



Barrel Protocol White Paper

© 2018 Barrel Inc. All rights reserved.

Table of Contents

TABLE OF CONTENTS	2
THE AI REVOLUTION IS FUELED BY DATA	3
WHAT IS "ALTERNATIVE DATA"?	3
HOW IS ALTERNATIVE DATA ACCESSED?	4
PII VS. NON-PII DATA	4
EXISTING BARRIERS IN THE CURRENT DATA ECONOMY	4
BARREL PROTOCOL OVERVIEW	6
THE ECO-SYSTEM PLAYERS	7
DISTRIBUTED DATA VERIFICATION	8
THE BRN TOKEN	10
TOKEN ALLOCATION	10
USE OF PROCEEDS	10
TEAM VESTING	11
DATA TYPES	12
INITIAL DATA TYPES AND DATASETS AVAILABLE AT LAUNCH	12
ADDING NEW DATA TYPES	13

The AI Revolution is Fueled by Data

Artificial Intelligence is undoubtedly one of the most important technological revolutions that will shape business in the coming years. By using AI, companies will be able to make decisions based on an enormous amount of data, at unprecedented speeds, and unlimited scale. Today's most successful companies are the ones already using AI as part of their business processes.

However, AI in itself does not bring any value, this value is created by combining AI with data. When people think of this AI revolution they usually think of traditional corporate data that is now analysed by these AI algorithms. Although this scenario is definitely common, much of the data that fuels the AI revolution comes from what is called **Alternative Data**.

What is "Alternative Data"?

Alternative data is information gathered from non traditional data sources, which can provide unique and timely insights into industries, investment opportunities and business as well as sector analysis in general. An alternative data set can be compiled from various sources such as financial transactions, sensors, mobile devices, satellites, public records, and the internet.

Examples of alternative data sets include:

- Credit card transactions
- Email Receipts
- Point of sale transactions
- Website usage
- Geo-location
- Obscure city hall records
- Satellite images
- Social media posts
- Online browsing activity
- Shipping container receipts
- Product reviews
- Price trackers
- Flight and shipping trackers
- and more..

Alternative data is usually consumed by hedge funds, who use the data to generate insights into investment opportunities. However, today we can see more and more companies using this data for business intelligence purposes. For example, by analysing email receipts, a company can learn about special offers that are being offered by its competitors.

How is alternative sold?

Alternative data is usually sold by intermediaries and aggregators, also known as **alternative data companies**. These companies are responsible for collecting the data from various sources such as credit card companies, Telcos, application providers, etc., packaging it and selling it to interested parties, while ensuring that the data is clean from Personal Identifiable Information (PII). In many cases the alternative data companies split the revenues with the raw data providers.

Over the past decade the alternative data commerce market has become a big business. In fact, the commerce of alternative data between businesses is projected to become a \$40B business in 2018 with a 50% annual growth. However, as described below, this market still suffers from substantial legal challenges concerning privacy, which need to be addressed in order to realize its full potential.

PII vs. Non-PII Data

Data can be sensitive and regulators around the world are trying to find the balance between privacy protection and the importance of data for businesses and government. One of the main initiatives to protect privacy is **GDPR (General Data Protection Regulation)**, a European legislation, which aims to give users control over their personal data, by unifying the regulation within the EU. GDPR and many other regulations make it difficult and expensive for companies to collect, store, and especially trade data.

In general terms, privacy laws' main concerns revolve around **PII** (Personally Identifiable Information). This is information that can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context. Although data that includes PII is very valuable, **alternative data is non-PII data** and can be also extremely valuable in many cases. For example, this data can be used to monitor special offers, find investment or expansion opportunities, and more.

So does this mean that non-PII data can be freely traded between companies and sold in masses to anyone who wants it? Well, this is not the case yet.

Existing barriers in the current data economy

In reality, data-related deals between companies are complex, and can take months—sometimes even years— to be realized. Both sides of the deal need to invest enormous amount of effort in ensuring:

- **Clean Data** – that the data does not include PII. If the data does include PII, which is very common, the alternative data company needs to clean this data from any PII. This process is extremely complex and lengthy. Any mistake can end-up costing a fortune.
- **Compliant collection** – that the data has been collected in a manner that is compliant with laws and regulations. This includes writing a "Terms of Service" that covers this type of activity, the users' consent during data collection, etc.
- **Security** – that the data remains secure both in transit (when data is moved around between the different parties) and at rest (when data is stored).

- **Data Integration** – that the data schemas and formats conform to the systems used by the data buyer. Very often this required complex IT projects, involving data transformation.

The result of these barriers is that ***the data economy is not yet widely available.*** Today, every potential data exchange requires extensive legal due diligence, data preparation, integration, and so on. In many cases this may take as much as a year or more.

Barrel's mission is to build an eco-system that will ensure and certify that data is clean and compliant with laws and regulations, thereby removing the barriers to a free-flowing data economy.

Barrel Protocol Overview

Barrel protocol aims to become the infrastructure that will facilitate a free-flowing and compliant data economy. Using the blockchain and smart contracts, Barrel automatically verifies that data is completely clean from any PII, ensuring data compliance, privacy, and security. After the data has been verified, it is sealed and certified by the Barrel Protocol.

In addition, through consensus-driven approach, the Barrel protocol creates regulatory-compliant data exchange standards, including definition of data schemas and formats for every participating data type.

The protocol focuses on the verification and the certification of the data, not the cleaning of the data or the trading of the data between alternative data companies and their clients. The protocol in itself is the foundation upon which developers can build specific data marketplaces, data exchanges, and through which alternative data companies can prove to their clients that their data is clean from any PII, using objective and unbiased technology to prove their claim.

The Eco-system Players

The Barrel protocol empowers an eco-system for the distribution of privacy-compliant data. This eco-system includes the following players:

- **Data Providers** – data providers are the suppliers of the data. These data providers can be a subsidiary of a credit card company that anonymized credit card transaction data . Data providers need to comply with privacy regulations regarding the collection of the data. For example, getting users consent on the relevant "Terms of Service" agreement, etc.
- **Data Resellers/Refineries** – these are today's alternative data companies. These companies aggregate data from Data Providers and resell this data to their clients. They are responsible for cleaning the data. The data can be cleaned manually or using sophisticated technology. As the party who is responsible for cleaning the data, they need to prove that the data is clean. Today, this requires long and tedious due diligence on their cleaning technology and practices, long legal on-boarding processes to establish the their legal liability, etc. Through the Barrel protocol this process can be streamlined or eliminated completely.
- **Validators** – validators offer computing power to run Barrel Protocol's verification software. If their hardware complies with the Barrel Protocol's hardware requirements, they simply download the Barrel Protocol software, and their node will automatically participate in the pool of validator resources. Their computing power will be used to verify and seal blocks of data. Validators will discard blocks of data that include PII.
- **Data Consumers** - data consumers are the buyer of the data. For example, hedge funds, retailers, etc.. As the users of the data, data buyers are usually more exposed to legal risks. They need to ensure that the data being purchased is completely clean from any PII and that the data providers, from whom they purchased the data, comply with privacy regulations concerning the data collection, storage and distribution. Using the Barrel Protocol they can ensure data is clean.
- **Barrel Protocol Developers** – The Barrel Protocol will be constantly developed by members of its community. These developers can be from any one of the eco-system members, described above, and/or data enthusiast wishing to participate in the advancement of the protocol, to ensure safe and compliant use of data.
- **Entrepreneurs/App Developers** – entrepreneurs and app developers will offer applications that use the clean data produced by the Barrel Protocol to generate insights based on non-PII data.

Distributed Data Verification

The Barrel Protocol technology is a network of data nodes that use the Barrel software to automatically verify and certify that the data is clean, all without the ability to "steal" or reproduce the data in any way. Below describes the process from the loading of the data by the Data Reseller/Refinery to the re-selling of the certified data to the data consumers:

1. **Loading the data to the network** - Using the Barrel Protocol client software, data is loaded into the network. The data contains predefined columns per data type (e.g. ride-hailing or food delivery).
2. **Breaking the data into chunks** - the data 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. Only the originator of the source, for example the Data Reseller/Refinery holds the original data and the order of the chunks. Since the breaking of the data occurs locally within the client, this data will never leave the premises of the Data Reseller/Refinery. The columns of a single file are distributed in such a way that N -anonymity is preserved. 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.
3. **Distributing the data to the nodes** - 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.
4. **Certifying a table to be clean** - when all nodes reach a consensus for all chunks of a table, the table is declared clean as well. The verification nodes are rewarded with BRN for the completed verification. A slow or rogue node may be penalized. Each verification node signs its decision using its private key for certification.
5. **Awarding the certification** - the Data Reseller/Refinery receives the SHA256 hashes of the distributed chunks and the hash of the complete table. Upon reaching consensus, a Merkle tree representing the chain from the original table to the verification decisions and signatures is persisted to the blockchain.
6. **Reselling the certified data** - when the Data Reseller/Refinery sells the data, the Data Consumer is able to confirm the data is PII-verified and consistent. The data can be resold without needing to repeat the process because the verification chain is hashed onto the Blockchain.

Throughout this entire process the original full data is maintained at the source (i.e. Data Reseller/Refinery). The network only holds fragments/chunks of the data, which cannot be reproduced in any manner. When selling the data, the Data Reseller/Refinery can simply point to the data hash in the blockchain to prove that it is verified, certified and clean from any PII.

Token Economics

At the heart of the Barrel Protocol economy is the BRN token. BRN token is used to pay for data validation and may be used to pay for the data itself. The total number of tokens issued is 10B tokens. The price for data validation will vary between data types based on demand and the required effort needed to validate the data type. Since the number of BRN token is limited, its value will be influenced by the demand for the service. So the BRN token establishes a network dynamic, where increased demand for any data type increases value for data tokens across the entire network.

BRN Token Usage

The following network players will use the BRN tokens as described below:

- **Data Refinery - Data Providers:** Although it is not mandatory, in some cases, such as a 'revenue share' model between Data Resellers/Refineries and Data Providers, the payment for the data to the Data Provider may be using BRN tokens.
- **Data Refinery - Data Validators:** The Data Reseller/Refinery pays the Data Validators for the verification and certification of the data using BRN tokens. The price is dependent on the data type. The payment to the Validators is managed by the Barrel network and the payout is performed upon completion of the verification process based on a smart contract.
- **Data Refinery - Data Consumers:** Data Consumers can use BRN tokens to pay for data, but they are not obligated to do so.
- **Barrel Protocol Developers** - we haven't discussed this.
- **Entrepreneurs/App Developers** - we haven't discussed this.

Collateral Requirement

To ensure the integrity of the data validation process, Validators will be asked to put up a certain amount of BRN tokens as collateral. This collateral will ensure that the verification process is legitimate.

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 2 years vesting schedule with a 6 month cliff.

Data Types

The Barrel Protocol will support a variety of data types and additional data types will be constantly added over time. Every data type requires specialized development. This development will be performed by Barrel at the beginning, but should be expanded to include Barrel Protocol Developers at a later stage.

Initial Data Types and Datasets Available at Launch

Data Type	Data Set	Description	Frequency	Geography
e-Receipts	Online and In-app Purchases.	Item-level detail about online and mobile application purchases from 10MM anonymized consumers	Weekly	Worldwide
e-Receipts	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

Adding new data types

The Barrel Protocol will be driven by a consortium of BRN token holders. The consortium will have a Council that will decide on the road map for development, specifically the addition of new data types. The council will meet every quarter to discuss new data type requests/submission from a list of data types that will be prepared in advance. Data types

can be suggested by non-token holders and there is no requirement for new data types to be sponsored.

Vision for the Future

We are at the very beginning of the data revolution. At this stage, the world is just beginning to express its privacy and security concerns. And regulators are responding with attempts to build guidelines that will find a balance between privacy and business agility. But once we get over these barriers, hopefully with the help of the Barrel Protocol, we can move to the next phase.

In this new phase, alternative data will be freely sold without any concern for privacy or regulatory requirements as it will be validated and certified by the Barrel network. This means that an endless amount of clean and compliant data will be widely available to a global community of companies, governments, researchers, and developers, who will use software and AI to find insights and unleash the true potential of this amazing data revolution, which many believe, can be as impactful as the industrial revolution.