- A.I.

- New to the scene

6/2/2023

2023 IMPACT OF AI

TODAY's AGENDA

- Introduction

- Understanding

- Risks and Costs

- Timeline

- Summary

Jorge Pereira
Modern Workforce Solutions Principal
Dell Technologies Services

INTRODUCTION – WHY ARE HERE?

Not an overnight thing

Artificial Intelligence field has been around since 1935

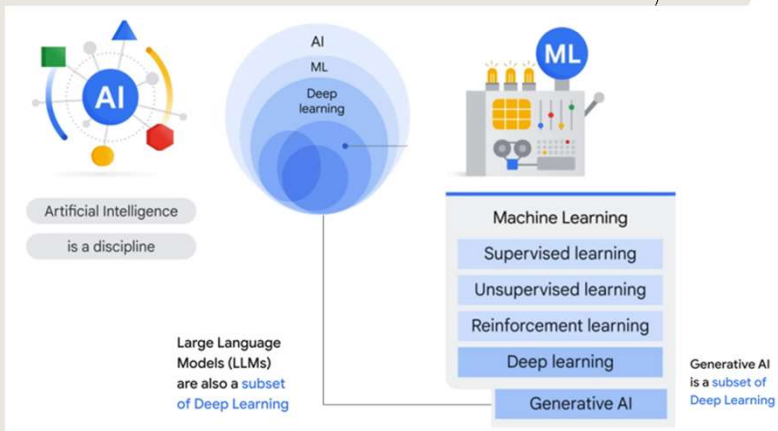**ChatGPT was released to the public November 30, 2022**;

- It achieved 100M users in January 2023

Welcome to ChatGPT

WHAT IS ARTIFICIAL INTELLIGENCE?

ARTIFICIAL INTELLIGENCE IS A MACHINE'S ABILITY TO PERFORM THE COGNITIVE FUNCTIONS WE USUALLY ASSOCIATE WITH HUMAN MINDS.
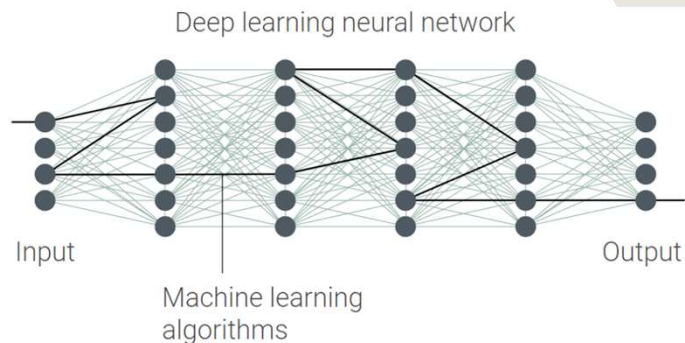
At its core, AI relies on advanced algorithms to perform two tasks:
1. Understand and respond to user queries: **language processing** and
2. Learn, catalog and process data: **machine learning**.

AI will / is embedded in everyday life, business, government, medicine and more.

DISSECTING CHATGPT

- **Training Data Sets**
- **Machine learning**
- **Natural language processing (NLP)**
- **Natural language models (NLMs)**
- **Deep Neural Network (DNN)**
- **Tokens**
- **Transformers**
- **APIs & Applications**

ChatGPT (is a web-based application that is built on a family of large language models (LLMs) collectively known as GPT-3

- ChatGPT was created with one main objective – to predict the next word in a sentence, based on what's typically happened in the gigabytes of text data that it's been trained on.
- ChatGPT has been trained on a vast amount of text covering a huge range of subjects, so its possibilities are nearly endless.

Bottom line: It is all about "educated guessing" ☺

ChatGPT stands for "Chat Generative Pre-trained Transformer"

- On March 14,2023, OpenAI announced that its next-gen language model, GPT-4, was available to developers and ChatGPT Plus subscribers – with Microsoft confirming that the new Bing is already running on GPT-4. The big change from GPT-3.5 is that the language model is multimodal, which means it can process both text and images.

Transformers

- A type of neural network developed by Google in 2017 that is particularly well-suited to natural language processing tasks, like answering questions or generating text.

ChatGPT APIs (and plugins) now available (Everywhere!)

- Microsoft announced CoPilot and Microsoft 365 products lines will integrate with OpenAI
- Copilot combines the power of large language models (LLMs) with your data in the Microsoft Graph—your calendar, emails, chats, documents, meetings, and more—and the Microsoft 365 apps

Deep learning neural network

Input

Machine learning algorithms

Output

CONSTANTLY LEARNING / UPDATING

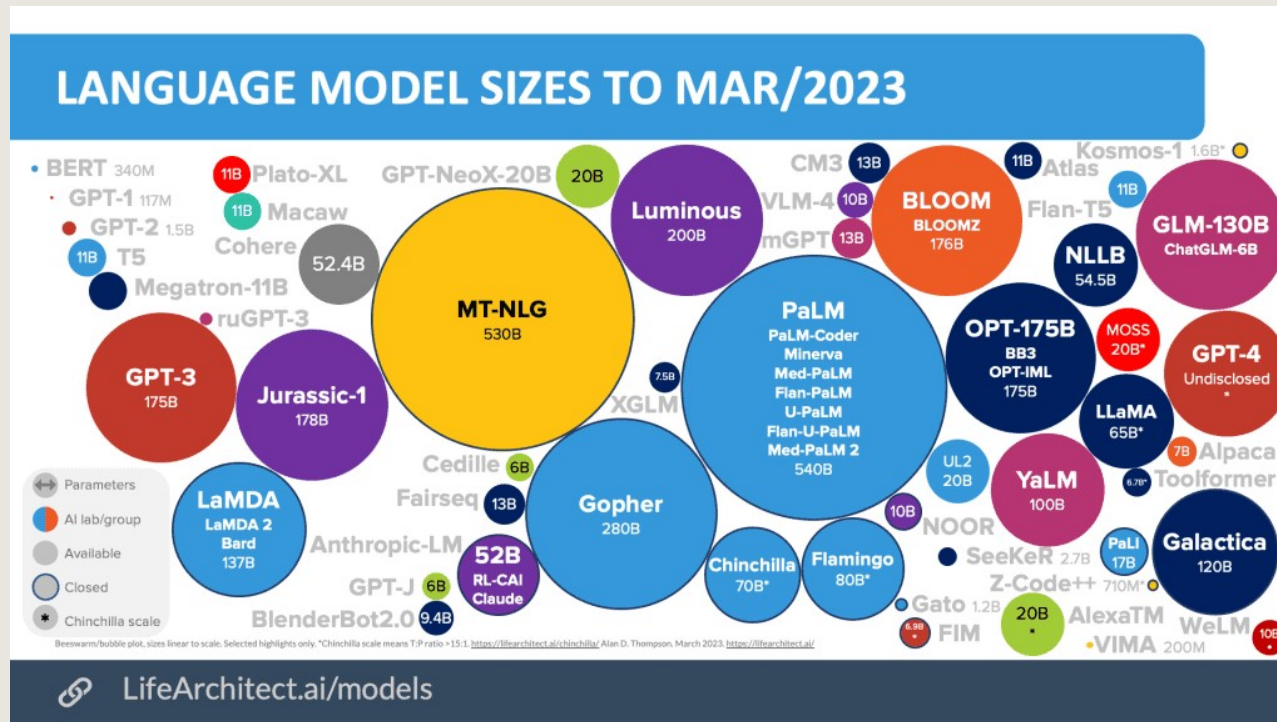|	GPT-1	GPT-2	GPT-3	GPT-3.5	CHATGPT
Release Date	Jun-18	Feb-19	May-20	Jan-22	Nov-22
Number of Models:	N/A	one model	one model	three models	
Parameters	N/A	1.5 billion	175 billion	1.3B, 6B, 175B	
Goal:	Release 1	produce a human-like text	Improved answers	the state of art with fewer toxic outputs	Training & Public Beta

Source: ChatGPT has reignited AI

"ChatGPT, a version of OpenAI's GPT-3.5 [family]... gained more than 100m users in its first two months and is now estimated to produce a volume of text every 14 days that is equivalent to all the printed works of humanity."
— Dr Thompson, Feb/2023, cited in report by the National Bureau of Economic Research (Scholes, Bernanke, MIT)

Note: GPT 3.0 used 175 billion parameters in their released Language model.
Microsoft Megatron-Turing model (already in production) has 530 billion parameters!

Getting better Every Minute!

NOT THE ONLY FOUNDATIONAL LANGUAGE MODEL

Source: Inside language models (from GPT-4 to PaLM) – Dr Alan D. Thompson – Life Architect

ChatGPT Advantages

GPT is an AI language model that uses deep learning techniques to generate human-like responses to natural language queries.

Natural language processing: ChatGPT is designed to understand and respond to natural language queries, making it easier for users to interact with the chatbot. This natural language processing capability is superior to traditional rule-based chatbots that require pre-defined responses.

Advanced learning: ChatGPT's deep learning AI tools enable it to continuously learn from user interactions, improving its responses over time. This makes it more accurate and efficient than traditional chatbots that require frequent updates and maintenance.

Context-awareness: ChatGPT can recognize and interpret the context of a conversation, allowing it to provide relevant and personalized responses to users. This is a significant advantage over traditional chatbots that provide generic responses irrespective of the conversation context.

Multi-language support: ChatGPT can support multiple languages, making it accessible to a broader audience. This is a significant advantage over traditional chatbots that are limited to a single language.

Flexibility: ChatGPT can be easily integrated into various applications and platforms, making it a flexible and versatile solution. This is a significant advantage over traditional chatbots that are often limited to specific platforms or applications.

The Cost of running ChatGPT

100 million active users in January!

LLM Development and training (tens of millions+) aside, estimations are that it costs Open AI $700,000 per day, or 36 cents per question, to keep the chatbot up and running. (~10 times more than a google search)

Microsoft already has this AI chip in the works. Code-named Athena, it is currently being tested internally with the brand's own teams. The chip is expected to be introduced next year for Microsoft's Azure AI services.

Nvidia released

Deploying current ChatGPT into every search done by Google would require 512,820.51 A100 HGX servers with a total of 4,102,568 A100 GPUs. The total cost of these servers and networking exceeds $100 billion of Capex alone, of which Nvidia would receive a large portion.

Google Search Cost Structure		
Metric	Current Google Search	ChatGPT Additional Costs
Revenue per query	$ 0.0161	$ 0.0161
Cost per query	$ 0.0106	$ 0.0142
Income per query	$ 0.0055	$ 0.0019
Query per second	320,000	320,000
Annual Revenue	$ 162.5 Billion	$ 162.5 Billion
Annual Costs	$ 107.0 Billion	$ 142.9 Billion
Operating Income	$ 55.5 Billion	$ 19.5 Billion

- What It Takes to Run ChatGPT
- How Much Does ChatGPT Cost to Run? $700K/day, Per Analyst (businessinsider.com)
- The Inference Cost Of Search Disruption – Large Language Model Cost Analysis (semianalysis.com)
- You won't believe how much ChatGPT costs to operate | Digital Trends

THE RISK OF AI

TACTICAL

Jobs

Data Security

Privacy

Data bias

Legal responsibility

AI programmed for harm

hal·lu·ci·na·tion

/hə͟ˌlo͞osəˈnāSH(ə)n/ • noun

an experience involving the apparent perception of something not present.

synonyms   antonyms   thesaurus

Definitions from Oxford Languages

LONGER TIMELINE

AGI stands for artificial general intelligence, which is a theoretical form of AI that is as capable or more capable than humans.
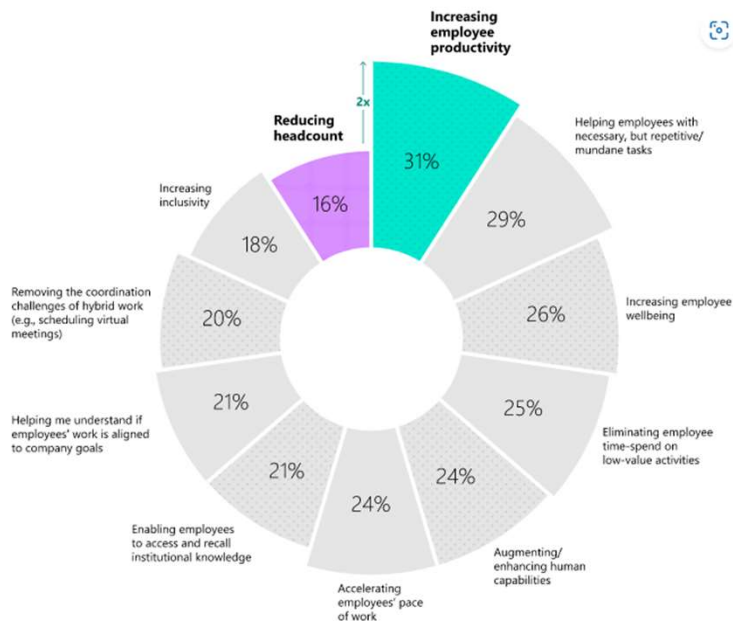
"Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."

Do the benefits outweigh the risks?

SUMMARY

AI's Productivity Promise

Amid fears of AI job loss, business leaders are 2x more likely to choose 'increasing employee productivity' than 'reducing headcount' when asked what they would most value about AI in the workplace.

Increasing employee productivity — 31%

Reducing headcount — 16%

Increasing inclusivity — 18%

Removing the coordination challenges of hybrid work (e.g., scheduling virtual meetings) — 20%

Helping me understand if employees' work is aligned to company goals — 21%

Enabling employees to access and recall institutional knowledge — 21%

Accelerating employees' pace of work — 24%

Augmenting/enhancing human capabilities — 24%

Eliminating employee time-spend on low-value activities — 25%

Increasing employee wellbeing — 26%

Helping employees with necessary, but repetitive/mundane tasks — 29%

Q: Will AI take my job?
A: Not likely but it will change it.

ChatGPT is an application
AI will be embedded into everything
It is a productivity booster / An assistant.

This is the first-time humans have created a "device" capable of creating content.   We do not know what these systems will create, as they are models.

This is much more than just ChatGPT – it is about the opportunity to use this technology in ways we can not yet fully explore.

THANK YOU

Jorge Pereira

Modern Workforce Solutions Principal

Jorge_Pereira@dell.com

Dell Technologies

RESOURCES

My Blogs on this topic:
- [ChatGPT has reignited AI](#)
- [Understanding how AI Works](#)
- [What It Takes to Run ChatGPT](#)

- [GPT-4 (openai.com)](#)
- [GPT-4 vs. GPT-3.5: how much difference is there? | Digital Trends](#)
- [GPT-3.5 + ChatGPT: An illustrated overview – Dr Alan D. Thompson – Life Architect](#)
- [GPT-2 vs GPT-3 vs GPT-3.5 vs GPT-4: A Comprehensive Comparison of OpenAI LLMs (opengenus.org)](#)
- [Frameworks and Tools – Meta AI (facebook.com)](#)
- [Inside language models (from GPT-4 to PaLM) – Dr Alan D. Thompson – Life Architect](#)
- [GPT-3.5 model architecture (opengenus.org)](#)