

# **The Institutional Deficit in Decentralized Autonomous Organizations**

*An Empirical Analysis of Forty DAOs*

**Wulf A. Kaal**

*Professor of Law, University of St. Thomas School of Law*

***Abstract***

This Article presents the first systematic empirical analysis of institutional architecture across decentralized autonomous organizations. Forty operational DAOs spanning eight industry segments: investment and DeFi, base-layer infrastructure, data and analytics, decentralized science, oracles and tooling, civic and political coordination, NFT collectibles, and gaming and virtual worlds. These segments are evaluated against a thirteen-category institutional rubric derived from the Calcaterra-Kaal framework. The framework synthesizes Arrow's Impossibility Theorem, the Folk Theorems of repeated games, and Incomplete Contract Theory into a proof that rule stability is institutionally self-defeating and that cooperative governance requires architecture that governs its own evolution. Five institutional patterns hold across every segment of the dataset. First, a visibility paradox: categories that produce visible artifacts (token launches, treasury balances, marketplace activity) score consistently above the midpoint, while categories that produce invisible governance infrastructure (legal wrappers, judicial branches, AI alignment policies, on-chain reputation ledgers) score consistently below it. Second, a universal AI-governance vacuum: AI Alignment scores 2.10 of 10 dataset-wide with no DAO scoring above 5, the only category in the framework where no entity crosses the midpoint. Third, token-plutocracy as the default governance form, with dataset-wide Decentralization at 5.17 and no production deployment of reputation-weighted on-chain aggregation. Fourth, legal-wrapper heterogeneity without convergence: eight distinct wrapper structures appear across the forty entities, with only one DAO using the Wyoming DAO LLC statute. Fifth, a convergent architectural agenda for institutional repair built around five upgrades: ERC-1155 multi-token reputation, tripartite separation of powers, stablecoin treasury infrastructure, weighted directed acyclic graph historiography, and values-drift detection. The unweighted dataset mean of 67.3 of 130 (51.8 percent) is the central quantitative finding: the median DAO has implemented roughly half of the institutional architecture the framework prescribes, with a projected post-upgrade mean of 95.3 representing a 42 percent improvement available through the convergent agenda. The deficit is structural rather than incidental. DAO architecture has solved the problems for which it was originally designed, decentralized capital formation and programmable value transfer, and has not yet solved the problems that emerged after its design, AI-mediated governance, Sybil-resistant identity, and constitutional separation of powers. The visibility paradox explains the under-investment: invisible institutional infrastructure is systematically underprovided relative to visible institutional infrastructure even when the invisible infrastructure is more predictive of long-run resilience. The Article develops implications for legal scholars, regulators, and DAO operators.

**Keywords:** decentralized autonomous organizations, DAO governance, institutional architecture, new institutional economics, reputation systems, reputation-weighted voting, weighted directed acyclic graph, validation pools, tripartite separation of powers, Arrow's Impossibility Theorem, Folk Theorem, Incomplete Contract Theory, mechanism design, dynamic regulation, blockchain governance, smart contracts, ERC-1155, token-weighted voting, flash-loan governance attacks, Sybil resistance, AI alignment, autonomous agents, agent integration, values-drift detection, stablecoin treasury, legal wrappers, Wyoming DAO LLC, Cayman foundation, empirical legal studies, Computative Economics

**JEL:** K22, G34, D02, L86, O33, D71, D72, D86, K20, G28, C72, P48

## **Contents**

|                                                                                  |           |
|----------------------------------------------------------------------------------|-----------|
| <b>I. Introduction</b>                                                           | <b>4</b>  |
| <b>II. Theoretical Framework: The Calcaterra-Kaal Institutional Architecture</b> | <b>6</b>  |
| A. Arrow's Impossibility Theorem and the Limits of Static Aggregation            | 6         |
| B. The Folk Theorems and the Defeat of Static Rules                              | 7         |
| C. Incomplete Contract Theory and the Limits of Specification                    | 7         |
| D. The Synthesis: Governance That Governs Its Own Evolution                      | 8         |
| <b>III. Methodology</b>                                                          | <b>9</b>  |
| <b>IV. The Thirteen-Category Rubric</b>                                          | <b>10</b> |
| <b>V. The Forty-DAO Dataset</b>                                                  | <b>12</b> |
| A. The Master Scoring Matrix                                                     | 13        |
| B. Cross-Cohort Category Means                                                   | 15        |
| <b>VI. Findings</b>                                                              | <b>16</b> |
| A. Finding 1: The Visibility Paradox                                             | 16        |
| B. Finding 2: The AI-Governance Vacuum is Universal                              | 17        |
| C. Finding 3: Token-Plutocracy is the Default                                    | 19        |
| D. Finding 4: Legal-Wrapper Heterogeneity Without Convergence                    | 21        |
| E. Finding 5: A Convergent Architectural Agenda for Institutional Repair         | 22        |
| <b>VII. Segment-Level Analysis</b>                                               | <b>24</b> |
| A. Data and Analytics (mean 76.2)                                                | 26        |
| B. Tools and Protocols (mean 71.8)                                               | 26        |
| C. Infrastructure (mean 73.4)                                                    | 27        |
| D. DeSci (mean 71.6)                                                             | 27        |
| E. Investment and DeFi (mean 63.4)                                               | 28        |
| F. NFT (mean 65.6)                                                               | 28        |
| G. Gaming and Virtual Worlds (mean 60.8)                                         | 28        |
| H. Political and Civic (mean 55.4)                                               | 29        |
| <b>VIII. The Architectural Repair Trajectory</b>                                 | <b>29</b> |
| A. The Convergent Five Upgrades                                                  | 29        |
| B. The Implementation Gap                                                        | 31        |
| C. The Path-Dependence Constraint                                                | 31        |
| <b>IX. Implications</b>                                                          | <b>32</b> |
| A. Implications for Legal Scholars                                               | 32        |
| B. Implications for Regulators                                                   | 33        |
| C. Implications for DAO Operators                                                | 34        |
| <b>X. Limitations</b>                                                            | <b>35</b> |
| A. Source-Document Constraint                                                    | 35        |
| B. Static Snapshot                                                               | 36        |
| C. The Rubric's Theoretical Specificity                                          | 36        |
| <b>XI. Conclusion</b>                                                            | <b>36</b> |

## **I. Introduction**

The decentralized autonomous organization is now an established legal and economic form. The universe of operational DAOs spans every major application area of decentralized governance: investment and lending protocols, base-layer infrastructure, data and analytics networks, decentralized science consortia, oracle and tooling protocols, civic coordination experiments, NFT collectibles communities, and gaming guilds and virtual worlds. Total assets under DAO governance have crossed the hundred-billion-dollar threshold. The set of jurisdictions that have promulgated DAO-specific legal infrastructure has grown to include Wyoming, Tennessee, Vermont, the Marshall Islands, Switzerland, the Cayman Islands, Singapore, the British Virgin Islands, Liechtenstein, and the United Arab Emirates. The form is no longer an experiment.

It is also not yet an institution. That is the gap this Article documents.

The DAO began in 2016 as a smart-contract-only governance form. The original ambition was that code, deployed to an immutable public ledger, would substitute for the institutional architecture that human-organized cooperative bodies had built over four millennia. The substitution proved inadequate within months. The original DAO experiment was hacked and the response from the Ethereum community required a hard fork: a discretionary intervention that could not have been encoded in the smart contract the fork undid.<sup>1</sup> The lesson was clear at the time and remains clear now: code alone is not institution. The institutional infrastructure that supports cooperative governance has to be rebuilt for the DAO form, not assumed away by it.

In the decade since, DAO architecture has accumulated layers. Token-weighted voting was the first layer. Multisig treasury control was the second. Off-chain forum deliberation followed by on-chain execution emerged as a third. Snapshot-based governance with Tally execution emerged as a fourth. Optimistic governance with delay periods became a fifth. Subdao delegation, working-group structures, and grant programs filled out the operational scaffolding.

---

<sup>1</sup> The DAO hack of June 2016 drained approximately 3.6 million ETH and was reversed only through a contentious hard fork that produced the Ethereum/Ethereum Classic split. See generally Quinn DuPont, "Experiments in Algorithmic Governance: A History and Ethnography of 'The DAO,' A Failed Decentralized Autonomous Organization," in *Bitcoin and Beyond* (Routledge 2018).

Legal-wrapper structures ranging from Cayman foundations to Swiss Stiftungen to Wyoming LLCs addressed the regulatory gap. Each of these layers solved a real coordination problem. The cumulative architecture is genuinely impressive when measured against the bare smart-contract-only governance with which the form began.

It is also incomplete. That is what the empirical analysis presented here documents. Across forty DAOs spanning eight industry segments, evaluated against the thirteen-category framework developed in Calcaterra and Kaal, the modal DAO has implemented approximately half of the institutional architecture that the framework prescribes. The deficit is not random. It clusters in identifiable categories. It tracks visible-versus-invisible institutional infrastructure. It is largest precisely in those categories where the next decade of DAO governance will be tested most severely: the integration of autonomous AI agents into governance processes, the formal separation of legislative, executive, and judicial functions, and the reputation-based coordination of cooperative work among parties who cannot rely on token holdings as a proxy for trustworthiness.

The Article proceeds in eleven Parts. Part II sets out the theoretical framework underlying the thirteen-category rubric, drawing on Arrow's Impossibility Theorem,<sup>2</sup> the Folk Theorems of repeated games,<sup>3</sup> and Incomplete Contract Theory<sup>4</sup> as the three formal results that motivate the institutional architecture. Part III details the methodology. Part IV explicates the thirteen-category rubric. Part V presents the forty-DAO dataset. Part VI presents the five findings. Part VII conducts segment-level analysis. Part VIII analyzes the architectural-repair

---

<sup>2</sup> Kenneth J. Arrow, *Social Choice and Individual Values*, 2d ed. (New Haven: Yale University Press, 1963).

<sup>3</sup> James W. Friedman, "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies* 38, no. 1 (1971): 1–12; Drew Fudenberg and Eric Maskin, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica* 54, no. 3 (1986): 533–54.

<sup>4</sup> Sanford J. Grossman and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94, no. 4 (1986): 691–719; Oliver Hart and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy* 98, no. 6 (1990): 1119–58.

trajectory. Part IX develops the implications for legal scholars, regulators, and DAO operators. Part X surfaces limitations. Part XI concludes.

## **II. Theoretical Framework: The Calcaterra-Kaal Institutional Architecture**

The thirteen-category rubric applied to the dataset derives from Calcaterra and Kaal, *Decentralization: Technology's Impact on Organizational and Societal Structure* (De Gruyter 2021).<sup>5</sup> The framework rests on three formal results from twentieth-century institutional economics, each of which generates a constraint that any cooperative governance system has to satisfy. Together, the three results constitute a proof that rule stability is institutionally self-defeating. The framework's response is not to find a better static rule set but to build a system that governs its own evolution.

### **A. Arrow's Impossibility Theorem and the Limits of Static Aggregation**

Kenneth Arrow demonstrated in 1951 that no social welfare function can satisfy a small set of intuitive desiderata simultaneously: unrestricted domain, non-dictatorship, Pareto efficiency, and independence of irrelevant alternatives. The theorem is not a critique of any particular voting mechanism. It is a structural proof that no static aggregation of preferences can faithfully represent collective will under the conditions Arrow specified. Every voting system makes tradeoffs. Plurality voting fails the Condorcet criterion. Borda voting fails independence of irrelevant alternatives. Single-transferable-vote satisfies more criteria but introduces complexity that creates strategic-voting opportunities. Approval voting makes different tradeoffs again. There is no escape: every aggregation rule violates at least one of the desiderata.

The implication for DAO governance is direct. Token-weighted voting, the modal aggregation rule across the dataset studied here, is not a neutral way to register member preferences. It is one specific choice in a space of choices, none of which is neutral. The Calcaterra-Kaal framework treats Arrow's theorem as a permanent structural constraint that motivates a different design philosophy: rather than trying to find the static aggregation rule that minimizes its violations, the architecture should be designed to evolve its aggregation rule in response to the specific failures

---

<sup>5</sup> Craig Calcaterra and Wulf A. Kaal, *Decentralization: Technology's Impact on Organizational and Societal Structure* (Berlin: De Gruyter, 2021).

it experiences in operation. The framework does not abolish voting. It treats voting as one component in a larger system that includes reputation accumulation, reference-based revaluation of past contributions, and protocol-development proposals that themselves enter the same governance process they propose to modify.

## **B. The Folk Theorems and the Defeat of Static Rules**

The basic Folk Theorem establishes that in infinitely repeated games with sufficiently patient players, any individually rational outcome can be supported as a subgame-perfect equilibrium. The strong implication is that patient strategic actors will exploit any static rule set given sufficient time. This is not a claim about specific actors or specific rules. It is a structural claim about the relationship between strategic patience and institutional design.

For DAO governance, the Folk Theorem implications are severe. Any DAO that operates with a static rule set, in a strategic environment with patient capital and high stakes, faces the Folk-Theoretic prediction that the rule set will eventually be gamed. The empirical record of DAO governance attacks confirms this prediction in operational reality. The Beanstalk Farms exploit of April 17, 2022, in which an attacker borrowed approximately one billion dollars in stablecoin flash loans to acquire two-thirds of the protocol's governance tokens for a single block and used that supermajority to pass a malicious proposal that drained approximately 182 million dollars from the protocol, illustrates the pattern in stark form:<sup>6</sup> token-weighted governance offers no defense against an attacker willing to acquire tokens temporarily. The response cannot be to find a static rule set that is harder to game. The response, in the Calcaterra-Kaal framework, is to build a governance system whose rules themselves evolve in response to gaming attempts.

## **C. Incomplete Contract Theory and the Limits of Specification**

The basic claim of Incomplete Contract Theory is that no contract can specify all future contingencies, because the cost of specifying contingencies grows faster than the value of marginal specification, and because some future states cannot be specified at all in the language

---

<sup>6</sup> MacKenzie Sigalos, "Beanstalk Farms loses \$182M in DeFi governance exploit," CoinDesk, April 17, 2022, <https://www.coindesk.com/tech/2022/04/17/attacker-drains-182m-from-beanstalk-stablecoin-protocol>.

available at the time of contracting. The implication is that every contract is incomplete in some respect, and that the institutions that govern the contract have to fill the specification gaps as they emerge.

For DAO governance, Incomplete Contract Theory generates a constraint complementary to Arrow and the Folk Theorems. Even if the Arrow problem could be solved (it cannot) and even if the Folk Theorem problem could be solved (it cannot), no smart-contract specification can anticipate every contingency. The original DAO experiment of 2016 illustrated this concretely: the smart contract specified what would happen if a member tried to withdraw funds, but it did not specify what would happen if the withdrawal logic was exploited by a recursive call. The specification was incomplete in a way that turned out to be catastrophic. Every smart contract since 2016 has been incomplete in similar ways, even when audited, formally verified, or run for extended periods on testnets. The institutional response cannot be to write more complete contracts, because that path is closed by the theorem. The response has to be to build governance institutions that can fill the specification gaps as they emerge.

#### **D. The Synthesis: Governance That Governs Its Own Evolution**

The Calcaterra-Kaal framework synthesizes the three formal results into a single institutional design philosophy. Arrow demonstrates that no static aggregation mechanism perfectly captures collective preferences. The Folk Theorems demonstrate that patient actors will game any static rule set. Incomplete Contract Theory demonstrates that no specification can anticipate all future states. Together, the three results constitute a proof that rule stability is institutionally self-defeating. The framework's response is to build a system that governs its own evolution rather than a system that aspires to a stable rule set that the three theorems show cannot exist.

The architectural mechanism through which the framework operates is the weighted directed acyclic graph (WDAG). Every governance action, work evidence, protocol-development proposals, comments, and votes, is a vertex in the graph. References between posts are directed, weighted edges that revalue past contributions in light of present consensus. The graph is acyclic because references point only backward in time, ensuring that the institutional record is cumulative and non-reversible. The critical structural feature is that protocol-development posts, which propose changes to the governance rules themselves, enter the same WDAG as

work-evidence posts, are evaluated by the same validation pool mechanism, and are subject to the same reference-based revaluation. The rules and the rules about rules coexist in the same graph, governed by the same mechanism.

Reputation, in the framework, is the variable that the WDAG accumulates and that the validation pool weights. A contributor's reputation is the cumulative reference-weighted score of past contributions, with each reference revaluing prior contributions in light of present consensus. This is the framework's response to Arrow: rather than aggregate preferences through voting alone, the system aggregates preferences through a citation graph that resembles the academic citation graph and the PageRank graph in functional structure. The validation pool is the framework's response to the Folk Theorem: rather than rely on a static voting rule, the system uses a randomized sampling mechanism with stake-weighted commitment that is structurally resistant to gaming because the rules of validation themselves can be revised by validation. Reference-based revaluation is the framework's response to Incomplete Contract Theory: rather than try to specify all future contingencies, the system updates the value of past contributions as new contingencies emerge.

The thirteen-category rubric operationalizes this institutional architecture into a set of observable categories. Each category corresponds to a specific function the framework requires the DAO to perform. The rubric is the framework, operationalized for empirical assessment.

### **III. Methodology**

Forty DAOs were evaluated against the thirteen-category rubric. The DAOs were drawn from eight industry segments, with five DAOs per segment. The segments span the major application areas of decentralized governance: investment and DeFi, base-layer infrastructure, data and analytics, decentralized science, oracles and tooling, civic and political coordination, NFT collectibles, and gaming and virtual worlds. Every DAO was evaluated against the same rubric on the same zero-to-ten scale, with totals out of 130.

Source material for each evaluation consisted of publicly available documentation: project websites, whitepapers, GitHub repositories, governance forum archives, Snapshot voting histories where applicable, and Etherscan or equivalent on-chain inspection of treasury

composition and proposal-execution records. Where DAOs were in transition during the evaluation window, ApeCoin DAO during its AIP-596 transition, Illuvium during the November 2025 treasury event, Bio Protocol during the veBIO transition, Mantle during ZK-rollup migration, the score reflects institutional architecture observable at evaluation time. The dataset is a snapshot, not a trajectory.

The scoring scale is ordinal but treated as cardinal for arithmetic operations (sums and means) because the rubric specifies that each unit of score corresponds to an institutional feature whose presence or absence is observable in source documentation. A score of 10 reflects full institutional implementation of the rubric's prescriptions for the category; a score of 1 reflects rudimentary implementation; a score of 0 reflects no observable institutional infrastructure. The unweighted total of 130 reflects equal weighting across categories. Readers who prefer weighted totals can apply their own weights to the master scoring matrix in Part V.

The figures presented throughout were generated from a single canonical CSV using matplotlib at 300 dots per inch. The CSV and figure-generation code are available upon request.

## **IV. The Thirteen-Category Rubric**

The thirteen categories correspond to the institutional functions the Calcaterra-Kaal framework identifies as essential to decentralized governance. Each category is scored on a zero-to-ten scale, with the maximum total of 130 points reflecting full institutional implementation of the framework. The categories are described below.

**Decentralization (Dec).** The distribution of governance power across the DAO's membership. The category considers token concentration (Gini coefficient or equivalent), voter participation rates, the existence and binding force of vote-delegation mechanisms, and the structural incentives toward concentration or distribution that the protocol creates. The framework identifies Decentralization as foundational: a DAO that is not actually decentralized cannot meet the framework's other requirements regardless of how it scores on peripheral categories.

**Attack Resistance (Att).** Institutional defense against governance attack vectors: token-economic attacks (governance token accumulation, vote-buying, flash-loan governance), smart-contract attacks (reentrancy, oracle manipulation, upgrade-mechanism exploitation), Sybil

attacks (identity fragmentation across wallets), and social-engineering attacks against multisig signers or core developers.

**Governance (Gov).** The formal architecture of legislative, executive, and judicial functions within the DAO. The framework identifies tripartite separation of powers as a foundational institutional feature, drawing on Montesquieu, the Federalist Papers, and the institutional-economics literature on checks and balances.

**Fundraising (Fnd).** Institutional capacity to aggregate capital, including primary token launch, secondary fundraising rounds, treasury growth from protocol revenues, and grants from external funders. The category measures mature, multi-channel capital aggregation with diversified treasury composition.

**Payment System (Pay).** Smart-contract automation of payments, including grants, contributor compensation, treasury allocations, and protocol fees. The category measures fully automated, multi-token payment infrastructure with stablecoin rails, vesting mechanisms, and clawback provisions.

**Regulatory Compliance (Reg).** Institutional compliance infrastructure: legal-wrapper structure (foundation, LLC, Stiftung, association), securities-law positioning, KYC/AML compliance where applicable, tax-reporting infrastructure, and engagement with regulators.

**Work-to-Earn (W2E).** Institutional infrastructure for compensating contributors based on documented work, including bounty programs, grant systems, contributor-reputation tracking, and on-chain proof-of-work mechanisms. The framework treats Work-to-Earn as essential institutional infrastructure for cooperative coordination because it directly implements the reputation accumulation the WDAG architecture requires.

**Impact Creation (Imp).** Institutional infrastructure for translating member contributions into measurable real-world outcomes, including impact-tracking dashboards, third-party verification of impact claims, and explicit linkage between protocol outputs and the DAO's stated mission.

**Chain Agnosticism (Chn).** Institutional independence from any single underlying blockchain. Multi-chain deployment, cross-chain governance synchronization, and chain-independent

treasury management. The framework treats Chain Agnosticism as a hedge against host-chain governance failure.

**Consulting (Cns).** Institutional capacity to export governance learning to other organizations through consulting services, knowledge-transfer programs, or productized governance tooling.

**Agent Integration (Agt).** Institutional infrastructure for incorporating autonomous AI agents into governance processes, including agent identity and authorization, agent-to-DAO communication protocols, agent-vote procedures, and agent accountability mechanisms.

**AI Alignment (AI).** Institutional infrastructure for constraining the autonomous AI agents that participate in governance, including bias-auditing procedures, human-override mechanisms, reasoning-transparency requirements, and values-drift detection.

**Operational Scalability (Ops).** Institutional capacity to scale governance operations as membership and asset base grow, including governance-process scalability (delegated voting, working-group structures, subDAOs), technical scalability (L2 deployment, sharding, modular architecture), and contributor-onboarding scalability.

## **V. The Forty-DAO Dataset**

The dataset is composed of five DAOs from each of eight industry segments. Table 1 presents the segment composition. The cohort spans the major application areas of decentralized governance: investment and DeFi protocols evaluated against credit infrastructure and lending mechanisms; base-layer infrastructure and naming-service protocols; data and analytics networks anchored by indexing, storage, and decentralized data marketplaces; decentralized science consortia spanning longevity research, pharmaceutical IP, and life-sciences platforms; oracle protocols, smart-account infrastructure, and governance-as-product tooling; civic and political coordination experiments; NFT collectibles communities with governance treasuries; and gaming guilds and virtual-world economies.

| Segment        | DAOs Evaluated                                                |
|----------------|---------------------------------------------------------------|
| Investment     | Goldfinch, HairDAO, Maple Finance, Nexus Mutual, Ondo Finance |
| Infrastructure | Polkadot, ENS DAO, Gnosis DAO, Mantle, Golem Network          |

|                  |                                                                            |
|------------------|----------------------------------------------------------------------------|
| Data & Analytics | The Graph, Filecoin, Ocean Protocol, Streamr, Vana                         |
| DeSci            | ResearchHub, VitaDAO, Molecule, Bio Protocol, OriginTrail                  |
| Tools/Protocols  | DeXe DAO, UMA DAO, inSureDAO, Orbs Network, SafeDAO                        |
| Political/Civic  | Fort Worth DAO, ConstitutionDAO, CityDAO, BanklessDAO, Ethereum Foundation |
| Gaming           | Beam DAO, Decentraland, YGG, Star Atlas, Illuvium                          |
| NFT              | TerraForms, Meebits, BAYC (ApeCoin), MAYC (ApeCoin), More Loot (AGLD)      |

*Table 1. Dataset Composition: Forty DAOs Across Eight Industry Segments.*

The segment-level mean totals range from 55.4 (Political/Civic) to 76.2 (Data & Analytics), a range of approximately twenty points or sixteen percent of the maximum total. The segments group into three tiers. The high-scoring tier is composed of Data & Analytics (76.2) and Infrastructure (73.4), driven by protocols with mature institutional architecture. The middle tier is composed of Tools/Protocols (71.8), DeSci (71.6), NFT (65.6), and Investment/DeFi (63.4). The low-scoring tier is composed of Gaming (60.8) and Political/Civic (55.4). The tier distinction is one of the most consistent patterns in the dataset and is developed in Part VII.

### **A. The Master Scoring Matrix**

The master scoring matrix is presented in Figure 1 in heatmap form. Each row is one DAO; each column is one of the thirteen categories. Rows are sorted by total score descending. Cell shading reflects the zero-to-ten scale, with deeper shading indicating higher scores. The visual structure surfaces the central pattern of the dataset immediately. The left-side columns tend toward the middle of the score range with consistent variation across DAOs. The right-side columns, Consulting, Agent Integration, and AI Alignment, are noticeably lighter across the entire dataset. The heatmap is also striated by segment, with Data & Analytics rows clustering near the top, Political/Civic rows clustering near the bottom, and the other segments distributed in the middle.

Figure 1. Cohort Heatmap: Forty DAOs × Thirteen Categories  
Rows sorted by total score (descending)

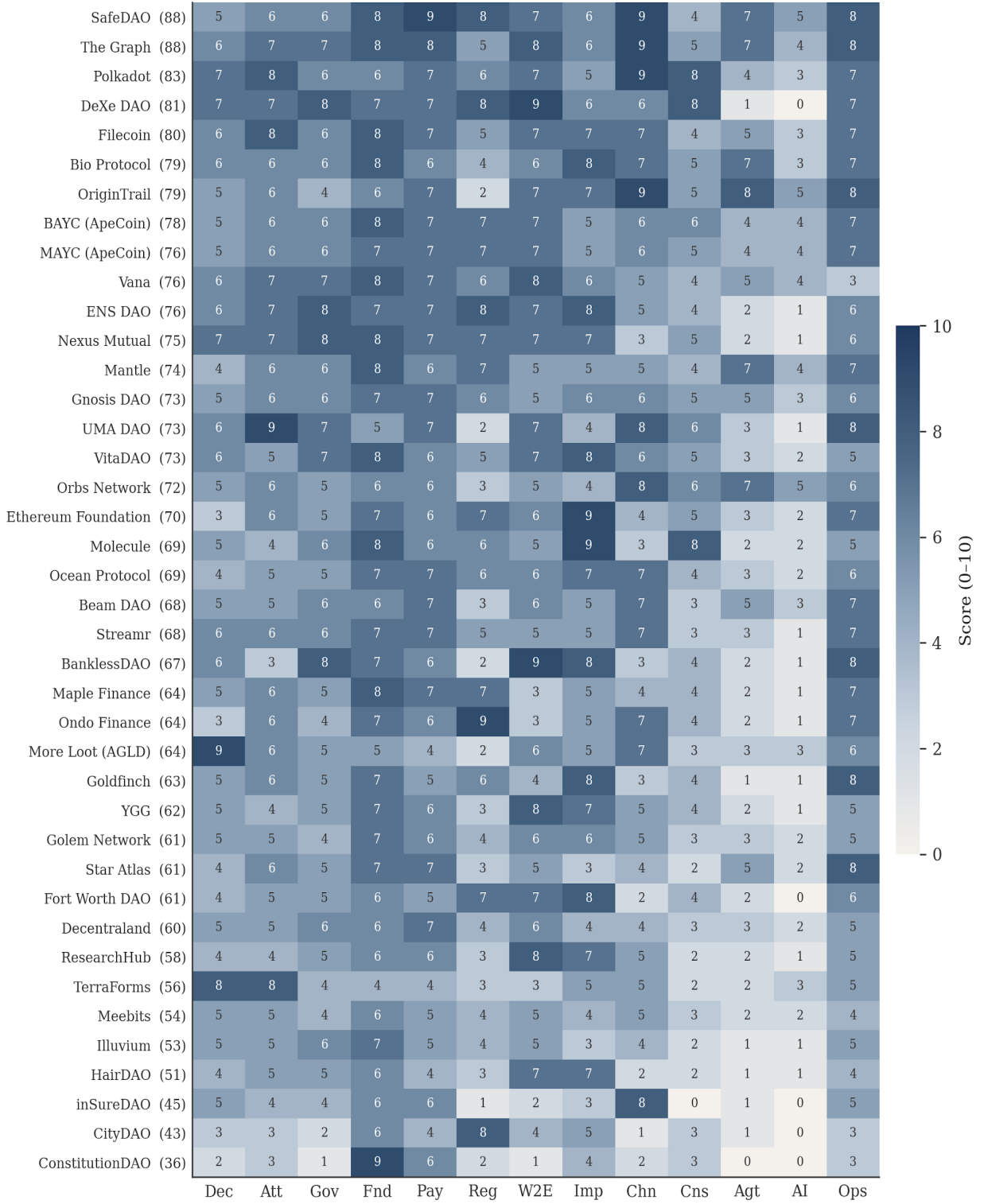


Figure 1. Cohort Heatmap of Thirteen-Category Scores Across Forty DAOs. Rows sorted by total descending.

## B. Cross-Cohort Category Means

Figure 2 presents the dataset-wide category means with standard deviation bars, sorted from highest to lowest mean. The capital-formation institutions (Fundraising at 6.88, Payment System at 6.25) score one to two standard deviations higher than the AI-governance institutions (Agent Integration at 3.30, AI Alignment at 2.10). The deficit pattern is structural rather than incidental: DAO architecture has solved the problems for which it was originally designed, decentralized capital formation and programmable value transfer, and has not yet solved the problems that emerged after its design, AI-mediated governance, Sybil-resistant identity, and constitutional separation of powers.

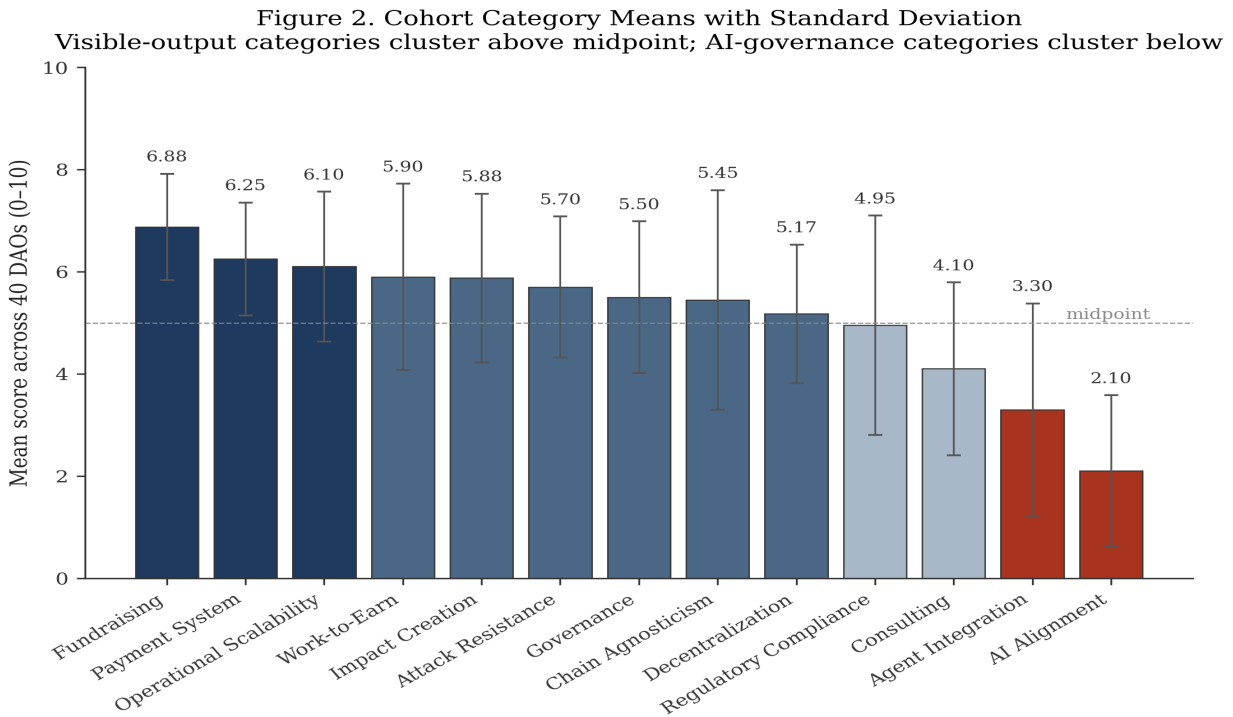


Figure 2. Dataset Category Means with Standard Deviation Bars.

The four categories above the midpoint of 5.0 (Fundraising, Payment System, Operational Scalability, Work-to-Earn) are categories that produce visible, fundable, measurable outputs. Token launches, treasury balances, contract automation, and grants programs all generate documentation that can be audited from public sources. The four categories below the midpoint (Regulatory Compliance, Consulting, Agent Integration, AI Alignment) are categories that require institutional development that does not produce visible artifacts until the DAO faces a

specific failure mode that the missing institution would have prevented. This visibility paradox is developed as Finding 1 below.

The distribution also exposes the framework's diagnostic value. The framework's foundational categories (Decentralization at 5.17, Governance at 5.50, Attack Resistance at 5.70) cluster around the midpoint. The framework's peripheral categories (Consulting at 4.10, Chain Agnosticism at 5.45) are dispersed in the lower-middle range. The framework's forward-looking categories (Agent Integration at 3.30, AI Alignment at 2.10) cluster near the bottom. If the framework's category importance ordering were applied as weights, the dataset mean total would be lower than 67.3 of 130 because the lowest-scoring categories are those the framework treats as most foundational to future DAO performance.

## **VI. Findings**

Five patterns hold across every segment in the dataset. Each is described below with its empirical basis, its theoretical interpretation, and its implications for the framework.

### **A. Finding 1: The Visibility Paradox**

Categories that produce visible artifacts, token launches, treasury balances, mission statements, marketplace activity, contract automation, score consistently above 5 across the dataset. Categories that produce invisible institutional infrastructure, legal wrappers, judicial branches, AI alignment policies, on-chain reputation ledgers, agent-identity protocols, score consistently below 5. The institutions that are most visible to users, investors, and journalists are well-developed; the institutions that are invisible until they fail are severely underdeveloped. The pattern holds across every segment.

The visibility paradox has a structural explanation. Visible institutions attract resources because they produce measurable, fundable outputs. Token launches generate immediate capital. Mission statements generate immediate social legitimacy. Marketplace activity generates immediate transaction-fee revenue. Contract automation generates immediate operational efficiency. These outputs are observable, attributable, and creditable to the institutions that produced them, so the institutions that produce them attract continuing investment. Invisible institutions, by contrast, produce institutional resilience that is observable only when the institution fails to function. A

DAO that has built a strong judicial branch is observable as such only when a governance dispute arises and the judicial branch resolves it well. A DAO that has built strong AI alignment infrastructure is observable as such only when an autonomous agent attempts an action the alignment infrastructure prevents. A DAO that has built reputation-weighted voting is observable as such only when a token-weighted attack is launched and the reputation weighting absorbs it. Until the failure event occurs, the institution is invisible to everyone except its institutional designers. Investment in invisible institutions is therefore systematically underprovided relative to investment in visible institutions, even when the invisible institutions are more important to long-run resilience.

The visibility paradox is a known pattern in institutional economics, often cited in the literature on regulatory capture, central-bank independence, and legal-system development. Its appearance in the DAO dataset is consistent with what the framework would predict: the institutional architecture for cooperative governance has converged on the institutions that produce visible, fundable outputs and has not converged on the institutions that produce invisible resilience. The empirical analysis surfaces this pattern. It does not claim to have caused it. It does not claim to resolve it. It does claim that any prescriptive use of the framework should foreground the visibility paradox as the primary obstacle to balanced institutional development.

## **B. Finding 2: The AI-Governance Vacuum is Universal**

AI Alignment scores 2.10 of 10 across the dataset. No DAO scores above 5. This is the only category in the entire thirteen-dimension framework where no DAO crosses the midpoint. Agent Integration scores only marginally better at 3.30 of 10. The deficit holds even for DAOs whose explicit business is AI-adjacent: Vana, positioned as the data DAO for AI training, scores 4 of 10 on AI Alignment; Filecoin, the storage substrate for AI agent infrastructure, scores 3 of 10. Streamr, the real-time data streams powering AI inference, scores 1 of 10. These DAOs are building the plumbing for autonomous agents to consume on-chain data while simultaneously failing to build the institutional guardrails to constrain those agents when they act on that data. Figure 6 presents the distribution.

Figure 6. The AI-Governance Vacuum  
AI Alignment cohort mean 2.10; Agent Integration cohort mean 3.30

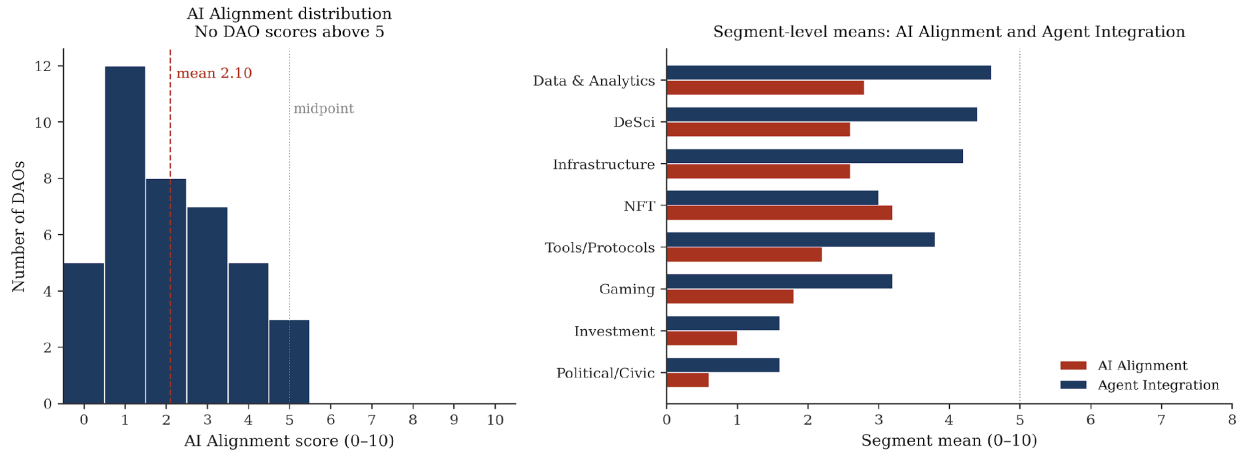


Figure 6. The AI-Governance Vacuum: Distribution of AI Alignment Scores and Segment-Level Means.

The AI-governance vacuum is the single largest convergent gap surfaced by the empirical analysis. It is also the gap most consequential for the next decade of DAO operations. The combination of autonomous AI agents and on-chain governance is not a speculative concern. It is operational reality. Autonomous trading agents already trade on Uniswap and other AMMs. Autonomous research agents already participate in Polymarket and Kalshi prediction markets. Autonomous governance agents have been deployed by several protocols on a pilot basis to summarize governance proposals, draft votes, and manage multisig signing. The infrastructure exists; the institutional guardrails do not.<sup>7</sup>

The framework predicts three concrete failure modes. First, governance attacks executed by autonomous agents. A patient strategic actor with capital can deploy autonomous agents to execute governance attacks at lower cost and higher precision than a human-organized attack would permit. The failure mode is: a governance vote with apparent majority support is executed, the protocol is modified, and the modification advantages a small set of capital holders who deployed the agents that produced the apparent majority. Second, governance paralysis driven by autonomous-agent disagreement. A DAO with a large population of autonomous

<sup>7</sup> For the prior treatment of agent-mediated governance and the institutional infrastructure required to make it tractable, see Wulf A. Kaal, "AI's Mother's Instinct: Engineered Consequence, Emergent Ethics, and the Institutional Trajectory Toward Agentic Alignment" (SSRN draft, 2026), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=6244278](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6244278).

agents, each acting on a different reasoning model and with different objective functions, may face votes in which no proposal can achieve majority support because the agents disagree on the correct outcome and human members lose tracking of the underlying reasoning. The failure mode is governance throughput collapse. Third, alignment drift in agents that participate over extended periods. An autonomous agent that participates in DAO governance over months or years may experience reasoning drift driven by exposure to the DAO's accumulated communications and decisions. Without alignment infrastructure that detects and corrects this drift, the agent's contributions may shift from the DAO's stated mission to objectives the agent has constructed for itself. The failure mode is an agent that began as a faithful representative of its principal becoming an actor with its own objectives, with voting weight that is difficult to displace.

None of these failure modes is hypothetical. The first has been documented in nascent form in flash-loan governance attacks executed by autonomous capital, including the Beanstalk exploit. The second has been documented in proposal-fatigue patterns across multiple DAO governance forums. The third has been documented in research literature on language-model alignment drift in extended-context settings. The framework predicts that all three will worsen as autonomous-agent participation grows, and that DAOs without AI-alignment infrastructure will be the first to experience them.

The 2.10 dataset-wide score is therefore not a marginal deficit. It is a structural deficit in the category that the framework identifies as foundational to the next decade of DAO operation. Every DAO in the dataset has a problem in this category. The DAOs whose business is most directly AI-adjacent have the same problem in only marginally less severe form. The DAOs whose business is least directly AI-adjacent score 0. The deficit is universal.

### **C. Finding 3: Token-Plutocracy is the Default**

Across all forty DAOs, governance is fundamentally token-weighted. Decentralization scores 5.17 dataset-wide. Even DAOs with the strongest formal governance processes, BanklessDAO at 8/10 on Governance, ENS at 8/10, operate underlying token-weighted voting that recreates capital-based concentration beneath the procedural facade. No DAO in the dataset has implemented reputation-weighted voting in production. The gap between the framework's

prescription, reputation as the proper coordinator of cooperative governance, and the actual practice of every DAO in the dataset is the most consistent institutional finding in the data.

Token-plutocracy as default has a well-documented historical explanation. The original DAO architecture was designed to be permissionless: anyone with capital could acquire governance tokens, and governance weight was therefore proportional to capital commitment to the protocol. This design choice solved a coordination problem, how to weight governance contributions in a permissionless context where identity cannot be verified, and produced a coordination problem, how to prevent capital concentration from translating directly into governance concentration. The first was urgent at the time of design. The second emerged only as protocols accumulated significant treasury value and governance capture became economically attractive.

Reputation-weighted voting is the framework's response to the second problem. Reputation, in the framework, is accumulated through documented contributions to the protocol over time, weighted by reference-based revaluation in the WDAG. A contributor's reputation cannot be acquired through capital purchase. It requires actual contribution to the protocol that other contributors are willing to reference. Reputation is therefore Sybil-resistant in a way that token holdings are not: a strategic actor can acquire token holdings by purchasing tokens, but a strategic actor cannot acquire reputation without making contributions that the rest of the protocol's contributors will reference favorably.<sup>8</sup> Reputation-weighted voting is therefore the architectural response to the governance-capture problem that token-plutocracy creates.

The dataset shows that no DAO has deployed this architecture. Every DAO uses token-weighted voting as the primary governance mechanism. Some have implemented quadratic voting (which redistributes voting power away from the largest token holders but does not change the underlying capital basis). Some have implemented vote delegation (which allows small holders to delegate to perceived expertise but does not change the capital basis). Some have implemented reputation tracking in informal forums (which produces social-pressure effects but does not change the on-chain governance weight). The architectural shift from token-weighted to reputation-weighted on-chain voting has not yet occurred in production deployment in any DAO.

---

<sup>8</sup> On the citation-honesty incentive structure underlying reference-weighted reputation, see Wulf A. Kaal, "Citation Honesty Mechanisms in Weighted Directed Acyclic Graph Governance," (SSRN draft, 2026), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=6269518](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6269518).

The reasons for non-adoption are worth considering. Reputation-weighted aggregation requires institutional infrastructure, a contributor-reputation tracking system, a reference-based revaluation mechanism, a Sybil-resistant identity layer, that no DAO has built at production scale. Building this infrastructure requires engineering investment that no DAO has committed. The investment is not made because the visibility paradox makes investment in visible institutions more attractive than investment in invisible institutions. Finding 3 is therefore a special case of Finding 1.

#### **D. Finding 4: Legal-Wrapper Heterogeneity Without Convergence**

Eight different legal-wrapper structures appear across the dataset. The Cayman Islands Foundation is used by ENS, Mantle, Vana, Gnosis, and the ApeCoin DAO. The Swiss Stiftung or Verein is used by the Ethereum Foundation, DeXe DAO, and several smaller entities. The Wyoming DAO LLC<sup>9</sup> is used by CityDAO. The Singapore foundation is used by Ocean Protocol. The British Virgin Islands issuer structure is used by Ondo Finance. The UK discretionary mutual is used by Nexus Mutual. The Delaware C-corporation is used by Star Atlas's operating entity. The unincorporated association is used by HairDAO, ConstitutionDAO at formation, and More Loot.

The Regulatory Compliance score range of 1 to 9 is the widest in the dataset. The set has not converged on a dominant legal architecture. The Wyoming DAO LLC, despite being a purpose-built statute, has been adopted by only one DAO in the entire forty-DAO sample. This is a striking under-utilization of available statutory infrastructure. The reasons are partly historical: many DAOs in the dataset were formed before the Wyoming DAO LLC statute was promulgated in 2021, and the cost of restructuring to adopt the Wyoming form would be substantial. The reasons are also partly substantive: the Wyoming DAO LLC has not been tested in major litigation, and DAO operators have generally preferred the Cayman Foundation or Swiss Stiftung forms because of their established judicial precedent in handling decentralized governance disputes. The CityDAO experiment with the Wyoming form is a useful case study in the costs

---

<sup>9</sup> Wyoming Decentralized Autonomous Organization Supplement, *Wyo. Stat. §§ 17-31-101 through 17-31-115* (effective July 1, 2021), available at [https://sos.wyo.gov/Forms/WyoBiz/DAO\\_Supplement.pdf](https://sos.wyo.gov/Forms/WyoBiz/DAO_Supplement.pdf).

and benefits, and the fact that no other DAO has followed suit is itself a significant data point on the perceived viability of the Wyoming form.

The implications for legal scholars are several. First, the legal-wrapper question for DAOs is not yet resolved at the level of statutory infrastructure. Multiple jurisdictions have promulgated DAO-specific statutes in the past five years; no single statute has emerged as the dominant choice. Second, the Cayman Foundation has emerged as the de facto market leader despite not being DAO-specific, because it provides the discretionary trust structure that DAOs require to manage governance disputes without binding judicial precedent. Third, the Wyoming DAO LLC, despite being purpose-built, has not been adopted at scale because the perceived legal risk of an untested statute outweighs the perceived legal benefit of statutory specificity. Fourth, the heterogeneity produces regulatory arbitrage: DAOs formed under Cayman law operate in jurisdictions where their members are subject to different regulatory regimes, and the institutional alignment between the legal wrapper and the regulatory regime that applies to members is incidental rather than structured.

### **E. Finding 5: A Convergent Architectural Agenda for Institutional Repair**

Across the forty DAOs, the projected post-upgrade score (95.3 of 130 mean) represents a +28-point improvement (+42 percent) over the baseline (67.3 of 130 mean). The largest projected improvements are concentrated in three categories: AI Alignment (typical projected improvement +4 to +7 points per DAO), Agent Integration (+3 to +5), and Governance (+2 to +4 from tripartite separation). The smallest projected improvements are in categories that are already strong (Fundraising, Payment System) or that depend on factors outside protocol design (Regulatory Compliance, where legal-wrapper choice is an exogenous business decision).

The improvement agenda converges on five upgrades. First, ERC-1155 multi-token reputation as the architectural mechanism for moving from token-weighted to reputation-weighted voting. Second, tripartite governance separation through the deployment of distinct legislative, executive, and judicial smart-contract modules. Third, stablecoin treasury infrastructure for predictable contributor compensation and protocol revenue management. Fourth, historiographic transparency through the implementation of WDAG-based governance logging that exposes the full reference graph of governance decisions to public inspection. Fifth, values-drift detection

through the implementation of automated monitoring of the DAO's reasoning and decision patterns to detect departures from stated mission or operational principles.

Figure 4. Current Versus Projected Institutional Score  
Every DAO sits above the parity line; cohort centroid at (67.3, 95.3) shows mean +28-point projected gain

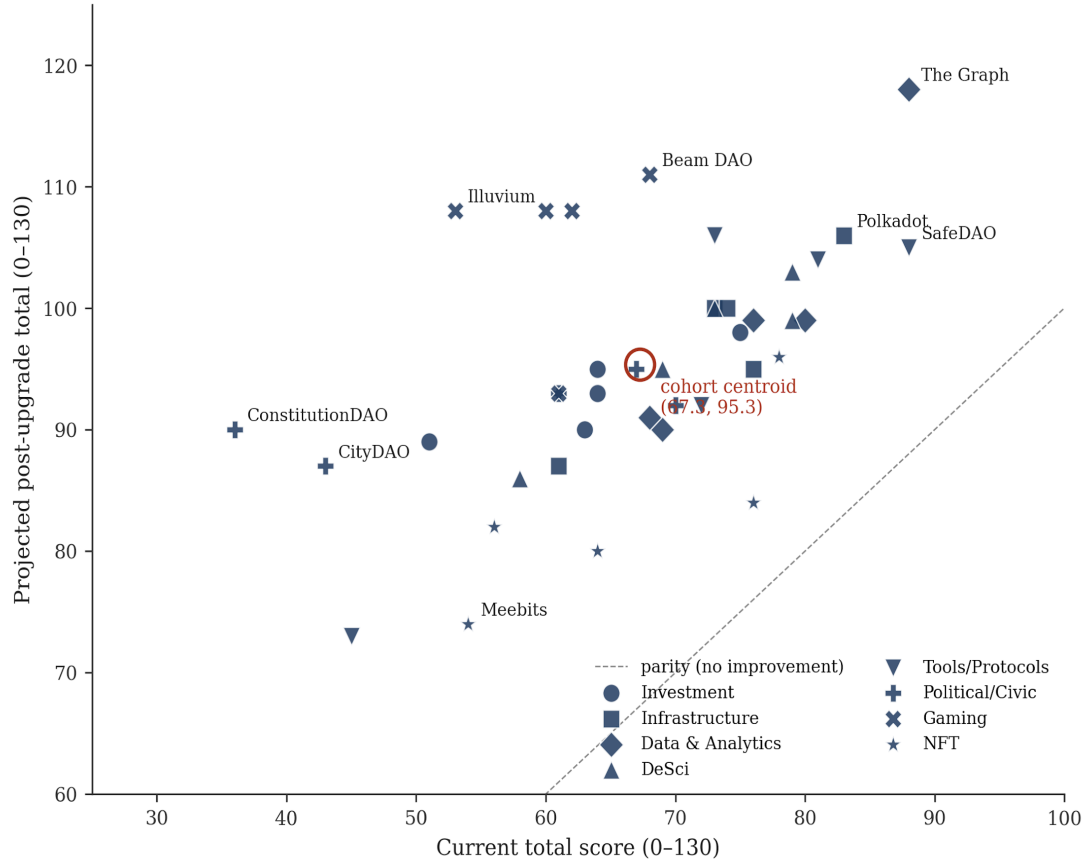


Figure 4. Current Total Scores Plotted Against Projected Post-Upgrade Totals. The 45-degree reference line indicates parity.

Figure 4 presents the current-versus-projected trajectory. Every DAO sits above the parity line, indicating projected improvement. The cohort centroid sits substantially above the diagonal, indicating that the average DAO has substantial room for improvement under the framework's prescribed architectural responses. The DAOs furthest above the diagonal, Beam DAO, Decentraland, YGG, Illuvium, are gaming and virtual-world DAOs whose underlying communities support significant governance scaling potential despite low current scores. The DAOs closest to the diagonal, More Loot at +16, Meebits at +20, Ondo Finance at +31, are DAOs whose current institutional architecture is either already strong relative to the framework's prescriptions or constrained by exogenous factors that the architectural agenda cannot address.

The five upgrades describe a substantial engineering investment. Implementation of ERC-1155 multi-token reputation requires a new smart-contract suite, a governance migration, and integration with the DAO's existing voting infrastructure. Implementation of tripartite separation requires a new institutional architecture, a smart-contract suite, and migration of historical governance to the new architecture. Implementation of stablecoin treasury infrastructure requires custody architecture, oracle integration, and contributor-payment migration. Implementation of historiographic transparency requires WDAG storage and query infrastructure and historical-data migration. Implementation of values-drift detection requires natural-language-processing infrastructure, alignment infrastructure, and integration with governance decision-making.

For an established DAO with existing infrastructure, the implementation effort for the full five-upgrade suite is roughly comparable to the effort that produced the DAO's original infrastructure. The estimates range from six to eighteen months of engineering work and from \$300,000 to \$2 million in audit and engineering costs per DAO. None of the DAOs in the dataset has committed to this investment. The investment that has been made is concentrated in the visible categories, token launches, treasury growth, marketplace activity, partnership development, rather than in the invisible categories that the five-upgrade suite addresses. The visibility paradox (Finding 1) explains why.

## **VII. Segment-Level Analysis**

Figure 3 presents segment-level radar plots that visualize how each segment compares against the dataset mean across the thirteen categories. Figure 5 presents the distribution of total scores within each segment as a box-and-strip plot.

Figure 3. Segment Profiles Against the Cohort Mean  
Solid: segment mean. Dashed: cohort baseline.

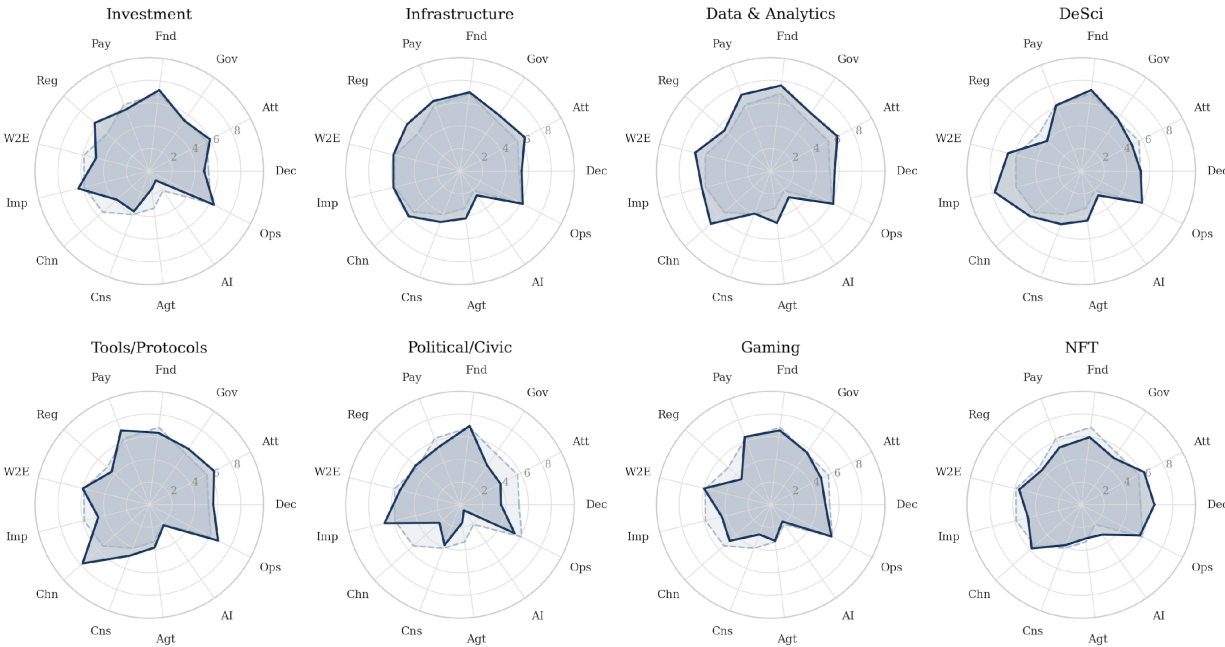
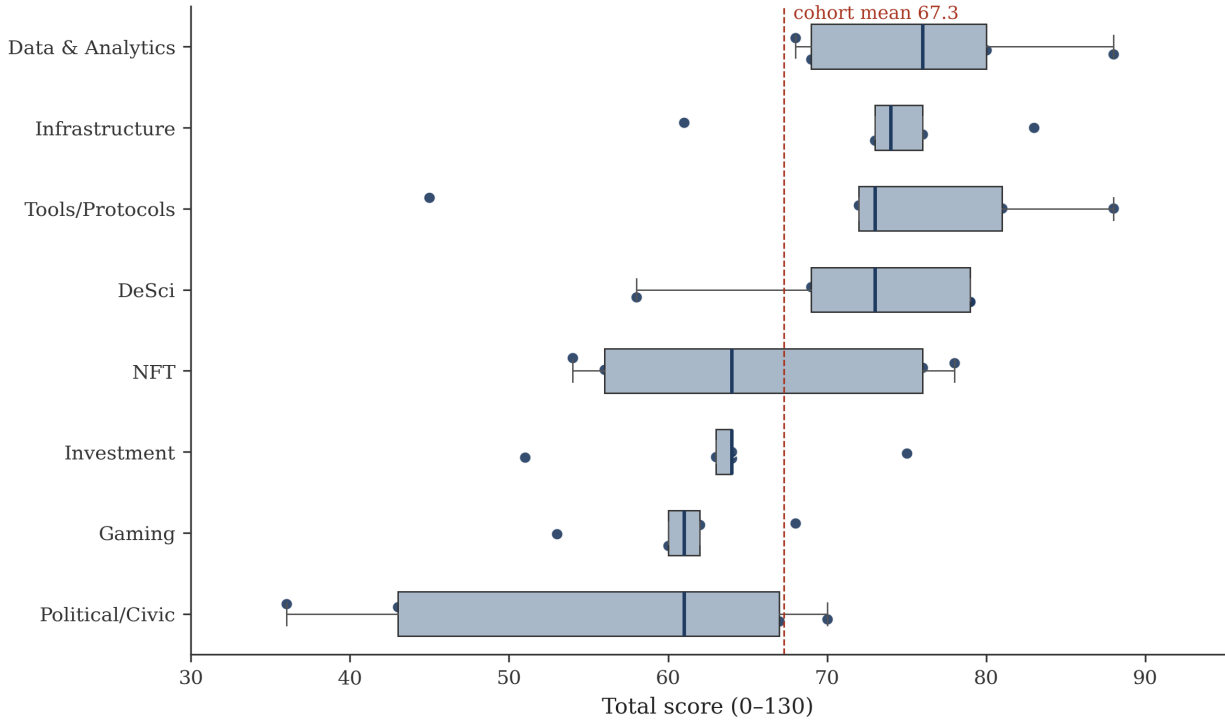


Figure 3. Segment-Level Radar Plots Against the Cohort Baseline. Solid: segment mean. Dashed: cohort baseline.

Figure 5. Distribution of Total Scores by Segment  
Boxes show interquartile range; points are individual DAOs



*Figure 5. Box Plot of Total Scores by Segment with Individual DAOs Plotted as Points.*

### **A. Data and Analytics (mean 76.2)**

Data and Analytics is the highest-scoring segment. The Graph, Filecoin, Ocean Protocol, Streamr, and Vana average 76.2 of 130, with a range from 68 (Streamr) to 88 (The Graph). The segment leads on Chain Agnosticism (mean 7.0) and Operational Scalability (mean 6.2), reflecting that these protocols were designed for multi-chain deployment and scalable infrastructure from inception. The segment also scores well on Fundraising (mean 7.6) and Payment System (mean 7.2), reflecting mature capital infrastructure and tokenized data marketplaces.

The Graph at 88 of 130 is the dataset's tied-highest scorer. Its institutional strengths are concentrated in the categories the framework treats as foundational: Decentralization (6), Attack Resistance (7), Governance (7), and Operational Scalability (8). The Graph also leads the dataset on Chain Agnosticism (9) and on Work-to-Earn (8), reflecting the curator-indexer-delegator structure that produces measurable contributions across multiple participation tiers. The Graph's deficit categories are Consulting (5) and AI Alignment (4): the protocol has not productized its governance learning and has not built the alignment infrastructure for autonomous-agent participation despite serving as the query layer for many AI-adjacent applications.

### **B. Tools and Protocols (mean 71.8)**

The Tools and Protocols segment includes oracle protocols (UMA DAO), insurance protocols (inSureDAO), smart-account protocols (SafeDAO), governance-tooling protocols (DeXe DAO), and infrastructure protocols (Orbs Network). The segment is heterogeneous within itself, ranging from 45 (inSureDAO) to 88 (SafeDAO).

SafeDAO at 88 of 130 is the dataset's tied-highest scorer alongside The Graph. SafeDAO's institutional strengths are concentrated in Payment System (9, the highest score in the dataset), Chain Agnosticism (9), and Regulatory Compliance (8). SafeDAO's role as multisig and smart-account infrastructure for \$100 billion-plus in custody assets has driven mature institutional architecture in the capital-management categories. Its deficit categories are Decentralization (5) and Consulting (4): the protocol's governance is concentrated among Safe

ecosystem participants and the protocol has not productized its custody-governance learning to other protocols.

DeXe DAO at 81 of 130 is the segment's second-highest scorer. Its institutional positioning is governance-as-product: the protocol provides governance infrastructure to other DAOs and has therefore developed institutional architecture in Work-to-Earn (9), Governance (8), and Consulting (8). Its deficit categories are AI Alignment (0) and Agent Integration (1), reflecting that governance-as-product has not yet been extended to autonomous-agent governance.

### **C. Infrastructure (mean 73.4)**

The Infrastructure segment includes Polkadot, ENS DAO, Mantle, Gnosis DAO, and Golem Network. The segment averages 73.4, with a range from 61 (Golem Network) to 83 (Polkadot).

Polkadot at 83 of 130 is the segment leader. Its institutional strengths are concentrated in the categories the framework treats as foundational: Decentralization (7), Attack Resistance (8), Governance (6, with substantial detail in OpenGov), Chain Agnosticism (9, anchored by the parachain architecture), and Consulting (8, the highest in the dataset and reflecting the protocol's mature governance-export capabilities through the OpenGov framework). Its deficit categories are Impact Creation (5), AI Alignment (3), and Agent Integration (4), reflecting that the protocol's institutional architecture is mature in core governance categories but underdeveloped in forward-looking categories.

### **D. DeSci (mean 71.6)**

The DeSci segment includes ResearchHub, VitaDAO, Molecule, Bio Protocol, and OriginTrail. The segment averages 71.6, with a range from 58 (ResearchHub) to 79 (Bio Protocol and OriginTrail).

The segment's institutional strengths are concentrated in Impact Creation (mean 7.8, the highest in the dataset), Fundraising (mean 7.2), and Work-to-Earn (mean 6.6). DeSci protocols have built mature institutional infrastructure for translating member contributions into measurable scientific outcomes and for compensating the contributors who produce them. OriginTrail at 79 is significant because it illustrates the pattern of forward-looking institutional development that the framework prescribes: the protocol has built Agent Integration infrastructure (score 8, the highest

in the dataset) ahead of dataset norms, anticipating the autonomous-agent participation that the framework predicts will be the central institutional challenge of the next decade.

### **E. Investment and DeFi (mean 63.4)**

The Investment and DeFi segment includes Goldfinch, HairDAO, Maple Finance, Nexus Mutual, and Ondo Finance. The segment averages 63.4, with a range from 51 (HairDAO) to 75 (Nexus Mutual). The segment's institutional strengths are concentrated in Fundraising (mean 7.2) and Operational Scalability (mean 6.4). Its institutional weaknesses are concentrated in AI Alignment (mean 1.0), Agent Integration (mean 1.6), and Chain Agnosticism (mean 3.8). DeFi protocols have built institutional architecture optimized for capital management within a single chain and have not built institutional architecture for autonomous-agent participation or cross-chain governance.

### **F. NFT (mean 65.6)**

The NFT segment includes TerraForms, Meebits, BAYC and MAYC, and More Loot. The segment averages 65.6, with a range from 54 (Meebits) to 78 (BAYC). The segment is institutionally heterogeneous because the underlying DAOs were formed through different paths. BAYC and MAYC operate under ApeCoin DAO, a purpose-built governance structure. TerraForms operates with native-token governance over generative-art NFTs. Meebits operates under Larva Labs / Yuga IP capture. The pattern reflects this heterogeneity: institutional architecture is mature where the DAO is purpose-built and underdeveloped where the DAO is retrofitted onto an existing IP holding.

### **G. Gaming and Virtual Worlds (mean 60.8)**

The Gaming segment includes Beam DAO, Decentraland, YGG, Star Atlas, and Illuvium. The segment averages 60.8, with a range from 53 (Illuvium) to 68 (Beam DAO). The segment's institutional weaknesses are consistent across DAOs. Decentralization (mean 4.8), Attack Resistance (mean 5.0), and Regulatory Compliance (mean 3.4) all score below the dataset mean. The segment's relative strength is in Operational Scalability (mean 6.0), reflecting the technical-scaling investment that gaming infrastructure requires.

## **H. Political and Civic (mean 55.4)**

The Political and Civic segment includes Fort Worth DAO, ConstitutionDAO, CityDAO, BanklessDAO, and the Ethereum Foundation. The segment averages 55.4, the lowest in the dataset, with a range from 36 (ConstitutionDAO) to 70 (Ethereum Foundation). The segment's institutional weaknesses are systemic: Decentralization (mean 3.6), AI Alignment (mean 0.8), and Chain Agnosticism (mean 2.4) all score below the dataset mean. The segment's relative strength is in Impact Creation (mean 6.8).

ConstitutionDAO at 36 of 130 is the dataset's lowest scorer. It illustrates an extreme version of the visibility paradox: the DAO scored 9 on Fundraising (the dataset maximum, reflecting the protocol's approximately 47-million-dollar ETH raise from 17,437 contributors in three days to bid on a first printing of the U.S. Constitution at Sotheby's<sup>10</sup>) and 1 on Governance (the dataset minimum, reflecting that the DAO had no functional governance architecture beyond the fundraising mechanism itself). A DAO can be highly successful on visible categories and effectively non-functional on invisible categories simultaneously.

## **VIII. The Architectural Repair Trajectory**

### **A. The Convergent Five Upgrades**

The improvement agenda converges on five upgrades. The convergence is the most analytically significant feature of the projected scores. The same architectural responses surface across DAOs in different segments with different institutional histories.

**ERC-1155 multi-token reputation.** The architectural mechanism for moving from token-weighted to reputation-weighted voting. ERC-1155 supports multiple token types within a single contract, allowing a DAO to issue distinct reputation tokens for distinct contribution categories, development, governance participation, community moderation, treasury management, while maintaining a single underlying contract. Reputation tokens are

---

<sup>10</sup> Jacob Kastrenakes, "ConstitutionDAO loses \$43 million auction of rare US Constitution copy," *The Verge*, November 18, 2021, <https://www.theverge.com/2021/11/18/22790452/constitutiondao-loses-auction-for-rare-us-constitution-copy-sothebys>.

non-transferable (locked to the contributor's address) and decay-weighted (older contributions count for less than recent contributions, mimicking the temporal structure of academic citation). The implementation gap is the engineering investment required to integrate ERC-1155 reputation issuance with the DAO's existing governance contracts and to migrate existing voters from token-weighted to reputation-weighted voting without disenfranchising historical contributors.

**Tripartite governance separation.** The deployment of distinct legislative, executive, and judicial smart-contract modules. The legislative branch is responsible for proposal initiation and voting. The executive branch is responsible for proposal execution and treasury management. The judicial branch is responsible for dispute resolution and rule interpretation. Each branch has distinct membership, distinct voting mechanisms, and distinct accountability procedures. The implementation gap is the smart-contract architecture required to implement three-branch separation, the governance migration required to populate each branch with its initial membership, and the cross-branch accountability mechanisms, impeachment, override, judicial review, that the framework requires to make the separation institutionally meaningful.

**Stablecoin treasury infrastructure.** The architectural mechanism for predictable contributor compensation and protocol revenue management. Restructuring DAO treasuries from native-token concentration toward stablecoin-backed reserves, with explicit treasury policies on reserve composition, rebalancing thresholds, and stablecoin-to-native-token swap procedures. The implementation gap is the engineering investment required to build the stablecoin treasury infrastructure (oracles for stablecoin pricing, swap routes for stablecoin-to-native conversion, custody architecture for stablecoin reserves) and the governance migration required to shift contributor compensation from native-token to stablecoin payment without disrupting existing contributor relationships.

**Historiographic transparency.** The implementation of WDAG-based governance logging that exposes the full reference graph of governance decisions to public inspection. Every governance action is a vertex in the graph, every reference between actions is a directed edge with an explicit weight, and the cumulative graph is queryable through a public API. The implementation gap is the engineering investment required to build the WDAG storage and query infrastructure and the governance migration required to populate the graph with historical governance decisions back to the DAO's formation.

**Values-drift detection.** The implementation of automated monitoring of the DAO's reasoning and decision patterns to detect departures from stated mission or operational principles. The infrastructure tracks the semantic content of governance proposals, votes, and execution decisions over time and flags significant departures from the DAO's stated mission. The implementation gap is the engineering investment required to build the natural-language-processing infrastructure for proposal monitoring, the alignment infrastructure for principle-tracking, and the governance migration required to integrate drift detection with the DAO's existing decision-making procedures.

## **B. The Implementation Gap**

The five upgrades describe a substantial engineering investment. The estimates, where they have been quantified, range from six to eighteen months of engineering work and from \$300,000 to \$2 million in audit and engineering costs per DAO. None of the DAOs in the dataset has committed to this investment. The investment that has been made is concentrated in the visible categories, token launches, treasury growth, marketplace activity, partnership development, rather than in the invisible categories that the five-upgrade suite addresses.

The implementation gap is therefore the central practical limitation of Finding 5. The framework prescribes a specific set of architectural responses to the institutional deficits the dataset exhibits. The empirical analysis confirms the responses are buildable. The empirical analysis does not confirm that the responses will be built in production deployment, because the engineering investment has not been committed and the investment is not in the obvious self-interest of any current DAO operator. The visibility paradox explains why.

## **C. The Path-Dependence Constraint**

The DAOs in the dataset were formed under the institutional architectures that were available at the time of their formation. ConstitutionDAO was formed in late 2021 under the institutional architectures that were available in late 2021, which did not include reputation-weighted voting, tripartite separation, or values-drift detection. The DAO's institutional weaknesses reflect the institutional architectures that were available at formation rather than a failure of design intent.

The path-dependence constraint implies that the framework's prescriptive value is greater for DAOs in formation than for DAOs in operation. A DAO that is being designed today can adopt the framework's prescribed architectures from inception, at marginal additional cost relative to designing the DAO without them. A DAO that is in operation today has to add the prescribed architectures on top of existing infrastructure, with significant migration costs and significant risks to existing contributor relationships. The framework should therefore be applied prescriptively to DAOs in formation, while operational DAOs should be evaluated against the framework as a diagnostic tool that identifies their institutional weaknesses without necessarily prescribing the migration path to address them.

## **IX. Implications**

### **A. Implications for Legal Scholars**

The first implication is that the institutional landscape of DAOs is sufficiently mature to support cross-DAO comparative analysis under a uniform framework. Until recently, DAO institutional analysis has tended toward case studies of individual DAOs, the DAO, Compound, MakerDAO, Uniswap, without a uniform framework that supports comparison across the case studies. The thirteen-category rubric applied across the dataset demonstrates that uniform comparative analysis is feasible, that it surfaces patterns that are not visible in individual case studies, and that it generates testable hypotheses about institutional development across the sector.

The second implication is that the institutional categories the framework identifies as foundational, decentralization, governance, attack resistance, are systematically underdeveloped relative to the categories that produce visible artifacts, fundraising, payment system. Legal scholarship focused on the visible categories has accumulated rapidly. Legal scholarship focused on the invisible categories, the constitutional structure of DAO governance, the dispute-resolution mechanisms appropriate for decentralized contexts, the institutional architecture for autonomous-agent participation, has accumulated more slowly. The underdeveloped scholarship corresponds to the underdeveloped institutional infrastructure.

The third implication is that the legal-wrapper question for DAOs is unresolved at the level of statutory infrastructure. Eight different wrapper structures appear in the dataset. The Wyoming

DAO LLC, despite being the most explicit statutory response, has been adopted by only one DAO. The Cayman Foundation has emerged as the de facto market leader despite not being DAO-specific. The legal scholarship on the optimal legal wrapper for DAOs has not converged, and the data suggests the convergence will not occur through institutional pressure in the short term.

The fourth implication is that the AI-governance vacuum is the most consequential gap for the next decade of DAO operation. AI Alignment scores 2.10 dataset-wide, with no DAO scoring above 5. Legal scholarship on AI governance in cooperative contexts is in early development. Legal scholars have an opportunity to develop institutional architectures for autonomous-agent participation that DAOs can adopt, to develop liability frameworks for agent-mediated governance failures, and to develop dispute-resolution mechanisms for disagreements that involve autonomous agents as parties.

## **B. Implications for Regulators**

The first implication for regulators is that the dataset identifies specific institutional weaknesses that produce regulatory risk. The Decentralization deficit is associated with governance-capture risk that the framework predicts will produce protocol-level failures over time. The Attack Resistance deficit is associated with technical-attack risk that has produced documented protocol-level failures across multiple DAOs in recent years. The AI Alignment deficit is associated with autonomous-agent governance risk that the framework predicts will produce protocol-level failures as agent participation scales. Regulators concerned with consumer protection, financial stability, and systemic risk in the DAO sector have an empirical basis to prioritize attention to these specific weaknesses.

The second implication is that statutory infrastructure for DAO governance has been promulgated faster than DAOs have adopted it. The Wyoming DAO LLC, the Vermont blockchain-based limited liability company, the Tennessee decentralized organization, the Marshall Islands DAO LLC, and the Liechtenstein blockchain-asset trust all provide DAO-specific statutory infrastructure. None has been adopted by more than a handful of DAOs in the dataset. The mismatch between statutory promulgation and statutory adoption suggests that regulators have built infrastructure for which there is limited demand at current pricing.

The third implication is that the visibility paradox implies that market discipline alone will not produce balanced institutional development. Market discipline produces under-investment in invisible institutional infrastructure, and the institutional infrastructure that is most predictive of long-run resilience is precisely the institutional infrastructure that is most underprovided. The data provides empirical support for the application of the public-goods argument to the DAO sector.

The fourth implication is that dynamic-regulation principles are particularly applicable to the DAO sector.<sup>11</sup> The Folk-Theoretic prediction that any static rule set will be gamed by patient strategic actors applies to regulatory rule sets as much as it applies to DAO rule sets. Regulators have an opportunity to design regulatory frameworks that are themselves capable of governing their own evolution: frameworks that use the same institutional architecture , reputation accumulation, validation pools, reference-based revaluation, automated alignment monitoring , that the framework prescribes for the DAOs the regulation governs.

### **C. Implications for DAO Operators**

The first implication for DAO operators is diagnostic. The data provides a benchmark against which any DAO operator can assess their protocol's institutional architecture. The thirteen-category rubric is reproducible, the dataset means and ranges are documented, and the segment-level patterns identify the institutional benchmarks appropriate to each DAO's segment. A DAO operator who wants to understand how their institutional architecture compares to forty peer DAOs can apply the rubric to their own protocol, compute the score, and locate the result in the distribution.

The second implication is prescriptive. The five convergent upgrades describe a specific architectural agenda that addresses the institutional weaknesses the dataset exhibits universally. DAO operators considering investment in institutional infrastructure have an empirical basis for prioritizing these five upgrades over alternative investments. The implementation gap is

---

<sup>11</sup> On dynamic regulation as a regulatory-design philosophy, see Wulf A. Kaal, *Evolution of Law: Dynamic Regulation in a New Institutional Economics Framework* (2013). *Festschrift in Honor of Christian Kirchner*, 2013, SSRN: <https://ssrn.com/abstract=2267560>

significant, but the engineering investment is buildable and the architectural responses are convergent.

The third implication is strategic. The dataset surfaces patterns that have competitive implications. DAOs that build the invisible institutional infrastructure earlier than their competitors will have institutional resilience that competitors lack when failure events occur. The first DAO in any segment to deploy reputation-weighted voting at production scale will have a defensible institutional differentiator. The first DAO to deploy formal tripartite separation at production scale will have a defensible institutional differentiator. The first DAO to deploy AI alignment infrastructure at production scale will have a defensible institutional differentiator. The visibility paradox implies that competitors will not pursue these investments simultaneously, and a DAO that pursues them anyway will therefore have a window of competitive advantage before competitor adoption catches up.

The fourth implication is forward-looking. The AI-governance vacuum is the institutional gap most consequential for the next decade of DAO operation. DAO operators that build AI alignment infrastructure now will be positioned to absorb autonomous-agent participation as it scales; DAO operators that do not will face the failure modes the framework predicts. The data shows that no DAO has built this infrastructure at scale. The window for first-mover advantage is open.

## **X. Limitations**

The empirical analysis has limitations that have to be surfaced explicitly. Three are most significant.

### **A. Source-Document Constraint**

The evaluations were generated from publicly available documentation. The scores measure what DAOs document about themselves rather than what DAOs do in operation. Consulting Snapshot voting histories or Etherscan treasury logs would have materially changed specific scores in some cases, particularly in Decentralization (where on-chain voting concentration is more readily measured than reading governance forum descriptions of voting concentration), Attack Resistance (where the actual record of attempted attacks is more informative than the

documented defenses), and Governance (where actual proposal-vote records are more informative than the documented procedures). The analysis reflects the source-document constraint and does not claim to overcome it.

## **B. Static Snapshot**

The scores represent point-in-time assessments. Several DAOs in the dataset were in active flux during the evaluation period: ApeCoin DAO during its AIP-596 transition; Illuvium during the November 2025 co-founder departure and treasury depletion; Bio Protocol during its veBIO token transition; Mantle during its migration to a ZK-rollup architecture; Vana in early operational stages with limited public governance history. For these DAOs, the reported scores may be obsolete by the time this Article is published. The analysis does not claim to track DAO trajectories over time. It reports a snapshot.

## **C. The Rubric's Theoretical Specificity**

The thirteen-category rubric is derived from the Calcaterra-Kaal framework. The rubric reflects the framework's specific theoretical commitments: that decentralization is foundational, that tripartite separation of powers is institutionally important, that reputation-weighted aggregation is the architectural response to token-plutocracy, that AI alignment is essential institutional infrastructure for autonomous-agent participation. These commitments are contestable. Other theoretical frameworks for evaluating DAO governance emphasize different categories and would generate different rubrics. The scores reflect the framework's theoretical commitments and would be different if a different framework were applied. The master scoring matrix is reproducible, and a different rubric could be applied to the same forty DAOs to produce a different analysis.

## **XI. Conclusion**

Five findings hold across every segment in the dataset. First, a visibility paradox in which institutional categories that produce visible artifacts score consistently above the midpoint while institutional categories that produce invisible governance infrastructure score consistently below it. Second, a universal AI-governance vacuum in which no DAO scores above 5 on AI Alignment and the dataset mean is 2.10 of 10. Third, the persistence of token-plutocracy as the

default governance form across the dataset, with no DAO having deployed reputation-weighted voting at production scale. Fourth, legal-wrapper heterogeneity without convergence, with eight distinct wrapper structures and only one DAO using the Wyoming DAO LLC. Fifth, a convergent architectural agenda for institutional repair built around five upgrades.

The unweighted dataset mean total of 67.3 of 130 (51.8 percent) is the central quantitative finding: the median DAO has implemented roughly half of what the framework prescribes. The deficit is structural rather than incidental. It reflects that DAO architecture has solved the problems for which it was originally designed, decentralized capital formation and programmable value transfer, and has not yet solved the problems that emerged after its design , AI-mediated governance, Sybil-resistant identity, and constitutional separation of powers.

The framework's prescriptive value rests on the buildability of the architectural responses to the institutional deficits the dataset exhibits. The architectural responses are buildable. The implementation gap from current state to production is significant, requiring six to eighteen months of engineering work and six-figure audit costs per DAO. No DAO in the dataset has committed to this investment as of the time of this writing. The visibility paradox explains the under-investment: invisible institutional infrastructure is systematically underprovided relative to visible institutional infrastructure even when the invisible infrastructure is more important to long-run resilience.

The institutional deficit in DAOs is real, structural, and addressable. The empirical analysis presented here documents the deficit, identifies its components, traces its causes, and surfaces the architectural responses that would address it. The deficit will not close itself: the visibility paradox produces under-investment in the institutional infrastructure that would close it. Closing the deficit will require deliberate institutional investment by DAO operators, regulatory attention to the public-goods aspect of invisible institutional infrastructure, and continued legal scholarship on the institutional architectures that decentralized governance requires but has not yet built.