

Introduction to Generative AI with Google Cloud



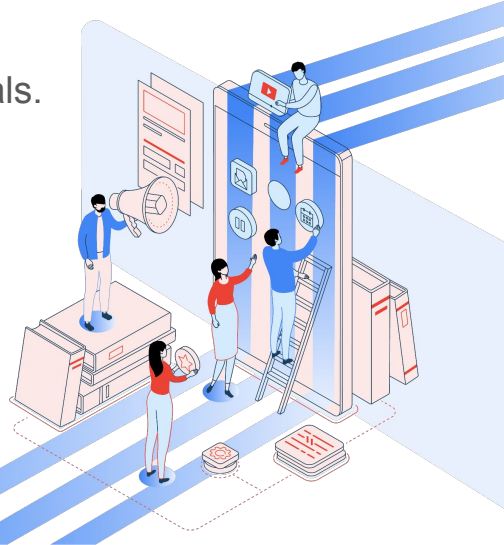
**Dr. Yogesh Haribhau
Kulkarni**

Google Developer Expert (Machine Learning)
AI Coach, Teacher, Speaker

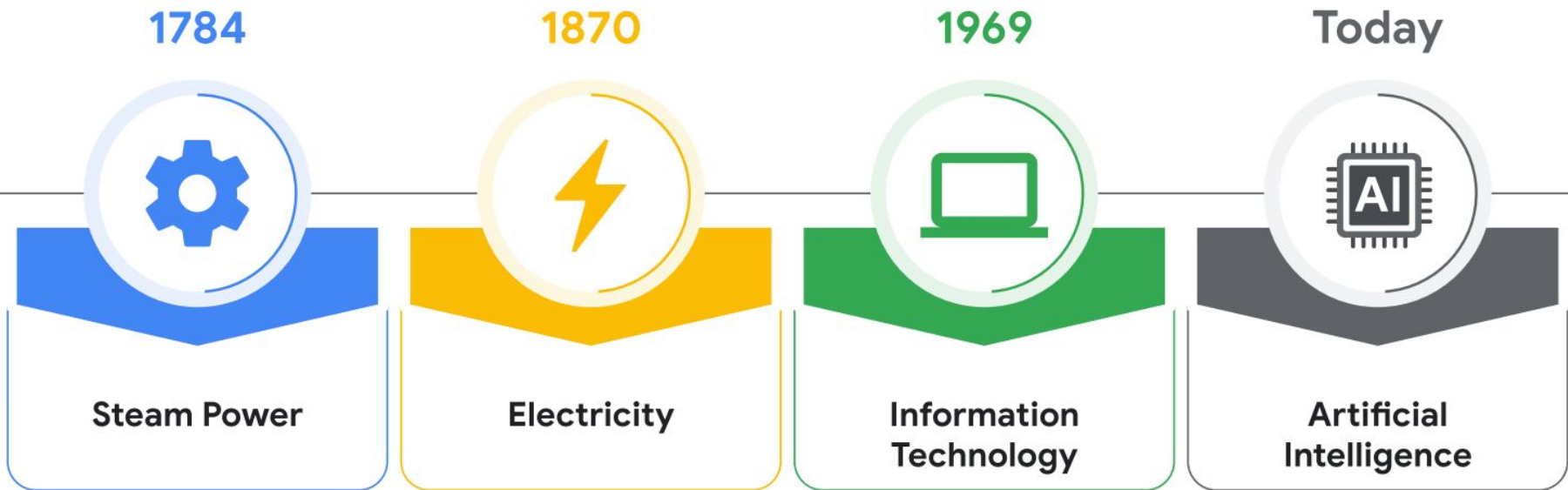


Era of Generative AI

- Generative AI: Branch of AI for creating original content based on existing data patterns.
- Applications: Images, videos, text generation.
- Examples:
 - Text Generation: PaLM - articles, stories, poetry.
 - Image Generation: StableDiffusion - realistic images of people, animals.
 - Music Generation: MuseNet - original music in various genres.



Where we came from ...



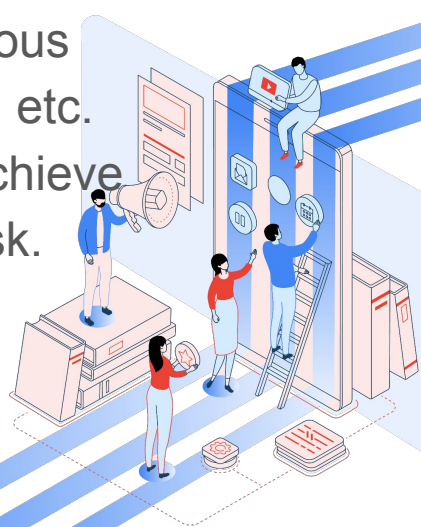
Classical ML Approach

- Data Collection: Obtain training and evaluation data for a specific use case (e.g., customer churn prediction).
- Train ML Model: Train a model from scratch for the identified use case.
- Deployment: Deploy the trained model to address the specific business problem.
- 1:1 Relationship: Each trained model corresponds to a single use case.

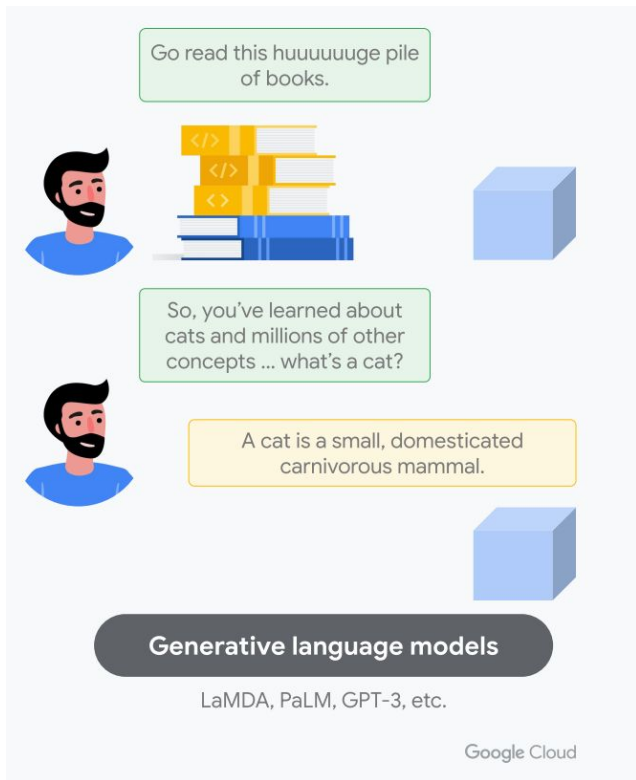


Foundation Models in Generative AI

- Paradigm Shift: Foundation Models (FMs) revolutionize AI work.
- One-to-Many Relationship: Foundation Models allow addressing multiple use cases without training or fine-tuning separate models.
- Versatility: A single Foundation Model can be used for various tasks like summarization, sentiment extraction, translation, etc.
- Simplified Development: Develop a single application to achieve multiple tasks instead of separate applications for each task.

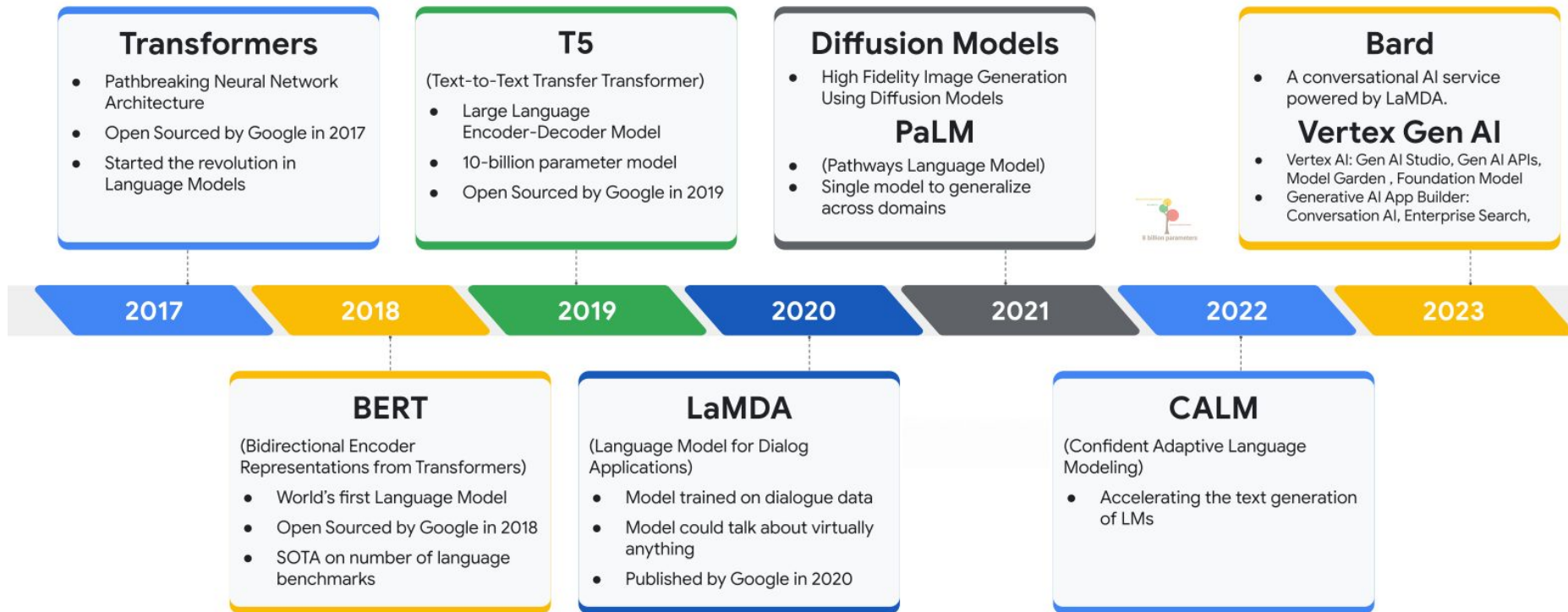


What are large language models?



- ML algorithms that can recognize, predict, and generate human languages
- Pre-trained on petabyte scale text-based datasets resulting in large models with 10s to 100s of billions of parameters
- LLMs are normally pre-trained on a large corpus of text

This revolution started at Google ...



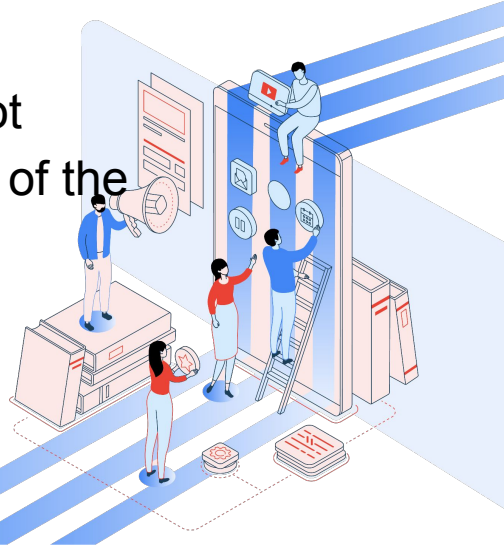
New Programming Language: English

- **Prompt:** The prompt is your text input that you pass to the model.
- **Prompt Design:** The art and science of figuring out what text to feed your language model to get it to take on the behavior you want.



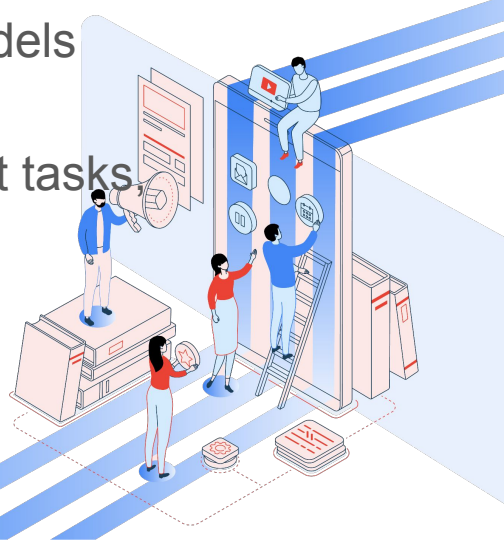
Shots

- **Zero-shot prompt:** The model is provided with no example when prompting for response.
- **One-shot prompt:** The model is provided with one example to the LLM within the prompt to give some guidance on what type of response you want.
- **Few-shot prompt:** Few-shot prompts are similar to one-shot prompts, but the model is given multiple labeled examples of the task.



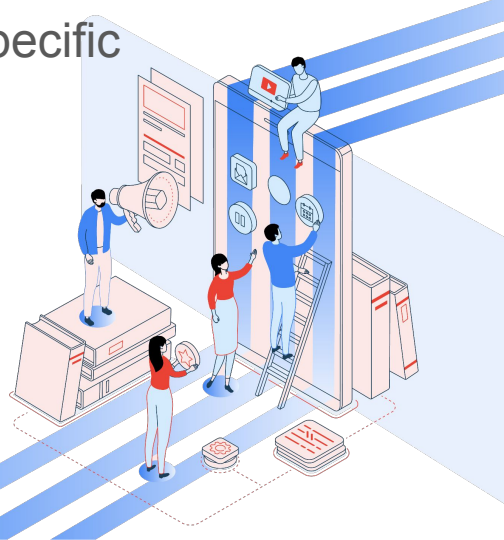
Evolution of Cloud ML APIs

- Previous Cloud ML APIs: Limited to specific applications, tied to one task per API.
- Example: Separate APIs for summarization, sentiment analysis, translation, etc.
- Foundation Models' Impact: Multi-purpose Foundation Models replace the need for task-specific APIs.
- Flexibility: Use the same Foundation Model for various text tasks, making development more efficient.



Reduced Barrier of Entry

- Drastic Reduction: Foundation Models lower the barrier of entry for AI applications in business.
- Focus on Applications: Companies concentrate on building applications, not training models.
- Prompt Engineering: An approach to tailor the model for specific tasks.



Gen AI on Google Cloud

- Google Cloud's new GenAI products for enterprise adoption.
- Google Cloud Solution: Secure environment for data with Generative AI capabilities.
- Customer Data Protection: Google doesn't use customer data to improve FMs, nor access or store it.



Gen AI Products/Services

- Generative App Builder: Offers out-of-the-box solutions for search and conversational experiences.
- Vertex AI Integration: GenAI services integrated with Vertex AI for end-to-end ML platform usage.
- Secure Development: Developers can securely develop and deploy models in production-ready environments.
- State-of-the-art Models: Access to cutting-edge models with enterprise security and deployment support.



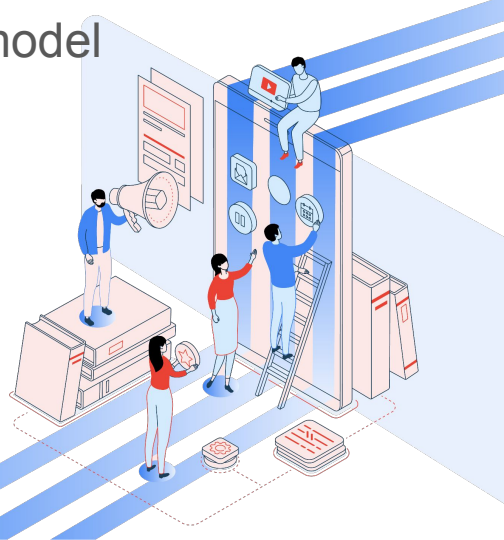
Gen AI App Builder

- Purpose: Empowers companies with ready-to-use solutions for search and conversational experiences.
- Benefits: Simplifies application development by leveraging GenAI capabilities.
- Use Cases: Enables efficient search experiences and interactive conversational interfaces.
- Instant Solutions: Quick implementation of GenAI-powered features without extensive development.

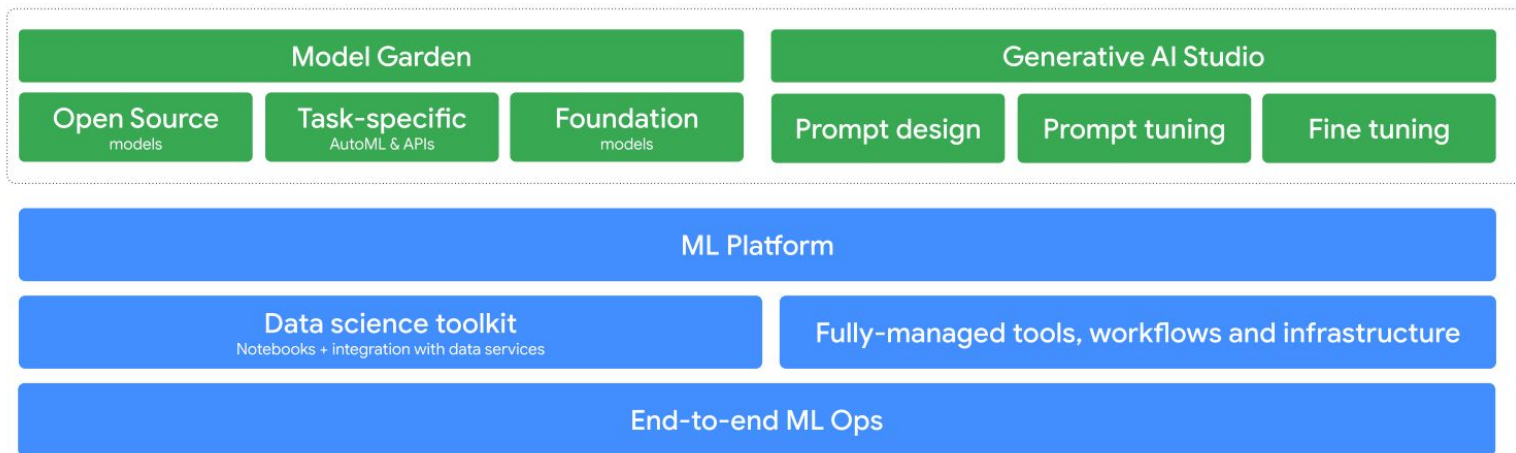


Vertex AI Integration

- Vertex AI Platform: Google's comprehensive end-to-end Machine Learning platform.
- GenAI Services Integration: Allows developers to leverage GenAI within Vertex AI.
- Secure Development: Ensures a secure environment for model development and deployment.
- Production-Ready: Enables deploying GenAI models in enterprise-grade production environments.



Vertex AI



Generative AI GitHub Repository

Sample code and notebooks for GenAI on Google Cloud



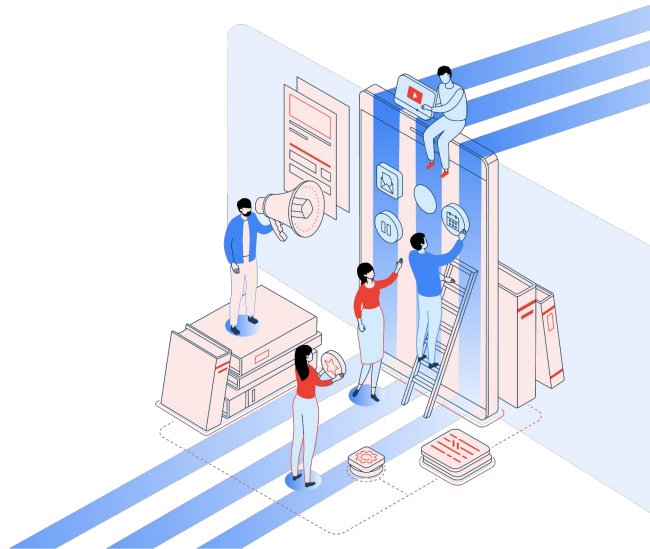
goo.gle/gen-ai-github

Table of Contents

- [Language/](#)
 - [Getting Started with Generative AI Studio without code](#)
 - [Intro to Vertex AI PaLM API](#)
 - [Intro to Prompt Design](#)
- [Examples/](#)
 - [Prompt Design/](#)
 - [Ideation](#)
 - [Question & Answering](#)
 - [Text Classification](#)
 - [Text Extraction](#)
 - [Text Summarization](#)
 - [Reference-architectures/ *NEW*](#)
 - [Product Description Generator from Image](#)
 - [Document Q&A/ *NEW*](#)
 - [Question Answering with Large Documents with LangChain](#)
 - [Question Answering with Large Documents \(without LangChain\)](#)
 - [Document Summarization/ *NEW*](#)
 - [Summarization with Large Documents with LangChain](#)
 - [Summarization with Large Documents \(without LangChain\)](#)
 - [LangChain-intro/ *NEW*](#)
 - [Getting Started with LangChain 🦜️🔗 + Vertex AI PaLM API](#)
 - [Tuning/](#)
 - [Tuning a Foundational Model, Deploying, and Making Predictions](#)

Google Cloud

Learn more about Generative AI at
goo.gle/generativeai



Thank You!

