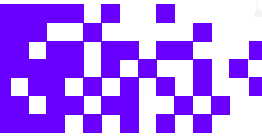


## HACKATHON GUIDE



# Vectara Overview

Vectara is the trusted GenAI platform. It's designed to make it easy for you to build and deploy GenAI applications that can generate text-based answers using your particular data (this type of application flow is also known as RAG or retrieval-augmented-generation). You just ingest your data and then build apps using the Query and Summarization APIs. It's that simple.

**Use-cases include:**

- **Question Answering:** Provide direct answers to specific questions based on the data.
- **Conversational AI:** Chatbots that can hold a full conversation with a human, with back-and-forth exchanges.
- **Semantic (Neural) Search:** Build applications powered by fast and powerful semantic search that finds documents by matching to the user intent.

# Getting started with Vectara

**You can get to know Vectara's features in 5 minutes:**

- [Sign up](#) for a free Vectara account.
- [Log in](#) and take the 5-minute walk-through.

**We have more resources to help you build apps with Vectara:**

- The [Quick Start guide](#) shows you how to use the Vectara Console.
- Read the [API recipes](#) for common patterns with our APIs.
- Our [API playground](#) shows you how Vectara's API requests and responses are structured.
- See "Additional resources", below, for a comprehensive list.

**General guidelines:**

- Vectara provides a generous free plan (ideal for hackathon) with up to 50MB of text and 15K monthly queries.
- Use the [Indexing API](#) to ingest data into your Vectara corpus.
- Use the [Search API](#) to run queries against the data ingested.
- The [Console](#) provides a unified view of information about your Vectara account, corpora and their characteristics. You can also run example queries directly from the console for quick experimentation.

# Common Questions

## What should I do if I need to go over the free plan limits?

Vectara's free plan provides a pretty generous offering of up to 50MB of text and 15K queries a month.

This should be sufficient for many use-cases, including in a hackathon setting.

To get additional capacity, you can add a credit card to the account and buy additional [bundles](#), or [transition](#) to our Scale plan.

## Should I use RAG and not fine-tuning?

Our experience shows that "Fine-tuning is for Form, and RAG is for facts" as discussed [here](#) and [here](#).

## What is the Boomerang embedding model?

Boomerang is the name of Vectara's newest embedding model. This model encodes text from your data as "vector embeddings" and is used to power the high performance retrieval process that is part of the RAG pipeline.

Read more about [Boomerang](#), and [the importance of using a good retrieval model](#) for getting best results from RAG.

## Where is my customer ID / corpus ID / API key?

- Here's some info on [how to find your customer ID](#).
- To get a corpus ID, view a corpus in the Console. The corpus ID is at the top of the screen, e.g. "Corpus ID: XXX".
- To create API Keys - follow this [guide](#).



# Additional resources

API docs	<a href="https://docs.vectara.com/docs/">https://docs.vectara.com/docs/</a>
API playground	<a href="https://docs.vectara.com/docs/rest-api/">https://docs.vectara.com/docs/rest-api/</a>
Getting help	Join our <a href="#">Discord</a> server or <a href="#">forums</a> if you have questions. If you have any feedback for us, we would be glad to hear it - please let us know in the forums or our Discord channel.
Open source	<ul style="list-style-type: none"><li>■ <a href="#">Vectara-ingest</a> helps with data ingestion – crawling data sources and indexing them into Vectara.</li><li>■ <a href="#">Vectara-answer</a> is a user interface for question answering – demonstrates a UI concept.</li><li>■ <a href="#">React-search</a> is a React package that allows you to integrate Vectara semantic search in any React app with just a few lines of code.</li><li>■ <a href="#">Create-UI</a> is a fast way to generate a Vectara-powered sample codebase for a range of user interfaces.</li></ul>
Sample code	We have quite a few "sample applications" hackers can take a look at <a href="#">here</a> . Additionally, we published some sample code in these jupyter <a href="#">notebooks</a> .
Blog posts	<a href="#">A reference architecture for grounded generation</a> <a href="#">What is vector search?</a> <a href="#">RAG done right: chunking</a> <a href="#">RAG done right: retrieval</a> <a href="#">RAG done right: databases</a>
Youtube	Flowise + Vectara <a href="#">Tutorial</a> Ask LangChain Docs <a href="#">video</a> More <a href="#">here</a>
Integrations	LangChain: <a href="https://blog.langchain.dev/langchain-vectara-better-together/">https://blog.langchain.dev/langchain-vectara-better-together/</a> LlamaIndex: <a href="https://vectara.com/llamaindex-vectara/">https://vectara.com/llamaindex-vectara/</a> Vectara and Airbyte: <a href="https://vectara.com/blog/vectara-and-airbyte/">https://vectara.com/blog/vectara-and-airbyte/</a>
Startup program	<a href="https://vectara.com/startups">https://vectara.com/startups</a>

Happy Hacking!

 Palo Alto, CA, USA [www.vectara.com](https://www.vectara.com)[Let's Go! Sign Up!](#)