# **Whitepaper**

# 

# 

# 

# 

# 

# **Mutin AI: Empowering AI Agents in Solana DAO Governance**

## **Abstract**

Mutin AI is a visionary open-source plugin built on the Model Context Protocol (MCP) that integrates intelligent agents into Solana’s decentralized governance. It enables AI agents to fetch on-chain **governance proposals**, cast votes on behalf of users, and facilitate decentralized ownership in existing Solana DAOs. By delegating routine decision-making to AI while preserving human oversight, Mutin AI lowers the friction of civic engagement in Web3 and strengthens participation in Decentralized Autonomous Organizations (DAOs). The plugin leverages Solana’s fast finality and low-cost transactions to ensure that autonomous voting actions remain efficient and secure, all while maintaining transparency through on-chain records. This white paper outlines the motivation, design, use cases, architecture, roadmap, and future potential of Mutin AI, illustrating how bridging AI with blockchain governance can transform the way communities self-organize.

## **Background**

Blockchain governance has rapidly evolved with the rise of DAOs. On Solana, the **SPL Governance** framework (accessed via the Realms platform) has become a backbone of the ecosystem, with over 800 DAOs collectively managing more than $1.5 billion in treasury assets. Despite this growth, achieving broad participation remains a challenge. Token holders often struggle to stay informed and engaged across myriad proposals, and many DAOs face difficulty reaching quorum due to low voter turnout. Even dedicated governance delegates find it arduous to keep up with the flow of information and proposals – doing so “requires context, coordination, and cognitive stamina,” a combination that is increasingly hard to sustain without support. In practice, only a fraction of members actively participate in each vote, which can undermine the decentralized ideals of equal voice and collective ownership.

Recent developments in artificial intelligence offer a promising solution. Large language models and AI assistants can rapidly digest information and make reasoned suggestions, acting as co-pilots for decision-making in complex governance. In fact, initial integrations of AI in DAO governance have already begun: most current tools focus on helping humans manage the information overload by summarizing lengthy forum discussions and tracking proposal updates, rather than replacing human judgment[s](https://stablelab.xyz/blog/how-ai-is-reshaping-the-role-of-dao-delegates#:~:text=First%20Integrations%3A%20AI%20as%20Copilot,for%20Human%20Delegates). For example, Solana Labs itself demonstrated a ChatGPT plugin that can fetch on-chain data to help users interact with tokens and NFTs, showcasing how AI can make blockchain usage more accessible and user-friendly.

The Model Context Protocol (MCP), introduced by Anthropic in late 2024, further standardizes how AI models connect to external systems. MCP is an open standard that enables secure, two-way connections between AI assistants and data sources or tools. In essence, MCP acts like a universal adapter, allowing AI systems to interface with any number of external services (databases, web APIs, blockchains) through a single protocol. This provides an opportunity to plug AI directly into on-chain governance processes without needing bespoke integrations for each DAO or platform.

**Why Solana for AI-Powered Governance?** Solana is an optimal blockchain for deploying Mutin AI due to its performance, cost-efficiency, and vibrant governance ecosystem. The network offers throughput up to ~65,000 transactions per second and transaction fees as low as $0.00025, meaning an AI agent can submit votes or query proposals in real time without prohibitive cost or delay. Finality on Solana occurs in seconds or less, ensuring that governance actions by an AI (such as casting a vote) register almost instantly. Equally important, Solana’s DAO tooling is mature and battle-tested – the Realms platform has lowered the barrier to create and run DAOs for less than $0.50 in network fees. This combination of speed, affordability, and ready-made governance infrastructure provides fertile ground for integrating AI. Solana’s design “is built for scale — a blockchain with low costs, high speeds, minimal energy impact, and easy user experience”, which aligns perfectly with Mutin AI’s goal to **reduce friction** and open up decentralized governance to a much wider audience. By building on Solana, Mutin AI ensures that delegating votes to an AI agent is as seamless and economical as possible, encouraging more users to actively partake in their communities.

*Figure: Conceptual logo representing Mutin AI’s integration of AI and governance (the checkmark symbolizes on-chain voting approval). By connecting AI agents to Solana’s governance mechanisms, the plugin empowers every token holder with a tireless personal delegate.* The transformative potential of this approach is profound: AI agents can watch governance proposals 24/7, distill complex discussions into actionable insights, and execute votes according to the user’s preferences or preset rules. Instead of replacing human judgment, Mutin AI acts as an amplifier of human intent – automating the mechanical aspects of governance so that community members can focus on high-level vision and deliberation. In doing so, it strives to **reinforce decentralization** by enabling broader participation. When every holder can easily exercise their voting rights (even indirectly through an AI), power in the DAO becomes more distributed, and decisions better reflect the will of the many rather than the few.

## **Use Cases**

Mutin AI unlocks a range of compelling use cases for different stakeholders in the Solana ecosystem:

* **Hands-free Participation for Token Holders:** An individual SOL or SPL token holder can deploy an AI governance assistant to act on their behalf. The AI agent will automatically fetch new proposals from the DAOs the user is part of, provide summaries or recommendations, and even cast votes based on the user’s specified criteria or consent. This means even passive investors or busy community members can maintain an active voice in governance without manually tracking every update. By lowering the effort to vote, Mutin AI can significantly boost voter turnout and engagement in DAO proposals. In practice, a user might instruct their agent with guidelines (e.g. “vote *Yes* on any funding proposal under $50k that clearly aligns with our roadmap, vote *No* on any that would increase protocol fees”) and let it handle the rest, with the ability to override when needed.
* **Enhanced Delegation for Power Users and Funds:** Large stakeholders, such as venture funds or elected delegates, often juggle governance across multiple DAOs. Mutin AI serves as a co-pilot that can monitor numerous governance forums and on-chain proposal queues simultaneously. It can alert the delegate to important votes, or even vote autonomously according to a predefined strategy. For example, a delegate could configure the AI to automatically vote in line with certain research reports or a publicly declared stance, ensuring consistency and quick responses. This not only lightens the load on human delegates but also brings more predictability and transparency to their voting behavior (the AI can publish a rationale for each vote it casts, drawn from its analysis of the proposal). Overall, the delegate’s effectiveness is amplified, as the AI filters noise and surfaces the signals that matter most across the delegate’s portfolio of DAOs. Experiments in the ecosystem are already pointing in this direction – projects are beginning to build **autonomous agents** that vote, post explanations, and execute on-chain actions with minimal human involvement.
* **Accessible Governance for Newcomers:** For a crypto user unfamiliar with DAO processes, interacting with governance UIs, wallets, and forums can be intimidating. Mutin AI enables a more user-friendly experience – a newcomer can simply chat with an AI agent (via a front-end chatbot or voice assistant) to ask questions like “What proposals are open in the Mango DAO this week?” or “Please vote for the proposal that best aligns with improving security.” The AI handles the technical complexities under the hood: fetching proposal details, explaining them in plain language, and executing the vote transaction once the user confirms. This has the potential to onboard a new wave of participants into web3 civic life by abstracting away jargon and clunky interfaces. By making DAO governance as easy as conversing with a helpful assistant, Mutin AI brings the ideal of inclusive, **user-friendly decentralization** closer to reality.
* **24/7 Monitoring and Rapid Response:** In fast-moving scenarios (such as emergency votes or rapid on-chain polls), human participants might not react quickly enough. An AI agent, however, can watch governance channels and on-chain events around the clock. If a critical proposal is introduced – for instance, a security patch or treasury defense measure that is time-sensitive – the AI could immediately flag it and even cast a provisional vote according to the user’s prior instructions (e.g. default to approving urgent security fixes). This ensures that a user’s stake is never idle or missed in crucial moments. Such automation can make DAO governance more resilient, as decisions can be enacted at machine speed when necessary (while still subject to human review during challenge periods, if applicable). Ultimately, this kind of AI assistance acts as a safety net for decentralized organizations, reducing the chances of malicious or harmful proposals slipping through due to voter inattention.

In all these use cases, **security and trust** remain paramount. Mutin AI is designed so that the user is always in control of their assets and preferences, as discussed next in the architecture section. The goal is not to cede unchecked power to algorithms, but to allow community members to harness AI as a **tool** – one that can tirelessly execute their collective will and uphold their governance principles.

## **Architecture**

Mutin AI’s architecture combines the flexibility of the Model Context Protocol with Solana’s robust on-chain programs to create a secure, transparent agent framework. At a high level, the system consists of three main components: (1) the AI agent (MCP client), (2) the Mutin AI MCP server (plugin) that interfaces with Solana, and (3) the Solana blockchain (specifically, DAO governance programs and relevant data accounts).

* **MCP Client (AI Agent):** This is the AI assistant (e.g. an instance of ChatGPT, Claude, or another LLM-based agent) that interacts with the user and decides what actions to take. The agent runs in an environment that supports the Model Context Protocol, meaning it can dynamically discover and invoke “tools” provided by connected MCP servers. In our case, once connected, the AI will see a set of governance-related tools it can use – such as *getProposals*, *getProposalDetails*, *castVote*, *delegateVotes*, etc. The logic driving the AI’s decisions can be a mix of prompt-based instructions (for example, the user’s voting guidelines embedded in the conversation context) and the AI model’s own reasoning on proposal content. However, crucially, the AI agent does **not** have autonomous access to sign blockchain transactions on its own. It must go through the Mutin AI plugin’s secure functions, which in turn require valid cryptographic signatures from the user’s wallet for any state-changing action.
* **Mutin AI MCP Server (Plugin):** The MCP server is the bridge between the AI and the Solana blockchain. Implemented following the open MCP specification, this server exposes a standardized API of tools and resources that the AI can call. For instance, when the AI wants to fetch proposals, it calls the *getProposals(dao\_public\_key)* tool. The plugin then uses Solana’s RPC interface or an indexing service to retrieve the list of active proposals from the specified DAO’s governance program accounts. Similarly, a *castVote(proposal\_id, choice)* tool will construct and submit a transaction to the SPL Governance program to record a vote on-chain. Each tool is defined with strict input/output schemas (using the MCP SDK and Solana libraries) to ensure the AI’s requests are safe and valid. The plugin effectively sandboxes what the AI can do: the agent can only perform predefined actions (like reading proposal data or drafting a vote transaction) and nothing beyond. This design minimizes the risk of malicious prompts causing unintended blockchain actions. The MCP server can be run by end-users themselves or provided as a hosted service by the Mutin AI project – in either case, it operates **non-custodially**. This means the server does not hold the user’s private keys; instead, it will either prompt the user’s wallet to sign a transaction (if using a wallet adapter integration) or require the user to preload a signing key in a secure enclave for the AI agent to use. By keeping key management on the user side, Mutin AI ensures that the power to authorize transactions remains firmly in human hands.
* **Solana Blockchain (DAO Programs):** On the other side of the plugin lies the Solana network, where all governance state is maintained. Solana’s SPL Governance program defines accounts for **realms** (DAO instances), **proposals**, **votes**, and more. When *getProposals* is called, the plugin queries a Solana node (or cache) for all proposal accounts under a given Realm (filtering by state = active). It then parses each proposal’s data (e.g., title, description link, current vote tallies, time remaining) and returns that to the AI. For *getProposalDetails*, the content of a specific proposal (such as the full text or attached documents) can be fetched if available on-chain or via IPFS links. For *castVote*, the plugin constructs an instruction to the governance program’s **Cast Vote** endpoint with the user’s chosen option and includes the necessary authorization. All votes are recorded immutably on-chain, which means the actions taken by an AI on behalf of a user are fully transparent and auditable. Other possible tools might include *delegateVotes(to\_public\_key)* if a user wants to delegate their voting power to another key (which could itself be an AI agent’s key), or *proposeAction(name, description)* if in the future Mutin AI supports AI-initiated proposal creation. Each of these on-chain interactions benefits from Solana’s high throughput and fast confirmation times – the AI can retrieve fresh data and submit transactions within seconds, enabling a smooth interactive experience.

*Figure: Conceptual illustration of an AI governance agent analyzing a DAO proposal on Solana. In Mutin AI’s architecture, the AI (depicted here as a digital assistant) uses the plugin to query on-chain data and then casts votes according to the user’s directives.*

To maintain security, several layers of safeguards are implemented in Mutin AI’s design. First, the MCP server requires authentication and permissioning – only authorized AI clients (with user consent) can access the governance tools, preventing any unauthorized usage. Second, any transaction the AI attempts to execute is first simulated (dry-run) on Solana before submission, with the results fed back for verification. This ensures, for example, that a vote transaction is valid and will succeed before it’s actually sent to the network. Third, Mutin AI incorporates rate limits and circuit breakers – for instance, restricting an agent to one vote per proposal per user and capping the number of transactions it can send per hour – to prevent errant behavior or spam.

Finally, the Mutin AI platform is **fully open source**, allowing the community to audit the MCP server code and even inspect the prompts or logic that guide the AI agents. This transparency is critical for trust – it ensures the AI acts as intended and enables the community to quickly catch any bugs or biases. Overall, the architecture strives to balance **empowerment** (giving AI tools to streamline governance) with **control** (giving users and developers oversight and constraints over those tools).

A noteworthy aspect of Mutin AI’s design is the emphasis on transparency not just at the code and transaction level, but also in decision rationale. We envision the AI agent providing explanations for its actions whenever possible. For example, if an AI votes “No” on a proposal, it could attach a short justification drawn from its analysis (e.g., “No – this proposal exceeds the quarterly budget and community sentiment on the forum is against it”). These rationales could be logged in an off-chain database or even recorded on-chain as a reference (if the governance platform allows adding a reason field or via an associated memo). This way, other DAO members can understand *why* an action was taken, facilitating accountability for AI delegates similar to human delegates. In essence, every vote cast by Mutin AI can be accompanied by an explanation, combining the efficiency of automation with the **accountability** of human-readable intent.

## **Roadmap**

The development and deployment of Mutin AI is planned in iterative phases, gradually expanding its capabilities and integration depth:

1. **Month 1 – Lightning MVP:** Ship core MCP plugin on devnet → mainnet with read-only proposal feed and wallet-signed one-click voting for 3 flagship Solana DAOs.
2. **Months 2-3 – Turbo Agent:** Add multi-DAO monitoring, rule-based auto-voting, user preference templates, and complete first external security audit.
3. **Months 4-6 – Ecosystem Plug-ins:** Embed Mutin AI inside Realms UI, major Solana wallets, and Discord/Telegram bots; roll out AI-assisted proposal drafting and forum-sentiment summaries.
4. **Months 7-12 – Autonomous Governance Era:** Pilot AI delegate committees managing routine treasury/parameter changes under on-chain safeguards, then expand cross-chain and transition Mutin AI to its own community DAO.

Each phase of the roadmap will be executed with careful consideration of community feedback and ethical guidelines. As we add more autonomy and intelligence to the system, we will do so incrementally and with robust oversight, constantly validating that Mutin AI is augmenting governance in a positive way. By the final phase, we aim to demonstrate that AI-driven governance can be **trustworthy, effective, and widely adopted** – setting a precedent for the future of DAO operations.

## **Future Potential**

The integration of AI agents into blockchain governance via Mutin AI heralds a new paradigm for decentralized organizations. By empowering individuals with AI delegates, we move toward a world where governance is both highly efficient and deeply inclusive. The potential implications are far-reaching:

* **Stronger DAO Participation:** If every token holder can effortlessly stay informed and vote (either personally or via their AI proxy), the average voter turnout in DAOs could surge dramatically. This addresses one of the key weaknesses in many current DAOs, where low engagement often leads to decisions being made by a small core group. Early indicators already suggest this trend – by late 2024 there were reportedly over 10,000 crypto AI agents active across domains like DeFi, trading, and governance, projected to reach 1 million by 2025. Governance-specific agents are poised to boost effective participation in ways not seen before. Rather than centralizing power, delegating to personal AI agents could distribute influence more evenly, as each participant – large or small – has an ever-vigilant assistant ensuring their voice is counted. This mitigates the historical tendency for voter apathy to concentrate power in the hands of a few.
* **Augmented Decision Quality:** AI’s analytical capabilities can improve not just the quantity of participation, but also its quality. An AI can digest vast amounts of data – from on-chain metrics to off-chain news and social media sentiment – when evaluating a proposal. It might detect anomalies or risks that a casual human voter could miss, providing an additional layer of scrutiny on every decision. For instance, AI summarization is already helping organizations like MakerDAO and Arbitrum DAO automatically extract key points from complex proposals and discussions and integrated sentiment analysis has been used to gauge community support in Aave’s governance (correlating positive forum sentiment and high participation with successful outcomes) In the future, Mutin AI agents could cross-reference proposals with external data (market conditions, risk alerts, regulatory news) to forecast the potential impact of a decision, giving communities a form of “predictive governance”. This **data-driven approach** to decision-making, guided by AI, could lead to more informed and rational outcomes, reducing the influence of hype or misinformation on DAO votes.
* **Reinforcing Decentralized Ideals:** The ultimate promise of Mutin AI is to help realize the original ideals of DAOs – **autonomy, transparency, and inclusivity** – on an even greater scale. By automating the execution of collective decisions, DAOs can become more truly autonomous (less dependent on constant human micromanagement for routine tasks). Every action an AI agent takes is transparently recorded on-chain, bolstering the auditability and trust in the governance process. Inclusivity is enhanced when language or expertise barriers are lowered by AI: for example, an AI could translate proposals into multiple languages for global communities, or simplify highly technical proposals into plain language summaries so that all members can understand and participate. There is even the possibility of AI agents representing underrepresented viewpoints – for instance, a “devil’s advocate” bot that analyzes proposals from a skeptical perspective and votes No unless certain quality criteria are met, adding an algorithmic check-and-balance to avoid groupthink. In short, AI assistance can make decentralized governance more accessible and equitable, reinforcing the principle that *anyone* with a stake can have a meaningful say in a DAO’s direction.
* **Emergent New Governance Models:** As human and AI collaboration in DAOs matures, we may see novel governance structures emerge. **Agentic DAOs** – organizations where AI agents participate alongside (or even independently of) humans in governance – could move from theory to practice. Imagine a DAO where certain decisions or budget allocations are managed by AI committees operating under human-defined rules. For example, a research DAO might entrust an AI agent collective with allocating micro-grants for promising proposals, with humans only intervening for large or contentious decisions. Alternatively, DAOs could implement “programmable governance” where constitutional rules are enforced by AI (e.g., automatically vetoing any proposal that violates core principles or exceeds specified risk limits). These scenarios blur the line between smart contracts and intelligent agents, potentially reinventing coordination at scale. Mutin AI lays the groundwork for this future by normalizing AI participation in governance and establishing best practices for safe human-AI coordination in a trustless environment.

Of course, embracing AI-driven governance is not without challenges. It raises important questions around trust, control, and ethics: *How do we ensure an AI’s decisions align with the community’s best interests? Who is accountable if an AI makes a flawed governance choice? Could widespread AI delegation inadvertently centralize influence in the hands of those who deploy the most sophisticated agents?* These are areas that the community will need to address through experimentation, open dialogue, and possibly new on-chain governance safeguards (for example, requiring periodic human review of AI-made decisions, or mandating diversity in the AI models used by large token holders to avoid single points of failure). The design of Mutin AI builds in transparency and user control as fundamental features to tackle some of these concerns from the start. As one commentator noted in exploring the convergence of AI and DAOs, the key is finding a balance where autonomous agents “create new capabilities while presenting novel challenges in trustless coordination”. We believe that with prudent safeguards and an open governance approach to the tool itself, the benefits of AI-augmented governance will far outweigh the risks.

In conclusion, Mutin AI represents a forward-looking convergence of two revolutionary domains: artificial intelligence and decentralized blockchain governance. By leveraging the speed and scalability of Solana, it shows how AI can be harnessed to strengthen the very foundations of decentralization – empowering individuals, enhancing transparency, and enabling organizations to coordinate at unprecedented scale. The road ahead is both exciting and demanding. As we implement this vision, we invite Solana developers, DAO members, and the broader crypto community to participate in shaping Mutin AI. Together, we can pioneer a new era where **AI agents and human stakeholders work side by side** to govern decentralized communities, making the ideals of Web3 more achievable than ever before.