

Generative AI and its Applications

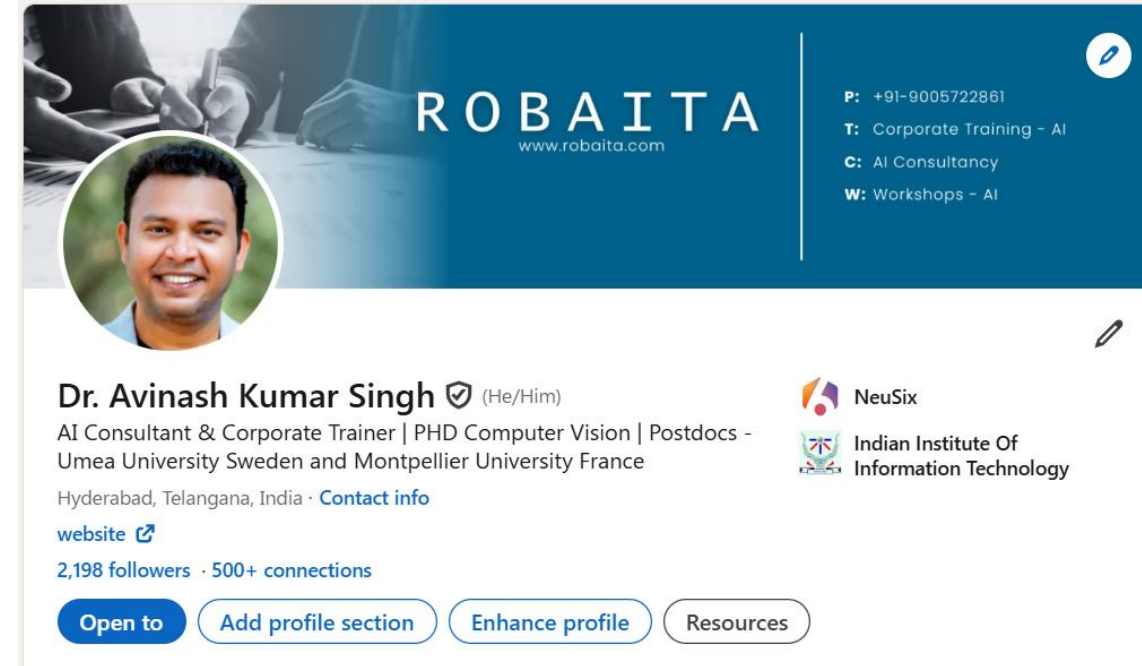
Dr. Avinash Kumar Singh

AI Consultant and Coach, Robaita



Dr. Avinash Kumar Singh

- ❑ **Possess** 15+ years of **hands-on expertise** in Machine Learning, Computer Vision, NLP, IoT, Robotics, and Generative AI.
- ❑ **Founded** Robaita—an initiative **empowering** individuals and organizations to **build, educate, and implement** AI solutions.
- ❑ **Earned** a Ph.D. in Human-Robot Interaction from IIIT Allahabad in 2016.
- ❑ **Received** postdoctoral fellowships at Umeå University, Sweden (2020) and Montpellier University, France (2021).
- ❑ **Authored** 30+ research papers in **high-impact** SCI journals and international conferences.
- ❑ Unlearning, learning, making mistakes ...



<https://www.linkedin.com/in/dr-avinash-kumar-singh-2a570a31/>



BRANE



Things to be discussed

- Introduction to Generative AI
 - Large Language Model
 - Vision Language Model
 - Audio Language Model
- Generative AI Use Cases
- Generative AI Application
 - Chatbot
- Agentic AI
 - Virtual Assistant
- Difference between Generative AI and Agentic AI
- Ethical use of AI

What is Generative AI

- Generative artificial intelligence (generative AI, GenAI, or GAI) is a subset of artificial intelligence that uses generative models to produce text, images, videos, or other forms of data.
- These models often generate output in response to specific prompts.
- Generative AI systems learn the underlying patterns and structures of their training data, enabling them to create new data.

llama3.2:latest

Here's an explanation of Generative AI that your grandfather can understand:

"Generative AI is like a super powerful pen that can draw anything you want. It uses computer magic to create new things, like pictures or stories, based on what you show it. Think of it like a very smart and fast artist!"

Prompt: You need to explain my grandfather who is a 60 year old about what is generative AI. My grandfather is studied till 10th standard and has no idea about AI and the technology. Explain in very simple terms, use example for better understanding. Don't cross more than 50 words.

Prompt: A bear, lion and rabbit talking in a dark green forest



Different Forms of Generative AI

Large Language Model (LLM)

- **Natural Language Understanding:**
 - LLMs comprehend text context, nuances, and intent, enabling them to answer questions, follow instructions, and carry on coherent conversations.
 - Example: Given a complex question like, "What are the effects of climate change on marine life?", the LLM provides a summary of relevant impacts, including ocean acidification and habitat loss.
- **Text Generation:**
 - LLMs can create new, contextually relevant text, such as essays, stories, code, or reports, based on prompts.
 - Example: A prompt like "Write a short story about space exploration" results in a creative narrative with characters, setting, and plot progression.
- **Summarization:**
 - LLMs distill long documents or articles into concise summaries, capturing key points and relevant information.
 - Example: Given a scientific paper on renewable energy, the LLM produces a brief summary covering main findings, like advancements in solar efficiency.
- **Translation & Language Adaptation:**
 - LLMs translate text between languages while preserving context, tone, and idiomatic expressions, making content accessible globally.
 - Example: A user inputs "Hello, how are you?" in English, and the LLM accurately translates it to "Bonjour, comment ça va ?" in French, maintaining conversational tone.

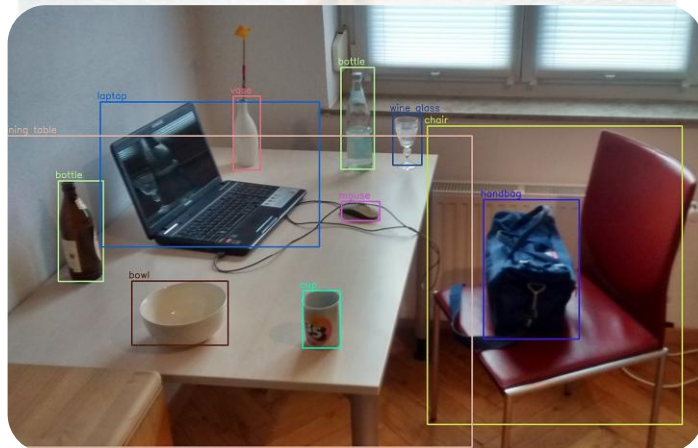
Different Forms of Generative AI

Vision Language Model (VLM)



A dog sitting on a rock in front of a lake

Image Captioning



Object Detection with Contextualization



Visual Question - Answer

See the image and tell me the total bill amount?



The total bill amount in the image is ₹5795.00.

<https://openai.com/index/sora/>



Visual Question Answering

what is this image all about, could you please describe this



This image depicts a bear, a lion, and a rabbit sitting together in a lush forest setting. The animals appear to be engaged in a calm, almost conversational moment, with the bear and lion looking at each other and the rabbit sitting attentively. The scene has a surreal, storybook quality, as these animals are not typically found interacting peacefully in nature, especially in a forest environment. The greenery and tall trees in the background create a serene, woodland atmosphere, adding to the feeling of a fantastical or symbolic meeting among the animals.

This type of imagery is often used to convey messages about harmony, unity, or storytelling, where animals of different species gather as characters with human-like qualities.

could you tell me the color of the rabbit and where it is in the image?



The rabbit in the image is light gray with hints of white, and it has large pinkish ears. It is positioned in the lower left area of the image, sitting on the ground in front of a log, close to the bear and lion. The rabbit appears to be looking up toward the bear and lion, as if it's part of their "conversation" or interaction.

Different Forms of Generative AI

Audio Language Model (ALM)

- Speech Recognition (Speech to Text):
 - ALMs can transcribe spoken language into text by converting audio signals into text data.
 - Example: Voice assistants like Siri and Alexa use speech recognition to understand and respond to spoken commands, such as "What's the weather today?"
- Speech Synthesis (Text-to-Speech):
 - ALMs can convert written text into spoken audio that sounds natural.
 - Example: E-book readers or accessibility tools read out text from an article or document to users with a natural voice, enhancing accessibility.
- Emotion Detection:
 - ALMs can detect emotions or sentiments in audio, helping understand the speaker's mood.
 - Example: In customer service calls, emotion detection helps identify if a customer sounds frustrated or satisfied, aiding in personalized responses.
- Speaker Identification:
 - ALMs can distinguish between and recognize different speakers based on voice characteristics.
 - Example: In a meeting transcription, the model can identify each speaker by their voice, making it easier to attribute statements to the correct person.

<https://suno.com/about>

Generative AI for HR

- **Candidate Screening Assistant:** Use prompts to analyze resumes and match them with job descriptions for shortlisting.
- **Employee Sentiment Analyzer:** Summarize feedback from surveys and generate insights about team morale.
- **Interview Q&A Generator:** Generate role-specific interview questions tailored to job profiles.
- **Policy Simplifier:** Convert complex HR policies into easy-to-understand summaries for employees.
- **Performance Review Generator:** Assist managers in drafting performance review comments based on employee data.

Generative AI for Finance

- **Financial Report Summarizer:** Generate concise summaries of complex financial reports for decision-makers.
- **Expense Anomaly Detector:** Prompt LLMs to detect irregular patterns in expense sheets.
- **Tax Compliance Assistant:** Explain compliance requirements and generate responses to regulatory queries.
- **Budget Recommendation Generator:** Analyze past data and suggest optimized budget plans.
- **Investment Summary Creator:** Summarize investment opportunities and risks based on uploaded reports.

Generative AI for Sales & Marketing

- **Customer Persona Generator:** Create detailed customer personas based on demographic and behavioral inputs.
- **Campaign Content Creator:** Generate engaging marketing copy and campaign taglines tailored to target audience.
- **Sales Pitch Refiner:** Transform raw sales inputs into polished, persuasive pitches.
- **Market Trend Summarizer:** Summarize market trend reports into key actionable insights.
- **Lead Qualification Assistant:** Score and prioritize leads by analyzing CRM data using LLM prompts.

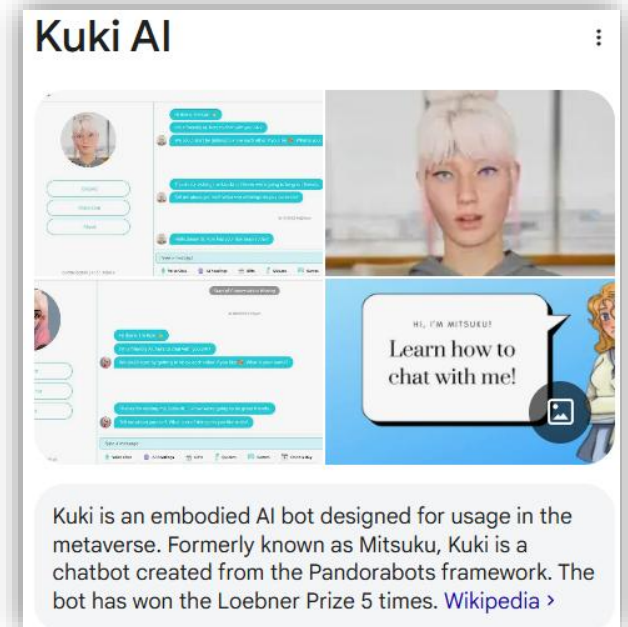
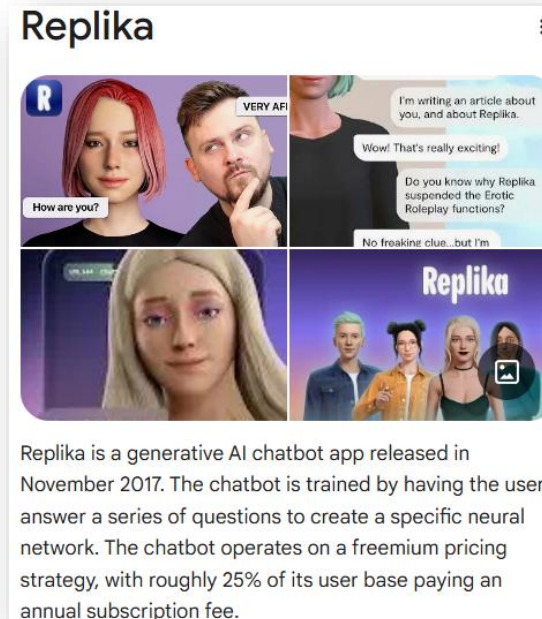
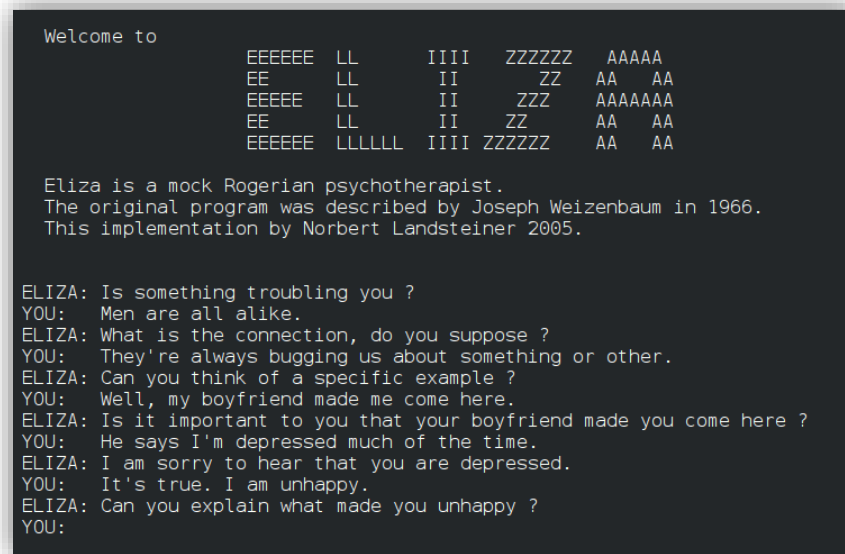
Generative AI for Admin

- **Meeting Summary Assistant:** Convert raw meeting transcripts into structured summaries and action items.
- **Policy Document Generator:** Draft internal administrative policies using prompt-based templates.
- **Travel Itinerary Planner:** Generate optimized business travel plans and schedules.
- **Office Memo Writer:** Compose clear, professional memos for administrative announcements.
- **Facility Issue Reporter:** Generate structured reports from free-text maintenance complaints.

Generative AI Application

Chatbot

A chatbot is a software application that simulates human conversation through text or voice interactions. It uses rule-based logic, natural language processing (NLP), or large language models (LLMs) to interpret user inputs and respond accordingly.

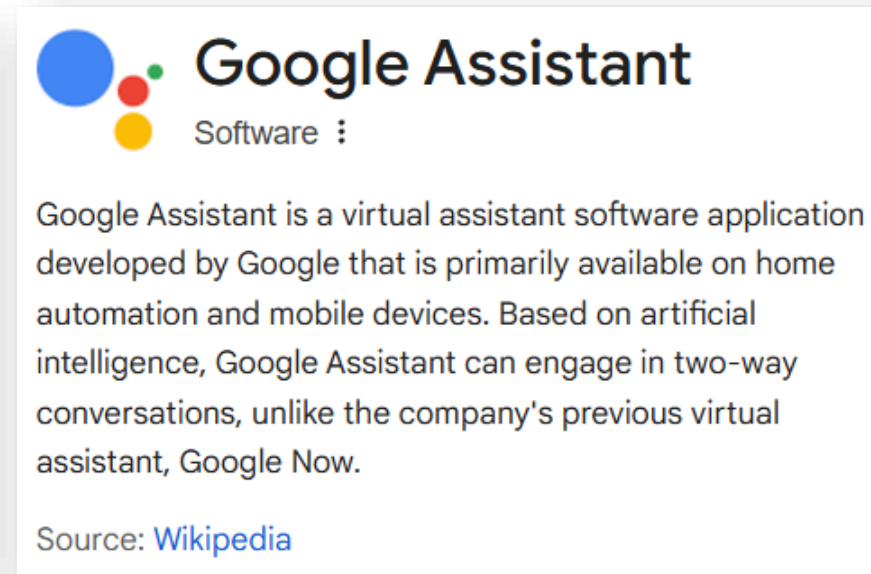
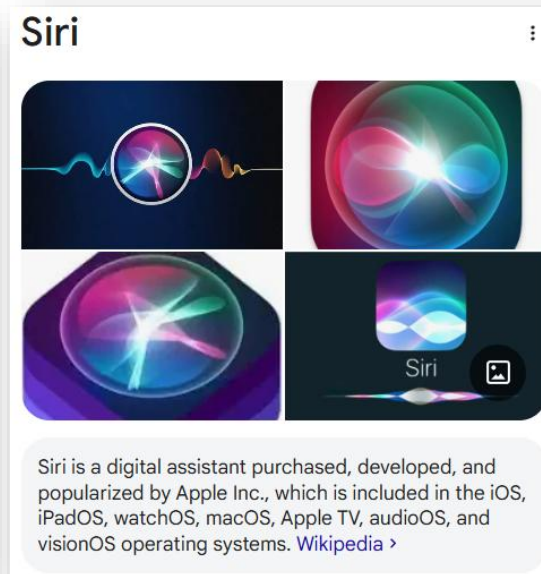
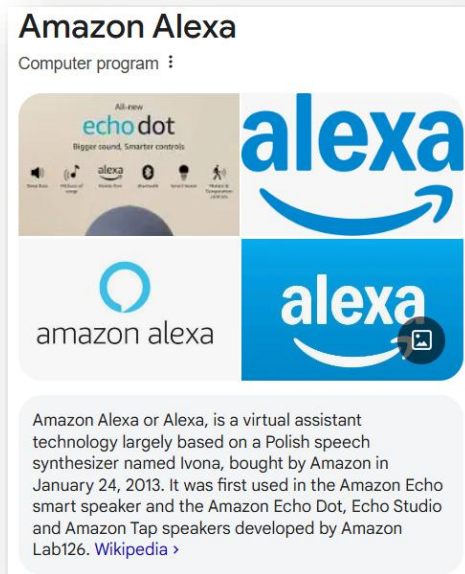


Agentic AI Application

Agentic AI refers to AI systems designed to autonomously plan, reason, and act to achieve goals, often interacting with tools, APIs, or external environments.

Virtual Assistant

A virtual assistant is a type of AI system that performs tasks or services based on user commands or preferences. It is often voice-enabled and designed to assist with daily activities, scheduling, reminders, and information retrieval.



Generative AI VS Agentic AI

Category	Generative AI	Agentic AI
Nature	Reactive, prompt-based	Proactive, goal-seeking
Initiation	Human-initiated	Self-initiated
Autonomy	Low	High
Context Handling	Limited, session-bound	Persistent, evolving memory
Decision-making	No long-term planning	Plans, adjusts, executes
Architecture	Single-task, stateless	Multi-agent, task-oriented
Generic Use Cases	Search, summarization, creative writing, chatbots, code assistants	Scheduling, multi-step task execution, autonomous research, personal AI assistants
Financial Services Use Cases	<ul style="list-style-type: none">Automated financial reportingVirtual customer assistantsRegulatory document summarizationPersonalized client outreachData structuring & enrichmentMarket sentiment analysisStress-test scenario modelingTransaction anomaly detectionKYC data processing	<ul style="list-style-type: none">Autonomous payment initiation, routing & retry logicDigital-asset settlement & token managementAI-driven digital wallet orchestrationProactive AML/fraud monitoring & alertingAutomated onboarding, KYC/ID verificationIntelligent loan origination & servicingAutomated regulatory reporting & SAR drafting
Learning	Pre-trained on large datasets; requires fine-tuning	Learns from experience, feedback, and evolving goals
Main Challenges	Hallucinations, lack of persistence, prompt sensitivity	Control, explainability, orchestration complexity, ensuring safety and compliance

Ethical Considerations & Challenges

Voice Cloning



Ethical Considerations & Challenges



Ethical Considerations & Challenges



Ethical Considerations & Challenges



**Can you
guess which
one is the
real person?**



Can Studio Ghibli Sue OpenAI Over AI-Generated Images? What Lawyer Says

The lawyer said that OpenAI could face legal challenges for training its AI models on copyrighted works belonging to Studio Ghibli

Edited by: [NDTV News Desk](#) | [World News](#) | Apr 02, 2025 14:42 pm IST ⓘ

Read Time: 3 mins

↪ Share



Image posted on X by @Zeneca



Via Email and Certified Mail

March 27, 2025

Gib Studio LLC
755 Dotings Ln
Miami, FL 33137

Re: Unauthorized Use of Studio Ghibli Intellectual Property – Cease and Desist

Dear Gib Studio LLC,

We are legal representatives of Studio Ghibli, Inc. ("Studio Ghibli"), the world-renowned animation studio responsible for original motion pictures such as Spirited Away, My Neighbor Totoro, Princess Mononoke, and others. Studio Ghibli is the exclusive owner of all associated intellectual property rights, including visual style, character likenesses, thematic elements, and trademarks.

It has come to our attention that your company has developed and publicly launched an application under the name "Gib", described as "a photo sharing app that turns all photos into Studio Ghibli style images." This product appears to heavily borrow from Studio Ghibli's artistic style, branding cues, and thematic presentation, and it references our studio by name in marketing and promotional materials.

Please be advised that this constitutes unauthorized use and misappropriation of Studio Ghibli's copyrighted works, trade dress, and brand name, and it is likely to cause confusion among consumers regarding endorsement, sponsorship, or affiliation — none of which exist.

Accordingly, we demand that you:

1. Immediately cease and desist all use of the Studio Ghibli name, references to its style, characters, or films in all promotional, descriptive, or functional aspects of your app or business;
2. Terminate the availability and distribution of the app "Gib" in all platforms, including but not limited to the Apple App Store, Google Play, and any third-party marketplaces;
3. Remove or take down all marketing materials, websites, and social media content that reference Studio Ghibli or utilize similar visual elements;
4. Provide us with written assurance within seven (7) days of the date of this letter confirming your compliance with the above demands.

Failure to comply will leave us no choice but to pursue legal remedies, including seeking injunctive relief, statutory and actual damages, and attorneys' fees, without further notice.

This letter is not intended to be a complete statement of Studio Ghibli's rights, all of which are expressly reserved.

Sincerely,
Sakura & Hoshino LLP
Legal Counsel for Studio Ghibli, Inc.
legal@sakura-hoshino.com
+1 (212) 555-7283

Ethical Considerations & Challenges

- ❑ Misinformation and deepfakes
- ❑ AI biases and responsible AI use
- ❑ Insecurity: The impact of AI on jobs
- ❑ Privacy and Data Protection
- ❑ Impact on Creativity and Originality
- ❑ Legal and Regulatory Issues

Checking AI Generated Content

For AI-Generated Image Detection:

1. [Hive Moderation AI-Generated Content Detection](#)

1. Detects AI-generated **images**.
2. Supports popular models like MidJourney, DALL·E, and Stable Diffusion.

2. [AI or Not](#)

1. Simple tool to upload an image and detect if it's AI-generated.

3. [Illuminarty](#)

1. Detects AI-manipulated images (e.g., face swaps, synthetic photos).
2. Browser extension available.

Checking AI Generated Content

For AI-Generated Text Detection:

1. [ZeroGPT](#)

1. Offers a percentage score for how much of the content is AI-written.
2. Supports multiple languages.

2. [Sapling AI Detector](#)

1. Offers real-time prediction scores for GPT-like content.
2. Free and lightweight.

3. [CopyLeaks](#) AI Content Detector

1. Strong educational use case.
2. Can integrate into LMS or CMS systems.

4. [Writer.com](#) AI Content Detector

1. Especially useful for marketing and business copy.

Generative AI Tools

Tool Name	URL	Modalities
ChatGPT	https://chatgpt.com/	Text, Image – Generation & Understanding – 5 images
Gemini	https://gemini.google.com	Text, Image – Generation & Understanding – not with people
Copilot	https://copilot.microsoft.com	Text, Image – Generation & Understanding
Meta	https://www.meta.ai/	Text, Image – Generation & Understanding
Perplexity	https://www.perplexity.ai/	Text
Claude	https://claude.ai/	Text, PDF, Dashboard creation, etc

Thanks for
your time