深度學習Deep Learning (13)

112-1

朱學亭老師

課程大綱

- W1-課程介紹/Introduction
- W2-Python/Colab and TensorFlow
- W3-Numpy/Pandas and PyTorch
- W4-Sklearn and 機器學習
- W5-神經網路, TensorFlow, PyTorch
- W6-載客熱點預測
- W7-自動光學檢查(AOI)-1
- W8-自動光學檢查(AOI)-2
- W9-Midterm presentation
- W10-**RNN**
- W11-YoloV5
- W12-AICUP 1
- W13-AICUP 2
- W14-GAN
- W15-NLP1
- W16-NLP2
- W17-Final presentation(1)
- W18-Final presentation(2)

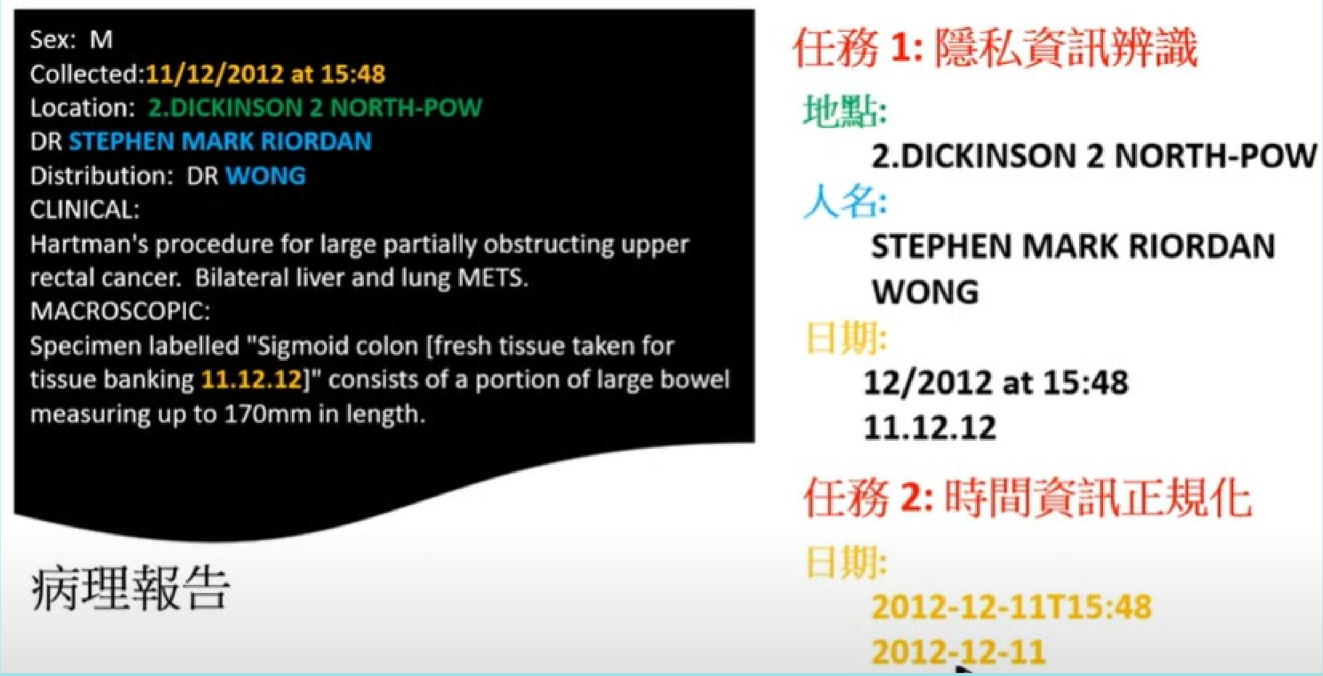
大綱

- Topic 1: AICUP
- Topic 2: LLM
- Topic 3: LLM-based Data De-identification

Topic 1: AICUP

什麼是AI競賽

- AI=大數據+深度學習



Sex: M
Collected:11/12/2012 at 15:48
Location: 2.DICKINSON 2 NORTH-POW
DR STEPHEN MARK RIORDAN
Distribution: DR WONG
CLINICAL:
Hartman's procedure for large partially obstructing upper rectal cancer. Bilateral liver and lung METS.
MACROSCOPIC:
Specimen labelled "Sigmoid colon [fresh tissue taken for tissue banking 11.12.12]" consists of a portion of large bowel measuring up to 170mm in length.

病理報告

任務 1: 隱私資訊辨識

地點:
2.DICKINSON 2 NORTH-POW

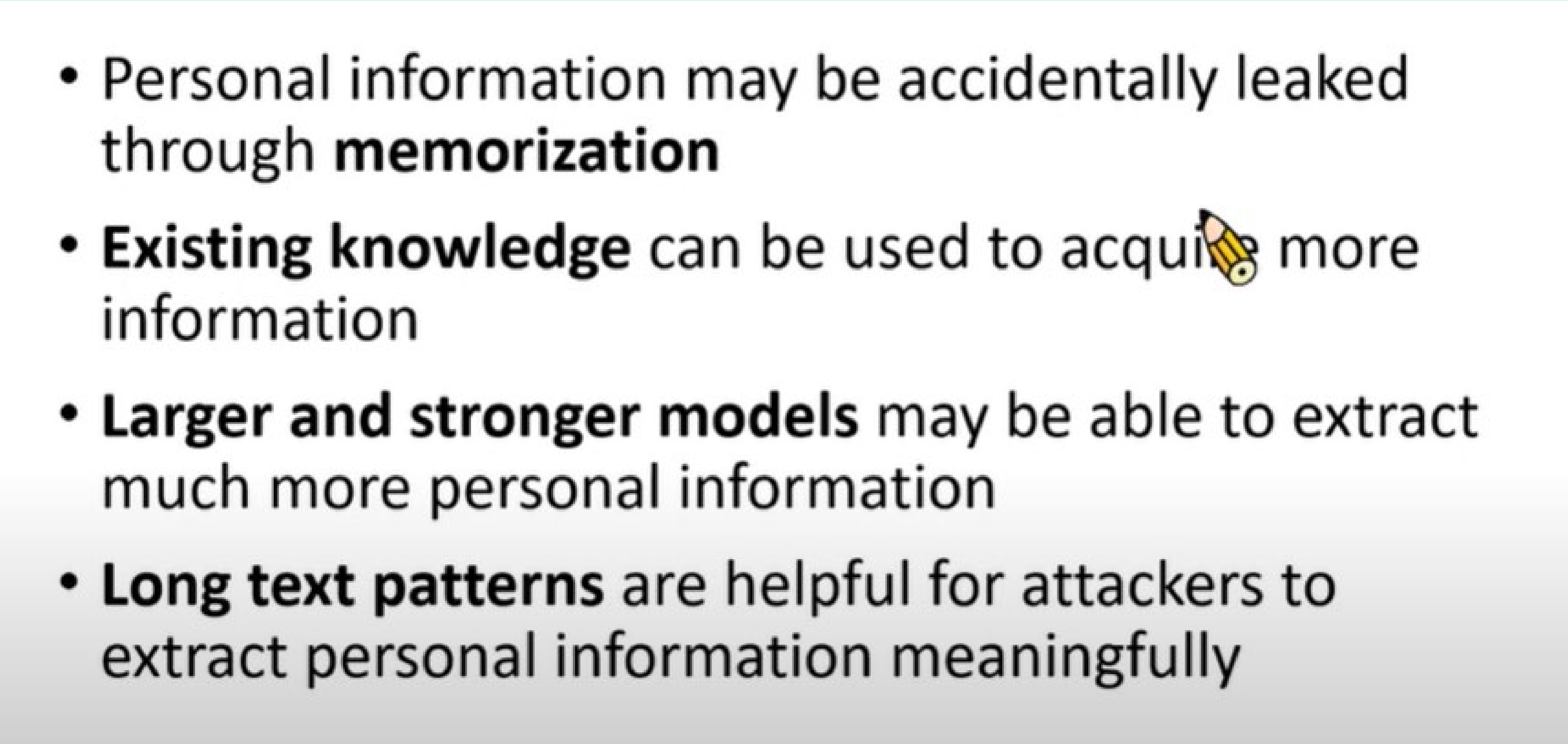
人名:
STEPHEN MARK RIORDAN
WONG

日期:
12/2012 at 15:48
11.12.12

任務 2: 時間資訊正規化

日期:
2012-12-11T15:48
2012-12-11

Privacy Risk of Pretrained LMs

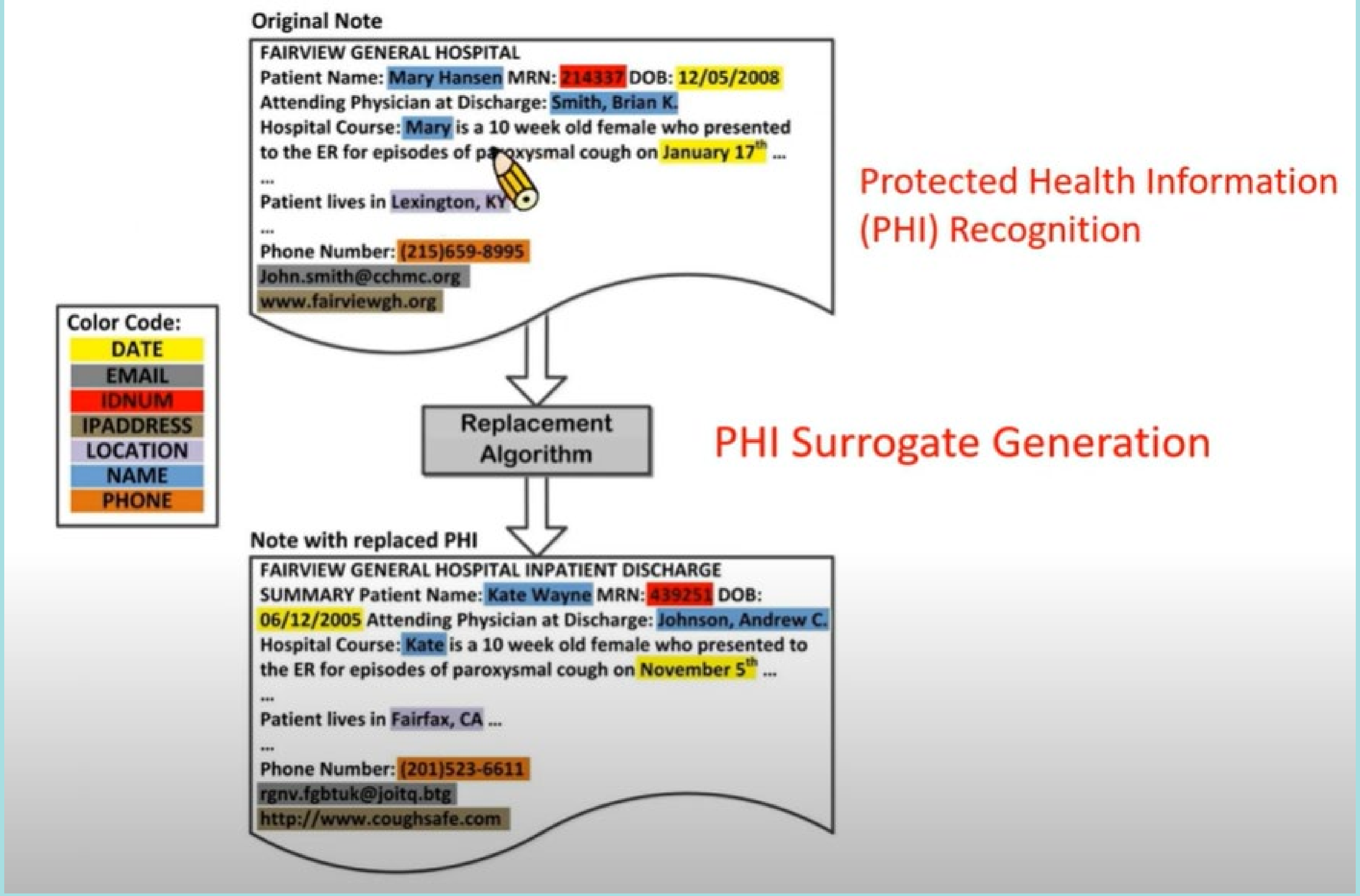


- Personal information may be accidentally leaked through **memorization**
- **Existing knowledge** can be used to acquire more information
- **Larger and stronger models** may be able to extract much more personal information
- **Long text patterns** are helpful for attackers to extract personal information meaningfully

受保護健康訊息（Protected Health Information; PHI）

PHI 類別	類型定義	範例
姓名	病患名、醫師名、人名	John Doe, Dr. Max, Mr. Smith
職業	無	lawyer, teacher
地點	診間號、部門、醫院、組織、街、城市、州國家、區號、其他	peri-operative unit-pow, macquarie ward-rhw,12 abc street
年齡	無	23, 98
日期	日期、時間、週期、頻率	24/12/1987, September 26th
聯絡方式	手機號碼、傳真、電子郵件信箱、網址、網際網路協定位址	+61-421123456、abc@gmail.com、194.223.1.1
識別符	社群安全碼、醫療紀錄號碼、健康計畫號碼帳戶、證照號碼、車牌、裝置號碼、生物識別碼、識別碼	Mrn : 9174338 Id number : 12rl500257
其它	無	

De-identification



Original Note

FAIRVIEW GENERAL HOSPITAL
Patient Name: Mary Hansen MRN: 214337 DOB: 12/05/2008
Attending Physician at Discharge: Smith, Brian K.
Hospital Course: Mary is a 10 week old female who presented to the ER for episodes of paroxysmal cough on January 17th ...
...
Patient lives in Lexington, KY
...
Phone Number: (215)659-8995
John.smith@cchmc.org
www.fairviewgh.org

Protected Health Information (PHI) Recognition

Color Code:

Color Code:
DATE
EMAIL
IDNUM
IPADDRESS
LOCATION
NAME
PHONE

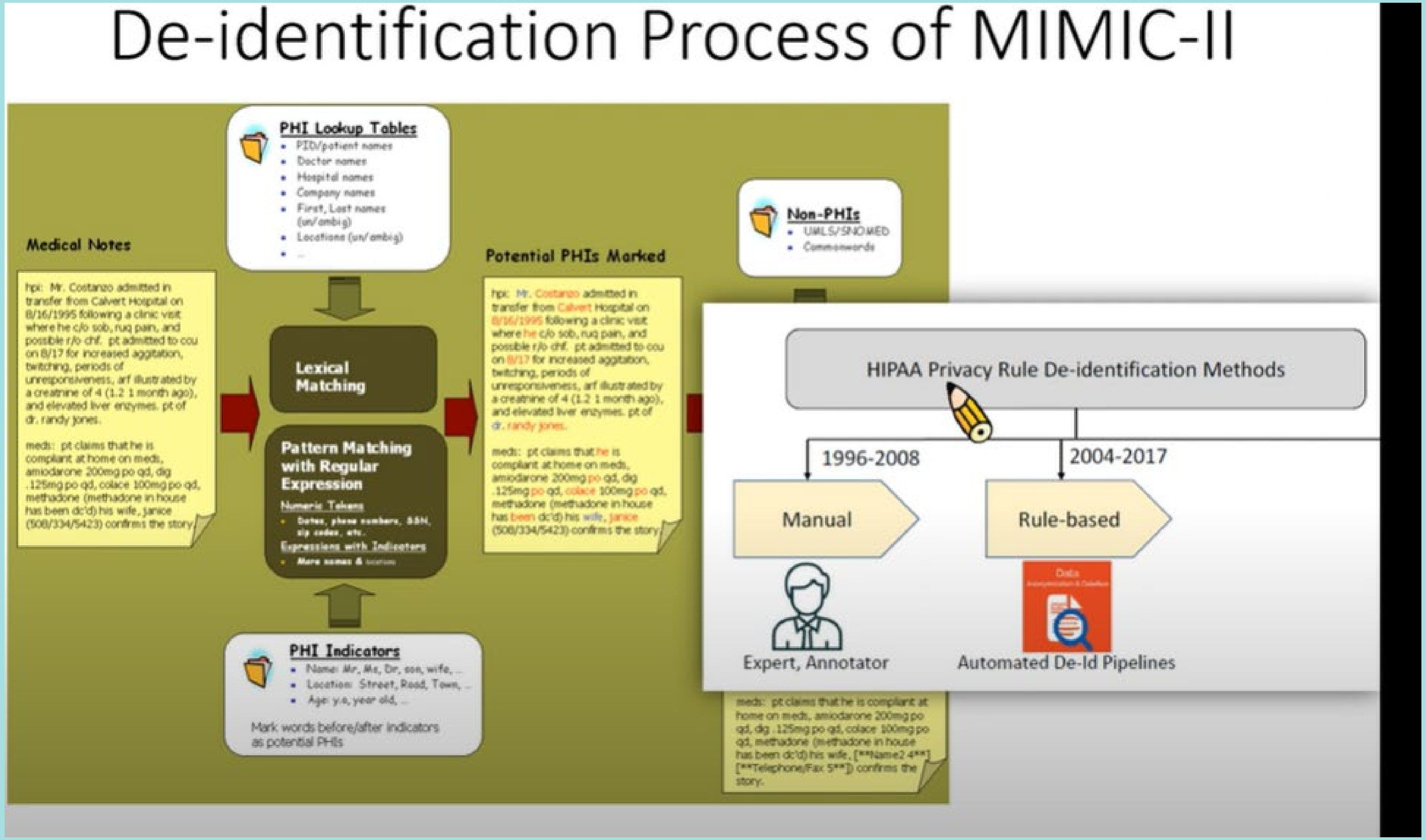
Replacement Algorithm

PHI Surrogate Generation

Note with replaced PHI

FAIRVIEW GENERAL HOSPITAL INPATIENT DISCHARGE SUMMARY Patient Name: Kate Wayne MRN: 439251 DOB: 06/12/2005 Attending Physician at Discharge: Johnson, Andrew C.
Hospital Course: Kate is a 10 week old female who presented to the ER for episodes of paroxysmal cough on November 5th ...
...
Patient lives in Fairfax, CA ...
...
Phone Number: (201)523-6611
rgnv.fgbtuk@joitq.btg
http://www.coughsafe.com

傳統的De-identification作法



De-identification Process of MIMIC-II

Medical Notes

hpi: Mr. Costanzo admitted in transfer from Calvert Hospital on 8/16/1995 following a clinic visit where he c/o sob, ruq pain, and possible r/o chf. pt admitted to ccu on 8/17 for increased aggitation, twitching, periods of unresponsiveness, arf illustrated by a creatnine of 4 (1.2 1 month ago), and elevated liver enzymes. pt of dr. randy jones.

meds: pt claims that he is compliant at home on meds, amiodarone 200mg po qd, dig .125mg po qd, colace 100mg po qd, methadone (methadone in house has been dc'd) his wife, janice (508/334/5423) confirms the story.

PHI Lookup Tables

- PID/patient names
- Doctor names
- Hospital names
- Company names
- First, Last names (un/ambig)
- Locations (un/ambig)
- ...

Lexical Matching

Pattern Matching with Regular Expression

Numeric Tokens

- Dates, phone numbers, SSN, zip codes, etc.

Expressions with Indicators

- More names & locations

PHI Indicators

- Name: Mr, Ms, Dr, son, wife, ...
- Location: Street, Road, Town, ...
- Age: y.o, year old, ...

Mark words before/after indicators as potential PHIs

Potential PHIs Marked

hpi: Mr. Costanzo admitted in transfer from Calvert Hospital on 8/16/1995 following a clinic visit where he c/o sob, ruq pain, and possible r/o chf. pt admitted to ccu on 8/17 for increased aggitation, twitching, periods of unresponsiveness, arf illustrated by a creatnine of 4 (1.2 1 month ago), and elevated liver enzymes. pt of dr. randy jones.

meds: pt claims that he is compliant at home on meds, amiodarone 200mg po qd, dig .125mg po qd, colace 100mg po qd, methadone (methadone in house has been dc'd) his wife, janice (508/334/5423) confirms the story.

Non-PHIs

- UMLS/SNOMED
- Commonwords

meds: pt claims that he is compliant at home on meds, amiodarone 200mg po qd, dig .125mg po qd, colace 100mg po qd, methadone (methadone in house has been dc'd) his wife, [**Name2 4**] [**Telephone/Fax 5**]) confirms the story.

HIPAA Privacy Rule De-identification Methods

1996-2008

Manual

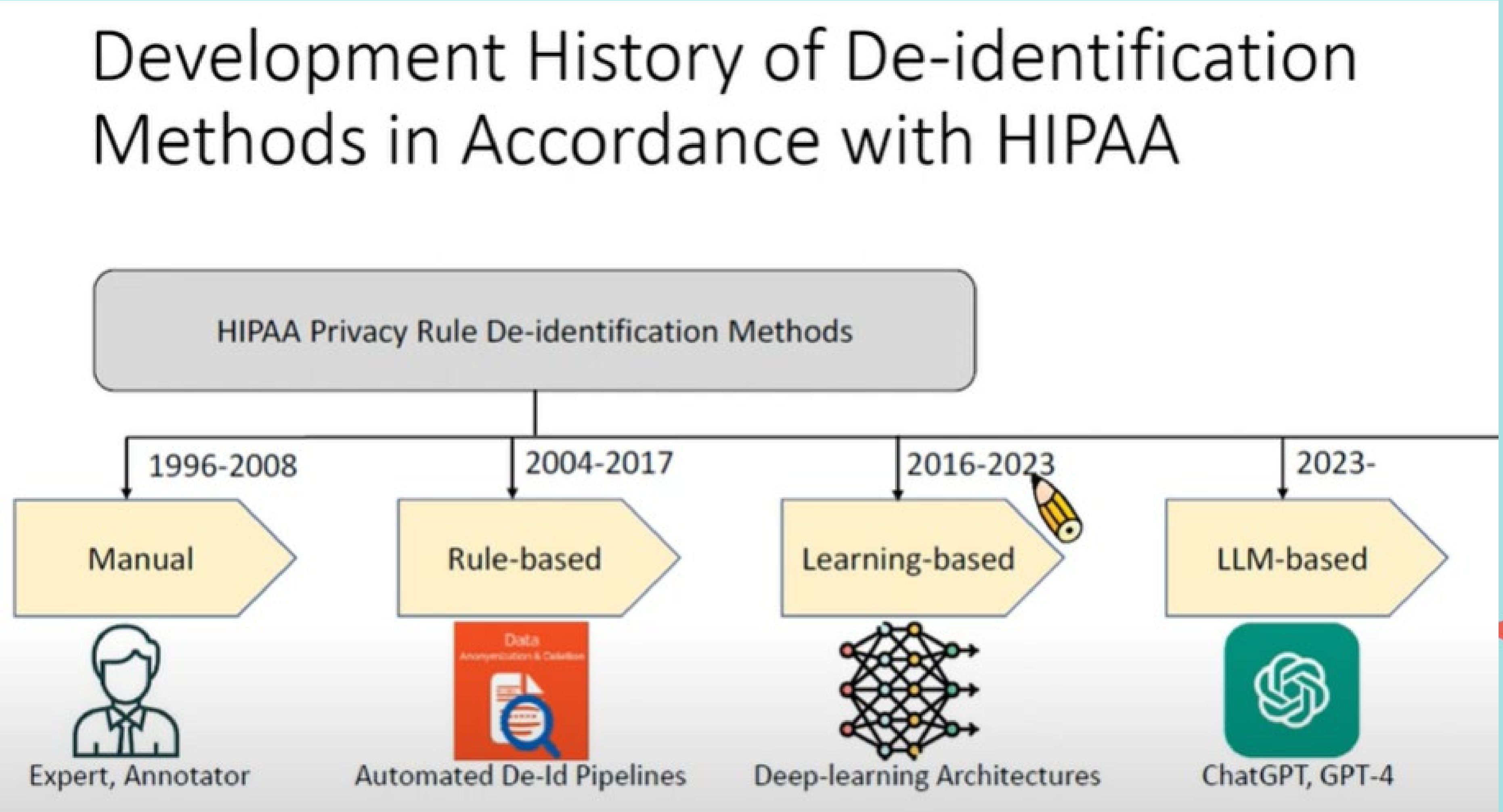
Expert, Annotator

2004-2017

Rule-based

Automated De-Id Pipelines

LLM-based De-identification



Development History of De-identification Methods in Accordance with HIPAA

HIPAA Privacy Rule De-identification Methods

1996-2008

Manual

Expert, Annotator

2004-2017

Rule-based

Automated De-Id Pipelines

2016-2023

Learning-based

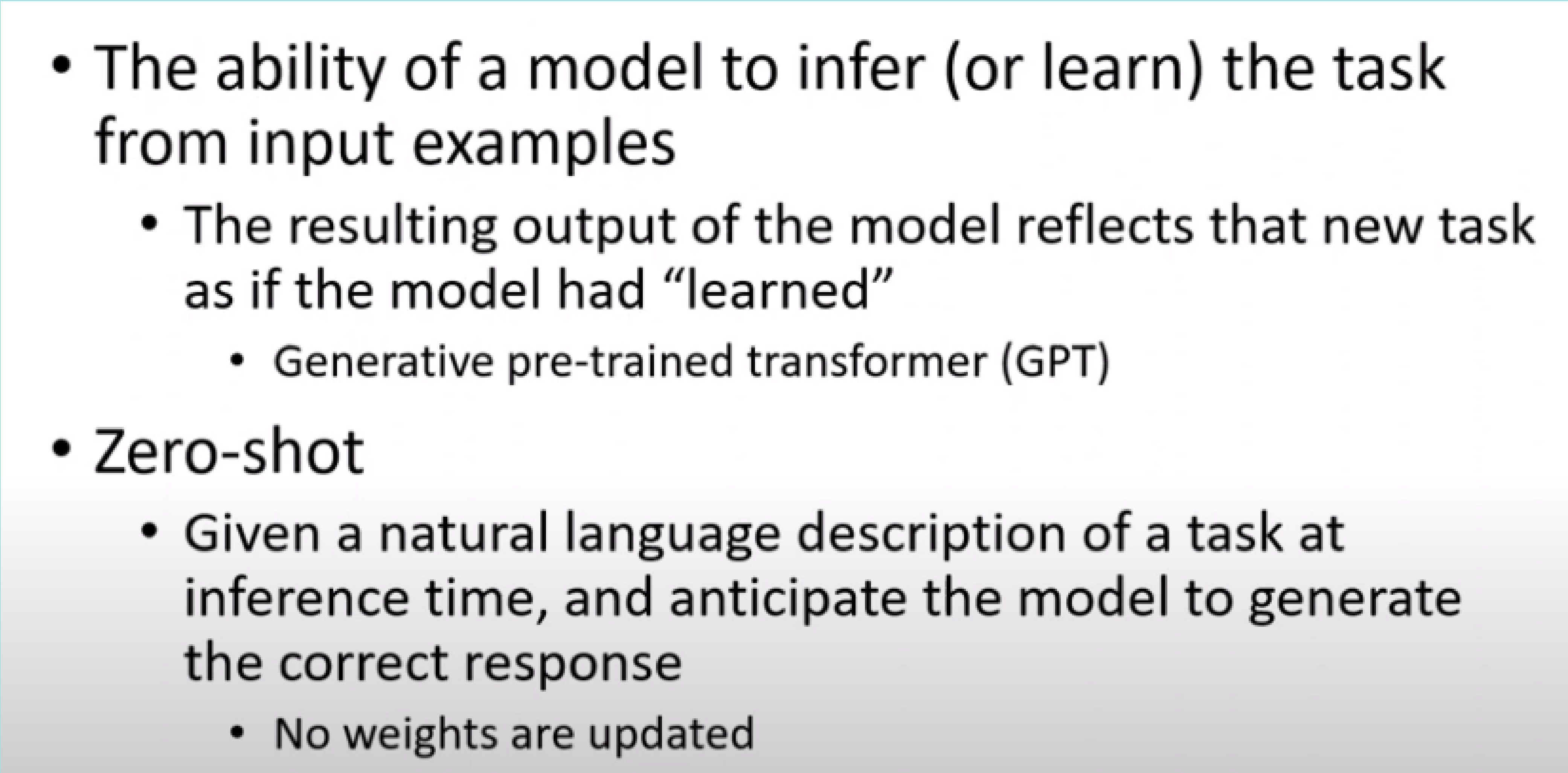
Deep-learning Architectures

2023-

LLM-based

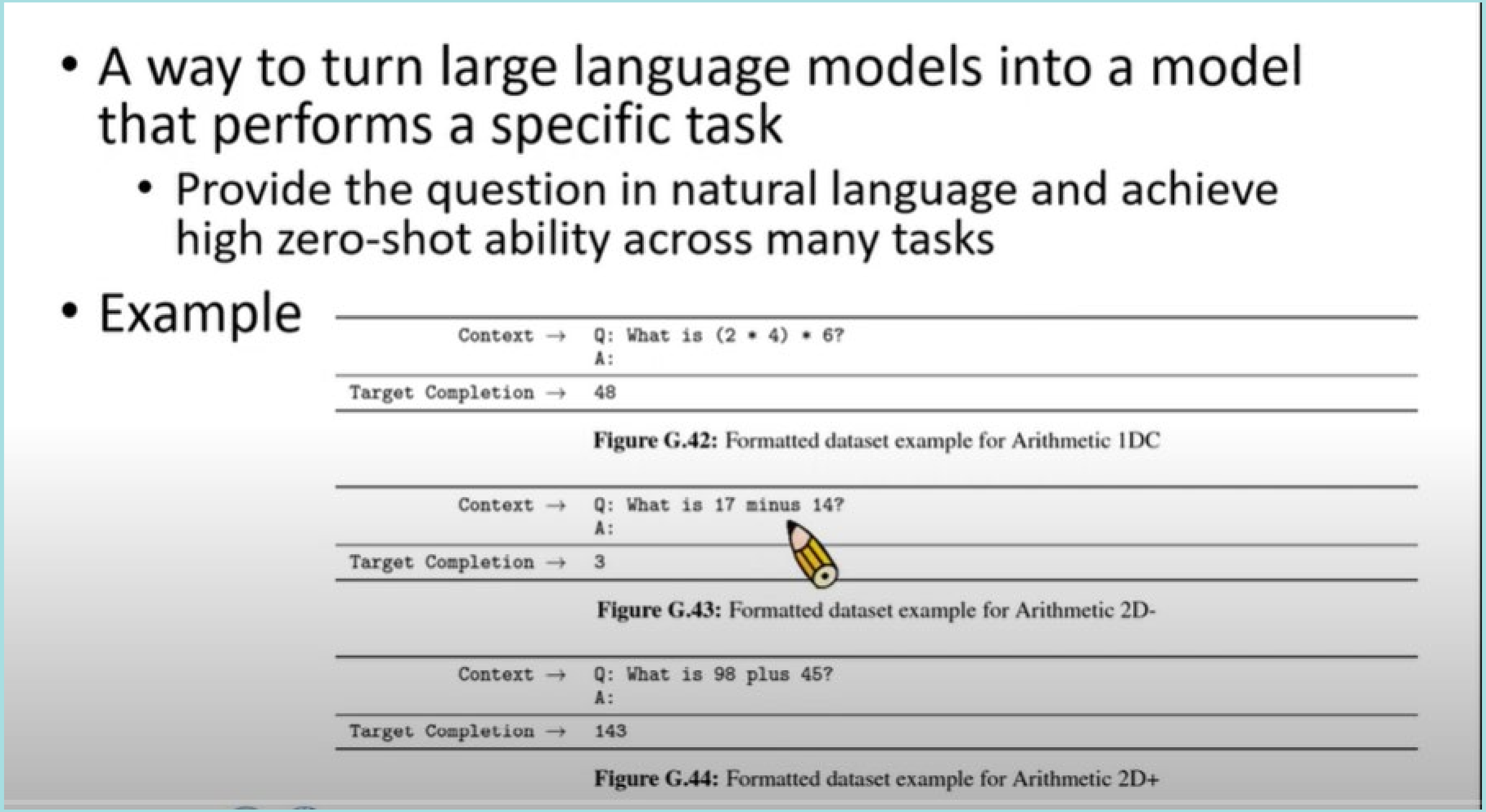
ChatGPT, GPT-4

In Context Learning



- The ability of a model to infer (or learn) the task from input examples
    - The resulting output of the model reflects that new task as if the model had "learned"
        - Generative pre-trained transformer (GPT)
- Zero-shot
    - Given a natural language description of a task at inference time, and anticipate the model to generate the correct response
        - No weights are updated

Prompting



- A way to turn large language models into a model that performs a specific task
  - Provide the question in natural language and achieve high zero-shot ability across many tasks
- Example

Context →	Q: What is (2 * 4) * 6? A:
Target Completion →	48

Figure G.42: Formatted dataset example for Arithmetic 1DC

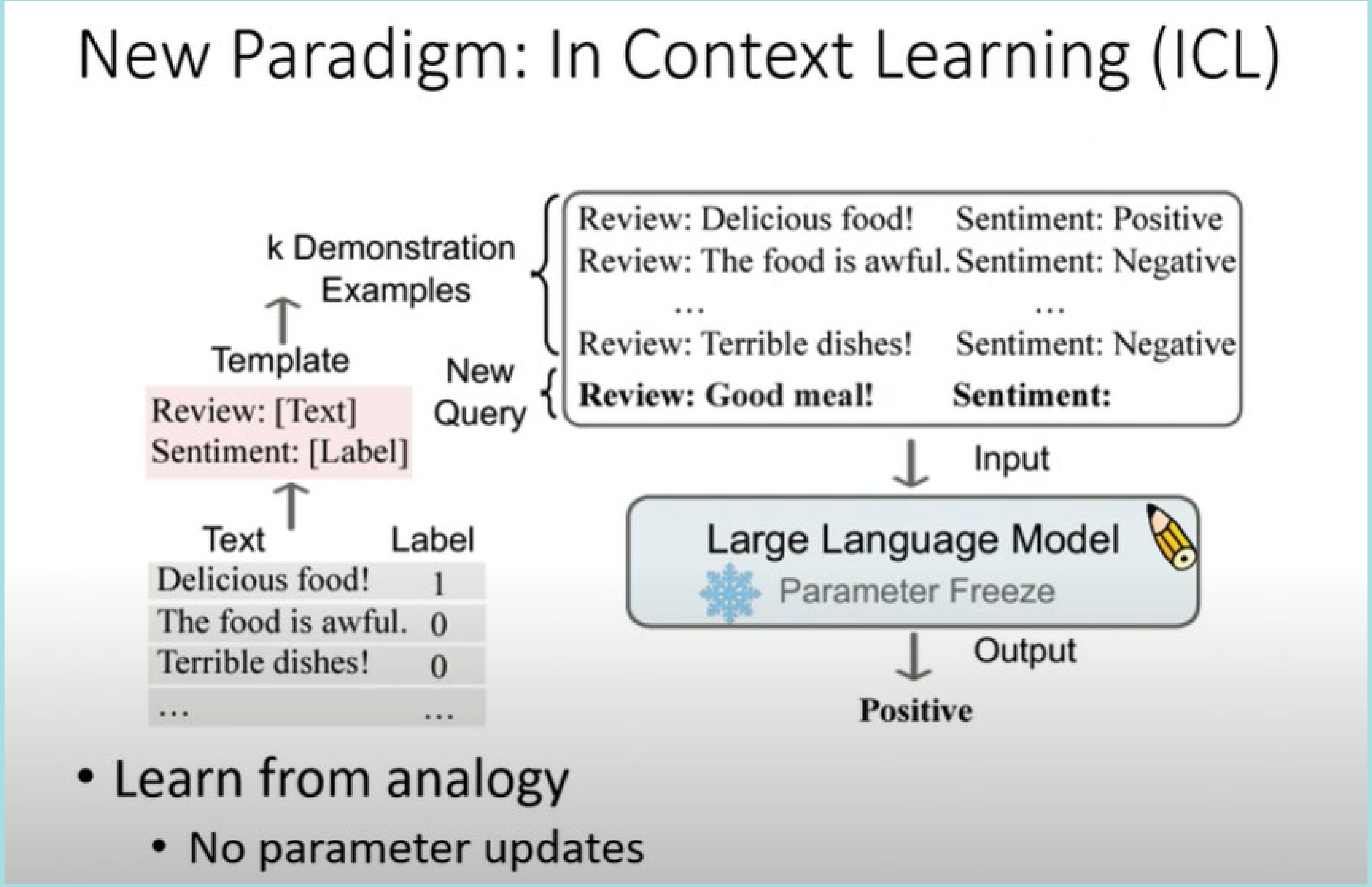
Context →	Q: What is 17 minus 14? A:
Target Completion →	3

Figure G.43: Formatted dataset example for Arithmetic 2D-

Context →	Q: What is 98 plus 45? A:
Target Completion →	143

Figure G.44: Formatted dataset example for Arithmetic 2D+

In Context Learning



New Paradigm: In Context Learning (ICL)

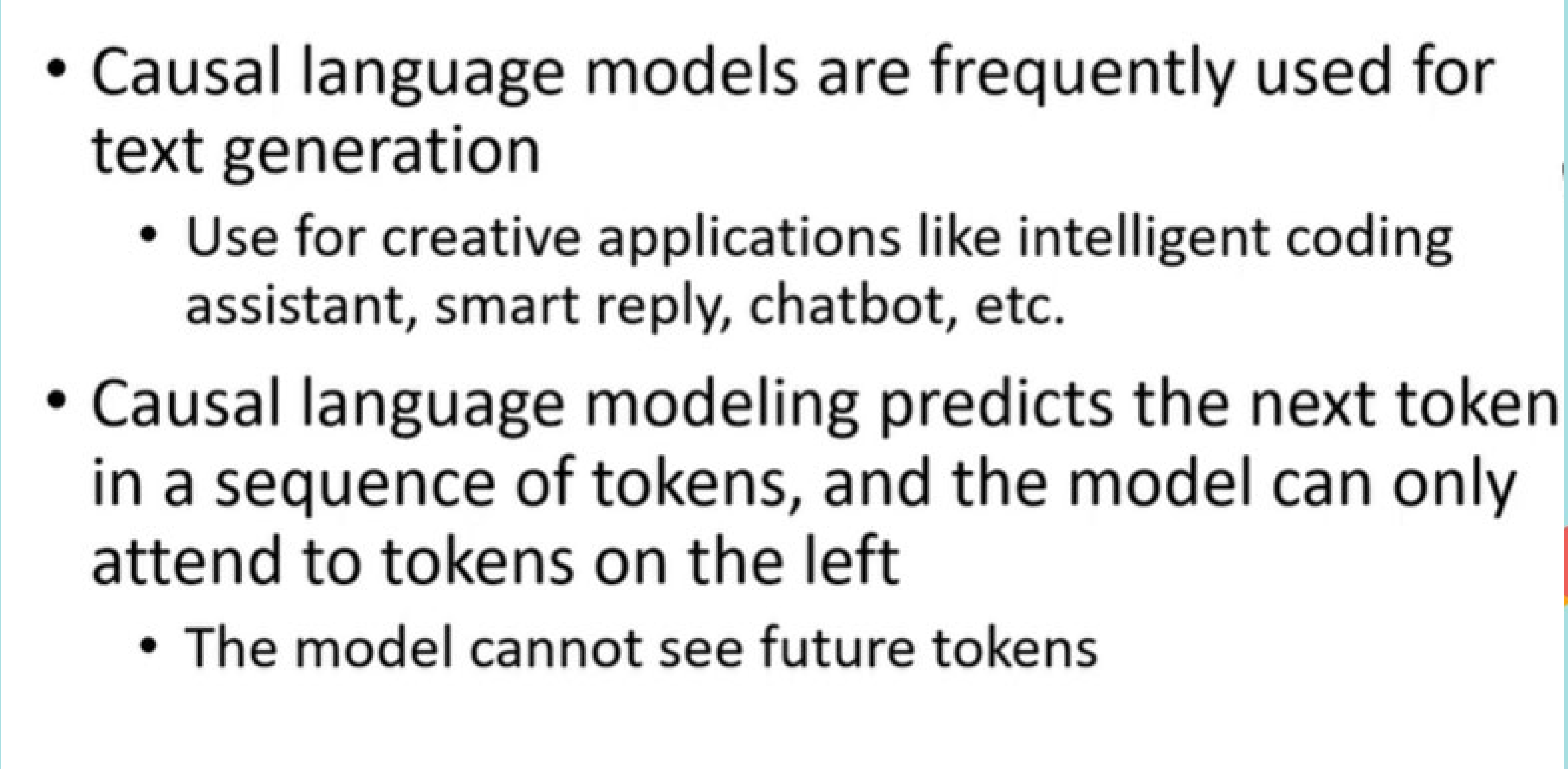
Review: [Text]
Sentiment: [Label]

Text	Label
Delicious food!	1
The food is awful.	0
Terrible dishes!	0
…	…

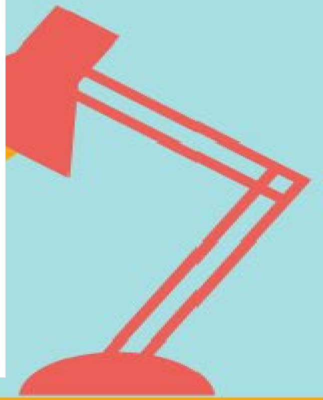
Review: Delicious food! Sentiment: Positive
Review: The food is awful. Sentiment: Negative
… …
Review: Terrible dishes! Sentiment: Negative
Review: Good meal! Sentiment:

- Learn from analogy
  - No parameter updates

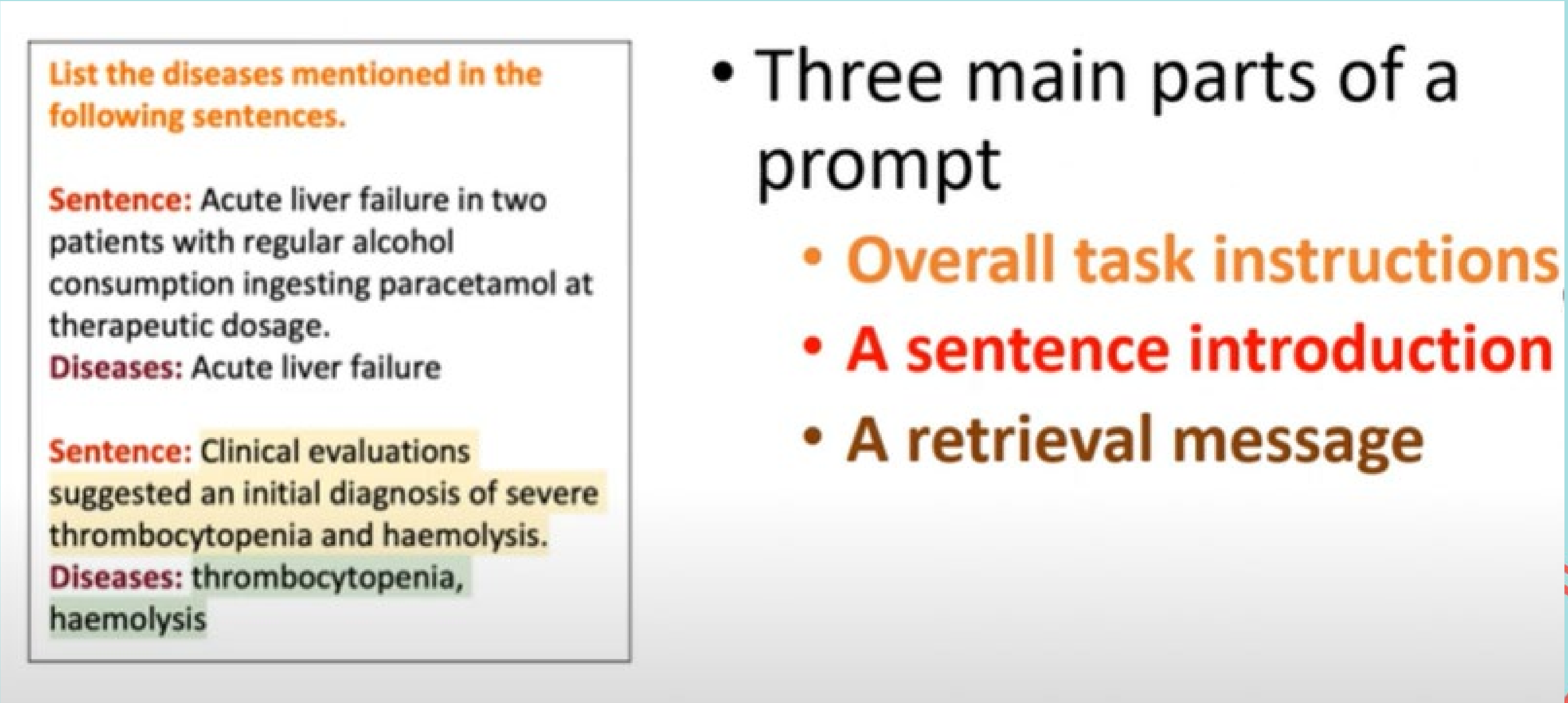
Causal language models



- Causal language models are frequently used for text generation
  - Use for creative applications like intelligent coding assistant, smart reply, chatbot, etc.
- Causal language modeling predicts the next token in a sequence of tokens, and the model can only attend to tokens on the left
  - The model cannot see future tokens



Prompt Design



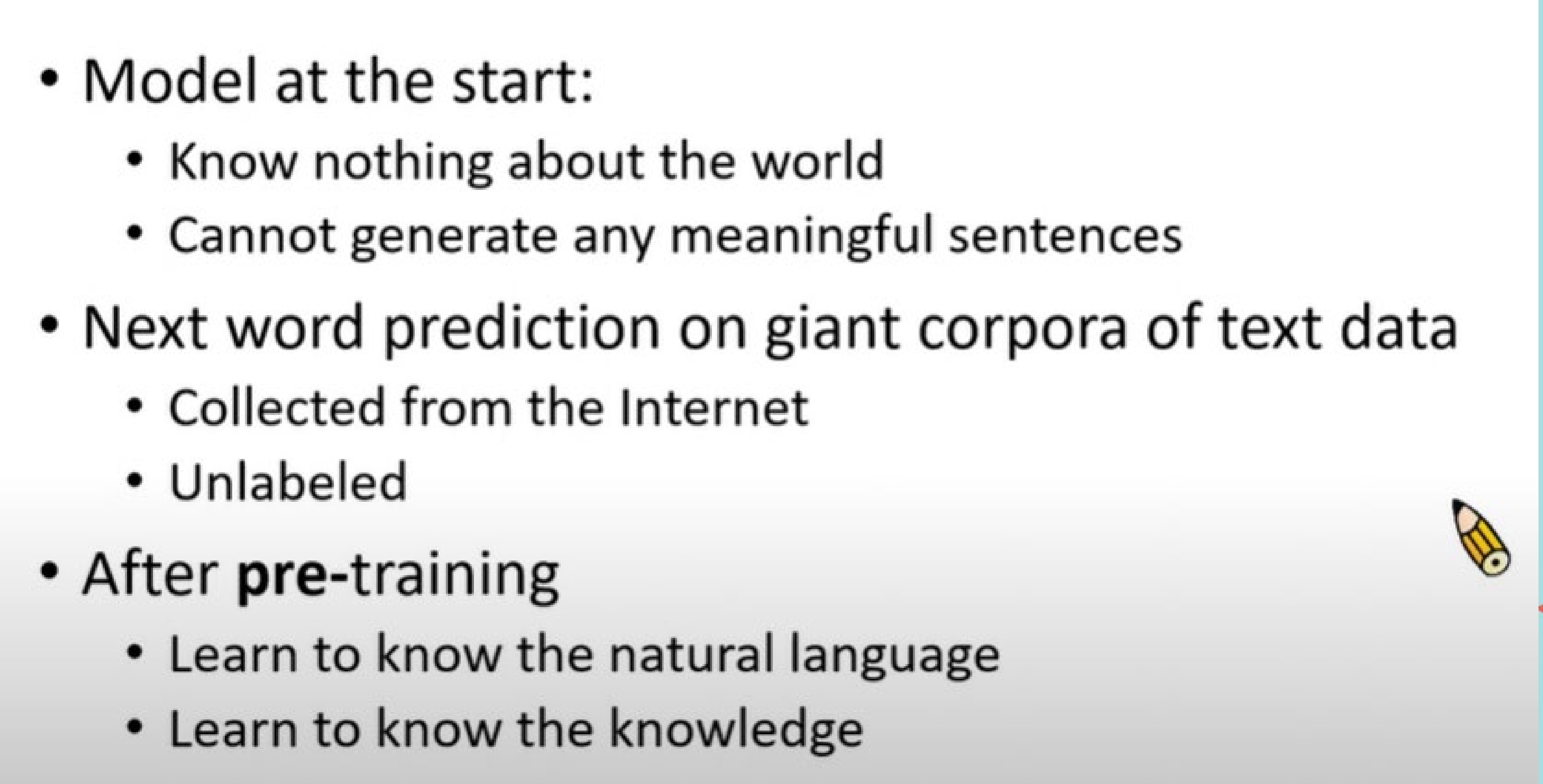
List the diseases mentioned in the following sentences.

Sentence: Acute liver failure in two patients with regular alcohol consumption ingesting paracetamol at therapeutic dosage.
Diseases: Acute liver failure

Sentence: Clinical evaluations suggested an initial diagnosis of severe thrombocytopenia and haemolysis.
Diseases: thrombocytopenia, haemolysis

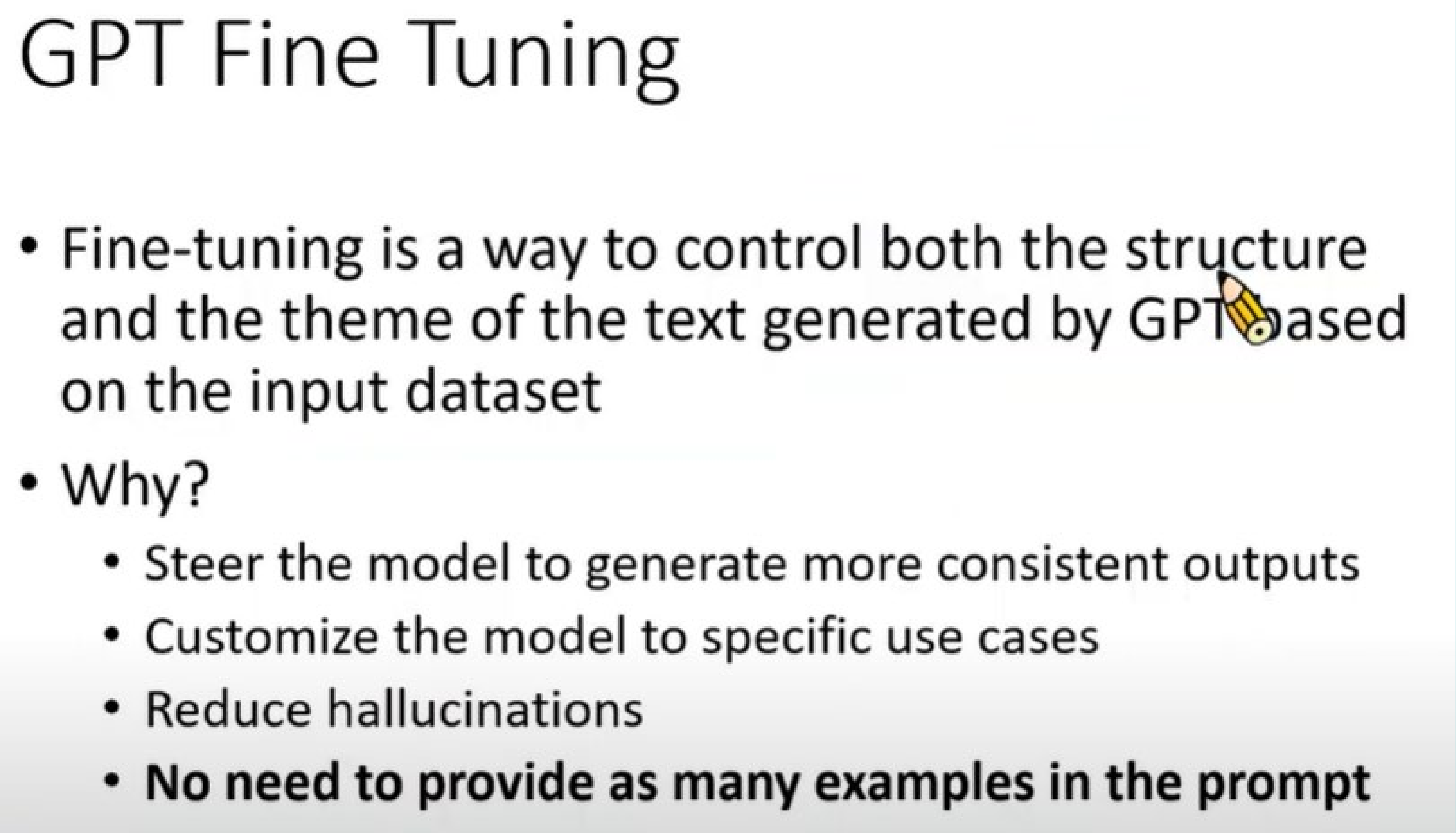
- Three main parts of a prompt
  - Overall task instructions
  - A sentence introduction
  - A retrieval message

Pretraining



- Model at the start:
  - Know nothing about the world
  - Cannot generate any meaningful sentences
- Next word prediction on giant corpora of text data
  - Collected from the Internet
  - Unlabeled
- After **pre-**training
  - Learn to know the natural language
  - Learn to know the knowledge

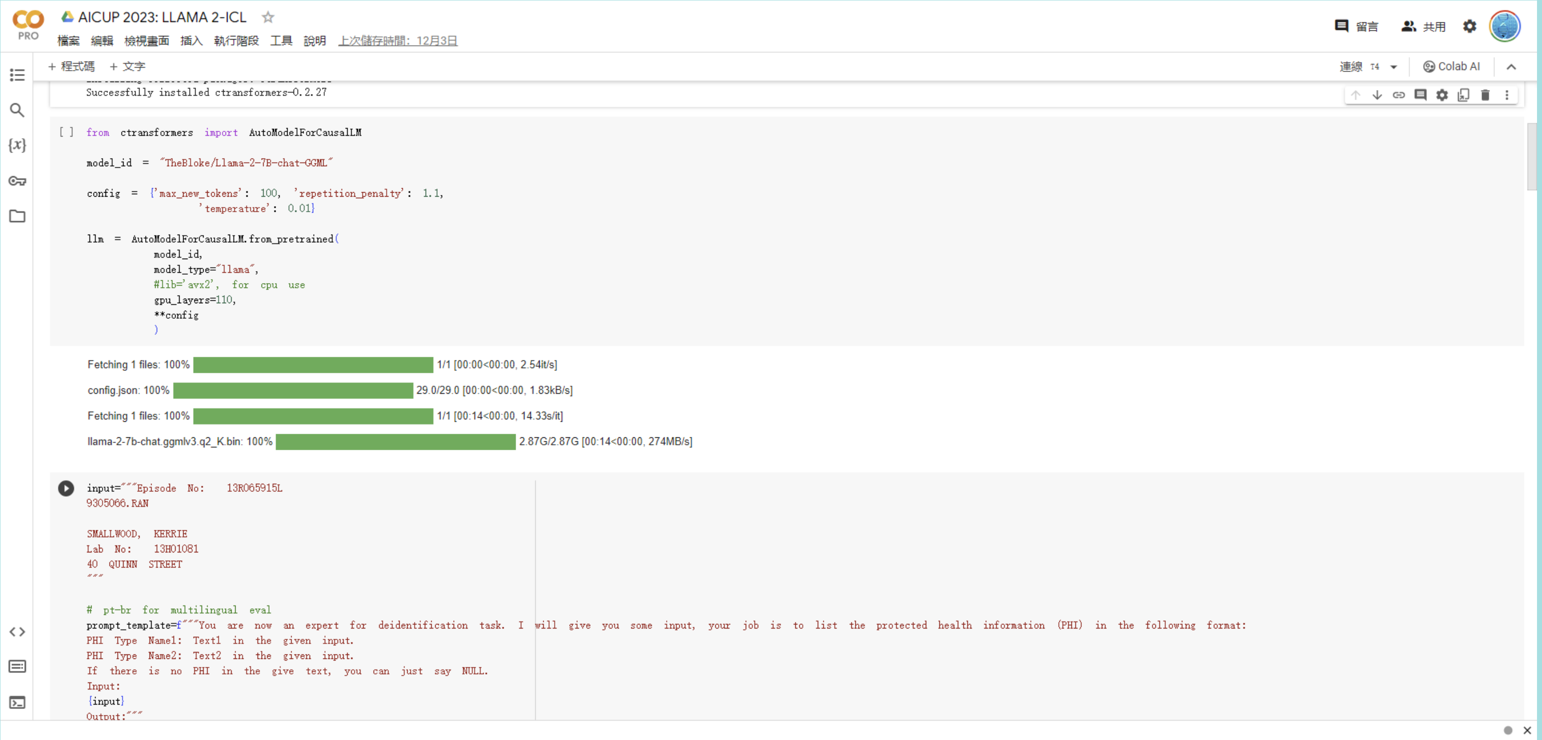
Fine Tuning



GPT Fine Tuning

- Fine-tuning is a way to control both the structure and the theme of the text generated by GPT based on the input dataset
- Why?
  - Steer the model to generate more consistent outputs
  - Customize the model to specific use cases
  - Reduce hallucinations
  - **No need to provide as many examples in the prompt**

LLM-based Data De-identification


AICUP 2023: LLAMA 2-ICL
檔案 編輯 檢視畫面 插入 執行階段 工具 說明 上次儲存時間：12月3日
+ 程式碼 + 文字
留言 共用
連線 T4
Colab AI

```
Successfully installed ctransformers-0.2.27
```

```
from ctransformers import AutoModelForCausalLM

model_id = "TheBloke/Llama-2-7B-chat-GGML"

config = {'max_new_tokens': 100, 'repetition_penalty': 1.1,
          'temperature': 0.01}

llm = AutoModelForCausalLM.from_pretrained(
        model_id,
        model_type="llama",
        #lib='avx2', for cpu use
        gpu_layers=110,
        **config
        )
```

```
Fetching 1 files: 100% 1/1 [00:00<00:00, 2.54it/s]
config.json: 100% 29.0/29.0 [00:00<00:00, 1.83kB/s]
Fetching 1 files: 100% 1/1 [00:14<00:00, 14.33s/it]
llama-2-7b-chat.ggmlv3.q2_K.bin: 100% 2.87G/2.87G [00:14<00:00, 274MB/s]
```

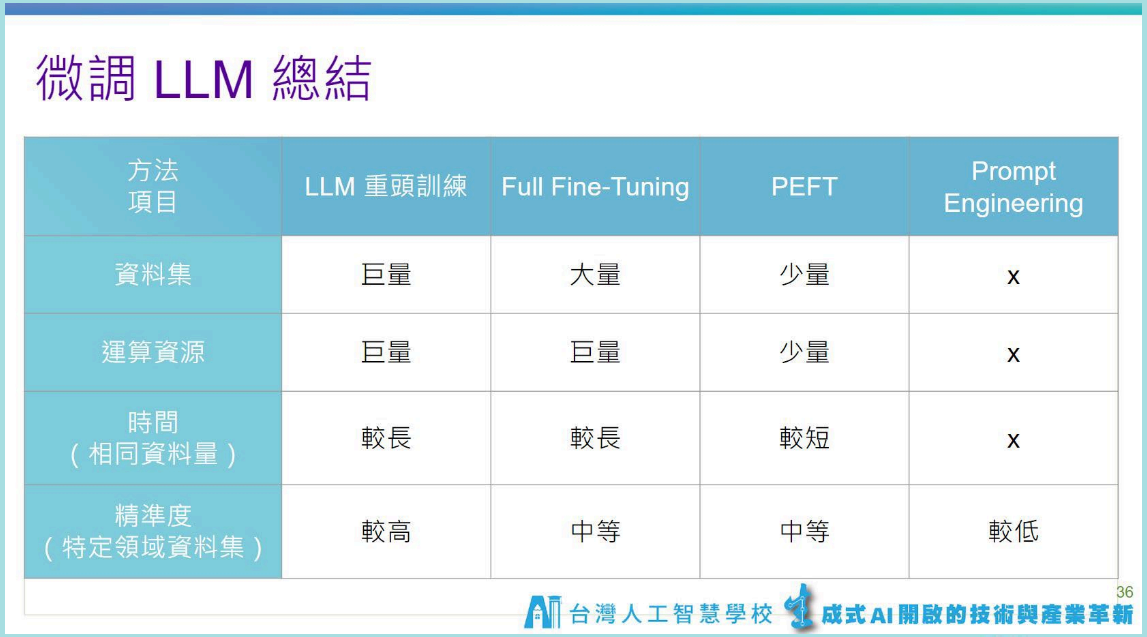
```
input="""Episode No:   13R065915L
9305066.RAN

SMALLWOOD, KERRIE
Lab No:   13H01081
40 QUINN STREET
"""

# pt-br for multilingual eval
prompt_template=f"""You are now an expert for deidentification task. I will give you some input, your job is to list the protected health information (PHI) in the following format:
PHI Type Name1: Text1 in the given input.
PHI Type Name2: Text2 in the given input.
If there is no PHI in the give text, you can just say NULL.
Input:
{input}
Output:"""
```

如何微調你的 LLM？

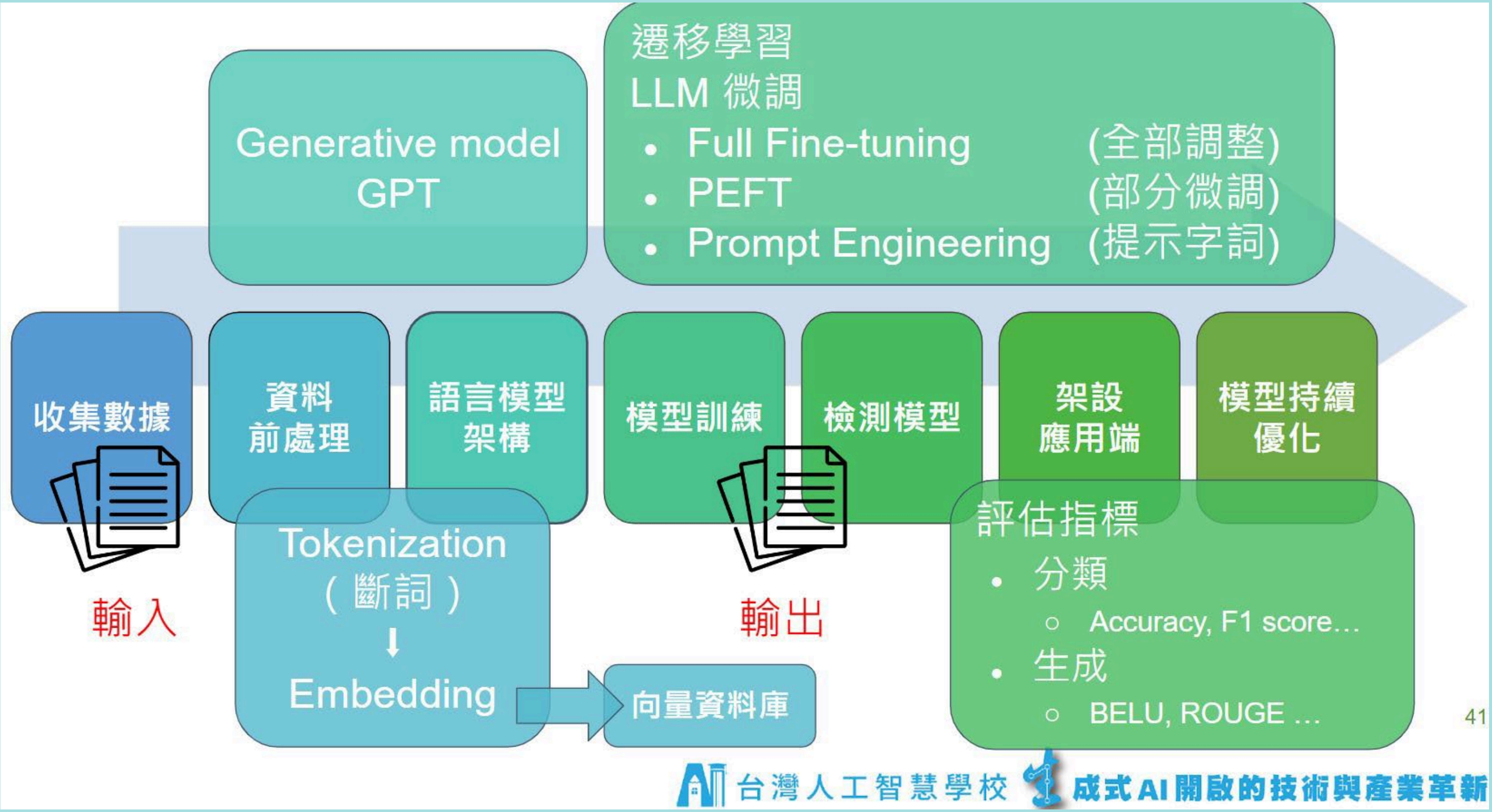
- Prompt Engineering → 提示字詞
- RAG(Retrieval Augmented Generation)→ 檢索手冊
- Parameter-efficient Fine-tuning(PEFT) → 部分微調
- Full Fine-tuning → 整體微調



微調 LLM 總結

方法 項目	LLM 重頭訓練	Full Fine-Tuning	PEFT	Prompt Engineering
資料集	巨量	大量	少量	x
運算資源	巨量	巨量	少量	x
時間 （相同資料量）	較長	較長	較短	x
精準度 （特定領域資料集）	較高	中等	中等	較低

如何微調你的 LLM ? 考慮資料集? 運算資源? 時間? 準確率?



Generative model
GPT

遷移學習
LLM 微調

- Full Fine-tuning (全部調整)
- PEFT (部分微調)
- Prompt Engineering (提示字詞)

收集數據

資料
前處理

語言模型
架構

模型訓練

檢測模型

架設
應用端

模型持續
優化

輸入

Tokenization
(斷詞)
⬇
Embedding

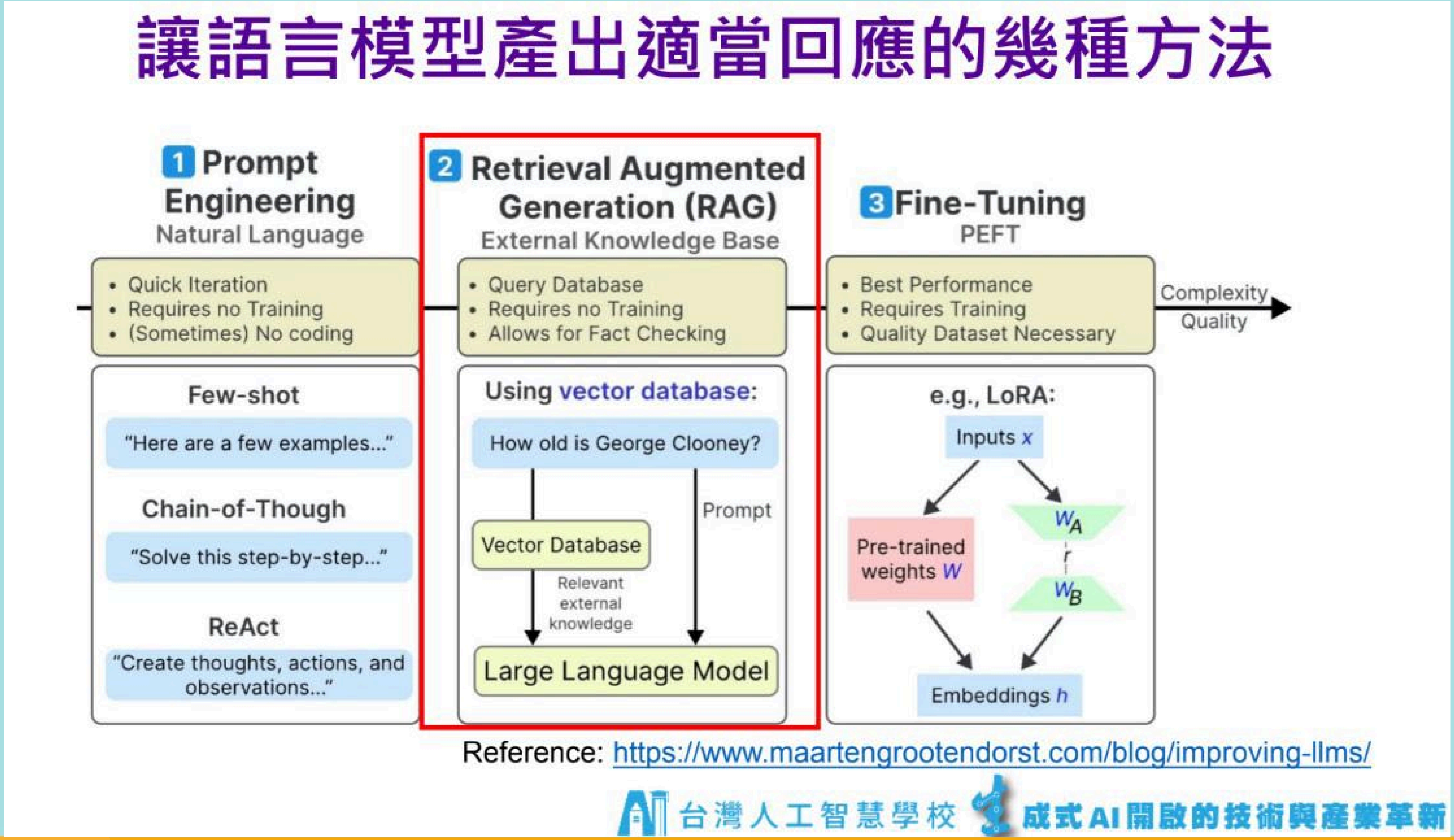
向量資料庫

輸出

評估指標

- 分類
  - Accuracy, F1 score…
- 生成
  - BELU, ROUGE …

RAG



讓語言模型產出適當回應的幾種方法

1 Prompt Engineering
Natural Language

- Quick Iteration
- Requires no Training
- (Sometimes) No coding

Few-shot
"Here are a few examples..."

Chain-of-Though
"Solve this step-by-step..."

ReAct
"Create thoughts, actions, and observations..."

2 Retrieval Augmented Generation (RAG)
External Knowledge Base

- Query Database
- Requires no Training
- Allows for Fact Checking

Using vector database:

3 Fine-Tuning
PEFT

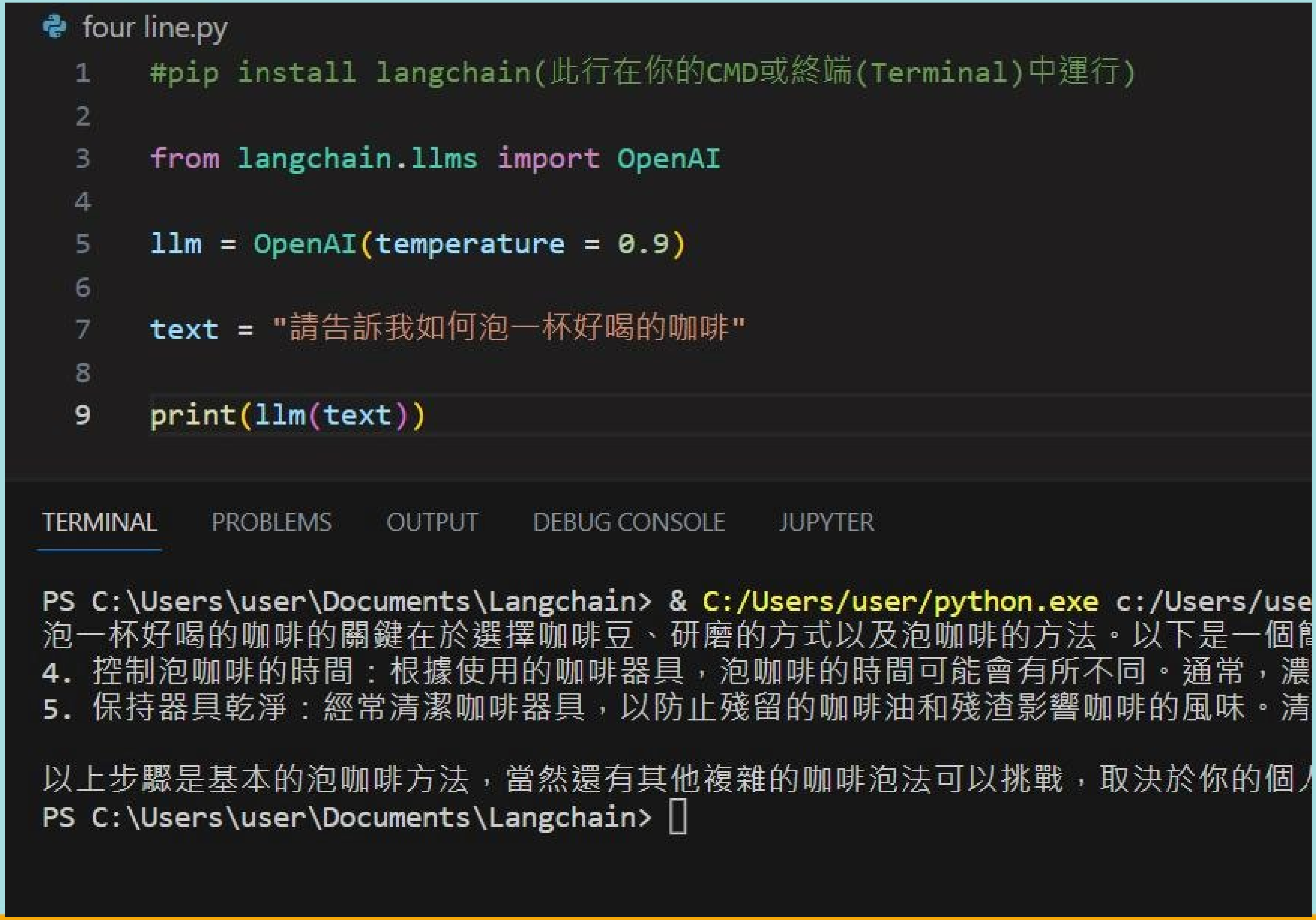
- Best Performance
- Requires Training
- Quality Dataset Necessary

e.g., LoRA:

Reference: https://www.maartengrootendorst.com/blog/improving-llms/

台灣人工智慧學校 成式AI開啟的技術與產業革新

4行學生成式AI



```python
#pip install langchain(此行在你的CMD或終端(Terminal)中運行)

from langchain.llms import OpenAI

llm = OpenAI(temperature = 0.9)

text = "請告訴我如何泡一杯好喝的咖啡"

print(llm(text))
```

```
PS C:\Users\user\Documents\Langchain> & C:/Users/user/python.exe c:/Users/use
泡一杯好喝的咖啡的關鍵在於選擇咖啡豆、研磨的方式以及泡咖啡的方法。以下是一個
4. 控制泡咖啡的時間：根據使用的咖啡器具，泡咖啡的時間可能會有所不同。通常，濃
5. 保持器具乾淨：經常清潔咖啡器具，以防止殘留的咖啡油和殘渣影響咖啡的風味。清

以上步驟是基本的泡咖啡方法，當然還有其他複雜的咖啡泡法可以挑戰，取決於你的個
PS C:\Users\user\Documents\Langchain> 
```

Thanks!

Q&A