

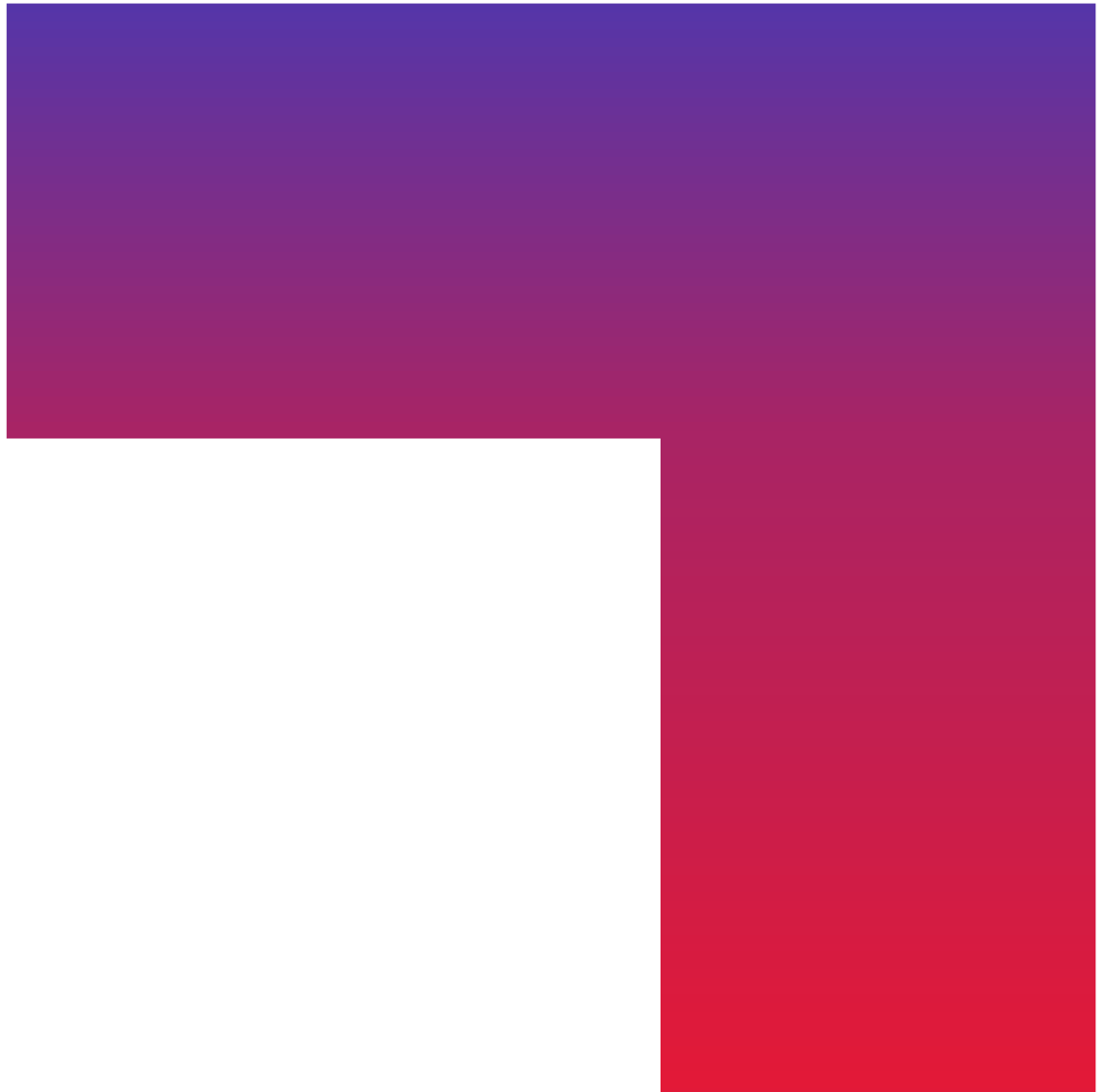
ChatGPT

& Generative AI

...A match made in Artificial Heaven

Canada Health Infoway
Health Analytics Forum

CGI



ChatGPT HYSTERIA.

ChatGPT has propelled AI into the spotlight. Companies are scrambling to get in front of it and avoid becoming redundant

1 M -> 5 DAYS

...Fastest company ever

CODE RED

Google declared code red, forced to release own ChatGPT-like model, BARD

\$10B

OpenAI-as-a-Service on Azure.
Bing search revived.

Microsoft invests into OpenAI



First of the big 5 to announce alliance for next gen contact centres, ad copy, financial advisory

ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users



* one million backers ** one million nights booked *** one million downloads

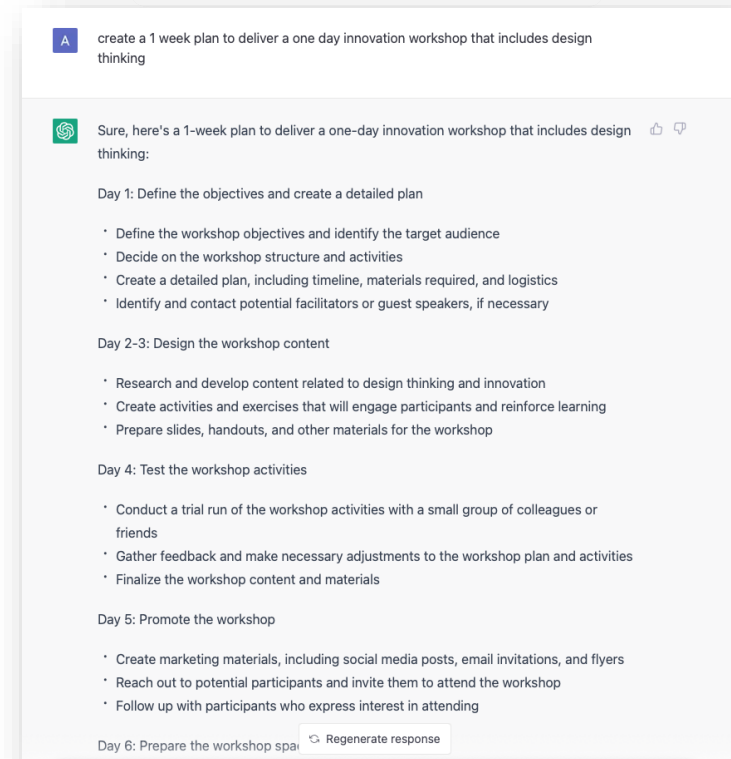
Source: Company announcements via Business Insider/LinkedIn



“Release a chatbot, fast!”

- OpenAI was conceived by tech leaders including Sam Altman, Peter Thiel, Reid Hoffman and Elon Musk over dinner talking about the future of AI
- Started in 2015 as a nonprofit research lab with plans to open-source code, pivoted to become for-profit in 2019 to better secure funding and scale
- ChatGPT was a complete left turn.
 - Company was on track to release GPT-4, with chatbots in early 2023.
 - Fearing loss of first-mover advantage, executives changed course, release quickly based on an old model. ChatGPT was born 13 days later.

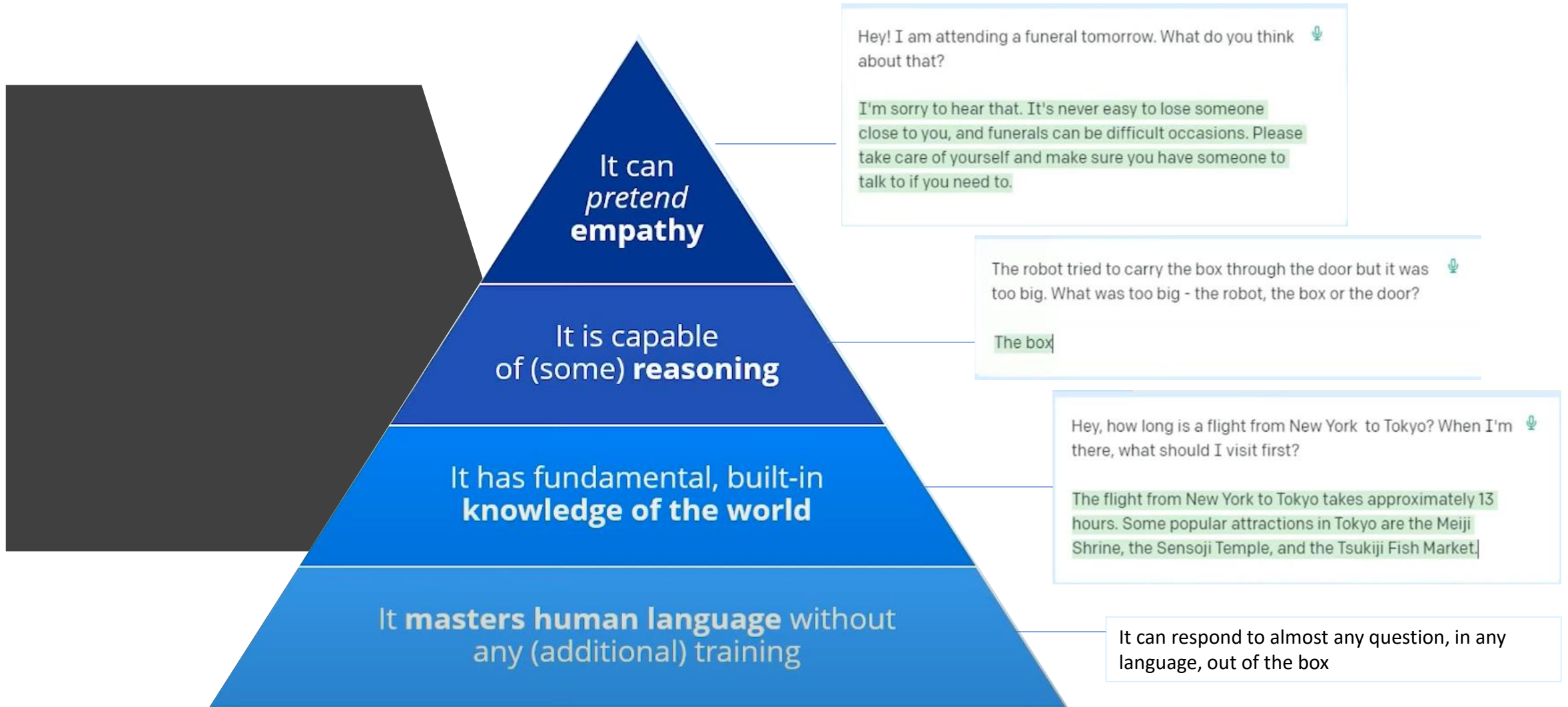
GPT = Generative Pre-trained Transformers



“Benefits for humankind could be so unbelievably good that it’s hard for me to even imagine. Worst-case scenario, A.I. could kill us all. ”

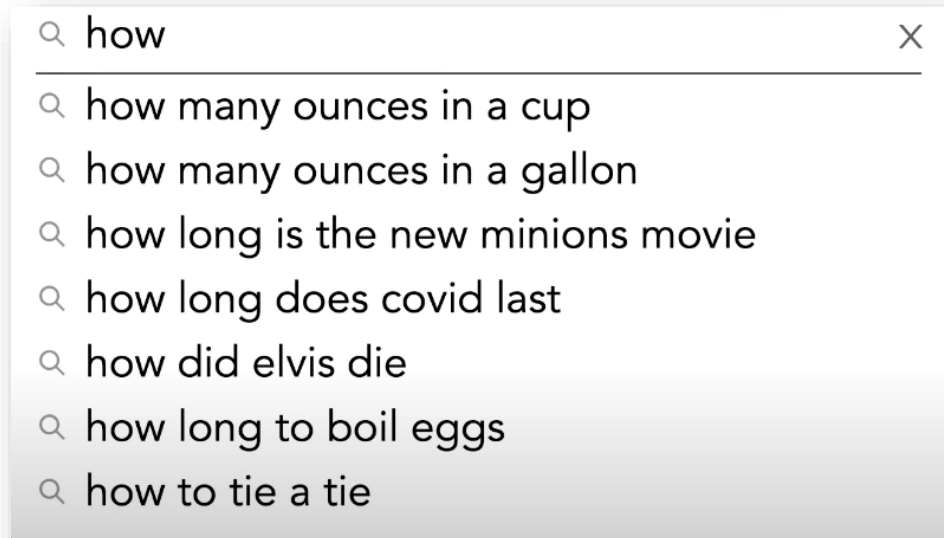
Sam Altman, CEO OpenAI

Why is ChatGPT so compelling?



How does ChatGPT work?

ChatGPT is an example of a Large Language Model - think of autocomplete on steroids



- There are ~ 40,000 commonly used words in the English language
- By analyzing a few million books and a few hundred billion words – we can estimate how common each word and word pairings are (60 trillion word combinations)
- To keep it manageable a neural network model is used to approximate the likely next words in the sequence
- As model, it's very close, but doesn't always get it right.
- Parameters such as Attention Weights, layers, context window size are added to make the model more rich, natural sounding like.

ChatGPT does not necessarily repeat content from the Internet. It uses Internet data to generate a giant list of permutations of the next best word in response to a specific question, hence the term "**Generative AI**".

What are Large Language Models (LLMs)?

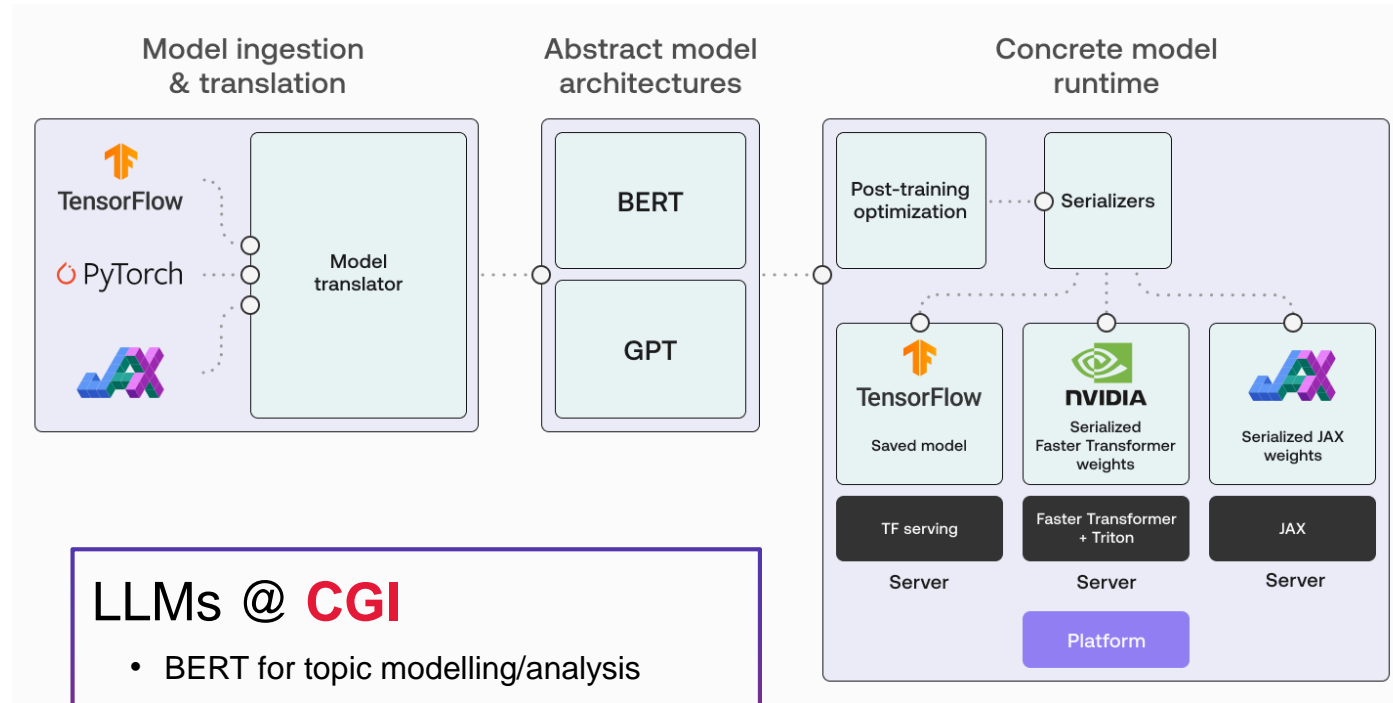
LLMs are a subset of models within the Natural Language Processing (NLP) area of data science.

- LLMs take advantage of powerful neural network frameworks such as TensorFlow (Google) or PyTorch (Meta)
- LLMs are made possible by new NLP architectures called Transformers

LLM Steps:

- 1) Pre-trained models are ingested, usually based on all text on the internet over a period of time
- 2) Trained models from step 1 are fed into transformer model architectures such as BERT or GPT
- 3) Queries are passed to the model and predictions/answers are served and exported to the appropriate end framework.

Cohere.ai TIF Framework



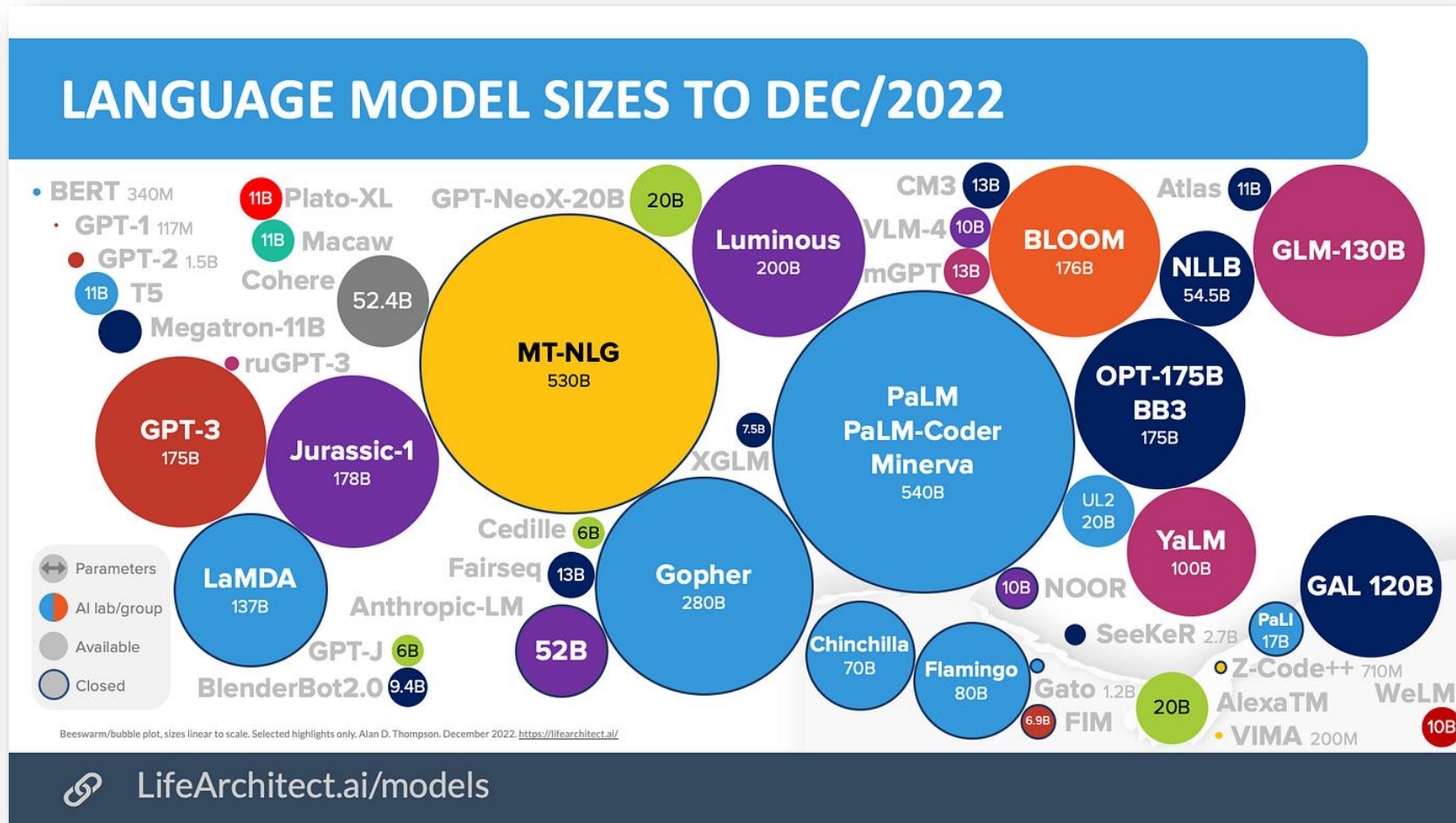
LLMs @ CGI

- BERT for topic modelling/analysis
- HuggingFace for pre-trained models
- Used in NLP projects for text analysis
- Used in Chatbot development and implementation

ChatGPT is based on Large Language Models – and there are many

The number of parameters is a proxy for model capability

Parameters are how the model represents patterns and relationships in the language data it is trained on.



Can LLM's such as **ChatGPT** be incorporated into solutions for our clients?

...possibly, but there are some prickly issues:

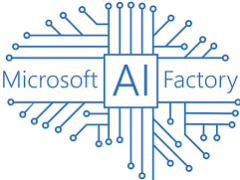
... out-of-date, data residency isn't clear, brand credibility may be at risk, provenance of answers is unclear...



<https://wp.technologyreview.com/wp-content/uploads/2023/02/AI-liar-search1.jpeg?resize=1272,714>



RESPONSIBLE AI GUIDELINES



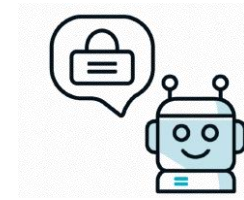
Canada

Bill 27 Digital Charter Implementation Act

CGI Ethical AI Framework

Privacy Protection & Security	Fairness	Transparency	Accountability
Access	Non - Bias	Data Provenance	Interpretation
Data Use	Inclusivity	Auditability	Explainability
<ul style="list-style-type: none"> Ensure data is secure and accessed by approved resources only Ensure privacy of users is never compromised through breaches or findings 	<ul style="list-style-type: none"> Ensure data and models are fair and not affected by human bias Ensure all population is fairly represented 	<ul style="list-style-type: none"> Ensure models and methodologies are visible to all Ensure provenance (i.e. source) of data is known Ensure, when needed, models can be audited 	<ul style="list-style-type: none"> Ensure data scientists are held accountable in how they interpret data sets and design questions Ensure accountability in how data and findings are explained

Potential Use Cases for **ChatGPT** (to name a few)



Government

- Automatically summarize complex government reports or legislation
- Consolidate complex data
- Automatically translate government documents into different languages
- Interpret sentiment from citizen social feeds
- Locate services

Utilities

- Frequently asked questions
- Providing personalized service recommendations
- Billing enquiries
- Tracking status of services

Retail

- Generate product descriptions, recommend product combinations
- Frequently asked questions
- Personalize recommendations to customers based on their past purchases.
- Order tracking

Tech

- Write and optimize computer code
- Structure data
- Debug code

HR

- Pre-screen resumes
- Generate interview questions
- Generate interview responses
- Respond to employee enquiries
- Navigate internal processes
- Assess performance
- Flag poor sentiment, brand vulnerabilities

Healthcare

- Wayfinding
- Distill complex medical information to improve health literacy
- Automate administrative tasks, patient reminders
- Handle customer inquiries and perform diagnostics

Contact Centres

- Analyze, optimize, supplement, and ultimately replace customer service and call-center agents
- Generate responses to complaints

Oil & Gas

- Anomaly detection
- Predictive maintenance
- Operations optimization - resource allocation
- Reservoir Modeling – training on geological data to predict reservoir behaviour



CGI is a Leader in Conversational AI

RECENT PROJECTS



COVID-19

- Nurse contact centre quickly became overwhelmed
- Website augment, navigational aid



COVID-19 Verifier

- Next gen conversational agent
- Supports COVID Vaccination Verification



Nurse Agent

- 17% improvement in time to resolution
- Frees up triage nurses to focus on complex, higher need cases



Student Portal

- Next gen conversational AI platform – LivePerson
- Contact centre solution



My Health Records

- Frequently asked questions



Service Desk

- Service Desk Support
- Live agent (future)



Citizen Support

- Wayfinding, Frequently Asked Questions
- Live agent (future)



Employee HR

- Frequently Asked Questions











Provider Registration









- Physician, Nurse Provider Registration
- Live agent (future)

A contrast worth noting – Generative vs Conversational AI

ChatGPT (GAI)

 Generic	 Use – case specific
 Can go 'Off Track'	 Stays on Task
 No System Integrations	 System Integrations
 No Channels	 Channel Integrations

Typical Chatbots (CAI)

 Feels Rigid	 Very Flexible
 Pre-Defined	 On the Fly
 High Effort	 Low Effort
 Not Human-Like	 Human Like

A contrast worth noting – GAI vs CAI

ChatGPT

Typical Chatbots



Use – case specific



Very Flexible



Stays on Task



On the Fly



System Integrations



Low Effort



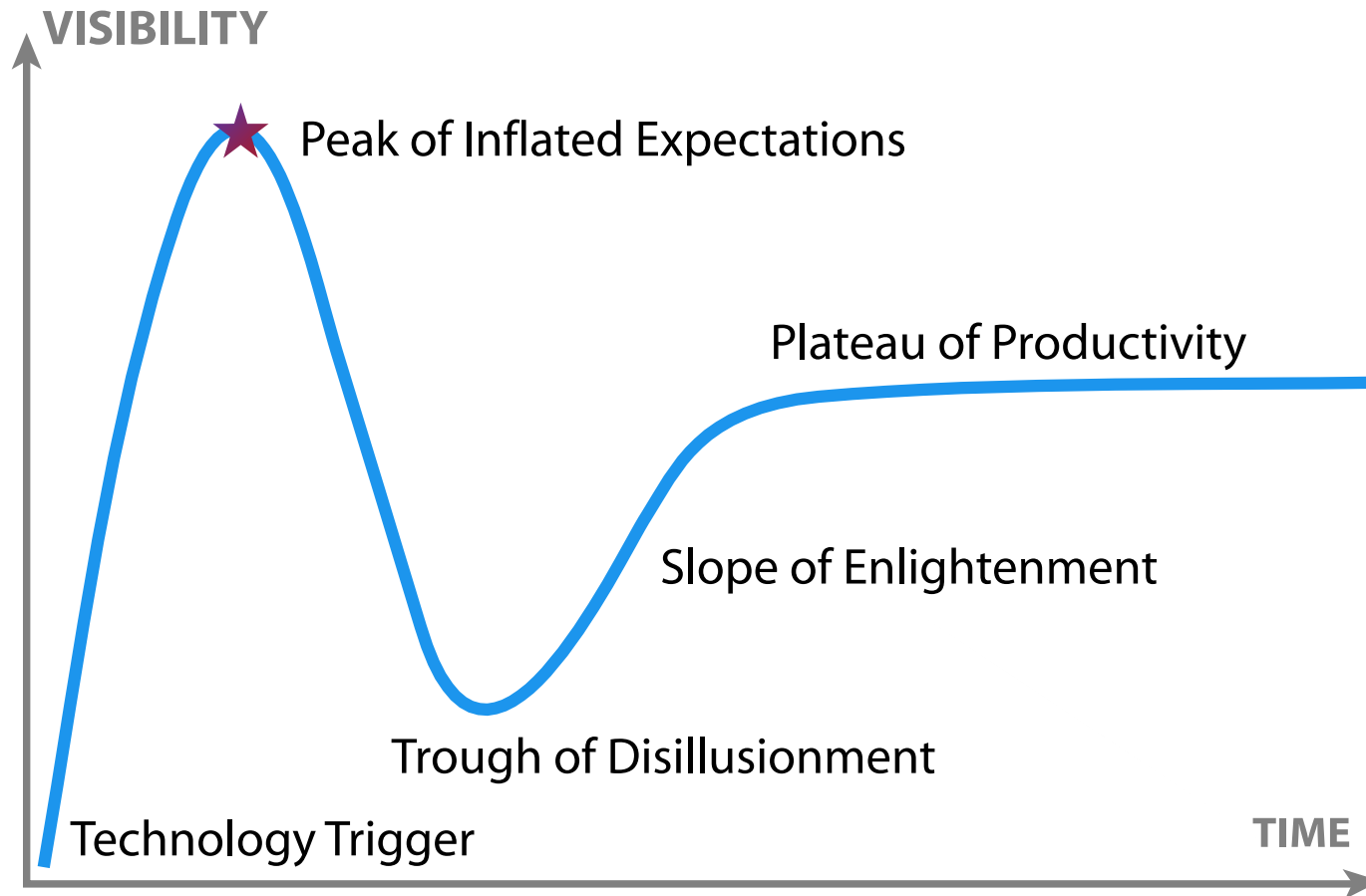
Channel Integrations



Human Like

ChatGPT Future of Generative AI

OpenAI released GPT-4 – able to process text, images, video, or audio ('multi-modal').



Companies using LLM's



Duolingo

Explain answers, clarify with examples. Students can role-play real-life conversations with characters from the app.



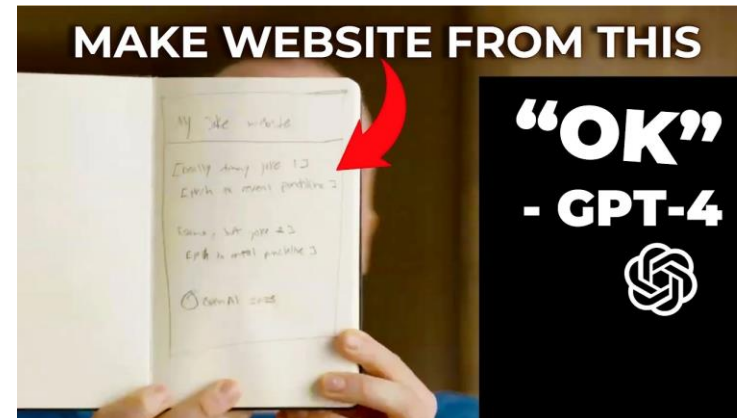
Uses GPT4 to identify, interpret and describe images for people who are hard-of-vision



Generate emails, summarize texts and coordinate calendars



Analyze massive amounts of data in real time and detect patterns in language and sentiment that could indicate fraud



Upload an image, construct the html...

fully functioning website from a napkin!

ChatGPT Demo and Examples

Prompts to test

Autogenerate a project plan

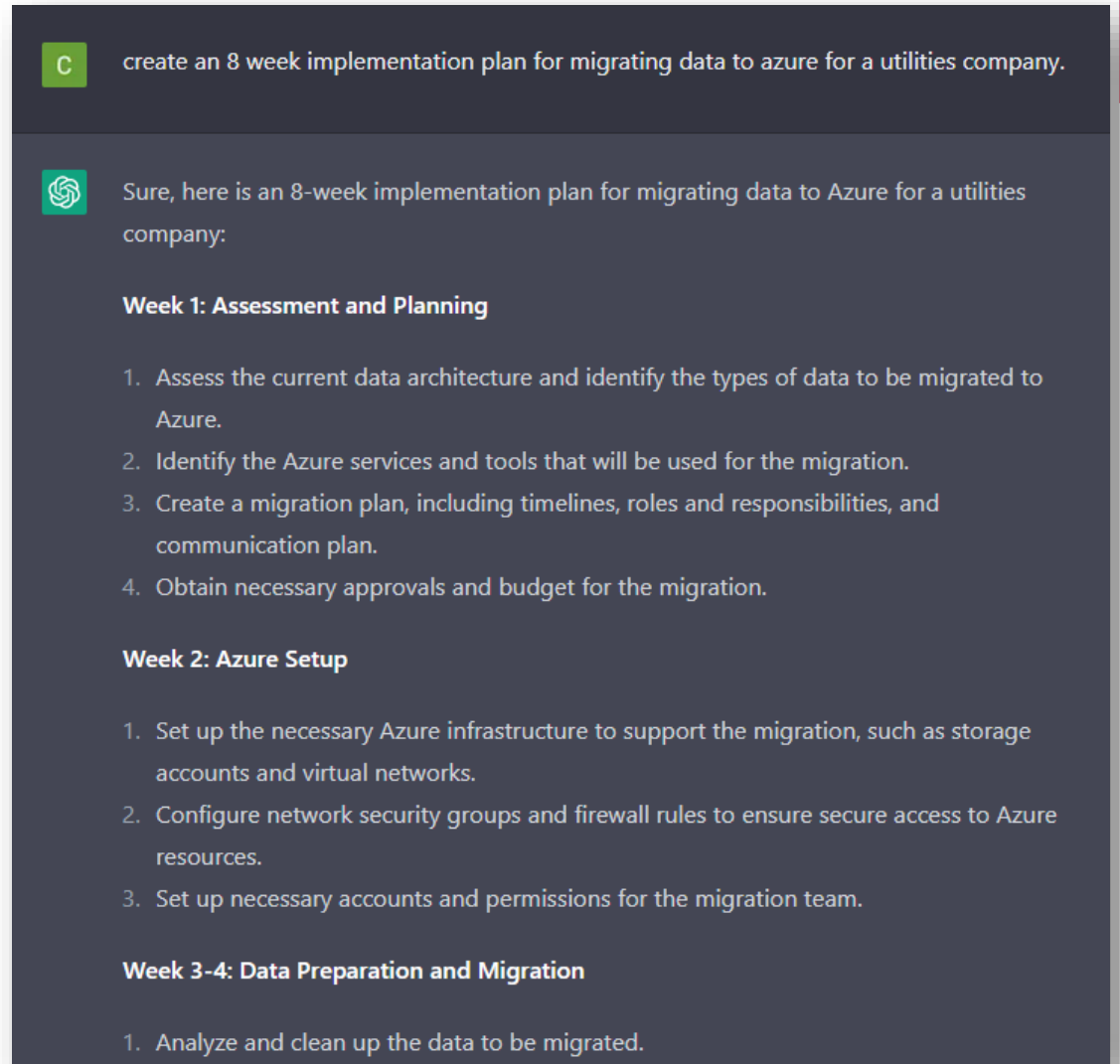
“Create an 8 week implementation plan for migrating data to azure for a utilities company.”

Write a Python program


“Write me a Python program that uses a large model to classify text using Spacy”

Specific proposal content generation

“Develop a proposal for a CIO of an energy company to respond to their request for a new oil & gas accounting system hosted on Microsoft Azure”



C create an 8 week implementation plan for migrating data to azure for a utilities company.

 Sure, here is an 8-week implementation plan for migrating data to Azure for a utilities company:

Week 1: Assessment and Planning

1. Assess the current data architecture and identify the types of data to be migrated to Azure.
2. Identify the Azure services and tools that will be used for the migration.
3. Create a migration plan, including timelines, roles and responsibilities, and communication plan.
4. Obtain necessary approvals and budget for the migration.

Week 2: Azure Setup

1. Set up the necessary Azure infrastructure to support the migration, such as storage accounts and virtual networks.
2. Configure network security groups and firewall rules to ensure secure access to Azure resources.
3. Set up necessary accounts and permissions for the migration team.

Week 3-4: Data Preparation and Migration

1. Analyze and clean up the data to be migrated.

How to get involved in LLMs

1

- Utilize ChatGPT & Embeddings API
- Prompt engineering and fine-tuning available
- Fine-tuning available on custom datasets
- Costs occur on a per-token basis for API requests
- Processing occurs on OpenAI infrastructure

Utilize Existing ChatGPT Models & Infrastructure

2

- Utilize base models of open-source LLM initiatives
- Open-source models are less capable
- Requires local hardware or cloud-based environments
- Requires heavy computing resources to scale
- Models can be fine-tuned on custom datasets

Utilize Open-source Models on Local Infrastructure or Cloud Environments

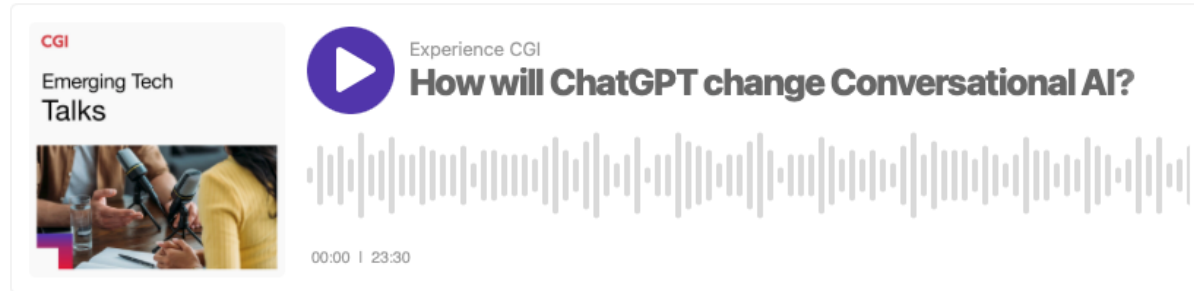
3

- Custom instances of ChatGPT will likely be released as a SaaS-type offering
- Allowing organizations to run their own ChatGPT separate from the public model
- No current dates set for launch of such a service
- Pricing is unknown

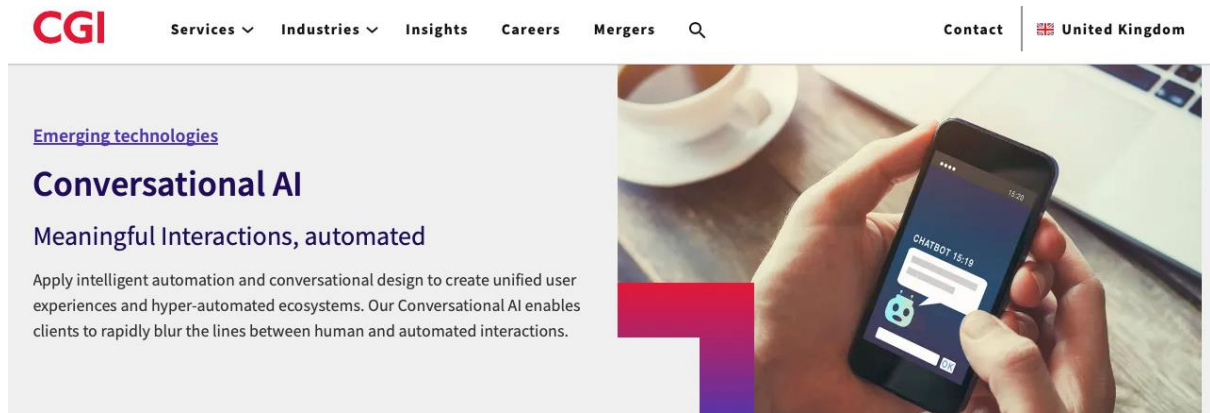
Utilize Cloud-based LLMaaS Offerings

FUTURE

More Resources:



<https://www.cgi.com/uk/en-gb/podcast/conversational-ai/how-will-chatgpt-change-conversational-ai>



<https://www.cgi.com/uk/en-gb/emerging-technologies/conversational-ai>

Additional Slides

What sets **ChatGPT** apart?

ChatGPT uses the transformer architecture GPT3.
Building on prior neural network frameworks used in NLP.

Training

- ChatGPT is trained on a massive amount of data
 - 570 GB of text from web pages and books. Over 300 billion words.
 - Almost the entire internet from 2011 to 2021

Output Accuracy

- ChatGPT is capable of answering complex and specific questions with impressive accuracy
- This includes answering queries with multiple questions and specific requirements, or writing/troubleshooting computer code.

Personality

- Answers are delivered in a very human-like way
- Often indistinguishable from human-written text

