OT-RFC-21 Collective Neuro-Symbolic AI

## "Show me the incentives and I'll show you the outcome."

Authors: OriginTrail Core Developers

Date: November 8th 2024

Since the inception of AI in the 1960s, two main approaches have emerged: neural network-based AI and symbolic AI. Neural networks are statistical systems that generate outputs by detecting patterns in training data, while symbolic AI employs deterministic models with explicit knowledge representations and logical connections. Today, transformers within the Large Language Model (LLM) group dominate neural networks, while knowledge graphs are the leading technology in symbolic AI for representing structured knowledge.

Used alone, each approach has limitations. Neural networks are probabilistic and can produce unwanted outputs (hallucinations), risk intellectual property issues, exhibit biases, and face model collapse with a growing amount of AI-generated (training) data online. Symbolic AI, meanwhile, is constrained by its rule-based reasoning, limiting creativity and user experience. Hybrid neuro-symbolic systems combine the strengths of both, leveraging neural networks' usability and creativity while grounding them in knowledge graphs. This approach can enhance reliability, mitigate biases, ensure information provenance, and promote data ownership over IP risks.

OriginTrail Decentralized Knowledge Graph (DKG), together with [NeuroWeb (the AI - tailored blockchain](#)) is surfacing as one of the key components of the symbolic AI branch, enhancing knowledge graph capabilities with the trust of blockchain technology, and powering , **Collective Neuro-Symbolic AI.**
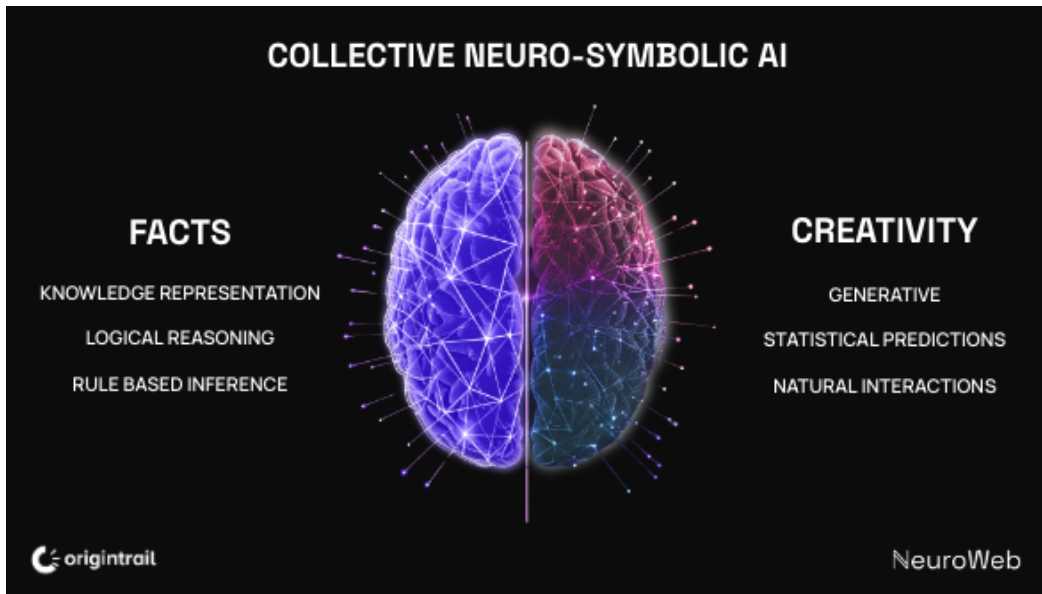
This RFC addresses the following key development milestones to further enhance the Collective Neuro-Symbolic AI and will serve as a basis for one of the most extensive roadmap updates to date:
- DKG V8 Testnet results and learnings,
- DKG Core and Edge Nodes Economics,
- Collective Programmatic Treasury (CPT),
- DKG V8 Mainnet launch in December

"Show me the incentives and I'll show you the outcome."

The quote by Charlie Munger speaks to the importance of setting the right incentives in any system. As the DKG network matures in scalability and adoption, the incentives can become more refined in their implementations and more aligned with supporting the key metric - growth of usage of the DKG network.

There are multiple roles in the OriginTrail ecosystem that are incentivized with both TRAC and NEURO. TRAC is incentivizing Core node operators and TRAC delegators while NEURO incentivises Neuroweb blockchain (Collator) node operators, NEURO delegators, and knowledge publishers (henceforth best represented by DKG EDGE node operators) for incentivized paranets.

The establishment of Collective Programmatic Treasury (detailed in a dedicated section below) will give the most active DKG paranets, by volume of new knowledge assets published to the DKG, an opportunity to take part in building the future of the technology.

The incentives updates and novelties will be released as a part of the DKG V8 mainnet release.
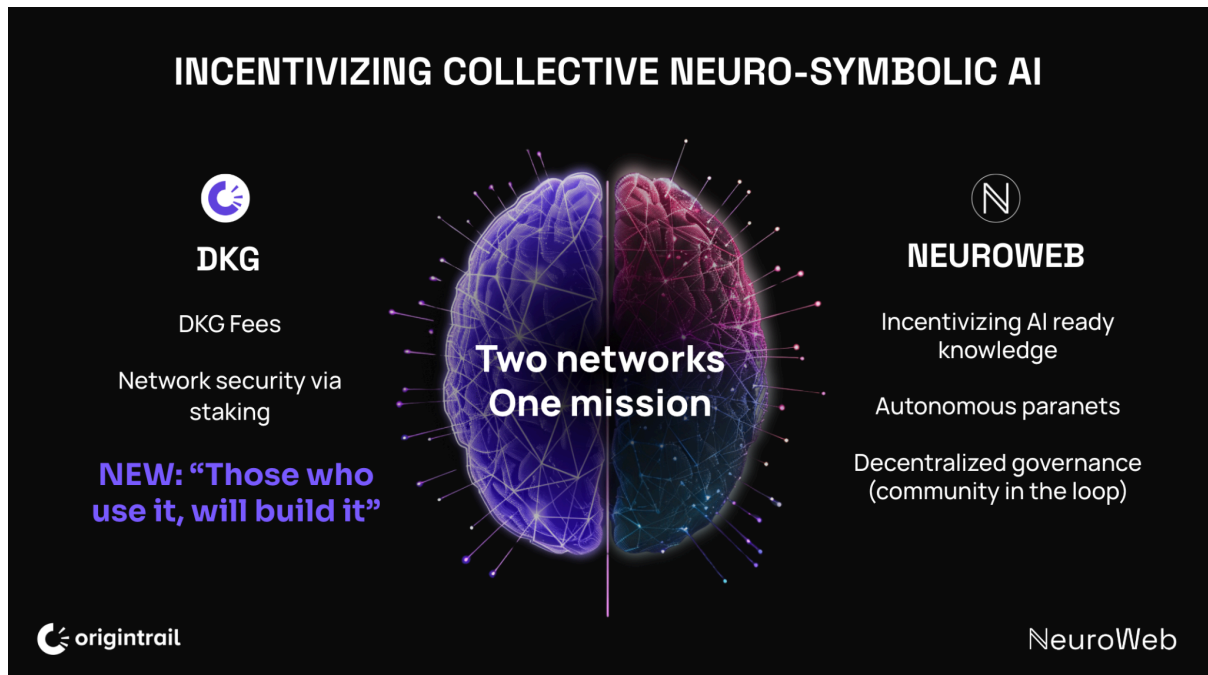
DKG V8 testnet results and learnings

In the first 5 weeks since the DKG V8 Testnet launch, the community has deployed over 500 V8 core nodes, which as part of the incentive program submitted **over 3.7 terabytes and 13.7B lines of core node operational logs,** and over 8 million Knowledge Assets published**.** These have proven very valuable inputs for the core developers who have introduced several optimizations to the DKG based on the submitted telemetry, including performance boosts on the new paranet syncing features, testing curated paranets, and other performance updates.
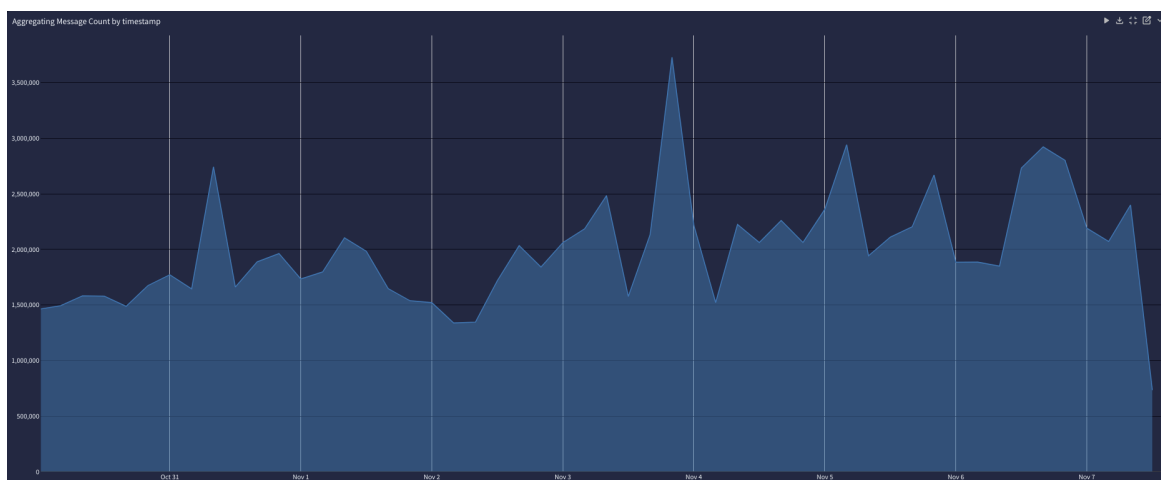

Chart of log lines submitted by V8 Core Nodes telemetry

The number of nodes on the V8 testnet highlights another key insight: even with a fixed reward budget of 100k TRAC, which was allocated to test the behavior of V8 Testnet Core Nodes, achieving an economically viable node count requires the full implementation of the DKG delegated staking feature. **Delegated TRAC acts as a market mechanism to balance the node count according to the rewards available in the network at any time.** This underscores the critical role TRAC delegators will play in maintaining stability and economic balance within the V8 DKG ecosystem.

As the initial phase of the V8 testnet wraps up, advancing V8 features and validating them **requires an environment where all economic incentives** are active to support the full deployment of the DKG V8. Key V8 components, such as the Edge Node and Core Node, will now continue to be **deployed and optimized on the V6 mainnet, with the V8.0 mainnet launch set for December this year.** This launch will initiate the Tuning Period, during which V8 will gain enhanced performance with features like Batch Minting, Random Sampling, and a new staking interface, all backed by real economic incentives.

In addition, synergistic effects between publishers (represented by DKG Edge Nodes, once the V8 network is deployed) and Core DKG Nodes will be fostered through horizontal scaling. This approach aims to **refine network signaling, enabling an optimal network size by aligning the number of nodes more precisely with network demands.**

The details in the following chapters of this RFC create a level playing field to prepare for **updates on existing incentives on the DKG Core node** and access to **Collective Programmatic Treasury (CPT)**.

DKG Core and Edge Nodes Economics

The DKG V8 has been designed with major scalability improvements at multiple levels, with a prototyped implementation tested in collaboration with OriginTrail ecosystem partners from data intensive sectors.

The major advancement that DKG V8 is making is in expanding the OriginTrail ecosystem's product suite to two key products:
- DKG Core Node V8 — highly scalable network nodes forming the network core, persisting the public replicated DKG
- DKG Edge Node V8 — user-friendly node applications tailored to edge devices (phones, laptops, cloud, etc)*

The expansion to more devices is intended to be based on ecosystem builders' capacity and market need.
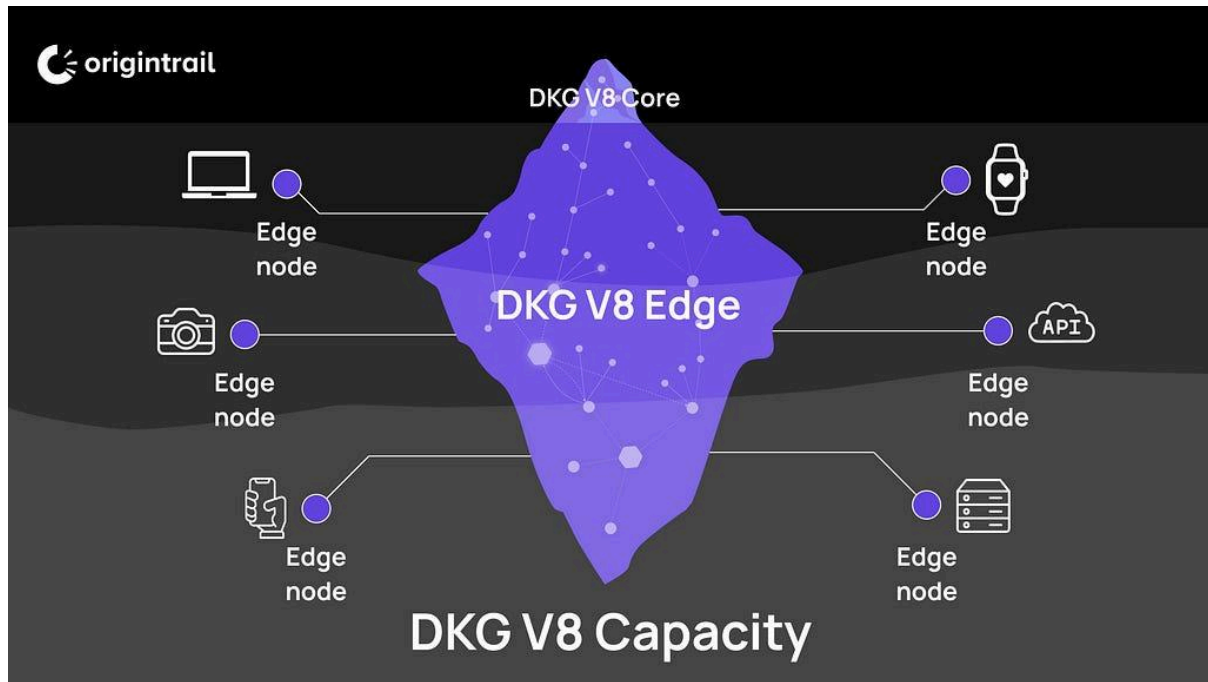
Internet scale with DKG Edge nodes

Edge nodes enable the DKG to reach every part of the internet we know today - any device, any user, any chain. Being a light-weight version of the DKG node, Edge nodes can support both accessing the private and public knowledge on the DKG as well as publishing new knowledge.

Having this capability, DKG Edge node is a very useful tool:
- for paranet operators to enable knowledge miners to publish new knowledge onto their paranets;
- for solution builders as a flexible interface for their neuro-symbolic AI products that can access both private and public parts of the DKG;
- for DKG Edge node operators that want to start publishing to the DKG so they could transform their DKG Edge node into a DKG Core node.

The continuation of V8 development **focuses on teams looking to deploy their paranets & Edge nodes on DKG Mainnet** to generate substantial usage. Therefore, the **DKG Edge Node Inception Program** budget of 750k TRAC is dedicated to builders launching paranets on both the V6 and V8 mainnet, with up to 100k TRAC per builder available as reimbursement for TRAC used for publishing to a particular paranet.

More details on how you can apply for the DKG Edge Node Inception Program can be found here.

Horizontal scaling with DKG Core nodes

The backbone of the DKG network in V8 is formed of **DKG Core nodes**, whose purpose is to ensure secure hosting of the public DKG and facilitate network communication in a decentralized fashion. DKG Core nodes are incentivised through competing for DKG publishing fees in TRAC tokens, which are distributed among the best performing nodes in the network.

The success of a Core node in capturing fees in DKG V6 is currently a function of 3 factors: *(1) node uptime and availability, (2) total TRAC stake delegated to a node* and *(3) network hash distance (enabling efficient knowledge content addressing)*.
Several learnings have been acquired in V6 through the period of the system running in production, most notably on how to improve scalability and further fine tune the incentive system for DKG growth, by updating the relevant parameters in the tokenomics formula.

Particularly, the community of node operators has been indicating the hash distance factor as the most problematic one, causing randomisation and impacting the system in an unpredictable and asymmetric way (the nodes with the same amount of stake and uptime could perform differently in terms of rewards due to a different hash ring network position).

On the other hand, the builders' feedback is that the friction to contributing to the DKG needs to be significantly lower, specifically in terms of publishing price per knowledge asset (addressed with scalability) and accessibility to publishing through available nodes, expressing the need for an approach similar to blockchain RPC services, which allow sending transactions to the blockchain without running a blockchain node.

Therefore V8 introduces an **updated Core node incentive system** with the following factors:

1. **Node uptime & availability,** in positive correlation, as nodes need to prove their commitment of hosting the DKG by submitting proofs to the blockchain (through the new V8 random sampling proof system)
2. **TRAC Stake security factor**, in positive correlation - the more stake a node attracts, the higher the security guarantees and therefore the higher chance of rewards (same as in V6)
3. **Publishing factor,** in positive correlation - the more new knowledge has been published via a specific core node (measured in TRAC tokens), the higher the chance of rewards,
4. **Node fee (formerly "ask"),** in negative correlation - the nodes with lower fees are positively impacting the system scalability, and therefore have a higher chance of rewards.

The illustrative incentive formula is therefore:

$$rewardChance = f_1(uptime,\ t)\ *\ (f_2(nodeStake,\ t)\ +\ f_3(publishingFactor, t)\ -\ f_4(nodeFee, t))$$

where the specific functions are to be validated on both the testnet (for technical functionality) and mainnet (for market functionality) during the V8 Tuning period.

This addition creates further alignment of Core nodes with the ecosystem growth as Core nodes that take up roles of driving adoption will become more successful. Importantly, it also creates an aligned horizontal scaling approach, since additional Core nodes in the DKG become required with growing adoption. This creates a positive self-reinforcing feedback loop: **new adoption leads to new nodes, which leads to increased scale, which unlocks further adoption.** We can imagine core nodes almost acting as a "solar panel" that allows publishers to capture TRAC fees from the network so they could use it for their publishing needs.

Network security via staking

TRAC delegators are using their TRAC to secure the DKG network by delegating it to selected Core nodes. In exchange for a delegation (and increasing the core node's chance of capturing rewards), the node operator splits a part of the captured rewards with the delegators. When selecting the core node to support, the delegators take all the key elements of a successful Core node into account which will, from DKG V8 onwards, include the amount of knowledge added to the DKG.

NEW: Those who use it, will build it: 60MM TRAC Collective Programmatic Treasury (CPT)

To achieve that those who use the network have incentives to build it in the future, the future development fund will be deployed as a 60MM TRAC **Collective Programmatic Treasury (CPT)**. The Collective Programmatic Treasury will be implemented with a programmatic release schedule emitting TRAC to eligible builders. The release schedule will follow the most famous example of emissions in the cryptocurrency space, that of the Bitcoin halving with minor alterations. The TRAC released from Collective Programmatic Treasury will be dedicated to (both conditions should be fulfilled) those who:

- **use TRAC tokens for publishing knowledge** (paranets spending the most TRAC for publishing knowledge), AND
- **have been confirmed eligible for incentives by the community** (paranets who have completed successful IPOs and are deployed on NeuroWeb).

The schedule

As mentioned above, the schedule draws inspiration from likely the most influential schedule process in Crypto, Bitcoin halvings. The halvings principle dictates that half of the outstanding amount is to be distributed in each following period in equal amounts throughout that period. While BTC halvings are set at 4 years, our schedule proposal is to set

this period for 2 years in the case of TRAC. That said, the emissions schedule would be as follows:

1st halving (year 1&2)	Total of 30,000,000 TRAC
2nd halving (year 3&4)	Total of 15,000,000 TRAC
3rd halving (year 5&6)	Total of 7,500,000 TRAC
4th halving (year 7&8)	Total of 3,750,000 TRAC
…	…

The Collective Programmatic Treasury will be deployed on the NeuroWebAI blockchain and will allow paranet operators to trigger *Collect reward* transactions which will calculate the amount of rewards they are eligible for and pay it out accordingly.

The distribution

The distribution amounts will be tied to the core principle of "Those who use it, will build it". The metric which will, therefore, define the amount of TRAC that a builder (represented by their paranet) will receive, is tied to their TRAC spending for creating knowledge on the DKG. A simple example would be as follows:

Paranet A spent 1,000 TRAC
Paranet B spent 2,000 TRAC
Paranet C spent 3,000 TRAC

Collective Programmatic Treasury amount for the period: 600 TRAC

Paranet A: 100 TRAC reward
Paranet B: 200 TRAC reward
Paranet C: 300 TRAC reward

all numbers are placeholders, just exemplifying the relationship between the spent and received amounts..

The Collective Programmatic Treasury will be observing DKG network usage on the innovation hub of OriginTrail ecosystem, the NeuroWebAI blockchain, thus applying only to NeuroWebAI hosted paranets.

The eligibility & humans in the loop

Not every paranet on NeuroWebAI is by default eligible for the TRAC dev fund emissions. In order to achieve that status, a paranet must have been voted in via the IPO process, gaining support by the NeuroWebAI community through a NEURO on-chain governance vote. In this way, the community collectively decides on the dev fund & NEURO incentive emissions, transparently implementing the "humans in the loop" system via on-chain governance.

The Collective Programmatic Treasury (CPT) is expected to be implemented in March 2025.

DKG V8 release timeline

- **November**
    - DKG V8 testnet layer 1 completed
    - OT-RFC-21 release
    - DKG V8 Edge node Inception program start
- **December**
    - **DKG V8.0 Mainnet and Tuning period launch**
    - Neuroweb collator staking
- **January 2025**
    - **DKG V8.1, Tuning period ends**
- **February 2025**
    - Neuroweb TRAC Bridge made available
- **March 2025**
    - DKG V8.2 release - Collective Programmatic Treasury