Appsin.Ai

SERVICE

# Generative AI Development Services for Everyday Business Work

Generative AI applications that draft, summarise, search and answer from your own content, built on the model that fits and tested before rollout.

https://appsin.ai · +91-9025555584 · sales@appsin.ai
AI development for businesses in India, the GCC and the USA

OVERVIEW

# Generative AI Development

Generative AI writes, summarises, answers and creates: replies to customers, first drafts of reports, answers from a policy library, code suggestions, product images. We build these tools for businesses, connected to your documents and systems, with checks that keep the output accurate enough to rely on and a person in the loop where it matters.

We work with hosted models from OpenAI, Anthropic, Google and Mistral, and open-source models such as Llama that you can host yourself. For most business tools, retrieval-augmented generation (RAG) is the core: the system finds the relevant passages in your content and the model answers from them, citing where each answer came from.

Generative models can be wrong with confidence, so evaluation is part of the build, not an afterthought. Each project gets a test set of real questions with approved answers, guardrails on what the system may say and do, and logging so you can see what it answered and why.

BUILT FOR

- Customer support and service teams
- Marketing and content teams
- Legal, finance and operations teams handling documents
- SaaS companies adding AI features
- Organisations with large document libraries

FEATURES

## What you get with Generative AI Development

01 **Use-case discovery**
We map the tasks where drafting, summarising or searching takes your staff the most time, and pick the first one to build with a clear measure.

02 **RAG knowledge assistants**
Assistants that answer from your manuals, policies and tickets, cite the source passage and say so when the answer is not in your content.

03 **AI copilots for staff**
Writing, research and coding helpers placed inside the tools your team already uses, from the CRM to the code editor.

04 **Intelligent document processing**
Extract, classify and check data from invoices, contracts and forms, with low-confidence fields sent to a person to confirm.

05 **Content generation**
Product descriptions, marketing copy, emails and social posts in your brand voice, reviewed and approved before anything is published.

06 **Image, voice and speech**
Image generation for product visuals and creative drafts, plus speech recognition, voice bots and text-to-speech.

07 **Chatbots and AI agents**
Customer-facing chatbots on web and messaging channels, and agents built with LangGraph, CrewAI or Microsoft Agent Framework that complete multi-step tasks.

08 **Model API integration**
OpenAI, Gemini, Claude, Llama and Mistral connected to your CRM, ERP, CMS or SaaS product through Python and FastAPI services.

09 **Fine-tuning and governance**
Fine-tuning where prompting and RAG fall short, plus usage policies, content filters and audit logs for privacy and compliance.

WHY IT MATTERS

## The difference it makes to your business

01 **Time back from routine writing**
First drafts, summaries and replies arrive in seconds for a person to edit, instead of being written from a blank page.

02 **Answers people can check**
Source citations let staff and customers see where an answer came from, and spot when the model has strayed beyond your content.

03 **Freedom to change models**
We design around a model layer you can switch, so you are not tied to one provider's pricing, limits or roadmap.

04 **Knowledge that is easy to find**
Years of documents become searchable in plain language, so new staff find answers without interrupting colleagues.

HOW WE DELIVER

## How we deliver Generative AI Development

01 **Requirements**
We agree the task, the users, the data sources, the constraints and the questions the system must answer well.

02 **Approach and model choice**
We choose between prompting, RAG, fine-tuning or an agent, and shortlist models by output quality, cost, speed and where they can be hosted.

03 **Design and build**
Architecture, retrieval pipeline, vector database such as Pinecone, Weaviate or Chroma, prompts and user interface are built in short iterations.

04 **Evaluate**
The system is scored against the test set, checked for invented facts and unsafe answers, and tuned until it meets the agreed standard.

05

### Deploy and support

We release to a pilot group, then more widely, monitoring answer quality, cost per request and user feedback, and updating as models change.

FAQ

## Questions about Generative AI Development

### What does a generative AI project cost?

The main drivers are the number and quality of data sources, whether RAG is enough or fine-tuning is needed, the integrations, the volume of requests (hosted models charge per token) and the hosting choice. A focused assistant over one document library costs far less than an agent working across several systems. We quote a fixed price for the agreed scope.

### How long does a generative AI project take?

Typically 4 to 16 weeks, depending on scope, integrations and how much customisation is needed. A pilot on one use case sits at the shorter end; a production system connected to several business systems at the longer end.

### Which models do you work with?

OpenAI GPT models, Google Gemini, Anthropic Claude, Mistral and open-source models such as Llama. We pick per task, comparing answer quality on your own test questions, cost per request, speed and whether the model can be hosted where your data rules require.

### Should we use RAG or fine-tuning?

For answering from your documents, start with RAG: it keeps answers current, shows sources and needs no training. Fine-tuning helps when you need a fixed format, tone or specialised behaviour that prompting cannot reliably produce. Many systems use both.

### How do you stop the AI making things up?

No method removes the risk entirely, so we reduce and contain it. Answers are grounded in retrieved passages with citations, the model is told to say when it does not know, outputs are scored against a test set before launch, and high-stakes actions need a person to approve them.

### Can it run without sending our data to an outside AI provider?

Yes. Open-source models such as Llama and Mistral can be fine-tuned and hosted on your private cloud or servers, so prompts and documents stay inside your environment. The trade-off is that you carry the hosting cost and may need a larger model for complex tasks.

### Can you add generative AI to our existing software?

Yes. We add it to web and mobile apps, CRM, ERP, SaaS products and internal tools through APIs, so users get drafts, answers and summaries in the screens they already work in.

NEXT STEP

## Pick one task to hand to generative AI

Book a free 30-minute call. Bring the task your team repeats most, and we will outline how a generative AI tool would handle it, which model we would test first and how we would measure it.

**Book a free consultation:** https://appsin.ai/book-a-consultation/

**WhatsApp or call:** +91-9025555584

## About AppsInAi

AppsInAi Private Limited is an AI development company in Coimbatore, India. We build generative AI assistants, AI agents, chatbots, machine learning models, computer vision and data platforms on our clients' own data, from a first proof of concept to production, with monitoring and support after launch.

**Contact**
sales@appsin.ai
+91-9025555584
Housing Unit, Sf No. 393, CMS School Backside, 3rd Street, Nehru Nagar, Ganapathy, Coimbatore 641006, India
https://appsin.ai