DataLink DAO: Monetizing the Data Frontier

WIP v0.1

Justin Gaffney

{ support@datalinkdao.com; in/jgaffney311; datalinkdao.com }

Abstract

DataLink DAO will present the opportunity for countless new business models to be extended across a multitude of verticals in the web3 space. The DAO will host the DataLink Dashboard, a powerful easy-to-use tokenization platform for institutions and users. The DataLink validator will use multiple technologies on the forefront of cryptography research that will aid in the validation of DAO data, documents, and identities. The DataLink Foundation will play a critical role as custodian of the DAO acting as support in onboarding, protocol maintenance, product development, and other DAO support and operational needs. The DataLink DAO will monetize anonymous data-sharing, and provide a powerful tokenization platform for credentials, documents, and other data.

Table of Contents

Introduction

For the information age to move forward, there needs to a be decentralized truth engine that can verify data, documents, identity, by anonymous, secure processes. DataLink DAO is a significant leap forward into this type of automated, truthful verification. The actors within the DAO have the ability to not have to trust the other actors, but know what data they are providing and what identity they claim is true, without comprising their privacy and data. Some identities that could be easily forged that have big social and monetary impact are medical doctors, lawyers, judges, or any other position that involves high trust in the entities' expertise. There is a need to remove this trust and know for certainty that the identity and their credentials are true.

The harvesting of user data is extensive and prominent throughout the SasS business model. In the past two decades, there have been countless hacks and infiltrations that have lost countless personal user data, at no expense or very limited liability to the company. If Alice gave Bob her information. Bob could give it to Charlie without Alice ever knowing. This data has no traceability beyond the initial transfer to the service, company, or person that requested the user's personal data. There are many nefarious groups that steal, fake, identities and documents from these 3rd party transfers of data behind the initial user's back. The user is never the wiser that their data is comprised until it is publicly announced and too late. This situation could be avoided with an oracle-based zk-proof system that keeps the user's identity private from the requestor while only providing permissioned data that is approved by that user. Requestors and users will both need to be within the DAO to access the DataLink verification service.

Verified data is absolutely necessary in the digital age as more and more value is coming onto the internet, and blockchain is used increasingly as the value layer. There always comes a time to take a step and back and evaluate how the current processes are operating, and restructure certain parts of that process because of new impacts to the

industry that become apparent and unavoidable. One of these processes is how user data and identities is stored and protected on the internet. It is obvious from the rampant hacks and user data loss that there needs to be research and new solutions offered. Data can be stored in a decentralized manner and zk-proofs could be implemented so there is a secure, user-friendly data and identity storage and verification solution.

Many organizations have amassed an incredible collection of user data that is now being stored in centralized cloud servers where ownership is not truly known. There is no incentive for keeping your data safe for these big companies that initially collect the data. It's all about the lowest cost and responsibility on there part which causes big, centralized data hubs that have a growing and growing surface attack area from hackers. With DataLink DAO's incentivization model organizations and users and rewarded for participating, providing value, and ultimately securing the DAO. These users provide value by using the DataLink Validator and providing throughput in the entirety of the DAO services.

There needs to be access control over your data, you can't trust the big companies to do the right thing, there can be no trust in these systems. The everyday people need to be protected from large agencies that amass large amounts of data and carelessly manage it. Privacy is becoming almost impossible nowadays because of our data being collected everywhere.. There is now a way for transparency into these systems when necessary. The truth system that DataLink enables can trustlessly validate data and documents without knowing each one another's identity.

Blockchain has presented a new opportunity for the internet to have value. The meaning of Web 3.0 is the value layer being added to the internet, and ownership and access control back to the user. The importance of the user owning or having explicit access control over their files cannot not be understated. To avoid large-scale personal data loss to the unknown abyss or the dark web, we can have a system of decentralized identities, having access control over their data and documents and being able to share only what is permissioned by the user.

Proffer Data Ownership Back to the User

Data is yours. Personal data, or data collected about you should be owned by you. There is a popular phrase in crypto, “not your keys, not your coins”. The importance of having ownership of your personal data cannot be underestimated. Datasets all over the world are only growing and we are seeing the first fully-tracked generation growing up right now. Data is being collected from when zoomers were adolescents and their millennial parents were giving their babies myspace pages. This data is invasive and when not controlled by the user could lead to nefarious uses that may not even been the intention in the first place. When data is not owned by individuals, then there is no control, accountability, or visibility into how that data is used or misused. The same holds true for user data. Hypothetically, if a company like Brave browser joined the DAO it could protect browsing habits of users by proving that their user’s data was truly owned by the user and they have access control over their data. They could see their data NFT and be able to view every time there was a request to that data.

A non-transferrable NFT would represent an individual’s profile or entity. A recently proposed primitive by Vitalik and team called Decentralized Society: Finding Web3’s Soul[1], explored this avenue of design. The primitive discussed in that paper was called “soulbound tokens”(SBTs), and discussed non-transferrable NFTs that would belong to a user and contain credentials and social, organizational affiliations. The non-transferrable NFT proposed in this paper would be represented as blockchain diplomas, or healthcare records, would be inspired by this SBT design but implemented on existing tech stacks such as ERC-721. The non-transferrable NFT that is owned by the user would contain verifiable documentation related to credentials, personal data, and any other use case decided in the future. The data NFT is stored in a decentralized manner so that security remains paramount. These challenges around decentralizing user data, and offering the users the chance to own their own data can be solved with tools that are

just now coming out into the Ethereum landscape. In the metaverse age, it will be paramount to have access control over your data, over your documents, over your images...or your identity could be stolen, and in a digital age that means you have the possibility to be cut off from the economy and society. The space is constantly evolving, and the main focus should be how to validate real user from AI, imposters from truthful users, and data including images and videos are true and not fakes.

Validating Data and Identity

The information age has allowed for numerous digital tricks that affect real life. Identity theft impacted almost 1.4 million people in Mexico in 2020. With digital manipulation of certified documents on the rise, it is crucial to verify people's identities and professional credentials when doing everyday business. Otherwise, you run into risk and harm that could be irreversible. People interact with tricksters every day and may not even know it. Some doctors have faked their educational certificates, some law enforcement could have fake credentials, and countless bots, this also contributes to lots of false data. Being able to identify a verified user or organization without exposing the entirety of their information is a crucial problem to explore. Blockchain and current data harvesting entities could turn data and identity validation into a very scary world. Think of AI automatically detecting your identity, what you are interested in, who you are talking to, allowing no privacy at all for the everyday human. This is already happening.

The information age is progressing quickly and as more data is collected about everything, having a protocol validate the data’s entire lifecycle is crucial to knowing whether the data or identity that you are accessing is true. With existing blockchain technologies we can build new ways for anonymous validation of data and identities. Some of these data NFTs would be soulbound and would be non-transferrable. Such use cases as diplomas, healthcare records, licenses would be included. Other data NFTs in our decentralized data-sharing organization could be transferrable such as mortgage

deeds. The user would have access control over their own private information. This is important in cases of healthcare records or other documents and data that hold your private information that the user wouldn't want controlled by a large corporation. Protocols such as Chainlink DONs can have all the sensitive data be hosted off-chain and only the hash of the data transfer transaction takes place on chain.

Bringing "Crypto" Back into Cryptocurrency

In 2022, there are a lot of fake companies, groups, projects that are still scamming people out of money. Using buzzwords like blockchain, web3, web5, to capitalize on an industry that is built around the core concept of cryptography. A handful of elite companies are paving the way for adoption of distributed ledger technology and the cryptographic means that enable it. DataLink DAO will be one of these pioneers to use the latest zero-knowledge proof, oracle, and other cryptography techniques to make the most trustless and secure data-sharing group available. Zero-knowledge proofs are key in the infrastructure of the DataLink Validator. The Validator will use an attestation model and zero-knowledge proofs to keep the user's identity anonymous from the requestor of his/her data. When there is a request for data or the validation of that data a zero-knowledge proof is created by the user and the requestor receives the verified information without knowing of all the user's info. This system combined with a system of non-transferrable NFT's is where this DAO structure gets exciting. If the user owns a set of documents with different personal information on each of those documents, that identity can be confirmed using a ZK-SNARK attestation system. Providing an automated, anonymous, decentralized, secure data-sharing system. This new way of partially revealing data will open immense use cases for automated validation of credentials while protecting and giving access control back to the user. The data is stored on a decentralized storage network that will host the data privately and will be

pinned so the location of the data NFT is kept permanent. Access control to this document would is absolutely crucial in keeping this knowledge private for the user.

There is a lot of novel research going on currently around anonymous data-sharing. Such technologies as Multi-Party Computation, ZK-proofs, plural privacy and some fantastic tools associated with those including async mpc, VDF, and garbled circuits. These on-going research tools will be tested and implemented into the DAO's data validation services. Tools such as async mpc and VDF will allow new design flows to be explored for intuitive user access control. A decentralized oracle network could add an even greater layer of security and decentralization to this DAO. When this DAO is successfully operational it would interesting to explore the possibilities of multi-party computation in the DAO. All users would have their private inputs trying to prove something is true about that dataset. This would work in the DataLink DAO because of the attestation model and having anonymous inputs can allow analytics to be shared as true even without showing all the personal data it retrieved to come to that analysis.

Truth-Based Implementation of Data Ownership

DataLink will have a powerful onboarding tool to allow legacy systems to onboard their data with ease. Many legacy systems cannot actively use blockchain technology because of cost, scaling concerns, list could go on. An easy-to-use dashboard with a suite of tools that eases the bootstrapping process for organizations to upload their data to the DAO space. When the organization uploads their legacy data to be traced by blockchain by hash the ability to make money off that data instantly arises. This verified data has value to 3rd parties including other locations, groups, that now use your data for free anyway. This data can be shared via a zk-proof request coming in, and rates are charged amount of requests or possibly size of requests. DataLink DAO will not hold any information. The issuer will remain in control of the user's personal information. DataLink Foundation can help in any support means needed from issuer to user. A request ticket is raised and

the issuer will fix the data error and re-generate a new credential and hash. Only the hash of the credential is published on-chain for cost-saving and privacy. By having the user generate the ZK-proof, the user essentially signs over read-only access of their data, giving access control back to the user. The tech stack will be geared towards a UX-first experience for easier onboarding and data management. The DataLink Dashboard will have easy look-up tables, and other external query tools to help manage their decentralized data and metrics, including rewards if applicable, throughout the data lifecycle. Support through onboarding templates for data sharing customizable by developers, it's just as easy as adjusting smart contract to your parameters. The DataLink "Template Builder" will allow internal teams within legacy organizations to collaborate with the DataLink Foundation team to help with any contract setup, api connection, template support, etc. There will be a set of default templates for different typical datasets, for various legacy industries. The organization will provide information to the contract based on available templates. To create a tokenized document, the organization admin will fill out the fields, upload any files, filters, etc. to create the NFT. This NFT will hold the payload that are to be shared within the DataLink DAO network. There will be a simple search on all public data within the DataLink. Public datasets that are in the DAO can be queried and monetized. The original file can also be validated because of the initial hash that is generated when the original document NFT is generated by the issuer. This data is traceable all the way back to creation and the file can be traced with every interaction involved. These private credentials that are generated can be attributed to a decentralized identity for anonymous and monetized data-sharing and tokenization.

New DAO in Town – DataLink DAO

DataLink would be an ideal example on DAOs can espouse the best practices for cybersecurity, data-sharing, and data validation. For organizations to share data

between them securely and trustlessly is an important issue that many legacy data-sharing organizations fall short. ISCO's are a great example of how organizations and governments can get together and share valuable data in a mutually beneficial way. There is an incentive problem with these organizations. Beyond being financed with membership fees, there is no real funding or incentive flows in these groups.

Bringing value to data is becoming incredibly important and profitable in this information age. A DAO could help with bringing value to every single data transfer within this new organization. Instead of data-sharing being a gray area between companies and groups, there should be absolutely, automated trust-less contracts that have pre-defined expectation for both parties. There can be no trust in data transfer when the new digital age supplies so many fake identities and attack vectors. There is no in-person handshake in most digital contracts nowadays. Individuals using the internet to communicate and transfer value are protected anonymously if they so choose, which must be protected. So, to fix this trust issue, there can be no trust. Trustless communication defined by cryptographic truth will be the preferred, if not the only offering in future business deals that hold tangible, and intangible assets. DataLink DAO can offer a group of organizations alongside the DataLink Foundation to form a powerful, DAO secured by the value of the data within and by network usage over time.

The DataLink DAO Foundation will be integral part of the DAO functionality. The Foundation will enable legacy organizations to on-board easier and with greater initial support. Legacy organizations could reach out to DataLink for api help or tokenizing their legacy data into the protected DAO ecosystem. Users and organizations have an official support channel if needed and is critical to the success of any DAO. The Foundation will not be the authority in the DAO, and will only execute authority in moments of extreme threats to the DAO in monetary or societal terms. Along with product development and support, the DataLink DAO Foundation will have legal parties, treasury custodians, operations, marketing, and other key operational pieces that will help maintain the longevity and legality of the DAO and all the participants.

A DAO can have countless positive impacts to the way companies and user's share their data. Industries can immensely benefit from this new type of collaboration including but not limited to; validated educational certificates, tracking and anonymous sharing of healthcare records, transparent validated tracking for supply chain, universal proven professional licenses, automating credit checks for mortgage and auto, and automating validation of documents for verification services. A data-sharing DAO can be structured in a way to be profitable for it's participants. There are lots of requests to data everyday and participants of the DAO could band together and say we want to be paid for access to our data. The organizations within the DAO can come together and vote on a fee increase or an entirely new fee structure. Other voting proposals that could come about in the data-sharing DAO would be votes to ban someone from the DAO, votes to reimburse a company or user in case of hack, percentage of revenue stream for all holders of the participant DAO NFT, maybe a vote to break up a miniDAO which has gained too much control, revenue, etc. A miniDAO is a group of organizations, users, or locations within those organizations, that come together on agreed set of terms within a smart contract. The smart contract is signed by both parties and the miniDAO is launched and permissioned data-sharing is enabled between those two parties. That miniDAO could possibly take in revenue from external calls to their permissioned database. A company like Brave, the popular crypto browser, could join the DataLink DAO, create a miniDAO, and only provide rewards to validated users that are known not be bots or fake accounts. The Brave browser could then have upsell to the users that join their miniDAO, the ability to sell their browsing data to receive more familiar advertisements. The user would consent to the sharing of their data, and in return the user would receive the fee revenue for a call to their data.

User reputation could be explored in this system to represent that each organization or user produces good verifiable data, and less potential for fake data to come into the DAO. A reputation system could be implemented in the public database with the

DataLink DAO to attest that this public data is good so it can be purchased by analytics companies who yearn for truthful data.

Incentivizing Data Ownership and Monetizing Data - Sharing

The organization who wants to use the DataLink DAO will have to purchase a technology license to access the validator services. This governance NFT will have the right to read-only for users and read-write for organizations to tokenize their assets to enable monetization of sharing data. DataLink DAO will have a fee for every external validation of data request that comes in. There will be additional fees for when there is a mini-DAO setup for group-to-group permissioned sharing within the DAO space. The users and organizations will have the option to pool their data into a public DAO database that is offered to 3rd parties for intake and use. This data will be extremely valuable due to the certainty of the validity of that data that is permissioned to be shared. Analytics, advertising, retail would be interested in this kind of verified public data. DataLink DAO users and organizations will have a chance to earn a share of the revenue generated by the DAO's fee collection structure. The small percentage of passive income will be based on engagement within the DAO. Users must interact with the DAO through various ways. For participation, simple captcha would be provided to show proof of human until SBT structure is implemented properly. Otherwise, greater rewards would be offered based on the amount of data tokenized, requests received, and/or sharing of public data to the DAO, the rewards would be additive in nature earning a certain yield for the user, paid out quarterly. The revenue or reward flow can be altered by DAO governance votes.

The DataLink DAO will be run by an on-chain democracy. This will be effective by incorporating soul-bound tokens into our protocol. A type of decentralized identity solution that takes in documents that are known to be verified by legit organizations and governments. When requests come in the documents verify the user's identity without the requestor knowing all the user's details or their identity at all. These soul-bound

NFT tokens are the exact solution DataLink is building their solution around. Documents verified using ZK-SNARKs, enabling identities, and credentials to be verified anonymously and securely. As the world turns more to digital processing of every kind of documents it is important to realize the cost of doing manual credentialing of documents. This old way is costly, introduces human error and fraud and this sort of process can be automated away thanks to new advances in ledger technology. The new way of decentralized data storage and a novel way of decentralizing the sharing of that personal data can be automated to enhance speed, security, and privacy.

Reputation is an important aspect of any reputable DAO. It will be no different in a data DAO like DataLink. Users and Organization need to have reputation attached to their decentralized identity to promote participating a secured system of data-sharing such as DataLink. Reputation could even have impact of the revenue stream received from the DataLink DAO. If reputation is tracked then, validity of documents and identity is enhanced and additional features can be constructed to protect against bots, nefarious actors in the DAO, and most importantly the overall operation and security of the DAO and the data living within. Illegal data could live in the DAO but be removed thanks to a reputation and governance system. DataLink DAO Foundation will also support legal affairs around activities that occur within the DAO. Reputation can increase with the amount of uses of the DataLink Validator or decrease if attestations come back in error or false once or repeatedly. Reputation could also have uses such as nodes validating a document or piece of data is repeatedly true and at every function of 10 the reputation increases, enabling research papers to be reviewed and scored in reputation, and rewards reviewers with reputation for their validated review.

Conclusion

Blockchain is the value layer of the internet. It is obvious from the rampant hacks and user data loss that there needs to be research and new solutions offered. Data can be

stored in a decentralized manner and zk-proofs could be implemented so there is a secure, user-friendly data and identity storage and verification solution. DataLink DAO will be a living, breathing truth engine powered by people’s goodwill and collaborative effort. The solution to verifying individual's credentials, and other data that needs to be publicly validated is DataLink.

Blockchain has presented a new opportunity for the internet to have value. The meaning of Web 3.0 is the value layer being added to the internet, and ownership and access control back to the user. The importance of the user owning or having explicit access control over their files cannot not be understated. To avoid large-scale personal data loss to the unknown abyss or the dark web, we can have a system of decentralized identities, having access control over their data and documents and being able to share only what is permissioned by the user. Imagine a database of files and documents across the internet that you know to be true. Amongst the sea of fakers and bots out in the digital world, verified data that has a digital stamp of truth that any requestor can verify themselves easily and instantly. DataLink is that stamp.

Acknowledgments

A special thanks for the inspiration, knowledge, collaboration, and drive to research the immense endeavor of building a trustless society in the information age. Thank you, Lindsay, LaunchCode, Chris, Josias, and other past hackathon teammates, Ethereum / Chainlink community, IC3, defi-learning.org and all the professors and mentors.