulty update could render voice assistants unresponsive or cause privacy leaks in millions of homes. The infrastructure therefore must support observability without compromising privacy, for example by collecting aggregate performance metrics that do not include raw audio or

household-specific transcripts. This tension between operational monitoring and privacy protection is a persistent structural challenge in intelligent home IoT systems.

6. Robustness, Fairness, and Socio-Technical Implications

Robustness in federated voice interaction refers to the ability of the system to maintain acceptable performance under changing acoustic conditions, device failures, adversarial manipulation, and distributional shift. Home environments are acoustically complex, with background music, television, conversations, and reverberation. Federated models trained on local data can adapt to these conditions, but they may also overfit to idiosyncratic noise and fail to generalize to new users or devices. Robust aggregation methods help filter anomalous updates, but they can inadvertently exclude genuine acoustic diversity. Evaluating robustness requires not only aggregate accuracy but also performance across different household types, languages, and noise profiles.

Fairness is closely related to robustness. Voice recognition systems have historically performed unevenly across demographic groups, with higher error rates for certain accents, dialects, and age groups. Federated learning can either reduce or amplify these disparities. If the coordinator weights updates by participation frequency or data volume, households with greater data production will dominate the global model. If differential privacy noise is calibrated uniformly, minority speech patterns may suffer higher error rates because their representations are less frequently observed. Fairness-aware federated deep learning requires explicit attention to participation balance, incentive design, and evaluation across subgroups [17]. System designers should not assume that privacy preservation and fairness naturally coexist; both require deliberate institutional commitment and participatory design with affected communities.

The socio-technical implications of voice interaction extend beyond individual privacy to domestic power dynamics. Smart speakers are often shared devices, but training data may be associated with a primary account holder, leaving other household members without meaningful control over their voice data. Children, guests, and caregivers may be recorded without full understanding or consent. Federated learning can reduce centralized storage, but it does not automatically resolve questions about who may opt out of local training or delete voice-derived model updates. The commercial context also matters. Surveillance capitalism has demonstrated that seemingly mundane behavioral data can be transformed into predictive commodities [18]. Voice interaction data are especially valuable because they reveal intent, emotional state, and social relationships. Privacy-preserving architectures must therefore be evaluated against the broader political economy of data extraction.

Fairness and transparency also require enforceable audit trails. Federated systems can be designed to record cryptographic commitments, update provenance, and aggregation decisions without exposing raw voice content. These records can support independent audits and enable households to understand how their devices contribute to model improvement. However, audit mechanisms must be meaningful rather than symbolic. They should reveal not only that data remained local but also whether updates were used for purposes beyond the original voice assistant function. The structural legitimacy of federated voice systems depends on public trust, which is difficult to build when terms of service are opaque and technical assurances are not independently verifiable.

7. Policy, Regulatory, and Governance Dimensions

The deployment of federated voice interaction in home IoT systems intersects with data protection law, consumer protection, cybersecurity regulation, and emerging artificial intelligence governance. The General Data Protection Regulation provides a benchmark for data minimization, purpose limitation, consent, transparency, and the right to erasure in contexts involving personal data [19]. Although federated learning reduces raw data transfer, voice interaction data are often personal data because they can identify individuals directly or indirectly. Model updates that encode speech characteristics may also be personal data under certain interpretations. Compliance therefore requires careful documentation of local processing, aggregation procedures, and retention periods.

Consent mechanisms in voice interfaces remain particularly challenging. Traditional notice-and-consent models are poorly suited to continuous audio environments where affirmative consent may interrupt interaction. Federated learning introduces an additional layer because home occupants may not understand that their device participates in model training. Policy frameworks should encourage just-in-time notices, household-level controls, and default configurations that minimize data collection. Regulatory guidance should also clarify whether aggregate model updates fall within data protection obligations and how the right to erasure applies when an individual requests deletion after their voice data have contributed to a trained model. Exact deletion from deep learning models is technically difficult, making this an important governance gap.

Beyond privacy regulation, cybersecurity rules for IoT devices increasingly require secure default settings, regular updates, and vulnerability disclosure. Federated voice systems can support these goals by enabling rapid model improvement without centralizing sensitive data. However, the same mechanism can be abused to distribute malicious model updates across many homes. Regulators should therefore recognize federated learning as a critical infrastructure that requires update integrity verification, participant authentication, and incident response coordination. International harmonization is also necessary because voice interaction models may be trained across households in multiple jurisdictions with conflicting data transfer rules.

Governance should extend beyond legal compliance to include organizational accountability. Device manufacturers, cloud providers, and platform operators should publish privacy impact assessments, model evaluation reports, and aggregate fairness metrics for federated voice systems. Independent oversight can help ensure that privacy claims are technically substantiated and that marginalized groups are not disproportionately harmed by model errors. Standards bodies should develop interoperability frameworks that allow voice models to be federated across different device ecosystems without creating lock-in or undermining privacy. The long-term viability of privacy-preserving voice interaction depends on aligning technical architecture with public governance expectations.

8. Future Outlook and Cross-Domain Comparisons

Future voice interaction systems will likely incorporate larger on-device models, multimodal inputs, and tighter integration with home robotics and health monitoring. These developments will intensify the need for privacy-preserving training because speech will be combined with visual, physiological, and location data. Federated learning can provide a foundation, but it must be extended with secure computation techniques such as secure multiparty computation to protect model updates during aggregation. SecureML and related systems demonstrate that secure computation can be practical for machine learning at moderate scale [20]. Applying

such methods to voice interaction will require further advances in efficiency and integration with heterogeneous edge hardware.

Cross-domain comparisons offer useful lessons. In healthcare IoT, federated learning has been explored to train predictive models across hospitals without sharing patient records. The governance challenges in healthcare are driven by formal regulatory obligations and strong institutional review processes. Smart homes lack equivalent oversight structures, which increases the risk of privacy erosion. In industrial IoT, federated approaches emphasize reliability, predictive maintenance, and safety rather than consumer privacy. Industrial systems often have more uniform devices and controlled network conditions, making aggregation simpler. Smart home systems must instead manage extreme heterogeneity, limited user expertise, and continuous environmental variability. The comparison suggests that home IoT federated learning requires more flexible trust models, stronger household-level controls, and greater tolerance for partial participation.

Another promising direction is the integration of federated voice models with on-device foundation models. Large language models are increasingly compressed and quantized for edge deployment. If speech recognition and natural language understanding can be performed locally, the need for cloud transmission decreases sharply. However, foundation models also raise concerns about memorization and unintended disclosure of training data. Federated fine-tuning with differential privacy may reduce these risks, but it introduces difficult trade-offs between model capability and privacy. Research on multilingual processing in smart homes suggests that localized adaptation across languages can improve usability without centralizing sensitive recordings [16]. Future work should evaluate how such adaptation can be governed across linguistic communities with different privacy expectations.

Interoperability standards will be crucial for federated voice systems to scale beyond single-vendor ecosystems. Without common protocols for update formats, aggregation rules, and security parameters, households may be forced to choose between incompatible platforms. Standardization can promote competition and reduce lock-in while allowing independent auditing. However, standards must not codify a single privacy model because local norms and legal requirements vary. A modular governance architecture that separates raw data handling, model update policy, and audit transparency would offer greater flexibility. The future of privacy-preserving voice interaction thus lies not in a single optimization target but in the coordinated evolution of technical systems, governance institutions, and user-facing controls.

9. Conclusion

Privacy-preserving voice interaction and federated learning represent a significant architectural response to the privacy risks of intelligent home IoT systems. By shifting training from centralized voice repositories to distributed local computation, federated learning reduces the direct collection of raw audio and supports data minimization. Yet privacy is not an automatic property of federation. Model updates may leak sensitive information, malicious participants may poison aggregation, and skewed participation may produce unfair models. System-level analysis reveals that effective privacy preservation requires coordinated attention to edge processing capacity, update authentication, differential privacy noise, aggregation robustness, deployment reliability, energy sustainability, and regulatory compliance. Voice interfaces also operate within social and domestic contexts that complicate consent and governance. Fairness must be evaluated across languages, accents, household types, and participation conditions. Policy frameworks such as the General Data Protection Regulation provide important foundations, but they require interpretation for

decentralized machine learning artifacts and continuous audio environments. Future progress will depend on integrating secure computation, on-device foundation models, interoperability standards, and independent audit mechanisms into a coherent governance model. The central conclusion is that privacy-preserving voice interaction is not merely a technical problem but a socio-technical design challenge. It demands structural choices that balance utility, privacy, fairness, and sustainability across the full lifecycle of intelligent home IoT systems.

References

1. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konecny, J., Mazzocchi, S., McMahan, H. B., Van Overveldt, T., Petrou, D., Ramage, D., & Roselander, J. (2019). Towards federated learning at scale: System design. *Proceedings of Machine Learning and Systems*, 1, 374-388.
2. McMahan, B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 1273-1282.
3. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., Rouayheb, S. E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., ... Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1-2), 1-210.
4. Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.
5. Leroy, D., Coucke, A., Lavril, T., Gisselbrecht, T., & Dureau, J. (2019). Federated learning for keyword spotting. *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6341-6345.
6. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 2, 429-450.
7. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1-19.
8. Ziegeldorf, J. H., Morchon, O. G., & Wehrle, K. (2014). Privacy in the Internet of Things: Threats and challenges. *Security and Communication Networks*, 7(12), 2728-2742.
9. Fernandes, E., Jung, J., & Prakash, A. (2016). Security analysis of emerging smart home applications. *Proceedings of the 2016 IEEE Symposium on Security and Privacy*, 636-654.
10. Granjal, J., Monteiro, E., & Silva, J. S. (2015). Security for the Internet of Things: A survey of existing protocols and open research issues. *IEEE Communications Surveys & Tutorials*, 17(3), 1294-1312.
11. Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 72-78.

12. Porambage, P., Ylianttila, M., Schmitt, C., Kumar, P., Gurtov, A., & Vasilakos, A. V. (2016). The quest for privacy in the Internet of Things. *IEEE Communications Magazine*, 54(2), 32-39.
13. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623.
14. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308-318.
15. Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 2938-2948.
16. Chen, X. (2025). Research on the Application of Multilingual Natural Language Processing Technology in Smart Home Systems. *Journal of Computer, Signal, and System Research*, 2(5), 8-14.
17. Madaio, M. A., Stark, L., Wortman Vaughan, J., & Wallach, H. (2020). Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-14.
18. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.
19. European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Union*, L119, 1-88.
20. Mohassel, P., & Zhang, Y. (2017). SecureML: A system for scalable privacy-preserving machine learning. *Proceedings of the 2017 IEEE Symposium on Security and Privacy*, 19-38.