Data Governance: Data Usage Labeling and Enforcement in Adobe Cloud Platform

Contents

The Adobe Cloud Platform (ACP) is a one-of-a-kind tool that includes services for wrangling and analyzing data and building predictive models and next best action. ACP makes the data, content, and insights available to experience-delivery systems to act upon in real-time, yielding compelling experiences in the relevant moment. With ACP, enterprises will be able to utilize completely coordinated marketing and analytics solutions for driving meaningful customer interactions, leading to positive business results.

An integral part is data governance, and how it enables the Adobe Cloud Platform.

Data in today's world is subject to government regulations, contractual restrictions and internal policies that limit what use can be made of it, and those come with potential penalties or adverse consequences for your non-compliance. The business value that can be derived from your data is very much controlled by your ability to know what data you have and where it came from, to catalog and categorize it, and to manage the myriad regulatory, contractual, and policy limitations on its use. Only once those problems are solved can data to be used, not only to derive business insights, but ultimately to continually improve your customers' experiences.

The Adobe Cloud Platform was built with these considerations in mind. That is why it provides a robust, powerful **data governance framework** that assists you in managing your compliance with the regulations, restrictions, and policies controlling the use of your data. With Adobe Cloud Platform, you can better catalog and categorize your data and respond to data usage requirements. ACP's open, extensible architecture will enable you to incorporate ACP and its data governance into your own integrations, building on top of ACP.

What is data governance?

The Data Management Association International (DAMA) offers a concise definition of data governance: *The exercise of authority, control, and shared decision-making (planning, monitoring, and enforcement) over the management of data assets.*

DAMA places data governance at the core of data management, viewing it as the discipline that serves as the foundation for all other aspects of data management.

But what does that mean to you? It means having a system and processes in place consisting of the people and digital tools required to exercise that authority, control, and shared decision-making. It means creating a team to control data governance and giving that team the tools they need to govern data effectively and efficiently in real time. It encompasses both the **strategies** and the **technologies** required to achieve the goal of data usage that complies with the regulations, restrictions, and policies applicable to that use. Data governance includes data cataloging, data lineage, data usage labeling, data access policies, and much more.

Why data governance?

Every organization today deals with data: individual bits of information about its operations, its products and services, its finances, its employees, its suppliers, distributors and competitors, and most significantly, its **customers**. It's become possible to not only learn a great deal about people individually and in groups, but also to create profiles of individuals that contain many dimensions of personal information. These profiles can be used to categorize individuals by various characteristics into groups, document their preferences and interests, and predict their behavior.

The increasing awareness of the value of such information has also led the organizations who collect and share it to place contractual restrictions on how it can be used by those with whom it's shared, both to protect its business value and limit risk. These laws, regulations, and contractual restrictions are augmented by internal policies of each organization.

Data governance is essential to data-driven success. Data governance starts with the following:

- Understanding the regulations, contractual obligations, and company policies that apply to your data ("Data Usage Policies")

- Classification of data by labeling it with appropriate meta-data

- Defining Data Usage Policies based on various compliance, legal and corporate policies

From here, a data governance system can help to enforce those policies so that, for any data usage request, it's possible to say "this use is permitted, or this use is prohibited, **and here's why**."

That *why* is critically important. Data governance, to be effective, has to encompass all parts of the organization. That's why Adobe is incorporating data governance features into Adobe Cloud Platform. Proper data governance activates data for these purposes and creates the foundation for the Adobe Experience Cloud.

Data governance roles.

Data Steward

"I ensure data usage complies with our policies."

Data Scientist

"I want to cleanse the data and find appropriate insights."

Data Engineer

"I need to understand requirements & build data platforms to fit business needs."

Marketer

"I want to use the right data for the right marketing needs."

Data governance is neither automatic nor occurs in a vacuum — it is a journey. You need people to effect data governance. What began as a role for one individual, today recognized as a data steward, has grown along with the data governance ecosystem. Let's look at some of the most essential actors in the data governance cycle:

Data stewards are at the heart of data governance. This role is responsible for interpreting the regulations, contractual restrictions, and policies and applying them to the data itself. Informed by their understanding of the regulations, restrictions, and policies affecting data, the role of a data steward includes the following:

- Review data and datasets, data samples, and metadata/labeling.

- Create policies and apply them to data at the connection, dataset, and field level; communicate policies to the rest of the organization.

Marketers are at the end point of data governance, in that they request data from the data governance infrastructure created by data stewards, scientists, and engineers. This role encompasses a number of different specialties under the marketing umbrella including the following:

- Marketing researchers make requests of data to enable them to understand customers, both individually and in groups (segments).

- Marketing specialists and Experience designers use data to design new customer experiences.

The Adobe Cloud Platform advantage.
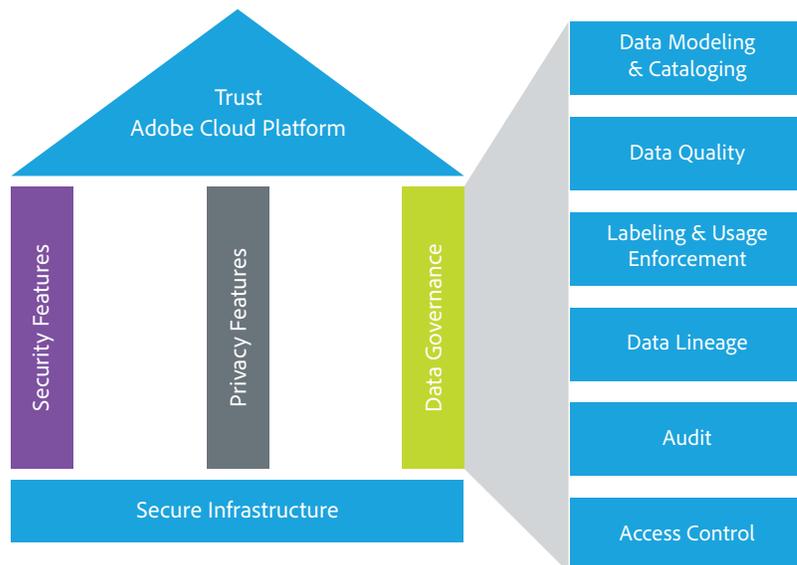
So what's Adobe's data governance advantage? End-to-end integration. To make data governance really work, you need an integrated solution that connects your data governance infrastructure with the tools you use to create and manage customer experiences. Adobe Cloud Platform is the only solution on the market that offers that kind of integration.

Data governance up to now has been implemented with third-party tools on top of database platforms, and sometimes with data governance features built into the platforms themselves. This does well, as far as it goes. But the breakdown in creating a practical, functional data governance practice that responds in real time to the needs of marketers and others who request use of data is that the governance model isn't reflected in the tools those professionals use. From analytics to campaign management to tools for building experiences and managing creative assets.

Data Catalog.

Adobe Cloud Platform is powered by a functionally rich and easily extensible Data Catalog. All the metadata about datasets including lineage, ownership, standard schemas, and subscriptions are maintained in the Catalog as objects. Customers and Partners can make use of catalog using the APIs available through Adobe IO.

Data Usage Labeling and Enforcement Framework.

The Adobe Cloud Platform introduces a new paradigm for aligning data to the usage restrictions associated with it: the Data Usage Labeling & Enforcement (DULE) Framework. The DULE Framework simplifies and streamlines the process of categorizing data and creating data usage policies so that marketing actions requiring data can immediately be evaluated based on the data usage policies applied to it.

Data Usage Labeling.

The core of Adobe Cloud Platform's system for data governance is DULE labeling. The DULE features enable data stewards to apply labels (metadata) to data, either as it's ingested or after, to categorize it according to what kind of Data Usage Policies apply to it.

ACP's DULE implementation includes predefined data usage labels that can be used to categorize data in four ways:

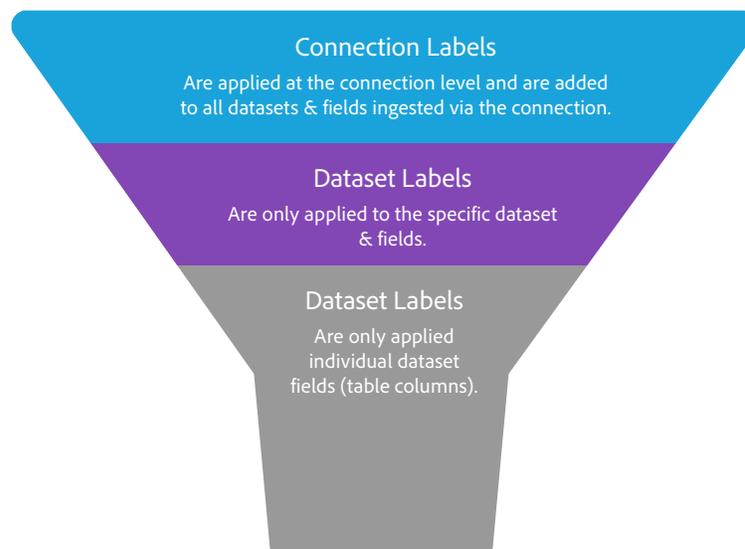| Contractual Data Labels	Identity Data Labels	Sensitive Data Labels	GDPR Data Labels
Used to label data that has contractual obligations or is related to a customer data governance policies.	Used to label and categorize data that can identify or contact a specific person.	Used to label and categorize sensitive data such as geographic data.	GDPR data labels are serviced in reference to data governance policies such as GDPR.

Data usage labels can be applied at three levels:

Connection Labels
Are applied at the connection level and are added to all datasets & fields ingested via the connection.

Dataset Labels
Are only applied to the specific dataset & fields.

Dataset Labels
Are only applied individual dataset fields (table columns).

Data usage labels are inherited. Labels applied at the connection level are propagated to all datasets and fields in the connection, while labels applied at the dataset level are propagated to all fields in the dataset.

Data Usage Policies.

Data often comes with restrictions on its use. These restrictions can be described by data usage policies, or rules created to provide guidance on allowing or prohibiting the requested use of the data. For example, given this dataset and the restrictions applied to it by data labels, can we use the data for an email targeting campaign? DULE labeling is the foundation for implementing Data Usage Policies. ACP's data governance framework offers predefined policies that cover most instances. In the future, data stewards will be able to create custom policies that can be applied to data labels as needed.

Marketing actions.

As mentioned earlier, marketing actions are tasks performed, on or with data, by marketers in Adobe Cloud Platform. In ACP, these are codified the same way the possible labels and policies are codified. Each defined marketing action represents a type of data use that might be requested. Typical marketing actions might be some of the following:
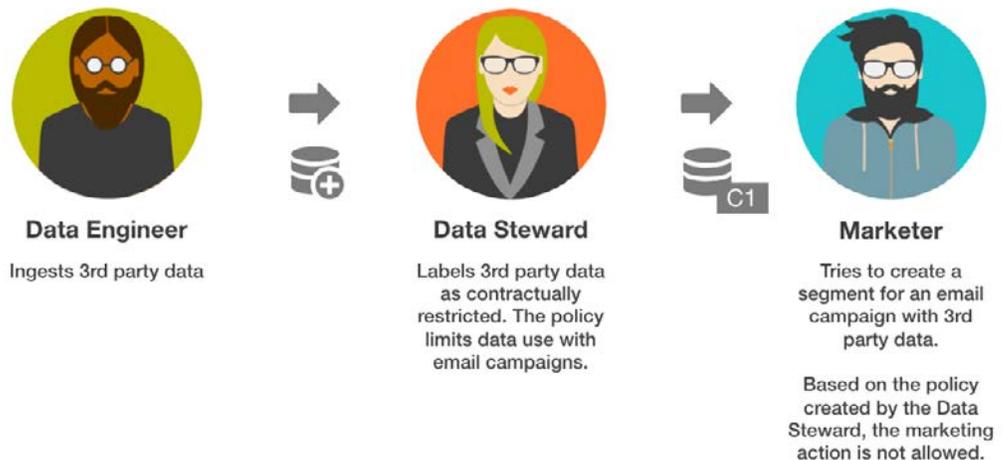
• Device-level analytics

• Targeted email campaign

• Customer segmentation

By codifying these actions as defined entities in the data governance framework, ACP makes them available for evaluation via the DULE Policy Service.

DULE Policy Service.

Adobe Cloud Platform's data usage governance is implemented as the DULE Policy Service, consumable by all solutions built on ACP. Once data has been catalogued and labeled, the DULE Policy Service will use those labels and the associated policies as input to its Policy Service to determine the permissibility of any marketing action.

Data governance in action.

Data Engineer

Ingests 3rd party data

Data Steward

Labels 3rd party data as contractually restricted. The policy limits data use with email campaigns.

Marketer

Tries to create a segment for an email campaign with 3rd party data.

Based on the policy created by the Data Steward, the marketing action is not allowed.

With Adobe Cloud Platform, data governance applies in real time, streamlining the process of determining the permissibility of data use.

Consider a hypothetical example: a digital marketing specialist working for a large retail company that operates a chain of stores as well as a shopping site. The marketing specialist wants to design a targeted email campaign for a special holiday promotional offer. For such a campaign, the marketing specialist reviews the ACP data relevant to targeting customers.

Because of the embedded workflows and in built data governance features, the marketer stays in compliance and the campaign proceeds without any lengthy data governance evaluation. This kind of efficiency is only possible when you can combine your data storage, data governance, and marketing and customer experience tools into a single, unified solution. That solution is Adobe Cloud Platform.

Conclusion.

With Adobe Cloud Platform, you have a solution to help you mange your data governance, embedded in the very same tool with which you create the analysis, marketing solutions, and customer experiences that make use of the data you need to manage. No other solution on the market provides this complete, end-to-end seamless integration of data storage, management, and governance applied at the very place you need it. Building your customer experience solutions on ACP gives you the data governance you need to derive business value and enable that continual, virtuous cycle of improving customer experience that makes your business an experience business.

References.

1. The Global Data Management Community, https://dama.org/.

2. "Adobe Cloud Platform: The Platform for the Experience Business," Adobe, https://www.adobe.io/apis/cloudplatform/overview.html.