Google Cloud Whitepaper
January 2023

Protecting healthcare data on Google Cloud

Google Cloud

Introduction

As more healthcare and life sciences organizations adopt cloud-based resources, huge amounts of data are being stored, processed, and analyzed across platforms. At Google Cloud, the privacy and security of customer data are primary design criteria that underpin all the services that we offer. Given the sensitive nature of individually-identifiable health information and other types of personally identifiable information (PII), the protection of healthcare data and the systems it resides on is of critical importance. While we help with support for regulations like the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the General Data Protection Regulation (GDPR), we also understand that at the core, customers want assurances around how we use, process, and handle customer data.

Google Cloud offers built-in data protection at scale, by default, designed to protect your organization from intrusions, theft, and attacks. Customer data stored in Google Cloud is encrypted at rest by default and, depending on the connection, Google applies default protections to customer data in transit. Google Cloud products like Cloud Key Management Service (KMS) allows customers to manage cryptographic keys for their cloud services, while other products like Access Transparency enable you to review logs of actions taken by Google staff when accessing customer data. We have also created the Google Cloud Healthcare Data Protection Toolkit; a series of scripts and procedures that walk you through the process of setting up the required controls needed for various data privacy obligations.

This whitepaper lays out the various questions healthcare and life sciences organizations have asked us about when determining whether to move their data onto Google Cloud. Our replies provide details about how we protect health information throughout its lifecycle as well as how we provide customers with transparency and control over their data in Google Cloud.

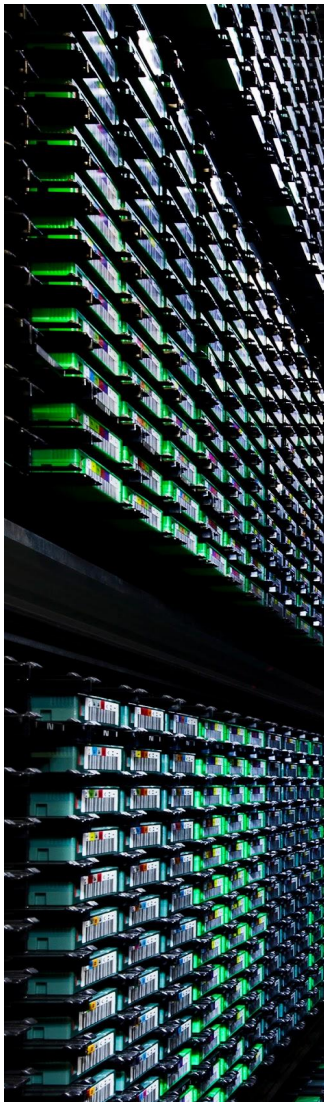
Although we reference the term “health information and data” throughout this whitepaper, it is important to note that we treat all customer data with the highest level of security and privacy, regardless of the type. If you’d like to learn more about how we define customer data, please refer to our Cloud Terms of Service.

Disclaimer

This whitepaper applies to Google Cloud products described at cloud.google.com. The content contained herein is correct as of January 2023 and represents the status quo as of the time it was written. Google’s security policies and systems may change going forward, as we continually improve protection for our customers.

I'm considering moving healthcare and life sciences data to Google Cloud, where should I start?

We encourage healthcare and life sciences organizations to read the commitments we've made to protecting the privacy of your data and your customers' data, which are detailed in our Google Cloud Trust Principles. We understand that the privacy of your data, as well as your customers' health information and data is of paramount importance to you. In addition to committing to our Google Cloud Trust Principles, we build privacy into our products from the earliest stages, and continually evolve our practices. We articulate those commitments in our Cloud Data Processing Addendum (CDPA); and undergo regular independent, third-party audits to verify these protections. In addition to the legal commitments outlined in the CDPA, customers who are subject to HIPAA and want to use Google Cloud for their business purposes involving personal health information (PHI), must enter into a business associate agreement (BAA) with Google that covers specific Google Cloud products and services. More information on HIPAA is discussed later on in this paper.



It is also important to understand the necessary controls that are required in order to align your workloads with data-privacy considerations (e.g. certification requirements, standard frameworks, etc.) and obligations. To help, we have created the Google Cloud Healthcare Data Protection Toolkit, a series of scripts and guidance documents that walk you through the process of setting up the relevant controls specific to the protection of healthcare information. You are not required to use this toolkit in order to meet your data security and privacy needs as a healthcare customer on Google Cloud; however we believe the toolkit can help automate control implementation and help accelerate your migration to the cloud.

To learn more, refer to the Google Cloud Privacy page and read about our dedicated privacy team.

How does the security, privacy, and compliance of Google Cloud compare to my existing on-premise environment?

Google runs on the same infrastructure that we make available to our customers. As a result, your organization can directly benefit from the security investments and protections that we use to protect our own services.

We focus on security, and protection of data is among our primary design criteria. Security drives our organizational structure, training priorities and hiring processes. It shapes our data centers and the technology they house. It's central to our everyday operations and disaster planning, including how we address threats. It's prioritized in the way we handle customer data. And most importantly, it's the cornerstone of our

account controls, our compliance audits and the certifications backed by independent third-party assessors we offer to our customers.

While healthcare and life sciences organizations have varying levels of IT expertise and resources, Google has highly dedicated and experienced security and privacy teams who are part of our software engineering and operations division. These teams provide project-specific consulting services to Google's product and engineering teams, monitor for suspicious activity on Google's networks, address information security threats, perform routine security evaluations and audits, and engage outside experts to conduct regular security assessments.

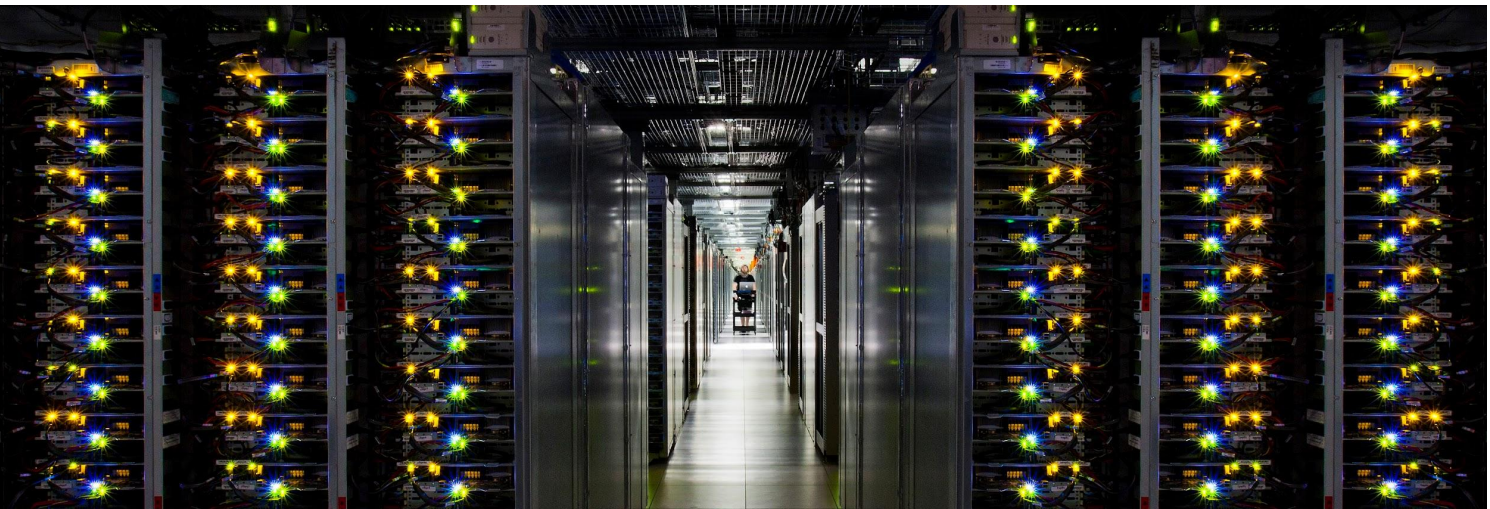
Our security and privacy practices are backed by several independent third-party audits that are performed on a regular basis to provide you with the assurance that the controls we've put in place actually protect your data. Some of those key standards are listed later on in the paper.

How are security and privacy responsibilities shared between Google Cloud and healthcare and life sciences organizations?

Google Cloud and its customers share responsibilities for managing the IT environment, including those related to security and compliance. In a traditional on-premise IT environment, all of these responsibilities would have to be managed by the user. Therefore, shared responsibility enables our customers to allocate resources more effectively by reducing the amount of effort needed to provision and support their IT environment. Shared responsibility does not remove the accountability and risk from customers using our services, but it does help relieve some of the burden as we manage and control system components and physical control of facilities.

Our role in the Shared Responsibility Model includes providing services on a highly secure and controlled platform and offering a wide array of features that customers can tailor to their needs. While Google Cloud provides secure and compliant offerings, the customer must ensure that the environment and applications that it builds on top of Google Cloud are properly configured and secured according to their design requirements and compliance needs.

For more information on our security features, please refer to the Trust and security page or view our Google security whitepaper.



Will my healthcare data be used by Google (or anyone else) if I choose to store it on Google Cloud?

Customers own their data, not Google; therefore, you decide how your data will be used on Google Cloud. At Google Cloud, we commit to never using your data for any purpose other than those necessary to fulfill our contractual and legal obligations. In addition, we've designed our systems to limit the number of employees that have access to customer data and to actively monitor the activities of those employees. Access to internal support tools is controlled via Access Control Lists (ACLs) and access authorization is enforced at all relevant layers of the system. To learn more, refer to Google's data access restrictions. We take our Google Cloud Trust Principles very seriously and commit to them in our Cloud Data Processing Addendum.

As part of Google's long-term commitment to transparency and user trust, we provide Access Transparency, a feature that enables customers to review logs of actions taken by Google staff when accessing customer data. For products integrated with Access Transparency, customers have the ability to view logs that capture when, how, and why our administrators access customer data (for example, viewing a label on a Google Cloud Compute Engine instance during a support call) .[1] Learn more about Access Transparency for Google Cloud.

Additionally, Google will retain, return, destroy, or delete customer data in accordance with the contract or service level agreements. To learn what happens when customer data is deleted in Google Cloud and how long it takes to complete Google's data deletion process, refer to the Data deletion on Google Cloud whitepaper.

How can I control who has access to my healthcare data?

Customers using Google Cloud can configure Cloud IAM permissions to control access to their cloud resources and limit access by their own administrators, curating the right amount of access at the project, folder or dataset level. This includes an extensive list of permissions and the predefined roles that grant them. You can also create your own custom roles that contain exactly the permissions you specify.

To help mitigate risks such as the misconfiguration of employee access controls or attackers taking advantage of compromised accounts, VPC Service Controls enables customers to define a security perimeter around Google Cloud resources, such as Cloud Storage, BigQuery, or Cloud Bigtable to prevent data exfiltration. Identity-Aware Proxy (IAP) enables customers to control access to cloud applications and VMs based on the user's identity and the context of their request. Customers can also control who has access to your Cloud Storage buckets and objects as well as the level of access. For more information, refer to the documentation page for Cloud Storage.

[1] There are some exceptions which are detailed in the Access Transparency documentation.

What about Google access to my healthcare data?

As stated above, Google Cloud is explicit in its commitment to customers: you own your data, and we will never use it for any purpose other than those necessary to fulfill our contractual/legal obligations. Google's internal data access processes and policies are designed to prevent unauthorized persons and/or systems from gaining access to systems used to process personal data. Google employs a centralized access management system to control personnel access to production servers, and only provides access to a limited number of authorized personnel.

Read the Google security whitepaper for more information.

How does Google safeguard my health data from unauthorized access?

The customer contract describes and governs Google's access to customer data. There are three ways customer data may be accessed in Google Cloud: 1) **Direct customer access**, 2) **Google support or administrator access**, and 3) **Google Cloud service access**.

Google has three types of controls in place to ensure that each of these access pathways function as intended:

Customer authorization: When services access data on behalf of a customer, they perform authorization checks to ensure the customer has appropriate permissions before proceeding.

Google administrator authorization: For services integrated with Access Transparency, Google uses a tool to validate that the business justification presented for access is valid, and log the justification to Access Transparency Logs. For some customers, even greater control is required. In this case, it may make sense to manage access by Google personnel on your own, and require your explicit approval every time your customer data is accessed. For this reason, we developed Access Approval[2] - a product that allows you to require your explicit approval whenever Google support and engineering need to access your customer data. It is important to note that there are a number of actions by Google that will not trigger an access approval request. Among these actions include access to lower-level storage systems designed to create access transparency logs. A full list of the exclusions can be found on the Overview of Access Approval page.

[2] Using Access Approval functionality will mean that Google may not be able to meet the SLAs for your chosen products, as any support response times may be increased. As such the SLAs do not apply to any service disruption to the extent impacted by the customer's use of Access Approval. We do not recommend that customers enable Access Approval for projects where you may require high service availability and rapid response by Google Support.

Service authorization: When the service accesses your data, Google uses technologies like Binary Authorization to validate the provenance and integrity of the software. Google Cloud service access is used in specific scenarios and must be authenticated and authorized to access data. To learn more about Google Cloud service accounts, refer to the Understanding service accounts documentation page. To learn more about how Google verifies code provenance and implements code identity, refer to the Binary Authorization for Borg whitepaper.

To prevent unauthorized access by other tenants sharing the same physical infrastructure, we logically isolate our customers' data. We have a variety of isolation and sandboxing techniques for protecting a service from other services running on the same machine. To read more, refer to the Google security whitepaper.

Can I choose where my health data is stored and where users can access it from?

Yes. Google Cloud provides you with the ability to control where your data is stored. Google Cloud provides services in locations across North America, South America, Europe, Asia, and Australia. To learn more about our locations, see our Geography and regions page, and for more information on the specific Google Cloud resources available within each location option see Cloud locations. Enterprises with data residency requirements can set up a Resource Locations policy that constrains the location of new resources for their whole organization or individual projects. For more details on Google Cloud's data location commitments, please read our Service Specific Terms.

What security measures does Google implement to protect my health data?

As more and more health information becomes digitized in an effort to provide better, more individualized care for patients, the need to protect the security, confidentiality, and integrity of such highly sensitive data has never been more important. Using the principles of "defense in depth," we've created an IT infrastructure that is more secure and easier to manage than more traditional technologies. We custom-designed our servers, proprietary operating system, and geographically distributed data centers. To read more about our "defense in depth" practices, refer to our security whitepaper.

In order to protect your health data, Google encrypts data at rest and encrypts data in transit, by default. The type of encryption used depends on the OSI layer, the type of service, and the physical infrastructure component. By default, we encrypt and authenticate all data in transit at one or more network layers when data moves outside physical boundaries not controlled by or on behalf of Google.

We offer several options for encryption key management. A fully managed encryption key service is provided by default that manages server-side encryption keys for customers; no setup or configuration is required. We also provide options for customers to supply their own keys and to fully manage their own encryption keys. Cloud Key Management System (KMS) is the service we provide for customers who choose to fully manage their own encryption keys.

Customers with stringent requirements for key storage can use Cloud Hardware Security Modules (HSMs), which allow customers to host encryption keys and perform cryptographic operations in FIPS 140-2 Level 3 certified HSMs.

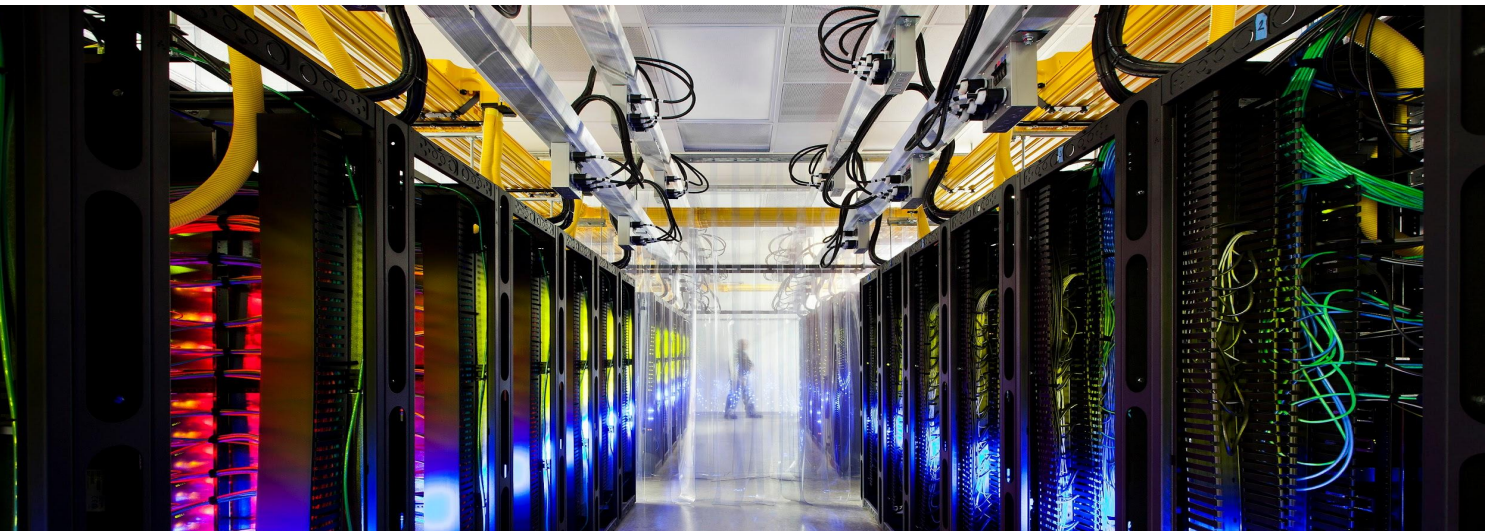
Additionally, we offer cutting-edge security tools to help customers manage their environments. For example, the Security Command Center for Google Cloud brings actionable insights to security teams, and VPC Service Controls help to establish virtual security perimeters for sensitive data.

To learn more, refer to our Security Products and capabilities page.

What other products and features can my organization leverage to protect the data processed on Google Cloud?

We understand that healthcare data often includes sensitive, personally-identifiable data. Even more, advancements in genomic data research can often reveal particularly sensitive insights into information such as if a set of patients is more likely to be predisposed to certain diseases.

In some situations, de-identification can be an effective way for researchers, data scientists, and healthcare and life sciences organizations to protect patient privacy. We have developed several tutorials for performing data de-identification on Google Cloud using the Cloud Healthcare API and Cloud Data Loss Prevention, including for medical images (including a deeper step-by-step example), clinical data, and other types of data within the organization.



How can the Google Cloud Healthcare Data Protection Toolkit help my organization confidently and quickly set up a secure workload on Google Cloud?

The Cloud Healthcare and Life Sciences team has created the Google Cloud Healthcare Data Protection Toolkit; a series of scripts and procedures that provide an end-to-end framework for deploying your entire organizational structure including central devops, auditing, monitoring, and data hosting projects.

Healthcare and life sciences organizations can use the Data Protection Toolkit to build HIPAA, GxP, GDPR, and other relevant compliance requirements and guidelines into templates to ensure that resources are deployed appropriately from the beginning of a project. Once these templates are created, they can be duplicated and reproduced across various deployments pertaining to data hosting and analysis, GxP, research & development, medical devices, genomics, and other healthcare scenarios.

Which parts of Google Cloud's Enterprise Compliance program are relevant to healthcare and life sciences organizations?

It is important for healthcare and life sciences organizations around the world to comply with regulations pertaining to the protection of patient, health, and other medical data and information. Google Cloud undergoes several independent third-party audits on a regular basis to provide this assurance to you. Some of the compliance requirements supported include:

HIPAA: Customers who are subject to HIPAA and want to use Google Cloud to process Protected Health Information (as defined in HIPAA) must enter into a BAA with Google that covers specific Google Cloud products and services. The BAA is an important contractual agreement that, among other things, prohibits business associates from using or further disclosing personal health information (PHI) other than as allowed by the written agreement or as mandated by law. The BAA also requires the business associate to implement appropriate safeguards to prevent improper use or disclosure PHI, mandating notification if such improper use occurs. To learn more about the BAA and how you can use Google Cloud in a HIPAA-aligned manner, you can refer to the Google Cloud HIPAA overview guide for more information. To help customers use our offerings in line with HIPAA, we recommend that they follow these best practices.

HITRUST CSF: Developed by the not-for-profit HITRUST, this framework contains a set of prescriptive security controls that relate to the organizational processes and technical controls for processing, storing, and transmitting sensitive data. Google Cloud is certified under HITRUST CSF.

UK NHS: The United Kingdom's National Health Service Department of Health (NHS DH) developed the Information Governance Toolkit (IG Toolkit) to provide best practices and guidelines around data governance for organizations that have access to, process, and store NHS healthcare data. We've published a Google Cloud whitepaper that discusses the compliance landscape for UK health data and, for organizations accessing patient data in England, an overview of NHS and the IG Toolkit.

FedRAMP: While not specific to the healthcare industry, FedRAMP requirements provide a comprehensive control check around information security. Google Cloud has recently achieved both FedRAMP Moderate and FedRAMP High ATO for a number of our products. Google's FedRAMP status is posted on the government's website: FedRAMP Marketplace.

For a complete listing of our compliance offerings, please refer to our Compliance resource center.

Conclusion

As healthcare and life sciences organizations increasingly adopt cloud services and reap the many benefits of this transformational change, you will undoubtedly have further questions. We encourage and welcome the dialog with customers and look forward to continuously working to ensure our customers are well positioned to not only meet their regulatory and compliance demands, but also the confidentiality, integrity, and privacy commitments inherent in each cloud deployment.

To learn more about our products, or to contact us, please visit cloud.google.com.