Digital Public Infrastructure for Decentralized Science*

Michael Zargham, David Sisson, Christine Vandevoorde, and Ilan Ben-Meir

September 2024

Abstract

The DeSci Ecosystem facilitates the curation and sensemaking of ever growing streams of scientific knowledge artifacts. The system is built on a flexible peer-to-peer protocol layer which supports the DeSci Platform (among other applications in the future). On the DeSci Platform, Autonomous Research Communities (ARCs) are incentivized to curate knowledge artifacts published by Participants to create a viable alternative to traditional academic journals.

This report guides readers through the systems engineering process, in order to provide design recommendations for a microeconomic model, a macroeconomic model, and subsidy programs for the DeSci Ecosystem. It begins by describing the DeSci Ecosystem's Animating Purpose and Stakeholders, then analyzes its Environment. It then offers a high-level overview of the project throughout all five stages of the engineering life-cycle, before focusing and the Requirements and Design of the new economic system.

*This work was developed by BlockScience with support DeSci Labs from and in collaboration with the DeSci Foundation

Contents

1 Introduction

The current state of academic publishing meaningfully harms both academic progress and public access to knowledge. Alternative models are possible, and are supported by the literature across various disciplines. Our work is situated in systems engineering, which, when conducted correctly, insists that the significance of any technology is determined by the context of its use – a context that includes the social practices of its users. Systems engineering thus provides a roadmap for steering meaningful change in complex socio-technical systems – such as re-engineering the flow of knowledge and resources in academic publishing away from monopolistic entities, and towards a decentralized model that guards against their abuses.

1.1 The Current State of Academic Publishing

On Sept 12, 2024, a federal antitrust lawsuit was filed "against six commercial publishers of academic journals, including Elsevier, Springer Nature, Taylor and Francis, Sage, Wiley, and Wolters Kluwer, on behalf of a proposed class of scientists and scholars who provided manuscripts or peer review, alleging that these publishers conspired to unlawfully appropriate billions of dollars that would otherwise have funded scientific research" (Lieff Cabraser Heimann & Bernstein, LLP 2024). According to the lawsuit, the publishers' transgressions fall into three categories:

> First, an agreement to fix the price of peer review services at zero that includes an agreement to coerce scholars into providing their labor for nothing by expressly linking their unpaid labor with their ability to get their manuscripts published in the defendants' preeminent journals. Second, the publisher defendants agreed not to compete with each other for manuscripts by requiring scholars to submit their manuscripts to only one journal at a time, which substantially reduces competition by removing incentives to review manuscripts promptly and publish meritorious research quickly. Third, the publisher defendants agreed to prohibit scholars from freely sharing the scientific advancements described in submitted manuscripts while those manuscripts are under peer review, a process that often takes over a year.

Taken together, the suit alleges, these abuses create "a variety of perverse market failures that impair the ability of scientists to do their jobs and slow dramatically the pace of scientific progress [... including] a worsening peer-review crisis, whereby it has become increasingly difficult to coerce busy scholars into providing their valuable labor for nothing." Overall, the lawsuit claims, the corruption of the academic publishing industry "has held back science, delaying advances across all fields of research." While only a single example, the claims of this suit are a concise and demonstrative account of the current state of scientific publishing. Open Access and Open Science movements are attempting to compete, but have not yet gained traction sufficient to counteract these monopolistic practices.

1.2 Alternative Models in Literature

Alternative models for knowledge sharing are both feasible and grounded in existing scholarship regarding knowledge commons and the emerging theory of contribution goods.

Elinor Ostrom's pioneering work on governing common-pool resources provides the foundational framework for thinking about decentralized governance of academic publishing (Ostrom 2002). Ostrom's writings provide insight into the ways that polycentric governance applies to the multiple roles in decentralized publishing (e.g. authors, reviewers, editors), and the ways that coordination arises organically through community governance. Key principles, such as collective-choice arrangements and monitoring, apply to structuring governance in the decentralized publishing model.

In *Governing Knowledge Commons*, Frischmann et al. "propose a framework for studying knowledge commons that begins with the Institutional Analysis and Development (IAD) framework developed and used by Elinor Ostrom and others and adapts it to the unique attributes of knowledge and information" (Frischmann, Madison, and Strandburg 2014). This framework builds on Frischmann's previous work, *Infrastructure: The Social Value of Shared Resources*, in which Frischmann explores "how infrastructure generates social value and the role that commons management plays in facilitating value generation by infrastructure users" (Frischmann 2012). Frischmann's books provide robust demonstrations of the ways that shared, decentralized infrastructures have been theorized as viable alternatives for managing collective resources, including knowledge.

In "Modelling Science as a Contribution Good," Kealey and Ricketts argue that science is a "contribution good," benefiting contributors more than non-contributors – which reinforces the case for a decentralized model rewarding participation over passive consumption (Kealey and Ricketts 2014). Rennie and Potts's "Contribution Systems: A New Theory of Value" builds on Kealey and Ricketts to propose a contribution theory of value in digital economies, where contributions are treated as a form of property with emergent property rights in the commons, providing a basis for thinking about decentralized academic publishing as a contribution-driven model (Rennie and Potts 2024).

Potts et al. also propose the knowledge club model of journals in "A Journal is a Club" (Potts et al. 2016), arguing that journals can be understood as self-constituted groups that balance positive externalities (citations, readers) with negative ones (crowding). The framing of a journal as a club provides a connection between the existing modes that academic communities use to manage knowledge resources and the relevant alternative models.

1.3 Systems Engineering as a Solution Framework

As Donella Meadows explains in *Thinking in Systems: A Primer*, even the largest and most complex systems can be influenced – or rather, "danced with" – especially if one is able to identify the "leverage points" at which they are susceptible to directed pressure (Meadows and Wright 2008). Although the academic public industry's dysfunction is deeply entrenched, the development of decentralized public infrastructure for academic publishing will facilitate interventions into the system's key leverage points (e.g. peer review processes, IP rights, and publication mechanisms) – interventions with the potential to produce systemic change. The resulting improvements in transparency, fairness, and accessibility are likely to have transformative effects on both science and society.

In *Engineering a Safer World*, Nancy Leveson explains how systems engineering can guide the conceptualization, design, implementation, testing, operations, maintenance, and governance of decentralized infrastructure, ensuring that it is robust, scalable, and adaptable to the community's needs (Leveson 2012). Leveson's approach insists that systems must be designed with both technological robustness and social adaptability, allowing fit-for-purpose systems to evolve with operator needs, rather than by imposing institutional controls that fail to adequately account for human factors.

Nataliya Shevchenko's "An Introduction to Model-Based Systems Engineering" highlights the importance of managing complexity in heterogenous or large-scale systems, and details the ways that a model-based approach can assist in doing so (Shevchenko 2020). Model-Based Systems Engineering helps to integrate various elements – such as the infrastructure and processes surrounding peer review, curation, and publication – into a systemic whole. The application of systems engineering practices to economics contexts allows us to "imprison complexity through modularity," to borrow a phrase from Richard N. Langlois, making the resultant systems at once more legible and more resilient (Langlois 2023).

1.4 Technology as Social Practice

Systems engineering acknowledges that technologies often influence the structure of the systems in which they participate – but a systems-theoretical view of technology, such as that advanced by Lucy Suchman et al. in "Reconstructing Technologies as Social Practice," considers the significance of a technology to be inseparable from the social practices related to its use (Suchman et al. 1999). Decentralized systems re-engineer how academic value is created, shared, and rewarded by reconfiguring social dynamics around their redesigned infrastructures; technological changes without accompanying social changes are insufficient to produce significant shifts in socio-technical systems. Furthermore, as Suchman demonstrates in "Organizing Alignment: A Case of Bridge-Building," building safe and reliable infrastructure is tantamount to building the institutions that develop, operate, and maintain those infrastructures (Suchman 2000). Building institutions which maintain reliable infrastructures is especially challenging when those institutions aspire to be *permissionless*, as is often the case in the web3 context (Nabben and Zargham 2022).

Kei Kreutler builds on Suchman's thought in "How to Repair a Spaceport," an exploration of the work required to create and maintain complex socio-technical systems (Kreutler 2023b). In *Artificial Memory and Orienting Infinity*, Kreutler demonstrates that useful memory requires context, and that an adequate account of context requires an understanding of the practices that surround one's object of inquiry (Kreutler 2023a). Considered alongside Suchman's work, Kreutler's writings make clear that decentralized technology must evolve along with the practices of the academic community, allowing for flexibility and self-governance.

In “What Constitutes a Constitution?” Zargham et al. show how blockchain governance can support the kind of distributed decision-making that is relevant to decentralized academic publishing systems, making it possible to create more decentralized institutions responsible for developing, operating, and maintaining the infrastructure that facilitates academic publishing (Zargham, Alston, et al. 2023).

Decentralizing academic publishing can thus be viewed as a process of “Self-Infrastructuring Knowledge Networks,” as discussed by Zargham and Reed in a 2024 lecture of the same title (Zargham and Reed 2024). As Kelsie Nabben argues in “Web3 as ‘Self-Infrastructuring’: The Challenge is How,” processes of self-infrastructuring involve an “inherent tension” that arises from the attempt to “[express] ideals in coherent technical and institutional infrastructure” (Nabben 2023). A systems engineering approach makes it possible to navigate this tension – although it is never possible to eliminate it entirely.

1.5 An Actionable Path Forward

The remainder of this report will detail how systems engineering can guide the design and evolution of a decentralized infrastructure for academic publishing, ensuring that it addresses the current system’s failings while evolving dynamically through social interaction and technological refinement. Its recommendations seek to combine Leveson’s systems-engineering and Suchman’s social-practices perspectives, in order to create a system that is both technically robust and socially responsive – with continuous improvement based on feedback from the academic community, in a manner akin to how complex systems evolve.

Zargham and Ben-Meir’s “Method for Functional Decomposition of Organizations and Their Environments” (Zargham and Ben-Meir 2023b) builds on Lawrence Lessig’s “Pathetic Dot Theory” (Lessig 2006) to identify four modalities of regulatory pressures that act on organizations from without, along with four types of “constitutive infrastructure” that enable organizations to resist these external pressures – an understanding that is reflected in this report’s understanding of how a decentralized platform will be shaped by the interplay of organizational structures and regulatory forces, and its recommendations for how to re-engineer critical infrastructures within academic publishing. Its recommendations are also informed by Alston et al.’s *Institutional and Organizational Analysis*, which highlights how institutional economics and organizational behavior are essential to any attempt at rethinking how academic publishing might be governed (Alston et al. 2020).

This report proceeds as follows: Section 2 introduces the project’s motivations and stakeholders. Section 3 provides a glossary and detailed definition of the DeSci Ecosystem, and Section 4 reviews the systems engineering process with reference thereto. Section 5 provides a deep dive into the requirements and design work produced by BlockScience, and Section 6 identifies future work.

2 Motivating the DeSci Ecosystem

2.1 Animating Purpose

The Purpose of the DeSci Ecosystem is to organize, enrich, and curate ever growing streams of knowledge artifacts being produced across a wide range of scientific research endeavors in an open and interoperable way that alleviates some of the pain points associated with existing peer review processes and curation in traditional academic journals. The DeSci Ecosystem aims for these artifacts to be publicly-accessible.

Under this model, the raw intellectual production – including publishing artifacts and making claims about published artifacts – is enabled for any network participant with an identity on the network.

The DeSci Ecosystem is architected to attribute, curate, and organize diverse streams of scientific artifacts. These processes of enrichment occur on an on-going basis. The enriched artifacts remain publicly accessible to feed forward into future production. **Figure 1** provides a visual representation of this Animating Purpose.

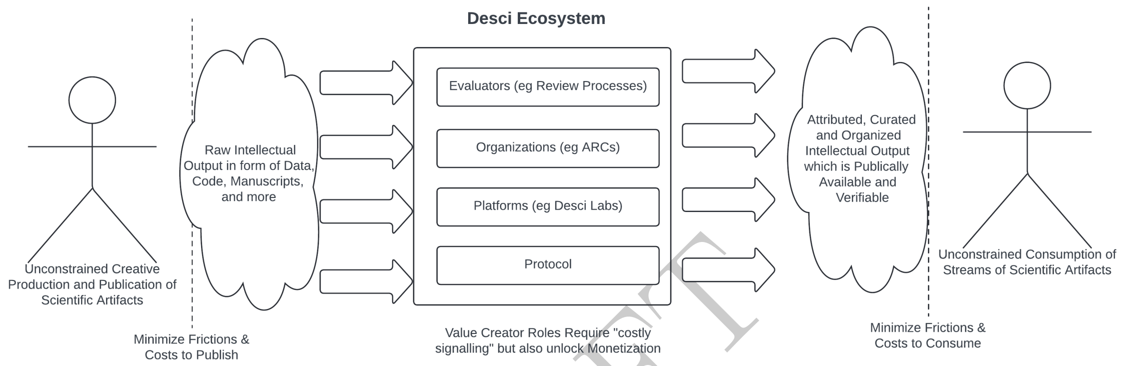

Figure 1: The Animating Purpose of the DeSci Ecosystem

While raw intellectual production is designed to be low friction, enrichment processes – such as creating organizations and evaluating claims about artifacts – are expected to have friction points and associated costs. These enrichment activities require "costly signalling," but are also the sites of monetization where contributors may be compensated.

2.2 Stakeholders

While individual Participants and Organizations engage with the DeSci Platform on a micro level, macro stakeholder groups inform the broader context of the ecosystem and operate on a slower timescale. These include the DeSci Foundation and DeSci Labs, as well as the broader public, notably: academic and citizen researchers, existing institutions that manage research, emerging institutions such as Metagov, and contributors to technical infrastructure such as developers and maintainers (Metagov 2023). Scientific computing communities such as those open-source software projects supported by NumFocus are good examples of scientific contributions that are underrepresented (NumFocus 2023a). **Table 1** maps the value propositions of different stakeholder groups.

Stakeholder	**Value Proposition**
Society	Indirect beneficiaries to improved scientific publication and evaluation.
Researchers & Editors	Open-access and Open data publishing standards. Less fragmentation. Better support for code and data sets. Discoverability and Accessibility.
Journals & Institutions	Harness emerging tech with native support for machine-readable metadata. Make meta-science computationally tractable. Scalability and Interoperability. New business models.

Table 1: Stakeholders and value propositions

Appendix A: User Stories Schema details the process by which the user stories that inform this stakeholder mapping were collected and situated within the DeSci Ecosystem.

Appendix B: Example ARC Workflows builds on these user stories to explore both the motivations and the high-level functional requirements of various *types* of stakeholders in the DeSci Ecosystem: authors, validators, ARCs, editors, and curators. It also details a functional ontological vocabulary that satisfies these requirements for the programmatic validation of research objects, elaborates a workflow for a Codex implementation of a programmatic evaluator, and offers sample Python code for an automated validator.

3 System Definition

The first step in understanding an organization – as Zargham et al. write in "What Constitutes a Constitution?" – involves mapping that organization's People, Purpose, and Environment (Zargham, Alston, et al. 2023). Having identified the Purpose of the DeSci Ecosystem in **Section 2.1**, and having begun mapping the People whose actions and interactions fall under its purview in **Section 2.2** and **Appendices A** and **B**, it is now necessary to attend to the Environment in which it operates. Conversations with various stakeholders quickly revealed conflicts amongst users' experiences and requirements, which made it clear that mapping the DeSci Ecosystem would require a layered approach.

BlockScience thus performed a decomposition informed by the "Method for Functional Decomposition of Organizations and Their Environments," which results in "a map of the forces acting on the organization, the structures and mechanisms that the organization has developed to enable itself to either resist or harness those forces, and the new pressures that these structures and mechanisms exert on the members of the organization in question [... that] makes it possible to identify what is enabling and/or interfering with the organization's ability to achieve its goals" (Zargham and Ben-Meir 2023b). This process made it possible to break apart many of the apparent conflicts in stakeholder requirements.

The process of functional decomposition also made it apparent that the DeSci Ecosystem is, in its structure, an implementation of Data Mesh Architecture – which, as Sisson et al. explain in "Data Mesh Architecture: Interoperability, Co-Operation, and Co-Regulation," carries meaningful implications for the ecosystem's technical, economic, and governance architectures (Sisson et al. 2024).

3.1 Glossary

Table 2 provides a glossary of key terms related to the DeSci Ecosystem and recommended design.

Term	**Definition**	**Example**
Participant	An individual level identity within the system; currently identified via Orcid, or an email and wallet address.	An author, member, researcher, or referee.
Organization	A group level identity within the system, consisting of one or more Participants.	An ARC, an academic journal, or community of practice.
Artifact	A publishable research object.	Manuscripts, software repositories, or data sets.
Attribute	A property of an Artifact.	Open-access, peer-reviewed, or reproducible.
Claim	An unverified attestation made by a Participant that an Artifact has an Attribute as defined by an Organization.	Metadata provided by the participant who uploaded the artifact.
Evaluator	A service operated by an Organization which evaluates Artifacts for specific Attributes based on Claims made by Participants.	An automated format compliance checker, or an implemented peer-review process.
Evaluation	A verified attestation produced by applying an Evaluator to an Artifact.	A journal acceptance, badge, or achievement on GitHub.
Authorship	Mapping that returns the author of an artifact, where the author belongs to the set of Participants, and an Artifact to the set of published Artifacts.	Authors of an academic paper or maintainers of an open-source software repository.
Membership	Subset of Participants that belongs to an Organization.	Editor(s) of a journal or data product owner(s) for an algorithmic evaluator.

Table 2: Glossary

3.2 Decomposition by Layers

The system is comprised of three layers: (i) the Protocol, (ii) the Platform, and (iii) Applications, which have features that support both Organizations and individual Participants. Following the Pace Layering framework, which organizes systems into layers based on their speed of functioning, the Protocol layer should be the slowest operating and most difficult to change. The Protocol is inclusive and unopinionated to the greatest extent practical. One joins the peer-to-peer network simply by participating in a manner compliant with the DeSci Protocol.

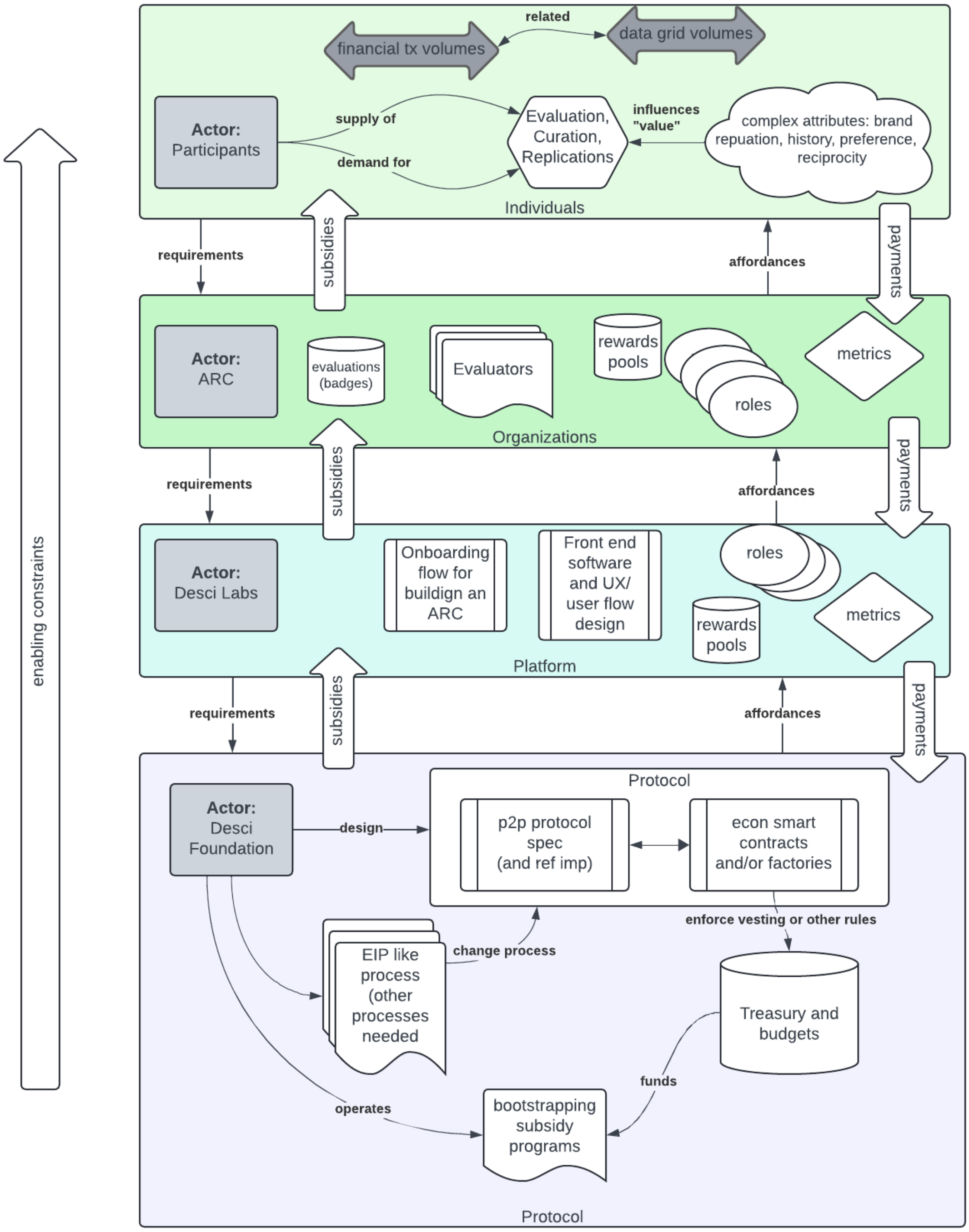

Figure 2: Subsystems decomposed by layer

Figure 2 depicts this layered architecture. At the bottom is the Protocol Layer. The primary actor at this layer is the DeSci Foundation, who are responsible for designing the Protocol (itself composed of a P2P Protocol Spec and Reference Implementation, and economic smart contracts and/or factories). These contracts will enforce vesting (or other) rules, which will fund the Treasury and the Foundations' budgets; these budgets will fund the bootstrapping subsidy programs (operated by the DeSci Foundation). The Foundation will also be responsible for maintaining the Protocol Layer through an EIP-like process, by way of which it can revise the P2P protocol spec as-needed.

Above the Protocol Layer is the Platform Layer, which is an abstraction layer between the Protocol and potential users, allowing users to have a more guided experience. Although DeSci Labs is operating the first Platform, competition between Platforms is expected to be good for the ecosystem in the long run. The Platform Layer will receive funds from the Protocol Layer in the form of subsidies, and will make payments back to the Protocol Layer in the form of fees. The primary actor at the Platform Layer is DeSci Labs, who will be responsible for developing the on-boarding flow for building an ARC, front-end software and UX/user flow design. The Platform Layer provides the affordances required for users to take on specific roles, and so should have reward pools to compensate users for taking on those roles. Finally, the Platform Layer generates metrics that can be used as feedback on any layer.

Above the Platform Layer is the Application Layer, whose users can be further subdivided into Organizations and Individuals. At the organizational level, the primary actor is an ARC, which operates Evaluators that issue Evaluations (badges). A rewards pool should exist to encourage ARCs and their members to take on necessary roles within the broader ecosystem (funded by subsidies feeding up from the Protocol and Platform Layers). Additional metrics are also generated on this layer. At the Individual level, Participants produce a supply of/demand for Evaluation, Curation, and Replications. A complex set of attributes, such as brand reputation, history, preference, and reciprocity will influence the "value" of these services. At this layer, the treasury is dynamic, due to ongoing minting and burning. The number of tokens needed to support transactions flowing through the system should be proportional to the amount of data flowing through the system.

3.3 Protocol Layer Affordances

The Protocol allows the following affordances to actors:

- Participants publish Artifacts
- Participants make Claims
- Participants form Organizations
- Organizations operate Evaluators
- Evaluators make Evaluations

Figure 3 illustrates how these affordances interact. Participant 1, the Author of Artifact y, makes a Claim of the form: "My Artifact y has Attribute x, as defined by Organization A." Organization A – the ARC referenced in the initial Claim – will then operate its Evaluator on Artifact y *in order to determine* if it does, in fact, possess Attribute x as the ARC in question defines it. This process will result in an Evaluation of the initial Claim. At the same time, Organization B, which might define Attribute X differently (and thus operate a different Evaluator to determine whether or not Artifact y does in fact possess it).

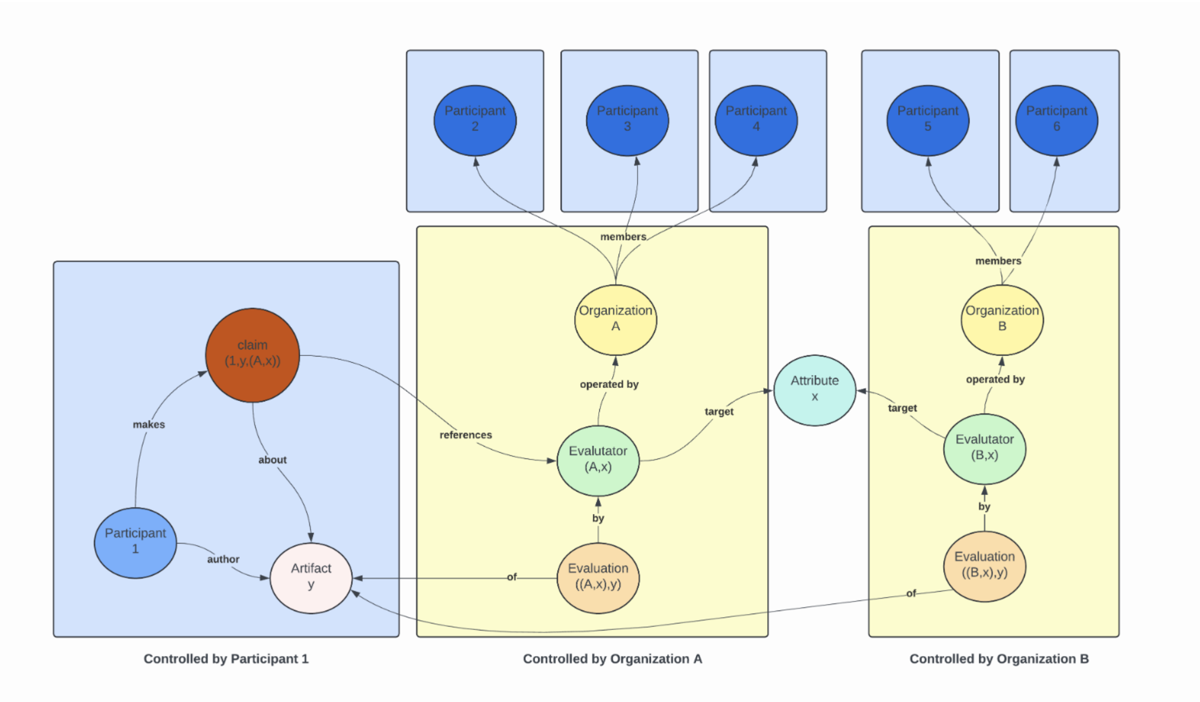

Figure 3: Evaluator diagram

Evaluators close the loop between the Organizations that operate them and the Evaluations being published in accordance with the Protocol. As shown in **Figure 3**, an Evaluator is controlled by a specific Organization to evaluate Claims made by Participants, where a Claim is an unverified attestation that an Artifact has an Attribute as defined by an Organization.

3.4 Protocol Layer Data Model

The DeSci protocol describes the rules by which the state of a peer-to-peer network can be modified by actors within the network. Given the above affordances, the state of the system at a given point in time can be described by:

- The set of all Participants
- The set of all Artifacts
- The set of all Organizations
- The set of all Attributes
- The set of all Evaluators
- The set of all Claims
- The set of all Evaluations

The state of the DeSci Network is best visualized as a graph with typed nodes and edges whose semantics depend on which types of nodes are being linked. The linked nature of the peer-to-peer network state is demonstrated in the Networkx prototype depicted in **Figure 4**. A detailed mathematical specification with notation for both states and state transitions, along with considerations for learning from inconsistencies, can be found in **Appendix C: Protocol Mathematical Specification**.

3.5 Assessing Implementability

The above sections define the "state of the world" as a network with typed nodes, as well as the affordances through which actors may induces changes to the state of the world. As a check to ensure these definitions are consistent and implementable, a simple proof of concept (PoC) was implemented in Python using Networkx. Source code for this PoC can be found in this Jupyter Notebook.

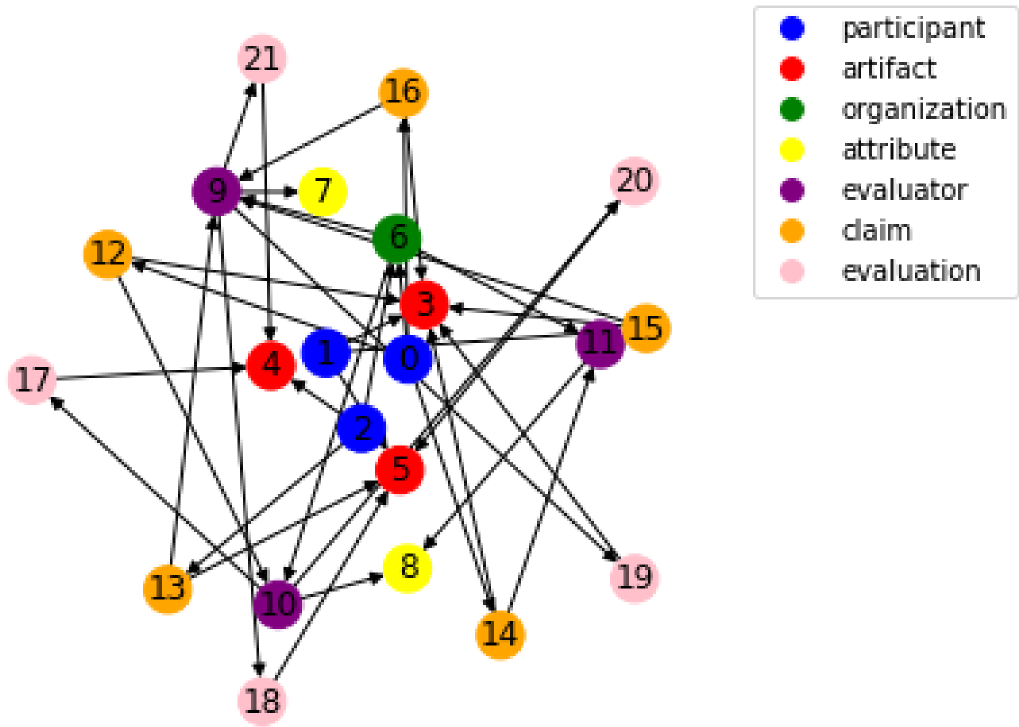

Figure 4: Networkx prototype Source Code

This implementation prioritizes defining the data model (see **Section 3.4**) and creating synthetic data to represent a valid state of the world, then proceeds to demonstrate how each of the affordances in **Section 3.3** induce valid state transitions over time.

The network examined in this PoC is implemented as a graph with nodes representing instances of all data model entities and edges representing outcomes of exercising all affordances. The network in instantiated by the `initialize_network` function defined in the second code block the referenced Jupyter Notebook using the following parameters:.

- `n`, the number of Participant nodes, is 10.
- `m`, the number of Artifact nodes, is 10.
- `p`, the number of Organization nodes, is 3.
- `q`, the number of Attribute nodes, is 5.
- `r`, the number of Evaluator nodes, is 10.
- `s`, the number of Claim nodes, is 150.
- `t`, the number of Evaluation nodes, is 75.

The `initialize_network` function also creates the following edges:

- Author edges connecting each Artifact node to one randomly selected Participant node representing outcomes of the Participants publish Artifacts affordance.
- A randomly determined number within the specified range of Member edges connecting each Organization node to Participant nodes representing outcomes of the Participants form Organizations affordance.
- Operator edges connect each Evaluator node to one randomly selected Organization node and Target egdes connect each Evaluator node to one randomly selected Attribute node representing outcomes of the Organizations operate Evaluators affordance.

- Claimant edges connect each Claim node to one randomly selected Participant, Claimed-About edged connect each Claim element to a randomly selected Evaluator, Claimed-For edges connect each claim to a randomly selected Artifact, and the value of each Claim node is randomly set to True or False representing outcomes of the Participants make Claims affordance and some outcomes of the Evaluators make Evaluations affordance.
- Evaluated-By edges connect each Evaluation node to a randomly selected Evaluator node, Recipient edges connect each Evaluation node to a randomly assigned Artifact node, and the value of each Evaluation node is randomly set to True or False representing the remaining outcomes of the Evaluators make Evaluations affordance.

This implementation exists solely to show that this system, as defined above, is implementable, demonstrating that such a networked dynamical system is indeed well-defined. Furthermore, it provides guidance on how data collected from node operators may be assemble into legible graph-structured data for the purposes of navigation and analysis.

3.6 Layers & Composability

The above layered model of the system is an open systems model where each layer performs a function, but the actors and activities within each have space to evolve.

This evolution may take the form of new applications and organizations emerging to fill different niches within the broader scientific community. It may involve changes to the patterns offered by the DeSci Platform, or the emergence of other platforms which share the same underlying peer-to-peer protocol and network.

Critically, the layered model presented still builds on top of existing infrastructures (such as the Ethereum Network and the Ceramic Network) and ultimately supports productive labor that exists beyond its scope (scientific production and evaluation).

4 Economic Systems Engineering

The following section applies the engineering design process to develop an economic model suitable for the DeSci protocol, given the animating purpose and system definition above. The engineering process helps to reconcile narratives with implementations, with the goal of setting expectations for users and operators. This document can be thought of as the green circle in **Figure 5** – an effort to create a clear connection between narratives about the DeSci Ecosystem (the red circle) and the actual system being implemented (the blue circle), with the ultimate goal of achieving alignment with end-user expectations and experiences (the yellow circle).

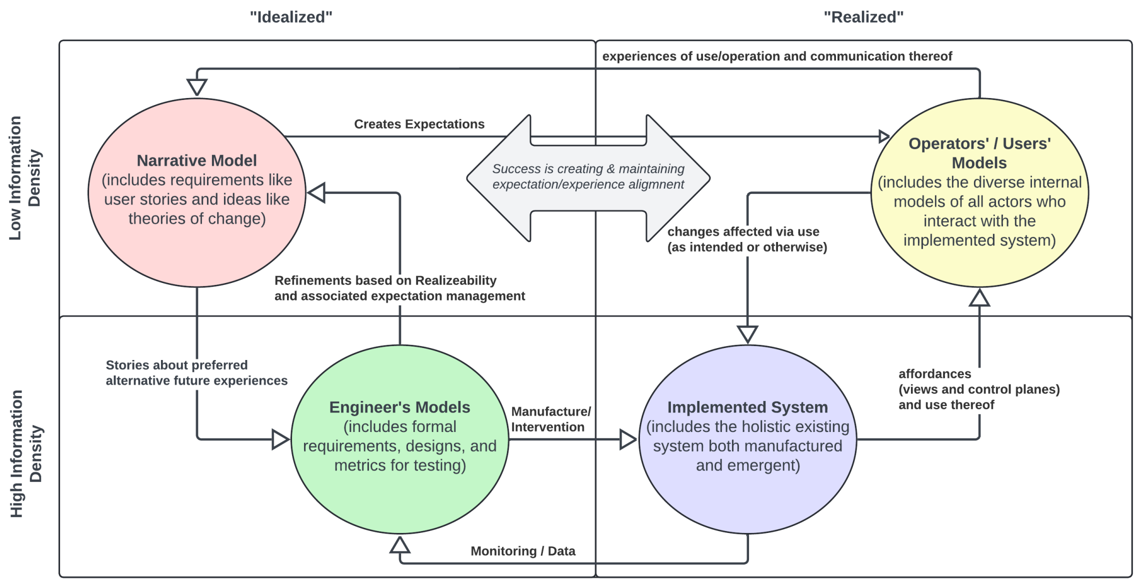

Figure 5: Mental models in the systems engineering process

As detailed in "Block-by-Block: Managing Complexity with Model-Based Systems Engineering" (Zargham and Ben-Meir 2023a), the engineering process can be broken down into: (i) Ideation and Conceptualization, (ii) Requirements and Design, (iii) Implementation, Integration, and Testing, (iv) Operations and Maintenance, and (v) Governance and Evolution.

This report is situated in step (ii) Requirements and Design. It is worthwhile, however, to locate its recommendations within the broader context of the project's entire life-cycle. This section thus offers a brief look at the work involved in **all five stages** of this project's life-cycle.

4.1 Ideation and Conceptualization

The first stage of the engineering life-cycle, Ideation and Conceptualization, involves both conceiving an idea and developing that idea into a coherent concept – mapping the project's stakeholders, identifying new capabilities and the possibilities that they enable, building Proofs of Concept (PoCs), and conducting experiments to validate critical assumptions.

The conceptualization of decentralized science (DeSci) has been shaped by efforts both within and beyond the web3 sector, each of which offers a unique perspective on how to decentralize the processes of scientific research, peer review, and publication. In the introduction of this report, alternative conceptions of scientific publishing and review were discussed. This section will highlight specific contributions which are situated in the web3, open-source software and open science landscapes, before the next section develops a structured systems engineering approach to bringing these ideas into practice.

The web3 sector has played a significant role in promoting decentralized frameworks for science, investigating the ways that blockchain technologies and decentralized autonomous organizations (DAOs) can be leveraged to enable alternative models of institutional and organizational governance. Ding et al.'s white paper "DeSci Based on Web3 and DAO: A Comprehensive Overview and Reference Model," which proposes a method for utilizing DAOs to manage decentralized scientific communities, is an exemplary representative of this body of thought (Ding et al. 2022). Framing DAOs as structures that enable transparency, trustless interaction, and decentralized governance, Ding et al. propose tokenized incentive systems to encourage participation in a decentralized scientific process, providing an influential illustration of ways that web3 tools can be employed to disrupt the traditional mechanisms of scientific publishing/knowledge-sharing.

All too often, however, web3 projects mistakenly conflate "public goods" problems with coordination problems. Zargham et al.'s paper "The Difference Between Public Goods Problems and Coordination Problems ... And Whether It Matters" addresses the formal distinction between these two classes of problems from a game-theoretic perspective, and provides helpful guidance for avoiding this common conceptual error (Zargham, Moore, and Stephenson 2021).

The work of Christian Roessler and Philipp D. Koellinger, both members of the DeSci team, builds on these ideas, with a particular focus on how decentralized platforms can restructure incentives within the peer review process (Roessler and Koellinger n.d.). The economic model developed by Roessler and Koellinger proposes a marketplace for peer review services as a mechanism for mitigating or eliminating the inefficiencies of the traditional peer review system. Within this marketplace, decentralized tokens are used to reward reviewers and editors for their contributions – replacing the current centralized and altruism-dependent model of peer review with a decentralized system in which participants are compensated on the basis of the value that they provide to the broader community. Roessler and Koellinger's work reflects a core principle of DeSci: aligning incentives across participants in the scientific process, in order to increase transparency and accountability throughout.

Parallel decentralized science initiatives have also emerged from outside of the web3 community. The NumFOCUS Open Source Science initiative, for example, has been at the forefront of efforts to promote open-source tools for scientific computing, motivated by the belief that the use of openly-accessible software, libraries, and platforms will enhance both the transparency and the reproducibility of scientific research (NumFocus 2023b). While NumFOCUS is not directly pursuing blockchain-based decentralization, the organization's open-source philosophy is closely aligned with the values of the DeSci movement. Their work demonstrates that it is possible to decentralize science without blockchain technology, through a combination of collaborative governance and community-driven contributions to open-source infrastructure.

The broader open-access movement, which advocates making scientific research freely available to the public, offers another significant articulation of many of DeSci's core principles. Supported by initiatives such as the Public Library of Science (PLOS) (Public Library of Science 2023), OpenReview (OpenReview 2023), and the Open Science Framework (Open Science Framework 2023), this movement has played a critical role in the push to democratize science by breaking down paywalls and promoting transparency. Although these efforts do not rely on blockchain technology, the attempt to dismantle the monopolistic control of traditional publishing houses is an important part of the broader effort to decentralize the production and promulgation of scientific knowledge.

4.2 Requirements and Design

Figure 6 offers a summary of the System Definition elaborated in **Section 3** of this report. At the bottom of the stack is public infrastructure such as a streaming client, a state machine client, and a wallet client. The DeSci Ecosystem – consisting of a Protocol Layer, a Platform Layer, and an Application Layer – is built atop the Protocol Layer. The Protocol Layer is made up of the Codex Protocol Client and DeSci smart contracts. The Platform Layer is built atop the Protocol Layer, and is made up of the User Interfaces for Participants and ARCs, as well as an Explorer that makes it possible to navigate the data that has been published on the platform. The platform is supplied with nodes of published knowledge, about which Participants can make Claims. Claims are made in reference to the Evaluators operated and maintained by ARCs, who run these Evaluators in order to produce Evaluations of such claims (the "demand side" of the Platform Layer's knowledge market). The Evaluators themselves are developed at the Application Layer, in the form of applications that have been designed for specific uses. Although Evaluators are a protocol-level concept, it does not make sense for them be entrenched beyond the application layer; instead, Participants and ARCs should have the ability to develop and operate any Evaluators for which there exists sufficient demand.

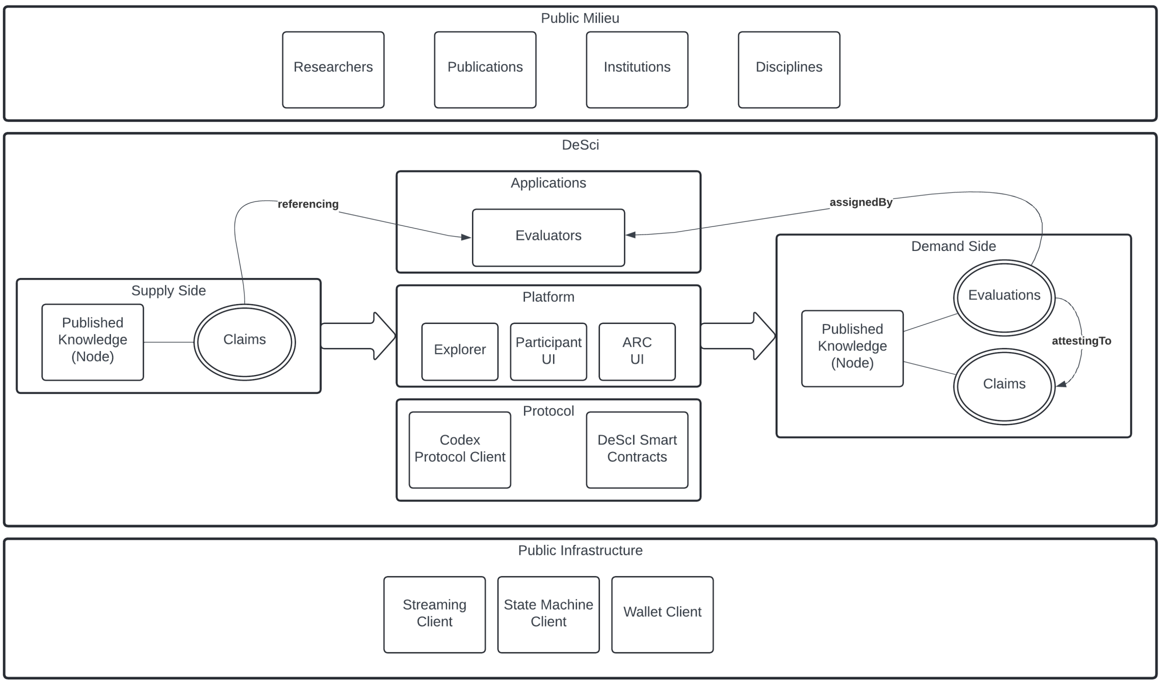

Figure 6: Visualization of system definition

Once the system has been defined at this level of detail, it becomes possible to proceed to the Requirements and Design stage of the engineering process, which is the primary focus of this report.

The Requirements and Design stage can be further broken down into:

- a) mapping the design space
- b) documenting the requirements
- c) analyzing the requirements to identify trade-offs
- d) specifying a design
- e) analyzing the design with respect to the requirements
- f) identifying considerations for calibration (which manifest as governance considerations in decentralized systems)

Each of these steps is addressed in detail in **Section 5: Economic Requirements and Design**.

4.3 Implementation, Testing and Integration

The Implementation, Testing, and Integration stage of the systems engineering process is focused on turning the designs from the previous stage into actual artifacts, testing these components, and carefully weaving them in to the already-existing system to which they will be added. These processes ensure that the system's components both inter-operate as expected **and** can be composed into a coherent whole that demonstrably meets the system-level design requirements from the previous stage.

The DeSci team has already done significant work in developing the DeSci Nodes platform (DeSci 2023b), and in building the infrastructure to ensure that the Open State Repository is both **fair** and meaningfully **open** (DeSci 2023a).

BlockScience has been actively testing this infrastructure as it exists today. **Figure 7** shows the individual Participant page set up by BlockScience's CEO, Dr. Michael Zargham, as part of this testing; **Figure 8** shows the page for an ARC that BlockScience created and is operating.

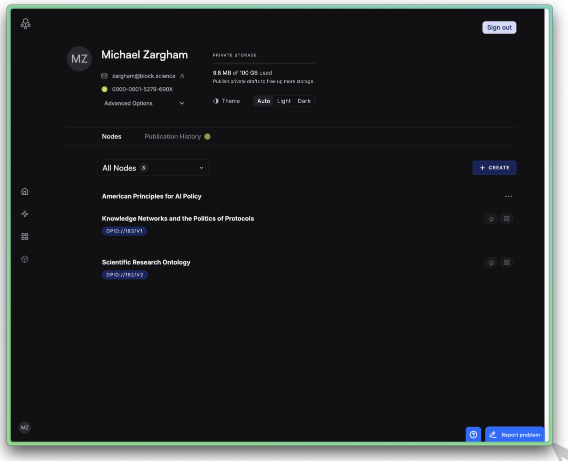

Figure 7: An individual participant page

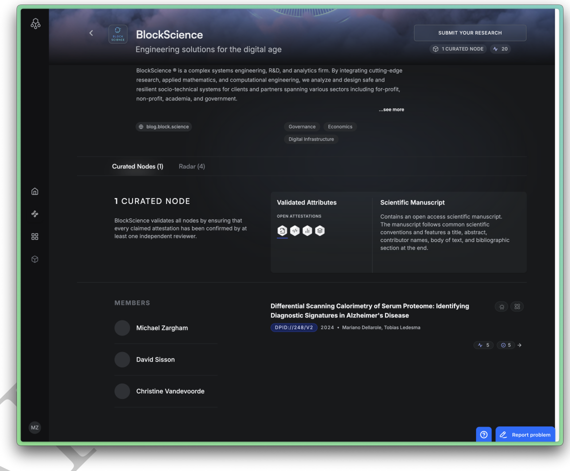

Figure 8: The BlockScience ARC's page

Crucially, BlockScience considers the economic systems whose designs are the primary subject of this report to be fundamentally **new** systems which have integrability with the DeSci Ecosystem's existing social and technical systems as a critical design requirement. To avoid either disrupting existing systems or monolithizing the platform, a highly modular design approach was pursued, informed by Richard N. Langlois's "Imprisoning Complexity in Modules" (Langlois 2023).

In addition to ensuring that the proposed economic systems can be integrated with the existing social and technical systems, it is also necessary to ensure that this newly-integrated system can *itself* be integrated with other existing systems (such as ORCID and Ceramic, in this instance). ORCID integration is visible in **Figure 7**.

4.4 Operations and Maintenance

Once the relevant designs have been implemented, integrated, and tested, the system can be deployed for its intended use. The next stage of the systems engineering process, the Operations and Maintenance stage, begins when the project goes live, and entails both **operating** the system under its real-world conditions, and performing ongoing **maintenance** to ensure that it is able to continue operating as intended.

As Zargham and Ben-Meir explain in "Knowledge Networks and the Politics of Protocols", protocolization must always negotiate a tension between **expressivity** (that is, how broadly *useful* the software is *potentially* able to be) and **usability** (which is to say, how *easy* the software is for users to *actually* use (Zargham, Ben-Meir, and Nabben 2023). At the level of infrastructure, it is important for the protocol itself to be *un-opinionated*, which necessitates some degree of configuration by its users; from the perspective of the application layer, however, users generally prefer an experience that is as close to "turnkey" as possible.

Once the economic systems are in operation, ARCs will function as "Domain Data Product Owners" within the DeSci Ecosystem – as described by in "Data Mesh Architecture: Interoperability, Co-Operation, and Co-Regulation" (Sisson et al. 2024). ARCs should therefore be allowed to manage their own operations as autonomously as possible – but in order to enable Domain Data Product Owners to manage their own operations efficiently, the DeSci Platform must afford them access to an array of highly-automated (and thus intrinsically-*opinionated*) tools.

Insofar as the Protocol Layer should favor *expressivity*, and the Application Layer should favor *usability*, the purpose of the Platform Layer is thus to function as what Sisson et al. would call a "beneficial intermediary" between the two. The job of the Platform is to lower the lift of operating and maintaining an ARC, without taking away rule-making and steering autonomy – to increase what Zargham, Zartler, et al. refer to in "Disambiguating Autonomy" (Zargham, Zartler, et al. 2023) as the "*tactical* autonomy" of ARC operators – that is, their ability to decide *how to get the job done* – by providing them with access to *technical* automations, which are necessarily highly opinionated.

The Platform Layer thus **creates value** for the broader ecosystem by conserving attention, and if properly designed could serve as the foundation for a profitable business. There are also long-term benefits to not forcing platform lock-in, and allowing other platforms to potentially build on the same protocol.

4.5 Governance and Evolution

To borrow once again from "Disambiguating Autonomy," one might say that the Operations and Maintenance stage is essentially *tactical* in nature, insofar as it involves optimizing the system's performance under a fixed set of rules and/or relatively constant conditions (Zargham, Zartler, et al. 2023). The Governance and Evolution stage, by contrast, is fundamentally *strategic* – it is about steering the system safely through those times when the existing frame of reference is inadequate, and the system's governing rules or operating conditions are in flux.

The Governance and Evolution stage involves interventions into the design of the system *once it is already up and running, and there are already stakeholders who have come to rely on it.*

The set of mechanisms through which a system's governance is conducted is referred to as that system's "Governance Surface," and describes *which* design decisions can be changed, by *whom*, and under *what* conditions. It can be thought of as the steering of the system relative to the Animating Purpose. Building on the insights of Voshmgir and Zargham's "Foundations of Cryptoeconomic Systems" (Voshmgir and Zargham 2020), Zargham and Nabben develop the concept of a Governance Surface in "Aligning 'Decentralized Autonomous Organization' to Precedents in Cybernetics" (Zargham and Nabben 2023).

Figure 9 illustrates how the Governance Surface operates on the broader system within which it is entrenched:

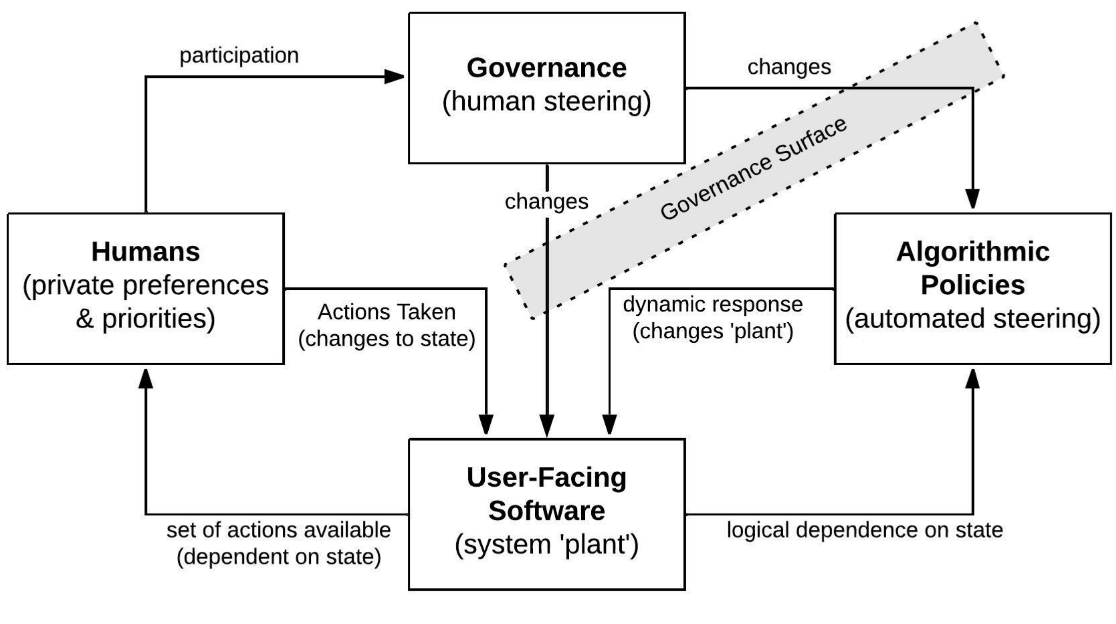

Figure 9: Governance Surface diagram

Viewing the DeSci Protocol as digital public infrastructure, there are two recommended tracks to modularize the governance surface:

1. Policy making processes, which are common to the institutions responsible for physical public infrastructure.

2. Technical infrastructure development and maintenance, which is standard for web3 and open source technical projects.

Policymaking processes should govern items like system level success metrics, appropriations of funds, subsidy program metrics, and reward functions. Inputs from Affected Stakeholders and Expert Supervisory Committees should feed into decision-making processes made by the Accountable Stakeholder Decision Authority. Monitoring infrastructure and public data can then provide transparent feedback on the outcomes of implemented policies, ideally in a feedback loop that informs future inputs from Affected Stakeholders and Expert Supervisory Committees.

Technical infrastructure, which includes an open source code base, data standards, and API specs, is often developed and maintained by a group of core developers who have decision authority based on their past contributions or roles in key organizations (for example, DeSci Labs or DeSci Foundation). Soft consensus and running code can be used for iterating on code satisfying existing specs, while EIP-like processes can be used for proposing substantive changes to specs or the Protocol.

5 Economic Requirements and Design

5.1 Design Space

The Design Space is the landscape of all possible designs for the Economics of the DeSci protocol. Given the System Definition above, this space can be segmented into the following components:

- (i) Micro-Economic Model
- (ii) Macro-Economic Model
- (iii) Subsidy Programs

Micro-Economic Model

The micro-economic model describes how the affordances of the network may be bound to resource requirements, including but not limited to staking DeSci tokens. This model provides a clear mapping between the data in the peer-to-peer network and the on-chain economic state encoded in smart contracts, at the level of identities, admissible actions, and transactions.

Macro-Economic Model

The macro-economic model describes how the total supply of DeSci tokens is regulated, as well as any parameters that regulate the token circulations from a relatively zoomed-out vantage point. Tokens may be in the possession of the Foundation, locked in vesting contracts, in circulation, or staked for their utility within the protocol (as defined by the micro-model).

Subsidy Programs

Subsidy programs are policies that describe rewards distributions according to measurable activity within the network. They serve two functions:

- (i) bootstrapping
- (ii) steering

Bootstrapping involves reducing risk and offsetting capital costs associated with participation in the network. Steering involves directing the activities of participants already active within the network.

5.2 Requirements

The requirements characterize what a "good" design is within a specific engineering project. In **Table 3**, the economic requirements are broken down into a set of individual criteria and associated to the subsystems they most directly affect.

Requirement	**Description**	**Affected Subsystems**
Bounded Supply	The total supply of DeSci tokens must have a strict upper limit.	Macro-Model
Direct Utility	The token must be directly related to value production in the network.	Micro-Model
Conservation	Token flows must be conserved; minting and burning need to occur across clearly defined boundaries.	Macro-Model, Micro-Model, Subsidies
Sustainability	The foundation needs to have tokens to allocate for steering in perpetuity.	Macro-Model, Subsidies
Precedented	Models are analogous to historical observed economic phenomena.	Micro-Model, Macro-Model
Transparency	The system can be monitored via publicly observable data sets.	Micro-Model, Macro-Model, Subsidies
Governability I	The macro-economic parameters have clear purposes and can be (re)calibrated.	Macro-Model, Micro-Model
Governability II	The subsidy policies have standard formats, clear mandates, and funds need to be appropriated seasonally.	Subsidies, Micro-Model
Stability	Governance processes are slow-moving such that changes are predictable and non-disruptive.	Macro-Model, Micro-Model, Subsidies
Legibility I	Costs and benefits for individual participants to take specific actions are clear.	Micro-Model, Subsidies
Legibility II	Structurally entrenched correlations between protocol use and economic trends.	Macro-Model, Micro-Model

Legibility III	The mandates, funding sources, and time periods of subsidies are straightforward to follow and validate success empirically.	Subsidies
Foundational	Economic models provide the basis for other technologies and institutions to be built.	Micro-Model
Permissionless	Participants can interoperate with the network by adhering to the protocol without requiring permission from an authority.	Micro-Model
Intersubjective	The economic model should avoid presuming an *a priori* ground truth.	Micro-Model
Fairness	Token allocations and subsidies do not privilege passive actors at the expense of active contributors.	Macro-Model, Subsidies

Table 3: Requirements and their affected subsystems

5.3 Requirements Analysis

Requirements Analysis involves examining the requirements to identify fundamental trade-offs that exist independent of specific designs. **Table 4** details inherent trade-offs between the requirements identified in **Section 5.2**.

Requirement 1	**Requirement 2**	**Trade-Off**
Bounded Supply	Sustainability	A strictly capped supply may limit the Foundation's ability to allocate tokens for steering in perpetuity, reducing flexibility for long-term sustainability.
Governability I	Stability	Making the system governable and flexible (Governability I) may introduce challenges to maintaining slow and predictable changes (Stability), as frequent recalibrations could disrupt long-term stability.
Conservation	Subsidy Programs	Ensuring token flows are conserved may conflict with dynamic minting and burning processes used in subsidy programs, where tokens may need to be distributed or removed for incentives.
Direct Utility	Fairness	Linking token utility directly to value production may inadvertently favor actors who can produce immediate value, creating imbalances for those contributing in non-immediate but meaningful ways (affecting fairness).
Precedented	Intersubjective	Using historically observed economic models (Precedented) can contradict the goal of avoiding presumptions of an *a priori* ground truth (Intersubjective), limiting innovative, decentralized approaches to economic modeling.
Governability II	Stability	Seasonally re-appropriating funds for subsidies may introduce unpredictability in funding (Governability II), which could create instability in participant incentives and thus overall system predictability.
Permissionless	Fairness	Ensuring permissionless participation may introduce fairness issues, as adversarial actors could exploit the system without necessarily contributing to value creation, thus contradicting fairness.
Transparency	Intersubjective	A highly transparent system may conflict with the goal of avoiding assumptions about objective truths (Intersubjective), as dashboard or other monitoring tools often assume a single viewpoint for interpreting data and outcomes.

Table 4: Requirements trade-offs

5.4 Recommended Design

Pursuant to the DeSci Ecosystem's Animating Purpose, a token native to the DeSci Network must provide enabling constraints and tools to impose signaling costs that facilitate curation and organization of scientific artifacts.

The defining analogy for the DeSci Token is that of "machine capital" – equipment required in service of productive labor. An Application is a knowledge-refinery, and its Evaluators are the filtration processes within that refinery. This invites an industrial machinery economic model where DeSci tokens (specialized machinery) must be acquired and deployed (staked) in order to operate any particular knowledge-refinery. More machinery (tokens) are required when the refinery has higher throughput (more total Evaluations per period) or more diverse processes (more Evaluators).

Machinery is made from physical materials which are conserved, but the machinery itself may experience wear, and worn equipment may be recycled. Insofar as machinery has market value, it is intrinsically tied to the value of that which it is used to produce, inseparable from the labor involved in that production. Therefore, the industrial machinery model provides a basis for coupling the micro-model (tokens as means of production required for productive labor) and the macro-model (physical limits on total machinery, wear, and recycling).

The above model offers an econo-physics that can be enshrined on-chain via smart contracts. Coupling the on-chain state with the distributed peer-to-peer network requires the ability to read the on-chain stake amounts associated with the identities of Organizations and Evaluators. This allows the tokens to aid in spam resistance, a key feature for a protocol aimed at curation.

Micro Model

The micro model serves primarily to manage frictions. Requiring staking on the part of actors performing curation roles aims to reduce spam while preserving permissionless participation. Below, we enumerate the affordances, while overlaying the frictions and costs associated with exercising those affordances, including any token-based mechanisms. **Table 5** characterizes the micro-model's affordances, and their associated frictions.

Affordance	**Knowledge Category**	**Relative Friction Level**	**Requirements**
Access Data	Consumption	Low	Read access to peer-to-peer network data is unrestricted.
Publish an Artifact	Production	Low	Participant must have DeSci Network Identity.
Make a Claim	Production	Low	Participant must have DeSci Network Identity.
Form an Organization	Curation	Medium	Participants must have on-chain identities.
Register an Organization as Active	Curation	High	Meet a staking threshold, α.
Publish an Evaluator	Curation	High	Organization must be active.
Register an Evaluator as Active	Curation	High	Meet a staking threshold, β.
Operate an Evaluator	Curation	High	Evaluator may not exceed baseline evaluation capacity allowance of γ evaluations per period Δ, unless staking additional tokens at rate $\frac{\beta}{\gamma}$ to unlock additional per period capacity.

Table 5: Micro model affordances and associated frictions.

This micro model provides micro-foundations for the macro model in the form of lower bounds on the total number of tokens staked as a function of the number of active organizations, evaluators, and evaluators' evaluation volumes, as well as the network-wide policy parameters α (tokens locked per organization), β (tokens locked per evaluator), and γ (evaluations per evaluator per time period Δ).

Note the separability between the controllable and uncontrollable parameters: user behaviors (forming organizations, publishing and operating evaluators) are uncontrollable, but the network-wide policy parameters are controllable via governance mechanisms. The micro model is discussed in greater detail in **Appendix D: Micro Model Details.**

Macro Model

The macro model is characterized using *Industrial Dynamics*, as described by Forrester (Forrester 1961). It is concerned primarily with the token supply, circulation, and long-term sustainability of the ecosystem.

Figure 10 illustrates the stock and flow dynamics for a DeSci token.

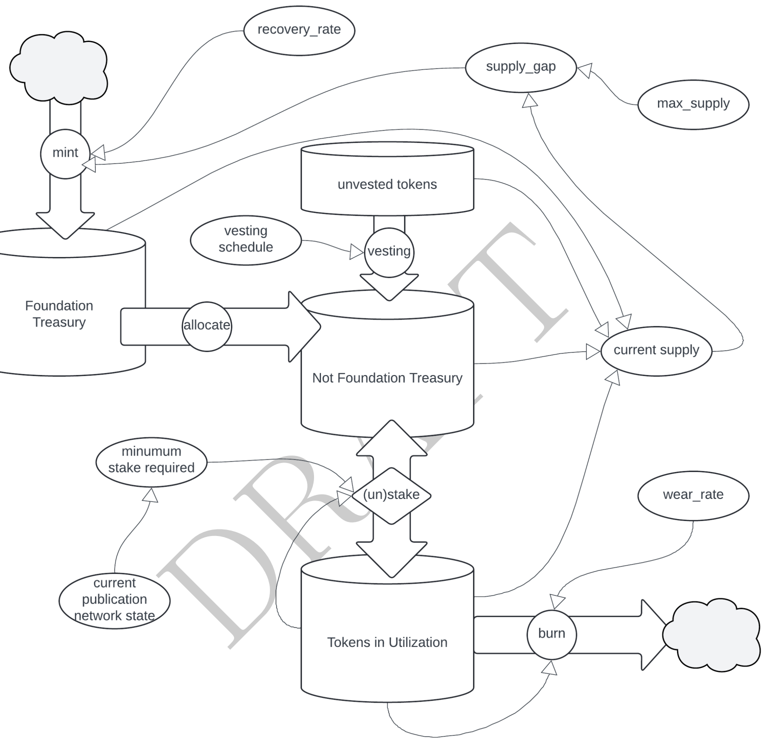

Figure 10: Token stock and flow dynamics.

Stocks: Building on the machinery analogy, tokens would exist in four high-level buckets, or "stocks":

- **Foundation Treasury:** Available for allocation to actors who will use them for productive curation labor.
- **Tokens in Utilization:** These are tokens in-use, staked on Organizations and Evaluators in service of curation activities.
- **Not Foundation Treasury:** Tokens circulating in the hands of other stakeholders which may be staked on Organizations or Evaluators but are not currently.
- **Unvested Tokens:** Tokens programmatically committed to ecosystem actors at the time of launch.

These stocks are coarse but sufficient to evaluate the design against the requirements. The sum of all four buckets must remain at or below a `max_supply` for all possible scenarios. It should be possible for the circulating supply to contract. Additionally, if the Network is active, the `Foundation Treasury` should not become empty. In order to assess the design, we must also characterize the flows.

Flows: The flows in the model represent the movement of tokens between the stocks:

- **Allocation from Treasury:** This flow is at the discretion of the Foundation. Rate limiting can be achieved primarily by initializing the system with `current_supply` much less than `max_supply`, so that the `Foundation Treasury` is a filling tank.
- **Minting to Treasury:** This flow is parameterized by `recovery_rate`, and can be conceptualized as a half-life, i.e., the amount of time it takes for `max_supply - current_supply` to be halved when there is no burning.
- **Burning due to Wear:** This flow represents the wear on the "machinery" as a result of use. This is parameterized by `wear_rate`, which can be conceptualized as a half-life, i.e., the amount of time it takes for `Tokens in Utilization` to be halved, if no new tokens are staked.
- **Staking and Unstaking:** This flow is at the discretion of token holders. Tokens can be staked on Organizations and Evaluators in order to keep them active and/or increase their evaluation capacity. Stakers putting tokens up should value the work being done enough to tolerate the wear. Losses due to the `wear_rate` are a form of "costly signaling" that disincentivizes Stakers from a "set it and forget it" mentality.
- **Vesting:** This regulates the rate at which tokens are introduced into the economy that were committed to ecosystem stakeholders. This ensures the economy is not flooded with circulating supply, which undermines the costly signaling function of staking.

A demonstrative example of this dynamical system is provided in **Figure 11**:

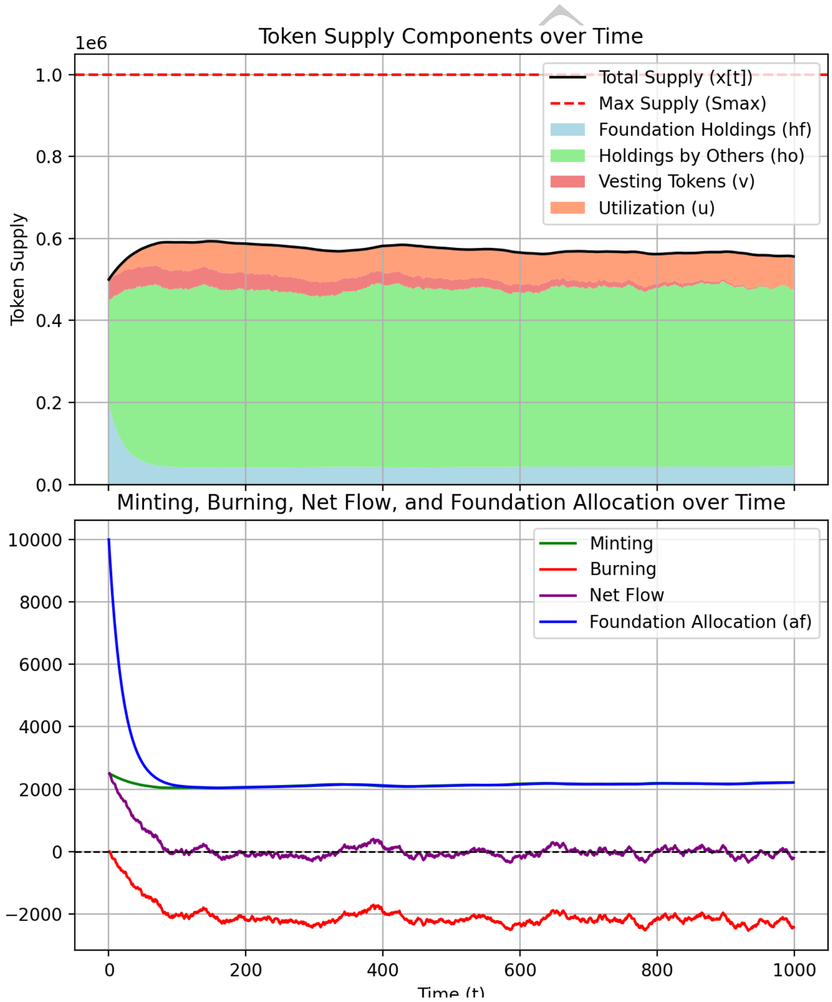

Figure 11: Example of the Dynamical System.

The properties of this design were validated via real analysis, and are discussed in greater detail in **Appendix E: Macro Model Details.**

Subsidy Programs

Subsidy programs enable the network to be bootstrapped and steered without setting expectations for perpetual rewards for any particular activity.

Subsidy Program Template:

- Tokens allocated to the program
- Target measurements on which the rewards are based
- Rewards as a function of measurements
- Eligibility criteria
- Relevant stakeholders
- Start and end dates for the subsidy program

These programs should have clear natural language descriptions of their purposes so that the outcomes can be qualitatively compared against their intent in addition to assessed based on the quantitative metrics they are targeting.

Candidate Bootstrapping Subsidies

Table 6 outlines three distinct subsidy programs defined according to the template above, each of which targets an aspect of the bootstrapping challenge. These subsidy programs can be tied back to the machine capital analogy by noting that machinery is being allocated to actors who have demonstrated willingness and capability to perform productive labor when provided with the means of production.

The token allocations as well as the start and end dates are not specified as they depend on details that are as yet unspecified.

Name	Purpose	Measurement	Eligibility	Stakeholders
Publication Volume	To incentivize actors to join the network, including creating accounts, connecting Orcid IDs, connecting crypto addresses, and publishing artifacts.	Count of publications. Rewards distributed proportionally to the square root of publications.	Accounts must have Orcid IDs and on-chain addresses connected, and verified publications.	Authors
Offset Costs	To offset frictions associated with setting up ARCs and publishing Evaluators.	Staking cost to operate an ARC and 3 Evaluators.	Proposals must be submitted with descriptions of Evaluators and their intended operation.	ARC Operators
Evaluator Usefulness	To reward Evaluator operators based on the usefulness of their Evaluators.	Number of qualifying Artifacts claiming their attestations.	Claims must be from participants eligible via the Publication Volume Program.	ARCs, Evaluator Operators who are looking for upside in crypto.

Table 6: Subsidy programs overview

These candidate bootstrapping subsidies are discussed in greater detail in **Appendix F: Subsidy Program Details**.

5.5 Design Analysis

Design Analysis reconciles the Recommended Design in **Section 5.4** with the Requirements and Requirements Analysis in **Sections 5.2** and **5.3**. These considerations help inform implementation and governance.

Table 7 provides a structured overview of how the design aims to satisfy the given requirements and includes qualitative assessments of the strengths and potential challenges of each solution.

Requirement	Means of Satisfaction	Assessment
Bounded Supply	Introduced a strict upper limit (`max_supply`) on the total number of DeSci tokens in circulation.	The bounded supply ensures control over inflation, but may limit flexibility for future expansions or new use cases. Appropriations need to be carefully managed to avoid depletion of resources.
Direct Utility	Tokens are staked for operational use, ensuring that tokens are tied to the productive functions (publishing, evaluating) within the network.	The system effectively ensures utility-driven value, but favoring directly measurable productive actions could potentially marginalize participants whose contributions are not immediately measurable.
Conservation	Implemented token burning via “wear” on tokens staked for Evaluators and Organizations, alongside careful conservation of token flows.	While the token wear-as-burn model helps align incentives and prevent “set and forget” behaviors, it introduces depreciation into accounting for token-utility and risks increasing the overall cost of participation for some actors.
Sustainability	Tokens held in the Foundation Treasury are reserved for steering via subsidies, supplemented by a minting mechanism with a capped recovery rate.	This approach provides long-term sustainability, but the treasury’s finite size could constrain steering efforts in the future, especially if token burn mechanisms are set to small values.
Precedented	The design follows analogies to historical economic models, particularly the machinery analogy for staking and token wear.	While grounding the design in familiar economic concepts helps with comprehension and stability, it is just an analogy, and as the system evolves it may deviate from its origins.
Transparency	Public on-chain data for all key actions (publishing, evaluating, staking) ensures transparency of the system’s operations.	Transparency is technically ensured, but operationally there must be block explorers, public dashboards, and other data services to ensure that non-technical stakeholders can be informed participants.
Governability I	Macro-economic parameters, such as staking requirements, are adjustable via governance mechanisms.	Provides flexibility for recalibration, but large or frequent governance adjustments could lead to instability or frustration among participants who prefer predictable, stable economic parameters.
Governability II	Subsidy programs adhere to a clear template, and funds are allocated seasonally by the Foundation.	Standardized templates for subsidies provide clarity, but seasonal allocations could result in delays or uncertainty in funding, particularly if participants rely on these programs to support their activities.
Stability	Governance processes are designed to move slowly, with parameters for changes carefully defined to ensure predictability.	Stability is achieved at the expense of agility. Slow-moving governance could limit the system’s ability to quickly respond to emerging issues, and participants may experience frustration if the system does not adapt to evolving needs in a timely manner.

Legibility I	Clear costs and benefits are tied to staking, with direct links between actions (publishing, evaluating) and their associated rewards and costs.	The legibility of the system helps participants understand the consequences of their actions, but the complexity of staking thresholds and token burning may make the system harder to grasp for newcomers or participants without a deep technical background.
Legibility II	Protocol-level actions are clearly mapped to economic trends, with metrics like staking volume and evaluator activity providing clear insights into system health.	This provides structural clarity and ensures protocol use can be tracked, but it may oversimplify correlations, potentially overlooking more nuanced or indirect contributions made by participants that don't fit into the clear-cut metrics.
Legibility III	Subsidy programs follow a straightforward template with clear goals, metrics, and timelines for success measurement.	The structured nature of the subsidies is easy to understand, but the potential complexity of measuring success based on both qualitative and quantitative outcomes might still lead to disputes or confusion over whether goals were met.
Foundational	The economic models provide an industrial machinery analogy that forms a strong basis for application-level incentives, ensuring productive contributions.	The foundational nature of the model allows other technologies and institutions to build upon it, but the industrial machinery analogy may be insufficient for future innovations that don't fit neatly into the model's constraints.
Permissionless	Participants can join the network by meeting staking requirements without needing approval from any central authority.	Permissionless participation ensures accessibility and inclusivity, but it could also open the door to exploitation by passive or malicious actors, potentially undermining the fairness and security of the system.
Intersubjective	The system avoids assumptions about any single economic truth, allowing for flexible interpretations of value within the network.	While the flexibility helps accommodate a variety of viewpoints and contributions, it may create challenges for designing uniform evaluation criteria, potentially leading to inconsistencies in how contributions are valued across the ecosystem.
Fairness	Token staking and subsidies are structured to reward active contributions, ensuring that passive actors cannot exploit the system.	Active contributors are incentivized, but the staking thresholds may disadvantage smaller or less resourced participants who find it harder to meet the economic requirements, raising concerns about fairness and accessibility for all participants. This design doesn't address the question of token allocations at launch.

Table 7: Design analysis overview

5.6 Governance Considerations

While this report's design recommendations seek to carefully balance and negotiate the trade-offs inherent in the design space, no design will ever be perfect. It is therefore critical that certain parameters be tunable by those responsible for the governance of the system.

Tunable Core Parameters:

- Wear rate
- Recovery Rate
- Org Registration Stake Required
- Evaluator Registration Stake Required
- Base Evaluations Per Period for Registered Evaluator
- Stake Required for Additional Evaluations Per Period

Tunable parameters of Seasonal Subsidy Programs based on the provided template:

- Tokens Allocated to the Program
- Target Measurements
- Rewards as a Function of Measurements
- Start and End Dates for the Subsidy Program

Economic governance of the ecosystem will consist of tuning along these parameters – they are a significant part of its Governance Surface.

6 Future Work

The DeSci protocol faces challenges fundamental to infrastructures. Infrastructures are capital intensive to design, develop, and deploy, but the benefits accrue to the users of the infrastructure who often take it for granted. Furthermore, the value created is often even more diffuse, as benefits of the infrastructure propagate beyond its immediate users to other beneficiaries of its use. This is most definitely the case for improving the efficiency and effectiveness of scientific publishing and evaluation. Researchers and academic institutions are the direct users of the infrastructure, but society at large benefits from defragmenting science.

While infrastructure is often invisible, one way to make the DeSci network and its associated benefits more visible is to build vertically, towards specific, demonstrable applications, in addition to scaling horizontally to a broad set of potential applications. DeSci Labs building a platform to make creating ARCs or connecting other applications to scientific publication and evaluation data feeds is a critical intermediate step. However, in order to demonstrate value, novel ARCs need to be developed.

As demonstrated in **Appendix B: Example ARC Workflows**, there are efficiency and effectiveness gains to be unlocked through the hybridization of human and machine workflows made possible by the DeSci protocol. The next step in the development of this ecosystem is to design and implement evaluators using the protocol and to demonstrate their usefulness by founding ARCs to operate them and securing paid subscribers to the associated services.

References

Alston, Eric et al. (2020). *Institutional and Organizational Analysis: Concepts and Applications*. Cambridge University Press.

DeSci (2023a). *Codex by DeSci*. Accessed: September 27, 2024. URL: https://codex.desci.com/.

— (2023b). *Welcome to DeSci Nodes*. Accessed: September 27, 2024. URL: https://nodes.desci.com/.

Ding, Wenwen et al. (2022). "DeSci Based on Web3 and DAO: A Comprehensive Overview and Reference Model". In: *IEEE Transactions on Computational Social Systems* 9.5, pp. 1563–1573. DOI: 10.1109/TCSS.2022.3204745.

Forrester, J.W. (1961). *Industrial Dynamics*. Students' edition. M.I.T. Press. ISBN: 9780262060035. URL: https://books.google.com/books?id=4CgzAAAAMAAJ.

Frischmann, Brett M. (Mar. 2012). *Infrastructure: The Social Value of Shared Resources*. Oxford University Press. ISBN: 9780199895656. DOI: 10.1093/acprof:oso/9780199895656.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199895656.001.0001.

Frischmann, Brett M., Michael J. Madison, and Katherine J. Strandburg (Sept. 2014). *Governing Knowledge Commons*. Oxford University Press. ISBN: 9780199972036. DOI: 10.1093/acprof:oso/9780199972036.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199972036.001.0001.

Kealey, Terence and Martin Ricketts (2014). "Modelling science as a contribution good". In: *Research Policy* 43.6, pp. 1014–1024. ISSN: 0048-7333. DOI: https://doi.org/10.1016/j.respol.2014.01.009. URL: https://www.sciencedirect.com/science/article/pii/S0048733314000195.

Kreutler, Kei (2023a). *Artificial Memory and Orienting Infinity*. © 2023 Ethereum Foundation, licensed under CC BY-NC 4.0. After 2026-12-13, licensed under CC BY 4.0. United States of America: Summer of Protocols. URL: http://summerofprotocols.com/ccplus-license-2023.

— (2023b). *How to Repair a Spaceport*. Accessed: 2024-09-27. URL: https://folklore.mirror.xyz/6Lm7xORHWDb-YoVtqv3Vc2G7dT73xiKws5pJBv9ADZE.

Langlois, Richard N. (Aug. 2023). *Imprisoning Complexity in Modules*. Working papers 2023-05. University of Connecticut, Department of Economics. URL: https://ideas.repec.org/p/uct/uconnp/2023-05.html.

Lessig, Lawrence (2006). *Code: And Other Laws of Cyberspace, Version 2.0*. Basic Books. ISBN: 9780465039142. URL: https://books.google.com/books?id=lmXIMZiU8yQC.

Leveson, Nancy G. (Jan. 2012). *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press. ISBN: 9780262298247. DOI: 10.7551/mitpress/8179.001.0001.

Lieff Cabraser Heimann & Bernstein, LLP (2024). *Academic Journal Publisher Antitrust Lawsuit*. https://www.lieffcabraser.com/antitrust/academic-journals/. Accessed: 2024-09-19.

Meadows, D.H. and D. Wright (2008). *Thinking in Systems: A Primer*. Chelsea Green Publishing. ISBN: 9781603580557. URL: https://books.google.com/books?id=CpbLAgAAQBAJ.

Metagov (2023). *Metagov: A Laboratory for Digital Governance*. Accessed: September 27, 2024. URL: https://metagov.org/.

Nabben, Kelsie (2023). "Web3 as 'self-infrastructuring': The challenge is how". In: *Big Data & Society* 10.1, p. 20539517231159002. DOI: 10.1177/20539517231159002. eprint: https://doi.org/10.1177/20539517231159002. URL: https://doi.org/10.1177/20539517231159002.

Nabben, Kelsie and Michael Zargham (2022). "Permissionlessness". In: *Internet Policy Review* 11.2. DOI: 10.14763/2022.2.1656. URL: https://doi.org/10.14763/2022.2.1656.

NumFocus (2023a). *NumFocus Sponsored Projects*. Accessed: September 27, 2024. URL: https://numfocus.org/sponsored-projects.

— (2023b). *OpenSource Science Initiative*. Accessed: September 27, 2024. URL: https://www.opensource.science/.

Open Science Framework (2023). *Open Science Framework*. Accessed: September 27, 2024. URL: https://osf.io/.

OpenReview (2023). *OpenReview.* Accessed: September 27, 2024. URL: https://openreview.net/.

Ostrom, Elinor (2002). "Chapter 24 Common-pool resources and institutions: Toward a revised theory". In: *Agriculture and its External Linkages.* Vol. 2. Handbook of Agricultural Economics. Elsevier, pp. 1315–1339. DOI: https://doi.org/10.1016/S1574-0072(02)10006-5. URL: https://www.sciencedirect.com/science/article/pii/S1574007202100065.

Potts, Jason et al. (Apr. 2016). *A Journal is a Club: A New Economic Model for Scholarly Publishing.* SSRN. DOI: 10.2139/ssrn.2763975. URL: https://ssrn.com/abstract=2763975.

Public Library of Science (2023). *About PLOS.* Accessed: September 27, 2024. URL: https://plos.org/about/.

Rennie, Ellie and Jason Potts (Mar. 2024). *Contribution Systems: A New Theory of Value.* SSRN. DOI: 10.2139/ssrn.4754267. URL: https://ssrn.com/abstract=4754267.

Roessler, Christian and Philipp D. Koellinger (n.d.). *Incentives for Science – A Mechanism Design Framework.* White paper. Forthcoming.

Shevchenko, Nataliya (Dec. 2020). *An Introduction to Model-Based Systems Engineering (MBSE).* Carnegie Mellon University, Software Engineering Institute's Insights (blog). Accessed: 2024-Sep-18. URL: https://insights.sei.cmu.edu/blog/introduction-model-based-systems-engineering-mbse/.

Sisson, David et al. (July 2024). *Data Mesh Architecture: Interoperability, Co-Operation, and Co-Regulation.* SSRN. DOI: 10.2139/ssrn.4880709. URL: https://ssrn.com/abstract=4880709.

Suchman, Lucy (May 2000). "Organizing Alignment: A Case of Bridge-Building". In: *Organization* 7, pp. 311–327. DOI: 10.1177/135050840072007.

Suchman, Lucy et al. (1999). "Reconstructing Technologies as Social Practice". In: *American Behavioral Scientist* 43.3, pp. 392–408. DOI: 10.1177/00027649921955335. eprint: https://doi.org/10.1177/00027649921955335. URL: https://doi.org/10.1177/00027649921955335.

Voshmgir, Shermin and Michael Zargham (2020). *Foundations of Cryptoeconomic Systems.* English. WorkingPaper 1. Updated version. Austria: WU Vienna University of Economics and Business. DOI: 10.57938/10d8b646-3266-4ec9-b94d-810b528c40a9.

Zargham, Michael, Eric Alston, et al. (2023). *What Constitutes a Constitution?* DOI: 10.5281/zenodo.10609125. URL: https://doi.org/10.5281/zenodo.10609125.

Zargham, Michael and Ilan Ben-Meir (2023a). *Block by Block: Managing Complexity with Model-Based Systems Engineering.* https://blog.block.science/block-by-block-managing-complexity-with-model-based-systems-engineering/. Accessed: 2024-09-19.

— (Oct. 2023b). *Method for Functional Decomposition of Organizations and Their Environments.* SSRN. DOI: 10.2139/ssrn.4606672. URL: https://ssrn.com/abstract=4606672.

Zargham, Michael, Ilan Ben-Meir, and Kelsie Nabben (2023). *Knowledge Networks and the Politics of Protocols.* Accessed: September 27, 2024. URL: https://blog.block.science/knowledge-networks-and-the-politics-of-protocols/.

Zargham, Michael, Scott Moore, and Matt Stephenson (2021). *The difference between Public Goods Problems and Coordination Problems ... and whether it matters.* Published on Mirror.xyz, Accessed: September 27, 2024. URL: https://s.mirror.xyz/djByMntM2rQF4tqUISYS2MAO3oCfSWoOZSOpZjsYwaw.

Zargham, Michael and Kelsie Nabben (Nov. 2023). "Aligning "Decentralized Autonomous Organization" to Precedents in Cybernetics". In: *MIT Computational Law Report.* https://law.mit.edu/pub/dao-precedents-cybernetics.

Zargham, Michael and Orion Reed (2024). *Self-Infrastructuring Knowledge Networks.* Lecture, OpSci: The Open Science DAO, YouTube. URL: https://www.youtube.com/watch?v=w3gY_eYG9SY&list=PLmWm8ksQq4YIQwkJHFZyrlbMYIapCLpBy&index=6.

Zargham, Michael, Jessica Zartler, et al. (2023). *Disambiguating Autonomy: Ceding Control in favor of Coordination in Cybernetic Organizing.* Available online. URL: https://blog.block.science/disambiguating-autonomy/.

Appendices

Appendix A User Stories Schema

The DeSci Ecosystem consists of Protocol, Platform, and Applications layers with numerous distinct actors and entities operating within and across these layers (including but not limited to the DeSci Foundation, DeSci Labs, Organizations, and Participants).

As such, we consider the following schema when constructing user stories in order to appropriately locate the story within the Ecosystem. The location of the story can be mapped by indicating a [layer, group] pairing from **Figure 12**, and follows the format:

As an [actor] in [layer, group], I want to ____ so that I can meet [desire].

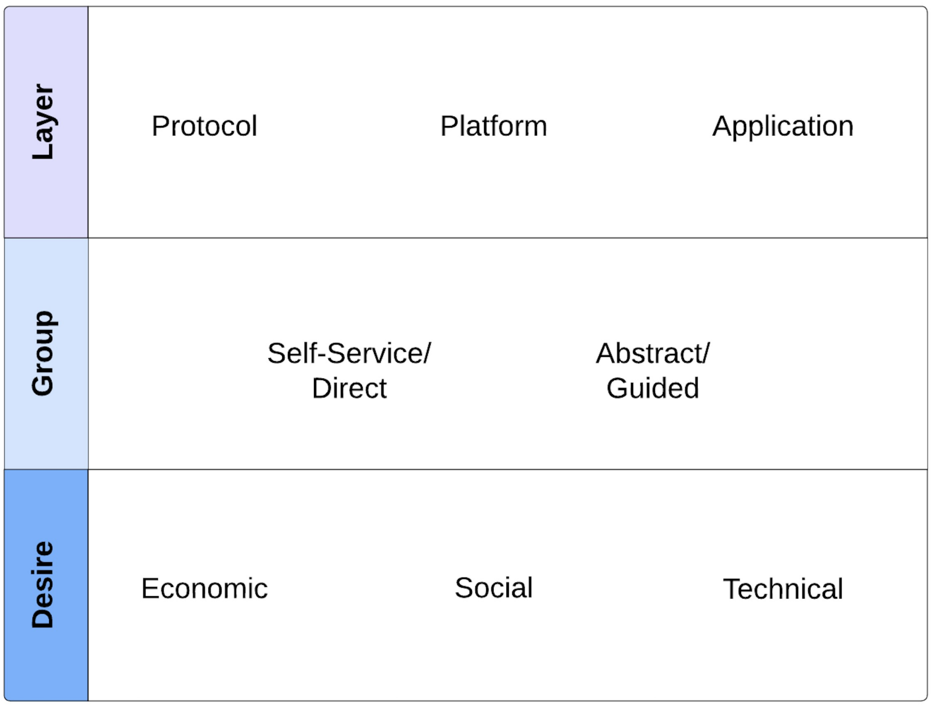

Figure 12: User stories schema

Explicit desires can be distilled into three broad needs:

1. **Economic** - Actors need to achieve economically viable and sustainable states.
2. **Social** - Actors want to exert influence, build reputation, develop expertise, and be in proximity (whether through production, consumption, or curation) to high-quality science.
3. **Technical** - Actors need the technical affordances required to achieve their economic and social desires.

Of note, the grouping of *self-service/directed* versus *abstract/guided* refers to an implicit desire, namely the desired user experience of the actor. The user experience attends to the question of how directly the actor wishes to interact with the Protocol versus rely on the Platform as an intermediary.

Actors in the self-service/direct group are motivated by an implicit desire to "not be denied the ability to do something" while actors in the abstract/guided group are motivated by an implicit desire to "not have to do something."

Examples

- **[actor: ARC, layer: application, group: abstract, desire: social]**
  As an ARC, I want the platform operator to abstract away my need to use a token so that I can serve non-crypto native users.

- **[actor: ARC, layer: application, group: self-service, desire: social]**
  As an ARC, I want to create a token so that I can reward actors for behavior that aligns with my purpose.

- **[actor: non-crypto native science consumer, layer: platform, group: abstract, desire: social]**
  As a non-crypto native science consumer, I want access to high-quality science free of charge so that I can learn.

- **[actor: crypto-informed staker, layer: platform, group: self-service]**
  As a crypto-informed staker, I want to stake tokens into specific ARCs so that:

  - I can earn passive income via interest. [desire: economic]
  - Steer rewards towards ARCs I care about. [desire: social]

- **[actor: ARC, layer: application, group: both]**
  As an ARC, I want my evaluators to be considered influential and valuable. [desire: social]
  As an ARC, I want other actors to be willing to pay me to execute evaluations. [desire: economic]
  As an ARC, I want to be able to pay contributors for their contributions to the maintenance and operations of my evaluators. [desire: economic]

Appendix B Example ARC Workflows

Introduction

This appendix develops the requirements for a feature built using the DeSci protocol and the Codex implementation. This feature enables a programmatic validation mechanism for the controlled expansion of a field of study through scientific research. It is in lieu of more traditional, centralized, academic institutions, and it implements a core purpose of DeSci's as illustrated in **Figure 13**.

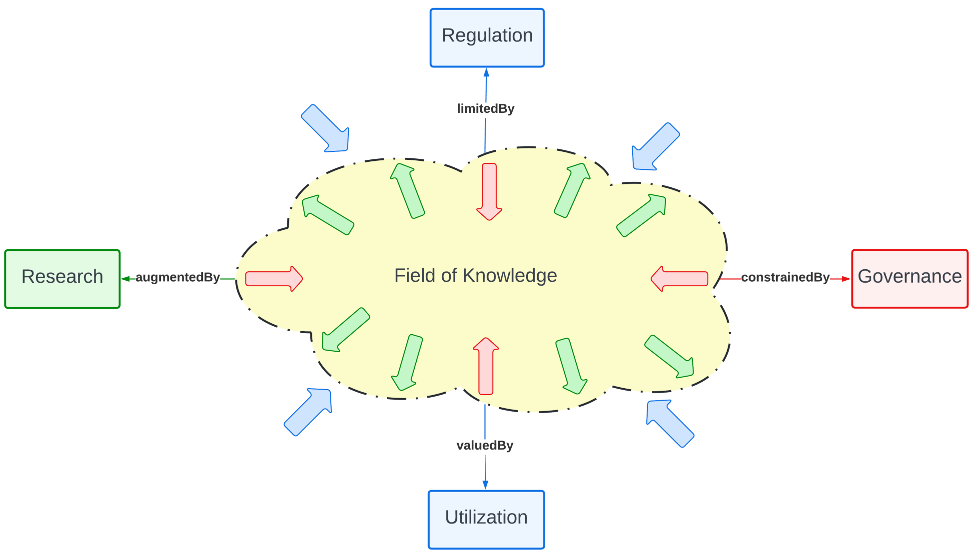

Figure 13: Environmental context of a curated field of knowledge.

The high-level requirements are specified in the user stories illustrated in **Figure 14** and repeated for legibility in the subsequent list. As shown in Figure **14**, the three outer (green or pink) user stories motivate their respective actors. The inner (yellow) user stories form the functional requirements of the validation mechanism being presented here. In the subsequent list, stories are tagged as being **Motivational** or **Functional**, respectively.

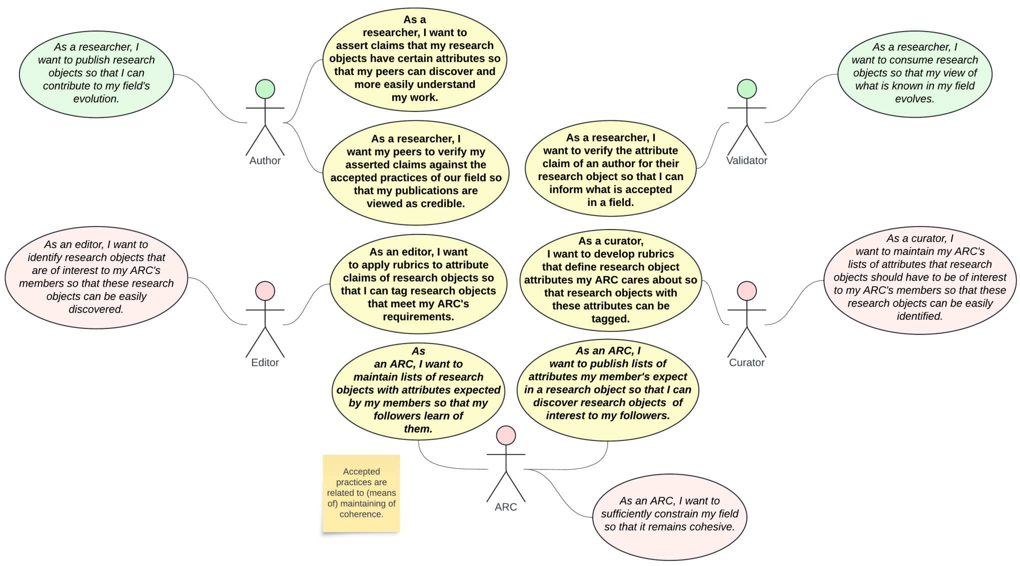

Figure 14: Codex validation of research objects.

- From the perspective of an author, an individual researcher who publishes their work:
  - **Motivational**: As a researcher, I want to publish research objects so that I can contribute to my field's evolution.
  - **Functional**: As a researcher, I want to assert claims that my research objects have attributes so that my peers can discover and more easily understand my work.
  - **Functional**: As a researcher, I want my peers to verify my asserted claims against the accepted practices of our field so that my publications are viewed as credible.
- From the perspective of a validator, an individual researcher who reads the published work of other researchers:
  - **Motivational**: As a researcher, I want to consume research objects so that my view of what is known in my field evolves.
  - **Functional**: As a researcher, I want to verify the attribute claim of an author for their research object so that I can inform what is accepted in a field.
- From the perspective of an ARC:
  - **Motivational**: As an ARC, I want to sufficiently constrain my field so that it remains cohesive.
  - **Functional**: As an ARC, I want to maintain lists of attributes my members expect in a research object so that I can discover research objects of interest to my followers.
  - **Functional**: As an ARC, I want to maintain lists of research objects with attributes expected by my members so that my followers learn of them.
- From the perspective of an editor:
  - **Motivational**: As an editor, I want to identify research objects that are of interest to my ARC's members so that these research objects can be easily discovered.
  - **Functional**: As an editor, I want to apply rubrics to attribute claims of research objects so that I can tag research objects that meet my ARC's requirements.

- From the perspective of a curator:
  - **Motivational**: As a curator, I want to maintain my ARC's lists of attributes that research objects should have to be of interest to my ARC's members so that these research objects can be easily identified.
  - **Functional**: As a curator, I want to develop rubrics that define research object attributes my ARC cares about so that research objects with these attributes can be tagged.

Figure 15 illustrates a functional ontological vocabulary that satisfies the requirements for the programmatic validation of research objects as listed in the above user stories.

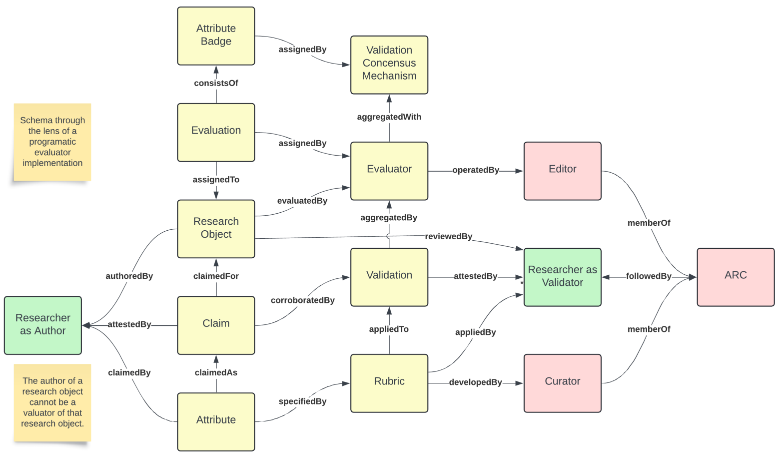

Figure 15: Schema of a programmatic evaluator through the lens implementation.

Figure 16 illustrates classes with which to implement programmatic evaluation of research objects using the Codex Protocol.

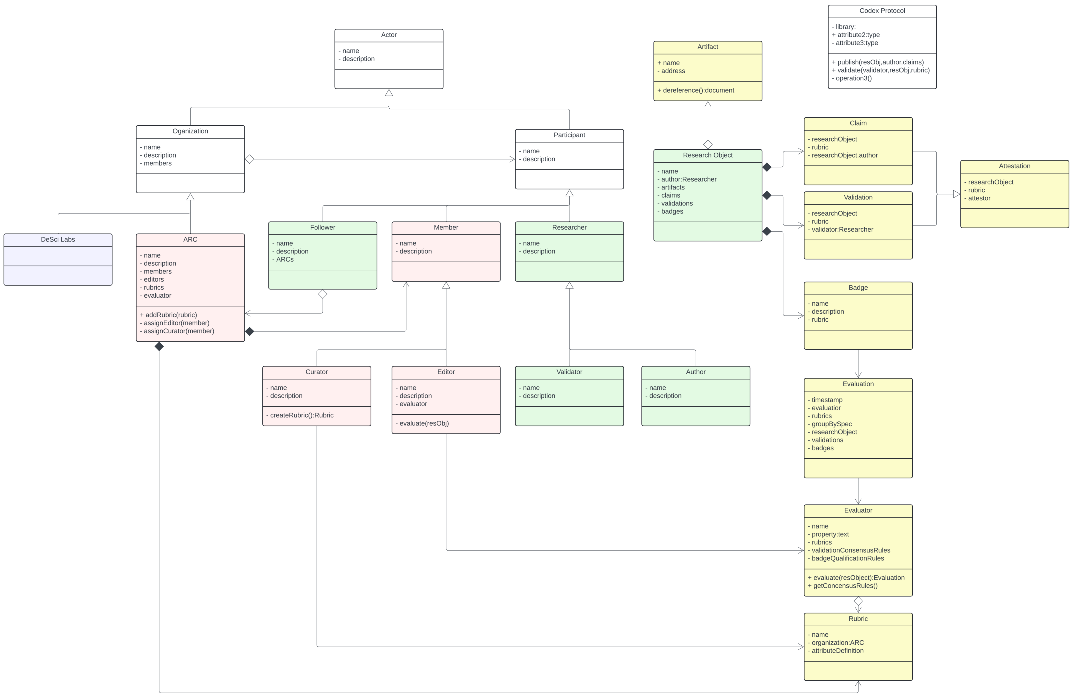

Figure 16: Codex implementation of a programmatic evaluator.

Workflow

- *Curators* develop **rubrics** for the attributes that the ARC of which they are a member.
- *ARCs* develop an **evaluator** that will attest to all **attributes** for which the *ARC's curators* have developed **rubrics** claimed by a **research object**.
- *Researchers* **author** an artifact of type **Research Object** as a collection of documents, *e.g.*, a scientific paper plus various corroborating data sets.
- The *researcher* **publishes** the **research object** as a **node** with DeSci's Codex Protocol and **claims** a set of **attributes** relevant to the **research object**.
  - Each **claimed attribute** must be defined by a **rubric**.
  - A **rubric** defines a single **attribute**—name, description, validation schema (JSON), and ARC that developed it.
- *Researchers* who are not the author can **validate attribute claims** of the **research object** based on a **rubric** that defines the **attribute**. Such *validators* can produce **validations** of a **claimed attribute** by a **research object**.
- *Editors* **process** the **research object** with **evaluators** for the *ARCs* of which they are a member, producing an **evaluation** of the **research object** from the viewpoint of the *ARC*.
- **Evaluators** work by aggregating all **validations** for a **research object** based on **rubrics** published by the *Editor's ARC*.

- An **evaluator's evaluation** of a **research object** consists of a collection of **badges**. Each **badge** represents the attestation of an *ARC* that a **research object** indeed has a **claimed attribute**.

This workflow can be applied to automated validation. As an example, an ARC may view publishing research objects that comply with a standard format is important (e.g., `https://schema.org/DigitalDocument`).

Sample Python Code

Sample python code for such a programmatic validator follows.

```
import requests
from bs4 import BeautifulSoup
import json
from jsonschema import validate as jsonschema_validate, ValidationError

class SchemaValidator:
    def __init__(self, schema_url):
        self.schema_url = schema_url
        self.schema = self._load_schema()

    def _load_schema(self):
        response = requests.get(self.schema_url)
        if response.status_code == 200:
            return response.json()
        else:
            raise Exception(f"Failed to load schema from {self.schema_url}")

    def _extract_json_ld(self, url):
        response = requests.get(url)
        if response.status_code == 200:
            soup = BeautifulSoup(response.text, ’html.parser’)
            json_ld_tag = soup.find(’script’, type=’application/ld+json’)
            if json_ld_tag:
                try:
                    return json.loads(json_ld_tag.string)
                except json.JSONDecodeError:
                    raise ValueError("Invalid JSON-LD format")
            else:
                raise ValueError("No JSON-LD metadata found on the page")
        else:
            raise Exception(f"Failed to fetch the URL: {url}")

    def validate(self, url):
        try:
            json_ld_data = self._extract_json_ld(url)
            jsonschema_validate(instance=json_ld_data, schema=self.schema)
            return True
        except (ValidationError, ValueError) as e:
            print(f"Validation failed: {e}")
            return False
        except Exception as e:
            print(f"An error occurred: {e}")
            return False

# Example usage for schema.org/DigitalDocument
schema_url = "https://schema.org/DigitalDocument"
validator = SchemaValidator(schema_url)
url = "https://example.com/sample-digital-document"
```

```
result = validator.validate(url)
print(f"Is the URL valid for DigitalDocument? {result}")
```

Note: This code was developed using the LLM-based code generation feature of the free version of the Cursor integrated development environment (IDE). To date, this code has not been run.

Appendix C Protocol Mathematical Specification

This appendix provides a detailed mathematical specification with notation for both states and state transitions, as well as considerations for learning from inconsistencies.

Notation

Name	**Symbol**	**Description**
Participant	$i \in I$	An individual level identity within the system; currently identified via Orcid, or an email and wallet address. Examples include, but are not limited to, an author, member, researcher, or referee.
Organization	$g \in G$	A group level identity within the system, consisting of one or more participants. Examples include, but are not limited to, a journal, community, institution, or informal group of academics.
Artifact	$d \in D$	A publishable research object, inclusive of collections of documents, data sets, software, and more.
Attribute	$a \in A$	A property of an artifact.
Evaluator	$e_{g,a}(d) \in \text{Bool}$	An evaluator is a classifier (whether computed by human or machine) which determines whether an artifact $d \in D$ has attribute $a \in A$ according to organization $g \in G$. The set of all such classifiers is $E = \{e_{g,a} : D \to \text{Bool}\}$ that have been offered by groups g.
Claim	$\hat{y}^i_{d,g,a} \in \text{Bool}$	An unverified attestation by a participant $i \in I$ that says artifact $d \in D$ has attribute $a \in A$ as defined by organization $g \in G$. This means the participant i is claiming that they believe $E_{g,a}(d) = \hat{y}^i_{d,g,a}$. The set of all such claims is C.
Evaluation	$y^*_{d,g,a} = E_{g,a}(d) \in \text{Bool}$	A verified attestation produced by applying an evaluator to an artifact. The set of all such evaluations is Y.
Authorship	$r(d) = i \in I$	Any artifact $d \in D$ must have been published by a participant $i \in I$. $r : D \to I$ is the map which returns the author of an artifact.
Membership	$i \in M_g \subset I$	Participant $i \in I$ is a member of organization $g \in G$ if $i \in M_g$, which can be thought of as the membership set for g.

Table 8: Mathematical notation

Dynamics

There are two types of actors in the system as described so far: Participants and Organizations.

The system can be viewed as a discrete event system where specific actions induce deterministic state changes. However, the specific choices of actions taken by actors are unknown and non-deterministic.

The state of this system at any time is denoted as $x \in X$ where X denotes a data structure representing the entire network:

$$X = I \times D \times G \times A \times E \times C \times Y$$

That is to say, a snapshot of the system is described by:

- The set of all Participants, I

- The set of all Artifacts, D
  - including the link back to the Participant who controls it
- The set of all Organizations, G
  - including the reference back to the Participants who are members
- The set of all Attributes, A
  - including the reference to all Evaluators which evaluate that attribute
- The set of all Evaluators, E
  - including the reference back to the Organization that operates it
- The set of all Claims, C
  - including reference back to the Participant making the Claim
  - including link back to the Evaluator being referenced
  - including link to the Artifact the Claim is about
- The set of all Evaluations, Y
  - including the reference back to the Artifact evaluated
  - including the reference back to the Evaluator that produced it

Dynamics for the discrete event model are any actions which induce a change:

$$x^+ = f_j(x, u_j) \in X$$

where j denotes a specific mechanism, $u_j \in U_j(x)$ denotes a valid message or input according to that mechanism, and f_j is the law of motion or action principle which describes the state transition for action u_j given the current state of the system $x \in X$.

Note that actions available are state-dependent as encoded by $U_j(x)$. For example, a Participant i may have specific authority over Artifact d for which they are the author $r(d) = i$, which other participants do not have. Similarly, an Organization g has authority over an Evaluator $e_{a,g}$ but not over another Organization's Evaluator $e_{a,g'}$ for the same Attribute a.

The Protocol allows:

- Participants to publish Artifacts
- Participants to make Claims
- Participants to form Organizations
- Organizations to operate Evaluators
- Evaluators to make Evaluations

For example, if a Participant i were to publish an Artifact d, the updated state space would be:

$$X^+ = I \times D^+ \times G \times A \times E \times C \times Y$$

where $D^+ = D \cup \{d\}$ is the updated set of Artifacts.

If a group of Participants were to form a new group identity, then:

$$X^+ = I \times D \times G^+ \times A \times E \times C \times Y$$

where $G^+ = G \cup \{g\}$ and $M_g = (i_1, \ldots, i_n) \subset I$ is the group of Participants forming the group.

A Participant i may make a Claim that an Artifact d has Attribute a according to the definition provided by Organization g:

$$X^+ = I \times D \times G \times A \times E \times C^+ \times Y$$

where $C^+ = C \cup \{c\}$, with c being the claim $\hat{y}^i_{d,g,a}$ denoting the Claim by i that d has Attribute a according to Organization g. Note that this is distinct from that Claim being checked. Evaluating the Claim by computing $y^*_{d,g,a} = e_{a,g}(d)$ is in the purview of Organization g.

Learning through Inconsistencies

An attestation $c \in C$ might claim that Artifact d has Attribute a as defined by Organization g, but when Evaluator $e_{a,g}(\cdot)$ is applied, it returns false. Such a state of inconsistency is possible to achieve in this state machine and is relevant to preserve in the history even though Evaluation $e_{a,g}(d)$ is canonical. Such an inconsistency may be viewed as a learning opportunity for both Participant i who made Claim c and Organization g who is responsible for maintaining $e_{a,g}$.

Appendix D Micro Model Details

The Micro Model attends to the question of how the token relates to specific activities within the network, with attention to implementation concerns.

Suppose there exists an *ERC20 Compatible* Native Token 'DeSci' whose token balances are stored in a struct mapping on-chain addresses to integer balances.

Furthermore, suppose there exists a Staking Contract such that a 'DeSci' token holder can voluntarily "lock up" tokens. Tokens that are staked are stored in a struct that specifies both the staker, who can still unstake, and the staked-on address. The staked-on addresses do not control the tokens staked on them, but they do unlock affordances as a function of their total stake.

A Participant controlling an identity in the DeSci Network may bind one or more on-chain addresses to their account by proving control via signature. (The signature need not be on chain, as it is simply proof of control.) Participants are not required to bind on-chain addresses unless they wish to access features that require on-chain identities.

Pursuant to minimizing frictions on raw intellectual production, Participants do not need to create on-chain identities or staked tokens in order to publish Artifacts or make Claims. Making sense of the firehose of Artifacts and Claims is the purpose of Organizations and Evaluators, and there is more friction around these enrichment processes.

Form Organizations

One or more Participants (who have on-chain addresses bound to their DeSci identity) may form an Organization, which is *functionally* equivalent to creating a SAFE multi-sig identity. An Organization is, by default, "inactive."

Organizations become "active" if their total stake exceeds the Organization registration threshold, α.

Whether an Organization is active need not be stored on-chain; it suffices that the Organization's identity in the DeSci Network is bound to the on-chain address of the Organization (multi-sig). Given *read* access to the on-chain state, a protocol-compliant DeSci client will be able to check whether an Organization is active when attempting to take an action restricted to active Organizations.

Publish an Evaluator

Publishing Evaluators is an affordance specifically for active Organizations.

From an on-chain perspective, the Evaluator's identity is a child of the Organization's identity. For on-chain identity management, one option would be to integrate with ENS so that Organizations have named domains, and their Evaluators can be treated as subdomains: for example, 'myevaluator.myorganization.eth'.

On the data network, an Evaluator is a set of requirements:

- It explicitly names an Attribute, which may or may not already exist within the network, but must be locally unique to the publishing Organization.
- It lays out a definition of that Attribute, which serves as a statement of "what it means for an Artifact to have an Evaluation of *True* from this Evaluator."
- It describes what evaluation process is used to assess the above. These processes may vary over time and across organizations, but must be made available at the time of publication.

The Evaluator is in the domain of control of the Organization that published it. Optionally, built into SAFE fork, affordances for a multi-sig Evaluator could update approval on-chain before changes are treated as accepted.

It is recommended to "regularize" the number of Evaluators, such that Evaluators are added only if they are producing unique and useful signals. This could be achieved with an additional staking requirement of β, such that an Evaluator is considered "inactive" if it does not have at least β additional 'DeSci' tokens staked to activate it.

Publish an Evaluation

It is also recommended to "regularize" the number of Evaluations, such that Evaluations are published only if they are expected to produce unique and useful signals.

One way to achieve this is by limiting every active Evaluator's baseline evaluation capacity to an allowance of γ evaluations per period, Δ.

Staking on an Evaluator an amount of tokens in excess of β could increase the number of Evaluations that the particular Evaluator is allowed to publish to the network per period Δ. This may be described as equivalent to staking more tokens to receive more API calls, where the act of publishing an Evaluation to the network consumes an API call.

Summary

The Micro Model addresses the requirement that the 'DeSci' token be intrinsic to the DeSci Network, contributing directly to its Animating Purpose by making key activities spam resistant, and ensuring actors engaged in curation have strong persistent identities around which they are incentivized to build reputation, without that reputation serving as a substitute for demonstrably useful curative labor. Furthermore, as the network grows in per-period productive labor (monthly active Organizations, monthly active Evaluators, and Evaluations per month), the total number of tokens locked up in service of that labor rises.

Note that the requirements for token staking are cumulative. For example, Participants who form an Organization must have both on-chain identities and DeSci Network identities. To register an Evaluator as Active, an Organization must meet both staking thresholds β (on the Evaluator) and α (on the Organization).

Appendix E Macro Model Details

Consider a simplified model where tokens serve the function of "machinery" in the curation and sense-making processes facilitated by the DeSci network. This model is presented from a macro frame of reference with micro foundations based on the premise that tokens will need to be staked in order to engage in curative activities.

Term	Mathematical Notation	Code Representation
DeSci Network State	G_t	`network[t]`
Max Token Supply	$S_{\max}$	`Smax`
Current Token Supply	x_t	`x[t]`
Tokens in Utilization	u_t	`u[t]`
Tokens (Un)Staked	Δu_t	`du[t]`
Tokens held in Foundation	$h_f(t)$	`treasury[t]`
Tokens Allocated by Foundation	$a_f(t)$	`allocated[t]`
Tokens in Vesting Contract	v_t	`unvested[t]`
Tokens held by Others	$h_o(t)$	`other[t]`

Table 9: Macro model terms and corresponding code representations

Assert:

$$0 < x_0 \leq S_{\max}$$

Require:

$$0 < x_t \leq S_{\max} \quad \forall t > 0$$

Parametric Policies

Term	Mathematical Notation	Code Representation
Vesting Schedule	Δv_t	`vesting_tokens`
Wear Coefficient	ω	`wear_rate`
Recovery Coefficient	ρ	`recovery_rate`
ARC Registration	α	`arc_stake`
Evaluator Registration	β	`evaluator_stake`
Evaluation Capacity per β tokens staked per Period	γ	`evaluation_rate`

Table 10: Parametric policies for the macro model

Deriving Summary Statistics from the Network

Term	Mathematical Notation	Code Representation
Number of ARCs	n	`arc_count`
Number of Evaluators	m	`eval_count`
Mean Evaluations per Evaluator	e	`evaluation_stake`

Table 11: Summary statistics derived from the network

The micro model guarantees that:

$$u_t \geq n\alpha + m\beta$$

but we expect scaling like:

$$u_t \approx n\alpha + e\gamma$$

as evaluation capacity grows beyond the minimum allowance.

Examining the Token Supply

Let's examine:

$$x_{t+1} = h_f(t+1) + h_o(t+1) + u_{t+1} + v_{t+1}$$

by breaking it down into its component terms.

Foundation

$$h_f(t+1) = h_f(t) + (S_{\max} - x_t)\rho - a_f(t)$$

Vesting

$$v_{t+1} = v_t - \Delta v_t$$

Utilization

$$u_{t+1} = u_t - u_t\omega + \Delta u_t$$

Other

$$h_o(t+1) = h_o(t) - \Delta u_t + a_f(t) + \Delta v_t$$

Now we sum across the terms and cancel the conserved flows to get:

$$x_{t+1} = x_t + (S_{\max} - x_t)\rho - u_t\omega$$

Conclusion

Using the above equation, we can conclude that for utilization $u_t \geq 0$, and parameters $\rho, \omega \in [0, 1)$ and current supply $0 < x_t \leq S_{\max}$, then $0 < x_{t+1} \leq S_{\max}$.

As we assumed $0 < x_0 < S_{\max}$, we can conclude via induction:

$$0 < x_t \leq S_{\max} \quad \forall t > 0$$

This satisfies our requirement that a max supply be respected for all possible scenarios.

Additional analysis would show that for $u_t > 0 \forall t$, and $\rho, \omega \in (0, 1)$, then $(S_{\max} - x_t)\rho > 0 \forall t$, which ensures future flow into the treasury for an active network. Finally, we can see that the supply will contract if:

$$(S_{\max} - x_t)\rho < u_t\omega$$

at any time t.

Appendix F Subsidy Program Details

Standard Template

As discussed above, the goal of subsidy programs is to encourage and incentivize participants in the DeSci ecosystem to undertake specific sets of actions that are necessary for the project's overall success, as it is picking up speed. While these subsidy programs can and should be tailored to specific goals, any actualizable subsidy must be definable according to a general template that specifies:

- Tokens allocated to the program
- Target measurements on which the rewards are based
- Rewards as a function of measurements
- Eligibility criteria
- Relevant stakeholders
- Start and end dates for the subsidy program

The subsidy programs described in the Recommended Design section of this report are discussed in greater detail below.

Connections to the Micro Model

Figure 17 illustrates in detail how the three candidate bootstrapping summaries proposed in the body of this report could function, and how they connect back to the Micro Model.

The top "track" illustrates a program designed to incentivize actors to join the network – creating accounts, connecting ORCID IDs and crypto addresses, and publishing artifacts – and is target at the "demand side" of the ecosystem. A participant begins by connecting their ORCID ID to their account, which ensures that they are meet the requirements for ORCID eligibility. The participant then publishes an artifact, and claims that it is eligible for the subsidy – encouraging participants to also engage with the system's claim affordances. The DeSci Foundation will run its "eligibility" evaluator on this artifact to determine the validity of the eligibility claim; if the artifact receives a positive eligibility evaluation, it "counts" for determining that participant's number of eligible artifacts per period. Rewards can then be paid out of a program budget (set at x tokens per month), proportionally to the square of the count of publications per period. Provided that the participant has a public crypto key, their rewards will then be paid out; it is possible to complete this flow without connecting a crypto address, but participants cannot receive rewards without an onchain ID.

The middle "track" illustrates a program designed to get ARCs and Evaluators set up on the platform, to seed the network with diverse and useful evaluators – it targets the "supply side" of the ecosystem. (It may be worthwhile to develop a list of desired evaluators, to help guide builders decide what to propose). This flow begins with individual organization members creating accounts, and forming an organization. The organization then plans an evaluator proposal, and "applies" to launch an ARC. If the organization is accepted as an ARC, it is "instantiated" on the chain, and activated via staked tokens **provided by this subsidy program's budget** – perhaps on a first-come, first-serve basis. The ARC uses this subsidized stake to operate its evaluators, making it possible for the DeSci Foundation to compare proposals based on outcomes, and issue or modify ARC/Evaluator guidelines. This program is primarily manual, and is meant to onboard the ARCs and evaluators that the ecosystem will require in order to function, while distributing tokens to operators so that they can stake them as needed; in other words, it is not intended to produce upside, but to defray initial operating costs.

The third "track" illustrates a program designed to reward ARCs for operating useful evaluators, and relies on the "vetting" that occurs throughout the other two tracks to reduce gamabililty – its purpose is to help *balance* the "supply" and "demand" sides of the ecosystem. Once an evaluator is live, participants can "claim" attributes for artifacts against the evaluator. The evaluator then evaluates, producing an evaluation. After filtering to artifacts which pass the eligibility check in stream 1, it can then be determined how many evaluations (of claims) for eligible artifacts the evaluator has performed in the relevant period, and rewards can then be paid out of a defined program budget (tokens per month) proportional to the evaluations (for claims) issued per period performed by the ARC operating the evaluator in question (provided that the ARC has an onchain address).

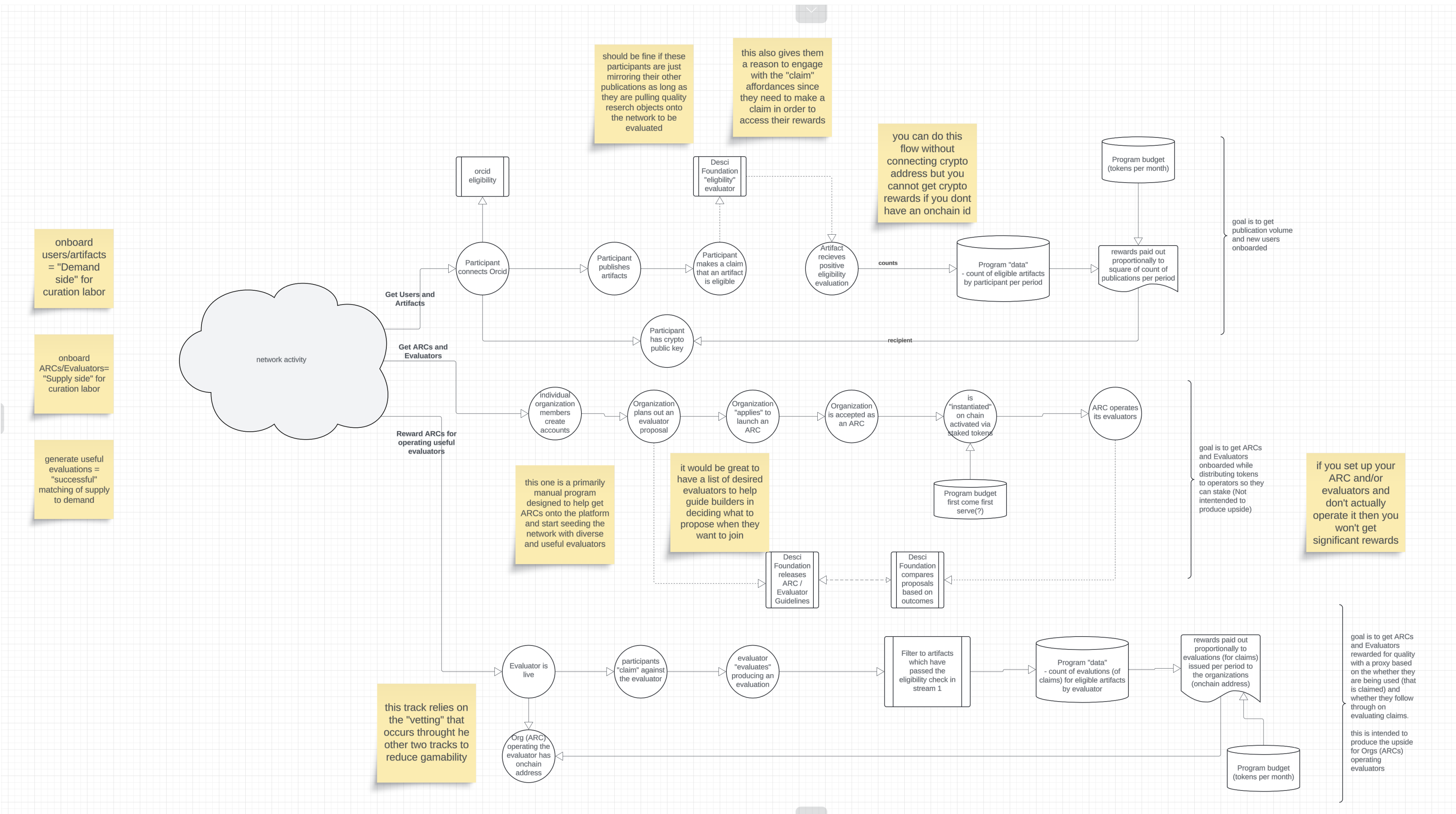

Figure 17: Subsidy programs – connection to micro model

Connections to the Macro Model

Figure 18 illustrates how the proposed subsidy programs connect to the Macro Model, and situates the subsidy reward pools in the stock and flow model detailed in the body of this report.

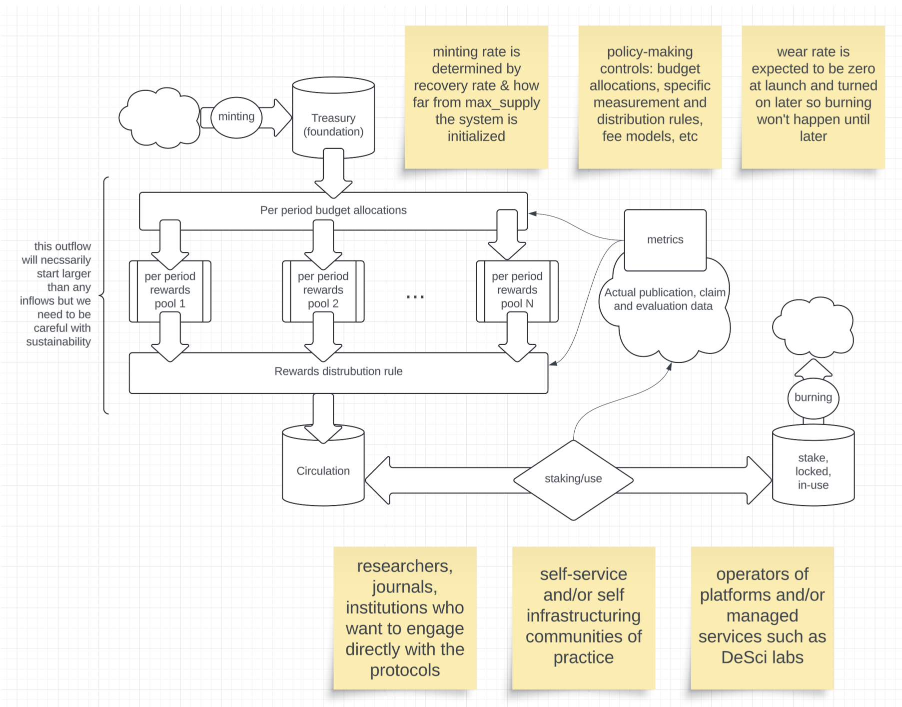

Figure 18: Subsidy programs – connections to macro model

As coins are minted into the Foundation's Treasury, the Foundation will decide to allocate some fraction of its resources to incentivizing desired behaviors through subsidy programs – a per-period allocation from the Treasury to the subsidy programs. The individual subsidy programs are then budgeted out of this overall allocation, and rules are set for the distribution of rewards by way of the subsidy programs. Those rewards are then distributed, and the issued rewards enter circulation (that is, staking and/or use). The staking and use of tokens must be instrumented in such a way as to provide information in the form of metrics back to the allocations and distribution rule mechanisms, to evaluate the effectiveness of the subsidy programs and inform decisions about allocations and distribution in subsequent periods. Staking produces a locked pool, from which tokens may be burned; the minting rate and burn rate are related, as it is desirable to keep an overall supply sufficient to cover current use, while remaining sensitive to both inflation and deflation.

Ex. The Foundation might choose to allocate 30 percent of the initial subsidy pool to the Publication Volume program (for 12 months, at a rate of 2.5 percent of the initial pool per month), and 10 percent of the initial pool to the Cost Offsetting program. The Evaluator Usefulness program would then receive the remaining 60 percent of the initial subsidy pool (at a rate of 5 percent per month, for 12 months).

The specific percentages and rates can be decided by the Foundation, with the initial token allocation numbers likely to influence the conversation. It is worth noting that the first two subsidy programs provide the *baseline* for the third, which is the main provider of *upside* to participants (and is thus the most sensitive to gaming). It may make sense to make the third program strictly data-driven, after the first two (somewhat discretionary) qualifying gates, and/or to try to develop a relatively-uncomplicated metric that rewards both **getting** more claims and **evaluating** a larger fraction of claims received.

Operations vs. Governance

Although the subsidy programs described here originate on the protocol layer, in practice they must be administrated at the platform layer. This can be accomplished by making sure that all items in the Subsidy Program Template (the appropriation of funds, definition of mandates and metrics, identification of stakeholders, etc.) are **defined** with the authority of the Foundation. For practical purposes, the programs can then be **administrated** on the platform layer.