Data Protection & Privacy Impact Assessment on the Private.Storage Application

Origin/Authors: Carey Lening, Principal Consultant, Castlebridge

Creation Date: June 2022

Revision Date: January 2024

Version: V2.1

Version Control

Date	Author	Version	Change Reference
March 2022	CL	0.1	Initial draft
May 2022	CL	1.0	Updates, adding images
December 2022	CL	1.1	Updates
August 2023	CL	2	Major revision
January 2024	CL	2.1	Final Version

Ratification

Status	Approved for Publication
Reviewed	Liz Steininger, President, Private.Storage
Attendees	Lauri Hmel, COO, Private.Storage Dororthee Landgraf, General Counsel, Least Authority Group

Table of Contents

Summary

Private.Storage commissioned Castlebridge to undertake a Data Protection & Privacy Impact Assessment (DPIA) of their new distributed storage platform. To demonstrate their commitments to preserving the privacy and security of their users, Private.Storage has agreed to make this DPIA public.

The Private.Storage platform consists of a distributed, open-source, privacy-preserving storage service that allows users to upload folders and files to a secure, end-to-end encrypted system. It relies on three underlying open-source platforms, Tahoe-LAFS, Gridsync, and Magic Folders. It was built from the ground up with privacy-by-design principles in mind, including robust client-side encryption, use-based payments that limit collection and storage of data, and a unique "accountless authentication" system that eliminates the need to create user accounts, IDs and passwords.

This Assessment was conducted by members of the team at Castlebridge, and relied on the following information:

- Private.Storage's public-facing documentation
- Review of the Tahoe-LAFS and Gridsync platforms and documentation
- Interviews with members of the development, engineering, security, and leadership teams
- An examination of the platform and codebase by members of the Castlebridge team.

The DPIA considers an assessment of the technical environment, the current state of privacy and data protection laws in the United States and the European Union, and includes recommendations for improvements to identified risks.

Ten risks were identified as part of this DPIA. There were three medium risks, three medium-low risks, and four low risks. Of those, four were identified as technology risks, three as process-based risks, and three as legal risks.

Update as of December 2022: Private.Storage has eliminated all but four risks (identified in Table 1a. The remaining risks primarily relate to technology risks which are not reasonably capable of being solved given the current state of security and technical controls available.

Table 1: Initial Identified Risks

Risk	Description	Risk Rating	Risk Category
No lawful basis for processing is declared in the Privacy Notice	Under the GDPR, controllers must provide a legal reason for why they process data. There are six legal bases for processing data: a) consent, b) performance of a contract, c) compliance with a legal obligation, d) legitimate interests of the controller, e) vital interests and f) public interest.	**Medium**	**Legal**
No designation of an EU Representative	Private.Storage has not formally designated a representative in the EU, as is required by Article 27 GDPR. Without an appointed representative / establishment in the EU, this leaves Private.Storage open for regulation from all EU Data Protection Authorities.	**Medium**	**Legal**
Details on where data is stored, and for how long are not well-defined	Under the GDPR, controllers must provide details on where data is stored physically, and any sub processors/third parties used, This information is currently not well-documented in the privacy notice.	**Medium**	**Process**
The voucher / capability string details are not exposed to the user	Vouchers are how the Private.Storage system verifies that a user has paid for service. At the payment screen, no details about the voucher code are provided to the user. If a user loses access to their local machine, they cannot gain access to their data, even if they save the recovery key, without paying for additional storage. This creates a loss of availability to their data. The risk here is less about data protection, and more about usability, particularly if multiple steps must be performed and the transaction process is not seamless, or errors not obvious to the user.	**Medium-Low**	**Technology**

Risk	Description	Risk Rating	Risk Category
Personal data of customers is shared with Stripe for payment processing	Currently, payments are processed using Stripe as a subprocessor. Stripe stores numerous pieces of personal information, including name, address, cardholder data, etc. This data is not maintained on Private.Storage servers, but it is accessible to a limited subset of Private.Storage employees (currently < 3), which means that it is currently possible to match the redemption of a voucher back to a customer who paid using Stripe.	**Medium-Low**	**Process**
Cookie details should be clearer	Currently, a Matomo session cookie is stored on a user's device and two persistent Stripe fraud prevention cookies are installed on a user's device. The Privacy Notice does not explicitly detail these cookies (though it does generally discuss the processing and purpose of Stripe and Matomo cookies in the abstract).	**Medium-Low**	**Legal**
Transparency regarding personal data collected via log files	There is a lack of clarity on what triggers a loggable event, what details are logged, where log details are kept, and reasonable grounds for reviewing logs. This risk is somewhat offset by the fact that log data is maintained for only 29 days.	**Low**	**Process**
IP addresses may be exposed	The baseline configuration of Private.Storage may expose IP address information. However, users can mitigate this by using an anonymising service such as Tor or I2P to connect.	**Low**	**Technology**
Availability of data cannot be guaranteed	If an attacker or third party initiates a DDoS or otherwise blocks access to the app or servers, a user loses their decryption key, enters an invalid voucher, or forgets to renew a lease or obtain a new voucher, they may lose access to any files stored on the system.	**Low**	**Technology**
The controller may not be able to strictly comply with a deletion / erasure request	Due to the sharded and encrypted nature of the application, it is currently impossible for Private.Storage to ensure that individual files or folders are completely deleted across all shards. While an individual data subject can delete their local copy or delete their private key, which would effectively make the data unrecoverable, there are at least some cases where a deletion request may be impossible.	**Low**	**Technology**

Table 1a: List of Mitigations Taken & Residual Risk

Risk	Mitigation Taken	Residual Risk Rating
No lawful basis for processing is declared in the Privacy Notice	This has been addressed in the Private.Storage Privacy Notice.	**N/A**
No designation of an EU Representative	As suggested, we have designated an EU Representative and this has been added to our Privacy Notice under section I., Data Controller and EU Representative.	**N/A**
Details on where data is stored, and for how long are not well-defined	This has been addressed in the Private.Storage Privacy Notice.	**N/A**
The voucher / capability string details are not exposed to the user	This was a bug that has been addressed by the team.	**N/A**

Risk	Mitigation Taken	Residual Risk Rating
Personal data of customers is shared with Stripe for payment processing	Private.Storage has implemented extensive mitigation efforts and organizational controls for the data received to process payments. This collection of data is an unfortunate part of the current payment processing system, fraud prevention, and 'Know Your Customer' requirements. Until that changes, or we can implement a more privacy-preserving payment system, we will continue to do whatever mitigation and controls we can in this area.	**Low**
Cookie details should be clearer	All Matomo cookies have been disabled and Private.Storage has limited the cookies used by Stripe as much as possible. Section II, 8 of the Privacy Notice has been updated to address the cookies used by Stripe.	**N/A**
Transparency regarding personal data collected via log files	This has been addressed in the Private.Storage Privacy Notice under Section II, 7.	**N/A**
IP addresses may be exposed	Tor integration is on the roadmap. In the meantime, Private.Storage have included suggestions in the documentation urging customers who wish more security to use a VPN.	**Low**
Availability of data cannot be guaranteed	Private.Storage will continue to look at and consider DDoS mitigation efforts and implement them when feasible. However, this is an issue that is possible for any type of service and is not truly solvable.	**Low**
The controller may not be able to strictly comply with a deletion / erasure request	Any data on any system is only deleted through physical destruction. When data is "deleted" from a hard drive or SSD only the name pointing to the data is removed. Thus, this is a risk of every system that relies on SSDs.	**Low**

Risk	Mitigation Taken	Residual Risk Rating
	We believe that we already have a best-case setup in our system: the data is encrypted on the customer's device before being stored on our servers and the customer retains sole control of the Recovery Key to decrypt the data. If the Recovery Key is destroyed by the customer then no one can ever access and decrypt that data. Effectively, this is equivalent to 'deleting' the data (e.g. by physically destroying the hard-drive). Most systems cannot give customers such control as they lack "end to end encryption".	

Conclusion

All systems that in any way touch user data will have risk; that said, the team has and continues to work diligently to mitigate these risks without sacrificing trade-offs to functionality and ease-of-use. The team has agreed to continue to pursue risk-mitigations for the remaining 4 low risks identified by Castlebridge.

Introduction

Why are We Here?

A Data Protection Impact Assessment is mandatory under the General Data Protection Regulation, when processing is "likely to result in a high risk to the rights and freedoms of natural persons."[1] For example, in cases where data processing involves a large number of data subjects, or where the data being processed includes Special Categories at a large scale. However, due to the risk-based approach present in many privacy laws, carrying out a DPIA is not always mandatory for every processing operation. A full list of instances where DPIAs are required can be found in Appendix 1: Criteria for Mandatory Data Protection Impact Assessment Under the GDPR.

The United States also does not currently require DPIAs or privacy impact assessments, however, the laws are ever changing. For example, California's passage of Prop 24 (the California Privacy Rights Act (CPRA), as well as the Colorado Privacy Act (CPA), and Virginia Consumer Data Protection Act (VCDPA), will require more limited 'data protection assessments' in certain situations starting in 2023.[2]

Notwithstanding legal obligation, controllers and processors may also wish to proactively undertake a data privacy or data protection impact assessment for a number of reasons. These include:

- To clearly define the scope of personal data processing and the processing activities being undertaken;
- To assess the potential issues and risks to data protection compliance and to the fundamental rights and freedoms of data subjects;
- To identify risks, and address those risks before a product or service goes live, ensuring that privacy-by-design and default principles are baked in;
- To provide transparency to data subjects, potential clients, partners, and other stakeholders.

This is especially true, when a company is touting their product or service offering as being "privacy-focused" or "privacy-enhancing."

And so is the case here: In a further effort to demonstrate their strong commitments to data protection and user privacy, Private.Storage has agreed to publish this impact assessment on behalf of its partners, customers, and users.

The goal here is to show, not just tell, why privacy matters to the organization.

---

[1] Article 35(1), General Data Protection Regulation, 2016/679, Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01) revised 4 October 2017.

[2] See: Sec. 1798.185(a)(15 - Data Protection Assessments under the CPRA; SecArticle 35, *supra* note 1; . 59.1.576 - Data Protection Assessments under the VCDPA; Sec. 6-1-1309 - Data Protection Assessments under the CPA.

Process & Scope

This Assessment was conducted by Carey Lening, CIPP-E, CIPP-US, CDPP of Castlebridge, and relied on the following information:

- Private.Storage's public-facing documentation
- A security audit of Tahoe LAFS and Gridsync conducted by IncludeSec[3]
- Interviews with members of the development, engineering, security and leadership teams
- The Private.Storage platform & codebase[4]
- Review of the Tahoe-LAFS and Gridsync platforms, codebase and documentation[5]

In Scope

For purposes of this DPIA, Castlebridge focused on the processing activities of the Private.Storage Platform, including the key-value store, file store, and application layers (discussed in more detail in the Review of the Technical Environment section). It also briefly discusses entity-level risks (namely around transparency and accountability) related to certain process-based gaps discovered as part of the assessment.

Out of Scope

This assessment did not generally cover the organization's overall data governance posture, other products developed by Private.Storage (such as the mobile app), or integrations with the Stripe platform, which is currently used for payments processing.

Methodology

The Risk Assessment Methodology Castlebridge applies works for both threshold (or what we refer to as 'triage') DPIAs and full DPIAs. It takes account of the clear requirement under Recital 75 of the GDPR[6], and the implied requirement in Recital 58 of Directive 2016/680/EU for organizations to

---

[3] Security Assessment of Least Authority's Gridsync Application and Tahoe LAFS Android Application (2021) at: 2021 Q2 Least Authority Gridsync Desktop and Tahoe LAFS Android App - Report.pdf.

[4] PrivateStorageIO GitLab repository: PrivateStorageio - GitLab.

[5] Tahoe-LAFS GitHub Repository: Tahoe-LAFS · GitHub; Gridsync Github Repository: gridsync · GitHub.

[6] Recital 75 GDPR:

> The risk to the rights and freedoms of natural persons, of varying likelihood and severity, may result from personal data processing which could lead to physical, material or non-material damage, in particular:
>
> where the processing may give rise to discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy, unauthorised reversal of pseudonymisation, or any other significant economic or social disadvantage;
>
> where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data;
>
> where personal data are processed which reveal racial or ethnic origin, political opinions, religion or philosophical beliefs, trade union membership,

assess risks from the perspective of the impact on the fundamental rights and freedoms of data subjects. The methodology Castlebridge applies also takes into account the competing rights and interests of Data Controllers for the purposes of informing decisions regarding safeguards and other mechanisms envisaged to ensure the protection of personal data and to demonstrate compliance with relevant legislation.

Variable Calculation Formulae[7]

We apply a variant of the Failure Mode Effects Analysis[8] methodology that is commonly used in quality management systems. Within this analysis, we rank the following variables to calculate a risk criticality score, which are set out in the table below. Rankings are based on a 1 to 10 scale.

Variable	Definition
Impact on Individual (**IoI**)	An assessment of the impact on the fundamental rights and freedoms or choice/agency of individuals arising from or as an outcome of the proposed processing activity.
Impact on Organization (**IoO**)	An assessment of the impact on objectives of the organization or on the brand or operations of the organization in the event that this risk materializes.
Likelihood of Detection (**LD**)	An assessment of how likely it is, in the normal course of operations and in light of the identified controls and mitigations that have been or will be implemented, that the occurrence of a risk would be identified in a timely manner sufficient to minimize impact on individuals or organizations.
Probability of Occurrence (**PO**)	An assessment of the probability that a given risk would manifest itself as an actual event impacting individuals or the organization.
Criticality of Risk (**CoR**)	A calculation of the severity of the risk without consideration of ease of detection.
Risk Priority (**RP**)	A calculation of the relative priority of a risk taking into account the likelihood of detection.

and where the processing of genetic data, data concerning health or data concerning sex life or criminal convictions and offenses or related security measures;

where personal aspects are evaluated, in particular analyzing or predicting aspects concerning performance at work, economic situation, health, personal preferences or interests, reliability or behavior, location or movements, in order to create or use personal profiles; where personal data of vulnerable natural persons, in particular of children, are processed; or

where processing involves a large amount of personal data and affects a large number of data subjects.

[7] Some details of Castlebridge's calculation process and methodology have been omitted from the public version of this report.

[8] For details of the history and origins of FMEA analysis see What is FMEA? Failure Mode & Effects Analysis | ASQ.

Defining the Risk

As this methodology is based on Failure Mode and Effects Analysis, it is important to identify the Failure Mode (event/incident/characteristic of processing) that would give rise to an impact and how that impact might arise (Effect).

As such, a single area of potential failure could have multiple effects associated with them which result in different impacts to data subjects or to the organization. The identification of the potential impacting events/outcomes is a key aspect of the Root Cause Analysis phase of the DPIA so that any correlation between potential impacts can be associated with the correct underlying failure mode.

A helpful way of approaching this is to apply a simple "*If X then Y resulting in Z*" logic test to help work from a failure mode (X) to potential outcomes (Z) or from outcomes (Z) to an underlying failure mode (X). In this way, failure modes (risks) that give rise to multiple impacts and effects within the system of processing personal data can be identified and "quick wins" for multiple risk factors can be identified.

Applying Quality Systems Methods

Quality systems methods and approaches such as a Fishbone Diagram and 5 Whys analysis[9] can also be used to define and cluster risks and potential modes of failure.

For the purposes of the Castlebridge DPIA Risk Framework, we cluster identified failure modes by categories which relate to the remedial actions that are most likely to have the most significant effect in addressing the underlying failure mode/root cause.

These are:

- **Governance** (internal data governance or decision-making processes and controls)
- **Process** (the definition of or execution of process)
- **People** (human factors including training, knowledge, awareness, and culture)
- **Technology** (technology features or functionality)
- **Legal** (legislative change or clarification of legislative basis).

[9] For more details on the Fishbone and Five Whys processes, see What is a Fishbone Diagram? Ishikawa Cause & Effect Diagram | ASQ and 5 Whys.

A breakdown of the risks identified by their type can be found in Identification and Assessment of Data Protection Risks. Further details, including a table of ranking scale guidance, can be found in Appendix 2: Risk Assessment Ranking Scale & Guidelines.

Figure 1: Omitted - Castlebridge Proprietary Information

About Private.Storage

The Private.Storage company was originally founded in 2017, as a joint venture product between Least Authority and Private Internet Access (a VPN service), based in Germany and the United States, respectively. Several iterations later, the final product evolved into the current PrivateStorage service. The company is now wholly owned by Least Authority and headquartered in Cranberry Township, Pennsylvania.

The Private.Storage platform consists of an open-source, distributed storage service that allows users to upload folders and files to a secure, end-to-end encrypted cloud-based system, with a user-facing application which hosts files locally on a customer's machine, and communicates with the server.[10]

As noted on their website:[11]

> ***PrivateStorage has been designed with privacy and security features so only you can access your data. No one else - not even us - can see your data when it is stored on PrivateStorage.***

The platform is currently publicly available.

The PrivateStorage application encrypts files locally on the user's machine, and syncs these files to the cloud, using a distributed, redundant sharding system. More details can be found in the Review of the Technical Environment section.

[10] The PrivateStorage Server repo can be found on Whetstone: https://whetstone.private.storage/privatestorage. The repo for the client-facing application can be found here: PrivateStorage.io, LLC · GitHub

[11] PrivateStorage | Private & Secure Cloud Storage

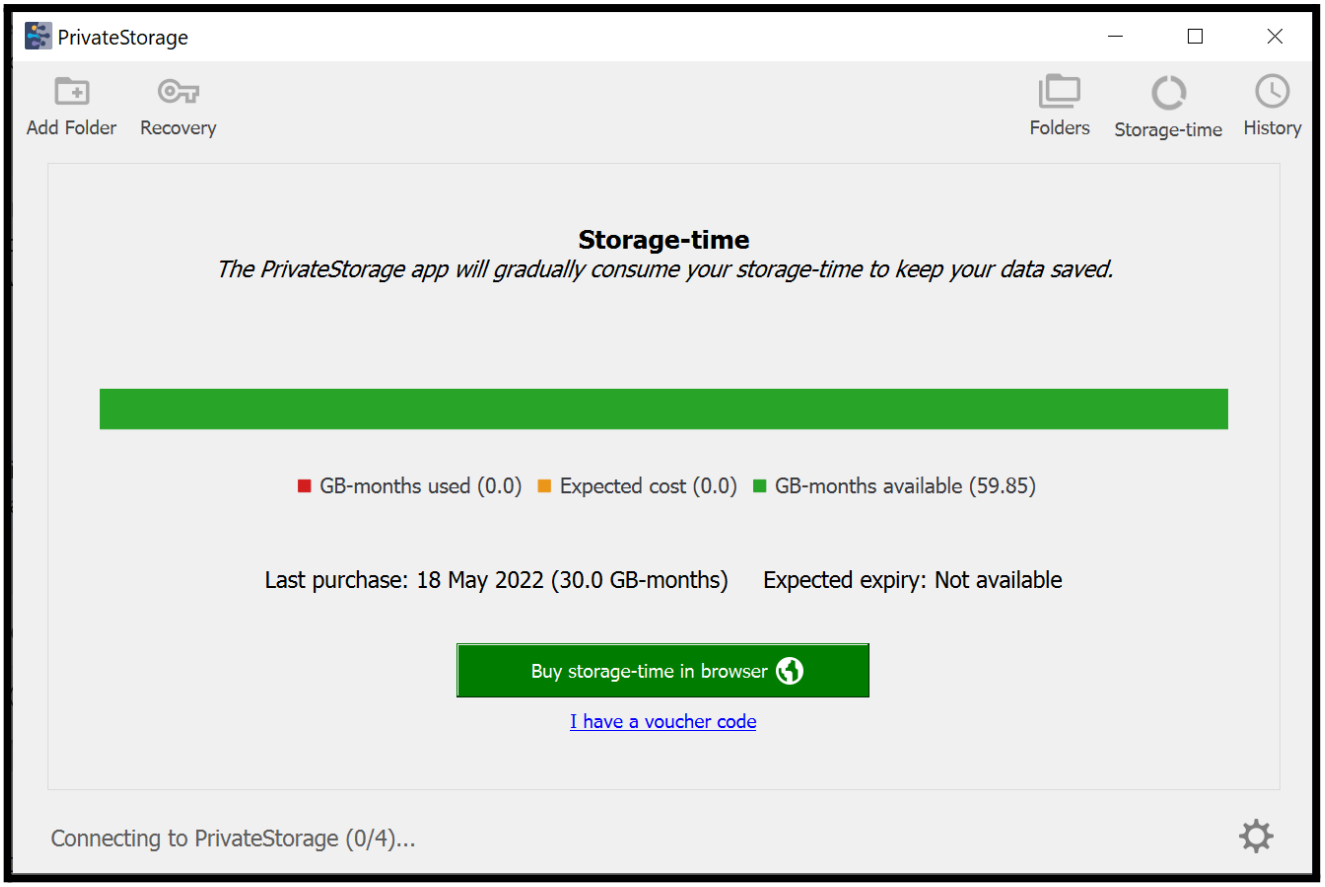


Figure 2: Current Version of the Platform

The system relies on an "accountless authorization" process, which leverages "capabilities" to authenticate a user, rather than a standard userid/password. More details on this process can be found in the Account Creation & Payments section.

A separate payment process, based on Zero-Knowledge Access Passes (ZKAPs), has been integrated into the Private.Storage platform. ZKAPs eliminate the need to expose payment card information to Private.Storage directly. Currently, the platform accepts payments via Stripe using an iframe, but in the future, Private.Storage also plans to offer a crypto-based payment system to fully eliminate the need to share information with third party payment providers (e.g., Stripe). More details on the payment processing can be found under the Zero-Knowledge Access Pass Authorizer (ZKAP) section.

Description of the Business Need and Benefits

Description of Business Need

Private.Storage aims to compete with other proprietary cloud storage services that claim to add additional privacy and security features (Tresorit, SpiderOak One and pCloud) beyond the usual cloud storage options (e.g., DropBox, Google Drive, Box, OneDrive), by providing a solution that offers the same core functionality, without collecting, storing or processing user personal data.

Description of Anticipated/Intended Benefits

The anticipated benefits arising from the proposed processing activities include:

1) Providing users with a secure, cross-platform, end-to-end encrypted data storage solution.
2) Giving users explicit, granular control over their data.
3) Implementing effective accountless authentication that limits what is exposed during the payments and voucher creation processes.
4) Ensuring that data is both available and securely stored.
5) Building on existing security controls already implemented by two open-source solutions, Tahoe-LAFS and Gridsync.

Review of Personal Data & Subjects In Scope

Categories of Data Subjects

The Private.Storage application and website capture data primarily about three different categories of individuals whose personal data may be processed:

Category of Data Subject	Description of Category
Customers	Users of the Private.Storage platform, including individuals who buy access/pay for storage, and those who receive shareable links to the service from paying users.
Website Visitors	Individuals who visit the Private.Storage website, but do not create an account with the company.
Third Parties	Individuals whose data may be stored by users of Private.Storage.
Employees/Contractors	Private.Storage employees & contractors.

Categories of Personal Data Processed

The proposed processing activity will involve the processing of the following categories of personal data.

Category of Personal Data	Description of Category	Source	Data Subject(s)
IP Address	Stored in log files and collected by Stripe. The Private.Storage website only stores the first octet of an IP address: e.g., 91.xxx.xxx.xxx.	Website & platform & when processing payments	Customers, Website Visitors
Email (if provided)	If provided by a data subject to be notified of the GA release (**prior to launch in 2023**), and when communicating with technical support through CDRLink. Email of employees will also be collected in some cases.	Website & direct communications from users	Customers, Website Visitors, Employees

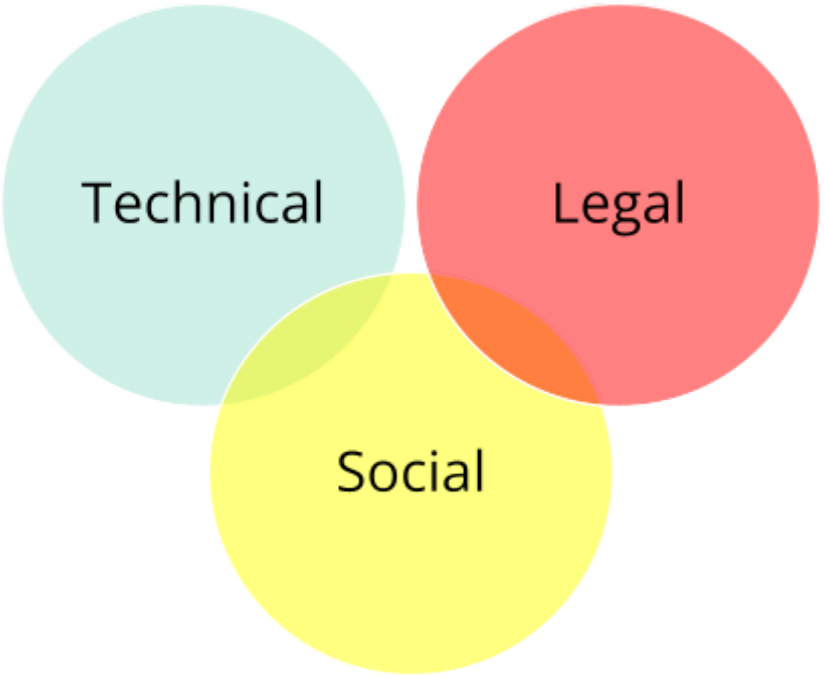
Category of Personal Data	Description of Category	Source	Data Subject(s)
Name	Collected as part of the 'Buy Storage-time' process workflow and stored by Stripe for payment purposes. Private.Storage does not store this information. Private.Storage employee names may be processed in correspondence with Customers for technical support purposes.	When processing payments to buy storage-time	Customers, Employees
Address	Collected as part of the 'Buy Storage-time' process workflow and stored by Stripe for payment purposes. Private.Storage does not store this information.	When processing payments to buy storage-time	Customers
Bank Details / Cardholder data	Collected as part of the 'Buy Storage-time' process workflow and stored by Stripe for payment purposes. Private.Storage does not store this information.	When processing payments to buy storage-time	Customers
Stripe Payment Token[12]	Collected as part of the 'Buy Storage-time' process workflow and maintained by Stripe for payment purposes. Private.Storage does not store this information, as the process/verification step is done instantaneously. Instead, Private.Storage receives tokenized data and converts this into a ZKAP which cannot identify a Customer.	When processing payments to buy storage-time	Customers

[12] Technically, a limited set of employees at Private.Storage and Least Authority will have access to this information by way of Stripe, which is why it has been included in this chart.

Category of Personal Data	Description of Category	Source	Data Subject(s)
Other data	Clients using the service may store other forms of personal data. However, these details are unavailable to Private.Storage or third party processors due to the technical implementation of the system.	Website, Platform & when processing payments, communications with company by customers	Customers, Third Parties

Review of the Environment

Below, we analyze the Private.Storage platform and related data processing from the technical, legal, and social perspectives.



Castlebridge undertook the technical assessment by reviewing all public-facing developer and customer documentation, as well as the source code directly. The author also installed and used the desktop application on a Windows 10 device, and reviewed proxy data using MitmProxy version 8.0.0 and database logs using DB Browser ver. 3.1.2.2. The website (including cookies and JavaScript) was assessed using a variety of scanning tools, including CookieBot. The author has not tested the mobile application or other software versions.

For the legal assessment, Castlebridge reviewed privacy and policy documents, the state of relevant legislation, case law and regulatory guidance. This analysis considers both alignment with national and sectoral privacy considerations, the nature and effectiveness of existing controls, and the potential impacts on fundamental rights and freedoms of individuals (data subjects and consumers under US law).

For the social assessment, we put ourselves in the shoes of a typical consumer, hoping to identify common challenges a user might experience when installing and using the application and data protection questions and concerns a user might have.

Review of the Technical Environment[13]

In this section we summarize the technical considerations for the implementation of this proposed processing. We also identify any relevant technical controls that are in place that ensure that privacy, security, and data protection principles are being met in practice.

The underlying technology for the Private.Storage application (Castlebridge reviewed v. 22.2.1) is a rebranded, modified version of the Gridsync GUI (Castlebridge reviewed v. 0.5.0), which sits on top of Tahoe-LAFS, a free, open-source secure, distributed and decentralized file system developed by Brian Warner and Zooko Wilcox.

Installation consists of a series of scripts that build Private.Storage assets and other files into the Gridsync source, add branding, set a total number of required shares (discussed below), and identify the servers to use.[14]

Private.Storage is built on three layers. The first two layers consist of the key-value store and file store, and are based on the Tahoe-LAFS system. The final application layer is based on Gridsync.

[13] This is a limited, high-level analysis of the underlying structure of Tahoe-LAFS and Gridsync. A more detailed architecture discussion can be found here: Tahoe-LAFS Architecture — Tahoe-LAFS 1.x documentation and on GitHub - gridsync/gridsync: Synchronize local directories with Tahoe-LAFS storage grids

[14] See: Private.Storage Desktop Readme: PrivateStorageDesktop/README.md

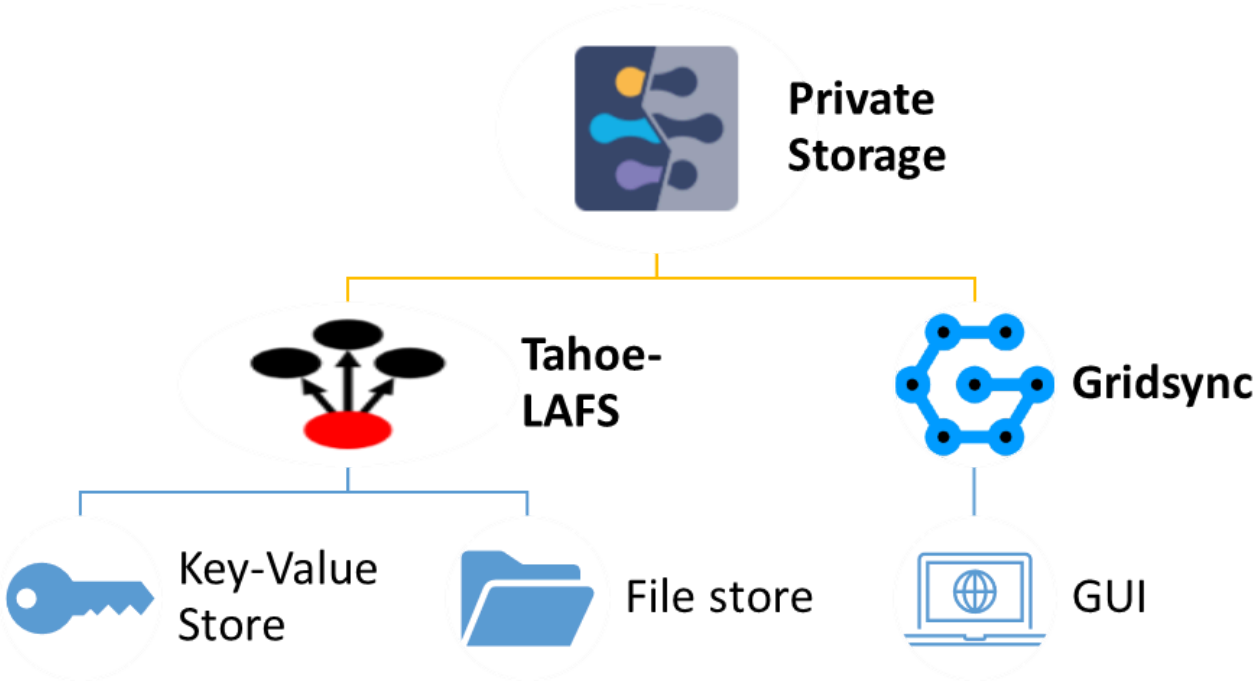


Figure 3: Private.Storage Components

Key-Value Store

The key-value store sits at the lowest layer of the platform. The key-value store is implemented by a grid of Tahoe-LAFS storage servers. Communication occurs over TCP. The storage servers (referred to as "nodes" or "storage nodes") store "shares" – small components parts of encoded files.

Shares are ordinarily broken up and distributed across multiple different nodes. A user can define the total number of shares they wish to use (e.g., 10) and the number of distinct storage nodes (e.g., 5). Tahoe-LAFS defaults to using 10 shares across 7 nodes, while the Private.Storage implementation defaults to 4 shares across 5 nodes.[15] While nodes can hold multiple shares, customarily, distribution is one share per server.[16]

Clients (discussed in the Application Layer section) are provided with a static list of nodes to connect to – currently five servers hosted by M247 in the United States.

Another component of the key-value store is the key. Files are encrypted client-side, broken into segments, which are erasure-coded and further segmented into blocks. More detail on this process is discussed in the Storage & Files section.

A hash of the encryption key is also used to form a "storage index" – which is used for server selection and as an index to shares within the selected nodes. The Private.Storage application

[15] Based on the period of assessment in May - July 2022.
[16] Further details on how servers are allocated can be found here: Tahoe-LAFS Architecture — Tahoe-LAFS 1.x documentation

computes secure hashes of the encrypted files and their shares. The hashes themselves are stored in a small data structure known as a "Capability Extension Block" which is stored on each storage server. Capability Extension Blocks are themselves stored in larger data structures known as "Capabilities" or "Capability Strings" which include the hash of the Capability Extension Block, permissions (read, write, verify), and any encoding parameters necessary to perform decoding. Importantly, no single share provides discernible information to a third party. Permissioning is derived exclusively from the Capability String. As the team has described, "The Capability String is like a physical key to a physical lock. If you have it, you can unlock the lock. If you don't, you can't present a warrant to the lock and have it capitulate to your authority and unlock itself."

File Store

The middle layer consists of the decentralized file store. This file store layer is responsible for mapping human-meaningful pathnames (directories and filenames) to pieces of data. The actual bytes inside these files are referenced by capability strings, but the file store layer is where the directory names, file names, and metadata are kept. Files have different permission-type capabilities – read-write and read-only. Directories also have a third permission type, 'verify-only', which is necessary for the integrity-checking step, although this is not exposed at the user-level (See: Integrity & Resiliency).

Storage & Files

Files stored can be either mutable or immutable. Immutable files, once uploaded to the storage nodes cannot be modified, whereas mutable files can be modified by someone who has read-write access. Users with read-write access to a file or directory can give other users similar (or lesser) access.

All files are encrypted prior to leaving the local device, ensuring that confidentiality and integrity are preserved. Currently, Private.Storage encrypts files based on the file's mutability status – immutable file content is encrypted using an AES128-CTR block cipher[17] and mutable files use 2048-bit RSA-PSS-SHA256[18] to asymmetrically encrypt a symmetric encryption key and apply AES128-CTR encryption to the file's contents. The encrypted files are broken into chunks or segments. This has the benefit of decreasing the lag between initiating a download and recovering the file and building in resiliency to the product. It follows a similar model to how BitTorrent services work.

The segment blocks themselves are erasure-coded and broken down into blocks, of which only a subset is required to reconstitute a segment. One block from each segment is sent to a given server/node. The set of blocks on a specific server constitutes a "share" and only a subset of these shares are required to reconstruct the file. Files are transmitted via an encrypted TLS session.

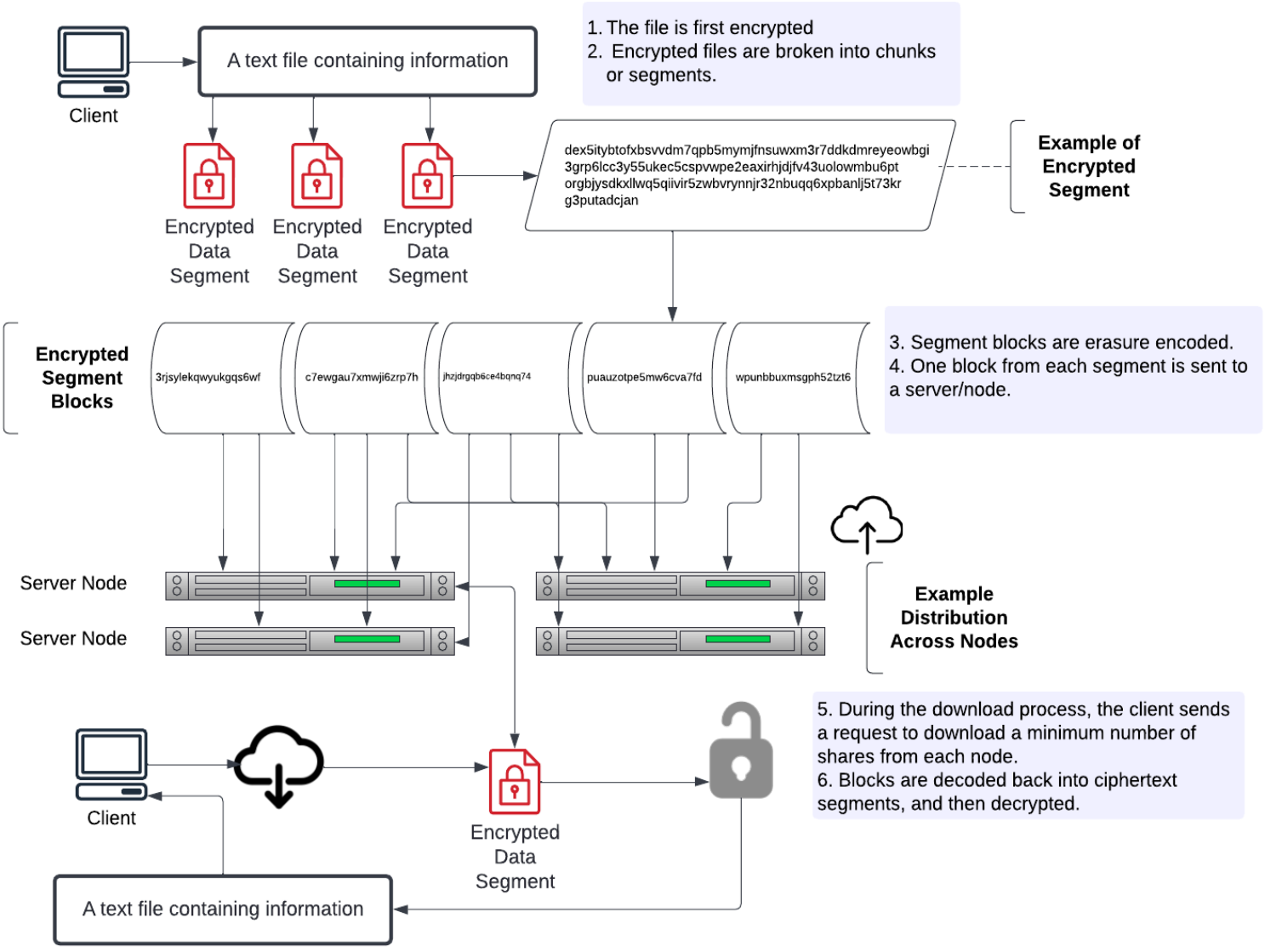
Extra shares are also created and saved on multiple storage nodes for redundancy and to ensure that data can still be accessed in case any particular server becomes unavailable. In the event of data loss or corruption, these extra shares can be used to reconstruct the data.

Due to the nature of the encrypted shards, Private.Storage (and the storage nodes hosted on M247) have no way to read or modify data stored on the system. This eliminates a core challenge that

[17] Block cipher mode of operation - Wikipedia
[18] Probabilistic signature scheme - Wikipedia.

confronts storage of data in the cloud – assurance of confidentiality and integrity does not rely on a third party. Tahoe-LAFS documentation refers to this as "provider-independent security."[19]

However, users must still rely on storage nodes to guarantee the *availability* of data, which is still subject to compromise, for example, if an attacker conducts a distributed denial of service (DDoS) event across M247 servers. A larger discussion on the technical and organizational measures in place to protect data can be found in the section Appropriateness of security measures over the processing of personal data.



[19] See: Welcome to Tahoe-LAFS!

Figure 4: Abbreviated View Of The Encryption, Storage & Decryption Process

Health Checks & Servers of Happiness

Before a file upload is considered successful, it has to pass an upload health check. For immutable files, this includes checking whether the 'servers-of-happiness' condition is met -- For mutable files and directories, a check is run to ensure that all encoded shares generated during the upload process are successfully placed on the grid.

The purpose of the servers-of-happiness check is to ensure that file availability will not be affected if a few nodes later fail. The current defaults (as of the date of assessment) require a total of 5 storage servers/nodes, where an upload is considered successful if it is shared with a minimum of 4 nodes (servers-of-happiness check), and a minimum recovery of 3 shares per file. The Tahoe LAFS system has the potential to support other algorithmic approaches to determine specific numbers of target servers for each share, but those are not currently implemented as part of the Private.Storage release.

To decode a file, the client downloads the required number of shares (e.g., 3) from each storage server, decodes the shares back into segments of ciphertext, and then uses the decryption key to convert the ciphertext into a plaintext file.

Logging

As stated in Section 2 of their privacy notice, Private.Storage does not log any details about normal use of their service. However, in some cases, such as if issues around data integrity are detected (e.g., if health checks repeatedly fail), an "incident log" file is created locally and may be created and stored on the storage servers where the incident occurred. The logs are also ingested into a logging system hosted on AWS located in the eu-central-1 region.

Details contained in the incident include file size, IP address information, storage index information, source code details, and information on the operations being performed when the incident occurred. Logs are stored unencrypted for 29 days and then deleted in accordance with the company's retention policy.

A client-side debug log (.dmp file on Windows) can also be created and sent by the user as part of a support ticket if requested by personnel for diagnostic analysis. This file contains machine information such as OS, system and process uptime, release information, stack trace information, and other diagnostic data including a FAILURE_ID_HASH. A review by Castlebridge using the WinDbg tool did not identify any instances where identifiable information was recorded.

Integrity & Resiliency

The Tahoe LAFS system (and by incorporation, Private.Storage) have built in some resiliency features, including a folder synchronization feature known as Magic Folders, and a file-repair function. Since shares may disappear if nodes/servers suffer a failure (temporary or permanent), the platform has built-in file checking and synchronization functions as well. Checking occurs at a poll_interval of 60 nanoseconds.

Folder synchronization occurs by use of the Magic Folder system, which was created by Least Authority. Magic Folders are created for each directory stored, and a process detects local changes between files and uploads those changes to the grid. It also detects remote changes made on the grid and downloads those changes to the local filesystem.[20] This is achieved by using a long-lived (running) subprocess (PrivateStorage-magic folder).

The team is working on additional features that would perform a more extensive check-and-repair process against files to ensure that the file integrity is still maintained, and regenerate and re-upload missing shares to new live servers.

Garbage Collection

Garbage collection is not enabled server-side by default: thus storage servers will not delete shares without being explicitly configured to do so. However, the client software does engage in a "lease maintenance process."

Under the lease process, a file is allocated or "leased" to a specific client/user for a fixed period. If the lease is not renewed, it can be marked for deletion. The Private.Storage application engages this lease maintenance process as a periodic check once a month if the client is running. If the software is not running during the scheduled time, then the check will occur at some point after the software is launched.[21] Data stored on the grid will be periodically inspected and incur "storage-time," which will use available tokens (described in the Zero-Knowledge Access Pass Authorizer (ZKAP) section) to renew any leases that are on the verge of expiration when data is still maintained in synced folders locally.

Customers are responsible for renewing their lease (i.e., purchasing more storage time) on a periodic basis at least frequently enough to prevent the lease from expiring before the next lease maintenance process occurs. However, Private.Storage currently does not have the technical capability in place to verify that a given lease share is fully erased server-side.

Sharing

Currently, the platform does not allow for sharing of files between users, though the team has prioritized this for future development.

Application Layer

The final layer consists of a desktop application itself. The Desktop application of Private.Storage is a branded package based on the open source Gridsync application, and incorporates a forked version of Tahoe-LAFS, Magic-Folder, and the ZKAP Authorizer. At the time this assessment was

---

[20] Magic Folder Github Repository: GitHub - LeastAuthority/magic-folder: Tahoe-LAFS-based file synchronization. The tableID current_snapshots includes the Tahoe-LAFS URI representing the most recent remote snapshot.

[21] The exact schedule is determined by the production-grid.json file. e.g., PrivateStorageDesktop/credentials/production-grid.json.

As an additional security feature, "Storage Time" checks are pseudo-randomly chosen from a uniform range. The developers have explained that the pseudo-randomization is "intended to take what might be a clear signal about client identity (eg, 'Client X checks its leases on the 3rd of each month, at 7:14am')" and add noise to these values. This makes it harder for an adversary to correlate user activity.

performed, releases for GNU/Linux, MacOS and Windows 10 were available.[22] The applications are primarily built in Python. The goal of the application is designed to make usability easier.

Three tools are installed as part of the desktop application:[23]

- **Tahoe-LAFS:** This includes installing all dependencies, including Python 3.
- **Magic Folder:** Provides bi-directional file synchronization – the application monitors local and remote directories, storing and retrieving new versions as they appear.
- **Recovery Keys:** Allows for connections and folders to be easily restored from a single file.
- **ZKAPAuthorizer:** A Tahoe-LAFS storage-system plugin which authorizes storage operations based on privacy-respecting passes.

Desktop Application

The high-level technical environment for this proposed processing is illustrated below.

[22] More details can be found at the Private.Storage Get Started page.
[23] A more detailed explainer on all features included as part of Private.Storage/Gridsync can be found here: GitHub - gridsync/gridsync: Synchronize local directories with Tahoe-LAFS storage grids.

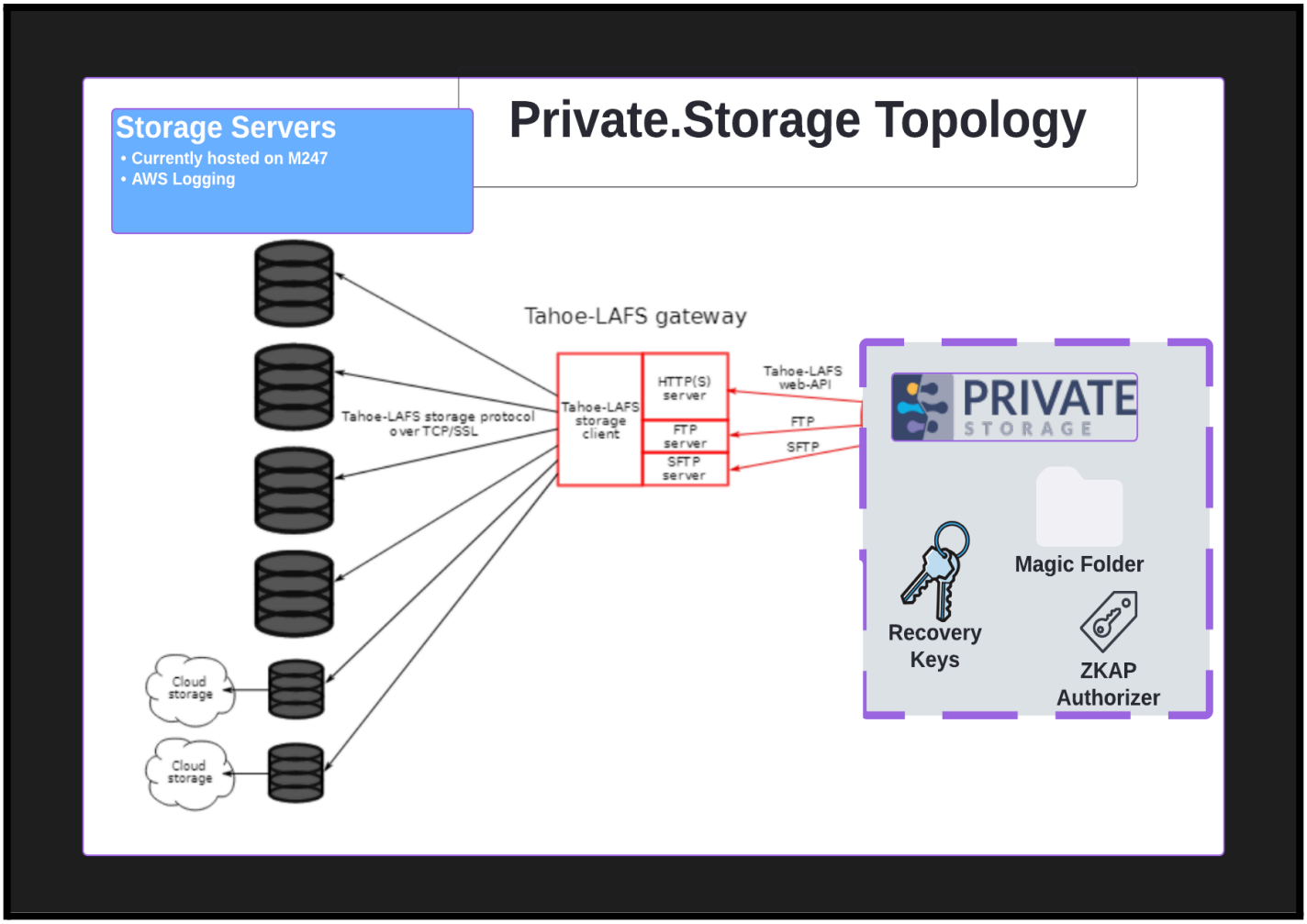


Figure 5 - Private.Storage Network Topology

Account Creation & Payments

Unlike other cloud-based file systems, Private.Storage does not rely on account credentials (e.g., email address/password). Instead, Private.Storage uses the concepts of "vouchers" and "storage-time". Vouchers link a purchase transaction (made via the website and Stripe payment gateway), with a uniquely-generated token. These tokens are designed to be cryptographically hard to link back to personal data gathered while processing the payment and creating the voucher. However, in some cases, the vouchers can be used to retrieve payment-related personal data. This is necessary for activities like processing refunds or confirming payment status.

However, there is no way to link a given file or folder uploaded onto Private.Storage servers with a voucher, payment token or user.

Private.Storage currently has two payment processes in development:

- a conventional payment process relying on Stripe as the payment processor, which is fully operational.

- a novel approach based on the use of Zero-Knowledge Proofs.

Stripe Payment Gateway

When a user first starts up the Private.Storage application, they are prompted to “Buy Storage-Time”. Clicking on the icon launches a new browser window to the https://private.storage/payment page where a user can purchase storage time in increments of 30 GB a month. Clicking on the link and accepting the terms of service then directs the user to a Stripe payment page. The URL is dynamic and includes includes a unique voucher string and checksum value: For example:

> https://buy.stripe.com/14keWS1wQaMS3285kk?client_reference_id=qaBce6FzZZKAf7HXlwebV9E8SAx8Dzg9IdZ7m2jjXzV8

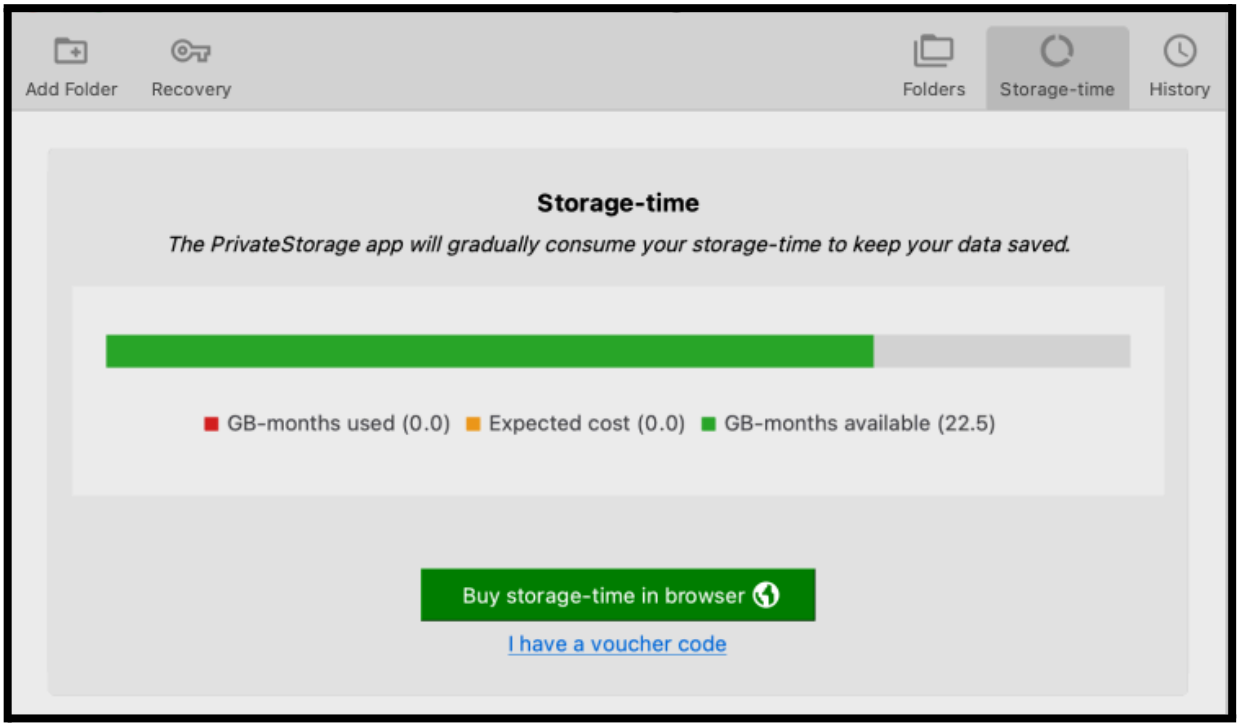


Figure 6: The Storage-Time UI

The Stripe Payments page collects customer details including name, country or region, credit card / debit number, CVV and expiration date, as well as an optional field for the cardholder’s mobile number. Once a payment token is received from Stripe, payment is considered complete, and a voucher is created and stored locally on the user’s machine. Personal data about the purchaser is never collected or sent to Private.Storage. Data on the purchase however, is exchanged with the Private.Storage servers, where the local voucher is exchanged for “Storage-Time” which cannot be linked back to an individual or their purchase.

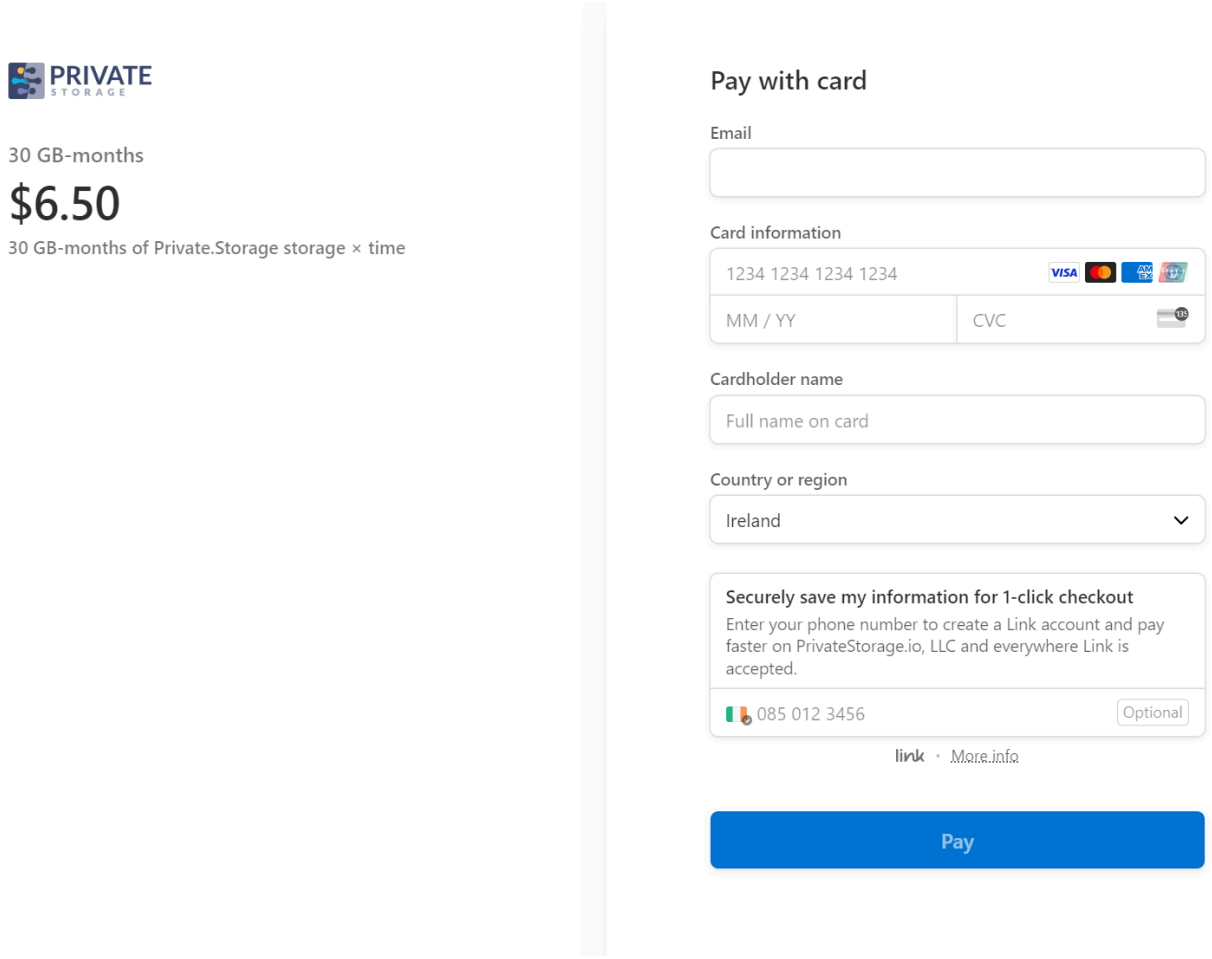


Figure 7: View of the Stripe Payments Page

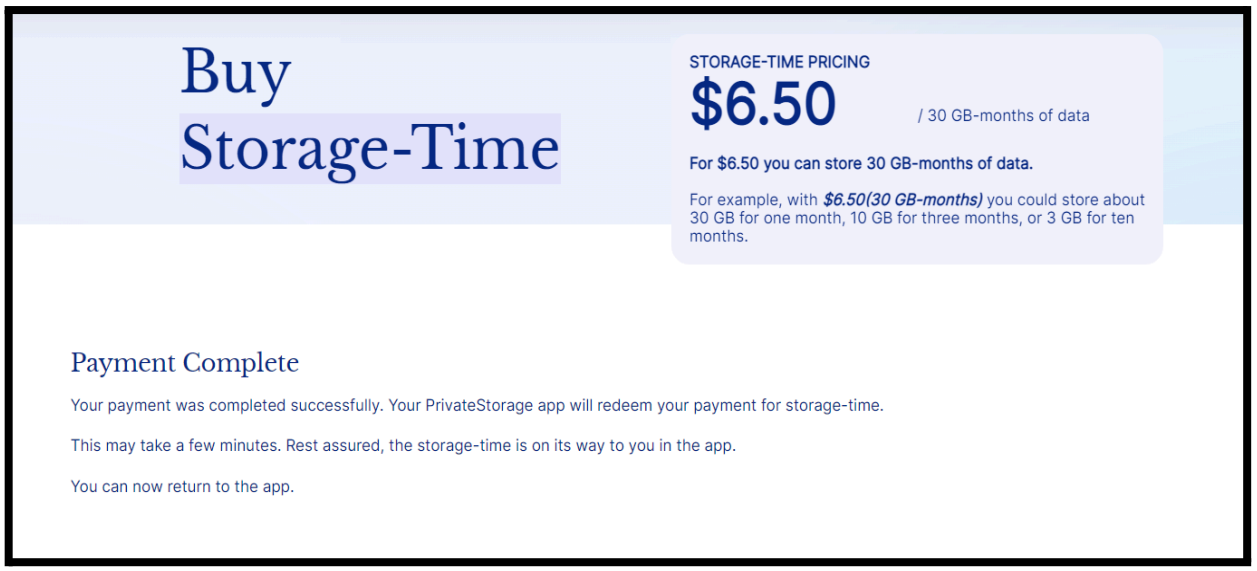


Figure 8: Conclusion of Payments Process via Stripe

Zero-Knowledge Access Pass Authorizer (ZKAP)

The Zero-Knowledge Access Pass (ZKAP) Authorizer was developed by the Least Authority and Private.Storage teams to provide a secure, privacy-respecting mechanism for receiving payments.[24] The use of ZKAPs helps to facilitate an online exchange of value, disconnected from payment and account or service data that is gathered about users. As noted by Least Authority, "[t]his is very helpful in use cases where mixing these data points is not in the best interest of the company offering the service," particularly those seeking to maximize user privacy to the fullest extent possible.

The ZKAP model is based on a variation of Privacy Pass—a zero knowledge cryptographic protocol for establishing trust. Under the existing Privacy Pass framework, 'proof-of-humanness' is checked and verified by use of CAPTCHAs. This allows individuals to provide proof (that they're not a bot, for example) without revealing information on where and when that trust was provided.[25]

Private.Storage ZKAPs rely on a modified version of the Privacy Pass protocol to verify that payment has been received via the Private.Storage payment server. Thus, it relies on 'proof-of-payment' rather than 'proof-of-humanness' to establish trust. Like the Stripe gateway process described above, this interaction is largely invisible to the user.

[24] More details on the process and history can be found here: ZKAPs - Least Authority. A more in-depth analysis of the underlying encryption framework used can be found here: The Ristretto Group. Currently, CloudFlare, Brave and Least Authority make use of the protocol.
[25] See for example CloudFlare's implementation of the Privacy Pass protocol: Supporting the latest version of the Privacy Pass Protocol. More details on the Privacy Pass framework can be found here: FAQ | Privacy Pass.

Importantly, there is no login or other form of user authentication involved. This privacy-preserving infrastructure maintains separation and simplicity.

Below is a diagram representing existing legacy payment systems (e.g., Stripe) versus the ZKAP Authorization framework.

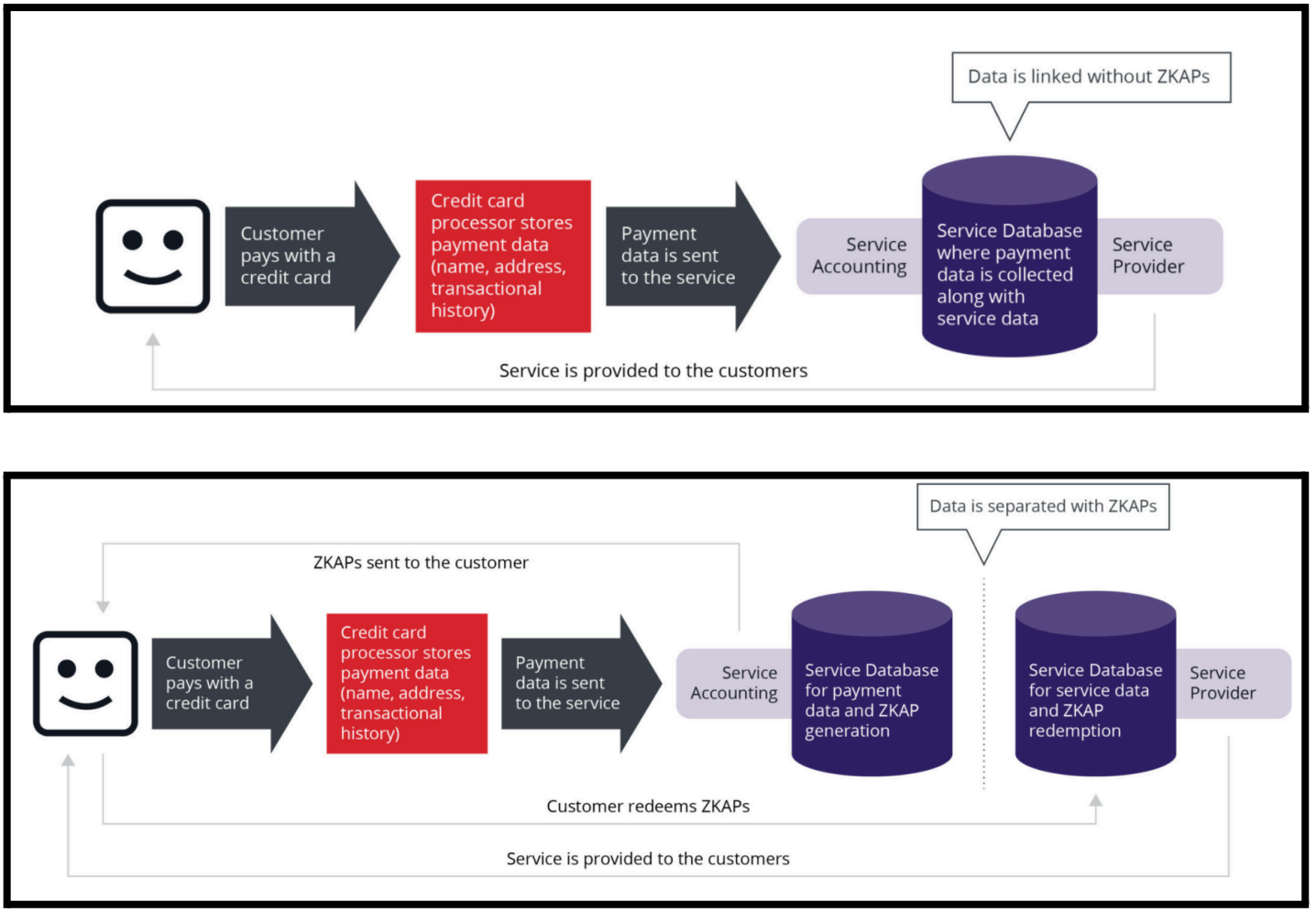


Figure 9: Comparison between traditional v. Zero Knowledge Access Pass Process - ZKAPs - Least Authority

Helpdesk Support (CDRLink)

Private.Storage uses CDRLink, a privacy- and security-focused helpdesk client, to address customer support queries.[26] CDRLink is an open-source application, built on the Zammad ticketing platform, and hosted on Greenhost in the Netherlands. The team opted for CDRLink in part due to strong privacy controls that are built-in to the application, including secure methods of ticket creation, integration with Signal, WhatsApp and other secure messaging platforms, strong permissioning, ticket deletion, and other features.

[26] CDRLink GitLab repository: Link · GitLab.

When a customer contacts support, through the website, Private.Storage has committed to delete the email address and message(s) after seven days, once an issue is resolved.[27]

Website

Private.Storage disabled all cookies (bar a strictly necessary cookie denoting a user's logged-in status) as of June 2022.

When users initiate a payment (by triggering the 'Buy Storage-Time' call in the application, Stripe adds the following third party cookies on the payments page for purposes of fraud prevention and detection. These are considered strictly necessary, as payments cannot occur without the cookies being in place:

Cookie ID	First Party / Third Party	Expiration	Category	Purpose
__stripe_sid	Third Party	Session	Strictly Necessary	Fraud Prevention / Detection
__stripe_mid	Third Party	1 year	Strictly Necessary	Fraud Prevention / Detection
m	Third Party	2 years	Strictly Necessary	Fraud Prevention / Detection

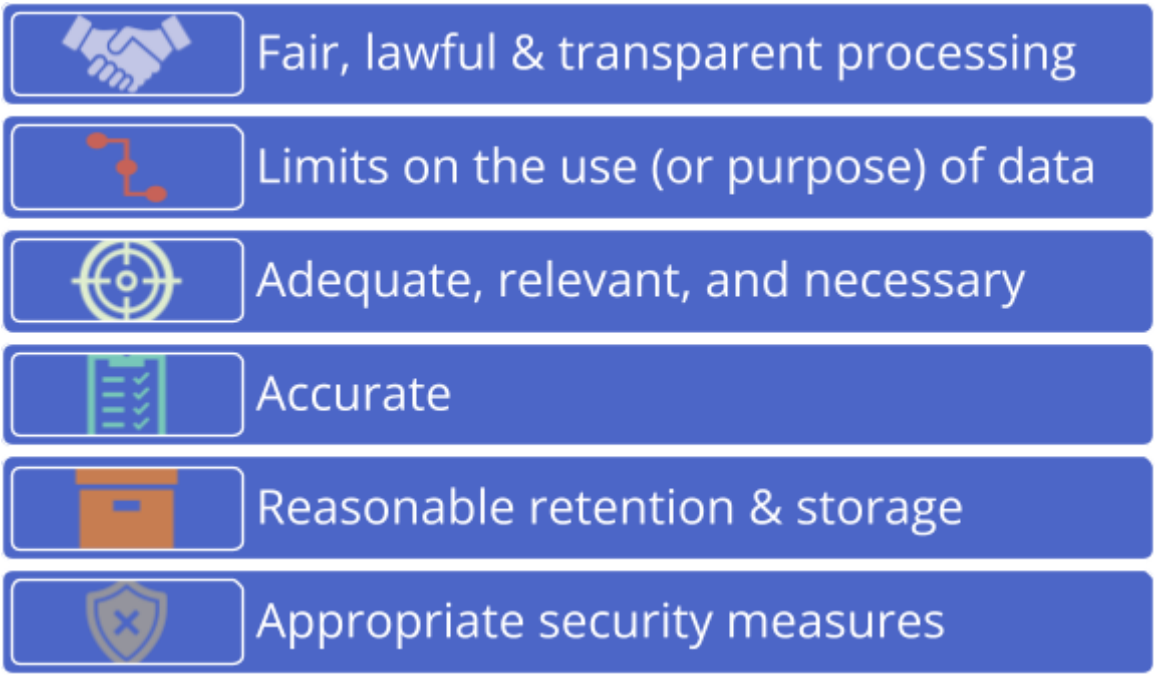
[27] See: Private.Storage Privacy Notice.

Review of the Legal Environment

This section of the DPIA examines the legal environment and related considerations associated with this processing activity.

Compliance with Data Protection Principles

Regardless of the legal framework considered, most laws governing data protection (or privacy) are grounded on a number of core principles. These include:



1. Fair, Lawful, and Transparent Processing / Notice

Fairness & Transparency

The principles of fairness and transparency require that any processing of data is obtained and used fairly, that the reasoning for processing that data is transparent, and that this information is clearly presented to the data subject. "Processing" covers a wide range of operations performed on personal data, including collection, storage, retrieval, use, transmission, disclosure, alteration, erasure and deletion. It encompasses activities by the controller as well as downstream processors and other third parties (such as Private.Storage).

Transparency also requires that controllers and processors not only provide information on what personal data is being processed, but how the data is processed, why it's processed, how long data is retained, where it's being stored, how it's being secured, who else has access to the data, as well as information about the controller or processor, and how a data subject can assert their rights.

The Private.Storage Privacy Notice and documentation reasonably defines what data is collected, processed, and stored, how it's collected and used, how long it's retained, and data subject rights. The policy is clear, detailed and easy to understand. It covers both processing that occurs on the website and via the application, as well as processing that is undertaken by Private.Storage personnel, or disclosed to third parties who will have access or use of that personal data (i.e., support & payment processing).

A chart has been provided below breaking down data subject rights that could be easily incorporated in the Privacy Notice. [Table 3: Breakdown of Subject Access Rights Across Select Legal Frameworks]

Lawful Basis

To be permissible under the GDPR, processing of personal data must be based on a legal ground identified under Article 6 (1) GDPR, or what lawyers refer to as having a 'lawful basis' for processing.

Lawful bases include consent, contract, compliance with a legal obligation (e.g., Know-your-Customer), or the controller's legitimate interests amongst others.

Private.Storage relies on the following lawful bases for processing personal data:

- **Consent** -- Customers of the platform consent to providing their email for notification purposes or to communicate with support personnel to address technical problems. Consent is explicit and informed, in that a user must voluntarily provide this information in order to communicate with Private.Storage employees, and this collection is only for the respective, stated purposes as described in the Privacy Notice. This information is not conditional on use of the product - it's entirely optional.

  For support tickets, email and information provided by data subjects are kept for seven (7) days after the closure of the issue. After 7 days, the closed issue and the corresponding email address are erased.

- **Contract** – Private.Storage is storing encrypted files on M247. To the extent that this non-personal data requires a lawful basis, the clearest basis would be that of contract with the data subject. Similarly, Private.Storage stores incident logs on M247. This could be covered by contract, or alternatively, legitimate interests as a legal basis.

  Finally, Private.Storage currently relies on Stripe to process payments information in order to obtain payment for the contracted-for service. Currently, this requires the collection of personal data, which is necessarily stored on Stripe, and is accessible by a limited number of employees at Private.Storage. Some of the data collected by Stripe is to comply with US and other national "Know Your Customer" laws. Therefore, in addition to contract, to the extent that data collected is for this purpose, compliance with legal obligations would be an appropriate lawful basis.

- **Overriding Legitimate Interests** – This covers review of data transmitted in logs (namely, IP addresses) and customer support investigations of the same. Legitimate interests may also be applied in the context of employee information that may be collected and stored by AWS and GitHub, which are used to store source code. Private.Storage works with the least amount of personal data needed – to fulfill this process, and takes measures to avoid including personal data that is not relevant (e.g., user details, passwords) in their source code.

  In an effort to balance the fundamental rights and freedoms of data subjects, Private.Storage must justify how their legitimate interests override those rights. This is referred to as the 'necessity' test.[28]

  Here, Private.Storage endeavors to collect the least amount of personal information necessary – primarily IP addresses and email addresses (when users contact customer support) and login details and pseudonyms of employees who have access to GitHub and AWS.

  Other elements (e.g., voucher IDs, browser identification and errors) do not directly identify a data subject. Private.Storage's interest is in maintaining a functioning, secure, and available system, not processing data about their customers or employees. Therefore, data is only collected in cases where an incident occurs (and a log of the incident is created), where a customer reports an incident directly to the team, or when employees create an account to perform development tasks.

  On their Privacy Notice, Private.Storage details the exact situations where personal data (such as IP address) is logged:

  > *In addition, in normal use of PrivateStorage, **we do not log anything** about your use of the service.*
  >
  > *In exceptional circumstances, such as when a potential problem with data integrity is detected, our server may automatically create an "incident log", which can contain information like file size, and your IP address. We use this information to understand and address possible errors in our service. We keep these logs for 30 days and then delete them.*[29]

There is a clear link between the collection and processing of this data, and most would reasonably understand why such data was collected. Moreover, the data is not particularly sensitive, and the risk to the data subject of a breach is minimal. Reasonable security controls (relative to the risk) are employed – namely limited retention (of email and IP addresses) and the use of sub-processors governed by strong contractual obligations to protect personal data.

---

[28] See: Article 6(1)(f) and Recital 47 & 49 GDPR.
[29] Privacy Notice.

2. Purpose Limitation

Many data protection legal frameworks also recognize that data controllers must have a valid purpose for processing personal data, and be limited to that purpose or purposes. Processors are further limited to the purposes imposed on them by the controller. Under the GDPR this is represented in Article 5(1)(b) which states that personal data must be

> *collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.*[30]

In general, the Private.Storage Privacy Notice clearly explains how personal data is used and the limited purposes for that use. The company processes data for four specific uses:

- **Using the service:** If a problem occurs when using the product, an Incident log may be created and may contain information such as a user IP address.
- **Visiting the website:** Users who sign up to receive information about the product release may provide their email addresses for that purpose. Email addresses will only be stored until the notification email is sent.[31]
- **Support requests/contacting the company:** Users who reach out to support, information provided (e.g., email address and text that may contain personal data provided by the user), will be processed to answer the user's query or troubleshoot an issue.
- **Processing payments:** When users purchase Private.Storage credits, payment information is sent to Stripe.

This is done at the time of collection, and is strictly limited to those purposes by various technical and organizational controls, including limits on data collection, short retention periods, transmission of information using TLS 1.3, and the ongoing development of privacy-enhancing tools such as Zero-Knowledge Access Passes (ZKAPs). To the extent that Private.Storage is acting as a processor (on behalf of their customers and Users), they are largely processing encrypted ciphertext – not personal data.

3. Adequacy, Relevance, and Necessity

The principles of data minimisation and privacy by design (which are articulated in most legal frameworks) require that the processing of personal data be limited to what is 'necessary'. Data processing is necessary when it is "adequate, relevant and limited" to the described purposes of processing.[32]

In order to function as a file storage and sharing system, Private.Storage must collect some personal information about the machine(s) connecting. Due to the nature of how the internet works, this may at times include IP address information, for example, when logging an incident. This

[30] Article 5(1)(b) GDPR.
[31] This was the case at the time of the assessment in 2022; now that Private.Storage is out of beta, this collection is no longer in place.
[32] Article 5(1)(c) GDPR.

information may also be necessary to ensure that Private.Storage networks remain resilient, available, and secure, and to defend against denial of service attacks.

Other information collected, including email addresses of interested customers, personal data provided by users directly to Private.Storage via the support process, and payments information, are adequate, relevant, and necessary for the company to meet its contractual obligations to customers and legitimate business interests – namely, addressing customer inquiries, , troubleshooting incidents, and processing payments information.

The team from the start has adopted 'privacy-by-design and default' as its baseline. This is reflected throughout the code, decisions made by the company, and technical controls that are in place.

The team continues to actively work towards collecting the least amount of information necessary to fulfill these purposes. For example, one area where this is actively being worked on is in regard to payments. Currently, Stripe potentially collects more information than is necessary and the team are working on ways to limit this, namely through the use of the ZKAP framework. Additionally, the application itself has implemented a process where accounts are not necessary, eliminating the need for users to create accounts (and store more information).

In May 2023, the team also disabled analytics tracking and cookies on its website by disabling Matomo tracking and anonymizing IP address information.

4. Accuracy

Due to the limited nature of what information is being kept or maintained about a data subject (namely, IP addresses of customers who have an incident, email addresses and information communicated to the support team, and information provided to Stripe), concerns around data accuracy are limited. Private.Storage limits the amount of data being held on a given user in regard to support tickets, and generally enforces short retention periods for any data it collects. Only two individuals have access to Stripe data.

5. Retention and Storage Limitation

As noted in this section, most data processed by Private.Storage, including data stored with third party processors, is stored in line with a reasonable retention schedule, and these details are provided to data subjects on the company's Privacy Notice. Namely:

- Log data for tracking and resolving incidents is maintained for 30 days and then deleted.
- Support details (including email) are kept for 7 days after closure of an issue. After 7 days the closed issue and corresponding email address is erased.

Voucher data and other information such as the Stripe token, are maintained indefinitely. However, this information cannot be linked back to individuals using their Private.Storage credits.

With regard to the Private.Storage platform, a fixed retention policy cannot strictly be complied with in regard to file storage on the platform itself. This is a known problem due to constraints on how the system enforces Garbage Collection. Currently, the only mechanism in place is that if a user fails to renew a lease, data may eventually be overwritten.

The team is actively working on addressing this risk. However, this risk is substantially mitigated by the fact that data is not stored in an identifiable way – in effect, it is no longer personal data as defined by data protection laws [See: Personal Data or Personal Information]. Thus, the main consideration from a retention period is compliance with a data subject access or deletion request.

6. Appropriateness of security measures over the processing of personal data

Many legal frameworks, including the GDPR (Art. 32), California Consumer Privacy Act (CCPA), Virginia Consumer Data Protection Act (VCDPA), and Colorado Privacy Act require controllers or businesses to ensure that adequate security controls (often referred to as 'technical and organizational measures', or TOMs) are in place. This section will analyze the platform and underlying applications Private.Storage is built on to assess those controls.

Currently, Private.Storage implements a number of technical and organizational measures to limit and mitigate risks from a security and data protection perspective. Notably:

- All data is end-to-end encrypted in storage.
- Users are given explicit, granular control over their data.
- "Accountless authentication" is in place which eliminates the need to create userids/passwords (eliminating the need for Private.Storage to collect or process that data).
- Sharded, redundant distribution across multiple systems ensures that availability of data is preserved.
- Files are broken into encrypted, segmented blocks. This makes it nearly impossible for an attacker (using current methods) to meaningfully identify, much less reconstitute, a file without access to the client's decryption keys and/or physical access to their system.
- The platform is built on an open protocol, using free/open source software libraries and code. This means the code is reviewable and auditable by third parties, and that vulnerabilities and exploits can be identified and addressed in a timely manner. Private.Storage also offers a whitehat disclosure (bug bounty) program.[33]
- Private.Storage and Least Authority engaged IncludeSecurity to conduct an external security audit and penetration test of the Tahoe-LAFS and Gridsync applications in May 2021.[34] That report found only minimal security risks related to the applications, some of which have been highlighted in the Risk Matrix (see: Table 5: Initial Risk Matrix).[35]
- The creators of Private.Storage have signed a commitment to not build in backdoor processes for governments or hostile regimes to access client data.[36]
- Log data is stored for the minimum amount of time necessary, in line with the company's retention policies.

---

[33] Private.Storage Whitehat Program.

[34] The assessment team focused on identifying security and privacy concerns in various areas including, but not limited to, cryptographic issues, data leakage, insecure file permissions, and denial of service issues. More details can be found here: 2021 Q2 Least Authority Gridsync Desktop and Tahoe LAFS Android App - Report.pdf.

[35] Most risks identified related to the Tahoe-LAFS Android application, which is not compatible with the Private.Storage application.

[36] Statement on Backdoors: tahoe-lafs/backdoors.rst at master.

- Support queries can be communicated and handled via Signal, an end-to-end encrypted messaging platform. Users do not strictly need to submit an email address.

Compliance with Data Subject Rights

Most privacy or data protection laws also provide certain rights and remedies to data subjects. These are broadly defined as data subject rights and include:

- information
- access
- rectification (correction of inaccurate data)
- objection (or opting-out of processing)
- deletion/erasure (the 'Right to Be Forgotten')
- data portability
- prohibition on automated decision-making or profiling
- accountability / redress, including a private right of action or a right to regulatory review / enforcement.

A breakdown of the relevant jurisdictional posture in these areas is included below:

Table 3: Breakdown of Subject Access Rights Across Select Legal Frameworks

Right	GDPR	CCPA/CPRA	CPA	VCDPA	UCPA
Information	Yes	Yes	Yes	Yes	Yes
Access	Yes	Yes	Yes	Yes	Yes
Rectification/Correction	Yes	Yes	Yes	Yes	No
Objection/Opt-out	Yes	Yes	Yes, limited to certain processing activities	Yes, limited to certain processing activities	Yes, limited to certain processing activities
Use Limitation	Yes	Sensitive information only (CPRA)	Yes	Yes, with exceptions	Yes
Deletion/Erasure	Yes	Yes	Yes	Yes	Yes
Data Portability	Yes	Yes	Yes	Yes	Yes
Automated Decision-Making	Yes	No	Limited (profiling opt-out)	Yes	No

Right	GDPR	CCPA/CPRA	CPA	VCDPA	UCPA
Private Cause of Action / Enforcement by Regulator	Yes	Limited to data breaches; otherwise only via Attorney General	No private right of action; limited via Attorney General	No private right of action; limited via Attorney General	No private right of action; limited via Attorney General

Due to the limited nature of personal data processed by Private.Storage, compliance with data subject rights is reasonably achievable. The one primary area where this deviates is with regard to deletion/erasure of data stored on servers.

Currently, due to the inability to identify or associate data stored on the servers with a specific customer, it would be impossible for a data subject to request or otherwise execute immediate and complete deletion of customer information stored across the platform. Users should be advised that a 'deletion' request can only necessarily include data that is identifiable by Private.Storage, and would therefore be constrained to details such as their email address, IP information, technical support details, and payment details processed by Stripe.

Although users can delete data client-side, this action may not propagate out to all servers or locations where data segment blocks are stored – at least not immediately. Eventually, storage blocks will be overwritten (once storage time credits are expired), but if an individual loses access to their encryption key or otherwise is incapable of logging in and redeeming those credits, it is not possible for Private.Storage to effectuate a deletion request on the individual's behalf.

Private.Storage cannot identify which storage location is associated with any given user, which is by design. Similarly, if, and when sharing functionality is developed, a user with access to a shared folder on the platform, could not compel Private.Storage to delete this data, or have them notify the controller (the customer) and request that they do so.

However, the risk here is more a theoretical & economic one for Private.Storage rather than a legal one. Since the data stored and distributed across its servers is not identifiable (at least given current limitations on encryption), it is no longer personal data, and is not covered by the scope of current data protection laws. However, the inability to delete the data does impose a cost on Private.Storage – as the data still exists or lives so long as an account maintains storage credits.

Contractual Relationships Between Private.Storage and Subprocessors

In most cases, Private.Storage is acting as a processor of data on behalf of its customers. However, in some cases, Private.Storage is acting as a controller in their own right. This includes:

- Collecting email addresses to contact customers who make a request;;
- Personal data that is collected for addressing technical support;
- Personal data that may be processed (e.g., Ids) as part of the software development process.

At this time, the highest-risk processing activity is via Stripe, which is used to facilitate payments. PrivateStorage is currently working with Stripe to restrict the amount of data collected and retention period of data stored.

Table 4: Subprocessors

Processor	Purpose of Processing	Details on Data Processed	Location of Processing	Legal Basis for Processing	Agreement in Place?
CDRlink	Handling customer support requests (Hosted)	Email address, phone number, other details provided by customer.	United States	Consent / Contract	Yes
Greenhost	Cloud storage/ hosting	CDRlink data	The Netherlands	Contract / Legitimate Interests	Yes
AWS	Internal Cloud storage/ hosting	Whetstone	Europe (Frankfurt)	Contract / Legitimate Interests	Yes
Stripe	Payment processing	Payment details – name, address, bank/credit card information, Stripe Token	United States	Legitimate interests	Yes
Whetstone	Gitlab	Code repository - team facing only.[37]	AWS (Frankfurt)	Legitimate Interests	N/A
M247	Hosting sharded data	End-to-end encrypted data storage; no identifiable personal data is stored on M247.	United States (New York, Miami, LA)	Contract	Yes

[37] Customer data is not stored in the code repository, but it is possible that identifying information (e.g., ids of engineers involved in the product development) may inadvertently be collected during the development process.

Identification and Assessment of Data Protection Risks

Below, we provide a full breakdown of risks identified during this assessment. For the purposes of the Castlebridge DPIA Risk Framework, we cluster identified risks by categories which relate to the remedial actions that are most likely to have the most significant effect in addressing the underlying risk and its root cause.

These are:

- **Governance** (internal data governance or decision-making processes and controls)
- **Process** (the definition of or execution of process)
- **People** (human factors including training, knowledge, awareness, and culture)
- **Technology** (technology features or functionality)
- **Legal** (legislative change or clarification of legislative basis).

Risk Category	
Governance (internal data governance or decision-making processes and controls)	
Process (the definition of or execution of process)	
People (human factors including training, knowledge, awareness, and culture)	
Technology (technology features or functionality)	
Legal (legislative change or clarification of legislative basis).	

Table 5: Initial Risk Matrix

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
No lawful basis for processing is declared in the Privacy Notice	Under the GDPR, controllers must provide a legal reason for why they process data. There are six legal bases for processing data: a) consent, b) performance of a contract, c) compliance with a legal obligation, d) legitimate interests of the controller, e) vital interests and f) public interest.	4	7	10	10	280	2800	**Medium**	Since the product will be marketed to individuals around the world, including Europe, lawful bases for processing should be specified in the privacy notice. In most cases, this will likely be grounded under consent, contract and legitimate interests. Contract or legitimate interests seems the most appropriate basis for use of the product, and the provisioning of hosting servers, the customer support portal and payment processing via Stripe.
No designation of an EU Representative	Private.Storage has not formally designated a representative in the EU, as is required by Article 27 GDPR. Without an appointed representative / establishment in the EU, this leaves Private.Storage open for regulation from all EU Data Protection Authorities.	3	8	10	10	240	2400	**Medium**	Private.Storage should nominate a representative in the EU.

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
Details on where data is stored, and for how long are not well-defined	Under the GDPR, controllers must provide details on where data is stored physically, any subprocessors/third parties used, This information is currently not well-documented in the privacy notice.	5	6	10	7	210	2100	**Medium**	Provide better transparency to users by detailing (ideally in an easy-to-read chart) the nature of processing activities, and the geographic locations where data is stored (e.g., the United States), particularly where that data is stored in cleartext, or shared with providers who will process personal data (CDRLink, Stripe).
The voucher / capability string details are not exposed to the user.	Vouchers are how the Private.Storage system verifies that a user has paid for service. At the payment screen, no details about the voucher code are provided to the user. If a user loses access to their local machine, they cannot gain access to their data, even if they save the recovery key, without paying for additional storage. This creates a loss of availability of data.[38]	5	7	5	10	350	1750	**Medium-Low**	Castlebridge recommends exposing the voucher / capability string to the customer and informing customers that this must be preserved in order to have access to the data. For example, a popup window exposing the voucher/ZKAP details could be provided, and customers encouraged to record this information in case the data is lost, they use another machine, or they wish to share access. The team noted that this is an ongoing, known challenge, and something actively being investigated. A technical plan has been developed to mitigate this risk. In the

[38] This risk was discovered by Castlebridge personnel when testing the platform. While a user can save their recovery key (and thus authenticate), there is currently no way to physically access data stored within the Private.Storage ecosystem without storing the voucher or paying for access again. Private.Storage has addressed this bug and this risk is no longer present (as of 2023).

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
	The risk here is less about data protection, and more about usability, particularly if multiple steps must be performed and the transaction process is not seamless, or errors not obvious to the user.[39]								interim, customer support can assist customers with processing refunds and issuing credits for those who encounter problems.

[39] When Castlebridge attempted to set up an account, for example, it was not clear at first that the transaction completed successfully. While a positive confirmation occurred on the payments page, no signal seemed to be sent to the application itself. It still displayed a 'no storage time' message. The author had to restart the application before it registered that a token had been received.

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
Personal data of customers is shared with Stripe for payment processing.	Currently, payments are processed using Stripe as a subprocessor. Stripe stores numerous pieces of personal information, including name, address, cardholder data, etc. This data is not maintained on Private.Storage servers, but it is accessible to a limited subset of Private.Storage employees (currently < 3), which means that it is currently possible to match the redemption of a voucher back to a customer who paid using Stripe..	3	5	8	10	150	1200	**Medium-Low**	Private.Storage has engaged in extensive mitigation efforts and organizational controls to limit the amount of data it receives from Stripe. Payment details are collected and processed by Stripe over a secure channel (https) and not stored by Private.Storage. Instead, Private.Storage works with the unique Stripe token generated for each transaction, which it immediately exchanges for a non-identifiable token id. This risk of employee access and re-identification is largely mitigated by internal organizational measures, and the use of technical controls (namely, least privilege access to a limited number of employees). Additionally, while it may be possible to verify that a voucher was purchased and potentially associated with a capability, it would be impossible for a Private.Storage employee to match that to specific use of the storage system. The capability string cannot be matched to access, or uploads/downloads on the system. The company is working on steps to move

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
									away from Stripe entirely and to instead allow payments through cryptocurrencies and other mechanisms. This is on the roadmap, but will not be available before the MVP goes live.
Cookie details should be clearer	Currently, persistent Stripe fraud prevention cookies are installed on a user's device. The Privacy Notice does not explicitly detail these cookies (though it does generally discuss the processing and purpose of Stripe cookies in the abstract).	3	4	9	10	120	1080	**Medium-Low**	The Privacy Notice should be updated to include more details (e.g., a chart) of specific details on the cookies stored on the site. An example has been provided as part of this DPIA in the Website section.
Transparency regarding personal data collected via log files.	There is a lack of clarity on what triggers a loggable event, what details are logged, where log details are kept, and reasonable grounds for reviewing logs. This risk is somewhat offset by the fact that log data is maintained for only 29 days.	3	4	8	8	96	768	**Low**	Update documentation and the privacy notice to be more explicit around what constitutes a loggable event, what details are being kept in logs, who has access to those logs, if logs are secured/encrypted and where logs are stored.

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
IP addresses may be exposed	The baseline configuration of Private.Storage may expose IP address information. However, users can mitigate this by using an anonymising service such as Tor or I2P to connect.	2	3	8	3	18	144	**Low**	The team has integration with Tor/i2P on their roadmap, but this is not yet implemented.[40] Users with an elevated threat model should be encouraged to access the platform via a VPN, Tor or i2P node.
Availability of data cannot be guaranteed.	If an attacker or third party initiates a DDoS or otherwise blocks access to the app or servers, a user loses their decryption key, enters an invalid voucher, or forgets to renew a lease or obtain a new voucher, they may lose access to any files stored on the system.	4	3	4	2	24	96	**Low**	Users should be encouraged to adopt sound security practices, including backing up private keys and ensuring that they timely renew leases. Private.Storage should consider DDoS mitigation efforts and to increase the total number of 'servers-of-happiness' to reduce the risk of unavailability. In response to a security audit conducted in May 2021, the maintainers of the Gridsync application added better error checking to the application to make customers aware of risks identified in the audit.

[40] This has been proposed as part of the Tahoe-LAFS process. See: Tahoe-LAFS Github repo: tahoe-lafs/anonymity-configuration.rst at master regarding certain tradeoffs exist by using Tor/I2P. Notably, traffic sent via Tor/I2P increases latency, and may be more prone to server loss.

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
The controller may not be able to strictly comply with a deletion / erasure request.	Due to the sharded and encrypted nature of the application, it is currently impossible for Private.Storage to ensure that individual files or folders are completely deleted across all shards. While an individual data subject can delete their local copy or delete their private key, which would effectively make the data unrecoverable, there are at least some cases where a deletion request may be impossible.	3	3	4	2	18	72	**Low**	Due to the nature of how data is stored within the system, and Private.Storage's 'accountless authorization' process, Private.Storage employees cannot identify what data is stored on its systems. customers should be encouraged to delete any information they no longer wish to be stored on the system locally, and if they wish to delete their account, to destroy or revoke the private key created during the installation process. At that point, the data is no longer 'personal data' under most legal standards as the information no longer can be tied to an identifiable person. For example, Article 11 GDPR explicitly excuses data controllers from the obligation to comply with erasure obligations under Article 17 GDPR, where, as here, "the controller is able to demonstrate that it is not in a position to

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
									identify the data subject."[41] There remains an extremely small risk that a third party discovers that personal data about them is stored by a customer on Private.Storage. As a processor, they would be required to forward information on to the controller/customer in this case, but that is effectively impossible.

[41] Article 11 GDPR:

1. If the purposes for which a controller processes personal data do not or do no longer require the identification of a data subject by the controller, the controller shall not be obliged to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with this Regulation.

2. Where, in cases referred to in paragraph 1 of this Article, the controller is able to demonstrate that it is not in a position to identify the data subject, the controller shall inform the data subject accordingly, if possible. In such cases, Articles 15 to 20 shall not apply except where the data subject, for the purpose of exercising his or her rights under those articles, provides additional information enabling his or her identification.

Article 17 GDPR:

1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:

(a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;
(b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing;
(c) the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2);
(d) the personal data have been unlawfully processed;
(e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;
(f) the personal data have been collected in relation to the offer of information society services referred to in Article 8(1).

Risk	Description	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence	Criticality of Risk	Risk Priority	Risk Rating	Suggested Mitigation(s)
									This fact should be emphasized in the privacy notice, along with an explanation as to why such requests cannot be fulfilled.

**

***

Management Response

PrivateStorage has a goal of building privacy into our product wherever possible, therefore addressing the risks noted in this report is very important to us. Below are our responses to the risks, and where necessary what we have done to address or mitigate those issues.

Risk 1 - No lawful basis for processing is declared in the Privacy Notice

Response: As suggested, this has been added to our Privacy Notice.

Risk 2 - No designation of an EU Representative

Response: As suggested, we have designated an EU Representative and this has been added to our Privacy Notice under section I. Data Controller and EU Representative.

Risk 3 - Details on where data is stored, and for how long are not well-defined.

Response: As suggested, this has been added to our Privacy Notice.

Risk 4 - The voucher / capability string details are not exposed to the user.

Response: The issue noted in the report is actually a bug in the system rather than an issue that needs a potential workaround as is suggested. Additionally, revealing the voucher to the customer will not actually solve this issue because the voucher has already been redeemed in the case mentioned. Therefore, instead of implementing the suggestion, we have fixed the bug so that neither vouchers nor ZKAPs are lost in the event the user machine is lost. This should no longer be a risk.

Risk 5 - Personal data of customers is shared with Stripe for payment processing.

Response: As stated, we have implemented extensive mitigation efforts and organizational controls for the data received to process payment. This collection of data is an unfortunate part of the current payment processing system, fraud prevention, and 'Know Your Customer' requirements. Until that changes or we can implement a more privacy preserving payment system, we will continue to do whatever mitigation and controls we can in this area.

Risk 6 - Cookie details should be clearer.

Response: All Matomo cookies have been disabled and we have limited the cookies used by Stripe as much as possible. Section II, 8 of the Privacy Notice has been updated to address the cookies used by Stripe.

Risk 7 - Transparency regarding personal data collected via log files.

Response: As suggested, the Privacy Notice has been updated to include this information in Section II, 7.

Risk 8 - IP addresses may be exposed.

Response: As stated, Tor integration is on our roadmap. In addition, we have included in the Privacy Notice the users' IP address may be exposed.

Risk 9 - Availability of data cannot be guaranteed.

Response: We will continue to look at and consider DDoS mitigation efforts and implement them when feasible. However, this is an issue that is possible for any type of service and is not truly solvable.

Risk 10 - The controller may not be able to strictly comply with a deletion / erasure request.

Response: Any data on any system is only deleted through physical destruction of the hardware on which it is stored. When data is "deleted" from a hard drive or SSD only the name pointing to the data is removed. Thus, this is a risk of every system and not unique to PrivateStorage. We believe that we already have a best-case setup in our system: the data is encrypted on the customer's device before being stored as shards on our servers and the customer retains sole control of the Capabilities (e.g. as included in the Recovery Key) to reassemble and decrypt the data. If all copies of the Capabilities are destroyed by the customer then no one can ever access the data since it is required to reassemble the shards and decrypt that data. The result is equivalent to 'deleting' the data (e.g. by physically destroying the hard-drive). Most systems cannot give customers such control as they lack "end to end encryption".

Appendix 1: Criteria for Mandatory Data Protection Impact Assessment Under the GDPR

✅ Use of personal data on a large-scale for a purpose(s) other than that for which it was initially collected pursuant to GDPR Article 6(4).
✅ Profiling vulnerable persons including children to target marketing or online services at such persons
✅ Use of profiling or algorithmic means of special category data as an element to determine access to services or that results in legal or similarly significant effects.
✅ Systematically monitoring, tracking or observing individuals' location or behavior.
✅ Profiling individuals on a large-scale.
✅ Processing biometric data to uniquely identify an individual or individuals or enable or allow the identification or authentication of an individual or individuals in combination
✅ Processing genetic data in combination with any of the other criteria set out in guidance on DPIAs issued by the EDPB
✅ Indirectly sourcing personal data where GDPR transparency requirements are not being met, including when relying on exemptions based on impossibility or disproportionate effort.
✅ Combining, linking or cross-referencing separate datasets where such linking significantly contributes to or is used for profiling or behavioral analysis of **individuals**, particularly where the data sets are combined from different sources where processing was/is carried out for difference purposes or by different controllers.
✅ Large scale processing of personal data where the Data Protection Act 2018 requires "suitable and specific measures" to be taken to safeguard the fundamental rights and freedoms of individuals.

Appendix 2: Risk Assessment Ranking Scale & Guidelines

Below we set out the guidelines for establishing the objective relative weighting of the variables which are included in the calculation framework outlined above.

Table 6: Risk Assessment Ranking Scales

Scale	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence
10	Critical impact on privacy or other rights and freedoms; May result in discrimination or similar impacts	Critical impact on organizational goals. A key business strategic business objective will not be met; Significant reputation damage occurring at national and international level/medium term financial losses likely	Immediately detectable while happening without special controls (Preventative/Detective controls)	Is happening/Has already Occurred
9	Significant impact on individual's rights and freedoms; Could result in discrimination or similar impacts	Significant impact on organization; A key business objective highly unlikely to be met or significantly impacted; Substantial impact on reputation likely at national & international level / Medium term financial losses possible	Immediately detectable while happening with some basic controls (Preventative/Detective controls)	Certainty of Occurrence in the short term unless preventative action taken immediately
8	High Impact on individual's privacy or other rights; Could result in discrimination or similar impacts	High Impact on organization; Key Business objectives are significantly impacted but tactical solutions may be possible; reputation damage likely at national level / Potential for financial loss	Immediately detectable with basic controls (Reactive/Detective Controls)	High Probability of Occurrence in immediate term (within 30 days of go-live)
7	High Impact on individual's privacy or other rights; May cause individual high degree of distress, upset, or inconvenience	High Impact on organization; Business objectives are impaired but can be resolved; Some reputation damage likely / moderate financial impact possible	Immediately detectable with advanced controls (Reactive/Detective controls)	High Probability of Occurrence in medium term (within 3 months of go live)
6	Moderate impact on privacy or other	Moderate impact on organization; May require process(s) to be temporarily	Detectable retrospectively using basic	Moderate Probability of Occurrence in

Scale	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence
	fundamental rights/freedoms. May cause individual high degree of distress or upset or inconvenience	suspended to address an incident; May lead to reputation impacts at national level; Potential financial loss arising from incident.	controls (Reactive controls)	immediate term (within 30 days of go live)
5	Moderate impact on privacy or other fundamental rights/freedoms. May cause individual distress or upset or inconvenience	Moderate impact on organization; May require process(s) to be temporarily amended to address an incident; May lead to minor reputation impacts at national level; Potential financial loss arising from incident.	Detectable retrospectively using advanced controls (Reactive Controls)	Moderate Probability of Occurrence in medium term (within 3 months of go live)
4	Minor impact on privacy or other fundamental rights; May cause some distress, upset, or inconvenience	Minor impact on organization; May require process(s) to be temporarily amended to address an incident; May lead to minor reputation impacts at national level; Low/No risk of financial loss	Potentially detectable using desk audit/review process	Low Probability of Occurrence in immediate term (within 30 days of go-live)
3	Very Minor impact on privacy or other fundamental rights; May cause some distress, upset, or inconvenience	Very minor impact on organization; No change required to process; Very limited risk of minor reputation impacts at national level; Affects only a very small number of individuals; Low/No risk of financial loss	Potentially detectable using detailed audit process /process review	Low Probability of Occurrence in medium term (within 3 months of go-live)
2	Very Minor impact on privacy or other fundamental rights; May cause some minor distress, upset, or inconvenience	Very minor impact on organization; No change required to process; Low/No risk of reputation impacts; Impact only affecting very small number of individuals; No risk of financial loss	Detectable only if notified by an affected person	Unlikely to occur over a 12 month period post go-live
1	No impact on privacy or other	No impact on business operations; No changes	Impossible to detect - reactive	Negligible probability of

Scale	Impact on Individual	Impact on Organization	Likelihood of Detection	Probability of Occurrence
	rights; No tangible or material effect on the individual	required to processes; Impact only to one individual; No risk of financial loss	controls only possible	occurrence in any timeframe

Note that we apply a 1 to 10 scale to all variables to allow for the reflection of nuance in risks and to minimize the risk of "in-between" risk factors leading to an over or under estimation of the risk.

Impacts arising from legislative basis, potential for sanctions if processing not undertaken, or potential breaches of legislation will be assessed as impacts on the organization. Impacts arising from the outcomes that may manifest as a *result* of processing or as a result of a risk manifesting itself will be assessed through the impact on the Individual.

Risks are assessed in the context of the controls and safeguards identified through the assessment of the Information Environment.

A rescoring of all risks is undertaken as part of the assessment of recommended remedial actions to determine, once remediations have been applied, what the residual risk remaining is in respect of the proposed processing activities.

Appendix 3: Definitions Under Select Legal Frameworks

When conducting any kind of privacy or data impact assessment, it's important to understand the language of what is in scope. Namely:

- what is (and is not) considered personal data (and special categories of personal data);
- what constitutes processing of personal data;
- what defines a data subject;
- what defines a controller or processor of data;
- distinctions between anonymous, pseudonymous, de-identified and aggregate data.

Since legal frameworks around the world differ, we need a clear understanding of these terms across these frameworks.

Personal Data or Personal Information

Law / Regulation	Definition
Article 4(1), General Data Protection Regulation, 2016/679/EU (GDPR)	***Personal data*** is any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
Sec. 1798.140, California Consumer Privacy Act (CCPA) / California Privacy Rights Act (CPRA)	***Personal information*** is any " information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."[42]
Sec. 6-1-1303, Colorado Data Protection Act (CPA)	***Personal data*** is **"**any information that is linked or reasonably linkable to an identified or identifiable individual, but excludes de-identified data or publicly available information."
Sec. 59:1-571, Virginia Consumer Data Protection Act (VCDPA)	***Personal data*** is any information that is linked or reasonably linked to an identified or identifiable natural person. Excludes de-identified data or publicly available information (a separately defined term).
Sec. 13-61-101, Utah Consumer Privacy Act (UCPA)	***Personal data*** is any "Information that is linked or reasonably linkable to an identified individual or an identifiable individual. … Personal data does not include deidentified data, aggregated data, or publicly available information."

Special Categories or Sensitive Personal Information / Data

Law / Regulation	Definition
Article 9, GDPR	***Special categories*** data refers to any information relating to an identified or identifiable natural person identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical,

[42] There are broad exceptions to what personal information does not include, but they are not included here.

Law / Regulation	Definition
	physiological, genetic, mental, economic, cultural or social identity of that natural person.
Sec. 1798.140, CPRA	***Sensitive personal information*** includes information that reveals a consumer's social security, driver's license, state Identification card, or passport number; account log-In, financial account, debit card, or credit card number in combination with any required security or access code, password, or credentials allowing access to an account; precise geolocation data; racial or ethnic origin, religious or philosophical beliefs, or union membership; the contents of a consumer's email, and text messages; unless the business is the intended recipient of the communication; genetic data; biometric information for the purpose of uniquely identifying a consumer; personal information collected and analyzed concerning a consumer's health; personal information collected and analyzed concerning a consumer's sex life or sexual orientation.
Sec. 6-1-1303, CPA	***Sensitive data*** means personal data revealing racial or ethnic origin, religious beliefs, a mental or physical health condition or diagnosis, a person's sex life or sexual orientation, citizenship, or citizenship status, as well as genetic or biometric data that may be processed for the purpose of uniquely identifying an individual. The definition also includes personal data from a known child.
Sec. 59:1-571, VCDPA	***Sensitive data*** means a category of personal data that includes data revealing racial or ethnic origin, religious beliefs, physical or mental health diagnosis, sexual orientation, or citizen or immigrant status, as well as processing of genetic or biometric data for identification, precise geolocation data, and personal data collected from a known child.
Sec. 13-61-101, UCPA	***Sensitive data*** is any data that reveals an individual's "racial or ethnic origin; religious beliefs; sexual orientation; citizenship or immigration status; or medical history, mental or physical health condition, or medical treatment or diagnosis; genetic personal data or biometric data; specific geolocation data."

Data Subject / Consumer / Identifiable Individual

Law / Regulation	Definition
Article 4(1), GDPR	*A* ***Data subject*** is any identifiable natural person located in the EU, or by a controller or processor located in the EU.

Law / Regulation	Definition
Sec. 1798.140, CCPA	*A **Consumer*** "is a natural person who is a California resident, … however identified, including by any unique identifier."
Sec. 6-1-1303, CPA	*A **Consumer*** is an individual who is a Colorado resident acting only in an individual or household context. Does not include individuals acting in a commercial or employment context, as a job applicant, or as a beneficiary of someone acting in an employment context.
Sec. 59:1-571, VCDPA	*A **Consumer*** is " a natural person who is a resident of the Commonwealth acting only in an individual or household context. It does not include a natural person acting in a commercial or employment context."
Sec. 13-61-101, UCPA	*An **identifiable individual*** means "an individual who can be readily identified, directly or indirectly."

Processing

Law / Regulation	Definition
Article 4(2), GDPR	***Processing*** refers to any "operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction."
Sec. 1798.140, CPRA	***Processing*** refers to "any operation or set of operations that are performed on personal information or on sets of personal information, whether or not by automated means." This includes the sale of personal information.
Sec. 6-1-1303, CPA	***Processing*** refers to "the collection, use, sale, storage, disclosure, analysis, deletion, or modification of personal data and includes the actions of a controller directing a processor to process personal data." This includes the exchange of information for monetary or other valuable consideration.
Sec. 59:1-571, VCDPA	***Processing*** refers to "any operation or set of operations performed, whether by manual or automated means, on

Law / Regulation	Definition
	personal data or on sets of personal data, such as the collection, use, storage, disclosure, analysis, deletion, or modification of personal data." This includes the exchange of personal data for monetary consideration.
Sec. 13-61-101, UCPA	***Process*** refers to an "operation or set of operations performed on personal data, including collection, use, storage, disclosure, analysis, deletion, or modification of personal data."

Controllers, Processors & Businesses[43]

Law / Regulation	Definition
Article 4(7) & (8), GDPR	A ***Controller*** is "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data" A ***Processor*** is "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller."
Sec. 1798.140, CCPA, CPRA	A ***Business*** is any legal entity that is "organized or operated for the profit or financial benefit of its shareholders or other owners, that collects consumers' personal information, or on the behalf of which such information is collected and that alone, or jointly with others, determines the purposes and means of the processing of consumers' personal information, that does business in the State of California." A ***Service Provider*** is a "legal entity that is organized or operated for the profit or financial benefit of its shareholders or other owners, that processes information on behalf of a business and to which the business discloses a consumer's personal information for a business purpose pursuant to a written contract." **Note:** As amended by the CPRA, a business must meet one of the following thresholds to be covered: i) has an annual revenue of $25,000,000 or more; ii) alone or in combination, annually buys, receives, sells, or shares, the personal

[43] Currently, Private.Storage is unlikely to be subject to any of the US privacy laws, as it does not meet the minimum thresholds regarding data subjects or business size.

Law / Regulation	Definition
	information of 100,000 or more consumers[44]; or iii) derives over 50% of their annual revenue from the sale or sharing of consumers' personal information.
Sec. 6-1-1303, CPA	A ***Controller*** is "A person that, alone or jointly with others, determines the purposes for and means of processing personal data." **Note:** A controller must meet threshold requirements of conducting business in Colorado or produces or delivers a commercial product or service targeted to 100,000 or more Colorado residents; or derives 50% of revenue or receives discounts from selling personal data and processes or controls personal data of 25,000 or more Colorado residents.
Sec. 59.1-571, VCDPA	A ***Controller*** is "A person that, alone or jointly with others, determines the purposes for and means of processing personal data." **Note:** A controller must meet threshold requirements of conducting business in Virginia or producing a product or service targeted to consumers in the Commonwealth that i) processes personal data of 100,000 or more consumers; or ii) derives over 50% of their gross revenue from the sale of personal data of at least 25,000 or more consumers.
Sec. 13-61-101, & 102 UCPA	A ***Controller*** is "a person doing business in the state who determines the purposes for which and the means by which personal data are processed, regardless of whether the person makes the determination alone or with others." A ***Processor*** is a "person who processes personal data on behalf of a controller." **Note:** A controller or processor must meet threshold requirements of conducting business in Utah or producing a product or service targeted to consumers in the State of Utah with an annual revenue of $25,000,000 or more and i) processes personal data of 100,000 or more consumers; or ii) derives over 50% of their gross revenue from the sale of personal data of at least 25,000 or more consumers.

[44] Beginning 1/1/2023.

Anonymous, Pseudonymous, Deidentified & Aggregate Data

Law / Regulation	Definition
Article 4(5), Recital 26 GDPR	***Anonymous data*** is data that cannot identify an individual. Anonymous data is **not** considered personal data. Although not explicitly defined, Recital 162 GDPR gives wider allowance for the processing of personal data for statistical purposes. ***Pseudonymous data*** refers to data that is processed "in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person." ***De-identified data and aggregate data*** are not defined under the GDPR.

Law / Regulation	Definition
Secs. 1798.140(b), (m), (aa), CPRA	***Anonymous data*** is not defined under the CCPA/CPRA. ***Pseudonymous data*** refers to "the processing of personal information in a manner that renders the personal information no longer attributable to a specific consumer without the use of additional information, provided that the additional information is kept separately and is subject to technical and organizational measures to ensure that the personal information is not attributed to an identified or identifiable consumer." ***De-identified data*** means "means information that cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer provided that the business that possesses the information: (1) Takes reasonable measures to ensure that the information cannot be associated with a consumer or household. (2) Publicly commits to maintain and use the information in deidentified form and not to attempt to reidentify the information, except that the business may attempt to reidentify the information solely for the purpose of determining whether Its deidentification processes satisfy the requirements of this subdivision (3) Contractually obligates any recipients of the information to comply with all provisions of this subdivision.The CCPA does not restrict the use, sale, retention or disclosure of deidentified data. ***Aggregate Consumer Information*** "means information that relates to a group or category of reasonably linkable to any consumer or household, including via a device. 'Aggregate consumer information' does not mean one or more individual consumer records that have been deidentifled."
Sec. 6-1-1303 (11), (22), CPA	***Anonymous data*** is not defined under the CPA. ***Pseudonymous data*** " means personal data that cannot be attributed to a specific natural person without the use of additional information, provided that such additional information is kept separately and is subject to appropriate technical and organizational measures to ensure that the

Law / Regulation	Definition
	personal data is not attributed to an identified or identifiable natural person." ***De-identified data*** "means data that cannot reasonably be used to infer information about, or otherwise be linked to, an identified or identifiable individual, or a device linked to such an individual, if the controller that possesses the data: (a) takes reasonable measures to ensure that the data cannot be associated with an individual; (b) publicly commits to maintain and use the data only in a de-identified fashion and not attempt to re-identify the data; and (c) contractually obligates any recipients of the information to comply with the requirements of this subsection. The CPA does not restrict the use, sale, retention or disclosure of deidentified data. ***Aggregated data*** is not specified under the CPA.
Sec. 59.1-571, VCDPA	***Anonymous data*** is not defined under the VCDPA. ***Pseudonymous data*** "means personal data that cannot be attributed to a specific natural person without the use of additional information, provided that such additional information is kept separately and is subject to appropriate technical and organizational measures to ensure that the personal data is not attributed to an identified or identifiable natural person. ***Deidentified data*** is "means data that cannot reasonably be linked to an identified or identifiable natural person, or a device linked to such person. A controller that possesses 'de-identified data' shall comply with the requirements of subsection A of § 59.1-577. . ***Aggregated data*** is not specified under the VCDPA.
Sec. 13-61-101, & 102 UCPA	***Anonymous data*** is not defined under the UCPA. ***Pseudonymous data*** "means personal data that cannot be attributed to a specific individual without the use of additional information, if the additional information is: (a) kept separate from the consumer's personal data; and

Law / Regulation	Definition
	(b) subject to appropriate technical and organizational measures to ensure that the personal data is not attributed to an identified individual or an identifiable individual. ***Deidentified data*** means "data that: (a) cannot reasonably be linked to an identified individual or an identifiable individual; and (b) are possessed by a controller who: (i) takes reasonable measures to ensure that a person cannot associate the data with an individual; (ii) publicly commits to maintain and use the data only in deidentified form and not attempt to reidentify the data; and (iii) contractually obligates any recipients of the data to comply with the requirements described in Subsections (14)(b)(i) and (ii). ***Aggregated data*** means "information that relates to a group or category of consumers: (a) from which individual consumer identities have been removed; and (b) that is not linked or reasonably linkable to any consumer.

Other Terms

The following table identifies the terms referred to in this DPIA.

Personal Data Breach	A breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed;
Data Flow	A graphical representation of the flow of information or data through a system.
Encryption	A technical process to convert data into an unreadable format which cannot be unconverted by an unauthorized individual.
Subject Access Request / "SAR"	A request made under Article 15 of the GDPR
Privacy	A fundamental right of individuals to be left alone which is recognised by jurisdictions around the world and the EU Charter of Fundamental Rights.
Data Protection Impact Assessment	A tool which can help organizations identify the impact of the envisaged processing operations on the protection of personal data. It should help an organization to identify and reduce the privacy risks of a project, with a focus on data.
Data Protection by Design	The embedding of data protection features and data protection enhancing technologies directly into the design of projects at an early stage. This includes implementing appropriate technical and organizational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.
Data Protection by Default	Implementing appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons
Third Country	A country which is not a member of the European Economic Area (EEA)