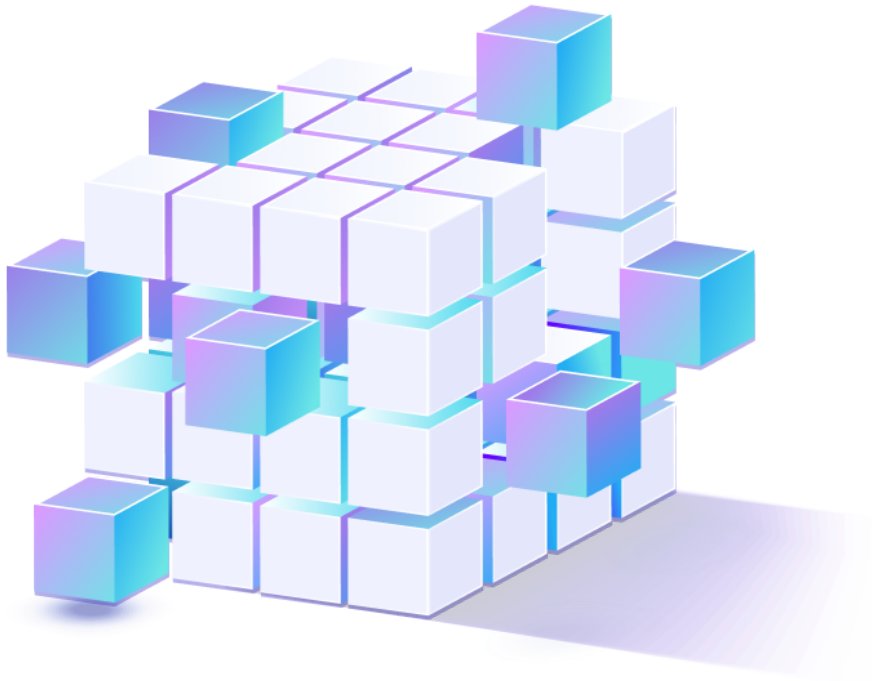
Using blockchain in a modern data and application strategy

A guide for data leaders, innovators, and practitioners



Version 1.4

Vendia

Executive summary

In 2021, Deloitte surveyed global leaders and found that seven in ten revealed they struggle to use data and technology effectively.[1] But, there's unlimited potential to leverage the vast volume of business and customer data to modernize data sharing and unlock real-time outcomes. For instance, as reported by McKinsey, as soon as 2025, the most innovative organizations will be organizing and supporting data assets as products, regardless of whether they're used by internal teams, external customers, or a permission-based partner network.[2]

The challenge facing organizations is discovering how to share business and customer data across applications, departments, partners, and clouds—in a *secure*, *trusted*, and *accurate* way. And that challenge is complex: How can organizations scale solutions for immediate impact and easy adoption? How can they ensure seamless integration with enterprise systems, transact quickly and securely, and guarantee uptime? And how can organizations achieve that outcome without exhausting internal resources and IT budgets or by building systems that require costly ongoing maintenance, compliance, and operational overhead?

During the last ten years, blockchain technology has offered companies the promise of a shared data ecosystem. But too many of those first-generation blockchains—whether public or private—fell short of their potential. They lacked the enterprise readiness, scalability, and cloud compatibility required for success.

Next-gen blockchains offer a way forward and help organizations overcome the problems and limitations of those early systems. Unlike earlier solutions, these newer offerings are built for business use cases and meet stringent enterprise requirements. The remainder of this document covers:

- The benefits and limitations of different blockchain types, including best practices for integrating them with existing centralized IT architectures
- Examples of how next-gen blockchains can accelerate business results and lower delivery risks for data-centric projects
- How next-gen blockchains such as Vendia Share bring a new level of trust and real-time data sharing to enterprises and their ecosystems
- How Vendia's unique technology combines the trust of blockchain with the scale of the cloud, offering the fastest time to market for secure, operational data sharing

[1] Deloitte. 2022. "2022 Banking and capital markets outlook." Accessed July 19, 2022. https://bit.ly/3zjy16v

[2] McKinsey. 2022. "The data-driven enterprise of 2025." Accessed July 19, 2022. https://mck.co/3yPNpWx

Contents

Strengths and shortcomings of blockchains and centralized IT

In the early days of blockchain, private and public blockchains both positioned themselves as a one-size-fits-all replacement for the entire IT stack. With some time and perspective, it's become clear that neither public nor private chains are a true substitute for the other, nor can they replace centralized (conventional, non-blockchain) IT architectures. In order to leverage the best of next-gen blockchain and understand these new design patterns better, it's helpful to begin by comparing the strengths and weaknesses of public chains, private chains, and conventional, centralized IT solutions. *Figure 1* summarizes these tradeoffs visually.

FIGURE 1

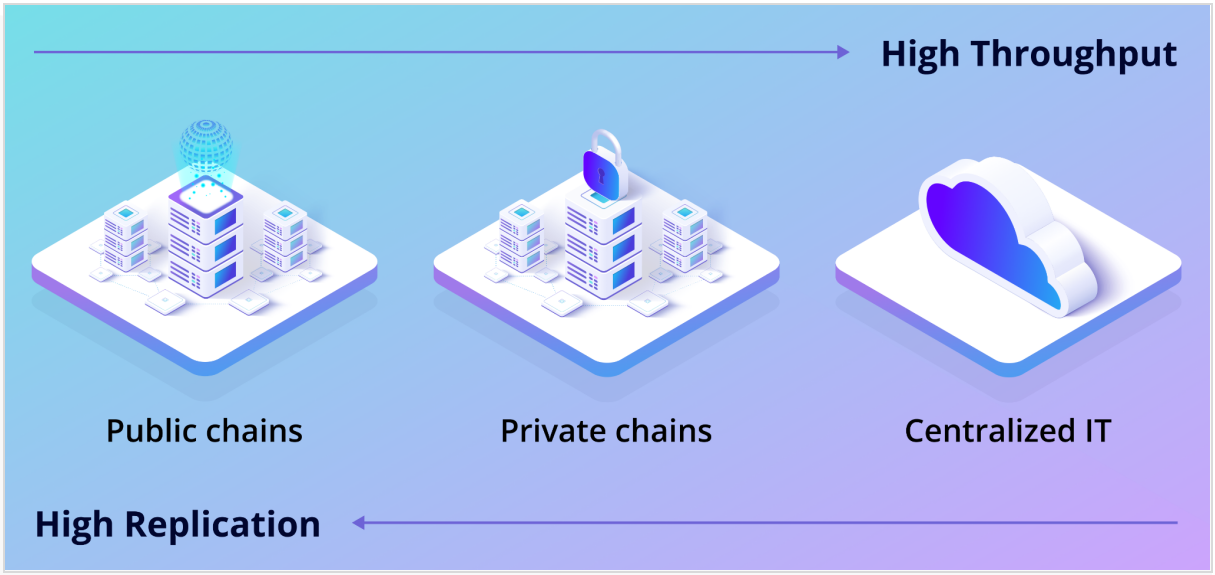


Figure 1: The continuum of capabilities for securely sharing data. Public chains offer the broadest distribution model for limited amounts of data at the highest transfer and storage cost. Private chains optimize for reliably sharing large amounts of data with a fixed population; this makes them ideal for building secure data sharing solutions. Centralized IT implementations will always offer the highest throughput, lowest cost, and easiest governance for high-bandwidth data.

Public blockchains: The best for archiving public data

Not all applications or solutions require placing data into a planet-spanning public record. But when they do, public blockchains offer the broadest degree of sharing and require the least amount of trust to achieve this outcome.

Public chains are most valuable when organizations need to bring multiple parties together with equal access and visibility. They're also at their best when the network can exist independently without needing to integrate closely with enterprise systems:

- **Public chains are best for sharing public and permanent information.** With broad replication across geopolitical boundaries, nation-states, cloud service providers, and a mix of corporate and individual ownership, public chains such as Ethereum[3] effectively create a public and permanent (irreversible and irrefutable) archive. These facts are held in the same order everywhere they're stored, so it's also possible to establish cause and effect (i.e., whether one fact came before or after another fact). This makes it possible to model and audit changes in ownership, such as transfers of value or the ownership of real and virtual objects—serving use cases such as the archival storage for ownership information (e.g., NFTs, digital home titles, etc.), self-service ticket transfers among consumers, and use cases that require broad access to, or auditing of, data digests.

- **This planet-wide dispersal of information comes with both literal and figurative costs.** Ethereum, the current gold standard in public chains that support data and smart contracts, has transaction rates that are 7-8 orders of magnitude higher than similar centralized IT transactions. Latencies can be 7 orders of magnitude higher, placing the price/performance ratio at up to 15 orders of magnitude worse than a typical cloud database, e.g., Amazon DynamoDB. InterPlanetary File System (IPFS) storage costs, especially when coupled with Ethereum to store hashes of files, are also orders of magnitude higher than public

[3] This article focuses on the use of public chains as elements of an operational IT solution, not for purposes of currency speculation.

cloud storage. This is particularly true when similar levels of durability and availability are required, making public chains a good solution only when an application's needs mandate the broadest degree of sharing.

- **Public chains innovate slowly and offer minimal integration solutions.** Integration with cloud services is minimal to nonexistent because, by design and intent, *public chains deliberately avoid any form of deep integration with the public cloud* or other "centralized" technologies, including SaaS services. Changes to the system are generally slow since open source developers, miners, and node operators all have to agree to any changes taking place in that complex arrangement of on-chain incentives—all while keeping the codebase secure, available, and backward compatible. This complex interplay of economic and technological incentive structures makes public chains slower to innovate and migrate than private ones and typically slower than a centralized IT solution, which can be adapted or modified with just a single company's authority.

- **Support for heterogeneous data types is limited and often complex to achieve.** Public chains often require stitching together multiple on- and off-chain solutions to represent common business artifacts (i.e., files). Also, because they must service essentially *any* data model from *any* customer anywhere in the world, public chains are typeless; they offer only key/value stores that require adopters to erect data models atop them. It's akin to creating a complex object-relational mapping (ORM) abstraction on top of an underlying database (but one with limited query and update functionality. *This is often a costly and overlooked element of interacting with a public chain in any real-world application.*

- **Public chains vary in their environmental footprint, but the gold standard (Ethereum) currently has a very poor environmental record.** Newer "Proof of Stake" approaches improve the cost of minting transactions. But, all public chains are structurally required to spend a higher percentage of their compute (and, thus, of their carbon footprint) achieving Byzantine resistance and worldwide consensus than is possible with centralized or private chain approaches. This overhead makes them more compute- (and carbon-) intensive on a per-transaction basis.

Due to the limitations described above, most organizations that do employ public chains will use some form of digest (Merkle Tree) approach to store only a *synopsis* of their data, rather than treating the public chain as a real-time operational data store, allowing

them to amortize the cost, environmental impact, and latency of public chain overhead.

Private blockchains: The best for partner data sharing and the environment

Private blockchains occupy an essential space between fully public chains and conventional, centralized IT implementations. They support a *controlled* population of business partners in a conventional chain relationship (e.g., manufacturers and suppliers, airlines and joint venture partners, local and global banks, lenders and underwriters, etc.) They also make data sharing much easier than centralized IT, where creating a single source of truth from the raw building blocks of cloud or on-prem infrastructure among many different companies and IT architectures represents a very heavy lift.

- **Private chains offer partners a single source of truth for operational data—without sacrificing privacy, compliance, or control.** Unlike public chains, private chains can support replicating large amounts of data in real time, making it easy to share clean, controlled operational data. Smart contracts make it easy for those same partners to share common policies, workflows, and data integrity constraints, too.

- **Private chains offer the most sophisticated data privacy, protection, and governance.** Because private chains are designed from the ground up with selective information sharing as a key feature, they generally have far more sophisticated (and well-tested) data controls than "roll your own" centralized solutions built in-house or glued onto public chains. Private chains delivered as SaaS solutions get all the conventional SaaS economic benefits of amortizing the costs of building, securing, and maintaining complex governance and data access controls over many customers—a benefit unavailable to one-off centralized solutions and incompatible with public chain mechanisms that require everyone to see all chain data, all the time.

- **Private chains offer IT teams rapid development time.** It may seem counterintuitive, but private chains usually offer the fastest time to market:

- Private blockchains provide a complete data modeling, storage, replication, and encryption platform, requiring less manual buildout than a built-from-scratch centralized application.
- They also offer more built-in capabilities (especially when offered in a SaaS deployment model) compared to a public chain; this makes them the closest to a working solution out of the box. Access controls, API generation, support for files, etc. are included, rather than being "DIY" challenges for adopters.
- Private chain users can innovate as fast as their joint decision-making process allows. Plus, they can build and manage without the overhead of buy-in from miners, node operators, and open source developers who could be anywhere on the planet—and who may not align with customers' needs particularly well.

- **Private chains can offer the lowest infrastructure costs and environmental footprint of all three categories.** Private chains have a unique advantage over public and centralized IT chains: *They can be delivered as SaaS*. As a result, private chains offer fast innovation, minimal operational burdens, minimal infrastructure costs, and take advantage of multi-tenanted approaches that offer enterprises and their partners significant economies of scale. They have the lowest carbon footprint of any of the three approaches, because they leverage multi-tenancy as well as modern "serverless" implementation techniques.

Centralized solutions: Classic IT wins on throughput and governance

Centralized solutions include virtually all existing applications and leverage a mix of on-premise and cloud services.[4] These solutions span the classic range of IT infrastructure: Operational databases, data warehouses for analytics, compute, file storage, and more.

[4] Exceptions, of course, include existing production deployments of private chains such as Hyperledger Fabric, Ethereum-based apps, or use of cloud services explicitly designed to target blockchains, such as Amazon Managed Blockchain. Today, these make up less than 0.01% of IT codebases.

Cloud-based centralized services and application designs will continue to be the majority of every enterprise's IT portfolio for durable, structural reasons:

- **Centralized solutions will always offer the highest throughput and lowest latency and cost structure for in-house data.** Sharing anything has a price, so unsurprisingly compute, storage, and network requirements and costs are minimized when an application or IT solution is "talking to itself." Latency and cost are minimized, throughput and bandwidth are maximized, and costs are tightly enveloped to just what is required for the underlying storage and compute needed to perform the actual function.

- **Centralized solutions have trivial solutions for many privacy, security, and governance controls that are designed to avoid sharing data in the first place.** A large percentage of data—and the vast majority of systems and services—within a company are never exposed to the outside world. Without intending to minimize the challenges of securing modern IT infrastructure, keeping wholesale systems data from being shared is far easier than operating public or shared infrastructure. It's also easier than making row-by-row and API-call-by-API-call decisions regarding what to expose or hide. Centralized IT will always offer easier ways to keep private things private than it will trying to make select data public. Conversely, trying to keep "private" data hidden while simultaneously storing it on a public chain represents a long-term exposure risk. Even advances in complex cryptography, such as homomorphic encryption, will never fully replace straightforward firewalls because they are inherently more costly and complex.

- **Centralized solutions require the least amount of external collaboration.** If an organization in a business network needs to seek permission from other parties to make a change, changes are likely to be slower. When the only entity an organization needs to consult with is itself, actual decision making as well as the resulting IT updates will typically be faster than when other parties are involved, with public chains being the slowest and private chains representing an intermediate point.

- **Centralized solutions are intermediate in their environmental impacts.** Because they typically serve only their owner and need to scale to maximum potential usage rates, centralized systems often suffer from low utilization, with 10-15% (i.e., 85-90% waste) being typical for enterprises. This waste makes them

worse than private chains but better than legacy public chains based on environmentally destructive Proof of Work (PoW) algorithms.

A better together approach

No single approach is ideal for modeling the complex, multi-party data sharing challenges inherent in real-world, business applications and use cases. Neither public nor private chains are a true substitute for the other, and neither is a replacement for conventional, centralized IT architectures. What has emerged instead is a playbook for how all three technology categories will interoperate to create best-of-breed IT outcomes over the next 20 years. As a best practice, centralized IT, private chains, and public chains are all employed together, making the most of their best advantages. (See *Figure 2* for a quick look at their complementary advantages.)

FIGURE 2

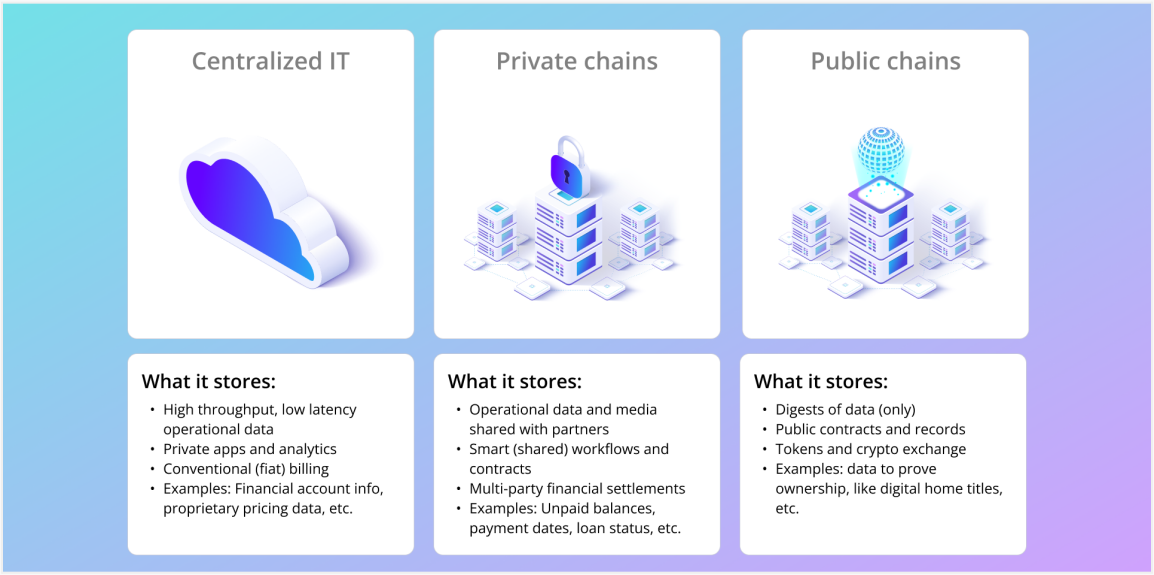


Figure 2: **Best practice for secure data sharing.** Employ all three service categories to their best advantage: Centralized operational data held in conventional storage and compute solutions (typically in the cloud); partner data sharing accomplished through high-speed, highly governable private networks; and, public chains used for digests and currency management.

Centralized IT solutions increasingly built in the cloud and often on top of managed cloud services, will continue to offer organizations the highest throughput, lowest latency, and easiest governance indefinitely.

- They're the right place to locate truly *private* applications and data, such as analytics and business intelligence (BI) exclusive to the owning company.
- Most systems continue to be internal-only, with no need to pay for the overhead (or security concerns) of outside-the-firewall distribution or coordination.
- These systems may include business, financial, and customer data like proprietary pricing data (e.g., data required to arrive at a rate), PII (personally identifiable information), customer account info, etc.
- They're also ideal for conventional fiat (non-cryptocurrency) billing, payment, and other money-moving systems.
- They will continue to form 80% or more of the foundational layer of enterprise IT going forward.

Private blockchains give organizations a key advantage: They allow firms to innovate, reduce costs, and bring new products and services to market quickly, even when doing so requires interacting with a complex ecosystem of data partners. Private blockchains can integrate seamlessly with the cloud solutions of their controlled populations; this makes private blockchains ideal for creating partner networks that enable businesses to share data with rigorous compliance, privacy, and governance capabilities.

- Partner data networks and shared applications center on private chains for their simplicity of governance, security, and throughput/cost ratio advantages.
- Over time, distributed ledgers will replace conventional public APIs; ledgers offer an automatically maintained and cryptographically provable single source of truth for cross-company use cases such as financial ledgers, supply chain tracking, business data file sharing etc. Recreating these complex capabilities, in every participant, is painfully expensive to design, develop, deploy, secure, and maintain over time without the assistance of a blockchain-based platform.
- Private chains are also ideal for cross-party financial settlements and payments, since they can automate reconciliation and ensure finality, thereby eliminating the errors and inefficiency of manual or redundant processes.
- Private chains will grow to eventually represent 15-20% of IT spend once adoption is fully matured.

Public blockchains will continue to offer very limited throughput and will continue to exhibit high latency, limited data type support, and modest innovation rates (due to their open source/consortium nature). Because of their storage and transaction costs, public chains will perennially represent a small percentage of business, financial, and customer

data (and corresponding IT spend), but they will continue to play a role in creating trusted, planet-wide stores of record:

- Public blockchains will play a key role in applications that require publicly available disclosure or non-repudiation.
- They will also play a key role where public data markets exist, serving key end user scenarios such as the archival storage for ownership (e.g., NFTs, digital home titles, etc.) and self-service ticket transfers among consumers.
- Cryptocurrencies and related on-chain token and payment services will continue to operate largely as public chains.
- Over the long term, public chain integration will comprise 1-2% of typical enterprise IT stacks.

Sharing data with next-gen blockchains

Centralized IT and blockchain technologies should not be at odds with one another. Rather, in order to leverage a "better together" approach organizations should look to next-gen blockchains. Next-gen platforms like Vendia Share give companies the *best of blockchain*, plus *the best of central IT* to empower transformation, securely share real-time data, and eliminate errors and the need for massive reconciliation.

Next-gen offerings combine the decentralization and trust of blockchains, the scale of public cloud services, and the TCO benefits of being delivered as a service. They are designed for rapid time-to-market and easy adoption. They also deliver the compliance and security features that enterprises require, including secure operational data sharing across partners, high throughput and low latency, integration with existing enterprise systems, and full lineage and traceability for all data types.

Example: Mortgage servicing using next-gen blockchain

Outside the industry, a mortgage payment may seem like a simple payment from a borrower to a lender. Financial service professionals, however, recognize that, behind-the-scenes, critical data is continuously being shared and reconciled among various parties:

- Lenders need to reconcile the status of debt and interest payments.
- Servicers must ensure that property taxes and insurance premiums are paid to the right parties.

- Organizations such as Fannie Mae and Freddie Mac need to keep track of mortgage transfers between servicers.
- Secondary markets that securitize mortgages must determine how to package mortgages into securities, including making necessary adjustments when borrowers delay or default on payments and providing appropriate disclosures.

Because these are financial transactions, every detail matters: The order of operations, the precision of shared calculations, and most of all the fact that all parties need to agree on a single source of truth regarding the status of every transaction. Accuracy, completeness, and up-to-date information are all key to the system's correct functioning … and the health of each participating organization's balance sheet.

Operating in this sophisticated system without real-time data sharing can require weeks or months of manual reconciliation work across disjointed systems and can expose financial service providers to a range of both bottom-line and regulatory risks.

How next-gen blockchains help

Next-gen blockchains offer the various parties in a business network, such as mortgage servicing and securitization in the example above, a *single source of truth* that's accessible and visible to everyone—with a clear record of every transaction that occurred. Records cannot be tampered with or deleted, making disagreements over the source of truth impossible and rendering costly and repetitive dispute resolution processes unnecessary, bringing about unprecedented levels of trust.

Every party in the process can be connected to a shared set of records, visible and auditable by everyone participating in the network. This connection removes the delays and errors caused by disjointed and manual processes as records are shuffled from one party to another, often with errors or omissions being introduced. A next-gen blockchain approach also makes it possible to get a real-time view of risk exposure and cash flow movements without waiting for data reconciliation delays: Every party knows the status of every transaction (to which it has access) within seconds of that transaction taking place … even if it originated in the IT architecture of another company.

The structural benefits of blockchain include its immutability, transparency, and the ability for every party and system involved to trace every transaction and change that's taken place. The new enabler is the scalability, cloud integration, and enterprise readiness of next-gen solutions. Together, these emergent benefits provide the foundation for an entirely new level of trust that can unlock innovation, new products, and competitive

advantages for organizations—from tracking mortgage servicing across providers to helping enterprsies provide accurate and auditable environmental, social, and governance (ESG) data.

Creating a single source of truth without sacrificing control

Organizations that want to investigate how next-gen blockchain can fit into their operational data strategy should explore Vendia Share—Vendia's next-gen blockchain platform. The Vendia Share platform brings all data-sharing parties together around a single source of truth updated in real time.

Indisputable, single source of truth (SSOT): Vendia automatically reconciles data to ensure only the latest data appears on everyone's copy of "The Truth" while giving everyone visibility of current and past states. This single source of truth makes it possible to audit and quickly settle disputes or gain new insights, regardless of the origin of the data.

Secure data exchange: With Vendia, companies control exactly what data they share (and don't share), eliminating the risks associated with sharing business, financial, and customer data. The Vendia Share platform is SOC2 and PII certified.

Powered by the cloud: Unlike on-premise solutions or public blockchains, Vendia Share is powered by the cloud and delivered as a SaaS offering that scales up and down with demand. Vendia Share customers get virtually infinite computing and scale available at a moment's notice, with the majority of costs based only on the capacity used.

Easy integration through a single access point: Every partner gets an access point (a Node) to the Vendia Share platform with access to their copy of the latest data across the data-sharing partnership. Every partner can easily connect their system (whether cloud or on-premise) to Vendia, and organizations can be up and running in a matter of days. Easy to use data modeling tools make it easy to represent existing systems, and Vendia's automatic API generation can create sophisticated, mobile- and cloud-ready APIs in minutes from any data model.

Faster innovation: Vendia Share makes it possible to create and spin up new Nodes in minutes and connect partners with ease. There's no need for painful infrastructure deployments or ongoing server maintenance. With low latency and guaranteed up-time, organizations can focus on what matters to them, their partners, and their customers.

Conclusion

Next-gen blockchain helps organizations transform into data-driven market leaders by equipping them with the best approaches for data sharing ecosystems. Functionally, next-gen blockchain solutions allow enterprises to share real-time data securely, eliminate errors, and erase the need for massive reconciliation.

When applied correctly as part of a complete data and application strategy, next-gen blockchains can accelerate innovation, unlock new revenue, and create new market leaders.

Dos and don'ts for defining an operational data strategy

Looking back on technology adoption trends, there are common pitfalls and best practices:

Avoid these temptations

- Don't assume a web3 cloud will replace the public cloud. Carefully analyze "web3" solutions for performance, security, and enterprise-grade readiness.
- Don't assume public chains can be made to perform adequately; the price/performance gap to private chains and centralized IT represents a "forever problem" in the construction of public chain technologies.
- Don't forget about data modeling challenges when considering public

Adopt these approaches

✔ Optimize tech selection for the best outcomes, rather than pursuing a "pure play" approach (i.e., only centralized/only private/only public).

✔ Exploit modern platforms, such as Vendia Share to avoid "DIY plumbing" that recreates multi-party and multi-cloud data replication, data sharing, and data security in every application; refocus that time and energy on differentiated user experiences and new business capabilities.

✔ Leverage the best of cloud, central IT, public, and private chains together with

chain-based applications or integrations.

an understanding of the unique advantages and limitations of each technology category.

Getting started with next-gen blockchain

Vendia offers a proof of concept (POC), which includes help from experienced solutions architects. In just five days, organizations get a working application that demonstrates proof of value, including:

- Multi-party data sharing
- Easy partner invitations
- Fine-grain data access controls
- A resilient global serverless architecture with no operational overhead

Request a proof of concept

Visit vendia.net/poc or reach out to the Vendia team directly to get started.

About Vendia Vendia is a real-time data cloud company that equips enterprise IT and innovation teams with real-time, multi-party data-sharing solutions. Organizations and their partners gain a single source of truth without sacrificing control, security, and compliance. Learn more at vendia.net.