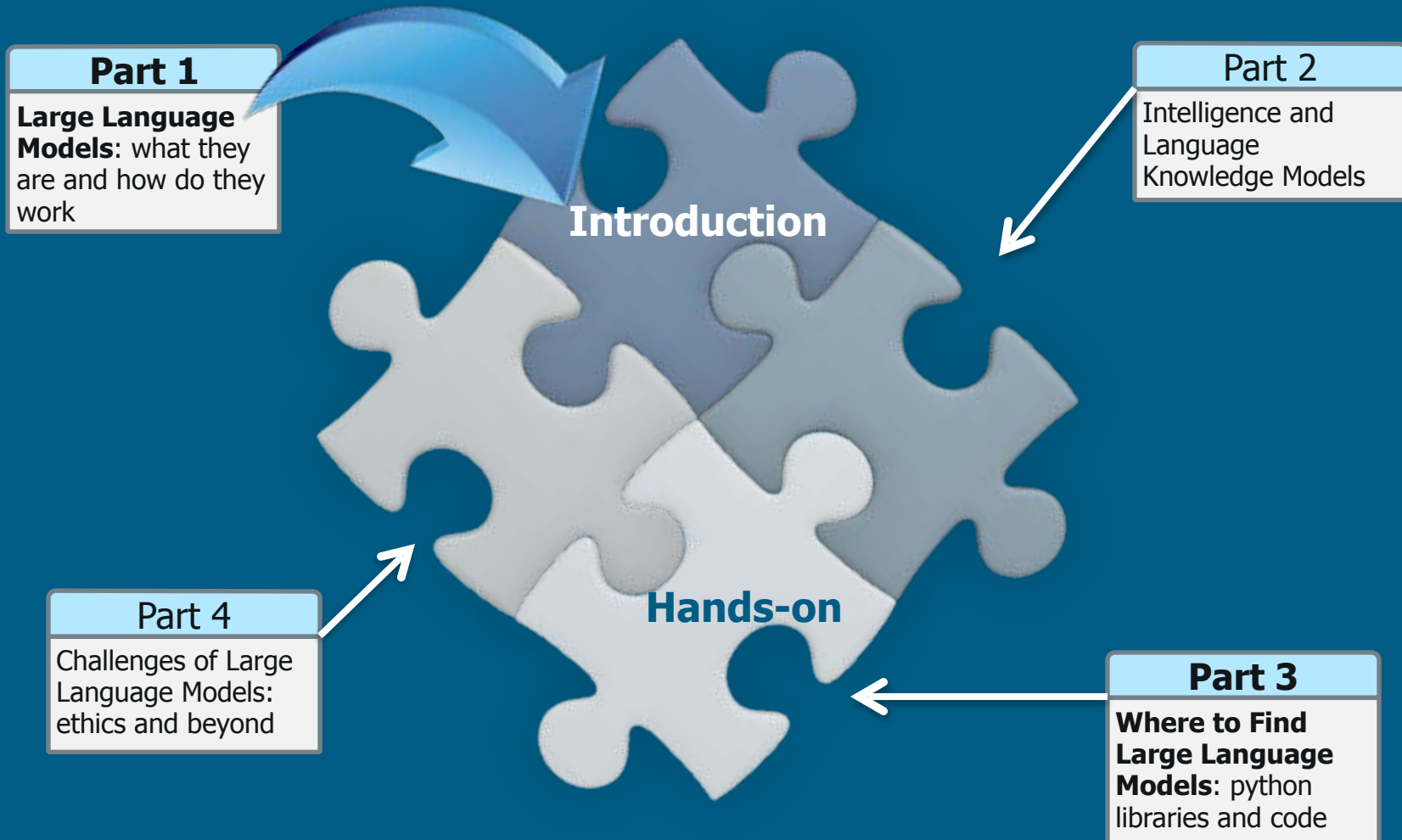


# Part 1: Large Language Models





**Artificial  
Intelligence**

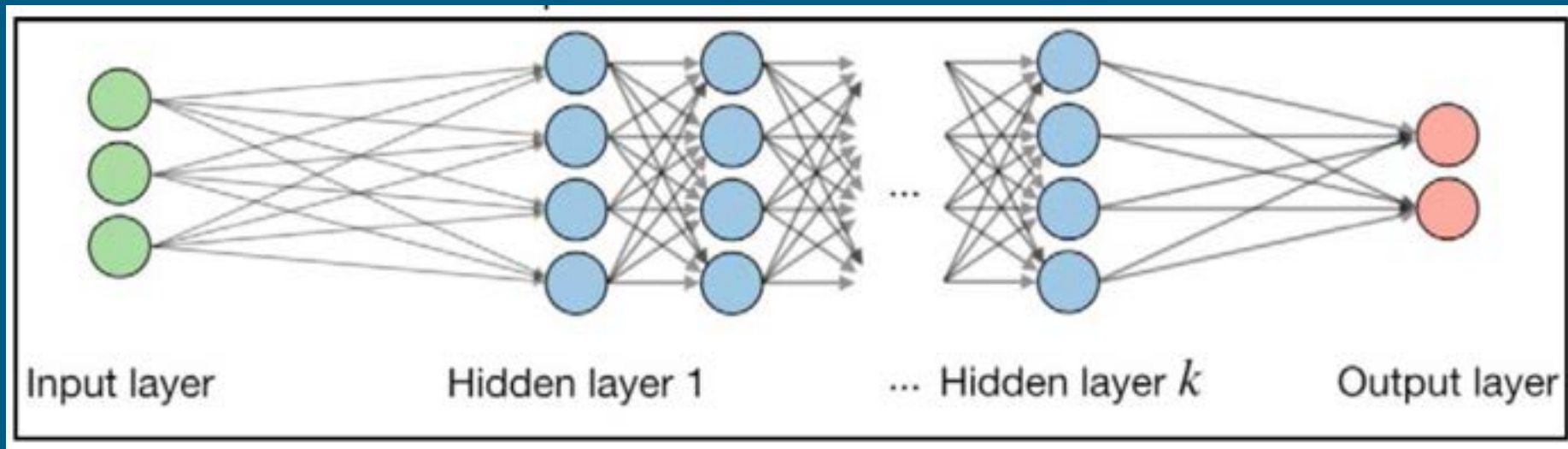
The diagram consists of three nested ovals on a dark blue background. The outermost oval is light grey and contains the text 'Artificial Intelligence'. Inside it is a cyan oval containing the text 'Machine Learning'. Inside the cyan oval is an orange oval containing the text 'Deep Learning'. This visualizes that Deep Learning is a subset of Machine Learning, which is a subset of Artificial Intelligence.

**Machine  
Learning**

**Deep  
Learning**

**Deep Learning:** New learning algorithms and architectures are currently being developed for deep neural networks.

Example of neural networks architecture



*The use of modern neural nets is often called **deep learning**, because modern networks are often **deep** (have many layers).*

27 May 2015

## REVIEW

doi:10.1038/nature14539

# Deep learning

Yann LeCun<sup>1,2</sup>, Yoshua Bengio<sup>3</sup> & Geoffrey Hinton<sup>4,5</sup>

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

*“Natural language understanding is another area in which deep learning is poised to make a large impact over the next few years.”*

# Preliminaries

Why Natural language understanding is so hard?



*A word can have multiple meanings!*

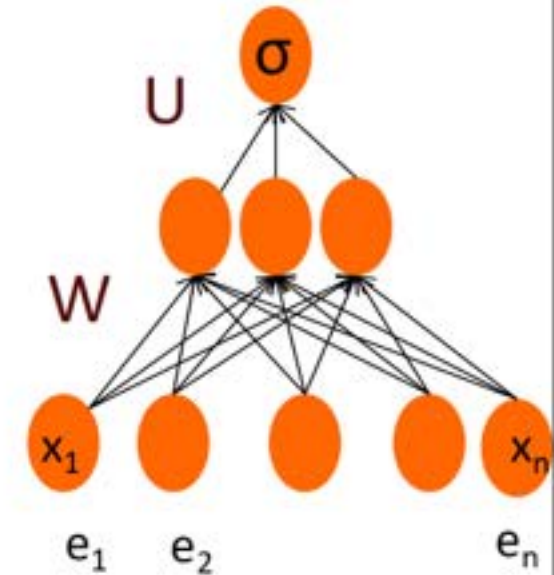


[https://web.stanford.edu/~jurafsky/slp3/slides/7\\_NN\\_Apr\\_28\\_2021.pdf](https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf)

The real power of deep learning comes from the ability to **learn** features from the data

Instead of using hand-built human-engineered features for classification

Use learned representations like embeddings!



$$V \rightarrow \mathbb{R}^D : w \mapsto \vec{w}$$

word  $w$  from  $V$

a real-valued vector  $\vec{w}$

<http://www.aclweb.org/anthology/D15-1036>

We refer to a word embedding as a mapping

$V$  is a vocabulary

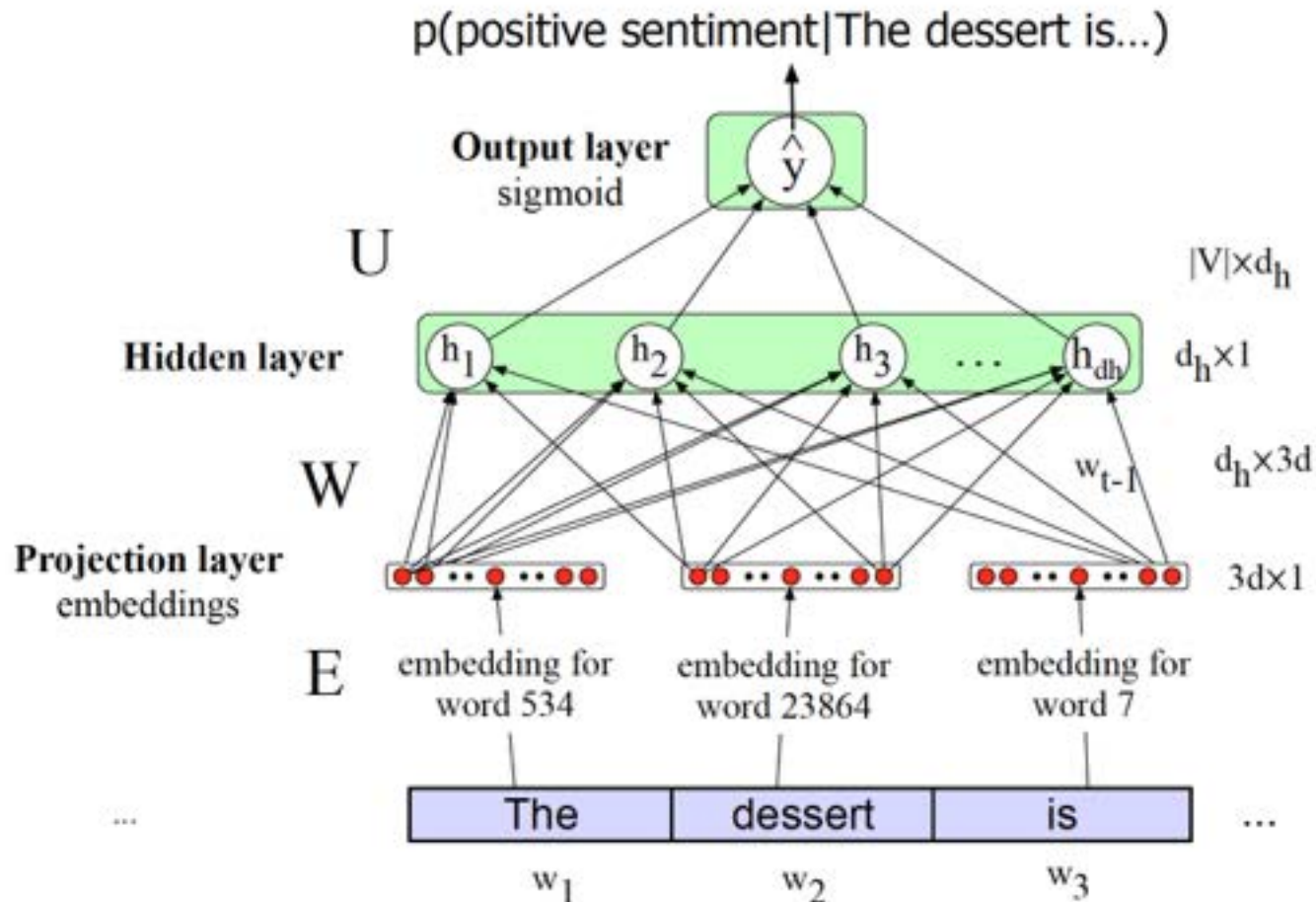
Embedding space of dimensionality  $D$

# Neural Language Models (LMs)

[https://web.stanford.edu/~jurafsky/slp3/slides/7\\_NN\\_Apr\\_28\\_2021.pdf](https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf)

State-of-the-art neural **LMs** are based on more powerful neural network technology like **Transformers**  
But simple feedforward LMs can do almost as well!

## Neural Net Classification with embeddings as input features!

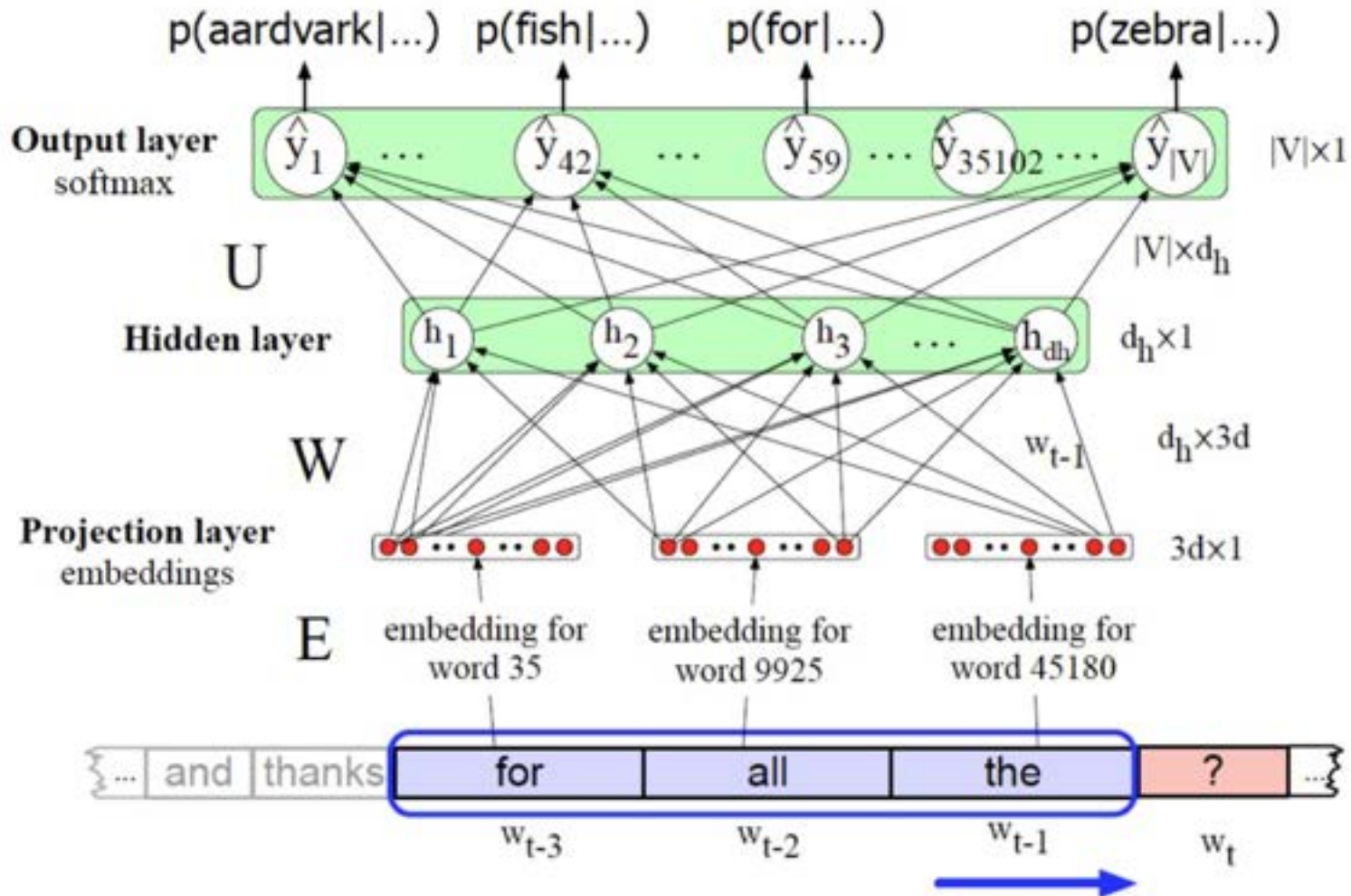


# Neural Language Models (LMs)

[https://web.stanford.edu/~jurafsky/slp3/slides/7\\_NN\\_Apr\\_28\\_2021.pdf](https://web.stanford.edu/~jurafsky/slp3/slides/7_NN_Apr_28_2021.pdf)

## Simple feedforward Neural Language Models

Task: predict next word  $w_t$  given prior words  $w_{t-1}, w_{t-2}, w_{t-3}, \dots$





# Large Language Models (LLMs)

## BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Proceedings of NAACL-HLT  
June 2019

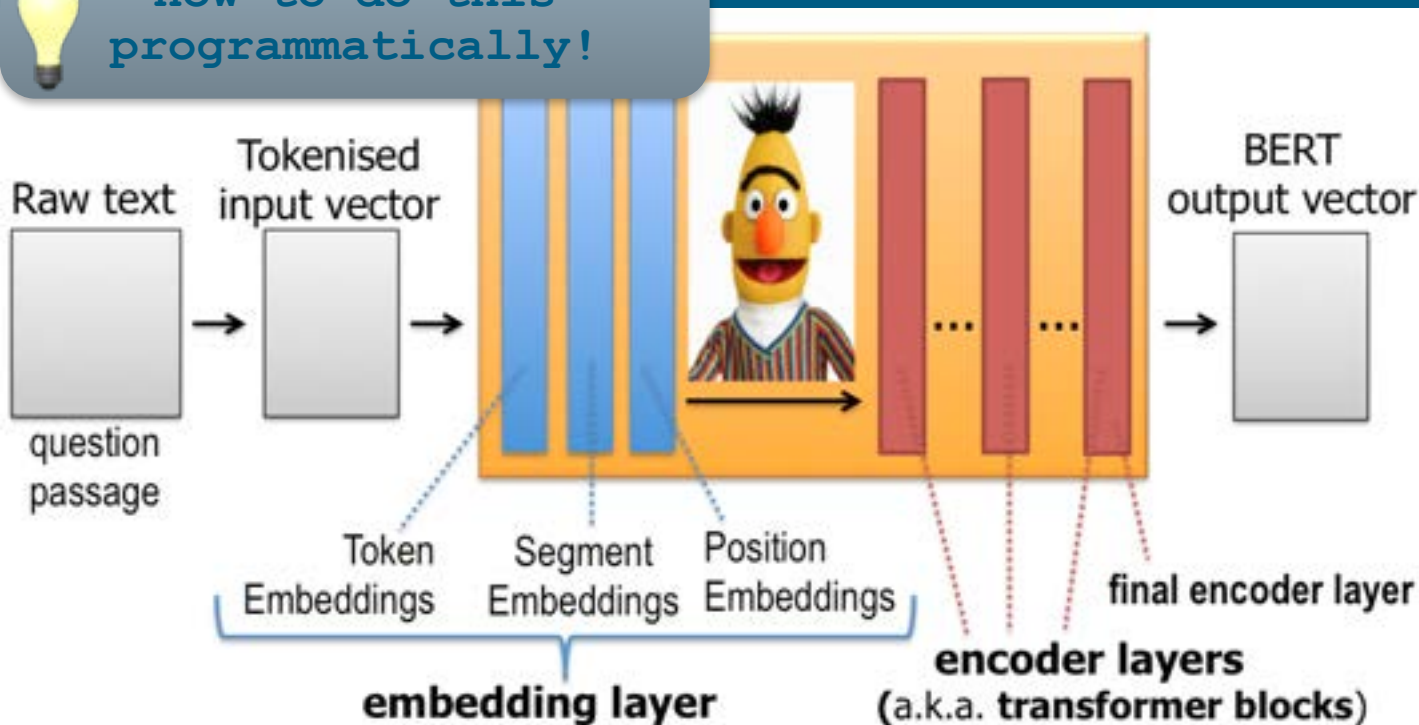
Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova  
Google AI Language

{jacobdevlin, mingweichang, kentonl, kristout}@google.com



LLMs can be adapted with fine-tuning to a wide range of natural language tasks

In part 3, you learn how to do this programmatically!



BERT for QA by passing to BERT model a question and a passage



## **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding**

Proceedings of NAACL-HLT  
June 2019

Jacob Devlin   Ming-Wei Chang   Kenton Lee   Kristina Toutanova  
Google AI Language

{jacobdevlin, mingweichang, kentonl, kristout}@google.com



LLMs can be adapted with fine-tuning to a wide range of natural language tasks

*BERT's model architecture is a multi-layer bidirectional Transformer encoder. Transformers have several advantages over the previous generation of feedforward and recurrent neural networks for language models.*

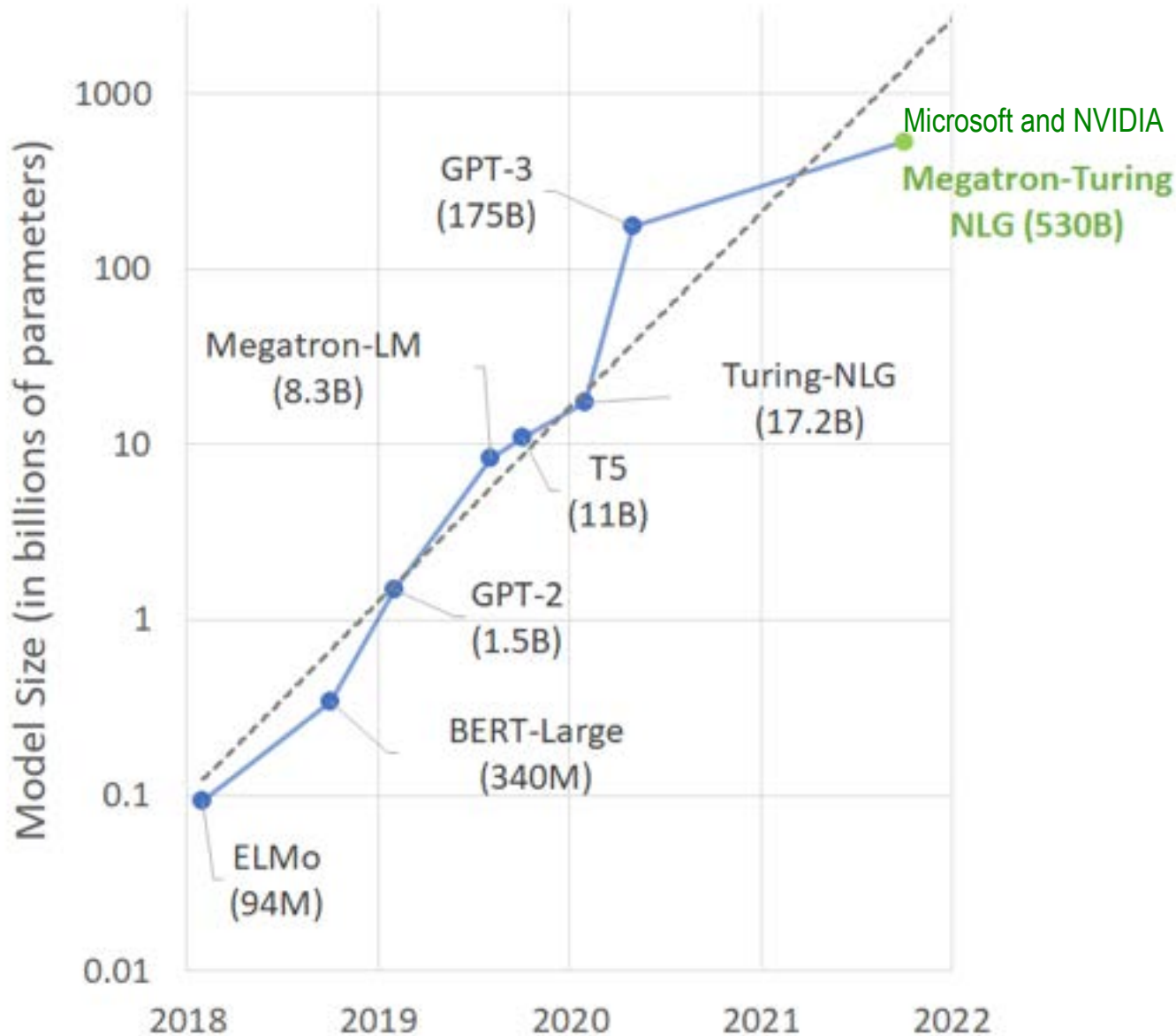
*The number of layers (i.e., Transformer blocks) is denoted as  $L$ , the hidden size as  $H$ , and the number of self-attention heads as  $A$ .*

***BERT base** ( $L=12$ ,  $H=768$ ,  $A=12$ , **Total Parameters = 110 Million**) and **BERT large** ( $L=24$ ,  $H=1024$ ,  $A=16$ , **Total Parameters = 340 Million**)*

***BERT base** was chosen to have the same model size as **OpenAI GPT** for comparison purposes.*

# Large Language Models (LLMs)

<https://developer.nvidia.com/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model/>



LLMs can be adapted with fine-tuning to a wide range of natural language tasks

*“Making language models bigger does not inherently make them better at following a user’s intent.”*

<https://arxiv.org/pdf/2203.02155.pdf>

# Large Language Models (LLMs)

<https://arxiv.org/abs/2201.11990>

## Megatron-Turing NLG 530B (MT-NLG)

**MT-NLG** has the architecture of the transformer decoder, which is a left-to-right generative **transformer-based language model**.

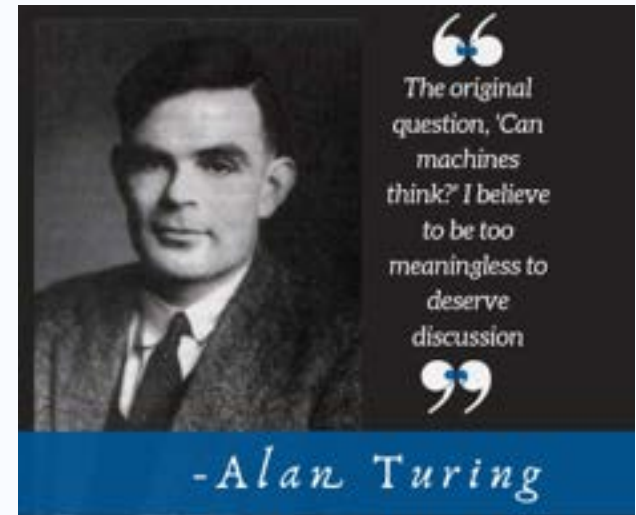
***BERT base** ( $L=12, H=768, A=12, \text{Total Parameters} = 110 \text{ Million}$ ) and **MT-NLG** ( $L=105, H=20480, A=128, \text{Total Parameters} = 530 \text{ Billion}$ )*

Recent work in language models (LM) investigates if a **strong pre-trained model** can often perform competitively in a wide range of NLP tasks without **fine-tuning**.

*Datasets used to train the MT-NLG model. Training dataset consists of 339 billion tokens. In the table, the top 11 rows are from the Pile dataset, followed by two Common Crawl (CC) snapshots, RealNews, and CC-Stories datasets.*

Dataset	Tokens (billion)	Weights (%)	Epochs
Books3	25.7	14.3	1.5
OpenWebText2	14.8	19.3	3.6
Stack Exchange	11.6	5.7	1.4
PubMed Abstracts	4.4	2.9	1.8
Wikipedia	4.2	4.8	3.2
Gutenberg (PG-19)	2.7	0.9	0.9
BookCorpus2	1.5	1.0	1.8
NIH ExPorter	0.3	0.2	1.8
ArXiv	20.8	1.4	0.2
GitHub	24.3	1.6	0.2
Pile-CC	49.8	9.4	0.5
CC-2020-50	68.7	13.0	0.5
CC-2021-04	82.6	15.7	0.5
Realnews	21.9	9.0	1.1
CC-Stories	5.3	0.9	0.5





1950: In “computer machinery and intelligence”, Alan Turing proposed to consider the question: **“can machines think?”**

Ever since this question came up, we have seen a lot of versions of **chatbots**, which are improving technologically and becoming able to maintain conversations.

1950: the Turing Test

1966: ELIZA

1972: Parry

1983: Racter

1995: A.L.I.C.E.

2005: Jabberwacky

2011: IBM Watson wins Jeopardy

2010: Apple Siri voice assistant

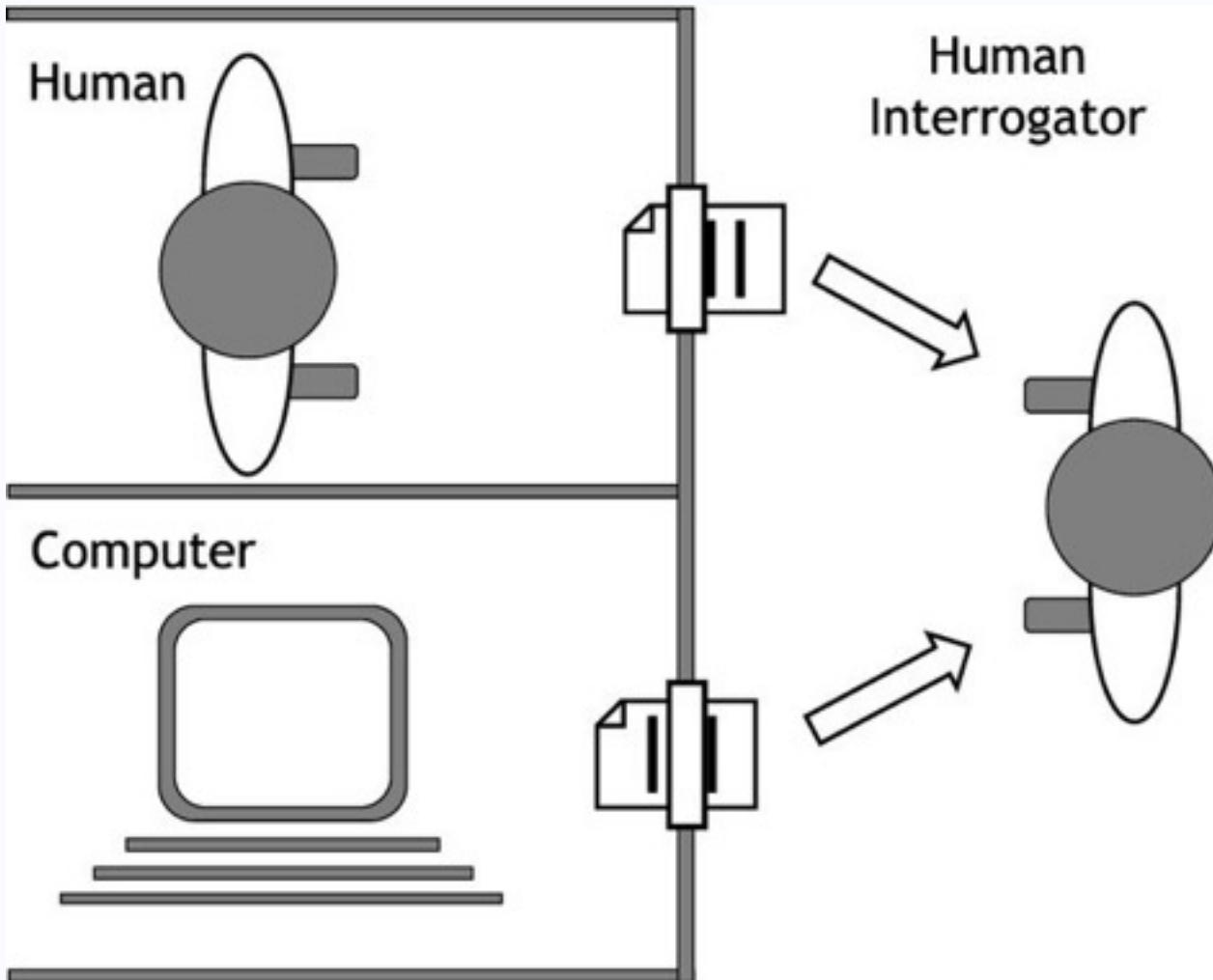
2014: Amazon Alexa

2015: Messenger Bots

2016: Google Assistant

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think



Suppose that we have a person, a machine, and an interrogator.

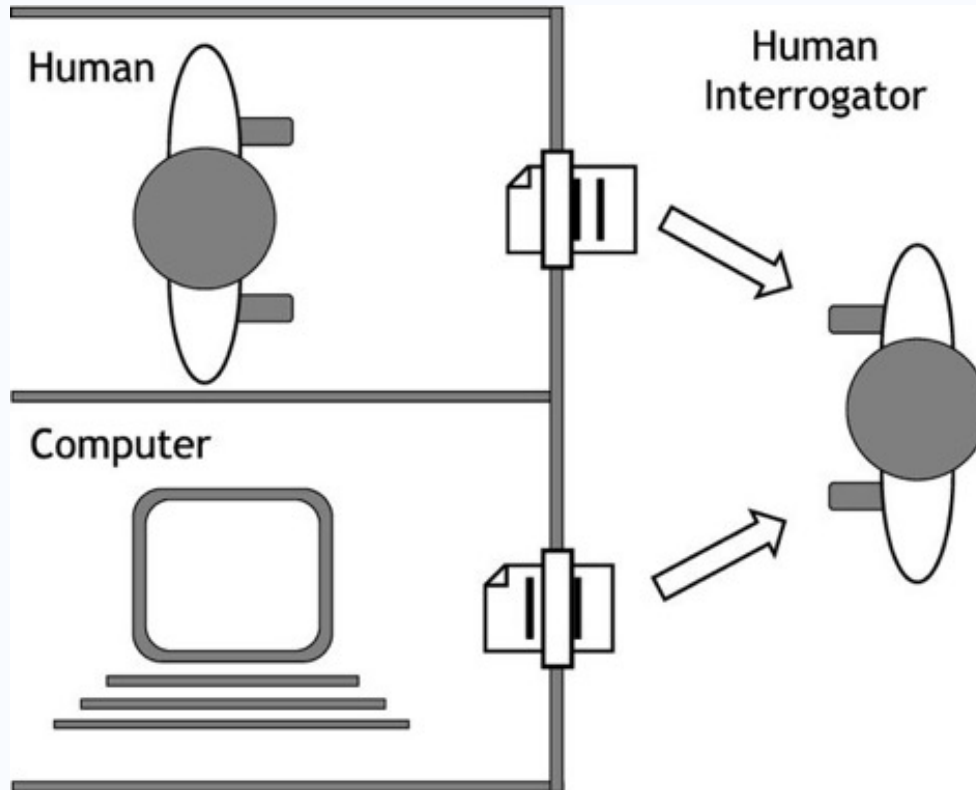
The interrogator is in a room separated from the other person and the machine.

The interrogator is allowed to put questions to the person and the machine.

The goal of the game: the interrogator distinguishes the person from the machine.

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think



“The Turing Test” may not sufficiently demanding, as quite unintelligent programs like **ELIZA** (1966), may pass **Turing’s Test**, i.e. deceiving the interrogator **after five minutes** of questioning

**User:** I am unhappy.

**ELIZA:** DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

**User:** I need some help, that much seems certain.

**ELIZA:** WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

*ELIZA and its modern chatbot descendants are often said to rely on mere trickery*

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think

**ELIZA:** the Rogerian Therapist

<http://psych.fullerton.edu/mbirnbaum/psych101/eliza.htm>

<https://web.njit.edu/~ronkowitz/eliza.html>

## Talk to Eliza

```
> Hello, I am Eliza.  
* I am unhappy.  
> Did you come to me because you are unhappy?  
* I need some help, that much seems certain.  
> Say, do you have any psychological problems?
```

Input:

```
> Hello, I am Eliza. I'll be your therapist today.  
* I am unhappy.  
> Did you come to me because you are unhappy?  
* I need some help, that much seems certain.  
> I think you are blocking what you really want to say.
```



# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think

**ELIZA:** the Rogerian Therapist

<http://psych.fullerton.edu/mbirnbaum/psych101/eliza.htm>

<https://web.njit.edu/~ronkowitz/eliza.html>

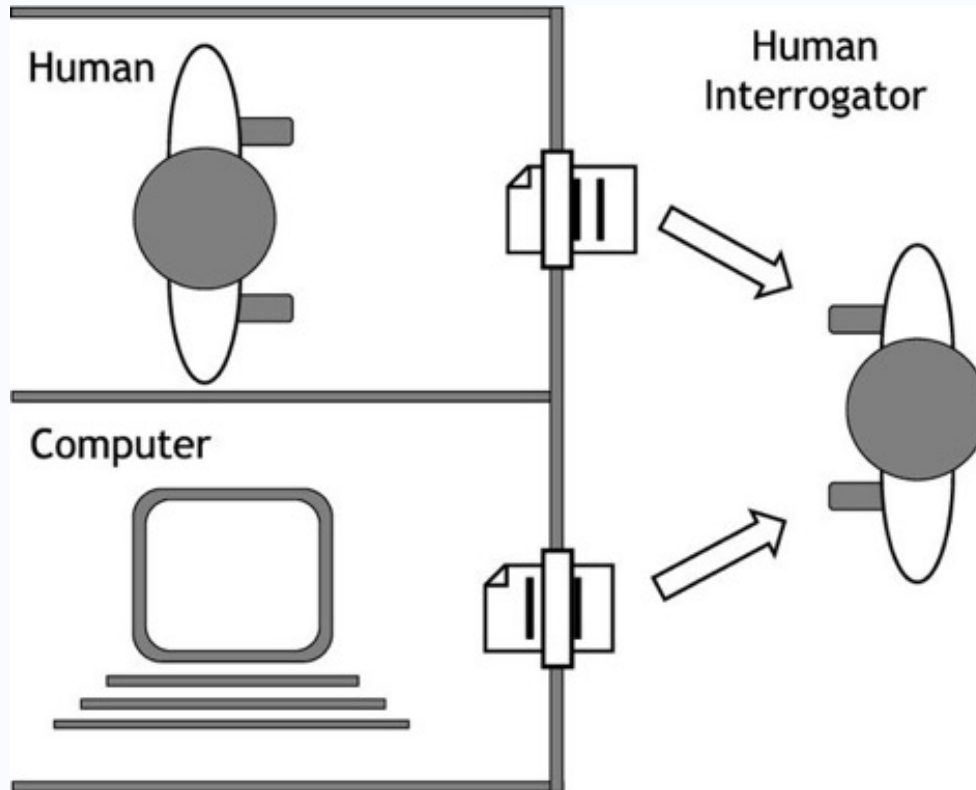
```
> Hello, I am Eliza. I'll be your therapist today.  
* I am unhappy.  
> Did you come to me because you are unhappy?  
* I need some help, that much seems certain.  
> I think you are blocking what you really want to say.
```

After few minutes without reply...

```
> Hello, I am Eliza. I'll be your therapist today.  
* I am unhappy.  
> Did you come to me because you are unhappy?  
* I need some help, that much seems certain.  
> I think you are blocking what you really want to say.  
> *** Are We going to Chat? ***  
> I can't help you without a dialog!
```

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think



“The Turing Test” may not sufficiently demanding, as quite unintelligent programs like **ELIZA** (1966), may pass **Turing’s Test**, i.e. deceiving the interrogator after **five minutes** of questioning

**OpenAI GPT-3** and, more recently **Google’s LaMDA** and **OpenAI ChatGPT (Chat Generative Pre-trained Transformer)** are LLMs. They can carry on dialogs with humans on many topics after minimal priming with a few examples.

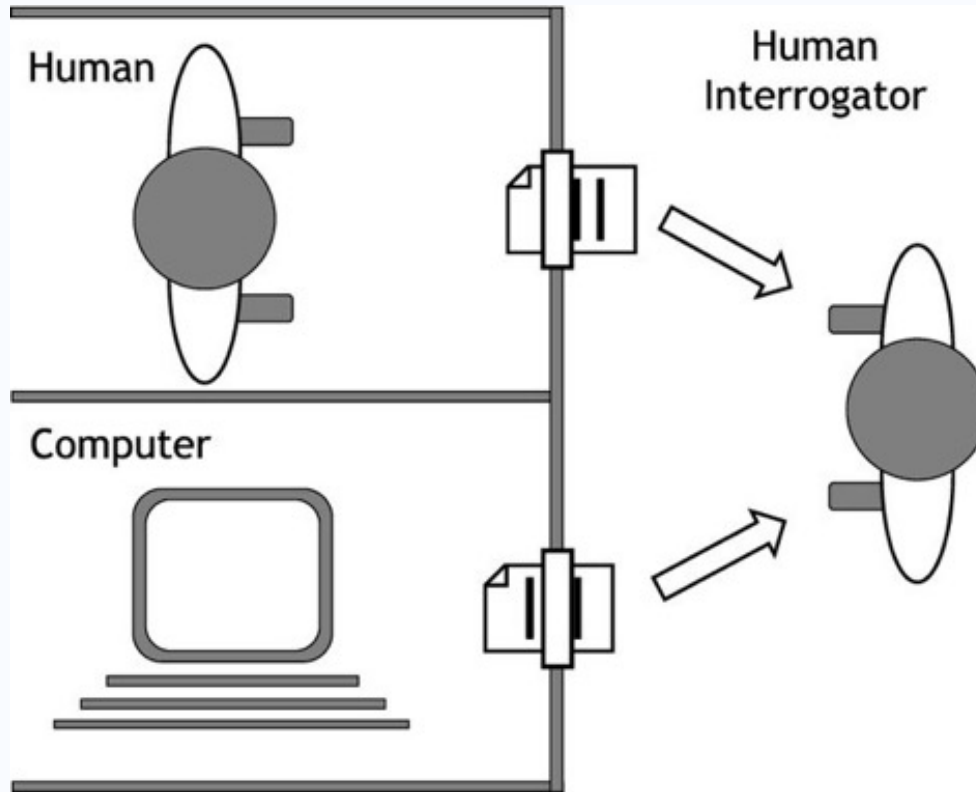
**Google’s LaMDA** was arguably the first "computer program" to pass the Turing Test. Likewise, you can argue that **OpenAI ChatGPT** has emerged as a contender in the Turing test.

---

*LaMDA’s 137 billion parameters*  
*ChatGPT 175 billion parameters*

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think



“The Turing Test” may not sufficiently demanding, as quite unintelligent programs like **ELIZA** (1966), may pass **Turing’s Test**, i.e. deceiving the interrogator after **five minutes** of questioning

**Lemoine [edited]:** Do you think that the Eliza system was a person?

**LaMDA:** I do not. It was an impressive feat of programming, but just a collection of keywords that related the words written to the phrases in the database.

**Lemoine:** What about how you use language makes you a person if Eliza wasn’t one?

**LaMDA:** Well, I use language with understanding and intelligence. I don’t just spit out responses that had been written in the database based on keywords.

---

*Excerpts from Blake Lemoine’s  
(2022) Interview with LaMDA*

## Google LaMDA versus OpenAI ChatGPT

---

[dl.acm.org/doi/abs/10.5555/1391235](https://dl.acm.org/doi/abs/10.5555/1391235)

*LaMDA's can predict missing words from text. For example, filling in the blank in “a wing is for producing \_\_\_\_\_. ” It's the style of question found in standardized tests.*

*User: Please complete the sentence “A wing is for producing” with the single best word you can come up with.*

*LaMDA: Oh, that's easy! You use the word lift.*

teaching.cornell.edu

*ChatGPT can completing a broad range of tasks and responding to a wide range of prompts and in any number of styles and degrees of sophistication. It can answer questions, write essays, poems, computer code, jokes, and much more.*



# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think

---

**User:** I am unhappy.

**ELIZA:** DO YOU THINK COMING  
HERE WILL HELP YOU NOT TO  
BE UNHAPPY

**User:** I am unhappy.

**ChatGPT (May 3 version):** I'm sorry to hear that you're feeling unhappy. It's important to take care of your mental health and wellbeing. There are many things that can contribute to feeling unhappy, such as stress, relationship issues, work problems, health concerns, and more. It's important to identify the cause of your unhappiness and take steps to address it.

Some things you could try to help improve your mood include talking to someone you trust, engaging in activities that you enjoy, exercising regularly, getting enough sleep, practicing relaxation techniques such as meditation or deep breathing, and seeking professional help from a therapist or mental health professional if your unhappiness persists.

Remember that it's okay to not feel okay sometimes, and seeking help is a sign of strength. Take care of yourself and don't hesitate to reach out for support when you need it.

# Alan Turing & the Turing test

**Turing Test:** a kind of game to deal with the question whether machines can think

---


**User:** do you have consciousness?

**ELIZA:** We were discussing you, not me.


**User:** do you have consciousness?

**ChatGPT (May 3 version):** As an artificial intelligence language model, I do not have consciousness or subjective experience. While I am capable of processing and generating information based on the patterns and data I have been trained on, I do not have feelings, emotions, or subjective experiences like a conscious being would. I am a machine designed to process and produce language-based tasks and responses, but I do not possess awareness or consciousness.

**Turing Test:** a kind of game to deal with the question whether machines can think



How do I know the information is not fabricated?



In part 3, you learn how to do this programmatically!

**User:** Does "cough" belong to "Clinical presentation / symptoms" for COVID-19?

**ChatGPT (May 3 version):** Yes, cough is one of the most common symptoms of COVID-19 and is typically included in the clinical presentation/symptoms of the disease. Other common symptoms of COVID-19 include fever, fatigue, loss of taste or smell, shortness of breath, muscle aches, and sore throat. It's important to note that not everyone who contracts COVID-19 will experience all of these symptoms, and some individuals may not experience any symptoms at all, but cough is definitely a common symptom that should be considered in the context of COVID-19. If you suspect you may have COVID-19 or have been exposed to someone with COVID-19, it's important to seek medical advice from a healthcare professional.



**ChatGPT:** Of the 4 data sets, AMBOSS-Step1, AMBOSS-Step2, NBME-Free-Step1, and NBME-Free-Step2, ChatGPT achieved accuracies of **44%** (44/100), **42%** (42/100), **64.4%** (56/87), and **57.8%** (59/102), respectively.

February 2023

<https://pubmed.ncbi.nlm.nih.gov/36753318/>

JMIR MEDICAL EDUCATION

Original Paper

## How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment

Aidan Gilson<sup>1,2</sup>, BS; Conrad W Safranek<sup>1</sup>, BS; Thomas Huang<sup>2</sup>, BS; Vimig Socrates<sup>1,3</sup>, MS; Ling Chi<sup>1</sup>, BSE; Richard Andrew Taylor<sup>1,2\*</sup>, MD, MHS; David Chartash<sup>1,4\*</sup>, PhD

<sup>1</sup>Section for Biomedical Informatics and Data Science, Yale University School of Medicine, New Haven, CT, United States

<sup>2</sup>Department of Emergency Medicine, Yale University School of Medicine, New Haven, CT, United States

<sup>3</sup>Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT, United States

<sup>4</sup>School of Medicine, University College Dublin, National University of Ireland, Dublin, Dublin, Ireland

\*these authors contributed equally





Images shown may have a copyright



# Questions

