



The State Of Digital Asset Data And Infrastructure: 2023 Edition

Introduction

Fourteen years after the emergence of Bitcoin, thousands of digital assets spanning layer-1 native tokens, stablecoins, and non-fungible tokens (NFTs) have emerged. Digital assets are employed to facilitate a variety of use cases such as payments, real-world asset tokenization, and borrowing and lending. In conjunction with this proliferation of use cases, these assets are traded on dozens of centralized and decentralized exchanges in staggering volumes compared to five years ago.

Naturally, this evolution renders navigating and operating within the digital asset landscape more complex than ever.

The Block Research's Digital Asset Data and Infrastructure report analyzes two pillars of the emerging digital asset economy – data and infrastructure – to shed light on how firms are making the increasingly intricate digital asset economy navigable. It aims to answer the two following questions: (i) How do data providers help industry participants extract value from a seemingly endless deluge of digital asset data?, and (ii) How do infrastructure providers make operating blockchain networks and accessing the applications built on top of them feasible for everyday users?

The report builds on The Block Research's [2021 Data and Infrastructure Report](#) and is structured in four parts:

- **Part 1** provides an overview of the different components of digital asset data and infrastructure. It segments their respective landscapes into several subcategories and highlights key considerations and developments relevant to firms and customers operating therein.
- **Part 2** outlines the landscape of firms focused on providing digital asset data. It identifies the different classes of data that providers specialize in and compares the products and services of leading firms.
- **Part 3** sheds light on the landscape of firms that operate core blockchain infrastructure. It identifies which components of the “web3 stack” they power and compares the products and services of leading firms.
- **Part 4** provides closing thoughts.

Table of Contents

Introduction	2
Part 1: Overview of Digital Asset Data and Infrastructure	7
Digital Asset Data	9
The Role of Data Providers in the Digital Asset Data Lifecycle	9
Market Data	11
On-chain Data	13
Forensics and Market Surveillance Data	18
Digital Asset Infrastructure	22
Staking-as-a-service	23
Node-as-a-service	26
Wallets	28
Part 2: Digital Asset Data Providers	31
Market Data Providers	33
On-chain Data Providers	37
Forensics and Market Surveillance Providers	45
Digital Asset Data Fundraising and M&A	47
Part 3: Digital Asset Infrastructure Providers	51
Staking-as-a-service Providers	53
Node-as-a-service Providers	56
Wallet Providers	59
Digital Asset Infrastructure Fundraising and M&A	61
Part 4: Conclusion	65
Appendices	68

Sponsored By



Amberdata is the leading provider of digital asset data. Our institutional grade infrastructure eliminates the integration challenges and maintenance headaches to access digital asset data, reducing cost and time to market. We deliver comprehensive data and insights into blockchain networks, crypto markets, and decentralized finance, empowering financial institutions with data for research, trading, risk, analytics, reporting, and compliance. Amberdata is the only provider offering historical and real-time fundamental (on-chain) and market data, eliminating the need to source and integrate data from multiple vendors. Our deep expertise makes Amberdata the trusted provider to leading financial institutions entering the asset class.

For more information about Amberdata, visit amberdata.io

Researched By



[The Block Pro](#) is The Block's premium product portfolio designed to help institutions evaluate opportunities in digital assets. Pro's research, news, and data products are powered by teams of subject matter experts deeply entrenched in the digital asset ecosystem who deliver actionable intelligence so businesses can make informed decisions.

[The Block Research](#) produces research content covering the digital assets, fintech, and financial services industries.

Contact

Email: research@theblockcrypto.com

Twitter: [@theblockres](https://twitter.com/theblockres)

To avoid perceived conflicts of interest, The Block is not included in the sample set of data providers analyzed in this report. To learn more about The Block Pro Data, please refer to www.theblock.pro.

Acknowledgments

We would like to thank Amberdata for sponsoring this research report and making its production possible. We are also grateful to those who shared their perspectives and were interviewed for this report, including individuals across the following organizations:

- **Amberdata:** Shawn Douglas, Tongtong Gong, Steve Henning, Lisa Hadrack
- **CryptoCompare:** Charles Hayter
- **Dune:** Hannah Smith
- **Kaiko:** Clara Medalie, Bediss Cherif, Louis Ramat
- **Lukka:** John Pagano
- **Nasdaq:** Michael Karbouris, Tony Sio, Thomas Neill, Rob Norris, Bill Dague, Mitchell Tessier
- **Quicknode:** Dmitry Shklovsky

We would also like to thank everyone at The Block who assisted with this report - research team: Andrew Cahill, Dipankar Dutta & Hiroki Kotabe; data team: Simon Cousaert, Edvin Memet, and Rebecca Stevens; funding team: John Dantoni, Atharv Deshpande, Edvinas Rupkus, and Jae Oh Song; design team: Zoe Ellyse Del Rosario.

Finally, while this report aims to provide a comprehensive overview of the data and infrastructure landscape, it does not provide an exhaustive analysis of all firms providing data and infrastructure services. If you believe that your company or project was miscategorized or would like your project to be considered for next year's report, please contact research@theblockcrypto.com.

Authors

Marcel Bluhm, Research Analyst – Reports

[Twitter](#)

[Linkedin](#)

Saurabh Deshpande, Research Analyst – Reports

[Twitter](#)

[Linkedin](#)

Jason Michelson, Research Analyst – Reports

[Twitter](#)

[Linkedin](#)

02/16/23

The State Of Digital Asset Data And Infrastructure:
2023 Edition

1

Overview Of Digital Asset Data And Infrastructure

Part 1: Overview of Digital Asset Data and Infrastructure

Dozens of layer-1 and layer-2 platforms have expanded the breadth of the blockchain landscape. Digital asset activity spans a wide range of disciplines, including payments, decentralized finance (DeFi), and NFT issuance and trading. Platforms enabling these use cases spring forth from a number of development environments, each with different methods for deploying applications and storing their data.

This Cambrian explosion of on-chain activity adds significant complexity to digital asset data and infrastructure. It also presents opportunities. A comprehensive taxonomy of blockchain data fields is being created in real-time by digital asset data firms as composable, and increasingly complex digital protocols continue to emerge. The demystification of the vast sea of digital asset data is unlocking value for the investment community, regulatory bodies, and financial institutions alike.

Likewise, infrastructure-as-a-service firms enable market participants to overcome the high technical and financial barriers associated with operating blockchain networks. Although different blockchains have different infrastructure requirements, they all require around-the-clock operation as digital assets do not have an ‘off switch’. Powering these blockchain platforms and the applications built on top of them requires technical expertise and enterprise-grade risk management processes. Digital asset infrastructure providers that offload these tasks from builders and users have emerged as critical – though sometimes overlooked – service providers.

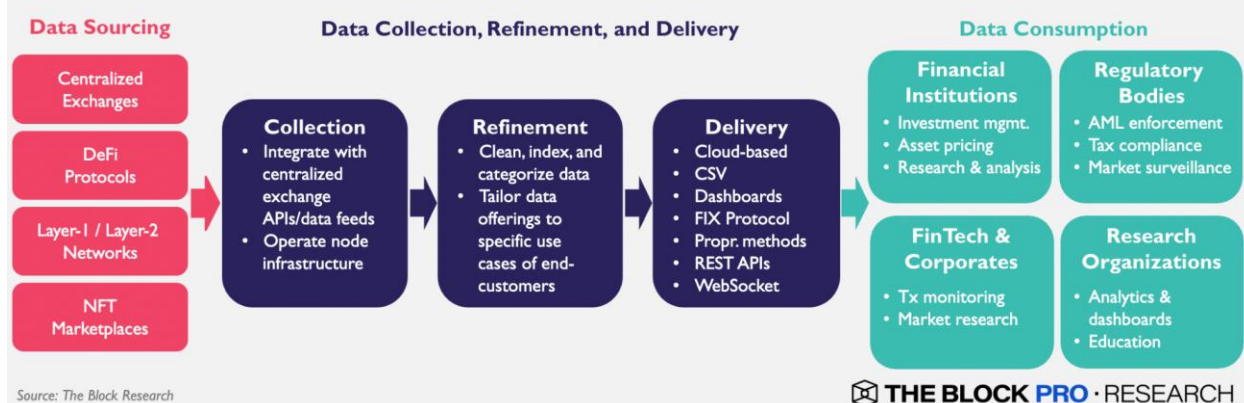
The following section of this report provides an overview of the market segments that compose the digital asset data and infrastructure landscape. It also highlights key developments and considerations for firms that conduct business within their respective verticals.

Digital Asset Data

The Role of Data Providers in the Digital Asset Data Lifecycle

Data providers play a crucial role in the value chain of digital asset data by acting as a bridge between (i) the raw data generated by digital asset trading venues and other on-chain protocols, and (ii) data consumers that employ digital asset data for a range of use cases.

The Digital Asset Data Lifecycle



At the highest level, data providers start with harvesting raw data from off-chain and on-chain sources spanning centralized digital asset exchanges (CEXes), DeFi protocols, layer-1 and layer-2 blockchain networks, and NFT marketplaces. For example, obtaining data from CEXes is relatively straightforward via established interfaces.¹ In contrast, extracting data from multiple decentralized exchanges (DEXes) is resource-intensive and requires data firms to run (or obtain access to) blockchain node infrastructure.

This data is then cleaned, refined, and standardized, tailoring it to end-users' requirements. For example, since automated market maker (AMM) DEXes bear no resemblance to traditional order book matching exchanges, data providers play a key role in normalizing DEX data such that it can be analyzed in a similar format as CEX data.

Finally, the data is delivered to end-consumers via a variety of methods suitable for specific data formats and customers. End-users ranging from financial institutions to research organizations consume the data for a variety of use cases. For example, financial institutions require high-quality data feeds to make investment decisions and manage their portfolios. Their sophisticated models require real-time and historical tick-level trade data² as well as order book data as inputs. This allows them to analyze metrics such as liquidity across CEXes and DEXes, and dynamics between different futures and spot markets. Further use cases for digital asset data end users are explored throughout the following parts of this report.

¹ Data can be accessed, for example, via an application programming interface. See further below for an outline of the most common data delivery methods.

² Throughout this report, 'trade data' or 'trading data' refers to the data related to trades taking place on exchanges.

What are Key Methods/Formats for Digital Asset Data Delivery?

Different users and their use cases may require different means to access data. For example, while sophisticated investors may require highly granular historical and real-time data (e.g. order book data) via application programming interfaces (APIs), forensic analysts may prefer dashboard solutions that visualize, for example, the provenance of a user's funds. The most relevant delivery methods for digital asset data consist of:

- **Cloud-Based Delivery:** Storing historical order book data can require petabytes of storage space, thus resulting in onerous infrastructure requirements for end users. This necessitates data delivery (into cloud buckets set up by customers) via cloud solutions such as Amazon S3 (Simple Storage Service), Azure Cloud, or Alibaba Cloud.
- **Comma Separated Values (CSV):** CSV is a format that enables data storage in structured tables for use in applications such as Microsoft Excel. It is typically used for historical data and is updated daily. CSV files are easily accessed for historical analysis, but managing several files for the same data (i.e., different data vintages) is inefficient. The CSV format is also not suitable for real-time and high-frequency trade data.
- **Dashboards:** Dashboards are popular tools for visually analyzing data. One considerable drawback of dashboard-based data delivery (when the dashboards are not user-generated) is that it adds friction when users want to manipulate data for analyses. Institutions that build proprietary models and frequently backtest their investment strategies are unlikely to opt for dashboards as the only form of data delivery.
- **FIX Protocol:** The Financial Information eXchange (FIX) Protocol is a vendor-neutral open message standard for trade-related data delivery. The FIX Protocol standardizes the communication of capital markets data such as execution reporting, order submissions/changes, and trade allocation.
- **Proprietary data delivery methods:** In combination with the standard data delivery methods listed here, several digital asset data providers have developed proprietary methods that aim to provide clients with higher levels of customizability while minimizing infrastructure costs. [Kaiko Stream](#) and [Delta Sharing](#) are two such examples that digital asset data companies Kaiko and Nasdaq (in partnership with Amberdata) provide, respectively. With these solutions, clients can filter data (i.e., choose which data points they are interested in) at emission, instead of receiving data and then filtering it, allowing them to reduce infrastructure requirements.
- **Representational state transfer (REST) APIs:** APIs are definitions and protocols that allow computer programs to communicate with each other. REST APIs conform to a specific design architecture. Most data providers offer data through APIs that are easily integrated into third-party applications. Data that frequently updates, such as market data, on-chain data, DeFi data, derivatives data, and analytics are available on a real-time and historical basis through REST APIs. However, clients using APIs need to frequently send data requests to the server to continuously receive (real-time) data updates.
- **WebSockets:** WebSockets are bi-directional, meaning that a server can push data without clients making requests, making them ideal for real-time data delivery. Both clients and servers can send and receive messages via WebSockets as long as the connection is

open. Maintaining the connection, however, can be resource-intensive and difficult to scale.

In addition to these delivery mechanisms³, data providers also use blockchain oracles as a means to deliver data (typically to applications on-chain). Different oracle solutions are explored in Appendix A of this report. Given the immense quantity of data emanating from the heterogeneous digital asset ecosystem, it is sensible to categorize data into separate classes.

What Classes of Data do Companies Provide?

This research report segments the digital asset data landscape into three distinct, though not mutually exclusive, categories:

- (i) **market data**: relates to data that is primarily used by market participants to make investment decisions. Market data pertains to all data emanating from digital asset trading venues including spot and derivatives markets. This data includes, for example, trade data and order book data.
- (ii) **on-chain data**: relates to data that captures the activity occurring on blockchain platforms and the applications that reside on top of them. Every confirmed blockchain transaction and every pending transaction in mempools⁴ falls within the realm of on-chain data. For example, data dashboards use on-chain data, such as the total number of active addresses on a layer-1 network or the total value locked (TVL) in a protocol, to gauge a blockchain's activity.
- (iii) **forensics and market surveillance data**: relates to data used to identify and eradicate illicit activity, including use cases such as money laundering and market manipulation.

The following subsections explore each category in more detail.

Market Data

“Clean and reliable crypto prices are essential to institutional investors and enterprises holding or trading digital assets, as they require independent and accurate tools to assess their custody. This is only the beginning, as [enterprise-grade price and reference] rates will pave the way to develop multi-asset indices tailored to crypto assets.”
- Ambre Soubiran, CEO of Kaiko (Kaiko Blog, October 2022)

Unlike traditional finance (TradFi) markets, where a designated handful of venues (e.g., NYSE, NASDAQ, CBOE, and CME in the United States) generate the vast majority of market data, dozens of centralized and decentralized exchanges generate digital asset market data. Sourcing and standardizing raw digital asset market data is complex and resource-intensive. Hence, digital

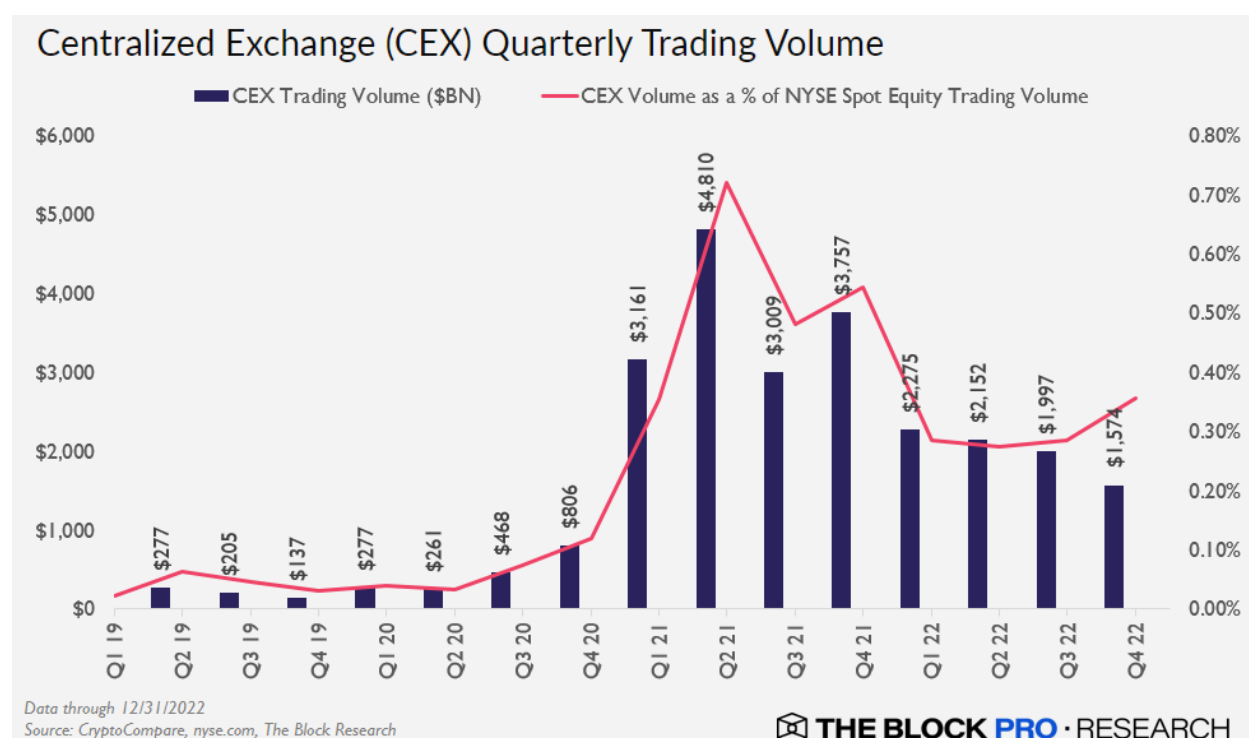
³ Please refer to Appendix B for additional information on which specific data delivery methods are supported by leading market data providers.

⁴ When a user initiates a transaction, it is first stored in a memory pool (mempool), where all pending transactions compete (usually via fee payment) for inclusion in an upcoming block. If accepted, this block gets appended to the tip of the blockchain, and the contents of the block are public for everyone to see. There is not one canonical mempool – each node in a blockchain network has its own.

asset data providers that handle the entire data value chain from procurement to refinement to delivery have emerged as mission-critical for financial institutions.

How Large is the Centralized Digital Asset Exchange Market?

The vast majority of digital asset trading still occurs on CEXes (e.g., Binance, Coinbase). Trading volumes on these CEXes have grown significantly in the past few years – on a quarterly basis, volumes have increased from ~\$100 billion in Q1-19 to ~\$1.6 trillion in Q4-22. While these volumes still pale in comparison to traditional spot equity exchanges such as the NYSE, digital asset exchange volumes have increased considerably across a fragmented landscape of ~30 venues.



The market data emanating from these exchanges can be further segmented into four major classes:

- **Aggregated data** is collected over different time intervals across venues. Examples of aggregated data include open, high, low, and close prices and volumes for a given interval (OHLCV), reference rates, and volume-weighted average prices (VWAP).
- **Trading data** consists of executed trades such as quantity, price, and timestamp. Trading data is critical for market analysis, profit and loss analysis, audit, and tax compliance.
- **Order book data** is the most granular form of data that can be used to analyze liquidity at different exchange venues. Smart order routers scan for order book data before determining the optimal venue and path for executing a trade. Typically, order book data gives insight into the following variables:
 - **Market depth** represents the total value of bids and asks on either side of the current market price.

- **Bid-ask spread** is the difference between the highest bid and lowest ask price.
- **Slippage** is the difference between a trade's expected price and the actual price at which the trade is executed.

The storage of order book data is very resource-intensive⁵. Thus, exchanges do not typically retain complete records of all order book data. However, for some financial institutions, such as quantitative trading funds and institutional liquidity providers, historical order book data is a critical input for investment decisions.

- **Derivatives data** consists of futures, options, and swap data. Examples of derivatives data include open interest (OI), volume, liquidations, funding rates, expiry-related data, and options-specific data such as put-call ratios, realized versus implied volatility, and options greeks.

On-chain Data

“Accounting for every address, every wallet, is massively complex, and a financial institution would need to spend millions of dollars and invest years of time just to learn how to do this properly.” – Shawn Douglass, CEO and co-founder at Amberdata (Hedgeweek Interview, August 2022)

The transparent and public nature of blockchain data is what differentiates the digital asset industry from its TradFi counterpart. With on-chain data, fundamental network metrics (e.g., account balances, DEX trading activity, fund flows, token supply, transaction fees, yields, etc.) can all be monitored in real time and used as an input for decision making. In contrast to CEX data which is stored off-chain in centralized databases, on-chain data is persistently generated and publicly available.

How is On-chain Data Used?

On-chain data has a plethora of applications including, but not limited to, the following use cases:

- Analyzing fundamental usage of different layer-1 and layer-2 networks. Metrics such as active addresses, on-chain value settled, transaction fees, and TVL provide a “look under the hood” of different networks.
- Analyzing the profitability of participating in DeFi protocols. Metrics such as nominal staking rewards, trading fees on DEXes, liquidity on DEXes, and collateralization ratios of lending protocols all provide insight into the opportunities and risks associated with participating in DeFi.
- Assessing the financial standing of digital asset intermediaries. While a formal framework for proof-of-reserves is [still being developed](#), on-chain data can provide insight into the digital asset holdings and, by extension, the financial stability of centralized digital asset companies.

⁵ For instance, a single instrument, such as BTCUSDT, generates several gigabytes of order book data every hour if every change to the order book (insert, delete, and update) is stored.

- Monitoring the fund flows of major digital asset market participants. On-chain data can be used to track the holding patterns of digital asset investment firms, miners, and the general digital asset investment community.

Why is Gathering and Streamlining On-chain Data so Complex?

Processing raw blockchain data into a format that institutions and investors are familiar with is a daunting task. Different blockchains have different data structures (e.g., Ethereum's account-based data structure vs. Bitcoin's unspent transaction output (UTXO) data structure), different execution engines (e.g., Ethereum's Ethereum Virtual Machine (EVM) vs. Solana's Sealevel runtime), and even different classifications of what constitutes a transaction. Gathering on-chain data and delivering it on a block-by-block or aggregated basis requires deep technical expertise – especially when this data is collected across heterogeneous blockchains.

In order to understand the on-chain data-related offerings of various data providers, three major categories of on-chain data are worth considering:

- **Network data** comprises raw metrics related to layer-1 and layer-2 networks, including statistics such as active addresses, mempool data, transaction counts, transacted value, and calculations of the circulating supply of native tokens.
- **DeFi data** comprises metrics and analytics related to applications deployed on top of different layer-1 and layer-2 networks. Such data includes, among others, liquidity pool metrics, DEX trading volume, and DeFi lending statistics.
- **NFT data** comprises metrics such as NFT sales, mints, and secondary market trading activity.

Network Data

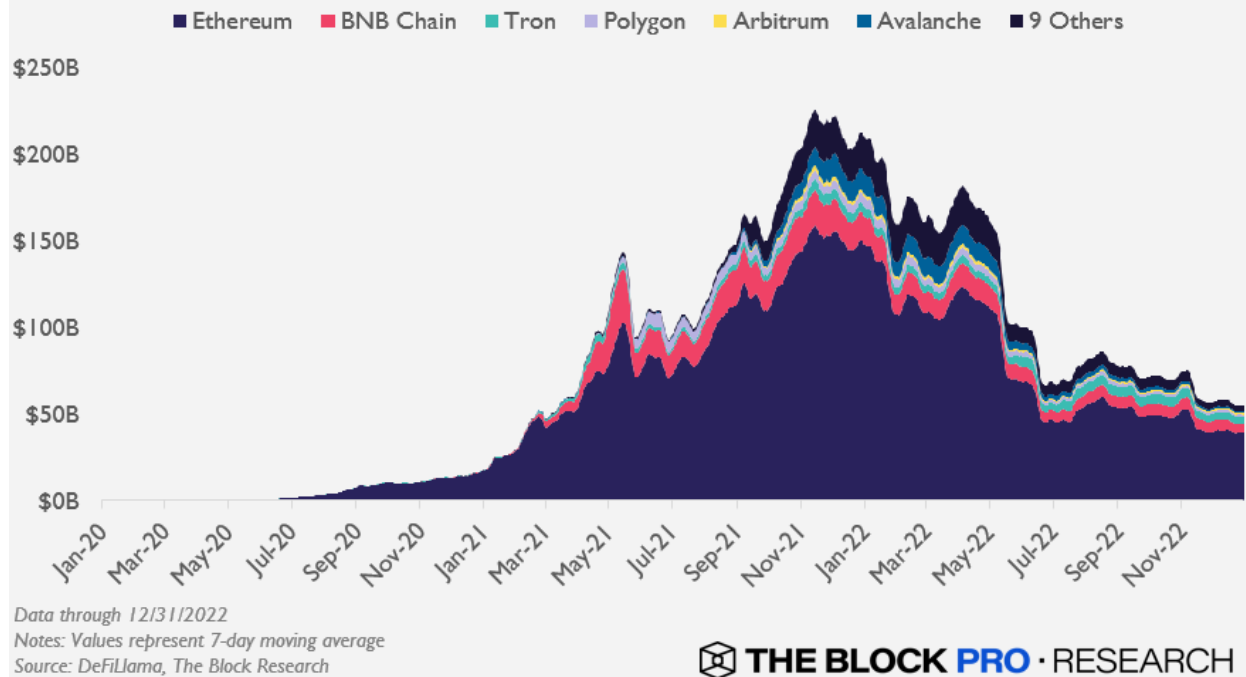
On-chain network data is critical for understanding the fundamental health and functioning of a blockchain. It is useful for analyses that involve network economics, usage, supply, and miner or validator metrics. For more information regarding on-chain data, please refer to The Block Research's [Digital Asset Data and Infrastructure: 2021](#) report.

DeFi Data

“DeFi is a radically transparent system that allows you to have visibility and build telemetry into mechanisms that incentivize the behavior of market participants.” - Shawn Douglass, CEO of Amberdata, (CoinDesk Webinar - Quantifying Opportunities and Risks in Liquidity Protocols, October 2022)

As of December 2022, [\\$54 billion](#) worth of value was locked across ~15 layer-1 and layer-2 blockchains.

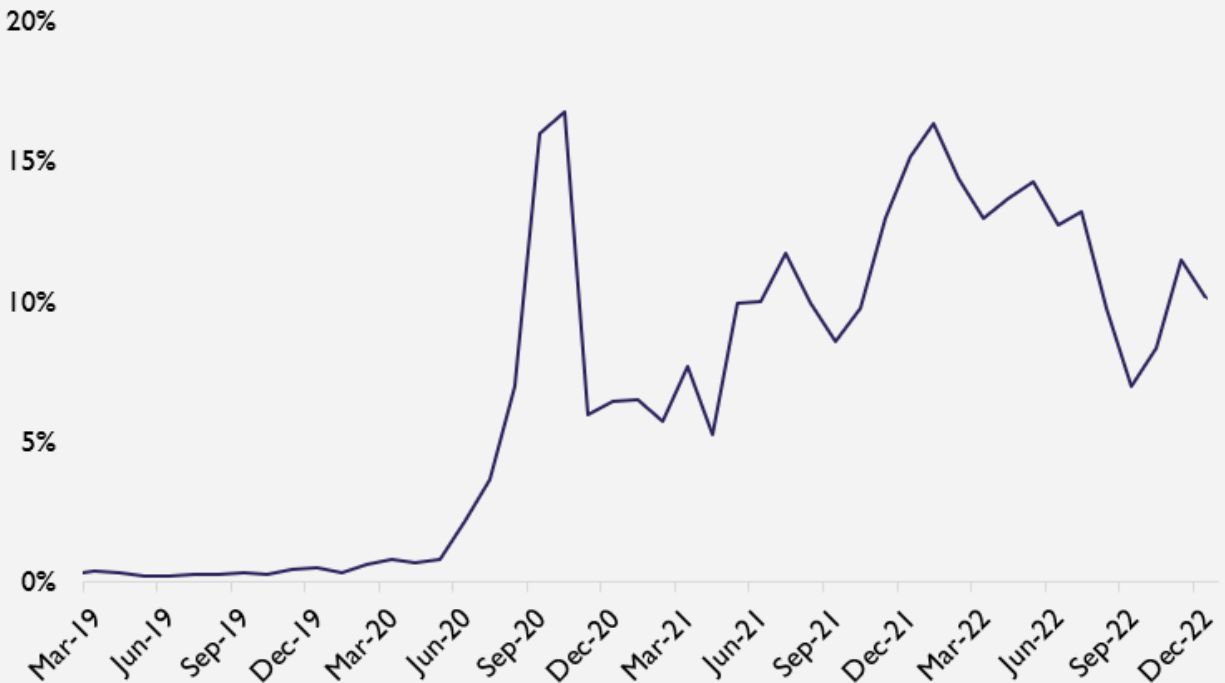
Total Value Locked in DeFi



DeFi protocols built on top of these layer-1 and layer-2 networks span a wide range of categories including spot and derivatives DEXes (e.g., Balancer, dYdX, Uniswap), DeFi lending platforms (e.g., Aave, Compound, Maker), and yield aggregators (e.g., Yearn Finance).

While the aggregate amount of financial activity in DeFi remains low relative to TradFi, DEXes have shown signs of sustained product-market fit and captured the attention of institutional trading firms. As displayed in the chart below, DEXes currently process [~12% of the trading volume](#) that is processed by CEXes.

DEX Volume as a Percentage of CEX Volume



Data through 12/31/2022

Notes: Represents monthly values

Source: CoinGecko, The Graph, The Block Research

 **THE BLOCK PRO** · RESEARCH

Accordingly, as will be discussed in Part 2 of this report, the raw amount of DEX data and the value which can be extracted from it has increased considerably over the past ~24 months.

NFT Data

While DeFi protocols are ushering in a new peer-to-peer financial paradigm, NFTs are redefining digital property rights for individuals and institutions alike. An NFT is a blockchain-based identifier (asset) that is verifiably unique and cannot be copied or substituted. NFTs hold great promise, for example in the realm of tokenizing real-world assets⁶, but come with their own set of unique challenges. For example, given their non-fungible nature, NFTs are far more illiquid (akin to the art market) when compared to fungible tokens.

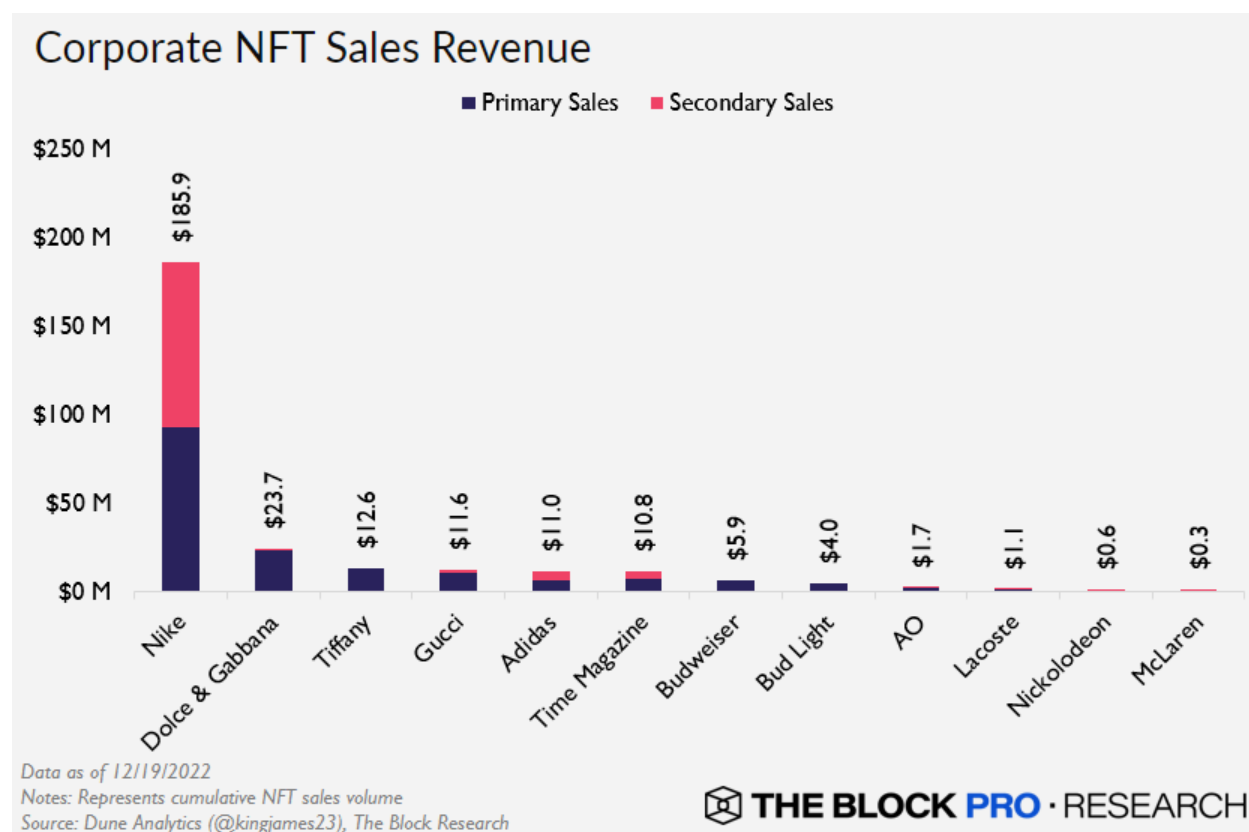
Over the past ~24 months NFTs have been employed to tokenize assets across a range of disciplines spanning art, collectibles, music, and gaming. While artists and content creators have been among the first to experiment with NFTs, institutions are well underway with integrating NFTs into their business models.

Institutional NFT Adoption

⁶ For a discussion on the tokenization of real estate and its potential benefits see “Dentons, [An introduction to fractional real estate investment](#)”.

Financial institutions have a long-lasting relationship with art collections. Leading financial institutions such as [Bank of America](#), [JPMorgan Chase](#), [Royal Bank of Canada](#), and [UBS](#) are all prolific art collectors and companies such as Visa have already [purchased NFTs](#).

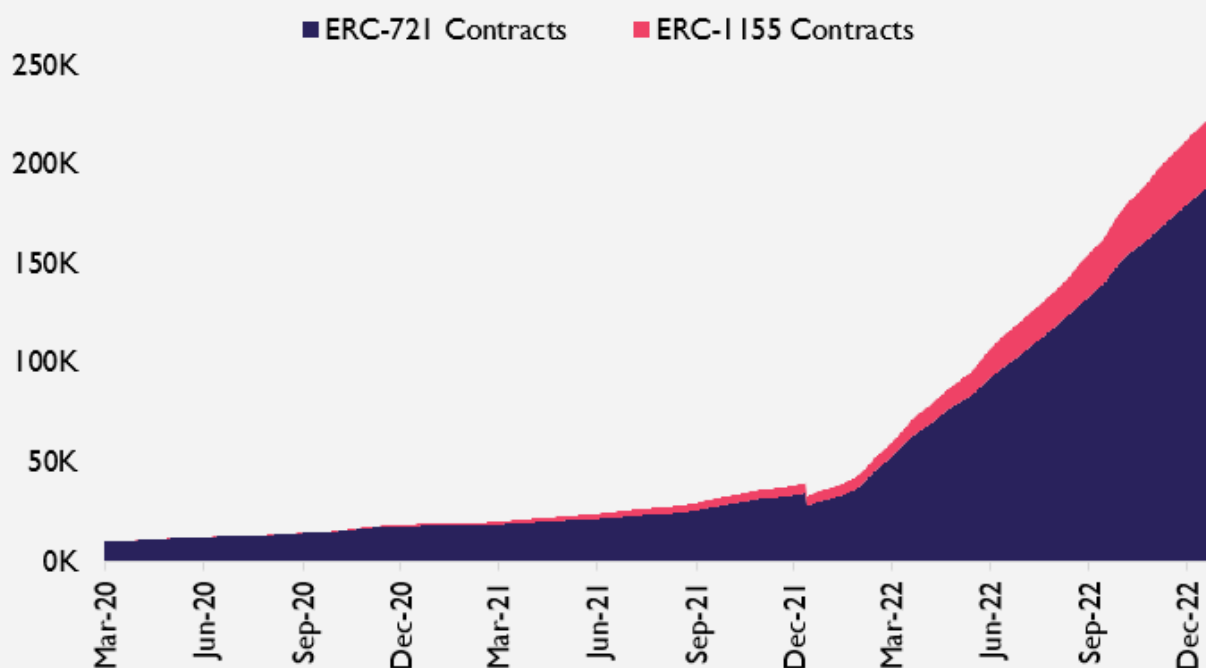
Furthermore, several corporations have already begun directly integrating NFTs into their business models. For example, in December 2021 Nike [acquired NFT startup RTFKT](#). As displayed in the chart below, the company has already generated ~\$186 million of cumulative NFT revenue across primary market sales and secondary market royalties.



In conjunction with the secular growth of NFTs, the quantity and complexity of data surrounding them have increased meaningfully over the past ~24 months. The chart below provides one approximation of the growth of Ethereum's NFT ecosystem. It shows the number of ERC 721 and ERC 1155 contracts⁷ deployed on Ethereum - which increased ~500% year-over-year in 2022.

⁷ ERC 721 and ERC 1155 are non-fungible token-related standards on Ethereum.

NFT Contracts on Ethereum



Data through 12/28/2022

Source: Coin Metrics, The Block Research

 **THE BLOCK PRO** · RESEARCH

Notably, several other layer-1 networks have also emerged as popular venues for NFT commerce. For example, Solana saw [\\$1.8 billion](#) worth of NFTs trade on its platform in 2022. Unlike fungible tokens that typically follow standardized formats (e.g., ERC-20 for Ethereum-based tokens), NFT contracts lack standardization and can have several distinct properties. Therefore, generating standardized and high-fidelity NFT data across different collections and different blockchains has proven challenging and resource-intensive.

Nonetheless, several firms focussed primarily on NFT data and analytics, such as DappRadar, CryptoSlam, NonFungible.com, and icy.tools are pushing the pace of NFT data provision. These providers are discussed in Part 2 of this report.

Forensics and Market Surveillance Data

The open and permissionless nature of blockchain technology provides individuals with an alternative to centralized financial infrastructure. However, this permissionless nature also poses new risks and challenges for institutions required to comply with financial regulations which aim to identify and eradicate illicit behavior.

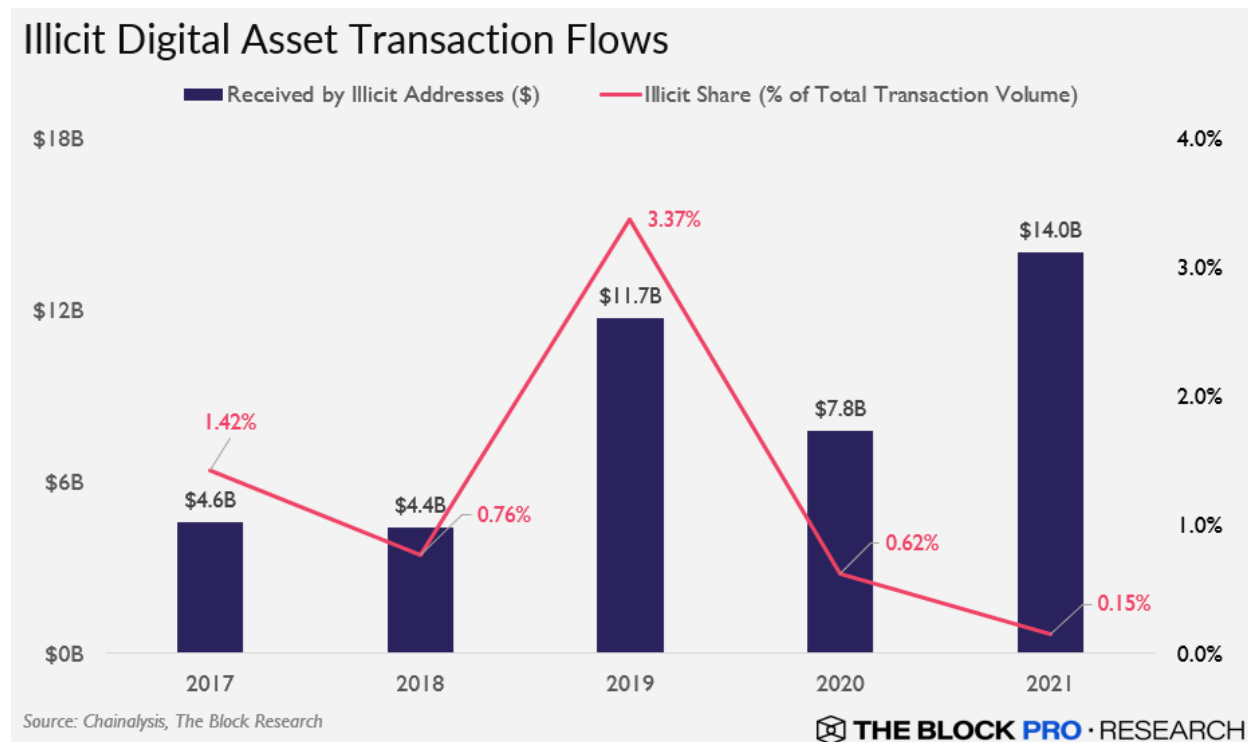
Accordingly, firms whose offerings facilitate transaction monitoring and market surveillance (and, by extension, help companies fulfill regulatory compliance requirements) have emerged as important players in the digital asset data landscape. This subsection provides insights into digital

asset intelligence companies' use of blockchain data to uncover/track illicit activity and carry out market surveillance.

Digital Asset Data and Forensic Analysis

While the illicit use of digital assets is often sensationalized, it is nevertheless clear that financial institutions, law enforcement agencies, and regulatory bodies all require tools and data analyses to monitor the digital asset space. Public entities such as Europol, the FBI, and IRS require capabilities to track and identify illicit actors.

Financial crime in digital assets is on the rise. According to [data](#) from blockchain analysis firm Chainalysis, the value of digital assets received by illicit addresses increased from \$7.8 billion in 2020 to \$14.0 billion in 2021. It is worth noting that despite this rise, the share of illicit transactions in the total transaction volume is decreasing, as displayed in the figure below.



The illicit use of digital assets can be manifold, ranging from being the payment rail of choice on dark net markets to terrorist financing. Based on new types of data available on blockchains, intelligence companies provide tailored insights that help navigate this relatively uncharted environment. The table below outlines key areas of illicit use of digital assets and how crypto intelligence companies help address these by using new tools and artificial intelligence, as well as the distinct advantages of (mostly pseudonymous) on-chain data.

Using Digital Asset Data to Prevent Illicit Use Cases

Illicit Use	Description	Role of Digital Asset Data / Analytics
Criminal balances	Criminal activity may lead to considerable balances (and proceeds thereof) being held outside the reach of law enforcement/the traditional financial system.	Identify, track and seize proceeds associated with criminal actors. Even if it cannot be established that funds were illicitly obtained, on-chain forensics may establish that they are controlled by criminal entities.
Darknet markets	Darknet markets offer illicit items such as drugs, weapons, hacking tools, or child pornography.	Identify and track actors and their sources of funds.
High-Risk jurisdictions	State actors, such as the Lazarus Group associated with North Korea, increasingly use crypto to evade sanctions.	Track and seize illegally obtained funds that are used by sanctioned entities.
Market manipulation	Activities such as 'just-in-time liquidity', which lead to worse trading prices for buyers on decentralized exchanges or spoofing on centralized exchanges.	Market manipulation can be uncovered by analyzing transaction and mempool data.
Money laundering	Money laundering requires significant liquidity and is therefore typically concentrated around a select few centralized (e.g., crypto exchanges) and decentralized venues (e.g., Tornado Cash).	Leverage data to identify venues to be targeted by law enforcement.
(NFT) Wash trading	Trading NFTs at artificially inflated prices with oneself may affect tax liabilities. For example, sell an NFT to yourself at a highly inflated price from a pseudonymous but self-funded address. If you need to generate a loss to lower tax liability, sell the NFT for a much lower price to another self-funded (but pseudonymous) address.	Identify whether trading takes place via self-funded addresses.
Scams and theft	Billions worth of USD have been stolen from victims via crypto scams and theft.	Track flows of illicitly obtained crypto to seize/restitute (e.g., funds extorted from Colonial Pipeline hack).
Terrorist financing	Terrorist organizations increasingly aim to use crypto as funding sources due to its relative anonymity and censorship resistance.	Identify funding sources of terrorist organizations and seize/blacklist their funds.

Source: Chainalysis, Elliptic, The Block Research, TRM Labs



High-Quality Granular Crypto Data Allow for New Approaches to Tackle Crime

In principle, on-chain data are: (i) constantly available in real time, (ii) highly granular, up to the individual transaction level, (iii) 'global' in nature as they contain the interactions of all network participants and are therefore not siloed, and (iv) available forever with the blockchain serving as a "universal source of truth".

Despite unique challenges stemming from chain-hopping techniques to disguise activity, chain outages, privacy tools, and pseudonymity, these aforementioned characteristics allow for the creation of market monitoring solutions that can be valuable for entities responsible for identifying and rooting out illicit activity.

It is important to note that most digital asset activity does not take place peer-to-peer or on decentralized protocols. The largest part of this activity is still carried out on CEXes. Therefore, to get a more complete picture of digital asset activity, the data from digital asset exchanges also must be taken into account. Since CEXes increasingly seek to be regulatory compliant, they require their customers to provide Know Your Customer (KYC) information to meet Anti-Money

Laundrying (AML) standards. Crypto intelligence companies also help them to screen customer transactions in real time and automatically flag high-risk customers or transactions. For example, a customer's funds may be frozen for further investigation if they are found to be associated with a blacklisted address.

Even if illicit actors are skilled at covering their tracks over time, they only need to make a single mistake, such as funding wallets from an address that had previously submitted KYC documentation, to identify and assign, with high probability, all of their criminal activity. Accordingly, law enforcement agencies have had a number of successes in identifying illicit activity that would not have been possible without these novel data, techniques, and tools. From 2016 to the present, United States government spending with forensics-focused firms Chainalysis, CipherTrace, and Elliptic has increased from a negligible amount to over \$50 million⁸, cumulatively.

Digital Asset Data Open New Possibilities for Market Surveillance Tools

In addition to new tools that facilitate the identification, tracking, and eradication of criminal activity, other areas of the emerging cyber economy may benefit from new possibilities of automatic on-chain data analysis and market surveillance. One of these is regulation. Regulatory compliance could become much more efficient when using tools such as '[embedded supervision](#)' ([ES](#)), which is defined as "a framework that lets compliance with regulatory goals be automatically monitored by reading the market's ledger, thus reducing the need for firms to collect, verify and deliver data." In October 2022, the European Commission put out a [tender that closed on December 1, 2022](#), to study how to automatically monitor the Ethereum network through ES and collect regulatory data in real time.

Blockchain intelligence companies are well-positioned to develop such tools based on their know-how and existing tech stack for market surveillance. Using novel approaches, which make use of the rich data and infrastructure in the nascent crypto ecosystem, has the potential to also significantly lower the cost of compliance which can be exceedingly high. For example, the total projected cost of financial crime compliance in the United States and Canada for 2022 is a [staggering \\$56.7bn, up 13.6% from 2021](#).

Lastly, while digital asset intelligence and forensic companies are helpful in using high-quality on-chain data to efficiently implement [micro-prudential regulation](#), they may also provide supervisors with tools to monitor systemic risks of the crypto ecosystem as a whole ('macro-prudential' surveillance). For example, interconnections and exposures between protocols could be monitored in real time with on-chain data. This allows supervisors to simulate how well the digital asset ecosystem is able to absorb financial shocks.

The following section provides an overview of the areas of digital asset infrastructure that are explored in this report.

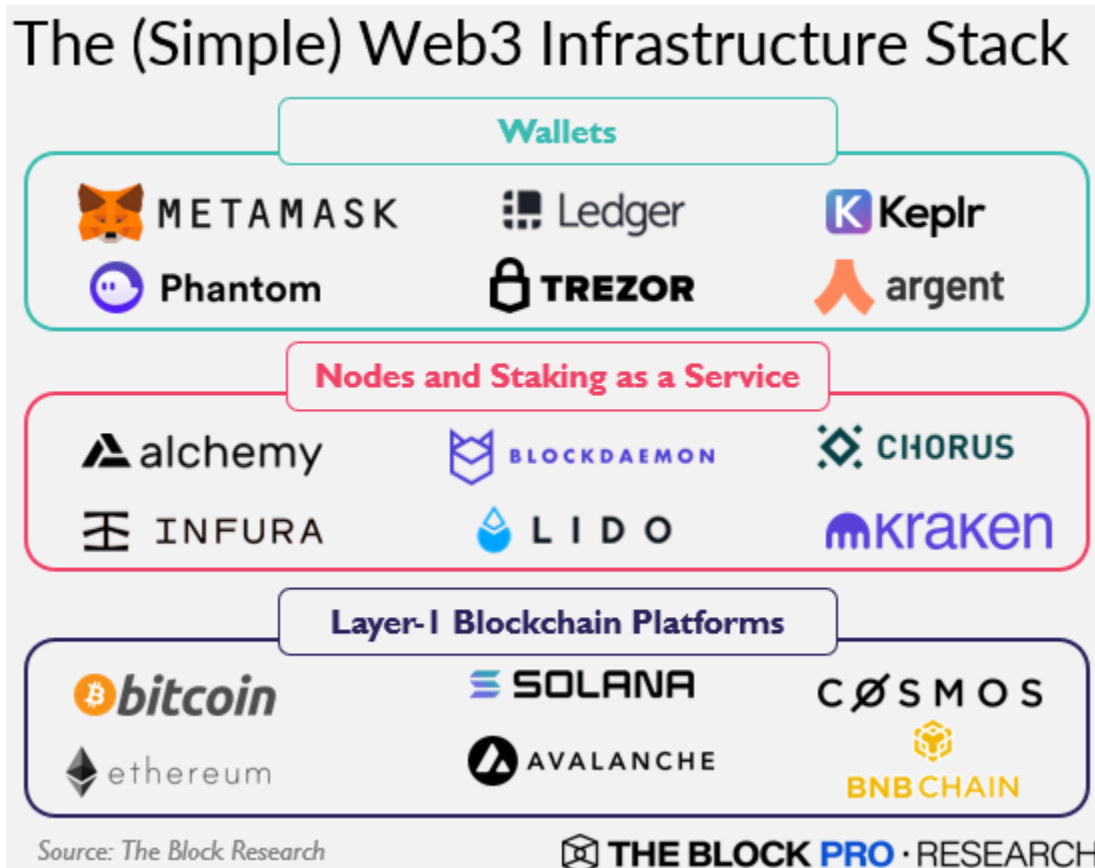
⁸ Source: Federal Procurement Data System.

Digital Asset Infrastructure

Infrastructure is a broad term that is often overused in the context of digital assets. It encompasses verticals such as [layer-1 platforms](#), [layer-2 scaling solutions](#), [interoperability networks](#), [custody solutions](#), and [storage services](#). For the scope of this report, infrastructure refers to the critical products and services that let users operate and interact with public blockchains on a day-to-day basis. These products and services are provided by three main classes of companies:

- (i) **Staking-as-a-service firms** provide off-the-shelf services that enable users and organizations to directly participate in securing different proof-of-stake (PoS) blockchain networks through validation.
- (ii) **Node-as-a-service firms** provide access and maintenance to shared and dedicated nodes across layer-1 and layer-2 networks.
- (iii) **Digital Wallet providers** provide software and hardware products that enable end users to interact with blockchains and submit transactions for approval.

Before diving into the specifics of each of the above categories, the following figure shows the simple web3 infrastructure stack from a bird's eye view.



Layer-1 blockchains sit at the base of the web3 infrastructure stack. They perform critical network functions such as achieving consensus on the validity of transactions and ensuring that blockchain data remains available. Broadly speaking, staking requirements and node architecture

(i.e., infrastructure requirements) are largely a function of the technical design of underlying layer-1 and layer-2 networks. For example, whether a blockchain network is monolithic⁹ or modular¹⁰ impacts the computational resources required to stake (i.e., participate in validation) and operate node infrastructure. For more information regarding the network architecture and validator node requirements, please refer to The Block Research's [Comparing Layer-1 Platforms: 2022 Edition](#) report.

Staking-as-a-service

The term “staking” refers to locking up tokens in a smart contract to participate in blockchain networks via different methods such as block production, governance, and validation. So long as staked tokens originate from a sufficiently distributed base of token holders, the cost of attacking a PoS network (which can be viewed as a proxy for overall network security¹¹), increases with the amount of capital staked by a network’s validator set. In return for locking up their capital, stakers are rewarded with native protocol tokens.

There are four widely recognized, though not mutually exclusive, avenues through which interested parties can participate in staking:

- (i) **Individual staking** whereby users manage the staking process from end-to-end. This entails sourcing and maintaining computer hardware, running blockchain software, and depositing native tokens in smart contracts to meet network requirements.
- (ii) **Pure play staking firms** are focused on providing individual and institutional users with staking-related services at their own facility or at co-located facilities. For example, staking firms such as Chorus One and Figment run node infrastructure on users’ behalf.
- (iii) **CEXes** which own their own staking hardware or outsource staking to a pure-play staking firm. Although most CEXes stake assets on users’ behalf, firms such as Coinbase, allow individuals to access dedicated validator nodes.
- (iv) **Liquid staking protocols** are community-owned and operated networks (e.g., Lido and Rocketpool) that are governed by decentralized autonomous organizations (DAOs) and facilitate the issuance of liquid staking tokens.

With the exception of the first avenue of staking (individual staking), all other avenues typically involve outsourcing infrastructure operation¹² to providers such as Amazon and Google Cloud. The decision to outsource infrastructure operation should be carefully considered, as it can make staking firms vulnerable to the practices of their related infrastructure providers. For example,

⁹ Monolithic refers to blockchains that handle all functionalities on the same layer - i.e., execution, settlement, and data availability all occur on the same layer.

¹⁰ Modular blockchains handle specific tasks such as data availability or execution while offloading other responsibilities to different layers.

¹¹ This is because most PoS networks slash the stakes of bad actors. With a higher total amount staked, the cost to attack the network also increases and the ability to slash means that bad actors stand to lose more in the case of an unsuccessful attack.

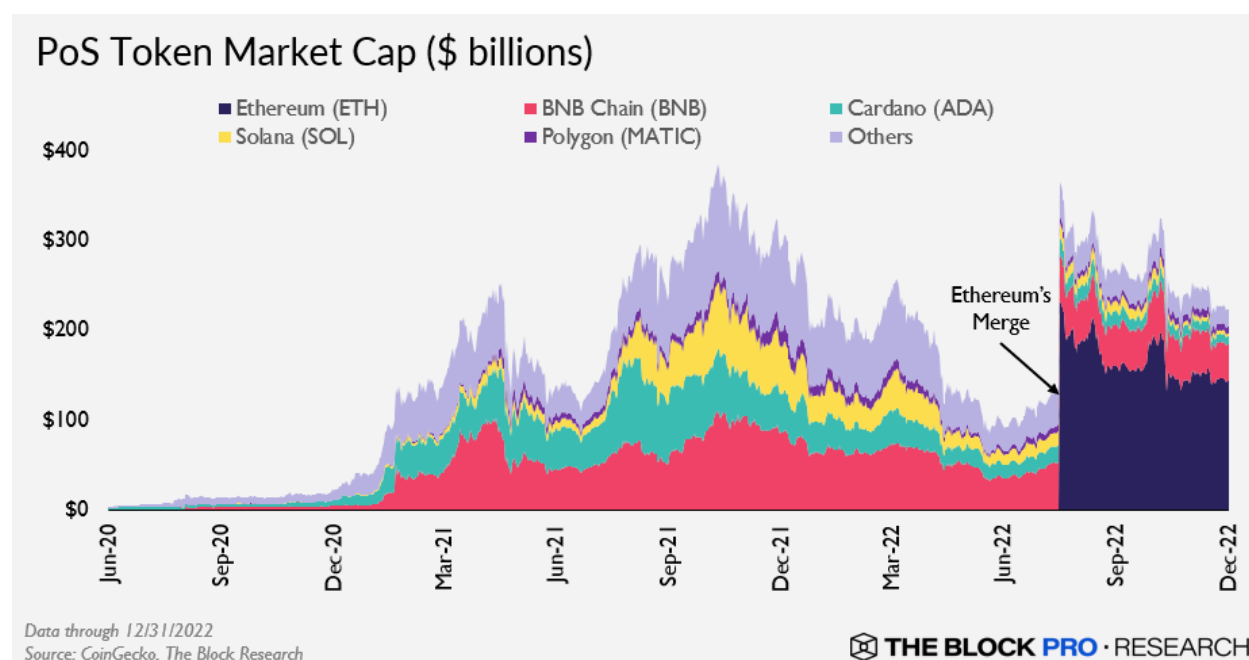
¹² Some of the staking service providers like Rocketpool allow users to run validator nodes on their own infrastructure.

Hetzner, a Germany-based cloud service provider, [blocked all Solana network activity](#) on its servers in November 2022, thus causing ~1,000 Solana validators to go offline.

Major Developments in the Staking Landscape

Since 2021, there have been two major developments in the staking market. Firstly, In September 2022, Ethereum moved from proof-of-work (PoW) to PoS consensus, making all ETH (collectively worth ~\$145 billion as of 12/31/2022) eligible for staking¹³.

As displayed in the chart below, the total market cap of PoS assets (which serves as a proxy for the total addressable market for staking-as-a-service providers) has increased by a factor of ten from ~\$22 billion in December 2020 to \$223 billion by December 2022.



Secondly, liquid staking derivatives created a new avenue for staking by allowing users to stake their assets, yet still “re-use” them in other DeFi activities. With liquid staking derivatives, DAOs and staking-as-a-service providers source and stake PoS assets (e.g., ETH) on behalf of their users and issue synthetic tokens (e.g., Lido’s stETH) against these stake assets. Accordingly, users can deploy these synthetic tokens for yield-generating strategies (e.g., trading, lending, providing liquidity on DEXes, etc.) while still receiving staking rewards.

What Role do Staking-as-a-service Providers Play?

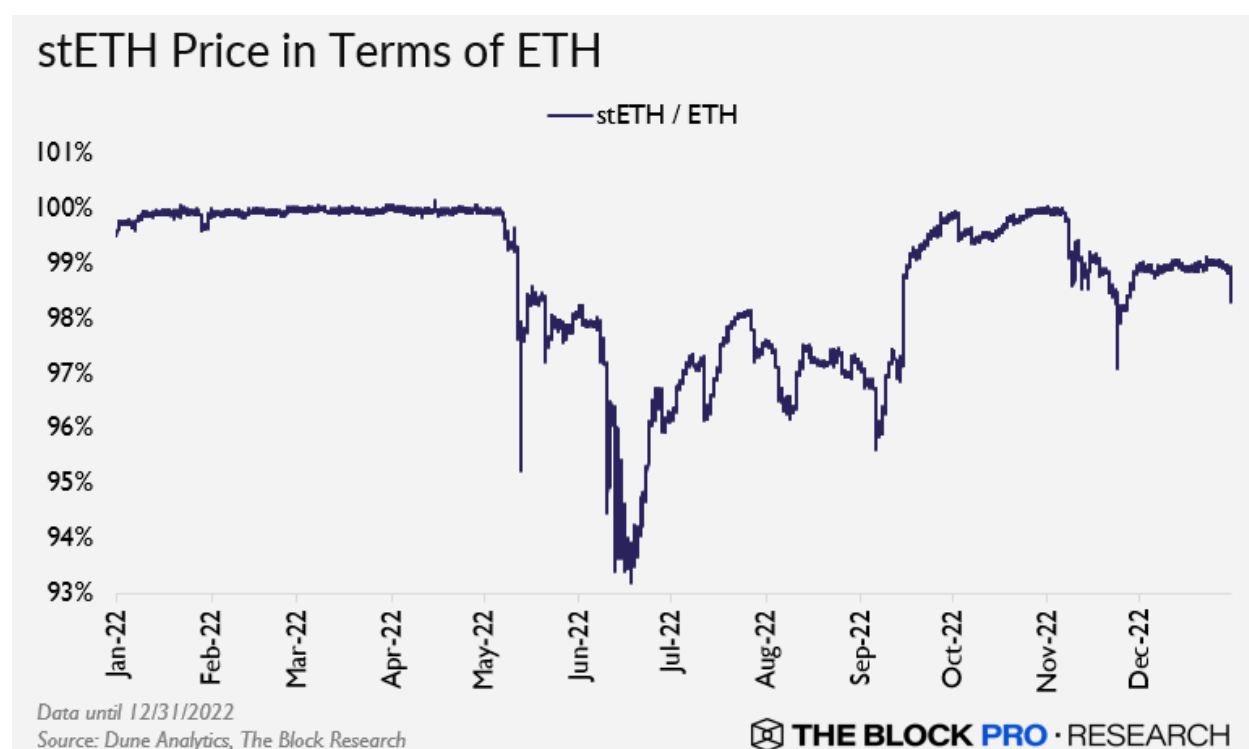
In the context of staking ETH on the Ethereum blockchain, staking-as-a-service providers address three major issues faced by users. Firstly, they eliminate the need for users to procure and operate computer hardware around the clock. Secondly, they eliminate staking minimums, by aggregating the ETH balances of many users into batches. In the absence of this pooling, an individual user

¹³ Although the beacon chain launched in Nov 2020, there was uncertainty around when ETH would move to PoS from PoW consensus, which took place in September 2022.

would need to own or obtain 32 ETH (which translates to an upfront investment of ~\$38,400 as of 12/31/2022) to participate in staking. Thirdly, as previously mentioned, liquid staking service providers allow users to generate staking rewards yet retain the ability to re-deploy staked assets into other investment strategies.

What Risks Come with Staking?

Benefits offered by staking service providers are not entirely without risks. For example, there is no guarantee that synthetic staking tokens (e.g., stETH) will trade at the same value as their underlying staked assets (e.g., ETH) on the secondary market. In the face of adverse market conditions and forced liquidations, staking derivatives have traded at steep discounts. For example, during the Three Arrows Capital (3AC) contagion in June 2022, a liquidity crunch for stETH drove its discount to ETH as low as 7%.



Furthermore, outsourcing staking operations to a third party can also result in custody risks should users no longer have signing power over their transactions. In the wake of digital asset intermediary bankruptcies (e.g., BlockFi, Celsius, FTX, Genesis), users should be well aware that there are significant risks associated with entrusting their assets to centralized counterparties.

Finally, staking-as-a-service provider downtime or adverse behavior could potentially result in a user's stake getting [slashed](#). Individuals that employ staking-as-a-service firms entrust these firms to perform all of their respective duties in the networks they are active in. Should these firms fail to meet network staking requirements, their customers could suffer financial repercussions.

Node-as-a-service

While infrastructure requirements can vary considerably across different layer-1 and layer-2 networks, all of these platforms rely on distributed networks of nodes. These nodes perform different functions such as i) processing and attesting to the validity of transactions, ii) storing transaction data, and iii) acting as interfaces to submit transactions and access their related data. For example, remote procedure call (RPC) nodes are essential for decentralized application (dApp) development since they allow for dApps to communicate with the blockchain¹⁴.

At the highest level, nodes can be subdivided into full nodes, which maintain all transaction records and have voting rights, and light nodes, which store and provide the necessary data to accommodate users' daily activities. Full nodes can be further broken down into several sub-subcategories such as pruned full nodes and archival full nodes¹⁵. By employing node-as-a-service firms, users ranging from dApp developers to institutional investors to data analytics firms can connect to a blockchain without the need to set up, run and maintain the necessary underlying node infrastructure, which can be very resource intensive¹⁶.

Self-hosted Nodes vs. Node-as-a-Service

Every dApp needs to connect to a blockchain to send and receive data. It can connect to a blockchain in two ways:

Via **self-hosted nodes** – In this case, the dApp developer manages the end-to-end infrastructure of a full node capable of pushing transactions into a blockchain. Self-hosted nodes reduce latency and give developers more control over their infrastructure but require individual operators to possess significant technical expertise. Although self-hosting can theoretically be the most reliable method to ensure an uninterrupted connection to a blockchain, maintaining a full node that always functions optimally can be costly and time-consuming.

Via **node-as-a-service** providers – Developers can outsource node operations to third-party services to focus on application development. Node service providers can further be divided into i) centrally hosted node networks and ii) peer-to-peer node networks. Centrally hosted node networks typically rely on third-party cloud hosting companies such as Amazon and Google, which provide fully managed services to host nodes. Peer-to-peer (p2p) node networks can be used by developers who neither want to run their own node nor rely on a centrally-hosted node. Node providers such as Ankr and Pocket Network (discussed in Part 3 of this report) incentivize individuals or organizations to run full nodes for multiple blockchains.

Important Attributes of Node-as-a-Service Providers

¹⁴ Remote Procedure Calls (RPCs) allow programs on one machine to call a function on another machine without knowing that it is remote. Using RPCs, one machine can execute certain tasks (which the machine itself may not be able to carry out) on a different machine.

¹⁵ For more detailed information on nodes and their subcategories, the reader is referred to the article "[Blockchain Nodes: An in-depth guide](#)".

¹⁶ A specific example for the associated cost of running a node, the reader is referred to our previous [report](#).

Node-as-a-service providers compete across the following dimensions to meet the needs of their customers:

- **Blockchain and node support** – Node-as-a-service providers usually specialize in a set of blockchains. Users must select providers that offer services on the required blockchains. Different users may require different types of nodes depending on their performance needs. Shared nodes are used by multiple customers simultaneously, whereas dedicated nodes are used exclusively by a single user. Scaling an application with a single shared node can prove difficult depending on bandwidth constraints. Users with very high performance and stability requirements may also use node clusters, which usually have at least two dedicated nodes that help with failover protection and load balancing (both described in more detail below).
- **Customer Support** – Node providers must have support teams with enough personnel distributed across different time zones to ensure that queries can be answered expediently. Most node-as-a-service companies have dedicated channels on applications such as Discord, Slack, and Telegram to handle such queries.
- **Data Accuracy** – Since blockchains operate in a decentralized fashion, there is a risk that data can be inaccurate. For example, users may not see transactions in their wallets. Such problems can arise because of a node's inability to pull accurate data due to, for example, traditional load balancing that ineffectively routes traffic. Benchmarking tools like [Alchemy's Data Accuracy Benchmark](#) can help compare the data accuracy of different node service providers.
- **Decentralization** – Decentralization of nodes can be important for an application to maintain minimum redundancy that ensures 100% uptime, and, as argued by purists, censorship resistance. Node decentralization can be further broken down into decentralization vectors, such as, for example, blockchain client diversity, cloud provider diversity, or geographical diversity.
- **Failover Protection** – Failover protection is an operational mode that automatically switches a network to a redundant server in the event of a system failure. It is used to prevent excessive loss of data and downtime from system failures in blockchain and 'web2' servers. Failover protection is critical for blockchains due to their need for constant accessibility but also for node operators to prevent being penalized when the consensus algorithm is violated due to a node's temporary downtime. For example, Ethereum node operators may be slashed if a node signs a transaction multiple times, which can happen if the node goes offline and comes back online.
- **Response Time** – Response time is the time it takes for a node to return requests. It can be a critical consideration for users that employ time-sensitive operations, such as traders, when choosing a provider.
- **Scalability** – Scalability is one of the most critical qualities of a node provider. Node scalability can be affected by two factors, i) load balancing and ii) autoscaling. dApps without load balancers rely entirely on a single node's performance. Load balancing is a scaling mechanism that directs client requests to nodes with the lowest workload in order to prevent a single node from being overwhelmed. Autoscaling refers to an automated mechanism that increases or decreases the resources allocated to a network based on its

usage patterns. This allows blockchains to provide consistent performance for users in periods of high demand while reducing costs when they experience low traffic.

Wallets

Wallets are user-facing applications that allow users to send transactions to blockchain nodes. Wallet types span non-custodial wallets, such as hardware and browser wallets, and custodial wallets such as exchange wallets. At their core, wallets are a set of keys: a private key and a public key. The private key proves ownership of the digital assets associated with the public key, while the public key maps to a public address that receives transactions. Non-custodial wallets require users to manage their keys, while custodial wallets offload private key management and its responsibility to entities like custodians and CEXes. This report focuses on 'off-the-shelf' wallet solutions and their providers. Readers interested in 'Wallet-as-a-service', which provides institutional customers with tailored solutions are encouraged to refer to last year's data and infrastructure [report](#).

Custodial Solutions

Given the hundreds of billions of dollars worth of client assets on their platforms, CEXes such as Binance and Coinbase account for a large share of balances held in custodial wallets. Additionally, institutional custody firms such as BitGo, Fireblocks, and Copper have emerged as popular service providers that have won the trust of many institutions. While it is difficult to pinpoint the total dollar value of all digital assets under custody, the value of assets secured by institutional custodians (e.g. BitGo, Fireblocks, Copper) is estimated to be over \$200 billion. Interested readers can dive deeper into custody solutions in The Block's [Institutional Custody for Digital Assets Primer](#).

Non-custodial Wallet Providers

Non-custodial wallet providers allow users to take custody of their private keys. Self-custody removes the counterparty risk involved in entrusting funds with a third party, which has once again proved to be critical in the wake of the collapse of FTX. But it also presents its own set of risks, such as private keys getting lost or stolen and hardware wallets getting damaged.

Wallets can be divided into two broad categories – hot wallets and cold wallets. While both hot and cold wallet solutions are available in custodial and non-custodial implementations, this report discusses them in the context of non-custodial implementations.

Hot Wallets

Hot wallets such as MetaMask, Trust Wallet, and Exodus allow users to take custody of their private keys through digital interfaces. These hot wallets are connected to the internet 24/7, which allows for quick access to funds for on-chain traders and exchanges processing customer transactions and withdrawals. This accessibility, however, comes at the cost of increased

exposure to malicious actors. Hot wallets are vulnerable to attack techniques, including phishing attacks, clipboard malware¹⁷, keyloggers¹⁸, and man-in-the-middle¹⁹ attacks.

MetaMask, the most popular hot wallet, injects a global API called window.ethereum into web3 sites allowing users to access a set of Javascript functions to interact with the blockchain through a node. This allows websites to push transactions and messages to and from the user, read data from the blockchain that the wallet is connected to, and identify the user. This is critical for users to be able to interact with decentralized applications. Hot wallets also typically allow users to adjust gas limits, which enables them to prioritize their transactions for inclusion in a block in return for a higher transaction fee or to pay a lower gas fee for transactions that are not time sensitive.

Cold Wallets

Cold wallets, though they can be as primitive as private keys written on a physical sheet of paper, typically employ small hardware devices to store users' private keys offline. These cold wallets are generally more secure than hot wallets because they are only connected to the internet when they are in use. Even when connected to the internet, cold wallets typically sign transactions within their respective devices and broadcast that signature via the internet – meaning that the wallet's private keys are never exposed to the internet. Although cold wallets can still be subject to man-in-the-middle attacks, this makes them extremely difficult to compromise. This security advantage, however, translates to a disadvantage in terms of the speed of transaction execution.

Extra layers of security add more friction for executing transactions and are, thus, unsuitable for active trading strategies. For example, the slower execution flow typical of cold wallets may result in failed transactions on automated market makers (AMM) if the price moves outside the predetermined range before the transaction is picked up.

In addition to distinguishing between custodial and non-custodial as well as hot and cold wallets, Ethereum's Ethereum Virtual Machine (EVM) supports both externally-owned account (EOA)-based wallets and smart contract-based wallets.

EOA Based Wallets vs. Smart Contract Wallets

Ethereum Virtual Machine (EVM) compatible chains have [two types of accounts](#) - EOAs and smart contract accounts. EOAs are controlled by any user who holds private keys, whereas smart contract accounts are controlled by code. Smart contracts open up a new design space for wallets and can provide a more secure user experience than EOA-based wallets. Smart contract wallets allow native functionalities like social recovery, multisig security, multi-factor authentication,

¹⁷ When a user is trying to copy and paste their address to transfer funds, the attacker's address gets copied instead.

¹⁸ Keyloggers are one of the oldest tools used to conduct cyber attacks. Keylogger viruses recognize keystrokes and patterns on touch screens. An attacker can use the information a keylogger produces to gain access to passwords or seed phrases.

¹⁹ Man-in-the-middle attacks occur when an attacker intercepts the communication between the user and a web application.

whitelisting contracts and addresses, bundled transactions²⁰, daily transaction limits, and emergency account freezing.

²⁰ For example, the wallet can let users bundle, currently separate, approve and spend transactions as one transaction.

02/16/23

The State Of Digital Asset Data And Infrastructure:
2023 Edition

2

Digital Asset Data Providers

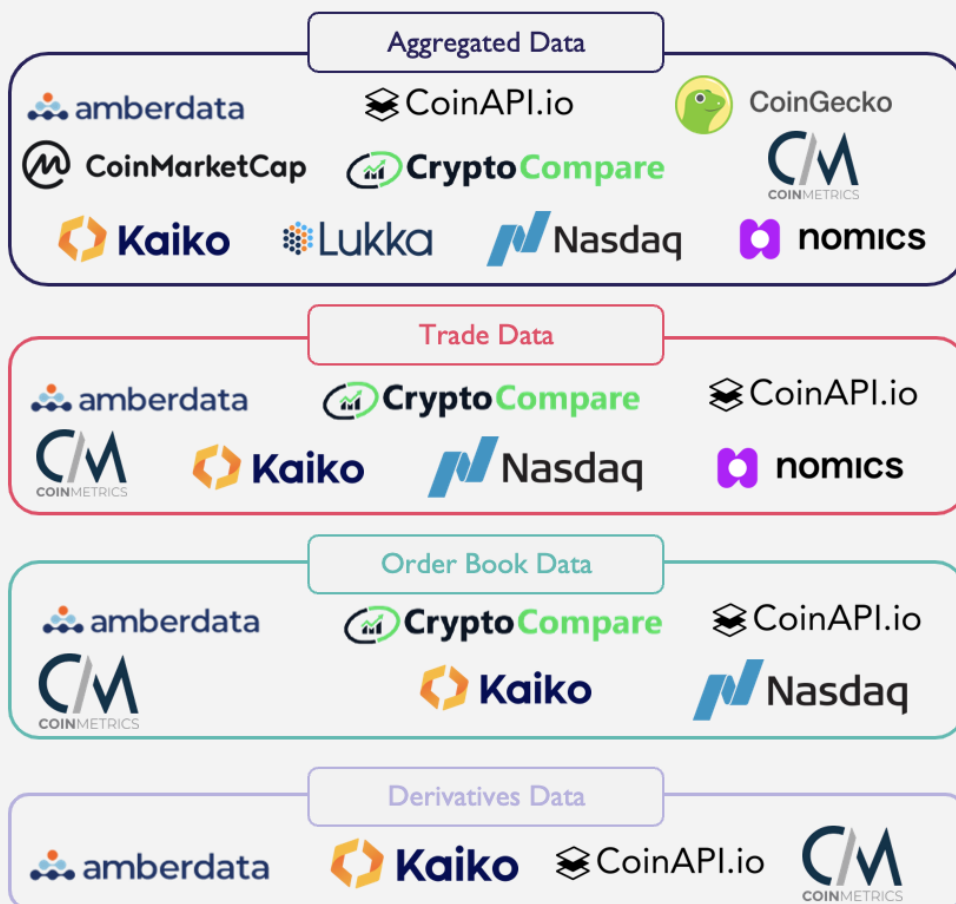
Part 2: Digital Asset Data Providers

“Data is the new oil. It’s valuable, 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, analyzed for it to have value.” – Clive Humby, Chief Data Scientist, Starcount (2006)

Raw data is an essential foundation for value creation and economic activity in the digital asset space. Much like oil, digital asset data is effectively unusable in the state in which it is collected. If raw data is the crude oil of digital assets, then entities that collect, process, and standardize this data into consumable formats are comparable to refineries. Financial institutions, developers, traders, and investors all rely on the services of these data providers to make sense of the ever-growing sea of digital asset data. This part of the report explores key digital asset data providers for market, on-chain and forensic data. It finishes with an overview of data providers’ fundraising and M&A activity.

Market Data Providers

Digital Asset Market Data Landscape



Source: The Block Research

 **THE BLOCK PRO** · RESEARCH

The following table compares leading market data providers across a range of metrics pertaining to aggregated, trade, order book, and derivatives data, as well as the number of exchanges and features they support. While some data providers may seem similar at a glance, there is a large degree of variance in the specific features and the depth of their offerings.

Comparison of Select Market Data Providers

Organization	Target Customer	Aggregated Data			Trade Data			Order Book Data				Derivatives Data		
		Exchanges	Reference	OHLCV	Tick-level	Real	Historical	Agg.	Bid-Ask	Real	Historical	CEX	CEX	On-chain
		Supported (#) ⁽²⁾	Rates	Candlesticks	Trade Data	Time Data	Data	Depth	Spreads	Time Data	Data	Futures	Options	Derivatives
Amberdata	Institutional	37+	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CoinAPI	Retail, Institutional	337	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
CoinGecko	Retail, Institutional	200+	×	✓	×	✓	✓	×	×	×	×	✓	×	×
CoinMarketCap	Retail, Institutional	200+	×	✓	×	✓	✓	×	×	×	×	✓	×	×
Coin Metrics	Institutional	40	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
CryptoCompare	Retail, Institutional	200+	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×
Kaiko	Institutional	106	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
Lukka	Institutional	10+	✓	✓	✓	✓	✓	✓	✓	×	✓	×	×	×
Nasdaq	Institutional	15+	✓	✓	✓	×	✓	✓	✓	✓	✓	×	×	×
Nomics ⁽¹⁾	Retail, Institutional	200+	×	✓	✓	✓	✓	-	-	-	✓	✓	×	×

Notes: (1) Based on publicly available data, Nomics provides historical order book data – the granularity of which is not disclosed. (3) Represents the total number of exchanges supported inclusive of spot and derivatives trading venues. Data offerings across individual trading venues and trading pairs could vary meaningfully, and the table above is intended to provide a high-level overview of data offerings.

Source: Company websites and documentation, The Block Research

THE BLOCK PRO · RESEARCH

Amberdata offers comprehensive coverage of all the market data categories. Amberdata also provides historical data related to liquidations, order books, and trades that are usually not provided by exchanges (due to the aforementioned high storage requirements for such data). The company also offers three types of reference rates: (i) weighted averages such as Volume Weighted Average Price (VWAP) and Time Weighted Average Price (TWAP), (ii) pre-trade reference rates describing the best takable price across a number of exchanges, and (iii) single asset index reference rates which calculate average prices across several fiat currency and stablecoin pairs. Amberdata provides snapshots (one minute apart) of entire order books across spot, swap, futures, and options markets. Although some data providers offer snapshots at a lower frequency, Amberdata records all the updates (gapless data) to order books, making snapshots less relevant. As Amberdata records what happens between snapshots, it allows clients to replay the order book tick-by-tick. Following its acquisition of Genesis Volatility in [October 2022](#), Amberdata's derivatives data expanded coverage to on-chain options protocols such as Ribbon Finance and Lyra. Layer-2 scaling platforms have made derivatives trading more practical on Ethereum, and as such, the on-chain derivatives space is likely to see increased user adoption. Amberdata delivers data in a variety of formats such as APIs, cloud-based delivery formats like Amazon S3 (Simple Storage Service), and WebSockets. Amberdata also offers FIX Protocol delivery for trades, OHLCV, and order book snapshots and updates for swaps, futures, and spot trading.

CoinAPI's data offering spans aggregate data such as reference rate and OHLCV candlesticks, trade data, order book data, and derivatives data. Its reference rate is a 24-hour VWAP for a number of exchanges. Along with real-time data, CoinAPI offers historical trade data via REST APIs. The organization provides current order book data that covers the entire depth (all bids and asks) and historical order book data with the top 50 bids and asks (Level 2). CoinAPI provides futures and options data from CEXes like Binance, Deribit, and Bybit. CoinAPI delivers data via REST and FIX APIs and WebSockets.

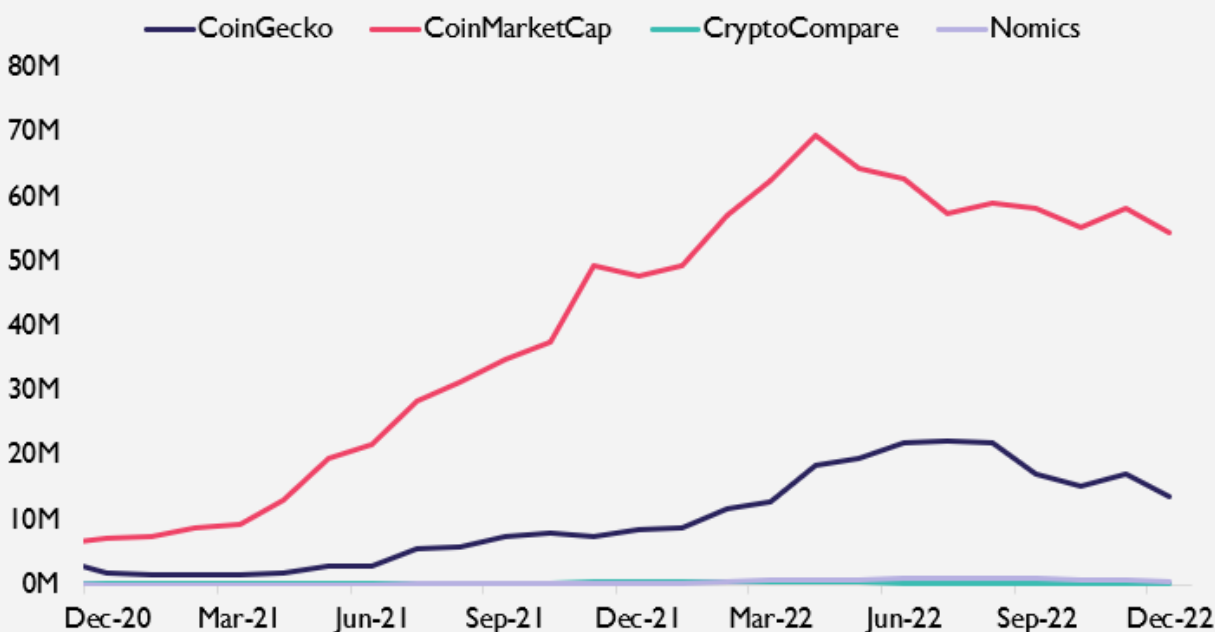
CoinGecko is a data provider with products similar to those of CoinMarketCap. It offers price and volume data (spot/derivatives) as well as charts for thousands of tokens across hundreds of exchanges. CoinGecko also hosts its own proprietary API with endpoints that return various types

of market data including live trading volume, live prices, exchange volumes, historical data, and derivatives data.

CoinMarketCap has found a strong niche in providing retail market participants with tools and services to deepen their understanding of the digital asset landscape. CoinMarketCap offers price, volume, and market cap data for nearly 9,000 digital assets and allows users to rank them across any of those metrics. Its platform also allows users to view the trading volume, average liquidity, and the number of markets and coins supported for over 200 CEXes and DEXes across spot, derivatives, and lending protocols. Additionally, CoinMarketCap hosts blockchain explorers for Bitcoin, Ethereum, Litecoin, and BNB Chain that allow users to search blocks, addresses, and specific transactions. For more sophisticated users, CoinMarketCap offers an API consisting of a suite of RESTful JSON endpoints that return cryptocurrency, exchange, global metrics, and fiat currency data.

As the market data provided by companies such as CoinMarketCap is freely available, their websites have emerged as popular destinations for retail investors. While a number of other providers such as CoinGecko, CryptoCompare, and Nomics provide similar services, CoinMarketCap still remains the most popular destination for free, retail-facing dashboards - since the beginning of 2022, monthly page views of its website have ranged between 40 million and 70 million according to data from semrush.com.

Monthly Page Views



Source: semrush.com, The Block Research

 **THE BLOCK PRO** · RESEARCH

Coin Metrics hosts several unique digital asset market data products. Its market data feed provides historical and real-time data for data types including order book snapshots, OHLCV candlesticks, open interest, funding rates, and liquidations delivered via WebSocket APIs. Additionally, Coin Metrics offers a suite of single-asset, multi-asset, and unique crypto asset benchmarks, [CMBI](#) (Coin Metrics Bletchley Indexes) used by crypto and traditional financial institutions. In collaboration with MSCI and Goldman Sachs, it recently announced [datonomy](#), a classification system for digital assets. Datonomy aims to create a standard method for market participants to analyze digital assets to enhance risk management, reporting, and strategization. Coin Metrics's [order book data](#) provides 10-second snapshots for +/-10% depth and top 100 bids and asks. It also gives snapshots of the entire order book that are one hour apart. Coin Metrics offers calculation agent services for institutions seeking to administer their custom indexes as well.

CryptoCompare offers a suite of data products suitable for both institutional and retail market participants. CryptoCompare delivers spot and derivative CEX market data via REST API in JSON and CSV formats. Its data is also standardized across instruments to allow for comparisons across markets, strategy backtesting, and market research. CryptoCompare also offers granular and standardized order book data (level 2) captured in snapshots once per minute, containing the top 30 bids and asks. The organization also provides indices in collaboration with MVIS (MarketVector Indices), Blockdaemon, and FT Wilshire.

Kaiko's centralized venue market data covers all market data categories except on-chain derivatives. Although Kaiko's trading data goes back to 2011, its order book data goes back as far as 2017. It offers snapshots of all the bids and asks within 10% of the best bid and ask twice a minute. It also provides tick-level order books daily via CSV files and the 5-week rolling history via Kaiko Stream. Kaiko offers quantitative data designed to produce actionable insights, including implied volatility, custom portfolio valuation, and a Value at Risk (VaR) estimator for single and multi-asset portfolios. Kaiko provides two types of rates: price and reference rates. While the company's price rates do not require exchanges to satisfy specific criteria, reference rates are derived from prices on exchanges that fulfill certain conditions, such as being in operation for at least five years, maintaining strong KYC/AML control, and offering APIs and WebSockets.

Lukka offers aggregate data, trade data, and derivatives analytics for digital asset markets. The aggregate data consists of reference rates calculated using methodologies like fair market value (FMV), VWAP, and TWAP, which align with International Financial Reporting Standards (IFRS) and Generally Accepted Accounting Principles (GAAP). A detailed methodology for how Lukka arrives at reference rates can be found [here](#). It also provides OHLCV candlestick data for historical trades along with the executed trades. The organization provides two types of order book data – (i) real-time data for best bid and ask, and (ii) historical data for the 10 bids and asks. Lukka offers derivatives analytics such as implied volatility and volatility surfaces that can be leveraged by funds employing quantitative trading strategies.

Nomics largely caters to retail users with its Nomics.com website, which features price data, market cap data, volume data, and charting for thousands of digital assets. It was built entirely

using the Nomics API, suitable for market research, intelligence, trade execution, and strategy backtesting. Nomics added support for live and historical order book data in 2021.

With several decades of experience in traditional finance, **Nasdaq**, in partnership with Amberdata, has begun to offer historical market data related to digital assets such as reference rates, trade data, and order book data. The order book data dates back to the beginning of 2019. Nasdaq partners with other data providers to deliver many digital asset data products such as daily price indices, hourly & end-of-day OHLCV data, and crypto asset ratios and sentiment data via Nasdaq Data Link.

The following subsection outlines key on-chain data providers and their service offerings.

On-chain Data Providers

Blockchain explorers like Etherscan are useful for analyzing on-chain activity on an ad-hoc basis. However, extracting time series of various data parameters from Etherscan can be burdensome. This is where data providers with services tailored towards institutions can be helpful, as they add structure to on-chain data.

The table below outlines the on-chain data offerings of several leading providers and the blockchains they support. As previously mentioned, on-chain data spans three general categories: (i) network data which captures the usage of layer-1 and layer-2 networks, (ii) DeFi data which captures the usage of different decentralized applications built on top of layer-1 and layer-2 networks, and (iii) NFT data which captures activity related to the issuance and trading of NFTs. Ethereum is the most widely supported network covered by on-chain data providers.

The On-Chain Data Provider Landscape

Organization	Est.	On-Chain Data Products	Select Chains Supported								
			Arbitrum	Avalanche ⁽²⁾	Bitcoin	BNB Chain	Ethereum	Optimism	Polygon	Solana	Tron
Amberdata ⁽²⁾	2017	Network, DeFi	×	✓	✓	✓	✓	×	✓	✓	×
Coin Metrics	2017	Network	×	×	✓	✓	✓	×	×	×	✓
CryptoSlam	2018	NFT	✓	✓	×	✓	✓	×	✓	✓	×
DappRadar	2018	DeFi, NFT	✓	✓	×	✓	✓	✓	✓	✓	×
Debank	2018	DeFi	✓	✓	×	✓	✓	✓	✓	×	×
Dune	2018	Network, DeFi, NFT	✓	✓	✓	✓	✓	✓	✓	✓	×
Flipside Crypto	2017	Network, DeFi, NFT	✓	✓	×	✓	✓	✓	✓	✓	×
Glassnode	2017	Network	×	×	✓	×	✓	×	×	×	×
icy.tools	2021	NFT	×	×	×	×	✓	×	×	×	×
Kaiko	2014	DeFi	×	×	×	×	✓	×	×	×	×
Nansen	2020	Network, DeFi, NFT	✓	✓	×	✓	✓	✓	✓	✓	×
NFTGo	2021	NFT	×	×	×	×	✓	×	×	×	×
NonFungible.Com	2018	NFT	×	×	×	×	✓	×	×	×	×
Rarity.Tools	2021	NFT	×	×	×	×	✓	×	×	×	×
Token Terminal	2020	Network, DeFi	✓	✓	✓	✓	✓	✓	✓	✓	×
Trait Sniper	2021	NFT	×	×	×	×	✓	×	×	×	×
Zapper	2020	DeFi, NFT	✓	✓	×	✓	✓	✓	✓	×	×

Notes: (1) Represents support for Avalanche C-Chain and does not account for other blockchains within the Avalanche ecosystem.

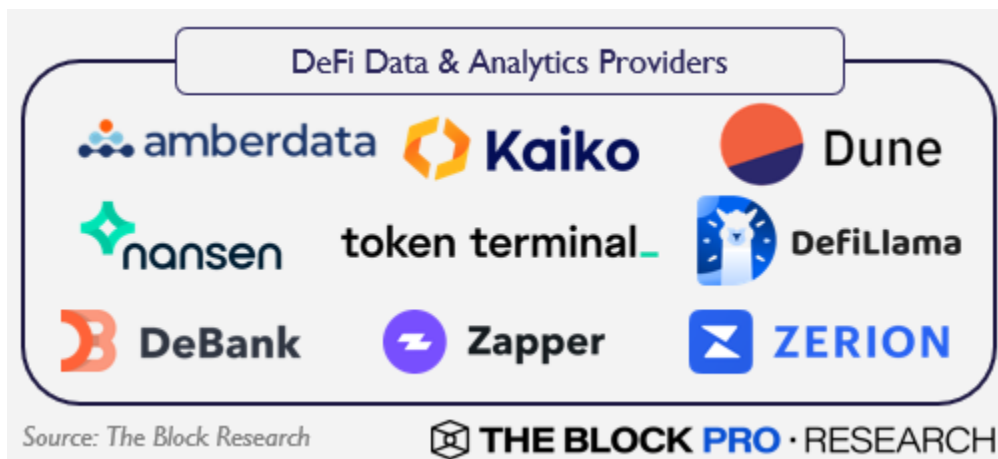
(2) Amberdata currently has beta support for BNB Chain and Solana.

Source: The Block Research

 **THE BLOCK PRO** · RESEARCH

Given recent developments in DeFi and NFT data provision, this research report focuses primarily on firms with offerings in these nascent categories. For more information on network data providers, please refer to The Block Research's [Digital Asset Data and Infrastructure: 2021](#) report.

DeFi Data & Analytics Providers



Amberdata provides institutions with comprehensive DeFi data pertaining to categories such as trading, liquidity pools, governance, and protocol-related data. Its offering takes block-by-block snapshots of DeFi protocols such as DEXes and lending/borrowing markets. In order to cover DeFi data holistically, Amberdata breaks it down into four lenses: protocol, asset/pool, wallet, and governance. The protocol lens gives an overview of a protocol using metrics including aggregate deposits (i.e., TVL), liquidations, borrowings, and repayments. The asset or pool lens retrieves information about a particular asset. The wallet lens provides data from the perspective of wallets (i.e., a particular address' interactions with individual protocols). This is helpful for accounting and taxation because it provides point estimates of address balances and their dollar values at any historical point. The governance lens provides information regarding protocol governance-related metrics such as voting power and resolution of prior votes. Amberdata provides this data for EVM-compatible chains Arbitrum, Avalanche, Ethereum, Optimism, and Polygon.

DeFiLlama's offering targets both retail and institutional participants and can be divided into two categories – dashboard and tools. The dashboard aggregates data such as liquidity pool statistics, yield overview, TVL, volumes, airdrops, and fundraises. Its data, which is fully open-sourced, spans [multiple](#) DeFi categories and [145](#) blockchains. It gives a [yield and volume overview](#) of several liquidity pools across 53 chains. Along with the pool yields, the company provides breakdowns of yield in the form of base APY (i.e., organic protocol revenue) and reward APY (i.e., inflationary token rewards). The dashboard includes DeFi liquidations and fees/revenues of protocols that are useful to analyze protocols.

Tools provided by DeFiLlama include token watchlists, trending contract tracking, and token liquidity metrics. Trending contracts show Ethereum and Polygon smart contracts that were used in recent (up to 100 minutes) transactions. Token liquidity lets users find the liquidity of a token in terms of another token and the most liquid venue for executing the trade.

Dune runs nodes on ten blockchains (Arbitrum, Avalanche C-Chain, Bitcoin, BNB Chain, Ethereum, Fantom, Gnosis Chain, Optimism, Polygon, and Solana) and stores the blockchain data in a database compatible with SQL, a programming language designed for manipulating relational databases. Dune's basic services offer powerful tools for researchers and retail

participants to flexibly query, curate, and analyze on-chain data. Dashboards created by non-paying accounts are free for everyone to view.

This flexibility, however, comes with certain costs. First, users must be proficient in SQL. Second, it demands customer awareness of on-chain data generation. For example, if a user wants to know the circulating supply of an ERC-20 token, they must know the token generating address, whether the token has a burning mechanism, etc. Third, since Dune doesn't update data in real-time it does not work well with models that take in real-time data.

In addition to this free offering, Dune also offers paid services that allow users to export the results of a query to CSV and keep queries and dashboards private.

Like Dune, **Flipside Crypto** collects data from a variety of blockchains and stores it in a SQL/Snowflake database that users can query for free. Unlike Dune, Flipside has no paid tiers – instead, protocols partner with the platform to issue bounties for specific dashboards to be queried using SQL. Users can access Flipside's API via its ShroomDK SDK²¹ by minting a ShroomDK NFT, which ties to an API key. Flipside provides data on [15 chains](#), including EVM-compatible chains such as Avalanche, Ethereum and Polygon as well as Algorand, Cosmos, and Solana.

Kaiko's DeFi data offering is geared towards institutions and covers two DeFi subsegments: DEXes and lending protocols. Lending protocols supported by Kaiko include Aave, Compound, and Cream. Kaiko's lending and borrowing data offers insight into total borrows, deposits, and APY for the three protocols mentioned above. Kaiko's DEX data consists of trade data and liquidity pool snapshots. Kaiko provides insights into all DEX trades and liquidity events (pool deposits and withdrawals). Kaiko also offers historical DeFi data that enables investors to run backtests on liquidity provisioning or trading strategies via REST APIs.

Nansen provides a global view of liquidity and staking pools across six EVM-compatible chains and a tool that shows contracts that recently received high inflows – thus allowing investors to stay up-to-date with newly deployed pools generating high interest. A 'wallet profiler' shows information about all addresses such as activity, held tokens, token-wise profit and loss, and frequent counterparties. Nansen also provides data covering inter-chain bridges with metrics such as total bridge and the number of depositors to bridge smart contracts.

Token Terminal is an analytics platform for DeFi. It monitors DeFi protocols and provides metrics such as fee earned, revenue, treasury holdings, expenses, fully diluted market capitalization, and price-to-sales ratio. These metrics allow users to evaluate DeFi protocols like companies and help compare their valuations.

DeBank, Zapper, and Zerion are DeFi portfolio tracker tools. **DeBank** allows users to track over 1,500 protocols across 39 chains. In addition to portfolio tracking, **Zapper** allows users to swap and manage their assets through its application which also provides data on liquidity pools and

²¹ SDK or Software Development Kit is a set of tools such as compiler, libraries, debuggers, etc. that allows a developer to build applications faster and in a more standardized manner.

staking opportunities across ~150 DeFi protocols. Like Zapper, **Zerion** allows users to manage their portfolio by letting them trade, lend, and borrow from the app with over 10 chains and 500 protocols.

Case Study: DeFi DEX Data

While DeFi data spans a wide range of disciplines, data related to DEXes has emerged as an area of focus for institutional and retail market participants alike. As displayed in the table below, Amberdata, DeFiLlama, Kaiko, and Nansen have all launched data products that provide varying levels of visibility into the emerging DEX landscape.

Case Study: DEX Data

Organization	Supported DEXes (#)	DEX Data Offerings				
		APR Monitoring	Impermanent Loss Tracking	OHLCV Series	Tick-Level Data	TVL Metrics
Amberdata	6 to 10	✓	✓	✓	✓	✓
DeFiLlama	200 to 250	✓	×	×	×	✓
Kaiko	1 to 5	×	×	✓	✓	✓
Nansen	50 to 100	✓	×	×	×	✓

Notes: Sample set of firms chosen based on availability of DEX data documentation. In addition to the firms listed in this table, Dune Analytics and Flipside Crypto facilitate DEX tracking via their user-generated dashboards

Source: Company documentation, The Block Research

 **THE BLOCK PRO** · RESEARCH

Supported DEXes: The quantity of DEX protocols supported by individual data providers varies meaningfully. Institutionally-focused data providers, such as Amberdata and Kaiko, have concentrated their offerings around the most widely used protocols. Both of these firms' DEX data covers Balancer, Curve, Sushiswap, and Uniswap (V2, V3) – which collectively account for the bulk of DEX trading volume. Retail-oriented firms such as DeFiLlama and Nansen have DEX data offerings that cover the longer tail of protocols, including those with relatively low levels of adoption, yet provide far less granular data.

APR Monitoring: DEX liquidity providers (LPs) earn token rewards (i.e., user trading fees and native token rewards) in return for committing capital to facilitate DEX trading. As displayed in the table above, the majority of providers publish annual percentage rate (APR) metrics which approximate the gross amount of yield LPs have earned relative to the total amount of capital that they have committed to liquidity pools.

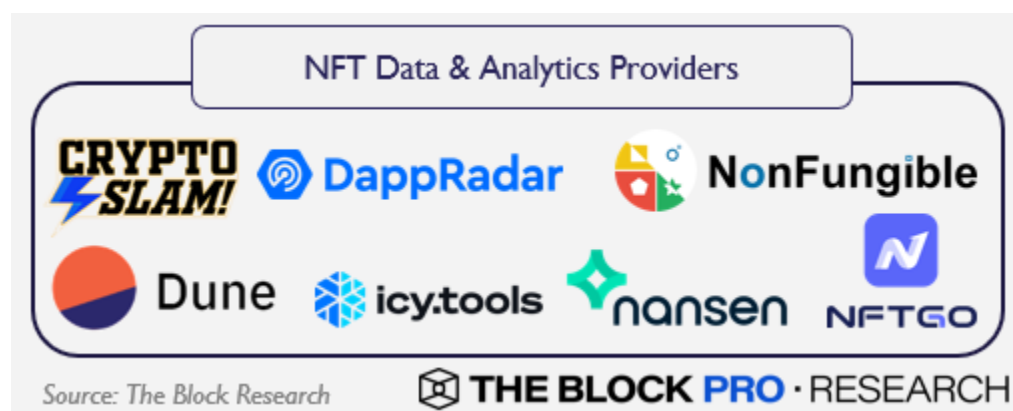
Impermanent Loss Tracking: While APR statistics approximate the gross amount of token rewards paid to LPs, they do not take into account [impermanent loss](#). As displayed in the table above, providers such as Amberdata have specialized in calculating impermanent loss across LP positions to provide market participants with a holistic view of LP returns that takes into account the implicit costs that come with providing liquidity.

OHLCV and Tick Level Data: Similar to CEX market data, DEX price and volume data are collected and aggregated across trading pairs and DEXes. Institutionally-focused firms Amberdata and Kaiko provide metrics that enable customers to track trading data at not only the aggregate level (i.e., OHLCV) but also at a granular trade level across the DEXes they support.

TVL Metrics: As displayed in the chart above, all DEX data providers support TVL metrics. Total Value locked allows users to analyze the liquidity for different trading pairs (and by extension implicit transaction costs such as slippage) across different DEXes and different layer-1 and layer-2 blockchains.

NFT Data and Analytics Providers

While DeFi has emerged as a major use case of layer-1 and layer-2 networks, NFTs had a breakthrough year in 2022 and have captured the interest of institutional and retail investors, alike. A new class of firms focused primarily on NFT data has emerged.



The following table compares leading NFT data and analytics providers across a range of metrics pertaining to trading volume, price, and other variables such as rarity scores.

Comparison of NFT Data & Analytics Providers

Organization	Supported Chains (#)	Historical Data			User Tools					Other Data		
		Floor Prices	Mint Data	Trade Data	API Alerts	Portfolio Availability	Rarity Tracking	Upcoming Score	Marketplace Mints	Marketplace Data	Ownership Distribution	Research Insights
CryptoSlam	20	×	×	✓	×	✓	×	×	✓	×	×	×
DappRadar	23	×	×	✓	×	×	✓	×	✓	✓	×	✓
Dune	4	✓	✓	✓	×	×	×	×	×	✓	✓	×
Flipside Crypto	13	✓	✓	✓	×	×	×	×	×	✓	✓	×
icy.tools	1	✓	✓	✓	✓	✓	×	×	✓	×	✓	×
Nansen	5	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
NFTGo	1	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
NonFungible.Com	3	×	✓	✓	×	✓	✓	×	×	×	×	✓
Rarity.Tools	1	×	×	×	×	×	×	✓	✓	×	×	×
Trait Sniper	1	×	×	✓	×	✓	×	✓	×	×	✓	×

Data through 01/19/2022

Source: Project/Company Documentations, The Block Research

THE BLOCK PRO · RESEARCH

CryptoSlam and **DappRadar** focus on aggregating NFT data across 15 and 23 blockchains, respectively. They showcase historical trading data of various NFT collections such as the number of buyers and sellers, volume, and unique active addresses. While DappRadar may not offer some of the more granular data products or APIs of other providers, it provides NFT trade data, upcoming mint info, portfolio tracking, and marketplace data across the largest number of chains. CryptoSlam covers the second highest number of chains at 19, but also offers an API that provides NFT data aggregation across supported chains.

Just as with other on-chain data, **Dune** and **Flipside Crypto** enable their communities to write SQL queries to generate NFT-related dashboards. Accordingly, metrics tracked across individual NFT collections and trading venues vary based on community participation for providers such as Dune and Flipside.

Nansen's offering includes NFT market overview data (market cap of all collections, average price, floor price, number of wallets holding a particular collection), latest mints, and undervalued NFTs based on their traits. The company also provides information including NFT exchange-related data, NFT holder distribution, rarity scores, breakdown of NFT mints, and sale listings by wallet types²².

Nansen offers a suite of unique NFT analytic tools. The platform's 'NFT Paradise' dashboard provides users with the mints of the latest NFT collections sorted by the number of minters in the previous 15 minutes, NFT transactions for addresses labeled as smart money based on their historical performance, and an NFT Sniper section which tracks the top 5 undervalued listings from each NFT collection on OpenSea. Nansen's 'NFT Item Profiler' and 'Rarity Profiler' allow users to view the trade history, rarity, price estimates, and the rarity of particular features of a specific NFT within a collection. Perhaps most uniquely, Nansen has created 6 NFT Indexes to track NFT market activity across the top 500 Ethereum NFT collections, the top 10 NFT collections, the top 100 social NFT collections, the top 50 gaming NFT collections, and the top 20 metaverse-related collections.

Icy.Tools and **NFTGo** provide NFT data such as ownership distribution, historical floor prices of collections, and historical trade and mint data. **Rarity.Tools** and **Trait Sniper** specialize in deconstructing metadata to capture different traits of NFTs in several collections and score NFTs based on the rarity of the traits within a collection. All four of the aforementioned firms provide data exclusively related to NFTs issued on Ethereum.

Although **NonFungible.com** offers historical trading and mint data readily available on its website, the company pivoted its model to on-demand data and consulting for NFT project developers, collectors, and investors. NonFungible.com delivers data for the three chains on which it runs proprietary nodes: Ethereum, Flow, and Ronin. As with market and on-chain data providers, NFT data providers tailor their offerings to the needs of different classes of market participants. While many NFT data providers create services with retail users in mind,

²² Smart Money, Smart Minter (those who minted popular collections), NFT Early Adopter, etc.

NonFungible.com offers multiple services for enterprises and institutions as well. In addition to NFT data, NonFungible.com provides businesses with existing NFT-related offerings and those looking to create new ones with go-to-market and tokenomics refining advisory services.

In summary, the following table displays the product offerings, target customers, and pricing models of all on-chain data providers analyzed in this section of the report. Notably, the vast majority of retail-focused firms provide standardized pricing for their offerings while institutionally-focused firms customize pricing based on the bespoke needs of their customers.

Comparison of Digital Asset Data Providers

Organization	Data Offering	Target Customer	Pricing Model ⁽¹⁾		
			Lowest Tier	Highest Tier	Free Tier?
Amberdata	Market Data, On-chain Data (Network, DeFi)	Institutional	Custom*	Custom	×
CoinAPI	Market Data	Institutional	\$79	\$599	✓
CoinGecko	Market Data	Retail, Institutional	\$103	\$799	✓
CoinMarketCap	Market Data	Retail, Institutional	\$29	\$699	✓
CoinMetrics	Market Data, On-chain Data (Network, DeFi)	Institutional	Custom	Custom	✓
CryptoCompare	Market Data, On-chain Data (Network, DeFi)	Retail, Institutional	\$80	\$200	✓
CryptoSlam	On-chain Data (NFT)	Retail	\$19	\$399	✓
DappRadar	On-chain Data (NFT)	Retail	\$199	\$999	✓
Dune	On-chain Data (Network, DeFi, NFT)	Retail, Institutional	\$69	\$1,337	✓
Flipside Crypto	On-chain Data (Network, DeFi, NFT)	Retail, Institutional	Free	Free	✓
Glassnode	On-chain Data (Network)	Retail, Institutional	\$29	\$799	✓
icy.tools	On-chain Data (NFT)	Retail, Institutional	Free	\$62	✓
Kaiko	Market Data, On-chain Data (Network, DeFi)	Institutional	Custom	Custom	×
Nansen	On-chain Data (Network, DeFi, NFT)	Retail, Institutional	\$116	\$2,500	✓
Nasdaq	Market Data	Institutional	Custom	Custom	×
NFTGo	On-chain Data (NFT)	Retail	\$99	\$299	✓
Nomics	Market Data	Retail, Institutional	Custom	Custom	✓
NonFungible.Com	On-chain Data (NFT)	Retail, Institutional	N/A	N/A	✓
Rarity.Tools	On-chain Data (NFT)	Retail	N/A	N/A	✓
Trait Sniper ⁽²⁾	On-chain Data (NFT)	Retail	\$125	\$125	✓

Note: (1) All of the firms listed in the table above that offer standard pricing packages also offer customized pricing for enterprises. All pricing plans are charged on monthly basis. (2) Trait Sniper prices its products in ETH (0.8 ETH). Price calculated in \$USD as of 1/31/22.

Source: Company documentation, The Block Research



The next subsection sheds light on the service offerings of digital asset data providers for forensics and market surveillance.

Forensics and Market Surveillance Providers

While the aforementioned firms' offerings span market and on-chain data, companies focused on forensics and market surveillance have also emerged as important players in the digital asset data landscape. Forensics and Market Surveillance firms source and analyze market and on-chain data, integrate this data into bespoke software solutions, and sell these products to financial institutions, regulatory bodies, and government agencies.



Blockchain Intelligence Group provides four core products. 'BitRank Verified' screens addresses and transactions, for example regarding AML compliance. 'QLUE' is a blockchain visualization engine used to analyze, identify, and track nefarious crypto activities. 'Crypto Investigations' is a platform to connect victims of scams with certified crypto investigators. 'Address Watch' is an automated crypto alert service to track addresses.

Chainalysis²³ offers a variety of products that can be of use for a wide range of activities such as monitoring an entity's activity or complying with sanctions. 'Chainalysis Reactor' identifies, labels, and visualizes addresses an individual or entity interacted with and hence can be used to monitor suspicious activities. 'Chainalysis KYT' is a compliance software for crypto businesses that can be tailored to a client's risk management and risk appetite. 'Chainalysis Cryptos' is a tool to carry out due diligence in the crypto space, such as, for example, tracking cryptocurrency businesses' on-chain activity. 'Chainalysis Business Data' is a software product that helps crypto businesses gain insights into customer behavior from engagement with their platform. 'Chainalysis Market Intel' provides investors with cryptocurrency market data. 'Chainalysis Crypto Sanctions Screening Tool' is a free sanctions compliance screening software.

CipherTrace features four core products that help users maintain compliance with AML laws. 'CipherTrace Armada' offers cryptocurrency risk and fraud controls for financial institutions. 'CipherTrace Inspector' allows clients to carry out financial investigations and blockchain

²³ More details about the company are provided in "The Block Research: [Chainalysis Company Intelligence](#)".

forensics. 'CipherTrace Sentry' is a tool for cryptocurrency anti-money laundering transaction monitoring. 'CipherTrace Traveller' helps customers comply with global travel rule regulations. CypherTrace was acquired by Mastercard in September 2021.

Elliptic²⁴ has four core products in the area of forensics and market surveillance. 'Elliptic Lens' allows for crypto wallet screening to identify the owner of a wallet and its source and destination of assets. 'Elliptic Navigator' is a tool for AML compliance monitoring. 'Elliptic Discovery' allows for due diligence and virtual asset service provider (VASP) tracking to assess financial crime risk when engaging with counterparties. 'Elliptic Investigator' provides cryptocurrency investigations to visualize and explore crypto wallets and transactions.

Eventus offers its 'Validus' platform which integrates trade surveillance, market risk, algo monitoring, and AML/transaction monitoring. In addition to its digital asset market coverage, its surveillance products span traditional equities, options, futures, foreign exchange (FX), and fixed-income markets.

Nasdaq provides its 'Nasdaq Trade Surveillance' tool for trade monitoring and crime detection. It covers monitoring illicit activities such as layering and spoofing, wash trading, and so-called 'pump and dump' schemes. As early as January 2017, the company had [seven digital asset exchanges](#) as customers for its market surveillance product.

Solidus Labs helps customers spot and report market manipulation. It also offers an "all-in-one" crypto risk monitoring suite 'Halo' to protect investors from threats in DeFi. It integrates this solution with key market players' offerings to carry out AML, bank account verification, KYC, KYT ('know your transaction'), and wallet screening.

TRM Labs²⁵ has three core products. 'TRM Forensics' allows users to trace the source and destination of crypto transactions. 'TRM Know-Your-VASP' is used for assessing risk for Virtual Asset Service Providers and other crypto businesses. 'TRM Transaction Monitoring' is used for monitoring digital asset deposits and withdrawals for AML/CFT ('Combating the Financing of Terrorism') compliance.

Verafin offers an integrated platform that provides tools for AML monitoring, fraud detection, blockchain analytics, risk management, and case management. Verafin was acquired by Nasdaq in 2021.

The following subsection provides an overview of data providers' fundraising and M&A activity.

²⁴ More details about the company are provided in "The Block Research: [Elliptic Company Intelligence](#)".

²⁵ More details about the company are provided in "The Block Research: [TRM Labs Company Intelligence](#)".

Digital Asset Data Fundraising and M&A

Fundraising and Investments

Fundraising of digital asset data providers increased significantly between 2020 and the first half of 2022. Chainalysis has led the way with \$536.6 million in total funding since its inception in 2014. In 2022, Amberdata (\$30 million Series B), Coin Metrics (\$35 million Series C), Dune (\$69 million Series B), Kaiko (\$53 million Series B), and Lukka (\$110 million Series E) all raised significant funding.

Fundraising in the Blockchain Data Provider Landscape

Organization ⁽¹⁾	Est.	HQ	Products	Latest Funding Round (in \$MN)	Latest Funding (date)	Latest Valuation (in \$MN) ⁽²⁾	Total Funding (in \$MN)
Amberdata	2017	US	Market Data, On-chain Data	\$30.0 (Series B)	May-22	\$330	\$47.0
Chainalysis	2014	US	Forensics	\$170.0 (Series F)	May-22	\$8,600	\$536.6
Coin Metrics	2017	US	Market Data, On-chain Data	\$35.0 (Series C)	Apr-22	Undisclosed	\$59.0
CryptoQuant	2020	South Korea	Market Data, On-chain Data	\$3.0 (Seed)	Aug-21	Undisclosed	\$3.3
CryptoSlam	2018	US	On-chain Data	\$9.0 (Seed)	Jan-22	Undisclosed	\$9.0
DeBank	2018	Singapore	On-chain Data	\$25.0 (Series A)	Dec-21	\$99	\$25.0
Dune	2018	Norway	On-chain Data	\$69.4 (Series B)	Feb-22	\$1,000	\$79.4
Elliptic	2013	UK	Forensics	\$60.0 (Series C)	Oct-21	Undisclosed	\$100.0
Eventus	2014	Canada	Forensics	\$30.0 (Series B)	Sep-21	Undisclosed	\$48.5
Kaiko	2014	France	Market Data	\$53.0 (Series B)	Jun-22	Undisclosed	\$83.0
Lukka	2014	US	Market Data	\$110.0 (Series E)	Jan-22	\$1,300	\$201.3
Nansen	2020	Singapore	On-chain Data	\$75.0 (Series B)	Dec-21	\$750	\$88.2
Zerion	2016	US	On-chain Data	\$12.3 (Series B)	Oct-22	Undisclosed	\$33.0

Notes: (1) Table only includes organizations with disclosed funding amounts. (2) Latest valuation represents valuation at latest funding date.

Source: Crunchbase, The Block Research

 THE BLOCK PRO • RESEARCH

TradFi institutional participation in digital asset markets has also risen dramatically over the past two years. [Citigroup Inc.](#), [Goldman Sachs](#), [Morgan Stanley](#), and [Wells Fargo](#) have invested considerable resources in blockchain and blockchain-adjacent projects, companies, and initiatives. Many of these firms have also invested directly in companies or initiatives that focus on digital asset data provision, for example, Citigroup Inc. with its [investment in Amberdata](#), Wells Fargo with its [investment in Elliptic](#), and Goldman Sachs by [partnering with MCSI and Coin Metrics](#) to create a cryptocurrency taxonomy called [Datonomy](#).

As established institutions foray into the digital asset space, they naturally look for high-quality data service providers. Chainalysis' size, history, and perceived safety were a factor in GIC's investment in its \$170 million [Series F](#) funding round in May 2022.

"Chainalysis is in a unique position of running a SaaS (software as a service) business in the cryptocurrency space due to the stability of its business model as well as the increasing demand for trust and safety in the overall industry." - Choo Yong Cheen,

Chief Investment Officer of Private Equity at GIC, (Series F funding round press release, May 2022)

How capital is used in these funding rounds largely depends on the size and strategies of the funded companies. Chainalysis' offerings are fairly established and have thus decided to use this capital to extend the reach of their services into new markets:

*"We are growing in regions such as Europe and Asia-Pacific. So far **up until now, we had basically no presence in South America, so that is another growth area for us.**" – Michael Gronager, Chainalysis CEO and Co-Founder, (Company's Series F press release, May 2022).*

Amberdata also held a funding round in May 2022. Having previously raised \$15 million in a Series A led by Citigroup, the firm raised \$30 million in its Series B led by Knollwood Investment Advisory. It intends to use this capital to accelerate its expansion and new product initiatives.

*"This financing will allow us to accelerate the expansion of our worldwide go-to-market activities, scale our industry-leading data infrastructure, and build out our world-class customer success team. It will also **allow us to accelerate new product initiatives including the expansion of DeFi depth and coverage, the development of digital asset indices, and market intelligence and risk analytics applications.** It is our intent to extend our leadership and continue to define the category for institutional digital asset data and research applications." – Shawn Douglass, CEO of Amberdata, (Series B funding round Press release, May 2022)*

Mergers and Acquisitions

Significant mergers and acquisitions deals in the data provider space have transpired since last year's data and infrastructure report. Mastercard's acquisition of CipherTrace in [September 2021](#) is an excellent example of a "web2" firm foraying into "web3" by acquiring a crypto-native company. Mastercard is no stranger to digital payments or cybersecurity; the firm has [2.9 billion cards](#) in global circulation and has launched a [Crypto Card Program](#) that allows users to spend digital currencies on everyday transactions. CipherTrace is a leading digital asset intelligence and forensics company that has historically provided cybersecurity solutions to large exchanges, banks, and government agencies.

*"We help companies – whether they are banks or cryptocurrency exchanges, government regulators or law enforcement to keep the crypto economy safe. **Our two companies [Mastercard and CipherTrace] share this vision to provide security and trust throughout the ecosystem.** We are thrilled to join the Mastercard family to scale CipherTrace's reach across the globe." – Dave Jevans, CEO, CipherTrace (Mastercard Press Release, September 2021)*

Notable Acquisitions in Blockchain Data Provision

Date	Acquirer	Sub-Category (Acquirer)	Target	Sub-category (Target)
Jan-21	NYDIG	FinTech/Financial Services	Digital Assets Data	Market Data
Apr-21	Coinbase	Trading/Brokerage	Skew	Market Analytics
Sep-21	Mastercard	Payments/Card Network	CipherTrace	Forensics
Oct-21	Chainalysis	Chain Analytics/Compliance	Excygent	Forensics
Oct-21	Lukka	Tax/Auditing	Blox Finance	Infrastructure
Nov-21	Blockdaemon	Staking/Node Infra	Anyblock Analytics	On-chain Data
Apr-22	Kaiko	Market Data/Analytics	Kesitys	Market Analytics
May-22	Quicknode	Mining/Staking Infra	Icy Tools	On-chain Data
May-22	Nansen	Data/Analytics	Ape Board	On-chain Data
Oct-22	Amberdata	Market Data/Analytics	Genesis Volatility	Market Analytics

Source: Crunchbase, The Block Research

 **THE BLOCK PRO** · RESEARCH

To expand its capability in monitoring on-chain illicit activity, Chainalysis acquired Excygent, a cybercrime investigation company that provided forensics services to government agencies, in [October 2021](#). Excygent was founded in 2018. Its expertise in cybercrime, open source analysis, and database forensics enables Chainalysis to expand the reach and quality of its offerings in supporting government investigations.

*"Cybercrime like ransomware is one of the biggest barriers to building trust in cryptocurrency. **The expertise brought by the Excygent team directly aligns with our mission to leverage the transparency of blockchains to weed out bad actors from the ecosystem and ultimately promote more financial freedom with less risk.**"* – Michael Gronager, Co-founder and CEO of Chainalysis, (Chainalysis press release, October 2021)

Kaiko, a leading market data provider, acquired quantitative analytics firm Kesitys in [April 2022](#) two months prior to its Series B. Kaiko had previously partnered with Kesitys in May 2021 to create quantitative products targeted at cryptocurrency traders and portfolio managers.

Amberdata [announced](#) its acquisition of Genesis Volatility, a data provider focused on options research and data analytics for financial institutions, in late October 2022. Genesis Volatility's offerings provide current and historical derivatives data from both CEXes and DEXes. Its set of tools is geared towards sophisticated digital asset options traders. This acquisition comes shortly on the heels of Amberdata's \$30 million Series B in May 2022. The integration of Genesis Volatility's data analytics capabilities into Amberdata's existing offerings is intended to provide financial institutions with a holistic view of DeFi.

*“If you look at TradFi as a roadmap to where crypto is going to go, derivatives rule the world in the rest of financial services. **In crypto specifically, options are so important because there are many regulated US entities that are uncomfortable holding crypto, perhaps because of regulatory concerns, or they're worried about getting hacked** So they're creating structured products and transacting in options because they're comfortable with those and their status as a derivative means that they don't have to deal with custody. Genesis Volatility has built the best analytics platform for vol surfaces, skews, and correlations ... it brings us an analytics platform to start to bring our deep insights into DeFi into on-chain markets to provide market intelligence.”* – Shawn Douglass, CEO of Amberdata, (In an interview with The Block, October 2022)

02/16/23

The State Of Digital Asset Data And Infrastructure:
2023 Edition

3

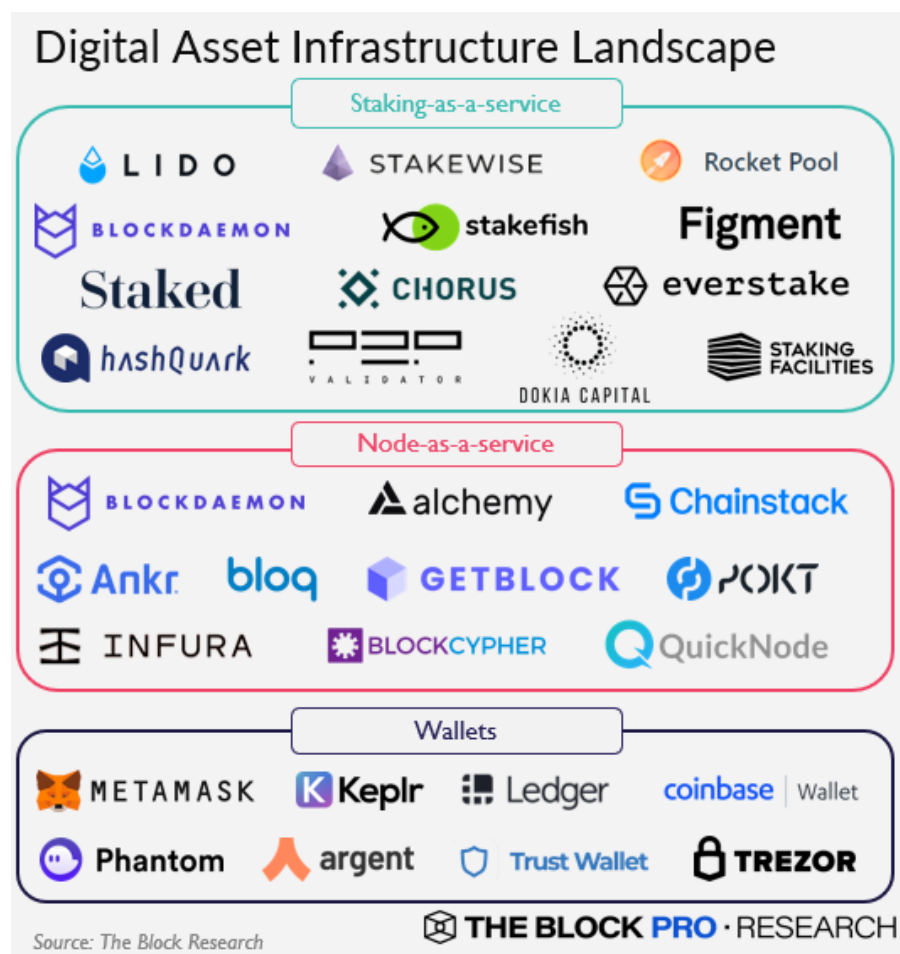
Digital Asset Infrastructure Providers

Part 3: Digital Asset Infrastructure Providers

“Infrastructure creates the form of a city and enables life to go on in a city.” – Paul Goldberger, Author and 1984 Pulitzer Prize winner (May 2009)

Digital asset infrastructure is the bedrock upon which all digital asset activity occurs. Just as the roads, buildings, bridges, and railways of a city shape the lives of its residents, digital asset infrastructure shapes a blockchain’s functionality, the applications on top of it, and how users interact with it. Businesses do not typically build the physical infrastructure that supports their operations; in much the same way, it is often impractical for developers of digital asset projects to construct the digital infrastructure necessary for their activities. Delegating their infrastructure needs to specialized providers allows them to concentrate on their core competencies.

This section focuses on node-as-a-service, staking-as-a-service, and wallet solutions. The following figure displays the key players that are covered.



While staking-as-a-service and node-as-a-service providers both offer node-related products, it is important to note the difference between the functionality of the nodes they provide. Staking-as-a-service providers facilitate access to PoS validator nodes, which execute consensus protocols, verify transactions, and append transactions to the blockchain. These nodes are integral to upholding the security of the blockchain, and products that provide access to them are useful for individuals and institutions who seek to earn the block rewards that incentivize entities to participate in maintaining this security.

In contrast to staking-as-a-service firms, node-as-a-service firms provide RPC nodes that allow developers to access blockchains. Unlike the validator nodes operated by staking-as-a-service firms, RPC nodes typically do not participate in consensus and earn block rewards. Instead, they focus on submitting new transactions and field requests to the blockchains they are a part of. RPC nodes are of particular importance to dApps that field a high volume of transactions and requests to operate at sufficient speed and scale, as well as to data analytics companies that purely seek to query blockchains for information. RPC and validator nodes are effectively identical in their architecture; however, it is most practical for a node operator to separate them, as splitting resources between them could prevent a node from doing either adequately.

Staking-as-a-service Providers

As displayed in the chart below, leading staking-as-a-service firms have emerged as major participants in the PoS landscape. Several firms stake hundreds of millions of dollars worth of PoS assets, while others now stake billions. The subsection below provides an overview of these firms and their product and service offerings.

Staking-as-a Service Provider Overview

Staking Provider	Target Userbase	Supported Protocols (#)	Total Stake Managed (\$) ⁽¹⁾
Blockdaemon	Retail, Institutional	60	\$1 - \$10 billion
Chorus One	Retail, Institutional	34	\$500 million - \$1 billion
Dokia Capital	Retail, Institutional	6	\$100 million - \$499 million
Everstake	Retail, Institutional	42	\$1 billion - \$10 billion
Figment	Institutional	47	\$100 million - \$499 million
HashQuark	Retail, Institutional	31	\$100 million - \$499 million
Lido	Retail, Institutional	5	\$1 - \$10 billion
p2p.org	Retail, Institutional	43	\$1 - \$10 billion
Rocketpool	Retail, Institutional	1	\$100 million - \$499 million
Staked	Institutional	26	\$500 million - \$1 billion
Stakefish	Retail, Institutional	23	\$1 - \$10 billion
Stakewise	Retail, Institutional	1	\$100 million - \$499 million
Staking Facilities	Institutional	9	\$100 million - \$499 million

Notes: Represents estimated values as of 12/31/2022.

Source: Company Documentation, stakingrewards.com, The Block Research

 **THE BLOCK PRO** · RESEARCH

Blockdaemon provides five staking-related products: Staking API, public validators, whitelabel validators, custodian staking, and liquid staking. The API for staking providers allows their clients to automate staking transaction signing and scale up validator nodes on PoS networks over time. Retail or institutional users can delegate their tokens to public validators run by Blockdaemon and earn passive rewards. Institutional custodians such as Coinbase and SEBA Bank use Blockdaemon's staking integration through which they offer staking products to their customers. Under its liquid staking product, Blockdaemon partnered with Stakewise to offer institutional-grade, KYC-compliant staking solutions so that institutions have access to liquidity while staking their assets.

Chorus One's whitelabel staking solution offers clients (with assets exceeding \$10 million) managed nodes for their users to stake with. The firm also provides liquid staking services which provide access to liquid staking derivatives such as Lido's stETH and stSOL. Chorus One's staking API for institutions, OPUS, simplifies staking while providing guarantees against slashing, double-signing, and security breaches.

With ~\$6.2 billion in staked assets, **Everstake** is one of the largest staking service providers. It manages its validator nodes across over 40 networks and allows users to delegate their tokens. With the help of liquid staking protocols such as Lido, Everstake lets users take advantage of

staking derivatives. In November 2022, Everstake [announced](#) a strategic partnership with Nexus Mutual, a decentralized insurance protocol, to insulate users from the risks associated with Ethereum staking.

Figment is an institutional staking service provider. It manages staking infrastructure for clients and offers staking-related services such as staking dashboards with analytics, diligence reports on PoS networks, analyst calls, and integrations with third-party custody solutions. Figment has partnered with Liquid Collective, a liquid staking protocol with multi-chain capabilities, to offer staking derivatives such as LsETH. Figment also provides insurance against downtime, double signing, and missed rewards.

HashQuark is the web3 infrastructure-focused subsidiary of Hashkey Group, a digital asset exchange and investment platform. HashQuark runs public validators of all major PoS networks and users can delegate their stake to these validators. The organization has built a solution called EtherPocket Pro, a node validation software that offers staking in the form of dedicated or shared validators to institutional and individual users. It also offers insurance against slashing and low activity.

Lido is the largest ETH staking service provider - approximately 30% of all staked ETH is staked through Lido. Lido's flagship ETH staking product stakes ETH on behalf of users and issues them a liquid staking derivative (stEth) in return. Although Lido launched with support solely for Ethereum, it has added support for KSM (Kusama), DOT (Polkadot), MATIC (Polygon), and SOL (Solana). Lido operates as a DAO and is governed by its native token, LDO.

P2P Org offers non-custodial staking infrastructure for individuals, funds, exchanges, and custodians. The company's products currently support staking for ~40 PoS networks. Its whitelabel node product enables individuals, custodians, and exchanges to avail themselves of having to operate high-performance PoS nodes. P2P also offers insurance against slashing.

Rocketpool is an Ethereum-focused liquid staking derivative protocol. It offers two kinds of ETH staking: i) "stake and run node" allows users to stake and run their own validator nodes and ii) "stake ETH" lets users stake while Rocketpool manages the end-to-end infrastructure. Like Lido, Rocketpool also has its governance token, RPL. When users want to run their validator nodes, Rocketpool allows them to do so with a minimum of 16 ETH along with RPL instead of the standard 32 ETH requirement. It does so by assigning the other 16 ETH from the pool where users, who do not wish to run validator nodes, deposit their ETH.

Staked offers staking services to its institutional clients for over 25 PoS networks. The non-custodial nature of Staked's products means that users maintain control over their assets and can have a say in the governance process of different networks via staked tokens. Staked also offers ETH staking in the form of a trust for accredited investors.

Stakefish manages end-to-end infrastructure while allowing users to maintain custody of their assets. The primary drawback of this model is that users must stake at least 32 ETH if they want

to use Stakefish's services to leverage Ethereum staking. Although Stakefish supports Lido (i.e., operates validator nodes on behalf of Lido), it does not offer staking derivatives when ETH is staked via Stakefish.

Like Rocketpool, **Stakewise's** products exclusively support ETH staking. Stakewise allows users to stake any amount of ETH and offers them liquid staking derivatives similar to Lido and Rocketpool. Stakewise's staking derivative serves as the underlying for Blockdaemon's liquid staking product, meaning that when users elect to use Blockdaemon's liquid staking services, they receive Stakewise's ETH derivative.

Staking Facilities is operationally similar to Stakefish in that the firm runs the infrastructure and users delegate tokens to their validators. However, instead of enforcing the 32 ETH minimum requirement for Ethereum staking, Staking Facilities uses Lido's liquid staking derivative, which allows for the staking of any denomination of ETH.

Node-as-a-service Providers

While node-as-a-service providers may seem fairly similar on the surface, as displayed in the table below, there is significant variation in their products and services, as well as their pricing models.

Comparison of Node-as-a-Service Providers

Organization	Products/Services						Pricing Model ⁽¹⁾		
	Supported Chains (#)	Debug and Trace APIs	Dedicated Nodes	NFT APIs	Shared Nodes	Staking Support	Fee Tiers (#)	Lowest Tier (\$)	Highest Tier (\$)
Alchemy	7	✓	✓	✓	✓	×	3	\$49	Custom
Ankr	15	×	✓	✓	✓	✓	NA	NA	NA
Blockcypher	5	✓	×	×	✓	×	3	\$100	Custom
Blockdaemon	60+	✓	✓	✓	✓	✓	NA	Custom	Custom
Bloq	6	×	✓	×	✓	✓	NA	NA	NA
Chainstack	15+	✓	✓	✓	✓	×	4	\$49	\$990
GetBlock	40+	×	✓	×	✓	×	4	\$6	\$500
Infura	5	✓	×	✓	✓	×	5	\$50	\$1,000
Pocket Network	15	×	✓	✓	✓	×	NA	NA	NA
Quicknode	13+	✓	✓	✓	✓	×	4	\$49	\$299

Notes: - (1) All of the firms listed in the table above that offer standard pricing also offer customized pricing for enterprises. All pricing plans are charged on monthly basis.

Source: Company websites and documentation, The Block Research

 **THE BLOCK PRO** · RESEARCH

Alchemy is one of the leading infrastructure providers targeting developers in the digital asset ecosystem. Alchemy's Supernode is an enhanced Ethereum API with all the functionality of a node with high reliability, correctness, and scalability. Two of Alchemy's standout products are Transact and Spearmint. Transact is a smart transaction product that allows sending transactions with Flashbot's private relay to prevent arbitrage and frontrunning. Spearmint is a tool that allows NFT creators and developers to create allow lists. Allow lists enable NFT projects to manage their

community better by easily letting them set the rules that dictate address eligibility for NFT mints. For example, they can mandate that only addresses linked with discord profiles with more than a certain level of experience points can mint the NFT at a specific price.

Ankr aims to address the node-level centralization problem for blockchains. Ankr incentivizes individual node operators to run nodes via its token, ANKR, which is also the governance token of the network. Users pay fees in ANKR and can stake tokens to secure the network. Individual node providers elect to use their own hardware or Ankr's servers and must stake 50,000 ANKR (~\$1,300 as of 1/31/22) as collateral to offer their services to the network. The protocol also compensates node providers for their services using the ANKR token. Ankr has three categories of product offerings: node access for developers and enterprises, staking, and a game development kit compatible with Unity and Unreal engine that enables integration with web3 wallets.

BlockCypher focuses on providing infrastructure for developers. The company offers APIs through which developers can interact with the Bitcoin, Ethereum, Litecoin, Dash, and Doge blockchains. The company also operates a private test chain that closely resembles the Bitcoin main chain. BlockCypher's Bitcoin APIs host several unique features, including i) an unconfirmed transaction confidence factor that can accurately predict the probability of an attempted double-spend succeeding against a transaction, ii) push APIs that allow developers to quickly retrieve information regarding when a transaction was confirmed or relayed through the network, and iii) an address forwarding API that enables seamless payment acceptance and consolidation without requiring users to create new accounts. The company's Ethereum APIs are slightly less extensive than its APIs for Bitcoin and other UTXO blockchains; however, these Ethereum APIs still contain analogs to the aforementioned features excluding a transaction confidence factor. BlockCypher has also built block explorers for the five chains it supports.

Blockdaemon is among the leading infrastructure builders with a comprehensive product set that targets a broad range of potential clients, including developers, exchanges, banks, and custodians. Blockdaemon's node-related products span three verticals: access, connect, and secure. The 'access' vertical comprises products including the Ubiquity API suite, Native and NFT API, and dedicated nodes. Ubiquity allows users to query blockchain nodes and broadcast transactions without running nodes. It uses a single, unified Blockchain API format to access data across chains. With the help of Gem, a company that Blockdaemon acquired, the 'connect' vertical consists of a fiat on-ramp. In conjunction with its acquisition of Sepior, Blockdaemon offers an MPC wallet²⁶ under its 'secure' vertical.

Bloq's node infrastructure offerings are twofold: connecting APIs and nodes. Bloq Connect is a set of autoscaling APIs that support Algorand (Mainnet and Testnet) and Ethereum (Mainnet and Goerli Testnet). This offering provides developers with a suite of REST and WebSocket APIs, the quick retrieval of blockchain data, and security monitoring for suspicious block activity. 'Bloq

²⁶ MPC wallets, like multi-sig wallets, involve multiple parties. A key difference between them is that multi-sig wallets require more than one private key to authenticate a transaction, whereas MPC wallets split up one private key among multiple parties.

Nodes' allows customers to spin up private enterprise-grade node clusters to access Algorand, Avalanche, Bitcoin, Bitcoin Cash, Ethereum, and Litecoin.

Chainstack offers managed blockchain nodes and API services. One of the primary attractions of Chainstack's offering is the customizability it provides to developers. Chainstack's node hosting service enables clients to choose from cloud providers such as Amazon, Google, Chainstack, or their own cloud setup with a choice of the geographical location of the nodes. It also offers a global mempool view for enterprises. Regarding API provision, Chainstack offers NFT and IPFS APIs. The NFT API is standardized across smart contracts, which allows developers to avoid complex programming. The IPFS API allows for quick data retrieval and stores data across multiple Storj nodes. It complies with S3 (Simple Storage Service), allowing easy movement between S3 and IPFS. Finally, Chainstack also provides Subgraphs, an indexing²⁷ solution that improves sync times.

GetBlock provides three node services: shared, dedicated, and node clusters. Although Getblock hosts nodes, it also allows clients to set up their own hardware at their preferred location. GetBlock also provides dApp developers with a technical stack that helps them quickly integrate with web3.

Infura is arguably the most widely used Ethereum infrastructure provider. MetaMask, another ConsenSys product and perhaps the most popular web3 wallet, connects to Infura endpoints by default. Infura's web3 infrastructure-as-a-service suite comprises three broad services- i) access to Ethereum's archive data, ii) easy smart contract deployment, and iii) transaction management. Infura's APIs allow developers to connect to the Ethereum network and IPFS storage. Full compatibility with all IDEs²⁸ helps developers in building smart contracts. Infura also manages transactions to be published on the Ethereum network. It reduces the drop rate of transactions by using a dynamic gas price escalation algorithm and real-time fee adjustments to ensure that they are processed quickly. It also offers flexible payment schedules to institutions so that they can transact on Ethereum without holding ETH.

A blockchain ecosystem is only as decentralized as its most centralized component. Node infrastructure centralization has been a pain point for blockchains. Like ANKR, **Pocket Network** launched as an independent, application-specific blockchain solution trying to solve the node centralization problem by incentivizing individual node providers with its token, POKT, to join the network. Pocket Network created a marketplace for dApps and relay nodes. dApps send RPC requests to a global network of Pocket Network nodes which relay the blockchain data back to dApps to earn POKT. To ensure that a node's incentives are aligned with the network, every node must maintain a minimum stake of 15,000 POKT (~\$900 as of 1/31/22). Node providers can manage their own infrastructure or use [node hosting services](#) such as BlockSpaces, BlokHub, CODR3, and Chainstack.

²⁷ In the context of database management, Indexing is a method of optimizing a database by minimizing the number of disk accesses required to process a query.

²⁸ IDE (Integrated Development Environment) refers to a software package that consolidates basic tools required for writing and testing software.

Quicknode has three APIs – Core, Token, and NFT. Quicknode's 'Core' API offers developers read and write access to blockchains. This is useful for real-time analytics, access to archival data of smart contracts and balances, and tracing and debugging transactions. Quicknode's 'Token' API allows users to easily retrieve ERC-20 token metadata²⁹ without indexing. Finally, the firm's 'NFT' API allows for the retrieval of metadata of Ethereum and Solana NFTs and can be used by galleries to display NFTs and filter them by traits, creator, and collection.

The table below shows node-as-a-service firms' client tools.

Digital Asset Infrastructure Providers' Client Tools

Organization	Client Tools				
	Dashboard	JSON-RPC	Rest API	SDK	WebSocket
Alchemy	✓	✓	✓	✓	✓
Ankr	✓	✓	✓	✓	✓
BlockCypher	✓	✓	✓	✓	✓
Blockdaemon	✓	✓	✓	✓	✓
Bloq	✓	✓	✓	✓	✓
Chainstack	✓	✓	✓	✓	✓
Getblock	✓	✓	✓	✗	✓
Infura	✓	✓	✓	✗	✓
Pocket Network	✓	✓	✓	✓	✗
Quicknode	✓	✓	✓	✗	✓

Source: Company documentation, Crunchbase, The Block Research

 **THE BLOCK PRO** · RESEARCH

Wallet Providers

While node-as-a-service providers make it easy for development organizations to deploy and manage their dApps, wallet providers allow users to tap into these applications by submitting and receiving transactions on layer-1 and layer-2 networks.

²⁹ Extracting ERC-20 tokens' metadata by querying the blockchain is not always straightforward. Indexing this data may reduce these efforts.

Overview of Leading Digital Asset Wallets

Wallet Overview					Wallet Features Comparison				
Wallet Name	Supported Chains (#)	Estimated Users (#) ⁽¹⁾	Key Storage	Revenue Drivers ⁽²⁾	Fiat On-Ramp	In Wallet Swaps	MultiSig Support	NFT Display	Staking/DeFi Support
Coinbase Wallet	5	1 to 5 million	Hot	Swap Fees	✓	✓	✗	✓	✓
Exodus	6	1 to 5 million	Hot	Affiliate Revenue	✓	✓	✗	✓	✗
Gnosis Safe	9	N/A	Hot	Swap Fees	✗	✓	✓	✓	✓
Keplr	50+	500k to 1 million	Hot	Affiliate Revenue	✗	✗	✗	✗	✓
Ledger	50+	1 to 5 million	Cold	Hardware Sales	✓	✓	✓	✓	✓
MetaMask	50+	> 10 million	Hot	Swap Fees	✓	✓	✗	✗	✗
Phantom	1	1 to 5 million	Hot	Swap Fees	✓	✓	✗	✓	✗
Trezor	50+	N/A	Cold	Hardware Sales	✓	✓	✓	✗	✗
Trust Wallet	65	> 10 million	Hot	Affiliate Revenue	✓	✓	✗	✓	✓
Argent	3	50k to 100k	N/A	Swap Fees	✓	✓	✓	✓	✓

Notes: (1) User count estimates based on downloads counts for hot wallets and hardware device sales for cold wallets. (2) Swap Fees for individual providers are as follows: 1.00% fee on all swaps (Coinbase Wallet), 0.01% fee on all trades (Gnosis Safe), 0.875% fee on all swaps (MetaMask), 0.80% fee on all swaps (Phantom). All metrics relate to retail-facing non-custodial wallets and do not account for solutions tailored to institutional custody.
Source: Company websites and documentation, The Block Research



Argent Wallet is an Ethereum-based wallet with social recovery and multisignature security. Because Argent is a smart contract wallet, users can carry out transactions or set access parameters that are not possible using EOAs. For example, Argent allows users to set guardians (such as a hardware or software wallet or a trusted third party) that can authorize transactions, lock, or recover wallets without giving them access to the assets within or using a seed phrase. Argent generates revenue through a 0.5% crypto-to-crypto exchange fee for in-wallet swaps through its integrated DEX aggregator, Paraswap.

With the exception of multisig support, **Coinbase Wallet** offers all relevant wallet features and supports all EVM-compatible chains. The wallet features a dApp browser that allows users to access various DeFi applications such as Aave, Compound, PancakeSwap, Trader Joe, and Uniswap from within the wallet application. Coinbase wallet charges a 1% fee on all in-wallet swaps.

Exodus offers three kinds of wallets: browser, desktop, and mobile. Exodus's wallets allow users to stake multiple PoS coins and have an integration with Compound that enables users to lend assets without leaving the wallet application.

Keplr Wallet allows users to access all blockchains within the Cosmos ecosystem. It also provides users with staking on various chains such as Cosmos Hub, Secret Network, and Osmosis. It generates revenue by charging projects that want native integrations with the wallet.

With [over 20 million](#) monthly active users, **MetaMask** is one of the leading browser wallets and supports all EVM-compatible chains. MetaMask allows in-app swaps by aggregating prices from different DEXes. While MetaMask does not receive any network gas fees for the swap, it levies a

service fee of 0.875% on each in-app swap transaction. MetaMask also provides an [institutional](#) wallet in partnership with pure-play custody firms.

Phantom was among the first user-friendly Solana-based wallets. For its swap feature, Phantom fetches the best execution price after scanning multiple DEXes. Apart from just viewing NFTs in the app, Phantom allows users to burn unwanted NFTs and earn rewards in SOL. Phantom charges 0.85% transaction fees on in-app swaps. Phantom has also announced its intentions to [add support](#) for Ethereum and Polygon.

Safe (formerly Gnosis Safe) allows users to create contract addresses that are smart contracts capable of triggering transactions after a set number of valid owner signatures are met. It allows owners to use on-chain and off-chain signatures. With on-chain signatures, every signature is recorded on-chain and consumes gas. The off-chain signature flow allows owners to store transactions after signing only call methods to execute transactions upon submitting a final signature. Unlike other wallet providers, Safe does not have any formal revenue drivers. Instead, Safe is governed by SafeDAO through SAFE, an ERC-20 token. Potential value capture mechanisms for the SAFE token can be implemented at the discretion of the SafeDAO.

Trust Wallet facilitates in-app swaps [via 1inch's API](#). 1inch is a DEX aggregator that eliminates the need for integration with multiple exchanges. Similar to Keplr, Trust Wallet also generates revenue through native wallet integrations with third parties.

Hardware wallets like **Ledger** and **Trezor** are non-custodial cold wallets and considered to be a very safe way for users to store their assets. With these wallets, users need a physical (device-level) verification every time they want to initiate a transaction. These wallet providers generate revenue by selling hardware devices, charging in-app transaction fees, and providing institutional custody solutions.

Digital Asset Infrastructure Fundraising and M&A

Fundraising and Investments

Leading blockchain infrastructure firms announced significant funding rounds in the first half of 2022. Even on the heels of market turmoil in the back half of 2022, many institutions remained positive regarding the long-term impact of blockchain technology.

"If we are entering 'crypto winter,' it's unlike the bear markets we've seen before. The crypto market today has institutional adoption. They see the promise crypto holds. Many institutions are long-term bullish on the tech." – Konstantin Richter, Founder and CEO at Blockdaemon (Blockdaemon press release, February 2022)

Accordingly, several institutions have begun offering digital asset infrastructure products in-house. For example, BNY Mellon [launched](#) digital asset custody for BTC and ETH in October 2022 and Nasdaq [announced](#) that it intends to launch a digital asset solution in September 2022.

As discussed later in this section, many other institutions have adopted “picks and shovels” investment strategies and acquired ownership stakes in leading blockchain infrastructure companies.

The following table displays information on the most recent fundraising rounds across blockchain infrastructure and staking.

Fundraising in the Blockchain Infrastructure Provider Landscape

Organization ⁽¹⁾	Est.	HQ	Products	Latest Funding (in \$MN)	Latest Funding (date)	Latest Valuation (in \$MN) ⁽²⁾	Total Funding (in \$MN)
Alchemy	2017	US	Node-as-a-service	\$200 (Series C)	Feb-22	\$10,200	\$564
Argent	2017	UK	Wallets	\$40 (Series B)	Apr-22	Undisclosed	\$56
Blockdaemon	2017	US	Staking/Node-as-a-service	\$207 (Series C)	Apr-22	\$3,300	\$431
ConsenSys ⁽³⁾	2014	US	Node-as-a-service, Wallets	\$450 (Series D)	Feb-22	\$7,000	\$725
Figment	2021	Canada	Staking-as-a-service	\$110 (Series B)	Dec-21	\$1,300	\$165
Lido	2020	Cayman Islands	Staking-as-a-service	\$24 (Venture Round)	Sep-22	Undisclosed	\$167
Phantom	2021	US	Wallets	\$109 (Series B)	Jan-22	\$1,250	\$118
Pocket Network	2017	US	Node-as-a-service	\$10 (Venture Round)	Jan-22	Undisclosed	\$20
QuickNode	2017	US	Node-as-a-service	\$35 (Series A)	Oct-21	Undisclosed	\$41

Notes: (1) Table only includes organizations with disclosed funding amounts. (2) Latest valuation represents valuation at latest funding date. (3) ConsenSys provides node-as-a-service and wallet products through Infura and MetaMask, respectively.
Source: Crunchbase, The Block Research

 **THE BLOCK PRO** • RESEARCH

Alchemy announced a funding round in [October 2021](#) in which the firm raised \$250 million at a \$3.5 billion valuation. Just four months later, in February 2022, Alchemy announced a \$200 million “Series C-1” round which was conducted at a \$10.2 billion valuation. The Series C-1 was led by Silver Lake and Lightspeed Venture Partners. While Lightspeed Venture Partners has invested in multiple digital asset-related companies since its inception, Silver Lake, a technology-focused private equity firm with more than \$90 billion in AUM, has historically focused on investing in “web2” technology companies.

*“Alchemy will primarily use this **funding to reinvest in the growth of the ecosystem to further the mission of making blockchain development accessible globally.** This includes expanding access to its tools for all developers, especially those just starting out, building out its educational resources to help onboard new developers to Web3, and continued strengthening of its infrastructure to support the industry’s rapid growth. It will also continue its global expansion, including the opening of new offices in and outside of the US, and support for developers across new, emerging chains.” – [Alchemy blog](#), (Series C-1 Fundraise, February 2022)*

Blockdaemon’s \$207 million Series C, which was announced in January 2022, was conducted ~\$3.3 billion valuation. Blockdaemon has stated several planned initiatives following this round, including launching a DeFi fund to invest in institutional-grade DeFi offerings and supporting strategic partnerships with StakeWise and Obol Labs, a trust-minimized staking protocol for public blockchains. SoftBank Vision Fund, the largest technology-focused investment fund in the world, participated in Blockdaemon’s Series C and ConsenSys’ Series D fundraising rounds. Both

Blockdaemon and ConsenSys have raised hundreds of millions of dollars at multi-billion dollar valuations.

In March 2022, ConsenSys, an Ethereum-focused decentralized protocol software company with a product suite including Infura, MetaMask, Codefi, Diligence, Quorum, Truffle, and ConsenSys NFT, [raised \\$450 million](#) at a \$7 billion valuation. This Series D more than doubled the valuation set in ConsenSys' Series C in November 2021. ConsenSys has discussed its intentions to use this capital injection to hire over 600 new employees globally and to fund an expansion and redesign of MetaMask.

"I think of ConsenSys as a broad and deep capabilities machine for the decentralized protocols ecosystem, able to rapidly capitalize at scale on fundamental new constructs that emerge, such as developer tooling, tokenization, token launches, wallets, security audits, DeFi (1.0, 2.0 and beyond), NFTs, bridges, Layer-2 scaling, DAOs, and more. This view has resonated with our crypto native and growth investors in a Series D that will enable us to execute powerful growth strategies." – Joseph Lubin, Founder and CEO at Consensys, (Series D press release, March 2022)

Mergers and Acquisitions

The table below shows notable acquisitions in the blockchain infrastructure industry in 2021 and 2022.

Notable Acquisitions in Blockchain Infrastructure Provision

Date	Acquirer	Sub-Category (Acquirer)	Target	Sub-category (Target)
Mar-22	Blockdaemon	Staking/Node Infra	Gem	Onramp/Digital Identity
Jul-22			Sepior	Wallets
Jan-21	Coinbase	Trading/Brokerage	Bison Trails	Staking/Node-as-a-service
Apr-21			Skew	Market Data/Analytics
Sep-21			Polychain Labs	Staking-as-a-service
Sep-21			BRD	Wallets
Nov-21			Unbound Security	Wallets
Aug-21	Jump Trading	Financial Services	Certus.One	Staking-as-a-service
Dec-21	Kraken	Trading/Brokerage	Staked	Staking-as-a-service
Jan-22	OpenSea	NFTs/Gaming	Dharma Labs	Wallets

Source: Crunchbase, The Block Research

 **THE BLOCK PRO** · RESEARCH

Following its Series C, Blockdaemon has made two acquisitions: Gem and Sepior. Gem, a blockchain firm focusing on crypto onramp technology and digital identity solutions, was [acquired](#) by Blockdaemon in March 2022. Blockdaemon intends to leverage Gem's offerings to allow institutions to access its node stack and connect to blockchains with as little friction as possible.

Blockdaemon also [acquired Sepior](#), a data and digital asset security company, in July 2022. Sepior is known for creating multiparty computation (MPC) algorithms to remove single points of failure in digital asset management institutions.

Coinbase's acquisitions of Bison Trails and Polychain Labs have bolstered its staking capabilities and allowed the firm to extend these services to clients. Coinbase also [announced](#) that it had acquired Skew, a data platform geared to serve institutional clients, in April 2021. Coinbase currently counts over 7,000 financial institutions as customers. To incentivize these institutions to continue using Coinbase's platform, it seemed critical the company carried out investments in data analytics software that provides insights similar to traditional financial markets.

In December 2021, Kraken, one of the largest centralized CEXes by volume, acquired Staked, a non-custodial staking solution. This acquisition further cemented Kraken's ability to support staking for customers who prefer to maintain custody of their assets rather than place them in the possession of others.

OpenSea, the largest NFT marketplace by trade volume, acquired Dharma Labs in a deal speculated to be worth between [\\$110 million and \\$130 million](#) in [January 2022](#). Dharma Labs initially launched a protocol for tokenized debt agreements before offering stablecoin savings accounts and ultimately allowing automated clearing house purchases in multiple US states. Dharma Labs' wallet offering was shuttered upon the closing of the acquisition, and Nadav Hollander, former Dharma Labs Co-Founder and CEO, became OpenSea's CTO. No specific information has been given yet regarding how Dharma Labs' technology will be used to enhance OpenSea's platform. However, OpenSea stated that it acquired Dharma Labs to accelerate product development, expand trust, safety, and reliability, and meaningfully invest in the NFT and web3 ecosystem.

4

Conclusion

Part 4: Conclusion

Through bull and bear markets, the secular trend towards tens of trillions of dollars residing on public blockchains over the coming years remains intact. The value proposition of blockchain technology remains clear: to enable individuals and organizations to securely transact in an immutable, censorship-resistant, and trustless environment. The failure of opaque centralized entities and the damage they have brought on the broader digital asset ecosystem underscore the need for decentralized financial infrastructure that provides users with transparency and reliability.

Improvements in digital asset data and infrastructure continue to open up new avenues for institutions and individuals to integrate digital assets with confidence. In this report we have described (i) how data providers help industry participants extract value from a novel type of data stemming from the digital asset ecosystem, and (ii) how infrastructure providers make operating and accessing blockchain networks feasible for everyday users.

Data firms provide crucial information emitting from a diverse and increasing range of sources. They are pioneering new methods for understanding and monitoring the emerging digital asset economy. Although there are signs of consolidation within the data provider landscape, companies are focusing on different niches and multiple winners will likely emerge. For example, Amberdata, Coin Metrics, and Kaiko are focusing on the data needs of institutions and traditional finance asset managers by providing high-quality off and on-chain data. Providers such as Dune and Flipside Crypto are well underway with cultivating large and diverse digital asset data contributor networks. Providers such as Nansen are championing the model of tagging addresses to facilitate deep analyses of on-chain asset movement.

Infrastructure providers are emerging as the backbone of applications and wallets that make interacting with blockchains possible. Liquid staking-as-a-service protocols such as Lido are democratizing participation in PoS networks and have emerged as an attractive alternative to established pure-play staking firms and CEXes. Node-as-a-service firms such as Alchemy and Infura have captured an enormous market share of node-related infrastructure services and are bridging the gap between blockchain networks and developer communities. Wallet providers such as Consensys (via Metamask) and Phantom have been integral in onboarding and retaining millions of users to the Ethereum and Solana ecosystems, respectively.

Nonetheless, across all of these industry verticals, dozens of competitors have emerged to offer nuanced services across different chains and targeting different use cases. For the time being, the risk of infrastructure being a winner-takes-all race across any one vertical is diminishing as new players with innovative strategies enter the space.

Disclosures

This report is sponsored by Amberdata. The content of this report contains views and opinions expressed by The Block's analysts which are solely their own opinions, and do not necessarily reflect the opinions of The Block or the organization that commissioned the report. The Block's analysts may have holdings in the assets discussed in this report and this statement is to disclose any perceived conflict of interest. Please refer to The Block's Financial Disclosures page for author holdings.

Beginning in 2021, Michael McCaffrey, the former CEO and majority owner of The Block, took a series of loans from founder and former FTX and Alameda CEO Sam Bankman-Fried. McCaffrey resigned from the company in December 2022 after failing to disclose those transactions.

This report is for informational purposes only and should not be relied upon as a basis for investment decisions, nor is it offered or intended to be used as legal, tax, investment, financial or other advice. You should conduct your own research and consult independent counsel on the matters discussed within this report. Past performance of any asset is not indicative of future results.

© 2022 The Block Crypto, Inc. All Rights Reserved.

02/16/23

The State Of Digital Asset Data And Infrastructure:
2023 Edition

Appendices

Appendices

Appendix A: Oracles

Layer-1 blockchains like Ethereum are not designed to “talk” to other blockchains or real-world data sources. However, possessing the ability to trigger transactions based on off-chain data increases the potency of decentralized applications. For example, i) decentralized prediction markets need to access real-world information for validating users’ predictions, or ii) trading venues for tokenized commodities need to access prices of commodities from exchanges outside of their blockchains. Broadly speaking, oracles bring data from the outside world into an application or a blockchain.

How Do Oracles Work?

At a high level, oracles have two components – an on-chain contract and a node network that gathers off-chain data. The node network gets the off-chain data on-chain and the contract is responsible for receiving requests from other contracts and delivering the output to them.

An oracle’s design depends on its use case. **Immediate-read** type oracles are used when data is needed immediately and unlikely to change. For example, if a prediction market needs to resolve a bet like ‘who will be the Prime Minister of the UK in November 2022?’ (after the election took place), it can use an immediate-read type oracle. A **publish-subscribe** oracle is a data feed streamed to smart contracts. This type of oracle is used when the data frequently changes. For example, a decentralized futures exchange constantly needs to know the price of ETHUSD. Another type of oracle design is **request-response**, where a client contract requests arbitrary data. It allows clients to access data other than what is provided by publish-subscribe oracles. This kind of oracle’s on-chain contracts listen to requests from other on-chain smart contracts. Nodes on the oracle’s network then gather this data from different sources via API calls or other means and deliver the data to the oracle contract. Nodes may collect this data directly from sources or use API calls from data providers like Amberdata and Kaiko. The oracle contract either directly delivers the data to contracts that requested it or performs computations on the data, such as aggregation, before doing so. This oracle design is ideal when smart contracts require only a small part of the data at one point in time. For example, when a supply-chain-based smart contract requires geolocation, it may employ a request-response type oracle. Interested readers can refer to [this paper](#) to further explore blockchain oracle designs.

First-Party vs. Third-Party Oracles

Another important design distinction of oracles is whether they are first-party or third-party oracles. In the case of third-party oracles, as the name suggests, oracle nodes are not data sources. The oracle incentivizes a network of nodes to gather data, ideally, from multiple third-party sources. The oracle then normalizes data collected from multiple nodes using methodologies like weighted averages or VWAP before delivering the data to the end user. If a third-party oracle has sufficiently diversified data sources, it is more robust to individual sources having issues. For example, in December 2021, Band Protocol [reported an incident](#) of misquoting USDT prices when

CoinMarketCap and Brave New Coin, 2 of the 4 data sources used by its nodes, reported highly irregular prices. This resulted in inaccurate reporting despite the contract using the average of the two middle prices. In the case of third-party oracles, users trust the economic incentives for the nodes to deliver the correct data.

First-party oracles are provided by the API providers themselves. As the data provider publishes data on-chain, it doesn't have to pass through middlemen (nodes), reducing the trust a user needs to place in the nodes. Getting rid of the intermediate nodes that handle data allows first-party oracles to be cheaper than third-party ones and reduces the attack surface. Since the data sources are known in this type of oracle design, their credibility is at stake when misreporting data. Thus, they are more likely to play by the rules and provide accurate data.

In effect, there is a trade-off between the trust assumptions of the first-party and third-party oracles. In the case of first-party oracles, users trust the data provider and its security practices that ensure data integrity. Whereas, in the case of third-party oracles, users trust a network of nodes, their incentive model that assures data quality, and the oracle that normalizes and “vets” data from multiple providers. This boils down to whom users trust more - data providers themselves or an oracle that “vets” the data from multiple nodes. On the technical side, compared to third-party oracles, first-party oracles have faced more [implementation challenges](#). Some of the API or data providers were apprehensive of running oracle nodes because, for example, i) first-party node operations can be unstable, ii) they may require regular user intervention, and iii) data providers are unwilling to spend or get paid in digital assets.

Oracle Solution Providers



Chainlink is the most widely used oracle within the digital asset industry. It primarily operates third-party oracles but Chainlink also provides first-party oracles for AD Derivatives (options data) and CipherTrace (KYC-related data). Chainlink uses its LINK token to incentivize its node operators to aggregate off-chain data.

Band Protocol has built its own Cosmos-SDK-based chain, BandChain. BandChain is a blockchain-agnostic solution that can seamlessly query data from [interblockchain communication \(IBC\)](#) protocol-compliant chains. This design is similar to that of Chainlink, but it is built on Cosmos rather than Ethereum. Like Chainlink, Band uses its native token, BAND, to compensate node operators.

API3 has built a solution called Airnode, which is an API wrapper that allows web APIs to connect to blockchains and directly provide data. Data providers, which are reluctant to run oracle nodes themselves because these require constant maintenance and monitoring, can use this solution.

Similar to API3, **Pyth Network** allows financial market participants such as CBOE (Chicago Board Options Exchange) and Jane Street to publish data on Pyth Network. Pyth is focused on pricing data and currently provides data for commodities, digital assets, equity, and forex. Operating on Solana allows Pyth to update data every 400ms (each block). Pyth mandates data publishers to stake PYTH tokens and shares 20% of the fee revenue with them. If published data is found to be incorrect, the publisher's stake is slashed.

While most oracles aggregate data from different sources and process it themselves before delivering it to requesters, **Flux Protocol** allows data providers to control data flow from the source to the destination chain. Its open-source provider node constantly queries the data providers' API endpoints and then calls the first-party oracle contract to post the data on the destination chain. Flux Protocol also uses its token, FLX, to ensure data correctness.

Appendix B: Data Delivery Methods

The table below provides an overview of key providers' market data delivery methods.

Data Delivery Methods of Market Data Providers

Organization	Cloud Based			FIX Protocol	REST	
	Delivery	CSV	Dashboard		API	WebSocket
Amberdata	✓	✓	✓	✓	✓	✓
CoinAPI	✓	✓	✗	✓	✓	✓
CoinGecko	✗	✓	✓	✗	✓	✗
CoinMetrics	✗	✓	✓	✗	✓	✓
CoinMarketCap	✗	✓	✓	✗	✓	✓
CryptoCompare	✗	✓	✓	✗	✓	✓
Kaiko	✓	✓	✗	✗	✓	✓
Lukka	✗	✗	✗	✗	✓	✓
Nasdaq	✓	✓	✗	✗	✓	✗
Nomics	✗	✓	✓	✗	✓	✓

Notes: In addition to the methods listed in the table above, several providers have developed bespoke data-sharing methods. (1) Kaiko offers data streaming through Kaiko stream. (2) Nasdaq has entered a partnership with Delta Bricks for sharing large data sets.

Source: Company documentation, The Block Research

 THE BLOCK PRO • RESEARCH