

# Better Data Governance for Responsible AI

**Shuba Lall**

AI/ML Data Governance Lead  
Google Cloud

# The 4 key enterprise readiness questions customers are asking



## My data is my differentiator

... but how can I **protect my IP** while customizing foundation models?

---



## I want to be an AI-everywhere organization

... but how do I deploy AI to every employee and in every location, while keeping everything **secure and compliant**?

---



## I want to scale up to meet the moment

... but how can I scale up with GenAI-specific infrastructure **cost-effectively, resiliently, and sustainably**?

---



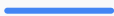
## I need to be effective, responsible, and safe

... but how do I make GenAI work for every end user while **mitigating harms**?

---



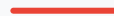
**Data  
governance  
and privacy**



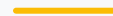
**Security and  
compliance  
support**



**Reliability and  
sustainability**



**Safety and  
responsibility**



# Enterprise readiness is at the core of Google Cloud's approach to AI



**Data  
governance  
and privacy**

---



**Security and  
compliance  
support**

---



**Reliability and  
sustainability**

---



**Safety and  
responsibility**

---

# You own and control your data, not Google

We ensure responsible use of data for Cloud AI/ML with these commitments to customers:



By default, we do not use customer data to train our models, in accordance with [GCP Terms](#) and [Cloud Data Processing Addendum](#).  
*Your customer data is only processed according to your instructions.*



You benefit from our privacy experience and commitments to transparency, compliance with regulations such as the GDPR, and privacy best-practices.



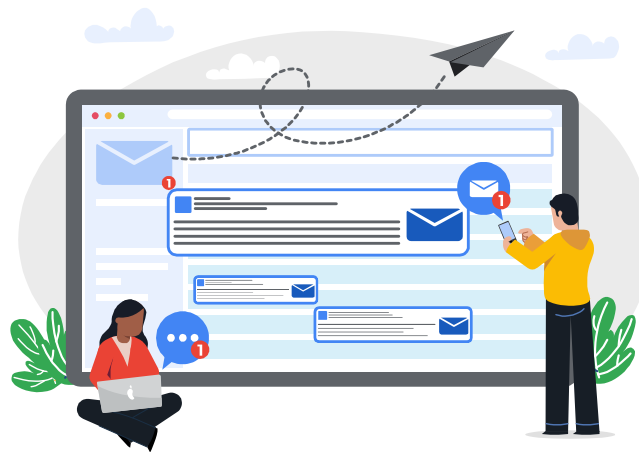
Our extensive data governance & security controls provide you with confidence in the protection of your data from other customers, users, attackers, and unauthorized access by Google employees.

# Data governance in model usage

Google processes prompts to provide the service.

Prompts input to the foundation model to generate a response, are encrypted in transit.

Google does not use prompt data to train its models without the express consent of its customers.



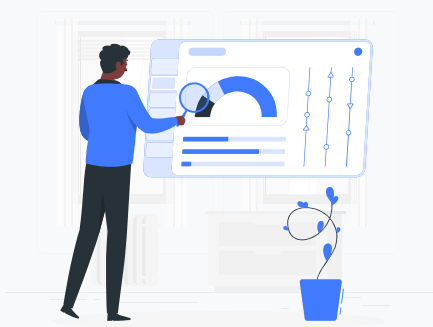
# Model Tuning



Parameter Efficient Fine Tuning (PEFT)

Customer specific adapter weights

Foundational model remains frozen during inference



Input data is secured at every step

Adapter weights are stored securely

Customer can delete adapter weights at any time

Customer Data will not be logged to train foundation models by default





# Our Approach to Data Governance & Privacy with our AI offerings

Policy Guardrails,  
Controls and Mitigations



Experience  
and Certifications



Privacy and Security  
as primary AI design criteria



Privacy by Design  
and by Default



Improving your  
AI Data Governance & Privacy



Transparency  
& DPIA Assistance



# 3 Actions Leaders Can Take Now for AI Data Governance

- 1 Record **data provenance**
- 2 **Establish & uphold policies** for data use, classification and handling
- 3 Take a **risk-based approach** to managing data use for AI





**Data  
governance  
and privacy**

---



**Security and  
compliance  
support**

---



**Reliability and  
sustainability**

---



**Safety and  
responsibility**

---

# Google AI Principles

## AI Should

---

- 01 be socially beneficial
- 02 avoid creating or reinforcing unfair bias
- 03 be built and tested for safety
- 04 be accountable to people
- 05 incorporate privacy design principles
- 06 uphold high standards of scientific excellence
- 07 be made available for uses that accord with these principles
  - Primary purpose and use
  - Nature and uniqueness
  - Scale
  - Nature of Google's Involvement

## Areas we will not pursue

---

- 01 likely to cause overall harm
- 02 principal purpose to direct injury
- 03 surveillance violating internationally accepted norms
- 04 purpose contravenes international law and human rights

# Google Cloud's approach to Responsible AI



## Product & use case reviews

Identifying, assessing, and mitigating potential impacts before they are generally available to customers.



## RAI Tooling, Enablement, and Support

Enabling our customers to identify, assess, and mitigate potential impacts within their use case and application(s).



## Education, research & best practices

Equipping our customers with recommendations, thought leadership, and transparency as they navigate Responsible AI.

# Responsible AI reviews identify, assess, and mitigate potential harmful social impacts based on our AI Principles.

**Example project:** A customer engaged Google Cloud's Professional Services to build an AI solution to detect objects in user-generated images to personalize a website experience.

**Identify  
potential harms** >

Users from subgroups might have a negative experience on the website based on how images are categorized.

**Assess  
risk levels** >

Unfair bias could be reinforced by in the user's experience on the website.

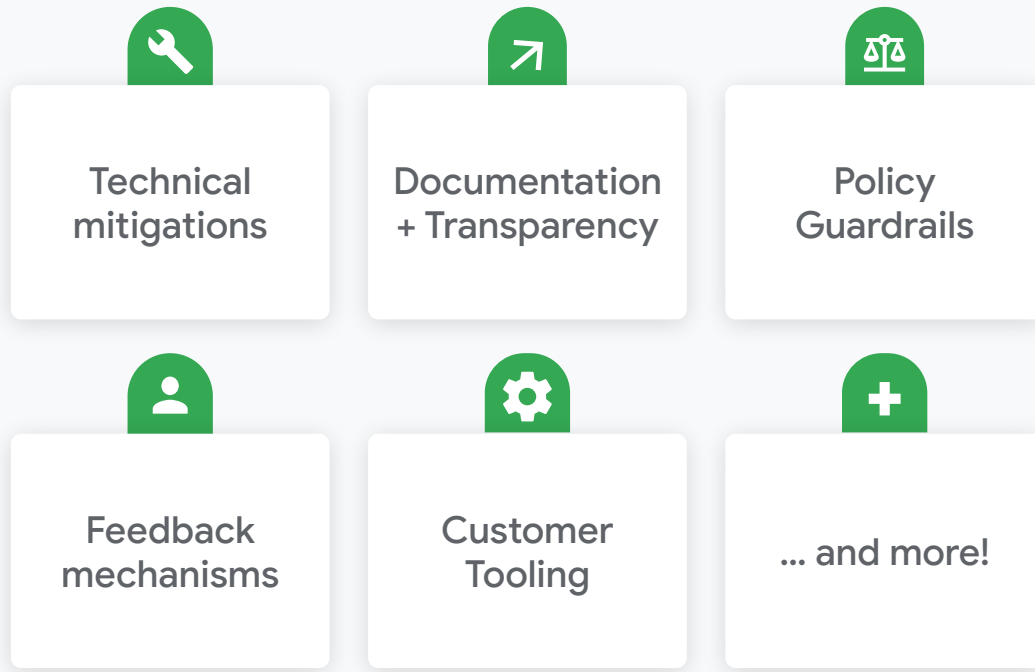
**Develop  
mitigation plans** >

Advise the customer on fairness testing for sensitive categories and specific subgroups.



## Responsible Generative AI Mitigation Plans

Our impact assessments and reviews shape **sociotechnical Responsible AI** mitigation plans for our generative products before they are in production.





# Responsible AI Tooling, Enablement, and Support

Enabling our customers to identify, assess, and mitigate potential impacts within their use case and application(s).



Technical  
Safeguards



Content  
Moderation  
API



Recitation  
Checker



Unfair Bias  
Evaluation  
Tooling





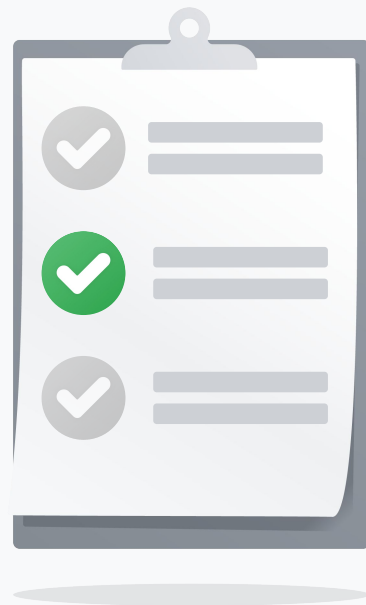


## Recitation checks

Recitation checks help ensure that our models do not replicate existing content at length

We've designed our systems to limit the chances of replicating existing content at length and we will continue to improve how these systems function.

Generally, if our API does directly quote at length from a webpage, it cites that page in our output from the model, or if over a certain length, it will block the output.





## Research, Education, and Best Practices

Equipping our customers with recommendations, thought leadership, and transparency as they navigate Responsible AI.



Leveraging  
Google  
Research



Model Cards



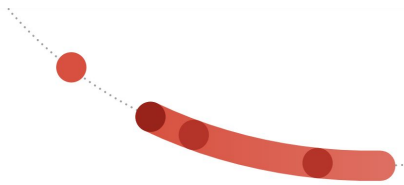
Responsible  
AI Guides



*... and more!*

# 3 Actions Leaders Can Take Now for Responsible AI

- 1 Be an **active and visible sponsor for RAI**
- 2 Establish **AI governance**
- 3 Build RAI **capacity and capabilities**



RESPONSIBILITY:

## Responsible AI practices

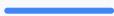
The development of AI has created new opportunities to improve the lives of people around the world, from business to healthcare to education. It has also raised new questions about the best way to build fairness, interpretability, privacy, and safety into these systems.

General recommended practices for AI    Fairness    Interpretability    Privacy    Safety

<https://ai.google/responsibility/responsible-ai-practices/>



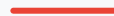
**Data  
governance  
and privacy**



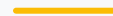
**Security and  
compliance  
support**



**Reliability and  
sustainability**



**Safety and  
responsibility**



# Thank You