Google Cloud

Google Cloud Whitepaper
August 2023

Generative AI, Privacy, and Google Cloud

Table of Contents

Disclaimer

This whitepaper applies to Google Cloud products described at cloud.google.com. The content contained herein is correct as of August 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.

Introduction

Across Google Cloud and Google Workspace, we've long shared robust privacy commitments that outline how we protect user data and prioritize privacy. Generative AI doesn't change these commitments — it actually reaffirms their importance.

We are committed to preserving our customers' privacy with our Cloud AI offerings and to supporting their compliance journey. Google Cloud has a long-standing commitment to GDPR compliance, and AI is no different in how we incorporate privacy-by-design and default from the beginning. We engage regularly with customers, regulators, policymakers, and other stakeholders as we evolve our offering to get their feedback for Google Cloud (formerly known as Google Cloud Platform), Google Workspace, or Google Workspace for Education (together, the "Cloud Services") AI offerings which process personal data.

As part of our privacy reviews, we have a special focus on the intersection of Cloud AI and Privacy when developing our offerings. When we bring them to market, Google Cloud AI's approach includes incorporating privacy design principles, designing architectures with privacy safeguards, and providing appropriate transparency and control over the use of data.

AI/ML Privacy Commitments for Google Cloud

Google Cloud customers benefit from:

- **Your data is your data.** The data or content generated by a Generative AI Service prompted by Customer Data ("Generated Output") is considered Customer Data[1], that Google only process according to customer's instructions[2].
- **Your privacy is protected.** We have always maintained that you control your data and we process it according to the agreement(s) we have with you. Furthermore, we will not and cannot look at it without a legitimate need to support your use of the service – and even then it is only with your permission.
- **Your data does not train our models.** We don't use data that you provide us to train our own models without your permission. And if you want to work together to develop a solution using any of our AI/ML products, by default our teams will work only with data that you have provided and that has identifying information removed. We work with your raw data only with your consent and where the model development process requires it.

At Google Cloud, we are committed to giving you increased control and visibility over your data. Transparency creates trust, and trust is necessary for any business to succeed in this arena.

[1] See the Generative AI Service terms as part of the Google Cloud Service Specific Terms
[2] See Cloud Data Processing Addendum

That's why we led the way in providing meaningful transparency into provider access to customer data and now we're extending that transparency to our AI and ML work. Helping you address global privacy and data protection requirements enables you to apply machine learning to accelerate your business with confidence.

- **You own and control your data and your data stays within your organization.** Whether it is in our Vertex AI Platform or Generative AI App Builder (Gen App Builder), we recognize that customers want their data to be private and not be shared with the broader Google or Large Language Model training corpus. Customers maintain control over where their data is stored and how or if it is used, letting them safely pursue data-rich use cases while complying with various regulations. Google does not store, read, or use customer data outside your cloud tenant.
- **Enterprise assurances for Privacy and Security. Your fine-tuned data is your data.** We are able to provide Cloud AI offerings such as Vertex AI and foundational models with enterprise-grade safety, security, and privacy baked in from the beginning. In our Gen AI implementation for enterprise customers, the data of the organization remains in their own instance, whereas our LLM is "frozen". The learning and finetuning of the model with customer's data is stored in the adaptive layer in the customer's instance.

Our privacy commitments to all Google Workspace users

We want to be completely clear that generative AI does not change our foundational privacy protections for giving users choice and control over their data. To that end, here are key facts about how Workspace data is handled:

- **Your data is your data.** The content that you put into Google Workspace services (emails, documents, etc.) is yours. We never sell your data, and you can delete your content or export it.
- **Your data stays in Workspace.** We do not use your Workspace data to train or improve the underlying generative AI and large language models that power Bard, Search, and other systems outside of Workspace without permission.
- **Your privacy is protected.** Interactions with intelligent Workspace features, such as accepting or rejecting spelling suggestions, or reporting spam, are anonymized and/or aggregated and may be used to improve or develop helpful Workspace features like spam protection, spell check, and autocomplete. This extends to new features we are currently developing like improved prompt suggestions that help Workspace users get the best results from Duet AI features. These features are developed with strict privacy protections that keep users in control. (See below for more detail on additional privacy, security, and compliance commitments we make for business customers).
- **Your content is not used for ads targeting.** As a reminder, Google does not collect, scan, or use your content in Google Workspace services for advertising purposes.

Security and compliance commitments for business, education, and public-sector customers

When Google Workspace commercial customers adopt Duet AI for Google Workspace Enterprise they get the same robust data protection and security standards that come with all Google Workspace services, with specific protections for businesses, education, and public-sector customers:

- **Your interactions with Duet AI stay within your organization.** Duet AI stores any prompts or generated content alongside your Workspace content and does not share them outside your organization.
- **Your existing Google Workspace protections are automatically applied.** Duet AI brings the same enterprise-grade security as the rest of Google Workspace, automatically applying your organization's existing controls and data handling practices, such as data-regions policies and Data Loss Prevention.
- **Your content is not used for any other customers**. None of your content is used for model training outside of your domain without permission.

Workspace's robust security and privacy commitments for customers can be found here, and business or public-sector organizations that want to learn more about Duet AI for Google Workspace Enterprise can find more here.

Further information

These are the privacy commitments that Google Cloud and Workspace extends to all of its users. But they aren't just words. To ensure we continually meet these high standards, independent auditors validate our practices against international standards and best practices. Our products have achieved a number of security and privacy certifications from independent auditors who assessed our services' compliance practices in those domains. We apply those practices also to our Generative AI offerings.

For further information about how enterprise customers can benefit from comprehensive privacy offerings of Google Cloud, please visit cloud.google.com/privacy.