Telecom Business Review
17 (1) 2024, 18-32
http://publishingindia.com/tbr/

Demystifying Decentralized Storage – A Critical Building Block of Future of Internet or Web 3.0

Sohag Sarkar

Business Development Lead (Codex), Director AI & Data (EY), Melbourne, Australia.
Email: sohag.sarkar@gmail.com

Abstract

In a parallel world, the contemporary internet ethos is being challenged by Web 3.0 (or Web3). The latter is a reimagined iteration of world wide web (or Web 2.0).

The fundamental construct of Web3 is decentralization that would transform the very fabric of interaction over the internet. Today's internet is centralized in nature; simply put the applications and data are hosted on centralized servers (or public cloud). This makes our 'data at risk' as it poses the classical problem i.e., a single-point-of-failure.

The centralized nature of data ownership signifies that there is limited or no sovereignty. And there are several challenges or treats pertaining to the centralized data storage: namely lack of data ownership, data censorship, cybersecurity breaches, and outages. Last few years have witnessed growing incidents of these treats happening in different parts of the globe and impacting millions of users.

The research paper titled "*Demystifying Decentralized Storage – A critical building block of future of internet or Web 3.0*" addresses this pertinent problem looming on humanity (individuals as well as businesses) i.e., how to protect our past, present and future data by leveraging Web3 technology? Is decentralization the solution to these data treats? Does Web3 storage provide economic advantage over contemporary cloud-centric (or centralized) data storage services? How do we solve the problem of data ownership and bring the empowerment back to the users (instead of few large IT giants who are hoarding data of the masses or enterprises)? Are there immediate challenges to store data on Web3? The research paper also brings to light the use cases (Artificial intelligence, IoT, Autonomous vehicles) that would be the early adopters of Web3 or decentralized storage.

Area—Emerging Technologies (Blockchain)

Keywords—Web3, Blockchain, Decentralized Storage, Artificial Intelligence

I. INTRODUCTION

A. Humanity and Information

Identity of modern-day humans (a.k.a. *home sapiens*) is derived from our ability to harness knowledge and experience. This is by means of stored information that serves as a foundation for learning, innovation, and progress; and plays a vital role in shaping human development and enhancing our capacity to create a better future.

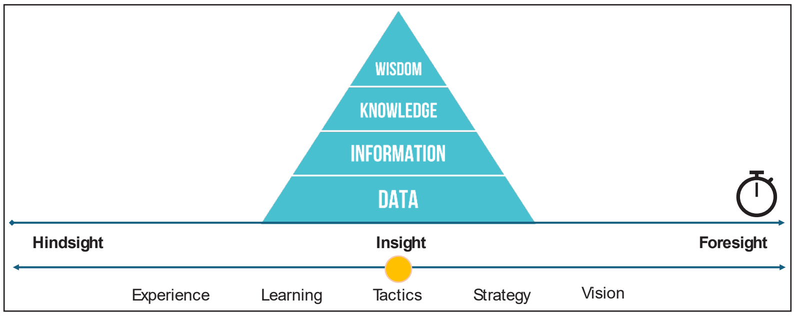


Fig. 1: DIKW Pyramid

B. Mechanism of Storing Information

The popular ways to store information (today) includes:

- *Biological Way:* (Naturally) Storing information in the one-billion neurons in the human brain that provides close to 2.5 petabytes (or a million gigabytes) of storage capacity (Paul Reber, 2010). (Artificially) Storing cold (or archival) information using synthetic DNA.
- *Physical Way:* Storing information using physical forms such as books, journals, newspaper, photograph or printed materials. Analog storage medium for audio or video such as microfilm, audio tape, VHS tape, vinyl records, etc.
- *Digital Way:* Storing information on electronic devices (computers, smartphones, cloud servers, network attached storage, etc.) or external drivers (USB, Discs – CD/DVD, etc.).

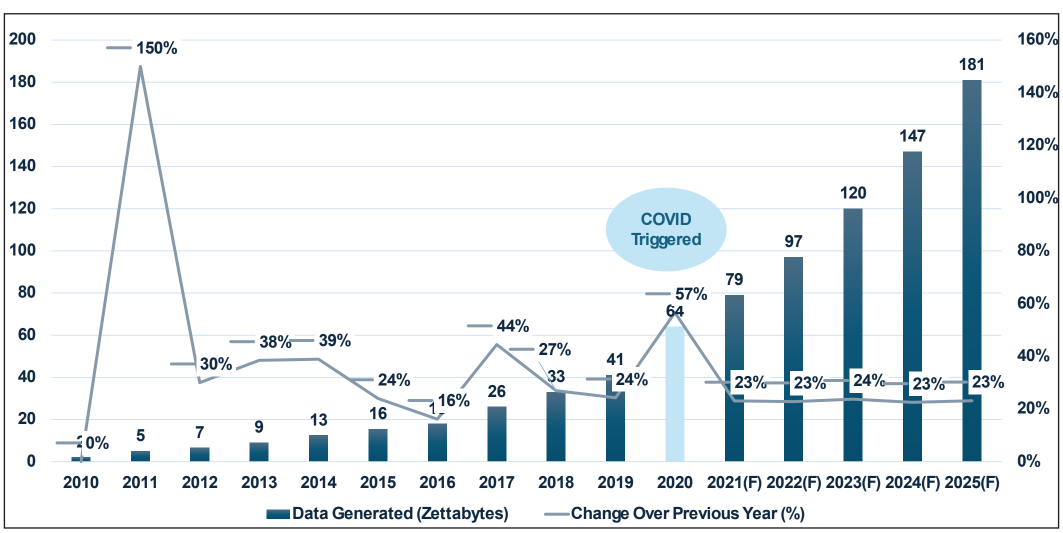
C. Sufficiency of Resources to Store Information for Future

The ability to store historical information or data for the future depends on several factors:

- *Technology Capability:* As evident from the hitherto popular Moore's law, technological advancement has increased our ability to store large amounts of data.
- *Digital Preservation Efforts:* There is growing impetus on preserving digital information by governments and organizations. These efforts include creating archives, using data redundancy techniques, and ensuring compatibility with future technologies.
- *Data Management and Organization:* Going beyond data storage and management and ensuring metadata and data integrity and durability over time.
- *Financial and Infrastructure Support:* Enabling budget and infrastructure for continuous data storage and preservation.
- *Challenges and Limitations:* Catering to known and future data challenges namely: outages, temporary or permanent loss, hardware malfunctions (or disk crashes), cybersecurity threats (like ransomware attack), ownership and cost of storage.

Big Bang of Data surge

Fabio (2024) states that more than 400 million terabytes of data are being generated daily. It would amount to 181 zettabytes in 2025, over half on account of videos. In the space of 13 years, data has increased by an estimated 74x from just 2 zettabytes in 2010.



Source: Statista, 2024.

Fig. 2: Volume of Data Generated Worldwide from 2010 to 2020, with Forecasts from 2021 to 2025

Factors that are contributing to this phenomenon growth includes:

- Proliferation of Internet and Mobile Devices (Smart Devices).
- Social Media and User-Generated Content (Platforms like Facebook, TikTok, Instagram, YouTube).
- Computing and Storage (Cloud services).
- Internet of Things (Wearables, Sensors).
- Big Data Analytics, Artificial Intelligence and Machine Learning (Generative AI).
- E-Commerce and Digital Transactions (Marketplace, Digital Payment).
- Digital Media and Streaming Services (Video, Music, Podcast).
- Regulatory and Compliance Requirements (Data retention).
- Remote Work/Education and Collaboration Tools (Virtual Meeting).
- Advancements in Digital Imaging and Sensors (HD Camera, Drones).

C. Is Cloud the Saviour?

Cloud service has emerged as a popular storage medium. It manages and stores resources and data in a single, centralized location or geography. Its popularity may be attributed to these factors:

- *Scalability:* Flexibility to increase or decrease storage capacity on-the-go without hardware (or capital) investment.
- *Accessibility:* Anywhere, anytime data access over internet driving globalization and collaboration.
- *Redundancy and Reliability:* Data replication to counter hardware failures or disasters.
- *Security:* Safeguarding data against malicious treats through encryption, access controls, software patching and compliance certifications.
- *Cost Efficiency:* Pay-as-you-go (or subscription based) model in comparison to higher investment and maintenance cost for on-premise solutions or data centres.

As per Data Journalist, Felix (2024), the top three cloud infrastructure providers own 67% of the worldwide market share. John Dinsdale, chief analyst at Synergy Research Group, predicts that the cloud service market will continue to expand substantially despite its current size (US$ 300 billion in 2024); and it may more than double in size (US$ 665 billion with CAGR of 22.4%) within the next four years.

However, (centralized) cloud storage has created its own challenges, namely:

- *Data Sovereignty:* Considerations and regulations around movement and storage of data outside of the country of origin.
- *Long Term Costs:* Growing data demands (historical, present-day as well as future) and reliance adds to the accumulated spend.
- *Vendor Lock-In:* Complexity around data portability and interoperability.
- *Censorship:* Restricting or denying access to information.

This paper analyses the 'state of data management' today to hypothesize the preferred data storage solution to sustain efficient upkeep of human data.

II. LITERATURE REVIEW

A. Evolution of Internet

The term World Wide Web (www) was coined by Tim Berners-Lee (W3C). It is a global web of information and resources interconnected through hypertext (Investopedia, 2024). Its inception version of 1990s, also known as *Web 1.0*, was characterized by read-only static webpages without any interaction feature (Cormode & Krishnamurthy, 2008). In 1999, *Web 2.0* was coined by Darcy DiNucci and popularized by O'Reilly Media. It transitioned us to a world of read-write internet that offered a two-way information exchange mechanism. It boosted user generated content, resulted in greater interactivity and collaboration, extensive network connectivity, popularised interactive platforms, and communication channels (Bruns & Bahnisch, 2009). In 2014, the term *Web 3.0* (or simply put 'Web3') was coined by Ethereum co-founder Gavin Wood. Web3 is characterized by the interplay of advanced technologies like Blockchain, Artificial Intelligence and Internet of Things (IoT). The centralization of control over transactions, content, and data in Web2, which paved the way for Web3 – synonymous with decentralization of power back to the users and removal of intermediaries. There isn't a distinct definition of future iterations of internet. For example: *Web 4.0* (also known as Symbiotic Web) is thought of as an advancement over Web3 with evolved decentralization and control, interplay of peer-to-peer (P2P) networks and more enhanced control of data by the user (Berners-Lee, n.d.). While, Jack Dorsey (ex-CEO of Twitter / X) has conceptualised *Web 5.0* (simply put Web2 plus Web3). The key characteristics include self-owned decentralized identifiers, verifiable credentials and decentralized web nodes for data retrieval and storage i.e. it would provide control of identity and data to the users (Patel, 2013).

B. Web 2 vs Web 3: Key Differentiators

Web 3.0 represents a significant evolution in how we interact with the internet, promising greater decentralization, verifiability, self-governance, permissionless, user control, ownership, and transparency (Algorand, 2023). Web 2.0, on the other hand, brought about unprecedented connectivity and content creation, however it led to centralization issues that Web 3.0 aims to resolve (Zerocap, 2024).

Table 1: Key Differences between Web2 vs Web3

Parameters	Web2	Web3
Organization structure	Centrally owned: Decisions are based on adding shareholder value	Community governed: Decisions are based on user consensus
Governance	Controlled by platform owners	Decentralized Autonomous Organizations (DAOs)
Architecture	Client-Server	Decentralized and Peer-to-Peer
User Interaction	Web browsers, mobile apps	Web3 browsers, decentralized applications (dApps)
Identity Management	Centralized authentication (e.g., OAuth)	Decentralized identity (e.g., Self-Sovereign Identity)
Content Creation	User-generated content on centralized platforms	User-generated content on decentralized platforms
Monetization	Advertising-based, subscription services	Token-based, decentralized finance (DeFi)
Data Ownership	Controlled by centralized entities	Owned and controlled by users
Data Storage	Centralized databases, cloud storage	Decentralized storage (e.g., IPFS)
Transactions	Traditional payment systems, credit cards	Cryptocurrencies, blockchain-based payments
Security	Security managed by centralized entities	Cryptographic security, smart contracts
Development Platforms	Web servers, APIs	Blockchain, smart contracts, decentralized apps (dApps)
Compliance and Regulation	Subject to local and international laws	Evolving regulatory landscape, jurisdictional challenges
Scalability	Scalable with server farms, cloud resources	Scalability challenges with blockchain, Layer 2 solutions
Interoperability	APIs and integrations controlled by platforms	Cross-chain interoperability, decentralized protocols

C. Evolution of Data Storage

Modern day data storage began with punch cards (1725) that were used to interact with textile looms. Almost a century later, these were used by Charles Babbage (Father of Computers) in his Analytical Engine. Magnetic tape was invented for recording sound in 1928. This gave way to video recording and data storage in the later years. Early 1960s saw the emergence of hard disk drives and floppy disk drives in 1970s, both introduced by IBM. Compact disc was introduced in 1980s together by Philips and Sony to store and play digital audio. The flash (USB) drive outmoded the floppy drives from computers in late 2000. Cloud computing concept dates to ARPANET in the 1960s (Capacity, 2019). In early 1980s, CompuServe was the first company to offer (128KB) disk storage space to the consumers, while AT&T offered web-based storage to individuals as well as businesses in 1994 (Computer History Museum [CHM], 2006). It was in 2006, when Amazon web services introduced commercial cloud storage services (Elastic Compute or EC2 / Simple Storage Service or S3). Decentralized data storage was introduced in mid-to-end 2010s: Storj (2014), Sia (2015), Arweave (2018), Bittorent File System (2019), Filecoin (2020).

D. Centralized vs Decentralized Data Storage

Data storage is a critical building block of web development (Web2 or Web3) that includes physical infrastructure and technologies to save and retrieve information digitally. Centralized (or cloud) storage systems dominate this space, however recent advancements in blockchain and peer-to-peer technologies have given rise to decentralized storage solutions.

'Centralized storage' practices storing of data in a single location or repository under the control of a single entity i.e. company or provider (Cryptobunq, 2023), like Amazon Web Services (AWS), Google Cloud Platform (GCP) or Microsoft Azure.

Table 2: Benefits and Challenges of Centralized Storage

Benefits	Challenges
• *Economies Of Scale:* Reduced cost of per unit of storage as the infrastructure scales up (Infrastructure, 2020), a significant saving over on-premise data centre. • *Operational Efficiency:* High reliability and performance due to sophisticated infrastructure and redundancy protocols (Endo et al., 2016). • *Integration and Interoperability:* Seamless connection and coordination of diverse applications, data, and services within or across multiple cloud platforms (Stephanie, 2023) • *Self-Service:* On-demand access to resources or services on the go (Spiegolski, 2024). • *Performance Assurance:* Offer dedicated customer support and higher service levels (Spiegolski, 2024)	• *Single Point Of Failure:* Making systems vulnerable to large-scale outages or data loss (*Unmasking Single Points of Failure: Vulnerabilities in Centralized Databases*, n.d.). • *Censorship And Privacy Concerns:* Users' data is controlled by a single entity (*Varying Standards of Censorship and Privacy in an Interconnected World*, n.d.) and may be subjected to censorship or unauthorized access. • *Regulatory Compliance:* These providers must navigate complex regulatory landscapes, which can be costly and complex (*Varying Standards of Censorship and Privacy in an Interconnected World*, n.d.). • *Vendor Lock-In:* Users are subject to (monopolised, non-transparent or hidden) pricing and policy changes owing to contractual obligations.

'Decentralized storage' leverages (blockchain and peer-to-peer) technology to distributes data across multiple nodes in a network, with no central authority or middleware (James, 2023).

Table 3: Benefits and Challenges of Decentralized Storage

Benefits	Challenges
• *Cost Efficiency:* Leverages unused capacity resulting in cost efficiency and advantage over centralized storage. Also, bandwidth may be cheaper because the data is already at the edge (Christine, 2021). • *Privacy, Security and Reliability:* System is resilient to cyberattack risk or reduced data breaches as no single entity has access to all the data (James, 2023), owing to no single point of failure. • *Censorship Resistance:* Decentralized storage is less susceptible to censorship and governmental control (Li et al., 2018) (He, Tang, & Wu, 2018). • *Fair Market Pricing:* No single node can demand premium (or non-monopolised pricing), thus ensures only good-quality nodes compete and survive (Anastasiya, 2021). • *Data Freedom:* Flexibility to move data across storage providers without being locked. Users face no risk of policy changes or termination of service (Cryptobunq, 2023).	• *User Experience:* Limitations around tools, proper SDKs, self-service capabilities, and User interfaces impacts the user experience and may act as a barrier to entry in the short term (Zachary, 2023). • *Performance Assurance:* Consistent Service Level Agreement (SLA) guarantees across all locations may be difficult (Zachary, 2023). Access time may be variable for edge vs distant nodes amongst other factors. • *Regulatory Uncertainty:* The legal framework for decentralized storage is still evolving, leading to potential risks (Johnson, 2019) (Zetzsche, Arner, & Buckley, 2020). • *Interoperability:* Being relatively new, the decentralized storage may lack the interoperability with existing storage solutions (James, 2023). Varying encryption and authentical methods across protocols may pose interoperability challenges.

To summarise, centralized and decentralized storage systems both offer unique benefits and face distinct challenges. It is pertinent to perform a critical analysis to ascertain which system is better placed to emerge as the future storage solution.

III. OBJECTIVES

"The risks of centralized storage (Web2 and cloud) highlight the need for decentralized solutions (Web3), empowering user control and security."

The clear and present danger witnessed within the centralized ecosystem (Web2 internet and Centralized cloud storage) is posing serious and imminent threats (privacy concerns, data breaches, and censorship) to individuals, communities, and businesses.

Web3 (or decentralized) storage leverages technology to address all these grave challenges. It is founded on the principle of placing control back to the users (and not select private entities). Therefore, the movement towards decentralized storage is organic and bound to accelerate.

To bring perspective to the above problem the paper addresses the following objectives:

- *Customer View:* To elaborate the key decision criterion that influences the buying (or switching) behaviour of individual and business towards data storage providers.
- *Business (Or Industry) View:* To explore the macro-environmental factors impacting storage players and the industry environment in which it operates.
- *Inflexion Points:* To hypothesize, the (natural, unnatural or unpredictable) events that accelerate the adoption or shift towards decentralized storage.

IV. RESEARCH METHODOLOGY

This analytical paper and uses facts or information already available (primary and secondary data sources such as reports, journals, blogs, research publications and market research). User survey (Respondents: 53) was carried out to ascertain the relative ranking of decision criterion for buying storage solution. PESTLE analysis was performed to study the key external factors (Political, Economic, Sociological, Technological, Legal and Environmental) that affect the success of a business and aid professional and executives to undertake strategic decisions.

A. Data Sources

Secondary Data: Thorough literature review on all the available literature; sources were both academics and industries. Key references are:

Reference drawn from the reports, journals, and articles from databases such as IEEE, ACM, Elsevier, and Springer to provide a theoretical framework and background context.

Blogs and market research reports were analysed in detail to find out the current trends, challenges, and opportunities within the decentralized storage landscape. Primary Data: In order to complement the secondary research, a user survey was conducted with 53 respondents. This survey helped find out the relative importance of various decision criteria while choosing storage solutions. The survey results provide valuable insights into consumer preferences and industry expectations.

B. Analytical Structure

PESTLE Analysis: This model was used to analyze how the extrinsic factors impacted the decentralized storage. The analysis encompassed:

Political: Evaluating government laws and regulations about decentralized storage.

Economic: Analysis of market trends, cost implications, and economic viability of decentralized solutions.

Sociological: Understanding consumer attitude and behavior towards data ownership and privacy.

Technological: Exploring technological advancements and their implications for decentralized storage solutions.

Legal: Analyzing the legal frameworks governing data storage and privacy rights.

Environmental: Environmental considerations of data storage and how decentralized solutions are much more sustainable.

This multifaceted methodology allows for a comprehensive analysis of the potential of decentralized storage as a foundational component of Web 3.0, informed by both empirical data and established literature.

V. KEY FINDINGS

A. Customer View

Top 10 decision criterion that influences the buying (or switching) behaviour of individual and business towards data storage providers, include:

Table 4: Decision Criteria for Buying Data Storage

#	Criteria	Criticality
1	**Capacity**	Adequate storage space to meet current and future needs is fundamental.
2	**Performance**	Fast data access and processing are crucial for efficient operations, especially in high-performance environments.
3	**Scalability**	The ability to easily expand storage as data grows without significant disruption or cost.
4	**Reliability**	Look for systems with features like RAID configurations, redundancy, and data integrity checks to minimize the risk of data loss.
5	**Data Protection**	Evaluate backup and disaster recovery capabilities to safeguard against data loss or corruption.
6	**Compatibility**	Ensure compatibility with your existing hardware, software, and infrastructure to avoid integration issues.
7	**Cost**	Compare initial purchase costs, ongoing maintenance expenses, and potential savings from energy efficiency or consolidation.
8	**Security**	Assess encryption capabilities, access controls, and compliance with relevant data protection regulations.
9	**Management and Ease of Use**	Consider how easy it is to configure, manage, and monitor the storage solution, including user interfaces and automation capabilities.
10	**Support and Vendor Reputation**	Evaluate the reputation of the vendor for reliability, customer support, and their track record with similar deployments.

Source: Gigaom (2024), Microsoft Learn (2023), DataCore Software (2021).

B. Business (or Industry) View

The macro-environmental factors impacting storage players and the industry are elaborated using PESTEL analysis below:

● Political

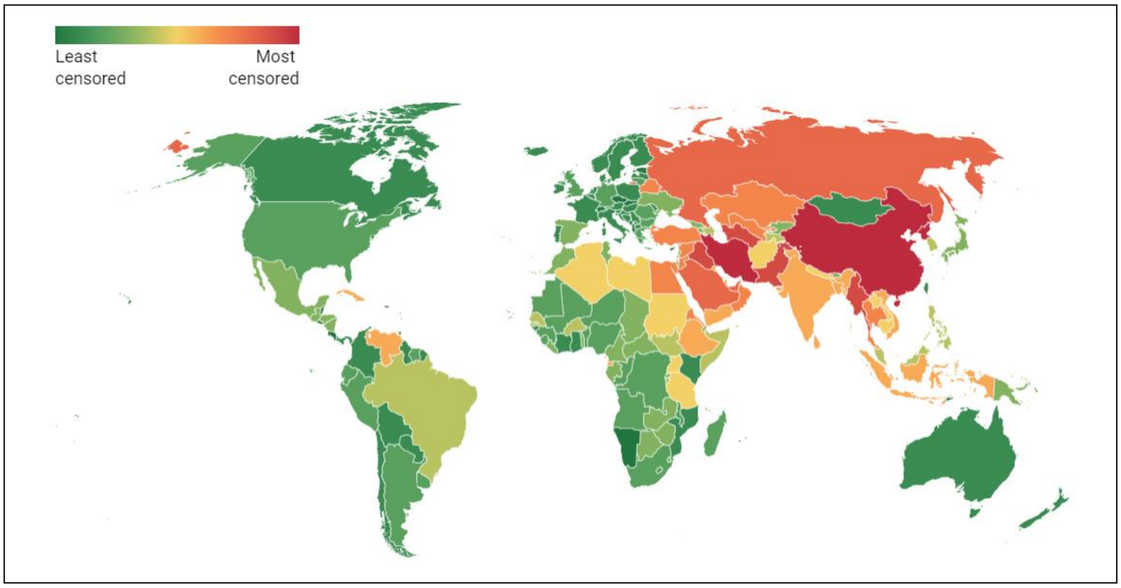
Geopolitical and political risks can significantly impact storage players, namely:

- *Data Sovereignty and Jurisdictional Issues:* Country-specific geopolitical tensions or changes in regulations such as data privacy, storage (locally), and access. Example: GDPR (EU), CCPA (California), HIPAA (US) or PDPA (Singapore)
- *Government Surveillance and Data Access:* Demands from governments for access to user data or for surveillance purposes (against contractual privacy).
- *Trade Restrictions and Sanctions:* Prohibiting service to country or entities.
- *Network and infrastructure Vulnerabilities:* Disruption to physical infrastructure including undersea cables resulting in downtime, data loss, or increased latency, impacting service reliability and customer satisfaction.
- *Market Instability and Economic Factors:* Result in reduced IT budgets, delayed projects, or increased scrutiny of technology investments.
- *Intellectual Property and Cybersecurity Concerns:* Risk of cyberattacks, espionage, or intellectual property theft; including Governments or state-sponsored attacks (Example: Stuxnet, SolarWinds) or political interference.
- *Regulatory Challenges:* Governments restrictions or regulations on specific technologies or monetization methods (cross-border transactions or payments).
- *Censorship Resistance vs. Compliance:* Clash with government regulations, including lawful intercept of user data.
- *Market Acceptance and Adoption:* Influence market confidence and adoption.

● Economic

The economic impact on data storage players can be substantial and multifaceted, influenced by various factors:

A Map of Internet Censorship and Restrictions



Types of online censorships: (1) Torrents restricted, (2) Torrents banned or shut down, (3) Adult sites restricted, (4) Adult Sites banned, (5) Political media restricted, (6) Political media heavily censored, (7) Social media restricted, (8) Social media banned, (9) VPNs restricted, (10) VPNs banned, (11) Messaging and VoIP App restrictions

Total Countries	Count of Countries with Extent of Censorships				Censored Countries
	1-3	4-6	7-10	>10	
195	107	41	25	2	175

Source: Comparitech, 2024.

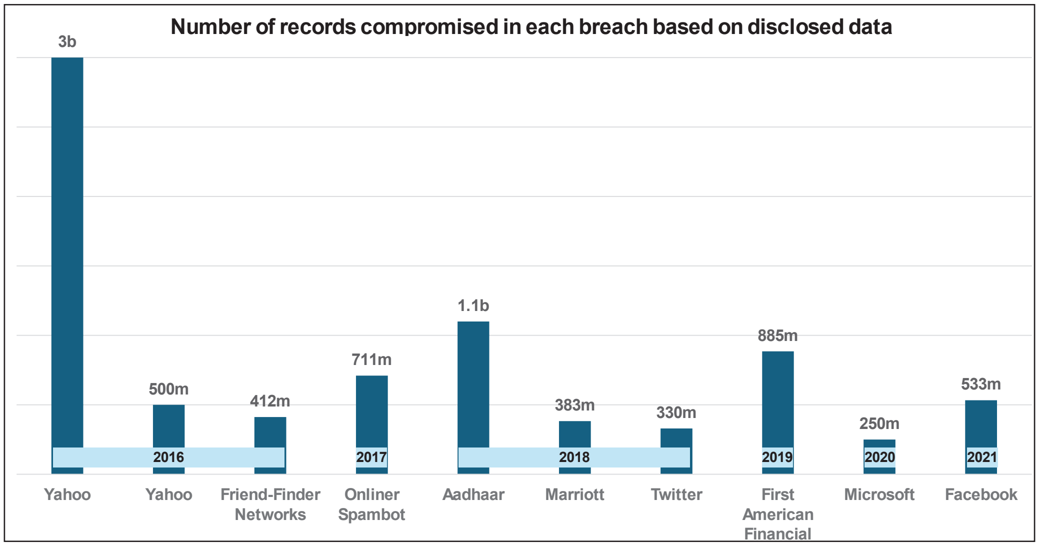
Fig. 3: Internet Censorship 2024

- *Market Demand and Growth:* During economic downturns, businesses may reduce IT spending, delay projects, or seek more cost-effective solutions.
- *Revenue and Pricing Pressures:* Competition intensifies during economic uncertainty, with users demanding innovate and more value-added services at competitive prices.
- *Capital Investment and Infrastructure Costs:* Significant investment ask to expand capacity, improve performance, and innovate with new technologies.
- *Currency Fluctuations:* Exchange rate changes impacts revenues and profitability.
- *Global Expansion and Regulatory Compliance:* Economic conditions impacts global expansion and profitability.
- *Workforce and Talent Management:* Economic cycles affect Labor markets and talent acquisition.
- *Customer Retention and Service Quality:* Economic uncertainty can impact customer retention rates.

● Social

Unique social risks that impact operations, user trust, and market acceptance of storage players:

- *Data Privacy and Security Concerns:* Scepticism on how personal data is collected, stored, and used. High-profile data breaches and scandals, such as the Cambridge Analytica, Equifax and Target, have heightened public awareness and demands for greater transparency and control over personal data.
- *Regulatory Compliance:* Failure to comply can result in hefty fines and legal consequences. Example: Google and Facebook vs GDPR fines.
- *Trust and Reputation:* Building trust is a major challenge.



Source: TechTarget, 2022.

Fig. 4: Biggest Data Breaches in History

- *User Adoption and Digital Literacy:* Providing clear, user-friendly experiences are crucial for wider adoption.
- *Cultural Attitudes Towards Technology:* In some regions, people may be more willing to share data, while in others, privacy may be a top concern.
- *Employment and Skills Development:* May demand shift in required skills and training.
- *Social Equity and Digital Divide:* Ensuring equitable access to their services can enhance social inclusion and corporate social responsibility profiles.
- *Community and User Engagement:* User engagement is critical to user satisfaction and brand loyalty.

• Technology

Advancements and innovations in technology impacts the strategies and operations of data storage players:

- *Advancement in Storage Technology:* Continuous advancement is warranted to drive storage capacities, faster data access speeds, improved durability, and advanced data security and encryption techniques.
- *Artificial Intelligence and Automation:* Significant data capacity will be required to support advanced AI use cases. Also, leveraging AI and automation to improve the efficiency and reliability of the storage solution.
- *Interoperability and Standards:* Interoperability and adherence to industry standards and protocols are crucial to ensure seamless integration with various applications, systems, and other storage providers.
- *Unstructured Data and Edge Use Cases:* Meeting the future demand unstructured data to empower advanced technologies, such as the Internet of Things (IoT), Artificial Intelligence (AI), analytics, automation, and others.

• Environmental

Operations and strategies of data storage players impact the natural environment:

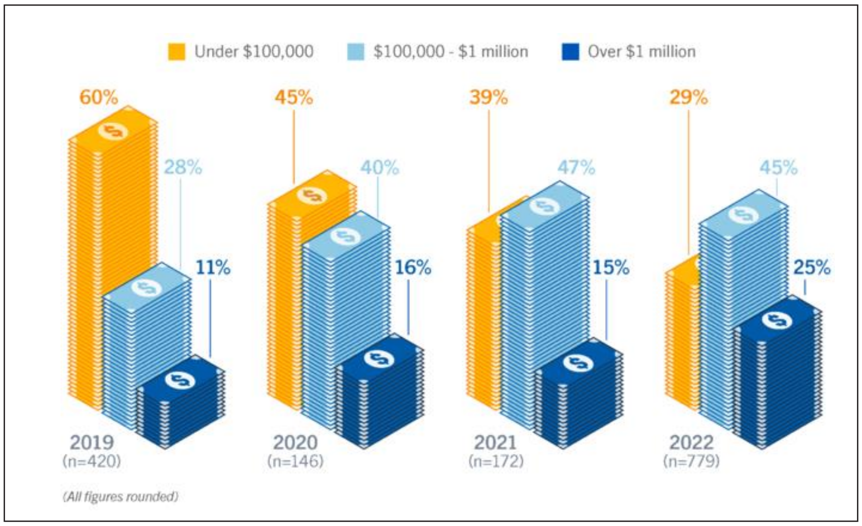
- *Energy Consumption:* Cooling systems, servers, and storage devices require continuous power, leading to high energy usage.
- *Carbon Footprint:* Progress or adherence to sustainability goals such as net zero or carbon neutrality.
- *Resource Usage and E-Waste:* Electronic wastes poses environmental and health risks if not properly managed or infrastructure is not efficiently utilized.
- *Renewable Energy Integration:* Substitution of fossil fuel generated energy with renewable alternatives: solar, wind or hydro.

- *Sustainable Practices and Innovation:* Commitments and actions to implementing sustainable practices, such as energy-efficient cooling systems, server virtualization, dynamic resource management or low-energy consensus mechanisms.

● Legal

Laws, regulations, and legal frameworks impact the operations and strategies of data storage players:

- *Data Protection and Privacy Laws:* Non-compliance can result in substantial fines and legal actions. Example: GDPR (EU), CCPA (California), HIPAA (US) or PDPA (Singapore). The GDPR, for instance, can levy fines up to 4% of a company's annual global turnover or €20 million, whichever is greater.
- *Intellectual Property Rights:* Proprietary or open-source technology, both need to adhere to the IP laws or licensing constructs.
- *Compliance and Regulatory Requirements:* Adherence to data retention, access controls, and auditing. Example: ISO/IEC 27001.



Source: Uptime Institute, 2023.

Fig. 5: Outages Costing over $1 million

- *Legal Risks and Liability:* Legal safeguards to counter data breaches, service outages, and contractual disputes.
- *Cross-Border Data Transfer:* Compliance with complex cross-border data laws. Example: EU-U.S. Privacy Shield, Standard Contractual Clauses (SCCs).

C. Inflexion Points

Natural, unnatural, or unpredictable events can accelerate the adoption or shift towards decentralized storage:

- *Major Data Breaches and Hacks:*
  - *High-Profile Incidents:* 'Repeated' large-scale data breaches affecting major companies or government entities could undermine trust on centralized storage.
  - *Ransomware Attacks:* Increased frequency and severity could highlight the vulnerability of these systems and push towards solutions less susceptible to single points of attack.
- *Government Censorship and Surveillance:*
  - *Authoritarian Measures:* Escalation of government censorship and surveillance in various countries could prompt user movement towards decentralized storage to protect their privacy and free access.
  - *Crackdowns on Digital Freedoms:* If governments impose stricter controls on internet usage and access to digital content, decentralized storage solutions that are less susceptible to control and censorship could become more appealing.
- *Regulatory Changes:*
  - *New Data Protection Laws:* Introduction of

stringent data protection and privacy laws might make centralized storage providers liable for data breaches and misuse, increasing operational costs and complexity.

- *Anti-Monopoly Actions:* Regulatory actions aimed at breaking up large tech monopolies could disrupt existing centralized storage models.

- *Technological Advancements:*
  - *Breakthroughs in Decentralized Technologies:* Significant advancements in blockchain technology, peer-to-peer networks, or other decentralized technologies that improve scalability, speed, and ease of use could accelerate adoption.
  - *Development of Killer Apps:* Creation of highly popular decentralized applications that demonstrate clear advantages over centralized counterparts could drive wider adoption of decentralized storage.
- *Economic Instability:*
  - *Financial Crises:* Economic instability or financial crises that undermine trust in traditional financial institutions and centralized storage providers could push individuals and businesses towards decentralized storage.
  - *Inflation and Currency Devaluation:* In regions experiencing high inflation or currency devaluation, decentralized solutions offering more stable and secure value storage could gain popularity.
- *Natural Disasters and Climate Events:*
  - *Infrastructure Damage:* Natural disasters causing significant damage to centralized data centres and infrastructure could highlight the fragility of centralized storage. Decentralized storage could be seen as more resilient and reliable in such scenarios.
  - *Environmental Concerns:* Increasing awareness of the environmental impact of large data centres could drive demand for more sustainable decentralized storage solutions that leverage distributed networks with lower overall energy consumption.
- *Cultural Shifts and Social Movements:*
  - *Digital Sovereignty Movements:* Growing movements advocating for digital sovereignty and control over personal data.
  - *Privacy Advocacy:* Increased awareness and advocacy for digital privacy rights.

VI. CONCLUSION

We (humans) are as good as the information we have; and data or information is the 'utmost asset' for human growth and development. Data is getting generated at an accelerated pace: expected to reach 181 zettabytes in 2025, 600+ zettabytes by 2030 and approx. 2,150 zettabytes by 2035. Less than 2% of the new data generated is getting stored. It is expected that the Global storage capacity will be growing at 19.2% CAGR and will be 9% of the volume of data generated in 2025.

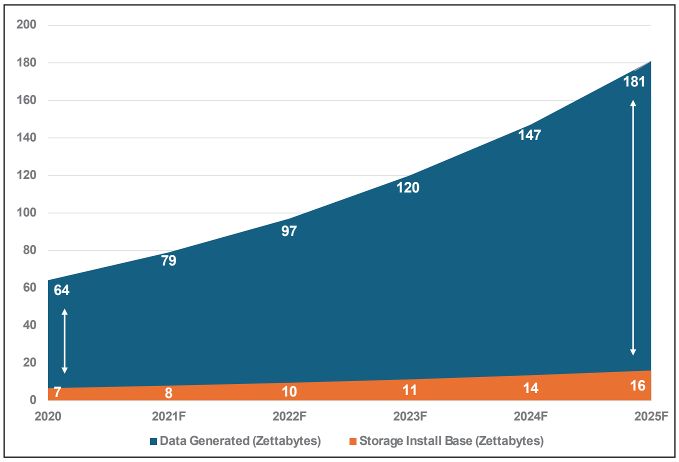


Fig. 6: Global Data Generated vs Storage Install Base (IDC, 2021)

Cloud storage alone cannot meet the future demands of data. Though a popular centralized storage mechanism in the Web2 era; it comes with risks that cannot be overlooked: single-point-of-failure, data security, ownership, privacy, and sustainability. Cloud consumes significant energy (4% of global electricity), not carbon neutral (2.5 to 3.7% of global GHG emissions i.e. more than aviation industry) and produce electronic waste (70% of toxic waste and 2% of solid waste) (Georgette, 2024).

Table 5: Cloud Storage Market by Segments (Analyst Reports)

Segment	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032
Consumer	4	4	5	5	5	6	6	7	8	8	9
Business	81	105	129	158	194	238	292	359	440	539	661
Overall	84	109	133	163	199	244	299	365	447	548	670

(All figures in USD billion)

- *Customer View:* Consumer storage ranges between 1-3% of the total cloud storage market size. The sensitivity around buying behaviour leans more towards pricing than reliability and security. Whereas business buying factors reliability, scalability, performance, and security over pricing.
- *Business View:* PESTEL analysis presents an equitable view on the critical factors to be considered by a storage player: centralized or decentralized. While, on some factors the centralized players have the advantage, while on others decentralized players have the edge. The viewpoint will drastically change in future and may lean towards decentralization.

Consumer	vs	Business
Cost		Reliability
Capacity		Scalability
Compatibility		Performance
Performance		Security
Reliability		Support & Service
Security		Capacity
Scalability		**Cost**
Support & Service		Compatibility

Fig. 7: Relative Ranking of Buying-Criteria for Data Storage Solutions (Survey)

Area	Factors	Player having edge	Rationale
Political	▪ Data Sovereignty and Jurisdictional Issues ▪ Government Surveillance and Data Access ▪ Trade Restrictions and Sanctions ▪ Network and infrastructure Vulnerabilities ▪ Market Instability and Economic Factors ▪ Intellectual Property and Cybersecurity Concerns ▪ Regulatory Challenges ▪ Censorship Resistance vs. Compliance ▪ Market Acceptance and Adoption	Decentralized	• Centralized storage faces high compliance costs; decentralized storage has regulatory uncertainty. • Centralized storage is more controllable by governments; decentralized storage resists censorship.
Economic	▪ Market Demand and Growth ▪ Revenue and Pricing Pressures ▪ Capital Investment and Infrastructure Costs ▪ Currency fluctuations ▪ Global Expansion and Regulatory Compliance ▪ Workforce and Talent Management ▪ Customer Retention and Service Quality	Decentralized	• Centralized providers benefit from large scale; decentralized can be cost-efficient • Centralized providers attract significant investment; decentralized networks rely on user participation
Social	▪ Data Privacy and Security Concerns ▪ Regulatory Compliance ▪ Trust and Reputation ▪ User Adoption and Digital Literacy ▪ Cultural Attitudes Towards Technology ▪ Employment and Skills Development ▪ Social Equity and Digital Divide ▪ Community and User Engagement	Centralized	• Centralized storage is trusted due to brand reputation; decentralized storage faces trust issues. • Centralized storage offers limited user control; decentralized storage provides greater data autonomy. • Centralized storage offers high service quality; decentralized storage may vary in service quality.
Technological	▪ Advancement in storage technology ▪ Artificial Intelligence and Automation ▪ Interoperability and Standards ▪ Unstructured data and Edge use cases	Both	• Centralized providers lead in tech advancements; decentralized storage drives innovation in blockchain. • Centralized storage has reliable, scalable infrastructure; decentralized storage faces scalability challenges.
Environmental	▪ Energy consumption ▪ Carbon footprint ▪ Resource usage and E-waste ▪ Renewable Energy Integration ▪ Sustainable Practices and Innovation	None	• Centralized data centres consume significant energy; decentralized storage utilizes existing resources but can have high energy consumption. • Major centralized providers invest in sustainability; decentralized storage's impact varies by network design.
Legal	▪ Data Protection and Privacy Laws ▪ Intellectual Property Rights ▪ Compliance and Regulatory Requirements ▪ Legal Risks and Liability ▪ Cross-Border Data Transfer	Centralized	• Centralized providers bear legal responsibility; decentralized storage has legal ambiguities. • Centralized storage faces high compliance costs; decentralized storage has complex multi-jurisdictional issues.

Fig. 8: PESTEL Analysis on Decentralized vs Centralized Storage Players

- *Accelerators:* The COVID-pandemic was the biggest market mover event that resulted in high growth across digital ecosystems. In storage industry alone, it accelerated the growth rate by 4.5% in 2020. Global phenomenon such as high impact data breaches (more than 100 million records), extent of censorship, outages (costing more than $1 million) such as Microsoft-CrowdStrike, natural calamities, technology use cases (AI, IoT, Autonomous cars or edge computing), and social (or behavioural) drifts – all can trigger accelerated adoption of decentralized storage over cloud.

Therefore, centralized storage alone cannot meet the future data demands of mankind and machines. Decentralized storage would come to the rescue not just as a sustainable alternative (utilizing underused storage resources) but also as a strong frontier for accessible, secured, persistent and censorship-resistant data storage.

"The future of storage is leaning towards decentralized, as centralized storage alone is not sustainable".

REFERENCES

Algorand Foundation. (2023). Web2 vs. Web3: What's the difference? *Algorand*. Retrieved from https://algorand.co/learn/web2-vs-web3

Anastasiya, H. (2021). Decentralized storage: The future of data storage? *PixelPlex*. Retrieved from https://pixelplex.io/blog/decentralized-storage/

Bischoff, P., & Bischoff, P. (2023, October 16). *Internet censorship 2024: A map of internet censorship and restrictions*. Comparitech. Retrieved from https://www.comparitech.com/blog/vpn-privacy/internet-censorship-map/

Bruns, A., & Bahnisch, M. (2009). *Social media: Tools for user-generated content: Social drivers behind growing consumer participation in user-led content generation,* Volume 1-State of the art.

Capacity. (2019, December 5). History of cloud storage. Retrieved from https://capacity.com/cloud-storage/history-of-cloud-storage/

Christine, A. (2021). Comparing the economics of centralized and decentralized cloud storage. *Storj*. Retrieved from https://www.storj.io/blog/comparing-the-economics-of-centralized-and-decentralized-cloud-storage

Computer History Museum (CHM). (2006). *Storage in the cloud*. Retrieved from https://www.computerhistory.org/storageengine/storage-in-the-cloud/

Cormode, G., & Krishnamurthy, B. (2008). Key differences between Web 1.0 and Web 2.0. *First Monday*.

Criteria for Choosing a Data Store. (2023, July 21). *Microsoft learn*. Retrieved from https://learn.microsoft.com/en-us/azure/architecture/guide/technology-choices/data-store-considerations

Cryptobunq. (2023). *What are centralized and decentralized storage systems*? Retrieved from https://www.cryptobunq.com/blog/centralized-and-decentralized-storage/

Duarte, F. (2024, June 13). Amount of data created daily (2024). *Exploding Topics*. Retrieved from https://explodingtopics.com/blog/data-generated-per-day

Endo, P. T., Rodrigues, M., Gonçalves, G. E., Kelner, J., Sadok, D. H., & Curescu, C. (2016). High availability in clouds: Systematic review and research challenges. *Journal of Cloud Computing, 5,* 1-15.

Felix, R. (2024). Amazon maintains cloud lead as Microsoft edges closer. *Statista*. Retrieved from https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/

Global DataSphere. (n.d.). *IDC: The premier global market intelligence company.* Retrieved from https://www.idc.com/getdoc.jsp?containerId=IDC_P38353

Harford, I. (2022). 10 biggest data breaches in history, and how to prevent them. *TechTarget*. Retrieved from https://www.techtarget.com/searchsecurity/feature/10-biggest-data-breaches-in-history-and-how-to-prevent-them

Harford, I. (2022, July 29). *10 biggest data breaches in history, and how to prevent them*. Security. Retrieved from https://www.techtarget.com/searchsecurity/feature/10-biggest-data-breaches-in-history-and-how-to-prevent-them

He, S., Tang, Q., & Wu, C. Q. (2018, November). Censorship resistant decentralized IoT management systems. In *Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (pp. 454-459).

Infrastructure, O. P. (2020). A paradigm shift towards on-premise modern data center infrastructure for agility and scalability in resource provisioning. *International Journal*, *9*(4).

James, H. (2023). What is decentralized data storage? *The Block*. Retrieved from https://www.theblock.co/learn/251865/decentralization-and-data-storage-in-cryptocurrency

Keary, T. (2023, December 5). Web4. *Techopedia***.** Retrieved from https://www.techopedia.com/definition/web4

Khan, T., Tian, W., Zhou, G., Ilager, S., Gong, M., & Buyya, R. (2022). Machine learning (ML)-centric resource management in cloud computing: A review and future directions. *Journal of Network and Computer Applications*, *204*, 103405.

Kilgore, G. (2024, May 16). *Carbon footprint of data centers and data storage per Country (per Calculator)*. 8 Billion Trees. Retrieved from https://8billiontrees.com/carbon-offsets-credits/carbon-ecological-footprint-calculators/carbon-footprint-of-data-centers/

Lawrence, A., Simon, L., & Uptime Institute. (2023). Annual outage analysis 2023. In *UII Keynote Report*. Retrieved from https://uptimeinstitute.com/uptime_assets/5f40588be8d57272f91e4526dc8f821521950b7bec7148f815b6612651d5a9b3-annual-outages-analysis-2023.pdf?mkt_tok=NzExLVJJQS0xNDUAAAGLOKD8DT_WKXcKBKyzfSYYl-Ln0amS5sNZenTtgi-NLyg8hLHFakxOayYi7wVYmE3jl7G4lpQOSeWkvyDai1ebeDT6IxNHsbbo5vmCJ_F2Bg

Mohan, V. (2024, April 2). *Top 10 evaluation criteria for software-defined storage*. DataCore Software. Retrieved from https://www.datacore.com/blog/top-10-evaluation-criteria-software-defined-storage/

Patel, K. (2013). Incremental journey for World Wide Web: introduced with Web 1.0 to recent Web 5.0 - A survey paper. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(10).

Paul, B. (2024). Internet censorship 2024: A map of internet censorship and restrictions. *Comparitech*. Retrieved from https://www.comparitech.com/blog/vpn-privacy/internet-censorship-map/

Ryan, K. (2024, May 22). Gigaom key criteria for evaluating enterprise object storage solutions. *Gigaom*. Retrieved from https://gigaom.com/report/gigaom-key-criteria-for-evaluating-enterprise-object-storage-solutions/

Spiegolski, T. (2024, February 6). *The power of centralized data storage infrastructure*. Open-E Blog. Retrieved from https://www.open-e.com/blog/the-power-of-centralized-data-storage-infrastructure/

Stephanie, C. (2023). What is cloud integration? Definition, examples, and benefits. *Digibee*. Retrieved from https://www.digibee.com/blog/what-is-cloud-integration/

Stevens, R. (2022, December 21). What is 'Web5' and how is it different from Web3? *CoinDesk***.** Retrieved from https://www.coindesk.com/learn/what-is-web5-and-how-is-it-different-from-web3/

Team, I. (2024, August 8). *Web 3.0 explained, plus the history of Web 1.0 and 2.0*. Investopedia. Retrieved from https://www.investopedia.com/web-20-web-30-5208698

Tim Berners-Lee. (n.d.). Retrieved from https://www.w3.org/People/Berners-Lee/

Unmasking Single Points of Failure: Vulnerabilities in Centralized Databases. (n.d.). Retrieved from https://inery.io/blog/article/unmasking-single-points-of-failure-vulnerabilities-in-centralized-databases/

Varying Standards of Censorship and Privacy in an Interconnected World. (n.d.). ISACA. Retrieved from https://www.isaca.org/resources/isaca-journal/issues/2021/volume-2/varying-standards-of-censorship-and-privacy-in-an-interconnected-world

Volume of Data/Information Created, Captured, Copied, and Consumed Worldwide from 2010 to 2020, with Forecasts from 2021 to 2025. (2024). Statista. Retrieved from https://www.statista.com/statistics/871513/worldwide-data-created/

What is Web3? (2023, October 10). McKinsey & Company. Retrieved from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-web3

Yasar, K. (2023, January 30). *Web 2.0*. WhatIs. Retrieved from https://www.techtarget.com/whatis/definition/Web-20-or-Web-2

Zachary, T. (2023). Exploring the decentralized storage landscape. *Chainstack*. Retrieved from https://chainstack.com/exploring-the-decentralized-storage-landscape/

Zerocap. (2024). Web 2 versus Web 3: Key differences. *Zerocap*. Retrieved from https://zerocap.com/insights/snippets/web2-versus-web3-comparison/

Zetzsche, D. A., Arner, D. W., & Buckley, R. P. (2020). Decentralized finance. *Journal of Financial Regulation*, *6*(2), 172-203.