**Professor Charles Belle,**
***Filecoin Foundation for the Decentralized Web Teaching Legal Fellow***
**and**
**Reuben Rosenberg,**
***Research Fellow***

**June 6, 2025**

# Government Onchain Open Data (GOOD) Project Report

UNIVERSITY OF SAN FRANCISCO | Center for Law, Tech, and Social Good

## 1 EXECUTIVE SUMMARY

Government agencies spend millions annually on fragmented, expensive cloud storage while struggling to provide transparent, accessible public data. Current solutions create vendor lock-in and require technical expertise that most civic organizations don't possess.

The Government Onchain Open Data (GOOD) Project proposes to create the first user-friendly platform enabling any civic technologist or government employee to upload public datasets to Filecoin's decentralized network with simple drag-and-drop functionality—no blockchain expertise required.

GOOD provides a living repository where data updates automatically while preserving immutable historical versions, supporting both real-time government operations and permanent public accountability.

Our technical evaluation of Filecoin storage solutions demonstrates that decentralized storage can reduce government storage costs while eliminating vendor dependencies. However, adoption requires accessible tools—the critical gap our platform addresses. A platform must enable, for example, a city clerk to preserve budget documents without understanding blockchain mechanics.

**Democratic Benefits:**

- **Transparency:** Citizens can independently verify data integrity through blockchain proofs
- **Permanence:** No single entity can delete or modify public records
- **Accessibility:** Standardized platform reduces costs for smaller municipalities
- **Sovereignty:** Government data remains under public control, not corporate platforms

**Funding:** Funding is not requested in this report, but would be required would to develop a production-ready prototype, conduct user testing with civic organizations, and host community data-thons to demonstrate real-world adoption.

A scalable platform democratizes access to decentralized storage for government entities of all sizes, creating a foundation for transparent, cost-effective, and permanently accessible public data infrastructure.

## 2 INTRODUCTION

### 2.1 The Goal

The GOOD Project creates a user-friendly platform enabling government entities to upload public datasets to Filecoin through simple drag-and-drop functionality. This living repository maintains immutable historical records while supporting real-time data updates, bridging democratic transparency values with practical data management needs.

### 2.2 Key Takeaways

- **Accessibility Gap**: Current Filecoin tools are too complex for everyday government users, creating an adoption barrier.
- **Cost Advantages**: Decentralized storage can offer significant long-term cost savings compared to traditional cloud providers, but there are significant upfront capital costs in time and tools.
- **Democratic Alignment**: Filecoin's transparency and immutability naturally align with democratic values of open government.
- **Standardization Benefits**: A unified platform can reduce costs and complexity for smaller entities while improving data discoverability.

### 2.3 Why move public data onchain?

**Current Pain Points:**

- Expensive vendor lock-in with centralized data storage providers can be cost prohibitive for many public budgets
- Technical barriers excluding non-technical government staff from data preservation
- Fragmented, costly systems are particularly challenging for smaller municipalities
- Lack of transparency and citizen verification mechanisms
- Complex version control without historical preservation

**Why move on-chain:**

- **Cost Efficiency:** Potential savings compared to traditional cloud storage
- **Accessibility:** User-friendly tools eliminate technical expertise requirements
- **Transparency:** Blockchain verification enables citizen accountability
- **Standardization:** Unified platform reduces complexity for smaller entities

## 3 RESEARCH METHODOLOGY

### 2.4 Research Goal

Our research focused on identifying technical and design requirements to enable civic technologists, especially those lacking substantial technical skills, to upload data easily to Filecoin through a simplified interface. We evaluated multiple storage solutions, including Lotus, Lighthouse, Web3.Storage, and Storacha, comparing trade-offs among ease of use, performance, and long-term cost to determine the most viable path for widespread adoption by public entities.

### 2.5 Next Steps

- Additional funding would be required to develop a web app prototype, conduct community testing, and do further research on the cost savings of using Filecoin versus existing centralized databases.
    - **Product Requirement Document**: Generate and publish on Github an open source document
    - **Prototype Development.** Develop a working dashboard that enables users to collect, automatically organize and label the data using agentic tools, and upload the data to the Filecoin network.
    - **Host a Data Jam.** Organize a datajam with a civic technology group and students to test the tool

### 2.6 Conclusion

Filecoin's decentralized protocol provides a possible solution for government data storage by eliminating vendor dependencies. The research shows that with proper tooling, decentralized storage can be made accessible to non-technical users while supporting the transparency and permanence requirements of democratic institutions. There are, however, several challenges that include whether there is competitive pricing and the tools to overcome existing legacy infrastructure.

Following creation of a website and Data-Thon, the website would serve as a platform for users to interact with the Filecoin network and support the initiative to decentralize government data. The platform could serve as a persistent gateway to decentralized government data.

### 3.1 Research Approach

Our research evaluated the existing Filecoin protocol and onramp solutions to determine what, if anything was missing to enable Filecoin to operate as an onchain solution for government data. Our research approach combined technical experimentation with user-centered design principles:

- **Technical Testing**: Evaluated existing Filecoin onramps including Storacha, Lighthouse, and Singularity through test uploads
- **Performance Analysis**: Assessed upload time, failure rates, cost structures, and user experience across platforms
- **Architecture Review**: Studied Filecoin's technical documentation to understand node operation and deal-making requirements
- **Usability Research**: Identified pain points and priorities for non-technical users interacting with decentralized storage
- **Cost-Benefit Analysis**: Compared long-term operational costs between centralized and decentralized solutions

### 3.2 Project Criteria

Our evaluation framework prioritized four key criteria:

**Simplicity**: Upload processes must be achievable within minutes by first-time users without coding experience or technical knowledge of blockchain systems.

**Transparency**: Users must receive content identifiers (CIDs) and storage confirmations, enabling direct verification through Filecoin or IPFS networks.

**Metadata Integration**: Datasets must include contextual information for discoverability, making stored data meaningful and searchable for future users.

**Cost Sustainability**: Long-term hosting must remain financially viable for public entities with limited budgets.

### 3.3 How Filecoin Works

Filecoin is a decentralized network that allows users to store and retrieve files securely. Unlike traditional cloud storage options that rely on a single company (e.g. Google or AWS), Filecoin uses a network of independent storage providers: individual users who are rewarded for offering storage space.

This decentralized storage format means no single entity controls all the data, while still maintaining security and a comparable low price for storage.

Filecoin's storage process

Users upload data onto the Filecoin network. Prices for storage are negotiated with individual groups who have offered to provide storage, called storage providers or SPs. Once a price and length of storage are agreed on, the storage provider will securely store the data while regularly proving that the data is secure.

To retrieve data, a user negotiates a price with a user who interacts with the Filecoin network to retrieve the file called Retrieval Provider (RP). Then the RP will retrieve the file and send it to the user.

Tools called storage onramps are available to avoid complicated aspects of decentralized storage. These tools allow a user to upload a file, and a third party will handle all storage and retrieval tasks behind the scenes.

SEE DIAGRAMS ON NEXT PAGE and APPENDICES A and B

#### 3.3.1 Issues Identified

- **Implementation vs. Cost Trade-offs**: Balancing ease of deployment against long-term operational expenses
- **User Experience Balance**: Maintaining simplicity while capturing necessary metadata for data discoverability
- **Infrastructure Requirements**: Managing intermediate storage needs and server capacity considerations
- **Technical Complexity**: Bridging the gap between Filecoin's technical requirements and user expectations

#### 3.3.2 Onramp Tools

Several tools currently enable data uploads to the Filecoin network, each designed for different user needs and technical capabilities. Storacha offers a developer-focused approach with subscription-based pricing ($10/100GB monthly) and automated storage deals through simple CLI commands, making it accessible for technical users but lacking a web interface for non-technical government staff. Lighthouse provides the most user-friendly experience with drag-and-drop web uploads and both annual ($99/150GB) and lifetime storage options, though it offers limited customization for specific government requirements. During testing, however, Lighthouse failed to upload data multiple times for either researcher. Singularity functions as a comprehensive data preparation tool that integrates with existing storage systems like S3 and Oracle, requiring users to manage their own storage provider relationships and deal negotiations—making it powerful for bulk operations but complex for everyday use. SEE APPENDIX C for more details.

The fundamental tradeoff across these solutions centers on ease of use versus cost optimization and control. While Lighthouse excels in simplicity with its intuitive web interface, it provides the least flexibility for customizing storage parameters or integrating with existing government systems. Storacha strikes a middle ground with good automation and developer API support, but it uses a CLI-based approach and subscription model that may not suit government users who need occasional uploads or lack technical expertise. Singularity offers the highest degree of customization and can achieve the lowest long-term costs through direct miner negotiations, but requires substantial technical knowledge and ongoing management overhead.

For government entities seeking to preserve public data, none of these existing tools adequately address the specific needs of non-technical civic users who require simple interfaces, transparent processes, and cost-effective solutions—creating the clear market gap that the GOOD Project addresses.

Figure 2: how miners operate on Filecoin. APPENDIX D compares types of servers

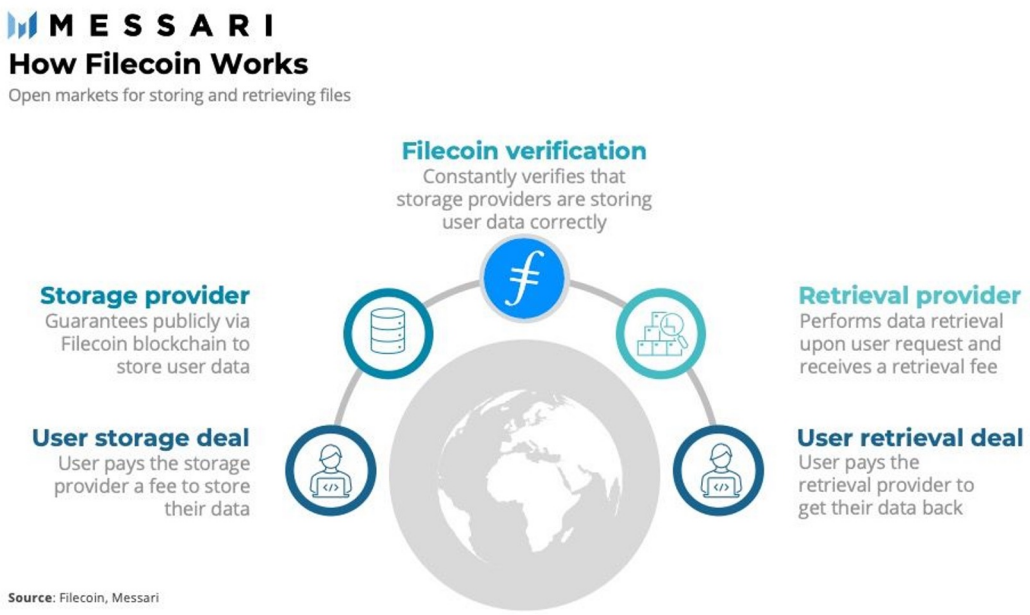


**Figure 1.**

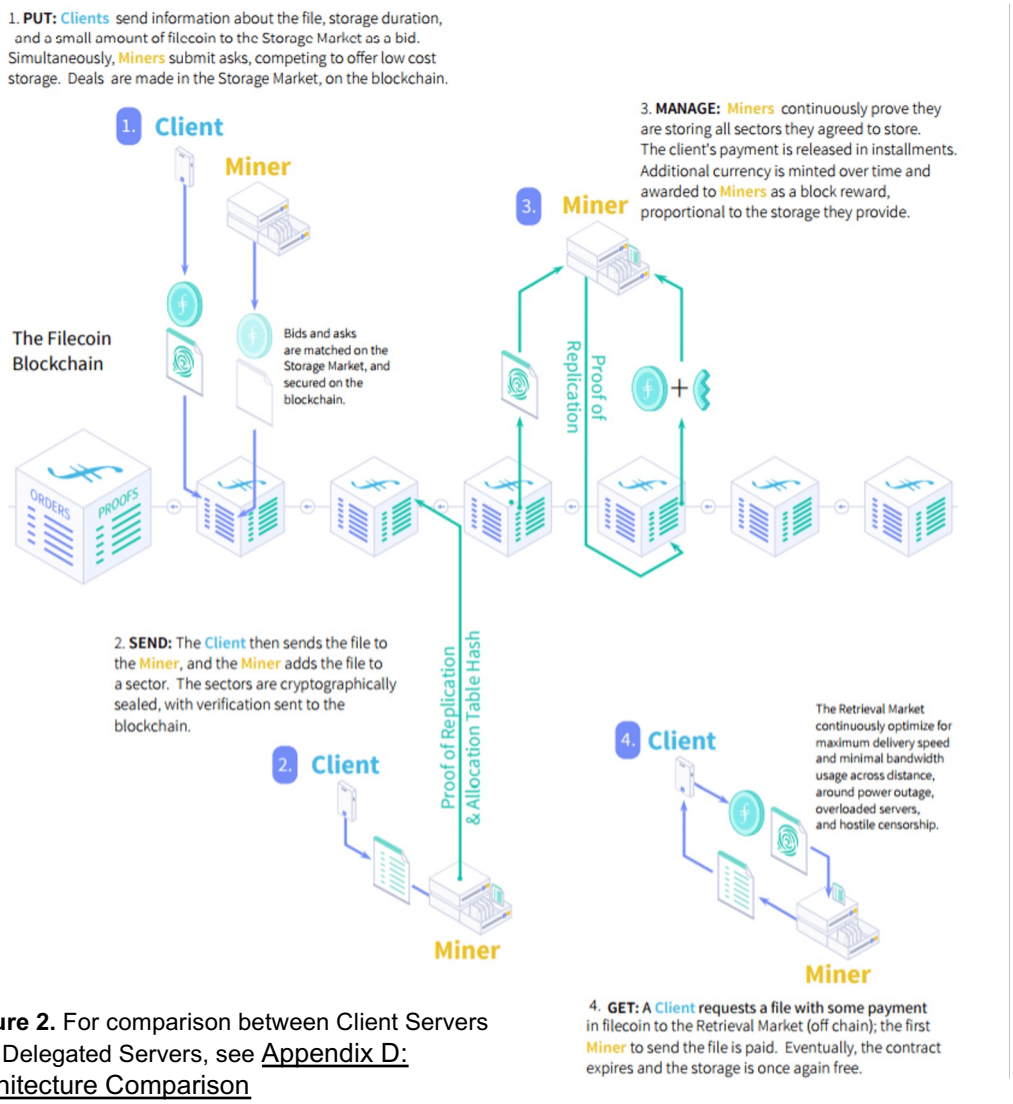


**Figure 2.** For comparison between Client Servers and Delegated Servers, see Appendix D: Architecture Comparison

## 4 DISCUSSION

### 4.1 Proposed Prototype Design

The final recommendation of The GOOD Project's research is a web application for civic technologists and government officials. A prototype can be built for testing with civic technology groups and students with grant support, through, for example, either Filecoin's Open Grant or a FIL Builder Next Step Grant.

Figure 3 on page 9 provides a demo version of what this application may look like. The link to the demo version can be found here: https://preview--sf-data-onchain-demo.lovable.app/

### 4..2 Core Takeaways

**Primary Finding**: Despite Filecoin's technical robustness, the ecosystem lacks accessible, cost-effective onramps for average users. The adoption bottleneck is usability, not storage capability.

**Democratic Alignment**: Filecoin's decentralized architecture naturally supports democratic values of transparency, accessibility, and public ownership of information. Unlike centralized providers, no single entity controls access to public data.

**Cost Efficiency**: Decentralized storage provides competitive pricing while eliminating vendor lock-in, particularly beneficial for smaller government entities with limited IT budgets.

**Living Repository Capability**: The platform enables frequent data updates while maintaining immutable historical records, supporting both current access needs and archival requirements.

### 4.2.2 Prototype Purpose

The prototype seeks to validate two critical assumptions: user accessibility and economic viability. Usability testing with civic technology groups and government employees will assess whether non-technical users can successfully upload datasets within minutes without blockchain expertise, measuring completion rates, error frequency, and user satisfaction. Economic testing will evaluate the platform's sustainability through a freemium model where basic uploads remain free for civic tech and government entities, with premium features (bulk uploads, advanced metadata tools, priority support) available through subscription tiers.

This approach ensures public data preservation remains accessible to all government entities, while generating revenue to support platform operations and potential transition from Storacha integration to native Filecoin node operation for long-term cost optimization.

### 4.2 Why Special Development for Civic Technologies?

Government transparency and civic engagement require tools designed specifically for public sector needs and non-technical users. Generic storage solutions fail to address the unique requirements of civic data management, including accessibility, transparency, and public accountability.

**Democratizing Data Preservation**: Current Filecoin tools require technical expertise that excludes the majority of government employees, elected officials, and civic organizations. A city clerk uploading budget documents or a county recorder preserving historical records should not need to understand blockchain mechanics or command-line interfaces. Our tool removes these barriers, enabling direct participation in decentralized preservation by the people responsible for creating and maintaining public data.

**Transparency and Public Trust**: Government data storage decisions directly impact public trust. When agencies rely on proprietary cloud services, citizens cannot independently verify data integrity or access policies. Our tool provides public CIDs and blockchain verification, allowing citizens to independently confirm that government data remains unaltered and accessible. This transparency strengthens democratic institutions by making government data practices auditable.

**Civic Infrastructure**: Just as governments invest in physical infrastructure like roads and utilities, they must invest in digital infrastructure. Decentralized storage represents a form of digital infrastructure that remains under public control while leveraging market efficiencies. The GOOD platform provides a foundation for this civic digital infrastructure.

**Emergency Preparedness**: Government agencies need resilient data storage for emergency planning, disaster response, and continuity operations. Centralized cloud services can fail during regional emergencies or cyberattacks. Decentralized storage ensures that critical government data remains accessible even when traditional infrastructure is compromised.

### 4.3 Supporting Filecoin's Data Preservation Ecosystem

The GOOD project aligns with and extends Filecoin's existing data preservation initiatives while addressing specific gaps in civic technology adoption.* Filecoin's data preservation programs have demonstrated the network's capacity for large-scale archival projects, but accessibility remains a significant barrier for government and civic users.

**Building on Established Success:** Filecoin's collaboration with organizations such as the Internet Archive and various research institutions has demonstrated the network's reliability in preserving valuable datasets. Our tool extends this success by creating pathways for smaller-scale but equally important civic data preservation.

**Addressing the Accessibility Gap:** While existing Filecoin preservation efforts often focus on large datasets and technical users, government agencies typically work with smaller, diverse datasets and have limited technical resources. The GOOD platform bridges this gap by providing the same preservation benefits through an accessible interface designed for civic users.

**Complementary Focus Areas:** Rather than competing with existing preservation programs, GOOD targets an underserved user base: local governments, civic organizations, and individual public servants who need straightforward data preservation tools. This creates a complementary ecosystem where large-scale archival projects and grassroots civic preservation coexist and strengthen each other.

**Scaling Decentralized Preservation:** By making Filecoin accessible to government users at all levels, from federal agencies to local nonprofits, the GOOD platform significantly expands the network's role in preserving public interest data, increases adoption, strengthens the overall Filecoin ecosystem, while advancing public transparency and accountability.

*See, e.g., https://github.com/orgs/data-preservation-programs/projects/18/views/1

### 4.4 Repository versus Archive

The GOOD project fundamentally differs from traditional digital preservation models like the Internet Archive in both technical architecture and philosophical approach. While the Internet Archive operates as a centralized repository with mirrors and backup systems, this platform would leverage Filecoin's decentralized network to create a truly distributed preservation model.

**Centralized vs. Decentralized Preservation:** Server outages, funding challenges, or legal actions can compromise access to millions of preserved documents. Governments (and the Internet Archive) have faced recent DDoS attacks that highlight these vulnerabilities. In contrast, the GOOD platform distributes government data across multiple independent storage providers on the Filecoin network, ensuring that no single entity controls access to critical public information,removing vulnerabilities to targeted attacks.

**Immutability and Verification:** Unlike traditional archives, where content can be modified or removed, data stored through cryptographic content identifiers (CIDs) makes tampering immediately detectable. This immutability is particularly crucial for government data, where maintaining the integrity of public records, election data, and policy documents is essential for democratic accountability.

**Economic Sustainability:** A decentralized approach leverages market-driven storage pricing on Filecoin, potentially reducing long-term costs while creating economic incentives for preservation. Storage providers are compensated directly for maintaining data, creating a self-sustaining preservation ecosystem. This means no reliance on donations and grants, which can often lead to ongoing funding uncertainty.

### 4.5 Mockup Demonstration: Early Implementation

To validate this approach and demonstrate feasibility, we developed an early mockup using Lovable.dev that showcases the core user experience and technical architecture of the GOOD platform. This mockup (click here), illustrates how complex Filecoin interactions can be simplified into an intuitive interface for government users.

**User Experience Validation:** The mockup features the drag-and-drop upload interface and minimal metadata entry forms described in our technical framework. Users can upload files, receive CIDs, and understand the preservation process without requiring blockchain expertise.

**Technical Architecture Proof:** The mockup demonstrates the feasibility of our proposed backend architecture, showing how Storacha integration can be abstracted behind a user-friendly interface. The implementation validates our assumption that government users can benefit from Filecoin's preservation capabilities without directly interacting with complex storage protocols.

**Iterative Development Approach:** This early mockup serves as a foundation for user feedback and iterative improvement. By deploying a working example early in the development process, we can gather input from potential government users and civic organizations to refine the interface and functionality before full-scale implementation.

**Scaling Pathway:** The mockup architecture is designed to support the transition from Storacha integration to native Filecoin node operation as usage scales. This dual-path approach ensures that the platform can launch quickly while maintaining the flexibility to optimize costs and performance as the user base grows.

Successful implementation of this mockup will demonstrate that the GOOD project's goals are not only necessary but technically achievable with current Filecoin infrastructure and tooling.

#### 4.5.2 Frontend Requirements

Prioritizes simplicity and accessibility:

- **Drag-and-Drop Interface**: Intuitive file upload without technical knowledge requirements
- **Minimal Metadata Forms**: Streamlined data entry captures essential information
- **Progress Feedback**: Real-time upload status and CID generation upon completion
- **Search Functionality**: Ability to discover and access previously uploaded datasets
- **Responsive Design**: Accessibility across devices and technical skill levels

#### 4.5.3 Backend Requirements

Systems handle all Filecoin network interactions:

- **Data Processing**: Automatic validation and CAR file conversion for network compatibility
- **Metadata Management**: Indexing and storage of dataset information for discoverability
- **Network Integration**: Automated storage deal creation and management on selected onramp
- **Version Control**: Tracking and maintaining relationships between data updates
- **Retrieval Management**: Simplified access to stored data through content identifiers

#### 4.5.4 Architecture Options

*Option 1: Storacha Integration*

- **Advantages:** Rapid deployment, well-documented APIs, automated deal management
- **Considerations:** Service fees per upload, dependency on third-party provider
- **Best For:** Quick implementation and testing phases

*Option 2: Native Filecoin Node + Dealmaker*

- **Advantages:** Lower long-term costs, full control over storage parameters
- **Considerations:** Higher initial complexity, requires Filecoin expertise
- **Best For:** Long-term sustainable operations with technical resources

**Recommended Approach**: Begin with Storacha integration for rapid deployment and user testing, with migration path to native implementation as usage scales and technical capacity develops.

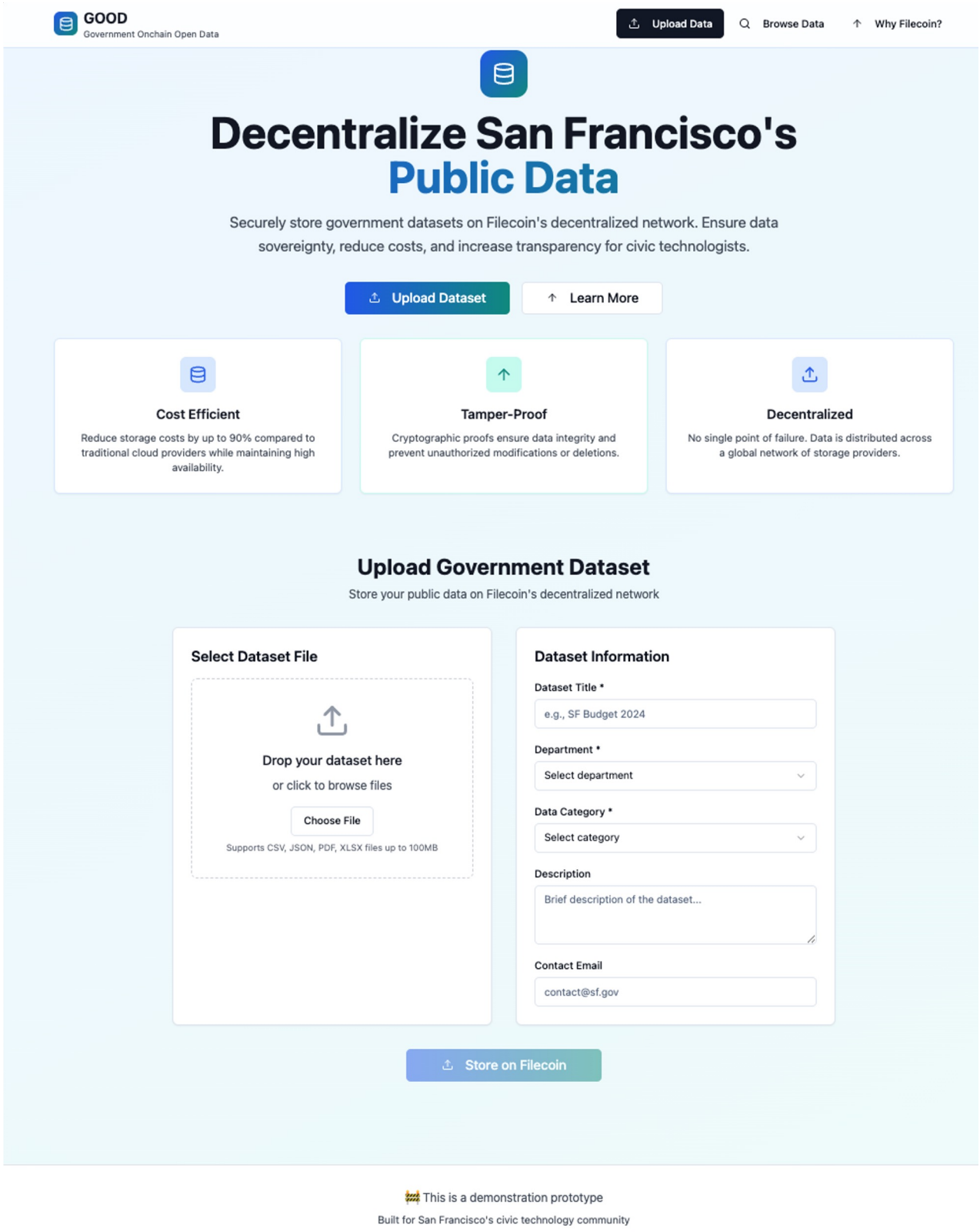


**Figure 3.**
See demonstrating website: https://preview--sf-data-onchain-demo.lovable.app/

## 5 CONCLUSION

The GOOD Project establishes that decentralized storage is technically viable for government entities. Our research demonstrates that with proper tooling, Filecoin can transform government data management from expensive, vendor-dependent systems into transparent, cost-effective, and citizen-verifiable infrastructure. A prototype would validate the core premise: complex blockchain interactions can be abstracted into simple drag-and-drop interfaces. This would enable citizens and government employees to preserve public data on decentralized networks without technical expertise, removing the primary barrier to adoption. Measurable benefits for democratic governance include cost savings by local government, increased transparency and accessibility of public data, data permanence, and local government data sovereignty. The next step requires technical development through a production-ready platform and partnerships with civic technology organizations for user testing.

The GOOD Project represents more than a storage solution—it's a pathway to digital sovereignty for democratic institutions. By enabling government entities to preserve public data on decentralized networks, we create infrastructure that serves citizens rather than corporate interests. This foundation supports long-term transparency, reduces public costs, and ensures that democratic accountability survives technological change.

**ACKNOWLEDGEMENTS**

We would like to thank Jefferson Sankara, Technical Product Manager at Filecoin for this assistance and support during the course of the research project.

**AUTHORS**

Professor Charles Belle

Charles Belle is the Filecoin Foundation for the Decentralized Web Teaching Fellow at the University of San Francisco School of Law's Center for Law, Technology, and Social Good. His research delves into decentralized governance, and he teaches a course on Decentralized Finance Law (DeFi). He is an adjunct professor at USF School of Law, where he has taught Decentralized Finance Law and Business Associations. Belle also holds a non-residential fellowship at Stanford University's Center for Internet and Society (CIS), and he was previously a fellow at UC Berkeley's Center for Technology, Society & Policy.

Reuben Rosenberg, Research Fellow

Reuben Rosenberg graduated cum laude from the University of California, Los Angeles (UCLA) in 2024 with a B.S. in Mathematics. He is currently pursuing a Master's in Computer Science at the University of Texas at Austin, specializing in machine learning. His research interests include decentralized technologies and artificial intelligence, with a particular focus on applying advanced computing techniques to expand public access to information and technology.

## 6 APPENDICES

### Appendix A: Data Journey Description

**Technical Data Flow (Storacha Implementation)**

Storage Process:

1. File uploaded to Storacha platform
2. Conversion to CAR (Content Addressable Archive) format
3. Upload to IPFS with immutable CID assignment
4. Automated Filecoin storage deal creation
5. Storage provider acceptance and block sealing with proof of storage
6. Indefinite storage on Filecoin network until deal expiration

Retrieval Process:

1. User provides saved CID for data access
2. Retrieval through IPFS HTTP gateway via Storacha
3. Direct access using formatted URL: https://[CID].ipfs.w3s.link/

**Non-Technical User Journey**

For Government Users:

1. Simple file upload generates unique, permanent identifier
2. Storacha handles all technical aspects of decentralized storage
3. Regular cryptographic proofs ensure data integrity and availability
4. File retrieval using provided permanent link

### Appendix B: Example/Required Metadata Fields

**Essential Metadata:**

- **Dataset Title**: Descriptive name for public identification
- **Source Agency**: Originating government entity
- **Publication Date**: Initial release timestamp
- **Update Frequency**: Expected refresh schedule
- **Data Format**: File type and structure information
- **Access Level**: Public availability classification
- **Contact Information**: Responsible official or department
- **Description**: Content summary and use case information
- **Keywords**: Searchable terms for discoverability
- **Related Datasets**: Connections to associated data

## Appendix C: Tools Comparison

| Feature | Storacha | Singularity | Lighthouse | Direct Filecoin Miner |
|---|---|---|---|---|
| **Primary Use Case** | General-purpose storage with developer APIs | Data preparation and deal execution tool | Decentralized storage with simplicity focus | Full control over Filecoin storage deals |
| **Cost Structure** | Subscription-based: $10/100GB/month, $99/150GB/year | No additional fees beyond direct deal-making | Annual and lifetime storage options | Market-rate based on miner negotiations |
| **Ease of Use** | Simple CLI, no website drag-and-drop | Medium/Hard - requires market interactions and complex CLI | Very Easy - drag-and-drop website interface | Hard - requires technical Filecoin knowledge |
| **Automation Level** | Automated storage and retrieval | Requires user-provided storage provider lists | Entirely automated upload process | No automation - manual deal execution |
| **Retrievability** | HTTP gateway, IPFS CLI, or curl access | CLI-based retrieval | IPFS JavaScript or website download | Manual market requests and retrieval provider interaction |
| **Customizability** | Medium - limited to platform features | High - customizable tool for bulk operations | Low - simplicity-focused with limited options | Very High - full control over all parameters |
| **Integration Support** | JavaScript SDK available | Easy integration with existing storage (S3, Oracle) | JavaScript API available | No built-in support, manual setup required |
| **Best For** | Balanced automation and developer support | Bulk storage with existing infrastructure | Easiest onboarding for new users | Advanced users optimizing storage costs |

## Appendix D: Architecture Comparison

Client/Server vs. Delegated Architecture

| Aspect | Client/Server | Delegated Server |
|---|---|---|
| **Setup Complexity** | Simple user uploads to server, but requires hardware for intermediate storage | Moderate setup with UCAN delegation, no additional hardware needed |
| **User Experience** | User-friendly with backend handling all Storacha calls | User-friendly with potential delegation complexity (manageable in frontend) |
| **Trust Requirements** | Users must trust server to handle data appropriately | No trust needed - direct user-to-Storacha uploads |
| **Server Costs** | Hardware costs plus Storacha fees | Storacha fees only |
| **Scalability** | Requires additional hardware and storage allocation | Easily scalable through Storacha allocation increases |