PRIVACERA

Democratizing Data Analytics in Financial Services

Whitepaper
October 2021

There are two reasons financial institutions are learning as much about their customers as possible by gathering—and processing—as much data as they can. Firstly, this information is invaluable for delivering more personalized products and services to increase revenues while making smarter decisions and improving operational efficiency. Better customer understanding can improve things like underwriting accuracy, for example, and many others.

Secondly, financial organizations face enormous regulatory pressure requiring them to be familiar with customers. Mandates like Know Your Customer (KYC) and Anti-Money Laundering (AML) call for detailed knowledge of customers and their financial habits.

Additionally, there are several business-specific drivers for fulfilling these imperatives. The right analytics on timely, comprehensive data supports fraud detection use cases in addition to driving recommendation engines for cross-selling and upselling opportunities. These deployments are regularly enhanced by machine learning technologies that involve huge amounts of data.

Other business drivers for this industry include customer 360 views for targeted sales and marketing opportunities, as well as customer loyalty programs to increase retention rates. Similarly, micro-segmenting customers and their transactions can lead to analytics to reduce churn.

Each of these drivers requires a host of different users, departments, and systems to frequently access data with sensitive information. For example, loan application processing requires reviewing a consumer's credit history, W-2 information, and a plethora of other sensitive personally identifiable information (PII). Employees motivated by the above business drivers are demanding data for these and other use cases.

Unfortunately, many of them encounter restrictions resulting in long delays for data access because of security, regulatory compliance, and data privacy needs. The conundrum is clear: as your data volumes grow, so must your guardrails for it, too. But how do you automate these critical facets of data governance across hybrid and multi-cloud settings in an expanding data landscape?

Industry Transformation: Neo Banking

The tremendous desire for data, as well as its effectiveness in fulfilling the above business drivers and other functions in this industry, has greatly contributed to the transformative impact of fintech in financial services. Fintech companies are born in the cloud, typically have no physical presence (which reduces capital costs), and are extremely data savvy. They rely on data—and a number of public data sources in particular—for almost all aspects of their

operations. Their business model is rooted in data, enabling them to automate several processes that typically required manual approaches. They'll utilize deep neural networks to micro-segment customers and identify specific attributes to make marketing campaigns in micro-lending, for example, deliver demonstrable results. Moreover, fintech players are known for extremely rapid cycle times, which positively impact customer service and customer satisfaction. They can issue loans and insurance policies in a fraction of the time traditional financial services companies can, doing in minutes or days what the latter does in weeks. Their reliance on online, public data sources prioritizes secure data sharing between organizations.

Not surprisingly, fintech organizations are transforming the financial services industry as a whole, spurring even traditional entities to keep pace with their innovations, swift processing times, and high customer satisfaction rates. This fact has redoubled the demands for data by conventional financiers seeking to match the agility of fintech upstarts. Therefore, traditional banks and insurers are migrating greater numbers of applications to the cloud while competing with fintech companies for engineering and data talent. Because fintech unicorns have a broader appeal among such employees, traditional entities must often resort to simplification and automation to overcome this talent shortage.

There are a couple immediate consequences of the increasing reliance on the cloud of traditional banks and their fintech counterparts. It's crucial these firms ensure regulatory compliance and proper centralized auditing and reporting for user access in hybrid and multi-cloud settings. This concern is always important, but becomes especially so when the data is no longer comfortably residing behind an organization's firewalls. When it was, firms had the comfort of familiar boundaries for their valued data. Thus, when transitioning to the public cloud, it's necessary to ensure there are data access boundaries there, too.

Key to Success:
Secure Access to Mass Data Quantities

Almost all the contemporary banking trends need tremendous amounts of data to succeed. The biggest challenge for this industry is granting secure access to large quantities of sensitive data with PII to satisfy the demands of business users and data scientists, while also adhering to security and regulatory requirements. Without enough data, it's almost impossible for data scientists to build accurate machine learning models to solve business problems requiring recommendations or fraud detection. Many organizations respond to this challenge by deliberately limiting data access to end users due to a lack of proper data governance compounded by inadequacies in related areas like data discovery, cataloging, and tagging sensitive data. Without a centralized governance platform, business units are severely limited in their access to the data they need to meet the business drivers for financial services. This limitation undermines the value of any investments in modern cloud and data technologies.

Organizations also contend with numerous operational inefficiencies because of a lack of access to enough data. For example, security and compliance personnel may choose to restrict a data scientist's access to only half of the available AWS S3 buckets because of the presence of sensitive data and a lack of centralized access control. Oftentimes, compliance and security employees are unaware of which specific datasets—or repositories—are used for particular business use cases. Because data scientists can't get access to all the data they need, their ensuing machine learning models produce suboptimal results, incur bias, and are built with insufficient features to identify outliers for determining money laundering activity, for example. The effects of such inefficacy stemming from a lack of access to enough data will eventually trickle down across use cases and departments, diminishing the performance of individual employees and the ROI of investing in data-driven initiatives.

The Privacera Approach:
Data Governance, Access Control, and Data Security in One

The solution to granting secure access to the universe of available data to fulfill financial service business drivers while upholding guardrails for regulatory compliance and sensitive data is to employ a unified data access governance/data privacy platform tailored for hybrid and multi-cloud deployments. This modern offering inherently supports the distributed data access governance model in which data stewards familiar with use cases in respective business domains grant end user access by enforcing centralized policies at the local level in individual sources. To further drive latency from the self-service analytics process, Privacera introduced a new data sharing approach called Governed Data Sharing that delivers a new level of flexibility and power to accelerate analytical initiatives by grouping functional data–such as sales, marketing, finance, and more–into Data Domains. Access policies are automatically applied to data sets inside Data Domains where data consumers, such as data scientists or business analysts, can browse through an inventory of data sets and request access to them. The use of Data Domains alleviates the operational burden of IT by putting data domain owners in direct contact with data consumers to manage access requests, greatly improving collaboration, flexibility, and responsiveness. This way, uniform policy enforcement occurs regardless of where users or data are, as delegated by stewards or data domain owners who know the data best.

01

IT organizers disparate data sets into logical *Data Domains* with underlying access policies.

02

Data owners can publish their data sets as part of their *Data Domains* to share with others.

03

Data consumers can easily find and subscribe to *Data Domains*.

04

Data can be **shared internally and externally** with authorized personnel for sanctioned projects.

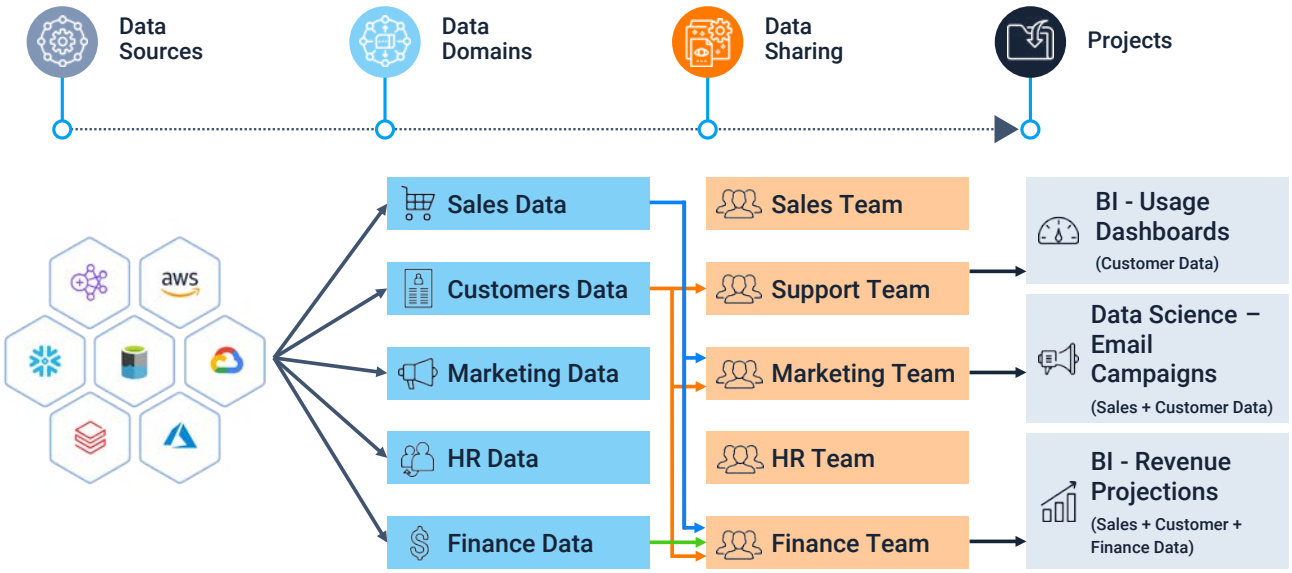

Exhibit: Framework for Governed Data Sharing

Privacera's access governance technology provides several controls that meet the dual data access mandate prevalent in finance. The first is to automate the data discovery process to identify sensitive elements in datasets. This step becomes the basis for obfuscating this information to increase user access to the data necessary to meet this industry's business drivers. Next are obfuscation mechanisms such as masking, encryption, and tokenization to ensure the right users access sensitive data without viewing any PII-related attributes. To return to the above-mentioned data science use case, masking sensitive data elements at the columnar level, for example, suffices to ensure these professionals can access all the S3 buckets they request for constructing a machine learning model. This method directly addresses the data quantity issue via secure access that complies with data privacy mandates. It produces desired effects like increasing the accuracy of machine learning models by enabling data scientists' access to the full array of features in organization's data.

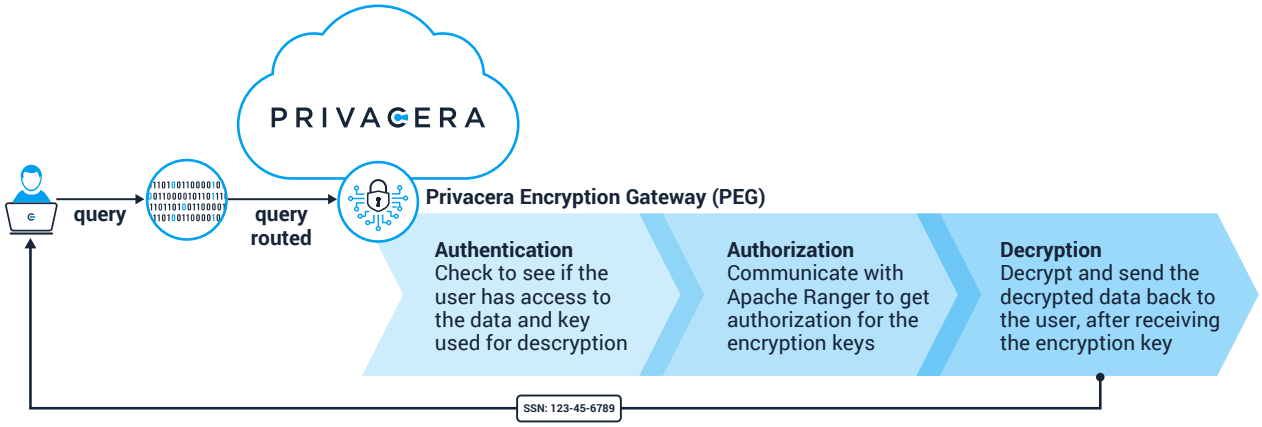

Exhibit: Decryption process for Privacera Encryption Gateway (PEG)

This solution also has centralized auditability so IT, stewards, and governance personnel can see in real-time who's doing what with which data. There's a central pane of glass for viewing all data sources, regardless of their location or cloud environment. Such visibility is critical for data lineage and reporting capabilities for regulators or enterprise users. This functionality is critical in financial services because of the surfeit of regulations—and audits—organizations undergo in this space. It's also vital for the contemporary focus on the cloud that fintech companies and traditional financiers now have because it gives them the guardrails necessary to safely expand to this environment to grant access to data to successfully compete in this vertical.

Case Study:
Global Banking Services Company Secures Collaborative FinTech Platform to Fulfill Rigorous Financial, Privacy, and Legal Regulations

An international financial services firm developed a platform for collaboration between fintechs and traditional players. Exchanging inter-party information requires rapidly sharing, aggregating, and anonymizing sensitive data between organizations to gain the trust of users and adhere to the numerous financial regulations and legal requirements they must fulfill. The greatest challenge was that the company had to enforce—and demonstrate—compliance with other companies' environments across hybrid and multi-cloud. Conventional governance approaches were too slow and couldn't provide data security at scale, impeding organizational agility and info security measures.

Privacera's unified data access governance technology rectified these problems with unparalleled visibility, auditability, data lineage, and reporting of who was accessing what data when—across source systems—to meet internal policies, comply with regulations, and satisfactorily demonstrate so to regulators. Its obfuscation approaches implemented policies

for controlled user access, building inter-organizational trust for customers of the international financier's data exchange platform while empowering them with automated data discovery, tagging, and obfuscation of sensitive data.

Moving Forward

The financial services industry can make good on the numerous business drivers requiring access to ever greater amounts of data for customer 360s, recommendation engines, micro-segmenting customers and more, while still maintaining adherence to a growing number of regulations. Privacera's centralized data access governance platform ensures these benefits while mitigating regulatory and data privacy risk by providing a single place to author policies, monitor data access, implement access controls, and maintain full auditability. It's a modern approach for meeting the demands of this rapidly changing industry today and tomorrow.

About Privacera

At the intersection of governance, privacy, and security, Privacera's unified data access governance platform maximizes the value of data by providing secure data access control and governance across hybrid- and multi-cloud environments. The hybrid platform centralizes access and natively enforces policies across multiple cloud services—AWS, Azure, Google Cloud, Databricks, Snowflake, Starburst and more—to democratize trusted data enterprise-wide without compromising compliance with regulations such as GDPR, CCPA, LGPD, or HIPAA. Trusted by Fortune 500 customers across finance, insurance, retail, healthcare, media, public and the federal sector, Privacera is the industry's leading data access governance platform that delivers unmatched scalability, elasticity, and performance.

Headquartered in Fremont, California, Privacera was founded in 2016 to manage cloud data privacy and security by the creators of Apache Ranger™ and Apache Atlas™.

Visit www.privacera.com or follow @Privacera on LinkedIn and Twitter.