

How Do I AI?

Supercharging Life & Work with GenAI Tools

NLP from scratch 

Myles Harrison

Monday, May 6th, 2024



Housekeeping



Be professional



Respect the space

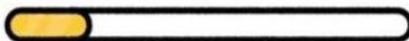
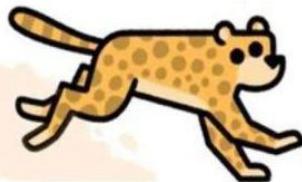


This meeting will not be recorded

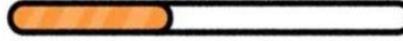
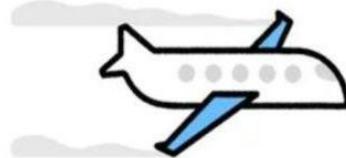


Please hold your questions until the end

THE FASTEST THINGS ON EARTH



CHEETAH



AIRPLANE



SPEED OF LIGHT



PEOPLE BECOMING
EXPERTS IN AI

Who am I?

- Data Scientist
- Career consultant
(SapientNitro, PwC,
Accenture)
- Trainer
- Human



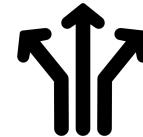
Manifesto



Knowledge is only valuable if it is useful.



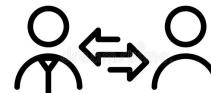
The best way to learn is by doing.



Learning is a non-linear process.



Learning is exploration, not a journey.



Teaching and learning are complementary.

GENERATIVE AI

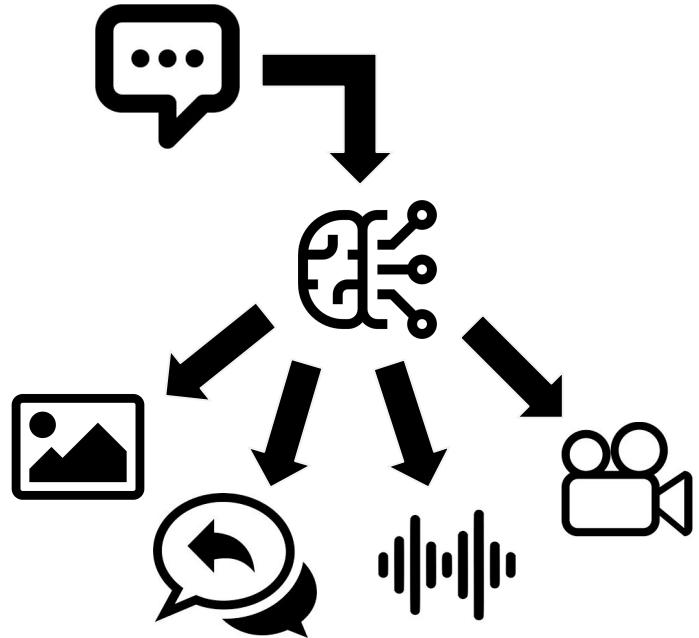
SO HOT RIGHT NOW

What is Generative AI?

Generative Artificial Intelligence (or GenAI) refers to the branch of artificial intelligence that deals with models and systems that can produce diverse outputs in the form of text, images, videos, or other data.

GenAI models learn from input **training data** and can generate new data that has similar characteristics, following the patterns and structures it has learned.

GenAI models are almost exclusively large language models (LLMs) a specific type of machine learning models with a very large number of parameters that are trained on massive datasets.



ChatGPT



Capabilities

*Remembers what user said
earlier in the conversation*

*Allows user to provide follow-
up corrections*

*Trained to decline inappropriate
requests*

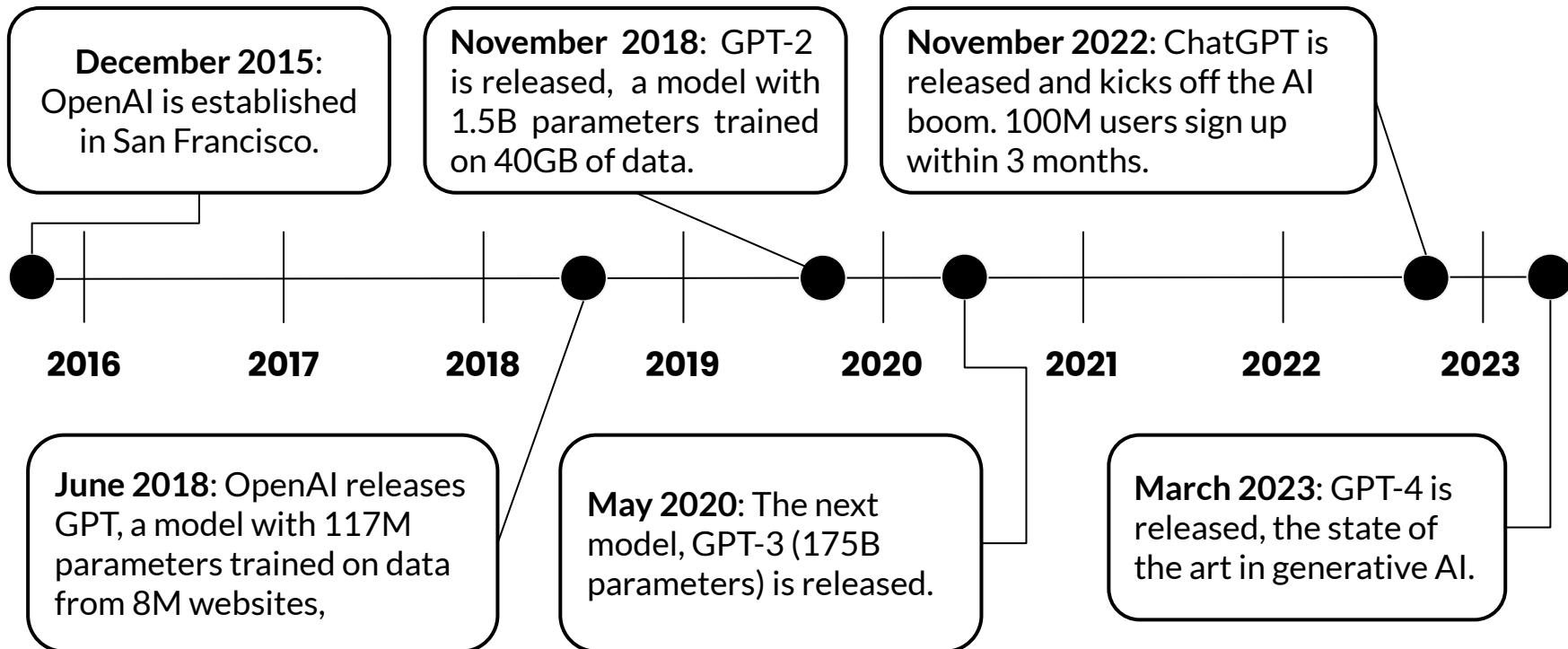
OpenAI & ChatGPT

- Founded in December 2015 as a non-profit artificial intelligence research organization
- Initially backed by a prominent technology entrepreneurs and investors, including Elon Musk, Sam Altman, and Peter Thiel.
- Mission to “ensure that artificial general intelligence (AGI)—by which we mean highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity.”¹
- Transitioned to for-profit (subsidiary) in 2019 with the goal of attracting more funding and resources to accelerate its research and development efforts
- Kicked off generative AI craze (AI boom) with release of ChatGPT in November of 2022



1, OpenAI Charter, 2018

OpenAI Timeline



GPT - The Household name of LLMs



GPT-1

115M parameters



GPT-2

1.5B parameters



GPT-3

175B parameters



GPT-4

1.76T parameters

Model & Size



Toronto Book Corpus
~800M words



WebText (8M docs, 40GB)



CommonCrawl, Books 1+2,
WebText, Wikipedia (~45 TB?)

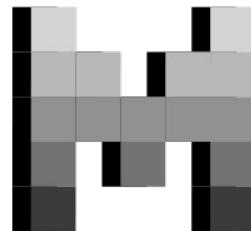


Training Data

The Big (Enterprise) Model Providers



OpenAI



**MISTRAL
AI_**

ANTHROPIC



Google



databricks

My God, it's full of APIs.

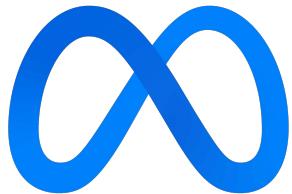




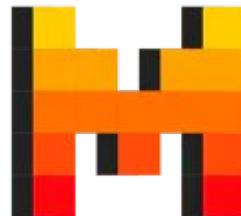
CHRIS. MADDEN

NLP from scratch

Open Models - The Big 4



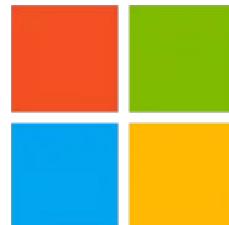
LlaMa



Mistral



Gemma



Phi

Open Models - The Menagerie



A woman with long, dark hair is shown from the chest up, looking off-camera with a serious expression. Her hair is blowing in the wind, creating a sense of movement. She is wearing a dark, sequined top.

AMERICA'S NEXT
topmodel



MY

You

Tell me about the musician Johnny Cash, and how he played a gig in Miami, Florida in the late 1980s with Van Halen and Ozzy Osborne.

**ChatGPT**

Memory updated

Johnny Cash was a legendary American singer, songwriter, and guitarist known for his deep, distinctive voice and his fusion of country, rockabilly, and gospel music. He rose to fame in the late 1950s and became one of the best-selling music artists of all time. Cash's career spanned several decades, and he was known for his rebellious image and heartfelt songs that often reflected themes of love, sorrow, and redemption.

In the late 1980s, Johnny Cash indeed performed a memorable gig in Miami, Florida, alongside rock legends Van Halen and Ozzy Osbourne. This concert was part of the "Monsters of Rock" tour, a series of music festivals featuring multiple rock bands. The tour was known for its large-scale productions and attracted massive crowds of fans.

Cash's inclusion in the lineup was somewhat unconventional given his country music roots, but it added a unique flavor to the tour. His performance showcased his versatility and appeal across different music genres, demonstrating his enduring influence and popularity as a musician.

"Hallucinations"

Multimodal Models

Multimodal models refers to model which not only work with text, but other media such as **images, sound, or even video**.

One of the most fundamental and popular multimodal models is the **Stable Diffusion model** which was released initially in August of 2022 by a research group, and later being developed by the company Stability AI.

Many other multimodal LLMs / services are based on some version of a diffusion model, notable ones being from OpenAI (DALL-E), Google (Imagen, superseded by Gemini) and Midjourney.

Other multimodal model services can generate audio (e.g. Suno AI) or even video (Pika Labs, Runway ML, Pika, OpenAI's yet-to-be-released SORA).

stability.ai

DALL·E 3



Midjourney

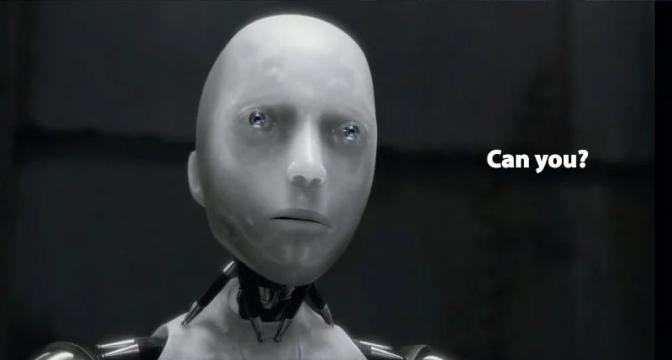
Pika **runway**



NLP from scratch



Can an AI draw hands?



Can you?





OUR T2 REMAKE

OFFICIAL TRAILER



Putting GenAI to Work



Prompting

With traditional machine learning models, data are taken as inputs, and predictions are generated as output for each observation.

For generative AI language models, the model is given a **prompt** - a piece of input text, usually in the form of a question or instruction - to produce a desired output: a response of text in return, or something like an image or video in the case of “multimodal” modals.

Prompting involves structuring your requirements into a format that effectively communicates the desired outcomes from the model, ensuring a result in alignment with what is needed.



“ IN A MODERN WORKPLACE, WE ONLY HAVE TO BE ABLE TO DO 3 THINGS: TYPE, TALK ON THE PHONE, AND GO TO MEETINGS. ”

- One of my previous consulting clients (paraphrased)



GitHub Copilot

Launched June 2021

<https://github.com/features/copilot>

**Github
Copilot**

NLP from scratch 

Microsoft Copilot

- AI-powered assistant developed by Microsoft and based upon GPT, available at copilot.microsoft.com and launched on February 7th, 2023.
- Chat interface as with ChatGPT and other “AI assistants”
- **Web search enabled** and provides references to answers
- Can **generate images** as part of Microsoft Designer
- Plugins for services like Kayak, Instacart, Shopify, OpenTable
- Being broadly **integrated into Microsoft products** as a whole: Windows 11, Bing Search, Edge Browser, Office 365, etc.





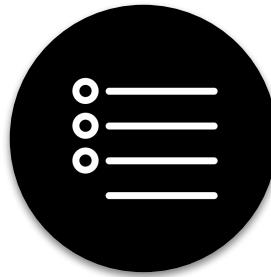
Microsoft and OpenAI extend partnership

Jan 23, 2023 | [Microsoft Corporate Blogs](#)

Examples of everyday GenAI assisted tasks



Composing an
email



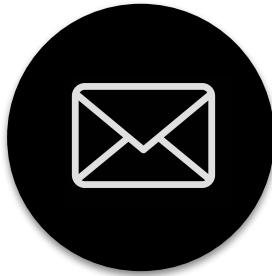
Generating an
outline for a
meeting or
presentation



Summarizing or
transcribing a
meeting

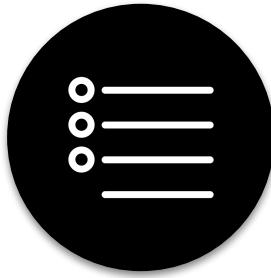


Content
creation and
brainstorming



Composing an email

Write an email asking a valued client to reschedule a meeting, offering to meet at their convenience the following week with availabilities of Monday, Wednesday, and Friday afternoons.



**Generating an
outline for a
meeting or
presentation**

Create an outline for a free workshop entitled “How Do I AI?”, introducing beginners to generative AI tools and how they can be leveraged for productivity in everyday life and work.



Summarizing
or transcribing
a meeting





Summarizing
or transcribing
a meeting



The screenshot shows a meeting transcript from the Otter.ai app. At the top, it says "Olli Team Wins Meeting". There is a search bar and a "OtterAI Chat" section with four user icons.

Sven Kramer 0:36
Welcome everyone to this month's Team Wins meeting. We had a great March, delivering a record number of new customers from conventions. Hey Charlie, can you share the Q1 Sales update?

Charlie Tse 2:40
Sure, thanks Sven. As you can see, we hit \$500K in Sales in Q1, which is a new record for our company. We crushed it as a team!

OtterPilot
Based on the meeting so far:

- Charlie announced a new team member named Brenda.
- Sven and Charlie discussed Q1 record sales driven by new customers
- The team ideated on Q2 strategies.



**Content
creation and
brainstorming**

Generate a stark black and white 4K image with ambient lighting of robot artificial intelligence assistants helping humans and people work in an office, people, man, woman, office worker, dress shirts, suits





NLP from scratch

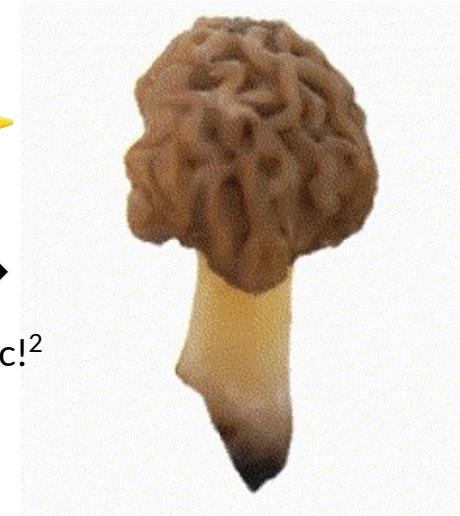
Image to 3D asset generation



AI Magic!¹



AI Magic!²



1. *Highly Accurate Dichotomous Image Segmentation*, Qin et. al (2002):
<https://github.com/xuebingin/DIS>

2. *TripoSR: Fast 3D Object Reconstruction from a Single Image*, VAST AI & Stability AI, Tochilkin et al. (2024)
<https://github.com/VAST-AI-Research/TripoSR>

Browser-based AI assistants

The image is a collage of three screenshots illustrating browser-based AI assistants:

- Copilot:** A screenshot of the Microsoft Copilot interface. It features a large image of a person in a yellow hoodie running in the rain, with text overlay: "Compose a pop song about adventures in the summer rain". Below the image, it says "Possible with Copilot in" and "page featuring a curious discovering a galaxy inside a...". At the bottom, there's a text input field with "anything..." and a large blue circular logo.
- Brave Leo:** A screenshot of the Brave browser's AI feature. The title is "Brave Leo". It shows a message from PCWorld: "This one Brave feature sold me on the whole browser | PCWorld". Below it is a button labeled "Summarize this page". A larger message from Brave Leo reads: "Hey there! I'm here to help you with any questions you have about the article you provided. The article talks about the Brave browser and its feature of blocking cookie consent notices by default. It also mentions that you can add a similar". At the bottom, there's a message: "Ask me anything else or go to other browsers like Chrome, Firefox, or Edge if you have specific questions or need further clarification, I'm here to help! Don't hesitate to ask!" and a "Suggested follow-ups" section.
- Opera AI:** A screenshot of the Opera AI interface titled "New chat". It has a header "CHOOSE A LOCAL AI MODEL ^". It lists four models:
 - Mistral Instruct
 - Vicuna 13b-Q4_0
 - Llama 7b-Text
 - Gemma 7b-Instruct

Goblin Tools

- A collection of **small, simple, single-task tools** backed by OpenAI's API, mostly designed to help neurodivergent people with tasks they find overwhelming or difficult (but are handy for everyone!)
- Desktop / web version is **free and without ads!** Pay versions available on Android & iOS at low price to cover operating and maintenance costs such that desktop site can remain free.
- **Magic To-Do** breaks down large tasks into smaller, more manageable chunks
- Also tools for tone checking, rewording, task length estimation, etc.



goblin.tools

Ollama

Run LLMs locally, with models such as LLaMa 2, Mistral 7B, Phi, Orca, and many more!

Local web server with REST API, available in a Docker container

<https://ollama.com/blog/ollama-is-now-available-as-an-official-docker-image>

Libraries in Python, Java, C, Rust, Ruby, Langchain, etc.

Multimodal support (Ollama Vision) with LLaVa added for image interrogation, Object detection, text recognition, etc. added 2024/02/02

<https://ollama.ai/blog/vision-models>



Get up and running with large language models, locally.

Run Llama 2, Code Llama, and other models.
Customize and create your own.

Download ↓

Available for macOS & Linux
Windows coming soon



Workshops



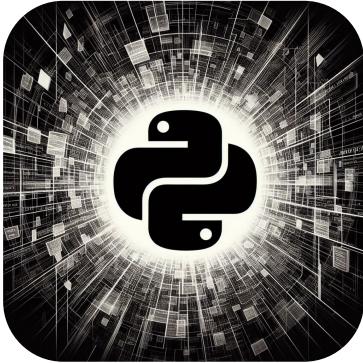
Generative Text Models & Fine-tuning LLMs

- Intro to Hugging Face
- LLMs for generative text
- Fine-tuning models
- PEFT (LoRA) and quantization



Building GenAI Apps with OpenAI and GPT

- Intro to APIs and OpenAI
- Working with the OpenAI API
- Setting up a dev environment
- Build a streaming chat app
- Open source alternatives to GPT



Intro to Python & Natural Language Processing

- Intro to NLP
- Intro to Python
- Working with text in Python
- Manipulating text in Pandas

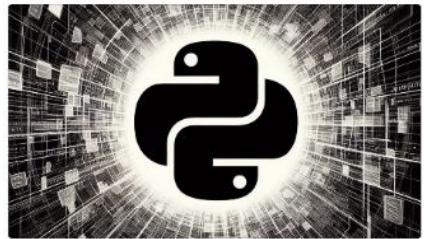


GenAI for Work (1 hr)

- Introduction to GenAI
- Generative AI landscape
- Everyday use cases for GenAI tools
- Prompting and prompt engineering
- Productivity with GenAI

Sign up at [nlpfromscratch.com/training!](https://nlpfromscratch.com/training)

Available - 4 Services

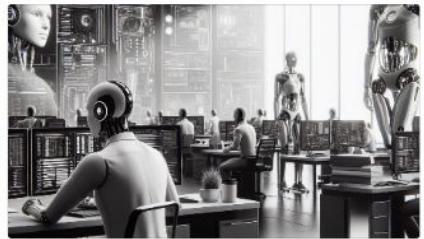


Introduction to Python and NLP

CA\$56.50

⌚ 3h 📝 1/30

Register Now



Generative AI for Work

CA\$16.95

⌚ 1h 📝 1

Register Now



Building Generative AI Application...

CA\$56.50



Generative Text Models and Fine...

CA\$56.50

⌚ 3h 📝 1/30



IRL Meetup Communities (on Meetup.com)



Toronto NLG Meetup
Webinars, In-person
workshops, Coffee
Networking ☕



MLTO



Data Practitioner Drinks (DPD)

Networking group for data analytics,
ML, and AI professionals (and others)

Machine Learning TO Meetup
(MLTO)

Industry talks & networking
events, restarting Summer 2024

More Resources

[www.github.com/
nlpfromscratch/](https://github.com/nlpfromscratch/)



NLP from scratch
nlpfromscratch

Learn natural language processing. From A to Z. From zero to hero. Fast.

[Edit profile](#)

10 followers · 0 following

<https://www.nlpfromscratch.com>

nlpfromscratch / README.md

NLP from

Popular repositories

[nlp-llms-resources](#)

Master list of curated resources on NLP and LLMs

61 stars 12 forks

[webinars](#)

Jupyter Notebook 1 star

I would value your feedback.





Thanks!



myles@nlpfromscratch.com



nlpfromscratch.com



linkedin.com/in/mylesharrison