



**BARREL**

Unlock the World's Data For AI

# A Next-Generation Smart Data Type and Decentralized Data Platform and Exchange

Vitalik Buterin development of Ethereum in 2015 focused on an important part of the Bitcoin experiment - the underlying blockchain technology as a tool of distributed consensus. Ethereum provides a blockchain with a built-in Turing-complete programming language that is used to create "smart contract" to build many of the distributed consensus systems Vitalik discussed in his white paper.

However, another complimenting component of the Bitcoin experiment, and missing from Vitaik's blockchain is a formal process for utilizing data. When Marc Andreessen writes about 'software' in Why Software Is Eating, software means 'data and algorithms' are eating the world. Data, is the medium of exchange - the currency.

As the word Software is increasingly replaced by AI. Barrel intends to provide a decentralized blockchain with built-in GDPR compliant protocol that can be used to create "smart data types" that can be used to encode various types of compliant data structures, to feed an AI driven world.

Commonly cited challenges in unlocking the world's data: Identifying and sourcing alternative datasets often buried behind corporate walls, ensuring the privacy and compliance of datasets, leveraging consistent structures and the creation of standards in highly fragmented vendors, rewarding the end user responsible for creating, often unknowingly, the data being sold, as well as more complex definitions involving scoring of the datasets - like backtesting data - business intelligence tools, data cleansing protocols or even blockchain-based decentralized data exchanges.

The Barrel Protocol provides for the creation of new Smart Data Types allowing users to harness the power of distributed consensus to solve any of the problems described above, as well as many others that we have not yet imagined, simply by writing up the logic in a few lines of code.

By Jonathan Meiri 2018 v2.8

**Key changes include**

**Implementation Details of Protocol Alpha Phase - Distributed Data Verification**

# The Principles of Barrel

**1/ Built-In Privacy, Security, Compliance**

**2/ Smart Data Tokens**

**3/ Consistent structure, and quality assurance (optional)**

**4/ Pegged Pricing through Smart Contract**

## Abstract

Abstract: Barrel facilitates the transaction between buyers and sellers of massive datasets. Barrel provides a protocol for creation of new Smart Data Types and Smart Contracts. Companies, apps and various services will be able to take part in the value generated by their data and can pass that value to their customers.

Most important Barrel intends to provide a blockchain with built-in Privacy, Security and Compliance measures. These features are intended to ensure the anonymity of the data, protect it from hacking or breakage, and eliminate the high cost associated with the existing compliance process.

Built-in Privacy by Design is an approach to systems engineering which takes privacy into account throughout the whole engineering process. Barrel Protocol extends this notion to include security, as well as compliance of the data.

## Barrel Network Tokens

Data is the currency is not a new idea. Almost a decade ago, Clive Humby made headlines when he declared that "data was the new oil but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down and analyzed for it to have value."

This network of democratized Smart Data Types that will have pegged (and in some instances continuous) liquidity. Oil is a great metaphor, in its rawest form, raw data and crude oil are almost useless. But when they are refined they can be turned into something much more valuable. The Barrel Protocol intends to build a protocol for data "refineries" these are the Barrel Network Tokens, and become a market maker for their valuation.

## Examples of Smart Data Types

- Anonymized Credit Card Transactions
- Web-browsing history
- Healthcare profile
- In-store Foot Traffic
- Maritime Ship Location
- Connected Car Data
- Audience Data used by Data Management Platforms (DMPs)

Each of these datasets is bought and sold millions of times a day to the tune of \$300B dollars annually and that number is expected to increase 100 fold as more organizations increase their usage of data. Each transaction undergoes scrutiny to ensure the data is collected, stored and handled appropriately. Further concerns regarding, structure, quality and compliance have created a mountain of friction and scores of middlemen that increase the cost and reduce liquidity of the assets. Finally many corporations are increasingly aware of the value of their data - and are looking for secure ways to monetize it, while keep the regulators at bay. Barrel can solve that.

## Proof of Work / Proof of Anonymization

### **Proof of Work - definition**

A proof of work is a piece of data which is difficult (costly, time-consuming) to produce but easy for others to verify and which satisfies certain requirements. It is one of two Blockchain systems (the other is Proof of Stake) that govern how transactions are verified on the decentralized network.

### **Proof of Anonymization**

Barrel uses a Proof of Anonymization model. Refiners are empowered to anonymize blocks of data based on agreed upon protocols, and nodes can then verify the anonymity and seal the block. Similarly to Bitcoin hashing, anonymizing the data is hard while verifying it is easy.

Each data type will have a different anonymization protocol. Multiple miners - Refiners - can fight to approve a certain block of data and will be incentivized to do so.

## Barrel Smart Data Types - The Data Token

The BRN is built on top of NEO and will hold a single reserve in NEO. Companies, apps various services can easily create new Data Tokens, and use BRN as their reserve. They can connect to

the BRN network, essentially making their dataset exchangeable or keep it within a Consortium blockchain.

The goal is to form a global, decentralized, highly liquid exchange with various types of company-generated data tokens (BRNs) and token changers. The data itself and its transmission live off chain.

The BRN establishes network dynamics where increased demand for any of the network's data tokens increases demand for the common BRN, benefiting all other data tokens holding it in reserve.

The BRN will be sold in a fundraiser scheduled for June 4th, 2018. Additional details will be announced in mid April.

## Token Economics

The token, BRN, will be used to exchange data on the network. Users earn tokens when they sell data, refineries earn tokens for refining data, and developers earn tokens when they provide smart contracts that service the network.

Applications that represent multiple users will have to hold BRN in reserve, equal to the amount of users they represent.

The consistent structure and quality assurance scoring will enable data buyers to easily merge data sets, compare datasets and create an intrinsic flight to quality, well refined datasets

The total number of tokens issued is 10B tokens. No additional tokens will be issued. Each Barrel contains 1,000,000 Drops which is the smallest unit of measurement.

## Implementation Details of Protocol Alpha Phase - Distributed Data Verification

A data provider loads into the Barrel client a CSV file containing predefined columns per data type (e.g. ride-hailing or food delivery). The table is split into chunks of size  $C \times 1$ ,  $C$  being the number of lines per chunk, e.g. 100. The columns are split in order to preserve k-anonymity of the data across verification nodes. The data provider alone holds the original ordering of the chunks.

Anonymity is preserved thanks to the combined effect of the distribution mechanism: first, the columns of a single file are distributed in such a way that N-anonymity is preserved.

Secondly, even if a rogue node was to obtain multiple columns of the same original table, it would not be able to reconstruct the file since verification nodes are left agnostic of the original ordering of the file and of which file the columns originated from. At most, a rogue node may try to construct all possible permutations of all the columns it receives from all possible original tables.

The data is distributed to the nodes using the BitTorrent protocol. Each Barrel data verification node runs a predefined PII sanitation algorithm and verifies that each chunk is PII-safe. When all nodes reach a consensus for all chunks of a table, the original table is declared safe as well. The verification nodes are credited with NEO GAS for the verification beforehand, in order to finance the verification itself, and are rewarded with BRN for the verification afterwards. A slow or rogue node may be penalised. Each verification node signs its decision using its private key for certification.

The data provider holds the SHA256 hashes of the distributed chunks, the hash of the complete file, and the original ordering of the chunks. Upon reaching consensus, a Merkle tree representing the chain from the original table to the verification decisions and signatures is persisted to the NEO blockchain. The data buyer is then able to confirm the data is PII-verified and consistent, and re-assemble it. The data can be resold without needing to repeat the process because the verification chain is hashed onto the Blockchain.

## Token Allocation

- 50% Fundraiser
- 20% Long-term Foundation Budget
- 15% Community Grants, Partnerships and Bounties
- 15% Team, Advisors, Early Investors

## Use of Proceeds

- 40% Software Development
- 20% Barrel Reserve
- 12% Marketing and Business Development
- 10% Seeding Refineries and Smart Contract Development
- 8% Operational Expense
- 5% Legal Expenses
- 5% Misc and Unexpected

## Team Vesting

In order to ensure the long term alignment of founders, team members and token holders, all founders and team members will have a 3 years vesting schedule with a 6 month cliff.

## Initial Data Types and Datasets Available at Launch

Data Type	Dataset	Description	Frequency	Geography
eReceipts	Online and In-app Purchases	Item-level detail about online and mobile application purchases from 10MM anonymized consumers	Weekly	Worldwide
eReceipts	Ride-hailing and Food Delivery	Item-level detail about online and mobile application purchases from 10MM anonymized consumers	Weekly	North America, Europe, Australia, South Africa, Southeast Asia, India, Brazil, Russia
SMS Receipts	Online Purchases	Purchase confirmation from 5MM anonymized consumers	Daily	India
SMS Receipts	Ride-hailing and Food Delivery	Purchase confirmation from 5MM anonymized consumers	Daily	India

## The Barrel Protocol Ecosystem

Different parties can take on different roles in the Barrel network ecosystem. The primary forms of participation are as follows:

- End-Users can receive, hold, transfer, request, purchase and liquidate smart data types
- Data vendors can join the public Barrel Blockchain offering data in exchange for coins. They can also join existing Consortium Blockchains or set up new Consortiums that will govern how their data is being handled
- BRN creators can launch or sponsor Smart Data Types to handle analysis, Quality Assurance, or manipulation of datasets.

- Data buyers interested in unique datasets can discover new datasets, buy specific types of Smart Data Types on the Barrel network or hold on to the underlying Barrel coin for flexibility.
- Corporations looking to understand the value of their in-house data can sponsor a creation of a Smart Data Type on the network to engage developers and potential partners. For example mobile operators, can sponsor a Smart Data Type on the Barrel Network. They can ensure the privacy and of the data and let developers innovate on top of it.

### **Scenario 1 - Barrel Smart Data Type for anonymized credit card datasets**

A credit card company launches a new Barrel Smart Data Type BRNCC, a Token that represents an anonymized standard set of US historical credit card transactions. This is not new, credit card companies have been selling their data for over 30 years. Only now there is a standard format with built in privacy and compliance measures. By introducing the new coin to its existing network of buyers it can materially reduce the cost of compliance and ensure privacy.

Now a developer can augment the Smart Data Type to include anonymized transactions to track the performance of say retail companies. Together they can now sell this data to retailers, analytics companies or hedge funds. A second developer can offer to build a credit scoring mechanism on top of the same credit card dataset. The value of BRNCC will increase as more developers build software on top of it. The credit card company will initially buy Barrel (with NEO) and in turn provide them to developers to build the software. Other credit card companies may join in and increase the value of BRNCC. As a next step an innovative credit card issuer may decide to share some of the profits with their most active and loyal users. Now for each transaction a user receives Barrel coins, like credit card points.

## **Consortium Blockchains - Definitions and Use Case**

### **Blockchains and Public Blockchain**

A Blockchain facilitates secure online transactions. It is one of the building blocks of Bitcoin and NEO and considered a major technological innovation in recent history.

A Blockchain was designed to securely cut out the middleman in any exchange of asset scenario. It does this by setting up a block of peer-to-peer transactions. Each transaction is verified and synced with every node affiliated with the blockchain before it is written to the system.

Anyone with a computer and internet connection can set up as a node that is then synced with the entire blockchain history and as such was considered public.

A public blockchain is most appropriate when a network needs to be decentralized.

### **Private Blockchain**

Barrel protocol does not support private blockchain as it essentially lets the middleman back in. A company writes and verifies each transaction. Though it does not offer the same decentralized security as its public counterpart, trusting a business to run a blockchain is no more dangerous than trusting it to run a company without blockchain. The company can also choose who has read access to their Blockchain transactions, allowing for greater privacy than a public blockchain.

### **Consortium Blockchain**

Consortium blockchain is partly private. There has been some confusion between Consortium and purely Private Blockchains. Vitalik Buterin, co-founder of Ethereum provides a definition:

“So far there has been little emphasis on the distinction between consortium blockchains and fully private blockchains, although it is important: the former provides a hybrid between the ‘low-trust’ provided by public blockchains and the ‘single highly-trusted entity’ model of private blockchains, whereas the latter can be more accurately described as a traditional centralized system with a degree of cryptographic auditability attached.”[2]

Instead of allowing any person with an internet connection to participate in the verification of transactions process or allowing only one company to have full control, a few selected nodes are predetermined.

For example, say a consortium of hospitals need to collect CT scans from a number of independent clinics, they can set up a consortium where a trusted set of hospitals serve as nodes and confirm the secure and private transfer of the CT scan.

Similarly hedge funds can ensure that data collected from a set of data providers is compliant without having to go through the expensive compliance process themselves or expressly state that they are using the data.

Barrel protocol, and Consortium Blockchains ensure that new Data Tokens have the specific privacy, security and compliance measures for handling the data.

Consortium blockchain platforms have many of the same advantages of a private blockchain, but operate under the leadership of a group instead of a single entity.

Think of it as subject matter experts, a trusted council that are known entities within their industries and can decide who has access to the blockchain ledger.

Council members can chose to include quality assurance measures, so that datasets within network can be easily priced.

The possibilities are endless.

## Appendix

### Framework for checking Barrel Protocol Validity

Is this doing something useful  
Is this the right technology solution  
Is it being adopted  
Is the infrastructure there

### References:

1/ Privacy-Preserving Computation, Florian Kerschbaum,  
<http://www.fkerschbaum.org/apf12.pdf>

2/ Decentralizing Privacy: Using Blockchain to Protect Personal Data, Guy Zyskind, Oz Nathan,  
Alex 'Sandy' Pentland, [http://web.media.mit.edu/~guyzys/data/privacy\\_slides.pdf](http://web.media.mit.edu/~guyzys/data/privacy_slides.pdf)

### Council Member

David Baggett - Early Bitcoin Developer, founder ITA Software (sold to GOOG)  
Vish Shatry - Vice President, Visa Digital Products, former Corp Dev at PayPal  
Dr. Moshe Bareket - Chief Finance and Strategy Officer at Tshuva Group, former Chairmen The Phoenix  
Jonathan Meiri - CEO Superfly Insights

### Token Refineries

Web Scraping - James Moran - CEO Yipit Data  
Anonymized Credit Card - Michael Babineau - CEO Second Measure  
Anonymized Location Data - Jeff Glueck - CEO Foursquare  
Anonymized Purchase receipts - Jonathan Meiri - CEO Superfly Insights

### Other

Tomer Barel - EVP, Global Chief Enterprise Services Officer at PayPal  
Yossi Barnea - CIO Delek Group (confirmed)  
Yudi Levi - CTO Bancor (Confirmed)