Executive Summary

AI is Reshaping Society—But Its Foundations Are Broken   Artificial Intelligence is becoming the foundational infrastructure for communication, decision-making, governance, and economic systems. Yet today's AI platforms are closed, opaque, and unaccountable:

- Users cannot verify how models were configured, what data was used, or how outputs were generated.
- Outputs are mutable, unverifiable, and depend entirely on centralized, proprietary infrastructure.
- Trust is fragile and based on promises, not proofs.

This creates a dangerous asymmetry where AI is a black box controlled by corporations, leaving users, enterprises, and regulators powerless to audit or verify critical decisions.

The Future of AI Must Be Decentralized, Verifiable, and Open   NeuroSpace is pioneering the first fully decentralized, verifiable AI protocol—where trust is built into the infrastructure itself:

- Every interaction is hashed, signed, pinned to IPFS, and committed on-chain.
- Model parameters, document sources, and session metadata are fully public and tamper-evident.
- Users—not providers—own the data, the proofs, and the conversations.

What NeuroSpace Enables Today

- Verifiable chat and RAG tools deployed on Base mainnet.
- Moderation and decentralized feedback live.
- 10 free requests per wallet (off-chain), with ETH and NeuroCoin payments fully integrated.
- NeuroCoin (NSPACE) launched on Base mainnet: 100M supply, 10M minted.

The Vision: Provable AI for Everyone   NeuroSpace is not just an app—it is the foundation for a future where:

- AI systems are transparent, auditable, and user-owned.
- Model benchmarking, evaluation, and reputation are community-driven.
- AI governance and moderation are protocol-based, not corporate-controlled.
- Agent reasoning and multi-step execution traces are fully verifiable.

This is not just a technical shift—it is a societal imperative. In the era of ubiquitous AI, only decentralized, provable infrastructure can create systems we can trust. NeuroSpace is building the future of AI that is open, accountable, and credibly neutral.

NeuroSpace: Building the future of provable, transparent, and user-owned AI infrastructure.
`https://neurospace.chat`

NeuroSpace: A Protocol for Verifiable and Traceable Language Model Inference

https://neurospace.chat

May 15, 2025

Abstract

Artificial Intelligence is becoming critical infrastructure for communication, decision-making, and computation. Yet modern language model systems operate as closed, opaque platforms: their outputs are unverifiable, their behaviors untraceable, and their deployment controlled by centralized entities. This creates systemic trust, accountability, and auditability challenges, particularly as AI systems increasingly mediate high-stakes human and institutional processes.

NeuroSpace introduces a decentralized protocol for verifiable and traceable language model inference. By combining cryptographic hashing, decentralized storage (IPFS), digital signatures, and on-chain commitments, NeuroSpace transforms each interaction into a tamper-evident, publicly auditable artifact. Every prompt, response, and associated metadata can be independently verified, attributed to a wallet address, and retrieved via content-addressed storage.

The protocol supports verifiable chat sessions, Retrieval-Augmented Generation (RAG) with on-chain proof of document sources, and traceable multi-step agent execution workflows. A native utility token, NeuroCoin (NSPACE), powers payments, staking, and decentralized governance—aligning incentives across users, model providers, validators, and the broader community.

NeuroSpace is not just a protocol for AI interaction—it is foundational infrastructure for a future in which machine intelligence is decentralized, transparent, and accountable by default.

1 Introduction

Large language models (LLMs) have transformed how information is generated, queried, and consumed. From summarizing documents to answering technical questions, they now mediate vast amounts of digital cognition. Yet despite their power, modern LLMs are opaque: their outputs are unverifiable, their behavior untraceable, and their deployment largely centralized.

This lack of transparency presents both technical and ethical challenges. How can users trust a model's answer? How can third parties audit what was said, when, and why? And how can we attribute responsibility for critical decisions made by algorithmic agents?

NeuroSpace addresses these challenges by introducing a cryptographically secure protocol for verifiable machine intelligence. At its core, NeuroSpace transforms each LLM interaction into a formally verifiable event. Prompt, response, and metadata are hashed, signed, and pinned to decentralized storage. These cryptographic commitments are then submitted to a smart contract, producing a permanent, auditable record on-chain.

This architecture enables users and third parties to:

- Reconstruct and verify the integrity of any message
- Attribute outputs to specific wallets or agents
- Audit entire chat sessions or agent execution traces

- Ground model outputs in immutable external documents (RAG)

In parallel, NeuroSpace introduces a native token, **NeuroCoin (NSPACE)**, to power its economic and governance layers. The token serves as the medium for payments, staking, and future coordination.

Together, these components form a new substrate for trusted AI infrastructure—one in which language model outputs are no longer ephemeral, but permanent, verifiable, and provably accountable.

2 Technical Protocol Overview

2.1 Verifiability Layer

The Verifiability Layer in NeuroSpace ensures that each interaction with a language model is publicly auditable and cryptographically secured. This section defines the interaction structure, commitment scheme, and verification protocol.

Interaction Model

Each user interaction is modeled as a structured tuple:

$$m = (\text{prompt}, \text{response}, \text{metadata}) \in \mathcal{M}$$

where:

- `prompt`: The user-provided natural language input.
- `response`: The LLM-generated output.
- `metadata`: Model configuration (name, temperature, token limit), session ID, timestamp t, and wallet address w.

Hash Commitment

A cryptographic commitment is computed using SHA-256:

$$h = H(\text{Serialize}(m)), \quad H : \{0,1\}^* \to \{0,1\}^{256}$$

where `Serialize` is a deterministic, bijective encoding function over $\mathcal{M}$. This ensures:

$$\text{Deserialize}(\text{Serialize}(m)) = m$$

Thus, any honest verifier can reconstruct h from the published message.

Digital Signature

Let $\text{Sign}_{sk}(h)$ denote an EIP-191-compatible digital signature using a private key sk. The user computes:

$$\sigma = \text{Sign}_{sk}(h)$$

This binds the message to a unique Ethereum address, enabling cryptographic attribution. The corresponding wallet address w can be recovered using:

$$w = \text{RecoverAddress}(h, \sigma)$$

On-Chain Commitment

A smart contract $\mathcal{C}$ maintains a public, append-only ledger of commitments:

$$(h, w, t_{\text{block}}) \in \mathbb{H} \times \mathbb{W} \times \mathbb{T}$$

where:

- h: SHA-256 commitment of the interaction.
- w: Submitting wallet address.
- t_{block}: Timestamp of the block when stored.

Verification Procedure

To verify a message m with signature σ and claimed signer w, the verifier performs:

1. Compute $h' = H(\text{Serialize}(m))$
2. Recover $w' = \text{RecoverAddress}(h', \sigma)$
3. Confirm $w' = w$
4. Check that $(h', w, t) \in \mathcal{C}$

If all checks pass, the interaction is deemed *verifiably committed*.

Security Guarantees

Assuming the collision resistance of SHA-256 and unforgeability of ECDSA, the protocol guarantees:

- **Integrity**: Adversaries cannot forge alternate $m' \neq m$ with $H(m') = H(m)$.
- **Non-repudiation**: Signers cannot deny authorship of valid σ.
- **Immutability**: On-chain records cannot be retroactively modified or deleted.

2.2 Storage and Retrieval

In NeuroSpace, content persistence is achieved through decentralized file storage, ensuring that every prompt-response pair is content-addressable, censorship-resistant, and retrievable for independent verification. The system utilizes the InterPlanetary File System (IPFS) to store serialized message data.

IPFS-Based Storage Model

Given a structured message tuple:

$$m = (\text{prompt}, \text{response}, \text{metadata})$$

the backend serializes m deterministically into a byte array $b \in \{0, 1\}^*$. This serialized payload is uploaded to IPFS, resulting in a unique content identifier (CID):

$$\text{CID}_m = \text{IPFS.Put}(b)$$

The CID is a multihash representing the SHA-256 digest of the content and serves as an immutable, globally addressable pointer.

Content-Addressable Integrity

Let:

$$h = H(\text{Serialize}(m))$$

Then under the default IPFS hashing function:

$$\text{CID}_m \equiv \text{Multihash}(h)$$

Any deviation in the content m results in a different CID, preserving strong integrity guarantees. The CID is recorded in the metadata and is optionally stored on-chain alongside the commitment hash for additional auditability.

Metadata Schema

Each stored message is bundled with metadata fields to allow full contextual reconstruction and reproducibility:

- **Model configuration:** name, ID, temperature, max tokens
- **Timestamps:** local timestamp t, block timestamp t_{block}
- **Identity:** wallet address w, session UUID s
- **Verification:** hash h, signature σ, IPFS CID

This metadata is used by verifiers to ensure consistent hash recomputation and signature validation.

Retrieval and Verification Flow

On retrieval, the client performs the following:

1. Fetch raw content via IPFS:

$$b = \text{IPFS.Get}(\text{CID}_m)$$

2. Deserialize to obtain m
3. Compute $h' = H(\text{Serialize}(m))$
4. Verify digital signature σ and wallet w
5. Check on-chain presence of $(h', w) \in \mathcal{C}$

This round-trip provides a full verification circuit from input to content to on-chain proof.

Availability Considerations

IPFS provides decentralized distribution but not persistence guarantees. NeuroSpace addresses this through:

- Redundant pinning across independent pinning services
- Optional integration with Filecoin or Arweave (future work)
- Frontend caching and fallback storage

The design ensures long-term data retrievability without relying on centralized infrastructure.

Security Properties

- **Tamper resistance:** IPFS CID is bound to content hash.
- **Censorship resistance:** Distributed peer-to-peer network.
- **Transparency:** Anyone can resolve CID and verify integrity.

2.3 Session Identity and Traceability

To support structured, longitudinal interactions, NeuroSpace introduces the notion of verifiable sessions. Each session represents a temporally ordered sequence of user-model interactions, cryptographically bound to a user identity and persistently linked across messages.

Session Definition

Let a session be defined as:

$$S = (s, w, \{m_1, m_2, \ldots, m_n\})$$

where:

- $s \in \mathcal{U}$: A globally unique session identifier (UUID).
- $w \in \mathbb{W}$: Ethereum wallet address initiating the session.
- $m_i \in \mathcal{M}$: Ordered messages generated within the session.

Each message m_i is defined as in Section 2.1 and is independently verifiable.

Session Binding

All messages within a session share the same session ID s and wallet address w. The metadata for each message contains:

$$\mathtt{metadata} = \{\ldots, s, w, t_i\}$$

This allows for:

- Temporal reconstruction of the session history.
- Identification of the originating wallet.
- Hash and signature consistency checks across session scope.

Deterministic Lineage and Replay

Let $\mathcal{H}_S$ be the ordered set of committed message hashes in a session:

$$\mathcal{H}_S = [h_1, h_2, \ldots, h_n] \quad \text{where } h_i = H(\text{Serialize}(m_i))$$

This sequence represents the immutable record of the session. Assuming the LLM is deterministic given fixed configuration, the session can be replayed or independently simulated.

Traceability Guarantees

By design, the session structure provides the following guarantees:

- **Causal ordering:** Each message is timestamped and ordered by submission time.
- **Attribution:** All messages are signed by the same wallet address w.
- **Replayability:** Full session reconstruction is possible given stored messages and metadata.
- **Non-equivocation:** Forking or altering session history requires a hash/signature mismatch.

Applications of Traceability

The verifiable session structure supports higher-order use cases:

- Auditable multi-turn conversations
- Agent memory logs with verified steps
- Cross-session reputation systems
- Forensic analysis and provenance tracking

In the future, sessions may be extended with external tools, agent state transitions, or encrypted embeddings, further enriching their traceability potential.

2.4 Verifiable Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) enhances LLM performance by incorporating external documents into the generation process. In NeuroSpace, we extend RAG with cryptographic verifiability, ensuring that all cited sources are immutable, content-addressed, and independently inspectable.

RAG Formal Model

Let:

- q: A user query (prompt)
- $D = \{d_1, d_2, \ldots, d_n\}$: A retrieved document set
- $r = \text{LLM}(q, D)$: The model-generated response

Each document $d_i \in D$ is stored on IPFS and associated with a unique content identifier:

$$h_i = H(d_i) = \text{SHA-256}(d_i) \quad \text{and} \quad \text{CID}_i = \text{Multihash}(h_i)$$

The set of document hashes used in a RAG call is denoted:

$$\mathcal{H}_D = \{h_1, h_2, \ldots, h_n\}$$

Binding Sources to Response

The response r is generated with explicit dependence on D. NeuroSpace binds r to the retrieval set by extending the verification payload:

$$m = (q, r, \mathcal{H}_D, \text{metadata})$$

This tuple is serialized and hashed as in Section 2.1:

$$h = H(\text{Serialize}(m))$$

The resulting hash is signed and submitted on-chain. The associated IPFS upload includes:

- The raw documents $\{d_1, \ldots, d_n\}$
- The computed response r
- The query q
- Verification metadata

Verification Workflow

To verify a RAG response, a client must:

1. Retrieve and hash each source document to confirm $h_i = H(d_i)$
2. Confirm each $h_i \in \mathcal{H}_D$
3. Recompute the verification hash from $(q, r, \mathcal{H}_D, \text{metadata})$
4. Verify the signature and on-chain hash commitment

This ensures that every cited source used in generation is publicly auditable, fixed at the time of response, and resistant to tampering.

RAG Security Guarantees

Under standard hash assumptions:

- **Source integrity**: No document can be swapped or altered without breaking the hash.
- **Provenance transparency**: Every fact in r is traceable to a specific document in D.
- **Tamper detection**: Any modification to D or r is detectable by recomputation.

Applications of Verifiable RAG

- Fact-checkable AI answers with cited, immutable sources
- Regulatory compliance and audit trails for enterprise AI
- Trusted medical, legal, and scientific assistance
- Verifiable customer support or document summarization

Future versions of the protocol may integrate document ranking scores, embeddings, and source attribution weights, all stored verifiably as part of the RAG context.

2.5   Agent Extensions (Future Work)

While NeuroSpace currently supports verifiable single-turn interactions and session-level traceability, future protocol extensions will enable verifiable autonomous agents. These agents will perform multi-step reasoning and tool use, with each internal action and output logged as part of a cryptographically traceable execution graph.

Agent Execution Model

Let an agent $\mathcal{A}$ be defined as a stateful process:

$$\mathcal{A} : \Sigma \to \Sigma$$

where Σ represents the agent's internal state, including memory, tools, and contextual history.

An execution trace of the agent is modeled as a sequence of steps:

$$T = [\tau_1, \tau_2, \ldots, \tau_k]$$

Each step τ_i represents a single action and has the form:

$$\tau_i = (\text{input}_i, \text{action}_i, \text{output}_i, \text{metadata}_i)$$

The metadata may include:

- Timestamp of execution
- Tool or function invoked
- Model configuration
- External context or API results

Trace Hash Commitments

Each step τ_i is serialized and hashed:

$$h_i = H(\text{Serialize}(\tau_i))$$

The entire trace is recorded as a directed acyclic graph (DAG), where edges represent causal or functional dependencies between steps. A final root hash h^* can be computed via Merkleization or hash chaining to produce a succinct, verifiable commitment:

$$h^* = \text{RootHash}(T)$$

This root hash may be submitted on-chain and signed by the agent's controlling wallet, or in some cases by a validator node or verifier.

Verification and Auditing

Any third party can retrieve the agent trace from IPFS, verify:

- Internal consistency of the steps (hash chain or Merkle path)
- Signature over the root hash
- On-chain existence of the commitment

Given that each step is tied to a specific model invocation, tool call, or environment interaction, this enables full transparency into agent behavior.

Use Cases for Verifiable Agents

- Auditable reasoning paths for complex questions
- Regulatory transparency in AI-assisted decision-making
- On-chain market of agent tasks with trust guarantees
- Agent-specific performance tracking and reputation scores

Design Considerations

Future versions of the protocol will explore:

- Lightweight proof systems for agent state transitions
- Zero-knowledge variants for privacy-preserving traces
- Staking mechanisms for agent honesty
- Incentives for third-party validators to audit traces

These extensions aim to generalize the NeuroSpace protocol from verifiable messages to verifiable computation, enabling composable and accountable AI agents.

3 Decentralized Model Feedback Layer: Core Protocol Primitive

A critical challenge in AI system design is collecting reliable, transparent, and unbiased feedback on model behavior. In traditional AI platforms, feedback pipelines are centralized, opaque, and entirely controlled by the model provider. This limits community participation, introduces systemic bias, and hides model shortcomings from public scrutiny.

NeuroSpace introduces a universal, decentralized feedback layer as a protocol primitive. This enables users to submit verifiable feedback on model outputs, creating an open, immutable corpus of critiques, performance data, and moderation signals.

Feedback as Verifiable, On-Chain Data

All users interacting with a model on NeuroSpace may submit structured, signed, and timestamped feedback on any response:

$$f = (\text{message\_id}, \text{feedback\_type}, \text{note}, w)$$

where:

- `message_id`: Reference to the specific output under review
- `feedback_type` $\in \{\text{hallucination, inaccuracy, harm, bias, other}\}$
- `note`: Optional free-form explanation
- w: Wallet address of the submitting user

All feedback is stored on-chain or within cryptographically verifiable off-chain logs, forming a tamper-resistant, decentralized repository of model critiques.

Incentive-Aligned Staking and Peer Review

To reinforce the quality and credibility of the feedback, users may stake NSPACE tokens to support or contest feedback submitted by others. This mechanism enables:

- Peer review of flagged responses and critiques
- Economic incentives for honest and high-quality reporting
- Community consensus on the accuracy, safety, and appropriateness of model outputs

Validated feedback increases a user's reputation score and may unlock protocol rewards, while bad-faith or malicious feedback may be penalized through slashing or decreased credibility.

From Private Labs to Public Consensus

By turning model feedback into a public, verifiable protocol primitive, NeuroSpace:

- Removes the moderation and feedback loop from centralized control
- Creates an open, community-driven signal layer for AI models
- Establishes the foundation for decentralized moderation, performance evaluation, and governance of AI systems

This **universal feedback layer is the foundation upon which decentralized moderation and performance alignment are built**, enabling trustless AI ecosystems that are accountable to their users, not just to private companies.

4 Decentralized Content Moderation

Traditional AI platforms tightly control moderation using proprietary filters, moderation APIs, and hidden system prompts. Users have no visibility into these mechanisms and no recourse to dispute or shape moderation policies. This creates a trust deficit, particularly in sensitive or controversial domains.

NeuroSpace transforms moderation into an open, decentralized, and community-governed process, built entirely on the verifiable model feedback layer introduced in Section 3.

Transparent, User-Driven Moderation Signals

Users on NeuroSpace can flag model outputs for:

- `harm`: Harmful, unsafe, or offensive outputs
- `bias`: Evidence of systemic or contextual bias
- `other`: Violations of community-defined policies

These flags are publicly recorded, signed by the reporting wallet, and linked to the message ID, creating an immutable moderation ledger.

Staking-Based Dispute and Resolution Mechanism

To reinforce credible moderation:

- Users may stake NSPACE tokens to support or dispute flags.
- Disputed feedback enters a public review phase, where the broader community votes on the outcome.
- Honest flagging is rewarded; dishonest or malicious flags may result in slashing.

This creates a decentralized moderation system where:

- No single entity dictates moderation rules.
- All moderation decisions are transparent, verifiable, and open to dispute.
- The community evolves its own standards over time, through staking, voting, and governance proposals.

Governance of Moderation Policies

Through DAO governance:

- The community can propose and adjust moderation categories, penalties, and rewards.
- Voting power can be weighted by reputation and stake.
- New moderation workflows or rules can be deployed transparently, ensuring the moderation system remains flexible yet community-controlled.

Removing Black-Box Moderation

By making moderation a protocol-layer primitive, NeuroSpace:

- Removes opaque moderation APIs and hidden system prompts.
- Ensures all moderation decisions are traceable, auditable, and contestable.

- Aligns incentives for honest participation, reducing the risk of censorship or manipulation by centralized actors.

This creates the first fully **decentralized, transparent, and user-governed content moderation protocol for AI outputs**.

5 Decentralized Model Performance Evaluation and Alignment

Beyond moderation, the NeuroSpace feedback layer also enables **transparent, community-driven model performance evaluation and alignment**. In traditional AI systems, model performance is measured and tuned exclusively by the provider, using private datasets, internal metrics, and proprietary evaluation pipelines. This creates a closed ecosystem where users and external developers have no insight into how models perform or how they are optimized over time.

NeuroSpace opens this black box by leveraging the decentralized feedback layer as a public, verifiable data source for model performance monitoring, comparison, and alignment.

Community-Driven Performance Metrics

By aggregating structured feedback on outputs (e.g., hallucination, inaccuracy, bias), NeuroSpace enables:

- Identification of frequently flagged failure modes
- Tracking of model behavior across versions and deployments
- Cross-model benchmarking on safety, trustworthiness, and hallucination frequency
- Open datasets for supervised fine-tuning, reinforcement learning, or retraining pipelines

This shifts the evaluation of model quality from private labs to the public domain, enabling community visibility and participation.

Model Reputation and Leaderboards

All feedback, staking outcomes, and dispute resolutions feed into an on-chain reputation system for:

- Individual model instances and deployments
- Model providers and developers
- Fine-tuned model forks or agent variations

This allows the creation of **public, tamper-resistant model leaderboards**, ranked across various community-defined dimensions (accuracy, safety, helpfulness, etc.).

Feedback-Conditioned Model Optimization

Model developers can leverage the community feedback corpus for:

- Fine-tuning models on real-world critiques and flagged outputs
- Adjusting agent behaviors, prompts, or configurations in response to community-identified issues
- Routing users to models or agents with higher reputation for specific use cases

This creates a decentralized, incentive-aligned ecosystem where model improvement becomes a collaborative, transparent, and community-owned process.

Decentralized Alignment as a Protocol Primitive

Through NeuroSpace, **feedback is not only a moderation mechanism but a core primitive for decentralized AI alignment**, unlocking:

- Open, trustless performance benchmarking
- Community-weighted system prompt adjustments or model routing
- Feedback-driven research and development pipelines

This empowers the broader ecosystem to steer models toward **greater utility, safety, and trustworthiness**, outside the control of any single organization.

Future-Proofed Feedback Modalities

While the initial implementation of the NeuroSpace feedback layer focuses on structured flagging of model outputs (e.g., hallucination, inaccuracy, harm, bias), the protocol is designed to support a **progressive expansion of feedback types and modalities**.
These include:

- **Binary Ratings:** Simple thumbs-up / thumbs-down reactions to model outputs, enabling lightweight positive or negative reinforcement signals.
- **Scalar Ratings:** Numerical ratings (e.g., 1-5 or 1-10 scales) capturing user perceptions of response quality, helpfulness, or satisfaction.
- **Multi-Dimensional Feedback:** Structured forms allowing users to rate outputs across several dimensions simultaneously (e.g., accuracy, informativeness, safety, tone).
- **Feedback-Linked Metadata:** Optional metadata such as user role, context, or use case, enriching the dataset for model alignment pipelines.

All feedback remains cryptographically signed, timestamped, and verifiable, ensuring transparency and auditability regardless of feedback type.

Protocol Extensibility

The feedback schema is designed to be:

- **Extensible:** New feedback types can be proposed and adopted via DAO governance.
- **Composable:** Feedback data can power a range of downstream use cases — moderation, performance evaluation, fine-tuning, and model routing.
- **Flexible:** Developers can integrate the NeuroSpace feedback primitives into third-party apps, agents, or frontends, creating a shared, decentralized feedback commons across platforms.

By supporting both **binary and nuanced feedback mechanisms**, NeuroSpace creates a flexible foundation for decentralized model alignment and community-driven evaluation at scale.

6   Token Economics: NeuroCoin (NSPACE)

NeuroCoin (symbol: **NSPACE**) is the native utility token of the NeuroSpace platform. While initial deployments support Ethereum-based pay-per-query payments, the long-term design centers NSPACE as the economic and governance backbone of the protocol.

6.1 Token Parameters

- **Symbol:** NSPACE
- **Total Supply:** 100,000,000 tokens (fixed)
- **Decimal Precision:** 18
- **Network:** Ethereum-compatible (Base L2 mainnet)

6.2 Utility Functions

NSPACE facilitates secure, programmable, and incentive-aligned interaction with the protocol. Core utility includes:

- **Query Payments:** End users pay in NSPACE to access LLM inference or invoke agent actions.
- **Staking:** Token holders may stake NSPACE to:
  - Prioritize computation requests
  - Guarantee agent execution integrity (future work)
  - Back storage and verification nodes
- **Governance:** NSPACE may be used to vote on protocol parameters, supported models, storage policies, and incentive structures.

6.3 Token Allocation (Tentative)

To bootstrap protocol utility and growth, the initial allocation of NSPACE may follow:

- 40% – Community and Ecosystem Incentives
- 25% – Founding Team and Core Developers (vesting over 3-4 years)
- 20% – Contributors and Advisors
- 15% – Protocol Treasury and Grants

6.4 Economic Design Principles

The NSPACE token adheres to the following design principles:

1. **Sustainability:** Token payments are used to fund verifiers, pinning nodes, and compute infrastructure.
2. **Decentralization:** No centralized authority controls the token supply or protocol evolution after launch.
3. **Incentive Alignment:** All stakeholders—users, model providers, validators, and auditors—are economically incentivized to act honestly.

6.5 Token Lifecycle and Circulation

- **Initial Phase:** ETH-based pay-per-query (current MVP) with soft transition to dual ETH/NSPACE support.
- **Adoption Phase:** Native NSPACE payments required for most core functionality (staking, priority execution, RAG extensions).
- **Maturity Phase:** Governance, model reputation markets, and protocol grants administered entirely in NSPACE.

6.6 Future Considerations

Future versions of the protocol may explore:

- Dynamic pricing mechanisms for token-based model access
- Burn-and-mint cycles tied to compute usage
- ZK-commitments for private token-based queries
- Cross-chain support for broader token composability

7 Protocol Roadmap

The NeuroSpace protocol is designed to evolve through progressive phases of decentralization, feature expansion, and community governance. The roadmap below outlines the key milestones toward a fully trustless, transparent, and user-owned AI infrastructure.

Phase 1 — Core Protocol Launch (Completed)

- Verifiable chat and Retrieval-Augmented Generation (RAG) tools deployed on Base mainnet.
- Moderation and decentralized feedback system live.
- ETH and NeuroCoin (NSPACE) payments integrated.
- NeuroCoin token contract deployed on Base mainnet with 100M total supply, 10M tokens minted.
- 10 free requests per wallet (off-chain) implemented.

Phase 2 — Community Infrastructure (In Progress)

- Reputation scoring for models, sessions, and users.
- Public dashboards for model evaluation, feedback statistics, and audit trails.
- Verifiable agent traces and replayable multi-step reasoning.
- Enhanced on-chain verification UX and developer SDKs.

Phase 3 — Decentralized Governance (Planned)

- Protocol DAO activation with NSPACE-based voting.
- Model registry and open submission framework for new models.
- Grants and protocol treasury governance mechanisms.

Phase 4 — Fully Verifiable Agents and Zero-Knowledge AI (Future Research)

- DAG-based agent execution traces with Merkleized commitments.
- Zero-knowledge proofs for privacy-preserving verifiable AI outputs.
- Agent staking, slashing, and honesty guarantees.
- Incentive mechanisms for third-party validators and auditors.

8 Why Now

The risks of unaccountable AI are no longer hypothetical—they are active, growing, and global. With rapid adoption of LLMs in high-stakes settings, society urgently needs infrastructure to make AI outputs verifiable, attributable, and trustworthy by default.

NeuroSpace provides that foundation—decentralized, transparent, and built for the public good.

Join us at neurospace.chat to explore, contribute, and help define the future of provable AI.

9   Conclusion

NeuroSpace introduces a protocol for verifiable and traceable machine intelligence. By combining cryptographic commitments, decentralized storage, and on-chain attestations, it transforms black-box LLM interactions into transparent, auditable computational artifacts.

Through the layered architecture of prompt hashing, digital signature binding, IPFS-based content addressing, and smart contract verification, NeuroSpace establishes the infrastructure for publicly verifiable AI inference. Its support for session tracking, document-grounded generation, and future agent-level traceability extends this integrity across increasingly complex AI workflows.

The NeuroCoin (NSPACE) token powers this ecosystem by aligning incentives across users, model providers, and protocol validators. NSPACE enables permissionless participation in the platform's economic and governance layers, anchoring its long-term decentralization and sustainability.

What has been built is more than a chat application—it is a live demonstration of Vitalik Buterin's original vision for decentralized applications: transparent, trust-minimized, and credibly neutral. NeuroSpace is not merely a platform; it is a foundation for a new standard of machine intelligence—one in which outputs are not just plausible, but provable.

NeuroSpace: Verifiable. Traceable. Decentralized Intelligence.