

THE DIGITAL ASSET DATA GUIDE FOR FINANCIAL PROFESSIONALS

GETTING THE DATA
YOU NEED TO
SUCCEED IN THE
DIGITAL ASSET CLASS

BY SHAWN DOUGLASS
AND TONGTONG GONG



INTRODUCTION

Less than a decade ago, virtually no financial professionals had heard of the term “blockchain”, with the technology confined to the domain of Bitcoin and its cypherpunks.

60%

of professional investors at institutions actively manage a portfolio of digital assets, or plan to, in the next 12-24 months



Figure 1. Financial Institutions rapidly adopting digital assets

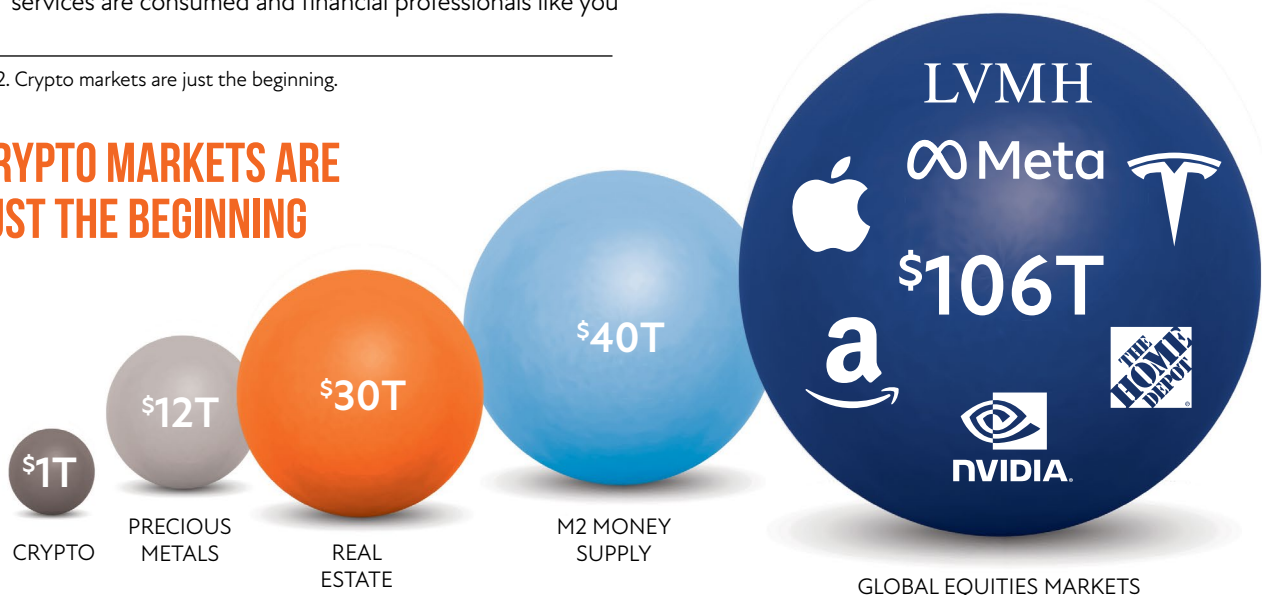
Fast forward to today and digital assets have grown to represent a significant market capitalization. Institutional adoption continues to grow rapidly with most either actively managing a portfolio of digital assets or planning to in the near term. Financial institutions can no longer ignore this asset class. However, while entering this complex and highly volatile asset class provides tremendous opportunity, it is also fraught with new risks and challenges. The digital asset class is no longer just about Bitcoin. There are now many blockchains, cryptocurrencies and centralized exchanges (CEX) that make up the crypto economy. Add to that the advent of non-custodial, decentralized exchanges (DEX) that leverage self-executing smart contracts to allow peer-to-peer (P2P) trading facilitated by automated market makers (AMM), and access to the emerging sector of decentralized finance (DeFi). DeFi is redefining how financial products and services are consumed and financial professionals like you

are realizing the need to innovate in this environment to compete with crypto-native market entrants. As digital assets become pervasive and widely adopted, the total market size and applications are unimaginable, but could represent trillions of dollars in new businesses being created as every individual and business around the world adopts this technology. Crypto native assets are just the beginning, as precious metals, real estate, equities, and more start to trade as digital assets, disrupting traditional financial services. (Figure 2). For example, security tokens have been used to create digital assets that represent ownership in physical assets, such as gold, sports cars, and even assets like equity and debt. Additionally, decentralized exchanges and DeFi are redefining how financial primitives are structured and traded, allowing new financial products to be built and deployed with instant global reach.

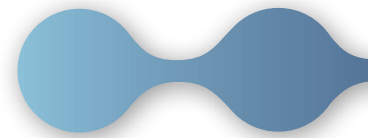
It has also led some to wonder: what data should we be looking at to enter the digital asset class? And how can we best analyze this data to gain insights on the industry’s growth potential, investment risks, and possible investment opportunities? Successful crypto exchanges, asset managers, arbitrageurs, and providers are netting [billions of dollars in profits](#), so it’s no wonder that many are looking for ways to tap into this emerging industry. In this guide, we’ll explore why you need digital asset data, the challenges with accessing it, and how to overcome them to gain an advantage in the crypto economy.

Figure 2. Crypto markets are just the beginning.

CRYPTO MARKETS ARE JUST THE BEGINNING



DIGITAL ASSETS, BLOCKCHAIN AND DEFI ARE REDEFINING FINANCE



58% multinational firms today are using crypto, while over 80% of banks have invested in the technology.

With former ETF traders taking advantage of crypto arbitrage opportunities to make billions of dollars and hedge funds aggressively increasing exposure to blockchain technologies and cryptocurrency, it's no wonder that every financial institution is looking to enter the digital asset class.

Financial professionals that ignore this massively growing industry today do so at their own peril. In addition to the return potential of this developing asset class, there is another reason that traditional financial institutions can't afford to ignore blockchain and that is the rise of DeFi. DeFi is redefining the four core financial primitives (liquidity, leverage, risk, and arbitrage) that move financial markets for the crypto economy. DeFi applications being built on these redefined primitives include:

- **Decentralized exchanges (DEXs):** Online exchanges that connect users directly on a blockchain so they can trade cryptocurrencies without an intermediary.
- **Lending platforms:** Smart contract-based lending applications that replace intermediaries (e.g., banks).
- **Prediction markets:** Markets for betting on the outcome of future events, such as sporting events.

Additionally, DeFi has spawned other new concepts not seen in traditional finance:

- **Yield farming:** Yield farming (or liquidity mining) refers to the practice of staking/lending crypto assets, bootstrapping economies that will generate interest, fees, and rewards in the form of additional tokens.
- **“Wrapped” coins:** A tokenized cryptocurrency that facilitates interoperability between different blockchain protocols. Wrapped coins allow users to earn interest on the coins they lend via decentralized lending platforms.
- **Stablecoins:** A cryptocurrency that's tied to an external asset (e.g., the dollar or euro) to stabilize the price.

- **Composability:** Because DeFi apps are open source, they can be inspected and forked (a completely independent line of development based on the original open-source code) to innovate and create new applications. Token standards (like ERC-20), smart contracts and the open source ecosystem empowers rapid innovation.
- **Money legos:** As an analogy, think of DeFi apps as Legos, the toy bricks that snap together to construct buildings, vehicles, etc. DeFi apps are like “money legos” that can be snapped together to build new financial products.

Traditional financial applications are built on centralized, expensive, and proprietary infrastructure and software that must be maintained and managed. DeFi is recreating traditional financial systems for the crypto economy and has the potential to replace them. However, financial institutions can innovate in this environment, taking advantage of the tremendous opportunities presented by digital assets, at massive cost savings over the traditional financial systems they've built and maintained.

The numbers speak for themselves: these are the early innings of what is already the next great technology disruption, and it's important that financial institutions understand how to get involved before it's too late. Companies that fail to innovate litter the history books, while those that do achieve massive success. From Kodak to Blockbuster, companies that don't innovate eventually become obsolete. The same can be said for financial services with respect to the digital asset and blockchain revolution: firms that don't evolve will be replaced by those that do.

Digital assets, blockchain and DeFi are already changing the way people save, invest, trade, own, and conduct business globally. This is not hype—this is reality.

WHY YOU NEED DIGITAL ASSET DATA

There are many questions that financial professionals will need to answer related to research, trading, risk, custody, analytics, reporting, and compliance to enter and succeed in the digital asset class.

And to answer those questions, they must have reliable data. Data fuels modern financial products and services, and in the digital asset class, it's no different. Without comprehensive, granular and unopinionated digital asset data, financial professionals wouldn't be able to make informed investment decisions, exchanges and wallets wouldn't provide useful features, and developers wouldn't be able to build the next generation of DeFi applications.

Let's dig into some of the use cases and questions that will need to be answered with digital asset data.

PRE-TRADE

Generate alpha, develop research, mitigate risk, get operational context

Development and backtesting of cryptoasset algorithmic/programmatic trading strategies

Financial professionals employ algorithmic execution across all asset classes to generate returns when risk is a factor. In crypto, you can execute algorithmic trading on spot markets, derivatives exchanges, or decentralized exchanges (DEX), so you'll need data across them. With fundamental (on-chain) data and price movement market data, you can build a crypto alpha generation platform to develop quantitative financial models or trading strategies that generate consistent alpha, or absolute returns.

Identification of arbitrage opportunities

Whether discretionarily acting or using quantitative data models executed through a bot (statistical arbitrage), once again, you'll need comprehensive data across markets, centralized exchanges, and decentralized exchanges (on-chain/blockchain fundamental data) to identify opportunity and manage risks like price impact, slippage, price movement and transfer fees. There are many opportunities for arbitrageurs to explore:

- On or across spot markets
- On or across derivative/futures exchanges
- Between spot markets and derivative exchanges
- Triangular arbitrage on a single exchange
- Defi arbitrage across multiple decentralized exchanges as well as the liquidity pools below those decentralized exchanges

Development and backtesting of liquidity provider/market maker models

In cryptocurrency, market making involves providing liquidity on a specific cryptocurrency by submitting both bid and ask limit orders on centralized exchanges. Market makers make a profit by collecting the bid-ask spread over multiple trades, as well as rewards. Market making can be done on an exchange or across multiple exchanges. Comprehensive data will be required to develop and train quantitative data models to:

- Understand volatility and liquidity on exchanges to adjust bid-ask spreads
- Determine correlations between implied volatility and spread width
- Determine when to pull bid/ask and pay penalties to avoid steep losses
- Get a view of your liquidity across exchanges in real-time (operational content and relationships)

Decentralized exchanges supply additional opportunities for market makers to provide liquidity. By participating in DEXs, liquidity providers can earn fees, rewards and other incentives. Context across DEXs, Pools and CEXs is critical to maintain a comprehensive vantage point. To operate effectively and efficiently, a liquidity provider must know:

- LP positions
- % of pool
- Fees earned and fees earned and unclaimed
- Protocol incentives
- Impermanent loss



Development of original research on digital assets

Research analysts require data to study cryptocurrencies to generate investment ideas and recommendations to traders as well as general digital asset strategies. They'll need to:

- Understand what digital assets are moving, determine how to rank them, and find trends in the data.
- Understand operational context and have a global perspective on markets.
- Understand both centralized and decentralized exchanges. How many participants are there in an exchange? What is the transaction volume? Who is providing the liquidity? Who held what, at what price, and for how long? Where did they send it? For wallets that hold a particular asset, what else did they hold? When they sold an asset, what did they buy? When a smart contract was deployed, who interacted with it?
- Identify the participants in a cryptocurrency ecosystem. Determine how much money is backing it, and what the flows into it and out of it are. Determine if it is risk-on or risk-off in the market. Analyze flow into and out of assets. Determine if people are trading one crypto for another or for a stable coin?
- Build risk models and identify opportunity.
- Analyze cryptoeconomic mechanisms.
- Understand the NFT market, and in the future, when almost every physical asset will be tokenized, identify risk and opportunity in those digital assets.

POST-TRADE

Manage risk and ensure compliance

Risk Management

Digital assets present opportunities for high returns, but they are also subject to extreme volatility. In order to monitor your firm's digital asset trading and identify and manage risk, you'll once again need data. As you enter the digital asset class and your quants develop algorithmic trading, arbitrage, lending and market making bots trying to generate returns, you'll need

to understand total exposure to volatility and other factors. While risk management professionals are used to these challenges, the complexity of digital assets exacerbates them. With thousands of cryptocurrencies and hundreds of exchanges, the amount of data you'll be dealing with to manage risk surpasses any other asset class by a wide margin. You'll also need to deal with unique risks in this asset class:

- **Valuation** – Traditional methodologies to quantify and determine exposure don't apply to cryptocurrencies. Whether you are using absolute models such as cash flows, staking, rewards, revenue earned, yields on cash flow, or relative valuation models like Stock to Flow and hash rates to production, you need comprehensive data you can trust.
- **Modeling** – modeling future cryptocurrency exposures and risk is complex and will require market pricing data from centralized exchanges as well as blockchain transaction data to gain insight into the mechanisms managing scarcity, the number of unique participants or the Total Value Locked (TVL) in a protocol.
- **Trading costs and liquidity** – Price discovery is challenging because of the high volatility and limited liquidity in the crypto market. Trading costs are high and can vary widely based on time of day and exchange.
- **Settlement** – Accounting and settlement of a trade has a significant impact on quantifying exposure valuation and adds complexity and risk. Since methodology can be substantially different across exchanges, it is complex to estimate risk exposure and counterparty credit risks are increased.
- **Derivatives** – Crypto futures and options trading, as in traditional markets, represents significant volume. Due to the leverage offered by exchanges, many of them have dynamic liquidation policies to manage risk. These liquidations in derivatives markets can propagate across exchanges and into the spot market.

Compliance

Of course, maintaining tax and regulatory compliance will require data, and once again, digital assets are more complex than traditional assets because they are living software and may trade across many trading venues.

It's clear that as institutions become more active in digital assets, regulation and mitigating exposure to enforcement actions, ensuring compliance and supporting audit requirements will be mandatory. In order to stay compliant and able to respond to regulators and auditors, you'll need to access data stored on multiple blockchains, centralized and decentralized exchanges, and wallets.

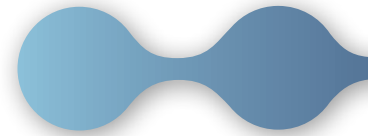
The complex nature of cryptocurrency transactions, and the fact that some network and digital assets protocol data is ephemeral, complicates accounting and tax compliance. A transaction can include a token

transfer, staking tokens to a liquidity pool in a DEX (swapping for LP tokens), obtaining fees and receiving rewards. On every single one of these sub-transactions, there's the potential to make or lose money.

The decentralized and pseudonymous nature of much of the crypto economy make KYC/AML compliance difficult. However, for regulated financial institutions to enter the asset class and adopt DeFi, they'll need to ensure and prove compliance. As an example, many financial institutions are building "blacklists" of non-compliant wallets. Prior to participating in a particular liquidity pool as a market maker, comprehensive and granular on-chain data can be used to determine the address of every wallet participating so it's ensured that no "blacklisted wallets" are also participating. Having historical on-chain transactions for "blacklisted" addresses can allow you to track their counterparties.



THE COMPLEXITIES OF BUILDING DIGITAL ASSET DATA INFRASTRUCTURE



Due to the highly decentralized, globalized, and round-the-clock nature of the digital asset ecosystem, the data is substantially more complex than in traditional finance, and collecting it is more challenging.

Digital assets trade in spot markets, derivatives markets, and on decentralized trading venues across the world and around the clock. Institutions need to connect to all these venues, stay up to date with fast-moving APIs, and normalize the data. And, in addition to all the trading venues, digital assets are built on global decentralized networks and protocols, presenting considerable operational challenges.

needs to be collected from hundreds, if not thousands, of sources. Given the sector's rapid growth, new data sources will need to be added frequently. Decentralized exchanges and DeFi increase the challenge considerably by necessitating a deeper understanding of on-chain data and more comprehensive data collection. Further, the advent of Layer 2 means there will be webs of blockchains and cross-chain assets in the future, requiring interoperability that will be even more complex to build and operate.

The number of blockchains, cryptocurrencies, and centralized exchanges already in existence means that data

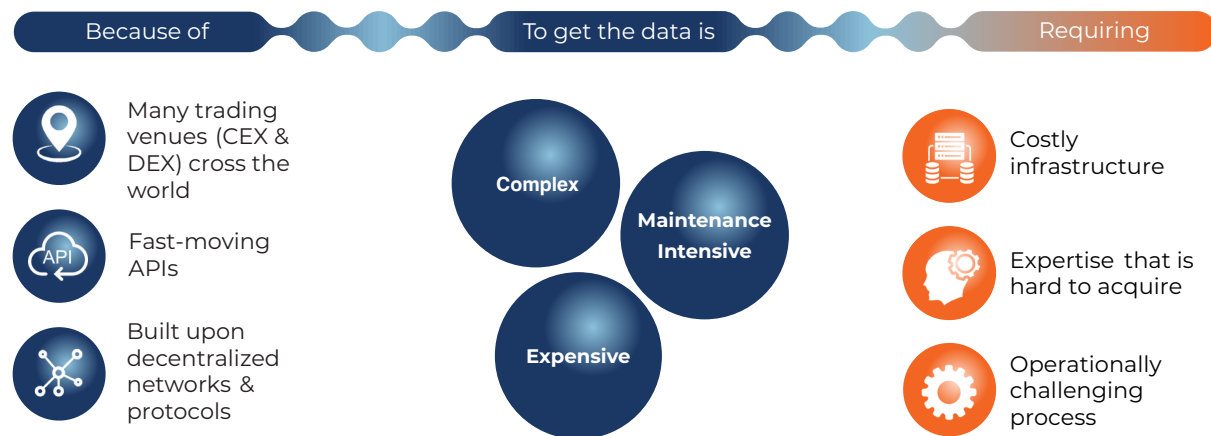


Figure 3. Financial institutions need reliable data infrastructure in order to offer digital assets, products and services.

DIGITAL ASSET DATA INFRASTRUCTURE REQUIREMENTS

In order to get a comprehensive understanding of digital assets, you need both market data and on-chain (blockchain) fundamental data for all the venues you plan to operate on. Each centralized exchange has its own market data APIs. Getting data from each exchange requires building and maintaining a dedicated infrastructure to collect, ingest, store and serve the data in a highly available and scalable service institutions can depend on.

In addition to comprehensive market data from centralized exchanges, institutions also need on-chain data to build the complete picture even if they have market data from every centralized digital asset exchange in existence.

On-chain data is the key to gaining visibility into DEX transactions and DeFi activities. Without this data, developing DeFi strategies, managing risk, and maintaining regulatory compliance is impossible. Accessing on-chain

data requires building and maintaining highly available nodes for each blockchain you need data for. Each blockchain has its own specific requirements, including the technology used, procedures, protocols and data. Each one also has its own security risks. Your team will need an expert for each blockchain you wish to work with to support its node(s).

Further, raw blockchain data is not indexed, searchable, or normalized to time series. A block is merely a point in time state of what's happening. Each block can have between a few hundred to a million transactions, but there's no context as to what happened in the previous block. Even simple questions like "what was the balance of my Ethereum wallet for the year," require multiple steps accessing a significant amount of data to answer.

Performing the needed analysis may require calling hundreds of thousands, if not millions, of blocks to extract details like balance, logs, and events. Additionally, it requires getting pricing data from a third-party source and reconstitution of the transaction in an external database. Analyzing compound transactions that occur across multiple blockchains, exchanges, and tokens is even more challenging.

As a result, before on-chain data can be used, it needs to be processed to create datasets in a format your data teams are familiar with by someone with deep expertise in the blockchain being accessed. To be usable for making investment decisions, these data sets need to be comprehensive, granular, unopinionated, and unfiltered. For institutional activities, it is also necessary to extract multiple types of information from the raw blockchain data to have the right context for each activity.

Adding to the challenge is the fact that certain transactional data, like mempool data, is ephemeral, which makes it impossible to replay exactly what happened on-chain unless the data is captured in flight. To capture this data, financial institutions need to actively and reliably track blockchain activity 24/7.

WHY INSTITUTIONS SHOULDN'T BUILD THEIR OWN INFRASTRUCTURE

The open-source nature of digital assets makes it theoretically possible for any entity to build its own digital asset data infrastructure. As a practical matter, however, financial institutions are not well-equipped for taking on this challenge.

Building the institutional-grade infrastructure needed to acquire and process all of the digital asset data needed for making investment decisions will require a considerable and time-consuming investment in talent and technology before revenue-generating activities can commence. The complexity would effectively require your financial institution to enter the data infrastructure business, as you'd have to:

- Recruit a team of experienced technology professionals from a limited talent pool. This team would need specialized blockchain developers and data infrastructure management experts, among others.
- Implement and support data pipelines transforming raw data into actionable information from blockchain nodes and exchange connections with 24/7 reliability.
- Process and integrate large amounts of data collected from multiple sources into actionable information usable by your financial teams.

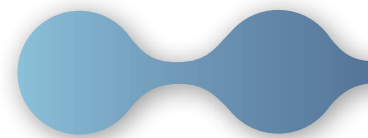
This process could take many months (if not years!) and cost millions of dollars. But it is not necessary for institutions to take on the challenge of building digital data infrastructure thanks to dedicated digital asset data vendors that provide solutions and services similar to those offered by companies like Bloomberg, Factset, and Refinitiv in traditional asset classes.

Financial institutions don't want to be in the data Infrastructure business if they don't have to be. They want to concentrate on their core business, which is providing financial services to their customers. What they really need are the fundamentals, market structural insights and signals to inform and catalyze your decisions on cryptos and other digital assets.

Between the potential returns of entering digital assets quickly and the downside risk of having incorrect or incomplete data for this data-driven asset class, it is hard to make a case for an institution to take the do-it-yourself approach.



WHAT TO LOOK FOR IN A DIGITAL ASSET DATA PROVIDER



By partnering with a digital asset data provider, institutions can begin generating investment returns without struggling to hire talent, learning how to deliver actionable information from raw blockchain data, troubleshooting and maintaining countless API connections, and wrestling with disparate data from numerous sources.

To get to market as fast as possible, it is important to pick a provider that can meet all of your needs for digital asset market and on-chain data. Trying to stitch together data from multiple vendor integration points introduces many of the operational challenges you would face if you were building the data infrastructure yourself.

The provider should have years of experience in the digital asset space and the expertise necessary to turn raw, disjointed crypto economic data into data sets that are indexed, searchable, and normalized to time series to make it easy to answer any question about digital assets. It's also important that you are getting granular, undocumented, comprehensive data that isn't smoothed out or filled in, so you can make your own decisions on digital assets.

Here's what you should look for in a digital asset data vendor when it comes to market and on-chain data.

Choose a vendor that provides comprehensive digital asset market data. They should provide streaming Level I and Level II price data to enable your real-time decision-making. You'll want access to live & full historical prices for any specified crypto asset. The vendor should deliver real-time and historical reference data to meet your trading and compliance needs. For example, they should provide access to VWAP and TWAP on and across exchanges, as well as OHLCV. They should provide Best Bid and Best Offer data (BBO) to give you visibility into spreads and potential slippage. Your chosen vendor should provide visibility into global price discovery and market depth on all major spot and derivative exchanges. They'll need to cover thousands of digital asset pairs and all the exchanges that matter to you, so your traders and liquidity providers can rely on the data to develop strategies and meet regulatory obligations. You'll need them to provide low-latency access to real-time and historical market microstructure data so you can effectively train models, discover, and manage risk, or even select the most appropriate venues to trade on. Make sure you can get access to Tick-by-tick order book events and one minute order book snapshots in a comprehensive API.

MARKET DATA

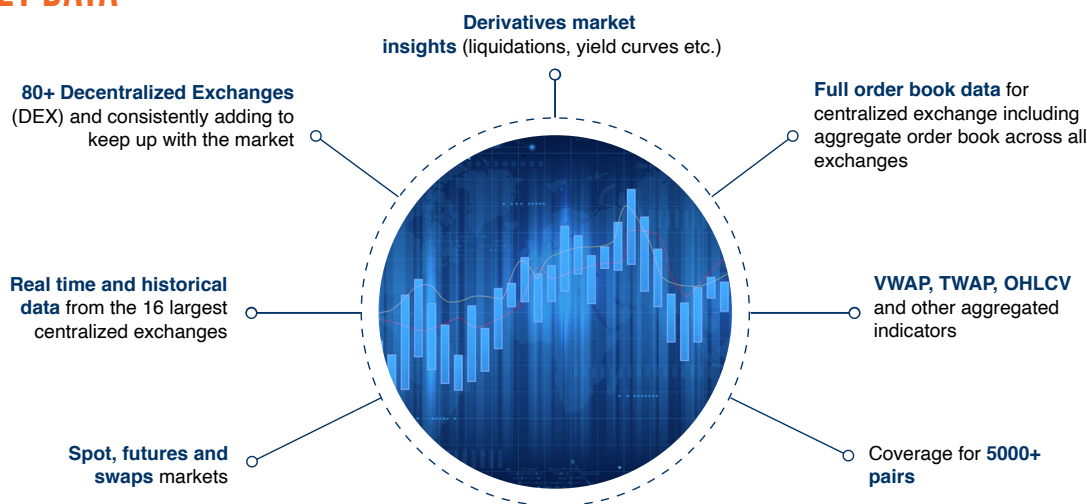


Figure 4. Digital asset market data.

ON-CHAIN DATA

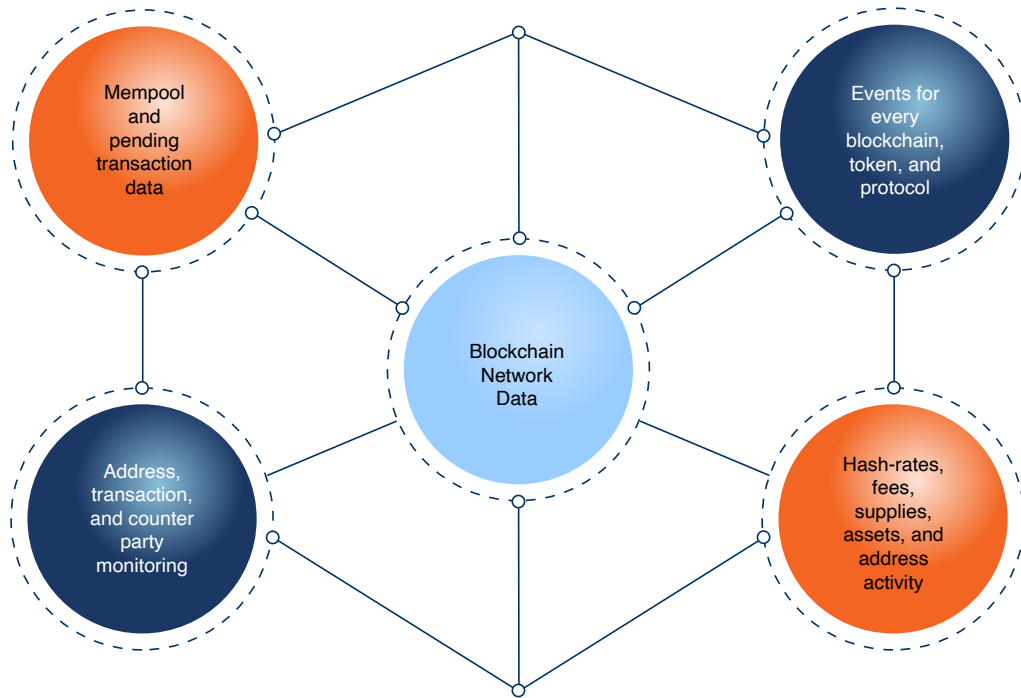


Figure 5. On-chain data.

For on-chain data make sure the provider's blockchain network data sets are built for your application specific use cases. Your chosen vendor should provide low latency access to blockchain network data, transactions, contracts, events, and logs for every digital asset, so you can analyze the smart contract mechanisms that power crypto economics. You'll need full transparency into every input and the resultant effects and market impact for every protocol in real-time as well as historically. Make sure your vendor provides mempool and pending transaction data so you can watch transactions in-flight before they are incorporated into a block and identify large movements to or from exchanges, DeFi protocols, counterparties, or an address of interest. They should provide on-chain events and data for every token combined with Level I

and Level II market data across the exchanges you care about so you can develop operational intelligence. You'll need the full historical activity for all accounts required for accounting, tax, and compliance purposes. Make sure they provide access to blockchain streams including validation data, so you can build scalable applications on top of on-chain data. You'll need deep insight into the state and health of the networks, assets, and activity such as hashrate, issuance, miner activity, fees, supply, asset, and address activity. You'll also want key indicators like NVT, asset adoption, uniques, risk adjusted returns, volatility, and asset velocity to help inform your decision making. You'll want access to data for filtering and ranking digital assets and valuation models like stock-to-flow to gain context.



CONCLUSION

Whether you've already started your journey into digital assets, or you're just in the planning stages, you need to start thinking about how you're going to get the data you need to support research, trading, risk, custody, analytics, reporting, and compliance to power your success.

Hopefully, this eBook has made a few things clear:

- You'll need comprehensive, granular, unopinionated market and on-chain data for effective decision making.
- Unfortunately, this means building and maintaining a massive data infrastructure that connects to multiple integration points with constantly changing APIs, as well as gaining the expertise to transform the raw data into actionable information. This is a daunting task that will be costly, time consuming and require highly skilled personnel that are in demand and hard to find.
- Rather than building and maintaining this complex data infrastructure yourself, partner with a digital asset data vendor who can meet your needs for both market and on-chain data. Doing so will allow you to concentrate on your core business and speed time to market for your digital asset offerings.



LOOKING TO ENTER DIGITAL ASSETS?

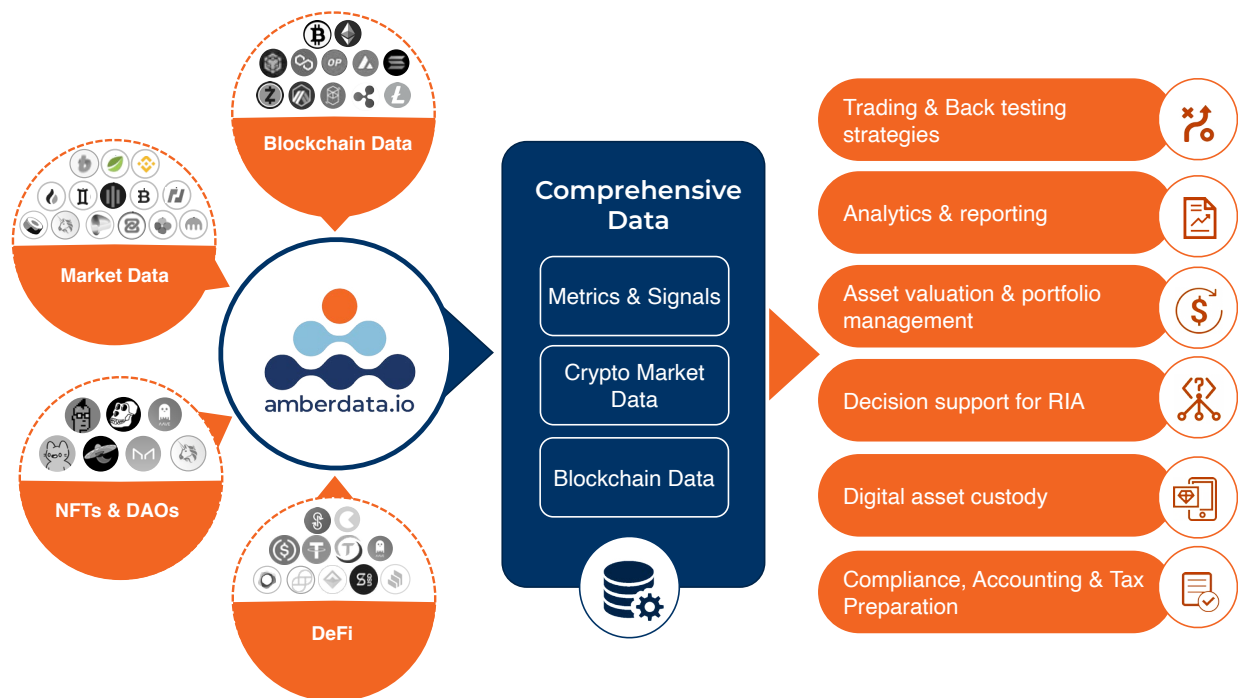


Figure 6. The Amberdata platform.

If you're looking to enter the digital asset space, you need Amberdata.

Our platform connects to all the blockchains and markets that matter today, allowing a comprehensive view of crypto markets, blockchain networks, NFTs, DAOs, and DeFi. We provide real-time and historical transparency into markets and price discovery across spot, derivative and decentralized exchanges, as well as on-chain data from the most active cryptocurrency networks and protocols.

Our data solutions support all pre- and post-trade functions. We provide deep market data, down to Level 2 order books, facilitating backtesting of quant trading strategies. And our blockchain data provides transparency not seen with other asset classes, allowing you to track pending transactions and wallet balances over time across various blockchain networks, as well as market

cap and total value locked. You can also create analytics dashboards with fundamental data to track network health and understand DeFi data like liquidity and lending rates. For fund accounting and administration, you'll know what was in a wallet at any time and what it was worth in any currency. For institutions that want to do custody themselves rather than outsource it, we provide the on-chain data needed.

With Amberdata, you get a single integration point for market and on-chain data, eliminating the need to integrate offerings from multiple vendors and allowing you to accelerate time to market for your digital asset products. We've built our data sets with institutional use cases in mind, providing the easy to consume formats and reliability you receive with traditional asset classes.

Request a demo to find out how the Amberdata platform solves digital asset data challenges and enables institutions to enter the digital asset space quickly, easily, and reliably. amberdata.io/contact-us



SHAWN DOUGLASS (CEO AND CO-FOUNDER)

Shawn Douglass is a business and technology executive with over 20 years of experience in the industry. Mr. Douglass is Co-founder and CEO of Amberdata, a leading provider of digital asset data infrastructure and market intelligence to financial institutions. Prior to founding Amberdata, Mr. Douglass served as President of Software and CTO at Unified, building and operating the company's SaaS offerings in cross-platform data management and analytics, while assisting in raising \$64M in venture funding. He also served as CTO at ServiceMesh (acquired by CSC for \$282M) and drove strategy and vision for the ServiceMesh Agility Platform. Prior to ServiceMesh, he was Managing Director at EMC Ventures, where he led strategic investments in cloud, security, big data/ analytics. He has held roles as Board Member, Operating Executive, Technologist, Advisor, Venture Investor, and derivatives trader. Mr. Douglass is a graduate of Harvard Business School.

TONGTONG GONG (COO AND CO-FOUNDER)

Tongtong Gong is Co-founder and COO at Amberdata. Ms. Gong has an extensive background in building and managing distributed systems at scale for enterprises and has led advanced development teams focusing on incubating new businesses for major software players. Her teams have successfully built high-performance computing platforms and marketing analytics software for both offline and digital markets. Prior to co-founding Amberdata, Ms. Gong served as VP of Engineering at Unified, and a variety of engineering and leadership roles at Acxiom, including establishing a Global Service Center in China. Ms. Gong was named one of CloudNow's Top Women Entrepreneurs and has been featured in Forbes, TechCrunch, and Nasdaq.

ABOUT

[Amberdata](#) is the leading provider of digital asset data and infrastructure. 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 serves as a critical piece of infrastructure for financial institutions entering the asset class and participating in digital asset markets. Learn more at amberdata.io.



amberdata.io

docs.amberdata.io

hello@amberdata.io