

# The Data DeFi Protocol: A Deep Sea Dive Into Ocean Protocol

Karim Halabi

Jul 29, 2021

As we move further into the 21st century, we become more reliant on a new resource: data. This resource is used to inform decisions and advance our understanding of both the natural and man-made world. Data is used to upgrade our systems, develop drugs, and engineer machinery - among countless other applications.

The sheer amount of data that now exists is staggering, and even more staggering is the exponential growth of the rate at which it is generated.

Similar to many other resources, the value of data is unlocked when it is refined and put to use. The issue faced by many industries is that much of this precious data cannot be accessed in the first place to draw insight from it. According to the [International Data Corporation](#), a mere 0.5% of the world's data is actually used and analyzed, with the remainder sitting dormant in silos.

Data lies dormant due to the fact that no real incentives exist for owners to share this resource - aside from a healthy dose of game theory - and in fact, they may be discouraged from doing so for fear of giving away their IP, aiding the competition, or privacy and security concerns.

Although modern legal and accounting standards have not yet recognized it as such, data is an asset - and an extremely valuable one at that. Information can be leveraged to serve the aforementioned functions and contribute to the value of any organization where it is relevant.

The tokenization of data allows for it to be treated as a tangible, liquid asset. Protocols such as Ocean Protocol are making the financialization of data a possibility.

## **Ocean Protocol Making Waves**

At the simplest level, Ocean Protocol has three core functions:

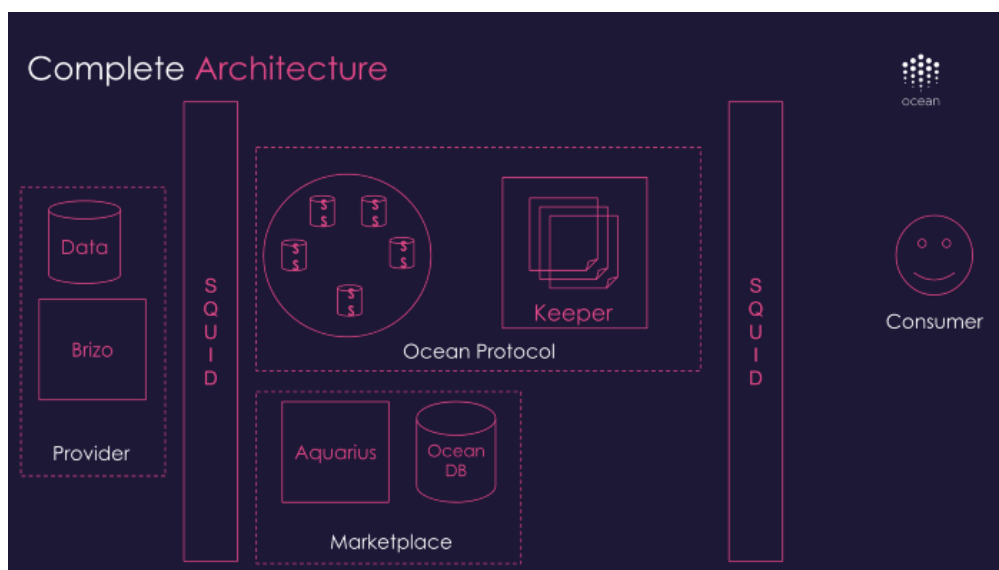
1. Encourages the public sharing and monetization of data
2. Permits individuals or entities to securely share sensitive data
3. Enables better price discovery by creating a more liquid market for data

How exactly these are achieved, and what the implications are, shall be examined and discussed.

## Exchanging Public Data

Ocean Protocol is comprised of several core components including:

- Provider - the data owner, who sells access to an asset and runs an access control proxy (Brizo) which accepts, verifies, and executes requests
- Ocean Protocol - a collection of EVMs and smart contracts, including the contract for the Ocean token, that maintains a registry of asset ownership, provides decentralized storage, and facilitates secure data exchange
- Ocean Marketplace - a decentralized exchange for data assets that uses Aquarius and OceanDB to store asset metadata
- Squid - the tech stack that ties it all together



Source: [Ocean Protocol](#)

The [Ocean marketplace](#) is a product of the protocol where data owners can publish their data sets and list them for sale. Since the Ocean protocol consists of smart contracts built on Ethereum, the sharing of data executes on a peer-to-peer basis without any overseeing third party. This ensures owners retain full ownership and control over their asset(s) until selling.

This marketplace and sharing of data is made possible through tokenization, however, it is not the data itself that is tokenized, but rather access to it.

When a data asset is published for sale, a corresponding "data-token" is also created - this is what is actually sold on the Ocean market. A data-token is an ERC-20 token that grants the holder access to a given set or stream of data. The owner of the data has the ability to decide how many data-tokens they'd like to issue and at which price they wish to sell access to the information.

Data providers are also able to set a host of other conditions including:

- Granting one-time access vs. perpetual access
- Time period limits for redeeming data

- Read vs. write access

This is akin to creating a token that can be bought on a decentralized exchange (DEX) and redeemed for a specific product/service. Ocean is an infrastructure where actors can do this easily and brings together data sellers and buyers.

With the ability to exchange data, Ocean enables individuals or entities to sell their valuable datasets to interested parties or individuals.

## **Exchanging Sensitive Data**

Ocean has a 'compute-to-data' function, wherein data scientists, researchers, and other interested actors can run algorithms and machine learning programs on valuable data sets without actually having the data itself. With this, sensitive data that cannot be made available for public consumption on the market can still be monetized and put to use.

The protocol facilitates this by sending the algorithm to the owner of the data set for approval, which if granted, is run on the data remotely. The results of the program (e.g. how the algorithm performed) are then sent back to the requester.

Therefore, data that may contain personally identifying information, IP, or other private information can be utilized for others' gain as well as allow the owner to profit.

What this would look like:

1. Bob owns private data that he doesn't wish to share, but also wishes to monetize
2. Alice is interested in improving her AI model using Bob's valuable data
3. Alice sends her algorithm to Bob so he can review it before it runs on his data
4. Bob approves/denies the algorithm. If approved:
5. Alice's algorithm runs on Bob's data remotely without it ever leaving Bob's digital premises
6. The enriched algorithm is returned to Alice
7. Bob's privacy is maintained

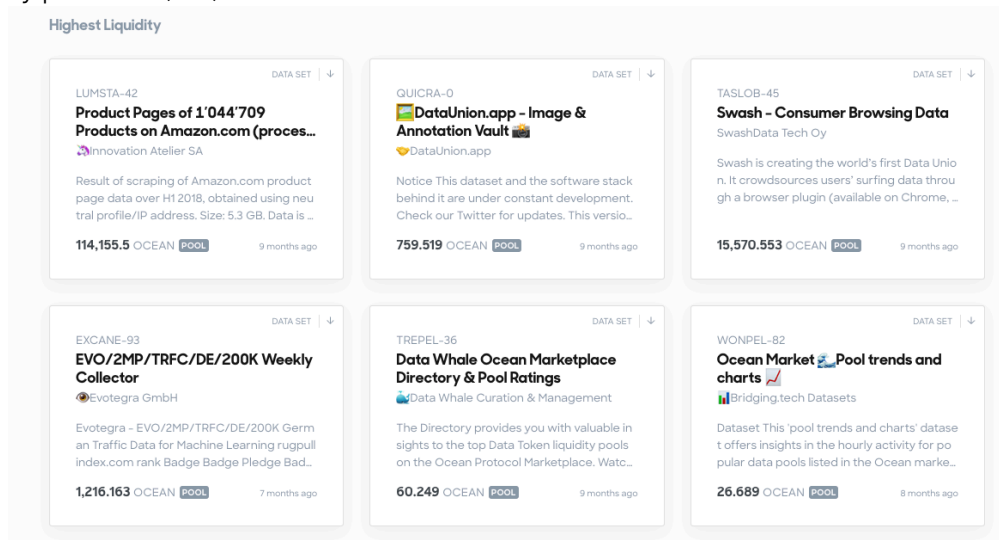
## **The Decentralized Data Economy**

When a new data-token is listed for sale on the market, the publisher can set their own price or let the market decide the value of the dataset through a "data offering". This data offering functions similar to an initial DEX offering on Balancer in which the data-tokens are listed and immediately available for purchase at the market price.

The open access to data offered by Ocean Protocol and increased price discovery of datasets through the Ocean Market allows data to be treated similar to a tangible asset due to its market price. If data is valued similar to other tangible assets (e.g. stock), organizations will be able to list data on their balance sheets and borrow against it to fund growth. Open data markets will be the catalyst that unlocks a greater value from data, and opens the door for its inevitable recognition as an asset class.

## Dive Into the Ocean Marketplace: a DEX for Data

The ocean market is built using [Balancer](#) infrastructure to operate as a peer-to-peer market. This allows it to function as any other DEX: with an automated market maker (AMM) and liquidity providers (LPs).



Source: [Ocean Marketplace](#)

When a new data set/service is published, a corresponding liquidity pool is also created. LPs earn their title of liquidity providers by doing exactly that: staking tokens in a pool in order to be rewarded with a cut of the trading fees, proportional to the size of their stake.

## Product Pages of 1'044'709 Products on Amazon.com (processed data)

DATA SET | [Luminous Starfish Token – LUMSTA-42](#)

Published By [Innovation Atelier SA](#) – Profile | Explorer

9 months ago

Result of scraping of Amazon.com product page data over H1 2018, obtained using neutral profile/IP address. Size: 5.3 GB. Data is available for 1'044'709 Products in 11'291 product categories. The data scraped has been processed to extract the main components of the product offering on the product page.

The following data is available:

- A .pickle file summarizing the tree of the dataset, which matches the structure of folders containing the data.
  - The last folders in the tree are the ones with data (product categories/segments)
  - Not all folders contain data
- Picture of product main picture thumbnail (JPG files in Image folders)
- Key information on products in CSV file for Top100 products in each product category:
  - category: the product category/segment according to Product tree (text following tree schema)
  - asin: the unique code of the SKU (text)
  - brand name: the name of the brand of the product (text)
  - title: the name of the product (text)
  - price: the price of the product (in USD)
  - rating: the average rating of the product (float)
  - is\_prime: whether the product is part of Amazon prime or not (1=yes, 0=no)
  - number reviews: the number of customer reviews of the product (integer)
  - description: the text describing the product on the product (text)
  - discount percentage: the % discount on price included in price (integer)

USE POOL TRADE

1 LUMSTA-42 = 114,155.5 OCEAN

Pool Datatoken

**Your Liquidity**

0 OCEAN	0 pool shares
0 LUMSTA-42	0 % of pool

**Pool Creator Liquidity**

435,114.884 OCEAN	100 pool shares
42.226 LUMSTA-42	84.48 % of pool
= €1,716,263.23	

**Pool Statistics** 10/50

**Liquidity** **Price**

515,073.242 OCEAN	118.376 pool shares
49.986 LUMSTA-42	10 % swap fee
= €2,230,808.52	

Fetching every 10 sec.

**ADD LIQUIDITY**

Your Pool Transactions [SHOW](#)

### Source: Adding Liquidity to a Pool

Trading fees are set by the pool creator (the owner of the data) and can range from 0.1% to 5%. In the Ocean market, providing liquidity is known as 'curating'. This is because stakers are incentivized to stake on datasets they believe to be of high value - as they will have higher sales, and LPs will subsequently receive greater rewards.

Through this, they are also encouraged to find relatively undiscovered offerings that are of value, which would give them a larger proportion of the pool.

Therefore, a greater stake on an offering can signal credibility and quality; helping buyers decide between offerings.

Ocean Protocol stats as of 24th July 2021:

- \$2,973,853 in total value locked (50% OCEAN tokens and 50% datasets)
- 419 data-token pools

Ultimately, a marketplace is created where people and organizations are incentivized to share their data. Ocean Protocol allows data owners to easily securitize data and sell it in a trustless, peer-to-peer fashion that does not compromise security, ownership, or privacy.

Ocean protocol can unlock the value of data in two ways. First Ocean can be used to enrich algorithms by providing training data (from real sources) that aids in algorithm development (increases accuracy, reduces bias, etc). Secondly, Ocean allows owners to monetize their data which is otherwise a dormant asset.

## The Transcendence of Data-Tokens

Although their *raison d'être* is to represent access to data sets and data services, data-tokens play a part beyond the Ocean market. As mentioned earlier, they are ERC-20 standard, and each data-token has a contract address.

In essence, this enables data-tokens to be held in Ethereum wallets, traded on other DEXs or CEXs outside the Ocean Market, and used with other Web3 protocols on Ethereum. The open nature ensures that data-tokens can be integrated into DeFi protocols, such as this [example](#) of data-tokens being used to create a Balancer Pool. Other protocols in which they could be used include [Enzyme Finance](#), where baskets of data-tokens can be treated like index funds, or even [UMA](#), where synthetics based on data-tokens can be created.

The composable nature of Web3 and native financial rails of DeFi expands the potential of valuing data and information. Data-tokens can be used as collateral for loans or stablecoins, and can be insured using DeFi insurance protocols expanding the potential for value creation and yield generation.

## Ocean Protocol Developer Tools

Ocean Protocol was originally formed as the brainchild of [BigChainDB](#), an organization that enables distributed databases to be used with blockchain and protocols like Ocean.

A core component of the Ocean protocol ecosystem is its robust developer tooling:

- Aquarius – off-chain storage of listed data assets’ metadata
- Squid – The Ocean tech stack, which connects all components together
- [Manta Ray](#) – an introduction to the protocol and how to use it for data scientists
- Ocean.js – the protocol’s javascript library

A suite of repositories and developer tools have been created by the core team to sustain the growth of the protocol. This activity can be observed on the [Github page](#), where two of the main repositories - Ocean.js and Aquarius - are consistently updated.

*Ocean.js*



Source: [Github](#)

*Aquarius*



Source: [Github](#)

Additionally, there have been various tools created by Ocean ecosystem developers that help navigate the market including:

- [Rug-Pull Index](#) - helps ecosystem participants identify which liquidity pools are at higher risk of being 'rugged'
- [Alga](#) - a phone app to track price action and history of datatokens, as well as other features such as portfolio tracking and an integrated wallet. It can be described as a 'Coin-MarketCap for data-tokens'

## What About the \$OCEAN Token?

The OCEAN token possesses three main functions:

1. Curation of data (staking)
2. Means of exchange: to purchase datatokens
3. Governance of the grants DAO

### **Data Curation**

Since the Ocean market uses Balancer pools to create data sets, this also presents the ability to use other DeFi primitives to incentivize behavior and liquidity. Users are rewarded with newly minted tokens for providing high quality, relevant data and keeping it available.

### **Means of Exchange**

Similar to many cryptonetworks, Ocean Protocol, also utilizes its own token to control monetary policy and prevent an individual or group from gaining a disproportionately large influence.

### **Governance**

Control of the protocol and development currently resides solely with the core team, however token holders have the ability to propose and vote on additional developments which, if successful, are granted treasury funds. Additionally, Ocean token holders have influence over [OceanDAO](#) - a grants DAO for the Ocean ecosystem.

## Token Economics

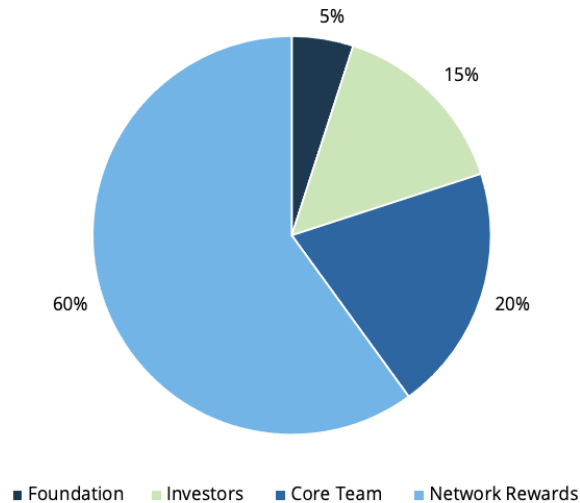
### **Supply**

- Current circulating supply: 434,026,836
- Total supply: 613,099,141

### **Distribution**



## Ocean Protocol Token Distribution



Source: Ocean Protocol

Founders have lock-up periods of 5 years since the founding of the network. This means there is a larger token vesting unlock in Q1 2022 that could result in downward price pressure. The 20% allocation to the core team includes the 100,000 tokens that are allocated to each adviser as well - of which there is a respectable amount.



## Ocean Token Distribution and Minting Schedule

Years	0	0.5	1	1.5	2	2.5	5	10	15	20	25	50
Percent Distributed	0%	0.83%	3%	6.2%	14.0%	21.1%	48.9%	78.5%	91.0%	96.2%	98.4%	99.98%
Percent Remaining	100%	99.17%	97%	93.8%	86.0%	78.9%	51.1%	21.5%	9.0%	3.80%	1.60%	0.02%

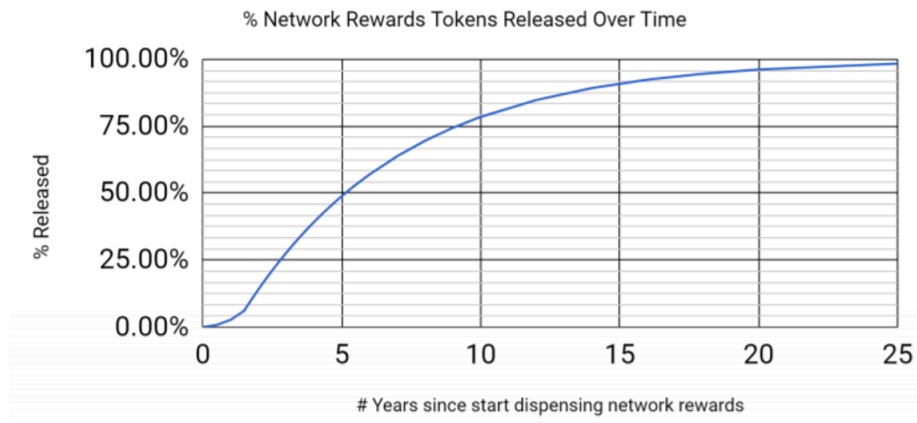
Source: Ocean Protocol

The Ocean distribution and minting table shows that the final token unlocks will occur by mid-2022, meaning that new tokens introduced to the supply after this point are purely network rewards.





## Ocean Network Rewards Distribution

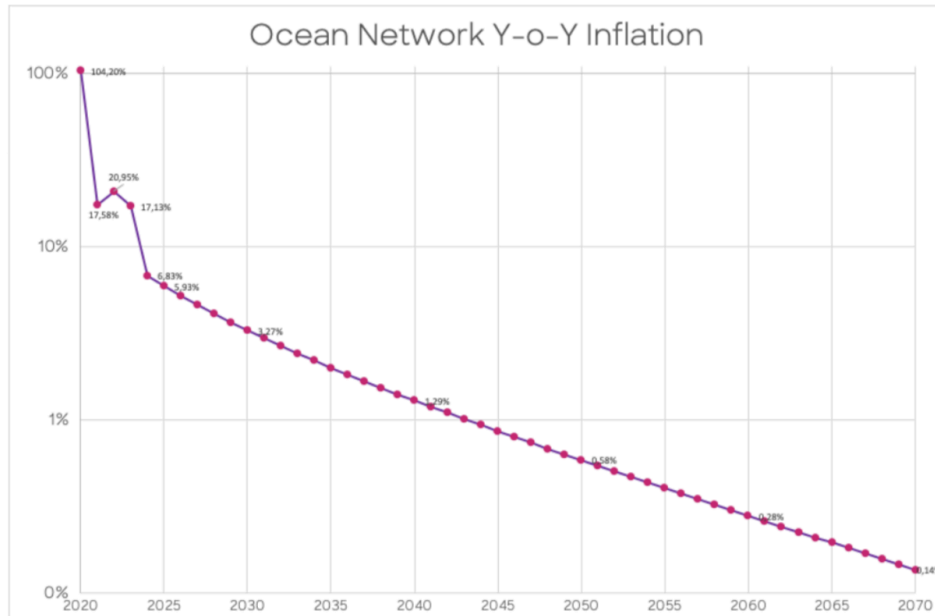


Source: Ocean Protocol

The second table (network rewards distribution) also shows that half of the network rewards will be released 5 years after they begin being released, at which point the release will slow significantly.



## Ocean Network Year over Year Inflation Schedule



Source: Ocean Protocol

The high release and unlock rates in \$OCEAN tokens results in inflationary mechanics for at least the first few years, after which the tokenomics become deflationary when the supply injections slow and are matched by token burns.

- 5% of all network revenues are burned
- All non-allocated funds during the DAO grants voting are burned

## A Rising Tide of Use Cases

There are a few different ways in which the Ocean Protocol is currently being utilized.

### Building On The Ocean

As the project is open-source, forks may be made to use the proprietary technology without using the Ocean market.



## Projects Leveraging Ocean Protocol's Technology

Project	Leveraging Ocean Protocol
<a href="#">Big Data Protocol</a>	The big data marketplace is a fork of Ocean Protocol which allows for the same functionality as the Ocean marketplace. In this use case, a percentage of transaction fees accrued on the Big Data platform are sent to the Ocean network.
<a href="#">WiSeKey</a>	WiseKey is a publicly listed company that has collaborated with Ocean to create a decentralized marketplace called WISEART for physical art and NFTs. WISEKey has developed chips to link physical objects with NFTs - the aim is to use this alongside Ocean's decentralized marketplace technology for p2p trading of these items.

Source: Messari

## Ocean Protocol as a Service

Ocean Protocol is also being utilized as a service by other multinational corporations that require the ability to confidentially share or distribute information. Other protocols aim to leverage Ocean protocol's compute-to-data functionality for their own uses or for their customers.



## Projects Using Ocean Protocol as a Service

Project	Ocean Protocol as a Service
Deutsche Bundesbank	The Central Bank of Germany has enlisted Ocean Protocol in order to develop a "decentralized network approach, which allows participants a permanent, symmetrical data access for the realization of real-time updates of statistics while preserving data sovereignty."
Daimler AG	The multinational manufacturer is creating an enterprise data marketplace using Ocean Protocol to unlock data sharing within the organization and with service partners. This is done so that data can be put to work and value can be extracted, and to make operations more efficient.
Zuelhke	Zuelhke's goal is to facilitate the creation of solutions for the health and pharma industry through data sharing. The company is leveraging Ocean technology to do so without compromising privacy or ownership, and to incentivize open sharing.
Oraichain	One of Oraichain's products is an AI marketplace where AI providers can sell their services to interested parties. Oraichain is using Ocean's compute-to-data function in their AI marketplace to increase security and functionality of this product. In addition, Oraichain publishes data generated by their AI services to the Ocean marketplace to monetize it.

Source: Messari

## Ocean Market: Monetization of Data

Various projects and companies are using Ocean Protocol for the monetization of data which as discussed earlier can be shared confidentially with third parties.



### Projects Using Ocean Protocol For Data Monetization

Project	Monetization of Data
DOVU	DOVU is a platform which allows people to get rewarded for offsetting carbon emissions. Using this platform, farmers measure how much carbon is present in their soil and sell it on the DOVU marketplace. The farmers now monetize the data itself by selling it on the Ocean marketplace to researchers and other interested parties, generating extra income.
Swash	A service that collects and bundles data generated by individual consumers whilst browsing the internet and sells it on the marketplace, with revenues being distributed among users.
Roche Diagnostics	Using Ocean Protocol, Roche aims to unlock data which is traditionally locked up and unavailable in the health industry due to privacy concerns. This will mean that data can be shared from patients' monitoring devices with professionals without compromising security or ownership of this data. This means that responses to critical conditions can be carried out faster, and insights can be achieved by leveraging the vast amount of untapped data.

Source: Messari

As the network grows, the demand for the OCEAN token will also grow for various reasons including:

- More tokens will be used for providing liquidity for different datatoken markets. Also, when tokens are being used for this purpose, they are effectively taken out of the circulating supply
- More data will be for sale, which means more OCEAN will be in demand to carry out datatoken transactions and buy/access data
- There will be more value in the protocol, and therefore parties will be more invested in its continued existence. The more invested a party is, the more they are incentivized to participate in governance and do what they can to make sure optimal decisions are made - tokens are required to participate in this governance

## Uniting the Seven Seas

As the world progresses towards a multichain future, Ocean protocol possesses the potential to be a key piece of middleware. Currently, Ocean Protocol runs on Ethereum but is currently operating as a side-chain with a Proof-of-Authority consensus mechanism in order to reduce gas fees. Ocean will be migrating to a Proof-of-Stake consensus mechanism in October 2021.

As of June 30th 2021, Ocean is also [deployed on BSC](#). However, this has yet to see much traction with only \$481 USD in total value nearly one month after the launch (24th July 2021).

The protocol is EVM compatible and thus can operate on different blockchains if scalability demands. The Ocean marketplace is already accessible via the Polkadot network due to an integration with [Moonbeam](#) – a parachain that facilitates easy re-deployment of smart contracts.

The protocol has also integrated other tech infrastructure to improve functionality. This includes Balancer - which allows the market to operate in a decentralized fashion, and [IPFS](#) to allow data sellers to store their data in a non-custodial way.

Another example of an important integration is the use of [Secret Network](#), which enables data sellers to share data without revealing the decrypted data URL as would normally happen using Ocean Protocol middleware. This increases convenience and allows stakeholders to cooperate in a completely trustless way.

The [roadmap](#) shows notable upgrades such as integrating [Balancer V2](#), allowing stakeholders to extract more value from the protocol. Furthermore, enabling users to view data assets across multiple chains at once is planned - reducing siloing and increasing accessibility.

The ability to operate cross-chain and the added functionality positions Ocean Protocol well to capture value as more and more data is generated, and owners are given greater incentives to share it.

## The Blue & Red Ocean Market

With data set to be the most important asset class of the next decade, there is no shortage of companies competing to capture this valuable market.

### Non-crypto Data Markets

#### [Quandl](#)

Quandl is a platform developed by NASDAQ which serves the financial market: providing data to investment and trading professionals. It allows third parties to sell their data and monetize it, but only if it meets certain requirements.

#### [Data & Sons](#)

Data marketplace where parties can buy data or monetize their own data by selling it. The company retains custody over all data sets and charges a 30% fee on all sales.

#### [Nokia Data Marketplace](#)

A private, permissioned blockchain launched by Nokia to facilitate the secure sharing of data and federated machine learning programs,

### Crypto Data Markets

#### [IOTA data marketplace](#)

Although not a blockchain project, it is also decentralized and crypto-powered. It has a similar value proposition in that it offers a p2p/b2b data marketplace. This project already has several partnerships with big-name enterprises and a strong working product.

#### [District0x](#)

District0x provides infrastructure for parties to build decentralized marketplaces and/or communities. Some existing examples include a decentralized ENS (Ethereum Name Service) website domains market and a decentralized p2p memes market, where people can create, sell and share memes. This provides the groundwork for another party to build a data marketplace using the District0x infrastructure.

#### [Streamr](#)

A very similar value proposition: a decentralized marketplace for live data streams where the means of exchange is also done via a native ERC token; [DATA](#). Actors have the ability to tokenize data and offer it for sale on the platform. The main point of differentiation is that Ocean is currently positioned more towards the AI and machine learning aspect of the data ecosystem.

## Smooth Sailing or Choppy Waters: Ocean's Strengths and Risks

### Industry Partnerships

The founders are [Trent McConaghy](#) and [Bruce Pon](#), and have been working on Ocean since 2017 and have been able to secure some partnerships over the past four years. Some of these partnerships include large institutions like the [Web3 Foundation](#) and the [International Telecommunication Union \(ITU\)](#), a branch of the UN, that have helped the protocol establish a level of trust within the data market ecosystem.

### Compute to Data

While we already mention the importance of compute-to-data, it's important to grasp the significance of this primitive. Compute-to-data - this function provides a competitive edge that no other protocols or organizations can boast.

The market for this kind of function is untapped and is increasing in both importance and size. Using private data to train and develop machine learning models is a service that only Ocean currently provides. It unlocks a plethora of data for researchers, developers, and engineers that were previously inaccessible.

### Ocean Community

While most protocols boast a vibrant community of developers, Ocean Protocol has prioritized developer tooling for data scientists that have culminated into Ocean's core offerings today. Importantly, Ocean has consistently executed on their roadmap over the past several years. Notably, the Ocean Protocol experienced an exploit during the [KuCoin theft](#) which resulted in a [hard fork](#) in order to redistribute the stolen tokens back to their original owners.

### Data-Token Composability

Finally, and most importantly, data-tokens are the paradigmatic shift that really sets Ocean apart from any other data tokenization service or data marketplace. The permissionless composability of data-tokens with DeFi allows for the use of AMMs in the market and automatic price discovery of data. New mechanisms of yield are possible including stakeholders curating (staking) tokens to pools and generating returns.

As data-tokens - or any version of tokenized data - gains traction, it enables the classification of data as a formal asset. This means companies can list it on their balance sheets and extract greater value from dormant data. Although Ocean has some strong differentiators, it has some potential threats as well.

### Vampire Attacks

The Ocean Protocol is open-source, and can thus be copied and tweaked to create a similar protocol. In the event of this happening successfully, market share could be lost. While

Big Data Protocol allocated a percentage of fees to Ocean Protocol, not all subsequent forks may be as generous.

### **The FAANG Guard: Existing Data Hegemonies**

Data is already collected and hoarded by big tech, who will be understandably reluctant to relinquish any of the control they have over externally generated data. The emergence of Ocean Protocol does not help parties achieve data sovereignty. It means that organizations collecting data produced are able to monetize it.

Ocean does not help prove legitimacy of ownership of data already generated, it merely allows for greater ownership of data once it is already in a party's hands.

Although distrust has been sown in big tech companies, it is not yet at the point where individuals and organizations will lobby for the deletion of data that centralized authorities possess. Therefore, it is important to consider whether the decentralization of the Ocean marketplace brings real business advantages. It will appeal to crypto and blockchain enthusiasts, although the majority is currently content with trusting their data to legitimate third parties.

### **Regulatory Risk**

The legal and accounting infrastructure around data as an asset is not ambiguous - it is nonexistent. Since data-tokens essentially act as securities within the Ocean ecosystem, the SEC and other regulatory bodies may deem them as such. There is a lack of clarity around this at this stage and remains to be seen how data-tokens will be classified, and what effect this may have on the marketplace.

## **Conclusion**

Data is a precious resource that is being produced in abundance. Ocean Protocol is a tool that can be used to unlock the value of this resource and encourage a wider distribution and financial distribution that results from this asset class. sharing for great mutual gain.

As data is becoming increasingly valuable, other data marketplaces may launch to capitalize on the untapped value. It remains to be seen whether Ocean Protocol's current advantages will make it the market of choice where parties go to monetize this asset or whether a new model will emerge.

The importance of data goes beyond enabling broader collaboration and sharing in a trustless manner. It is only through the monetization and tokenization of data (datasets, machine learning algorithms, etc) that a new asset can be created. The tools and primitives created by DeFi, Ocean Protocol, and other cryptonetworks are setting the stage for a new, decentralized data economy.

*Special thanks to Mason Nystrom for his feedback and valuable input.*