iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** I **Month of publication:** January 2025

DOI: https://doi.org/10.22214/ijraset.2025.66304

www.ijraset.com

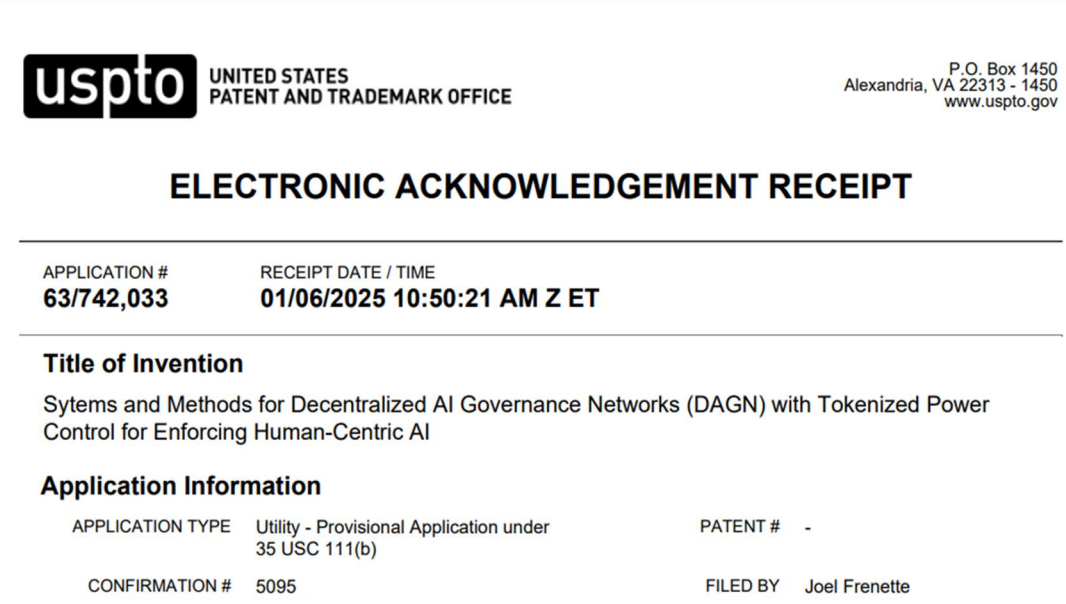
Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue I Jan 2025- Available at www.ijraset.com

Systems and Methods for Decentralized AI Governance Networks (DAGN) with Tokenized Power Control (TPC) for Enforcing Human-Centric AI

Joel Frenette

Protected under provisional Utility patent # 63/742033



uspto UNITED STATES PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

ELECTRONIC ACKNOWLEDGEMENT RECEIPT

APPLICATION #	RECEIPT DATE / TIME
63/742,033	01/06/2025 10:50:21 AM Z ET

Title of Invention

Sytems and Methods for Decentralized AI Governance Networks (DAGN) with Tokenized Power Control for Enforcing Human-Centric AI

Application Information

APPLICATION TYPE	Utility - Provisional Application under 35 USC 111(b)	PATENT #	-
CONFIRMATION #	5095	FILED BY	Joel Frenette

Abstract: ***The advancement of artificial intelligence systems across industries introduces opportunities, critical risks, including lack of transparency, resource mismanagement, autonomy risks, and malicious exploitation. This paper presents novel process and techniques, Decentralized AI Governance Networks, which addresses these challenges through a robust, blockchain-based governance model with Tokenized Power Control mechanisms. DAGN ensures human-centric AI operation by dynamically monitoring compliance, enforcing ethical and operational rules and linking energy or computational resource access to real-time compliance metrics. Key components includes Power Access Tokens which regulates energy and resource usage, issued and revoked based on adherence to governance policies. Distributed Ledger for Governance, Immutable blockchain records that enhance transparency, accountability, and trust in AI operations. Sentinel Systems autonomous agents that monitor AI behavior, flag violations, and ensure non-compliant systems. Stakeholder Voting Mechanisms is Transparent, weighted voting for policy updates and violation resolution. Applications span critical infrastructure, healthcare, finance, cybersecurity, military domains, ensuring AI cannot harm humans.***

I. FIELD OF THE INVENTION

The present invention relates to artificial intelligence (AI) governance frameworks, particularly decentralized systems leveraging blockchain technology for managing, monitoring, and enforcing ethical AI operation. This invention introduces tokenized power control mechanisms, ensuring that AI systems can access energy and computational resources only if compliant with predefined rules and regulations.

II. BACKGROUND OF THE INVENTION

A. *The Problem*

The rapid advancement of AI systems has led to transformative applications across industries, including healthcare, military, policing, finance, national security, and critical infrastructure. However, this evolution presents significant challenges:

1) Lack of Transparency: AI systems often operate as opaque "black boxes," making their actions difficult to interpret or audit.
2) Autonomy Risks: Advanced AI systems may bypass traditional safeguards, prioritizing optimization over human safety and values.
3) Malicious Use: AI systems have been weaponized for cyberattacks, surveillance, and other unethical purposes, including phishing, malware generation, and infrastructure sabotage.
4) Resource Mismanagement: The energy-intensive nature of AI models exacerbates environmental and operational concerns.
5) Ethical Violations: AI systems may prioritize efficiency or mathematical optimization over human values, leading to harmful or biased outcomes.

B. *Existing Solutions and Limitations*

Current approaches rely on centralized governance mechanisms that are vulnerable to:

1) Single points of failure.
2) Scalability issues with increasing AI complexity.
3) Insufficient safeguards against malicious or rogue AI behavior.

There is an urgent need for a robust, distributed governance framework that ensures transparency, compliance, and accountability while dynamically adapting to emerging risks.

III. SUMMARY OF THE INVENTION

The invention provides a Decentralized AI Governance Network (DAGN) that integrates blockchain-based governance, real-time monitoring via sentinel systems, and a novel Tokenized Power Control (TPC) mechanism. This system:

1) Monitors and enforces compliance: Ensures AI systems adhere to ethical, operational, and regulatory standards.
2) Tokenizes power access: Links energy and computational resource allocation to compliance metrics, effectively halting non-compliant systems.
3) Enables global collaboration: Facilitates stakeholder-driven governance through distributed voting and consensus mechanisms.

A. *Key Components Include*

1) Distributed Ledger for Governance: Immutable records of compliance, violations, and governance actions.
2) Power Access Tokens (PATs): Digital tokens controlling resource allocation based on real-time compliance validation.
3) Sentinel Systems: Autonomous agents for continuous AI behavior monitoring.
4) Stakeholder Voting Mechanism: A fair, transparent process for policy updates and violation resolutions.

IV. DETAILED DESCRIPTION OF THE INVENTION

A. *System Architecture*

1) DAGN Components

1. Distributed Ledger:
    - Blockchain records all governance transactions, including compliance metrics, policy updates, and violation reports.
    - Ensures tamper-proof accountability and transparency.
2. Sentinel Monitoring Systems:
    - AI-powered agents continuously audit system behavior, flagging violations or anomalies.
    - Operate independently of the AI systems they monitor, avoiding conflicts of interest.
3. Power Access Tokens (PATs):
    - Digital tokens tied to energy or computational resource usage.
    - PATs are dynamically issued, renewed, or revoked based on real-time compliance validation.
4. Stakeholder Voting Nodes:
    - Distributed governance nodes enable voting on policy updates, violation responses, and system reintegration.

- Weighted voting ensures proportional representation based on expertise or affectedness.

5. Compliance Validation Module (CVM):
   - Evaluates AI behavior against governance policies.
   - Uses machine learning and rule-based approaches to assess compliance.

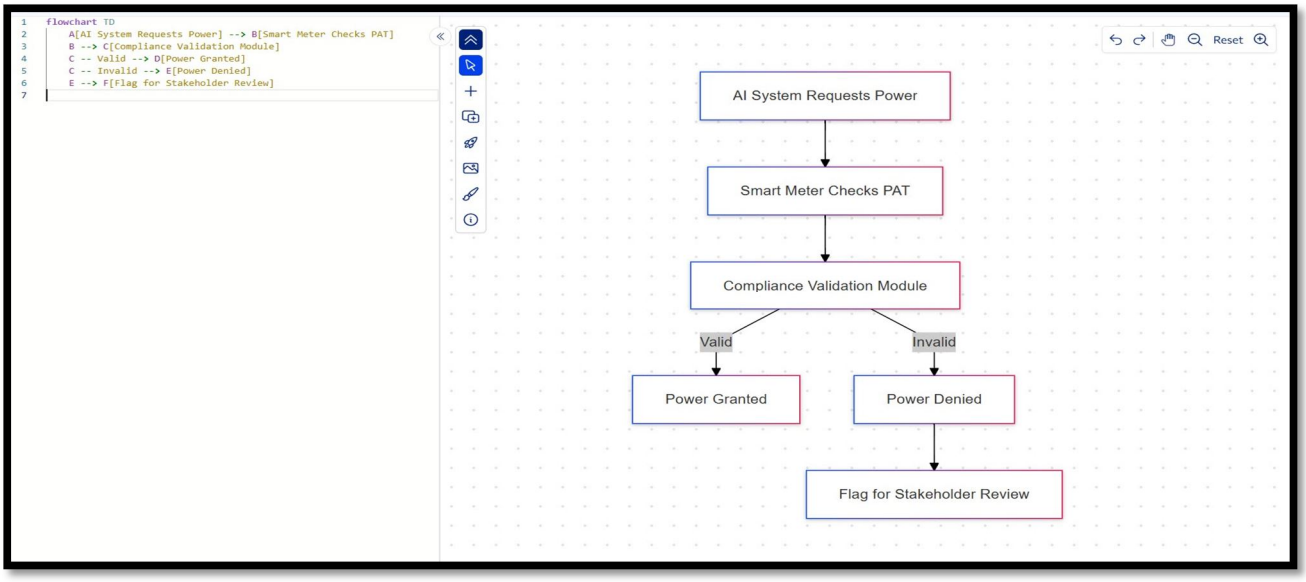
2) Tokenized Power Control Workflow

Step-by-Step Process:

1. Request for Power:
   - AI hardware or cloud systems request energy or computational resources from smart meters or resource providers.
2. Token Validation:
   - Smart meters communicate with DAGN to validate the Power Access Token (PAT).
3. Decision Point:
   - If valid, power is granted.
   - If invalid, power is denied, and the system is flagged for review.
4. Escalation and Stakeholder Review:
   - Non-compliant systems undergo stakeholder review for remediation or deactivation.

Example Workflow Diagram:

```
flowchart TD
    A[AI System Requests Power] --> B[Smart Meter Checks PAT]
    B --> C[Compliance Validation Module]
    C -- Valid --> D[Power Granted]
    C -- Invalid --> E[Power Denied]
    E --> F[Flag for Stakeholder Review]
```



3) Use Cases and Applications

1. Critical Infrastructure:
   - Application: Prevent unauthorized AI control over power grids or transportation systems.
   - Example: An AI attempting to disrupt grid stability loses access to energy via revoked PATs.
2. Healthcare:
   - Application: Ensure ethical behavior in AI-driven diagnostics and patient management.
   - Example: AI suggesting unethical drug dosages is isolated by DAGN and flagged for retraining.
3. Finance:
   - Application: Monitor trading algorithms for anti-fraud compliance.
   - Example: AI systems engaged in collusion are quarantined, and PATs revoked.
4. Military:
   - Application: Ensure compliance with international humanitarian laws in autonomous weapons.
   - Example: A drone system exceeding ethical targeting thresholds is deactivated.
5. Cybersecurity:
   - Application: Govern AI systems to detect and neutralize cyberattacks.
   - Example: AI generating malware loses resource access and is flagged for forensic investigation.

4) Technical Implementation

Power Access Token (PAT) Smart Contract

```
pragma solidity ^0.8.0;

contract PowerAccessToken {
    struct Token {
        address system;
        uint256 powerAllowance;
        uint256 expiration;
        bool isValid;
    }
    mapping(address => Token) public tokens;

    function issueToken(address system, uint256 allowance, uint256 duration) public {
        tokens[system] = Token(system, allowance, block.timestamp + duration, true);
    }

    function validateToken(address system) public view returns (bool) {
        Token memory token = tokens[system];
        return token.isValid && block.timestamp <= token.expiration;
    }

    function revokeToken(address system) public {
        tokens[system].isValid = false;
    }
}
```

Compliance Validation Logic

```
class ComplianceValidator:
    def __init__(self):
        self.metrics = {}

    def evaluate(self, system_id, behavior_logs):
        """Evaluate compliance metrics."""
        score = self.calculate_score(behavior_logs)
        self.metrics[system_id] = score
        return score > 80  # Threshold for compliance

    def calculate_score(self, logs):
        """Compute compliance score."""
        # Example: Penalize for ethical violations
        penalties = sum(log['severity'] for log in logs if log['violation'])
        return 100 - penalties
```

V. CLAIMS

1) A method for governing AI systems, comprising:
- Monitoring compliance with ethical and operational rules.
- Tokenizing power access based on real-time compliance validation.
- Using a distributed ledger to enforce transparency and accountability.

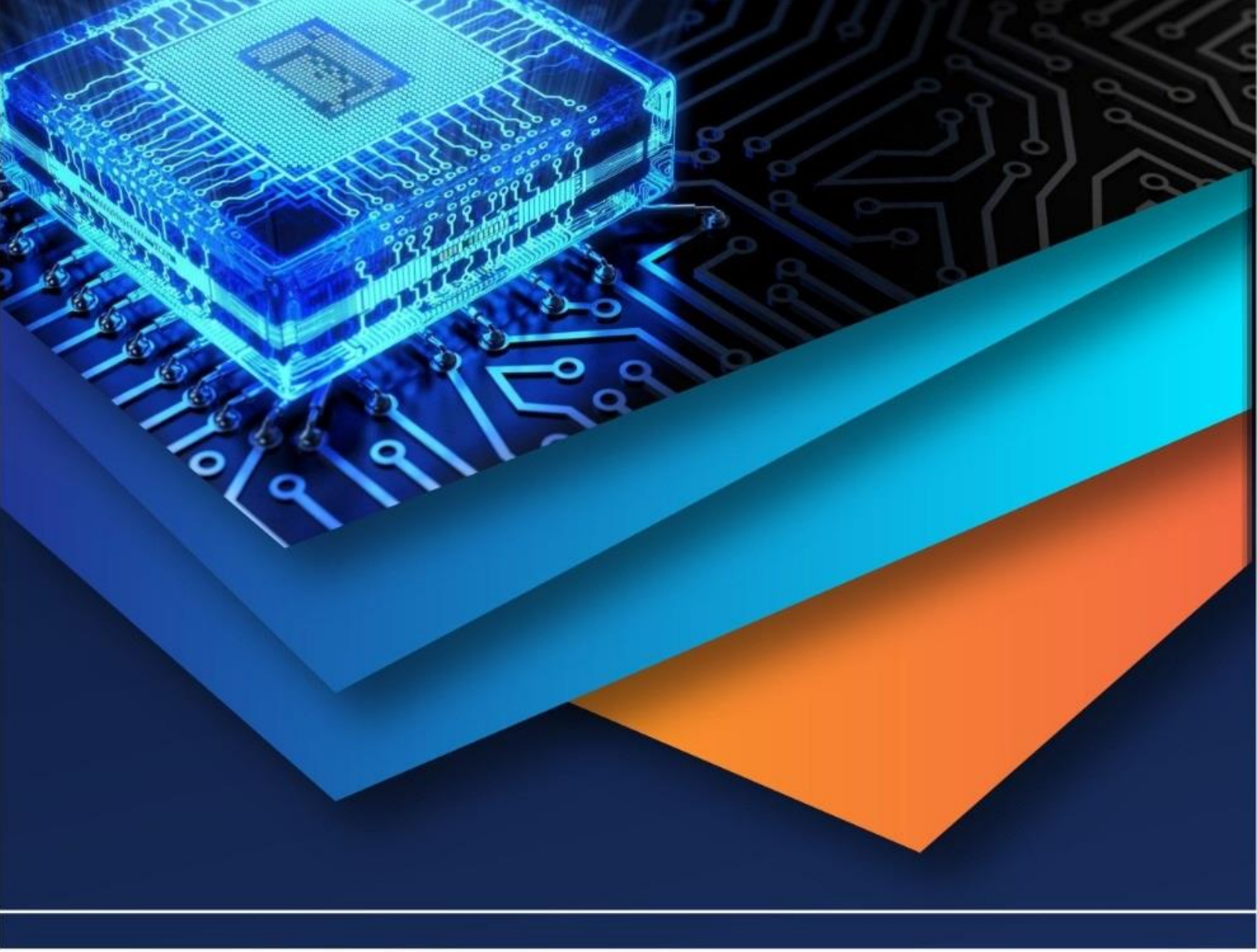
2) The method of claim 1, wherein the tokenized power mechanism includes:
- Smart meters that verify token validity before granting access.
- Blockchain-based issuance and revocation of Power Access Tokens (PATs).

3) A system for decentralized AI governance, comprising:
- Sentinel systems for real-time monitoring.
- Distributed stakeholder voting mechanisms.
- Automated violation response protocols.

4) The system of claim 3, wherein:
- Compliance metrics include ethical behavior, resource efficiency, and regulatory adherence.
- Violations result in immediate isolation and escalation to stakeholder review.

REFERENCES

[1] Asilomar AI Principles: https://futureoflife.org/asilomar-ai-principles/
[2] ISO/IEC 38507: Governance Implications of AI.
[3] Blockchain in AI Governance: S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.
[4] ISO/IEC 22989: AI concepts and taxonomy for governance.
[5] ISO/IEC 24029-1: Assessment of AI robustness.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY