



MEMO — Concept for Consideration: Governable Decentralized Neural Networks (G-DNNs)



Introduction

This memo introduces a proposed architecture called **G-DNNs** — *Governable Decentralized Neural Networks* — inspired by Sutton's concept of **neurons as goal-seeking agents** and building upon the real-world infrastructure of **InitialS-AI**, a decentralized auditing system for AI models.

We believe this direction presents a concrete pathway to operationalize Sutton's DNN vision, with added layers of **auditing, verification, and dynamic evolution**, all governed via a decentralized incentive system rooted in blockchain.

1. Foundational Mapping: DNNs × AI Governance

Your Concept: DNNs	InitialS-AI Framework	G-DNNs Integration
Neurons are goal-seeking autonomous agents	AI models require audit, scoring, and repair	Each neuron (agent) has an interface for governance & scoring
Network grows via dynamic structure (backbone + fringe)	Model versions must be auditable and trackable	Fringe neurons undergo on-chain review to become backbone
Neurons strive to be "heard"	Nodes rewarded for high-quality contributions	Neuron "survival" is based on evaluation + use frequency
No fixed global goal	Multi-dimensional scoring: safety, bias, robustness	Neurons adapt in a governance-defined quality space
Learning is continuous and online	Auditing is continuous and task-driven	Evaluation feedback loops back into learning process

2. Bridging with Blockchain: Neuron ↔ Node Mapping

DNNs Concept	Blockchain Perspective
Neuron (Agent)	On-chain compute node / modular AI agent
Neuron's goal	Agent-level objective or task payoff
Neuron connections	Message-passing / call graphs between nodes
Backbone neurons	High-credibility, high-weighted core nodes
Fringe neurons	Newcomer nodes subject to validation

3. Lifecycle of a Governable Neuron

Fringe-to-Backbone Pathway (governance-enabled):

1. New neuron/module is registered on-chain

2. Assigned domain-specific validation tasks (e.g., classification, code generation)

3. Outputs are audited by decentralized experts and validators

4. Metrics include:

◦ Safety class (toxicity, privacy)

◦ Quality class (accuracy, bias, novelty)

◦ Reward-to-Stake scores

✓

If successful, the neuron is integrated into the backbone, callable in broader inference chains.

4. Governance Mechanisms

Mechanism	Proposed Design
KYA (Know Your AI)	On-chain auditing tasks triggered upon neuron registration, using InitialS evaluation framework
Backbone-fringe evolution	High-performing fringe neurons challenge for backbone position via tokenized staking or voting
Listen-to-earn	Neurons gain survival points by being used/invoked in live inference
Performance decay	Dormant or low-score neurons face demotion or replacement
Multi-dimensional review	Decentralized metrics across safety, bias, hallucination, OOD behavior, etc.
Transparent versioning	Full training history, evaluation logs, and usage stats are stored immutably on-chain

5. Implementation Feasibility (MVP Suggestions)

Functionality	Technical Implementation
Modular neuron agents	Agentic LLMs with autonomous goals (e.g., LangChain, ReAct frameworks)
Audit task routing	Off-chain coordination + on-chain result settlement
Network structure tracking	Graph database (e.g., Neo4j) + Graph NFT representations
Governance and staking	Use InitialS token logic: audit rewards, staking, and governance voting

6. Key Insight

The view that “each neuron should seek to be heard” can be operationalized through a decentralized audit system — where **being heard** = **being validated, useful, and trusted**.

This G-DNNs proposal transforms Sutton's philosophical framework into a **living, evolving, and governable network of AI agents**, bringing **plasticity, autonomy, and verifiability** together in one

system.

7. Why This Matters Now

- LLMs are growing, but remain fragile, centralized, and non-transparent
 - Real-time learning and continual plasticity are not yet achieved
 - Governance frameworks for AI are emerging (EU AI Act, UN, NVIDIA), but lack architectural anchor points
 - G-DNNs provides both **philosophical alignment** with Sutton's work and a **technically grounded roadmap** to build next-generation autonomous AI systems
-

Appendix & References

1. Sutton, R. (2024). *Decentralized Neural Networks*. <http://incompleteideas.net/Talks/DNNs-Singapore.pdf>
2. Silver, Sutton et al. (2021). *Reward is Enough*. <https://web.eecs.umich.edu/~baveja/Papers/RewardIsEnough.pdf>