Primary AI

Blockchain & AI for Patient-Centric Medical Record Data and Research Monetization White Paper

Shaun Pritchard and Chris Maxwell

Florida Atlantic University Boca Raton, Florida

Spritchard2021@fau.edu , cmaxwell2021@fau.edu

ABSTRACT

This whitepaper presents a novel distributed computing platform and client application with a novel patient-centric distributed computing architecture for patient medical and health data, which is built on top of a modern, decentralized blockchain-based technology. The proposed technology application platform addresses three main issues related to patient medical records, sharing medical and health information and utilizing the data for scientific research, thereby reducing data scarcity while also encouraging issues such as interoperability, instant data authorization, digital identity of patent records, and real time health monitoring. Most importantly the evolution of fault tolerant security persistence with a data sharing ecosystem for patient-centric healthcare and medical technology.

With the advent of two foundational cryptographic tools, distributed computing platforms have gained new viability: secure hashing algorithms, and public-key encryption[1]. As a result of these technologies and the proposed frameworks, architecture, we have outlined the scope of a new platform protocol which will implement communication over an immutable distributed hash table peer-to-peer network consisting of a blockchain-based distributed version control ledger system known as the source chain. In contrast to having one master ledger like a blockchain for all users and transactions linked together on the network. Primary.ai implements a source-chain where each user: u, $\forall u \in A$ where A is all nodes or authorized applications such as providers. It is the concept of having individual ledgers for each user on the network consisting of only transactions and events with those in the network for which they are authorized, which are also recorded in the source-chain on their ledger (source-chain) and on the user who authorized them. Using a public-key encryption system which is protected by hashing algorithms in the source-chain, this system can encrypt and track historical and real-time transactions that build a source chain of peer user data. Our protocols have been developed in conjunction with these key technologies to store encrypted shards and rulesets of transactions among peers across the distributed hash table network. This ensures the integrity and validation of all data and the actors on the network.

This whitepaper will provide more information about the inner workings of the system, its protocols, and technology. sections and topics will be discussed to put the proposed technology in context and assess its viability as follows: Objectives and goals, Proof of concept, Issues with legacy EHR systems, Current need for medical data and EHR solutions, Market data analysis, Current blockchain technology, Current blockchain in healthcare solutions, Issues with current blockchain technology in healthcare, Introduction to Primary.ai platform technology, Key technologies and architecture of Priamry.ai, Overview of blockchain and distributed version control (DVC) systems, Formal overview of Holochain protocols and technology, Formal overview of Primary.ai, High level overview of Priamry.ai technology, Proposed artificial Intelligence systems for Primary.ai, Proposed technology stack, and a market evaluation.

Objectives and Goals

1. Bridge the gap between patients, providers, and scientific research with an immutable, non-forgeable, and non-repudiation interoperability platform.
2. Make healthcare more patient-centric by empowering individuals to control their own healthcare data and making their own healthcare decisions. Giving them the opportunity to advance medicine and science while monetizing their data[2], [3].
3. Assuring interoperability and security of patient data.
4. Solving the medical data scarcity issue, enabling individuals to use their data for life-changing medical, health, and scientific research, which can lead to breakthroughs and new discoveries.
5. Allow health and medical providers to securely share and authorize medical data on behalf of patients instantly. Providing them with the ability to provide optimal health care and better patient care.
6. Incorporate direct compliance support for relevant healthcare regulations (e.g., HIPAA/HITECH) and privacy regulations (e.g., GDPR).
7. Provide better health outcomes by combining real-time mIoT health data with historical records.

Additional Proof of Concept (PoC) Goals

1. Provide a full scope to the development of a minimal viable product (MVP) of the proposed technology.
2. ShowcasePrimary.ai capabilities, architecture, and design patterns of Primary.ai.
3. This Proof of Concept (PoC) is intended to serve as a learning vehicle and foundation to the agile design of a robust application platform.

4. Encourage collaboration with other data, computer scientists in the field medical applications in AI and cryptology.

5. Bring about business solutions and opportunities for Venture capitalist (VC's), backers and supporters of the platform.

Issues with legacy EHR and medical data systems

Understanding the problem, we are trying to solve with Primary.ai is important for context. It is also important to examine the scope of current technology in contrast and the technologies that are being implemented for the platform, as well as the issues and downfalls surrounding them. In addition to solving a number of problems, current EHR and EMR data management technology and sharing systems also have some very significant downfalls and limitations that affect their adaptability and reliability, among these limitations are.[4]–[6]:

- *Financial cost and return on investment:* The costs of purchasing, implementing, supporting, and maintaining such a system are unaffordable, especially for small hospitals and clinics. Even if they are given the system for free, there are other financial costs related to the management of the interface, customization for flexibility, training, maintenance, and upgrades.

- *Upgrading the workforce:* Currently, the workforce is addicted to paper-based records. Training patients, doctors, pharmacies, hospitals, and so on to adopt the solution is a difficult and time-consuming task. Sometimes it may require changing the workforce. For example, when banks started implementing computers, transitioning from record-keeping books to digital records, a lot of people could not understand and adopt it and therefore lost jobs.

- *Integrity of data input:* Accidental data entry errors, such as selecting the wrong patient or clicking on the wrong choice in a menu of dosages, may occur.

- *Security and privacy:* This are one of the most important concerns. Health records need to be stored securely as eHealth databases are always a target of hackers. Health records contain very sensitive information, and leakage could result in catastrophe. Strong access control should be implemented, and regular feedback should be taken. Without the patient's consent, their records should not be shared with anyone.

- *System downtime:* There are chances of regular system downtime due to network or hardware related issues. The inability to use the system is of great concern.

- *Lost patient access:* In the event of a development beyond the control of the patient, such as a software malfunction in the healthcare provider's office, the patient can no longer ask the care provider for a paper script to take to a pharmacy in order to obtain needed medicines. This leaves the patient at the mercy of technicians or other undiscoverable workers.

Current need for medical data and EHR solutions

According to global market research the use of stolen or compromised credentials remained the top cause of a data breach in the 2022 report, accounting for 19 percent of all analyzed breaches[7], [8]. A healthcare data breach cost a health or medical institute $10.1 million on average, according to IBM Security's in March 2021 to March 2022 the annual Cost of a Data Breach Report[7], [9]. IBM looked at data breaches from the cost of a breach in the healthcare industry went up 42% since 2020[9]. That is up 9.4% from the same timeframe a year earlier. Healthcare has had the highest breach-related financial damages for 12 consecutive years, according to IBM's report. 13% of this issue is caused by centralized third party software on legacy EHR systems. The amount of patient data on the black market, dark web, and underground is irrepressible. It is imperative that new solutions be developed[8].

Market data analysis

According to market data, Electronic Health Record (EHR) Market size was valued at over USD $29.5 billion in 2020 and is expected to grow at a CAGR of more than 6.4% between 2021 and 2027 with an evaluation of $30.1 billion-dollar industry[10]. According to market research, The global Medical Internet of Things (mIoT) in the healthcare market was valued at USD $73.5 billion in 2021 and is all set to surpass USD 190 billion by 2028, exhibiting a CAGR of 25.9% during the forecast period 2022-2028[11]. Research shows that global artificial intelligence in Healthcare Market Size & Share to Surpass $95.65 Billion by 2028[12]. The Global Blockchain Technology in Healthcare Market size is expected to reach $5.3 billion by 2028, rising at a market growth of 39.9% CAGR[13].

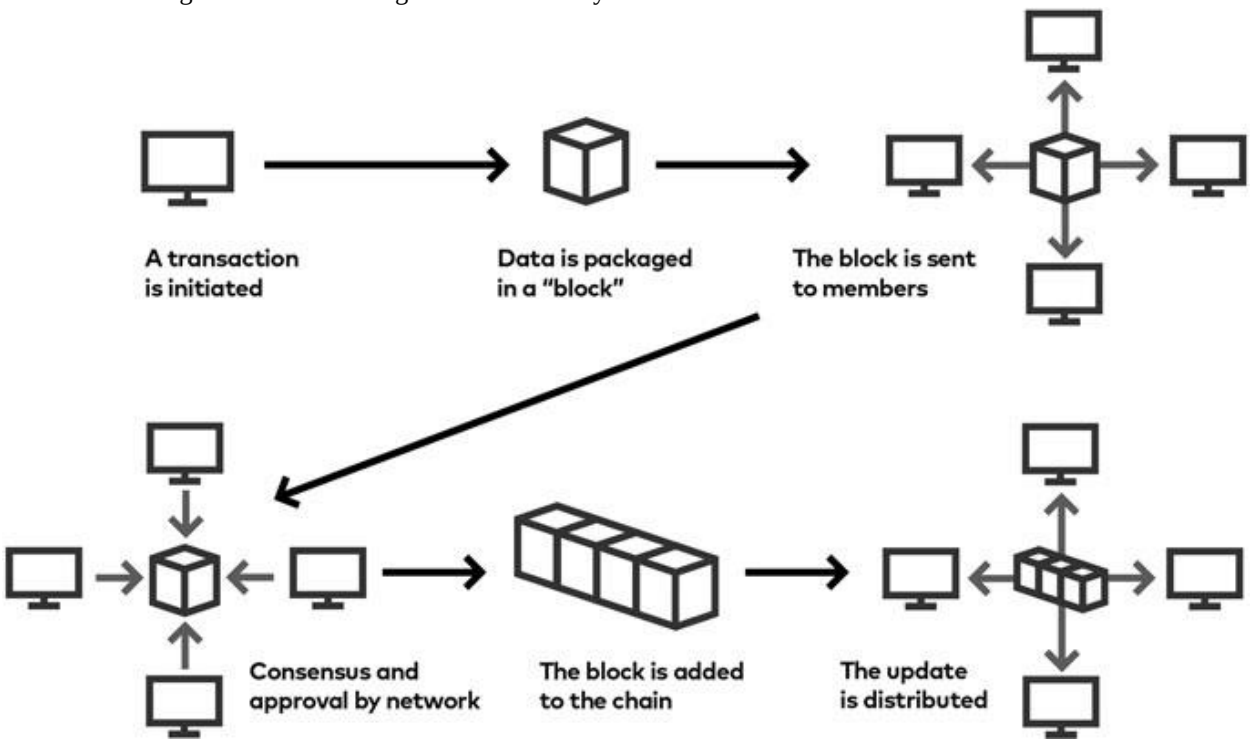
Current blockchain technology

Blockchain technology is a peer-to-peer technology that is decentralized and distributed. As it implements censuses of actions and events to a distributed ledger

system, it is mostly governed by algorithms that solve complex cryptographic puzzles[14]. It is then followed by the Proof of Work (PoW) mechanism. A Proof-of-Work algorithm is basically used to verify a transaction before confirming it in a "blockchain." Upon successful verification, a new block is added to the chain. This process is necessary for security and privacy. In PoW, miners compete to find a random string, a new hash is generated once the string is hashed with the previous block's hash and transactions. An input string is converted to an output string of a certain length through the use of hashing. There are x leading zeros in the newly generated hash, where x is dependent on the difficulty level. A description of blockchain's working mechanism is given in Figure 1.

The purpose of this paper is not to disclose the full scope of the Blockchain technology, but to define it in order to outline its context to the proposed *Priamry.ai* technology platform. The following resources go more into depth on the topic of blockchain technology[15]–[18].

Figure. 1



Current blockchain solutions for healthcare

The sharing of healthcare data among various institutions, hospitals, and healthcare providers today is very complex. Medical and health providers face issues related to the sharing of patient data. Organizations must pay particular attention to both individual consent and data sharing under regulatory systems such as *Hipaa* and *GDPR*, which protect individual data, information, and privacy[19]–[21]. Blockchain can play a crucial role in healthcare and resolving regulatory concerns by providing fault-tolerant encryption-based medical and health technology, which can be used to secure and distribute patient data[18]. By removing both geographical and system barriers, blockchain technology can improve interoperability in the medical tourism sector. A major problem is that it is difficult to have access to the clinical data accumulated by patients, in other words, their medical history, so that better care can be delivered within their country[22]. The use of blockchain technology may be the ultimate solution to creating a decentralized, uniform worldwide EHR system. However, this proposed technology is not intended to compete with or become an EHR software platform, but rather to integrate with existing technologies. Furthermore, Blockchain enables members/patients to have global mobility, since their medical history data can be securely accessed by any provider, anywhere in the world, via the Internet. An interoperable healthcare system allows clinics to share healthcare information without limits and optimize their procedures [20]. There are three categories of interoperability, which allows different information systems to talk and comprehend information passed to each other which include syntactic, structural, and semantic data. The goal of interoperability is to ensure seamless communication and processing of patient

data[23]–[25]. Unfortunately , there are many issues with interoperability in current legacy health systems being used today as follows:

Current interoperability issues in medical and health

- Managing inconsistent information from multiple sources.
- Establishing the validity of electronic patient information requests.
- The process of overcoming organizational resistance to the sharing of data.
- Managing interoperability requires hiring specialists at a high cost.
- Facilitating the accessibility of data.
- A number of solutions arise to these issues as a result of blockchain technology. A blockchain offers features that are not available in current legacy systems that try to provide interoperability as follows:
- *Decentralization:* As blockchain-based networks replicate data end-to-end, they offer fault-tolerance architecture. With the implementation of decentralization, security and privacy have been managed more efficiently.
- *Immutability*: Since blockchain technology is immutable and tamper-proof, it ultimately provides fault-tolerant security. As a result of the hash function, the data is resistant to tampering of any kind. Some hashing algorithms, such as SHA-256, RSA, and RIPEMD-16, are used to calculate hash values [1].
- *Mechanism of consensus:* In exchange for doing this work, the winner of the contest gets a financial incentive for doing so [13]. The winner distributes the block to all other nodes in the network who confirm and validate the block, adding it to their chain [14]. Several consensus algorithms are used to ensure data integrity and to validate blocks, such as proof of stake and proof of burn.
- *Traceability:* Blockchains are ledgers that continually grow as the number of blocks increases. Each block contains a list of all transactions that have been completed. In this chain of blocks, every block has a parent block. The genesis block, or the 0th block in the chain, is the first block in the chain. The hash code for the 0th block is added to the header of the following block, and then the hash code for the second block is calculated. In turn, the hash of the second block becomes the parent of the third block, and so on. A key feature of blockchain is its ability to provide provenance of data, which allows investigators to keep track of activities chronologically and trace the chain backward if necessary. As a result, each block is linked to the other, and has a timestamp to identify it. The origin of this link and all transactions applied to the link can be traced back to the 0th block.
- *Security:* Cryptographic techniques and mathematical models of behavior and decision-making are used to secure blockchains. A blockchain is immutable and impossible to duplicate or destroy. In addition to providing fault-tolerant security and privacy, blockchain technology provides an advanced level of encryption to protect every transaction and exchange that takes place due to the complexity of its nature.

Data management in healthcare is becoming more advanced as a result of the ability to access and share personal health information. In the medical and health industry, blockchain's various superior characteristics, such as decentralization, immutability, consensus mechanisms, traceability, and security, offer solutions that provide a profound impact. Furthermore, senior physicians need to exchange clinical domain information with data scientists. Unstructured data collected from mobile phones and health data devices that monitor patients constantly for diagnostic information requires advanced standards for data preparation. A substantial number of health information sources are difficult for information systems to comprehend, making the integration of clinical knowledge and data standards that exchange this knowledge from various case studies critically important[26].

New privacy protection rules along with cloud-based storage for health data and patient data have opened up new possibilities for managing health data, enabling patients to easily access and forward their health records. It is essential for any data-assisted organization to ensure data security, transactions, storage, and continuous integration, especially in healthcare services where blockchain has the capability of taking care of these open issues in a more profound, secure, and efficient manner [27]–[29]. According to Figure 2, there are currently proposed blockchain models for health and medical data, which present a number of challenges due to their complexity, which will be discussed in due course.

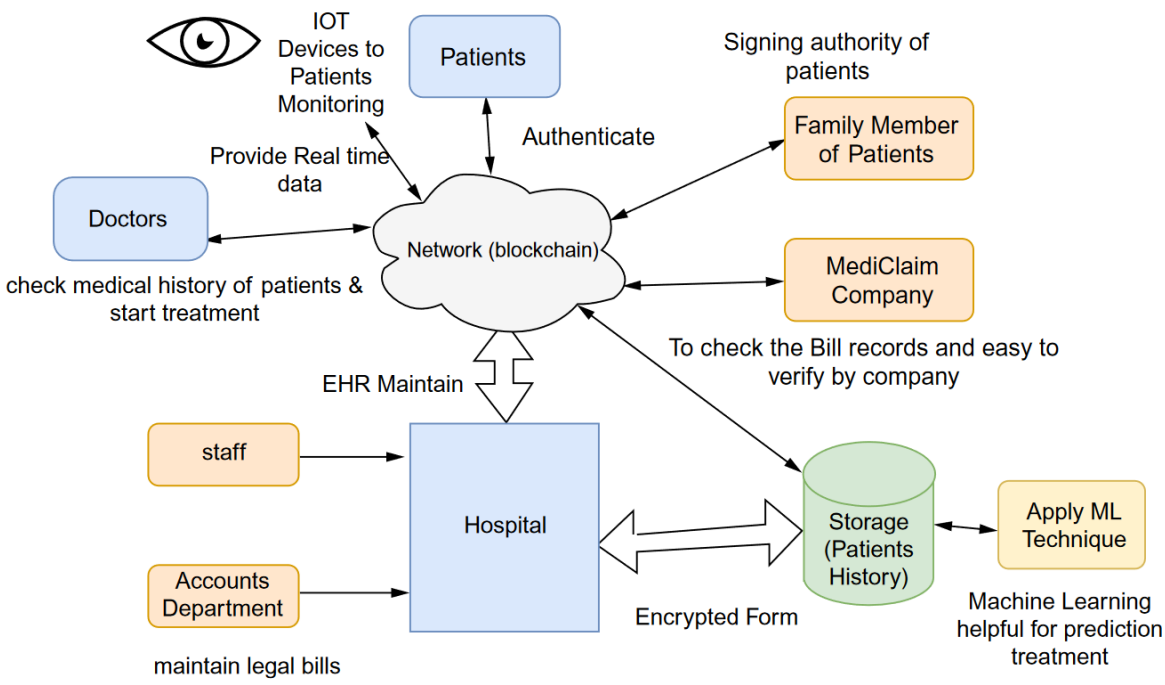


Figure. 2

Issues with current blockchain technology in healthcare

While blockchain technology is a rapidly growing industry that brings new advents of fault tolerant, secure, and technology that has the potential to meet and exceed all three inoperable requirements in the health and medical industry. The blockchain technology does have some constraints that do cause concern for adaptation of the technology into current legacy systems.

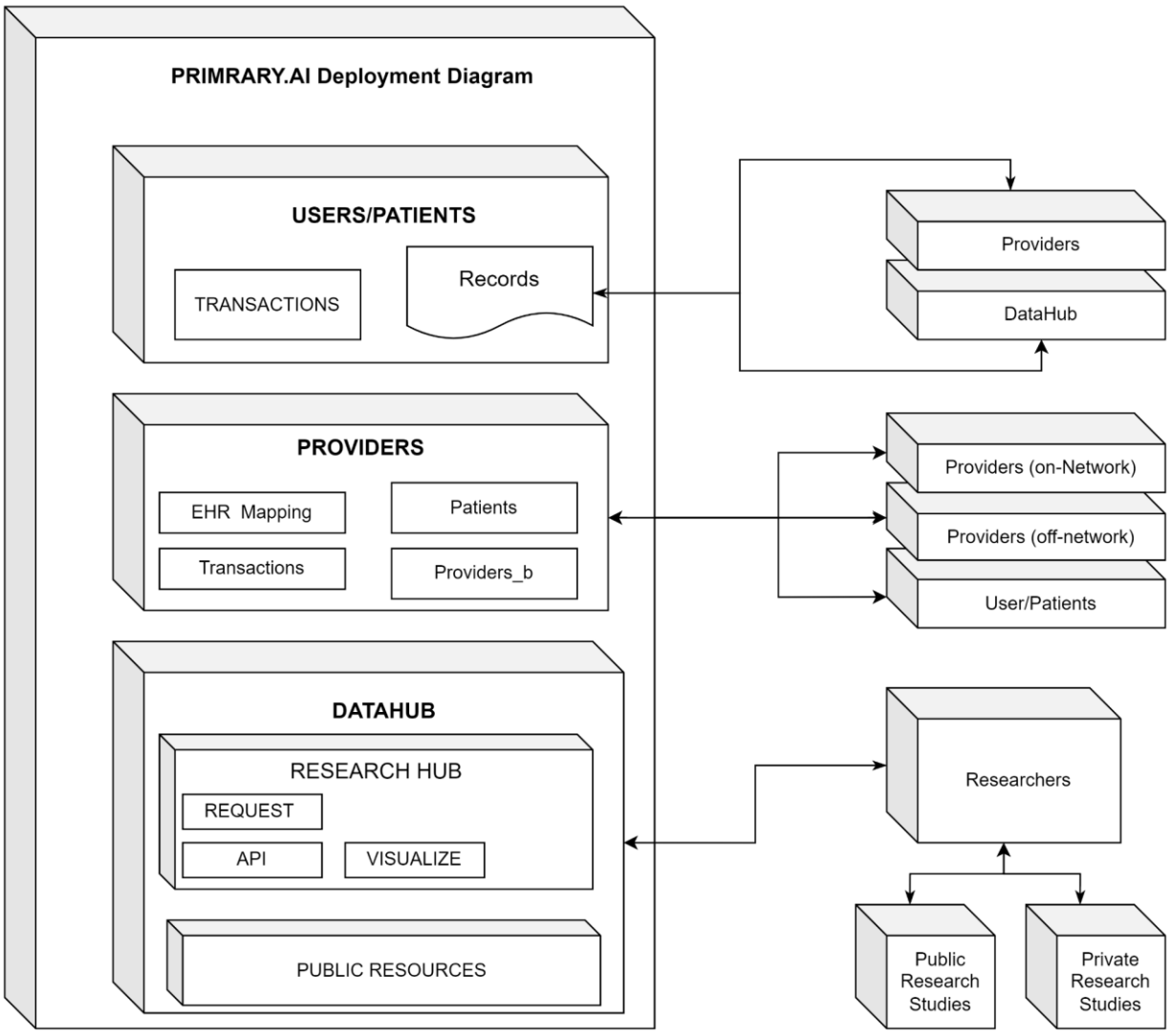
- *Scalability:* Scalability is less of a concern due to the decentralized architecture[30]. Nevertheless,
- It is important to note that private clinics, healthcare centers, rural hospitals, enterprise research organizations, insurance companies, individual patients, and IoT startups all have millions of users with varying infrastructures. All of them are unlikely to be able to maintain the same blockchain decentralized architecture. A higher computation power is also required by blockchain technology, resulting in a higher electricity consumption by network equipment [75]. In order to make blockchains popular, the scalability issue must be addressed.
- *Storage:* EHR/EMRs and sensor data generated by wearable IoT devices and medical records produce a tremendous amount of data in healthcare and medical records. The blockchain architecture, on the other hand, only supports a very limited amount of data storage on the chain. Due to its decentralized and hashed architecture, blockchain has too high a cost for storing data. It is also possible for blockchain data access, management, and operations to be costly if the size of the data is large. As a result, blockchain applications must be designed with this consideration in mind.
- *Privacy and regulations:* The blockchain maximize the level of security of its content in many ways. An architecture based on cryptography, decentralization, independence, and immutability can provide the highest level of security for its contents. Big data in healthcare pertains to sensitive information about the patient, by the patient, and for the patient. It may therefore be risky to keep a copy of these data in each node. For current blockchain technology, the most critical issue is the long-term storage of personal information and electronic health records. This practice is not followed by several countries and standardized organizations according to the (GDPR).
- *Modification:* As a result of blockchain's attributes of immutability, the system is secure; however, as a consequence of its immutability, there is no option for data modification or deletion, and data modification or changes are inevitable. In this case, either a new block must be created by consensus between all nodes or a new chain must be generated. Neither of these approaches is

feasible or cost-effective. Consequently, blockchain applications must be developed in a manner that minimizes the need for data modification.

Introduction to Primary.ai Platform

While we have discussed the advantages and disadvantages of blockchain technologies, we have also discussed the limitations of current legacy systems. Our aim now is to propose an application that leverages the best attributes of this technology in a novel, robust way using the best attributes of the technology. As well as being a platform, Primary.ai distributes inverted blockchain technology and agent-centric distributed peer-to-peer network. In contrast to a single source of immutable truth with a robust ledger that is not scalable. We propose solutions through a shared consensus and traceability mechanism, the immutable ledger is based on a shared repository of each user's data, actions, and events. Which allows us to use the best features of the innovative blockchain technologies. This is a description of the proposed technology and an overview of the technology we are currently developing. Figure.3 illustrates a user's high-level view of the platform.

Figure. 3



Key technology architecture overview

Before we review the underlying scope of the application, let us discuss the technical specification and protocols that will make this application work in the proposed architecture, technologies.

Understanding Blockchain and distributed version control (DVC) technology

As discussed in Holochain [31], they present a formal specification of the Holochain protocol system as well as an analysis of its systemic integrity, capacity for evolution, total system computational complexity, implications for use cases, and current state of implementation. Holochain is a scalable, agent-centric distributed computing platform. That goes on to explain and characterize in comparison to the average distributed systems and demonstrate the benefits of shifting from the common paradigm of a data-centric model to what they refer to as an agent-centric model. To understand these concepts, they outline the characterization of several common canonical distributed system approaches commonly in use today: GitHub , Bitcoin, and Ethereum. The use of these cryptographic tools for distributed computing platforms has brought new opportunities to solve several key problems in distributed computing, such as secure hashing algorithms, and public-key encryption: data that can be verified and tampered with across nodes in a distributed system, and data provenance that can be verified using digital signature algorithms.

The latter is achieved in git, where all nodes can update their hash-chains as they see fit. The degree of overlapping shared state of chain entries (known as commit objects) across all nodes is not managed by git but rather explicitly by action of the users making pull requests and doing merges. The characterization they have labeled this approach is called agent-centric i.e. node-centric, because of its focus on allowing nodes to share independently evolving data realities. Hash-chains are implemented in Git to make monotonic data-stores intrinsically tamper-proof (and thus shareable across multiple nodes confidently).

Whereas in comparison In Bitcoin (and blockchain in general), hashes of data are encrypted cryptographically, and public keys are used as addresses, allowing other agents to mathematically verify the source of the data. They note, the “problem” is understood to be that of figuring out how to choose one block of transactions among the many variants being experienced by the mining nodes (as they collect transactions from clients in different orders), and committing that single variant to the single globally shared chain. We call this approach data-centric because of its focus on creating a single shared data reality among all nodes.

Although both methods are sound technologies, the main problem with git is that anyone can commit, copy, add new data, access the previous data, to any repository. The data is cryptographically secure and immutable; however, anyone can access the data and all of its versions. It can be copied and mutated to whatever the new author determines.

A blockchain system, such as Bitcoin, ensures that every transaction related to the user's new copies, access, and ownership of a block is recorded on a master ledger, making it one of the most secure systems through census where all nodes carry identical values. As the ledger containing the transactions grows, the complexity of the ledger increases exponentially. As a result, the system is not scalable and requires more complex resources to operate.

Holochain[31] claims the assumption that untrusted nodes, i.e., independently acting agents(nodes) solely under their own control, and an insecure channel.do this because the very justification of the cryptographic tools mentioned in [31] is to allow individual nodes to trust the whole system under this assumption.

The cryptography immediately becomes visible in the state data when any other node in the system uses a version of the functions different from itself. This property is often referred to as a trustless system. However, because it simply means that the locus of trust has been shifted to the state data, rather than other nodes, we refer to it as systemic reliance on intrinsic data integrity.

Combining git and blockchain paradigms and features to create what is called a trustful system. They provide a system that ensures users, regardless of who they interact with, can do so with absolute confidence in the possible outcomes and how those outcomes might manifest. It enables a range of decentralized application architectures.

This approach is referred to as being agent-centric because of its focus on allowing nodes to share independently evolving data realities. The agent-centric distributed generalized computing system, where nodes can still confidently participate in the system as whole even though they are not constrained to maintaining the same chain state as all other nodes. In broad strokes: a Holochain[31] application consists of a network of agents maintaining a unique source chain of their transactions, paired with a shared space implemented as a validating, monotonic, sharded, distributed hash table (DHT) where every node enforces validation rules on that data in the DHT as well as providing provenance of data from the source chains where it originated. In comparison, the data-centric approach to distributed computing comes from the fact that if you can prove that all nodes reliably have the same data then that provides a strong general basis from which to prove the integrity of the system as a whole.

Holochain and the agent centric protocol

I will begin by describing the framework complexity and protocol algorithms from an overview perspective. Holochain based application Ω_{hc} is defined as:

1. Call X_n the *source chain* of *n*.

2. Let *M* be a virtual machine used to execute code.

3. Let the initial entry of all X_n in *N* be identical and consist in the set

DNA{e_1,e_2,...,f_1,f_2,...,p_1,p_2,...} where e_x are definitions of entry types that can be added to the chain, f_x are functions defined as executable on *M* (which we also refer to as the set F_{app} = {app_1,app_2,...}), and p_x are system properties which among other things declare the expected operating parameters of the application being specified. For example, the resilience factor as defined below is set as one such property.

4. Let $\imath_n$ be the second entry of all X_n and be a set of the form {*p*,*i*} where *p* is the public key and *i* is identifying information appropriate to the use of this particular Ω_{hc}. Note that though this entry is of the same format for all X_n it's content is not the same. Call this entry the *agent identity* entry.

5. $\forall e_x \in DNA$ let there be an $app_x \in F_{app}$ which can be used to validate transactions that involve entries of type e_x. Call this set F_v or the *application validation functions*.

6. Let there be a function $V_{sys}(ex,e,v)$ which checks that *e* is of the form specified by the entry definition for $e_x \in$ DNA. Call this function the *system entry validation function*.

7. Let the overall validation function $V(e,v) \equiv {}^{W}x\ F_v(e_x)(v) \wedge V_{sys}(e_x,e,v)$.

8. Let F_I be a subset of F_{app} distinct from F_v such that $\forall f_x(t) \in F_I$ there exists a *t* to *I*(*t*) that will trigger $f_x(t)$. Call the functions in F_I the *exposed functions*.

9. Call any functions in F_{app} not in F_v or F_I *internal functions* and allow them to be called by other functions.

10. Let channel *C* be *authenticated*.

11. Let *DHT* define a distributed hash table on an authenticated channel as follows:

(a) Let Δ be a set {δ_1,δ_2,...} where δ_x is a set {*key*,*value*} where *key* is always the hash *H*(*value*) of *value*. Call Δ the *DHT state*.

(b) Let F_{DHT} be the set of functions {dht_{put},dht_{get}} where:

i. *dht*put(δkey,value) adds δkey,value to Δ

ii. *dht*get(*key*) = *value* of δkey,value in Δ

(c) Assume $x,y \in N$ and $\delta_i \in \Delta_x$ but $\delta_i \notin \Delta_y$.

Allow that when *y* calls $dht_{get}(key)$, δ_i will be retrieved from *x* over channel *X* and added to Δ_y.

DHT are sufficiently mature that there are a number of ways to ensure property 11c. For our current alpha version, we use a modified version of [Kademlia] as implemented in [LibP2P].

12. Let DHT_{hc} augment *DHT* as follows:

(a) $\forall \delta_{key,value} \in \Delta$ constrain *value* to be of an entry type as defined in DNA. Furthermore, enforce that any function called $dht_x(y)$ which modifies Δ also uses $F_v(y)$ to validate *y* and records whether it is valid. Note that this validation phase may include contacting the source nodes involved in generating *y* to gather more information about the context of the transaction, see IVC2.

(b) Enforce that all elements of Δ only be changed monotonically, that is, elements δ can only be added to Δ not removed.

(c) Include in F_{DHT} the functions defined in A.

(d) Allow the sets $\delta \in \Delta$ to also include more elements as defined in A.

Let $d(x,y)$ be a *symmetric* and *unidirectional* distance metric within the hash space defined by *H*, as for example the XOR metric defined in []. Note that this metric can be applied between entries and nodes alike since the addresses of both are values of the same (a) hash function *H* (i.e. $\delta_{key} = H(\delta_{value})$ and $A_n = H(pk_n)$).

(b) Let *r* be a parameter of DHT_{hc} to be set dependent on the characteristics deemed beneficial for maintaining multiple copies of entries in the *DHT* for the given application. Call *r* the *resilience factor*.

(c) Allow that each node can maintain a set $M = \{m_n,...\}$ of metrics m_n about other nodes, where each m_n contains both a node's direct experience of *n* with respect to that metric, as well as the experience of other nodes of *n*. Enforce that one such metric is uptime which keeps track of the percentage of time a node is experienced to be available. Call the process of nodes sharing these metrics *gossip* and refer to IVC3 for details.

(d) Enforce that $\forall \delta \in \Delta_n$ each node *n* maintains a set $V_\delta = \{n_1,...,n_q\}$ of *q* closest nodes to δ as seen from *n*, which

are *expected by n* to also hold δ. Resiliency is maintained by taking into account node uptimes and choosing the value of q so that

$$\sum_{i=0}^{q} uptime(n_i) \geq r$$

with $uptime(n) \in [0,1]$.

Call the union of such sets V_δ, from a given node's perspective, the *overlap list* and also note that $q \geq r$.

(e) Allow every node n to discard every $\delta_x \in \Delta_n$ if the number of closer (with regards to $d(x,y)$) nodes is greater than q (i.e. if other nodes are able to construct their V_δ sets without including n, which in turn means there are enough other nodes responsible for holding δ in their Δ_m to have the system meet the resilience set by r even without n participating in storing δ). Note that this results in the network adapting to changes in topology and DHT state migrations by regulating the number of network wide redundant copies of all $\delta_i \in \Delta$ to match r according to node uptime.

Call DHT_{hc} a *validating*, *monotonic*, *sharded* DHT

13. $\forall n \in N$ assume n implements DHT_{hc}, that is: Δ is a subset of D (the non hash-chain state data), and F_{DHT} are available to n, though note that these functions are NOT directly available to the functions F_{app} defined in DNA.

14. Let F_{sys} be the set of functions $\{sys_{commit}, sys_{get}, ...\}$ where:

(a) $sys_{commit}(e)$ uses the system validation function $V(e,v)$ to add e to X, and if successful calls $dht_{put}(H(e),e)$.

(b) $sys_{get}(k) = dht_{get}(k)$.

(c) see additional system functions defined in B.

15. Allow the functions in F_{app} defined in the DNA to call the functions in F_{sys}.

16. Let m be an arbitrary message. Included in F_{sys} the function $sys_{send}(A_{to},m)$ which when called on n_{from} will trigger the function $app_{receive}(A_{from},m)$ in the DNA on the node n_{to}. Call this mechanism *node-to-node messaging*.

17. Allow that the definition of entries in DNA can mark entry types as *private*. Enforce that if an entry σ_x is of such a type then $\sigma_x \notin \Delta$. Note however that entries of such type can be sent as node-to-node messages.

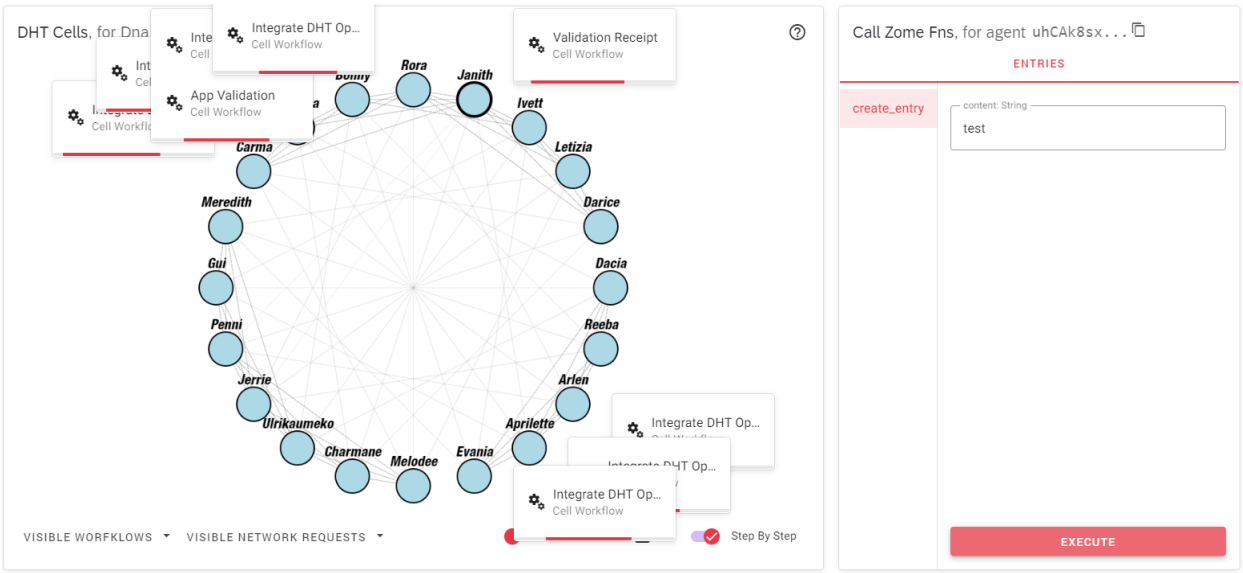
18. Let the system processing function $P(i)$ be a set of functions in F_{app} to be registered in the system as callbacks based on various criteria, e.g. notification of rejected puts to the DHT, passage of time, etc. The agent-centric solution to these requirements, according to[31], involves holographically managing system integrity through application-specific validation routines within every agent/node. All decentralized applications are governed by these sets of validation rules, which vary based on context. The agent keeps a close eye on the portion of reality that is important to him or her - within the context of a given application - to balance high confidence thresholds with a low need for resources and complexity in a given application.

Transactions can be used in two different ways, for example:

1. receipt of an email message where we are trying to validate it as spam or not and
2. commit a monetary transaction where we are trying to validate it against double-spend.

Agents may wish to evaluate these contexts differently and expend differing levels of resources to validate them. Holochain allows such validation functions to be set contextually per application and exposes these contexts explicitly. One could, therefore, build a Holochain application that implements all or partial Blockchain characteristics in its validation functions. As a framework, Holochain enables a spectrum of decentralized application architectures, with Blockchain as a specific instance on one end of the spectrum. Just for added context, the core difference between $\Omega_{bitcoin}$ and Ω_{git} lies in the former's constraint of $\forall n,m \in N : X_n =^! X_m$. One direct consequence of this for $\Omega_{bitcoin}$ is that as the size of X_n grows, necessarily all nodes of $\Omega_{bitcoin}$ must grow in size, whereas this is not necessarily the case for Ω_{git} and in it lies the core of Bitcoin's scalability issues. Holochain solves this issue., as shown below in Figure. 4 we can see the interactions between nodes X_n and transactions between DNA$\{e_1,e_2,...,f_1,f_2,...,p_1,p_2,...\}$

Figure. 4.



Holochain uses gossip for nodes to share information about their own experience of the behavior of other nodes. Informally we call this information the node's *world model*. Recall that each node maintains a set M of metrics m about other nodes it knows about. Note that in terms of the discussed formalism, this world model is part of each node's non-chain state data D $\forall m \in M$ let the function $G_{\text{with}}(m)$ return a set of nodes important for a node to gossip with defined by a probabilistic weighting that information received from those nodes will result in changing *m*other. $\forall m \in M$ let the function $G_{\text{about}}(m)$ return a set of nodes important for a node to gossip about defined by the properties of m. the protocol that defines subsets of $G_{\text{with}}(m)$ according to a correlation with what it means to have low vs. high confidence value c. This allows the protocol to push or pull the node based on the overall ranking in which the confidence score infers.

The definition of Holochain can be as simplified as defining an application, understanding that the user interface is the client level, and that the complex backend level is defined by the scope of the technology mentioned above. While what is considered a backend server which defines the singleton-based validation rules set that are used to receive and request encrypted data; building a source chain of all events and actions while communicating as a conductor to talk to the authorized nodes in the Distributed Hash Table DHT network. Like git, this mechanism signs and validates all data while creating versioned data that is cryptographically signed and verified. The data is stored on the source chain which resides in the user's application. Like Bitcoin, each action and transaction are tracked to an internal ledger, which is signed with the user's public and private keys in the chain commits providing the flexibility of a distributed peer to peer system with security of blockchain.

It should also be noted that once a user/patient/agent authorizes the transmission of data in Primary.ai to a provider or to a research request in the DataHub which will be explained further, an encrypted shard of that data is distributed amongst each node that has authorization to validate new transmissions and records. This is a built-in security feature as described above in the *gossip & world model* that allows nodes in the network who have been authorized to have a user's data to ensure the rules and integrity of the user stay intact or else it is flagged and no transaction or future event can take place.

A Holochain network is a DHT (Distributed Hash Table) of peers, operating together by the ruleset. The DHT is the main way in which agents share data in Holochain. Every agent runs its own node: if a node wants to participate in the DHT, they have to run their own instance, have authorization and compliance with the rules-sets, and join the network. Usually, DHTs are just key-value stores: you can create some piece of data which gets hashed, and after that anyone can query its contents with that hash. Zomes are the singleton rules sets written in web assembly to handle the transactions and rules between two nodes.

Initially, we will focus on building three applications using Primay.ai's AI platform. We refer to this as a patient-centric distributed blockchain platform application. Due to the fact that it has 3 different properties, all of which provide key functionality over the network, we consider it a platform. There are three main application modules, which are composed of microservices, rules, and permissions that differ from one

another. User applications (uApps) run on a user's device and connect only to data hubs and providers connected to the secure DHT network.

By hosting and curating Datahubs (dh), the Primary.ai team allows vetted authorized institutes and researchers who meet Hipaa and GDPR guidelines to post submissions and requirements in relation to the data they are interested in using[20], [21]. Essentially, the Providers application (pApp) integrates with existing EHR systems via file loaders over the DHT network and APIs to share user-authorized data with only other providers and patients. Applications are subject to a variety of constraints based on their role and transactions in network:

Formal overview of Primary.ai platform protocols

Each node is composed of (DNA) protocols for publishing and subscribing based on nodes that interact with Zomes (rules logic to authorize access to receive and transmit data over the distributed hash table(DHT)). The unique public key of the node allows a node to record all transactions in the user block ledger and validate data and transactions with the private key. Asynchronous data streams controlled by the validation access to microservices that can send and receive EHR, EMR, and/or connect to user. Patient MIoT medical device data; records locally on user/patient/actor device or proposed future security record storage service we plan to provide on the datahub DHT network. The following is an overview diagram of a Primary.AI node in Figure.5.

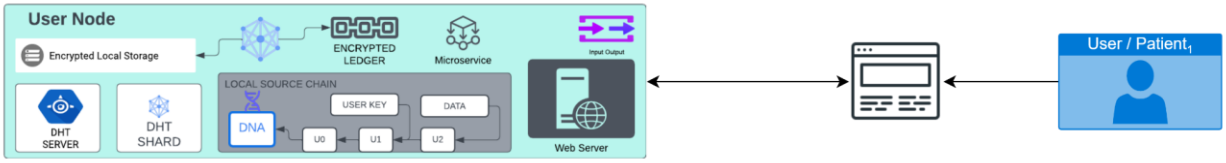


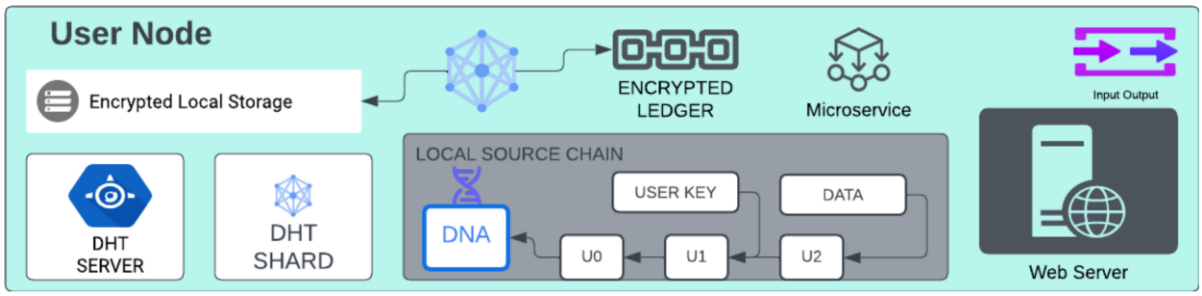
Figure.5

1. Let sn = source chain cryptographic block ledger.
2. Let *Provider app* be *pApp where* ∀*pApp* ∈ *DNA* ⊆ sn.
3. Let data hubs be *dh where* ∀*pApp* ∈ *DNA* ⊆ sn.
4. *Let C = pApp* ∉ *dh* ⊕ *dh* ∉ *pApp* .
5. Let user's application be *uApp* where *uApp* ⊆ *pApp* × *dh* | ∉ *uApp* : *where* ∀*pApp* ∈ *DNA* ⊆ sn.
6. Let N be the set of elements nodes {n1, n2, . . . nn} ⊇ *uApp* where N ∈ dh | N ∈ pApp.
7. Let $p_{providers}$ ∈ {p1,p2…p) where, pApp ∃p∈ pApp ⊕ uApp ∨C
8. Where uApp can send digital signature authorization to allow $p_{providers}$ to share their data. $p_{providers}$ can then send data to other providers.
9. If provider is on the Primary.ai network that $p_{providers}$ sends the data and transactions are recorded on the source chain of the provider and the uApp user.
10. If $p_{providers}$ do not exist on a network uapp will send quick-sign validation via TLS encryption email the transaction will appear on both the provider and uApp user. The forwarding provider will receive records via EHR being used by the initial provider.
11. Let $r_{researchers}$ ∈ {r1,r2…r) where dh ⊇ $r_{researchers}$ researchers can than publish requirements for studies to ∀pApp
12. Where: $r_{researchers} \subseteq pApp \notin p_{providers}$
13. ∀pApp can determine if they want to participate in the study and submit their response based on the *dh* requirements.
14. Using machine learning models for schema matching where J(A,B) = 1 - $\frac{A \cap B}{A + B - A \cup B}$: Users of uApp can make the choice if they want to participate.
15. *dh* will verify computational rules of *uApp* are valid and *uApp* will verify computation rules for *dh* are valid.
16. Upon request uApp will authorize data and send notify dh
17. *dh* will then determine using machine learning schema matching models if it meets the requirements. Upon success of meeting requirements, a final authorizations request submission will be sent to the uApp user. Where uApp user will have an open web-socket channel to the event table of the study which relays △event data. This allows the user to pull their data if they no longer participate and monitor how their data is being used. All transaction data will be recorded in the immutable source chain ledger of both *uApp* and

dh. I should note the user does not get to see the actual inferences of the study from the researcher only if the data is being used. This is to protect the integrity of the research study itself and the research being conducted.

18. Let N be the set of elements {n1, n2, . . . nn} participating in the system. Call the elements of N nodes ∈ uApp .

Let each node n consist of a set Sn with elements {σ1, σ2, . . . }. Call the elements of Sn the state of node n. Where ∀σi ∈ Sn : σi = {Xi , Di} with Xi being a hash-chain and D a set of non-hash chain data elements. Let nd be encrypted data record events of the node element where ∀σi ∈ Sn⇔ nd : σi = {Xi , Di}



The main goal is outlined in the elaboration charts below, but we also plan to implement *proprietary A.I models as follows:*

1. Schema matching
2. Importing of raw data files and image data
3. Manage peer connections
4. Recommendation system for studies

A more detailed explanation of the AI models will follow. Additionally, we need to discuss a payment portal for monetizing patient users. Currently there are several APIs that are very easy to integrate, such as Stripe and PayPal[43], [44]. We propose to use Plaid API service to let users authorize their own payment gateways from within the Priamry.AI platform[45].

In order to realize our goal, we are interested in exploring the option of using an actual blockchain token that would be used as a stable currency and allow users to choose between earning USD and primary.ai crypto currency. We recommend implementing something that has breached the trial of discrepancy and has time-tested protocols and advanced security measures to ensure the integrity of our systems using Solna or Ethereum. Just for added context, the core difference between $\Omega_{bitcoin}$ and Ω_{git} lies in the former's constraint of $\forall n,m \in N : X_n =^{!} X_m$. One direct consequence of this for $\Omega_{bitcoin}$ is that as the size of X_n grows, necessarily all nodes of $\Omega_{bitcoin}$ must grow in size, whereas this is not necessarily the case for Ω_{git} and in it lies the core of Bitcoin's scalability issues.

A high-level overview of Primary.ai framework:

In order to discuss the application framework, we must discuss the primary computational architecture and protocols used to build a fully distributed peer to peer blockchain based platform for sharing medical data. The first step will be a brief overview of the application, followed by a discussion of the computational architecture, and then the anonymity of the application itself. As part of our solutions, we are implementing robust technologies in order to create a never-before-seen unbreakable cryptographic network for storing patient data in a distributed way, using a secure hash algorithm and private keys for access, and utilizing a new version control system to process transactions with patient data.

We characterize this software as a decentralized platform centered around agent-centric /patient-centric approach with a client applications which persist digital identity, access, and local encrypted storage of their medical records and real-time health data giving them total access and anonymity to distributed nodes on the network referred to as data hubs and providers.

Figure. 6.

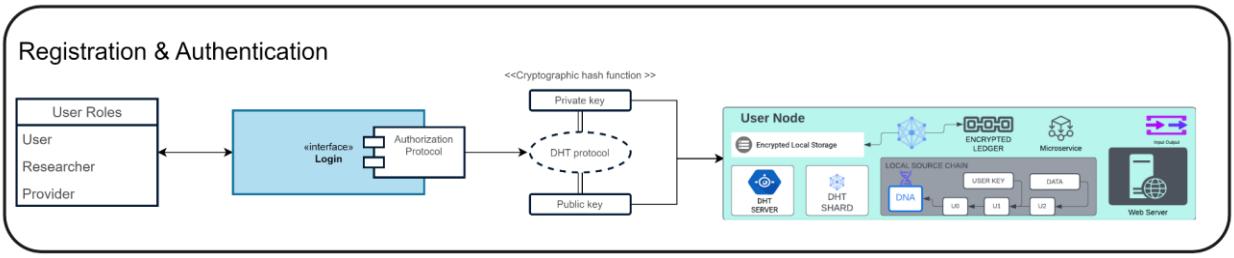


Figure. 6

The following Figure.7, is a high-level overview of the network process for Primary.ai where transactions occur on the DHT.

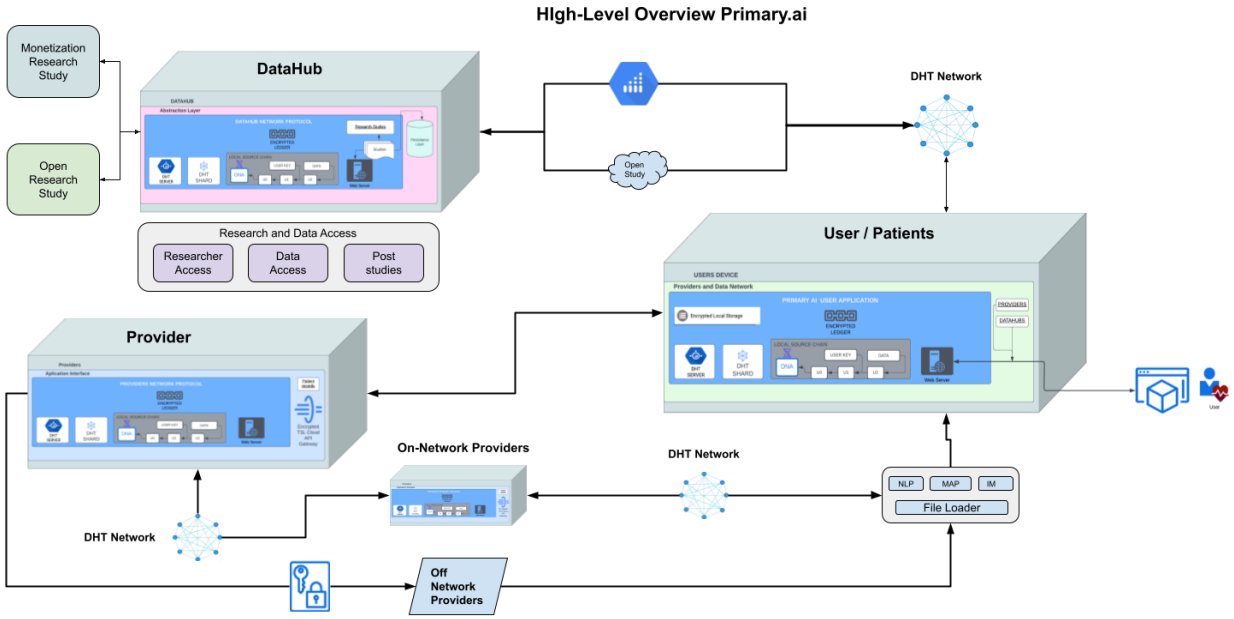


Figure. 7.

Datahubs(dh)

DataHub are the home base of Primary.ai, while Priamry.ai DataHub will act as its own node on the main DHT network providing authorization and access. Each patient/user, provider, researcher, and research project will be their own instance of a node on the Primary.ai closed network. This will essentially allow users/patients to opt-in and allow access to patient data for scientific research and study purposes, access and distribution of medical data, and authorization of medical data for us in scientific research. Studies may be faith-based or may have a monetization component that allows the user to earn while their data is subject to the study offer that they consent to. Using the Primary.ai security and novel protocols, a user can opt-out and withdraw their data from any active study and from any data set that is being used as part of the study. Furthermore, datahubs provide validation of data regardless of whether the user pulls it or not. Once a user opts in, a fragmented copy of the user's data validation rules is available on the datahub. The importance of this component in terms of security and integrity of a platform will be discussed in further detail.

Scientists and researchers can post studies in the datahubs and define the requirements for the research they are conducting. Patients can be notified securely so that they are able to decide whether to participate in the study or not. Most often, the criteria outline the duration of the study, the data requirements, and whether the data can be monetized by the user. Consideration criteria for authorized institutes and researchers. Figure.8 , shows the entire DataHub research project interface.

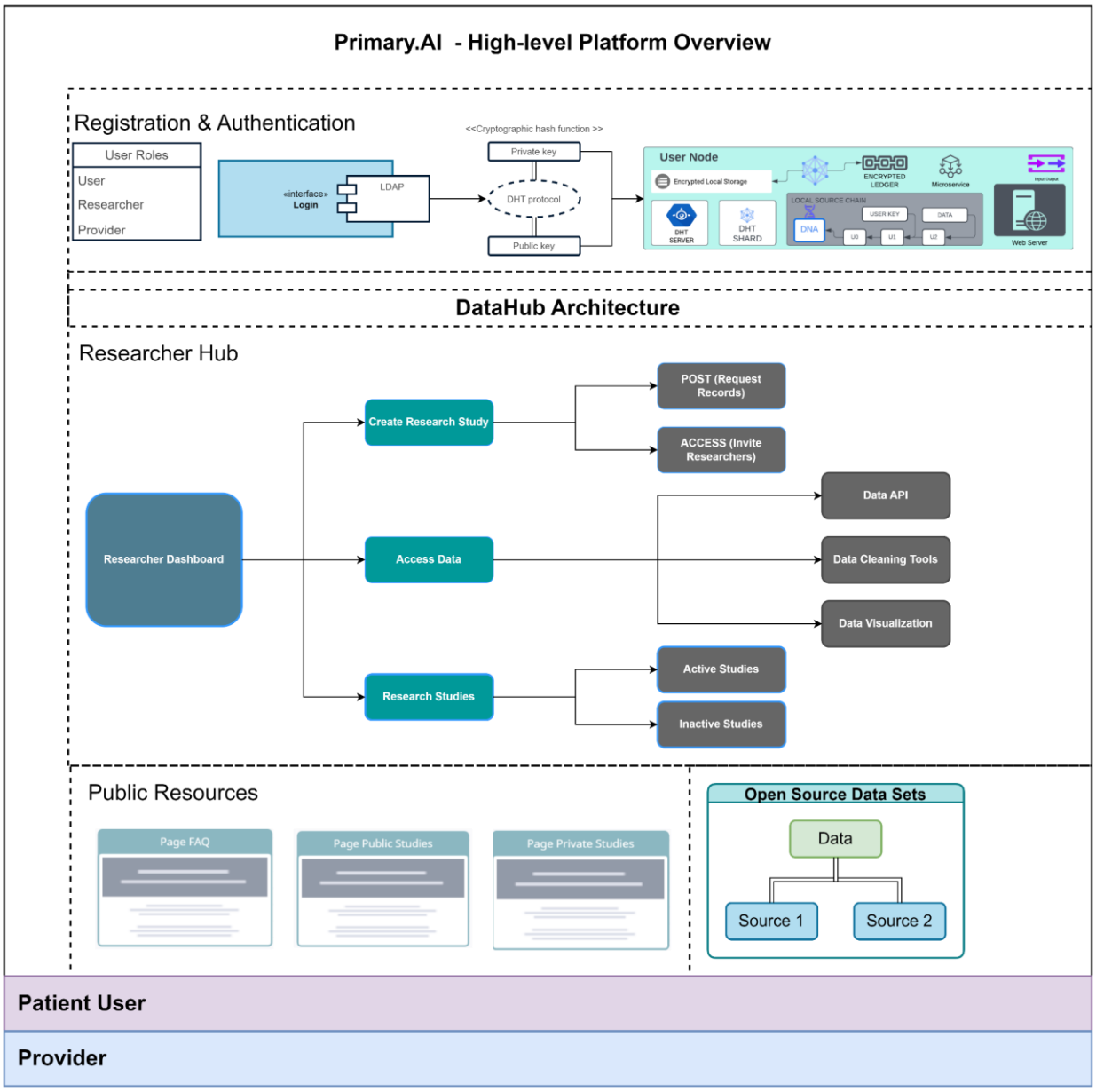


Figure. 8

Below in Figure.9 shows the high-level overview of the Primary.ai Data Hub architecture and access flow within the platform.

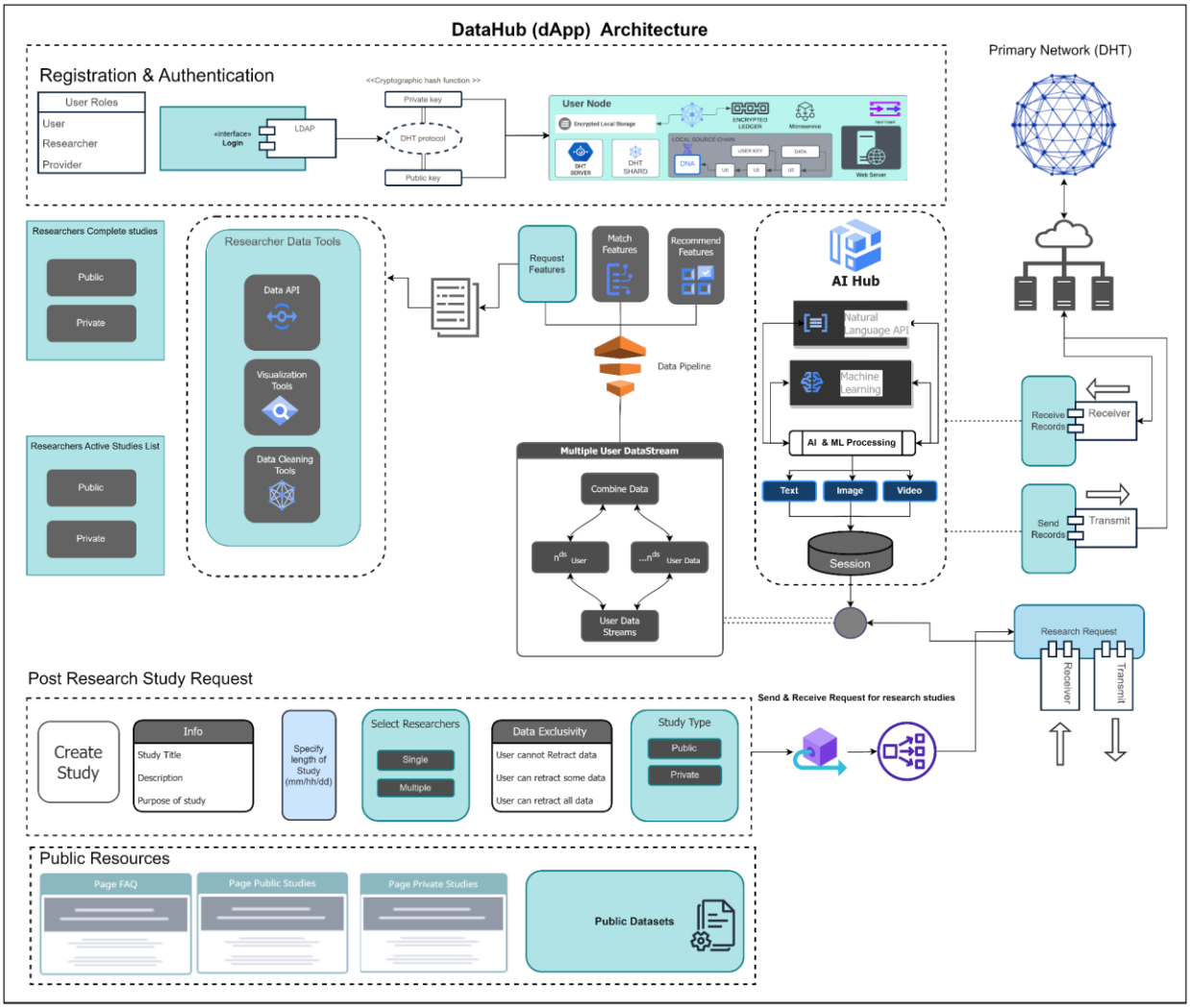


Figure.9

The following Figure.10 shadows the sequence and process diagrams showing the behavior and network communication request between datahub researchers and User/patients

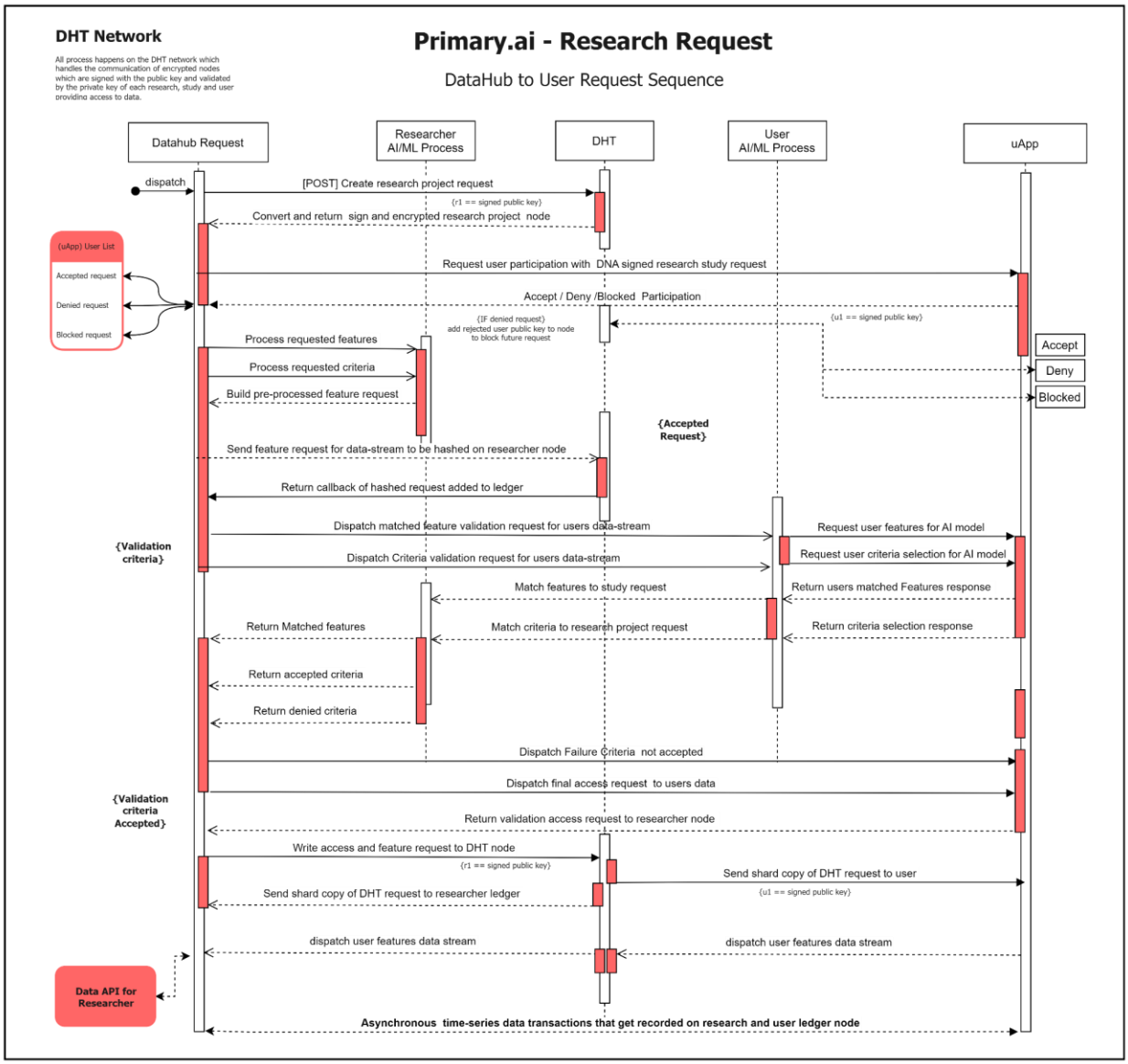


Figure.10

Providers(pApp)

The medical industry is diverse, traditional hospitals and private practitioners are not the only viable end users. Primary.ai categorizes insurance companies, governments, healthcare providers, non-profits, and many other companies in the medical industry as providers. One of the challenges is that, first of all, patients need access to their medical data not only for the purpose of conducting medical research and studies, but also for reasons relating to their medical health and well-being. EHR/EMR systems require patients to sign a paper release form and then send an encrypted fax to a patient's authorized recipient if they wish to release their data to another provider or institution. If the patient wishes to release their data, they must go to their primary provider. Aside from the fact that this can take an insurmountable amount of time, which may have a negative effect on the health of the patient, it may also be an outdated practice that is susceptible to security risks as well as breaches in security. Using Primary.AI, the patient is able to easily submit a verified authenticated digital signature that is verified by the patient, which is a primary and public encrypted file. Patients will have the ability to send their medical records through the primary as shown in Figure 11.

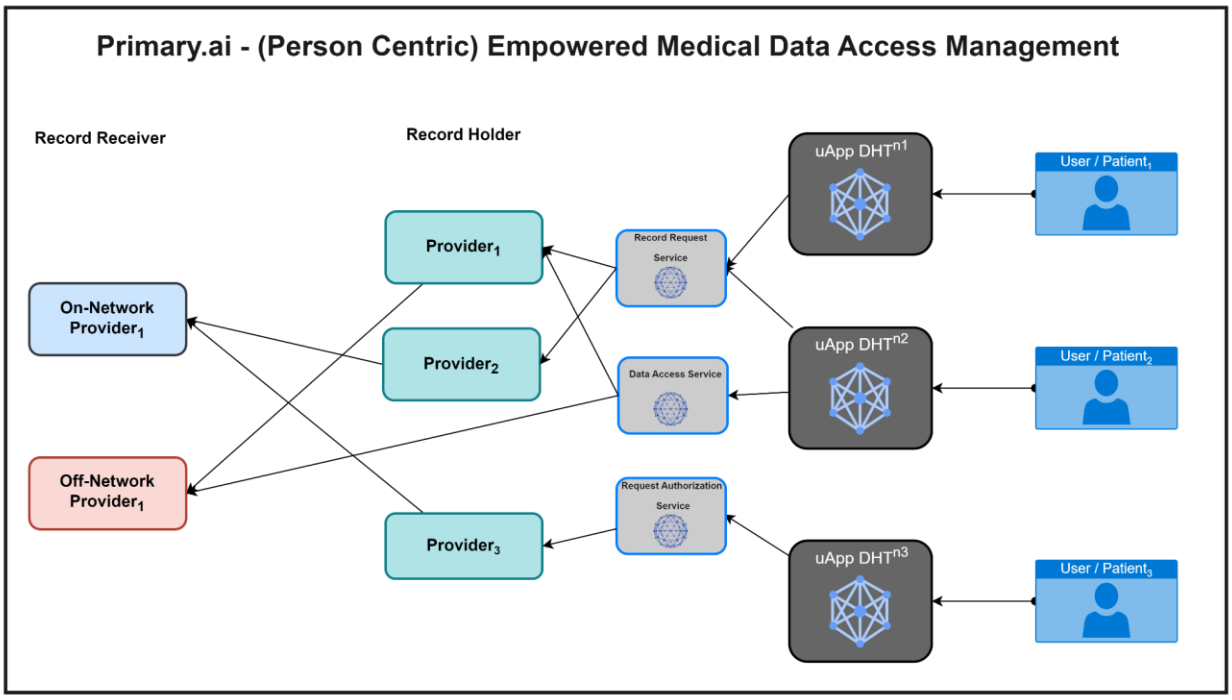


Figure 11

AI provider application for EHR and EMR, and providers will be able to send the medical records on their behalf if they use primary.AI. In addition to giving patients quick access to their medical data, this will also solve one of the primary problems associated with the medical industry: interoperability. As the patient already has their data and is able to request it when needed, they will not need to rely on silos of data when it comes to their health care. As part of this data, all patient records, imaging, and personal health information can be included. Moreover, we plan to incorporate features that will allow real-time patient health data tracking through mIoT devices, such as fit-bits, and other wearables. The data will be stored directly on a patient's device and accessible on the patient's device accessible to their node (agent network, thus providing greater insights to patients' providers and caretakers and, ultimately, helping them provide better patient care. In contrast to traditional methods and current practices, Primary.al aims to make provider medical and health records transactions instantaneous, interoperable, and more secure. Furthermore, all transactions regarding patient data records will be recorded in the patient's application with alerts. Providing them with full transparency over their medical records and who has them.

Generally, providers have their own Electronic Medical Records (EMRs) or Electronic Health Records (EHRs). It is important to note, it would be a significant barrier to entry if we required them to adopt the primary.ai application architecture for their internal systems to replace their current ones. It may be necessary for providers to provision nodes that are configured to join Primary.ai app instances or initially implement an API-based approach to integrate with pApp providers application to the user's data.

As part of future endeavors, Priamry.ai extensions will build out extensions to more popular and trusted EHR systems such as Cerner Medical EHR and Epic EHR. For example, these integrations would be tightly coupled with major big tech providers of digital identity for medical and health, such as Imprivata, Broadcom, IBM. We are proposing a proponent: Major health digital identity companies have already overcome the integration barriers to these current systems. Through collaboration we hope to ensure an easy transition of our product for providers to have access to patients on the Primar.ai network giving patients the ability to quick sign digital medical authorizations and gain access to their own records. Due to the immutability, non-forgeability, and non-repudiation of each user's records, healthcare practitioners can be assured that integrity of the records meet compliance.

As a result of the architecture of Primay.ai's current framework, mapping between the internal systems of Healthcare Providers and the Primary.ai application will also be a capability as well, since it uses APIs and file-mappers as the method of collecting health data. In addition to converting health records into structured data, artificial intelligence is also being used

as a tool to assist in the ETL process of digital, image-based, and paper-based medical records. Detailed explanations of these topics will be provided in the following sections.

Medical providers using the Primary.ai framework application will also have access to all transactions of patient data including those from other providers. This will ensure the integrity of the patient records over time. Figure 12 illustrates the internal access and architecture of the Provider application.

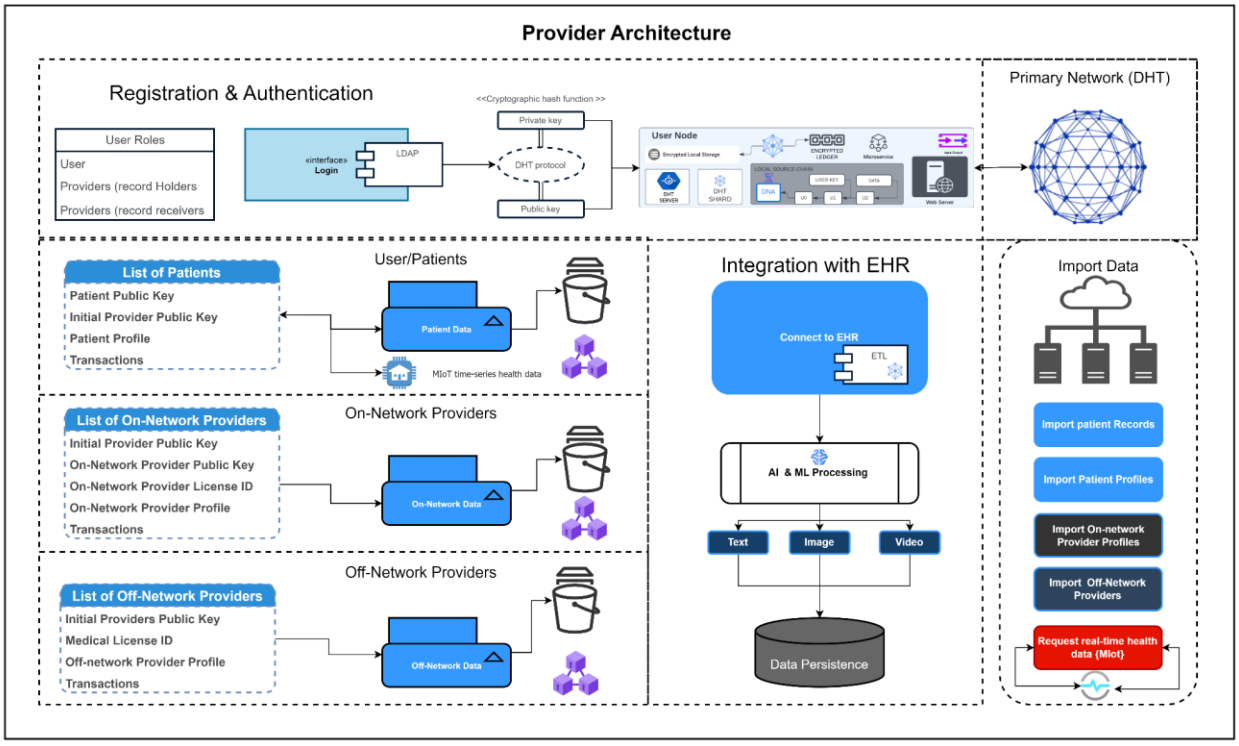


Figure 12

Provider application Transactions and sequences associated with the release of user requests:

2. User dispatches request to release records *{General Authorization for Use or Disclosure of Health information}*
3. *Digital document is signed with date, expiration default (90 days) [is selectable], type of information, purpose of medical release (Options for all classifications and categories must be selected by user), number of authorized persons or facilities to receive records, authorized person facility to receive*
4. *Authorized persons or facilities to receive records*
    a. If records are personal = = *self*
    b. If record is being sent to another provider on-network (signed public key and medical id is provided from database in form
5. User selects provider from on-network (primary.ai provider Node)
6. If no provider is available off-network enter
7. User will use search or manual input provider info
8. If the request to the Provider (record holder) is off network; request is sent via TLS encrypted email.
9. If Request is sent to Provider holding the records on-network (record holder)
10. Signed user public key validation request is dispatched to provider (record holder) through DHT
11. Providers (record holder) send validation request to provider (records receiver)
12. provider (records receiver) responds with public key and medical id validation response to Providers (record holder) .
13. Provider node validates and accept request (can be set to auto request validate or manual validation)
14. Signed transaction is recorded on providers block ledger in initial version of request

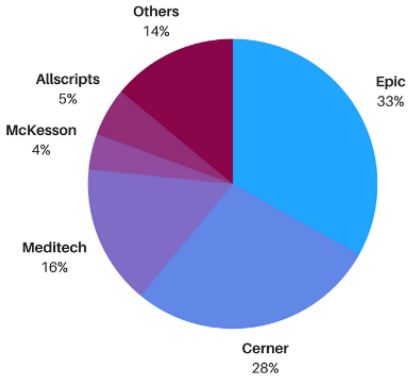
15. Signed transaction is sent to user and Providers (record receiver) to store on their respective block ledger through the DHT network (DHT also stores shard of transaction)
16. Provider app implements integrates with providers EHR system
17. ETL process map records into a sharded version-controlled record
18. Records are then digitally (Crypto) signed, encrypted, and sent to the user node IF only a personal request for the records.
19. Personal records are received public key and medical provider ID signed transaction that is recorded on users block ledger.
20. Personal records and transaction headers are locally stored.
21. If records are being sent to on-network provider (record receiver) Digital encrypted transformed records are versioned, encrypted, and dispatched through the DHT network to block the ledger of the provider (record receiver).
22. Users app records all transaction with public key of of provider (record receiver) provider (record holder), medical license id, address, and provider contact info

Providers EHR Integration:

Currently these providers hold the top market shares in global EHR systems.

and features.

1. Integration tool to existing EHR systems
    - Cener
    - Epic
    - AllScripts
    - athenahealth
    - CareCloud
    - DrChrono
    - eClinicalWorks
    - GE healthcare
    - greenway Health
    - Kareo
    - Meditech
    - McKesson
    - NextGen
    - Practice Fusion



The following Figure 13 displays the sequence diagram for the medical record release and request process i.e. *General Authorization for Use or Disclosure of Health information*.

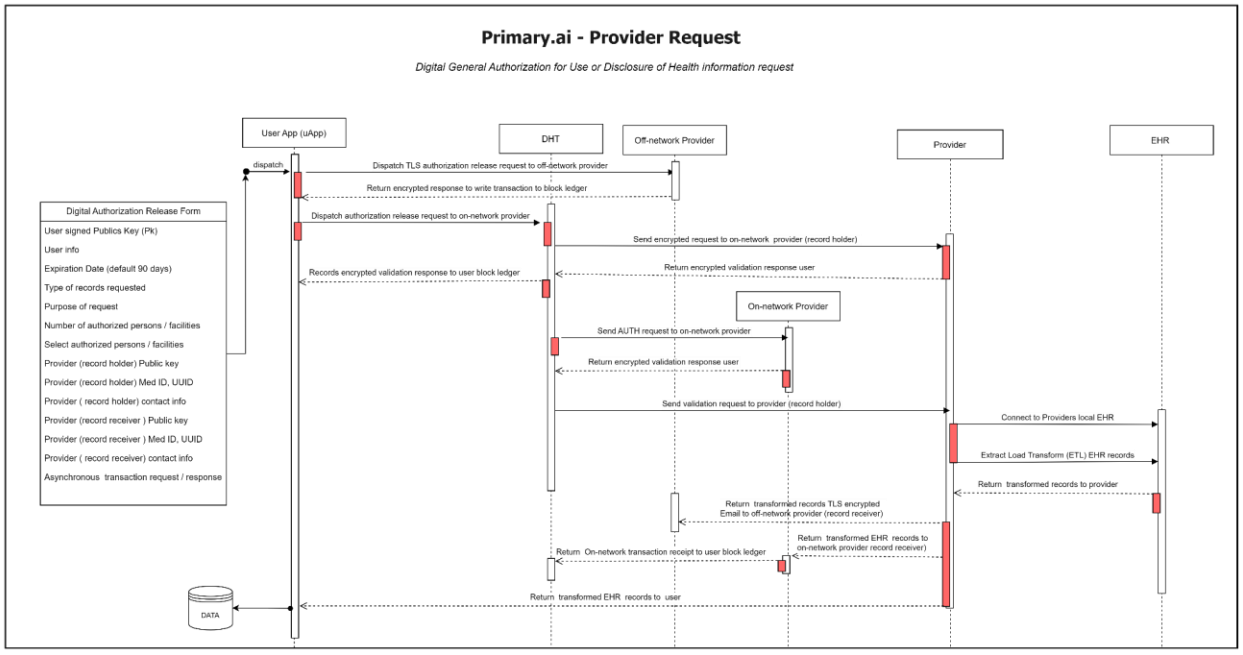


Figure 13

Provider request real time health data from patient Medical IoT wearable devices log:

The following describes the logical process to handle access to patient wearable devices real-time / time-series health data access. This feature can help providers set custom metrics to monitor and get alerts on their patients allowing them to provide optimal patient care. This feature must be authorized by the user/patient as shown in Figure 14 below.

MIoT Authorization sequence:

1. Provider dispatches MioT records API access request to user (UUID, medical ID, and Public Key signed request).
2. User/patient receives a dispatch request with a signed public key and medical UUID.
3. User accepts authorizes access to the dispatch request.
4. User sends validation key and encrypted public key API access to the provider.
5. Provider receives response and validates signature of response hash and meta.
6. Provider returns request to access user MIoT API time series stream.
7. If successful, the request will open a live stream channel to all health data features being. recorded by the user in real time via the provider interface.
8. If not successful, a response is sent to the provider denying access or throws an error.

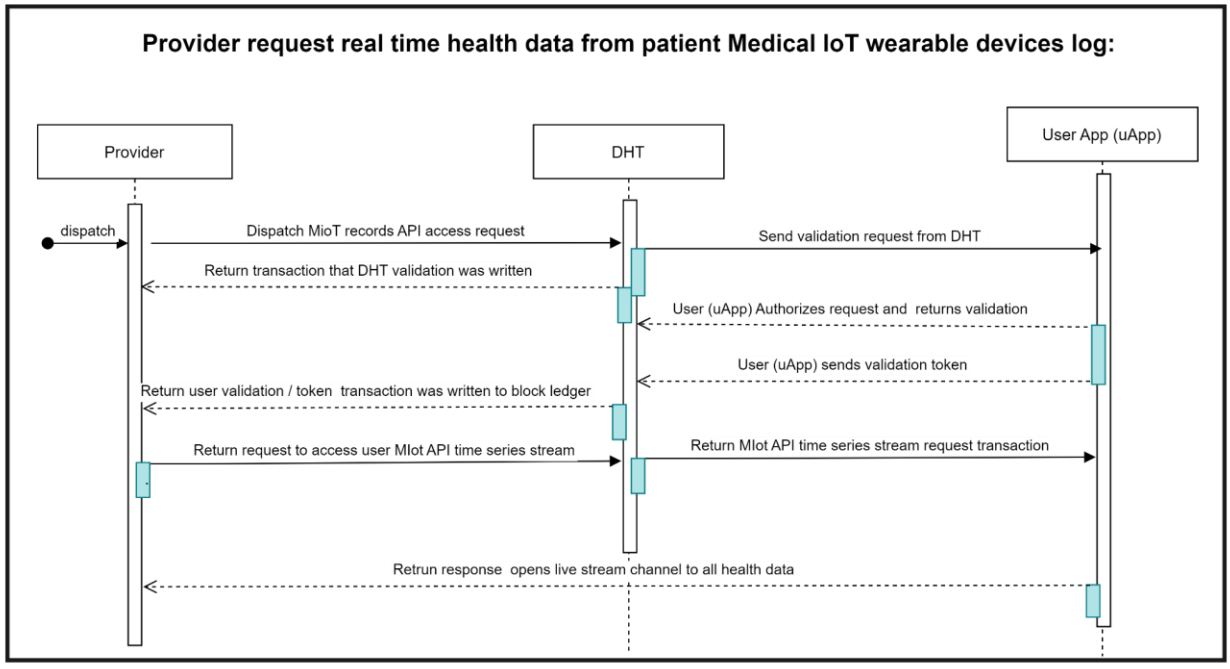


Figure 14

Provider application UI

The following outlines the UI/UX screen and behaviors of the provider application on the Primary.ai network.

1. Initial provider setup screen
    a. Sets up public and private key
    b. Medical ID is hashed and displayed as unique UUID
2. Integration with EHR system screen
    a. Can connect, discover EHR and prepare to import data
    b. Can import Patient profiles
        i. Import patients reference info if they have no app or node?

ii. Import bass names and info of patients from current EHR for discovery

c. Can import patient medical records

i. Providers can discover their patients who are on the network

ii. Once patients are discovered on the Provider app the provider app will send a request to patient, patient can match or click a setting in their app Then patient and provider will be connected.

d. Can do Same discover process above for providers on=network

3. Dashboard to access all patients screen

a. Each patient card is signed with user's public key

b. Each patient can submit a digital signed verification of medical record request.

c. Record gets sent with patient's public key

d. Record is used and shared on Providers node and ledger * version transactions immutable.

e. Provider can send records upon request to on-network providers

f. Providers can send records to off-network on behalf of patient via TLS email

4. Single patient screen

a. Patient profiles

b. Profile shows all transaction of medical records

i. DHT transactions (on-network)

ii. TSL transactions (off-network via encrypted fax)

c. Shows signed digital signature of digital release document with encoded user public key and hashed verify Q-code.

i. Shows history, transactions, and times

d. Outside health stream data (only if patient grants access) MIoT device data

e. Patient notifications

f. Number of records sent

g. Number of records accepted

h. Number of records errors

5. List of on-network providers screen

a. Can see list of on-network provider cards

b. Each Provider has a public key and node on the network

c. Only records shared with other providers can shard the records

6. List of off-network providers screen

a. List of providers off network

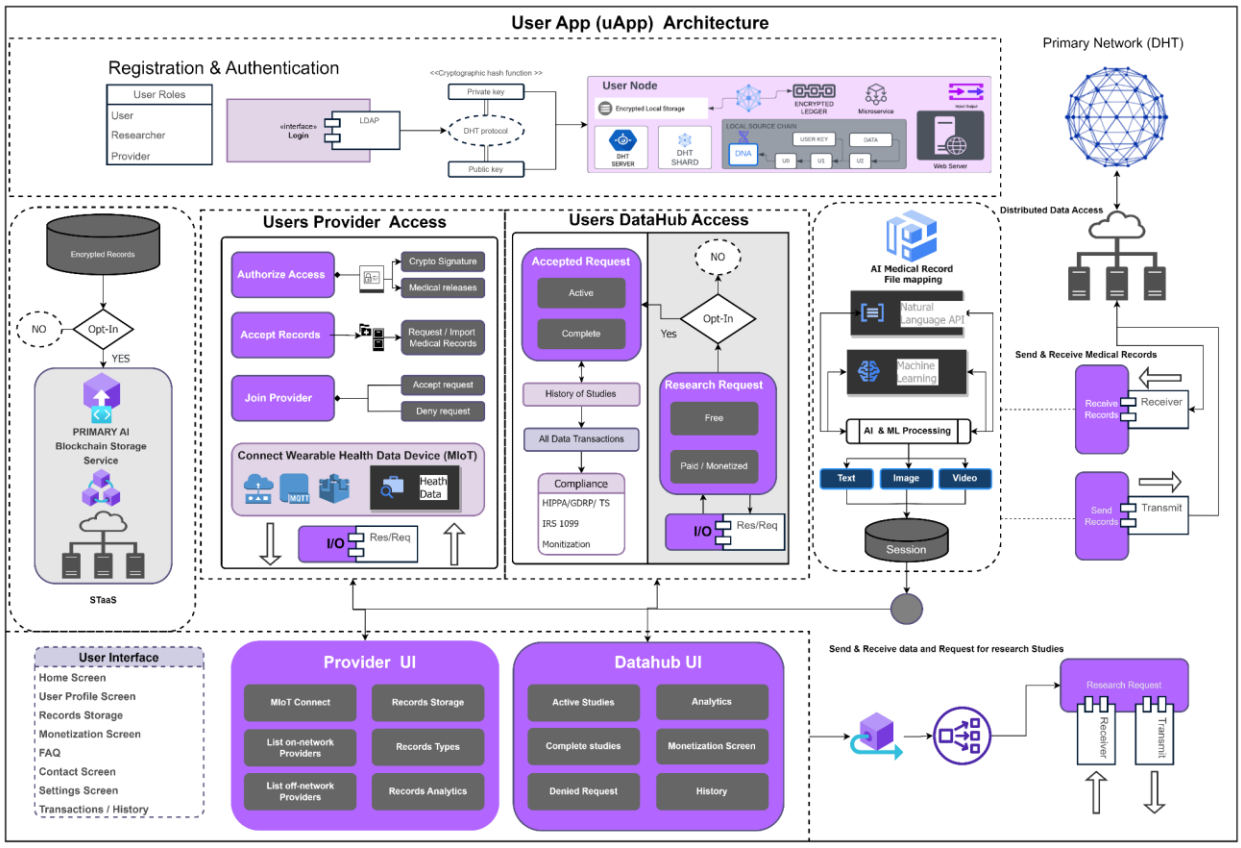
b. Providers card will have signed by the provider who sent the records

7. Provider data access usage profile screen
8. Provider account screen
9. Stats and analytics screen
10. DHT transaction screen

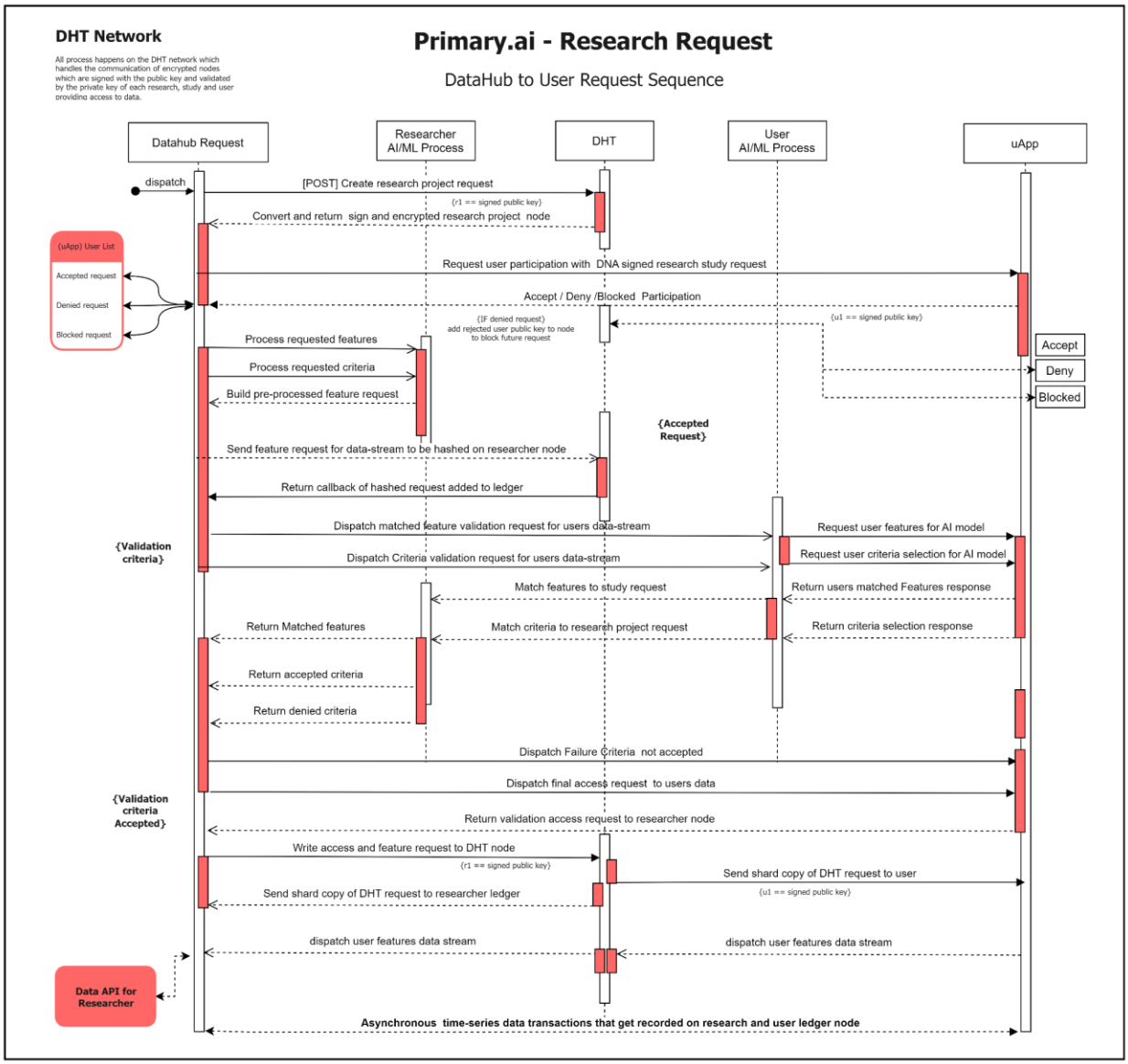
User Application (uApp)

The user-centric application relies on several key technologies to store, share, and provide access to user's personal health and medical information. For each transaction, i.e. adding data to records, adding new provider information, deleting data from records, etc., users are assigned encrypted private and public key sets. These keys are used to sign and verify every transaction in the data and for each action assumed on the data i.e. sharing data. As in version control systems, such as GitHub, nothing is ever deleted; this is defined by immutable commit actions to the application chain that are updated only by being added upon the main records creating a new version on top of the master record. The following Figure 15 shows the architecture for the user application and logical flow diagram of resources available in the user app (uApp). This includes the artificial intelligence systems process, API access dataflow, records storage, request processing, access to

Providers, access to DataHubs, and MIoT Medical IoT synchronous data interface.



By verifying the user's private and public keys, the uApp maintains the persistence and integrity of the original data. It is important to note, there is never any deletion of real data unless the user deletes all of their records. Let us say, if a new record is added or a provider removes something from the patient's record like a prescription. That data will still exist on the internal ledger system of the application. It creates a new version by adding it to the user's internal block and validating it with the user's keys and shared authorized access rules the user gave permissions and access to. As part of the internal ledger system, all transactions within the history of data are also saved as immutable records. The following Figure 16 shows the sequence for authorization and access of the user app (uApp)



Each user has their own instance of this application, meaning all of their devices can join this app, making their health information accessible to providers or datahubs. In order to preserve the integrity of the system, we do not permit users to share their data directly with other users. The P2P Distributed hash table (DHT) network will only allow providers(pApp) and datahubs (dh) to accept incoming user data based on the user's role; users will not be able to accept requests from other users. In order to validate access on other user devices, users will need to use their private keys.

As a result, key barriers to adoption are addressed by establishing clear data integrity and traceability for users' data from official sources, including physicians, scientific studies, and researchers. There is a concern among health care practitioners and researchers that health records that have been in patient custody may be incomplete or have been altered, making them hesitant to trust them. Due to the immutability, non-forgeability, and non-repudiation of each user's records, healthcare practitioners can be assured that the observations or diagnostic reports are accurate and complete.

uApp Dashboard

Users/patients will have access to the following tools and resources within the uApp.

- Import sign and verify medical records with AI
- Datahub central
- Datahub transactions

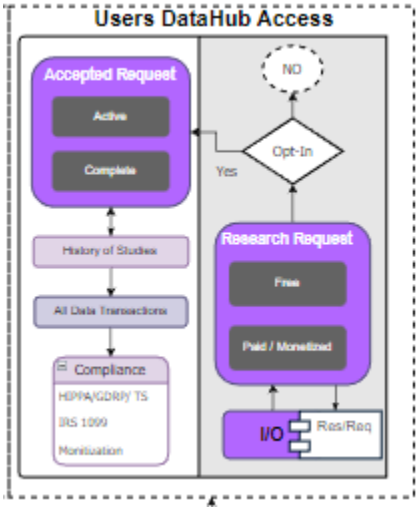
- Providers and provider transactions
- Wearable and mIoT integration
- Data visualization feedback
- Basic profile and settings
- Analytics dashboard

uApp cryptographic verification

Medical records can be requested from providers currently using Pirmary.ai and verified digital copies of the records can be imported. A majority of medical providers may not have access to Primary.ai or will not be using it at the moment. Because of this, we have integrated artificial intelligence and file mapping to import, sign, and verify raw and formatted data into the medical record application. With an artificial intelligence model, the application extracts text, image, and binary data and converts it to structured data while maintaining data integrity and adding mIoT health data preemptively. Detailed explanations of the AI models and process will be provided in the following sections.

Users will also have access to all transactions to and from both providers and datahubs for research studies, and including all versions of the initial data records, versions, and mutations of outside data. This is an important concept and like version control systems this is essentially what we are doing with patients records while ensuring the integrity of the data through blockchain-based signature process and internal app ledger that is immutable, non-forgeable, and non-repudiation.

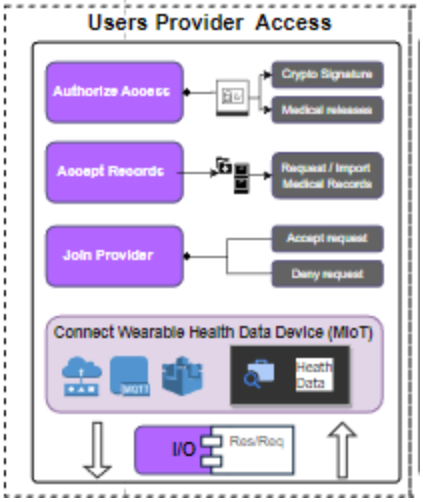
uApp Datahubs UI



From the datahubs, the users will have access to an in-app feed of open studies where they can read the criteria and monetization level for the studies and determine if they want to submit an inquiry to the study. Since the user and study are not sure if the data will be a fit for the study. We propose an AI based recommendation system which matches the users structured data by key categories specified by the researchers. Since we are trying to maintain and ship an initial minimal viable product, the recommendations system will be a future endeavor. In the meantime, a sample of the user's data will be selected, and the user can check or uncheck what fields of that data will be sent to the study to see if the data will be a fit. Once the study confirms the sample a handshake will be sent to the user to accept manually in the form of an alert. Once the alert is accepted, the requirements for which data points needed for the study will be extracted into a data secure dataset cryptographically signed with the researchers public key and the user public key and transaction will appear on the dashboard screen of both researcher and user. One issue has been made present, that a user only has to manage their data which is a one to one relationship with the datahub and researcher.

While the data hub will need to validate many users to contrive a complete data set for scientific research and study. This issue will be solved by the verified researchers setting each field of data they need for both the samples and the data set. If a sample is accepted but does not forward all the fields required for the studies data set it will be rejected. This will be an automated process on the researcher's application to ensure only valid data that meets the requirements is coming through. Also, the researcher must specify how big the sample of data they need is and how long they will need it for. Provisions will be put in place that ensure the researchers set the study criteria correctly or else the study will remain unpublished until it is met. Alert notification will be sent to users over the encrypted network to notify them if they have not met all criteria and fields. Since the mapping of data and patient records into the Primary.ai controls the mappings of data columns the uniformity of data from researcher to user will be predetermined.

uApp Providers UI



Similarly, the user will have access to the providers screen that will show all transactions from each provider in their own separate module per transactions per provider and a screen to view them collectively. The idea of this technology from a provider's perspective is to have instant access and authorization as described previously. From the user perspective it allows the patient to grant access with a single click which also acts as a recorded , verified, and signed transaction to ether accept records coming in to be transformed or to allow providers authorization to send the records to other providers on the Primary.ai network or to have verified authorization to submit records to providers without it. The user will have a Hipaa compliant authorization form signed with their public key that gets sent to the provider requesting authorization. An encrypted copy of this digital authorization will also be sent via email to the provider for their records regardless if they participate on the Primary.AI platform or not. The users will also be able to manually add providers requesting and receiving access to their records by entering in providers basic info and medical license information and email. This will allow the user to send an instantly signed authorization form to that provider granting them access to records from an off-network provider or allowing them to send the records. This functionality will use ISO 27001 certified TLS encryption to send the users verification and transmit it via email from the Priamry.ai networks while creating an audit trail in the DNA of the user's application.

IoT devices, especially medical IoT devices, are being used more and more because of Bluetooth and Wi-Fi technology advancements. A medical device with Wi-Fi and Bluetooth is used to transmit health and medical information via the mIoT, which relies on machine-to-machine communication. By 2026, this market is expected to account for $94.2 billion dollars, according to Market data. If the user currently uses mIoT devices such as Fitbit, smart watches, and other wearable devices to track health data. A user can opt in, connect their device, and transmit data to the primary.ai application in real-time. Primary.AI will enable medical providers to access real-time patient data (RTPD) such as body temperature, blood oxygen level, systolic and diastolic blood pressure, heart rate, sleep patterns, etc. By creating a separate, verified and signed medical record in the users block ledger, Primary.ai acts as a trusted source of data transformation for vitals and advances in patient care. The user has full control over granting access to that data to providers and data hubs. By doing so, providers and patients are able to access new opportunities that are not currently available outside of a hospital or medical facility and provide providers with access to critical patient data in real time, which can be used to diagnose and prognosis future events, thus bringing more value to both parties. mIoT devices will be able to record user records for

uApp Data visualization feedback

Priamry.ai will also provide data visualization feedback to show history versions and graphical charts and visualizations for data usage, transactions, history, data usage, errors, and other important user benchmarks.

uApp Basic profile and settings

Initially, the uApp will include a basic user profile and account information. Basic user information will be displayed along with obfuscated private and public keys. It is important to note that each user will only get one set of private and public keys. Afterwards, these encrypted keys cannot be restored and must be stored in a safe place designated by the user. The user should store these keys in a hardware wallet or cold storage wallet in order to ensure their safety. Users may hard reset the application if they lose their devices, have them stolen, or misplace their keys in order to delete all medical data and transactions, which requires them to create a new account. Due to the fact that the hard data is stored locally on the user's device, they will be responsible for ensuring backups of their medical data. This could be a potential future service for users to secure and securely store encrypted backups that can only be decoded by the users themselves. Furthermore, it is recommended that users not only set up their accounts on one device if they have a home computer or table, but set them up on two devices so that they can continue to manage their data without losing it. A built-in feature that can be accessed from any other user device also ensures the integrity and security of their data in the event of a stolen device by hard erasing all data on the device, similar to self-destructing mechanisms. With basic settings, users can toggle between features, notifications, in-app alerts, delete

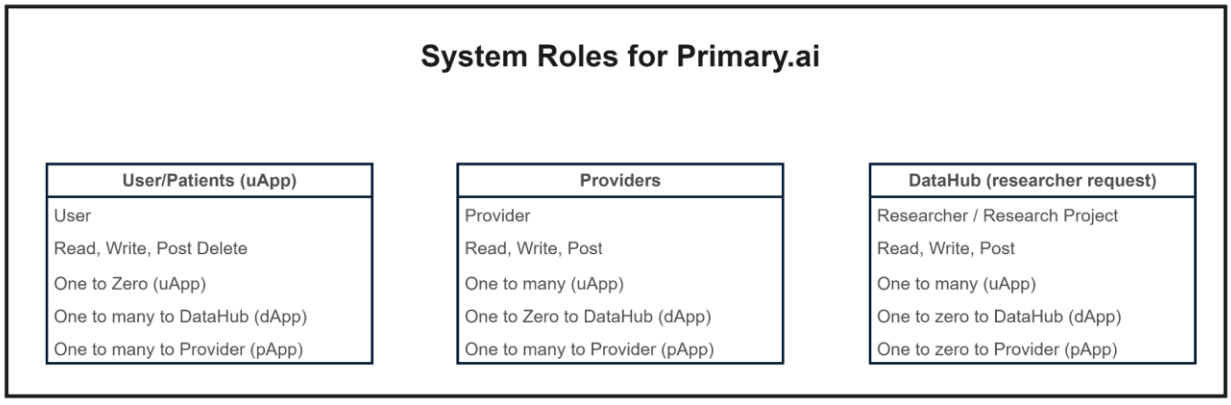
accounts, and self-destruct a lost or stolen device that cannot be recovered.

Scope of Primary AI

The following is a simple scope of the pages and screens that will be implemented for each user type. Also, the following shows the roles based for each user

1. **Home page** - access
2. Login auth - private & public key-based node
    a. **Patient**
        i. Access public and private keys/node/DHT/personal block ledger
    b. **Provider**
        i. Access public and private keys/node/DHT/personal block ledger
    c. **Researcher**
        i. Research project
            1. Access public and private keys/node/DHT/personal block ledger
        ii. Access public and private keys/node/DHT/personal block ledger
3. **User/Patient portal**
    a. User account
    b. Profile
    c. Public/private Keys set
    d. Settings
    e. Dashboards
    f. Tools
    g. Data Storage
    h. Analytics
    i. MIoT Sync dashboard
    a. Transactions history (per connection)
2. **Provider Portal**
    - User account
    - Profile
    - Public/private Keys set
    - Settings
    - Dashboards
    - Analytics
    - On-network providers
    - Off-network Providers
    - MIoT Patient data streams
    - Transactions history (per connection)
3. **Research portal**
    - User account
    - Profile
    - Public/private Keys set
    - Settings
    - Dashboards
    - Create research study project
        - Request data access users page
        - Data API
        - Data tools
        - Transactions history (per user connection)
    - Invite researchers
    - Transactions history (per connection)

Roles and access management



1. Providers and Researchers will have a private individual node upon assignment

2. With public key
3. With private key
4. Providers and Researchers can create groups
5. Groups will accept multiple user nodes
6. Roles will be creator, editor, publisher, reader
7. Groups will be a multithreaded node with its own DHT
8. Each individual will also write to the DHT both individual and group will have blocks on blockchain
9. Each transaction from an individual or group will become a new version of the latest record.
10. Providers upon patients/user authorized access verification hash. Will be able to transmit data records to other providers on or off network
11. Providers upon patients/user authorized access verification hash. Will be able to accept incoming transmission of patient medical or health records.

Primary.ai Artificial Intelligence use-case

With the proposed technology, we aim to solve several business problems utilizing state of the art machine learning and artificial intelligence algorithms and processes. As a result, we have defined the following business problem.

Business Problems

- Importation and classification of medical records
- Research project recommendations system
- Managing rulesets and transactions

Keeping patient records accessible seamlessly is the first business challenge. Since there are hundreds of electronic health record systems that have different formats for medical data images, videos, and sounds that could be attributed to patients' medical records, it is unlikely that all patient data will be structured and easily accessible.

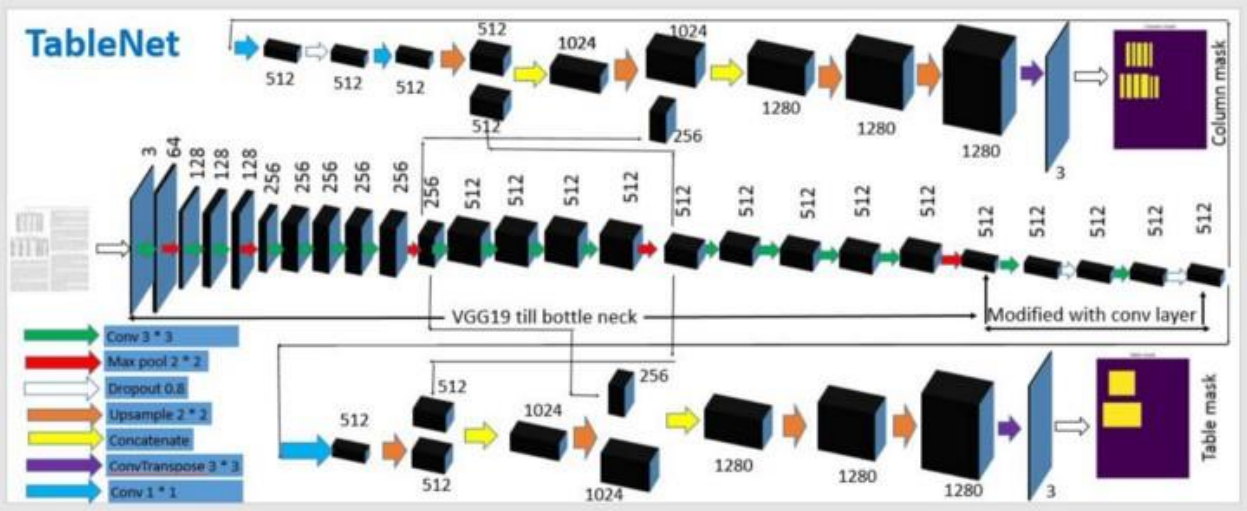
In a database (DB) format, the EHR/EMR stores and manages patient health information, personal information, medical history, drug reactions, health status, medical examination, and admission/discharge records systematically. Generally, EMR data can be classified into structured, semi-structured, and unstructured depending on the degree of structuring[32].

A Structured data set refers to data that has been structured to follow a predetermined format and structure during storage.

A semi-structured data set is a set of data whose format and structure can be altered. In addition to the actual data, it provides structural information about the data.

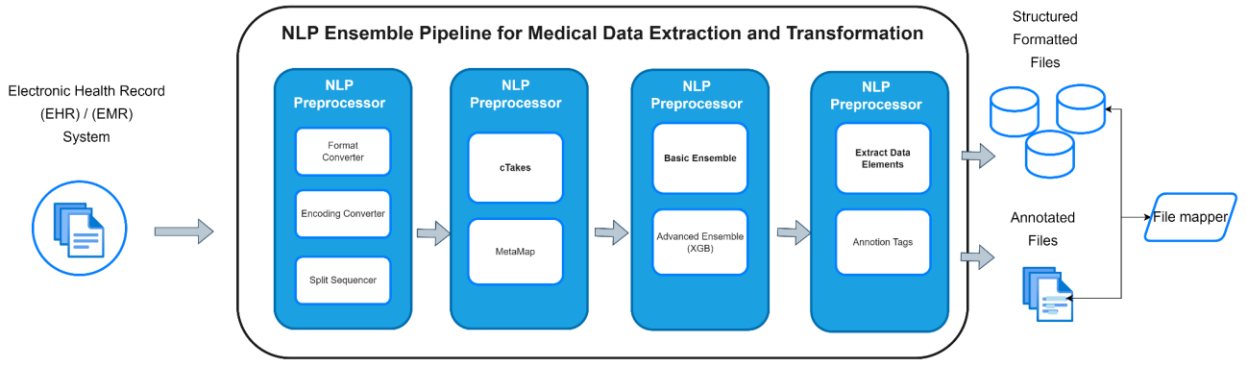
An unstructured data set is one that does not have a defined structure. As a result of their irregular shapes, they are hard to define. Unstructured data is typically represented by text and images. Unstructured data may include radiographic images, videos, or photographs.

As we cannot predict every outcome depending on the type of patient record, we have designed a logic system to handle the internal structure format through our file mappers that structure the data into our database. The ability to extract text, image, and video data with reasonable accuracy can be achieved with several tools, but the task of labeling and piecing the information back together in a machine-readable format is much more challenging. We propose using TableNet deep learning model for the detection and data extraction of scanned document images[33]. TableNet uses a single input image to generate two distinct semantically labeled output images for tables and columns. For the base network, the model uses pretrained VGG-19 Convolutional Neural Networks with two decoder branches based on VGG-19 features. Table region segmentation is carried out by one decoder branch, while column region segmentation is carried out by another. Both table and column detectors use VGG19's encoding layer, but the decoders differ between them. During the training process, both the table and column detectors are used to calculate gradients, while the decoders are learned independently [34].



Not all data will be unstructured so we want to be able to detect when certain electronic data sources sent to the patient have a structured format, we can use to preprocess the incoming data source. Simple machine learning and feature extraction can be used to implement this pre-processing. In order to evaluate the accuracy of incoming records, we created baseline models of our internal data structure and compared the classification algorithm. To classify and categorize incoming semi-structured and unstructured data, we use entity extraction methods and custom fields. Future pre-processing can be sped up by adding more categories[32], [35], [36].

Once patients have successfully imported, transformed, and processed their records into the correct internal structure, viable data will be allocated for query responses to the datahub upon the user's authorization per request. Using NLP model with classifier algorithm to match Datahub research categories (similarly to job matching algorithms) to match best fit research requests to the user's data. NLP, or natural language processing, is a branch of artificial intelligence that allows computers to understand text and speech like humans [32], [37]. Among the most efficient methods for fitting linear classifiers and regressors under convex loss functions, such as linear Support Vector Machines and Logistic Regression, we have chosen Stochastic Gradient Descent (SGD). Additionally, we consider XGBoost and random forest implementations[36], [38]. Upon granting access to the datahub and opting into research projects, the datahub will have separate internal ledgers of the user and their key data features available for future requests. These features will be allocated to a distributed collaborative recommendations system unique to each user and the records acquired. The recommendation systems can be thought of as an AI microservice for each user[39]–[42]. Our future plan is to use NLP to create a personal assistance feature in future beta versions of the platform to manage rules, notifications, and alerts. In addition to this whitepaper, more detailed research analysis and formal papers involving the chosen algorithms and implementation will be provided. The following Figure 17 shows the AI and ML process with NLP for Primary.ai



Technology Stack

This technical specification describes the technology stack used in the current alpha build of the Primary.AI application. The application currently consists of a blockchain-based P2P network that interacts with a user device or a designated server for a backend. The current front-end is web-based and mobile-based. It is not necessary to mention all internal packages in order to secure risk and vulnerabilities. A UX/UI design system allows us to avoid relying on third party libraries that may pose a security risk. For brevity I allocated links to the technologies resources as opposed to referencing vague citations.

Backend technologies:

- Linux
- Rust
- Nix
- Node.js
- Holochain
- Spring Boot
- Spring Boot security with LDAP

Data Persistence:

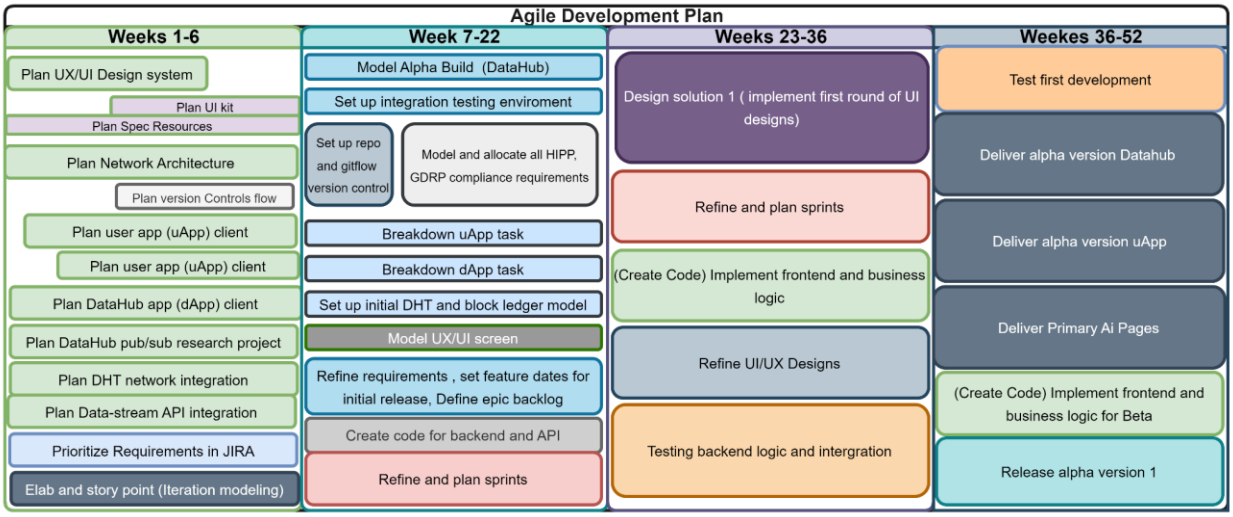
- P2P DHT based chain protocol
  - SQLite
  - Postgres
  - Influx Data
  - SupaBase

Front End Web technologies

- React
- Next.js
- React Native (mobile device)
- Rust
- SCSS
- Axios

Agile Workflow

Here is our proposal and plan for achieving alpha release version one within the first 52 weeks. Currently, the prototype version is being modeled and hosted locally on a private repository locally. intend to focus on strategic planning and deployment of an alpha version one within the first year following the completion of the initial prototype.



Market Evaluation

According to market data, Electronic Health Record (EHR) Market size was valued at over USD $29.5 billion in 2020 and is expected to grow at a CAGR of more than 6.4% between 2021 and 2027 with an evaluation of $30.1 billion-dollar industry[10]. According to market research, The globa❖ Medical Internet of Things (mIoT) in the healthcare market was valued at USD $73.5 billion in 2021 and is all set t❖ surpass USD 190 billion by 2028, exhibiting a CAGR o❖ 25.9% during the forecast period 2022-2028[11]. Researcl❖ shows that global artificial intelligence in Healthcare Market Size & Share to Surpass $95.65 Billion by 2028[12]. The Global Blockchain Technology in Healthcare Market size is expected to reach $5.3 billion by 2028, rising at a market growth of 39.9% CAGR[13].

Total Addressable Markets:

Electronic Health Record: $29.5 billion

Medical Internet of Things (mIoT): $73.5 billion

Artificial Intelligence in Healthcare $95.65 billion

❖ Blockchain in healthcare $5.3 Million billion

Primary.ai Services

Currently, these services are hypothetical and will most likely be refined over time to determine market need and success. We have 3 potential sources of revenue, user/patients, Researchers, and Providers, we also have the potential for partnering with services providers and affiliates in-app. For graffiti we will only address the 3 main categories of revenue in the business model.

Here we will define service and baseline evaluations of the following:

- Total Available Market(TAM) (add up all product sales across the market)
- Share of relevant Market (SAM) (add up only relevant product sales across the market)
- Potential Market (SOM) (divide last year's revenue by last year's SAM)

Researchers services

Services we provide are based on 15%-25% of research studies grant revenue for monetization, user fee based on tier level research access as follows. Researchers with grants will allocate funds to monetize individual studies and that offer monetization to users and their personal medical data. There are approximately 7.8 million active researchers, 141,740 medical researchers, and 300,000 AI researchers' practitioners. The average grant for each research project according to {} is approximately $581,293 dollars [46]–[48].

- **Tier 1** - free access (limited access to research tools, timed research studies, and minimum request for research studies.
- **Tier 2** - $25.99 per month of $300 a year for full access.
- Special research studies based on rate and quantity of data being requested $0.05 per sec/hour or customized data fees.(based on server load resource adjustments)

Based on the following data we estimate, and evaluation as follows:

Evaluate (TAM) for Researchers

Total Market size of viable researchers = 7.8 million

Annual Contract Value * Number of Possible Customers = $2,340,000,000 (2.3 billion)

Evaluate (SAM)

Share of relevant Market size of viable researchers = 7,23,033

Annual Contract Value * Number of Possible Customers = $216,909,900

Evaluate (SOM)

No data is available for this calculation

Provider services

The total addressable market has a potential value of $203.95 Billion for providers who could potentially utilize the service provided by Priamry.ai not to mention the security benefits which will potentially save these providers millions of project revenue from security loss as mentioned above previously[49], [50].

The service we plan to offer is full integration, maintenance, and development into custom EHR and present EHR systems. This will require custom integration and development on our part. Our main focus will be the following EHR providers. Custom integration will have to occur with hospitals EHR. The average value spent by Hospitals is $162,000 for development and $85,000 per year for maintenance. We are estimating an average integration fee of $80,000 and maintenance fee $45,000 on the low end. For this analysis we will calculate the average. As a software service that is fully integrated with EHR mapping. We will have a custom fee between $12,000 and $4,000 per month to providers which includes maintenance and integrations, helpdesk, and customization of data fields. Monthly fees for per patient basis would be an initial $10, 000 setup and integration fee and $60,00 per patient. This includes private practice physicians, insurance, specialty, and other classifications of a provider or owner of medical data[50].

Provider Evaluations

Evaluate (TAM) for Providers

Annual Contract Value * Number of Possible Customers = $1.218 trillion globally

US hospitals 6,093 * $162,000

US Annual Contract Value * Number of Possible Customers = $987,066,00

US providers 1,073,616 * $10,000

US providers Annual Contract Value * Number of Possible Customers = $10,736,160,000

Evaluate (SAM)

No data is available for this calculation at this time

Evaluate (SOM)

No data is available for this calculation at this time

User Evaluations

Currently we have no data until we reach the alpha release, we initially discussed taking a surcharge percentage of monetization fee for maintenance, resources i.e (servers, time, space) and operations of the applications. The initial estimate is 10% off each user patient who is currently active in a monetized study. Once we on-board user we facilitate this calculator and review the evaluation further. We also plan on hosting a secure DHT based records server that will be impenetrable and un-hackable with obfuscated shared distributed persistence hosting that can only be decrypted and accessed by user's private key. These are future speculations as of now.

Potential competitors

The following are the current startups disrupting the healthcare industry with blockchain-based technology. According to global healthcare blockchain marketing data Blockchain in healthcare is a newly formed but large market that is expected to grow to $5.61 billion in size by 2025, maintaining double-digit growth along the way. As we speak, the blockchain is already conquering the healthcare industry and making it more customer-centric, personalized, pleasurable to interact with, and thereby, much more effective in treating diseases and disabilities.

These are a few blockchain-based platforms or startups which could prove invaluable as potential competitors or sources.

- *BurstIq – Medical record data security and early warning drug abuse blockchain technology to safely and securely manage patients' data.(this is one to watch)
- Chronicled – The startup leverages blockchain and the Internet of Things (IoT) to power smart, secure supply chain solutions.
- Accenture – Blockchain Mobile App and Web Application
- DD KOIN – Leading Cryptocurrency
- Medicoin – Healthcare Blockchain-Based Web Application
- Clinico – Builds a patient-centered clinical trial data-sharing community.
- Coral – Provides technology-based solutions to accelerate care delivery, automate multiparty administrative processes and improve health outcomes.
- Curisium – Tracks contract performance in real-time, and prospectively simulates alternatives.
- iSolve – Complements existing systems and processes to ensure data provenance and create an interoperable and high-performance environment focused on improving patient outcomes.
- Medicalchain – Utilizes blockchain technology to securely store health records.
- Patientory – A global population health management software giving users access to their health data.
- Pokitdok – Develops APIs for healthcare verticals such as claims, pharmacy and identity management. Their platform, DokChain, is a distributed network of transaction processors operating on both financial and clinical data across the healthcare industry.
- Dentacoin – is a blockchain-based cryptocurrency that connects patients and dentists.
- BlockPharma – Blockchain App that tracks drugs throughout the supply chain, from manufacturing to the final user.
- BlockPill – is a scalable distributed ledger technology to provide you a safer medicines prescription and facilitates health professionals' with blockchain technology.
- Encryption – free market for genetic testing DNA data
- Clinicoin – is an open source wellness platform that rewards users with cryptocurrency for engaging in healthy activities.
- Iryo – first open healthcare protocol for the secure and private exchange of medical data.
- Solve.Care – Blockchain platform for digit health networks
- Doc.ai – is an enterprise AI platform that unlocks the value of health data.

Scalability Evaluation

Based on our current estimates with proper backing and support we believe we can have a minimal viable product within 7-12 months with the ability to scale future feature enchantments while opting to all marketable channels through a scaled business plan.

Future work and challenges

In the beginning, we will use Rust assembly language to complete the blockchain system model, then we will design user interfaces and UX/UI models for the client applications to interact with the backend and API of the applications. We will develop a minimum viable product that includes a web/mobile user/patient application, a Datahub portal, and a test provider portal that adheres to current digital identity and EHR standards. There will be basic functionality and the ability to manually import both textual and image-based records in both digital and hard copy formats through the user application. As a result, they will have the ability to grant access to providers, or to Datahub researchers with basic visualization tools for monitoring all events and actions regarding access and governance of their medical data. Through the request-response handshake protocol, Datahubs

will provide users with an asynchronous connection to their data, and users will be able to post basic requests for data. As this will be an alpha version, we will initially allow only a limited number of researchers who have expressed interest in using the platform from the early stages of development to help test its effectiveness.

In the first phase of the alpha version release, we propose to develop the provider's application, which we will model so that it can be retrofitted onto existing EHR and EMR systems. In the field of EHR systems, Cerner and Epic are the leading names; therefore, our focus will be on integrating file mapping with these platforms, as well as others if necessary. Additionally, we will advance our AI systems for managing data streams and importing medical records and images of patients. When testing and market viability have been established, we will preclude the initial beta release phase. We also propose to build a DHT based storage server which will be a secure lock and key storage facility for patient records who do not want to them on their device. Our plan is to use a greenfield Agile methodology. Challenges will include building our team and gaining public interest in our core mission and product. We would like to partner with established tech companies in the medical space to help drive confidence and leadership in this project. This whitepaper will act as a living document allowing us to add additional features and research to optimize the full solution of the Primary.ai platform.

Resources

[1] T. Lakshmanan and M. Muthusamy, "A Novel Secure Hash Algorithm for Public Key Digital Signature Schemes," *The International Arab Journal of Information Technology*, vol. 9, no. 3, 2012.

[2] "Would you sell your own health data? These companies want to help you." https://www.fastcompany.com/90409942/would-you-sell-your-own-health-data-theres-a-market-for-it-but-ethical-concerns-remain (accessed Nov. 06, 2022).

[3] "How Can Patients Make Money Off Their Medical Data?" https://news.bloomberglaw.com/pharma-and-life-sciences/how-can-patients-make-money-off-their-medical-data (accessed Nov. 06, 2022).

[4] "HHS cyber arm warns of EHR vulnerabilities | Healthcare IT News." https://www.healthcareitnews.com/news/hhs-cyber-arm-warns-ehr-vulnerabilities (accessed Nov. 20, 2022).

[5] F. F. Ozair, N. Jamshed, A. Sharma, and P. Aggarwal, "Ethical issues in electronic health records: A general overview," *Perspect Clin Res*, vol. 6, no. 2, p. 73, 2015, doi: 10.4103/2229-3485.153997.

[6] J. Adamu, R. Hamzah, and M. M. Rosli, "Security issues and framework of electronic medical record: A review," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 2, pp. 565–572, 2020, doi: 10.11591/eei.v9i2.2064.

[7] "Average Healthcare Data Breach Costs Surpass $10M, IBM Finds." https://healthitsecurity.com/news/average-healthcare-data-breach-costs-surpass-10m-ibm-finds (accessed Nov. 20, 2022).

[8] "Unrelenting cyber attacks cost health systems $10M per breach." https://www.fiercehealthcare.com/health-tech/healthcare-data-breach-costs-reach-record-high-10m-attack-ibm-report (accessed Nov. 20, 2022).

[9] "Cost of a data breach 2022 | IBM." https://www.ibm.com/reports/data-breach (accessed Nov. 20, 2022).

[10] "EHR EMR Market Size is projected to reach USD 47.6 Billion." https://www.globenewswire.com/news-release/2022/09/07/2511881/0/en/EHR-EMR-Market-Size-is-projected-to-reach-USD-47-6-Billion-by-2030-growing-at-a-CAGR-of-5-5-Straits-Research.html (accessed Nov. 06, 2022).

[11] "IoT Medical Devices Market - Global Forecast to 2026 | MarketsandMarkets." https://www.marketsandmarkets.com/Market-Reports/iot-medical-device-market-15629287.html?gclid=Cj0KCQiAyMKbBhD1ARIsANs7rEEDh42ErUUg3SpsZsszqdG7U-5eRFfnkHe_81Q9yAFjoTAABvPHwOMaAkmpEALw_wcB (accessed Nov. 12, 2022).

[12] "Artificial Intelligence In Healthcare Market Size Report, 2030." https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-healthcare-market (accessed Nov. 20, 2022).

[13] "Global Blockchain in Healthcare Market: Focus on Industry Analysis and Opportunity Matrix - Analysis and Forecast, 2018-2025." https://www.researchandmarkets.com/reports/4519297/global-blockchain-in-healthcare-market-focus-on (accessed Nov. 20, 2022).

[14] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", Accessed: Nov. 20, 2022. [Online]. Available: www.bitcoin.org

[15] "What Blockchain Could Mean for Your Health Data." https://hbr.org/2020/06/what-blockchain-could-mean-for-your-health-data (accessed Nov. 20, 2022).

[16] L. A. Linn and M. B. Koo, “Blockchain For Health Data and Its Potential Use in Health IT and Health Care Related Research”.

[17] Y. Lu, “The blockchain: State-of-the-art and research challenges,” *J Ind Inf Integr*, vol. 15, pp. 80–90, Sep. 2019, doi: 10.1016/J.JII.2019.04.002.

[18] A. Hasselgren, K. Kralevska, D. Gligoroski, S. A. Pedersen, and A. Faxvaag, “Blockchain in healthcare and health sciences—A scoping review,” *Int J Med Inform*, vol. 134, p. 104040, Feb. 2020, doi: 10.1016/J.IJMEDINF.2019.104040.

[19] “Health Information Privacy Law and Policy | HealthIT.gov.” https://www.healthit.gov/topic/health-information-privacy-law-and-policy (accessed Nov. 20, 2022).

[20] H. Office for Civil Rights, “HIPAA Privacy Rule and Sharing Information Related to Mental Health Background”.

[21] “General Data Protection Regulation (GDPR) – Official Legal Text.” https://gdpr-info.eu/ (accessed Nov. 20, 2022).

[22] A. Farouk, A. Alahmadi, S. Ghose, and A. Mashatan, “Blockchain platform for industrial healthcare: Vision and future opportunities,” *Comput Commun*, vol. 154, pp. 223–235, Mar. 2020, doi: 10.1016/J.COMCOM.2020.02.058.

[23] “What Is Interoperability and Why Is It Important? - tokenex.” https://www.tokenex.com/blog/what-is-interoperability-and-why-is-it-important/ (accessed Nov. 20, 2022).

[24] “Top 5 Challenges with Interoperability in Healthcare.” https://www.healthjump.com/blog/5-challenges-with-healthcare-interoperability (accessed Nov. 20, 2022).

[25] “Healthcare remains costliest industry for data breaches | Healthcare Dive.” https://www.healthcaredive.com/news/healthcare-breach-costs/628344/ (accessed Nov. 20, 2022).

[26] “Fewer than 4 in 10 health systems can successfully share data with other hospitals, survey finds | Fierce Healthcare.” https://www.fiercehealthcare.com/tech/fewer-than-4-10-health-systems-can-successfully-share-data-other-hospitals (accessed Nov. 06, 2022).

[27] C. Stergiou, K. E. Psannis, B. B. Gupta, and Y. Ishibashi, “Security, privacy & efficiency of sustainable Cloud Computing for Big Data & IoT,” *Sustainable Computing: Informatics and Systems*, vol. 19, pp. 174–184, Sep. 2018, doi: 10.1016/J.SUSCOM.2018.06.003.

[28] A. Alanazi and Y. al Anazi, “The Challenges in Personal Health Record Adoption,” *Journal of Healthcare Management*, vol. 64, no. 2, pp. 104–109, Mar. 2019, doi: 10.1097/JHM-D-17-00191.

[29] N. E. Cho, W. Ke, B. Atems, and J. Chang, “How Does Electronic Health Information Exchange Affect Hospital Performance Efficiency? the Effects of Breadth and Depth of Information Sharing,” *Journal of Healthcare Management*, vol. 63, no. 3, pp. 212–228, 2018, doi: 10.1097/JHM-D-16-00041.

[30] “Opportunities and Challenges of Blockchain Technologies in Health Care BLOCKCHAIN POLICY SERIES,” 2020.

[31] “Holochain | Distributed app framework with P2P networking,” 2022. https://www.holochain.org/ (accessed Nov. 20, 2022).

[32] S. Lee and H. S. Kim, “Prospect of Artificial Intelligence Based on Electronic Medical Record,” *J Lipid Atheroscler*, vol. 10, no. 3, p. 282, Sep. 2021, doi: 10.12997/JLA.2021.10.3.282.

[33] L. Wen, X. Li, X. Li, and L. Gao, “A new transfer learning based on VGG-19 network for fault diagnosis,” *Proceedings of the 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design, CSCWD 2019*, pp. 205–209, May 2019, doi: 10.1109/CSCWD.2019.8791884.

[34] “TableNet: Deep Learning Model for End-to-end Table Detection and Tabular Data Extraction From Scanned Document Images | by Devi Prasad | Analytics Vidhya | Medium.” https://medium.com/analytics-vidhya/tablenet-deep-learning-model-for-end-to-end-table-detection-and-tabular-data-extraction-from-1961fb2f97e1 (accessed Nov. 20, 2022).

[35] S. Khalid, T. Shehryar, S. Nasreen, and T. Khalil, “A survey of feature selection and feature extraction techniques in machine learning Enhanced Framework for recognizing indoor daily life activities View project Bio Medical Imaging View project A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning,” 2014, doi: 10.1109/SAI.2014.6918213.

[36] A. Cayir, I. Yenidogan, and H. Dag, “Feature Extraction Based on Deep Learning for Some Traditional Machine Learning Methods,” *UBMK 2018 - 3rd International Conference on Computer Science and Engineering*, pp. 494–497, Dec. 2018, doi: 10.1109/UBMK.2018.8566383.

[37] S. Y. Christopher Stahl, “DeepPDF: A Deep Learning Approach to Analyzing PDFs,” Tennessee , 2020.

Accessed: Nov. 20, 2022. [Online]. Available: https://www.osti.gov/servlets/purl/1460210

[38] R. Ransing, A. Mohan, N. B. Emberi, and K. Mahavarkar, "Screening and Ranking Resumes using Stacked Model," *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, ICEECCOT 2021 - Proceedings*, pp. 643–648, 2021, doi: 10.1109/ICEECCOT52851.2021.9707977.

[39] P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," *Procedia Comput Sci*, vol. 167, pp. 2318–2327, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.284.

[40] A. T. G Adomavicius, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Trans Knowl Data Eng*, vol. 17, no. 6, pp. 734–749, 2005.

[41] B. Shapira, F. Ricci, P. Kantor, and L. Rokach, "Recommender systems handbook." Springer, 2011.

[42] Q. Zhang, J. Lu, and Y. Jin, "Artificial intelligence in recommender systems," *Complex and Intelligent Systems*, vol. 7, no. 1, pp. 439–457, Feb. 2021, doi: 10.1007/S40747-020-00212-W/FIGURES/1.

[43] "Stripe | Payment Processing Platform for the Internet." https://stripe.com/ (accessed Dec. 05, 2022).

[44] "Powerful tools for your business." https://www.paypal.com/us/home (accessed Dec. 05, 2022).

[45] "Plaid APIs - Reference Docs & Support | Plaid." https://plaid.com/docs/api/ (accessed Dec. 05, 2022).

[46] N. S. Fleming, S. D. Culler, R. McCorkle, E. R. Becker, and D. J. Ballard, "The financial and nonfinancial costs of implementing electronic health records in primary care practices," *Health Aff*, vol. 30, no. 3, pp. 481–489, Mar. 2011, doi: 10.1377/HLTHAFF.2010.0768.

[47] "Market Sizing: Measuring Your TAM, SAM, and SOM | Similarweb." https://www.similarweb.com/blog/research/market-research/market-sizing/ (accessed Dec. 05, 2022).

[48] Source: BRDPI and NIH IMPAC, "Research Project Grants: Average Size," *NIH Data Book Report ID: 155*, Dec. 04, 2022.

[49] "Fast Facts on U.S. Hospitals, 2022 | AHA." https://www.aha.org/statistics/fast-facts-us-hospitals (accessed Dec. 05, 2022).

[50] "U.S. physicians - statistics & facts | Statista." https://www.statista.com/topics/1244/physicians/ (accessed Dec. 05, 2022).