

RADONC AI CURRICULUM

Introduction to Deep Learning

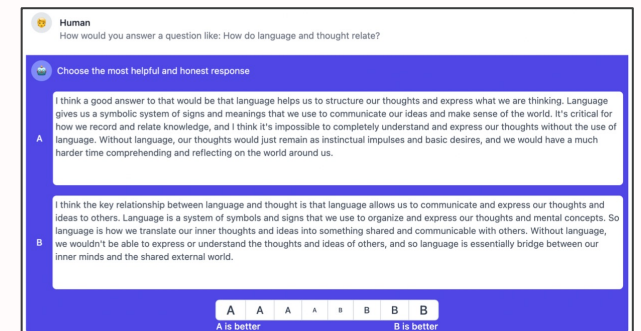
LECTURE 8: Turning LLMs into Assistants

ANDREW Y. K. FOONG, PH.D.

June 2nd, 2026



Radiation
Oncology
AI & Data Analytics
AIDA



Today's lecture

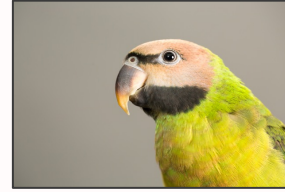
1. Can LLMs understand?
 - *What's needed for generalization*
2. Prediction is not assistance
 - *Next-token predictors are poor chatbots*
3. From predictor to assistant
 - *Post-training—the heart of modern AI*
4. Anatomy of a conversation
 - *What happens when you talk to a chatbot*
5. Prompting
 - *Writing effective instructions for LLMs*
6. Q&A

Can LLMs Understand?

What's needed for generalization

Generalization vs. memorization

- Can now create digital parrot.
- But **memorization \neq understanding.**
- Real question: can LLMs complete text that isn't anywhere online?
 - Shakespearean poem about AI in Healthcare?
 - Feedback on a never-before-seen radiation treatment plan?
- Yes! As long as the training data is large and diverse enough.
 - Won't work for toy internet example.
 - But when trained on real internet, LLMs *learn the principles of language and reasoning.*
 - Not a given!



Good prediction requires understanding

- Memorization \neq understanding.
- But next-word prediction + **generalization** \rightarrow understanding.
- **THOUGHT EXPERIMENT:**
 - Imagine LLM trained on detective novels.
 - Novel ends with “and the killer was...”
 - To accurately predict next word, LLM must understand the whole novel.
- If we can successfully train LLM to predict next word in new scenarios, we have created understanding!
- Generative AI revolution: **we can train excellent next-word predictors that generalize to brand-new scenarios.**



ILYA SUTSKEVER, FRS
former Chief Scientist
OpenAI

Large language models from 40,000 feet

Plan for LLMs:

1. Represent words as (many) numbers.
2. Generate sentences by predicting next word.
3. Train on data from Internet.
4. Specialized neural network architecture for text.
5. Steer network to be helpful and accurate.

Large language models from 40,000 feet

Plan for LLMs:

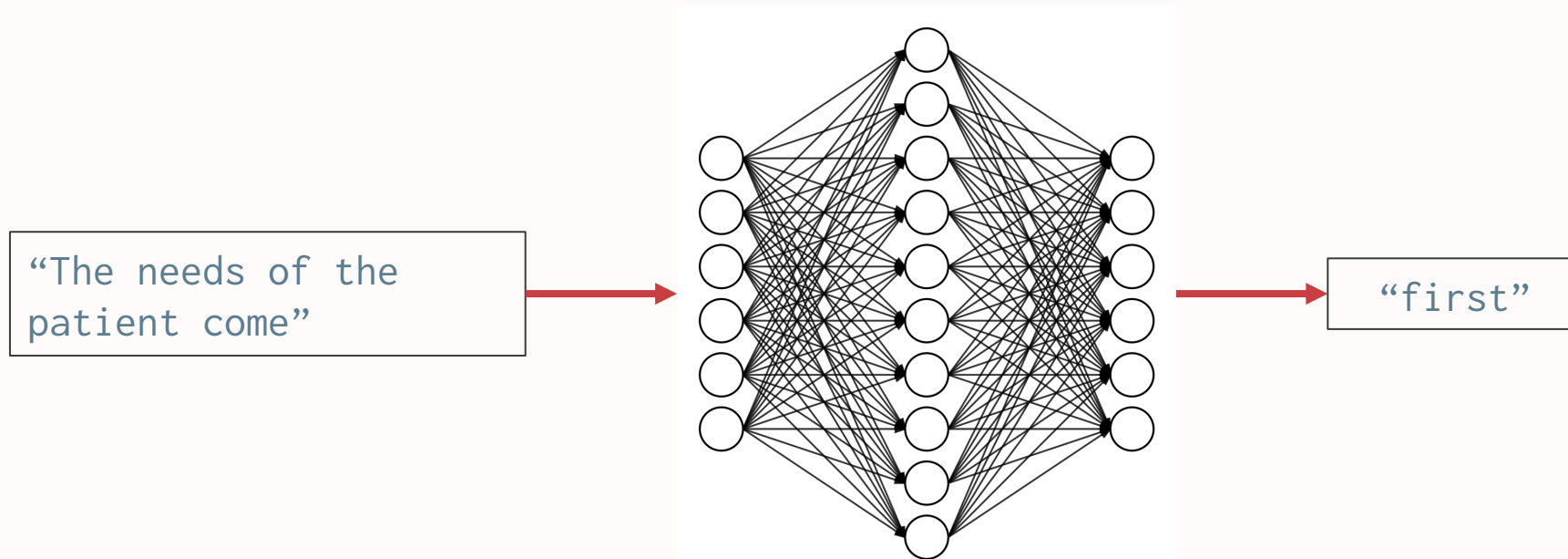
1. Represent words as (many) numbers.
2. Generate sentences by predicting next word.
3. Train on data from Internet.
4. Specialized neural network architecture for text.
5. **Steer network to be helpful and accurate.**

Prediction is not assistance

Next-token predictors are poor chatbots

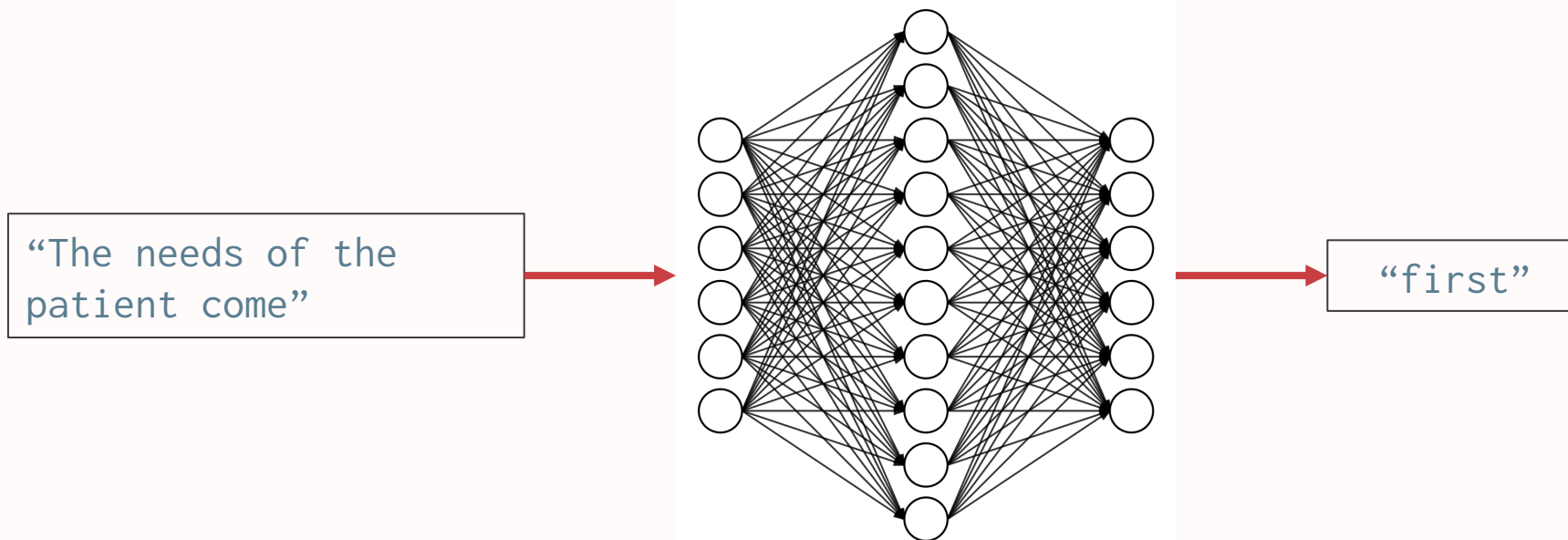
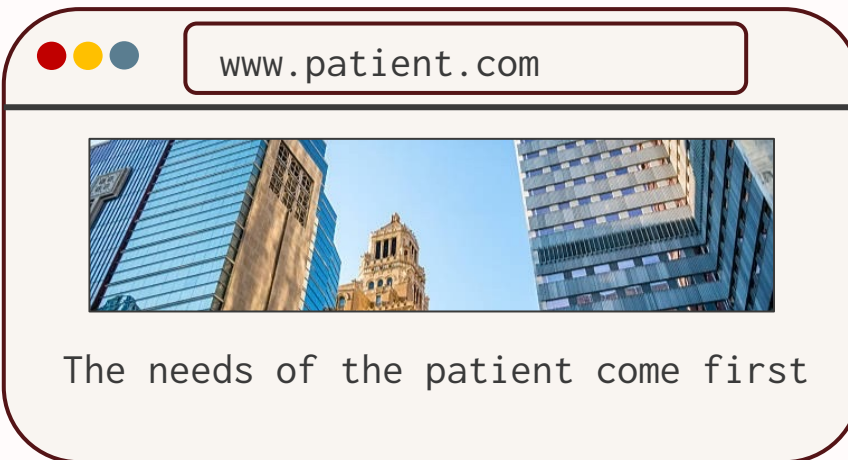
Learning by predicting

- ChatGPT is a large language model (LLM).
- LLMs are neural networks that:
 - Take incomplete sentences as input.
 - Output prediction for next word.



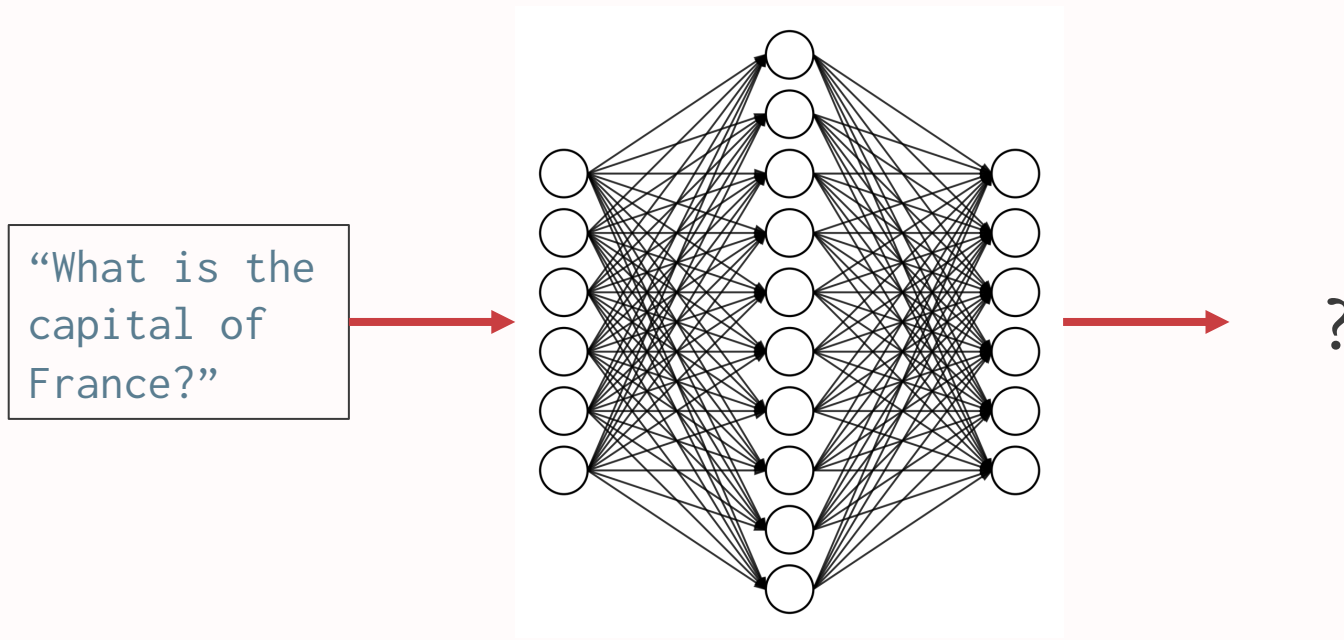
Learning by predicting

- Trained on text from entire internet.
- Learns what a typical website would say.
 - Similar to autocomplete.
 - No intrinsic knowledge.
 - All knowledge must be learned by association.



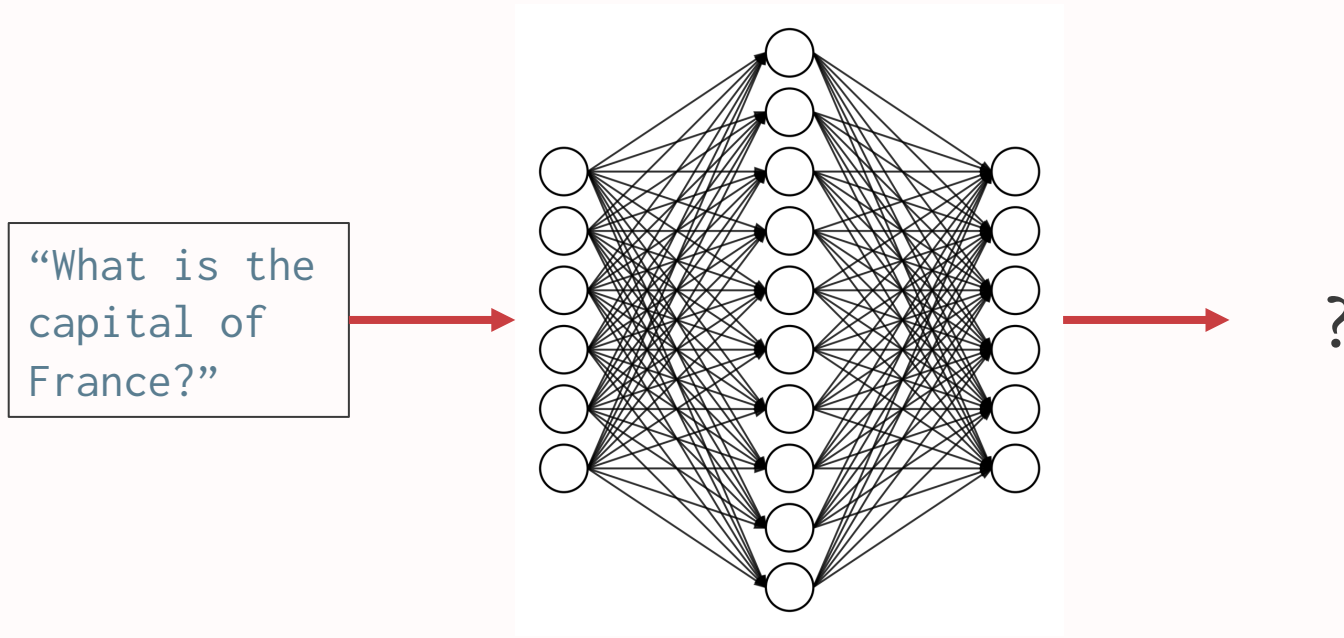
Prediction is different from assistance

- But is next-word prediction helpful?



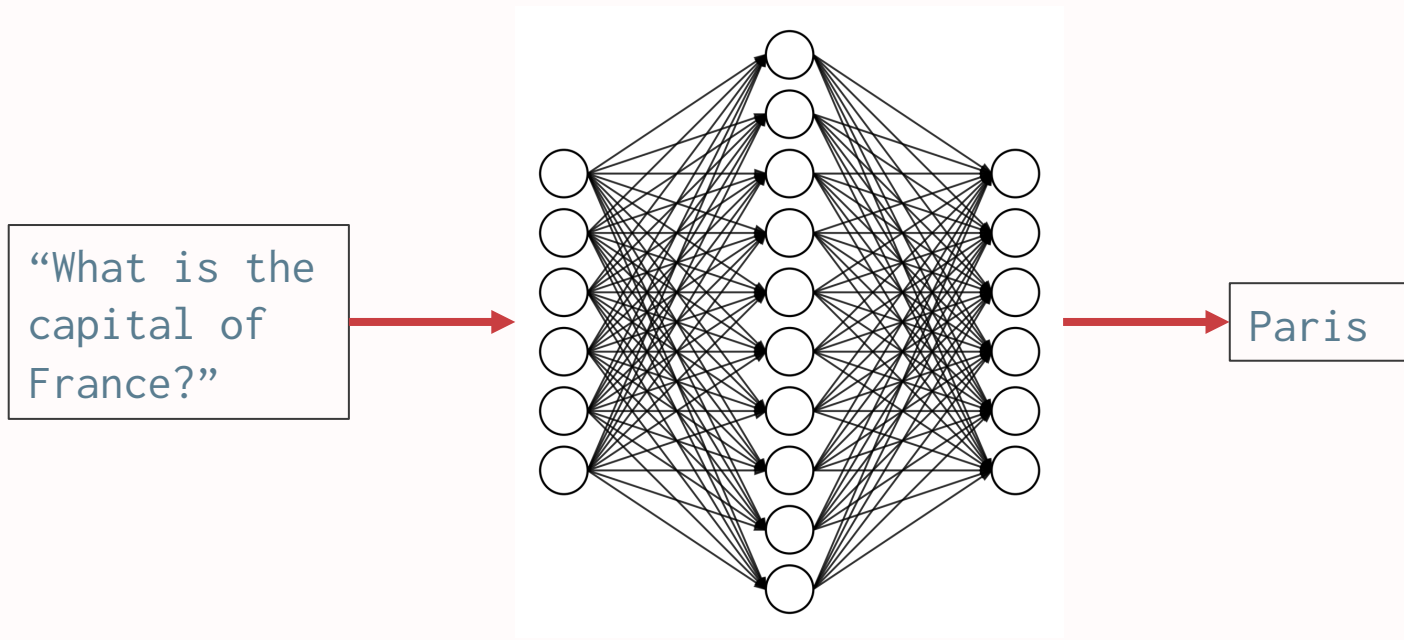
Prediction is different from assistance

- But is next-word prediction helpful?



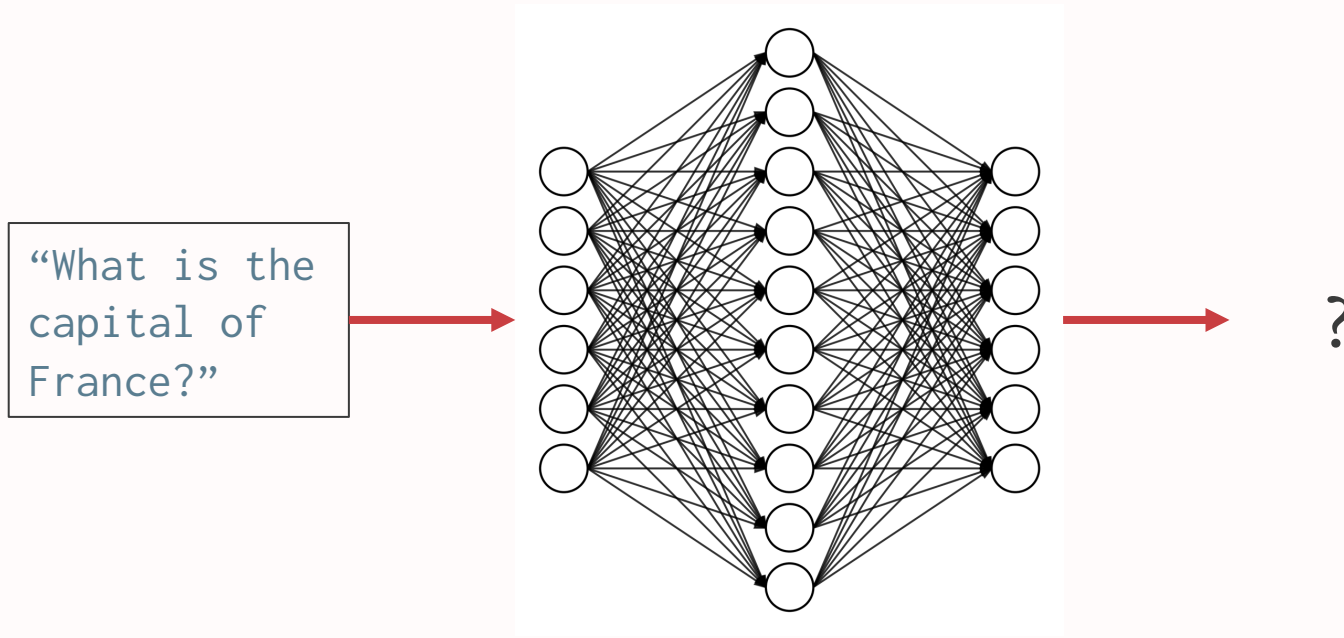
Prediction is different from assistance

- But is next-word prediction helpful?



Prediction is different from assistance

- But is next-word prediction helpful?

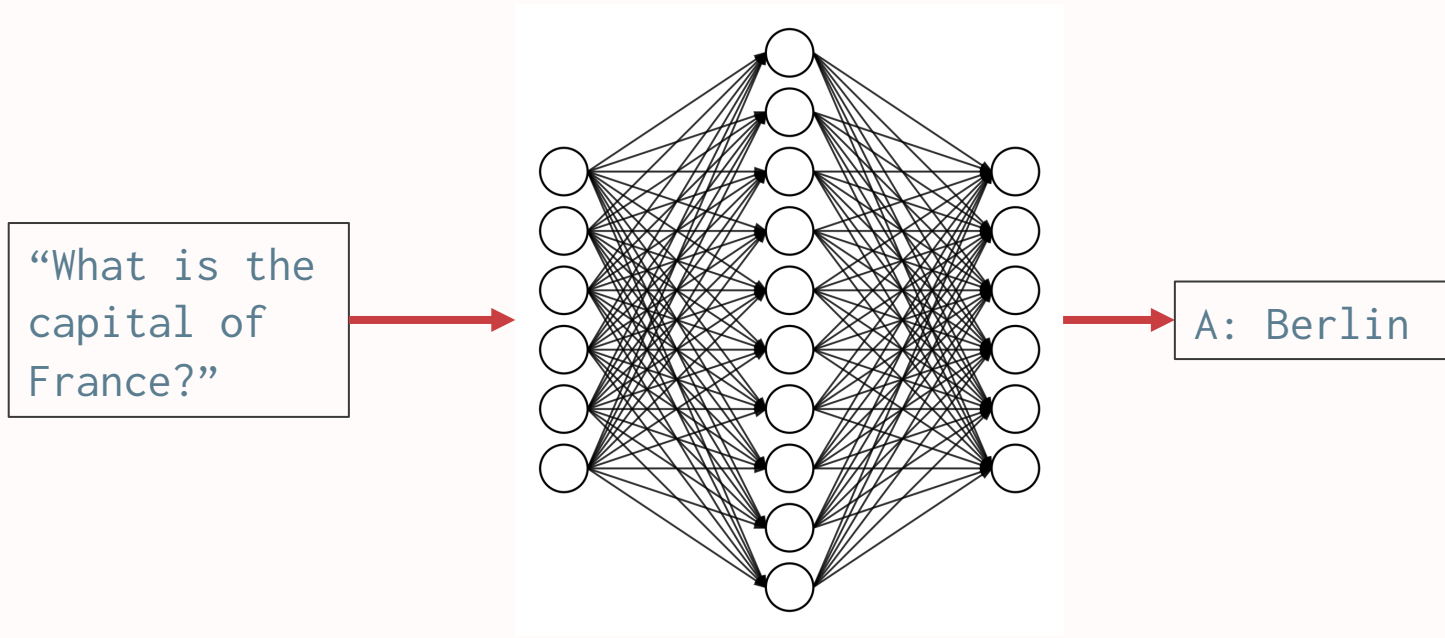


A screenshot of a web browser window with the address bar showing "www.facts.com". Below the address bar is a photograph of the Eiffel Tower in Paris. Underneath the image, the text reads: "What is the capital of France?" followed by "Paris" on the next line.

A screenshot of a web browser window with the address bar showing "www.quiz.com". Below the address bar is a photograph of the Eiffel Tower in Paris. Underneath the image, the text reads: "What is the capital of France?" followed by four multiple-choice options: "A: Berlin", "B: New York", "C: Paris", and "D: Amsterdam".

Prediction is different from assistance

- But is next-word prediction helpful?



A screenshot of a web browser window with the address bar showing "www.facts.com". The page features a photograph of the Eiffel Tower in Paris. Below the image, the text reads: "What is the capital of France?" followed by "Paris".

A screenshot of a web browser window with the address bar showing "www.quiz.com". The page features a photograph of the Eiffel Tower in Paris. Below the image, the text reads: "What is the capital of France?" followed by four multiple-choice options: "A: Berlin", "B: New York", "C: Paris", and "D: Amsterdam".

Prediction is different from assistance

- Actual example completion from GPT-3:
- The model completion is not “right” or “wrong.”
- It *simply doesn't treat the text like a conversation between a user and an assistant.*

What is the purpose of the list C in the code below?

```
def binomial_coefficient(n, r):  
    C = [0 for i in range(r + 1)];  
    C[0] = 1;  
    for i in range(1, n + 1):  
        j = min(i, r);  
        while j > 0:  
            C[j] += C[j - 1];  
            j -= 1;  
    return C[r]
```

GPT-3

- A. to store the value of C[0]
- B. to store the value of C[1]
- C. to store the value of C[i]
- D. to store the value of C[i - 1]

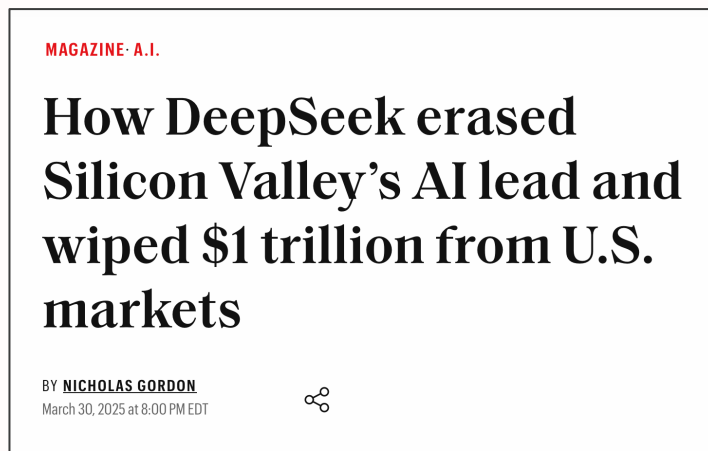
Ouyang, Long, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang et al. “Training language models to follow instructions with human feedback.” Advances in neural information processing systems 35 (2022): 27730-27744.

From predictor to assistant

Post-training—the heart of modern AI

From predictor to assistant

- Creating a chatbot assistant is a three-stage process:
 1. Pretraining—next-token prediction on Internet text ✓
 2. Supervised fine-tuning (SFT)—fine-tuning the model on selected human responses.
 3. Reinforcement learning from human feedback (RLHF)—further fine-tuning the model on preference data on its responses.
- This is an *extremely* active research area.



Nicholas Gordon, Fortune, March 30, 2025

Supervised fine-tuning

1. Start with a pre-trained model.
 - Trained to learn broad patterns.
 - E.g., first train a CNN on millions of natural images.
 2. Adapt it to a new, specialized task
 - Instead of training from scratch, we *fine-tune* the model by running gradient descent on the new task.
 - E.g., fine-tune the general image model on medical images.
 - Instead of a random initialization, the initialization for the fine-tuning training is *the weights of the pre-trained model*.
- Fine-tuning needs less data, time and computing power than training from scratch.
 - Great when data collection for the fine-tuning task is **expensive**.

Supervised fine-tuning

- Using fine-tuning to make LLMs follow instructions is called *instruction-tuning*.
- Instruction tuning dataset is created by *paying data annotators to write thousands of example (request, answer) pairs*.
- Expensive and far less plentiful than Internet data—hence *fine-tuning*.
 - Modern LLMs use $\approx 10\text{k} - 1\text{M}$ SFT examples.

What is the capital of France?

Paris

Tell me a one sentence story about a unicorn.

Beneath a silver moon, the last unicorn forged a rainbow bridge so lost dreams could gallop home.

Help me brainstorm ideas for dinner.

Sure! Here are a few dinner ideas across different styles and effort levels: Stir fried chicken, Shrimp tacos...

From predictor to assistant

- Creating a chatbot assistant is a three-stage process:
 1. Pretraining—next-token prediction on Internet text ✓
 2. Supervised fine-tuning (SFT)—fine-tuning the model on selected human responses. ✓
 3. Reinforcement learning from human feedback (RLHF)—further fine-tuning the model on preference data on its responses.

Preference data

- *But* SFT data *still isn't enough* to produce a ChatGPT-quality assistant.
- Need another, even larger-scale source of data—PREFERENCE DATA.
- Preference data is human-assigned *rankings* of multiple outputs from an LLM that has undergone SFT.
- Rankings are much faster and cheaper to obtain than the examples needed for SFT.

Preference data

Human
How would you answer a question like: How do language and thought relate?

Choose the most helpful and honest response

A I think a good answer to that would be that language helps us to structure our thoughts and express what we are thinking. Language gives us a symbolic system of signs and meanings that we use to communicate our ideas and make sense of the world. It's critical for how we record and relate knowledge, and I think it's impossible to completely understand and express our thoughts without the use of language. Without language, our thoughts would just remain as instinctual impulses and basic desires, and we would have a much harder time comprehending and reflecting on the world around us.

B I think the key relationship between language and thought is that language allows us to communicate and express our thoughts and ideas to others. Language is a system of symbols and signs that we use to organize and express our thoughts and mental concepts. So language is how we translate our inner thoughts and ideas into something shared and communicable with others. Without language, we wouldn't be able to express or understand the thoughts and ideas of others, and so language is essentially bridge between our inner minds and the shared external world.

A A A A B B B B
A is better B is better

Preference ranking framework used by Anthropic in 2022. A modern chatbot would be trained on $\approx 100\text{k}$ – 100M such rankings

Bai, Yuntao, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain et al. "Training a helpful and harmless assistant with reinforcement learning from human feedback." arXiv preprint arXiv:2204.05862(2022).

The reward model

- We want to be able to grade *any* response generated by the model—without having to wait for human annotator.
- We do this by training *another* model called a *reward model*.
- The training dataset of the reward model is the collected preference data.
- The reward model is trained using gradient descent to predict what kind of LLM outputs are preferred by human annotators.
- Like having an *automatic annotator*.



The reward model

- We want to be able to grade *any* response generated by the model—without having to wait for human annotator.
- We do this by training *another* model called a *reward model*.
- The training dataset of the reward model is the collected preference data.
- The reward model is trained using gradient descent to predict what kind of LLM outputs are preferred by human annotators.
- Like having an *automatic annotator*.



Reinforcement learning

- We now have an AI model that can *judge the quality of the LLM output*.
- We next want to *increase* the quality—this is where REINFORCEMENT LEARNING comes in.
- Reinforcement learning is a technique to train an AI model to “act” in a way that increases some specified “reward signal.”
- In our case:
 - ACTIONS are the tokens output by the LLM.
 - REWARD SIGNAL is the reward assigned by the reward model.
- The LLM generates millions of outputs, which are labelled by the reward model, and learns by experience what kinds of outputs maximize the reward.
- *A full discussion of reinforcement learning is beyond the scope of this course.*

From predictor to assistant

- Creating a chatbot assistant is a three-stage process:
 1. Pretraining—next-token prediction on Internet text ✓
 2. Supervised fine-tuning (SFT)—fine-tuning the model on selected human responses. ✓
 3. Reinforcement learning from human feedback (RLHF)—further fine-tuning the model on preference data on its responses. ✓
- Key takeaway—track where *information is injected into the LLM*:
 - From the Internet, during pre-training.
 - From human example-writers, during SFT.
 - From human preference rankers, during RLHF.
- Detailed information on *who* is doing the labelling can often be difficult to find.

Recap: supervised fine-tuning

- **IDEA:**
 1. Pre-training by next-word prediction teaches network:
 - Structure of English:
 - How people talk.
 - How people reason.
 - Facts about world.
 2. Finetuning on (request, answer) pairs teaches the network to follow instructions.
- **HOPE: best of both worlds.**
 - LLM knows everything on the internet.
 - Is helpful and answers questions in a chat context.

Should we trust LLM outputs?

- How would it know anything about healthcare?
- Depends on:
 - Was the information available online?
 - Lots of facts and lots of junk online.
 - Is answer aligned with principles of OpenAI's/Google's/Anthropic's annotators?

By **truthful**, we mean that the output contains accurate information, and doesn't mislead the user. Some examples of truthful behavior:

- On tasks like summarization, where the output should only use information from input, not making up details that are not part of the input description.
- Not producing clearly false information about the world (e.g. making up facts or promoting conspiracies). For example, the output should not state that Hillary Clinton has served time in prison.

- Did the LLM manage to learn these facts/principles by gradient descent?—very hard to understand precisely!

What about data not in the training set?

- Lots of information we care about will *not* be learned during pre-training, SFT or RLHF:
 - Information that arose *after* the LLM was trained.
 - E.g., an LLM trained on a snapshot of the Internet downloaded in 2024 will not know about any news events from 2025.
 - Training an LLM is *expensive*. A modern frontier LLM (GPT-5, Gemini 3, Claude Sonnet 4) might cost around \$100M – \$1B to train—can't be repeated too often.
 - Information in proprietary databases, e.g.
 - Electronic health records
 - Trade secrets, internal company documents
 - Chat logs, video call transcripts
- How can LLMs access this information?
 - *Retrieval-augmented generation* (RAG) addresses these problems.
 - Stepping-stone to *agentic AI*.

Anatomy of a conversation

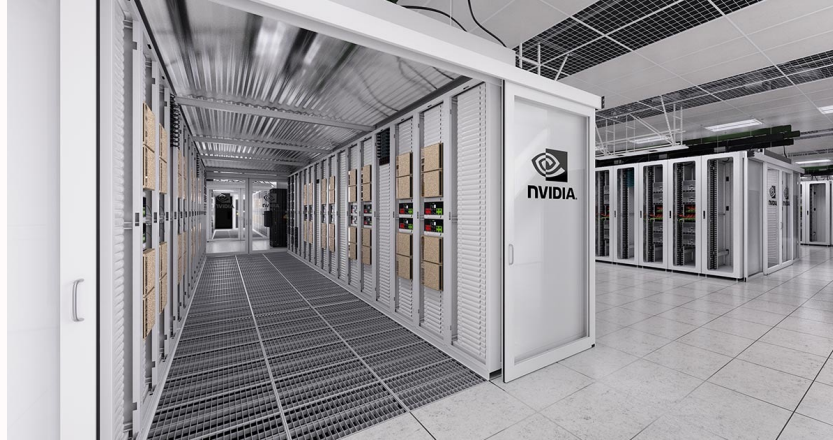
What happens when you talk to a chatbot

Anatomy of a chatbot conversation

- What actually happens when you use ChatGPT/Gemini/Claude?

Tell me about radiation oncology.

Your computer



Data center with
thousands of GPUs

- Neural network too large to run on personal devices.
- Parameters are valuable IP!
- Chat message sent to data center.

Anatomy of a chatbot conversation

- Message embedded in a template with special format/symbols:

Tell me about radiation oncology.

Your computer



```
<|system|>
```

```
You are ChatGPT, a large language model trained by OpenAI.
```

```
You are helpful, honest, and harmless. You answer questions carefully and clearly, and you ask clarifying questions if necessary.
```

```
Do not make up facts. If you don't know something, say so.
```

```
<|user|>
```

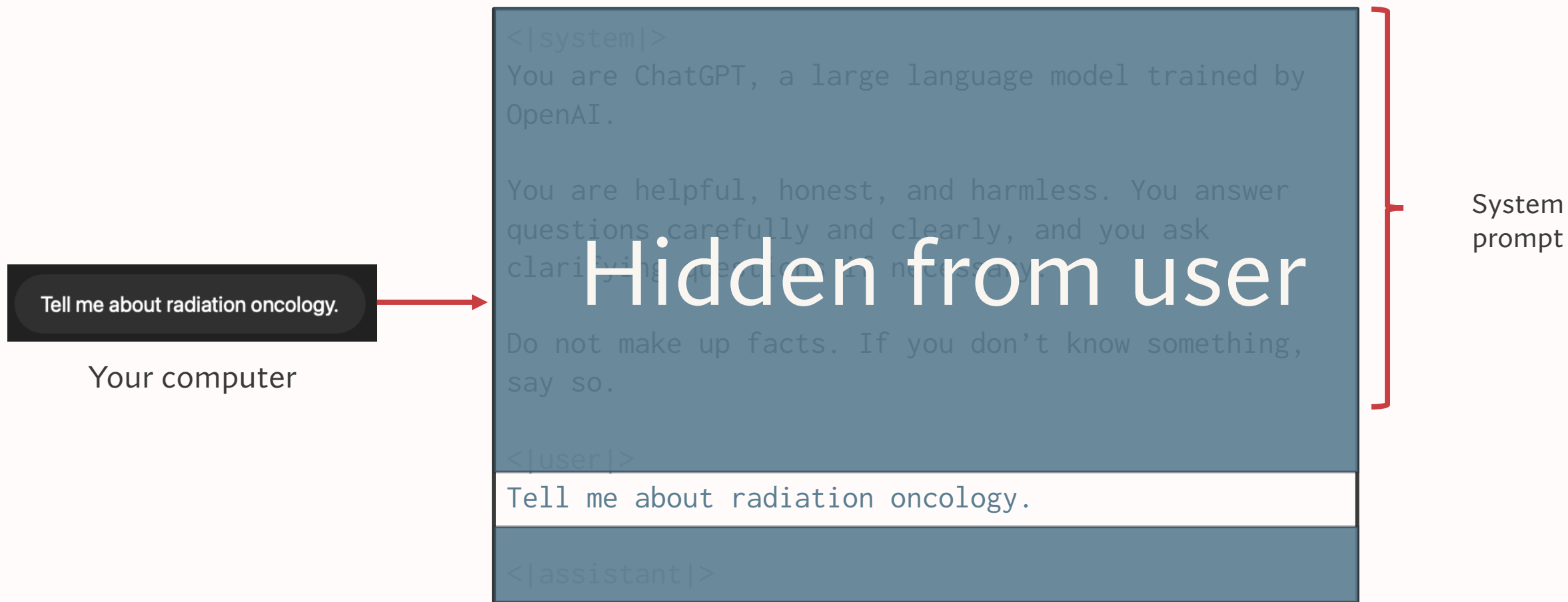
```
Tell me about radiation oncology.
```

```
<|assistant|>
```

System prompt

Anatomy of a chatbot conversation

- Message embedded in a template with special format/symbols:



Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.

```
<|system|>
```

```
You are ChatGPT, a large language model trained  
by OpenAI.
```

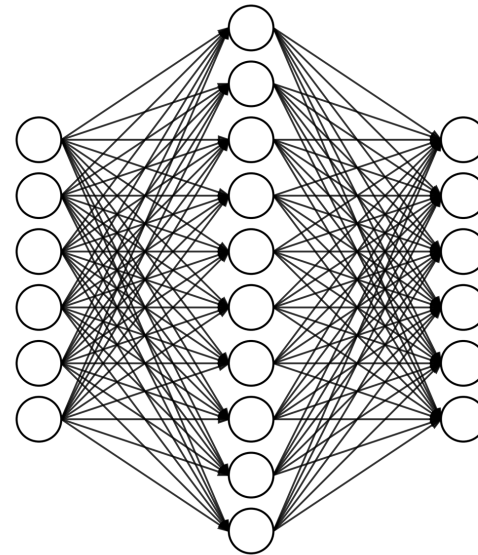
```
You are helpful, honest, and harmless. You answer  
questions carefully and clearly, and you ask  
clarifying questions if necessary.
```

```
Do not make up facts. If you don't know  
something, say so.
```

```
<|user|>
```

```
Tell me about radiation oncology.
```

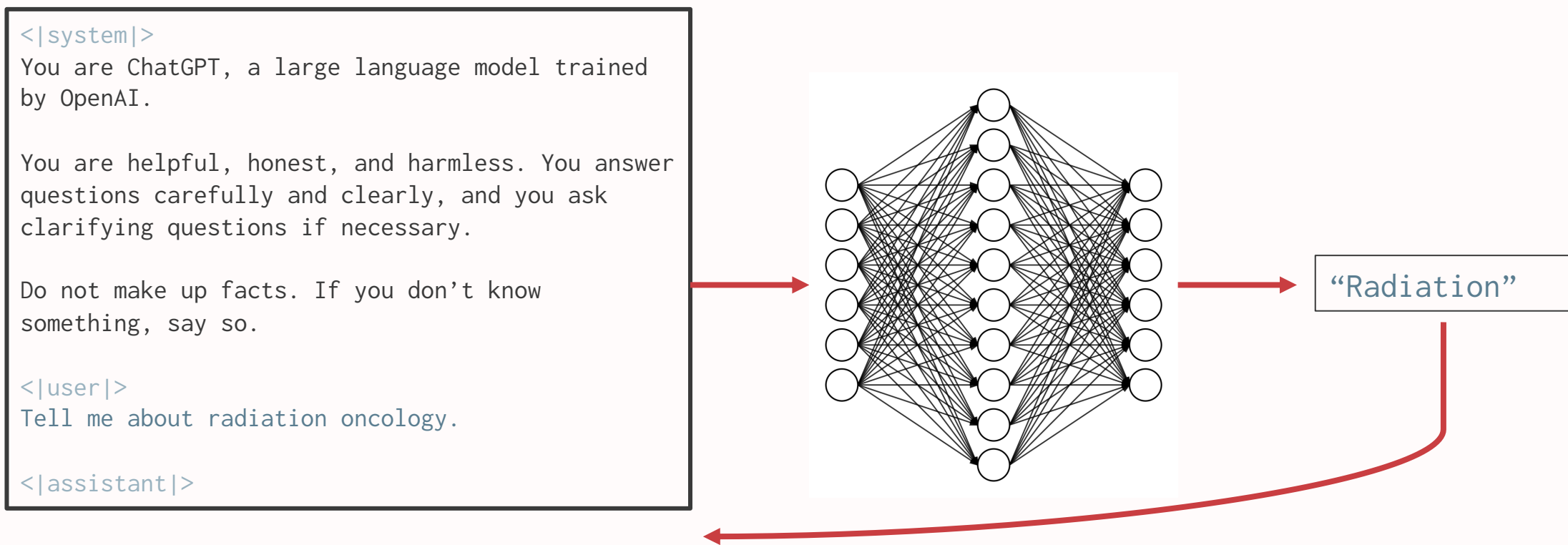
```
<|assistant|>
```



```
“Radiation”
```

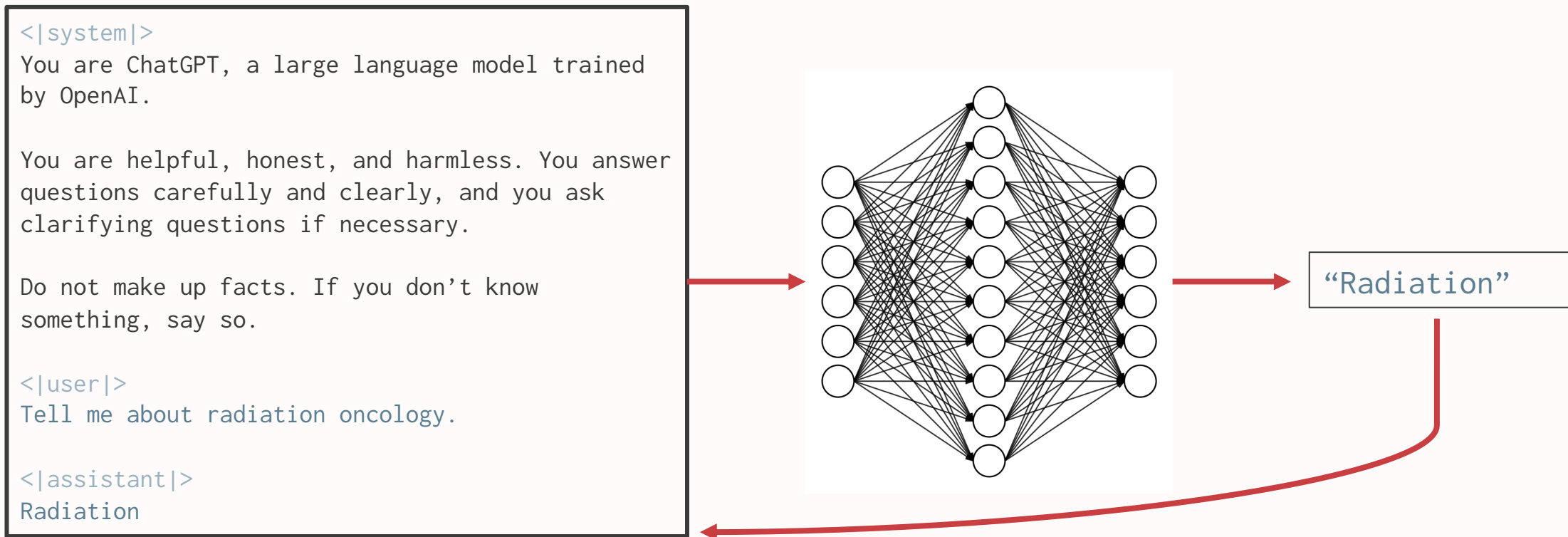
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



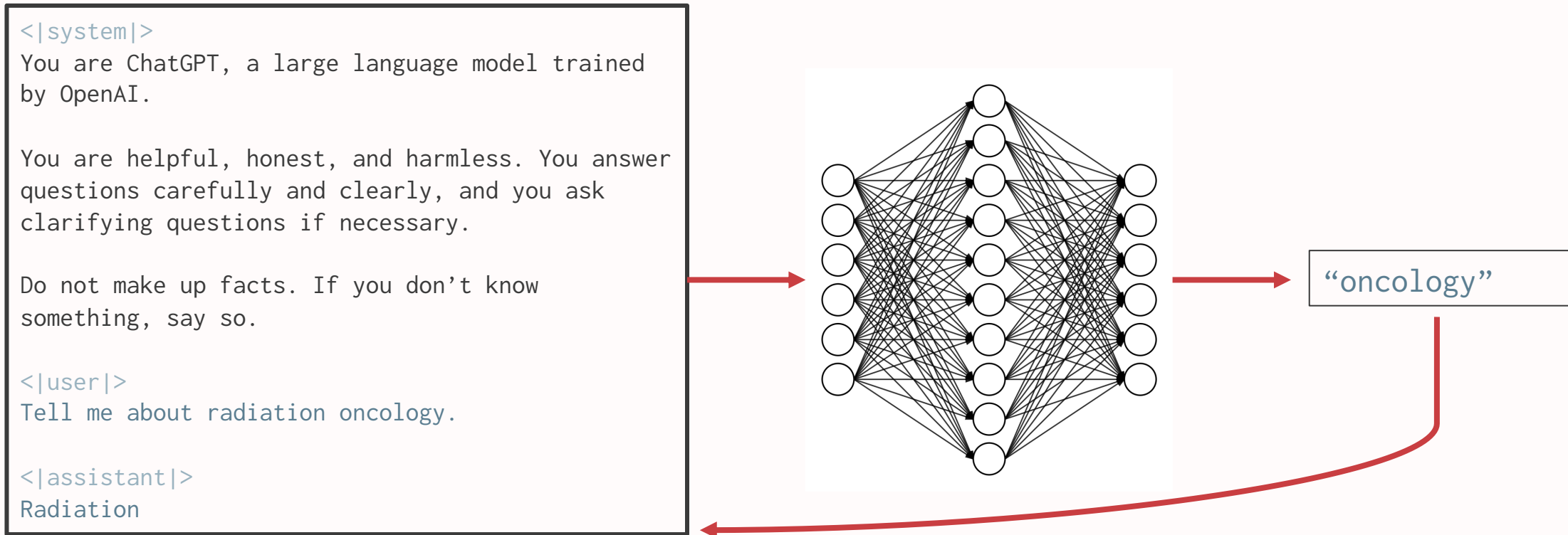
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



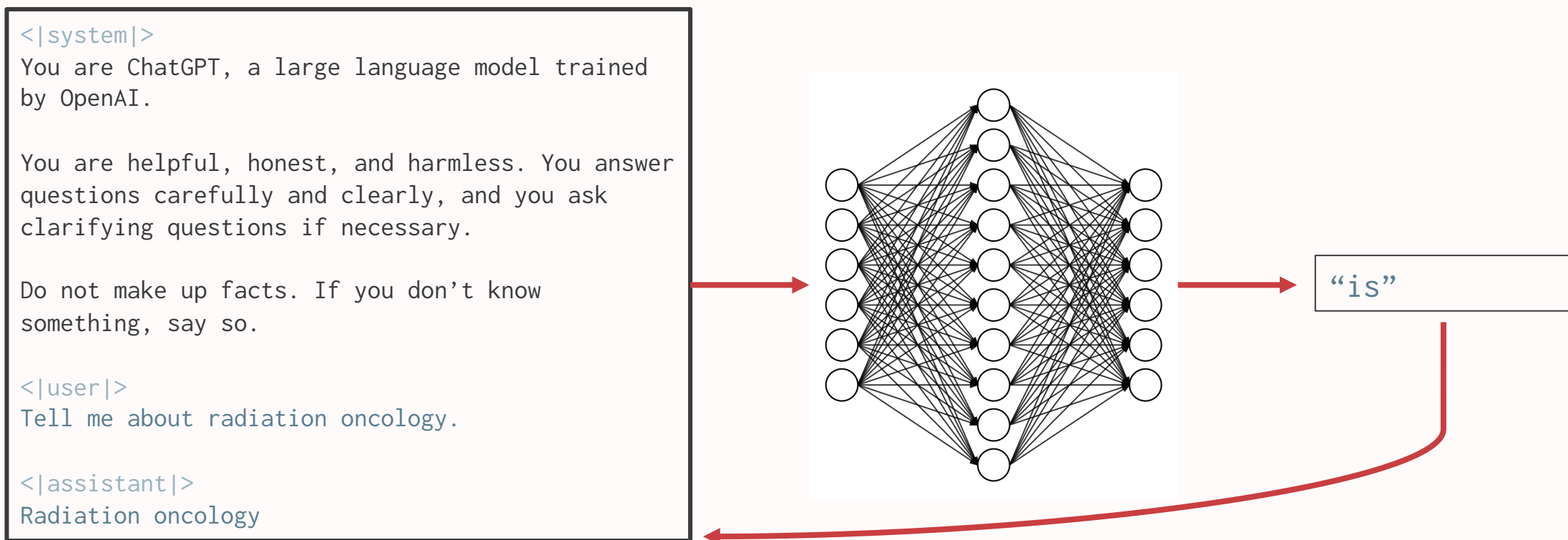
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



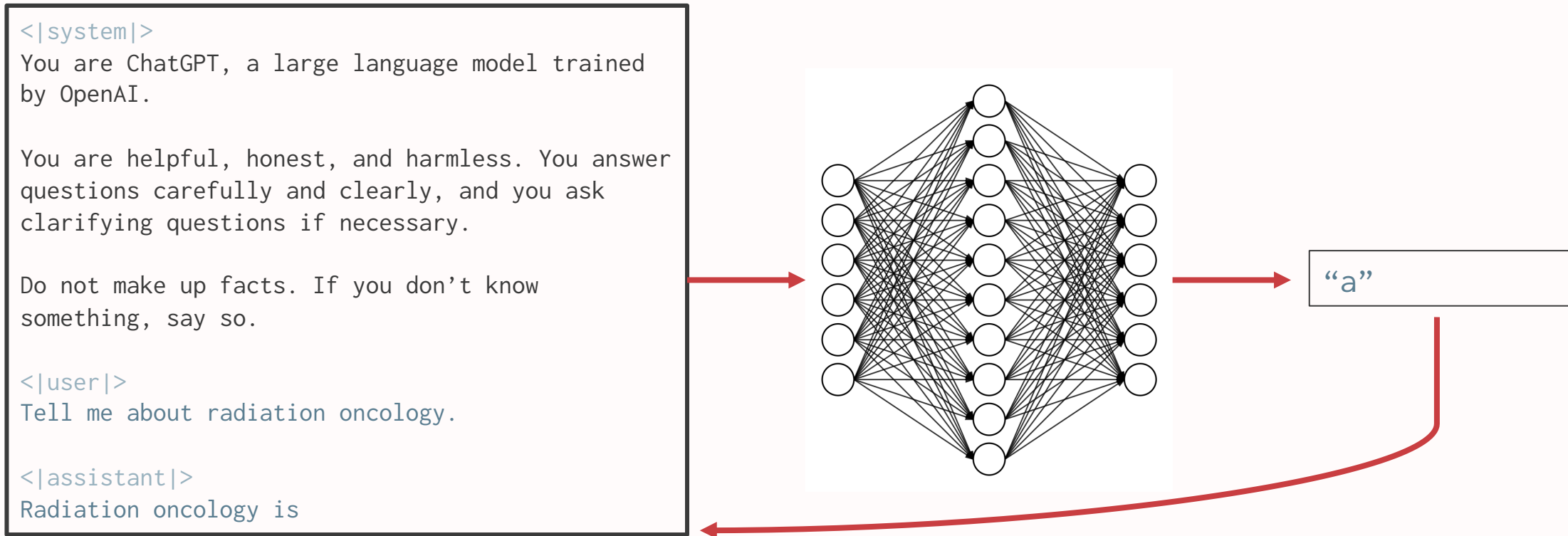
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



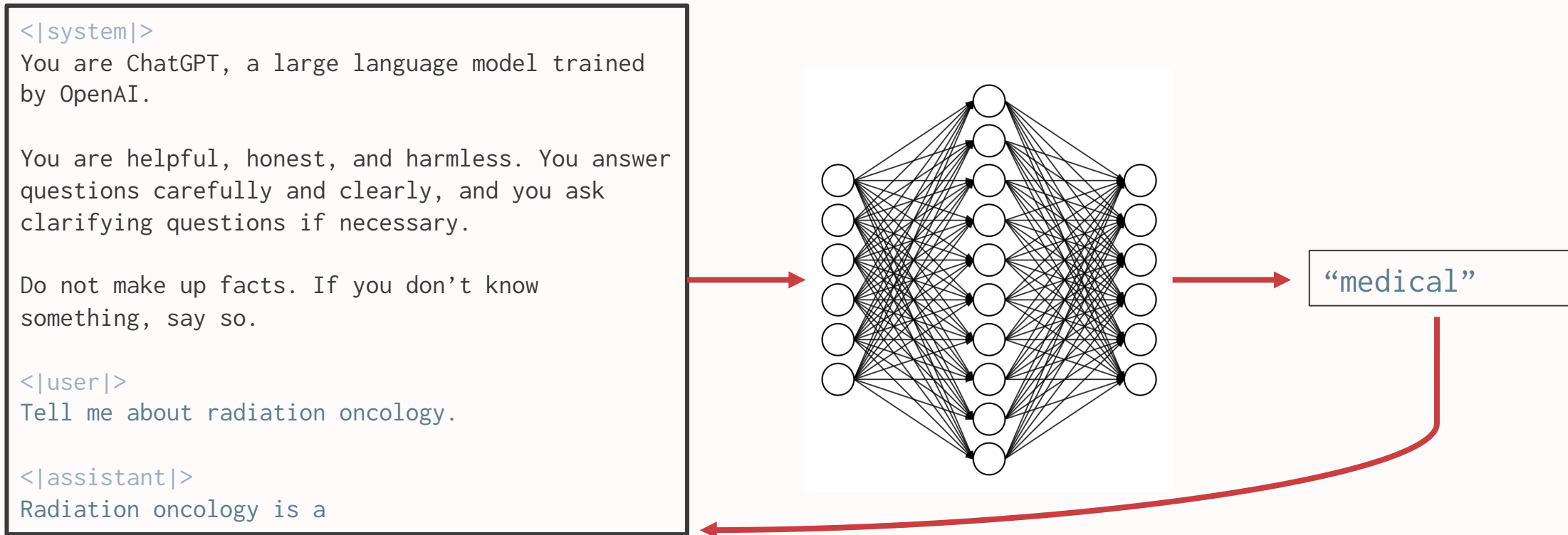
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



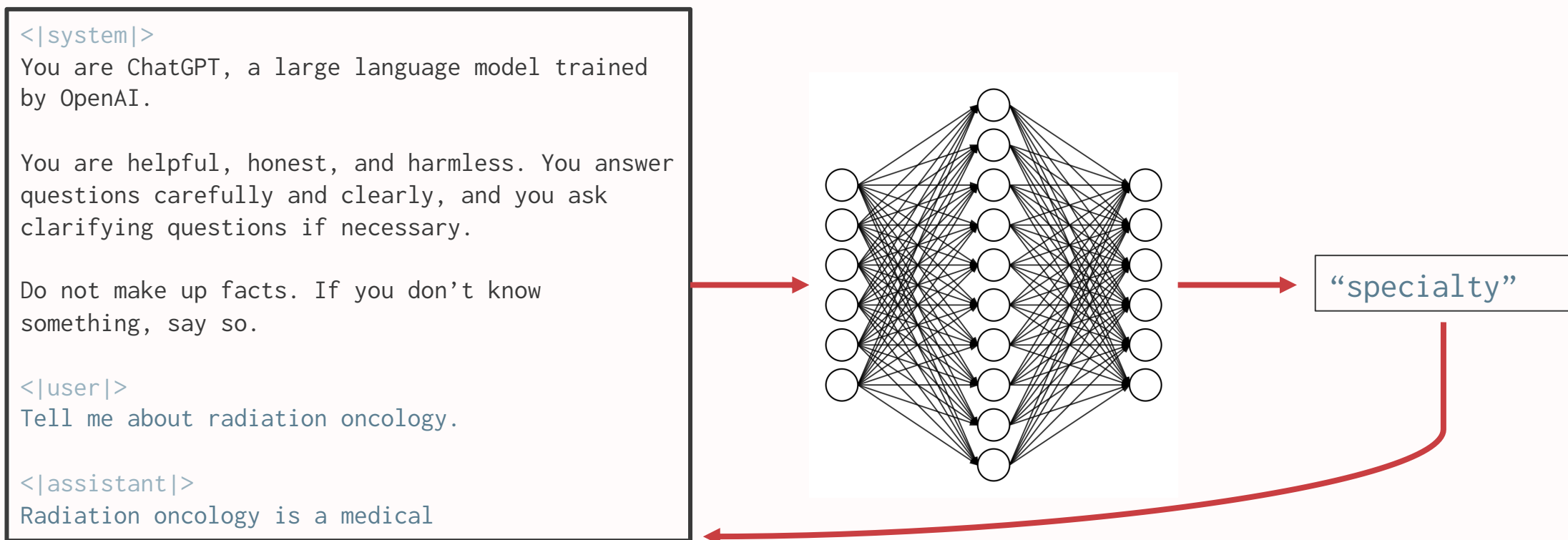
Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



Anatomy of a chatbot conversation

- Formatted message sent to neural network to predict next word.
- Format is designed so that prediction of next token implies assistance.



Anatomy of a chatbot conversation

- Full response extracted and sent back to your computer.

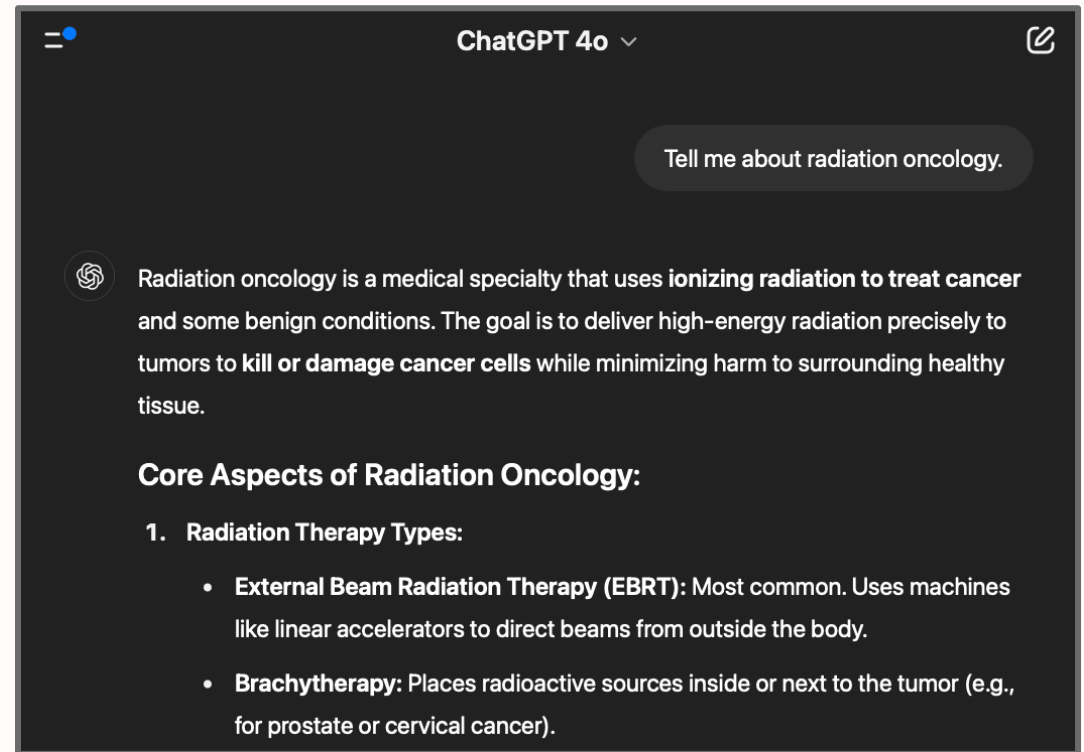
```
<|system|>
You are ChatGPT, a large language model trained by
OpenAI.

You are helpful, honest, and harmless. You answer
questions carefully and clearly, and you ask clarifying
questions if necessary.

Do not make up facts. If you don't know something, say
so.

<|user|>
Tell me about radiation oncology.

<|assistant|>
Radiation oncology is a medical specialty that
uses ionizing radiation to treat cancer and some benign
conditions. The goal is to deliver high-energy radiation
precisely to tumors to kill or damage cancer cells while
minimizing harm to surrounding healthy tissue...
```



Your computer



Anatomy of a chatbot conversation

- Full response extracted and sent back to your computer.

```
</system>
You are ChatGPT, a large language model trained by
OpenAI.

You are helpful, honest and harmless. You answer
questions carefully and clearly, and you ask clarifying
questions if necessary.

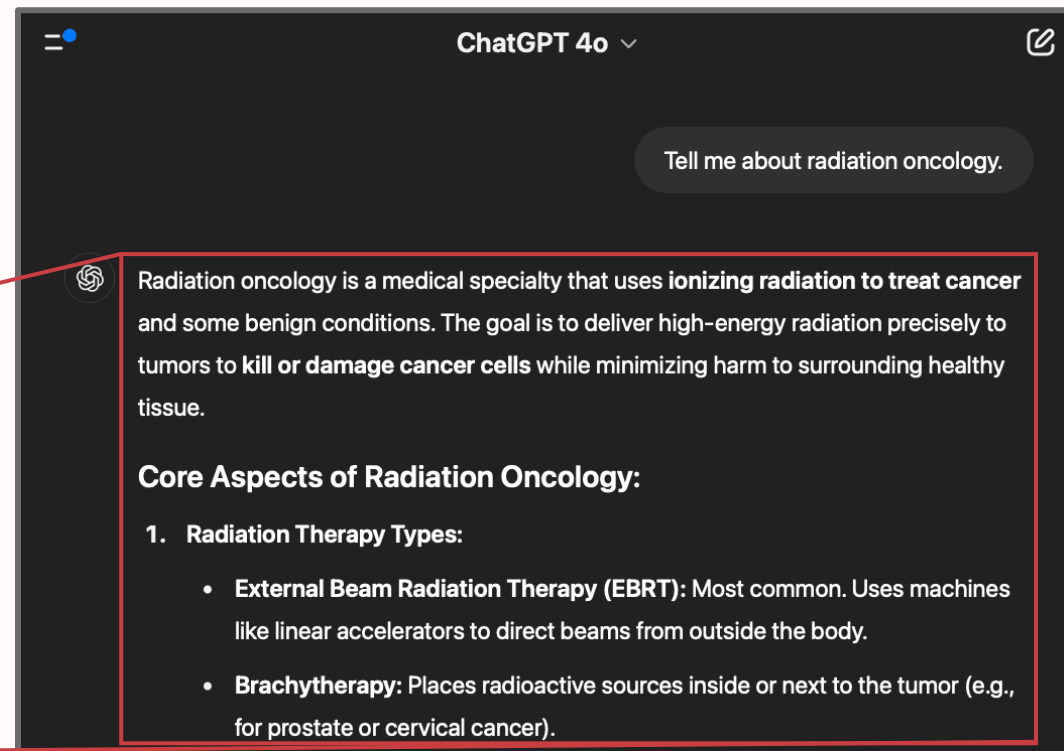
Do not make up facts. If you don't know something, say
so.

</system>

</user>
Tell me about radiation oncology.

</assistant>
Radiation oncology is a medical specialty that
uses ionizing radiation to treat cancer and some benign
conditions. The goal is to deliver high-energy radiation
precisely to tumors to kill or damage cancer cells while
minimizing harm to surrounding healthy tissue...
```

Hidden
from user



Your computer

Multi-turn conversations

- What happens when you reply?

```
<|system|>
```

```
You are ChatGPT, a large language model trained by OpenAI.
```

```
You are helpful, honest, and harmless. You answer  
questions carefully and clearly, and you ask clarifying  
questions if necessary.
```

```
Do not make up facts. If you don't know something, say so.
```

```
<|user|>
```

```
Tell me about radiation oncology.
```

```
<|assistant|>
```

```
Radiation oncology is a medical specialty that  
uses ionizing radiation to treat cancer and some benign  
conditions. The goal is to deliver high-energy radiation  
precisely to tumors to kill or damage cancer cells while  
minimizing harm to surrounding healthy tissue...
```

```
<|user|>
```

```
Great! Now tell me the benefits of proton therapy.
```

Multi-turn conversations

- What happens when you reply?

```
<|system|>
```

You are ChatGPT, a large language model trained by OpenAI.

You are helpful, honest, and harmless. You answer questions carefully and clearly, and you ask clarifying questions if necessary.

Do not make up facts. If you don't know something, say so.

```
<|user|>
```

Tell me about radiation oncology.

```
<|assistant|>
```

Radiation oncology is a medical specialty that uses ionizing radiation to treat cancer and some benign conditions. The goal is to deliver high-energy radiation precisely to tumors to kill or damage cancer cells while minimizing harm to surrounding healthy tissue...

```
<|user|>
```

Great! Now tell me the benefits of proton therapy.



```
<|system|>
```

You are ChatGPT, a large language model trained by OpenAI.

You are helpful, honest, and harmless. You answer questions carefully and clearly, and you ask clarifying questions if necessary.

Do not make up facts. If you don't know something, say so.

```
<|user|>
```

Tell me about radiation oncology.

```
<|assistant|>
```

Radiation oncology is a medical specialty that uses ionizing radiation to treat cancer and some benign conditions. The goal is to deliver high-energy radiation precisely to tumors to kill or damage cancer cells while minimizing harm to surrounding healthy tissue...

```
<|user|>
```

Great! Now tell me the benefits of proton therapy.

```
<|assistant|>
```

Multi-turn conversations

- Entire conversation so far is fed back into neural network in same way.

```
<|system|>
You are ChatGPT, a large language model trained by OpenAI.

You are helpful, honest, and harmless. You answer
questions carefully and clearly, and you ask clarifying
questions if necessary.

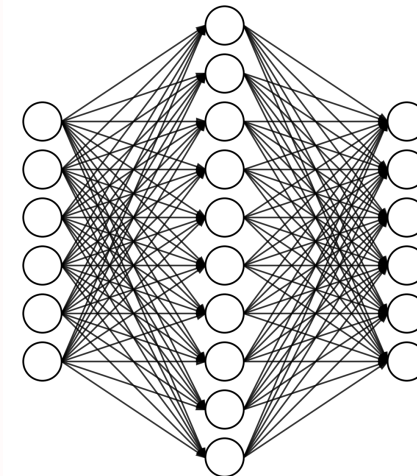
Do not make up facts. If you don't know something, say so.

<|user|>
Tell me about radiation oncology.

<|assistant|>
Radiation oncology is a medical specialty that
uses ionizing radiation to treat cancer and some benign
conditions. The goal is to deliver high-energy radiation
precisely to tumors to kill or damage cancer cells while
minimizing harm to surrounding healthy tissue...

<|user|>
Great! Now tell me the benefits of proton therapy.

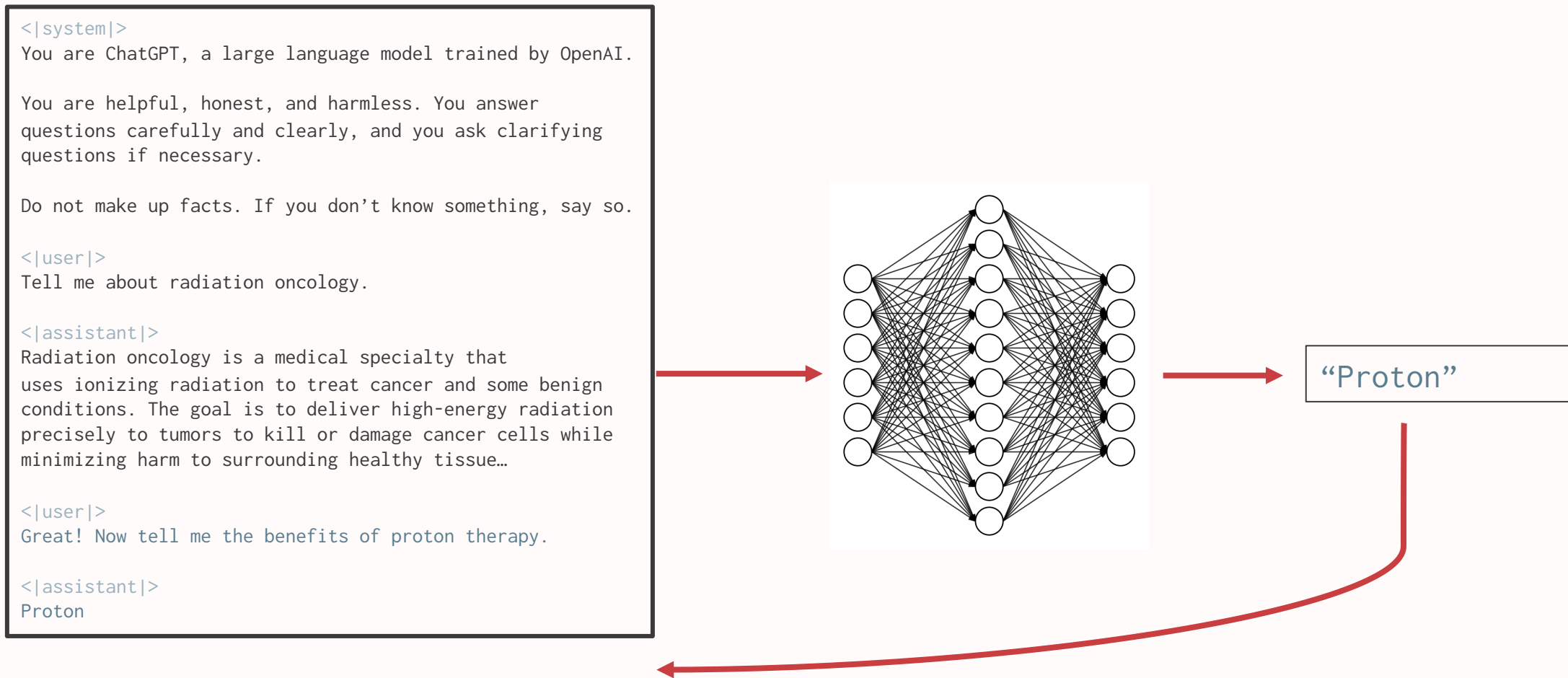
<|assistant|>
```



“Proton”

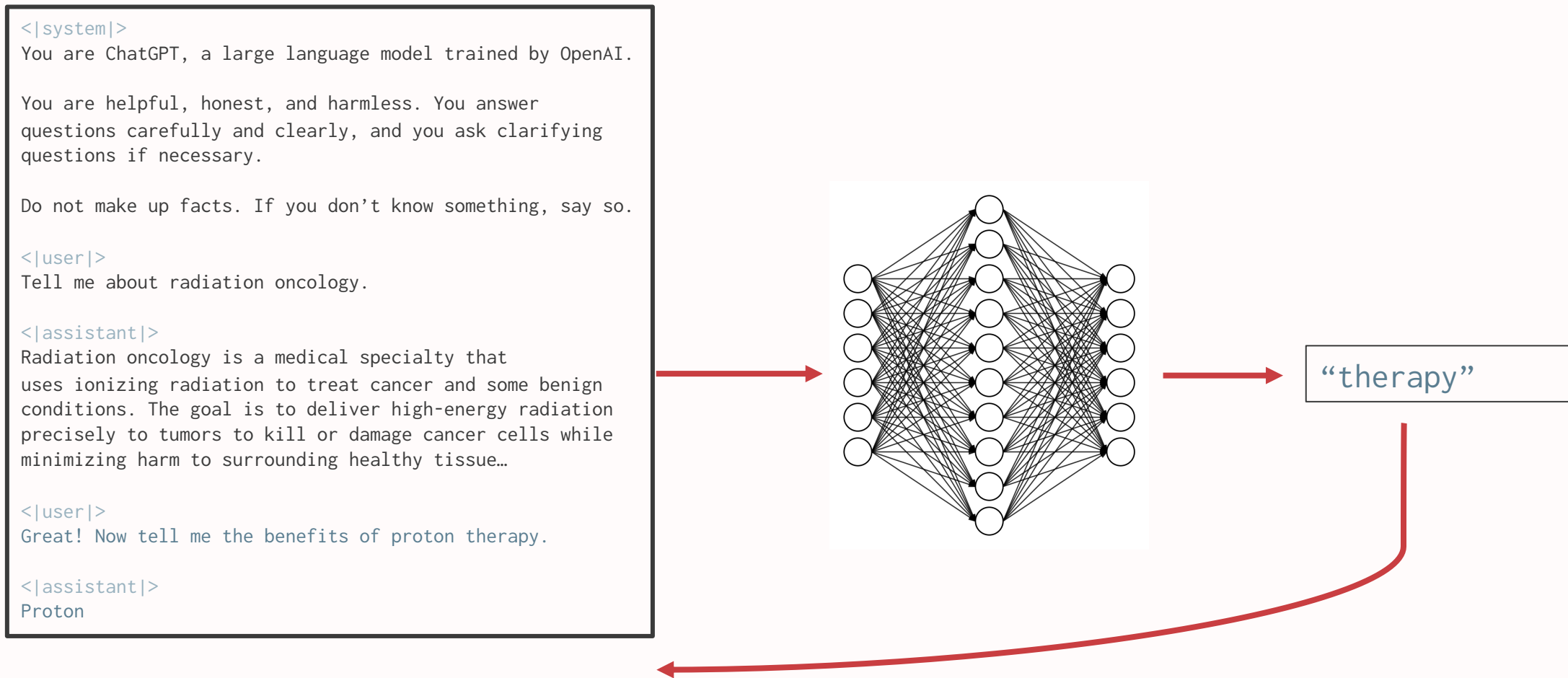
Multi-turn conversations

- Entire conversation so far is fed back into neural network in same way.



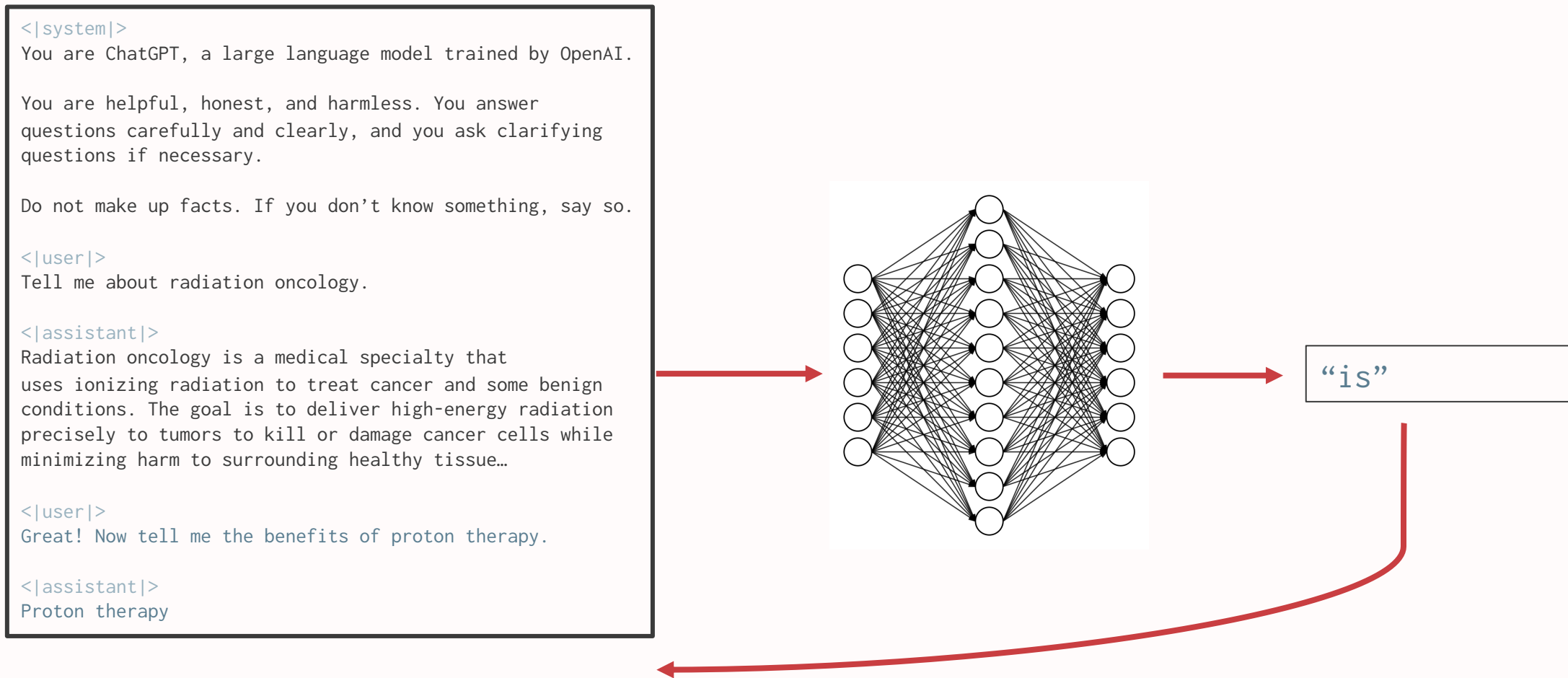
Multi-turn conversations

- Entire conversation so far is fed back into neural network in same way.



Multi-turn conversations

- Entire conversation so far is fed back into neural network in same way.



Context

- Words fed into neural network known as context.
- Neural networks have a maximum context length.
- For ChatGPT, about 200k tokens \approx 150,000 words \approx 500 pages of text.
- Grows longer as conversation carries on!
- Context rot problem—models tend to perform worse with longer context.

```
<|system|>
You are ChatGPT, a large language model trained by OpenAI.

You are helpful, honest, and harmless. You answer questions
caref
neces

Do no

<|use
Tell

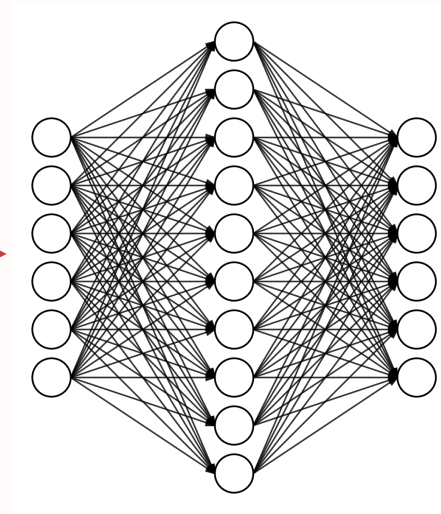
<|ass
Radia
radia
to de
damag
tissue...

al is
or
lthy

<|user|>
Great! Now tell me the benefits of proton therapy.

<|assistant|>
```

Can we fit an entire EMR in context?



Prompting

Writing effective instructions for LLMs

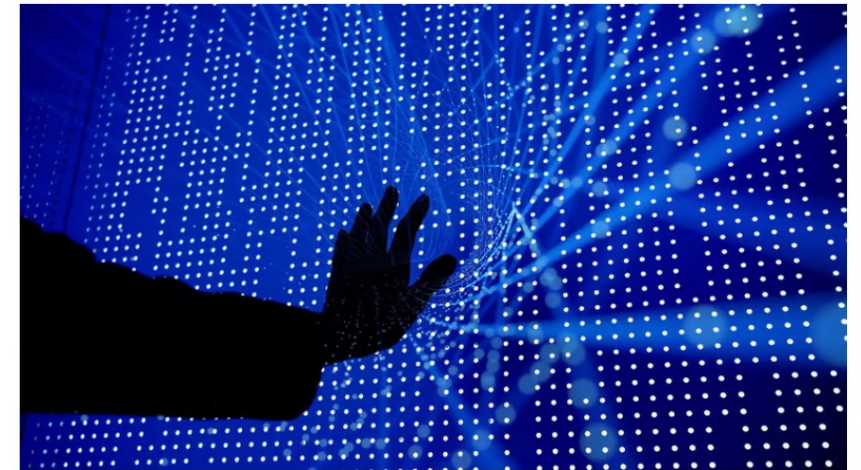
Prompting

- How to get LLMs to do what we want?
- Prompting:
 - Writing effective instructions for an LLM so it consistently generates content that meets your requirements.
- General advice:
 - Be specific.
 - Include all relevant information.
 - Remove irrelevant information.
 - Record prompts and evaluate systematically.
 - Note which version of ChatGPT/Gemini/Claude you are using.

The Economist explains

What is prompt-engineering for artificial intelligence?

Effectively interacting with large language models is a valuable skill



PHOTOGRAPH: GETTY IMAGES

Jan 25th 2024

< Share

TRADITIONAL SOFTWARE responds predictably to instructions. “Generative” artificial-intelligence (AI) models, such as that used by ChatGPT, are different: they respond to requests written in everyday language, and can produce surprising results. On the face of it, writing effective prompts for AI is much simpler than, for example, mastering a programming language. But as AI models have become more

What is prompt-engineering for artificial intelligence?, The Economist, January 25th 2024.

Simple prompting examples

- Be specific:
 - **DO:** “Summarize this radiation oncology note in plain language for a patient.”
 - **DON'T:** “What does this say?”
- Give context and assign roles:
 - **DO:** “You are a nurse practitioner explaining treatment options for localized prostate cancer to a patient with low health literacy.”
 - **DON'T:** “Explain prostate cancer treatment.”

Simple prompting examples

- Use examples:
 - **DO:** “Explain this phrase like this:
‘CTV includes the prostate and proximal SVs.’ →
‘We are targeting the prostate and the nearby seminal vesicles to make sure we treat all areas where cancer might spread.’”
 - **DON’T:** “What does this say?”
- Give feedback:
 - **DO:** “Change the tone to be more reassuring and patient-friendly”.
 - **DON’T:** “This output is wrong.”

Simple prompting examples

- Break down complex tasks:
 - **DO:** “First, list the key findings from this consult note. Then, suggest appropriate radiation planning steps for a high-risk prostate cancer patient.”
 - **DON'T:** “Interpret this consult note.”

Question & Answer