

Collibra AI Copilot

AI feature description	This is a textual chat interface, where the user can pose numerous questions regarding the customer's environment. Within certain parameters, the feature will find answers to those questions and deliver responses accordingly.
Feature optional?	Yes
Model	Model Type: Google Vertex AI Gen AI?: Yes Owner: Google LLC
Intended use	Answering user questions, enabling easier discovery of relevant assets.
Model training/retraining data	Collibra relies on the training data leveraged by Google Vertex AI as shipped. For more information, please see the Google Vertex AI Product Documentation .
Model input	User queries, as well as any supporting data (Collibra assets or information from Collibra's public documentation) will be sent to the LLM. When the Copilot feature is first enabled, metadata from customer data assets and business terms are sent to an embedding model in order to generate embeddings. These embeddings are necessary to populate a RAG (retrieval-augmented generation) database so that the LLM can search across many assets.
Model output	Freeform response, potentially containing links to either public documentation or customer assets.
Is input data used to train the model for general product improvement purposes?	No
Data storage and retention	All Collibra storage locations are referenced here . Retention is concurrent with customer environment retention.
Human in the loop?	Yes. Usage of suggested data is within the user's discretion.
Is personal data processed?	Personal data is not intended to be included in the model input data, but if it is, its processing is subject to the customer's data privacy terms and conditions with Collibra under its Master Cloud Agreement. You can find our standard, customer facing privacy terms here .
Other considerations	Any customer data assets a user does not have access to as part of their platform permissions (as designated by the customer's account administrator) will not be included in any input or output for that user.