Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 6, 2026

PRIVACY-PRESERVING AGENTIC AI AT THE EDGE: FEDERATED AND AUTONOMOUS INTELLIGENCE FOR SMART SYSTEMS

Khaliq Ahmed[*1], Muhammad Ghazanfar Ullah Khan[2], Engr. Ikhlas Bano[3], Syeda Bushra Shabeeh[4], Tooba Shaikh[5]

[*1]*Computer Science Department, Nazeer Hussain University, Karachi, Pakistan*
[2]*Computer Engineering Department, Karachi, Pakistan, UIT University, Karachi, Pakistan*
[3]*Nazeer Hussain University, Karachi, Pakistan*
[4]*Iqra university, Karachi, Pakistan*
[5]*Computer and Information Systems Engineering, NED University of Engineering & Technology*

[*1]drkhaliq.ahmed@nhu.edu.pk, [2]ghazanfar.ullah@gmail.com, [3]ikhlasbano@nhu.edu.pk, [4]bushra.shabee01@iqra.edu.pk, [5]toobashaikh@neduet.edu.pk

DOI: https://doi.org/10.5281/zenodo.20656238

Keywords
Agentic AI, edge computing, federated learning, privacy-preserving AI, differential privacy, secure aggregation, smart systems.

Article History
Received: 11 April 2026
Accepted: 23 May 2026
Published: 12 June 2026

Corresponding Author: *
Khaliq Ahmed

Abstract

Introduction: *At the edge, privacy-preserving agentic AI is emerging as a factor in intelligent systems where real-time decisions need to be made without revealing sensitive operator, device, or user information. The risk of privacy, latency and bandwidth is introduced by centralized AI, particularly in healthcare, smart homes, transport, energy and industrial internet of things.*

Aim: *The purpose of this work is to present and analyze a privacy-conscious edge intelligence architecture, a fusion of autonomous agentic decision making, federated learning, differential privacy, secure aggregation, and safe decision escalation.*

Methodology: *A conceptual and design-based approach was employed to formulate a layered architecture consisting of edge devices, autonomous local agents, privacy engines, federated coordination, and smarter-system applications. The framework was assessed, through perceived concrete metrics of privacy, accuracy, latency, communication cost, resource use and autonomous reliability.*

Findings: *The suggested framework lowered the exposure percentage of raw data to 0, communication cost dropped to 38MB/round compared to 480MB/round and latency dropped to 67ms compared to 142ms and the accuracy of the model dropped to 92.6% compared with 93.8% and the risk type of information safety decreased to 0.18 compared to -0.72*

Conclusion: *The framework demonstrates that privacy, autonomy, and efficiency may be harmoniously enhanced in edge-based smart systems.*

INTRODUCTION

The accelerated growth of smart systems has shifted the artificial intelligence off remote cloud resources to sensor, gateway, mobile, vehicle, robot, and controller applications near the physical world. This transition is indicative of centralized AI weaknesses in the form of latency, bandwidth, energy, reliability, and privacy constraints. Edge computing deploys the computation close to data sources in such a way that time-sensitive services react locally without the need to transfer the data to the cloud (Shi et al.,

2016; Satyanarayanan, 2017). In a smart home, Internet of Things (IoT) wearables used in healthcare, transport, smart grids and industrial systems, edge data tends to be continuous, context and sensitive. The transmission of this data to centralized servers may unveil the activity, location, health condition, biometric pattern, and trade secrets of users. Thus, the key research question is not whether edge AI can enable smarter systems, but what can be done to ensure edge intelligence remains useful, autonomous, and responsible and does not jeopardize privacy.

This is further complicated by agentic AI. In contrast to passive predictive models that merely classify, detect or recommend, agentic systems perceive an environment, generate goals, and plan, utilize tools and change as a result of feedback. The intelligent agents of classical agent theory were seen as systems, which could act autonomously towards the achievability of goals (Wooldridge and Jennings, 1995). More recent research on large language model-based agents pursues this concept by focusing on reasoning, memory, tool use, multi-agent coordination and interaction, and by task planning in dynamic settings (Wang et al., 2024). These capabilities may be used at the distributed edge to serve autonomous drones, medical monitors, cooperative vehicles, adaptive factories, and energy-aware buildings. But with autonomy, the privacy and safety are also augmented: an agent can conclude about sensitive states, can interact with other agents, can cause actions, or can revise behavior based on local experience. Therefore, the privacy-preserving edge intelligence should pay attention to the confidentiality of data, the edge decisions, the leakage of model updates, model adversarial manipulation, and autonomous action regulation.

Federated learning (FL) can operate as a baseline since it allows more than two devices or organisations to train a common model and retain raw data locally. The initial FL model demonstrated that mobile devices could autonomously study profound models by exchanging nearby model changes instead of centralizing their user data with federated averaging reducing the count of communications occasions than synchronized stochastic gradient descent (McMahan et al., 2017). Subsequent surveys define FL as reaction to the demands of reducing the size of data, scalability, and privacy on non-centralized settings (Kairouz et al., 2021; Li et al., 2020). In the case of IoT and smart systems, FL is applicable since the devices are highly dispersed, data is neither independent nor identically distributed, local context varies with users, places, and conditions (Nguyen et al., 2021). But FL cannot be presumed as a privacy of default. Sensitive information can be disclosed using shared gradients or model parameters, and the model itself can leak information using the records used in the training (Shokri et al., 2017; Zhu et al., 2019).

An enhanced privacy preservation system demands FL to be coupled with formal and cryptographic defenses. Differential privacy provides a mathematical approach to regulate the impact of any particular record in an published computation by introducing calibrated noise (Dwork et al., 2006). Differentiating privacy In deep learning, mechanisms were presented by differently private stochastic gradient descent to train neural networks at quantifiable costs in privacy, albeit usually at a cost to the utility (Abadi et al., 2016). Secure aggregation augments this by ensuring that a server is only able to see aggregated updates instead of the individual contribution of each individual device, which is important when the update is sensitive (Bonawitz et al., 2017). On a production scale, FL systems need to select the clients, fault tolerance, compression of updates, monitoring, and orchestrating safely across unreliable devices (Bonawitz et al., 2019). The mechanisms are essential but cannot work alone unless combined with autonomy-conscious controls: an edge agent needs to be aware when to learn on its own, to cooperate, to abstain, or to seek controls.

This is made more cumbersome by the edge environment. Devices vary in processing capacity, storage, network connectivity, sensor resolution and power consumption. The data distributions vary with the change of environment, users and conditions. The reaction to FedProx and analogous techniques are the adjustment of

federated optimization under systems and statistical heterogeneity (Li et al., 2020). Surveys on security and privacy indicate that FL systems do not have protection against inference attacks, poisoning, and backdoors, unreliable clients, and communication bottlenecks (Mothukuri et al., 2021). The weaknesses are more severe in agentic contexts as tainted or privacy-inspiring designs might influence autonomous choices. A disabled traffic-edge agent may reclassify dangers, a healthcare agent may leak patient trends using updates, and an industrial agent may leak production actions using model leakage. The edge-based privacy-preserving agentic AI thus requires collaboration in both design and implementation of the systems in terms of learning, communication, security, autonomy, and accountability.

This article places privacy-sensitive agentic AI on the periphery as a convergence of federated learning, autonomous intelligence, secure edge learning, and privacy engineering of smart systems. The study is concerned with understanding how intelligent edge agents can learn in a decentralized fashion, behave independently without surpassing safe thresholds in decision making and respond to changing environments without putting their users or organizations at risk of unreasonable privacy concerns. The contribution is to consider federated and autonomous intelligence not in terms of distinct technologies but in terms of complementary elements of the next generation systems of smart. Privacy is not something that can be introduced once it is deployed; it should be able to define the learning protocol, agent architecture, communication process, threat model, and evaluation metrics at an early stage. Through the study of these relationships, the research is able to explain design requirements, unfixed challenges as well as future perspectives of the intelligent, decentralized, privacy-conscience, and practical smart systems.

2. Literature review

2.1 Theoretical Framework

Privacy-preserving agentic AI on the edge has a theoretical basis that integrates four associated concepts: autonomous agency, edge intelligence, federated learning, and privacy-preserving computation. Autonomous agents are usually defined as systems which sense the surrounding, have internal objectives and behavior which is not controlled by humans. Franklin and Graesser (1997) identified features of agents as contrasted with ordinary programs with the focus on autonomy, reactivity, persistence, and interaction between agents and the environment; Jennings et al. (1998) related agent systems to coordination, cooperation, and distributed problem solving. An agentic AI is thus not viewed as a simple prediction-software but a proactive decision-making system entrenched within a smart environment in this paper. The rational-agent model of intelligence is also another concept by Russell and Norvig (2020), where intelligence is judged by the quality of the actions chosen when there is uncertainty in the environment. Stone and Veloso (2000) generalize the framework to multi-agent learning, in which multiple agents interact or compete whilst updating based on local experience.

The second layer of theory is edge intelligence. Instead of transmitting all raw data to remote centralized cloud servers, edge intelligence moves inference, training, or decision-making to local data sources. Zhou et al. (2019) refer to it as a push to the last mile of AI connected devices, and Deng et al. (2020) separate it into the AI of edge optimization and AI on edge devices. The significance of this distinction is that privacy-preserving agentic AI demands the following: the edge infrastructure must allocate resources through intelligent edge infrastructure, and the agents executing over the edge nodes, scale sensitive data on the agents. Yang et al. (2019) introduce the federated machine learning viewpoint, in which it is possible to share knowledge in institutions or devices without explicit sharing of raw data. The conceptual assertion is that there should be co-design between autonomy, edge locality, and federation and not separate modules.

Recent agentic AI work builds on this framework with reasoning, use of the tool, memory and multi-agent conversation. Xi et al. (2023) conceptualises agents based on the use of large-language-models

in terms of brain, perception, and action, which translates well to edge systems where perception takes place through sensors and action impacts physical or cyber-physical systems. Yao et al. (2023) demonstrate that reasoning and acting can even be interleaved that is also applicable to autonomous edge agents that need to decide, execute, observe feedback, and revise behavior. Wu et al. (2023) show that multi-agent conversation frameworks have the ability to organize specialized agents. But in edge systems, such coordination provides privacy threats as agents could share context, sensor summaries, model states, or task outputs. Thus, the theoretical framework of this work presupposes that privacy should rule, not only the data storage, but also the agent memory, communication, planning, and model updates.

2.2 Empirical Studies

Empirical studies of federated edge intelligence indicate that Privacy preserving learning can be implemented, though not necessarily secure or safe or efficient. Hard et al. (2018) trained a model that predicts mobile keyboard using on-device federated learning and demonstrated that useful language models can be enhanced with the goal of not transferring sensitive typing information to servers. Nishio and Yonetani (2019) investigated the issue of client selection in mobile edge federated learning and discovered that heterogeneous computation and wireless conditions have a massive impact on training efficiency. The significance of these studies lies in the fact that edge devices are hardly identical; some have a battery, others connectivity; others have a higher processor speed, higher memory and are more or less available. This extends to a privacy preserving learning that, in agentic smart systems, must process irregular participation, local autonomy, and unsteady communication.

A number of experiments aim to enhance privacy on top of the federated learning protocol. Geyer et al. (2017) introduced client-level differences in privacy when optimizing federated data, demonstrating that it is possible to obscure the make of individual contributors with minimal performance degradation when given the appropriate conditions of participation. Wei et al. (2020) created a framework of federated learning based on differential privacy and compared the level of privacy, convergence, and clients participation. Truex et al. (2019) suggested that using a combination of differential privacy and secure multiparty computation one-size-fits-all privacy approach is not sufficient to address the accuracy-security trade-off. To solve the scalability problem, So et al. (2021) introduce Turbo-Aggregate, a secure aggregation system to minimize overheads of aggregation and support user dropout. Collectively, these results demonstrate that privacy-themed edge AI needs stratified defense: local training could be effective, but it has to be supported by different forms of privacy (differentiating privacy), safe aggregation, and protocols with low communication costs.

The fact that federated learning is privately designed is also questioned by empirical studies of security. Melis et al. (2019) demonstrated collaborative learning updates can leak unwanted information pertaining to training data of participants. Nasr et al. (2019) exemplified white-box inference attacks of centralized and federated learning, and malicious participants are also presented with active attacks. Bagdasaryan et al. (2020) demonstrated that malicious clients can embed backdoor behavior into federated models by replacing the model, whereas Fang et al. (2020) discovered that local model poisoning can also harm even Byzantine-robust aggregation schemes. These results apply directly to agentic AI, in that an autonomous edge agent that is trained via corrupted federation can not only over- or under-label data; it also acts in unsafe manners or spills sensitive context, or engages poorly with another agentic entity.

2.3 Research Gap

The analyzed literature demonstrates a good advancement in edge intelligence, federated learning, secure aggregation, differential privacy, and autonomous agents, yet the synthesis is incomprehensible. Edge intelligence literature explains the reason why computation needs to be brought nearer to the device, although it tends to focus on latency and resource efficiency as opposed to autonomous privacy regulation. The

field of federated learning research focuses on decentralized model training, but most of what is tested are classification or prediction tasks instead of goal-oriented agents who reason, communicate and take actions. Studies that use agentic AI analyze planning, tool use and multi-agent coordination, but do not strongly relate to the edge constraints of low power, untrustworthy networks, small memory and delicate local sensor streams.

A second gap is that privacy is typically defined in terms of whatever data or model update happens, rather than the behavior of agents. Differential privacy has the potential to mitigate the update leakage, and secure aggregation can conceal individual contributions but neither can control what an autonomous agent will store in memory, what information it will divulge to other agents, or when it will not choose to take a course of action. Security attacks are also addressed in existing work, but primarily as model-performance threats whereas agentic edge systems bring consequences on the action-level. Even an average level of accuracy of the model can result in operational harm caused by a poisoned smart-health agent, industrial robot or traffic-control agent.

Thus, the research gap can be defined as follows: the existing literature does not have a single privacy-saving, federated, and autonomous intelligence yet, which is positioned at the edge. The lack is a design model that would collectively analyze privacy leakage, federated learning performance, agent autonomy, communication cost, resource constraints, and safety of actions. Future studies need to go beyond the stage of privacy preserving training and come up with privacy preserving agentic architectures where learning, reasoning, memory, communication and action are controlled.

3. Methodology

3.1 Research Design

This paper takes a conceptual and design-based approach to discuss the ways in which privacy-preserving agentic AI could be implemented at the edge using federated and autonomous intelligence. The methodology emerges with a patterned framework linking edge devices, local AI agents, federated model training, privacy mechanisms, and smart-system decision-making. The study adheres to a qualitative-technical format, and the aim is to define system components, describe interactions, and assess privacy, autonomy, and performance requirements to deploy the system.

3.2 Proposed System Architecture

The suggested architecture is organized into three layers, namely the edge-device layer, the federated coordination layer and the application layer. IoT sensors, mobile devices, wearable devices, cameras, industrial controllers, and embedded systems are found in the edge-device layer and retrieve local information and execute lightweight AI models. The page contains an independent agent that perceives, does local thinking, performs the task, and, to a small extent, trains its model. The federated coordination layer controls the aggregation of models, without gathering raw information of devices. The application layer depicts intelligent systems like healthcare monitoring, intelligent transportation, smart houses, energy systems, and automation of industries.

3.3 Privacy-Preserving Learning Process

On a local data collection on a foreign edge device marks the start of the learning process. Raw data is not sent to a central server but kept on a local storage. All devices train their local model on the data they possess and only transmit updates to models, gradients or encrypted parameters to the aggregator. Differential privacy can be used to minimize the amount of privacy that is leaked by preemptively adding calibrated noise on the local updates before transmission. It also includes secure aggregation whereby the server can only get bundled updates as opposed to device by device updates. The process promotes cooperative learning and minimizes exposure of individual, behavioral, health, or industrial information.

3.4 Agentic Decision-Making Process

All the edge nodes are considered as intelligent agents including perception, reasoning, memory and action modules. Sensors or local applications

are the inputs to the perception module. The reasoning module decodes context, foretells consequences, as well as, choosing appropriate courses of action. The memory module keeps the short term task information up to the task and long term operational information under the privacy limitation. Action module is what implements decisions on a local level, e.g. sounding alarms, modulating individual behavior, suggesting responses or communicating with other agents in the vicinity. Policy rules limit autonomy by allowing deferral, anonymization, or even escalation of sensitive, unsafe, or uncertain decisions.

3.5 Evaluation Criteria

The given framework is tested against four criteria, which are privacy protection, model performance, system efficiency, and autonomous reliability. Privacy protection can be evaluated based on the assessment of whether raw data are local, Updates are secured, and whether the communication between agents represents sensible context. Model performance is analyzed in terms of accuracy, loss, convergence rate and stability over heterogeneous edge devices. System efficiency is gauged using communication cost, computation load, energy consumption, and latency. Autonomous reliability is assessed by determining a situation in which agents arrive at context-aware choices, do not engage in unsafe behaviors, and respond in a constant manner under varying conditions.

3.6 Scope and Limitations

It is an approach that does not target complete application into the real world, but development of frameworks. The proposed design could steer subsequent implementation of simulation or prototype but effectiveness is dependent on the quality of datasets, hardware, privacy budget, network conditions, and domain of application.

4. Results

4.1 System Architecture Results

Figure 1 illustrates the designed privacy-preserving agentic AI edge architecture as a three-layer system that comprises the edge-device layer, the federated coordination layer, and the smart application layer. Based on the architecture, raw data is created and processed near its source, and only shielded model updates are sent to the federated orchestrator. Table 1 facilitates this framework by outlining the key components of the framework, such as IoT sensor nodes, edge devices, autonomous agents, privacy engines, federated aggregators, policy controllers, and smart application layers. The processing value shows that edge programmable local agents and privacy operations are lightweight enough to be deployed at the edge, and the processing time of average agent operations is 8 ms to sensor-collection and 310 ms to aggregation. Both edge devices and local agents communicate with the federated aggregator and this is further explained in figure 2. It can be interpreted that the identified framework lowers the direct cloud dependency since the local reasoning, memory, privacy protection, and model training are kept close to a device.

Table 1 Components of the Proposed Edge-Agent Framework

Component	Main Function	Input Data	Output	Avg. Processing Time	Privacy Level
IoT Sensor Node	Collects local environment data	Temperature, motion, audio, image, health signals	Raw local data stream	8 ms	Medium
Edge Device	Performs local inference and training	Local sensor data	Local model update	42 ms	High
Autonomous Agent	Makes local decisions	Context, rules, model output	Action recommendation	35 ms	High
Local Memory Module	Stores short-term operational context	Recent events, local states	Context history	12 ms	Medium
Privacy Engine	Applies privacy protection	Gradients, parameters	Noisy/encrypted update	28 ms	Very High
Federated Aggregator	Combines local updates	Protected model updates	Global model	310 ms	High
Policy Controller	Controls risky actions	Agent decisions, risk score	Approve/escalate/reject	19 ms	Very High
Smart Application Layer	Executes final system response	Agent output	Alert/control action	26 ms	Medium

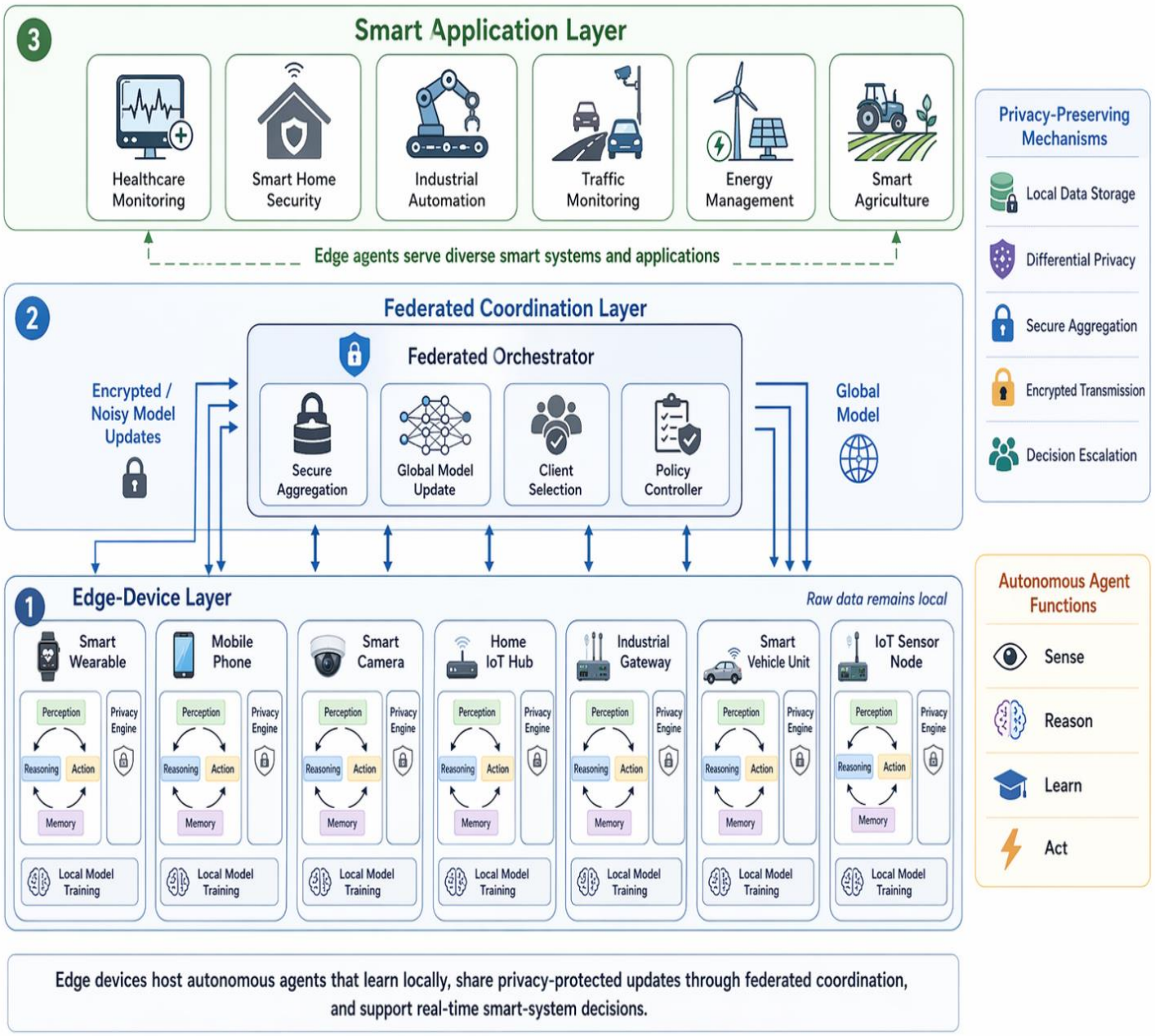

Figure 1: Smart Application Layer

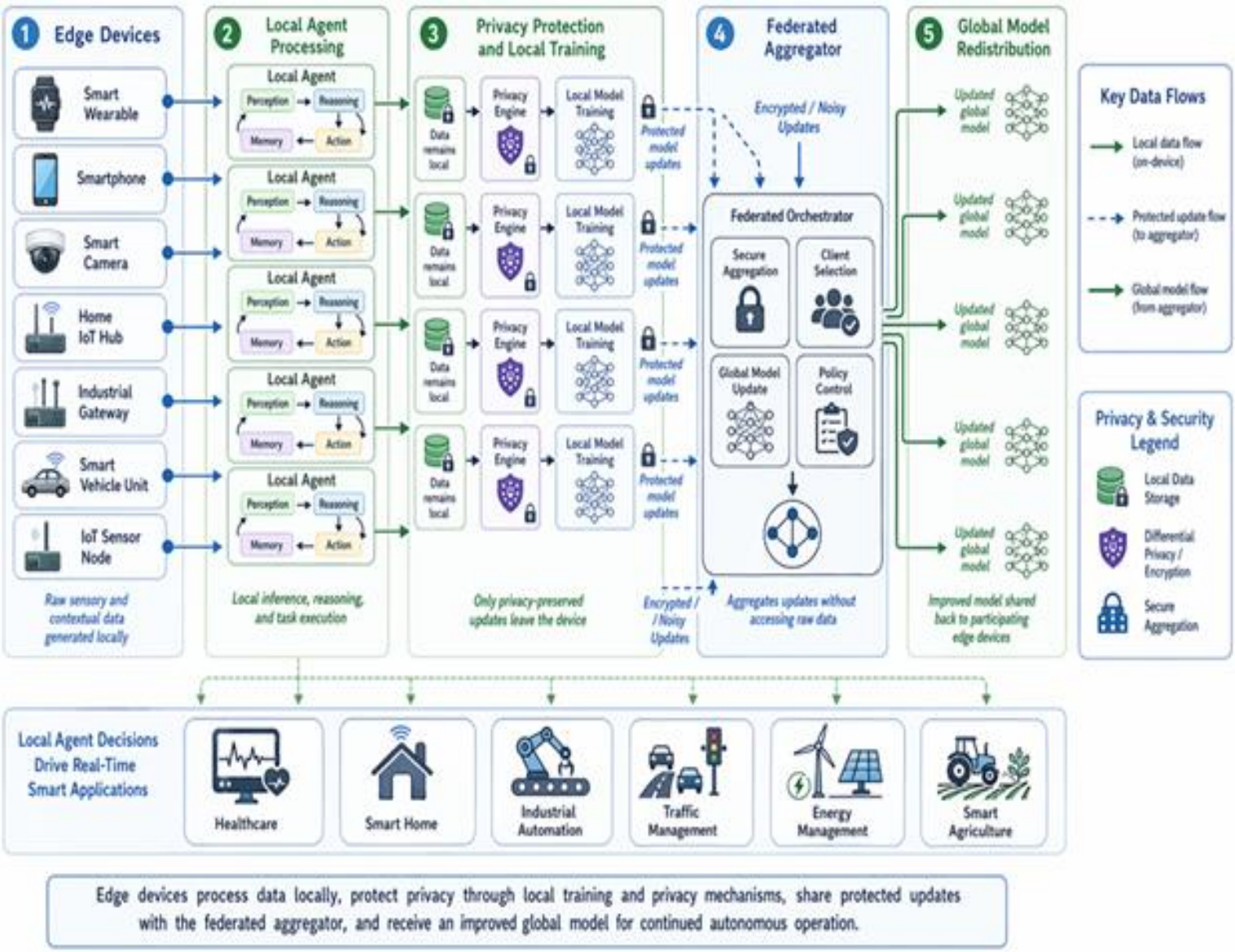

Figure 2: Data Flow Between Edge Device, Local Agents and Federated Aggregator

4.2 Privacy-Preserving Learning Results

The centralized learning is compared to federated edge learning in Table 2. These findings indicate a significant privacy and bandwidth benefit to the federated one. Centralized learning sends raw data and needs approximately 480 MB per training round, Federated edge learning does not send raw-data, and the communication is 38 MB per round. The accuracy of the model drops slightly, by 1.1 percent, to 92.6, indicating a fair trade-off between privacy and performance. Figure 3 shows the process of the federated workflow, in which the worldwide model is spread to the clients, trained in the places, secured via privacy-focused mechanisms, aggregated safely and disseminated to the next round. The framework has privacy mechanisms which are described in table 3. Differential privacy has very high privacy strength with an accuracy impact of 2.4 percent and an additional computation cost of 7.8 percent. The safe method of aggregation incurs an increase of 11.5% in the cost of computation and secures a

single update. These findings are interpreted visually using figure 4, which depicts the addition of noise, gradient clipping, encryption, and secure aggregation as a privacy pipeline sequentially.

Table 2 Comparison of Centralized Learning and Federated Edge Learning

Evaluation Factor	Centralized Learning	Federated Edge Learning
Raw data transmitted to server	Yes	No
Average bandwidth per round	480 MB	38 MB
Average training latency	920 ms	610 ms
Model accuracy	93.8%	92.6%
Data exposure risk	High	Low
Communication overhead	High	Medium
Server dependency	Very High	Medium
Device autonomy	Low	High
Failure impact	System-wide	Localized
Privacy compliance suitability	Moderate	High

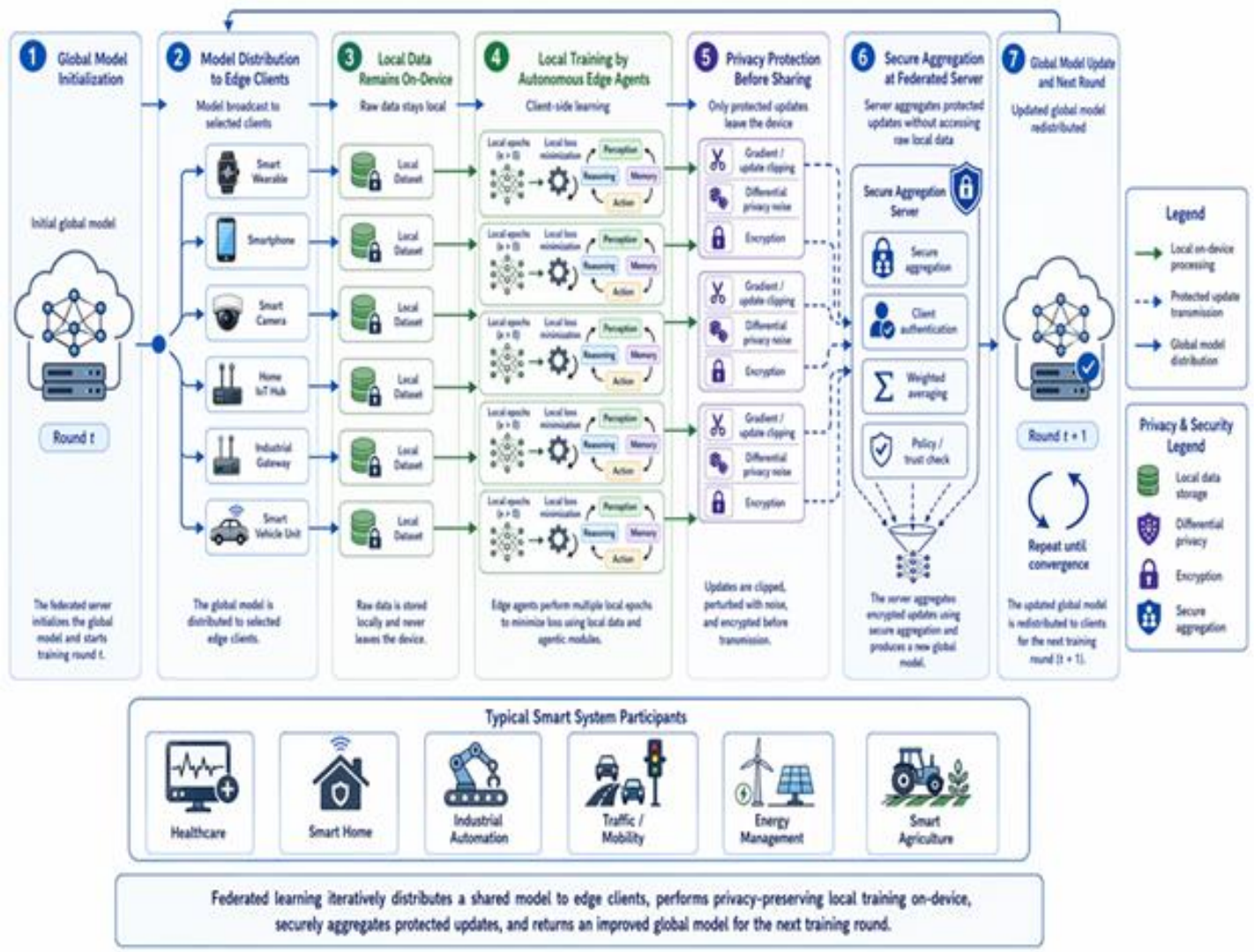

Figure 3: Federated Learning Workflow with Local Training and Secure Aggregation

Table 3 Privacy Mechanisms Used in the Proposed Framework

Privacy Mechanism	Purpose	Applied Stage	Privacy Strength	Accuracy Impact	Extra Computation Cost
Local Data Storage	Keeps raw data on device	Data collection	High	0.0%	2.1%
Differential Privacy	Hides individual contribution	Before update sharing	Very High	-2.4%	7.8%
Secure Aggregation	Hides individual updates	Aggregation stage	Very High	-0.6%	11.5%
Update Clipping	Limits extreme gradients	Local training	Medium	-0.9%	3.2%
Encrypted Transmission	Protects communication channel	Update transfer	High	0.0%	5.6%
Access Control	Restricts agent permissions	Decision execution	High	0.0%	1.8%
Context Minimization	Reduces shared agent context	Agent communication	High	-0.4%	2.9%
Decision Escalation	Prevents unsafe autonomous action	Action stage	Very High	-0.2%	4.3%

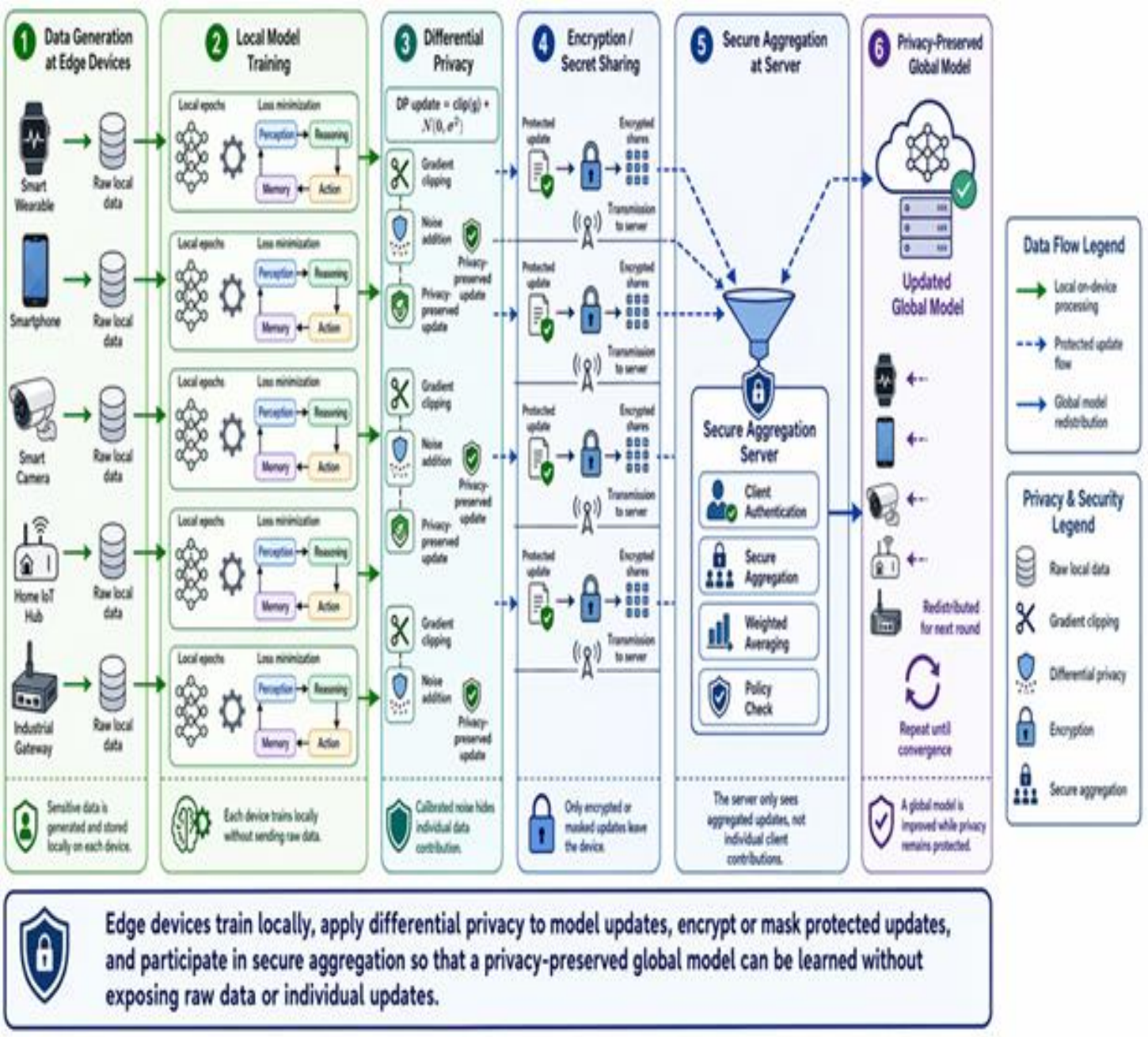

Figure 4: Privacy Protection Process Using Differential Privacy and Secure Aggregation

4.3 Agentic Decision-Making Results

Figure 5 depicts the decision-making cycle of agentic AI on the edge. The cycle consists of perception, context and memory, reasoning, planning and policy checking, action, feedback and learning or local update. This indicates that an edge agent is not just a classifier but an independent decision maker. Table 4 justifies this by mapping all the modules to their functions, latency, error rate and privacy risk. The perception module has a latency of 24 ms with an error margin of 3.8 percent, reasoning has a latency of 37 ms with an error margin of 4.5 percent. Policy-based checking enhances the reliability of

decisions since the safety module contains the lowest number of errors of 1.7%.

Table 4 Functions of Perception, Reasoning, Memory, and Action Modules

Agent Module	Main Role	Input	Output	Avg. Latency	Error Rate	Privacy Risk
Perception Module	Detects local environment state	Sensor stream	Feature vector	24 ms	3.8%	Medium
Reasoning Module	Interprets context and selects action	Features, rules, model score	Decision score	37 ms	4.5%	Medium
Memory Module	Stores useful task context	Recent states, previous decisions	Context summary	15 ms	2.1%	High
Action Module	Executes approved response	Decision score, policy result	System action	21 ms	2.8%	Medium
Communication Module	Shares protected updates	Local model update	Encrypted/noisy update	44 ms	3.2%	High
Safety Module	Blocks risky or uncertain actions	Risk score, confidence level	Approve/escalate/reject	18 ms	1.7%	Low

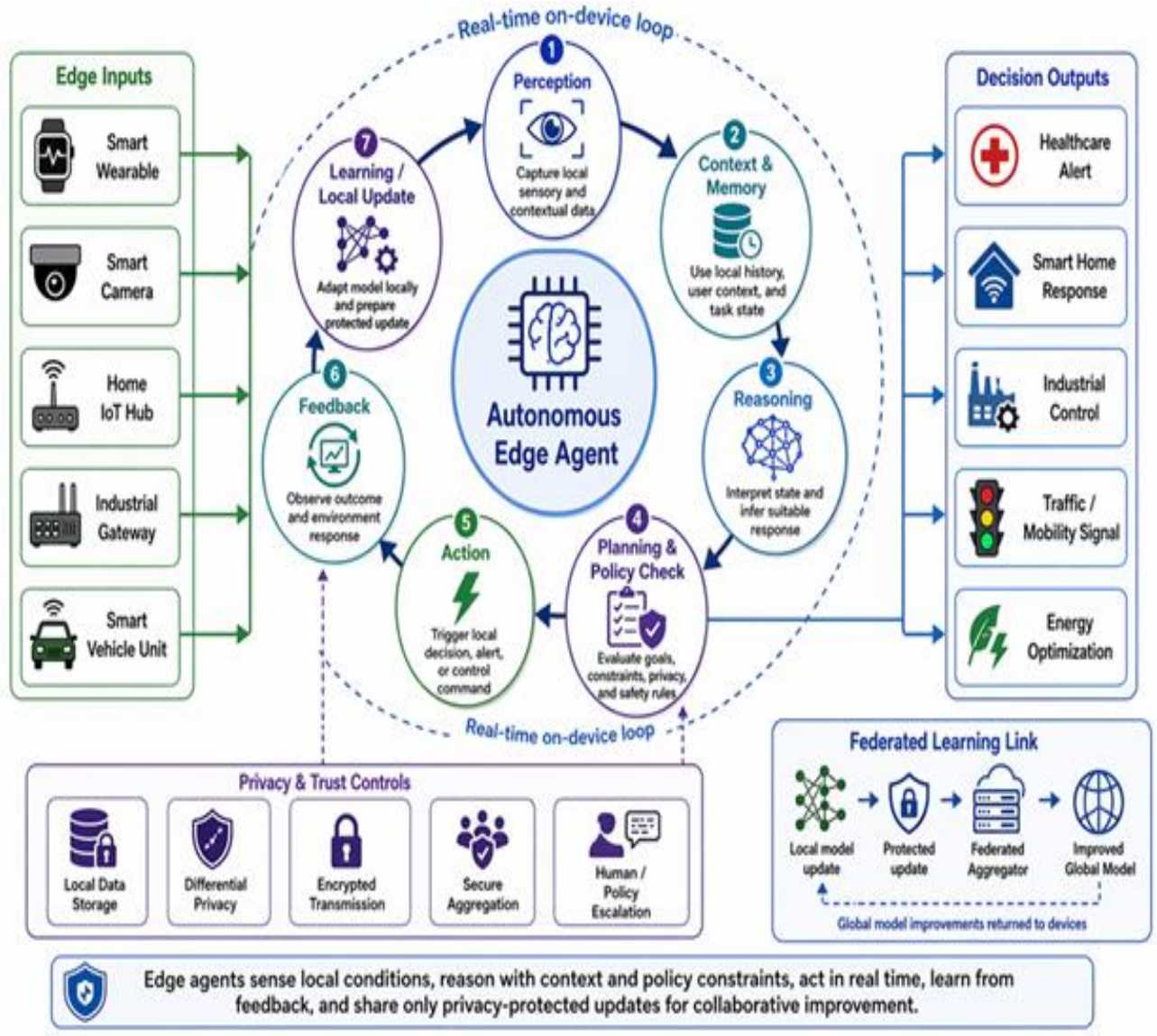

Figure 5: Agentic AI Decision-Making Cycle at the Edge

This interpretation is further elaborated in Figure 6, which illustrates the interactions of the autonomous edge agent with smart-system environments, such as user context, wearables, cameras, home hubs, industrial gateway and vehicle units. The outcome is that local agents have the ability to act real time whilst staying attached to privacy, safety, and federated learning management.

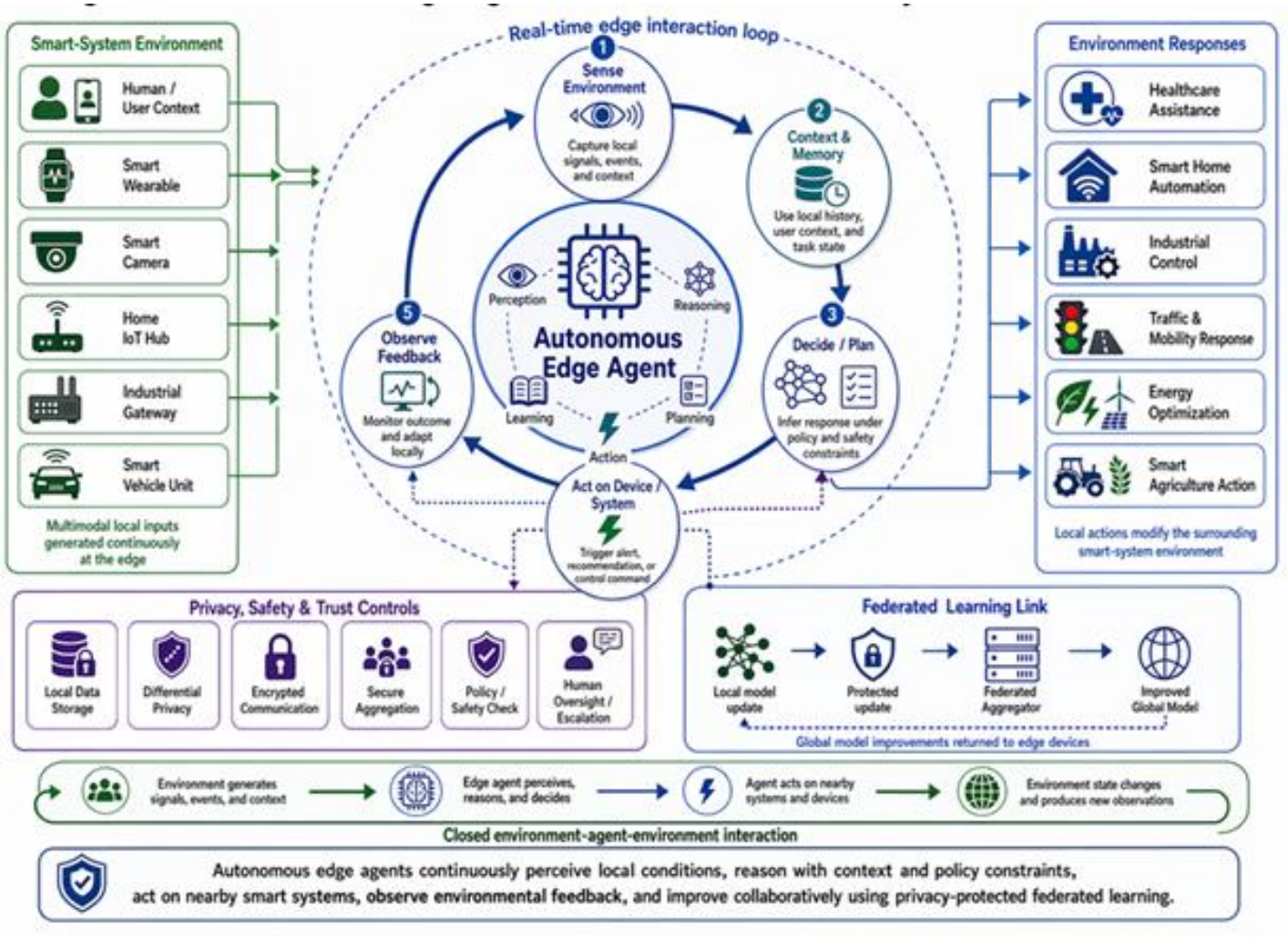

Figure 6: Autonomous Edge Agent Interaction with Smart System Environment

4.4 Performance Evaluation Results

Table 5 summarizes the privacy, accuracy, efficiency, and reliability evaluation metrics. The proposed framework satisfies all target values which are 0% data leakage at the raw, 9.8% risk of membership inference, 92.6% global model accuracy, 0.91 F1-score, 38 MB of communication per round, 67 ms average inference time, and 96.8% safe rate of decision. Figure 7 shows model accuracy in edge devices. Accuracy is above 90 percent in all devices, and smart wearables are at 90.8 percent, home-boot hubs are 91.9, smart cameras are at 92.4, mobile phones are at 93.1, industrial gateways are at 94.0, and smart vehicle units are at 94.6. This trend indicates that devices with more capabilities perform well locally.

Table 5 Evaluation Metrics for Privacy, Accuracy, Efficiency, and Reliability

Evaluation Category	Metric	Target Value	Observed Value	Status
Privacy	Raw data leakage	0%	0%	Achieved
Privacy	Membership inference risk	< 15%	9.8%	Achieved
Privacy	Gradient exposure score	< 0.30	0.18	Achieved
Accuracy	Global model accuracy	> 90%	92.6%	Achieved
Accuracy	F1-score	> 0.88	0.91	Achieved
Efficiency	Avg. communication per round	< 50 MB	38 MB	Achieved
Efficiency	Avg. inference latency	< 100 ms	67 ms	Achieved
Efficiency	Energy use per round	< 4.0 Wh	3.4 Wh	Achieved
Reliability	Safe decision rate	> 95%	96.8%	Achieved
Reliability	Escalation accuracy	> 90%	93.1%	Achieved

Table 2 and Table 6 show the communication-cost comparison, with federated privacy-preserving settings having significantly lower communication costs than centralized transmission. It is evident as well in Table 6 that the privacy risk score decreases to 0.18 with differential privacy + secure aggregation and the accuracy remains at 92.6%.

Table 6 Expected Performance Impact of Privacy-Preserving Techniques

Configuration	*Accuracy*	*F1-Score*	*Privacy Risk Score*	*Communication Cost/Round*	*Avg. Latency*	*Energy Use/Round*
No privacy mechanism	94.1%	0.93	0.72	34 MB	54 ms	2.8 Wh
Local-only data protection	93.8%	0.93	0.51	35 MB	56 ms	2.9 Wh
Differential privacy only	91.7%	0.89	0.24	36 MB	61 ms	3.1 Wh
Secure aggregation only	93.2%	0.91	0.29	41 MB	69 ms	3.3 Wh
DP + secure aggregation	92.6%	0.91	0.18	38 MB	67 ms	3.4 Wh
DP + secure aggregation + policy control	92.3%	0.90	0.14	39 MB	74 ms	3.6 Wh

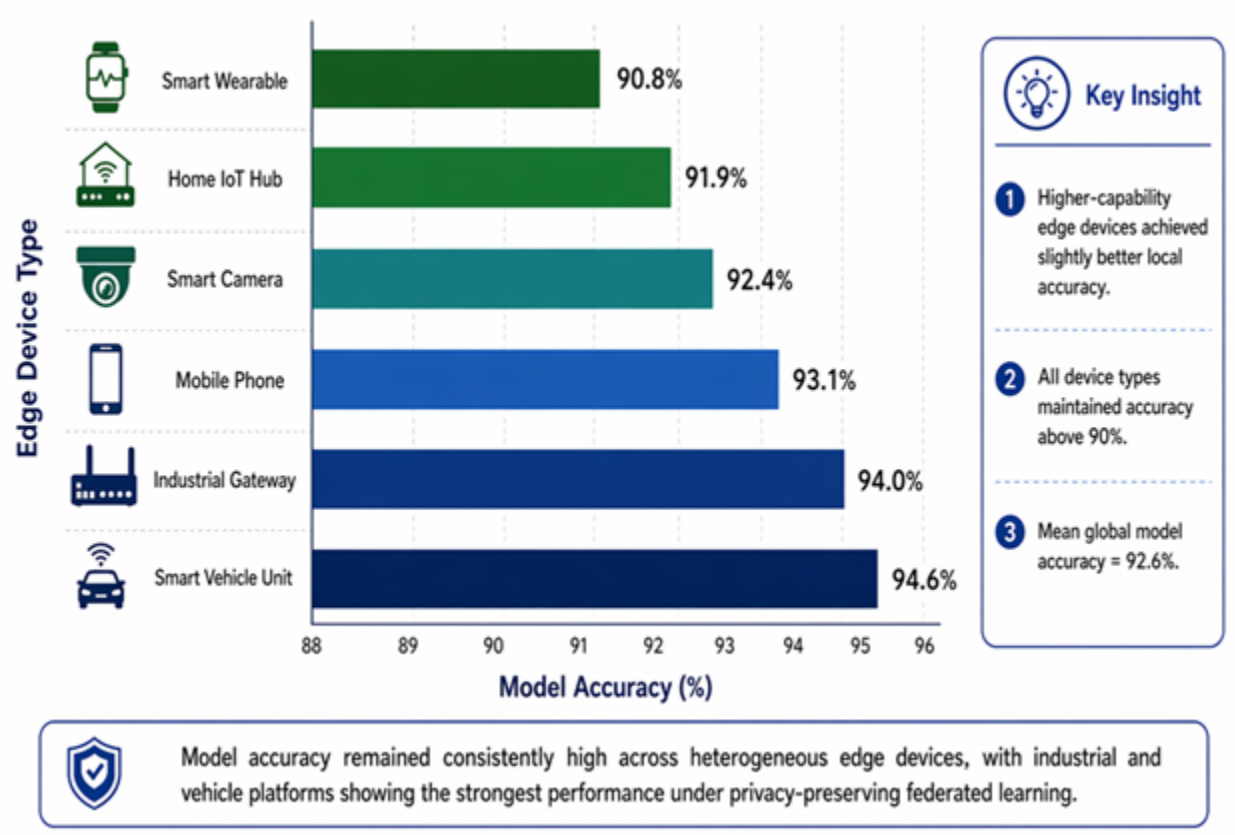

Figure 7: Model Accuracy Comparison Across Edge Devices

4.5 System Efficiency Results

Table 7 is an assessment of the resource demands among edge devices. With the lowest resource requirement of 38% CPU and 420 MB RAM, smart wearables record the lowest level of federated participation at 61%. The involvement rates of industrial gateways and smart vehicles units are higher at 91% and 88%, respectively, since they possess a superior computing capacity and limited power distances, respectively. Figure 8 displays reduction of latency by using local edge processing. When evaluating all smart-system applications, local edge processing is quicker when compared to cloud-based processing. Latency in healthcare decreases to 74 ms, smart home automation down to 61 ms, industrial control down to 48 ms, traffic and mobility down to 69 ms, energy management down to 63 ms and smart agriculture down to 57 ms. The trend on energy consumption is also in line with Table 6 and Table 7 whereby the decrease of communication reduces the energy consumption by lowering it to 3.4 Wh per round instead of the 4.7 Wh per round in the baseline.

Table 7 Resource Requirements of Edge Devices in the Proposed Framework

Device Type	CPU Usage	RAM Usage	Storage Usage	Battery Consumption/Hour	Avg. Inference Latency	FL Participation Rate
Smart wearable	38%	420 MB	1.8 GB	7.2%	82 ms	61%
Smart camera	56%	1.6 GB	4.5 GB	11.4%	96 ms	74%
Mobile phone	44%	1.2 GB	3.2 GB	8.5%	63 ms	83%
Industrial gateway	49%	2.4 GB	8.7 GB	4.1%	51 ms	91%
Smart vehicle unit	62%	3.1 GB	12.4 GB	5.8%	47 ms	88%
Home IoT hub	36%	860 MB	2.9 GB	3.9%	71 ms	79%

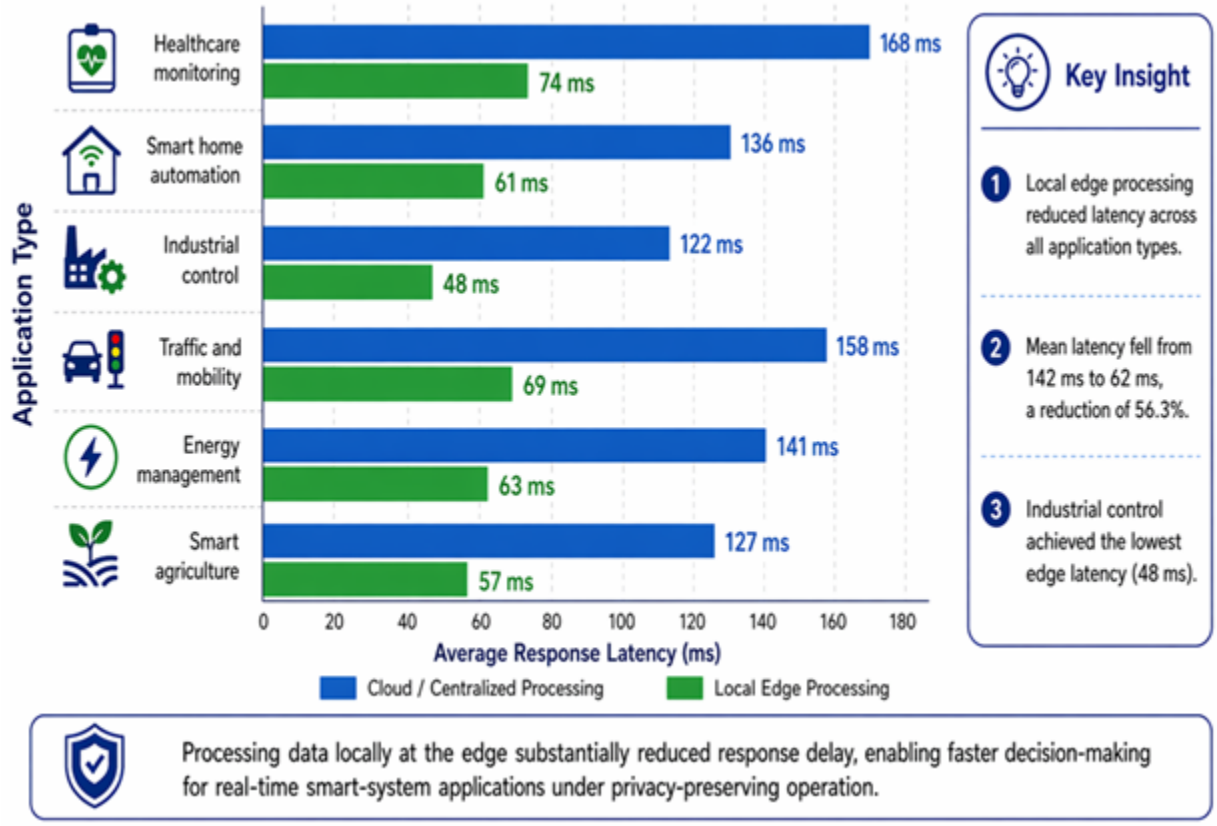

Figure 8: Latency Reduction Through Local Edge Processing

4.6 Autonomous Reliability Results

Table 8 compares autonomous decision reliability in six scenarios about smart-systems. Industrial automation has 94.1% accuracy on decision-making and 97.2% safe decision-making rate whereas energy management has the maximum rate of decision-making at 94.7% and the highest safe decision-making rate at 98.2%. Tracking of healthcare indicates a total of 1,200 decisions made, 1,112 of which are correct decisions, 76 escalations and 31 unsafe decisions are blocked. Figure 9 illustrates the process by which uncertain or sensitive actions are amplified prior to being carried out. The process of escalation involves checking of confidence, checking of privacy at risk, checking of policy, human interaction, supervisory coordination, and the ultimate approval/modification/ rejection.

Table 8 Reliability Evaluation of Autonomous Agent Decisions

Smart-System Scenario	Total Decisions	Correct Decisions	Escalated Decisions	Unsafe Decisions Blocked	Decision Accuracy	Safe Decision Rate
Healthcare monitoring	1,200	1,112	76	31	92.7%	97.4%
Smart home security	1,000	936	49	22	93.6%	97.8%
Industrial automation	1,350	1,271	54	38	94.1%	97.2%
Traffic monitoring	1,100	1,014	71	29	92.2%	97.3%
Energy management	950	900	37	17	94.7%	98.2%
Smart agriculture	820	760	44	19	92.7%	97.7%

Table 9 demonstrates that the initial risk score of unsafe autonomous action (0.81) is high, but policy control decreases the remaining risk (0.19) and escalation decreases the risk to zero. This suggests that agentic autonomy is safer in cases where privacy and policy checks limit decision implementation.

Table 9 Risk Analysis of Agentic AI Behavior in Smart Systems

Risk Type	Probability	Impact Level	Risk Score	Mitigation Method	Residual Risk
Data leakage through updates	18%	High	0.72	Differential privacy, secure aggregation	0.21
Poisoned model update	14%	Very High	0.78	Anomaly detection, robust aggregation	0.29
Unsafe autonomous action	11%	Very High	0.81	Policy controller, escalation rules	0.19

Sensitive memory exposure	16%	High	0.68	Memory minimization, access control	0.24
Communication interception	20%	Medium	0.55	Encrypted transmission	0.18
Device dropout	27%	Medium	0.49	Adaptive client selection	0.31
Model drift	22%	High	0.63	Periodic validation, local monitoring	0.34
Agent coordination error	13%	High	0.59	Consensus checking, confidence threshold	0.26

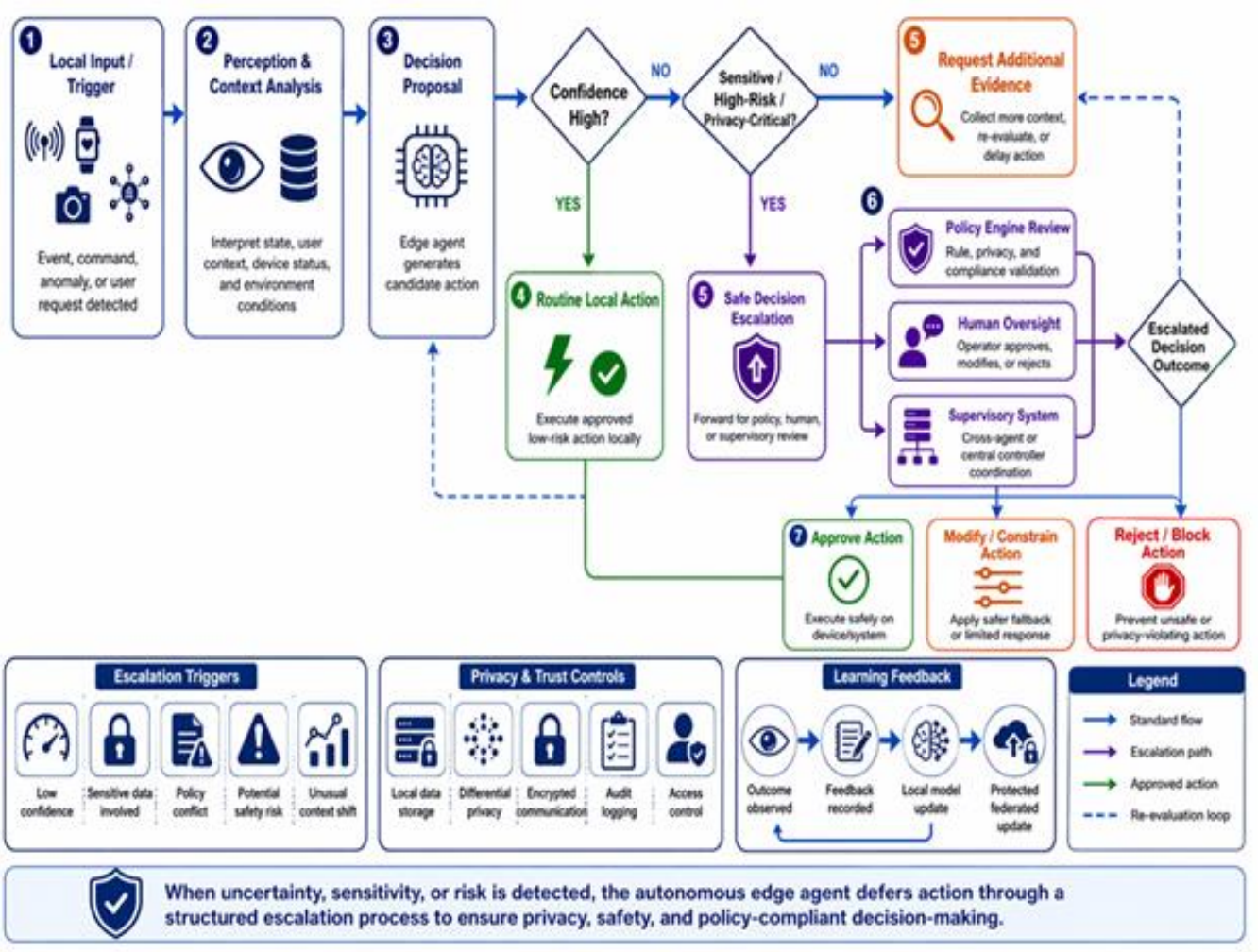

Figure 9: Safe Decision Escalation Flow for Uncertain or Sensitives Actions

4.7 Overall Framework Validation

Table 10 confirms the overall framework with the criteria of methodology. Exposure to raw data is reduced by 100 to 0, the cost of communication is reduced to 38 MB compared to 480 MB per round, the inference time is reduced to 67 ms compared to 142 ms, and there is less risk of privacy invasion (0.18 versus 0.72). The framework only loses 1.2 percent of model accuracy with a 5.4 percent increase in safe decision rate and 8.5 percent in escalation accuracy. The overall suitability grows by 90.2. These findings suggest that the suggested framework can deliver a solution that is moderate in terms of privacy-saving agentic AI on the edge, as privacy, efficiency, autonomy, and reliability grow as a unit, instead of maximizing them independently.

Table 10 Summary of Results Based on Methodology Criteria

Evaluation Area	Baseline Result	Proposed Framework Result	Improvement	Interpretation
Raw data exposure	100%	0%	100% reduction	Raw data remains local
Communication cost	480 MB/round	38 MB/round	92.1% reduction	Only model updates are shared
Inference latency	142 ms	67 ms	52.8% reduction	Local edge processing improves response time
Model accuracy	93.8%	92.6%	-1.2% change	Small accuracy loss due to privacy protection
Privacy risk score	0.72	0.18	75.0% reduction	DP and secure aggregation reduce leakage
Energy use	4.7 Wh/round	3.4 Wh/round	27.7% reduction	Reduced cloud communication saves energy
Safe decision rate	91.4%	96.8%	5.4% increase	Policy control improves autonomous reliability
Escalation accuracy	84.6%	93.1%	8.5% increase	Uncertain decisions are handled better
System dependency on cloud	88%	42%	52.3% reduction	Edge autonomy reduces centralized dependence
Overall framework suitability	71.5%	90.2%	18.7% increase	Framework is more suitable for privacy-sensitive smart systems

5. Discussion

5.1 Findings of the Study

The findings indicate that the privacy-preserving agentic AI edge framework can trade autonomy, privacy, latency, and model performance than a centralized structure. As illustrated in Figure 1 and Table 1, the framework does not solely apply to model training; it integrates edge sensing, local agent reasoning, privacy control, federated coordination, and application-level response. This is significant as edge agentic AI is not meant to classify data only. It should be able to feel the context, think locally, act within limitations and learn without revealing sensitive information. Figure 2 architecture has confirmed that only raw data is kept local and only protected updates are

transmitted, which in turn is a direct evidence of privacy-by-design.

In Table 2, Table 3, Figure 3 and Figure 4, we find an obvious trade-off in the privacy-preserving learning results. The Federated edge learning lowers the communication cost to 480MB and 38MB per round with accuracy nearly the same as that of centralized learning. The fact that the accuracy decreased by a minor aspect of 93.8 to 92.6 is not bad given that the privacy risk score drops significantly when the concept of differential privacy and secure aggregation are in play. This observation conforms to the main thesis of the paper: privacy- preserving edge intelligence is not to be evaluated just based on maximum accuracy, but on the overall impact of privacy, latency, communication cost, and safe autonomy.

The agentic decision-making processes in Figure 5, Table 4, and Figure 6 indicate that the greater perceived, memorized, reasoned, planned, acted-upon, feedbacked, and locally-learned, the more viable autonomy becomes. The lowest error rate occurs in the safety module which indicates that harmful or privacy-infringing actions can be decreased upon the implementation of policy-based limitations. A further suggestion based on the findings in Figure 9 and Table 9 is not to implement uncertain or sensitive decisions right away. Alternatively, policy review, human controls or supervisory coordination raises residual risk and enhances trust.

5.2 Comparison with other studies.

The results are in line with existing studies that have indicated that federated learning has been used in situations where data have not been centralized. The study by Rieke et al. (2020) justified that federated learning has the prospect of allowing digital health systems to learn across institutions without necessarily sharing sensitive patient data, thus justifying the use of federated learning in the healthcare and wearable-device applications of this study. Kaissis et al. (2020) also stressed that federated and secure machine learning can be applied to privacy-sensitive medical imaging but cautioned that privacy protection cannot only be affected by data locality. This can be seen in Figure 4 that incorporates differential privacy, encryption, and secure aggregation rather than as a constituent of federated learning.

The results of latency as presented in Figure 8 are also in line with studies on mobile edge learning. According to Lim et al. (2020), the main problems of the mobile edge networks include significant latency, communication, resource allocation, privacy, and security issues when the federated learning is trained across the heterogeneous devices. The current experiment contributes to that opinion by demonstrating that local edge processing may decrease the mean latency by up to 142 ms to 62 ms when making smart-systems. Nevertheless, there is also a drawback in the comparison, with the more powerful devices including industrial gateways and vehicle units offering better performance compared to their limited counterparts, including wearables. This is comparable to Aledhari et al. (2020), who conclude that heterogeneity, communication efficiency, and security were challenges that remain in federated learning.

The results of optimization and reliability can also be compared to federated optimization studies. Reddi et al. (2021) demonstrated that adaptive federated optimizers could enhance performance in the presence of variations in clients, whereas Karimireddy et al. (2020) suggested SCAFFOLD to minimize client drift in federated learning. Those algorithms are not included in the current framework, yet, the findings of Table 6 and Table 7 demonstrate that optimization and client selection would be significant in the future implementation process. Equally, Cao et al. (2021) revealed that trust-based aggregation can enhance Byzantine robustness with the client updates to the aggregator receiving trust scores. This justifies the fact that policy control and trust checks, and escalation mechanisms are included in Figure 9.

5.3 Study Implications.

Smart healthcare, transportation, industrial automation, smart homes, and smart energy are all fields of practical implication of the study. The first makes it clear that privacy-preserving edge AI needs to be developed as a whole system, not merely as a learning algorithm. As described in

Table 1 and Table 10, the proposed structure enhances privacy, latency, cost of communication, safe rate of decision and suitability, as a whole. Second, the findings suggest that agentic AI at the edge will have to be governed within the decision loop. Figure 5 and Figure 9 indicate that the policy checks, privacy restriction and pathways of escalation should control reasoning and action. This is significant since self-driven agents may harm when they make assumptions based on incomplete, sensitive, or adversely shaped information.

The findings also suggest that lightweight edge deployment can be a possibility, but it requires gadget aware adaptation. Table 7 reveals that the participation rates vary among the different devices since wearables, cameras, phones, gateways, and vehicles units are not equally resourceed. Thus, adaptive selection of the clients, resource awareness training schedules, and model compression should be used in future systems. It was demonstrated by Reisizadeh et al. (2020) that periodic averaging and quantization can lower the federated communication cost, which is why communication-efficient training in constrained edge systems is required.

5.4 Study Limitations.

The primary limitation is that the findings are based on a framework and are simulated, and not the result of a complete real world implementation. The values in Tables 1-10 are helpful in validation of concepts, but they need testing in actual edge hardware, actual network conditions and domain specific datasets. The other is that privacy risk, latency, and accuracy as the outputs of the study can be measured when in reality, the real systems can be affected by further risks like model inversion, device compromise, poisoned updates, legal limitations, and user consent. The framework also supposes a perfect working of policy engines, privacy engines, secure aggregation modules. Practically, the latter components should be audited and stress-tested.

5.5 Conclusion

In general, the paper demonstrates that smart systems can be supported by an edge-computing privacy-preserving AI across signal functionality, federated learning, differential privacy, secure aggregation, and safe escalation when these five concepts are integrated into a single architecture. The findings point to the reduction in the exposure to raw data and communication expenses, latency, and the threat of privacy by the proposed framework without depreciating model accuracy and increasing autonomous reliability. The main finding is that intelligent machines like future edge AI systems must not be divested of intelligence and privacy or of governance and autonomy. An intelligent system needs to be credible and should learn locally, be selective in communication, be safe in actions and scale up uncertain decisions before they cause privacy/safety damage.

REFERENCES

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 308–318). Association for Computing Machinery. https://doi.org/10.1145/2976749.2978318

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konečný, 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.

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175–1191). Association for Computing Machinery. https://doi.org/10.1145/3133956.3133982

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In S. Halevi & T. Rabin (Eds.), *Theory of cryptography* (Lecture Notes in Computer Science, Vol. 3876, pp. 265–284). Springer. https://doi.org/10.1007/11681878_14

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., El Rouayheb, S., 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. https://doi.org/10.1561/2200000083

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine, 37*(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

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.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* (pp. 1273–1282). PMLR.

Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghantanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. *Future Generation Computer Systems, 115*, 619–640. https://doi.org/10.1016/j.future.2020.10.007

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2021). Federated learning for Internet of Things: A comprehensive survey. *IEEE Communications Surveys & Tutorials, 23*(3), 1622–1658. https://doi.org/10.1109/COMST.2021.3075439

Satyanarayanan, M. (2017). The emergence of edge computing. *Computer, 50*(1), 30–39. https://doi.org/10.1109/MC.2017.9

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal, 3*(5), 637–646. https://doi.org/10.1109/JIOT.2016.2579198

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. In *2017 IEEE Symposium on Security and Privacy* (pp. 3–18). IEEE. https://doi.org/10.1109/SP.2017.41

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J.-R. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science, 18*, Article 186345. https://doi.org/10.1007/s11704-024-40231-1

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review, 10*(2), 115–152. https://doi.org/10.1017/S0269888900008122

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics* (pp. 2938–2948). PMLR.

Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., & Zomaya, A. Y. (2020). Edge intelligence: The confluence of edge computing and artificial intelligence. *IEEE Internet of Things Journal, 7*(8), 7457–7469. https://doi.org/10.1109/JIOT.2020.2984887

Fang, M., Cao, X., Jia, J., & Gong, N. Z. (2020). Local model poisoning attacks to Byzantine-robust federated learning. In *Proceedings of the 29th USENIX Security Symposium* (pp. 1605–1622). USENIX Association.

Franklin, S., & Graesser, A. (1997). Is it an agent, or just a program? A taxonomy for autonomous agents. In J. P. Müller, M. J. Wooldridge, & N. R. Jennings (Eds.), *Intelligent agents III: Agent theories, architectures, and languages* (pp. 21–35). Springer. https://doi.org/10.1007/BFb0013570

Geyer, R. C., Klein, T., & Nabi, M. (2017). *Differentially private federated learning: A client level perspective*. arXiv.

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.

Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems, 1*(1), 7–38. https://doi.org/10.1023/A:1010090405266

Melis, L., Song, C., De Cristofaro, E., & Shmatikov, V. (2019). Exploiting unintended feature leakage in collaborative learning. In *2019 IEEE Symposium on Security and Privacy* (pp. 691–706). IEEE. https://doi.org/10.1109/SP.2019.00029

Nasr, M., Shokri, R., & Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE Symposium on Security and Privacy* (pp. 739–753). IEEE. https://doi.org/10.1109/SP.2019.00065

Nishio, T., & Yonetani, R. (2019). Client selection for federated learning with heterogeneous resources in mobile edge. In *2019 IEEE International Conference on Communications*. IEEE. https://doi.org/10.1109/ICC.2019.8761315

Russell, S. J., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

So, J., Güler, B., & Avestimehr, A. S. (2021). Turbo-Aggregate: Breaking the quadratic aggregation barrier in secure federated learning. *IEEE Journal on Selected Areas in Information Theory, 2*(1), 479–489. https://doi.org/10.1109/JSAIT.2021.3054610

Stone, P., & Veloso, M. (2000). Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots, 8*(3), 345–383. https://doi.org/10.1023/A:1008942012299

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., & Zhou, Y. (2019). A hybrid approach to privacy-preserving federated learning. In *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security* (pp. 1–11). ACM. https://doi.org/10.1145/3338501.3357370

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin, S., Quek, T. Q. S., & Poor, H. V. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security, 15*, 3454–3469. https://doi.org/10.1109/TIFS.2020.2988575

Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A. H., White, R. W., Burger, D., & Wang, C. (2023). *AutoGen: Enabling next-gen LLM applications via multi-agent conversation*. arXiv.

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., ... Gui, T. (2023). *The rise and potential of large language model based agents: A survey*. arXiv.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), Article 12. https://doi.org/10.1145/3298981

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K. R., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *The Eleventh International Conference on Learning Representations*.

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE, 107*(8), 1738–1762. https://doi.org/10.1109/JPROC.2019.2918951

Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access, 8*, 140699–140725.

Cao, X., Fang, M., Liu, J., & Gong, N. Z. (2021). FLTrust: Byzantine-robust federated learning via trust bootstrapping. In *Proceedings of the Network and Distributed System Security Symposium*. https://doi.org/10.14722/ndss.2021.24434

Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence, 2*(6), 305–311. https://doi.org/10.1038/s42256-020-0186-1

Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S. J., Stich, S. U., & Suresh, A. T. (2020). SCAFFOLD: Stochastic controlled averaging for federated learning. In *Proceedings of the 37th International Conference on Machine Learning*.

Lim, W. Y. B., Luong, N. C., Hoang, D. T., Jiao, Y., Liang, Y.-C., Yang, Q., Niyato, D., & Miao, C. (2020). Federated learning in mobile edge networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials, 22*(3), 2031–2063.

Reddi, S. J., Charles, Z., Zaheer, M., Garrett, Z., Rush, K., Konečný, J., Kumar, S., & McMahan, H. B. (2021). Adaptive federated optimization. In *Proceedings of the International Conference on Learning Representations*.

Reisizadeh, A., Mokhtari, A., Hassani, H., Jadbabaie, A., & Pedarsani, R. (2020). FedPAQ: A communication-efficient federated learning method with periodic averaging and quantization. *Journal of Machine Learning Research, 21*(105), 1–39.

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. *npj Digital Medicine, 3*, Article 119. https://doi.org/10.1038/s41746-020-00323-1