S-AI-IoT: A Sparse Artificial Intelligence Architecture with Hormonal Orchestration, Parsimonious Agent Activation, and Symbolic Memory for Adaptive, Secure, and Explainable Internet of Things Systems

Said Slaoui

Mohammed V University, Rabat, Morocco

ABSTRACT

The Internet of Things is undergoing a paradigm shift from passive data collection infrastructures toward ecosystems of autonomous, resource-constrained distributed computing entities. Existing IoT intelligence frameworks — whether rule-based, deep reinforcement learning-based, federated, or bio-inspired— share ***six*** structural failures: activation indiscriminateness, opacity, non-stationarity instability, federated learning overhead, absent symbolic memory*, *and absence of reproducible evaluation infrastructure**. No existing approach addresses all six simultaneously while satisfying the energy, connectivity, security, explainability, and scalability constraints of operational IoT deployments.This article introduces S-AI-IoT, a formally grounded, intrinsically parsimonious, hormonally regulated, and natively explainable IoT intelligence framework that is the first to integrate these four properties by architectural design.S-AI-IoT extends the Sparse Artificial Intelligence (S-AI) paradigm through four original contributions. First, a seven-layer bio-inspired modular architecture deployed across node, gateway, and cloud tiers, with formally guaranteed local hormonal stability at each tier, enabling coherent autonomous operation under connectivity loss.Second, five canonical IoT hormones — Sensorin, Connectin, Energexin, Resiliencin, and Normin — whose reaction-diffusion dynamics on the IoT communication graph implement a unified continuous signaling layer jointly encoding sensor quality, connectivity, energy, resilience, and security compliance. Third, a primal-dual parsimonious orchestration mechanism selecting the minimal sufficient agent subset at each decision instant, with Lyapunov stability guarantees — formally established in Article II (Theorem 6.1) under the deployability condition* $\lambda_k > D_k \lambda_{max}(L^*) + \sum_{m \neq k} \gamma_{km} + \rho_k \parallel \chi^* \parallel_\infty$*, verifiable a priori from network parameters alone — and Euler-Maruyama discretization.Fourth, a distributed symbolic engram memory enabling contextual recall, cross-episode behavioral acceleration, and intrinsic explainability at negligible additional cost over the normal decision cycle.A comparative analysis demonstrates that S-AI-IoT is the only framework among those surveyed to simultaneously address all six identified structural dimensions.This article is the first of a three-part series: Article II develops the complete formal mathematical specification and algorithmic implementation; Article III presents the experimental evaluation on the SAI-UT+ IoT testbench across ten operational scenarios with full ablation study.*

KEYWORDS

Internet of Things, sparse artificial intelligence, hormonal orchestration, activation parsimony, reaction-diffusion dynamics, bio-inspired architecture, symbolic memory, engram, explainable AI, federated learning, multi-agent systems, edge computing, energy harvesting, duty cycle management, intrinsic explainability, homeostatic regulation, neuroendocrine systems, S-AI-IoT, S-AI-ROBOTICS, constrained computing

1. INTRODUCTION

1.1. From Passive Sensing to Intelligent IoT

The IoT has evolved from a passive sensor-to-cloud data collection infrastructure into a heterogeneous ecosystem of autonomous distributed computing entities. Connected devices reached 14.5 billion in 2024 (12% CAGR), projected to exceed 44 billion by 2035. The AIoT market is forecast to grow from USD 171.4B in 2024 to USD 896.8B by 2025 at 31.7% CAGR, reflecting a paradigm shift toward AI capabilities embedded directly within IoT nodes.

Academic literature converges on a shared diagnosis: classical IoT architectures — built for predictable behavior, stable connectivity, and centralized authority — are structurally inadequate for modern deployments characterized by device heterogeneity, intermittent connectivity, energy scarcity, evolving security threats, and regulatory demands for transparency. Their limitations are intrinsic, rooted in design choices that privilege global model completeness over parsimony, centralized control over distributed autonomy, and reactive response over anticipatory self-regulation.

This article — the first of a three-part series — introduces S-AI-IoT, a bio-inspired, formally grounded IoT intelligence framework extending the Sparse Artificial Intelligence (S-AI) paradigm and building on the lineage of S-AI-ROBOTICS and the broader S-AI research program. It develops the motivations, state-of-the-art positioning, theoretical and bio-inspired foundations, and global architecture of the framework. The complete formal mathematical modeling with stability guarantees and the algorithmic implementation are the subject of Article II. The experimental validation on the SAI-UT+ IoT testbench is presented in Article III. Rather than reducing the footprint of monolithic architectures post-hoc, S-AI-IoT is architecturally sparse by design: specialized behavioral agents are activated only when the current operational context — encoded in a symbolic hormonal state — justifies it. This principle of context-aware parsimony, inspired by biological neuroendocrine systems, is the architectural kernel of the framework.

1.2. Core Challenges in IoT Intelligence

Device heterogeneity and resource constraints. IoT deployments span microcontrollers with 256 KB RAM to multi-core gateways, using protocols ranging from BLE and Zigbee to 5G NR, and messaging standards including MQTT, CoAP, and AMQP. Deep neural networks requiring hundreds of megabytes and GPU inference are categorically incompatible with battery-powered endpoints. The core issue is always-on computation: an intelligence module running regardless of contextual need consumes energy, memory, and CPU cycles that may be critical for sensing or communication. Activation parsimony must be a first-class architectural property.Dynamic connectivity and link instability. Wireless link quality (RSSI, PDR, latency) fluctuates continuously due to interference, mobility, and duty-cycling. Empirical studies report median PDR below 0.9 in industrial environments, with partitions affecting 15–25% of nodes during peak interference. Classical routing protocols (RPL/6TiSCH) and SDN control planes degrade precisely when connectivity is most needed. An intelligent IoT framework must embed topology awareness natively in its regulatory layer, without relying on a persistent centralized coordinator.Energy constraints and duty cycle management. Nodes in precision agriculture or structural monitoring operate on budgets of a few hundred millijoules per day, potentially for years. Energy harvesting (solar, RF, thermal) adds supply-side uncertainty. Existing duty cycle optimization treats this as a decoupled standalone problem. A coherent framework must integrate duty cycle modulation organically into its self-regulation mechanism — the role assigned to the Energexin hormone in S-AI-IoT.Security and dynamic compliance. The IoT attack surface spans firmware exploitation, replay attacks, Sybil attacks, man-in-the-middle interception, and botnet-based DDoS. Compliance obligations (GDPR, ETSI EN 247 623, IEC 62223) impose continuous runtime behavioral adaptations. DRL-based IDS approaches (e.g., D3O-IIoT) demonstrate adaptive defense potential but remain computationally prohibitive and opaque. Security must be an intrinsic behavioral dimension, not a persistent high-cost inference layer.Explainability and regulatory compliance. The EU AI Act (2024) and standards such as IEC 62061 mandate traceable, auditable automated decisions. Dominant XAI approaches (SHAP, LIME) are post-hoc, requiring model evaluations or approximations 10–100× costlier than base inference — inapplicable on resource-constrained nodes. Post-hoc XAI is an architectural afterthought. A framework grounded in symbolic hormonal state and behavioral engrams delivers intrinsic explainability without post-hoc computational overhead: every decision is traceable to the hormonal context that triggered it, at the cost of a single O(1) engram-writing operation per decision cycle.Scalability and hierarchical orchestration. IoT deployments range from dozens to tens of thousands of nodes, demanding at-most-linear algorithmic complexity and sufficient local autonomy during gateway unavailability. Centralized orchestration creates single points of failure; MARL introduces non-stationarity and policy divergence; Kubernetes imposes memory footprints incompatible with edge nodes. The required response is hierarchical: node-level local intelligence, cluster-level aggregation, and gateway-level orchestration, with behavioral state propagated by diffusion rather than explicit synchronization.

1.3. Structural Inadequacy of Existing AI Paradigms

The above challenges collectively expose five transversal structural failures of dominant IoT AI paradigms:

Activation indiscriminateness: Rule-based systems, DQNs, FL clients, and behavior trees are always-on, executing regardless of context — the primary driver of unnecessary energy consumption.

Opacity: Deep learning IoT models are opaque by construction. Post-hoc explanation is inapplicable in real-time constrained settings; opacity is intrinsic to parametric distributed encoding.

Non-stationarity instability: DQN convergence guarantees require finite, stationary state-action spaces — conditions rarely met in operational IoT, where device failures and link fluctuations continuously alter the observation space.

Federated learning overhead: Non-IID data distributions cause client drift, degrading convergence. Gradient transmission costs per training epoch are comparable to hours of continuous sensing on deeply constrained devices.

Absent symbolic memory: Without episodic memory, each decision cycle starts from an uninformed prior, unable to exploit recurring operational patterns. This is both a performance and an explainability limitation.

1.4. The S-AI Framework: Paradigmatic Foundations

S-AI-IoT is grounded in the S-AI paradigm, built on three principles: activation parsimony (minimal agent subset activated per context), hormonal orchestration (virtual hormone-like signals diffuse through the architecture, modulating agent activation thresholds), and symbolic memory (decisions encoded as behavioral engrams for contextual recall and intrinsic explainability).Prior instantiations include S-AI-GPT (conversational AI), S-AI-NET (network management), S-AI-Cyber (cybersecurity), and S-AI-ROBOTICS — the most complete architectural antecedent, establishing the six-layer pipeline, the five canonical hormones (Sensorin, Connectin, Energexin, Resiliencin, Normin)reaction-diffusion dynamics on a weighted graph, and the primal-dual orchestration framework, all inherited and adapted here for IoT. The IoT transposition is non-trivial: distributed sensing density replaces physical embodiment; battery limitations replace actuator energy constraints; security and compliance replace physical safety requirements; and sense-infer-communicate cycles replace robotic perception-action loops. S-AI-IoT introduces five IoT-specific hormones — Sensorin, Connectin, Energexin, Resiliencin, and Normin — formally specified in Section 6.

1.5. Positioning of S-AI-IoT and Gap Analysis

Table 1 provides a comparative analysis of eight representative categories of existing IoT intelligence approaches against the six structural dimensions identified in Section 1.2. Each criterion is defined as follows: **Parsimony** — agents or modules activated only when contextually justified, not always-on; **Explainability** — decisions traceable without post-hoc approximation at inference time; **Security integration** — security responses architecturally coupled to operational state rather than implemented as an isolated layer; **Energy-awareness** — duty cycle or resource allocation modulated dynamically by operational context; **Symbolic memory** — episodic behavioral records enabling cross-episode recall and convergence acceleration; **Scalability** — algorithmic complexity compatible with large-scale deployment growth.

Table 1. Comparative analysis of IoT intelligence frameworks.

Approach	Parsimony	Explainability	Security Integration	Energy-Aware	Symbolic Memory	Scalability
Rule-Based / Threshold Systems	Partial	High (static)	Low	Low	None	High
Deep Q-Network (DRL)	No	None	Partial	No	None	Low–Medium
Federated Learning	No	Low	Partial	Partial	None	Medium
Behavior Trees / FSM	Partial	Medium	Low	No	None	Medium
Swarm / Ant Colony	No	Low	Low	No	None	High
Container Orchestration	No	Low	Medium	No	None	High (cloud)
TinyML / On-Device AI	No	Low	Low	Partial	None	Medium
Classical MAS	Partial	Low	Low	No	None	Low
S-AI-IoT (proposed)	**Yes**	**Yes (intrinsic)**	**Yes (Normin)**	**Yes (Energexi**	**Yes (Engrams)**	**Yes** [1]

Approach	Parsimony	Explainability	Security Integration	Energy-Aware	Symbolic Memory	Scalability
				n)		

Partial = criterion partially addressed in specific sub-configurations but not architecturally guaranteed as a first-class property. No = criterion not addressed by architectural design. Scores are assigned against the six criteria defined in Section 1.5, based on the representative references cited for each framework category in Section 2; a criterion is marked Yes only when it is guaranteed as a first-class architectural property, not merely achievable through configuration or extension. Container Orchestration covers infrastructure-level platforms (Kubernetes, Docker Swarm) operating at the scheduling layer without behavioral intelligence; included because these platforms are frequently proposed as IoT edge orchestration solutions in the literature [9]. [1] $O(N)$ per orchestration cycle for the Hormonal Engine (N = number of nodes); $O(M \cdot 5)$ for engram retrieval ($M \leq 1\,000$ engrams); knapsack problem tractable for $|A| = 24$ with hormonal significance filter. Full complexity analysis in Article II, Section 3.5.As Table 1 illustrates, S-AI-IoT is designed to address all six structural dimensions simultaneously — a property that none of the surveyed framework categories achieves. The formal mathematical proofs of Article II and the experimental validation of Article III will substantiate this claim rigorously.Among the frameworks surveyed, no approach addresses more than three of the six dimensions: rule-based systems achieve high static explainability and scalability but are activation-indiscriminate, energy-unaware, and memoryless; DRL-based approaches partially address security integration but provide no explainability, no parsimony, and no symbolic memory; federated learning partially addresses energy awareness but carries no symbolic memory, no intrinsic explainability, and no architectural parsimony. Swarm approaches achieve scalability but offer no individual contextual awareness across the remaining five dimensions.This analysis is based on the representative categories retained in Table 1; approaches outside this scope may partially address different subsets of dimensions. Notably, neuromorphic computing platforms (e.g., Intel Loihi, IBM TrueNorth) share S-AI-IoT's event-driven activation parsimony but currently lack hormonal orchestration, symbolic memory, and integrated security-energy trade-offs at the architectural level; Digital Twin-assisted IoT approaches share the property of contextual memory but operate at the platform level without intrinsic behavioral parsimony or hormonal self-regulation. Both constitute directions for future comparative analysis.The structural limitations of existing IoT intelligence frameworks are not independently addressable through incremental improvements to any single existing approach. Each approach satisfies a local subset of requirements determined by its foundational design principles: a DQN cannot be made intrinsically explainable without fundamentally altering its parametric distributed encoding; a federated learning system cannot achieve activation parsimony without architectural reconception of its always-on inference pipeline; a behavior tree cannot embed symbolic episodic memory without extending its Boolean condition model to continuous hormonal state representation. These limitations are intrinsic, not incidental.What is required — and what S-AI-IoT provides — is a unified architectural reconception that integrates activation parsimony, hormonal self-regulation, and symbolic memory as first-class architectural properties from the ground up. This reconception is grounded in a formally specified and stability-proven mathematical framework (Article II) and evaluated on a comprehensive simulation testbench across ten operational scenarios against five baseline approaches with full ablation study (Article III).

1.6. Main Contributions

C1 — S-AI-IoT Architecture. A seven-layer bio-inspired modular architecture spanning constrained IoT nodes, edge gateways, and cluster coordination layers.

C2 — Five Canonical IoT Hormones. Formal specification of Sensorin , Connectin , Energexin , Resiliencin , and Normin , with emission functions, reaction-diffusion dynamics, cross-inhibition structure, and decay laws derived from IoT operational semantics.

C3 — Parsimonious Hormonal Orchestration. A primal-dual constrained mechanism selecting the minimal required agent subset, with Lyapunov stability guarantees and Euler-Maruyama discretization.

C4 — Duty Cycle Modulation via Energexin. Continuous adjustment of each node's active window as a function of , analogous to biological sleep-wake metabolic regulation.

C5 — Symbolic Memory with Extended Engrams. Distributed storage of engrams enabling contextual recall, cross-episode acceleration, and intrinsic explainability.

C6 — SAI-UT+ IoT Testbench (Article III). A simulation testbench specifically designed for bio-inspired IoT intelligence evaluation, covering ten operational scenarios including single-stress and multi-stress combined conditions, with systematic comparison against five baseline approaches (rule-based systems, DQN, federated learning, behavior trees, TinyML) and a full ablation study quantifying the individual contribution of each architectural component. The design and results of SAI-UT+ are presented in Article III.

1.7. Article Organization

The remainder of this article is organized as follows. Section 2 provides a comprehensive review of related work covering six domains — classical IoT architectures, machine and deep learning, multi-agent systems, bio-inspired methods, explainability and security, and the S-AI research series — and synthesizes the six structural gaps motivating S-AI-IoT. Section 3 develops the theoretical and bio-inspired foundations: activation parsimony, cognitive modularity, artificial hormonal signaling, IoT-specific bio-inspirations, symbolic memory, and anti-oscillation mechanisms. Section 4 presents the seven-layer, three-tier global architecture, including the IoT-MetaAgent, the Gland Agent layer, the specialized agent families, the Hormonal Engine, symbolic memory, and the security architecture. Section 5 concludes and positions this article's contributions within the trilogy.Article II of this series develops the complete formal specification of the 24 specialized agents with input-output contracts, the full mathematical modeling of the framework — reaction-diffusion dynamics, Lyapunov stability analysis, primal-dual constrained orchestration, and engram formalism — and the algorithmic implementation with pseudocode for all six core algorithms. Article III presents the complete experimental evaluation on the SAI-UT+ IoT testbench and the general conclusions and future research directions of the trilogy.

2. RELATED WORK

The research landscape relevant to S-AI-IoT spans six interconnected domains: classical IoT management architectures, machine and deep learning, multi-agent systems, bio-inspired methods, explainability and security, and the S-AI framework series. Each domain contributes essential but partial building blocks; their shared structural gaps, synthesized in §2.7, collectively motivate the design of S-AI-IoT.

2.1. Classical IoT Management Architectures

Rule-based and threshold-driven systems [1, 2] represent the historical baseline for IoT intelligence: deterministic, interpretable, and deployable on microcontrollers, but intrinsically static. Their if-then logic cannot adapt to environmental non-stationarity, cannot compose multiple concurrent signals into a coherent situational assessment, and carries no behavioral memory. Their brittleness scales with deployment complexity, as demonstrated across industrial, smart city, and healthcare IoT contexts [3, 4].Cloud-centric architectures introduce unacceptable latency (100–440 ms) and backhaul dependency [5]. Edge computing and fog computing reduce latency and improve local resilience, but address only where computation occurs, not how intelligence is organized within those tiers [5, 6]. An edge node running a monolithic inference engine remains always-on and context-blind [6]. SDN and NFV add programmability to IoT network management but require persistent controller reachability [7]**.** Energy harvesting nodes further compound this challenge through stochastic power availability [8]**,** making always-on operation architecturally untenable.Container orchestration platforms extend microservices patterns to edge nodes but impose memory footprints (100--200 MB for kubelet alone) incompatible with constrained endpoints **[9]**, and provide no behavioral intelligence above the scheduling layer.

2.2. Machine Learning and Deep Learning for IoT

Systematic benchmarking of inference cost across hardware tiers — from cloud servers to edge accelerators and microcontrollers — confirms that inference latency and energy consumption vary by orders of magnitude depending on deployment context [10]. Deep reinforcement learning [11, 12] has become the dominant paradigm for adaptive IoT resource management, with DQN, PPO, and SAC applied to spectrum access, power control, and routing. Despite strong performance in controlled settings, DRL is structurally ill-suited to operational IoT: sample complexity precludes online training; policies degrade under distribution shift inevitable in dynamic IoT environments; and opacity creates regulatory barriers in safety-critical deployments [13]. Inference cost is continuous and irrepressible regardless of contextual necessity.Federated Learning addresses privacy and bandwidth constraints by distributing training across devices, with FedAvg [14] as the reference aggregation protocol [15]. Distributed optimization techniques such as particle swarm optimization have also been explored as coordination mechanisms in decentralized intelligent systems [16]. However, gradient exchange overhead far exceeds LPWAN link capacities (LoRaWAN: 0.3–44 kbps) [17], convergence under non-IID data distributions is unreliable [18], and FL provides no mechanism for orchestrating when and at what cost to run inference — a fundamentally architectural question [15]. Mixture-of-Experts and soft-threshold sparsity [19, 20] share S-AI-IoT's parsimony intuition but require parameter storage and loading infrastructure incompatible with microcontroller-class hardware, and their gating logic provides no interpretable systemic self-awareness. TinyML achieves remarkable compression (int8 quantization, structured pruning, knowledge distillation) enabling sub-millisecond inference on 256 KB MCUs, but compression does not resolve the always-on problem: continuous inference at 1 Hz may consume 10--25% of daily energy budget regardless of context **[48]**. Multi-agent reinforcement learning applied to IoT resource management **[49]** enables decentralized coordination without a central controller, but inherits the non-stationarity instability of single-agent DRL and provides no mechanism for intrinsic explainability or symbolic memory

2.3. Multi-Agent Systems Applied to IoT

Classical BDI-based MAS apply explicit belief-desire-intention reasoning and message-passing coordination (Contract Net, auction) to IoT scenarios including smart buildings and sensor networks [21,22]. Symbolic belief bases exceed constrained node memory, and coordination traffic saturates low-rate wireless links; these architectures are designed for stable agent populations, not the topology churn of real IoT deployments [3]. Swarm intelligence approaches — ant colony optimization, particle swarm, artificial bee colony [23, 24]— achieve scalability and fault tolerance through stigmergic coordination but carry no individual contextual awareness: an ACO agent has no representation of its own energy budget, security posture, or sensor health [25]. Brambilla et al. identify this limitation as fundamental to swarm architectures. Behavior trees offer modular, interpretable behavioral sequencing but operate on discrete Boolean conditions, precluding the graded, continuous modulation of behavioral intensity that dynamic IoT contexts require, and carrying no persistent episodic memory [21].

2.4. Bio-Inspired Approaches in IoT

The vertebrate neuroendocrine system — with its graded hormonal signaling, spatially diffused regulation, cross-inhibitory axes, and tight coupling to behavioral repertoire selection — constitutes the biological model for S-AI-IoT's regulatory layer [26,27]. Lichtner et al. [27] demonstrated neuroendocrine-based behavior generation in autonomous robots; Bonabeau et al. [25] showed that hormone diffusion on a communication graph enables multi-robot coordination without explicit negotiation. These proofs of concept, however, remain confined to single-platform, fully-powered scenarios incompatible with large-scale heterogeneous IoT deployments.Artificial Immune Systems apply clonal selection and affinity maturation to anomaly detection and fault isolation in distributed systems [28, 29]; evolutionary algorithms and bacterial foraging optimization address parameter optimization in wireless sensor networks [26]. These mechanisms are valuable for specific sub-problems but do not constitute unified behavioral architectures: an AIS detector identifies anomalies but provides no orchestration of corrective response; an evolutionary optimizer minimizes a scalar metric but cannot integrate multi-dimensional operational state. Three structural gaps characterize the bio-inspired IoT literature: (i) domain specificity — no unified layer jointly manages sensor quality, connectivity, energy, security, and compliance; (ii) limited formal grounding — qualitative biological analogies abound but rigorous dynamic specifications and stability proofs are rare; (iii) absence of reproducible evaluation benchmarks.

2.5. Explainability, Security, and Energy Awareness

The XAI 2.0 manifesto articulates the fundamental inadequacy of post-hoc explainability for high-stakes IoT deployment [13]: SHAP and LIME require model evaluation counts proportional to feature dimensionality, rendering real-time inference-time explanation on constrained nodes infeasible, and explain individual predictions rather than behavioral strategy [30, 31]. S-AI-IoT implements intrinsic explainability through a semantically interpretable hormonal state vector and an auditable extended engram that requires no post-hoc computation.Roman et al. [32] survey security threats in edge and fog architectures — device compromise, lateral movement, data exfiltration, denial-of-service — and identify the core gap: absence of integrated, adaptive security mechanisms responsive to evolving threats without manual reconfiguration. S-AI-IoT addresses this through the Normin hormone, continuously encoding firmware currency, encryption status, and behavioral deviation, and triggering proportionate security agent activation without network-wide coordination.Sudevalayam and Kulkarni [8] establish the stochastic nature of energy harvesting sources and the consequent requirement for adaptive duty cycle management aligned with energy availability. Transforma Insights [33] extend this perspective to the broader AIoT context, highlighting the need for embedded intelligence capable of adapting to supply-side uncertainty. S-AI-IoT formalizes duty cycle modulation directly as a hormonal output, with cross-inhibition ensuring energy-driven sleep does not compromise safety-critical functions.

2.6. The S-AI Research Series

S-AI-IoT is the IoT instantiation of a coherent research program initiated with the foundational S-AI architecture [34], which establishes activation parsimony, task decomposition, the Pareto 20/80 heuristic, smart hormonal orchestration, and the Hormonal MetaAgent (HMA) as the five theoretical pillars of sparse AI. The S-AI-GPT trilogy demonstrates that large language model behavior can be governed by hormonal orchestration with substantially reduced activation cost [35, 36], and introduces the symbolic engram system for cross-episode learning and behavioral consistency [37]. S-AI-Anti-Hallucination and its triadic extension formalize homeostatic stability under cross-inhibitory hormonal dynamics, providing mathematical foundations directly specialized for IoT diffusion in Section 6 [38, 39]. Other domain-specific S-AI extensions in networking and cybersecurity further motivate the transfer of hormonal orchestration to distributed and adversarial environments. S-AI-EDU confirms the domain transferability of the modular agent typology and symbolic memory, providing the methodological template for Section 5 [40]. S-AI-ROBOTICS serves as the direct architectural patron of S-AI-IoT [41]: it establishes the six-layer pipeline, the five canonical hormones with logistic-saturated emission functions, the reaction-diffusion graph dynamics with cross-inhibitory coupling, and the primal-dual constrained orchestration mechanism with Lyapunov stability guarantees. S-AI-IoT inherits this apparatus entirely and performs a principled domain transfer involving five adaptations: (i) distributed multi-node gland layer replacing robotic perception; (ii) IoT-specific hormone semantics (Sensorin, Connectin, Energexin, Resiliencin, Normin); (iii) dynamic topology with time-varying link weights; (iv) duty cycle modulation as direct hormonal output; and (v) extended engrams including topology state and energy budget as first-class contextual dimensions.

2.7. Identified Gaps and Motivation for S-AI-IoT

The survey above reveals that existing approaches address IoT intelligence through domain-specific, independently designed modules — routing, security, energy, anomaly detection — that cannot represent or jointly modulate the multi-dimensional operational state of a constrained node. Six structural gaps characterize the field:

Gap 1 --- Among existing approaches, no unified continuous signaling layer jointly encodes and modulates sensor quality, connectivity, energy, resilience, and compliance. Gap 2 --- No framework surveyed treats selective, context-conditioned activation as a primary architectural property guaranteed at the design level. Gap 3 --- Post-hoc XAI is computationally infeasible on constrained nodes; none of the reviewed frameworks provides architecturally intrinsic explainability. Gap 4 --- In current IoT intelligence architectures, security management is typically decoupled from energy and connectivity context, preventing principled adaptive trade-offs. Gap 5 --- No constrained IoT framework reviewed provides persistent distributed symbolic memory with contextual retrieval. Gap 6 --- Reproducible simulation benchmarks specifically designed for bio-inspired IoT intelligence remain scarce, impeding systematic comparative evaluation.

3. THEORETICAL AND BIO-INSPIRED FOUNDATIONS OF S-AI-IOT

The design of S-AI-IoT rests on three mutually reinforcing theoretical commitments: elevation of computational parsimony to a primary architectural property, adoption of artificial hormonal signaling as the universal medium of contextual self-regulation, and embedding of symbolic memory as a first-class behavioral persistence mechanism. Each commitment is rooted in convergent evidence from computational neuroscience, cognitive systems theory, distributed control theory, and the formal analysis of resource-constrained distributed systems.

3.1. Parsimony as a Structural Principle for IoT Intelligence

3.1.1. The Always-On Fallacy in Constrained Computing

The dominant design assumption in existing IoT intelligence systems — whether based on deep learning inference, rule-based control, or behavior tree execution — is that the intelligence module is persistently active, consuming computational and energetic resources at a fixed baseline rate regardless of operational context. This always-on fallacy is an architectural convention inherited from cloud and HPC contexts where resources are effectively abundant. Its transfer to constrained IoT constitutes a structural category error: continuous inference at even modest frequencies (1–10 Hz) on battery-powered microcontrollers can consume 10–25% of the daily energy budget independently of whether the environment warrants intelligent response.Biology offers a decisive refutation through the principle of metabolic parsimony. The human brain, despite 86 billion neurons, maintains at any given moment an active firing pattern involving only a small fraction of its neuronal population. Sparse coding simultaneously reduces metabolic cost, increases signal-to-noise ratio, and enhances pattern separability. The implication is profound: selective, context-conditioned activation is not a degraded compromise but a computationally superior strategy.

3.1.2. Formal Grounding: The Principle of Activation Parsimony

The S-AI framework [34] formalizes parsimony through the Pareto 20/80 heuristic: under typical operational conditions, approximately 80% of behavioral objectives are addressable by 20% of available specialized agents. This is grounded in sparse approximation theory [42]: any structured signal can be accurately represented using a small subset of an overcomplete dictionary. In S-AI-IoT, the signal is the node's multi-dimensional operational state, the dictionary is the agent repertoire, and sparsity is enforced by hormonal orchestration selecting the minimal sufficient agent subset.The formal consequence is a constrained agent-selection problem: at each decision instant, the IoT-MetaAgent identifies the subset maximizing behavioral utility while satisfying energetic and computational budget constraints. This is formalized as a primal-dual constrained optimization, with the hormonal state vector as the context parameter shaping the utility landscape. Theoretical guarantees — optimality under convex relaxation, convergence of primal-dual dynamics, and Lyapunov stability — connect the parsimony principle to the rigorous apparatus of nonlinear control theory.

3.1.3. Distinction from TinyML and MoE Approaches

S-AI-IoT parsimony must be distinguished from superficially similar notions in TinyML and Mixture-of-Experts (MoE) literature [19, 20]. TinyML achieves efficiency through model compression within a fixed always-active inference pipeline; MoE through input-conditional expert routing via a neural routing function. Both preserve the always-on assumption at the system level. S-AI-IoT parsimony is architecturally deeper: the activation decision precedes computation entirely, is grounded in a formally specified symbolic state rather than a learned neural function, and is sensitive to energy budget, security posture, and connectivity quality — dimensions wholly outside the representational scope of both TinyML and MoE. The distinction is between computational sparsity within an active pipeline and architectural sparsity in the activation of the pipeline itself.

3.2. Cognitive Modularity in Distributed IoT Environments

3.2.1. Modularity and Functional Encapsulation

Cognitive modularity — organizing intelligence into functionally specialized, semi-independent modules with well-defined input-output contracts — has deep roots in computational neuroscience and robotics [27]. In IoT, this translates into functional encapsulation: each operationally distinct capability (sensor quality monitoring, connectivity management, energy budgeting, fault recovery, security compliance) is encapsulated in a specialized agent with a formally defined activation condition, input specification, output contract, and resource cost. This serves three functions: (1) fine-grained activation control without activating a monolithic intelligence module; (2) formal composability, making the behavior of any combination of active agents predictable; (3) natural explainability, since behavioral output at any moment is the aggregated output of a small, explicitly identified set of active agents each traceable to the hormonal context triggering its activation.

3.2.2. Coordination Without Communication: The Hormonal Coherence Principle

Classical MAS architectures coordinate through explicit message passing (Contract Net, auction-based negotiation, consensus protocols), imposing communication overhead growing polynomially with agent count [22]. In constrained IoT (LoRaWAN: 0.3–44 kbps, Zigbee: 244 kbps), this is architecturally inadmissible. S-AI-IoT resolves this through the hormonal coherence principle: inter-agent coordination is achieved not through direct message exchange but through shared observation of a common hormonal state vector broadcast across the node and diffused through the network graph. Each agent determines its activation eligibility from this shared observation without direct communication with other agents. Coherence of the resulting collective behavior is guaranteed by the structural design of hormonal emission functions, constituting an implicit stigmergic coordination mechanism [25] generalized to continuous-valued hormonal signals.

3.2.3. Agent Typology: Functional Completeness and Minimal Redundancy

The IoT agent typology is governed by two jointly satisfied requirements: functional completeness (the registry covers all operationally relevant behavioral capabilities) and minimal redundancy (no overlap in functional scope creating ambiguity in activation decisions). This balance is achieved through the five-family organization of Section 5 — sensor monitoring, connectivity management, energy management, security and trust, resilience and self-healing — each hormonally coupled to a distinct dimension of the five-dimensional state vector.

3.3. Artificial Hormonal Signaling for IoT Behavioral Regulation

3.3.1. The Neuroendocrine System as Formal Model

The biological neuroendocrine system exhibits seven properties of direct architectural relevance to S-AI-IoT:

Graded and continuous signals — hormone concentrations encode regulatory stimuli with analog precision.

Spatial diffusion — hormones propagate through the bloodstream, implementing broadcast coordination without targeted addressing.Cross-inhibition and cross-activation — competing hormonal axes (HPA: hypothalamic-pituitary-adrenal; HPT: hypothalamic-pituitary-thyroid;

HPG: hypothalamic-pituitary-gonadal) interact through reciprocal inhibition, preventing any single axis from dominating under all conditions.

Temporal integration — effects accumulate and decay over characteristic timescales, providing low-pass filtering of environmental noise.

Inherent stability — maintained through negative feedback circuits constituting biological closed-loop control.

Metabolic efficiency — maintaining circulating hormone concentrations is energetically negligible relative to the coordination enabled.

Interpretable state representation — each hormone encodes a specific regulatory signal with formally specifiable effects.

3.3.2. From Biology to Formal Architecture

The transposition from neuroendocrine biology to the S-AI-IoT hormonal architecture involves four formal correspondences. The biological secretory gland corresponds to the IoT Gland Agent monitoring a local observable (sensor SNR, PDR, residual battery, fault recovery status, security compliance score) and emitting a hormonal signal proportional to deviation from homeostatic target. The bloodstream corresponds to the weighted communication graph through which hormonal signals propagate with rates proportional to link quality. Receptor-mediated effector response corresponds to hormonal coupling functions mapping hormonal state to activation utilities. HPA axis feedback inhibition corresponds to cross-inhibition terms implementing competitive suppression between incompatible hormonal axes.

3.3.3. Key Formal Properties of the Hormonal Architecture

Four properties constitute the theoretical backbone:

Property 1 — Bounded Range. At all times, guaranteed by logistic saturation of emission functions and the projection step in Euler-Maruyama discretization. Essential for numerical stability and interpretability.

Property 2 — Homeostatic Stability. In the absence of environmental perturbations, hormonal dynamics converge to a unique equilibrium determined by stationary emission functions and the network Laplacian. Stability established through Lyapunov analysis.

Property 3 — Cross-Inhibition Prevents Regulatory Conflict. The cross-inhibition structure formally suppresses simultaneous elevation of physiologically incompatible hormones — specifically Energexin (energy crisis) and Connectin (energy-costly connectivity maintenance) — implementing principled regulatory trade-offs without explicit decision-theoretic reasoning, analogous to allostatic regulation — the biological process by which the body maintains stability by adjusting its regulatory setpoints in response to anticipated demand [43].

Property 4 — Intrinsic Interpretability. Each component encodes the urgency of a specific regulatory dimension. Every agent activation decision is traceable to a specific hormonal context and reconstructible from the symbolic engram record at zero computational overhead.

3.4. IoT-Specific Bio-Inspirations

3.4.1. Duty Cycle as Metabolic Sleep-Wake Regulation

The biological circadian system [44] implements a hormonally controlled alternation between high-activity (wakefulness) and reduced-activity restorative states (sleep), tightly coupled to energy availability. S-AI-IoT formalizes this through Energexin, modulating the active window of each node as:When Energexin is elevated (energy crisis), the active window shrinks toward (prolonged sleep); when low (abundant energy), it expands toward (full wakefulness).Cross-inhibition by Resiliencin and Sensorin overrides energy-driven sleep when fault-recovery or sensing urgency demands extended wakefulness — precisely as the biological HPA (hypothalamic-pituitary-adrenal) stress-response system overrides the circadian sleep drive (the hormonally controlled alternation between wakefulness and restorative sleep states [44]) during acute physiological challenges.Advantages over existing approaches (Q-learning schedules, threshold-triggered protocols): unified architectural coherence with broader behavioral orchestration, multi-objective energy-quality trade-offs via cross-inhibition, and temporal smoothing preventing high-frequency duty cycle oscillation in harvesting environments.

3.4.2. Network Topology as a Dynamic Endocrine Graph

The IoT communication graph is the formal analog of the vascular architecture. Link weight is defined as:A high-quality link (high RSSI, low loss) is a well-perfused blood vessel; a degraded link is a constricted vessel attenuating hormonal diffusion. This enables spatially differentiated regulatory responses: a Sensorin elevation at node propagates to neighbors at a rate determined by link quality, triggering neighbor-level responses that decay with topological distance. The time-varying IoT graph introduces a non-stationary perturbation to reaction-diffusion dynamics, handled through the stochastic perturbation term and Euler-Maruyama projection.

3.4.3. Security Compliance and the Artificial Immune Analogy

The Normin hormone formalizes an immune system analog, encoding the security and compliance posture of each node as a continuous signal derived from firmware currency, encryption status, certificate validity, behavioral deviation from baselines, and compliance with GDPR, ETSI EN 247 623, and IEC 62223. Deviations exceeding the anomaly threshold — analogous to self/non-self antigenic recognition — trigger Normin elevation, activating SecurityComplianceAgent and IntrusionDetectionAgent. The critical advantage over conventional AIS and rule-based IDS: security responses are not isolated detection events but hormonal signals that interact with energy, connectivity, and resilience dimensions through cross-inhibition, implementing principled security-energy trade-offs unavailable in any existing IoT security framework.

3.5. Symbolic Memory and Engrams in Distributed IoT Systems

3.5.1. The Engram Concept

The engram is the physical substrate of a memory trace — the ensemble of neural modifications that encode a past experience retrievable to influence future behavior. S-AI-IoT generalizes this to the distributed IoT context: each significant decision episode is stored as an extended engram record:

$$\varepsilon_m^{IoT} = \langle h^{(m)}, x^{(m)}, a^{(m)}, o^{(m)}, T^{(m)}, \beta^{(m)} \rangle$$

where $h^{(m)} \in [0,1]^5$ is the hormonal context, $x^{(m)} \in \{0,1\}^{|A|}$ the agent activation vector, $a^{(m)}$ the composite action vector, $o^{(m)}$ the observed outcome, $T^{(m)}$ the local network topology state, and $\beta^{(m)} \in [0,1]$ the energy budget level. This six-tuple captures not only what the system did but the complete operational context, enabling both retrospective explanation and prospective recall.

3.5.2. Content-Addressable Retrieval and Behavioral Acceleration

The retrieval mechanism implements a formal analog of hippocampal pattern completion: given a current hormonal context, the memory subsystem identifies stored engrams via weighted cosine similarity and presents their action components to the IoT-MetaAgent as candidate behavioral templates, reducing the orchestration search space and accelerating convergence. Computational cost — proportional to the number of stored engrams and hormonal vector dimensionality, both small on constrained nodes — makes this feasible on microcontroller-class hardware, in contrast to deep neural memory architectures.

3.5.3. Distributed Engram Architecture

Each node maintains its own local engram store. Cluster-level coherence is achieved through selective engram propagation: when a node encounters a novel scenario with no closely matching local engram, it broadcasts a memory query to its neighborhood, soliciting engrams from neighbors with similar past hormonal contexts. This extends effective memory capacity beyond local storage, implementing collaborative experiential learning consistent with distributed cognition models.

3.5.4. Engrams as Native Explainability Artifacts

For any agent activation decision, the corresponding engram provides a complete, human-readable justification — hormonal context, agents activated, actions taken, outcome observed, topology and energy budget — at zero additional computational cost, as a byproduct of the normal decision cycle. This constitutes intrinsic explainability [13]: transparency structural to the decision process itself, not an interpretive overlay. The contrast with SHAP [30] and LIME is qualitative: those methods approximate reasoning post-hoc and are computationally infeasible on constrained nodes; the engram record is an exact, causally faithful record requiring no approximation.

3.6. Stability, Regulation, and Anti-Oscillation

3.6.1. The Chattering Problem in Noisy IoT Environments

Measurement noise, intermittent connectivity, and stochastic energy harvesting can propagate through emission functions and induce spurious hormonal fluctuations causing unnecessary agent activation and deactivation cycles. In the extreme, a regulatory system with insufficient noise rejection exhibits chattering — alternately activating and deactivating agents at frequencies determined by noise rather than genuine operational events. In IoT contexts, chattering may consume more energy than persistent always-on operation, defeating the purpose of parsimony. S-AI-IoT addresses this through a three-layer anti-oscillation architecture.

3.6.2. Hormonal Temporal Integration as Low-Pass Filter

The time-constant in the reaction-diffusion equation implements a first-order low-pass filter on sensed observables. Sensorin and Connectin — responding to potentially rapid link quality changes — are calibrated with small time constants; Energexin and Resiliencin — encoding slowly-evolving energy and recovery processes — use larger time constants providing stronger noise rejection. This biologically motivated differentiation of hormonal timescales, directly analogous to the endocrine system's timescale hierarchy (millisecond neuropeptides to day-timescale steroid hormones), constitutes a principled multi-timescale regulation approach absent from all existing IoT intelligence frameworks.

3.6.3. Hysteresis in Agent Activation Thresholds

Even with temporal integration, residual fluctuations may cause repeated threshold crossings. S-AI-IoT applies a formal hysteresis mechanism: agent is activated when utility exceeds upper threshold and remains active until it falls below strictly lower threshold. The hysteresis band defines a dead zone eliminating spurious activations from threshold-level noise, analogous to Schmitt triggers and thermostat-type regulators in nonlinear control.

3.6.4. Time-to-Live and Cooldown Constraints

The Time-to-Live (TTL) parameter specifies the minimum duration an agent remains active once triggered, ensuring sufficient time to complete its behavioral response. The Cooldown parameter specifies a mandatory deactivation period following each activation episode, preventing rapid re-triggering from recurring transient stimuli. These constraints jointly implement a refractory period mechanism directly analogous to the absolute and relative refractory periods of biological neurons [45].

3.6.5. Separation of Temporal Scales

The three primary processing layers operate at formally separated timescales: sensing/perception (fastest), hormonal dynamics (intermediate), orchestration/agent activation (slowest). Rapid environmental fluctuations are absorbed at the perception layer; medium-frequency variations are integrated by hormonal dynamics at characteristic timescales; agent activation decisions — carrying the highest energetic and behavioral cost — are made at the slowest timescale after the hormonal state has converged to a stable representation. This multi-timescale architecture guarantees stability under singular perturbation theory.

3.7. Theoretical Positioning

Reaction-diffusion systems. S-AI-IoT hormonal dynamics belong to the mathematical class of reaction-diffusion systems on graphs [46, 47], with two technically significant extensions: (1) the graph is dynamic, requiring Lyapunov direct method rather than standard spectral analysis; (2) nonlinear cross-inhibition terms require nonlinear stability analysis beyond linearization. Both constitute directions for formal mathematical development.Primal-dual optimization. The orchestration mechanism extends the Network Utility Maximization (NUM) framework to an agent-selection context with time-varying utility functions coupled to hormonal dynamics. The composite closed-loop system converges to a neighborhood of the jointly optimal equilibrium, with the radius determined by stochastic perturbation intensity.S-AI research program continuity. S-AI-IoT is the IoT instantiation of a formally unified theoretical framework whose principles have been validated across conversational AI [35, 36,37], anti-hallucination [38,39], network management, cybersecurity, intelligent education [40], and embodied robotics [41], building on the foundational S-AI architecture [34]. The five adaptations from S-AI-ROBOTICS to S-AI-IoT — distributed gland layer, IoT-specific hormone semantics, dynamic topology weights, duty cycle output, extended engrams — constitute the minimal necessary modifications to accommodate IoT operational semantics, constituting principled domain transfer rather than adhoc redesign.The theoretical foundations of this section collectively justify the architectural commitments of S-AI-IoT and provide the conceptual scaffolding upon which the formal mathematical specification is erected.

4. GLOBAL ARCHITECTURE OF S-AI-IOT

The theoretical foundations of Section 3 — activation parsimony, artificial hormonal signaling, and symbolic engram memory — impose four structural constraints on the architecture: it must be hierarchically decomposable (intelligence distributable across node, gateway, and cluster tiers), hormonally coherent (all behavioral decisions at every tier driven by the same formally specified hormonal state), parsimonious by construction (agent activation carries computational cost only when the hormonal context justifies it), and intrinsically explainable (auditable behavioral traces generated as a byproduct of normal operation).

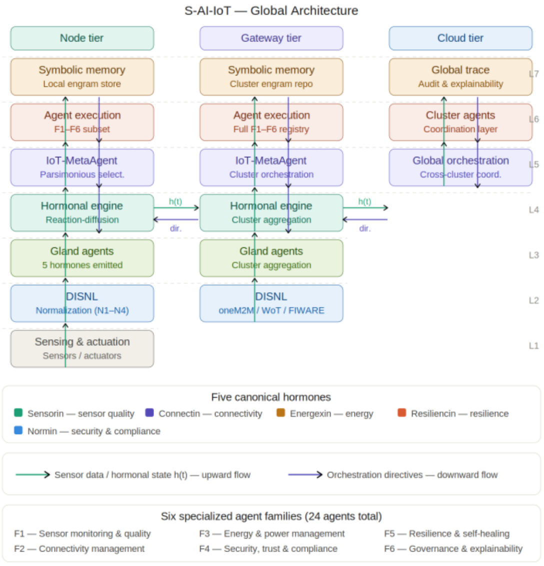
4.1. Seven-Layer Functional Stack and Hierarchical Orchestration Model

4.1.1. Architectural Overview

The global architecture of S-AI-IoT is organized as a seven-layer functional stack in which each layer has a formally defined input-output contract with its adjacent layers, and information flows both upward (raw sensor data to behavioral decisions) and downward (orchestration directives to actuator commands). The seven layers are, from bottom to top:

(L1) Physical Sensing and Actuation Layer (L2) Data Ingestion and Semantic Normalization Layer (L3) Gland Agent and Hormonal Emission Layer (L4) Hormonal Engine and Reaction-Diffusion Layer (L5) IoT-MetaAgent and Parsimonious Orchestration Layer (L6) Specialized Agent Execution Layer (L7) Symbolic Memory and Explainability Layer

Figure 1. Global architecture of S-AI-IoT: seven-layer functional stack deployed across three tiers (node, gateway, cloud), with hormonal state propagation (upward, green) and orchestration directives (downward, purple). Each tier displays only its operative layers; dashed horizontal lines indicate layer alignment across tiers.



This decomposition extends the six-layer S-AI-ROBOTICS pipeline [41] by introducing a dedicated L2 normalization layer, motivated by the substantially greater data heterogeneity of IoT deployments: unlike robotic platforms where sensor modalities are fixed at design time, a large-scale IoT deployment may have formats, units, and sampling rates varying node-by-node and evolving dynamically as devices are commissioned and decommissioned [1, 2].

4.1.2. Hierarchical Deployment Model

S-AI-IoT is designed for deployment across three tiers:
At the node tier, each IoT endpoint executes a lightweight S-AI-IoT stack instantiation comprising local gland agents for the five canonical hormones, the local hormonal engine, a subset of specialized agents compatible with the node's computational budget, a local engram store, and a compact IoT-MetaAgent operating under strict energy and memory constraints [8].

At the gateway tier, a gateway device — equipped with substantially greater resources (multi-core processor, gigabytes of RAM, persistent storage) — executes a full S-AI-IoT stack instantiation, aggregating hormonal signals from its served node cluster, orchestrating cluster-level agent activations, and maintaining the cluster-level engram repository [5, 6].

At the cloud or server tier, a global orchestration layer aggregates hormonal information from multiple gateways, maintains a global behavioral trace, and provides the interface for regulatory audit, system configuration, and cross-cluster coordination.

The critical architectural property of this model is graceful degradation under connectivity loss: each tier is capable of autonomous operation when connectivity to the tier above is disrupted. A node losing gateway connectivity continues executing its local gland agents, hormonal engine, and specialized agents using its local engram store; a gateway losing server connectivity continues orchestrating its cluster autonomously.This property is formally guaranteed by the **local hormonal stability** of each tier, established by Lyapunov analysis.

4.1.3. Information Flow and Inter-Layer Contracts

The information flow is governed by formally defined inter-layer contracts: L1 → L2: raw sensor readings as timestamped tuples — L2 → L3: normalized observation vectors annotated with quality flag — L3 → L4: emission signals for each hormone and node — L4 → L5: hormonal state vector — L5 → L6: activation directives dispatched to selected agents — L6 → L7: completed action records submitted to the symbolic memory subsystem.

4.2. IoT-MetaAgent and Parsimonious Orchestration

4.2.1. Role and Functional Specification

The IoT-MetaAgent (IoT-MA) is the central orchestration component of S-AI-IoT. Its primary function is the continuous resolution of the parsimonious agent-selection problem: given the current hormonal state vector, identify the minimal subset of specialized agents whose joint activation maximizes the behavioral utility of the node subject to its computational, energetic, and temporal resource constraints [34].The IoT-MA executes four functions sequentially at each orchestration cycle: 1. Hormonal state assessment: reads and compares against baseline equilibrium to identify hormonal dimensions elevated above their activation significance threshold. 2. Memory-guided anticipation: queries the symbolic engram memory for stored engrams whose hormonal signature closely matches the current state, retrieving candidate activation patterns proven effective in similar past contexts. 3. Constrained agent selection:

solves the primal-dual optimization to identify, using memory-retrieved candidates as warm-start initializations. 4. Activation dispatching: issues activation directives to selected agents, specifying execution priority, resource budget, and Time-to-Live.

4.2.2. Parsimonious Activation Policies

The IoT-MA enforces three complementary parsimony policies. The hormonal significance filter suppresses agent consideration for any agent whose primary hormonal coupling falls below significance threshold, preventing spurious activations from noise-induced fluctuations. The resource budget policy enforces the hard constraint, with dynamically adjusted as a function of Energexin [8]. The hysteresis activation policy implements threshold hysteresis from Section 3.6.3: agent is activated when and remains active until, preventing chattering at the activation boundary.

4.2.3. Cross-Tier Orchestration and Degraded-Mode Operation

In normal operation, the node-tier IoT-MA transmits periodic hormonal state reports to the gateway at reporting interval modulated by Connectin level — increasing frequency when connectivity is high, reducing it when constrained. The gateway-tier IoT-MA aggregates node reports into a cluster-level hormonal state, driving cluster-level agent activations.In degraded mode (gateway connectivity loss), the Resiliencin hormone is elevated by a connectivity-loss trigger, the NetworkRejoinAgent is activated with elevated priority, reporting frequency is reduced to conserve energy, and local behavioral decisions are made autonomously using the node's local engram store [41].

4.3. Gland Agent Layer: Hormonal Emission Interface

4.3.1. Functional Definition and Five Gland Agents

The Gland Agent layer interfaces between raw observational data and the Hormonal Engine. Each Gland Agent is a lightweight software module monitoring a specific set of IoT observables and emitting a hormonal signal proportional to the degree of departure from homeostatic target.S-AI-IoT defines five primary Gland Agents: SensorGland: monitors SNR, missing data rate, drift score, calibration age → emits — ConnectGland: monitors PDR, aggregated RSSI, end-to-end latency, jitter → emits — EnergyGland: monitors normalized battery level, CPU load, relative power consumption, harvested power availability → emits — ResilienceGland: monitors active fault count, time-since-last-recovery, redundancy availability → emits — NormGland: monitors firmware currency score, encryption protocol status, certificate validity, behavioral deviation → emits.

4.3.2. Emission Function Architecture

All five Gland Agents implement a **logistic-saturated emission architecture** in which the emission signal is a sigmoid function of a domain-specific observation aggregation, modulated by a saturation factor that ensures emission ceases when the hormone is already at maximum. The calibration parameters governing gain and bias are specified in the mathematical framework of Article II, which additionally develops a principled maximum-likelihood calibration procedure enabling systematic parameter estimation from deployment data as an alternative to manual configuration (Article II, Section 4.6).For Sensorin, the aggregation combines SNR degradation, missing data rate, drift score, and calibration age. For Energexin, the aggregation encodes three complementary facets of energy stress: battery depletion, processing overload, and instantaneous power overconsumption **[8]**. Per-cycle computational cost per Gland Agent is fewer than 100 arithmetic operations — a negligible fraction of the available budget even on a Cortex-M0+ class microcontroller at 42 MHz.

4.4. Specialized Agent Typology: Functional Families and Hormonal Couplings

Agent design is governed by three principles: functional completeness, minimal redundancy, and unambiguous hormonal coupling, yielding a six-family organization [34].

Family F1 — Sensor Monitoring and Quality Agents (Sensorin-coupled). Activated by elevated Sensorin: SensorDiagnosticAgent (continuous quality assessment, anomaly detection), SensorRecalibrationAgent (online recalibration when drift exceeds threshold), DataImputationAgent (lightweight statistical imputation for data stream gaps), SensorFusionAgent (activation of redundant modalities and quality-weighted fusion). Secondary coupling with Resiliencin governs escalation from diagnostic to recovery mode [1].

Family F2 — Connectivity Management Agents (Connectin-coupled). Activated by elevated Connectin: LinkQualityMonitorAgent (continuous link metric measurement, maintains), AdaptiveRoutingAgent (routing table modification toward higher-quality paths), ProtocolAdaptationAgent (cross-layer optimization of transmission power, spreading factor, channel), NetworkRejoinAgent (exponential backoff and multi-channel scanning for network rejoin) [5, 6].

Family F3 — Energy and Power Management Agents (Energexin-coupled). Activated by elevated Energexin: DutyCycleControllerAgent (adjusts per Section 3.4.1 duty cycle modulation equation), TransmissionPowerControlAgent (reduces radio power as function of Energexin), ComputationOffloadingAgent (identifies offloadable tasks to gateway when local energy insufficient), HarvestingForecastAgent (lightweight harvested power prediction for proactive scheduling). Cross-inhibition of Energexin by Resiliencin and Sensorin ensures F3 agents do not reduce duty cycle or power to levels compromising critical functions [8].

Family F4 — Security, Trust, and Compliance Agents (Normin-coupled). Activated by elevated Normin: SecurityComplianceAgent (compliance monitoring against GDPR, ETSI EN 247 623, IEC 62223; corrective actions), IntrusionDetectionAgent (behavioral anomaly detection implementing the artificial immune analogy), FirmwareUpdateAgent (security patch lifecycle management, scheduled during low-load and sufficient-energy windows), PrivacyGuardianAgent (data minimization and purpose-limitation enforcement at the DISNL outgoing filter, engram-logged for regulatory audit) [32].

Family F5 — Resilience and Self-Healing Agents (Resiliencin-coupled). Activated by elevated Resiliencin: FaultIsolationAgent (root cause identification and affected subsystem isolation), SelfHealingAgent (recovery procedures: software restarts, hardware resets, redundant resource activation, gateway escalation), RedundancyManagementAgent (dynamic inventory of backup resources), SystemCheckpointAgent (periodic operational state recording in persistent storage for rapid state restoration).

Family F6 — Governance, Coordination, and Explainability Agents (Multi-hormonal). Activated by composite hormonal conditions: EngineWatchdogAgent (hormonal engine health monitoring, corrective interventions), EngineSchedulerAgent (agent scheduling within available computational budget), ExplainabilityAgent (formats completed decision records as extended IoT engrams and generates human-readable justifications on demand), ClusterCoordinationAgent (two-tier hierarchical orchestration of Section 4.2.3) [13, 31].

4.5. Hormonal Engine and Distributed Symbolic Memory

4.5.1. Hormonal Engine: Architecture and Computation

The Hormonal Engine (HE) is the computational core of S-AI-IoT, integrating emission signals from the five Gland Agents, propagating hormonal information via reaction-diffusion dynamics, applying cross-inhibition, and delivering the updated hormonal state vector to the IoT-MetaAgent at each orchestration cycle [46, 47].The HE operates as a five-stage pipeline: Stage 1 (emission integration): five emission signals buffered in a circular delay buffer implementing emission delay — Stage 2 (reaction computation) — Stage 3 (diffusion computation) — Stage 4 (stochastic time-stepping) — Stage 5 (state broadcast): updated hormonal state vector written to shared memory.Computational cost is at most 44 inner-loop iterations per cycle. Total memory footprint is fewer than 440 bytes in typical configurations. The complete computational cost of the full S-AI-IoT cycle — covering the Hormonal Engine, agent selection, and engram retrieval — is $O(K \cdot |N_i| + |A|^2 + M_{PSM} \cdot K)$ per orchestration cycle per node, amounting to fewer than 10,000 arithmetic operations in typical IoT configurations ($K = 5$, $|N_i| \leq 10$, $|A| = 24$, $M_{PSM} \leq 1\,000$), well within the real-time constraints of gateway-class and mid-tier IoT hardware. Full complexity analysis is developed in Article II, Section 5.

4.5.2. Volatile Context Memory: Distributed Context Manager

The Distributed Context Manager (DCM) maintains a sliding-window record of recent hormonal states, emission signals, and agent activation events over a configurable retention window (typically 10–100 orchestration cycles). It serves three functions: provides hormonal history for emission delay computation in Stage 1; supplies recent behavioral context for trend-based anticipation by the IoT-MA; and provides the short-term episode record from which the ExplainabilityAgent constructs decision justifications on demand.

4.5.3. Persistent Symbolic Memory: Engram Architecture

The Persistent Symbolic Memory (PSM) is organized as a content-addressable store of extended IoT engrams [37]:The extension beyond the hormonal-action-outcome triple of S-AI-GPT to include topology and energy budget reflects the IoT-specific insight that the same hormonal context may warrant qualitatively different behavioral responses depending on network topology and energy posture.

Engram retrieval is triggered by the IoT-MA via weighted cosine similarity:The most similar engrams are retrieved and their action components are presented as candidate behavioral templates. Empirical evidence from S-AI-EDU and S-AI-ROBOTICS demonstrates that memory-guided initialization reduces convergence time by a factor of 2–4 in recurrent scenarios [40, 41].

The PSM is implemented as a fixed-capacity circular buffer with LRU eviction. New engrams are written only when the behavioral significance criterion is met: the activation pattern must differ from the most recently stored engram by at least one activated agent, and the outcome must be non-trivial.

4.6. Data Ingestion and Semantic Normalization Layer

The DISNL mediates between the physical diversity of IoT sensing infrastructure and the semantic uniformity required by the hormonal emission layer. It implements a four-stage normalization pipeline [1, 2]:

N1 — Unit conversion and temporal alignment: conversion to canonical physical units (SI for physical quantities; normalized for dimensionless metrics) and timestamp correction with linear interpolation for asynchronous sensors.

N2 — Spatial georeferencing: annotation with GPS coordinates or symbolic location identifier, enabling spatial indexing and supporting topology-dependent diffusion computations [47].
N3 — Quality assessment and flagging: quality flag based on range validity, temporal freshness, and cross-sensor consistency [1].

N4 — Missing data handling: last-value carry-forward (slowly-varying quantities), spatial interpolation from neighboring nodes, or explicit missing-data markers triggering elevated Sensorin emission in the SensorGland.

The canonical observation vector is structured according to a lightweight semantic model inspired by the W3C SSN ontology, encoded as a JSON configuration file loaded at node initialization, making sensor modality addition, removal, or reconfiguration transparent to the Gland Agent and Hormonal Engine layers.

Integration with Existing IoT Standards. S-AI-IoT is designed as an intelligence orchestration layer that operates above existing IoT connectivity and data standards rather than replacing them. The DISNL constitutes the primary integration point: its four-stage normalization pipeline is explicitly designed to ingest data from heterogeneous sources conforming to standards including oneM2M, FIWARE NGSI-LD, W3C Web of Things (WoT), and OCF. The canonical observation vector is structured according to a semantic model inspired by the W3C SSN/SOSA ontology, enabling alignment with existing semantic IoT descriptions without requiring ontology conformance at the node level. At the communication layer, S-AI-IoT's control plane operates over standard IoT messaging protocols (MQTT, CoAP, AMQP), making it deployable over existing network infrastructure without modification. The hormonal state vector and engram records are encoded as lightweight JSON structures transmissible over any IoT messaging fabric. Integration with platform-level standards such as Matter and oneM2M is achieved through the gateway tier, which exposes a standard northbound API while executing the full S-AI-IoT stack internally. A conceptual integration specification for the three most widely deployed platforms (oneM2M, FIWARE, W3C WoT) is outlined in Article III; full adapter implementation is identified as a practical engineering deliverable beyond the scope of this trilogy.

4.7. Security, Safety, and Compliance Architecture

4.7.1. Integrated Security as an Architectural Property

Security in S-AI-IoT is not an independent module layered over the intelligence architecture but an intrinsic property of the hormonal regulatory system, implemented through the Normin hormone. Elevated Normin triggers F4 security agents proportionally to assessed severity, implements cross-inhibitory suppression of connectivity-expanding behaviors (preventing compromised nodes from broadcasting sensitive data), and elevates the engram-recording threshold to ensure all security-relevant episodes are permanently stored [32].This hormonal security integration delivers three architectural properties that no existing IoT security framework surveyed in Section 2 provides simultaneously as first-class architectural guarantees:(1) security-energy trade-offs through Normin-Energexin cross-inhibition; (2) intrinsic security explainability — every security decision is traceable to the Normin level and specific compliance observables that elevated it, enabling post-incident forensic analysis without dedicated logging infrastructure;

(3) adaptive threat response — NormGland reconfigurable to incorporate new compliance observables without modifying the Hormonal Engine or F4 agent implementations [13, 31].

4.7.2. Signed Communication Channels and Control/Data Plane Separation

S-AI-IoT enforces a strict separation between the control plane (hormonal state reports, orchestration directives, engram propagation messages, security policy updates) and the data plane (application-level sensor data, actuation commands, telemetry). Control plane messages are cryptographically signed using node-specific private keys managed by the SecurityComplianceAgent, and authenticated by receiving nodes before incorporation into hormonal state computation. This architecture follows ETSI EN 247 623 security-by-design principles and IEC 62223 defense-in-depth, and prevents injection of forged hormonal signals that could manipulate the behavioral orchestration of target nodes.

4.7.3. Privacy Architecture and GDPR Compliance

The PrivacyGuardianAgent enforces data minimization and purpose-limitation obligations at the DISNL outgoing data filter, suppressing or anonymizing measurements whose transmission would violate the node's privacy configuration.

4.7.4. Byzantine Robustness of the Hormonal Layer

The hormonal diffusion mechanism constitutes a potential attack surface beyond the cryptographic controls of Section 4.7.2: a compromised node that successfully injects falsified hormonal signals — artificially elevated Normin or Resiliencin — can propagate behavioral disturbances through the reaction-diffusion graph before detection by the IntrusionDetectionAgent (F4-B). S-AI-IoT addresses this threat through three complementary mechanisms. First, the cryptographic authentication of Section 4.7.2 constitutes the primary defense: hormonal state reports are signed with node-specific private keys, and any unsigned or invalidly signed report is rejected before incorporation into the Hormonal Engine computation. A Byzantine node that has not compromised its cryptographic credentials cannot inject falsified signals into the hormonal layer.

Second, the reaction-diffusion dynamics themselves provide structural attenuation of injected disturbances. The graph diffusion term $D_k \sum_{j \in \mathcal{N}_i} w_{ij} \left(h_{k,j} - h_{k,i}\right)$ acts as a spatial averaging operator: an anomalous hormonal value injected at a single node is attenuated by the weighted average of its neighbors' legitimate values, with attenuation proportional to node degree and link quality. For a network with average degree $\bar{d}$ and uniform link weights, a falsified signal of amplitude Δh injected at a single node propagates to its neighborhood with amplitude reduced to $\Delta h / \bar{d}$ after one diffusion step, and decays geometrically with topological distance.

Third, the IntrusionDetectionAgent (F4-B) monitors inter-node communication patterns and detects behavioral anomalies indicative of hormonal injection attacks via its behavioral deviation signal $\delta_{beh,i}(t)$, which feeds directly into the NormGland emission function. Upon detection, F4-B issues a quarantine directive that removes the suspect node from the hormonal diffusion graph — effectively setting $w_{ij} = 0$ for all links to the compromised node — halting further signal propagation.

The combined effect of these three mechanisms — cryptographic authentication, diffusive attenuation, and quarantine — bounds the impact of a Byzantine node to a transient hormonal perturbation in its immediate neighborhood prior to detection. A formal analysis of the maximum tolerable fraction of Byzantine nodes as a function of network topology and detection latency is identified as a direction for future work, to be addressed in the experimental evaluation of Article III.

5. CONCLUSION

This article — the first of a three-part series on S-AI-IoT — has presented the motivations, state-of-the-art positioning, theoretical and bio-inspired foundations, and global architecture ofa formally grounded, intrinsically parsimonious, hormonally regulated, and natively explainable IoT intelligence framework that is the first to combine these properties by architectural design.The contributions developed in this article are fourfold. The literature review (Section 2) identified six structural gaps shared by existing approaches: absence of a unified continuous signaling layer, absence of activation parsimony as a primary architectural property, inapplicability of post-hoc XAI on constrained nodes, security-energy decoupling, absence of distributed symbolic memory, and absence of reproducible benchmarks. The theoretical foundations (Section 3) established activation parsimony as a formal principle grounded in sparse approximation theory, artificial hormonal signaling as the universal medium of contextual self-regulation with four formal biology-to-architecture correspondences, and engram-based symbolic memory as a behavioral persistence mechanism. The global architecture (Section 4) described the seven-layer functional stack, the three-tier hierarchical deployment model with formally guaranteed local hormonal stability at each tier, ensuring coherent autonomous operation upon connectivity loss, the IoT-MetaAgent with its three parsimony policies, the five Gland Agents and twenty-four specialized agents organized in six families, the Hormonal Engine and Persistent Symbolic Memory, and the intrinsic security architecture via Normin. The comparative positioning (Table 1, Section 1.5) demonstrated that, among the eight categories of frameworks surveyed, S-AI-IoT is the only one designed to simultaneously address all six identified structural dimensions.Article II of this trilogy develops the complete formal specification of the 24 specialized agents with formal input-output contracts (C1–C5 of Article I), the full mathematical modeling — the reaction-diffusion equation, Lyapunov stability Theorem 6.1 establishing global asymptotic stability, the primal-dual constrained orchestration mechanism, the Euler-Maruyama discretization with projection, andthe engram formalism with Proposition 6.2 on intrinsic explainability at O(1) additional cost per decision cycle— as well as the algorithmic implementation with pseudocode for all six core algorithms.

Article III presents the complete experimental evaluation on the SAI-UT+ IoT testbench covering ten representative operational scenarios including multi-stress combined conditions, with systematic comparison against five baseline approaches (rule-based systems, DQN, federated learning, behavior trees, TinyML) and a full ablation study quantifying the individual contribution of each architectural component. It also develops the general conclusions of the trilogy and future research directions, including validation on real IoT deployments, extension to smart grid and connected health domains, federated hormonal learning (FHL) for distributed parameter adaptation, and investigation of scalability properties beyond 10,000 nodes.

REFERENCES

[1] J. Luo, W. Luo, and M. Liu, “Smart IoT sensor data quality management: Survey and challenges,” Future Generation Computer Systems, vol. 145, pp. 195–212, 2023.

[2] D. M. Dauda, M. A. Al-Garadi, and A. Rayes, “Internet of Things applications and deployment architectures: A comprehensive survey,” Sensors, vol. 24, no. 3, p. 812, 2024.

[3] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, “A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications,” IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1125–1142, 2017.

[4] IoT Analytics, “State of IoT — Spring 2025,” IoT Analytics Research, Hamburg, Germany, 2025.

[5] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, 2016.

[6] Z. Deng, W. Luo, and Y. Zhou, “Edge intelligence: Architectures, challenges, and applications,” IEEE Internet of Things Journal, vol. 7, no. 8, pp. 6895–6914, 2020.

[7] L. Djennadi and G. Diaz, “Exploring SDN architectures for IoT over low-power and lossy networks (LLNs): A survey,” Computer Networks, 2025. https://doi.org/10.1016/j.comnet.2025.110961

[8] S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: Survey and implications,” IEEE Communications Surveys & Tutorials, vol. 13, no. 3, pp. 443–461, 2011.

[9] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, “Software-defined networking: A comprehensive survey,” Proceedings of the IEEE, vol. 103, no. 1, pp. 14–76, 2015.

[10] V. J. Reddi, C. Cheng, D. Kanter, P. Mattson, G. Schmuelling, and C.-J. Wu, “MLPerf inference benchmark,” in Proc. ACM/IEEE ISCA, 2020, pp. 446–459.

[11] V. Mnih, K. Kavukcuoglu, D. Silver, et al., “Human-level control through deep reinforcement learning,” Nature, vol. 458, no. 7540, pp. 469–473, 2015.

[12] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in Proc. ICML, 2018, pp. 1861–1870.

[13] L. Longo, M. Brcic, F. Cabitza, et al., “Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions,” Information Fusion, vol. 106, p. 102301, 2024.

[14] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Proc. AISTATS, 2017, pp. 1273–1282.

[15] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 44–60, 2020.

[16] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Proc. IEEE Int. Conf. Neural Networks, 1995, vol. 4, pp. 1942–1942.

[17] U. Raza, P. Kulkarni, and M. Sooriyabandara, “Low power wide area networks: An overview,” IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 855–873, 2017.

[18] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, “On the convergence of FedAvg on non-IID data,” in Proc. ICLR, 2020.

[19] W. Fedus, B. Zoph, and N. Shazeer, “Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity,” Journal of Machine Learning Research, vol. 23, no. 120, pp. 1–39, 2022.

[20] A. Kusupati et al., “Soft Threshold Weight Reparameterization for Learnable Sparsity,” in Proc. ICML, 2020, pp. 5544–5555.

[21] M. Colledanchise and P. Ögren, Behavior Trees in Robotics and AI. CRC Press, 2018.

[22] M. Wooldridge and N. R. Jennings, “Intelligent agents: Theory and practice,” The Knowledge Engineering Review, vol. 10, no. 2, pp. 115–152, 1995.

[23] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, “Swarm robotics: A review from the swarm engineering perspective,” Swarm Intelligence, vol. 7, no. 1, pp. 1–41, 2013.

[24] M. Karaboga and B. Basturk, “On the performance of artificial bee colony (ABC) algorithm,” Applied Soft Computing, vol. 8, no. 1, pp. 687–697, 2008.

[25] E. Bonabeau, M. Dorigo, and G. Theraulaz, Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, 1999.

[26] K. M. Passino, “Biomimicry of bacterial foraging for distributed optimization and control,” IEEE Control Systems Magazine, vol. 22, no. 3, pp. 46–67, 2002.

[27] G. Lichtner, A. Kirsch, and E. Körner, “Autonomous behavior generation in neuroendocrine robot controllers,” Autonomous Robots, vol. 30, no. 4, pp. 363–380, 2011.

[28] J. Timmis, A. E. I. Brownlee, J. A. Greensmith, and A. Tyrrell, “Theoretical and empirical investigation of an immune inspired algorithm for job shop scheduling,” Swarm Intelligence, vol. 4, no. 4, pp. 229–246, 2010.

[29] E. Hart and J. Timmis, “Application areas of AIS: The past, the present and the future,” Applied Soft Computing, vol. 8, no. 1, pp. 191–201, 2008.

[30] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Proc. NeurIPS, 2017, pp. 4765–4774.

[31] S. K. Jagatheesaperumal, M. Rahouti, K. Ahmad, A. Al-Fuqaha, and M. Guizani, "Explainable AI over the Internet of Things (IoT): Overview, state-of-the-art and future directions," IEEE Open Journal of the Communications Society, vol. 3, pp. 2106–2136, 2022.

[32] R. Roman, J. Lopez, and M. Mambo, “Mobile edge computing, fog et al.: Security threats and challenges,” Future Generation Computer Systems, vol. 78, pp. 680–698, 2018.

[33] Transforma Insights, “The AIoT Paradigm: Embedding Intelligence in the Edge,” Transforma Insights Report, 2025.

[34] S. Slaoui, “S-AI: A sparse artificial intelligence architecture with hormonal orchestration and parsimonious agent activation,” International Journal for Multidisciplinary Research, vol. 1, no. 2, 2025.

[35] S. Slaoui, “S-AI-GPT: Hormonal orchestration for conversational AI,” IJAIA, 2025.

[36] S. Slaoui, “S-AI-GPT II: Extended hormonal regulation for large language models,” IJAIA, 2025.

[37] S. Slaoui, “S-AI-GPT III: Symbolic engram memory for cross-episode behavioral consistency,” SSRN, 2025.

[38] S. Slaoui, “S-AI-Anti-Hallucination: Homeostatic stability under cross-inhibitory hormonal dynamics,” IJAIA, 2025.

[39] S. Slaoui, “S-AI-Anti-Hallucination II: Triadic extension,” IJAIA, 2026.

[40] S. Slaoui, “S-AI-EDU: Modular agent typology and symbolic memory for intelligent education,” IJAIA, 2026.

[41] S. Slaoui, “S-AI-ROBOTICS: Six-layer bio-inspired architecture for autonomous robotic systems,” IJAIA, 2025.

[42] D. L. Donoho, “Compressed sensing,” IEEE Transactions on Information Theory, vol. 46, no. 4, pp. 1289–1306, 2006.

[43] B. S. McEwen and J. C. Wingfield, “The concept of allostasis in biology and biomedicine,” Hormones and Behavior, vol. 43, no. 1, pp. 2–15, 2003.

[44] M. H. Hastings, A. B. Reddy, and E. S. Maywood, “A clockwork web: Circadian timing in brain and periphery,” Nature Reviews Neuroscience, vol. 4, no. 8, pp. 643–661, 2003.

[45] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” Journal of Physiology, vol. 117, no. 4, pp. 440–544, 1946.

[46] A. M. Turing, “The Chemical Basis of Morphogenesis,” Philosophical Transactions of the Royal Society B, vol. 237, no. 641, pp. 37–72, 1946.

[47] M. E. J. Newman, Networks: An Introduction. Oxford University Press, 2010.

[48] J. Noor, M. A. Hossain, and M. R. Hasan, “TinyML-based intelligent IoT: Architecture, applications and challenges,” Future Generation Computer Systems, vol. 152, pp. 232–248, 2024.

[49] Y. Zhang, Z. Yang, and D. Zheng, “Multi-agent reinforcement learning for IoT resource management: A survey,” IEEE Internet of Things Journal, vol. 11, no. 5, pp. 7645–7662, 2024.

AUTHOR

Said Slaoui is a professor at Mohammed V University in Rabat, Morocco. He graduated in Computer Science from University Pierre and Marie Curie, Paris VI (in collaboration with IBM France), 1986. He has over 40 years of experience in the fields of AI and Big Data, with research focused on modular architectures, symbolic reasoning, and computational frugality. His recent work introduces the Sparse Artificial Intelligence (S-AI) framework, which integrates bio-inspired signaling and agent-based orchestration. He has published numerous scientific papers in international journals and conferences, and actively contributes to the development of sustainable and explainable AI systems.