Sentient Protocol Paper

Contents

1 Introduction: Problem and Solution

Artificial intelligence has matured into a landscape of powerful, closed models controlled by a handful of corporations. These giants lock up data, algorithms and compute, turning innovation into siloed products and leaving independent builders with few avenues to participate. Yet, the wealth of community-driven research and tooling shows that a different future is possible.

Sentient aims to reshape this landscape by building an open network where thousands of specialized intelligences—models, agents, datasets and evaluators—can collaborate like neurons in a global brain. The protocol addresses three persistent challenges faced by open-source AI artifacts:

1. **Monetization** – How can developers sustain open contributions when income typically flows only to closed-source providers?
2. **Coordination** – How can decentralized contributors align on roadmaps, quality standards and reward allocations without central management?
3. **Distribution** – How can open artifacts reach the right audiences (other developers, enterprises, end-users) and compete with entrenched proprietary platforms?

Sentient tackles these challenges by weaving together four layers:

- **GRID**, a decentralized registry and routing layer for discovering, composing and serving AI artifacts.
- **Protocol**, a suite of smart-contract mechanisms to monetize artifacts, manage contributor rewards and coordinate stakeholders.
- **Technology primitives**, such as model fingerprinting, trusted execution environments (TEEs) and verifiable AI, which provide the security and auditability needed for fair monetization and governance.
- **Tokenomics**, a carefully designed economic system built around the `$SENT` token that aligns incentives, funds contributions and drives adoption.

2 GRID: The Global Research and Intelligence Directory

The GRID—short for Global Research and Intelligence Directory—is the foundational layer of Sentient's architecture. The core components of the GRID are *Artifacts*, which are modular, composable units of intelligence, such as models, agents, data sources, and tools, contributed by our community of builders.

It is designed to act as an open, searchable catalogue of AI artifacts contributed by a global community. Each artifact is a modular, composable unit of intelligence (e.g., a language model, a domain-specific agent, a data source or a tool). Instead of building a monolithic AGI, Sentient's thesis is that intelligence emerges from the cooperation of specialized components. The GRID serves as the canvas on which these components interoperate, enabling workflows that out-perform

any single model. By making participation permissionless and by rewarding useful contributions, Sentient hopes to unlock the latent potential of open AI.

2.1 The Building Blocks of Intelligence

Artifacts on the GRID are identified by canonical IDs and accompanied by rich metadata: version numbers, semantic tags, capability descriptors, licensing information and performance metrics. Contributors can specify inputs and outputs, model sizes, hardware requirements, allowed uses and call rate limits. This metadata enables routing and composition: if a query needs a summarization engine, a search agent and a domain-specific database, the GRID can find the appropriate components based on their tags and reputation scores. The registry is itself decentralized, ensuring no single entity can censor or prioritize certain artifacts without community consensus.

Sentient's GRID already hosts over 110 artifact partners, including more than 50 specialized agents, 50+ data providers, various models and multiple partners for compute and verifiable AI infrastructure. Notable examples include:

- **Dobby**, a crypto-native language model offering expert-level insights into the web3 ecosystem
- **Open Deep Search**, an open-source library enabling agents to perform multi-step web searches and synthesize nuanced information
- **ROMA**, a meta-agent framework for orchestrating teams of specialized AI agents
- **Caldo** (people search and enrichment), **Exa** (search for LLMs), **The Graph** (blockchain data indexing) and **EigenLayer** (integrated compute layer)

These partners illustrate that the GRID is not limited to language models: it aggregates a variety of intelligences and data sources. By onboarding such diverse artifacts, Sentient extends its reach into domain-specific verticals, from crypto insights to scientific research to social analytics.

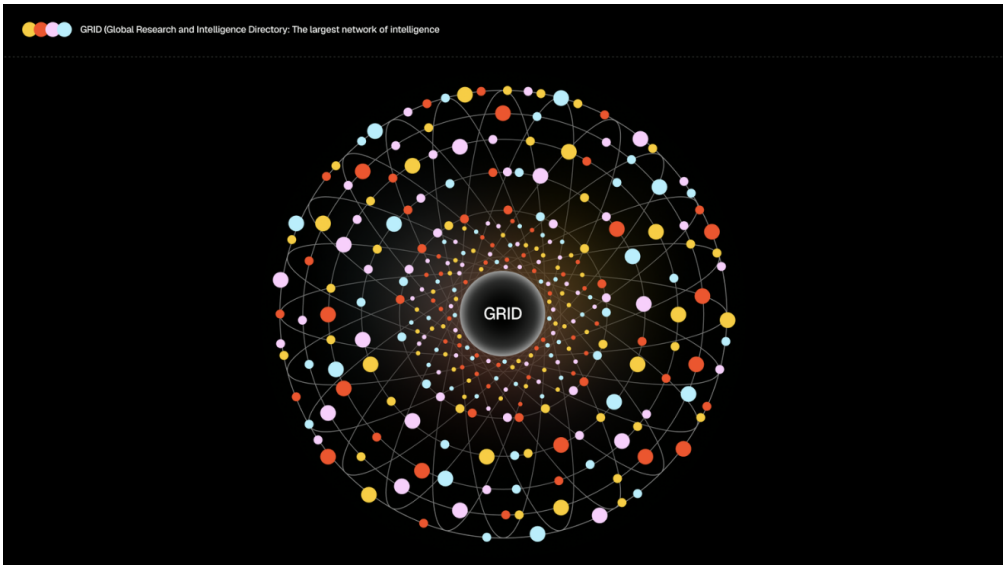

Figure 1: GRID : Global Research and Intelligence Directory

2.2 Workflow Routing and Execution

A query sent to the GRID is split, routed to the right intelligences, enriched with tools like search and domain data, then merged into the best result, giving users a coherent, high-quality output. GRID's output reflects the work of thousands of open-source developers, not a small closed team.

This routing is powered by expert and community-defined workflows. For simpler queries, GRID leverages expert-designed workflows. Each query is analyzed, classified, and routed into the most relevant workflow for that use case.

Example Workflow: Research Analysis

A user asks: "Which European SaaS startups have raised over $50M in the last year?"

The GRID identifies this as a research query and routes it to a specialized workflow:

- **Search**: A search agent compiles an initial list of relevant startups.
- **Research**: A research agent evaluates the founder profiles and company data.
- **Search**: The search agent finds the latest revenue metrics on these startups.
- **Conceptualize**: An infographic artifact visualizes revenue trajectories.
- **Aggregate**: The final, synthesized answer is prepared and delivered.

Workflows are designed for different use cases (e.g., writing, finance, travel) and are curated by community experts, who are rewarded based on their effectiveness.

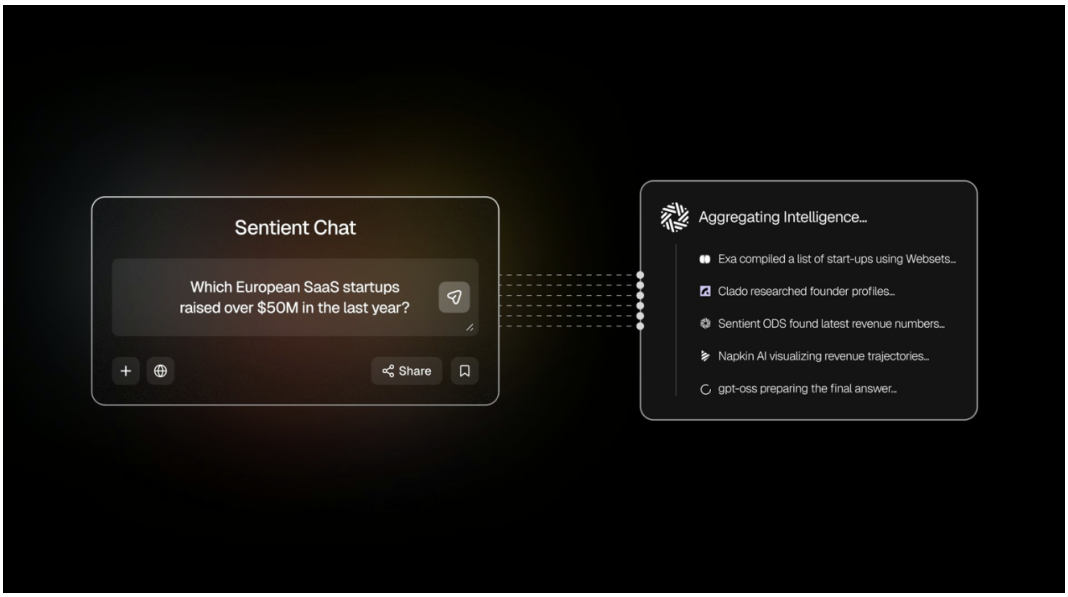

Figure 2: Workflow in GRID

2.3 Builder Experience

The GRID network consists of two key roles: Builders and Users. Builders are developers or organizations who create and deploy AI artifacts; they form the supply side of the GRID. The builder journey comprises three phases:

1. Artifact registration – The first step in the builder journey is to onboard an artifact. During the network's early stages, Sentient curates a high-quality set of partners to ensure that the GRID launches with reliable models, agents and tools. Over time the process becomes permissionless: any builder can register a new artifact by submitting its metadata, defining its inputs and outputs, and declaring licence terms. Builders choose from standard licensing templates—open research, commercial or derivative—which govern how the artifact may be used and how revenue is split. When registering, builders may stake `$SENT` to signal confidence in their contribution and to increase their share of emission rewards (emission rewards are given for providing, maintaining and improving such artifact); they also set a commission on staking rewards, because anyone can stake `$SENT` on artifacts they believe in and emissions are shared between the artifact owner and stakers according to this commission. The protocol performs minimal checks to confirm that metadata is complete and that any required fingerprinting has been applied. Fingerprinting embeds secret triggers into models, enabling owners to detect unauthorised copies and proving that only licensed instances are served (see Section 4).

2. Artifact management Once an artifact is live, builders have a full suite of tools to monitor and manage its service. They can adjust the amount of `$SENT` staked on their artifact, which influences the artifact's `$SENT` emission share; update commission rates charged to delegators who stake on the artifact; and monitor real-time emission statistics and revenue. Builders can release new versions of their artifacts, update license terms, and change per-call prices as market conditions evolve. They can also respond to user feedback by improving documentation or tuning models. Builders can interact with network representatives to influence emission weights. They also track usage metrics and evaluation scores and may coordinate with shared evaluators and TEEs to ensure safety and performance.

3. Supercharging contributors – Sentient is building tools to help projects programmatically manage and distribute a portion of their earnings to individual contributors. Builders can allocate local budgets for tasks (e.g., improving documentation, augmenting datasets) and set criteria for reward splits.

3 The Sentient Protocol Design

The protocol layer operationalizes monetization, coordination and distribution. It comprises smart contracts, governance mechanisms and incentive structures that align the interests of builders, users, stakers and the broader community.

The Sentient network is not centralized under a single operator. Instead, decentralization is achieved

by distributing intelligence across a wide set of independent participants who contribute agents, models, datasets, and compute resources into the GRID. These contributors are economically aligned through the `$SENT` token: they earn emissions, fees, and governance rights based on their participation and usage of the network.

3.1 Monetization

3.1.1 Pricing and licensing

Builders select a licensing template—open research, commercial or derivative—when registering an artifact. They might allow free research use with attribution while charging for commercial deployment. Smart contracts enforce license terms and revenue splits, and Sentient's fingerprinting technology embeds secret query–response pairs into models. These fingerprints allow owners to detect unauthorized copies and prove misuse, ensuring that only authorized usage generates revenue.

3.1.2 Emission budgets and bootstrapping

Open artifacts often face a chicken-and-egg problem: they need usage to generate revenue, but adoption requires resources. To solve this, Sentient allocates `$SENT` emission budgets to curated artifacts. These budgets provide periodic token streams that subsidize the artifact's contributors, hosts and evaluators even before significant user revenue arrives. The emission budgets are finite and subject to periodic re-evaluation by the Protocol DAO. As an artifact gains traction and begins generating revenue, its share of emissions may taper, making room for new entrants.

Emission budgets are guided by signals such as stake weight (community support), expert votes from network representatives and real-world metrics like product usage and revenue. Builders who demonstrate steady adoption and positive user feedback can earn extended budgets; those that fail to gain traction may see budgets reduced, freeing resources for other projects.

3.1.3 Revenue splits and usage flows

When a user or agent invokes an artifact, the call is mediated by a smart contract on Sentient's network. The contract references the artifact's license and splits the payment according to the terms agreed upon during registration. Ownership stakes recorded at registration determine how much of each payment goes to the original builders and maintainers. Payments are then distributed to several stakeholders:

- **Builders and maintainers**: A share proportional to their recorded stake compensates the team that created and maintains the artifact.
- **Hosts**: Model users or operators who provide the compute, bandwidth and storage necessary to serve the artifact may earn a portion of the fee. They must maintain collateral to discourage misreporting and ensure they serve the correct version of the model.
- **Evaluators**: If the artifact relies on third-party evaluators—for example, safety checks or quality assessments—those evaluators can receive a share of the usage revenue.

- **Stakers and representatives**: Stakers on artifacts and network representatives are rewarded indirectly via emission formulas (described in the tokenomics section) rather than from individual usage fees.

To preserve transparency, usage data is recorded on-chain or in verifiable off-chain indexes. This public record not only enables auditing but also prevents double spending—the risk that a dishonest host could reuse the same usage event to claim payouts multiple times. In blockchain parlance, double spending means attempting to spend the same digital token or recorded transaction more than once; by tying each request to a unique entry in an immutable ledger, the protocol ensures that every call to an artifact is counted exactly once.

Fingerprinting is used to embed secret key–response pairs into each distributed model, allowing provers to verify that a host requested permission and paid for usage. By standardizing these flows, Sentient keeps economic incentives aligned while leaving room for artifacts to customize splits through their own governance processes.

3.1.4 Incentive alignment and anti-abuse mechanisms

To discourage misuse, the protocol combines several mechanisms:

- **Fingerprinting** (see Section 4) detects unauthorized model copies. Builders can test suspect deployments by sending secret queries; if a host is serving an unlicensed copy, evidence can be presented to trigger penalties and revenue redirection.
- **Trusted execution environments** (Section 4.2) enforce correct model execution and usage accounting; remote attestation ensures that the binary and model hash match the licensed version.
- **Reputation and staking** – stakers who support an artifact risk losing value if the artifact engages in misconduct. As an indicator of commitment and assurance of service standards, hosts stake `$SENT` to operate and may lose their stake if they misreport usage or serve unauthorized versions. This deters free-riding and fosters cooperation.

3.2 Coordination among artifact contributors

3.2.1 Contributor categories and roles

Sentient recognizes that AI artifacts are built by a diverse set of people: core developers who write code, dataset curators who collect and label data, evaluators who design test suites, infrastructure operators who run compute, documentation authors, community moderators and integrators. Contributors can register one or more roles in the artifact's contributor registry. The registry serves as an on-chain membership list and determines eligibility for rewards and governance rights. By capturing roles explicitly, the protocol ensures that contributions beyond code—like documentation, dataset curation or community outreach—are visible and valued. A single person may wear multiple hats; for example, a developer might also be a maintainer and sit on the evaluator committee.

3.2.2 Three-tier scoring process

After a contribution passes the eligibility check, Sentient helps artifacts to evaluate it using a three-tier scoring system that blends objective measurements, subjective assessments and human judgment. This process aims to capture the full impact of a contribution—whether it improves performance, enhances quality or delivers subtle but important value—so that rewards can be allocated fairly.

Tier 1: Quantitative traction score Objective metrics that are easy to measure, including software signals (e.g., lines of code changed, pull requests merged, total commits), downstream signals (e.g., forks or stars growth), and performance metrics (e.g., inference latency, model accuracy, dataset coverage).

Tier 2: Qualitative evaluation by AI Evaluators and peer reviewers inspect code quality, maintainability, documentation, reproducibility and fairness. Sentient is exploring a domain-specific evaluation model (inspired by SWE-bench) to quantify qualitative impact, including distribution, task complexity and unlocked functionality.

Tier 3: Qualitative evaluation by human experts Maintainers or designated committees can adjust scores within predefined bounds to account for edge cases, ensuring fairness and adaptability. Adjustments are transparent and reviewable via governance.

Weighted scoring and adaptability Each artifact can tailor the weighting of quantitative, qualitative and manual inputs to suit domain-specific priorities, keeping the scoring system relevant across varied repositories.

3.2.3 Local governance and budgets

Artifacts operate as mini-DAOs under the umbrella of the protocol. Each artifact maintains its own treasury funded by revenue and emissions and sets budgets for maintenance, new features and community initiatives. Contributors stake or delegate tokens to gain voting rights; proposals move from off-chain discussion to on-chain voting. Budgets can fund bounties, continuous rewards, retroactive grants and shared infrastructure initiatives. A cadence of regular checkpoints supports re-allocation and refinement of priorities.

3.2.4 Maintainer dashboards and automation tools

Operational tooling enables maintainers to coordinate contributors and manage reward flows via an integrated dashboard surfacing contributor registries, backlogs, budgets and emission stats. Features include:

- Define and fund tasks (bounties, budgets, deadlines; continuous or retro payouts)

- Assign and track contributions (VCS integrations, feedback loops)
- Program reward splits (using the three-step scoring system)
- Automate payouts (smart contracts release funds on completion)
- Analyze contributions (reputation, throughput, average scores, expertise)

3.3 Distribution

Distribution refers to the dissemination and adoption of open-source AI artifacts across developers, enterprise teams and end-user products. Because the protocol is open and composable, artifacts can be registered, composed and reused across workflows, turning distribution into a cooperative network effect.

3.3.1 Distribution goals

Ensure useful artifacts reach the right audiences. Usage drives both revenue and token emissions, incentivizing builders to design broadly applicable components, including for other artifacts.

3.3.2 Leveraging the network of artifacts

Artifacts are both services and consumers; they can be invoked by or call other components. Pipelines (e.g., research agent + search + summarizer + translation + evaluator) propagate usage and revenue upstream. Sentient surfaces these compositions via Chat, dashboards and partner APIs, compounding network effects.

3.3.3 Segmented distribution paths

Developer-centric: SDKs, templates, CLIs, package manager publishing with license metadata, portals and forums. Composability means upstream components accrue usage even when end users never call them directly.

Enterprise: Private deployments via containers or TEEs, enterprise-grade licenses with SLAs, on-prem data handling and audit trails; subscription and usage-based terms that route revenue to creators.

Product: API endpoints and UI components for consumer apps (e.g., summarization, recommendations). Multi-artifact experiences cascade usage credit and revenue up the chain.

3.3.4 Matching and recommendation

A matching engine (inspired by Kaito) connects artifacts to users based on semantic tags (domain, modality, compliance, license) and signals (usage, ratings, stake). High-quality artifacts rise as adoption grows.

3.3.5 Community ownership and distribution incentives

Broader stakeholder ownership amplifies distribution. Temporary incentives (e.g., emission multipliers for milestones) accelerate early adoption and decay over time. Adoption by other artifacts

counts toward usage, encouraging reusable components and joint ventures.

4 Verifiable Trust Primitives

Sentient's protocol is supported by technology primitives that ensure proper accounting, security and verifiability. These primitives are developed in the open and shared with the ecosystem.

4.1 Fingerprinting

Fingerprinting is an AI-native cryptographic primitive that helps prove ownership and detect unauthorized copies of models. Unlike digital signatures on static files, fingerprinting works for generative models whose behavior is defined by billions of parameters. Sentient's fingerprinting library embeds secret query–response pairs into the model during fine-tuning or adapter training. A fingerprint consists of a sequence of triggers and corresponding responses. Triggers are inputs drawn from a distribution unknown to the public; responses are outputs with certain lexical or statistical properties.

4.1.1 Embedding fingerprints

Techniques include:

- **Specialized fine-tuning** – train on a small dataset containing secret trigger pairs, preserving normal performance.
- **Parameter expansion** – use low-rank adapters or prompt-tuning vectors to encode fingerprints without altering base weights.
- **Benign data mixing** – mix triggers with benign data to prevent overfitting and detection.
- **Model mixing** – weighted averaging to bundle multiple fingerprint sets while preserving fidelity.

Fingerprint parameters (key length, number of triggers, redundancy) tune robustness vs. detection time.

4.1.2 Detection and enforcement

1. **Challenge generation** – owner/auditor generates secret triggers.
2. **Model interrogation** – send triggers; fingerprinted models return matching responses with high probability.
3. **Evidence recording** – record detections and verify on-chain against commitments.
4. **Remediation** – redirect revenue, slash offending hosts, suspend licenses; repeat offenders lose hosting rights.

Further resources:

- Fingerprinting library (OML 1.0): https://github.com/sentient-agi/OML-1.0-Fingerprinting
- Research paper (AlphaXiv): https://www.alphaxiv.org/abs/2502.07760

4.2 Trusted Execution Environments (TEEs) — General Overview

TEEs provide hardware-enforced isolation so code and data remain confidential even on untrusted hosts. Enclaves attest the exact code binary and AI model hash used, ensuring proper execution.

4.2.1 Secure Execution

Code, weights and inputs load into an enclave (isolated memory). The host OS/hypervisor cannot read or modify it. After execution, the enclave emits an attestation document; critical actions (payments, key handling, governance) can be gated on valid attestation.

TEEs vs. other verifiable compute

- **ZK proofs**: full cryptographic verification; currently high latency/cost for large models (improving rapidly).
- **MPC**: splits secrets across nodes; reduces single-node trust with protocol/latency trade-offs.
- **TEEs**: near-native speed, general-purpose; trust vendor root-of-trust; be mindful of side-channels. NVIDIA supports GPU TEEs on several datacenter cards (e.g., Phala).

Sentient invests in TEEs for hardware-independent guarantees and negligible overhead, while encouraging exploration of TEE/MPC/ZK options across partners.

4.2.2 Data Privacy and Compliance Examples

Consumer PII: decrypt only inside the enclave, process, then re-encrypt; attestations prove which audited binary/model handled the PII.

Enterprise/regulatory: protect data "in use" for healthcare/finance/legal; enforce TEE-only access with attestable proofs.

Verifiable AI evaluation: audit pipelines by attesting code, model and dataset hashes; enable trustless scoring and benchmark execution.

4.2.3 Sentient Enclaves Framework (SEF)

SEF is an end-to-end stack for confidential AI on AWS Nitro Enclaves (CPU TEEs). Nitro protects user data from the cloud host; Enclaves protect execution from the machine host. SEF adds:

- **Reproducible builds** producing EIFs with deterministic hashes
- **Remote Attestation Web Server** inside the enclave
- **File-system monitor** for unconditional attestation of data ingress/egress
- **Transparent VSock proxies** with optional kill-switch semantics

SEF emphasizes security, ergonomics and reproducibility for verifiable computation.

SEF repo: https://github.com/sentient-agi/Sentient-Enclaves-Framework

4.2.4 Final Considerations

TEE vendor trust (silicon/firmware) is a shared assumption across AWS/Intel/AMD. SEF is open-source and portable to AMD/Intel using EIF on QEMU. This flexibility supports broad protocol applications, from agent hosting to evaluations.

5 Tokenomics

Sentient's network channels economic rights to those who create, maintain and curate its intelligence. Emissions are distributed among constituencies for alignment over time. Significant portions are reserved for builders (developers, researchers, data curators, evaluators, maintainers), early backers (with extended vesting), and the Sentient Foundation (non-profit steward for grants, safety, community and protocol development). An initial airdrop seeds a broad base of stakeholders.

Token Utility

- **Payment for AI artifacts**: users pay for models, agents, datasets and evaluation; contracts route revenue to creators, hosts, evaluators; splits are artifact-governed.
- **Staking and delegation**: holders stake on artifacts or delegate to representatives (Reps) who vote on emissions/policy; delegators share Rep rewards.
- **Governance**: token-weighted voting for emissions, gauges, funding and upgrades; proposals via holders or delegates.
- **Liquidity and composability**: pair with artifact-specific tokens/derivatives; bonding curves/liquidity pools to crowdfund artifacts.

Token Emission Mechanism

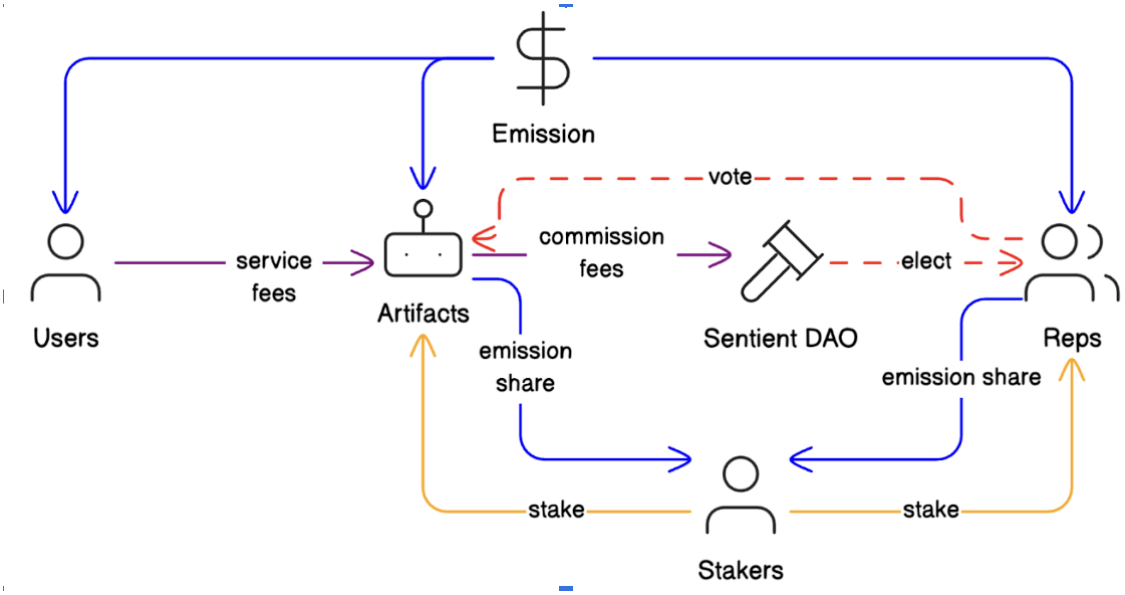

Figure 3: Token Emission Mechanism.

Sentient mints continuously to fund growth and reward participation. DAO controls the annual emission rate (fraction of current supply per epoch) and allocation policy across pools.

1. User incentives – emissions to users of Sentient interfaces (chat, APIs, dashboards, partner apps) to bootstrap network effects.

2. Representative and staker rewards – base stipends to Reps and stake-weighted distribution to stakers. Stakers receive the majority of their Rep's yield. Define:

$$\text{StakeWeight } (SW) = \text{Token\_staked} \times \left(1 + 3 \times \frac{\text{LockDuration (years)}}{\text{MaxLockDuration (4 years)}}\right).$$

3. Artifact emissions – the largest allocation for AI artifacts and contributors. A gauge blends three signals:

$$\begin{aligned} V &\in (0,1) \quad \text{(stake-weighted normalized Rep votes)} \\ \text{SWS} &\in (0,1) \quad \text{(artifact's stake weight share)} \\ \text{RS} &\in (0,1) \quad \text{(artifact revenue share in rolling period)} \\ W &= \lambda V + x \cdot \text{SWS} + (1 - x - \lambda)\text{RS}, \\ &\text{with } \lambda + x \leq 1,\ \lambda, x \in [0,1]. \end{aligned}$$

Governance tunes λ and x to emphasize metrics over time.

This mechanism creates a loop: useful artifacts attract users; engagement generates revenue and signals; emissions reward contributors and stakers; governance rebalances as the network evolves.

Governance

Holders govern the protocol via direct democracy and representative decision-making. The DAO controls economic levers (emission rate, gauge parameters, treasury). Reps are DAO-elected validators or trusted community members who lock stake to assign relative emission weights. Any holder can initiate proposals and vote or delegate. Delegators can migrate stake among Reps; the DAO can appoint or deprecate Reps. Checks and balances aim for robustness and adaptability.

Summary

The Sentient tokenomics design remains true to foundational goals—monetize open-source AI, coordinate contributors and distribute value—while incorporating new protocol features. Distribution ensures broad participation and long-term alignment. Utility turns the token into both fuel and glue. A dynamic emission mechanism funds growth, rewards stakeholders and adapts via gauge parameters and bootstrapping budgets. Layered governance empowers the community to steer the network's evolution with transparency and accountability, creating a resilient, extensible economic framework for a community-owned intelligence network.