

Palantir Privacy and Governance Whitepaper

PALANTIR.COM

COPYRIGHT © 2024
PALANTIR TECHNOLOGIES INC.

ALL RIGHTS RESERVED

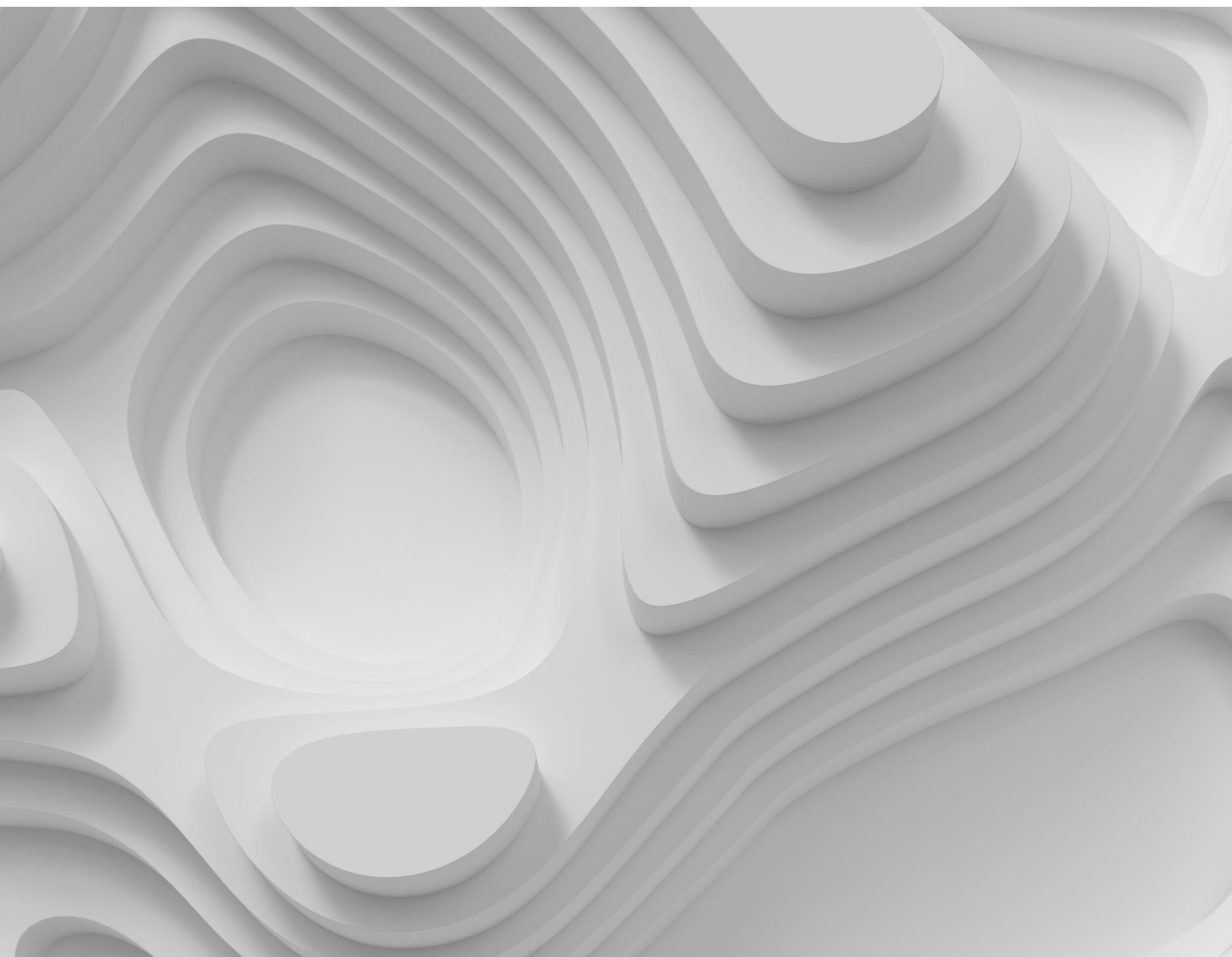
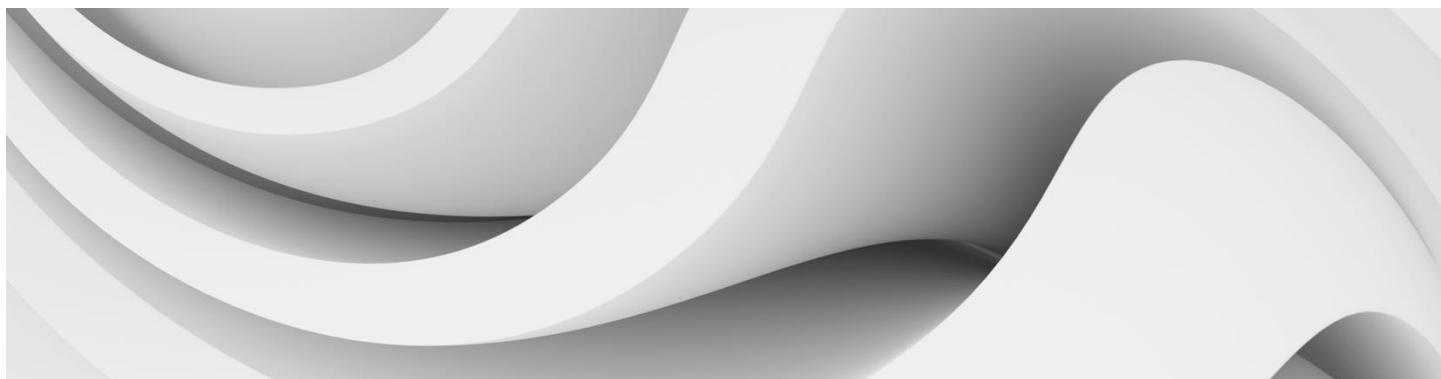


Table of Contents

1.	Introduction	03
2.	Privacy and Civil Liberties at Palantir	04
3.	Maximizing Governance Through Palantir Platforms	05
4.	Governance Principles and Platform Offerings	06
i.	Security and Integrity	06
ii.	Transparency	08
iii.	Use Limitation and Purpose Specification	12
iv.	Data Minimization	17
v.	Retention and Deletion	20
vi.	Accountability and Oversight	23
5.	Conclusion: Value in Governance Across Industries	25



1 Introduction

Palantir Technologies builds software platforms that empower organizations to effectively integrate their data, decisions, and operations to tackle real-world problems. For over two decades, these platforms have been employed by public, private, and non-profit institutions worldwide, operating in highly sensitive data environments, to support their most complex mission-critical objectives. By enabling data-driven insights, our software equips decision-makers with the tools to address pressing challenges, improve operational efficiency, and foster innovation, all while maximizing transparency, governance, and security over their data. Going back to our company's founding commitments, we believe that the institutional use of powerful information technologies should not come at the expense of the privacy, civil liberties, and fundamental rights of the individuals served by those institutions.

In this vein, privacy protection and responsible data governance consistently remain among our foremost priorities throughout the development and deployment of our software. Over the last 20 years, we have devoted significant and continuing resources to constructing comprehensive privacy, governance, and security features that enable our clients to fulfill regulatory and compliance obligations while improving their own privacy and security practices. Our steadfast commitment to investment in this area has helped us to establish a reputation for delivering the highest standard of data protection and security products, ensuring the protection of sensitive information while empowering organizations to fully utilize their data assets. As a result, Palantir has become a trusted partner for some of the most critical institutions and missions worldwide.

This whitepaper lays out how Palantir's products provide a comprehensive infrastructure for customer privacy, governance, and security needs - enabling organizations across industries to know their data is secure, have comprehensive oversight and control over how that data is accessed and used, and readily comply with regulatory data requirements, all within a single platform.

Palantir's three flagship platforms for enterprise data management are [Foundry](#), [Gotham](#), and [AIP](#) (Artificial Intelligence Platform). The features, capabilities, and product approaches described in this whitepaper exist in some form in all three and are available off-the-shelf for customers to deploy. Throughout this document, we'll refer to "Palantir platforms" to represent all three.

2 Privacy and Civil Liberties at Palantir

At Palantir, we strongly believe that effective data integration and analysis should go hand in hand with robust security, governance, and privacy controls. This belief is instantiated not only in the product investments we've made since our founding, but also in our company culture in many forms, including the longstanding [Privacy and Civil Liberties](#) (PCL) Engineering team. PCL, in close partnership with security, legal, compliance, data protection, technical architecture, and other teams at Palantir, focuses on translating ethical, policy, regulatory, and other important privacy and security principles into practical features and capabilities within our platforms.

Palantir's approach to privacy and governance engineering blends security, privacy, compliance, and regulatory standards alongside technical architecture. We develop and deploy privacy-enhancing technology to address increasingly significant and complex questions and are opinionated in adopting product development practices that give primacy to security- and privacy-by-design approaches. Moreover, we believe our ethical responsibility extends beyond incorporating existing formal regulatory requirements into our products; we should also be building towards future policy developments, as well as sharing our insights with policymakers to [help shape policies that](#) establish the most relevant and advanced standards for privacy and security.

This commitment to privacy and civil liberties is central to our company mission - we recognize that our customers include some of the most trusted institutions in the world who demand full control and transparency over their data in the information technology products they use to manage and make use of those assets. Our products are therefore built to simultaneously deliver core data integration and analysis features, while operationalizing governance and security. In Palantir's platforms, both mission and governance are treated as two sides of the same coin, enabling our customers to operate as effectively, efficiently, transparently, and responsibly as possible.

3 Maximizing Governance Through Palantir Platforms

Organizations today face a wide array of challenges with respect to protecting and governing their data environments. Not only must they meet continually evolving regulatory standards, but they also face disparate landscapes of legacy and new systems and data streams that often severely challenge their ability to attain a reliable view over how their data is being used. Palantir products help to address this challenge by **enabling comprehensive governance for our customers** - allowing these organizations to clearly define how their data is being accessed, shared, used, and analyzed across our platform, in addition to engendering confidence that it is safe and adhering to privacy standards.

In this way, we not only serve as the leading software provider for enabling customers with vast, disjointed data landscapes to unify their systems into a common operating picture and unlock the value of their data. Our platforms also act as a robust governance solution, with products and features that:

- **Maximize security:** meeting the highest standards of security accreditations and safety protocols for storing and handling data.
- **Give organizations transparency and full control** over how their data is being integrated, accessed, used, and interacted with, wherever it is stored.
- **Instantiate privacy, ethics, norms** into the backbone and UI of our products and applications.
- **Achieve streamlined compliance** with global privacy and data protection regulation.

Over the last twenty years we have continually focused on fostering these capabilities into our product development practices and culture. This creates immense value for our customers, primarily by limiting the need to procure separate platforms for privacy-enhancing technology, governance, and compliance.

In the following sections of this whitepaper, we illustrate how complex and interwoven governance topics are addressed within the Palantir platforms. These concepts include Security and Integrity; Transparency; Purpose Limitation; Data Minimization; Retention and Deletion; and Accountability and Oversight. By translating the larger ideals - security, privacy, governance - into concrete goals (policy/regulatory and operational) and product features, we hope to provide a clear roadmap on how organizations can leverage our platforms to facilitate governance capabilities.

4 Governance Principles and Platform Offerings

As mentioned in the overview of this whitepaper, Palantir has several platform offerings. In this section, we focus on capabilities and features within the Foundry platform, which is widely used by both private and commercial institutions. As a data integration and analysis engine, Foundry underpins our Artificial Intelligence Platform (AIP) and these two can be considered as interconnected. Foundry is also often deployed as the base layer for Palantir Gotham, which is our global operating system commonly used in defense, intelligence, disaster relief, and other domains. As such, it is the best vehicle for examining governance and privacy capabilities available for Palantir customers, and is the platform used by global commercial customers across industries.

4.1 Security and Integrity

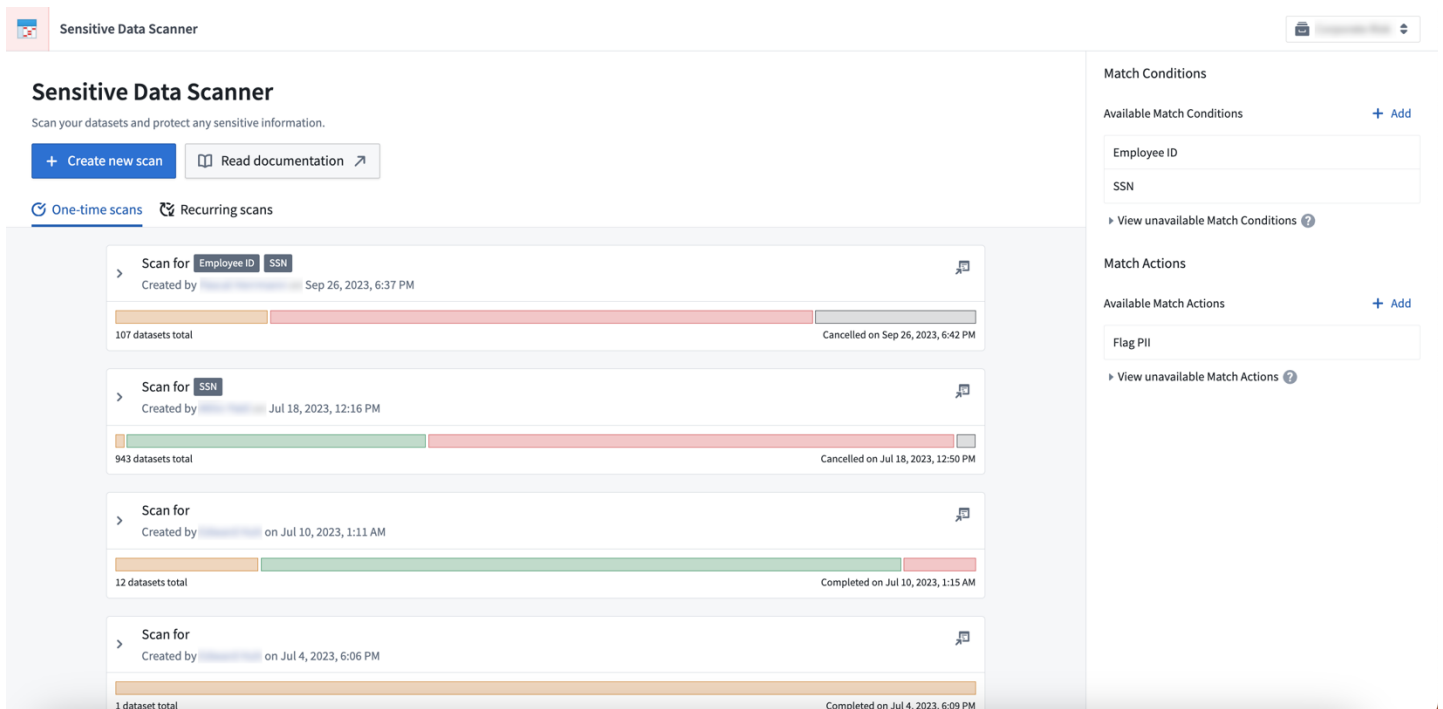
The security and integrity of a software platform are important foundational elements to enabling privacy protections and sound governance for an organization. Ranging from product to infrastructure to platform management systems - strong security standards are essential to protecting intellectual property, defending against threats, and enabling users to leverage the information technologies in a safe and trusted environment. For this reason, Palantir has always approached our platform architecture with a keen focus on maximizing security at every step. This systematic security approach, highlighted in our [security accreditations](#), is a key reason why we are able to operate in the most sensitive environments with customers who must maintain the highest levels of care for their data. The Palantir Platform Security Whitepaper, available upon request, provides a detailed guide on our architecture, and we also make available technical documentation on features as fine-grained as, for example, [native support for multi-factor authentication as well as intentional periodic re-authentication mandates](#) within short time intervals.

Security is a base layer that enables platform integrity, which we define as the ability for platform administrators to have high confidence in holding a comprehensive understanding of what data exists in the environment they are managing. These administrators often face the challenge of monitoring sensitive data activity across an organization and having to ensure compliance with a range of existing and emerging regulations.

Security and Integrity

In an ideal world, all privacy safeguards are taken proactively. But in reality, organizations are often fast-moving and trying to solve business-critical problems in a scalable, efficient manner. This is why we equip our products with off-the-shelf tools to enable data stewards to effectively scan for sensitive data at key points in the data lifecycle, including at the initial stage when data is integrated into our customer’s instance of the Palantir platform.

Example 1: [Sensitive Data Scanner](#): A tool that allows customizable scans to flag sensitive data and ensure compliance



Using readily configurable controls, data administrators can leverage this tool to define what sensitive data means for their organization, pair those definitions with data structures representing the sensitive data, and then perform scans across their organization’s data either ad-hoc or at set intervals. Administrators can rely on these continual backend scans to flag potentially noncompliant data and take appropriate actions to secure sensitive data.

With these and other platform-enabled security and integrity features activated to underpin Foundry, organizations can then focus on operationalizing Palantir’s governance and privacy products to meet their needs.

4.2 Transparency

Transparency is essential to enabling governance in software platforms. In a software and data context, transparency means understanding and being able to report on the flow of data across a platform - how it is ingested, accessed, transformed, and shared at scale within your organization.

From a regulatory and compliance perspective, transparency is a cornerstone of many privacy frameworks and legislative regimes - including the Fair Information Practice Principles (FIPPs), GDPR, and most state-level privacy laws such as the California Consumer Privacy Act (CCPA). This is because transparency is a prerequisite to accomplishing the other regulatory mandates such as disclosures, access controls, and the exercise of data subject rights (e.g., right to access and the right to be forgotten). Only when an organization is able to accurately identify and report on data within its operations can it then build additional functionality that facilitates more in-depth governance goals.

Operationalizing Transparency

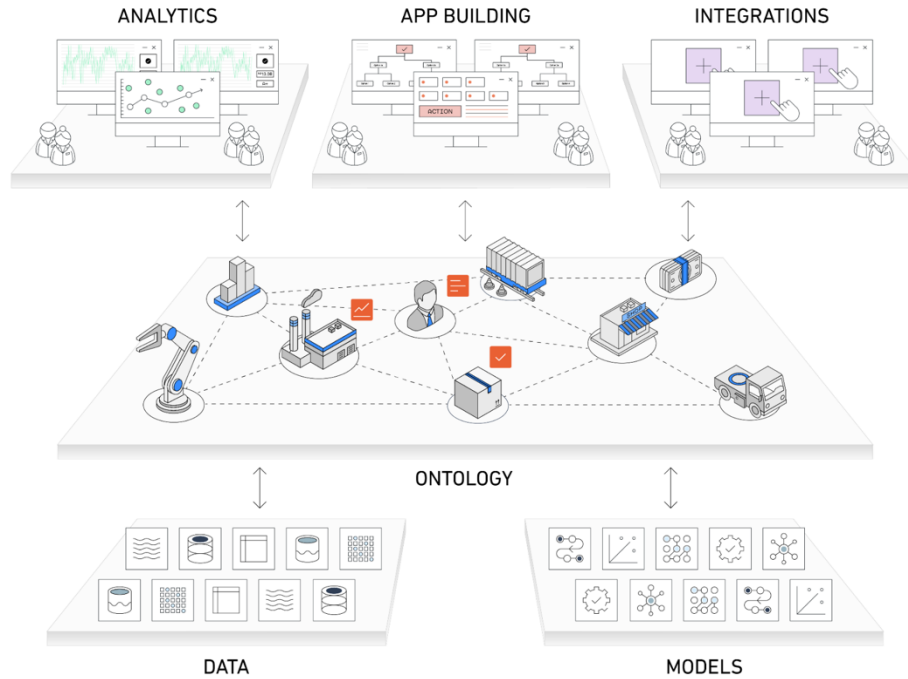
Within an organization, transparency can take many forms. At the platform level, transparency is generally distilled into tools and approaches that enable an understanding of the data landscape and offer the ability to report, at a granular level, on key actions related to that data. These action types could be the ingestion of data, users accessing data fields, transforming it for harmonization or analysis, exporting it to share with internal or external audiences, among other data interactions. This section highlights how Palantir platforms foster transparency for our customers via data integration, lifecycle visualization, and interaction checkpoints.

Palantir platforms begin enabling transparency by helping organizations achieve a unified data landscape, or common operating picture, via what we call an **“[Ontology](#)”** (*screenshot below*). Put simply, an Ontology is a semantic layer that integrates the data, logic, and actions of an institution, allowing them to represent the nouns and verbs of their organization’s operations, and interact with those as objects or representations, within Foundry.

Instead of users having to interact with their data in unstructured or raw formats like tables, the Ontology represents data and its uses in terms that relate to real-world concepts, that are readily understandable by domain experts, and that incorporate the language specific to that organization. The Ontology sits “on top” of an organization’s datasets and models, connecting these digital artifacts and sources to their real-world counterparts, ranging from physical assets, like equipment, to concepts such as transactions or logistics. Put simply, the Ontology takes an organization’s vast, disparate data and formats it into a common and relatable picture. This integration is the first step to achieving transparency.

Transparency

Example 2: Ontology: An Organization's Common Operating Picture



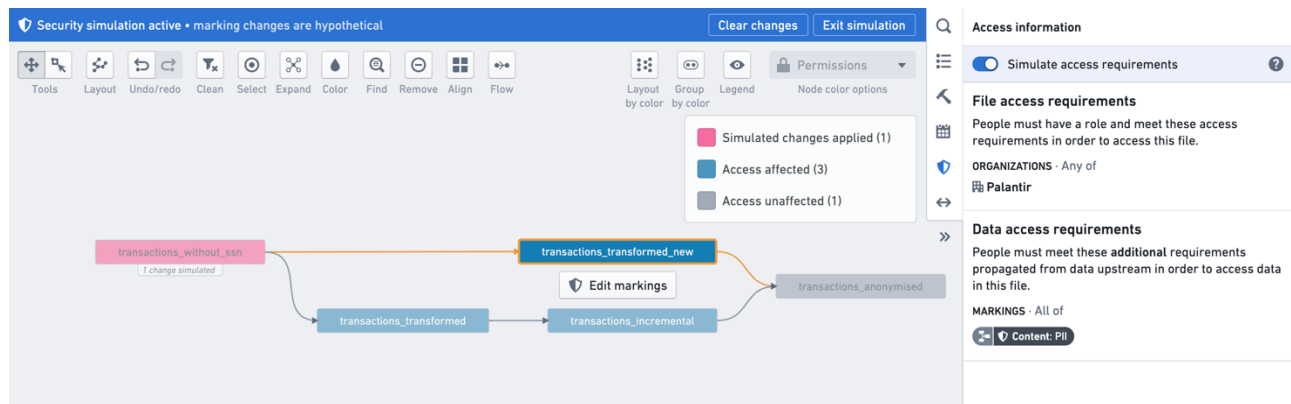
Palantir platforms further operationalize transparency through clear and digestible visualizations of a data landscape (screenshot below).

Example 3: [Data Lineage](#): A tool that facilitates easy understanding of data flows and access across an organization.

The screenshot shows the Palantir Data Lineage tool interface. The main view displays a data flow: **Example Data** (red node) feeds into **Example Contour** (orange node), which then feeds into **Example Report** (purple node). A permissions legend is visible, showing four categories: **View all data (0)** (blue square), **No access (1)** (red square), **No data (0)** (dark grey square), and **Permissions unknown (0)** (light grey square). A note states: "You are only able to view a user's access level if you have that access level". The permissions type is set to "Data access in datasets". The "View as:" dropdown is set to "Test User". A tooltip is visible over the dropdown menu, listing "Data access in datasets" and "Resource access".

Transparency

Example 3 Continued: [Data Lineage](#): A tool that facilitates easy understanding of data flows and access across an organization.



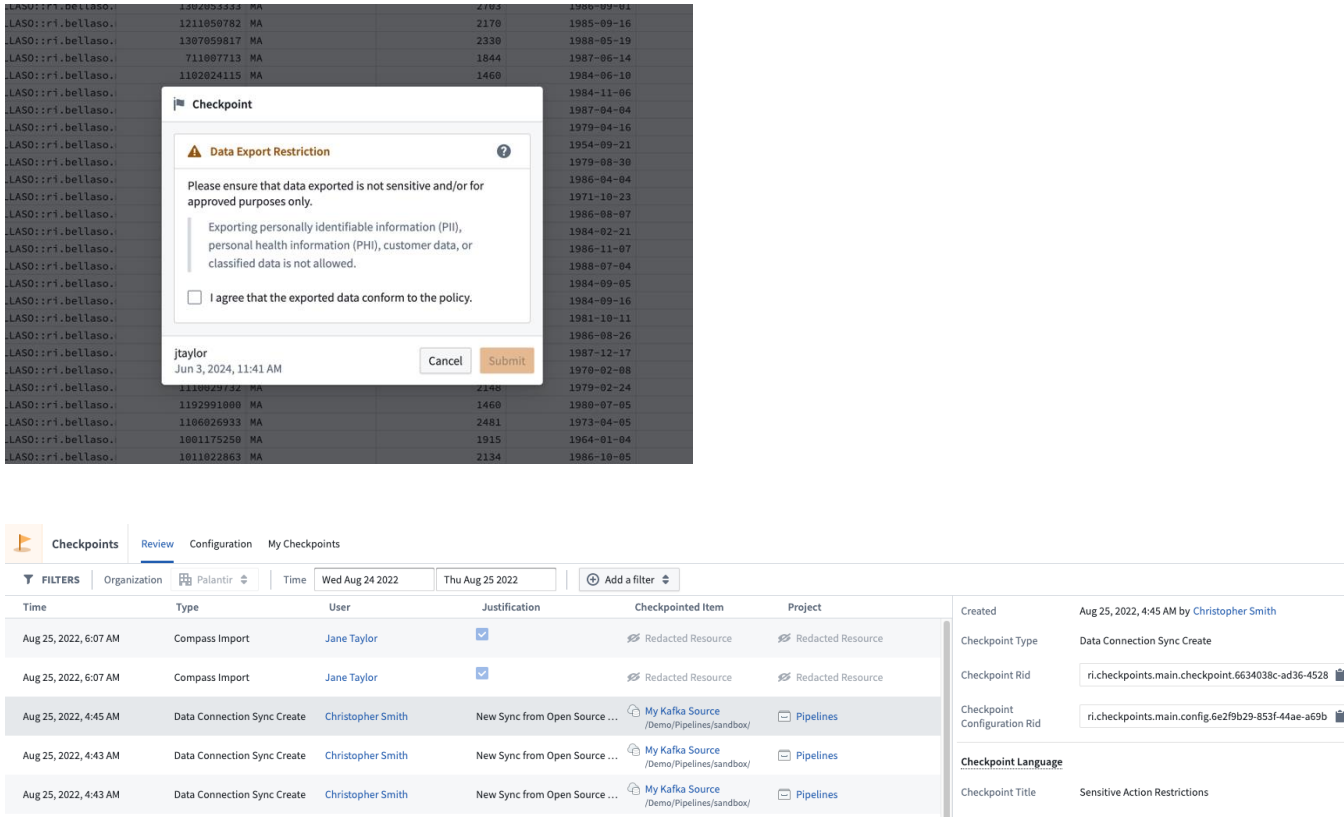
Understanding data lineage allows administrators within an organization to visualize the flow of their data across the platform. Reading the image from left to right, one can see data on the platform from ingest to data transformations to platform applications; overarching views of data lifetimes and interactions paint a clear picture at scale. Each of these views can be further analyzed, allowing users to understand data health. This enables them to schedule checks to keep data fresh and to investigate access and security either at a high level or in more granular detail.

Finally, to facilitate greater transparency, users and administrators need to be able to monitor data actions, such as access or exporting, especially with respect to sensitive data. Actions may be mapped to justifications in order to ensure compliance with data protection and governance policies, but also to understand what types of data are being interacted with, why, and by whom.

Palantir creates tooling that allow administrators to set highly customizable controls around actions within the platform. The [Checkpoints](#) product is configurable both in what data actions it applies to, as well as the nature of the checkpoint and justifications required to proceed with an action.

Transparency

Example 4: [Checkpoints](#): A tool that facilitates purpose justifications and auditability for sensitive actions.



As seen above, an example of this could be an organization with policies that require any data export from the platform to be formally acknowledged by the user, along with a justification and a corresponding audit log entry, due to its sensitivity. These checkpoints are highly configurable. Administrators can choose the format in which justifications should be provided (e.g., ticked box, free text) as well as the types of actions to which they are applied (e.g., import, export, sharing, transformation, accessing certain data fields, etc.). Importantly, linking actions to granular audit logs of these decisions - with user IDs and timestamps - allows an organization to build a transparent picture of data interactions at scale.

Using platform governance features allows customers to leverage the Palantir platform to solve their critical data integration and analysis problems while maintaining a transparent picture of their data in real-time. Governance features sit alongside other core functionality as a reminder that privacy, security, and responsible data best practices can be incorporated by design and need not present an undue burden for software use. They can be utilized in a single interface and maximize organizational impact while adhering to regulatory and compliance standards.

4.3 Use Limitation & Purpose Specification

Purpose specification is the concept of clearly stating what data is required for processing according to a specific purposes, while use limitation is the act of ensuring data and product uses, including those that involve personal data, are confined to predetermined and authorized applications. In both instances, it is essential to limit the processing of that data to a designated set of purposes and provide auditability and consent tracking if those purposes change over time and require new data to be accessed. These concepts present a challenge at the intersection of policy and operations: defining purposes and mapping required data, and technical architecture of rails to limit data processing and associated product use cases according to those purposes. At Palantir, we treat these concepts as interrelated, i.e. “purpose limitation”, in terms of how we design and execute platform features.

We enable organizations to achieve purpose limitation at scale through multiple products and features that provide granular security markings on data, controls over access, and the ability to link access to set purposes. An organization can use Palantir’s products to determine what data is needed to solve specific problems, ensure that security and access are tightly monitored according to institutional or regulatory needs, directly link purposes to data assets, and seamlessly provide records of this access history over time.

Capabilities that Enact Purpose Limitation

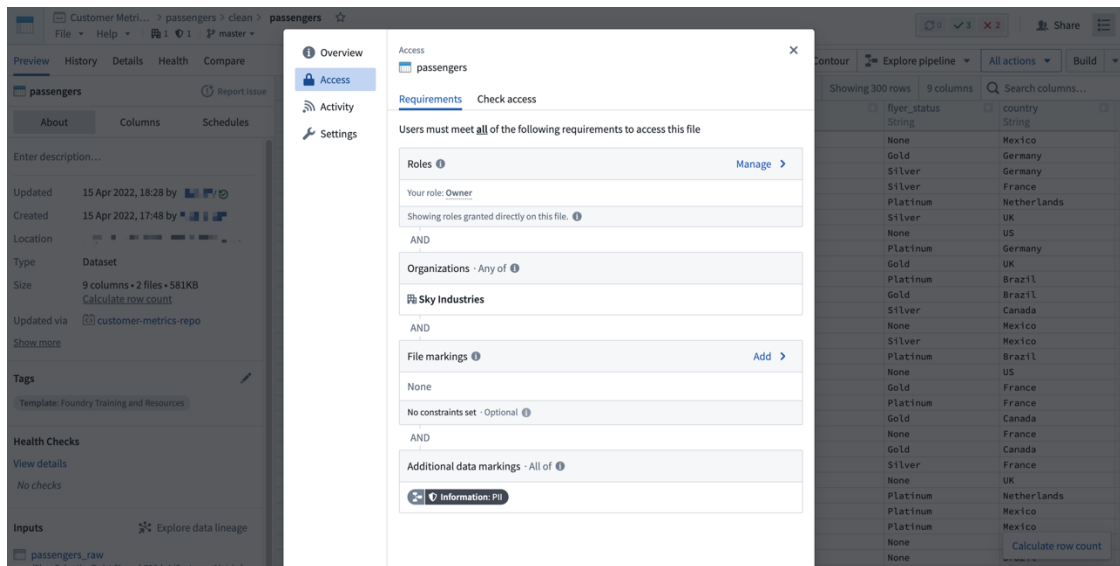
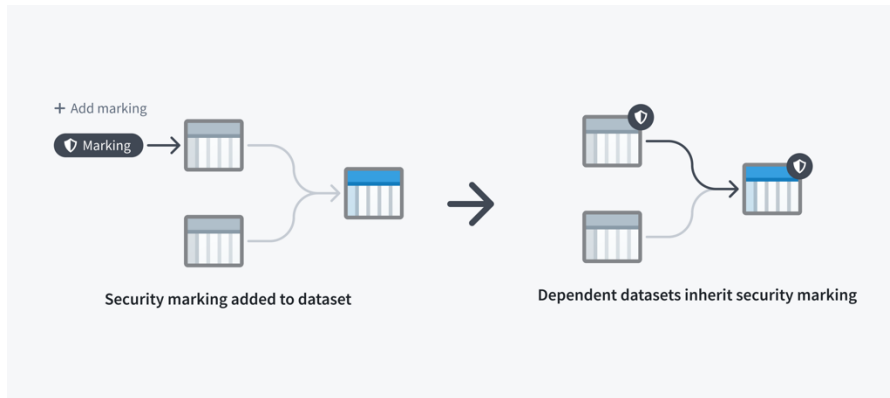
Purpose limitation starts with aligning specific problems to technical solutions using the least amount of required data. The effectiveness hinges on an organization's transparency regarding the data they use, along with its usage purpose and methods. This challenge becomes even more complex at scale - managing hundreds or thousands of purposes across a vast organization, each with their own attributes, timelines, and scope.

The solutions then become increasingly technically challenging: ensuring that you have fine-tuned markings on data, can attribute markings to specific access within your organization, and link access based on a specific purpose, thus closing the loop.

The Palantir platforms enable administrators to apply granular security markings to data the moment it is ingested. These markings can be applied at the most granular levels, such as the individual cell of a tabular data sheet, and indicate the sensitivity of the underlying data assets, map to user access, and apply downstream wherever that data may go within the platform. This means that even if data is transformed or changed for downstream analysis, administrators can be confident that the access control picture remains consistent.

Use Limitation and Purpose Specification

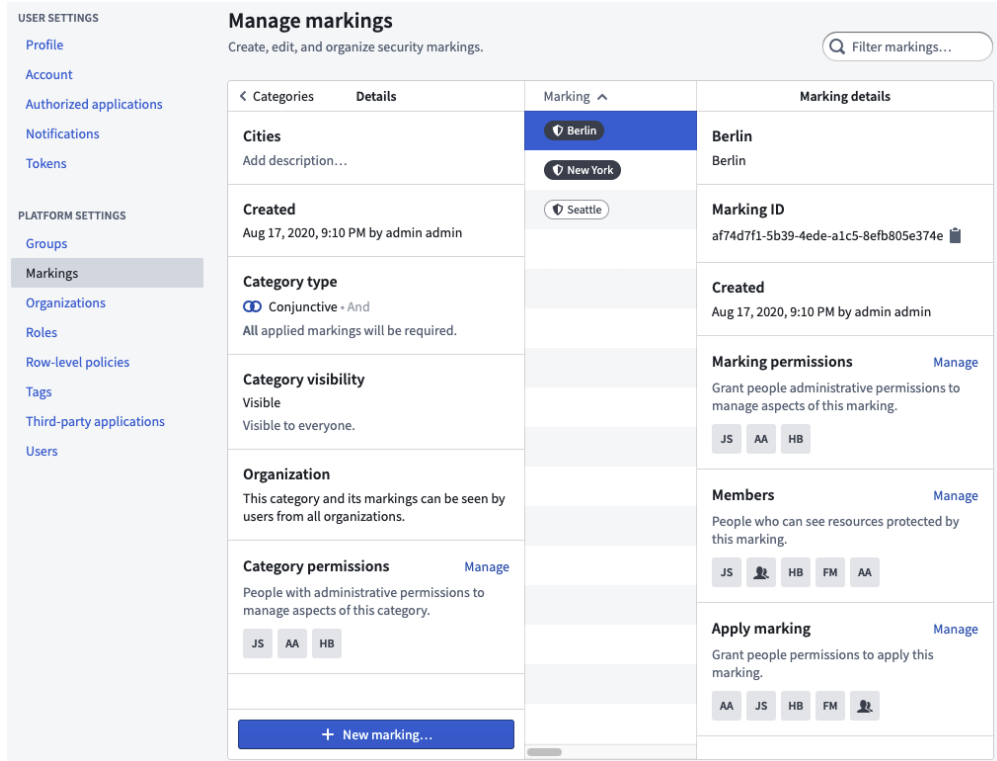
Example 5: Security Markings on Data Assets



Security markings can be attributable to specific use cases, workflows, projects, or entire organizations - enabling purpose limitation controls to be achievable wherever data may flow. These markings also form the basis of data access controls by determining who is able to access what data. Access controls within Palantir, seen below, are highly configurable sets of permissions around data that can be scaled to organizations, user groups, geographies, use cases, or whatever requirements an organization may have.

Use Limitation and Purpose Specification

Example 6: Security Markings and Creating Robust Access Controls



These access controls are the essential next step to achieving purpose limitation at scale. They allow an organization to take a purpose - for example, resource allocation within a given geography - and tie it to specific data assets. Data stewards can label which purposes justify access to data ingested onto the platform at a highly granular level, selecting for specific cells or subsets of a given data asset.

This approach, labelled **“Purpose Based Access Controls” (PBAC)**, goes beyond typical role based access control structures and allows for a higher level of control of data access that can be scaled and audited across an organization.

Use Limitation and Purpose Specification

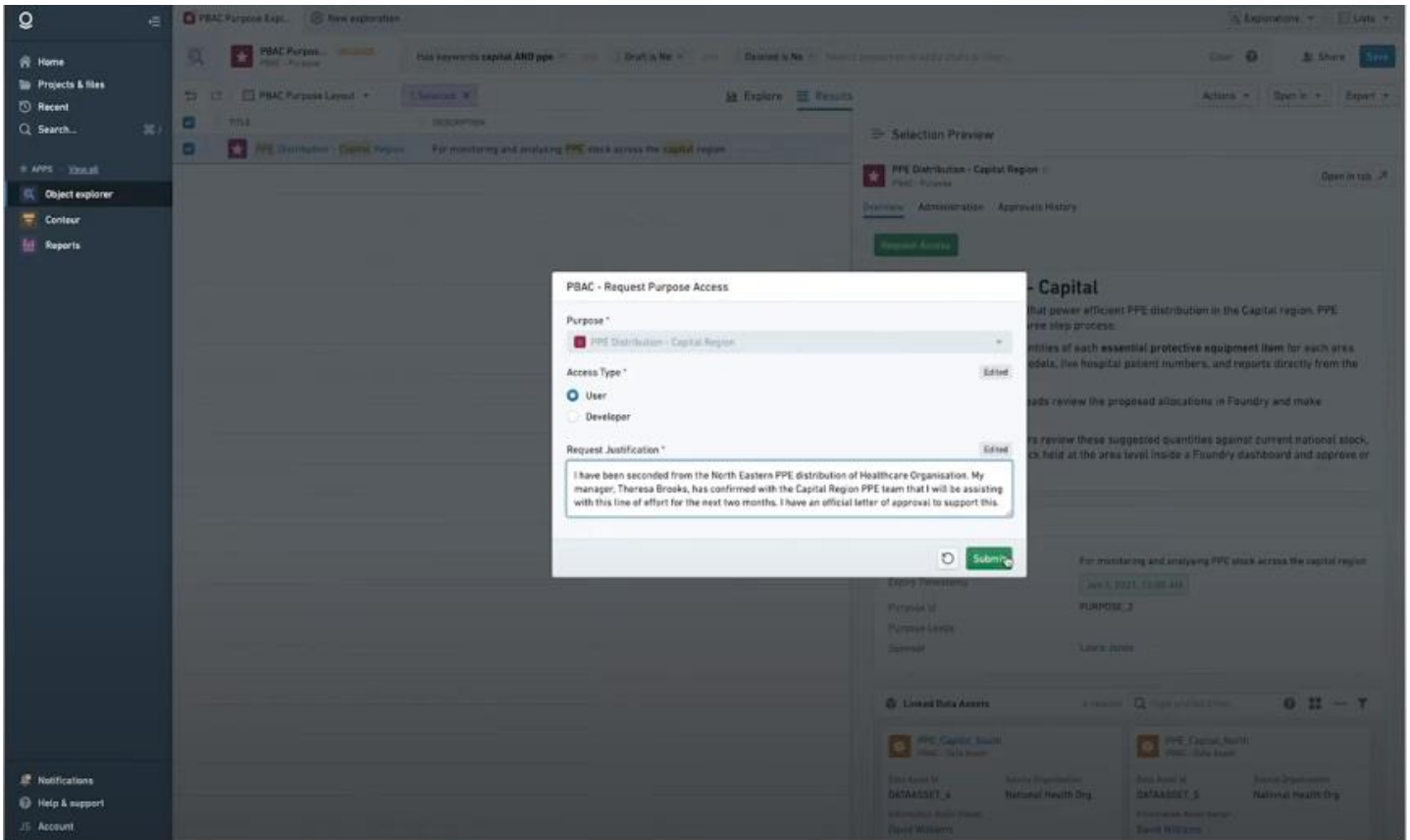
Example 7: [Purpose Based Access Controls \(PBAC\) Monitoring Dashboard](#)

TITLE	DESCRIPTION	SPONSOR	EXPIRY TIMESTAMP	PURPOSE ID
PPE Distribution - West Region	For monitoring and analysing PPE stock across the West region	Laura Jones	Jan 1, 2021, 1:00 AM	PURPOSE_10
Supply Chain - Consumables - Di...	ICU Consumables supply chain data at regional level	John Smith	Jan 1, 2021, 12:00 AM	PURPOSE_12
Supply Chain - Consumables - Di...	ICU Consumables supply chain data at regional level	John Smith	Mar 1, 2021, 12:00 AM	PURPOSE_14
Supply Chain - Consumables - Di...	ICU Consumables supply chain data at regional level	John Smith	Mar 1, 2021, 11:00 AM	PURPOSE_16
Emergency Services Calls	Call Data to emergency services	John Smith	Mar 1, 2021, 12:00 AM	PURPOSE_17
PPE Distribution - North East Re...	For monitoring and analysing PPE stock across the North East region	Laura Jones	Jan 1, 2021, 10:00 AM	PURPOSE_4
PPE Distribution - Central Region	For monitoring and analysing PPE stock across the Central region	Laura Jones	Mar 1, 2021, 12:00 AM	PURPOSE_8
Equipment 360 View	360 View gives providers a holistic view of consumables and PPE stock levels as well as a predictive forwar...	John Smith	Sep 1, 2021, 1:00 AM	PURPOSE_11
Supply Chain - Consumables - Di...	ICU Consumables supply chain data at regional level	John Smith	Jan 1, 2021, 12:00 AM	PURPOSE_13
Supply Chain - Consumables - Di...	ICU Consumables supply chain data at regional level	John Smith	Mar 1, 2021, 12:00 AM	PURPOSE_15
PPE Distribution - Capital Region	For monitoring and analysing PPE stock across the capital region	Laura Jones	Jan 1, 2021, 12:00 AM	PURPOSE_2
PPE Distribution - South East Re...	For monitoring and analysing PPE stock across the South East region	Laura Jones	Mar 1, 2021, 12:00 AM	PURPOSE_7
Supply Chain - Medicine	For collating and analysing data about the provision and allocation of medicines across the national network	Laura Jones	Jan 1, 2021, 12:00 AM	PURPOSE_3
PPE Distribution - East Region	For monitoring and analysing PPE stock across the East region	Laura Jones	Mar 1, 2021, 12:00 AM	PURPOSE_6
PPE Distribution - South West R...	For monitoring and analysing PPE stock across the South West region	Laura Jones	Jan 1, 2021, 12:00 AM	PURPOSE_5
PPE Distribution - North West Re...	For monitoring and analysing PPE stock across the North West region	Laura Jones	Mar 1, 2021, 12:00 AM	PURPOSE_9

Seen above in a notional example of how our products are used by healthcare organizations. More specifically, it depicts how a data administrator is able to leverage PBAC to monitor purposes for data across their institutions. The Palantir platforms enable administrators to have readily retrievable records on the details of those purposes including their linked data assets, current members, and when access expires. This also scales to individual users being able to request access to a given purpose for their role.

Use Limitation and Purpose Specification

Example 8: [PBAC](#) Requests: governing data access requests



Seen in the interface above, users have configurable requests to access data for a given purpose. These requests and their corresponding approvals and permissions are auditable by the data administrators within an organization, allowing them to view their access control landscape over time.

This approach and series of features, available within our platforms, foster purpose limitation at scale. By blending together policy, operational, and technical solutions, Palantir platforms enables streamlined governance and transparency, highly configurable and granular security markings and access controls, and the ability to tie all these concepts together and attribute them back to purposes.

4.4 Data Minimization

Unlocking the value of your organization's data should not come at the expense of revealing sensitive information. This is where data minimization comes into play. Data minimization is typically regarded as a privacy principle that relates to limiting the collection and retention of personal data that is necessary and relevant to a specific purpose. Since Palantir platforms focus on management and use of enterprise data rather than collection, our software helps reinforce data minimization at the additional level of usage, i.e., beyond the initial phase of acquisition. This helps organizations reduce the risks associated with storing sensitive data for extended periods of time without losing out on the potential insights derived through integration and analysis.

Beyond access controls, Palantir platforms also enable institutions to apply dynamic data minimization procedures such as pseudonymization and selective masking to meet the complex and often context-dependent requirements of data privacy statutes and other institutional data governance requirements. These capabilities empower enterprises to align their data processing operations with applicable data minimization principles, simultaneously reinforcing both purpose limitation and data security.

Methods for Reducing and Protecting Sensitive Information

There are numerous approaches to addressing data minimization requirements and governance objectives. In this section, we'll focus primarily on how Palantir's platforms enable anonymization, which encompasses various methods depending on context, each with its own advantages and disadvantages.

It is worth noting that much of the impact of data minimization hinges on what data an organization decides to use for given operational use cases. For example, even if data is aggregated, generalized, or pseudonymized, there is still a risk of re-identification depending on the context of the workflow and attributes of the data that is present. Palantir's easily configurable tools help meet customers wherever they are in their governance journey, while also providing expansive tools that we often work with customers in the field to implement to meet rigorous data governance and regulatory obligations.

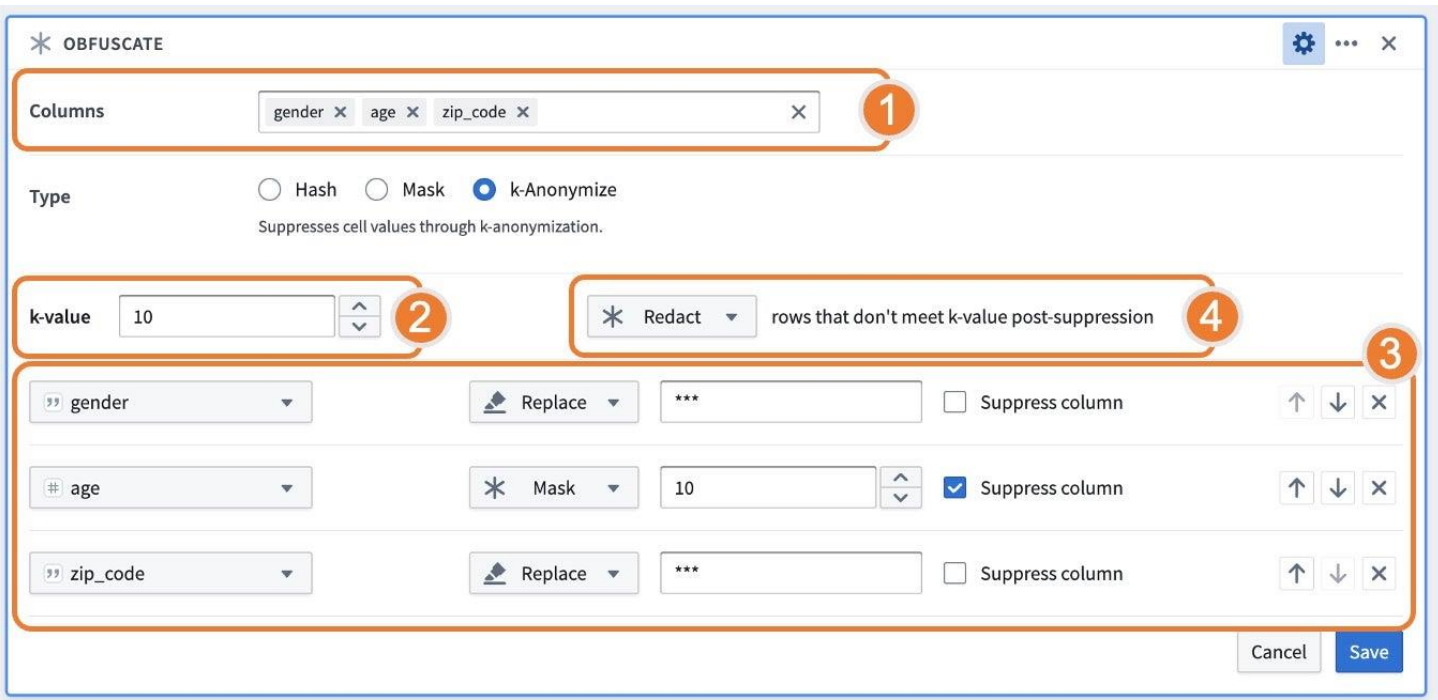
Data Minimization

Methods for Reducing and Protecting Sensitive Information

K-anonymization refers to a statistical process of transforming a dataset such that for each identifying property within the dataset, there needs to be at least k-1 records with the same properties in order for it to be displayed. It can be obtained by using a combination of techniques like “masking” and “generalization”.

Within Palantir platforms, organizations can [operationalize k-anonymization](#) by suppressing particular data elements on their datasets, as seen in example 9 below.

Example 9: K-Anonymization



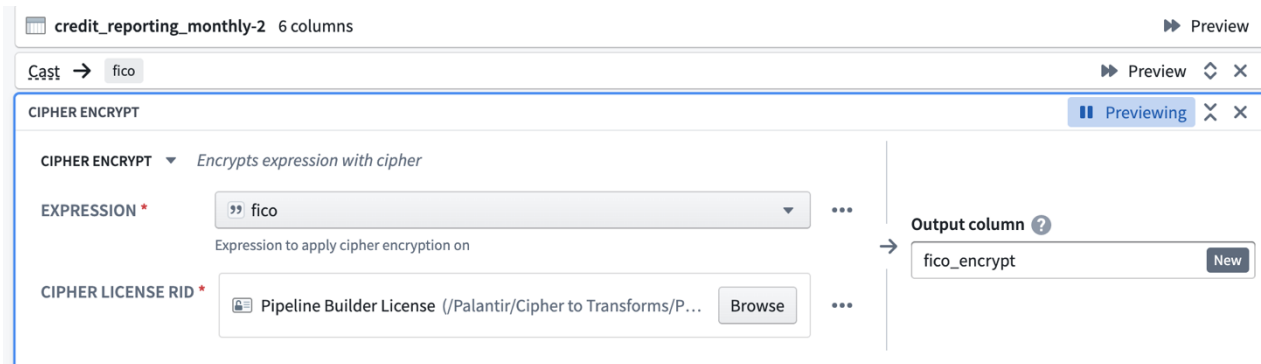
Moving into other data minimization techniques, Palantir platforms come with off the shelf obfuscation tooling. Obfuscation refers to the process of obscuring, or encrypting the meaning of data and replacing its visual representation with non-revealing and unreadable cipher representations. Masking, a prevalent obfuscation method, involves substituting data with realistic but fake information. Encryption, a more sophisticated technique, replaces identifiable data points with encrypted values that can be reverted through decryption. In contrast, hashing is an irreversible form of obfuscation; it's impossible to ascertain the original value from its properly hashed version.

Data Minimization

The Palantir platforms use sophisticated and recognized encryption standards at the storage and network levels to secure data in transit and at rest. Palantir platforms further offer an additional layer of protection that helps organizations achieve encryption through easy-to-use, no-code applications in order to configure privacy and governance protections in operational workflows.

As seen in the product example below, operational users can easily obfuscate information (either through encryption or hashing) by selecting columns within a dataset on which to apply these operations. This approach yields two key benefits: first, encrypting data does not hinder analysis, as both general data operations and AI interactions can still be executed while incorporating the encrypted information. Second, these workflows guarantee that while additional insights can be derived from further analysis, sensitive information remains undisclosed and is only accessible to users on a need-to-know, right-to-know basis.

Example 10: [Cipher](#): Selecting a Column to Encrypt

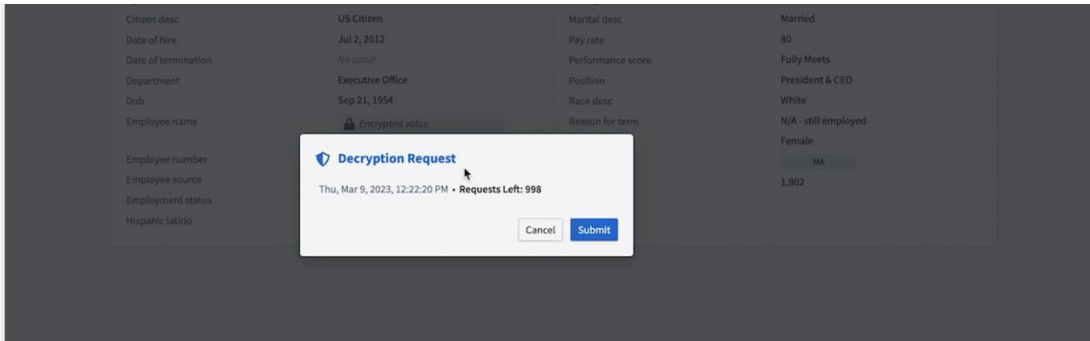


Example 11: [Cipher](#): Front-end interface of an encrypted column in object view

Properties			
Age	32	Manager name	Brandon R. LeBlanc
Citizen desc	US Citizen	Marital desc	Married
Date of hire	27 Oct 2008	Pay rate	28.5
Date of termination	No value	Performance score	Fully Meets
Department	Admin Offices	Position	Accountant I
Dob	24 Nov 1985	Race desc	Black or African American
Employee name	Encrypted value bBfPEqrXTFAUfVt+mSMzQglUGTez/3S4pWFiroF2	Reason for term	N/A - still employed
Employee number	1,103,024,434	Sex	Female
Employee source	Diversity Job Fair	State	France
Employment status	Active	Zip	1,450
Hispanic latino	No		

Data Minimization

Example 11 Continued: [Cipher](#): Front-end interface of an encrypted column in object view



Depending on their access control permissions, users can still interact and decrypt the encrypted values above. Those decryptions can then be given Checkpoints (see Transparency, Example 3 above), to ensure users consciously acknowledge they are accessing sensitive data. This also ensures decryptions requests can be logged for auditing by platform administrators down the line.

As demonstrated in these examples the Palantir platforms empower organizations to implement dynamic data minimization techniques beyond access controls, to accommodate the intricate and context-dependent requirements of clients across industries and jurisdictions. These methods allow institutions to align their data processing operations with data minimization principles, contributing to both purpose limitation and data security. To read more about anonymization, please refer to [our blogpost](#) and [whitepaper](#) dedicated exclusively to this topic.

4.5 Retention and Deletion

Data retention and deletion policies are critical and complex aspects of the data governance lifecycle. Retention (the time bounds set around how long data will live on a platform or be accessible) and deletion (the process of purging data after retention expires) pose significant technical challenges in distributed systems and cloud deployments. These are also mandated under the GDPR (General Data Protection Regulation) and other data privacy legislation such as the CCPA (California Consumer Privacy Act) as rights to deletion or the "right-to-be-forgotten." In complex information system terms, the exercise of these deletion rights often necessitates highly precise and expansive operations, including tracking down each instance of relevant information and carrying out targeted deletion of specific fields. In these cases, whether due to compliance requirements or other needs, organizations must be able to set firm retention policies, purge data with confidence, and be able to provide records of deletion.

Retention and Deletion

Palantir platforms incorporate policy engines for managing complex retention and deletion policies, allowing enterprises to confidently delete personal data at set or ad hoc intervals in compliance with legal requirements. However, aside from systemic deletion requests, organizations may sometimes be required to handle deletions at a more granular level. In those cases, Palantir platforms can confidently carry out reliable, granular deletion processes.

Step One: Understanding Data Lineage

Data deletion is often complicated due to factors such as data duplication, integration with additional data, storage in various formats to accommodate a range of use cases, among many other examples. Our platform features enable users to trace modifications and merges back to their origin, ensuring comprehensive examination of the data history both upstream and downstream. The structure that governs the flow of data through a system, starting from its source and extending to its various final forms, is commonly known as "data lineage" - as detailed in the section on Transparency.

Understanding the context of data integrations and transformations is an essential initial step before starting any kind of deletion processes. Palantir platforms ensure that organizations can effectively delete data not only from downstream datasets but also across its entire lineage.

Retention and Deletion Tooling: Meeting Regulatory Requirements

Moving into data retention strategy, by leveraging Palantir products such as [Data Lifetime](#), organizations can not only orchestrate retention policies but also access a dashboard that displays all operations performed on any given policy that users have created. The dashboard includes information such as new datasets added or removed from a policy, as well as policy configuration changes. The captured changes are not limited to a single individual but include actions taken by anyone with access to the application, creating a powerful oversight tool that enables checks and balances. Lastly, it is important to mention that deletion processes are highly sensitive actions. Access to such tools should be tailored to meet the specific administrative needs of any given organization, as described [here](#).

Retention and Deletion

Example 12: Deletion in [Data Lifetime](#): Creating and managing lineage-aware retention policies on datasets

Configure policies for automated, lineage-aware data deletion

Data Lifetime policies describe rules about when data should be deleted. Policies are lineage-aware, so when a policy is applied to a dataset, it will also apply to downstream datasets. All transactions in affected datasets will be marked for deletion according to the policy's rules.

All configured policies

NAME AND DESCRIPTION	TYPE
FDD Test Policy This is a test policy!	Fixed deletion date
Lorem ipsum asdf	Fixed deletion date

Policy operation history

Data Lifetime records policy operations to track when a policy configuration is edited, or a policy is directly applied to or removed from a dataset. Once an operation is created, it will be evaluated on impacted datasets and their datasets to update scheduled transaction deletion dates. Because a single operation can impact many datasets directly, and because this evaluation will also impact all descendants of these datasets, a single evaluation can take a while to complete.

- Latest 6 minutes ago by [redacted]
 - Updated policy configuration (Propagated to 2 datasets total)
- 13 minutes ago by [redacted]
 - Applied policy to 1 dataset (No downstream datasets impacted)
- 13 minutes ago by [redacted]
 - Applied policy to 1 dataset (No downstream datasets impacted)
- Policy created at Thu, May 2, 2024, 9:55:47 AM by [redacted]

Review transactions marked for deletion

Explore all transactions which have been marked for deletion by Data Lifetime. You can review both transactions which are scheduled for deletion in the future and transactions which have previously been deleted.

Filter by: Time range [Apr 25th, 2024 at 9:43 AM - May 31st, 2024 at 9:43 AM] Transaction ID

TIMESTAMP	DATASET	TRANSACTION	DELETION CAUSE	STATUS
4/30/2024, 8:34:04 AM	test copy 92 /PCL/ /Lots of test files	...e99b09810c26	FDD Test Policy	Deleted
5/2/2024, 9:40:07 AM	Notional Data_france_licenseplate /PCL/ Testing	...e955d777b47	FDD Test Policy	Pending deletion
5/31/2024, 12:00:00 AM	ssn /PCL/ Test Project/0_debug_post_scan_bug	...d310344c3362	Test_loremipsum	Scheduled

Retention and Deletion

Retention and deletion processes are not only limited to the dataset level - Palantir platforms include the ability to establish deletion protocols that can remove all resources across each layer of our products, starting with access controls, followed by the application, storage, backup, and cloud infrastructure layers.

With Palantir's deletion capabilities, organizations can effectively and readily delete data and seamlessly pull audit logs of these actions for compliance or other requirements.

Furthermore, because we understand that regulatory requirements evolve over time, we provide our clients with user-friendly tools that can be configured and adapted to cater to evolving use cases as well as to both non-technical and technical users. Our ultimate goal is to provide customizable tools that equip organizations with improved oversight and control over sensitive workflows. The following section will explore how Palantir empowers organizations with the necessary tools to enhance their accountability and oversight objectives.

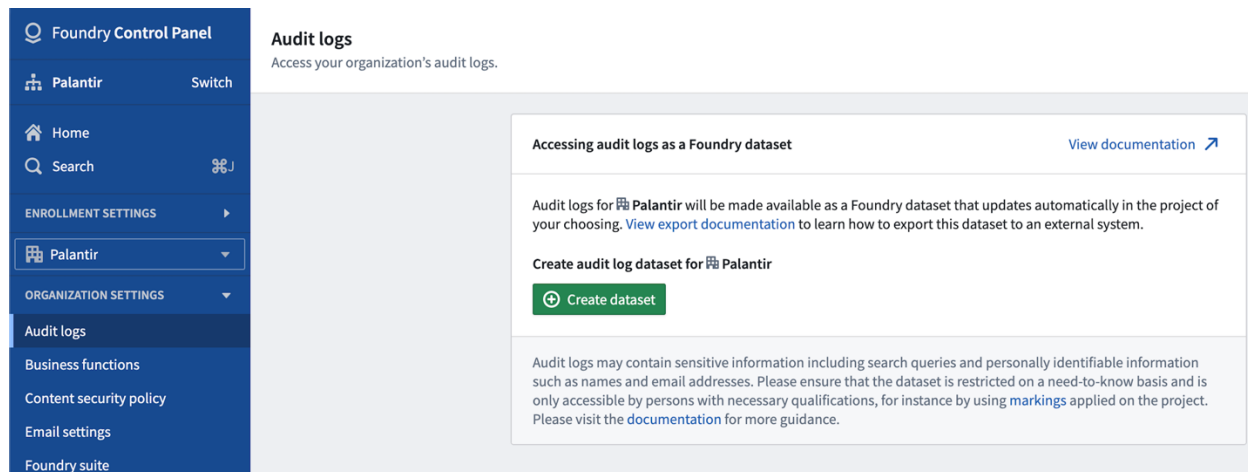
4.6 Accountability and Oversight

As regulatory requirements necessitate data controllers to actively demonstrate compliance with applicable provisions, data infrastructure must therefore support meaningful interrogation of all data processing activities. Palantir platforms, with their data provenance tracking, version control, audit logging, and policy enforcement capabilities, offer comprehensive records essential for auditing purposes.

Palantir platforms preserve an audit trail of all actions taken by users across its products, making it possible for data protection and other organization administrators to review per-user, per-resource data access events. Palantir software ensures that results can always be reproduced by preserving a record of the information and logic utilized to generate the output. The Palantir platforms further create and maintain complete records of all transformations that have been applied to datasets, with user attribution.

4.6 Accountability and Oversight

Example 13: Creating an [audit logs](#) report for an organization



Audit logs serve as the main method for supervisory or other designated governance or oversight authorities to investigate the actions performed by users across the platforms. In some cases, audit logs will contain contextual information about users including Personal Identifiable Information (PII), such as names and email addresses. It is also important to note that the workflows themselves may be considered highly sensitive in nature. As such, audit log contents should be considered sensitive and viewed only by persons with the necessary security qualifications. Audit logs are usually integrated into a dedicated system for security monitoring (a "security information and event management" or SIEM solution) operated by organizations. However, Palantir also enables data governance officers to review audit logs in-platform, based on strict access control measures.

Audit logs from all Palantir platforms are first written to disk and then archived to a stack-specific storage bucket within 24 hours of being written (AWS S3, Azure Blob Storage, or on-premises storage). Audit logs are engineered to be append-only throughout their lifecycle as to ensure their integrity for analysis. The platform's combined audit logging, analysis, and data provenance capabilities enable organizations to capture the necessary information for holding users accountable for their activity on the system and identifying behavior that may indicate data misuse or abuse. Investigators can then assess whether a law or policy violation occurred and take appropriate actions to hold users accountable according to their organizational policies. To find out more on how auditing in Palantir works, please refer to our [documentation](#).

5 Conclusion: Value in Governance Across Industries

Palantir builds platforms that incorporate cutting edge privacy, governance, and security principles as core features alongside our existing data integration and analysis tools. Far from needing a separate privacy-enhancing technology ([PET](#)) provider to meet regulatory or ethical standards with data use, we focus on building trust, transparency, and regulatory compliance into one unified platform. The value of these features is apparent across both public and commercial entities looking to leverage data for critical decisions. Our products, built on twenty years of privacy and security engineering, enable organizations to understand how and why they arrived at a specific outcome and adapt to the constantly evolving world of regulatory compliance. This product approach is a founding commitment of Palantir, and something we strive to maintain and build upon with the development and application of emerging technologies.