collab HEALTH

Light Paper

Authors: Aaron Lee-Zucker, Garrett Ruhland, Michael Kerr, Sam Parker, Simon Abizmil, Saurabh Mathur

Date: March 28, 2022

Status: Draft

1.0 WHAT IS COLLAB.HEALTH NETWORK

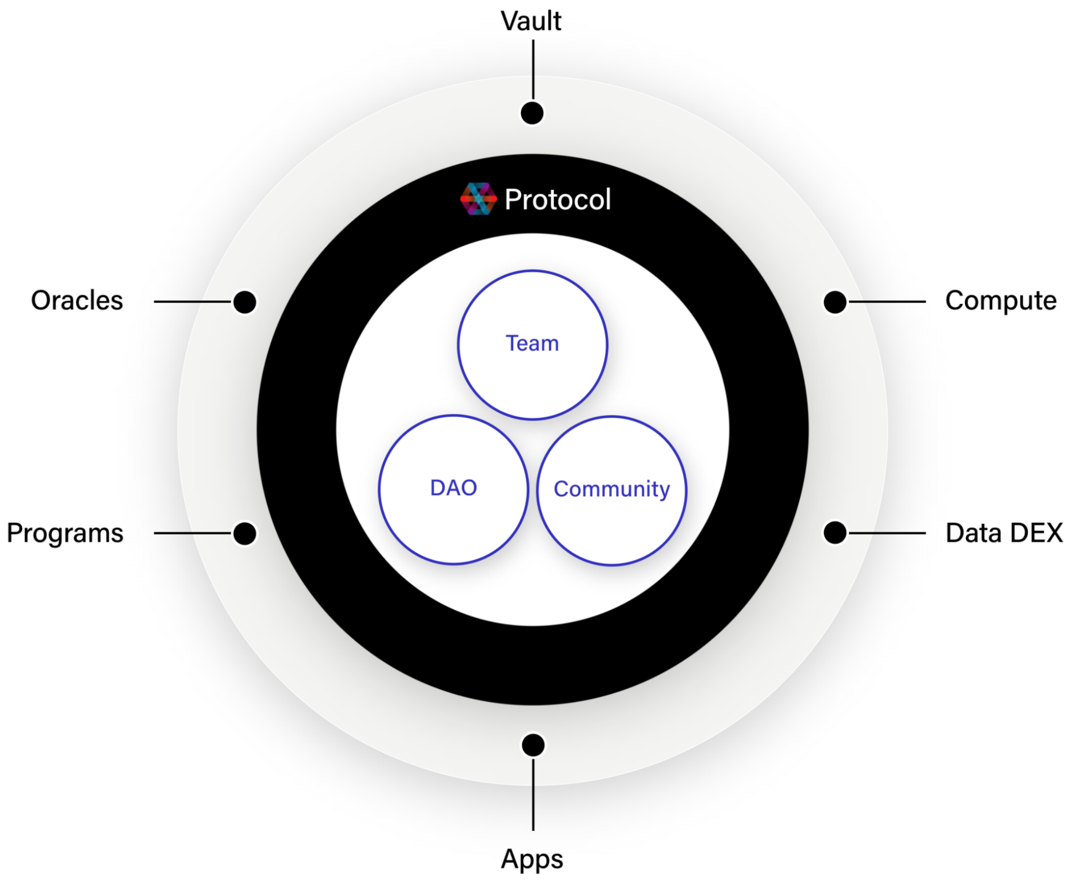
The Collab.Health Network (CHN) is a decentralized protocol that enables individuals to custody and share personal health data; receive interpretation and analysis; participate in research, and be rewarded for their participation and contributions. CHN realigns incentives for the network of stakeholders that participate in the creation of personal health data (providers, labs, wearables) such that it is more profitable to enable permissioned use of data than it is to hoard it.

- The Network offers recurring and ongoing rewards to all participants in a virtuous cycle that empowers individuals to optimize their own health. Consumers, researchers, clinicians, influencers, developers, data providers, and other participants are each aligned with an incentive structure that recognizes their contributions.
- Allows all participants to collaborate by sharing their strategies, plans, experiments, data, analyses, algorithms, and recommendations.
- Empowers and encourages individuals to collect their health and wellness data into self-sovereign custody for privacy-preserving research, application, and monetization.
    - Enable consumers to privately share their data with trusted collaborators and stakeholders to improve decision-making around their own health and wellness.
    - Enable consumers to monetize their data for commercial use or donate in the public interest while preserving their privacy.
- Enables creators to monetize strategies, plans, and data they provide to the community, and earn a stake in consumer-monetized data produced as a result of following their strategies and plans.
- Enables consumers to participate in pseudonymous or anonymous health and wellness research and development.
- Empowers private, powerful AI/ML for consumer health and wellness optimization.

CHN is powered by a community in which consumers contribute to the breadth, depth, and diversity of accessible data by sharing results of personal efforts to proactively optimize health and performance through n-of-1 experimentation.

Community members contribute to the development of the core network protocol directly and through decentralized governance of the network, wherein they can submit proposals for improvements to the network and governance structure.

The Collab.Health Team is bootstrapping this endeavor by launching the Network and Community, and by building the first Apps and dApps in the Collab Network. The Community will govern, build, maintain, and extend the Network after launch.



2.0 MOTIVATION

Personalized health cannot reach its full potential under an economic incentive structure that rewards data hoarding. The resulting fragmentation of information makes it too hard to aggregate and quantify the systems biology or whole-person perspective on individual health needed to develop personalized recommendations for every aspect of health and wellness.

Shallow and Narrow

The efficacy of *any* empirical science is directly proportional to the ***depth and breadth*** of data it has available for it to explore. By depth, we mean the comprehensiveness of the relevant data concerning an individual entity, and by breadth, we mean the size and diversity of the total set of entities. While an individual is unusual if the depth of a health record contains their entire prescription history, the depth of an individual's aggregate record of internet behavior for advertising and/or marketing purposes is approaching the entire history of every digital interaction they've had. Clinical studies are designed to limit sample size to that required to investigate a single interventional variable. Addressing systemic health outcomes requires much larger volumes of data (ex: Metformin, regarded as one of the safest drugs in the world, may cause birth defects when consumed by fathers prior to conception). In fact, few clinical studies gather data from more than 1 percent of the overall group concerned, while research conducted by Facebook and Google can easily access sample sizes larger than any country on earth.

Data Creation and Integration is Costly

The type of data available for marketing science (website visits, social graphs, credit card purchases, etc) is primarily produced as a side effect of consumer behavior, and typically does not require the consumer to intentionally enable the researcher. Data required for personalized health research is, by and large, not passively harvestable. Creating personal health data requires specialized apps, hardware, clinical services, wet labs, and purposeful behavior to capture and contextualize parameters of interest. Consumers pay out of pocket to produce most personal health data, and most data creation tools are narrow and purpose-specific. This combination of factors imposes high cost on the consumer:

1) Cost of **hardware** such as wearable sensors and mobile devices.

2) Cost of **software** - health apps.

3) Cost of **effort** - users must develop **new habits and routines** to receive benefits from tools.

4) Cost of **lost opportunity from inaccessible and incompatible data**. Data generated about consumers by most commercial tools is captured by the entity that provides the tools.

Such entities are incentivized to keep the data and insights that they generate private to maintain a competitive advantage. Their valuation is proportional to the volume and predictive utility of data that is captured and maintained in their **siloed walled gardens**. This not only results in a *scarcity* of integrated whole-person datasets, but that data that is accessible is gatekept by for-profit entities in widely variant formats.

As a result, each new personalized health and wellness product is required to engage in a repetitive process of acquiring ***the same*** data from users to personalize their interactions and recommendations. This leads to **redundant data labor** that the user is responsible for performing across services. This enhances profit and reduces the risk for the commercial operator but is at cross-purposes with, and limits the potential of knowledge gained by, the individual's need to aggregate, integrate, analyze and act on their own data for the sake of whole-person health. In addition, this adds friction to the user, both in terms of usage (they often collect duplicate data), but also monetary cost in that they have to buy many wearables.

Personalized Health Research Needs Data Liquidity

In order for an individual to assess the effect of an intervention on themselves, they must follow a protocol for measuring that effect. In order to analyze the effects of many individual experiments, the data produced by each observation must be readily comparable with others. There is a need for standards that provide a basis for analyzing the results of individual experiments at the population scale.

While there exist many examples of n-of-1 experimentation in narrow contexts, that data cannot be analyzed at scale. Ultimately, this fragmentation limits the potential data pool for investigating personalized health.

3.0 HOW IT WORKS

Collab.Health aims to enable health science to reach its fullest potential by creating an ecosystem powered by sufficient incentives for the collaborative creation, distribution, correlation, and monetization of health data - both for individuals and for the full range of actors and stakeholders operating in the health and wellness sector.

It is our thesis that if the cost associated with health data creation can be overcome, the amount and availability of health data will be exponentially increased, and result in a corresponding increase in the efficacy of health science as a whole.

The Collab.Health network consists of

- **Health Programs:** create, share, adopt, monetize, and track the results of open-source health programs. Returns value to the program creator when users follow their program and go on to monetize their data.

- **Data Vault:** access to secure, decentralized, sovereign, data storage for health data and tracking of program adherence and insight generation.

- **Data Verification:** an oracle capability to (a) automate the upload to users' vaults and assert the validity of health and wellness data and (b) provide credentialed verification of data and or actions from qualified providers. Data attestation returns value when users go on to monetize their data.

- **Data Exchange:** a data dex for the exchange of health data for network tokens. Value is returned to all participants involved in the data generation (users, program creators, oracles).

- **Algorithm Marketplace:** users can purchase algorithms for use within the trusted compute to generate insights from their health data

- **Trusted Compute:** access to secure, decentralized, trusted compute for algorithms to run against data stored within the users Vault.

- **Community Governance:** a DAO powered community to provide governance of the oracles, primitives, schemas, and network improvements.

- **Applications:** To power key management, data collection and verification, data exchange, and health program tracking.

3.1 CONCEPT OF OPERATIONS.

All users download and utilize the Collab Wallet for network transactions.

Users can then manage their decentralized storage and compute needs via the Collab wallet.

All users can become program creators by publishing open-source health programs to the network via the Collab dApp.

All users on the network can adopt these programs via the Collab dApp, or alternately directly in the Collab Wallet.

Users use the Collab wallet to link and upload data to their health vault directly or via network oracles.

Users following programs utilize the Collab Wallet to track actions specified by the programs.

Free and paid algorithms that are run inside of decentralized, trusted compute run against the data held in the user's vault to provide them insights on the current state of their health and wellness, and potentially the efficacy of any programs they follow.

DEX Behavior (a) Users can offer their data for sale in the dex utilizing zero-knowledge qualifiers, creating a health data pool attesting to the quality of the data, without showing the data itself. Data buyers then pay for access to this data. (b) data purchasers can put up bids for the type of data that they are willing to pay for and users for, and users respond to the advertised offer, completing ZK qualifications to ensure their data meets the offer prior to handing over access to the data itself.

3.2 INCENTIVES

Collab.Health proposes the following incentives to maximize Public Health Record (PHR) quality into two distinct categories:

1. **Health Improvements** through the creation and publishing of strategies, programs, and algorithms that will all intrinsically improve in efficacy for a user given a higher-quality PHR.

2. **Financial Gain** by having an efficient market for: (1) selling and buying personal health data and/or insights created from the same, from which the first point of sale will always be the individual who originally generated that data.

HEALTH

The more a user gains awareness of their own behavior through observation and analysis of observation, the greater their propensity for positive health outcomes.

The more a user generates high-quality data, the more useful the feedback they receive.

The more a user acts on valuable feedback, the more impactful and sustainable their health optimization results become.

The more impactful and sustainable the user's results, the more valuable the user's data becomes.

FINANCIAL

Collab.Health Network incentivizes creators to produce Strategy and Protocol content to receive tokens. Users are incentivized to produce data guided by Strategy and Protocol to receive personalized interpretation and recommendations.

While we foresee there being a great many non-financial incentives for individual participation - our philosophy of incentive design considers financial incentives fundamental to growth.

In keeping with Collab's foundational goal of providing sufficient incentives for individuals to actively engage in Health Data creation, Collab will support multiple different and complementary financial incentive mechanisms for individual participants in the shared data economy.

Direct Purchase

The simplest but arguably most ubiquitous form of financial incentivization within the network.

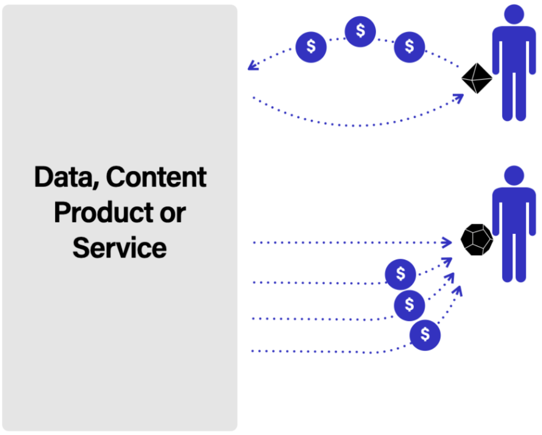
A buyer offers an amount to pay for some data or abstracted inference from said data using our "Data Dex" primitives and the provider agrees. We see this as being the primary financial incentive for individual actors within the network to continually provide valuable data, and we also see powerful incentives for communal pooling of data or "Data DAOs" whereby each contributor of a certain type of data to a communal pool earns a revenue share over said data proportional to their contribution. We foresee "Data DAOs" of this kind becoming a standard arrangement between participants in HPs as an additional value add both for the purchaser of data as well as to create an additional "moat" to protect Private HPs from bootlegging effects.

This same pattern will be utilized for algorithm or protocol purposes.

Downstream Participation - Micro shares

A powerful mechanic heavily inspired by the Hyperledger Bootleg prototype, Downstream Participation Micro shares incentivize both participation in, and early adoption of collaborative enterprises on the network that an individual perceives to have value. This mechanic works by allowing individuals to obtain by default with their mere participation a Microshare of all future revenue generated by a particular collaborative enterprise within the network. This mechanic also performs double duty as a disincentive for bootlegging of private “pay-walled” data sets or enterprises such as a Private Protocol whose creator wished to charge an “entrance fee” in order to gain access to the information and community surrounding it. We expect this to provide a network effect of incentivized community members to serve as a strong “moat” to such Private Enterprises, so that even when someone inevitably copies and/or “bootlegs” the rote instruction set of something like a protocol, the original protocol will retain the advantage of all the honest members who are incentivized to remain loyal and only add value by their help and community participation to the original due to their Microshare in its success.

This mechanic also serves as an “early price signal” system similar to bonding rate curves for curators employed by The Graph Protocol, as Micro shares will naturally tend to be worth more earlier on before the majority of downstream users pay their “cover charge” from which the user will obtain a “micro dividend” from.



User Pays for Item

User also receives a "micro-share" in future revenue for that item, paying out every time someone else pays for that same item.

Direct Token Minting - Subscription System

One of the most powerful incentive mechanisms available to a tokenized Web3 platform comes from the ability to directly mint tokens out of thin air directed towards a type of behavior on the network that is desirable. This can be thought of as an “inflation tax” placed upon the rest of the network that subsidizes a certain type of actor behavior that is deemed to be worth the cost imposed by said inflation on the rest of the network. While exploitative inflation taxes remain one of the main things driving new users into crypto, applied correctly they can be incredibly powerful so long as the inflation rate is codified within the protocol's monetary policy and openly used as an Overt, rather than a Covert tax.

The mechanism that Direct Token Minting will be applied to incentivizing various actors within the network will be heavily inspired by the system used by Dev Protocol, whereby an individual's staked tokens yield two distinct yield streams:

1. The native staking yield stream, which is by default directed to be paid out back to the Individual who is staking the tokens, in the usual manner you would find with the network validation staking payouts on any proof of stake network.

2. A secondary "subscription" yield stream that may NOT be directed back to the individual who staked the coins but must be directed towards a DAO-defined set of actors within the network that are performing the behavior that is desired to be incentivized.

The fact that the token staker is unable to utilize their "subscription" stream of tokens generated from staking for anything other than compensation for the health services of DAO-approved actors within the ecosystem. This provides a natural bootstrapping mechanic in the form of an immediately available regular income stream that may be tapped into by those who add value within the Collab ecosystem, while still maintaining buyer market intelligence to choose which actors will receive this income and in what amounts.

3.3 TECHNOLOGY

Financial incentives are often the first thing that comes to mind whenever one thinks of incentives in general and particularly within the blockchain & web3 space. However, **when it comes to creating economic incentives for the creation of any information that is meant to be freely traded on an efficient public market, the method of achieving that is anything but trivial** for the following reasons:

1. For a market to be considered efficient, its prices accurately reflect the value of the items within the market.

2. The foundation of any price system for a thing is the "First Point of Sale," which refers to the event when the ownership of an item is traded from the first actor that can claim valid property rights over it to an external buyer for the first time within the market.

3. If we want to maximize the incentives for individuals to create valuable data, **the individual who makes the data must be the First Point of Sale price-setter for that data, which infers that the individual must *truly* own that data.**

4. Ownership of information is equivalent to mere knowledge of said information and the freedom to transfer (communicate) it to anyone. In other words, **Information Ownership = Privacy + Censorship Resistance.**

It is this foundational need for Privacy and Censorship Resistance in order to allow individuals to truly own their data that is why Collab Health is being built as a Web3 platform. Especially when one considers the stringent regulatory requirements regarding privacy, security, and user control for Health-related data, there is no other paradigm that we believe is capable of supporting these ideals to the fullest extent.

When it comes to ensuring privacy for an individual's Health Data we believe the only valid solution is self-sovereign custody via cryptographic methods. In other words, we believe the value and volatility of individuals' Health Data warrant it having the same standard of custody as we would associate with something like a cryptocurrency private key. This being the case, the core User-Facing application Collab Health will offer is the Collab Health Wallet. The wallet is functionally similar to existing Web3 wallets like Metamask, as its primary role is a secure way for users to have custody of their private keys for encrypting, unencrypting, and signing (approving) data within the network.

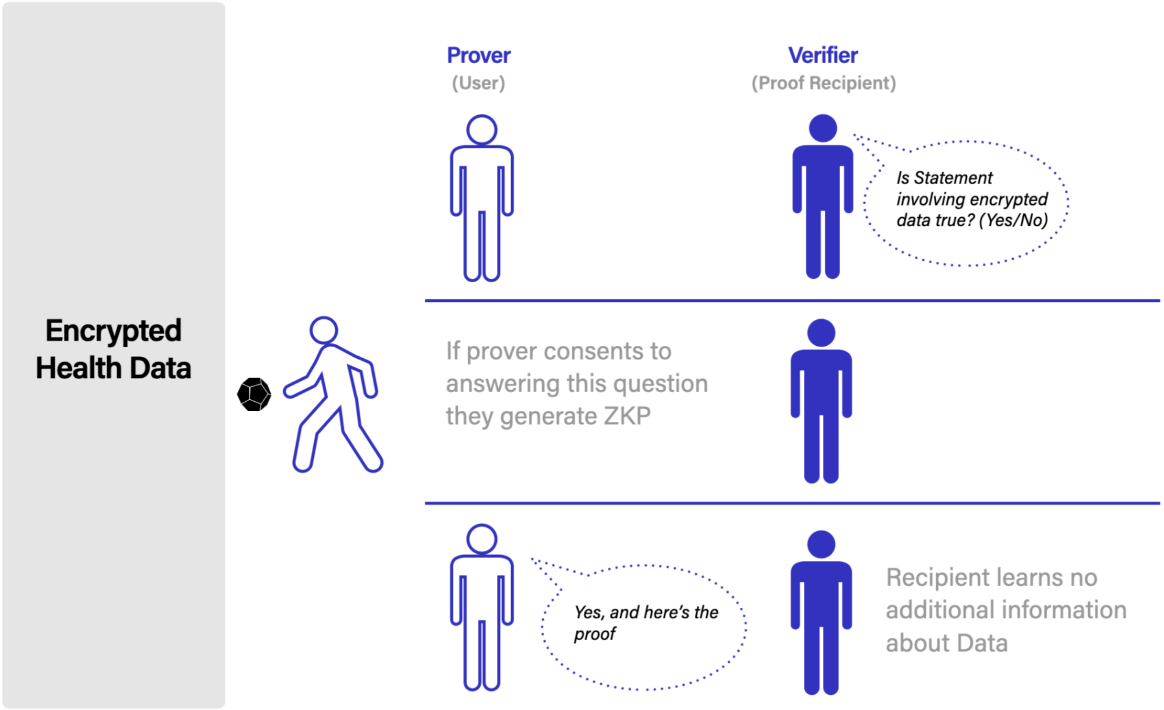
All data that a user creates via any means will be required to be entered via the wallet and will always be encrypted by default. This default action of forcing privacy whether or not the user wants or even cares to have it is to prevent accidental “opsec leaks”, and to assist users who may not fully realize how volatile that data might be in the wrong hands, or how valuable it would be to an interested party.

The majority of what goes into ensuring to the highest possible degree of certainty that user’s data will remain private unless they explicitly choose to divulge it comes down to the selection and arrangement of cryptographic schemes used to encrypt, unencrypt, sign, and create proofs. However, the design constraints for encryption standards don’t simply end with maximization of privacy, **all the encrypted Health Data in the world isn’t going to be interesting to potential buyers unless they are actually able to do useful things with it while still maintaining the privacy of the individual who provided it.** While this may seem paradoxical, there are a large number of cryptographic primitives currently being explored that offer the ability to do just that.

PRIVACY

With these constraints in mind, we will now outline the Cryptographic Primitives we see as being prerequisites to building a safe, useful, and compliant market for Health Data:

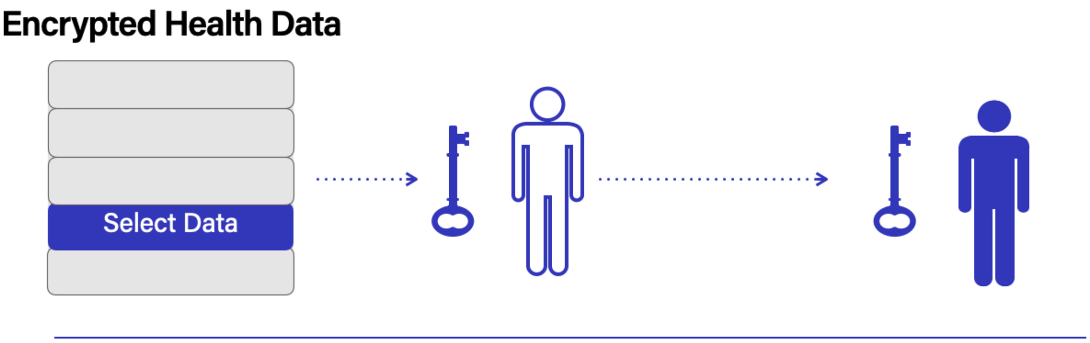
ZERO KNOWLEDGE PROOFS: The encryption standards employed by Collab Health will allow users to create privacy-preserving proofs of arbitrary high-level statements about their data, and share those with parties of their choosing. To give an example, this would allow a user to share proof to a sleep physician that they have averaged only 4 hours of sleep a night for the past month by creating a Zero-Knowledge Proof over the data created by their Smartwatch Sleep monitoring, without revealing a single bit of the "plaintext data" itself. This is the Gold Standard for privacy-preserving data sharing techniques and will be our recommended default for nearly all use cases.


Encrypted Health Data
Prover
(User)
Verifier
(Proof Recipient)
Is Statement involving encrypted data true? (Yes/No)
If prover consents to answering this question they generate ZKP
Yes, and here's the proof
Recipient learns no additional information about Data

ATTRIBUTE & IDENTITY-BASED ENCRYPTION: Collab's encryption standards will also have the ability to share chosen segments of "plaintext data". The interesting thing about this relatively new encryption technique is that it allows users to create a key that can only encrypt a segment of data of their choosing within a totally public but encrypted set of data, making it completely impossible to "leak" any more data than the precisely chosen data segment. This level of "opsec safety" is crucial because it is very easy even for relatively small segments of "plaintext" Health Data can be completely de-anonymizing. These same techniques also give the ability for users to allow certain Entities or even Entities with specific Attributes to access segments of their data. This feature is a must for certain applications such as sharing with a set of physicians of certain expertise or to an un-enumerated group of individuals meeting a certain standard of trust or authority. The activities required to run Attribute & Identity Based encryption in a decentralized manner will be a default responsibility of Network Staking Validators.

Fine-Grained Access Control

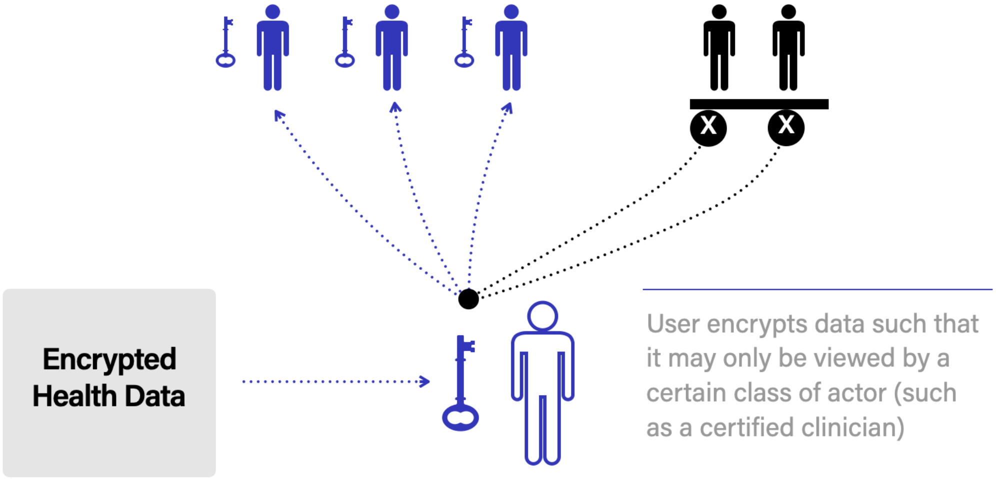
User creates a key that is only able to decrypt a particular segment of pre-encrypted data, and then sends that key to a third party, who is unable to view any data outside of that chosen segment.



Encrypted Health Data

Encrypted Key

User creates a key that is only able to expose a particular segment of data



User encrypts data such that it may only be viewed by a certain class of actor (such as a certified clinician)

REVOCABLE & QUANTUM RESISTANT ENCRYPTION: The encryption standards for Collab Health must be revocable not only for the safety factor of giving individuals an “undo” button for granting access but it is explicitly demanded as a matter of regulation surrounding giving individuals complete control over their Health Data. Like the previous primitives, Network Staking Validators will be responsible for being the decentralized administrators of the distributed proxy re-encryption required to achieve this feature. Finally, the encryption methods must be used by Quantum Resistant. **Due to the value and volatility of Health Data, it is unacceptable to use encryption techniques that have a**

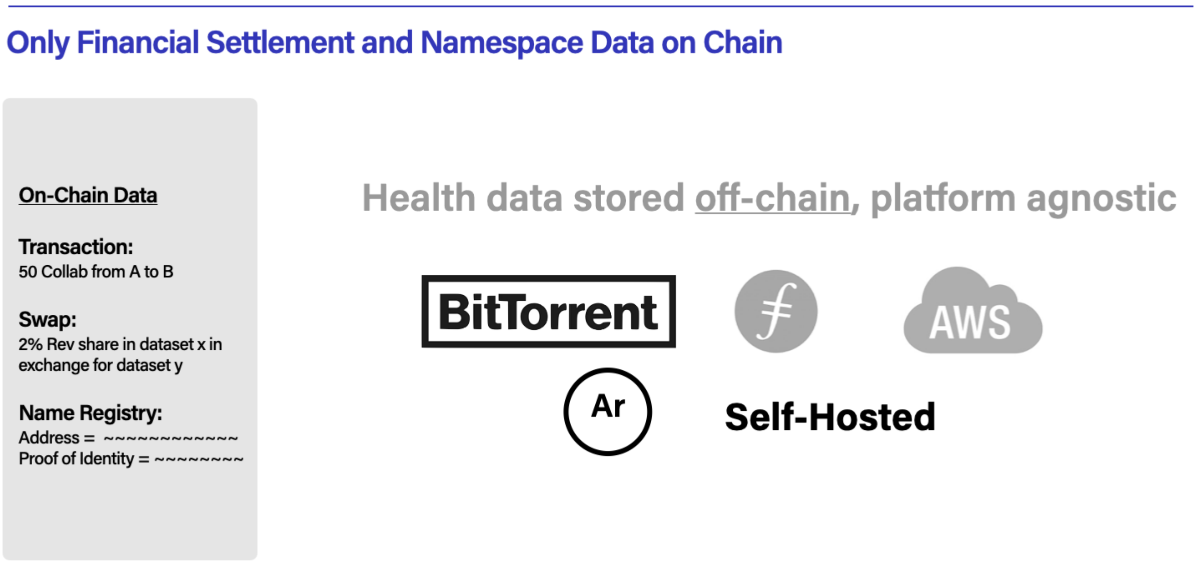
“quantum expiration date,” especially considering we expect Collab users will inevitably choose to store a significant amount of (encrypted) data in public decentralized storage mediums. It’s worth pointing out that we get this feature essentially for free as the technique likely to be used for Decentralized Attribute and Identity-based Encryption just so happens to be Quantum Resistant (NTRU Lattice Scheme).

Only when all of these cryptographic prerequisites are met will we consider the privacy of a user’s Health Data to be of a sufficient degree of security. These primitives cover the “Privacy” element of the “Privacy + Censorship Resistance” requirement for information ownership; let us now move on to what Web3 primitives we have chosen to ensure Censorship Resistance.

CENSORSHIP RESISTANCE

BLOCKCHAIN-BASED SETTLEMENT: when it comes to absolutely guaranteeing that a transfer request of a digital asset executes as intended, no matter what level of bad actors may attempt to censor it, nothing comes close to a blockchain. It is, after all, what Web3 is all about to one degree or another. The Collab Health Blockchain will be its own Layer 1 built on the IBC standard using the EVM execution model, giving it simultaneous interoperability and scalability advantages from day one while maintaining the highest possible developer pool for creation and validation of all code from a tried and true execution paradigm. The Cosmos team's work towards making certain Distributed Key activities a default responsibility of Network Staking Validators also lends itself to our similar goals for our cryptographic primitives that require similar schemes.

The Collab Blockchain will stick to very conservative Layer 1 philosophies. The amount of data written to the Layer 1 should be limited to simple transfers of tokens or NFT-style assets, and the registry of Namespace Data. The Health Data itself will not be stored on-chain, nor should every state change be involved in advanced computation and algorithms run on top of Health Data.



Only Financial Settlement and Namespace Data on Chain

On-Chain Data

Transaction:
50 Collab from A to B

Swap:
2% Rev share in dataset x in exchange for dataset y

Name Registry:
Address = ~~~~~~~~~~~~
Proof of Identity = ~~~~~~~~

Health data stored off-chain, platform agnostic

DECENTRALIZED PERMASTORAGE FOR DATA: Part of building the “ironclad information ownership” we are striving for requires the ability for users to store their data in a medium where it is infeasible for theft, natural disaster, hardware failure, or de-platforming to cause them to lose their precious Health Data. To this end, we will allow users the option to store their data on the Arweave Permaweb, the only decentralized storage platform to have a pay-once-store-forever economic model and by far the highest level of storage redundancy for files. This means for the user that so long as they pay a set fee to store the data, they can rest assured that nothing short of a global calamity could have a hope of causing them to lose their data. We also recognize that Arweave has certain shortcomings regarding transaction cost and settlement speed of stored files, which is unacceptable for many Health applications requiring a heavy volume of files or instantaneous settlement and file availability. Because of this, we will be using an innovative layer two network that will both batch file transactions to the Arweave network and allow for files to be hosted and made available immediately in a quick manner until they are confirmed on the Arweave network.

The use of the Arweave network will be essentially invisible to users. They will pay for the Arweave storage fees using the Native Collab Token. The files hosting location will be accessible via a standard URL web link as if they were on any other hosting provider. We also recognize that some users may not be comfortable storing their Health Data off-edge, even in a securely encrypted form, so storing data on the Permaweb will always remain optional. In general, the Collab Network will be designed such that a user’s data can be stored and archived wherever and however they wish, and the rest of the system will remain agnostic.

As previously mentioned, due to the liklihood of this uncensorable form of storage being popular for some users, the ability to have an element of revocability of data being stored on such mediums is crucial. This was the main motivation behind the utilization of a Revocable Encryption method.

ATOMIC DATA SWAPS: The ability for an individual to transfer some data to another party may be Censorship Resistant, but if the ability for them to receive compensation for that data transfer is not, then it wouldn't do them much good. Atomic Data Swaps piggyback on the idea of Atomic Swaps that have been used successfully for years to exchange two digital assets even across potentially incompatible blockchains. The word "Atomic" refers to the fact that either the trade executes as intended, or nothing happens at all, there is no way for one of the participants to "stiff" the other, and it doesn't require a centralized escrow for users to trust it.

Atomic Data Swaps will consist of a transaction standard for creating Buy and Sell offers of the form **"Here is the thing that I wish to trade with, and these are the conditions that must be met for me to execute this trade."** Once this Buy or Sell offer has been made and signed, it is sent to a traditional Order Book-style exchange where market participants can see it, choose to fulfill the opposing end of the trade, or even create a counter-offer.

The lowest-level framework for creating these trade execution conditions will allow for **arbitrary conditional logic as one would expect within any smart contract paradigm.** This is a must to allow actors to create "data deals" and interactions beyond what we as creators of the network would foresee as being a use-case and act as an accelerator for innovation just as it has for the rest of Web3 as a whole. However, there are always benefits to the leaders of decentralized platforms to provide useful defaults and "pre-built" scaffolding to help bootstrap the ecosystem and act as a benchmark for quality and scope. Some of these "default primitives" we intend to provide are:

- The ability to grant "revenue shares" on some data set such that the owner of the shares is given rights to some percentage of future trades or monetization made with that Data.
- The ability to grant "revenue shares" on higher-level entities such as Individuals, "Data DAOs," Oracles, Intervention Providers, etc., such that the owner of the shares is given rights to some percentage of future trades or monetization those entities gain.
- The ability to atomically trade Data and Revenue shares for Tokens
- The ability to atomically trade Data and Data shares for access to other Data or Revenue Shares
- Buy and Sell offers for Data or Revenue shares meeting specific requirements such as:

1. A zero-knowledge proof of some explicit result (i.e., a particular physical phenomenon measured after a particular type of intervention)
2. A selection of plaintext data that meets specific requirements (i.e., heart rate data taken a certain amount of time after strenuous exercise on a particular diet)
3. Comes from a data provider meeting a certain standard of trustliness
4. It was generated by a specific type of hardware data oracle
5. Was indexed by a particular processor of data or algorithm etc.
6. Some arbitrary smart contract logic was executed correctly

VALIDATION

Oracles

A critical part of any marketplace for data is the ability for data buyers to be given strong guarantees that the data they are purchasing is both accurate and authentic. One can't simply provide blind financial incentives for the creation of arbitrary data, otherwise the incentive is for bad actors to manufacture huge amounts of fake datasets that have no correspondence to reality simply to game the system for profit. Collab will combat this by employing the use of Oracles within the network, as well by ensuring that there are no such blind incentives.

An Oracle in Web3 is used to refer to any entity that makes an on-chain attestation that some Real World Data is true. Unlike most blockchain paradigms Oracles are NOT entirely trustless or decentralized, the trust model surrounding whether or not an Oracles is making true attestations is based on their reputation of past performance, and the presence of competing Oracles who will are continuously incentivized to elevate any incidents of fraud amongst their competitors. Sometimes to add to their trust model, Oracles may lock up some number of tokens as a form of collateral that may be "slashed" if they misbehave.

In Collab's case, an Oracle may be able to actually append some data to a user's health record (pending their approval), or they may merely "sign off" on some existing data in a user's health record, lending their credibility to signal the veracity and authenticity of the data. Examples of Oracles in the Collab Ecosystem could be, among others:

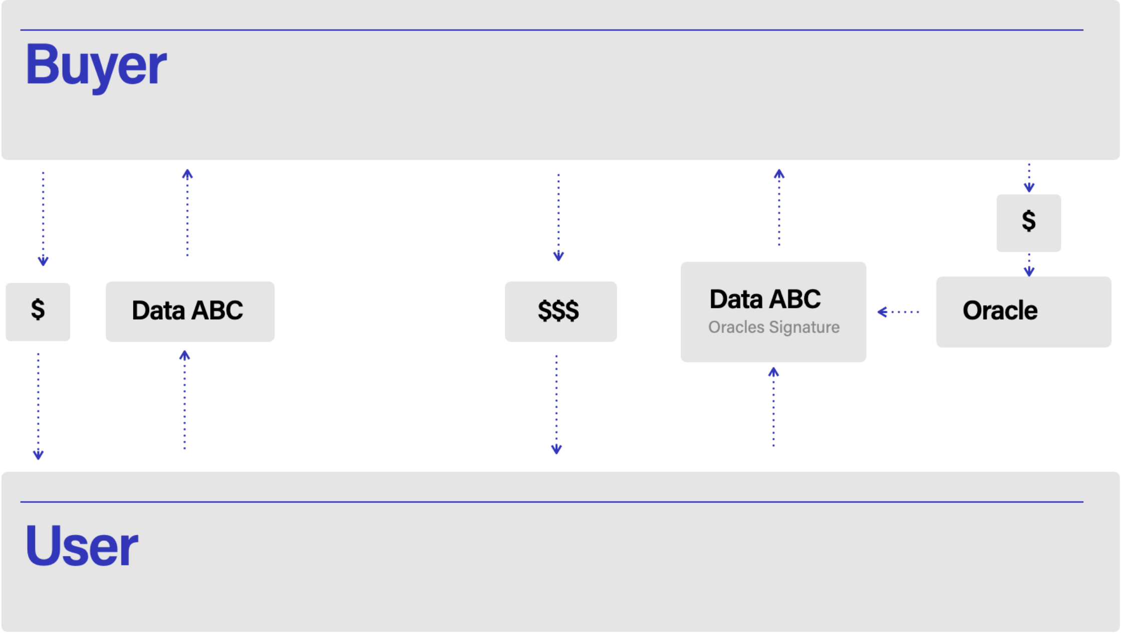
- **Wearable or Sensor Manufacturer**, verifying a user's claim that some data they provided did in fact originate from their device
- **A Personal Trainer**, attesting to the fact that a user did in fact complete a particular workout routine within a particular time

- **A Licensed Physician**, attesting to the fact that a user did in fact undergo a particular procedure

- **A Private Medical Test Company**, attesting that to the fact that a particular test a user engaged in returned a particular result

While the incentives for why an Oracle would want to engage in this activity on the network may vary, with the ability to form arbitrary agreements for Revenue Shares on a particular set of data, we expect a healthy competitive market for providing credibility for data to emerge. We are particularly interested to see how legacy Health Care Workers engage in this system. By law, they would be required to *truthfully* provide any and all data they may have created for a particular individual, if they also have the ability to make some percentage of all revenue created by that data merely by joining the network of Oracles we expect that the incentives will be more than great enough for them to do so. We also expect that early users of the platform will be actively incentivized to try to seek out and “convert” Health Care Workers to join in order to raise the value of their produced data by having it verified.

The other half of minimizing incentives for fraud is to ensure that **the only sources of financial incentives for providing data come from Market Buyers and not by blind network inflationary means.** If data buyers are the sole price-setters for purchased data, then the potential for fraudulent data will become naturally “priced in” to the market, with data being validated by trustworthy oracles demanding the highest price and unvalidated data demanding the very least. This will also serve as continual upwards pressure toward better, more verifiable data, and a more competitive and therefore trustworthy network of Oracles.



CONTENT FILTERING

On any decentralized Web3 system there always emerges the problem of Content Moderation and the technical, political, and philosophical ramifications associated with on one hand having to put up with inarguably harmful content, and on the other hand allowing for censorship and losing decentralization in the process. A system like Collab only accelerates this quandary given the context of individuals in theory being able to recommend ANY kind of unfounded “interventions” they wish. We are helped somewhat in this problem by the fact that the storage and hosting of data on the network is entirely platform agnostic, and no actual data aside from transactional and namespace data will be stored on the Collab Layer 1. But while we believe that the natural incentives the ecosystem provides will always be for individuals to NOT participate in any kind of harmful behavior, we believe it is critical given the use case of Collab to not simply default to pure Anarchy when it comes to Content Moderation. We also believe that a Web3 platform can not be said to be truly decentralized if there are network-level abilities for a central authority to censor anything. To balance this dichotomy, Collab will utilize the primitive of Content Filters, which pushes the role of (benevolent) censorship up to the level of the Application, while keeping the Layer 1 protocol entirely censorship-resistant.

Content Filters are related to the role of Oracles, and can be thought of as self-imposed censorship that individuals who use the network can subscribe to of their own volition. This concept has already started to be explored by Web3 platforms such as Arweave with their Portals, and we believe it to be the winning model in this quandary. A Content Filter is an application-level blacklist that an application user subscribes to that will block all content on the network that has been flagged by the administrator of the Content Filter. If a user for whatever reason does not wish to subscribe to a particular filter, they may choose to switch to a different one, or even choose to use none at all. These filters may include among other things:

- Harmful and illegal content
- Protocols that are obviously nonsense or unsafe
- Patent or Copyright infringing content
- Doxxing

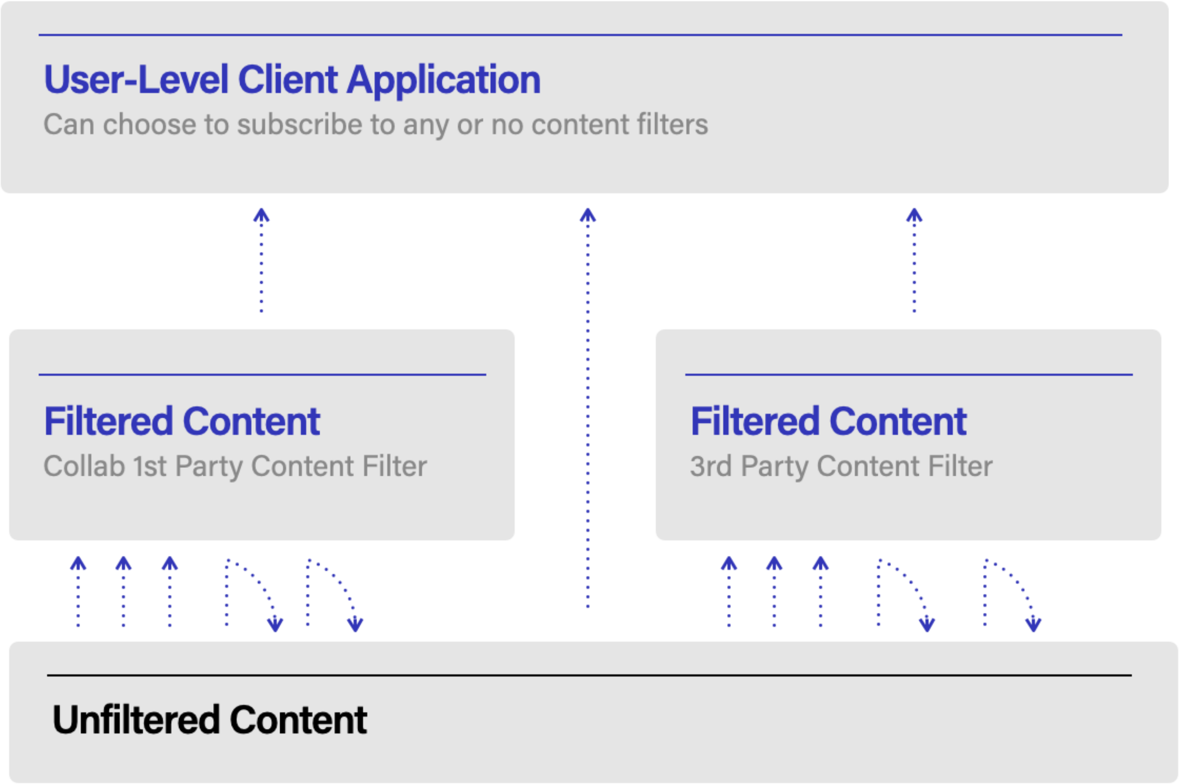
The Collab team will maintain our own Content Filter that will be shipped as default in our first-party clients. It is important to note that we as the developer of the default Collab Client for iOS and Android platforms will likely be required to lock our own Content Filter by default (user will be unable to unsubscribe) in order to comply with App Store policies surrounding Content Moderation. As stated earlier, however, Layer 1 will always remain 100% censorship-resistant and users may always choose to use a non-mobile client where our default Content Filter will not be enforced dictatorially.

NAMESPACES

NAMESPACE DEFINITION: A set of 1-to-1 mappings between permutations of a human-readable alphabet (names) and a set of data objects such that each unique name appears only once within the set, and is usually under exclusive control by a particular individual or entity.

EXAMPLES: Internet Domain names mapping to IP addresses, Twitter handles mapping to users, Passport numbers mapping to Citizens.

👉 Collab will have two distinct Namespaces at the protocol level, one for Applications, and one for Actors.


User-Level Client Application
Can choose to subscribe to any or no content filters
Filtered Content
Collab 1st Party Content Filter
Filtered Content
3rd Party Content Filter
Unfiltered Content

- **The Application Namespace**
  Used to identify applications that are developed using the Technical & Economic Primitives. While applications on Collab will have a hashed "contract address" similar to what is seen on Ethereum and other platforms, the sensitive nature of Health Data demands that users be given the strongest possible proof that the application they allow their data to be processed by is going to do exactly what they want it to. By having a Namespace that maps human-readable names to Collab Applications, this not only aids greatly in discoverability and trust, but it also makes it virtually impossible for potential bad actors to "impersonate" a known application with one that is somehow malevolent. This problem is solved (somewhat) on Ethereum by piggybacking on the existing DNS standard, whereby users first line of trust that an Ethereal dApp is the one they want is the domain name for the URL that it is hosted on, with Collab we want to bring this trust layer down to the level of the Layer 1 so that there are even fewer opportunities for bad actors to spoof dApps and also to allow for this trust layer to operate in environments other than a web browser, such as the native Collab Mobile App.
  It is acceptable, and even beneficial, for some elements of the Application

Namespace to be centralized, acting in a similar way to centralized App stores, whereby there is a level of central curation and quality control that goes into allowing applications to be entered into the Namespace. This level of centralization is acceptable because, in the event that an application is denied it doesn't mean that it is prevented from existing on the low-level network, it merely is denied access to what will ostensibly be the most popular indexing, discovery, and quality control system for the network.

- **The Actor Namespace**
  Used to identify Individuals and Entities within the network, essentially identical to how social "handles" work on existing Web2 platforms. These Identifiers will be for individuals, creators, companies & brands, research groups, DAOs, etc. This second namespace must be decentralized, as ownership of these "handles" will be required to interact with fundamental elements of the network, and therefore a centralized denial of the same would mean that individuals could be prevented from using features of the network by a centralized entity. There will, however, be benefits to a centralized **verification** service to enhance trust and provability particularly as it pertains to Influencers and Brands.
  The fact that this Namespace must be decentralized gives some additional requirements, namely:

  - Some small economic value must be placed on gaining access to a handle. This is a must to prevent spamming and name-hoarding/squatting. Functionally this means that a small number of Collab tokens will be required to "unlock" a name.

  - The Namespace will be required to launch with a significant number of names "reserved" in the centralized ownership of the Collab Foundation:

    - Names of Influencers, Brands, & other established Entities. These will be given to the entities by the Collab Foundation pending a verification process. This is essential to prevent Name-squatting on high-value names by individuals hoping to sell the word to their rightful owner at a high premium.

    - Many "default" outgunned human-readable names (sillymousepotato, funkyhorseromp, etc.). These will aid with onboarding so that new users can jump right into the ecosystem without purchasing Collab tokens to get a handle.

    - A list of names deemed inappropriate or illegal to be burned (made inaccessible) by the foundation.

- **These are the Web3 Primitives that give the Privacy and Censorship Resistance that Individuals need to truly own their Health Data in the strongest possible**

terms, thus giving them the most robust viable price negotiation platform within the market for their data, and therefore that they will receive the highest possible compensation for their data within the market. We believe that giving Individuals the most robust potential financial incentives within a market for Health Data will result in the early adopters producing the highest quality personal Health Data, which will, in turn, attract more numerous and diverse Buyers of data, which will in turn increase incentives for data provision, which will then lead to more adoption, etc

HEALTH 3 PRIMITIVES

Personal Health Record

PHR registry, custody, sharing, and staking functionality.

Ontology

Standards for generating and contextualizing personal health data.

Experiment

Interventions, protocols, procedures, data collection, and algorithms that guide their use.

Digital Twin

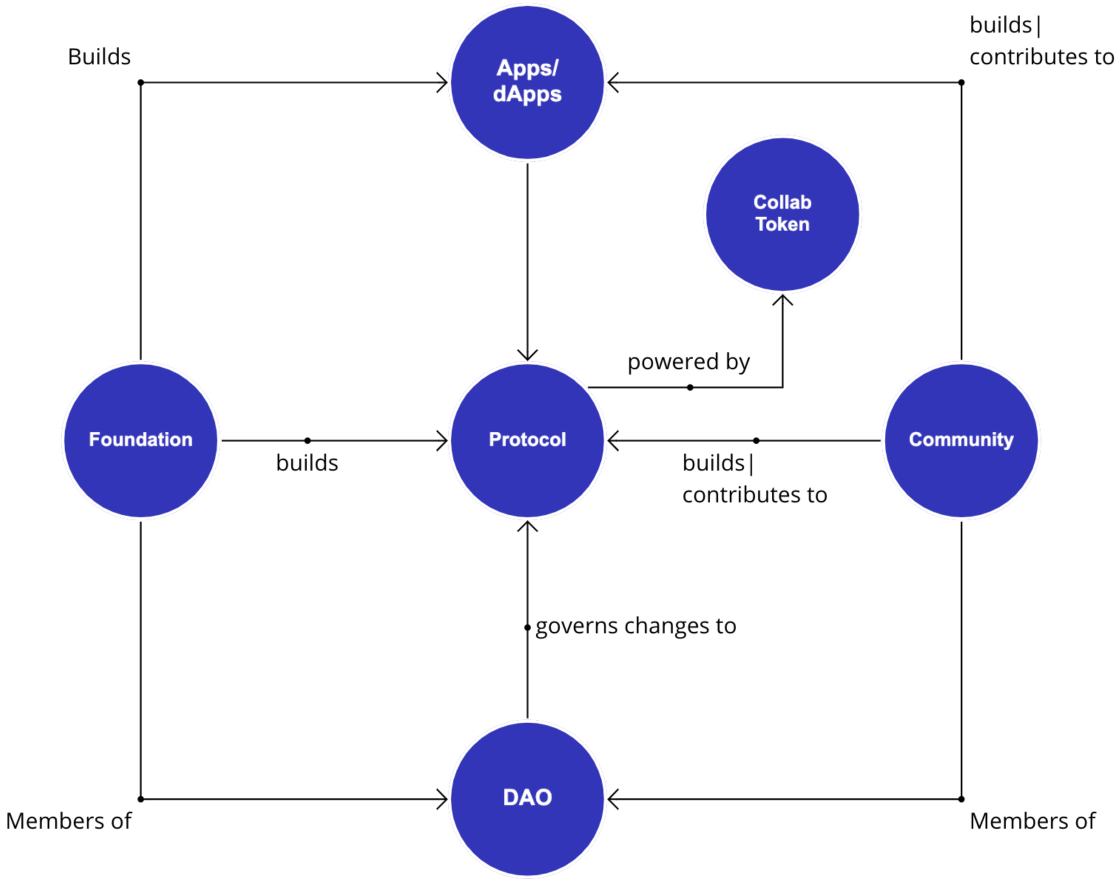
Converged, standardized, anonymized representation of user profile suitable as input for interpretation, prediction, and recommendation.

APPLICATIONS

Collab.Health will build the first, but not the only, applications in the web3 health and wellness ecosystem. These are aimed at bootstrapping the ecosystem by streamlining data ownership, privacy, & management; program creation & sharing; schedule integration & logging, and monetization functionality.

Collab **Wallet** - holds user's keys, signs/confirms all actions, management of user's data (import, export, delete, etc).

3.4 GOVERNANCE



The primary mechanism for governance regarding proposed changes and updates to the network will be simple network validator stake voting. Proof of Stake fundamentally is network validators staking tokens to “vote” on which of the potentially many proposed versions of the works ought to become canonical and “confirmed”. This mechanism is ideally suited to serve double-duty to also allow those members with the greatest stake in the success of the network to vote on proposed versions of the chain that have been given updated or altered features. In practice this would mean that a proposed update to the network codebase or monetary policy can only become canonical if it has the support of over 50% of the network stakeholders.

There are, however, certain applications for having decentralized voting for network stakeholders that don’t involve making a fundamental change to the protocol itself. These will be realized through a native Governance DAO that will allow any token holder to vote at a rate of **one vote per**

token even if their tokens are being staked. The issues that fall under governance of this DAO will mainly consist in four categories:

- **Voting on allocation of treasury funds to various improvement proposals or other value-adding enterprises.**This could include providing a grant to a team willing to develop some key application or piece of infrastructure, funding a marketing campaign, or even simply a donation to what the DAO considers a worthy cause.

- **Voting on amendments to a DAO-governed Content Filter.** While we feel that it is necessary to have a centralized Content Filter that is run by the Collab Team in order to be compliant with Apple App Store requirements. We also see the value in having a Content Filter that is completely democratic in its curation.

- **Voting on recipients of the Network Support Rewards.** Since the concept of Network Support Rewards fundamentally requires careful discrimination of potential recipients to prevent Sybil-gaming the system, this makes it an ideal candidate to fall under the purview of the DAO.

- **Voting on Public Scheme Governance.** Collab Network maintains namespaces for the following schemas. These schemas establish the metadata parameters that are used to interpret, share and buy or sell personal health data on the network.

  - *Scheduling*: Metadata for integrating actions by a calendar, event, workflow, goal completion, and other criteria into a calendar or similar applications.

  - *Measurables*: Measurable definitions standardize the exchange of all forms of biomarkers, self-reports, and other quantified properties related to health.

  - *Others*: Certain standardized definitions are employed as metadata on other schema objects. They include, but are not limited to, standard units of measure, goals, biomarkers, etc. These dictionaries will initially be centrally controlled but will transition to community/DAO management over time.

Additional responsibilities that the community wishes the DAO to own are possible, but these additions would be subject to network validator approval, and not within power of the DAO itself to impose.

4.0 LAUNCH STRATEGY

Following a period of collecting feedback on the litepaper and the concept, we envision the mainnet launch to include the release of the Collab Wallet, formation of the DAO, issuance of the

Collab token, and core protocol functions for inflation and governance. This may also include access to early prototypes of Collab developed applications.

This will be followed (tentatively, and subject to change) by several phases of development

- Phase 1: Vault and Protocols - Launch of the decentralized health data storage, functional encryption to power data sharing, health protocol creation and adoption.
- Phase 2: Oracles, Zero Knowledge, and DEX - Data validation via oracle attestation, privacy preserving zero knowledge capabilities for offer qualification, decentralized data exchange.
- Phase 3: Compute and Algorithms - Decentralized, trusted compute for algorithms to run inside plus an Algorithm DEX for discovery and purchase.

5.0 CONCLUSION

Collab.Health offers individuals, businesses, academia, government, and society an unprecedented opportunity to work together to produce better health outcomes by creating a collaborative ecosystem. The Collab health Network fuses together the best of personalized health and wellness, wearable & health technologies, and emerging Web3 offerings to deliver private, self-sovereign, consumer-centric health optimization.

6.0 APPENDIX

Individual A human person

Ecosystem An interconnected web of actors.

Health Goal A quantified objective related to lifespan, healthspan, or function/performance.

Health Strategy A systematic approach to fulfilling a Health Goal. Ex: Her strategy is first to eliminate sources of toxicity. Secondly to adopt and optimize interventions to improve metabolic function. In order to optimize whole-person wellness and function, she will integrate her metabolic strategy into her lifestyle and habituate it for sustainability.

Health Protocol Specific plan of action that implements a Health Strategy. Ex: She uses precision nutrition, hydration, exercise, sleep & recovery, gut health, digital interventions, and social activities to engage her body's own self-regeneration mechanisms and shift her homeostatic balance towards her Goal.

Health Optimization The process of achieving and sustaining peak physical, cognitive and emotional health.

DAO Short for “Decentralized Autonomous Organization”. An online democratic organization where proposals, voting, and rules enforcement can all be done in a decentralized manner using tokenized “voting shares” in the organization.

DEX Short for “Decentralized Exchange”. A decentralized platform or protocol that allows the trading of assets between non-trusting parties without requiring a centralized authority to administrate or custody of the assets in escrow.

Inflation An increase in the supply of a scarce resource, i.e. a token. Can be considered a tax or subsidy on existing holders of the asset to the benefit of those receiving the newly created supply of assets.

Encryption The process of converting some data into a form whereby the original data is unreadable to anyone who does not possess some private method of access, i.e. a Private Key.

Zero-Knowledge Proof A proof created about some Encrypted data that may convince a verifier that some high-level semantic statement about that data is true without revealing any additional information about said data.

Attribute-Based Encryption A form of encryption in which data may be encrypted such that only individuals with specific Attributes may decrypt the information, sometimes used interchangeably with Identity Based Encryption, which is a form of encryption in which data may be encrypted to certain Identities.

Revocable Encryption A form of encryption that allows one to encrypt information to a given Private Key, but then later revoke that Private Key’s ability to decrypt that information.

Post Quantum Encryption Forms of encryption that are not defeated by algorithms capable of running on realistic Quantum Computers, ie Shorr’s Algorithm.

Staking A method of “locking up” tokens in order to generate some passively produced yield based on the amount and duration of the tokens locked. Also used to describe the act of token holders voting on the canonical blockchain for Proof of Stake networks.

Oracle An entity within a blockchain ecosystem that attests to Real World Data on-chain in order to have Smart Contracts be able to execute based on said Real World Data. Oracles usually are required to stake assets on the veracity of the data that they provide and are at risk of having this stake “slashed” (lost) if they attest to false data.

Decentralized A file hosting protocol that allows individuals to trustlessly host files

Storage on a cloud-like network in which no centralized entity has control.

Web3 An umbrella term for technologies and platforms that are attempting to build a truly Decentralized and Self-Sovereign Internet.

Namespace A set of 1-to-1 mappings between permutations of a human-readable alphabet (names) and a set of data objects such that each unique name appears only once within the set, and is usually under exclusive control by a particular individual or entity.