Doctelligence: A Decentralized Intelligence Health Network

Abraham Nash

abrahamnash@protonmail.com

Abstract: Currently, health data is fragmented and siloed across providers and institutions. Such fragmentation of health data is problematic because of identification errors, problems with retrieval, and difficulty in training and validating artificial intelligence (AI) tools in healthcare. To address these problems, this paper outlines a novel framework that integrates 1) self-sovereign identity architecture coupled with a personal health record ; 2) a federated learning protocol on a public blockchain that facilitates the training of AI on personal health records; 3) a trustless rewards mechanism that incentivizes participation in the network, whilst ensuring the fair distribution of rewards. A patient receives these rewards into their digital wallets as an incentive to opt-in, with a long-term roadmap to funding decentralized insurance solutions.

Introduction

Access to accurate and up-to-date patient health data is fundamental in delivering efficient, safe, and quality health care to patients, critical for research (e.g., advancing our understanding of a disease), and developing new tools for diagnosis and creating new treatment pathways (e.g., with AI decision support tools). Currently, there are a number of issues surrounding the collection, storage, retrieval, maintenance, ownership, access, and use of health data. Most of patient health data is fragmented across healthcare providers due to the lack of standardization in representing patient identities and inconsistencies in how such data is collected across different health IT systems [1]. Such repetitive and over collections of health data are inefficient and tedious, creating fatigue for patients and verifiers (e.g., physicians, nurses, etc.) and error-prone in the verification process. According to a previous study, 10 out of 17 medical errors that contribute to approximately 195,000 deaths per year are related to the incorrect identification of patients in the US alone [1]. Furthermore, digital records of health data, known as electronic health records (EHR), are stored in siloed databases mainly due to institutional interests in shareholding and using health data as strategic assets [2]. Such large institutional databases constantly face security risks as a main target of hacking [3]. And the lack of EHR interoperability reinforces health data's siloed nature, hindering AI practitioners from training healthcare AI-tools and enhancing the delivery of healthcare pathways, such as the implementation of diagnostic tools and decision-support systems [4]. Furthermore, individual patients are excluded from the current ecosystem of valued health data exchange. Without ownership (i.e., access and management control) over their health data, patients have little input over the use of personal health data generated about them and receive no financial return when it is being used by third parties.

To address the issues caused by the fragmented and siloed health data, decentralized identity management (DecIdM) systems seek to remove much of the dependency on single third-party mediators between users and their requested service providers by building on top of decentralized blockchain technologies [1]. A blockchain is a record of computed blocks of information linked together using cryptography, each containing a cryptographic hash of the previous block, a timestamp, and transaction data [5]. DecIdM leverages the blockchain to return the ownership, management, and access control of identities and identifiers alike to the original, individual identity hosts [1]. Patients might be expected to install a digital wallet on their personal devices to create their blockchain-based IDs, which can then be used to communicate with the rest of the network using self-sovereign server technologies. Decentralized identity management also enables the creation and using blockchain-based identities to credential physicians by using a third-party (e.g. doximity) to match the credential with a public/private key pair [6]. Self-sovereign server technologies (SSST) consume only the blockchain-based ID and use it to secure and manage access to patient data located in PHR systems [6]. However, if decentralized identity architectures are to replace their centralized counterparts, a decentralized FL (federated learning) environment that preserves access to this information for learning is required. In FL, a blockchain can alleviate the fundamental limitations of trust, resiliency, and accessibility to FL protocols for learning on personal health data [7].

In addition, the blockchain can be directly used to address the need for a trustless rewards mechanism to bind participation in a decentralized network. Thus, for FL to be deployed in decentralized identity systems, it needs a reward mechanism, and this is an integral component of quantitative and evaluative enquiry in this field [8]. However, a specific utility of a reward mechanism in healthcare remains to be addressed.

Though FL techniques were initially developed to run at edge devices (e.g., mobile phones, laptops), research effort in the healthcare domain is focused on running on multiple siloed databases to preserve the interests of hospital networks that deploy centralized identity management systems. In fact, a workshop paper published by Consensys (an enterprise offshoot of Ethereum) specifically outlined an enterprise architecture to foster FL across healthcare consortia [9]. This is based on a centralized identity architecture still being employed, and in reality, preserves the interests of institutional stakeholders by minimizing the risk of financial penalties from data breaches [10]–[13]. As mentioned before, cross-silo architectures in healthcare regularly face security risks posed by maintaining large institutional stores of health data [14]–[17]. Furthermore, data owned by single institutions or a network of prepared data silos may be very homogeneous, producing over-fitted models that will be inaccurate when used on other inputs [18]. In addition, centralized healthcare provision leads to data being available to companies and researchers from relatively few institutions and geographic regions. Silos of institutional databases often contain unquantifiable bias due to their incompleteness with respect to co-variables such as co-morbidities, ethnicity, gender, and more [19]. Because of this, much of modern research enters into cross-institutional collaborations in order to retrieve health data owing to individual limitations in patient demographics. Such collaborations usually consist of centralized institutions of hospital networks (i.e., hospitals, clinics) [20]–[25], and individual patients are not included as part of the ecosystem.

Finally, if the data is decentralized and FL is is conducted on mobile devices, and a reward mechanism is programmed that enables patients to act as direct participants in a decentralized ecosystem — a secure and scalable auditing protocol is required to reach a consensus on how those rewards are distributed among participants. Currently, very few auditing protocols exist, which are often designed for siloed architectures and/or with a small number of participants. For example, one paper outlined a decentralized auditing protocol that specified that patients are not directly compensated when a hospital uses patient data, and it cannot enforce such compensation to be made since differential privacy prevents anyone from knowing whether a particular individual was included in the training dataset [26]. As a result, such frameworks work with a small number of participants, and consider every participant in a given FL round to be an Evaluator when evaluating the contributivity of each participant before a reward is allocated. For example, if the experiment includes 100 participants, then 100 participants are required to download N - 1 participant models to validate all of them to assess their respective contributivities (i.e., how good their model is). However, this is not scalable because the asymptomatic costs of validation increase in scale with a growing number of participants in the FL process. As decentralized health data mangement has the potential to overcome access to scalable health data for learning, then a scalable and secure auditing protocol to fairly allocate rewards is required.

This paper seeks to address the issues discussed above and outline a roadmap toward a learning health system by proposing a framework that integrates 1) a self-sovereign identity architecture, 2) a federated learning protocol on a permissionless blockchain, 3) a scalable decentralized auditing protocol for the rewards process. The integrated framework allows the health data to be owned by individual patients (i.e. determine access and management controls to health data generated about them), minimizing errors in the identification and verification process to enhance the safety and efficiency of healthcare and, at the same time, facilitate access, training, and validation of AI tools over individually managed and controlled (sovereign) health data.

In the framework proposed by this paper, an FL protocol is decentralized onto a public blockchain so that no authority assumes responsibility for accessing learning on patient health data. This is not possible with cross-siloed institutional FL, as a central authority employs a centralized identity management system to oversee the interests of each hospital, and subsequently to orchestrate the learning process between each siloed database. As healthcare providers remain competitors, a private blockchain that is permissioned gives third parties the ability to pick and choose who has access to siloed patient data for learning, which does not meet the conditions of the framework proposed by this paper. In fact, all examples of FL currently utilize a permissioned blockchain mimicking their centralized counterparts as stakeholders in the real world [20]–[25].

Instead, an FL protocol is written as a set of rules onto an *intelligence* smart contract and published (distributed) onto a public blockchain that orchestrates the coordination and rewards process. A public blockchain preserves a “trustless" system — meaning anyone can access the network to conduct training that preserves decentralization and avoids points of centralization. AI practitioners, or else, any entity seeking to train their models on patients' health data, can access an FL smart contract to orchestrate machine learning on stores of patient data via distributed protocols on a public blockchain. As data never leaves patients' PHR store, a patient only needs to opt-in to set the permissions of access once using a decentralized identity as self-sovereign technology to manage their permissions, instead of every time they receive healthcare.

In exchange for offering access to AI learning on their personal health data and computational resources, a cryptographic micropayment (e.g., Dai, etc) is transacted into a patient's digital wallet (without a third party requirement). A financial micro-payment is used to incentivize opt-in, with a long-term road map to the funding of decentralized insurance solutions. An *intelligence* smart contract protocol handles the FL coordination and rewards process, and specifies for a scalable and secure decentralized auditing protocol to ensure patients receive a fair allocation of their rewards. This is to protect against malicious participants that do not follow the protocol, and so a threat model needs to be in place to hold participants *accountable* for their contributions.

Evaluators assure this accountability, and stake an token (e.g. ERC-20 native token, etc), to gain the right to evaluate patient models in the rewards processes in a proof of stake (PoS) ecosystem. Evaluators found to be acting maliciously are slashed from the network and can loose some or all of their stake, and this maintains the security of the network. As an example, an ERC-20 is the technical standard for all smart contracts on the Ethereum blockchain and acts as a store of value that can be sent and received [27]. It provides a list of rules that all Ethereum-based tokens must comply with, such as how to transfer the tokens, and how transactions are approved [27]. Ethereum is an open-source blockchain platform with smart contract features that support Turing-complete operations [28]. This enables programming features capable of solving any computation problem that is selected. Furthermore, layer 2 solutions are building off of this decentralized blockchain protocol to offer the possibility of lower/non-existent transaction fees and/or privacy preserving transactions without compromising the fundamental capabilities of decentralization and consensus [29], [30]. Ethereum and/or it's layering solutions can thus be used for a wide (and open-ended) set of capabilities relevant to healthcare applications in this network [6]. This provides an environment to develop AI tools that are updatable and adaptable, suited for a scalable provision that can be used by healthcare professionals in their day-to-day practice.

This framework comprises three layers: A root layer, a real-world layer, and an intelligence layer. The root layer section provides an overview of self-sovereign server technology (SSST). Secondly, the real-world layer section outlines the flow of information in the delivery of healthcare provision. This is important to define as the utility of the reward mechanism which drives a closed feedback loop of value to foster the development of AI tools in healthcare, which can be potentially transformative in improving the delivery of healthcare. Third, the intelligence layer outlines the *intelligence* protocol for scalable and secure trustless FL process as well as the specific utility of a reward mechanism to fund decentralized healthcare insurance solutions.

1. Root layer

1.1 Self-Sovereign Server Technology

A decentralized identity itself is quite similar to an easily verifiable and secure public-private key pair – in public-key cryptography, a pair of mathematically related public and private keys is used to create digital signatures and encrypt data. Since it is computationally infeasible to obtain the private key given its paired public key, these public keys can be shared freely, thereby allowing users to encrypt content and verify digital signatures [31]. In order to ensure that a DID exists and is unique, the public address of that DID must be registered in the underlying blockchain framework. When a DID is first generated, it is only supplied with an empty address with little meaning about an identity [1]. It is a free choice of the DID holder to include any additional aspects of their profile as necessary, and all aspects can be accessed from a single location instead of being arbitrarily disseminated [1]. These aspects represent information that can be verified by other parties in order to offer services or handle requests based on user agreements or regulations. Gender, age, nationality, occupation, and educational level are examples of aspects to be associated with a DID [6]. Decentralization aims to eliminate intermediaries between users and service providers. However, this feature alone does not resolve the issue of users feeling burdened with multiple identifiers across different services.

Without standardization, DID (Decentralized Identifier) representations would remain implementation-specific, leading to incompatibility among DIDs generated by various systems. In such a scenario, a limited number of services could eventually monopolize the DID market.

ortunately, a community of experienced researcher teams, including Microsoft, IBM, MasterCard, and others, have developed open standards (e.g., W3C DID [1]) to guide services in designing new DID models and their operations. These standards focus on a multi-level and multi-module DID structure to accommodate the varying volumes and types of information associated with identities. An example from the W3C DID v1.0 standard showing the structure and a minimal set of attributes of a DID document can be found online [1].

In order to link any kind of a digital identity to its corresponding physical identity, a portion of its aspects requires an endorsement from designated authorities offline. In using a DID, this is commonly known and referred to as "verifiable credentials for the web of trust" [1]. In other words, the remaining aspects of a DID must be validated by other parties who can be held accountable for their authenticity [1]. In other words, aspects of a DID that involve official reputation, such as legal name and date of birth, must be directly attested by authoritative agencies, while other aspects can be verified by other DID holders who have been directly or indirectly endorsed by authorities. As proof of verification, a digital certificate is created and signed by the verifying DID holder. The certificate can then be shared or used to build a chain of related certificates with non-repudiation [1].

To attribute decentralized identities with various attributes, a sovereign technology can be used by an individual patient. A sovereign technology can be defined as one that does not offer a privacy policy because there is no counter-party to have a privacy contract with [32]. Examples of sovereign technologies are UMA authorization servers that fulfill the storing and assertion of policies, or the implementation of FIDO standards for the authentication of the owner of the sovereign technology (e.g., a password, biometrics), etc [32].

Therefore, a Self-Sovereign Server Technology (SSST) describes an identity container that uses a combination of a mobile user interface that controls the identity and an always-connected server that stores attributes, policies, and transaction receipts associated with that identity [33]. Attributes, typically patient health data like the contents of the prescription, are meant to be selectively shared. Policies are kept private, but they control external access to attributes. Receipts are the signed result of transactions stored in case of audit or dispute. As shown in Figure 1, an identity container comprising a hardware device, cryptographic identity, and server technology allows each participant (i.e., physician and patient) to interact [33]. In summary, the identity container is capable of storing attributes (e.g., personal health records), creating and/or interacting with policies (on-or-off chain), and storing receipts in health care provision.

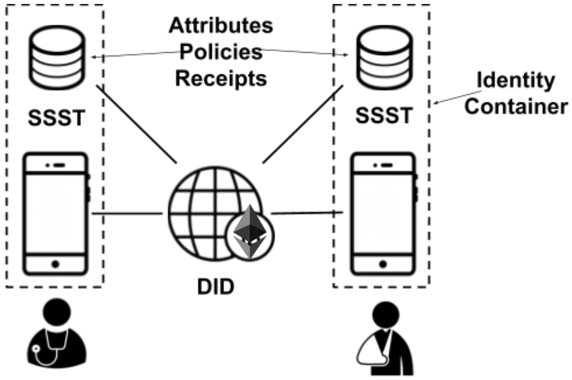

Figure 1 A transaction between a physician and a patient. Adapted from HIE of One [33].

Self-sovereign server technology in the set-up of health care provision as established by patients and physicians [33]. Public/private key pairs link each participant's identity store to a public blockchain ledger and this allows each stakeholder to enter themselves into a decentralized network using their local devices (e.g., mobile, laptop). A professional entity (e.g., physician) in control of their private key authenticates themselves into the network using a private key associated with a trusted reputation mechanism to prove their professional qualifications, i.e., license to practice [33]. For example, this could be a credentialing/licensing agent accepted in the local healthcare ecosystem, e.g., a medical licensing body, or medical society [33]. Physicians use sovereign technology to secure themselves in the authentication process, such as signing, encryption, payment keys, and biometrics (e.g., fingerprint) via their user interfaces [33]. Patients have the option to integrate their own selection of open-source servers to store their personal health records system and health information and to store their permissions of access on- or off-chain [33], [6].

A user interfaces application to manage DID keys, biometric access controls, and key recovery is made publicly available as open-source technology by a variety of vendors (e.g., SSST-built from the source) [33]. An open-source personal health record system is an integral component of SSST in the set up of 'sovereign' healthcare provision. To do this, a physician uses a sovereign (cryptographic) identity to verify a relationship with a patient, authenticate themselves, determine permissions for personal health records access, and also enable the ability to access wider services in the provision of health care - as captured by Figure 1 [33], [34]. Other healthcare services (e.g., pharmacies, laboratories, radiological services, etc.) can similarly adopt their own decentralized identifier and sovereign technologies to interact with a patient's SSST to access their personal health record in the delivery of healthcare. In a decentralized health data management setup, all such entities in a care pathway can potentially adopt self-sovereign support technology support depending upon their preferences to personalize their security, privacy, and economic interests when participating in a network [33].

1.2 Personal Health Record

A patient's sovereign server technology (SSST) can provide privacy-related server software which connects physicians to read and write information into their health records (PHR). Patient health record stores are accessed through a patient's SSST which can be held over their lifetimes. This is possible as a decentralized identifier (DID) has the potential to enable a lifelong practical and reliable identifier and attributes linked to that identifier under the self-sovereign control of the individual patient [1].

Access permission policies can be stored on a server as off-chain permissions [33]. However, this has also been demonstrated using access permissions stored on the blockchain, including specific custodianship-based relationships [6], [35]. It is important to note, that writeable permission of the patient into their own records requires situational judgment. On the one hand, a patient has more incentive to ensure accurate medical records than physicians do [36]–[38]. On the other hand, there are specific situations of custodianship which need to be considered (e.g., HIV, Psychiatry, Paediatric Parent Custodianship, Power of Attorney, etc). However, it has been shown that incorporating patient-based feedback increases the accuracy of physician-recorded health information and improves healthcare provision in the long run [37], [39]. Additionally, decoupling centralized identity management from health record systems creates the potential to unify a common format and store in which to record health care information in the delivery of healthcare provision.

It is possible to couple a SSST with an open-source personal healthcare record [38]. Both the physician and patients are provided a user interface for their healthcare which can be coupled as a software component as a part of their SSST. A clinician accesses a patient's health record through a one-way API portal governed by off-chain access permissions, as specified by the patient [33],[34]. As shown in Figure 2, an API is used by physicians in accessing and writing into a patient's health record [40]. A secure off-chain policy store can be implemented using authorization server components [33]. Other access permissions can specify for data exchanges to maintain policies of convenience (e.g. auditing, research) [33].

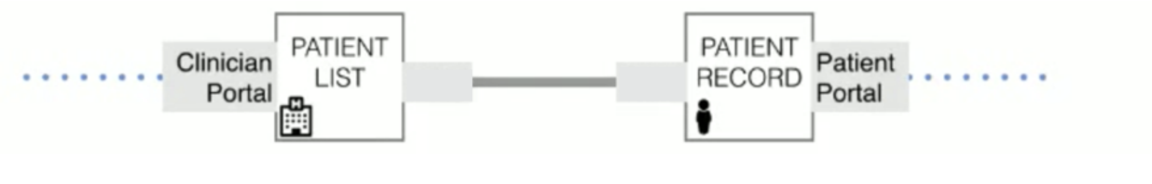

Figure 2 API access to personal health record. Adapted from NOSH & HIE of One [41].

Other practical solutions include the ability for on-chain policy stores which handle access to PHR stores, where viewing permissions may be recorded as encrypted pointers stored on a public blockchain ledger [6], [35]. For example, a mechanism called “sign then encrypt” can theoretically be used to govern access permissions to a patient’s health record on a public blockchain network [6].

In both scenarios (either on- or off-chain), once permission is granted, a physician may request access by signing with their public/private key pairs. When an issuer's signature is certified, the client (i.e., patient) then checks the blockchain contracts and verifies if the address issuing the request is allowed access to the query string. If the address checks out, it runs the query on the client’s (i.e. patients) preferred server, which grants access to their health record and returns the result over to the client (i.e. API clinician-portal) [6]. A data requester's access to a resource can be approved or revoked at any time by the patient as a state update in the policy store where all permissions (e.g., reading, writing) are logged [6]. This means that the acceptance, rejection, or deletion of access is controlled by the patient, though a physician can also send requests and delete their affinities to patients [6]. These processes of client checks can be automated by server components of a patient's SSST.

A patient-centered holding of personal health information using on or off-chain policy stores for permissions to determine access to personal health records held on a patient’s SSST is possible and demonstrated [40], [41]. This paper maintains the code of patients as owners of their own personal health information and acting as sovereign agents in accessing their own health networks. In this instance, a programming language such as solid pods… can be used to store and self-determine who has access to this data….

2. Real-World Layer

2.1 Patients and physicians

The primary architectures of authentication and the setup of healthcare provision surrounding patient-physician relationships in this network, are based on the principles of sovereign identities on the blockchain [33]. Such architectures allow patients to interact directly with physicians and other healthcare services to facilitate their healthcare provision. An overview of patient-physician interactions for patient-centered decentralized healthcare provision is outlined as the following:

A physician first attributes a cryptographic identity (i.e. public/private key pair) to their professional credentials/licensure, whereby the physician's license is confirmed by a medical society and/or a renowned third party (e.g. doximity) [41]. The physician's public/private key pair is attributed to these credentials using an identity management system (e.g., uPort) [41]. A physician can access their patient’s personal health record by clicking on the link to the patient’s PHR, which will prompt the physician to sign on with their decentralized identity (DID) [41]. The patient’s authorization server component of their SSST looks up the physician's medical license number attribute via their DID and verifies the status against the state medical society’s physician directory server [39]. The physicians’ own SSST authorization server component can determine policies of access to the directory to preserve their privacy. It is important to note that the medical society directory service is not self-sovereign technology [33].

A physician who is authenticated can subsequently use their public/private key pairs to sign and issue services for health care provision (e.g., pharmacy, imaging, referral). These can be deposited as a transaction into the patient's health record [5]. As shown in Figure 3, it displays the scenario of a patient requesting a pharmacy prescription to demonstrate an authentication architecture in practice. A physician sends and signs their public/private key pair along with their license/credentials to an agent who confirms the validity of these attributes (e.g., medical society, renowned third party, etc.) [41], [42]. A physician may then sign their service requests to be stored in the patient's health record to use in a sovereign manner in healthcare provision. The patient signs in to the service with their DID credentials. The DID then verifies that the patient that signed in to the pharmacy matches the patient in the prescription. A patient-centric architecture enables a patient to set up healthcare provisions without relying on an institutional centralized identity provider (e.g., a large EHR provider) [33].

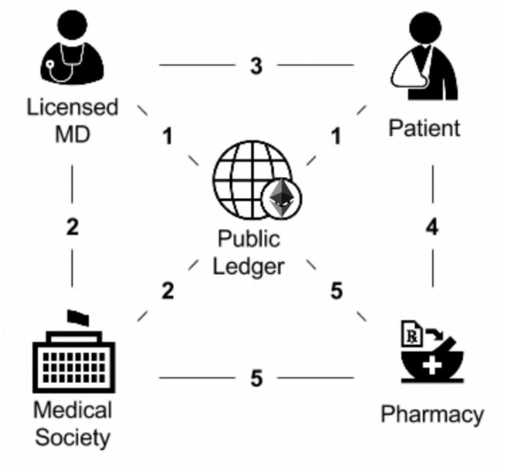

Figure 3 An example of patient use of a pharmacy service. Adapted from HIE of One [33].

This same architecture is true for all healthcare services (e.g., referrals, pharmaceuticals, immunizations, medical investigations, imaging services, consent forms, and more). A healthcare service may use its sovereign servers to interact with a patient's SSST in the network and to talk to the server components of institutions issuing licensing and credentialing attributes to a physician's cryptographic identity. A time limit may be set before re-authentication of identity is required to maintain access to a patient's health records, with a security policy store existing on- or off-chain [33]. In practice, these steps may be automated by server components of each physician using their SSST. Success will mean that no service has to trust a centralized identity provider for the patient or physician, as long as the jurisdiction they are in recognizes the authority of blockchain-based identity and reputation-based agent (e.g., medical society) [33].

2.2 Systems

This paper defines s*ystems* as those which store, retrieve, generate, send, and process health data (e.g. triage tools, diagnostic tools, AI decision support systems, etc.). A human-centered evaluation of implementing AI systems in clinical environments is an open area of research [35]. Beede [2020] et al. began in this direction by conducting a human-centered study of a deep learning system used in clinics for the detection of diabetic eye disease. The study characterized current eye-screening workflows, user expectations for an AI-assisted screening process, and post-deployment experiences in a traditional systems design (i.e., EHR and hospital provider). However, previous works have highlighted several obstacles in going from research and development environments to the hospital or clinical settings [36]. For example, these obstacles include the frequent lack of utility to clinicians and logistical hurdles that slow or block deployment [37]. AI tools are key systems for integration into clinical environments, however, implementing these systems is currently difficult as health data is stored across different provider silos. For example, these AI tools typically makes use of machine learning (e.g., multi-class classification) algorithms in order to predict the likelihood of an outcome (e.g. diagnosis) based on numerical data in a case (e.g., demographic data, clinical history, etc.). AI predictive models ideally require one source of truth to ensure that the machine-learning (ML) models are using the most accurate and latest data [6]. Given these restrictions, the immediate focus should be on possible interventions for changing the architecture of the system by integrating appropriate methods, and technological actions to provide quality health care [6]. For example, it is already known that recording health data in a single location allows for better coordination and use of AI tools [38]. In the framework of this proposal, all a patient's health data is recorded into one location (i.e., the patient-owned PHR). This overcomes the disputes between healthcare providers in accessing health data, and secondly the outputs generated are done so back into one location (i.e., the patient-owned PHR). From this perspective, AI tools are natural subjects that can benefit from the use of DIDs [1]. It is currently a relatively unexplored area of research; however, it is central to consider these systems in the context of the proposed framework, whereby AI tools use their own DID and read/write into a patient-owned PHR. A DID attribute to each device would correspond to a persistent and unique identity profile like any human user or entity. A software counterpart can be used to consume this DID so that its medical staff user can link its associated device with the appropriate patient being treated [1]. The pairing can be captured as an aspect or a credential of the DID profile and becomes easily verifiable [1]. An extension of this research inquiry is to study the operational contexts of integrating AI systems into clinical workflows according to the proposals framework, in the human-centered context of using patient-owned PHR systems: for example, in accessing authentic health data for AI systems functionality, attributing DIDs to AI systems, and in allying physicians with AI technologies, etc.

Most medical devices today are designed to have the ability to store control software and data. From this perspective, medical devices are natural subjects that can benefit from the use of DIDs [6]. These systems can therefore use their own DID and read/write into a patient's health record. The recording of patient health data into a single location (i.e., a patient's PHR) allows the use of systems to better coordinate and record personal health data for the provision of health care. AI tools to work by running their algorithm's on data recorded into a patient-owned PHR as shown in figure 4, by physicians (e.g. symptoms, signs, medical history, physical examinations, etc) and in combination with other inputs, such as investigations (e.g. imaging, histopathology, mobile health, etc). An AI system attributed with a physician then accesses the PHR to record its output into the patient's PHR. Such partnerships with machines support physicians and patients to enable open dialogue in steering healthcare together (e.g., expenditure, treatment choices, etc.). Other AI tools and services that work outside the PHR can improve the accuracy of specific tasks in the health care pathway (e.g., automated imaging diagnostic tool, mobile EKG device, omics studies, etc.).

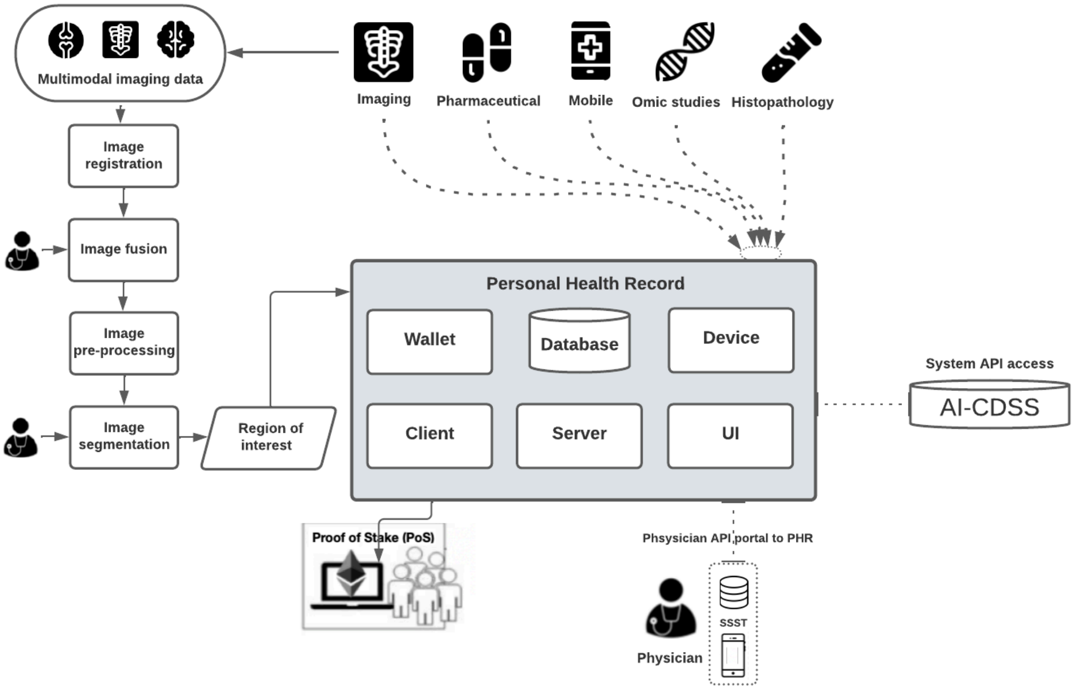

Figure 4. Interactions amongst physicians and supporting systems that read and write directly into the personal health record are mediated by the blockchain (e.g. Ethereum).

A decentralized identity assigned to each device would correspond to a persistent and unique identity profile similar to any human user or entity. A software counterpart can then consumes this DID so that its medical staff user can link its associated device with the appropriate patient being treated [1]. The pairing is captured as an aspect or a credential of the DID profile and becomes easily verifiable [1]. Additional information can be certified by one or more medical professionals too, for example, to specify a medication type and dosage to be delivered.

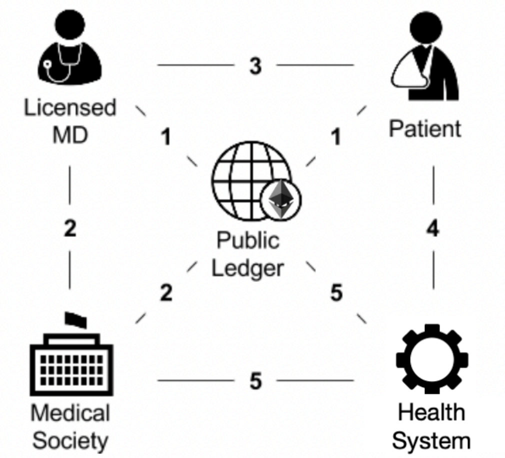

Figure 5 *An example of patient use of a system [2].*

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 5. 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 [43], although many open-source systems do not require licensing and authentication in practice. An open-source PHR system is well-suited to an ongoing prototyping between physicians, patients, developers, and other key stakeholders for adapting a PHR database for training and deploying AI tools.

However, if decentralized identity (DID) architectures are to replace their centralized counterparts, a decentralized FL environment that preserves access to patient-owned health data for AI tool development in the first place, and that simultaneously maintains patient ownership over their health data, is required. This leads us to our next layer, the intelligence layer.

3. Intelligence Layer

3.1 Federated Learning

Federated Learning (FL) is a distributed machine learning approach that enables training on decentralized data stores, on devices such as mobile phones. FL is a more general approach to learning on local data stores and addresses the fundamental problems of privacy, ownership, and locality of data [44].

In the traditional FL paradigm, a central server coordintes the aggregation of local model updates to update a global model. However, some potential problems exist in these approaches such as dishonest aggregation, accidental network connection failure, unexpected external attacks, etc [44]. Furthermore, the large number of participants in FL makes it difficult to ensure that all clients are honest and will train the local models according to the predefined FL protocols [44]. Therefore, there may exist dishonest clients/agents (e.g., individual patients) submitting false data about their local training results [44]. Also, in the traditional FL, there is a lack of incentives that encourage clients to follow the protocol honestly and provide reliable data—clients are contributing their computing powers without any rewards in return [44].

In the architecture proposed in this paper, a patient uses a decentralized identity (DID) stored by their SSST to opt into the *intelligence* smart contract protocol distributed on a public blockchain. Storing FL protocols on a public blockchain enables an immutable ledger to openly structure how patients' data is being used, while individual personal health information remains sovereign to each patient on their devices and is never exchanged. However, the main advantages of the blockchain in FL are computational in that it enhances round-delineation, model selection, and model aggregation and preserves these aspects in a decentralized manner [7]. A distributed consensus protocol enables transparency, fairness, and impartiality, enabling protocols of FL to build trust with patients [7]. Second, fault tolerance is innate to the blockchain owing to its peer-to-peer (P2P) design, provides resiliency in computation, and improves the integrity of the entire system. It can use smart contracts (SC) to orchestrate rounds for multiple FL tasks simultaneously from different sets of devices [7]. The need for an incentive mechanism for patients to opt-in to these FL protocols is also addressed, which is an integral component of quantitative and evaluative inquiries in the field of FL [8]. Most importantly, the architecture proposed in this paper is able to use a public blockchain to orchestrate a "trustless" process in the coordination and rewards process without a third party [7], [45]. Otherwise, a private blockchain relies on a trusted setup in which the orchestrator can collude with the Model Owner to act maliciously, especially in the case of unfairly favoring the rewards process e.g. stealing some or all the reward.

It is acceptable as a generalizable framework for Model Owners to make use of third-party services and/or their own services in the aggregation process itself. The important aspect is ensuring that access ownership for the provision and the FL process remains sovereign in the hands of the patients: this holds true so long as two specific elements (i.e. coordination and rewards) of orchestration are handled on a public blockchain. Another key area is to establish clear privacy-preserving techniques in the update and aggregation process (e.g., differential privacy, etc) to preserve scalability and enable the anonymity and privacy of patient's health data during the learning process. Differential privacy, the process of adding random noise to model parameters can prevent these attacks [46]. By the definition of differential privacy, with extremely high probability, N−1 colluding agents cannot infer any information about the remaining agent (when (∀N≥3)N−1>N2) [52].

Other proposals do exist that outline frameworks for FL over a public blockchain which could potentially meet the requirements of the proposed framework, specifically in the aggregation process. IPLS collectively trains a model in a peer-to-peer fashion without the assistance of a server in the aggregation process using an IPFS-based protocol [45]. In contrast to the centralized setting, where only the server is responsible for storing, updating, and broadcasting the model to the participating agents, the model is split in multiple partitions that are replicated on multiple agents [45]. However, the disadvantage of this framework is the extensive expertise required to deal with a variety of model types and model compression techniques, which makes it difficult to train more complex algorithms in the healthcare domain. Vincent [2020] et al. proposed "Blockchain-aided Federated Learning" (BC-FL) to learn over a public blockchain to replace the need for a central server in the aggregation process [7]. It considers that local model updates can be received by miners through a gossip protocol over the P2P network [47]. However, it has been shown that gossip-like protocols are notorious for diverging from the real value and not reaching consensus at all [48]. Ramanan [2020] et al. proposed "BAFFLE', an aggregator-free Fl protocol to replace the need for a central server during the FL process [49]. However, this requires splitting up and compressing machine learning models on the blockchain itself, and this encounters a large number of problems in relation to the complexity of model compression techniques, and the research required to make this possible is complex.

A disadvantage of the blockchain in the FL process is that the network topology may affect the performance of the learning process [49]. However, in the real world, this level of delay may be considered acceptable. Another key consideration is that the blockchain can be architected as a private chain that uses a central authority responsible for orchestrating FL protocols (similar to central servers) with the ability to select participating clients (e.g., hospitals in a network). Therefore, a public blockchain ledger is vital to the proposed framework to overcome the competitive interests of institutional stakeholders and the tendency of re-centralization of health data.

3.2 Intelligence Protocol

A patient opts in using their decentralized identity (DID) to distribute FL protocols as defined by the *intelligence* smart contract on the Ethereum public blockchain. An *intelligence* protocol uses a public blockchain to orchestrate the coordination and rewards process in the training of their ML model on health data held on patient-owned health records (e.g., NOSH). By ownership, this means the ability for patients to self-govern management and access control to their health data using their own SSST as described in Section 1. AI clinical tool developers or any entity seeking to train their models on patients' personal health data utilize the *intelligence* smart contract to specify the FL coordination of participation across a number of rounds of training. This interacts with a token-based smart contract on a public blockchain which is rewarded to patients based on specific value-based measures (e.g., model contributivity scores).

An on-chain FL protocol coordinates with an off-chain distributed file storage system called the InterPlantary File Storage System (IPFS) [50]. This optimizes the cost of participation, as it provides a location to upload and download model updates in the learning process. Ethereum measures computational costs in terms of gas, which incorporates the amount of storage used and CPU instructions executed [26]. While the amount of gas is independent of market conditions, such as the price of gas or the exchange rate of Ethereum, it is directly proportional to real-world costs [26]. Combining an off-chain decentralized file storage service, therefore, lowers the cost of participation and maximizes computational efficiency - additionally, there are numerous decentralized scaling solutions (e.g. layer 2s, etc) available in the Ethereum ecosystem. It is also assumed that clients' cost of computation is likely to be relatively small compared to the value of their personal health data, so every client is incentivized to train on their entire dataset in each iteration in order to maximize expected performance and reward [26].

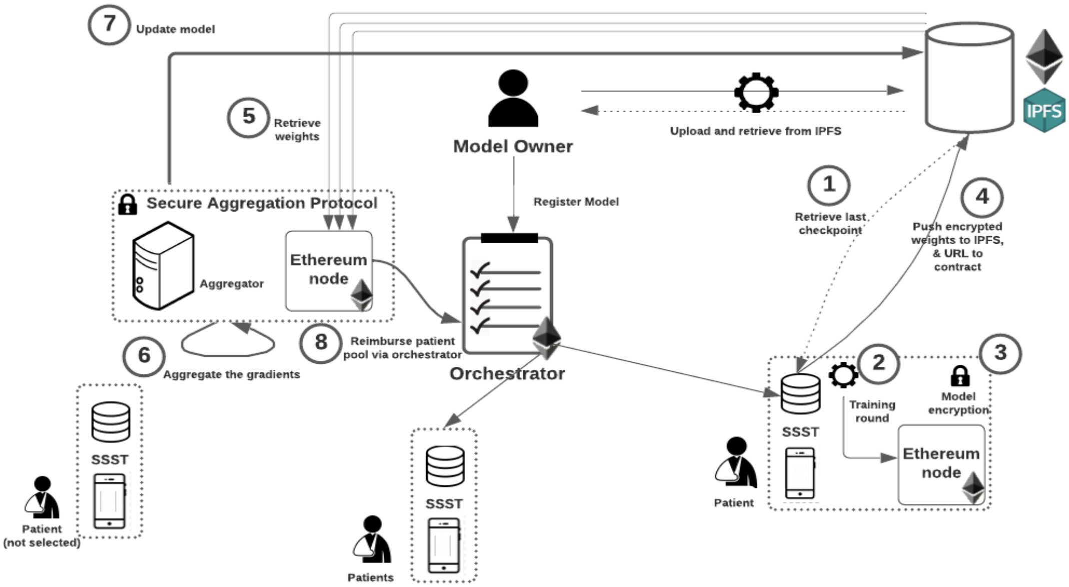

Figure 7. Global overview of a training round. Adapted from Consensys [9].

A Model Owner (e.g. AI clinical tool developer) deploys the *intelligence* smart contract to the blockchain, creates a genesis model, uploads it to IPFS, and records its CID on the contract. The *intelligence* smart contract recieves a reward from the Model Owner as a deposit which is allocated to each patient after the FL rounds are complete. A criterion by which the number of rounds are complete is set by the Model Owner who decides based upon some metric they are satisfied with (e.g. F1 score on a test set). Upon confirmation of the Model Owner's transaction (i.e. native token deposit), the patients can then see the genesis model and proceed to download it from IPFS [51]. Using their own personal health data, each patient runs their iterations of training on the model, encrypts and uploads the model to IPFS, and records their CIDs to the *intelligence* smart contract [51]. This is all done during a single FL round, and each patient then waits until the next available round. At the commencement of the next training round, the patients can see all the CIDs of the updates submitted in the previous training round. They each download these model updates from IPFS and calculate the mean aggregate of them. This is done independently, and they all arrive at the same result [51]. The patients start from their aggregate (this is considered the global model at the current training round) [51]. After the Model Owner is satisfied with the training and their criterion is met (e.g. F1 score, etc) the process is complete. Once the process is complete, the Model Owner uploads the secret IPFS CID location of their chosen test set, and Evaluators retrieve the test data set with which to evaluate patient models. A standardised fraction of evaluators to participants (e.g. 1:10) are assigned at random to each FL group to download all of the patients' models, and retrospectively evaluate patients' respective contributivity off-chain using the IPFS layer. However, an on-chain implementation can be used if any agent disagrees with the off-chain calculated scores (though this is more expensive) [51]. A select number of Evaluators are chosen by the *intelligence* smart contract to evaluate each patient's model updates. Evaluators can be anyone who has staked a native token in the network and performs evaluation duties. The results of the evaluation process are then communicated to the *intelligence* smart contract. Once they reach a consensus, a score is calculated, and each model is assigned a number of tokens, and this is recorded on the *intelligence* smart contract. These tokens (e.g. native token) are originally deposited into the smart contract pending evaluation of scores, before the training rounds begin.

3.3 Decentralized Auditing Protocol

In a trustless FL process that issues a reward (e.g., native token, etc), an evaluation metric is required to assess the relative contributions of each participant (e.g. patient). A number of works outlined various procedures for participant contributions, such as 2CP and Blockflow. 2CP employs an open-source repository, Substra, for measuring the contributivity of participants' data using a step-by-step evaluation [51], [52]. To get the contributivity score for client A using, one defines a performance metric v (eg. negative test loss, F1 scores, etc.) and records the marginal performance gain A makes to the model M in each iteration i, and sum these gains over every iteration in the training process as per: C(A) = X i v(Mi) − v(MA i+1) [51], [52]. However, the contribution procedure is performed on a small number of participants (e.g., 3-10 participants) and specifically targets cross-silos federated learning [51], [52]. Furthermore, the procedure does not elaborate on the security guarantees of the rewards process or address the issue of scalability [51]. On the other hand, Blockflow evaluate participants' overall scores, which reflect the minimum of 1) the median score reported for their model (as determined by all clients) and 2) the inverse of the maximum difference between one's reported score and the median score for each model [26]. However, similarly to 2CP, Blockflow uses a small number of participants and does not evaluate the potential for scale. This is because the underlying data-holding architecture itself is designed to consider siloed architectures only and/or with a small number of participants. In short, these frameworks consider all participants to be Evaluators in the entire FL process due to the small number of participants. For example, if an experiment includes 100 participants, then 100 participants are required to download N - 1 participant models and evaluate all of them. However, this is not scalable as its asymptomatic costs scale with a growing number of participants [26].

Instead of requiring all other participants to be Evaluators, the *intelligence* protocol delineates the role of participant and Evaluators. Instead of expecting all participants to be Evaluators themselves, the *intelligence* protocol uniquely defines the random selection of Q<<N number of Evaluators to evaluate each patient's work. And secondly, the FL process is prolonged over an increasing number of asynchronous rounds, which requires only a subset of total patients to evaluate thereby increasing scalability in large experiments and greatly reducing gas consumption on the blockchain. Whilst the role of the Evaluators does not have to the participants themselves (e.g., the patients), patients can still be considered as potential Evaluators. It is important to note that for the purposes of this paper a steady-state system is assumed, in which the number of Evaluators does not change in any FL round that is occurring. Although, in reality, the network is changing and is very dynamic, this paper assumes no disturbance in Evaluators, which should be addressed in future research.

Evaluators can be provided with a high-quality, well-distributed, and highly representative dataset to act as a control dataset. This will act as a benchmark that Evaluators can use to evaluate each patient models. This is issued by the model owner as a part of the FL process, and is a possible way of dividing the reward between patients. The reward is split between patients according to the value that each one contributes to the final performance of the model at the end of each round, i.e., the contributivity of their data [26]. This value is objective and can be calculated after the training process is complete. In this protocol, each patient contributes to the model with the knowledge that they will be fairly rewarded [26]. The datasets are evaluated retrospectively, rather than before the training process. However, each Evaluator is likely to reach different values for their scores during the evaluation process. Therefore, a patient's overall model score is going to be an average (or median, etc.) of all the Evaluators score and this is used to determine the share of the reward they will receive. BlockFlow[2021], explored a 1:1 ratio score allocation of Evaluators to participants, as each participant evaluates every other participant's score [26]. Their experiments recorded the average median agent F1 scores when varying the number of Evaluators for 1, 25, 50, and 100 agents [26]. They found that the average absolute difference between Evaluators scores at each of these levels of participation was <0.67% on their datasets (income data) [26]. However, the *intelligence* protocol defines a fraction of the total number of participants to act as Evaluators of patient scores, in order to accommodate scale over an increasing number of FL rounds. A sufficiently large number of Evaluators, with high probability, would lead to accurate results and be resistant to M<N2 malicious agents [50]. Therefore, future experiments in implementing the *intelligence* protocol should evaluate whether this difference in score allocation holds true when a larger ratio of participants to Evaluators is used, and secondly when validating for heterogeneous data sources such as health data.

The *intelligence* protocol can be specified to use any contributivity scoring procedure for Evaluators to perform their off-chain evaluations of patients' contributions. This depends on the context of the learning being conducted. All Evaluators use the CID recorded on the *intelligence* smart contract to locate and download all the patients' models off-chain (i.e., IPFS) for evaluation in a particular FL round. They perform their off-chain evaluations of each patient's model using their control dataset, encrypt their score, and report the encrypted score to the *intelligence* smart contract [26]. Each Evaluator first reports to the smart contract the set of patients whose models were successfully validated (i.e., within an acceptable bound specified by the Model Owner). Patients 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 smart contract - a basic outline of this workflow is shown in Algorithm 1 [26].

Algorithm 1:

```
for all evaluators $k \in n$ do
    $ark \leftarrow$ get patient k model’s IPFS address from smart contract
    $wrk \leftarrow$load and validate model ark
    if wrk is valid then
      report wrk as valid in intelligence smart contract
    end if
  end for
  wait for data retrieval deadline
  for all evaluators $k \in N$ do
    $vrk \leftarrow$ count number of patient models evaluators k marked as valid
    $zrk \leftarrow$ count number of evaluators who marked patient k’s model as valid
    if $vrk \leq N2$ or $zrk \leq N2$ then
      $N \leftarrow N \setminus k$
    else
      if patient i does not have valid model wrk then
        for all agents $k' \in N$ do
          $wrk \leftarrow$ request and retrieve model k from evaluator $k'$
        end for
      end if
      $sri,k \leftarrow$evaluatei(wrk)
      $bri,k \leftarrow$random encryption key
      $s'ri,k \leftarrow$encryptbri,k(wrk)
      report encrypted score $s'ri,k$ for agent k to smart contract
    end if
  end for
```

```
    wait for before encrypted score submission deadline
    for all evaluators k∈N do
        provide decryption key bri,k to provably reveal score sri,k to smart contract
    end for
    wait for score decryption submission deadline
    p← get scores from smart contract
    wr+1i←∑nk=1(wrk*pk)n*∑nk=1(pk) {all patients should now have identical wr+1i}
end for
```

Algorithm 1 - A decentralized auditing protocol for a single FL round. Adapted from Blockflow [26].

3.4 Threat Model

The intelligence protocol offers a 50% malicious threat model for validating the model integrity of patients' contributions when submitting updates to the *intelligence* smart contract during the FL process. In a proof of stake ecosystem, in a 51% attack scenario under PoS, an attacker would need to own over 50% of the total staked coins and so far this system has remained robust in protocols such as Ethereum blockchains [cite].

Firstly, in an experiment with N agents, it is resistant up to M∈[0,N2) agents neglecting to follow the protocol for the experiment to maintain its integrity [26]. Public/private key cryptography and a proof-of-stake consensus protocol secure the Ethereum blockchain [28]. 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 [53]. Secondly, as the Ethereum blockchain is public and anonymous, clients could theoretically enroll multiple times in an experiment and thus have a disproportionate vote. However, through decentralized identity verification or manual processes, agents can ensure that each other agent controls only one account [1], [6], [33].

Third, IPFS is immutable, meaning agents cannot change their model after submitting the cryptographic hash to the smart contract [50]. The *intelligence* threat model guarantees that there are strictly more than N2 honest patients in this way. Also, so long as N2 or more Evaluators who receive these models for evaluation are honest, which the *intelligence* protocol guarantees, then this remains resistant to N2 attacks. Since IPFS allows anyone to share any content, one or more honest parties would share the model with all other participants in the event that participants (i.e., patients and Evaluators) 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 [26], [50].

Fourth, there are several attacks possible on the contribution scoring procedure itself. Firstly, malicious models are defined as those which have weights that are not reflective of a truthful dataset. For example, models trained on randomly generated data or on inverted output features are considered malicious [26]. Naively averaging such models into a global model would likely harm the shared objective. The *intelligence* protocol can be used to specify contribution scoring procedures that penalize those who submit malicious models. BlockFlow[2021] demonstrates a contributivity score whereby lower scores result in less cryptocurrency received [26]. For example, in our network, evaluators can use a median score system bounded between 0 and 1, and any agents model that is evaluated to be more than 0.5 away from the median will receive an overall score of 0, and no share of the cryptocurrency pool [26]. In this case, the fabrication of scores will only penalize those who attempt it as the protocol limits a patient's overall score with the evaluation on which one was furthest away from the median [26].

Fifth, patients can collude during the training process to submit better models. For example, they can secretly share raw data or models among M<N2 colluding patients [26]. The *intelligence* protocol rewards patients who contribute strong models, and it is acceptable for multiple patients to submit identical models. Such collusion is no different from having many patients with strong datasets and is not considered to be an attack [26]. For attacks by Evaluators in the evaluation process, through encryption and a commit-then-reveal protocol (e.g., using Elliptic Curve Diffie Hellman keys), the *intelligence* smart contract prevents Evaluators from copying others' scores without collusion [54]. Or else, consider the case where a minority subset of malicious Evaluators using the BlockFlow contributivity scoring system report perfect 1.0 scores for a subset of models and 0.0 scores for all others

(e.g., models from honest agents) [26]. Since there are strictly less than half malicious Evaluators, and only the median model scores are used to determine one's overall score, the median score is guaranteed to be between the minimum and maximum scores reported by the honest agent [26]. Evaluators are incentivised to stake a native token to gain the right to evaluate patient models in the rewards processes in a proof of stake (PoS) ecosystem. Evaluators found to be acting maliciously are slashed from the network and can loose some or all of their stake, and this maintains the security of the network.

Lastly, a significant concern in federated learning is the potential misuse of the control dataset by evaluators, who risk downloading and sharing it with network participants, leading to model poisoning or unwarranted payouts for non-contributive training rounds. To address this, the evaluation of models on a test dataset must occur without actually exposing the data to the evaluators. The goal would be to ensure that the evaluators can determine the performance of each model (e.g., accuracy, precision, recall) without learning anything about the test data itself.

…Data-encryption
….Homomorphic computation
…Secure multi-party computation
…ZK-proofs
…Costs - computational costs are less of a concern to me regarding the blockchain components due to the increasing scalability solutions being made available. Work off this assumption.
… we delineate between agent and evaluator, because we cannot expect an agile system to have evaluators who meet all the same data-sepcific use cases e.g. healthcare, telecoms, etc. So, we need the test dataset to be able to specify the markup of the type of training data a particular model owner is interested in

Evaluators can prove to the system they correctly evaluated against the control dataset without accessing it, mitigating risks of misuse or leakage.

3.5 Decentralized Insurance Solution

In return for access to learning on patients' health data, it is possible to deposit a cryptographic micropayment (e.g. dai) into patients' digital wallets. A cryptographic micropayment is used by AI developers to access on-chain smart contracts which coordinate the federated learning process with patients' personal data stores.

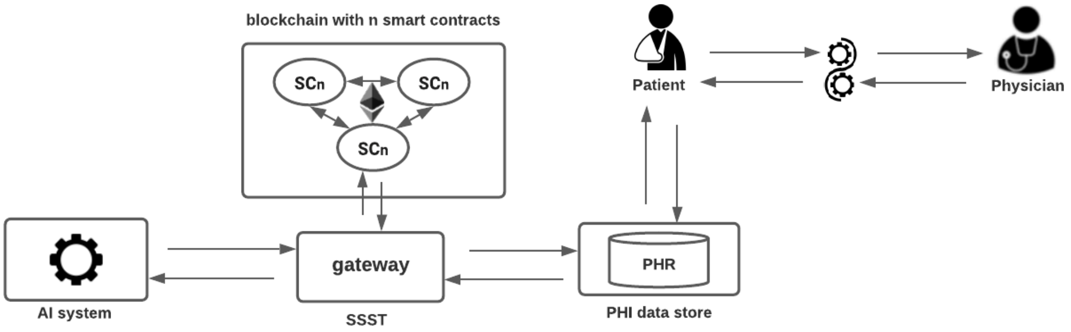

Figure 9. The rewards mechanism of cryptographic insurance in a learning health system.

A patient and physician establish the setup of healthcare and the physician records health data into the patient's PHR as described in sections 1 & 2. A cryptographic micropayment is used to access FL protocols on-chain to train a model on a patients health data for AI tools in healthcare. Micropayments (e.g. dai) are deposited directly into patients' decentralized identity wallets (e.g., Tally Ho, etc.) upon successfully submitting a model update to an FL round. A patient uses their financial reward to purchase decentralized insurance premiums, which fund and/or subsidize the provision of health care, and so the cycle continues.

This looped-mechanism enables a transition into a learning health system as tangible value is recirculated in the form of new health system technologies embedded into models of health care as described in Section 2. It is then used by patients to fund provision and this re-supplies value-based health data into a patient's PHR, which in turn, provides value-based health data to train AI tools in health care. While the formatting standards for health data entered into the record by physicians and systems remain inconsistent in centralized EHR systems and silos, patients can use an open-source personal health record (PHR) that consistently maintains these standards to meet criteria of compatibility with training on the network, ensuring compatibility across the network. This equips AI tools to learn from and inform patient event outcomes, preserving a patient-centered self-sovereign architecture and providing these systems permissionless access to up-to-date personal health data.

In addition to a value exchange for the use of a patient's computational resources and access to their health data, funding is required for health care services and treatment. A learning health system utilizing a fungible asset to resupply this value can simultaneously develop new systems technologies which can be used to improve models of health care provision. For the provision of decentralized healthcare itself, there is no incentive required - however, it is required to introduce a learning health system that creates value to fund decentralized insurance solutions for health care provision as described in section 3.2.

In FL 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 [26], [54], [55]. Patients with ongoing and/or chronic healthcare conditions are likely to have more health data in their PHRs, as they receive more healthcare. In turn, they are more likely to receive greater rewards in the learning process. Furthermore, learning on health data is not limited to one occasion or use-case, but multiple occasions and use-cases. This is suitable, as patients with ongoing conditions are likely to require more expensive healthcare insurance premiums.

The automated verification process of health care insurance claims can be handled better on the blockchain as it provides a reliable source of information with which to verify information and insurance credentials [1]. Decentralized insurance protocols are well suited to enhance the functions of more cost-effective and reliable coverage schemes. A long-term roadmap scales a reduction in the costs of healthcare insurance, lowering the cost of entry to provision and increasing access to healthcare.

3.6 Smart Contracts

intelligence.sol: Distributed FL protocol to orchestrate the coordination and rewards process to train AI tools on personal health data — number of rounds of training, model contrbutivity score, decentralized auditing protocol, etc.

token.sol: A native token in proof of stake ecosystem (i.e. on a L2 protocol), providing access for Evaluators to assess patients' model contributivity in which they report scores accurately to maintain security of the network.

3.7 Tokenomics

Designing a new ERC-Non-Speculative Design Standard for integrated ecosystems
The token aims to follow a non-speculative design, where only network contributors (e.g., Evaluators) are incentivized to purchase and hold tokens, as they earn rewards for their contributions. This design, known as Partial Common Ownership Tokens (PCOT), is a new standards proposal.

Solving the outlined problem requires a comprehensive understanding of the complex interplay between tokenomics, incentives, and decentralized governance. The focus here is on creating a non-speculative, sustainable ecosystem for federated learning, particularly one that respects patient data sovereignty and ensures the integrity and reliability of evaluators. Below, I'll provide an analysis that could guide the development of such a system, touching on mechanisms for rewarding contributors, ensuring fair valuation, and managing the token supply.

Rewarding Contributors and Developers
Dual-Token Ecosystem? Or Single-Token system?

Single

Utility Token - only staked assets are taxed, but then, only evaluators who stake can earn from the rewards of evaluating in the network. If they arefg a set of rules onto a smart contract and published (distributed) onto a public blockchain that executes these functions. It is important to use a public blockchain as this preserves a “trustless" system — meaning anyone can access the network to conduct training according to FL protocols which preserves decentralization and avoid points of centralization.
In this way, patients can preserve their own self-sovereign identity and contribute to building global models which address healthcare challenges common to a community of patients with similar conditions. As data never leaves patients’ PHR stores, a patient only needs to set permissions for information access once instead of every time they receive healthcare. AI clinical tool developers, or else, any entity seeking to train AI on patients’ health data, can then access a “permissionless” FL-smart contract to train on health data via accessing FL protocols distributed on a public blockchain. An incentive is deposited directly into a patient's digital wallet in return for access to these on-chain FL protocols with a road map to developing a decentralized insurance solution. A scalable and secure decentralized auditing protocol ensures the fair distribution of rewards.

In summary, this paper seeks to orchestrate FL on a public blockchain to develop AI tools in health care provision. The intelligence protocol is a generalizable framework for experiments that are seeking to conduct learning on decentralized data stores with a reward mechanism that incentivizes participation. This paper serves as the basis for research into a wider ecosystem of decentralized healthcare and other services.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] M. Di Pierro, “What Is the Blockchain?,” *Computing in Science Engineering*, vol. 19, no. 5, pp. 92–95, 2017, doi: 10.1109/MCSE.2017.3421554.

[6] 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.

[7] C. Ma *et al.*, “When federated learning meets blockchain: A new distributed learning paradigm,” *arXiv preprint arXiv:2009.09338*, 2020.

[8] S. Kit Lo, Q. Lu, C. Wang, H.-Y. Paik, and L. Zhu, “A Systematic Literature Review on Federated Machine Learning: From A Software Engineering Perspective,” Jul. 2020. Accessed: Apr. 19, 2022. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2020arXiv200711354K

[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] P. Kierkegaard, “Electronic health record: Wiring Europe’s healthcare,” *Computer law & security review*, vol. 27, no. 5, pp. 503–515, 2011.

[11] P. Heath, “Keeping Information Secure with Remote Users: Hospitals, HIPAA Restrictions and Telecommuting,” 2014.

[12] J. Q. Chen and A. Benusa, “HIPAA security compliance challenges: The case for small healthcare providers,” *International Journal of Healthcare Management*, vol. 10, no. 2, pp. 135–146, 2017.

[13] Y. He and C. Johnson, “Challenges of information security incident learning: an industrial case study in a Chinese

healthcare organization," *Informatics for Health and Social Care*, vol. 42, no. 4, pp. 393–408, 2017.

[14] A. Wright, S. Aaron, and D. W. Bates, *The big phish: cyberattacks against US healthcare systems*, vol. 31, no. 10. Springer, 2016, pp. 1115–1118.

[15] T. Floyd, M. Grieco, and E. F. Reid, "Mining hospital data breach records: Cyber threats to US hospitals," in *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, IEEE, 2016, pp. 43–48.

[16] L. Ayala, "Cybersecurity for Hospitals and Healthcare Facilities," 2016.

[17] K. Chinthapalli, "The hackers holding hospitals to ransom," *BMJ*, vol. 357, 2017.

[18] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, PMLR, 2017, pp. 1273–1282.

[19] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine Learning and Systems*, vol. 2, pp. 429–450, 2020.

[20] J. Lee, J. Sun, F. Wang, S. Wang, C.-H. Jun, and X. Jiang, "Privacy-preserving patient similarity learning in a federated environment: development and analysis," *JMIR medical informatics*, vol. 6, no. 2, p. e7744, 2018.

[21] Y. Kim, J. Sun, H. Yu, and X. Jiang, "Federated tensor factorization for computational phenotyping," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 887–895.

[22] D. Liu, D. Dligach, and T. Miller, "Two-stage federated phenotyping and patient representation learning," in *Proceedings of the conference. Association for Computational Linguistics. Meeting*, NIH Public Access, 2019, p. 283.

[23] I. Balelli, S. Silva, M. Lorenzi, and A. D. N. Initiative, "A Probabilistic Framework for Modeling the Variability Across Federated Datasets," in *International Conference on Information Processing in Medical Imaging*, Springer, 2021, pp. 701–714.

[24] "An open-source federated learning framework. - Fed-BioMed." https://fedbiomed.gitlabpages.inria.fr/ (accessed Apr. 19, 2022).

[25] A. Chakravarty, A. Kar, R. Sethuraman, and D. Sheet, "Federated Learning for Site Aware Chest Radiograph Screening," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2021, pp. 1077–1081.

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

[27] 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.

[28] "Ethereum whitepaper - whitepaper.io." https://whitepaper.io/document/5/ethereum-whitepaper (accessed Apr. 19, 2022).

[29] "Buy & Stake MATIC Token | Polygon's Native Token." https://polygon.technology/ (accessed Dec. 22, 2022).

[30] "SKALE | Zero Gas Fee EVM Network | Welcome to the SKALEverse." https://skale.space/ (accessed Dec. 22, 2022).

[31] A. J. Menezes, P. C. Van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*. CRC press, 2018.

[32] "10 Sovereign Technology Components," *Google Docs*. https://docs.google.com/document/d/1q5BGr4THTPai-PKJLZ0_N1ZTsKDzBD795XPicCQIvIs/edit?usp=embed_facebook (accessed Apr. 19, 2022).

[33] 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.

[34] M. S. C. MD, "NOSH ChartingSystem Installation Instructions." Apr. 13, 2022. Accessed: Apr. 19, 2022. [Online]. Available: https://github.com/shihjay2/nosh2

[35] 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.

[36] D. Blumenthal, "Implementation of the federal health information technology initiative," *New England Journal of Medicine*, vol. 365, no. 25, pp. 2426–2431, 2011.

[37] 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.

[38] "NOSH ChartingSystem | A new open source health charting system for doctors." https://noshemr.wordpress.com/ (accessed Apr. 19, 2022).

[39] "Latinos and Technology in the Fight Against Diabetes." https://www.softwareadvice.com/medical/industryview/latinos-diabetes-report-2014/ (accessed Apr. 19, 2022).

[40] *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

[41] *HIE of One Highlights*, (Oct. 15, 2017). Accessed: Apr. 19, 2022. [Online Video]. Available: https://www.youtube.com/watch?v=N_3DbDZUTIg

[42] "Consolidated CDA Overview | HealthIT.gov." https://www.healthit.gov/topic/standards-technology/consolidated-cda-overview (accessed Apr. 19, 2022).

[43] B. A. Fiedler, "Device failure tracking and response to manufacturing recalls," in *Managing Medical Devices Within a Regulatory Framework*, Elsevier, 2017, pp. 263–275.

[44] Z. Wang and Q. Hu, "Blockchain-based Federated Learning: A Comprehensive Survey," *arXiv preprint arXiv:2110.02182*, 2021.

[45] C. Pappas, D. Chatzopoulos, S. Lalis, and M. Vavalis, “Ipls: A framework for decentralized federated learning,” in *2021 IFIP Networking Conference (IFIP Networking)*, IEEE, 2021, pp. 1–6.

[46] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in *Theory of cryptography conference*, Springer, 2006, pp. 265–284.

[47] G. D. M. Serugendo, M.-P. Gleizes, and A. Karageorgos, *Self-organising software: From natural to artificial adaptation*. Springer Science & Business Media, 2011.

[48] L. Kiffer, A. Salman, D. Levin, A. Mislove, and C. Nita-Rotaru, “Under the hood of the ethereum gossip protocol,” in *International Conference on Financial Cryptography and Data Security*, Springer, 2021, pp. 437–456.

[49] P. Ramanan and K. Nakayama, “Baffle: Blockchain based aggregator free federated learning,” in *2020 IEEE International Conference on Blockchain (Blockchain)*, IEEE, 2020, pp. 72–81.

[50] J. Benet, “Ipfs-content addressed, versioned, p2p file system,” *arXiv preprint arXiv:1407.3561*, 2014.

[51] 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.

[52] “distributed-learning-contributivity/README.md at master · LabeliaLabs/distributed-learning-contributivity · GitHub.” https://github.com/LabeliaLabs/distributed-learning-contributivity/blob/master/README.md (accessed Apr. 19, 2022).

[53] *51 Attack*. https://www.coindesk.com/tag/51-attack/ (accessed Apr. 19, 2022).

[54] 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.

[55] 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.