Decentralized Health Intelligence Network (DHIN)

Abraham Nash
University of Oxford
abraham.nash@cs.ox.ac.uk

Abstract: *Decentralized Health Intelligence Network* (*DHIN*) is a theoretical framework addressing significant challenges of health data sovereignty and AI utilization in healthcare caused by data fragmentation across providers and institutions. It establishes a sovereign architecture for healthcare provision as a prerequisite to a sovereign health network, then facilitates effective AI utilization by overcoming barriers to accessing diverse medical data sources. This comprehensive framework leverages: 1) self-sovereign identity architecture coupled with a personal health record (PHR) as a prerequisite for health data sovereignty; 2) a scalable federated learning (FL) protocol implemented on a public blockchain for decentralized AI training in healthcare, where health data remains with participants and only model parameter updates are shared; and 3) a scalable, trustless rewards mechanism to incentivize participation and ensure fair reward distribution. This framework ensures that no entity can prevent or control access to training on health data offered by participants or determine financial benefits, as these processes operate on a public blockchain with an immutable record and without a third party. It supports effective AI training in healthcare, allowing patients to maintain control over their health data, benefit financially, and contribute to a decentralized, scalable ecosystem that leverages collective AI to develop beneficial healthcare algorithms. Patients receive rewards into their digital wallets as an incentive to opt-in to the FL protocol, with a long-term roadmap to funding decentralized insurance solutions. This approach introduces a novel, self-financed healthcare model that adapts to individual needs, complements existing systems, and redefines universal coverage. It highlights the potential to transform healthcare data management and AI utilization while empowering patients.

1. Introduction

Healthcare is undergoing a profound digital transformation, propelled by the widespread adoption of electronic health records (EHRs) and the growing potential of artificial intelligence (AI) in medical diagnostics and treatment [1], [2]. However, this shift faces significant challenges in data management, privacy, and interoperability. Fragmentation and siloing of patient health data across various providers and institutions not only undermine data sovereignty but also hinder the effective utilization of AI in healthcare.

To address these challenges, this paper proposes *Decentralized Health Intelligence Network* (*DHIN*), an innovative framework that builds upon the principles of *Decentralized Intelligence Network* (*DIN*) [3]. *DHIN* aims to redefine healthcare data management by leveraging blockchain technology, federated learning, and cryptographic techniques.

This comprehensive approach encompasses:

1. A sovereign architecture for healthcare provision, utilizing self-sovereign identity and personal health records.
2. A scalable federated learning protocol implemented on a public blockchain, enabling decentralized AI training while preserving data privacy.
3. A trustless rewards mechanism to incentivize participation and ensure fair compensation for data contributions.

By addressing the critical issues of data sovereignty, AI utilization, and patient empowerment, *DHIN* offers a promising solution to the current limitations in healthcare data management. This framework not only supports effective AI training but also introduces a novel, sustainably financed healthcare model that adapts to individual health needs.

The structure of this paper is as follows: **Section 2** presents a detailed problem statement, outlining the current challenges in healthcare data management. **Section 3** establishes the specific requirements of the *DHIN* framework. **Section 4** provides an overview of the systems architecture, detailing the root, real-world, and intelligence layers. **Section 5** delves into the methodology, explaining self-sovereign server technology, personal health records, and the decentralized federated learning protocol. **Section 6** discusses the threat model, addressing potential security concerns. **Section 7** explores the concept of a decentralized insurance solution. Finally, **Section 8** onwards concludes the paper, suggesting future research directions and discussing the broader implications of this work.

2. Problem Statement

The current healthcare landscape is marked by significant fragmentation and siloing of patient health data across various providers and institutions. This fragmentation undermines data sovereignty and hinders the effective utilization of artificial intelligence (AI) in healthcare. The ramifications are severe:

1. Patient identification errors contribute to approximately 195,000 deaths per year in the US alone [4].
2. Inefficient and error-prone data collection processes create fatigue for patients and healthcare providers [4].
3. Electronic Health Records (EHR) are often stored in siloed databases, becoming targets for security breaches [5], [6].
4. Lack of EHR interoperability impedes AI practitioners from developing and implementing diagnostic tools and decision-support systems [7].
5. Patients are excluded from the ecosystem of valued health data exchange, lacking ownership and control over their personal health information [8], [9].

These issues collectively obstruct the delivery of efficient, safe, and high-quality healthcare, hinder medical research advancement, and impede the development of innovative diagnostic tools and treatment pathways.

3. Requirements

***DHIN* adapts the core tenets of *DIN* to the healthcare domain, specifically addressing the requirements that:**

- No authority can resume access and management controls over Participants' data.
- Only Participants can decide who accesses their data for federated learning.
- No third-party broker determines reward allocations for Participants' contributions.

DIN is a cross-sector framework, originally designed to preserve participant sovereignty over data while fostering collaborative AI efforts across various domains [3]. In the context of *DHIN*, this generic framework is repurposed to address the unique challenges and opportunities within healthcare.

This healthcare-specific adaptation ensures that no single authority controls the FL process, preserving participant sovereignty over their health data while fostering collaborative AI efforts. Crucially, the *DHIN* framework prevents data other than model updates from needing to leave the Personal Data Stores (PDS), maintaining user privacy and control in a healthcare context.

Decentralized Health Intelligence Network (*DHIN*), as a healthcare application use case of *DIN*, acknowledges the ongoing existence of institutional silos in traditional healthcare. Designed to facilitate a transition towards decentralized, sovereign data stores controlled by individuals, *DHIN* offers flexible integration with Electronic Health Records (EHR) and institutional data systems, supporting a transition to decentralized, sovereign data stores. By allowing broad participation in the federated learning (FL) protocol, the *DHIN* protocol complements new avenues for access to scalable data for AI engineering in a decentralized fashion, specifically within the healthcare sector.

While enhancing data access, privacy, security, and monetization, *DHIN* acknowledges that institutional silos and centralized learning are likely to continue, enabling a truly global reach and inclusivity, fostering a decentralized and sovereign AI development landscape within the healthcare sector. The self-sovereign identity layer may address ongoing complexities such as healthcare custodianship and emergency care, allowing *DHIN* to either operate independently or integrate with existing systems; however, this discussion extends beyond the scope of this paper. This paper illustrates *DIN*'s adaptability to healthcare data sovereignty and AI challenges while promoting a fair and secure environment for AI development that respects individual rights and encourages standardization of data formats.

4. Systems Architecture: An Overview

This framework supports robust and ethical AI training in healthcare by allowing participants to maintain sovereignty over their health data while simultaneously benefiting financially and contributing to a decentralized, scalable ecosystem. This ecosystem harnesses collective AI capabilities to advance healthcare algorithms, demonstrating new possibilities in medical research and patient care.

To meet this frameworks requirements and overcome the challenges specific to healthcare, *DHIN* leverages three key components:

1. **Root Layer:** This foundational layer provides an overview of self-sovereign server technology (SSST), which underpins the system's ability to maintain individual data sovereignty in healthcare ecosystems.
2. **Real-World Layer:** This layer focuses on practical healthcare workflows—such as prescriptions, referrals, and investigations—conducted in a sovereign manner within the *Decentralized Healthcare Intelligence Network* (*DHIN*). It integrates principles from existing frameworks to ensure secure, patient-controlled exchanges. Additionally, the layer emphasizes the flow of information through the sovereign integration of Medical AI Devices, which interact with patient-owned health records. This enhances healthcare delivery by enabling efficient, transparent, and patient-centric care within the *DHIN* framework.
3. **Intelligence Layer:** This layer outlines the integration of the *DIN* protocol for a scalable and secure trustless federated learning process. It also explains the specific utility of the reward mechanism in funding decentralized healthcare insurance solutions.

By leveraging these decentralized mechanisms across these three layers, the *DHIN* framework not only addresses the current challenges in health data management but also catalyzes the adoption of scalable, sovereign data solutions. These solutions uphold individual rights, foster technological advancements, and pave the way for new use cases in healthcare. Ultimately, *DHIN* supports the ongoing transition towards decentralized data ownership, promising more equitable, efficient, and innovative healthcare ecosystems for the future.

Key stakeholders previously specified in the *Decentralized Intelligence Network* (DIN) ecosystem include [3]:

- **Participants:** Individuals who own and control their health data stores (i.e., patient's). They contribute data to the federated learning process while maintaining privacy and benefiting from collaborative AI training.
- **Model Owners:** Entities such as pharmaceutical companies, research institutions, or AI developers who utilize the FL protocols to enhance their models with decentralized health data, without compromising individual data sovereignty.
- **Evaluators:** Network-staked entities responsible for decentralized auditing. They ensure transparency and fairness in evaluating participant contributions and overseeing the distribution of rewards.

Building upon these three stakeholders introduced in the *DIN*, *DHIN* introduces two additional key ecosystem stakeholders:

- **Healthcare Professional:** Healthcare professionals (i.e., a physician) who use self-sovereign server technologies (SSST) to securely interact with the network, verify patient relationships, and access health records while maintaining high standards of privacy and security.
- **Medical AI Devices:** AI-powered tools and devices with their own public/private key identifiers, capable of interacting with patient data to support diagnosis, treatment, and healthcare delivery within the decentralized ecosystem.

This expanded set of stakeholders enhances *DHIN's* capability to address healthcare-specific challenges and opportunities, fostering a more comprehensive and efficient decentralized health system.

DHIN employs smart contracts (SC) to manage crucial processes, including for coordinating and rewarding AI training, and for enabling Evaluators to assess participant contributions in a secure, novel proof-of-stake ecosystem. By leveraging these decentralized mechanisms, the framework ensures scalable, sovereign data solutions that respect individual rights and drive technological advancements.

Participants receive cryptographic micropayments into their digital wallets as an incentive to opt into the system. This not only provides immediate value to individuals for their data contributions but also aligns with a long-term roadmap aimed at funding decentralized insurance solutions. This innovative approach could potentially transform the healthcare financing landscape.

By leveraging these decentralized mechanisms, the *DHIN* framework not only addresses the current challenges in health data management but also catalyzes the adoption of scalable, sovereign data solutions. These solutions uphold individual rights, foster technological advancements, and pave the way for new use cases in healthcare. Ultimately, *DHIN* supports the ongoing transition towards decentralized data ownership, promising a more equitable, efficient, and innovative healthcare ecosystem for the future.

5 Methodology

5.1 Root layer

5.1.1 Self-Sovereign Server Technology

In the root layer of the *DHIN* framework, we encounter the foundational technology that enables true data sovereignty: Self-Sovereign Server Technology (SSST). This innovative approach re-defines how individuals, particularly patients, manage their digital identities and health data in a decentralized ecosystem.

At its core, SSST functions as an identity container, combining a user-friendly mobile interface with a continuously connected server. This setup allows patients to securely store and control their health attributes (such as prescription details or lab results), access policies, and transaction receipts [4]. Imagine having a digital vault on your smartphone, where you not only keep your most sensitive health information but also dictate who can access it and when.

Blockchain technology is well-suited for decentralized identity (DID) that does not depend on a centralized root of trust [4]. DID extends blockchain methods to enable a lifelong practical and reliable identifier (e.g., a public/private cryptographic key pair) and attributes linked to that identifier under the self-sovereign control of the individual person [4]. A number of DID systems are emerging, and this technology under the total control of the physician (MD) and the patient can leverage DID to allow for a prescription or equivalent regulated transaction [10].

A key component of SSST is the use of decentralized identifiers (DIDs) [10]. With DID-based self-sovereign identity:

- Patients can selectively disclose only necessary health information to providers.
- Providers can authenticate themselves and access patient records with patient permission.
- Health data can be linked to patient DIDs while remaining under patient control.
- Interactions between patients and providers can be securely logged using DIDs.

The white paper "Powering the Physician-Patient Relationship with HIE of One Blockchain Health IT," presented by Adrian Gropper, MD, in 2016, provides an excellent overview of the real-world benefits SSST can offer to both individuals seeking healthcare and the professionals providing it [10]. The major benefit of self-sovereign support technology in this context is the re-decentralization of the trusted relationship between the physician and the patient. The full value of the medical consultation is now available to the two principal parties, with each managing their own policies to provide access to shared resources such as a physician directory or a pharmacy. Additionally, this approach enhances security through diversity, as patients and physician can adopt different self-sovereign technologies to support their respective security, privacy, and economic interests [10].

The beauty of SSST lies in its versatility and security. Each participant in the healthcare ecosystem – be it a patient or a physician – interacts through their own identity container. These containers can be linked to a public blockchain ledger via unique public/private key pairs, ensuring both accessibility and security [4], [10].

For physicians, SSST offers a secure method of authentication. Picture a physician using their smartphone to log into the system, their professional credentials verified through a trusted reputation mechanism – perhaps a digital version of their medical license [10]. This process employs various security measures, including digital signatures, encryption, and even biometric data like fingerprints, all managed through an intuitive user interface [10].

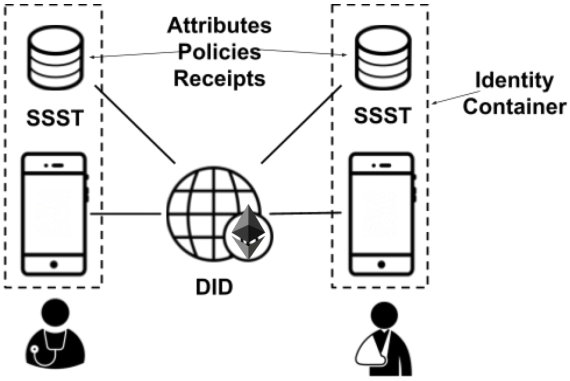

Figure 1. *A transaction between a physician and a patient. Adapted from HIE of One* [10].

Patients, on the other hand, gain unprecedented control over their health records. They can choose which open-source servers to use for storing their personal health information and manage access permissions with ease. It's akin to having a personal health safe, where the patient holds the only key and can grant or revoke access as they see fit.

This system facilitates seamless, secure interactions between patients and physicians. When a physician needs to access a patient's records, they use their sovereign identity to establish a verified relationship with the patient, request access to relevant health data (e.g., via API), and provide care based on comprehensive, up-to-date information [10]. These interactions can be securely logged using DIDs, creating an auditable trail of data access and sharing [4].

The root layer, powered by SSST, lays the groundwork for a healthcare ecosystem where data privacy and individual autonomy are paramount. It transforms the traditional model of centralized health records into a patient-centric system, where individuals are truly in control of their health data. This aligns with the goals of patient-centric, interoperable health IT systems, fostering trust, efficiency, and improved healthcare outcomes [2], [5], [11].

5.1.2 Personal Health Record

In the context of the *DHIN* framework, the Personal Health Record (PHR) functions as a Personal Data Store (PDS)—a concept that, while central to this discussion, will not be elaborated on here for brevity. See references [12], [13]. Building upon the foundation of Self-Sovereign Server Technology (SSST), the PHR is a critical component within the *DHIN* ecosystem, enabling patients to manage and share their health information securely and efficiently. Unlike traditional health records, the PHR is designed to be a dynamic, patient-controlled gateway to a lifetime of health information, protected by multiple layers of cryptographic security [10], [14].

At the core of this system is SSST, which acts as a secure bridge connecting healthcare professionals to the patient's health records. With the patient's consent, healthcare professionals can access and update these records through a secure, user-friendly interface [10], [14]. This process is similar to using a specialized, highly secure app, ensuring both ease of use and strong protection of sensitive health data.

A distinguishing feature of the *DHIN* framework is its ability to incorporate and expand upon established principles, such as those discussed in pre-existing works. The framework offers flexibility in managing access permissions by allowing patients to store these permissions either off-chain (on a private, secure server) or on-chain (directly on the blockchain). This approach ensures precise control over who can access their information and under what conditions [15], [16]. This system functions like a digital lock on a medical file cabinet, enabling patients to grant or revoke access to specific records for different healthcare providers as needed.

This level of control raises important considerations regarding the balance between patient autonomy and medical expertise. For example, while patients are generally motivated to maintain accurate records, certain scenarios—such as psychiatric care or pediatrics—may require professional oversight [17], [18], [19]. The *DHIN* framework is designed to accommodate these nuances, enabling customized access policies that respect both patient rights and the need for expert medical guidance.

The technical implementation is streamlined and secure. Healthcare providers access patient records through an API portal, governed by the permissions specified by the patient [20]. This process is akin to having a personal health data concierge, ensuring that only those with the correct credentials and permissions can access sensitive information.

Figure 2. *API access to personal health record. Adapted from NOSH & HIE of One [20].*

This patient-centric approach to health records not only enhances privacy and control but also has the potential to standardize health information formats across different providers. For instance, when moving to a new city, a patient's entire medical history could be instantly accessible to their new healthcare provider without the need for complicated record transfers or redundant testing [20].

Preliminary testing of these concepts has been conducted using open-source PHR systems like NOSH in family medicine settings [19]. Additionally, these ideas have been demonstrated through an integration with Gropper's SSST, providing a demo simulation of a real-life patient encounter between patients and physicians [20]. The principles and applications demonstrated in these settings can be extracted and applied agnostically to any Personal Health Record (PHR) system. This highlights the versatility and adaptability of these concepts within the context of the *DHIN* framework.

In summary, the PHR within the *DHIN* framework empowers patients to move from passive recipients of healthcare to active managers of their health information. This approach fosters greater engagement, accuracy, and continuity in medical care, placing patients at the center of their healthcare journey.

5.2. Real-World Layer

5.2.1 Interaction Layer: Redefining Patient-Doctor Relationships

The interaction between patients and physicians in the *Decentralized Healthcare Intelligence Network* (*DHIN*) framework represents a significant shift in healthcare provision, leveraging blockchain technology and self-sovereign identities to create a secure, patient-centric ecosystem. The objective is not merely to add a layer of technological sophistication but to fundamentally reshape how healthcare interactions are conducted, ensuring greater transparency, privacy, and patient control.

In the *DHIN* system, physicians establish their professional identity using a cryptographic key pair, functioning as a highly sophisticated, unforgeable digital signature [20]. Unlike traditional systems where credentialing might involve multiple intermediaries, in this framework, the physicians identity is verified by trusted entities such as medical societies or reputable third-party services [20]. This verification process ensures that only qualified professionals can access the system, aligning with the need for trust and integrity in healthcare interactions [20].

Key Benefits:

- **Credentialing by Trusted Entities:** Medical societies or established healthcare providers credential medical professionals, offering a trusted setup that patients are familiar with. This integration retains the confidence of patients while empowering doctors to operate within a decentralized framework, potentially recommending cost-effective treatments or referring patients to services outside traditional provider networks [10].
- **Enhanced Professional Autonomy:** Doctors can maintain full control over their practice, free from provider constraints, which can stimulate better market conditions and more personalized patient care [10].

When a doctor needs to access a patient's records, the process is designed to be seamless and secure. Rather than using a conventional username and password, physicians utilize their unique digital identity, which is verified against official medical directories [10]. This method ensures that only authorized professionals can access patient information, preserving both security and privacy.

Real-World Example: Imagine a scenario where a doctor prescribes medication or orders tests. Instead of a paper-based or email system prone to forgery and errors, the *DHIN* framework employs a digital identity system where each action by the physician – whether writing a prescription or ordering a test – is digitally signed [10]. This is akin to a digital prescription pad where each prescription is verifiably authentic and tamper-proof.

Consider the process of filling a prescription [10]:

- **Digital Prescription Transmission:** The doctor sends a digitally signed prescription directly to the patient's health record.
- **Patient Access and Verification:** At the pharmacy, the patient uses their own digital identity to access the prescription.
- **Pharmacy Verification:** The pharmacy's system verifies both the patient's identity and the doctor's digital signature, ensuring that the prescription is authentic and intended for the patient in question.

This streamlined process eliminates the need for physical prescriptions, reducing the potential for errors and fraud while maintaining a high level of security.

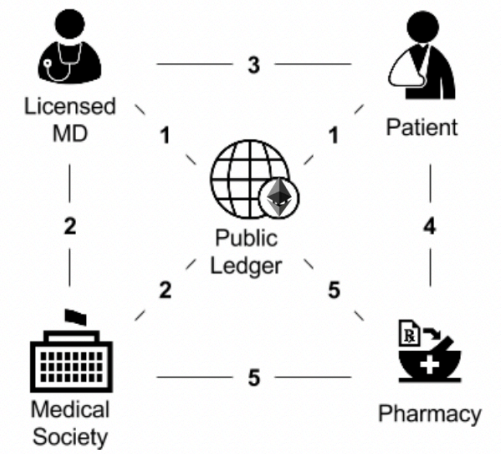

Figure 3. *An example of patient use of a pharmacy service. Adapted from HIE of One* [10].

Insights from Gropper (2020) [10]:

- **Credentialing Services:** The Medical Society provides valuable credentialing to its members without risking patient data breaches.
- **Physician Autonomy:** Licensed MDs have total control over their relationships with patients, maximizing the value of their professional licenses.

- **Patient Empowerment:** Patients can choose their pharmacies, preserving privacy and potentially saving money by working directly with their physician.
- **Pharmacy Innovation:** Pharmacies and other healthcare providers can innovate and add value independently of intermediary hospitals and EHR vendors.
- **Support Services:** Suppliers of SSST can offer value-added services to both physicians and patients without patient lock-in or disrupting the institutional trust chain.

The *DHIN* architecture extends beyond prescriptions to all aspects of healthcare, including referrals, imaging services, immunization records, and consent forms. Each interaction is secured, verified, and recorded, contributing to a comprehensive, patient-controlled health record.

Decentralization and Patient Empowerment: The decentralization inherent in the *DHIN* framework means that neither patients nor doctors need to rely on large, centralized electronic health record providers. Instead, each participant in the healthcare ecosystem – whether doctor, patient, or healthcare service provider – can interact directly and securely, with the patient maintaining central control.

In essence, the *DHIN* framework draws upon and extends the principles laid out by Gropper [10], particularly in establishing sovereign data stores that facilitate secure and efficient healthcare exchanges. While not limited to Gropper's work, these technologies are integral to the *DHIN* framework, embedding trust in every interaction, prioritizing privacy, and empowering patients to maintain true ownership of their healthcare journey.

5.2.2 Medical AI Devices

The *DHIN* framework introduces a theoretical model for integrating Medical AI Devices into a decentralized healthcare system. These systems are defined in this paper as entities that handle health data—whether by storing, retrieving, generating, sending, or processing it, including tools such as triage systems, diagnostic instruments, and AI decision support systems.

Currently, implementing AI systems in clinical environments faces major challenges [21]. A significant issue is the fragmentation of health data across various provider systems, which complicates the development and deployment of effective AI tools [4], [15], [21], [22], [23], [24]. These tools typically need accurate and up-to-date data from a single, unified source.

The proposed framework addresses this challenge with an approach that might seem ironic at first: it suggests centralizing all patient health data in one location. However, this centralization is not in the traditional sense of a single, centralized database. Instead, it centralizes data at the point where it is most relevant—the patient's own Personal Health Record (PHR). This means that while the data is distributed across various sources, it is centralized in the hands of the individual it concerns across their lifetime. This method resolves disputes over data access between healthcare providers and ensures that AI-generated outputs are recorded in the patient-owned PHR. A more complete PHR with information held over an individual's lifetime will enable more comprehensive information, thus enhancing the precision and accuracy of AI tools (e.g., decision support) as they work with more complete data about the patient.

In this model, Medical AI Devices would each be assigned their own Decentralized Identifiers (DIDs). These DIDs function similarly to unique digital IDs for people or organizations, allowing devices to interact directly with the patient's PHR [4]. This setup enables devices to read from and write to the PHR as needed.

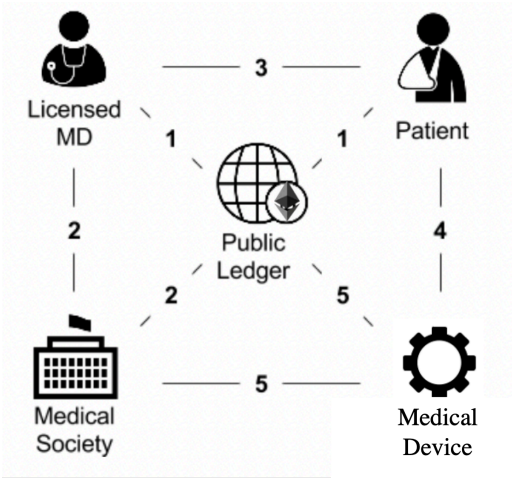

Figure 4. *An example of patient use of a Medical AI Device* [4]. *Adapted from HIE of One* [10].

Each system makes use of the same authentication architectures as physicians to attribute licensing/credentials of their system (e.g., FDA/CE, trusted third-party), as shown in **Figure 4**. Regardless of authentication, the incorporation of DIDs can reduce the manual effort of device labeling and provide a safer use of medical devices using their blockchain-based identities [25], although many open-source systems do not require licensing and authentication in practice. An open-source Medical AI Device is ideal for agile prototyping among physicians, patients, developers, and other key stakeholders, facilitating the adaptation, training, and deployment of AI tools.

Medical devices today are capable of storing control software and data, making them suitable for using DIDs [4]. By consolidating patient health data into a PHR, these devices can more effectively coordinate and manage health information. Medical AI Devices would analyze data from the PHR to enhance healthcare coordination.

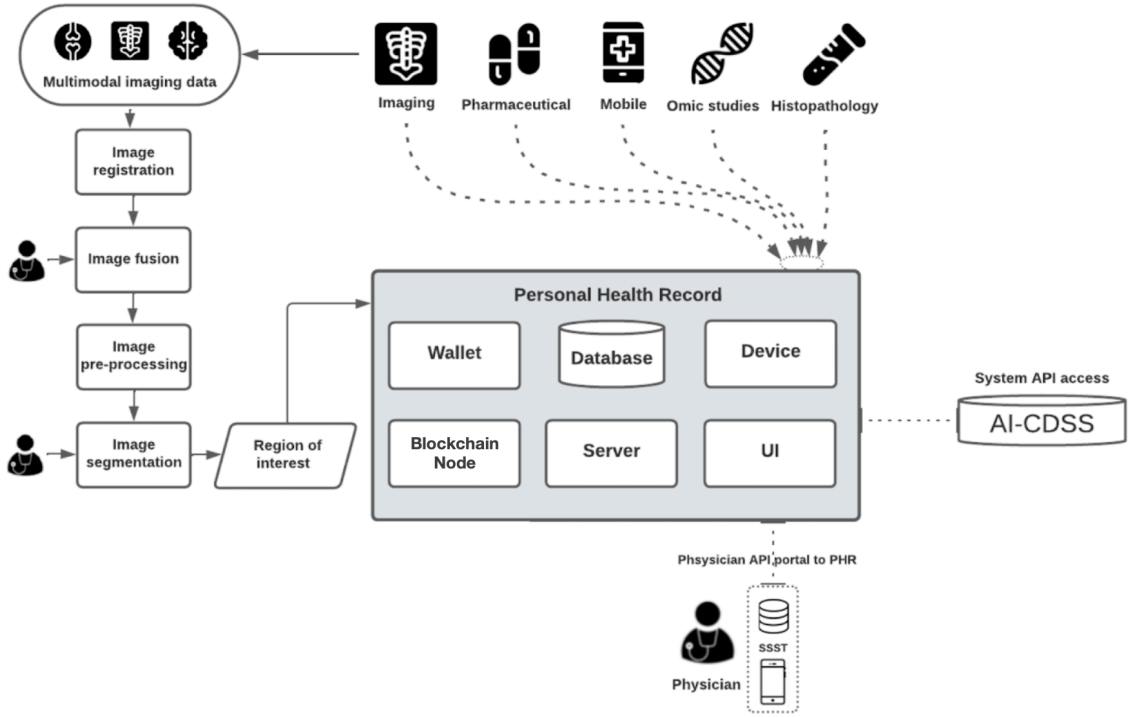

Figure 5. *Interactions amongst physicians and supporting systems that read and write directly into the personal health record are mediated by the blockchain.*

Medical AI Devices working with physicians would use the PHR to record their outputs, promoting better collaboration between physicians and AI tools. This approach aids in making informed healthcare decisions and enhances the accuracy of specific tasks in the healthcare process, such as diagnostic tools or mobile health devices.

The *DHIN* framework envisions Medical AI Devices running algorithms on the data stored in the PHR. This data could include details entered by physicians (e.g., symptoms, medical history, physical examination results) and additional information from sources like imaging or mobile health devices. The paper acknowledges that this is a relatively new area of research and calls for further study into integrating Medical AI Devices into clinical workflows using patient-owned PHR systems. This includes ensuring that these devices access accurate health data, assigning DIDs to Medical AI Devices, and effectively integrating AI with medical practice.

To implement this framework, a decentralized federated learning (FL) environment is essential. This environment would allow AI tools to develop using patient-owned health data while maintaining patient control over their information.

In summary, this theoretical framework sets the stage for the next layer of the system—the intelligence layer—focused on overcoming the challenges of deploying AI in a decentralized healthcare ecosystem.

5.3 Intelligence Layer

5.3.1 Decentralized Intelligence Network (*DIN*) Protocol

Decentralized Intelligence Network's (*DIN*) protocol operationalizes the federated learning (FL) architecture outlined in **Section 4** [3]. Built on a decentralized public blockchain infrastructure, the *DIN* protocol manages the coordination and rewards process for training machine learning models using data stored in Participant-owned Personal Data Stores (PDSs) [3]. In *DHIN*, this process occurs using patient-owned personal health records (PHRs). This approach ensures data sovereignty while enabling collaborative AI development.

Participants opt into FL protocols defined by *DIN* protocol smart contracts (SC) on a public blockchain. The protocol leverages the blockchain to coordinate the FL process and manage rewards, while ensuring that raw data remains within the Participant's sovereign datastore [3]. Only model updates are shared during the FL process, preserving privacy and control.

To enhance scalability and computational efficiency, the *DIN* protocol incorporates an off-chain decentralized file storage system, such as the InterPlanetary File System (IPFS) [26]. This system provides a location for uploading and downloading model updates during the learning process, optimizing participation costs and complementing the blockchain's transaction recording capabilities [3].

This setup ensures a fair and transparent reward system while maintaining data sovereignty and reducing reliance on centralized infrastructure. The following sections detail the specific methodologies and operational mechanisms of the *DIN* protocol, including the roles of key participants such as Model Owners, Participants, and Evaluators [3].

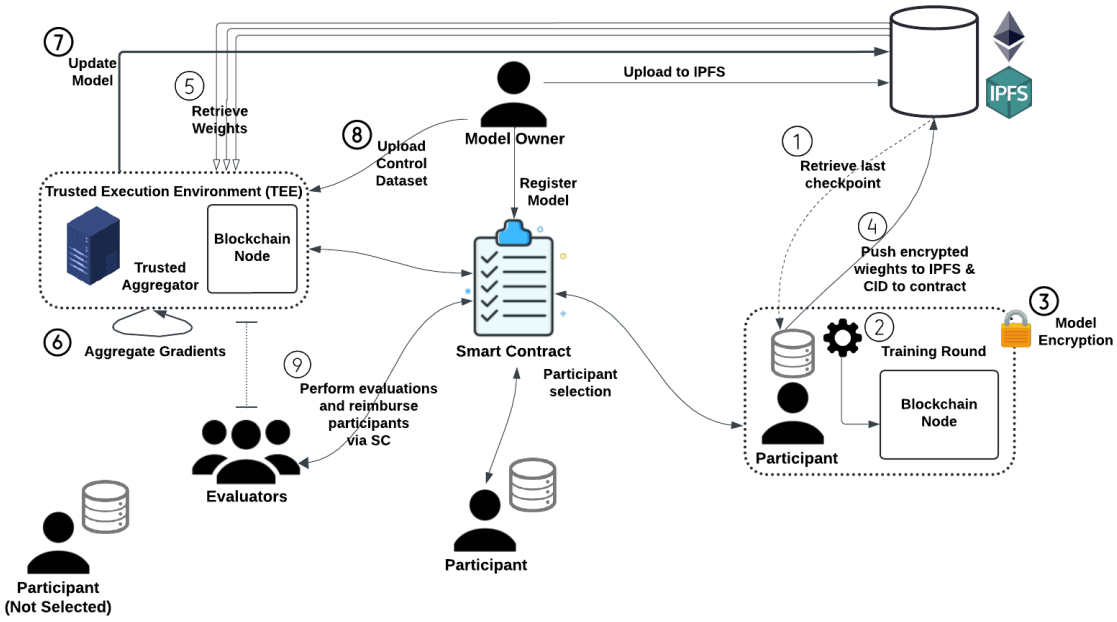

Figure 6. *Global overview of a training round. Adapted from Consensys Health [9].*

1. **Model Owner:** a. Deploys *intelligence* smart contracts (SC) on the blockchain. Although referred to as "*intelligence*," multiple contracts may work in unison on-chain, fulfilling different roles to enhance scalability and efficiency (e.g., DIN protocol, evaluator registry, evaluator staking, aggregator management, NFT staking for evaluators, reward distribution, etc.). b. Creates genesis model, uploads to IPFS, records its CID on the contract. c. Deposits reward amount in smart contract (SC) to be allocated to Participants after FL rounds.
2. **Upon Model Owner's transaction confirmation:** a. Participants can see the genesis model CID and download it from IPFS. b. Using their own data stores, Participants run training iterations on model. c. Participants encrypt their models using secure aggregation protocol (e.g. Bonawitz et al. 2019) [27]. d. Participants upload their encrypted updated models to IPFS, recording CIDs on *intelligence SC*. e. This completes one FL training round; Participants wait for the next round.
3. **Aggregation:** a. The secure aggregator server, operated by the Model Owner, fetches all encrypted weights from decentralized storage as instructed and verified on-chain. b. Standard Process: The Model Owner uses the secure aggregator to perform aggregation using secure aggregation protocol c. The secure aggregator server, using a blockchain node operated by the Model Owner, fetches encrypted model weights from IPFS via CIDs recorded on the blockchain, ensuring secure processing. d. To manage scalability, a master aggregator spawns multiple aggregators each round, as in Bonawitz et al. (2019) [27], with device management integrated into on-chain coordination by the *DIN* smart contracts (SC). The protocol's existing components handle device connections, make local decisions based on instructions from the Coordinator on-chain, and forward devices to the aggregators, ensuring efficient allocation while minimizing communication with the on-chain Coordinator. e. The aggregated results are confirmed on-chain before the global model update is revealed to the Model Owner. f. The Aggregator uploads the new global model update to shared storage and updates on-chain *intelligence* SC with the new pointer. g. Multiple secure aggregator servers or the preferred Model Owner environment for aggregation can be instantiated to handle larger workloads or enhance decentralization and reliability.
4. **Next Training Round:** a. A new global model is published by the Model Owner for the next training round. b. **Optional:** Participants can check all CIDs of the previous round's model updates in their aggregator subgroup. b. Download updates from IPFS and independently calculate the mean aggregate for their subgroup (all participants in a subgroup reach the same result). c. Each subgroup performs the average of their own subgroup's global model and then averages other aggregator groups' averages as published on IPFS located on-chain to reach global model published by Model Owner.

5. **Model Aggregation Continuation:** a. The Model Owner continues to aggregate FL round updates, averaging their updates as rounds progress. b. The Model Owner tests the average global model update against their control dataset to determine when they are satisfied with the training. c. The decision to end or continue training is communicated to the SCs either in advance or during training, depending on the Model Owner's availability of funds. d. The Model Owner signals the final round, after which one additional round occurs to evaluate Participants' rewards before the final global model is revealed to the Model Owner.
6. **Post-Training:** a. The Model Owner encrypts the control dataset and uploads it to the aggregation server and may open a Trusted Execution Environment (TEE) and/or utilize privacy-preserving techniques if desired. b. The *intelligence* SC randomly assigns a standardized ratio of Evaluators to Participants (e.g., 1:10) according to the Model Owner's needs [8]. c. Evaluators, who are randomly assigned to a specific subgroup FL aggregation round for scalability, benchmark all Participant models of that aggregator group inside the secure aggregator server, potentially using TEE and/or other privacy-preserving techniques for off-chain benchmarking (unless there is a disagreement, in which case benchmarking is conducted on-chain).
7. **Evaluators:** a. Anyone staking native token can be an Evaluator. b. Evaluate Participants' models against the control dataset in the secure server environment. c. Submit evaluation consensus scores with ZK-proofs to *intelligence* SC (see **5.3.2 Decentralized Auditing Protocol** for rewards).
8. **Distribute Rewards:** a. *Intelligence* SC calculates reward fraction for each participant based on objective scores (e.g., as Shapley Values, Substra Scoring, Median Scoring, and other open-preference scoring methods) [8], [28]. b. Distributes Model Owner's deposited reward accordingly, per recorded scores.

In the related works section, we briefly reference the protocol by Bonawitz et al. (2019), which employs synchronous rounds with subsets of devices for scalable federated learning. Their Secure Aggregation method ensures privacy by encrypting device updates, and intermediate results are aggregated by dedicated actors to manage computational costs. This framework builds on these principles, enhancing scalability and privacy while addressing sovereignty and decentralized participation. By integrating subgroup aggregation and subset evaluation, we extend Bonawitz et al.'s approach to support sovereign networks and broader applications [29], [30].

To incentivize participation in this scalable and decentralized FL process, we propose integrating a scalable, "trustless" rewards mechanism—one that does not require a third party for transactions. This mechanism is discussed in the following section.

5.3.2 Decentralized Auditing Protocol

Delineating the roles of Participant and Evaluator in the protocol raises concerns about the potential misuse of the control dataset by Evaluators in federated learning (FL) scenarios. In previous examples, each Participant evaluated every other Participant using their data and an objective scoring metric [8], [28]. However, once we distinguish between Participant and Evaluator and adjust the protocol to scale and generalize to other data types, a new issue emerges. Evaluators might download and illicitly share the control dataset published by the Model Owner with Participants in one or more aggregator subgroups in a FL rounds when benchmarking Participants' contributions after Model Owenr signals final round, potentially leading to harmful activities such as model poisoning or unfair compensation during the model training process.

To mitigate these risks, implementing secure evaluation mechanisms where the test dataset remains concealed from the Evaluators is essential. Evaluators can prove to the system they correctly evaluated against the control dataset without accessing it, mitigating risks of misuse or leakage. Evaluators can be provided with high-quality, well-distributed, and highly representative control datasets by the Model Owner. Evaluators can use this as a benchmark to evaluate each Participant's models as shown in **Figure 6**. The step-by-step protocol elaborates on the processes of Evaluators' involvement in the **decentralized auditing protocol** within the rewards process, as illustrated in **Figure 6**, as follows:

1. Model Owner selects secure computation method: a. Opens a TEE within the aggregation server, and/or b. Chooses alternative privacy-preserving computation meeting security standardsModel Owner distributes encrypted control dataset to chosen secure environment.
2. Within a secure environment: a. Evaluators compute performance metrics (such as accuracy, precision, and recall) on models using a control dataset and benchmark these metrics using objective scoring methods (e.g., Shapley Values of

Distribution, Substra, Median Scoring, etc) [cite] b. Evaluators utilize remote attestation mechanisms to prove secure execution. c. Evaluators generate Zero-Knowledge Proofs (ZKPs) for each metric.

3. Contingency for MPC: a. Participants encrypt model updates using MPC or other privacy-preserving technique to benchmark Participants' scores. b. Evaluator evaluates models on encrypted test dataset using privacy-preserving protocol.
4. Evaluators submit ZKPs and encrypted evaluations to blockchain *intelligence* SC.
5. *Intelligence* SCs verifies ZKPs using a consensus mechanism (e.g., majority agreement).
6. Based on verified scores, *intelligence* SC calculates and distributes rewards transparently.
7. All privacy-preserved transactions are recorded on the blockchain for an immutable audit trail.
8. Privacy-preserving techniques (e.g., MPC) along with TEEs and ZKPs, provide multiple layers of security. MPC, in particular, can conceal data, TEEs isolate it, and ZKPs verify correctness without revealing information.
9. Protocol protects against insider threats by computing within secure environments and exporting only ZKPs.
10. Final Global Model Update: a. Model Owner signals end of training process upon satisfaction with model performance. b. Final Global Model update revealed to Model Owner after completion of rewards process.

DIN can utilize any contributivity scoring procedure for Evaluators to perform their off-chain evaluations of Participants' contributions. The specific procedure depends on the context of the learning task being conducted. This evaluation process is triggered when the Model Owner is satisfied with the global model's performance metrics (e.g., F1 score, accuracy, etc.), and occurs after a given number of FL rounds have been iterated in the FL process.

Following the approach of Bonawitz et al. (2019), this protocol can employ synchronous rounds with subsets of devices [27]. Evaluators are randomly assigned to aggregator Federated Learning (FL) participant subgroups at a specific ratio (e.g., one Evaluator for every ten Participants). This approach enables high scalability, allowing the system to handle a large number of devices while maintaining efficiency.

Evaluators perform their off-chain evaluations of each Participant's model using the control dataset published by the Model Owner, encrypt their score, and report their encrypted scores with a Zero-Knowledge Proof (ZKP) to the *intelligence* SC. Each Evaluator first reports to the smart contract the set of Participants whose models were successfully validated (i.e., within an acceptable bound specified by the Model Owner). Participants who fail this test are eliminated in that particular FL round. Once all the scores are received, each Evaluator provides the decryption key to provably reveal their score to the *intelligence* SC.

This protocol ensures secure and reliable model evaluations within a secure server environment and can incorporate privacy-preserving techniques such as Multi-Party Computation (MPC), as well as third-party Trusted Execution Environments (TEEs) selected by the Model Owner, safeguarded by ZKPs and blockchain technology. These precautions ensure that the protocol uses encryption to make individual evaluations uninspectable by any central authority or even other participants in the FL process, as detailed in the threat model (**see Section 6**). These techniques could be applied in the context of assessing rewards, similar to how they're used in aggregation.

The protocol aligns with the Model Owner's incentives, who, despite bearing the cost, benefit from robust data contributions and trustworthy evaluation processes necessary for successful model training and improvement. Crucially, it preserves data sovereignty for Participants—no central authority determines access to their data, which never leaves its original storage. Rewards are determined in a decentralized manner by a public blockchain smart contract based on pre-defined, auditable, and transparent metrics, eliminating the need for a central authority to decide compensation. This approach overcomes the potential issues of centralized systems, such as dishonest aggregation or external attacks, while maintaining the scalability benefits demonstrated by Bonawitz et al. (2019) [27].

This decentralized auditing protocol maintains Participants' autonomy while enabling secure, reliable, and incentive-aligned model evaluations. It addresses the challenges of sovereignty and decentralized participation within a novel framework, enabling wider application of these tools and workflows within sovereign networks.

6. Threat Model

The papers threat model addresses potential risks in the federated learning (FL) process, ensuring robust security and privacy. The protocol is resilient to up to 50% malicious participants, leveraging public/private key cryptography and a proof-of-stake consensus mechanism. By using immutable storage on IPFS we ensure data integrity. Additionally, the use of Zero-Knowledge Proofs (ZKPs) and Trusted Execution Environments (TEEs) mitigates risks associated with model evaluation and reward distribution. This comprehensive approach ensures the security and reliability of the FL process, maintaining participant trust and data sovereignty.

Firstly, in an experiment with N agents, it is resistant up to $M \in [0, N/2)$ agents neglecting to follow the protocol for the experiment to maintain its integrity [8]. For example, public/private key cryptography and a proof-of-stake consensus protocol secure the Ethereum blockchain. Currently, there are no feasible attacks on the Ethereum Network, without controlling 50% of the computational power of the entire Ethereum network and such an attack has never been successful on the Ethereum mainnet [31].

Secondly, as a public blockchain is public and anonymous, clients could enroll multiple times in an experiment and thus have a disproportionate participation. However, through decentralized identity verification, verifiable credentialing, or manual processes, agents can ensure that each other agent controls only one account [15], [32], [33].

Third, IPFS is immutable, meaning agents cannot change their model after submitting the cryptographic hash to the smart contract [26]. Like in BlockFlow, the *DIN* protocol requires each agent to report if it can load strictly more than N/2 models, and have strictly more than N/2 agents report the same for their model. The *DIN* threat model guarantees that there are strictly more than N/2 honest Participants. Additionally, as long as N/2 or more Evaluators who receive these models for evaluation are honest, which the *DIN* protocol guarantees, the system remains resistant to N/2 attacks. Since IPFS allows anyone to share any content, one or more honest parties would share the model with all other Participants if they are unable to retrieve a model directly from the source (e.g., due to firewall restrictions). Therefore, each Participant would still be able to obtain all necessary models [8], [26].

Fourth, there are several possible attacks on the contribution scoring procedure itself. Malicious models are those with weights that do not reflect a truthful dataset, such as models trained on randomly generated data or inverted output features. Naively averaging such models into a global model would likely harm the shared objective. The *DIN* protocol can choose contribution-scoring procedures that penalize those who submit malicious models. For instance, BlockFlow (2020) uses a contributivity score system where lower scores result in less cryptocurrency received [8], [26]. In this system, any agent with an evaluation more than 0.5 away from the median score receives an overall score of 0 and no share of the cryptocurrency pool [8], [26]. This penalizes attempts to fabricate scores, as the protocol limits a Participant's overall score to the evaluation furthest from the median [8], [26].

Fifth, Participants can collude during the training process to submit better models by secretly sharing raw data or models among $M<N2$ colluding Participants [8], [26]. The *DIN* protocol rewards Participants who contribute strong models, and it is acceptable for multiple Participants to submit identical models. Such collusion is not considered an attack, as it is similar to having many Participants with strong datasets [8], [26]. For attacks by Evaluators in the evaluation process, the I smart contract can use encryption and a commit-then-reveal protocol (e.g., Secret Sharing MPC, Elliptic Curve Diffie-Hellman keys, etc.) to prevent Evaluators from copying others' scores without collusion [34]. If a minority subset of malicious Evaluators reports perfect 1.0 scores for certain models and 0.0 scores for all others (e.g., models from honest agents), the median score is guaranteed to be between the minimum and maximum scores reported by the honest agents, as long as there are strictly fewer than half malicious Evaluators [8], [26]. Evaluators are incentivized to stake a native token to gain the right to evaluate Participant models in the rewards process within a proof-of-stake (PoS) ecosystem. Evaluators found acting maliciously are slashed from the network, losing some or all of their stake, thus maintaining network security and incentivizing honest work.

Sixth, in this papers threat model, it is crucial that the control dataset provided by the Model Owner remains encrypted to prevent its misuse. If the control dataset were accessible to colluding Participants, Model Owners, Evaluators, or other entities, they could exploit it to skew the reward distribution. For example, colluding parties could use the control dataset to strategically improve their model performance or manipulate evaluation outcomes to gain undeserved rewards. Encrypting the control dataset ensures that it cannot be revealed or utilized by these entities to unfairly influence the results. To enhance security further, dual protection strategies can be employed. For instance, the preferred scoring method, such as median scoring used by BlockFlow (2020), can be integrated into the protocol [8]. In this approach, any score deviating significantly from the median—beyond a specified

threshold—can be penalized. BlockFlow's method maps any score differing by more than 0.5 from the model's median to a score of 0, with an a priori score set at 0.5 [8]. This mechanism encourages evaluators to provide honest assessments by penalizing scores that deviate substantially from the median. This method helps mitigate the risk of anomalous scores due to collusion and maintains fairness in the reward distribution process. Overall, encrypting the control dataset and employing robust scoring mechanisms collectively safeguard the integrity of the evaluation process and prevent potential manipulation by malicious actors.

Seventh, both the aggregation process involving trusted third-party secure aggregation servers and other such as Trusted Execution Environments (TEEs) with Zero-Knowledge Proofs (ZKPs) introduce distinct threat models. Concerns with aggregation hardware include potential data interception and manipulation, insider threats at third-party providers, hardware vulnerabilities such as side-channel attacks, and compliance issues with data protection regulations [35], [36]. TEEs, while isolating sensitive computations, face risks from hardware exploits, software vulnerabilities, and third-party trust issues. Additionally, implementations of ZKPs must be carefully managed to avoid cryptographic flaws that could undermine their effectiveness [37]. To counteract these threats, robust encryption, rigorous access controls, regular security audits, and compliance assurance are employed [37]. These measures ensure that data remains confidential and integral, reducing the dependency on trust by making processes transparent and verifiable through smart contracts on the blockchain. This strategy enhances security and stabilizes residual risks within a robust, transparent operational framework.

7. Decentralized Insurance Solution

The *DHIN* framework introduces an innovative approach to healthcare financing through a decentralized insurance solution, creating a self-sustaining cycle of value in healthcare.

At its core, this system uses cryptographic micropayments as a reward mechanism for patients who contribute their health data to AI development. Specifically, micropayments are used by Model Owners to access on-chain smart contracts which coordinate the federated learning process with patients' personal health records (PHRs), which serve as personal data stores (PDS). This approach ensures data sovereignty while enabling collaborative AI development. Imagine receiving a small digital payment each time your health information helps train an AI model to improve healthcare outcomes. These micropayments are deposited directly into patients' digital wallets (e.g., Taho [38]), seamlessly and securely.

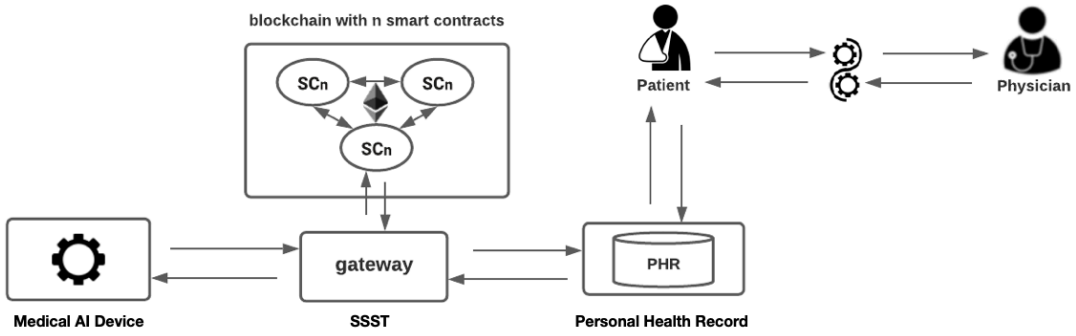

Figure 7. *The rewards mechanism of cryptographic insurance in a learning health system.*

Here's how the cycle works:

1. Patients and physician interact as usual, with health data recorded in the patient's Personal Health Record (PHR).
2. Model Owners (i.e., Medical AI Devic Developer) use cryptographic payments to access on-chain smart contracts, allowing them to train their models on patients' health data through federated learning protocols.
3. Patients receive these micropayments in their digital wallets when they successfully contribute to an AI training round.
4. Patients can then use these rewards to purchase decentralized insurance premiums, which in turn fund or subsidize their healthcare provision.

This cyclical mechanism creates a learning health system where value continuously circulates. New health technologies, embedded in AI models, inform and improve healthcare delivery. As patients use their rewards to fund their care, they generate more valuable health data, which then feeds back into AI development.

Importantly, this system preserves patient sovereignty over their data while providing AI tools permissionless access to up-to-date health information. It's a win-win scenario: patients benefit financially from their data, and AI systems get access to high-quality, diverse datasets.

The system also addresses a key challenge in healthcare: funding for services and treatments. By creating a fungible asset (the cryptographic micropayments) that can be used for insurance premiums, the framework establishes a sustainable funding model for decentralized healthcare.

In federated learning with non-homogenous data sources such as health data, data quantity and quality are the most valuable contribution to enhancing the accuracy of training machine learning algorithms [39]. Interestingly, this system may naturally benefit those with chronic or ongoing health conditions. These patients, who typically have more extensive health records due to frequent care, are likely to receive greater rewards in the learning process [39]. This aligns well with their potentially higher insurance needs. Furthermore, learning on health data is not limited to one occasion or use-case, but multiple occasions and use-cases.

Looking to the future, this framework lays the groundwork for more efficient insurance claim verification. By leveraging blockchain technology, the system could provide a reliable, immutable source of information for verifying insurance credentials and claims [4]. Decentralized insurance protocols are well suited to enhance the functions of more cost-effective and reliable coverage schemes [4].

The long-term vision is ambitious yet promising: a scalable system that could potentially reduce healthcare insurance costs, lower barriers to healthcare provision, and ultimately increase access to quality healthcare for all.

In essence, the *DHIN's* decentralized insurance solution represents a paradigm shift in healthcare financing, aligning patient interests, technological advancement, and healthcare provision in a novel, self-sustaining ecosystem.

8. Tokenomics, Governance, and Public Goods

This section introduces a novel protocol designed to incentivize and manage participation within a *Decentralized Intelligence Network (DIN)*. It highlights the significance of integrating economic and social dimensions into modern systems architecture, emphasizing the importance of tangible incentives and broader systems design. This approach aligns with the perspectives of early internet pioneers and recent literature.

The protocol incorporates a novel public goods funding mechanism that not only secures the network but also integrates seamlessly with existing public goods ecosystems. It leverages NFT staking, unique evaluation mechanisms, and principles of Partial Common Ownership (PCO) to foster a circular economy for the development and enhancement of global AI models. While the staking mechanism is central to the protocol, it does not necessarily require a native token or coin, though the potential for such integration could be explored.

NFT Proof of Stake (PoS) Mechanism for Evaluators: This component explores the novel application of combining innovative evaluation mechanisms and staking methods with Harberger taxation and PCO principles. It focuses on using NFT staking to secure the network and aims to establish standard, open-source implementations of Partial Common Ownership (PCO) of Ethereum ERC721 NFTs. By integrating these methods with Harberger taxation, this approach not only strengthens network security but also seeks to contribute revenue to public goods ecosystems, potentially enabling self-funding development of the *DIN* and other adjacent public goods ecosystems. Core components of the staking mechanism are detailed:

- **Network Fees**

        - The Model Owner, who trains the algorithm and pays Participants for the Federated Learning (FL) process on sovereign data stores, is assumed to pay this fee as part of the process.
        - **The fee distribution from rewards to participants and evaluators is dynamic and remains an open question.** It could be either an added tax on the reward or a distribution model, such as allocating 97% of the reward to participants and 3% to evaluators.
        - This fee is separate from and in addition to any blockchain-specific gas costs, and is a part of the estimated costs of the rewards process.
        - The exact calculation and distribution method for this reward is an open question for experimentation.
- **Fee and Reward Currency**
    - Fees and rewards are primarily paid in a stablecoin (e.g., USDC, etc) however new stablecoin assets prevent value depreciation (e.g., RAI) that is not pegged to centralized stablecoin assets are worth exploring.
    - This approach helps mitigate risks associated with inflationary measures or other external factors affecting centralized stablecoins, to fairly reward Participants.
    - The use of native tokens in lieu of stablecoins remains an open question for further exploration.
- **Evaluator Staking and NFTs**
    - Evaluators must stake an NFT to participate in evaluation processes i.e., to earn network fees.
    - These NFTs represent the Evaluator's stake and reputation in the network.
    - NFT values are self-assessed and based on the chosen stablecoin to ensure stability.
- **NFT Valuation and Taxation**
    - NFT values are subject to **Harberger taxation:**
        - Owners periodically self-assess their property and pay tax on its value.
        - Others are able to purchase the property from the owner at the taxed price at any time, forcing a sale .
    - Harberger taxation is priced in the fees paid to Evaluators (i.e., stablecoin, native token,etc) and is charged periodically based on the value of the owner's NFT asset.
    - Values can be adjusted either:
      a) Dynamically based on performance, or
      b) At set periods, after which they become open to auction (PCO mechanism).
- **Evaluator Incentive Structure**
    - As Evaluators perform more work, they:
      a) Receive more rewards from fees.
      b) Can assess their NFT stake at a higher value, as the profits from fees exceed the amount taxed.
    - This structure incentivizes high-quality evaluations and active participation.
- **Partial Common Ownership (PCO)**
    - Implements a mechanism where NFTs can be put up for auction after certain periods or at all times.
    - Helps maintain fair valuation and prevents monopolistic behavior; setting non-speculative asset pricing which reflects work done in the network.

This protocol aims to create a balanced, fair, and efficient system for decentralized machine learning model evaluation and improvement. It is also a novel proposal for the implementation of a public goods funding mechanism. By leveraging economic incentives and novel ownership structures, the protocol aligns the interests of all participants towards the common goal of advancing AI capabilities, while simultaneously contributing to the funding of public goods. It addresses potential challenges and areas for further refinement, ensuring a sustainable and equitable ecosystem for all involved. Alternatively, it considers leveraging existing ERC standards or exploring other traditional staking protocols—such as those used with a native coins—that do not aim to become part of a public goods ecosystem to meet the wider protocol requirements.

Public Goods & Governance in *DIN*: Projects like Gitcoin funding vision for community-driven proposals and public goods funding highlight the increasing focus on supporting shared resources and communal benefits. The proceeds from the taxation mechanism within our protocol are **allocated towards funding *DINs* open-source public goods infrastructure like or broader ecosystems**. This could involve supporting the network's own public goods or contributing to initiatives such as Gitcoin grants. **These approaches align with principles embraced by communities such as RadicalxChange (RxC), the Plurality Book, Ethereum blockchain ecosystems, and the Kernel Community. They also reflect the values of RDI Berkeley in DeAI, which emphasize openness, responsibility, and a democratized AI economy.** By integrating transparent, community-driven mechanisms and decentralized models, these approaches promote equitable participation and resource distribution . Their work underscores the need for **ongoing experimentation and a willingness to explore new ideas**, crucial for developing transparent systems that benefit the public.

DIN is an organizational network and public goods software by design, though its applications extend beyond the realm of public goods. It emphasizes self-sovereignty and may explore governance mechanisms to keep engaged, incorporating models such as Gov4Git (non-coin-based voting) and quadratic voting tools to engage contributors. The concept of sovereign networks is particularly relevant to personal data stores, as it delves into decentralized governance, commons-based peer production, and digital communities with shared values that operate independently of traditional structures. These sovereign networks explore the potential for decentralized models to reshape governance and resource distribution in novel ways, and may also contain relevance to experimental network states .

DIN may include **DPPs** (*DIN* Proposal Protocols) for drawing attention to proposals for improving the network. These proposals will be voted upon by contributors who are allotted non-coin voting credits, employing tools akin to those or including Gov4Git . A flexible, inclusive rollout driven by community input is essential to mitigate wealth concentration within the crypto ecosystems. Circulating financial value within ecosystems that benefit the public can stimulate economically advantageous societies. This approach, coupled with a commitment to continuous innovation, is vital for advancing these concepts in dynamic and impactful ways.

6. Conclusion and Future Directions

Decentralized Health Intelligence Network (*DHIN*) is a groundbreaking extension of the *Decentralized Intelligence Network* (*DIN*), designed to transform healthcare data management. *DHIN* builds on *DIN's* principles to address healthcare-specific challenges by utilizing personal data stores, decentralized federated learning, and a trustless rewards system.

Decentralized Health Intelligence Network (*DHIN*) extends the *Decentralized Intelligence Network* (*DIN*) framework to offer a transformative approach to healthcare data management. *DHIN* empowers patients with complete control over their health data, which remains securely within their Personal Health Records (PHR) stores. Patients can easily manage their data-sharing preferences by toggling a simple opt-in or opt-out switch, streamlining interactions with healthcare providers and researchers.

The system employs *DIN's* scalable federated learning protocol on a public blockchain, ensuring that patient data never leaves the PHR. Instead, only model parameter updates are shared, employing advanced privacy-preserving techniques to safeguard data. A decentralized, scalable auditing system fairly distributes cryptographic micropayments to incentivize patient participation, without intermediary control.

DHIN addresses the challenge of data silos, facilitating large-scale, collaborative healthcare research while maintaining individual privacy. The streamlined opt-in/opt-out functionality and one-time setup enhance overall efficiency in patient interactions with healthcare services and research entities. Rather than being constrained by siloed data stores, broader adoption of the *DHIN* allows the FL protocol to operate across various geographical regions. This enables patients to opt in and contribute

their health data, expanding the data pool beyond fragmented, patient-specific sources. As a result, AI tools can achieve greater accuracy, particularly for rare diseases, by leveraging a more comprehensive dataset that transcends individual patient silos.

Looking to the future, *DHIN* aims to address crucial nuances beyond the core framework by incorporating data privacy measures, such as protocol opt-ins to protect Participant identities, potentially using zero-knowledge proofs (ZK-Proofs). The network also plans to develop broader connections with sovereign data stores, explore non-speculative token designs, and work toward establishing a decentralized health insurance solution that offers wider coverage and additional financial incentives. These efforts will ultimately integrate *DHIN* into a comprehensive and innovative healthcare ecosystem.

We invite researchers, practitioners, and stakeholders to engage with the *DHIN* framework. By collaborating, we aim to advance scalable, sovereign data solutions that enhance healthcare technology while respecting individual rights and data ownership. Together, we can build a more equitable, efficient, and innovative healthcare system that benefits all participants.

7. Acknowledgments

Special thanks to Paritosh Ramanan, Rui Zhao, Harry Cai, Peng "Dana" Zhang, and Jesse Wright for their invaluable discussions on many of these ideas. Specifically, their contributions include scalable FL architectures (Paritosh), scalable decentralized auditing protocol (Paritosh, Harry), sovereign architectures, and decentralized identity management (Dana). Special thanks to Rui and Jesse, who provided significant insights across various aspects of the framework. Their insights and contributions have significantly shaped the development of the *DIN* framework. Additionally, works on *DIN* were selected for presentation as a speaker at the **Summit on Responsible Decentralized Intelligence - Future of Decentralization and AI**, hosted by **Berkeley RDI** on **August 6, 2024**, at the **Verizon Center, Cornell Tech Campus, Roosevelt Island, NYC**. This summit offered an exciting opportunity to share and further refine these ideas with a broader audience.

References

[1] T. Davenport and R. Kalakota, “The potential for artificial intelligence in healthcare,” *Future Healthc J*, vol. 6, no. 2, pp. 94–98, Jun. 2019, doi: 10.7861/futurehosp.6-2-94.

[2] V. Ehrenstein, H. Kharrazi, H. Lehmann, and C. O. Taylor, *Obtaining Data From Electronic Health Records*. Agency for Healthcare Research and Quality (US), 2019. Accessed: Jan. 21, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK551878/

[3] A. Nash, “Decentralized Intelligence Network (DIN),” Aug. 04, 2024, *arXiv*: arXiv:2407.02461. doi: 10.48550/arXiv.2407.02461.

[4] P. Zhang and T.-T. Kuo, “The Feasibility and Significance of Employing Blockchain-Based Identity Solutions in Health Care,” in *Blockchain Technology and Innovations in Business Processes*, S. Patnaik, T.-S. Wang, T. Shen, and S. K. Panigrahi, Eds., in Smart Innovation, Systems and Technologies. , Singapore: Springer, 2021, pp. 189–208. doi: 10.1007/978-981-33-6470-7_11.

[5] J. M. Grossman, K. L. Kushner, E. A. November, and P. C. Lthpolicy, “Creating sustainable local health information exchanges: can barriers to stakeholder participation be overcome?,” 2008.

[6] S. T. Argaw, N.-E. Bempong, B. Eshaya-Chauvin, and A. Flahault, “The state of research on cyberattacks against hospitals and available best practice recommendations: a scoping review,” *BMC medical informatics and decision making*, vol. 19, no. 1, pp. 1–11, 2019.

[7] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, “Key challenges for delivering clinical impact with artificial intelligence,” *BMC medicine*, vol. 17, no. 1, pp. 1–9, 2019.

[8] V. Mugunthan, R. Rahman, and L. Kagal, “Blockflow: An accountable and privacy-preserving solution for federated learning,” *arXiv preprint arXiv:2007.03856*, 2020.

[9] J. Passerat-Palmbach, T. Farnan, R. Miller, M. S. Gross, H. L. Flannery, and B. Gleim, “A blockchain-orchestrated federated learning architecture for healthcare consortia,” *arXiv preprint arXiv:1910.12603*, 2019.

[10] A. Gropper, “Powering the physician-patient relationship with HIE of one blockchain health IT,” in *ONC/NIST use of Blockchain for healthcare and research workshop. Gaithersburg, Maryland, United States: ONC/NIST*, 2016.

[11] P. Esmaeilzadeh and T. Mirzaei, “The Potential of Blockchain Technology for Health Information Exchange: Experimental Study From Patients’ Perspectives,” *Journal of Medical Internet Research*, vol. 21, no. 6, p. e14184, Jun. 2019, doi: 10.2196/14184.

[12] M. Van Kleek and K. OHara, “The Future of Social Is Personal: The Potential of the Personal Data Store,” in *Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society*, D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, and J. Stewart, Eds., in Computational Social Sciences. , Cham: Springer International Publishing, 2014, pp. 125–158. doi: 10.1007/978-3-319-08681-1_7.

[13] R. Zhao *et al.*, “Libertas: Privacy-Preserving Computation for Decentralised Personal Data Stores,” Sep. 28, 2023, *arXiv*: arXiv:2309.16365. doi: 10.48550/arXiv.2309.16365.

[14] Adrian Gropper, *Patient Centered Health Records - NOSH and HIE of One and beyond...*, (Jan. 21, 2016). Accessed: Apr. 19, 2022. [Online Video]. Available: https://www.youtube.com/watch?v=ehcJJMB3xvM
[15] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom, "FHIRChain: applying blockchain to securely and scalably share clinical data," *Computational and structural biotechnology journal*, vol. 16, pp. 267–278, 2018.
[16] A. Ekblaw, A. Azaria, J. D. Halamka, and A. Lippman, "A Case Study for Blockchain in Healthcare:'MedRec' prototype for electronic health records and medical research data," in *Proceedings of IEEE open & big data conference*, 2016, p. 13.
[17] D. Blumenthal, "Implementation of the federal health information technology initiative," *New England Journal of Medicine*, vol. 365, no. 25, pp. 2426–2431, 2011.
[18] E. Katsh, N. Sondheimer, P. Dullabh, and S. Stromberg, "Is There an App for That-Electronic Health Records (EHRS) and a New Environment of Conflict Prevention and Resolution," *Law & Contemp. Probs.*, vol. 74, p. 31, 2011.
[19] "NOSH ChartingSystem | A new open source health charting system for doctors." Accessed: Apr. 19, 2022. [Online]. Available: https://noshemr.wordpress.com/
[20] Adrian Gropper, *HIE of One Highlights*, (Oct. 15, 2017). Accessed: Apr. 19, 2022. [Online Video]. Available: https://www.youtube.com/watch?v=N_3DbDZUTIg
[21] E. Baylor *et al.*, "A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy," Jan. 2020. doi: 10.1145/3313831.3376718.
[22] P. Zhang and M. N. Kamel Boulos, "Chapter 50 - Blockchain solutions for healthcare," in *Precision Medicine for Investigators, Practitioners and Providers*, J. Faintuch and S. Faintuch, Eds., Academic Press, 2020, pp. 519–524. doi: 10.1016/B978-0-12-819178-1.00050-2.
[23] E. S. Berner, Ed., *Clinical Decision Support Systems*. in Health Informatics. New York, NY: Springer New York, 2007. doi: 10.1007/978-0-387-38319-4.
[24] H. S. Goldberg *et al.*, "A highly scalable, interoperable clinical decision support service," *J Am Med Inform Assoc*, vol. 21, no. e1, pp. e55–e62, Feb. 2014, doi: 10.1136/amiajnl-2013-001990.
[25] B. A. Fiedler, "Device failure tracking and response to manufacturing recalls," in *Managing Medical Devices Within a Regulatory Framework*, Elsevier, 2017, pp. 263–275.
[26] J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561*, 2014.
[27] K. Bonawitz *et al.*, "Towards Federated Learning at Scale: System Design," Mar. 22, 2019, *arXiv*: arXiv:1902.01046. Accessed: Jul. 29, 2024. [Online]. Available: http://arxiv.org/abs/1902.01046
[28] H. Cai, D. Rueckert, and J. Passerat-Palmbach, "2cp: Decentralized protocols to transparently evaluate contributivity in blockchain federated learning environments," *arXiv preprint arXiv:2011.07516*, 2020.
[29] "Secret Sharing Sharing For Highly Scalable Secure Aggregation," ar5iv. Accessed: Jun. 22, 2024. [Online]. Available: https://ar5iv.labs.arxiv.org/html/2201.00864
[30] D. Pereira, P. R. Reis, and F. Borges, "Secure Aggregation Protocol Based on DC-Nets and Secret Sharing for Decentralized Federated Learning," *Sensors*, vol. 24, no. 4, Art. no. 4, Jan. 2024, doi: 10.3390/s24041299.
[31] 51 Attack. Accessed: Apr. 19, 2022. [Online]. Available: https://www.coindesk.com/tag/51-attack/
[32] "Decentralized Identifiers (DIDs) v1.0." Accessed: Mar. 19, 2024. [Online]. Available: https://www.w3.org/TR/did-core/
[33] C. Brunner, U. Gallersdörfer, F. Knirsch, D. Engel, and F. Matthes, "DID and VC:Untangling Decentralized Identifiers and Verifiable Credentials for the Web of Trust," in *Proceedings of the 2020 3rd International Conference on Blockchain Technology and Applications*, in ICBTA '20. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 61–66. doi: 10.1145/3446983.3446992.
[34] R. Haakegaard and J. Lang, "The elliptic curve diffie-hellman (ecdh)," *Online at https://koclab. cs. ucsb. edu/teaching/ecc/project/2015Projects/Haakegaard+ Lang. pdf*, 2015.
[35] "PPFL: Privacy-preserving Federated Learning with Trusted Execution Environments," ar5iv. Accessed: Jun. 22, 2024. [Online]. Available: https://ar5iv.labs.arxiv.org/html/2104.14380
[36] T. Geppert, S. Deml, D. Sturzenegger, and N. Ebert, "Trusted Execution Environments: Applications and Organizational Challenges," *Front. Comput. Sci.*, vol. 4, Jul. 2022, doi: 10.3389/fcomp.2022.930741.
[37] "Enterprise Ethereum Alliance Off-Chain Trusted Compute Specification v1.1." Accessed: Jun. 22, 2024. [Online]. Available: https://entethalliance.github.io/trusted-computing/spec.html
[38] "Taho," GitHub. Accessed: Aug. 11, 2024. [Online]. Available: https://github.com/tahowallet
[39] F. Malandrino and C. F. Chiasserini, "Federated Learning at the Network Edge: When Not All Nodes Are Created Equal," *IEEE Communications Magazine*, vol. 59, no. 7, pp. 68–73, Jul. 2021, doi: 10.1109/MCOM.001.2001016.
[40] *pluralitybook/plurality*. (Jun. 22, 2024). Jupyter Notebook. Plurality: The Future of Collaborative Diversity and Democracy. Accessed: Jun. 22, 2024. [Online]. Available: https://github.com/pluralitybook/plurality
[41] Optimism, "Retroactive Public Goods Funding," Optimism PBC Blog. Accessed: Mar. 02, 2024. [Online]. Available: https://medium.com/ethereum-optimism/retroactive-public-goods-funding-33c9b7d00f0c
[42] "Gitcoin | Fund What Matters To Your Community." Accessed: Jun. 22, 2024. [Online]. Available: https://www.gitcoin.co/
[43] "Partial Common Ownership," RadicalxChange. Accessed: Jun. 22, 2024. [Online]. Available: https://www.radicalxchange.org/wiki/partial-common-ownership/
[44] P. Cuffe, "The role of the erc-20 token standard in a financial revolution: the case of initial coin offerings," in *IEC-IEEE-KATS Academic Challenge, Busan, Korea, 22-23 October 2018*, IEC-IEEE-KATS, 2018.
[45] *gov4git/gov4git*. (Jun. 18, 2024). Go. Gov4Git Foundation. Accessed: Jun. 22, 2024. [Online]. Available: https://github.com/gov4git/gov4git

[46] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”.
[47] E. G. Weyl, P. Ohlhaver, and V. Buterin, “Decentralized Society: Finding Web3’s Soul,” May 2022, Accessed: Jun. 22, 2024. [Online]. Available: https://www.microsoft.com/en-us/research/publication/decentralized-society-finding-web3s-soul/
[48] “Home Page,” P2P Foundation. Accessed: Aug. 11, 2024. [Online]. Available: https://p2pfoundation.net/
[49] A. S. Cheung, “From Data Subjects to Data Sovereigns: Addressing the Limits of Data Privacy in the Digital Era,” in *Data Sovereignty: From the Digital Silk Road to the Return of the State*, A. Chander and H. Sun, Eds., Oxford University Press, 2023, p. 0. doi: 10.1093/oso/9780197582794.003.0005.
[50] J. Ernstberger *et al.*, “SoK: Data Sovereignty,” *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*, pp. 122–143, Jul. 2023, doi: 10.1109/EuroSP57164.2023.00017.
[51] “The Network State: How to Start a New Country.” Accessed: Aug. 04, 2024. [Online]. Available: https://thenetworkstate.com
[52] “Plurality.Institute.” Accessed: Aug. 11, 2024. [Online]. Available: https://www.plurality.institute/
[53] “Home,” ethereum.org. Accessed: Jul. 01, 2024. [Online]. Available: https://ethereum.org/en/
[54] “Start | Kernel.” Accessed: Jul. 28, 2024. [Online]. Available: https://www.kernel.community/en/start/