KnowledgeArc.Network

Table of Contents

Abstract

Launched by the company KnowledgeArc, the KnowledgeArc.Network is a new digital asset management system for decentralized and democratized storage of the world's knowledge, using blockchain technology. It is an ideal platform for storing, permanently, digital assets and the metadata that describes them.

Blockchain technology offers a perfect solution to the problems faced by archiving; how to achieve immutability, persistence, longevity, uniqueness. Using the technology for asset ownership, storage and identification guarantees permanence. By its nature, a blockchain-based archive is impervious to data loss and attack by malicious parties, and is permanent.

Our Vision

At KnowledgeArc, we're passionate about ensuring the integrity and longevity of digital assets. We have been officially established in the archiving industry for five years now, providing infrastructure, support and customization to leading institutions, government organizations and universities.

In an industry controlled by expensive software solutions and closed corporations, it has always been our mission to provide high-quality, low-cost solutions which are open source and public access.

The blockchain offers a perfect solution to the problems faced by archiving; how to achieve immutability, persistence, longevity,

uniqueness. Utilizing blockchain for asset ownership, storage and identification guarantees permanence. By its nature, a blockchain-based archive is impervious to data loss and attack by malicious parties, and is permanent. We envision a fully distributed, democratized archiving ecosystem, free of corporate control and inexpensive to use. We call this the KnowledgeArc.Network, a platform for storing, permanently, digital assets and the metadata that describes them.

Stakeholders

- **Libraries**: KnowledgeArc.Network provides true immutability and permanency for libraries where users can access peer-reviewed, open access information. Current systems such as DSpace cannot provide this since their databases can be corrupted or lost, or have their information changed by users through the user interface.
- **Academic Archives:** Using KnowledgeArc.Network, academic archives can deploy their own local archive or connect with other repositories. Because of the nature of decentralization, archives can be easily disconnected from then reconnected to other networks. Unlike legacy archiving systems such as EPrints and DSpace, data will be immutable and permanent.
- **Research and Peer Review:** Currently peer review and academic publishing is controlled by a few very large private organizations. They exert a lot of power over what can and cannot be published. KnowledgeArc.Network breaks down these barriers of entry by making this middleman redundant. Institutions can gain direct access to peer review and peer reviewers and publishers can profit directly instead of being subject to a larger entity who takes most of their share.
- **Funders and Those Seeking Funding:** KnowledgeArc.Network is focused on academic information. Therefore, there is a need to incentivize the community to engage in more scientific discovery.
- **Smaller Institutions with Limited Budgets**: Smaller institutions usually find it difficult or impossible to compete with larger publishers and universities. Maintaining an archive can be an expensive proposition, and some institutions with private collections of information might have no budget at all.

KnowledgeArc.Network provides an ecosystem which enables anyone to share information at next to no cost.

In fact, because the KnowledgeArc.Network software is free and open source. Because it uses decentralized technologies to run, a user could potentially run an archive for free on their local computer. The institution can potentially generate revenue from the archive using KnowledgeArc.Network's micro-payments, where cited or re-used works can collect fees from usage. They can also use KnowledgeArc.Network to trade or exchange archived information.

The Archive (ARCH) Token

We have launched our new ERC20-based utility token, Archive (ARCH).

Designed as method of exchange across KnowledgeArc.Network, ARCH tokens will serve as a form of payment and will power the KnowledgeArc.Network ecosystem.

There are currently **1,025,000,000** ARCH tokens available.

Token Distribution

Tokens will be distributed as follows:
- **500,000,000:** For sale to users of KnowledgeArc.Network. Tokens acquired from this general sale can be used for purchasing services within the ecosystem, e.g., distributed IPFS backups, peer review, academic funding, paying authors through sharing and citation incentives.
- **100,000,000:** To be deployed every 12 months over a 5 year period. Our aim is to incentivize these token purchasers to use their tokens to fund the development of the ecosystem; pay for backups, fund academic projects, pay for peer review, micropayments for citations, sharing, social impact.

- **Up to 100,000,000:** To be airdropped to existing customers. Our aim is to incentivize our existing customers to move their information to the decentralized ecosystem. Our customers are large institutions with great influence over how information is archived.
- **400,000,000:** Organizational requirements are as follows:
  - Employee and contractor payments with vesting periods of up to 12 months to protect circulating supply.
  - Founders will be required to vest their tokens for up to 5 years and strict controls will be placed around how and when founders sell their tokens.
  - Provide a funding pool for 3rd party projects which contribute new features and integrations, and a focus on open source.
  - Funding of scholarly works and scientific discovery with a focus on academia in developing countries.
  - Incentivize participation by experts in various academic, archiving, journalling, peer review fields.

Incentivizing the Community

An obvious use-case for the ARCH token is accepting it as a form of payment for backing up assets to the decentralized cluster. However, we also envisage providers being able to connect their own IPFS nodes to our backup ecosystem and receive compensation in ARCH tokens. Users can rest assured that their data is truly distributed and providers can actively participate in backups and be rewarded for their efforts.

We envisage a marketplace where providers can register their IPFS Cluster nodes and users can rent backup space. When a provider is selected, they are added to the list of backup nodes the user controls and the user's data is written to the provider's infrastructure. Users will be able to select the provider nodes they want, allowing them to decide on how many replicas they need and in what locations they would like them stored to. Archive will be used to pay for resources and providers can price their nodes as they see fit. Users can choose the best price, whether it be based on the lowest cost or the most reliable provider.

The Next Iteration

The next step in putting the ARCH tokens to work will be to include a decentralized-based solution for backing up assets to an IPFS-powered cluster. Users will be able to pay for backups using ARCH tokens and providers can add IPFS nodes to the cluster and be reimbursed in ARCH tokens.

For the long-term, we envisage our entire ecosystem powered by ARCH tokens. Staking, peer review, sidechain support, authentication; everything archive-related which requires a payment will be paid for using ARCH tokens.

New use-cases will present themselves as the KnowledgeArc.Network project evolves, but we also recognize that usability needs to be something that happens as soon as the token is launched.

We want to make sure people can easily obtain ARCH tokens and so are engaging various exchanges. Because we are open source proponents and believe in the power of democratization of academic information, we are currently focused on decentralized exchanges. This does not however preclude us from approaching larger more centralized exchanges in the future as the opportunity arises.

KnowledgeArc.Network Ecosystem

KnowledgeArc.Network's ecosystem is already under development.

It will consist of smart contracts, designed to store critical information about an archived asset. The asset and its metadata will be stored on a decentralized system too, utilizing blockchain technology to ensure assets and metadata are permanently stored and easily retrieved.

The Decentralized Archive chart (figure 1) details how the archive is structured and how the system stores assets and maintains information:

Decentralized Archive

Archiving contracts form the basis of the decentralized archive.

Archiving Contracts

Item-Specific Contracts

Item-specific contracts store information about archived items, manage ownership, admin details, etc. and can be bought and sold.

Item

Embargoed Item

Item-specific contracts can allow for different types of functionality, e.g., a time-sensitive contract might allow access to information only after a particular amount of time.

Collection

A collection contract can store multiple items which group together logically.

Token

A token contract might provide a mechanism for buying and selling archived items or for powering decentralized services.

Metadata is stored outside the contract to work around the limitations of some blockchains. As with the file store, a decentralized database such as OrbitDB would store metadata per contract address.

Files are stored in infrastructure designed for rich formats. Ideally, this would use a decentralized system such as IPFS.
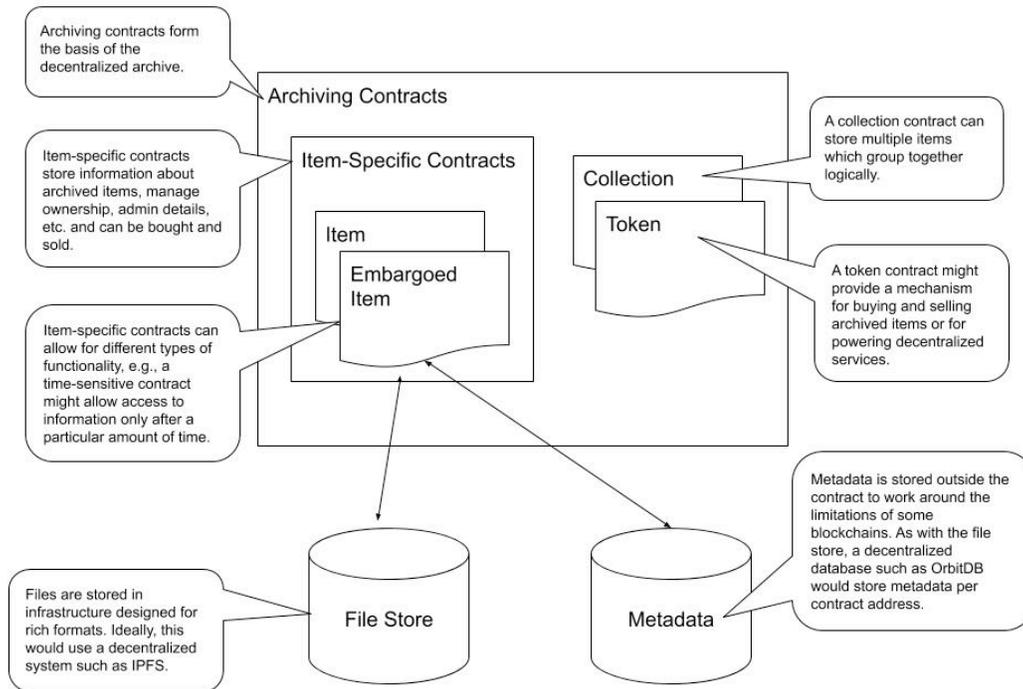
File Store

Metadata

Figure 1: Decentralized Archive

Owners of archived items will be able to trade their works in open, decentralized markets, transferring ownership seamlessly. Additional blockchains will be employed to provide services such as identity management, citations, persistence and other archive-related requirements.

The Archive Integration chart (figure 2) details how the decentralized archive interacts with programming and the user interface:
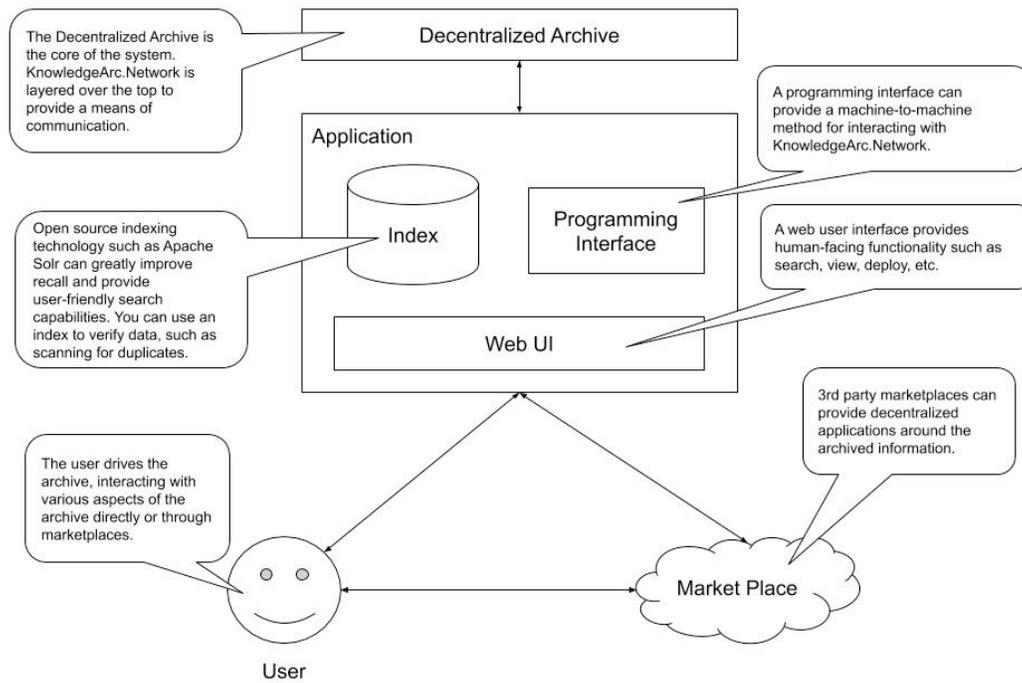
Figure 2: Archive Interaction

We have spent the past year tokenizing various academic material. Using a combination of ERC721 non-fungible tokens, IPFS, and OrbitDB, the team is now focused on developing an open-source decentralized archiving ecosystem that can be implemented by anyone.

Everything up until now has been research and prototyping. Now it is time to implement real world solutions which can be used in production scenarios.

Ecosystem Components

The KnowledgeArc.Network's ecosystem will be built and run by integrating the following tools:

DSpace

Our first implementation will be decentralized backups for our existing archiving platform. The KnowledgArc Platform is built upon multiple open source technologies, with the central pillar being DSpace, an open source digital asset management system that provides an out-of-the-box archiving solution.

Backup solutions available for this technology are centralized and often costly however. Our goal is to address this by providing a distributed solution which is cost effective and engages the community, both users and providers.

Development has only just begun, but we are already deploying a number of IPFS nodes coordinated through IPFS Cluster to provide persistent, decentralized storage for DSpace assets. We are also developing tools which integrate with the DSpace platform; these tools will push assets directly to IPFS Cluster in real time, ensuring they are backed up with minimal chance of data loss. More importantly, backups will not rely on one or two big data repositories, i.e. a single point of failure.

IPFS, OrbitDB, and Etherium Swarm

DSpace not only stores assets; metadata is key to describing what those assets represent. This metadata should also be backed up in a distributed manner but current backup solutions do not provide tools for robust metadata backup. Metadata is as important to the archive as the assets it describes and ensuring its permanence is vital to a fully-implemented backup plan. This is where the InterPlanetary File System (IPFS)-based OrbitDB plays a key role.

OrbitDB is a distributed database built on top of IPFS. Once we have implemented asset backups using IPFS, next on our roadmap will be pushing metadata to OrbitDB. We will also expand file storage to Ethereum Swarm.

Po.et

Successful management of copyright is integral to an academic repository. Attribution, citation and licensing all depends on clear terms of use as outlined in an archived item's metadata.

As we move assets and metadata to decentralized systems such as IPFS, Swarm and OrbitDB, we will also migrate licensing to Po.et, a decentralized solution for content ownership. Copyright and licensing can be tracked and attribution easily determined because Po.et is an open system. Manual or, in the case of KnowledgeArc.Network, automated methods can be easily utilized to verify ownership of assets on the distributed archive.

Current Implementations use centralized methods to store copyright and licensing terms and these terms can be easily changed or manipulated at any time. What is needed is an trusted, immutable ledger of time-stamped items, stored in a way that is accessible to all.

Because assets will be stored on a decentralized system such IPFS, a hash will be used to identify the asset uniquely.

This asset's hash, along with associated metadata, will be stored in a decentralized database, currently OrbitDB, and the owner will "sign" both the asset and the metadata.

Both the asset hash and metadata will be pushed to Po.et, where it will be time-stamped by the distributed licensing ecosystem, and a hash will be returned, identifying the copyrighted material. This hash will be stored against the item's metadata, again utilizing OrbitDB.

Items can change over time. However, assets on the blockchain are truly immutable. Even if the same asset changes, it is recognized as two separate pieces of data on the blockchain. Therefore, an item which changes will be linked with its previous versions and will be issued a separate Po.et hash.

Both the KnowledgeArc.Network and the Po.et ecosystems allow each platform's native token to carry out various tasks. We envisage incentivizing across both platforms using various methods of exchange.

Po.et's token POE and KnowledgeArc.Network's token ARCH both utilize ERC20 tokens— this means that various platforms exist that can allow for easy conversion from one token to the other.

Potentially, users will not even need to be aware of what token that they are using, but only that a small cost or other incentive is required to carry out a particular task.

LBRY

LBRY provides a single index of published content; E-Books, audio, video, all stored on a public blockchain. Content providers control LBRY, cutting out large corporate manipulation and 3rd party censorship. Providers can also generate direct income from their content without a single organization taking a chunk of the profits. Everything on LBRY is open source, decentralized and owned by the community.
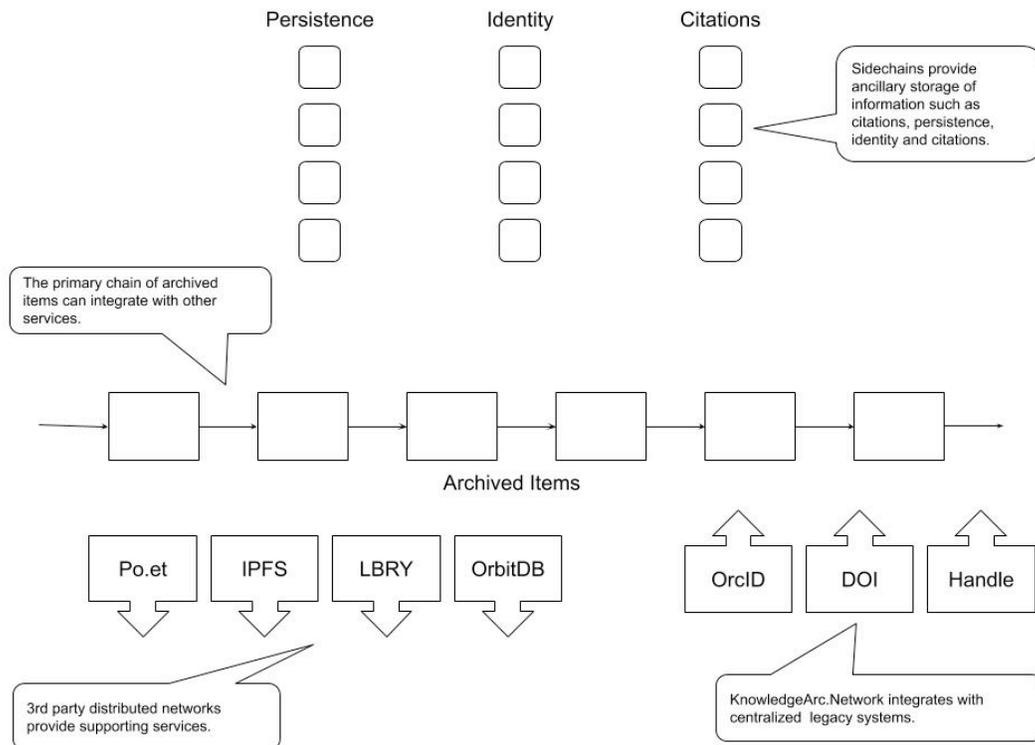

Figure 3: Sidechains & Integrations

The Sidechains and Integrations chart (figure 3) details how sidechains provide ancillary storage of information.

LBRY' philosophy of democratic, decentralized, open-source digital storage perfectly aligns with KnowledgeArc.Network, and provides the ideal solution for distributing open accessed academic material.

The Ultimate Goal

So what's the point of working on something like backups? Isn't this about blockchain and decentralization and democratization of data. Isn't this about removing the middleman (or middleware) and replacing it with something that works anywhere, anytime?

Academic institutions can be reluctant to embrace nascent technologies; understandable, as they invest huge amounts in the technology they implement and need to be able to commit to solutions which can last the test of time. Archives are a perfect example; organizations have to think in terms of decades when it comes to storing academic material; material which must be immutable and permanent.

Introducing a decentralized backup system based upon IPFS and OrbitDB allows institutions to "dip their toes" into the new distributed web and gives them time to evaluate the full impact of blockchain and cryptocurrency technologies on their archiving requirements.

They can then implement decentralized technologies without having to fully commit to them. When the time is right, these organizations will be able to migrate over to this new archiving ecosystem without the headache of a major shift in technology.

Project Roadmap

Ongoing solutions will present themselves as we continue to develop the KnowledgeArc.Network ecosystem. The following dates are estimates and dependent on technology, resources, and budget available:

Goal	Status
Deploy Archive (ARCH) ERC20 token	Deployed, January 2019
Implement IPFS-based OrbitDB as a decentralized delivery mechanism for JCar Assets	Released, March 2019
Develop JArchive, a new Digital Asset/Document Management System. The initial JArchive project will be centralized but will integrate the following decentralized elements to ease the move from centralized to decentralized applications: • IPFS OrbitDB • Ethereum • Po.et • Lbry	3rd Quarter 2020
Archive asset metadata to IPFS and OrbitDB from JArchive	3rd Quarter 2020
Integrate Po.et claims mechanism with DSpace, Open Journal Systems (OJS) and other existing open source academic systems	4th Quarter 2020
Use Po.et to claim decentralized archived items. Store hash on the decentralized database	4th Quarter 2020
Deploy IPFS cluster for KnowledgeArc Platform Decentralized Backup (also called the Decentralized DuraCloud).	4th Quarter 2020
Introduce BYO IPFS cluster node and connect consumers with providers. Consumers will pay for the service using the ARCH token.	2021

Goal	Status
Integrate other content management systems (such as Wordpress, Square Source, Drupal, etc.) with the decentralized archive.	2021
Permanence and immutability of assets and metadata through ERC721 smart contracts.	2021
Social sharing, citation and attribution micropayments.	2021