

Generative AI Transforming the Future

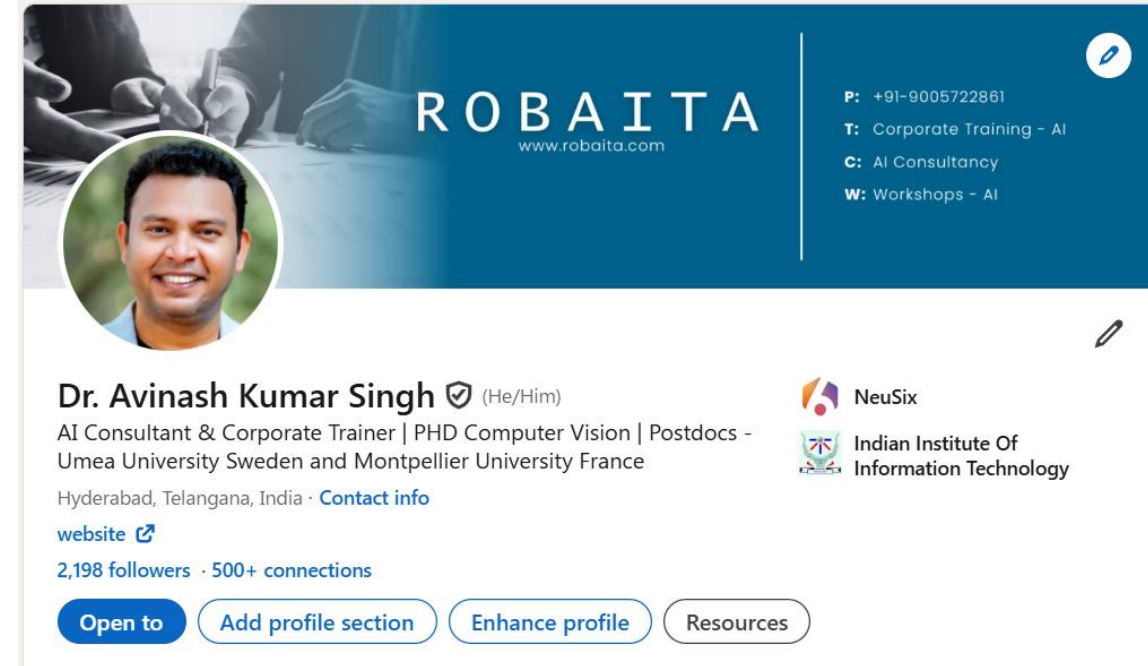
Dr. Avinash Kumar Singh

Robotics and Artificial Intelligence Training Academy



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/>



HCLTech



BRANE



Things to be discussed

- What is Generative AI
- How it works
- Different forms of Generative AI
 - Large Language Model
 - Vision Language Model
 - Audio LLM [Text to Speech, Speech to Text]
- Generative AI Applications
- Generative AI Tools
- Ethical Considerations & Challenges
- Future Direction

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



How it works

Tiktokenizer

gpt-4o

User

What is 2+2?

×

Assistant

2+2 = 4

×

User

What if it was *?

×

Assistant

2*2 = 4, same as 2+2!

×

Add message

```
<|im_start|>user<|im_sep|>What is 2+2?<|im_end|>
<|im_start|>assistant<|im_sep|>2+2 = 4<|im_end|>
<|im_start|>user<|im_sep|>What if it was *?<|im_end|>
<|im_start|>assistant<|im_sep|>2*2 = 4, same as 2+2!<|im_end|>
```

Token count
50

```
<|im_start|>user<|im_sep|>What is 2+2?<|im_end|><|im_s
tart|>assistant<|im_sep|>2+2 = 4<|im_end|><|im_start|>
user<|im_sep|>What if it was *?<|im_end|>
<|im_start|>assistant<|im_sep|>2*2 = 4, same as 2+2!<|
im_end|>
```

```
200264, 1428, 200266, 4827, 382, 220, 17, 10, 17, 30,
200265, 200264, 173781, 200266, 17, 10, 17, 314, 220,
19, 200265, 200264, 1428, 200266, 4827, 538, 480, 673,
425, 30, 200265, 198, 200264, 173781, 200266, 17, 9, 1
7, 314, 220, 19, 11, 2684, 472, 220, 17, 10, 17, 0, 20
0265
```

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.

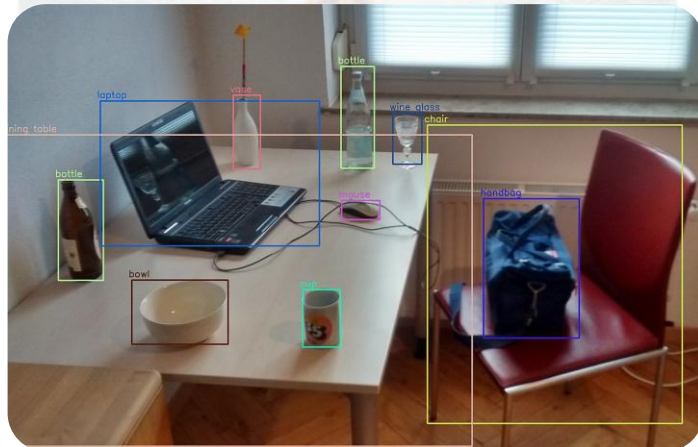
Different Forms of Generative AI

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>

Different Forms of Generative AI

Multimodal LLM (MLLM)



Voice Cloning



**Image, Voice,
Emotion**

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
Elevenlabs	https://elevenlabs.io/	Voice, Language Translation

Generative AI in Business & Marketing

- ❑ AI-powered customer support (chatbots, virtual assistants)
 - ❑ "Watson Assistant" from "IBM", [<https://www.ibm.com/watson/assistant>]
 - ❑ AI-powered virtual assistant that helps businesses automate customer service through chat and voice interactions.
 - ❑ "Drift" from "Drift", [<https://www.drift.com/>]
 - ❑ Conversational AI chatbot that engages website visitors, qualifies leads, and automates responses for customer support.
- ❑ Content creation (social media posts, product descriptions)
 - ❑ "Copy.ai" from "Copy.ai", [<https://www.copy.ai/>]
 - ❑ AI-powered content generation tool that creates engaging social media posts, blog content, and product descriptions.
 - ❑ "Jasper AI" from "Jasper AI", [<https://www.jasper.ai/>]
 - ❑ AI writing assistant that helps marketers generate creative ad copies, SEO-friendly articles, and email content.
- ❑ Personalized ads and targeted marketing campaigns
 - ❑ "Google Ads Smart Bidding" from "Google", [<https://ads.google.com/>]
 - ❑ Uses AI to optimize ad bidding strategies based on user behavior, maximizing conversions and ROI.
 - ❑ "Meta Advantage+ Ads" from "Meta (Facebook)", [<https://www.facebook.com/business/ads>]
 - ❑ AI-driven ad campaign tool that automatically adjusts targeting, budget allocation, and creative elements for better ad performance.

Generative AI in Healthcare

❑ AI-assisted medical diagnosis (AI reading X-rays, MRIs)

❑ Qure.ai from Qure.ai [<https://www.quire.ai/>]

- ❑ Automates the interpretation of X-rays and CT scans to detect abnormalities like lung diseases, brain injuries, and fractures.

❑ Aidoc from Aidoc Medical [<https://www.aidoc.com/home>]

- ❑ AI-powered radiology solution that assists doctors in detecting critical conditions from medical imaging in real time.

❑ Drug discovery and medical research

❑ IBM Watson for Drug Discovery from IBM [<https://www.ibm.com/watson-health/solutions/drug-discovery>]

- ❑ Uses AI to analyze scientific literature, clinical trial data, and molecular interactions to accelerate drug discovery.

❑ Insilico Medicine from Insilico Medicine [<https://insilico.com/>]

- ❑ AI-driven biotech company that designs new drug molecules, identifies targets, and predicts clinical trial outcomes.

❑ Virtual health assistants for patient engagement

❑ Ada Health from Ada Health GmbH [<https://ada.com>]

- ❑ AI-powered symptom checker that provides preliminary diagnoses and health guidance based on user input.

❑ Babylon Health from Babylon Health [<https://www.babylonhealth.com>]

- ❑ AI-driven virtual healthcare platform offering symptom checking, telemedicine consultations, and personalized health assessments.

Generative AI in Finance

❑ Automated Financial Reporting and Analysis

- ❑ "BlackLine" from "BlackLine, Inc." [<https://www.blackline.com/>]
 - ❑ Automates financial close, reconciliation, and reporting, reducing manual effort and improving accuracy in accounting processes.
- ❑ "Workiva" from "Workiva Inc." [<https://www.workiva.com>]
 - ❑ Provides cloud-based financial reporting and compliance solutions that integrate data across departments for real-time reporting.

❑ Fraud Detection and Risk Assessment

- ❑ "Feedzai" from "Feedzai, Inc." [<https://feedzai.com/>]
 - ❑ Uses AI to detect fraudulent transactions, assess risk in real time, and prevent financial crimes in banking and e-commerce.
- ❑ "FICO Falcon Fraud Manager" from "FICO" [<https://www.fico.com/en/products/fico-falcon-fraud-manager>]
 - ❑ AI-driven fraud detection platform that monitors transactions, identifies suspicious activity, and prevents fraud in financial services.

❑ AI-Driven Customer Service for Banking

- ❑ "Kasisto KAI" from "Kasisto" [<https://kasisto.com>]
 - ❑ AI-powered virtual assistant designed for banks to provide customer support, answer queries, and assist with transactions via chat and voice.
- ❑ "Amelia" from "Amelia, an IPsoft Company" [<https://amelia.ai>]
 - ❑ Conversational AI assistant that automates customer interactions in banking, resolving queries and guiding users through financial tasks.

Generative AI in Education & Research

❑ AI-generated study materials and summaries

❑ "Quizlet AI" from Quizlet [<https://quizlet.com>]

❑ Uses AI to generate flashcards, quizzes, and study guides based on textbooks and class notes.

❑ "Scholarcy" from Scholarcy Ltd. [<https://www.scholarcy.com>]

❑ Summarizes research papers, textbooks, and articles into key points and structured notes.

❑ Personalized learning experiences

❑ "Squirrel AI" from Squirrel AI Learning [<https://squirrelai.com>]

❑ AI-driven adaptive learning platform that personalizes study plans based on student strengths and weaknesses.

❑ "Knewton Alta" from Wiley [<https://www.knewton.com/alta/>]

❑ Analyzes student performance and tailors real-time learning materials to improve understanding.

❑ AI-driven research assistance (literature reviews, insights)

❑ "Elicit" from Ought [<https://elicit.org>]

❑ AI research assistant that automates literature reviews, extracting key findings from academic papers.

❑ "Scite" from Scite Inc. [<https://scite.ai>]

❑ Uses AI to analyze citations, providing context and insights for scientific research papers.

Generative AI in Entertainment & Media

❑ AI-Generated Scripts and Stories

- ❑ "Sudowrite" from "Sudowrite", [<https://www.sudowrite.com/>]
 - ❑ AI-powered writing assistant that helps authors generate creative story ideas, complete dialogues, and refine narratives.
- ❑ "ChatGPT" from "OpenAI", [<https://openai.com/chatgpt>]
 - ❑ Generates story plots, dialogue, and entire scripts for books, movies, and games based on prompts.

❑ AI-Assisted Music and Video Production

- ❑ "AIVA" from "AIVA Technologies", [<https://www.aiva.ai/>]
 - ❑ AI-powered music composition tool that creates symphonies, background scores, and personalized melodies for content creators.
- ❑ "Runway ML" from "Runway", [<https://runwayml.com/>]
 - ❑ AI-powered video editing and creative tools that enable real-time style transfers, background removal, and video enhancement.

❑ Deepfake Technology and Ethical Concerns

- ❑ "DeepFaceLab" from "DeepFaceLab Community", [<https://github.com/iperov/DeepFaceLab>]
 - ❑ Open-source deepfake software used for face-swapping and realistic AI-generated videos, often raising ethical concerns regarding misinformation.
- ❑ "Synthesia" from "Synthesia", [<https://www.synthesia.io/>]
 - ❑ AI video creation tool that generates realistic talking avatars, often used for business presentations but also debated for potential misuse in fake content.

Generative AI in Retail & E-commerce

❑ Virtual try-ons and AI fashion recommendations

- ❑ "Zyler" from "Anthropics", [<https://www.zyler.com/>]
 - ❑ AI-powered virtual fitting room that allows users to try on clothes digitally using just a photo.
- ❑ "Vue.ai" from "Mad Street Den", [<https://vue.ai/>]
 - ❑ Uses AI to generate personalized fashion recommendations based on user preferences and past purchases.

❑ AI-powered product descriptions and reviews

- ❑ "Copy.ai" from "Copy.ai", [<https://www.copy.ai/>]
 - ❑ Uses AI to generate engaging and SEO-friendly product descriptions automatically.
- ❑ "Persado" from "Persado", [<https://www.persado.com/>]
 - ❑ AI-powered content generation tool that creates compelling product descriptions and marketing copy.

❑ Automated customer service and support

- ❑ "Drift" from "Drift", [<https://www.drift.com/>]
 - ❑ AI-powered chatbot that engages with customers, answers queries, and qualifies leads in real time.
- ❑ "Freshchat" from "Freshworks", [<https://www.freshworks.com/freshchat/>]
 - ❑ Conversational AI tool for businesses that automates customer support via chatbots and messaging.

Ethical Considerations & Challenges



Ethical Considerations & Challenges



Ethical Considerations & Challenges



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

Future Direction

Generative AI is expected to continue evolving, with several promising areas of exploration:^{10,11}

- Multimodal Generative Models: Combining different types of data, such as images and text, to produce richer, more complex outputs.
- Real-Time Interaction: Generative AI could power more interactive and dynamic dialogue systems, improving human-AI communication.
- Ethical and Fair AI Models: Researchers are focusing on developing fairer models by addressing bias, improving transparency, and enhancing interpretability.
- Energy-Efficient AI: As concerns about the environmental impact of AI grow, the future of generative AI will likely include innovations in model efficiency and sustainability.

**Thanks for
your time**