

自然語言處理的新篇章： ChatGPT與文字智慧的結合

*A New Chapter in Natural Language Processing:
The Fusion of ChatGPT and Textual Intelligence*

Yung-Chun Chang, Ph.D.

Associate Professor

Deputy Chief Data Officer, Office of Data Science

Graduate Institute of Data Science, Taipei Medical University, Taiwan.

Email: changyc@tmu.edu.tw

Profile



張詠淳 (Yung-Chun Chang)

<http://nlp.tmu.edu.tw/>

- **Work Experience**

- Deputy Chief Data Officer, Taipei Medical University, Office of Data Science, (2020/8–present)
- Associate Professor, Taipei Medical University, Graduate Institute of Data Science, (2020/8–present)
- Assistant Professor, Taipei Medical University, Graduate Institute of Data Science, (2017/8–2020/7)
- Postdoctoral Research Fellow, Academia Sinica, Institute of Information Science, (2016/8–2017/7)
- Research Assistant, Academia Sinica, Institute of Information Science, (2008/1–2016/7)

- **Research Interests**

- Natural Language Processing
- Information Extraction and Retrieval
- Business Intelligence
- Biomedical Text Mining
- AI in EHRs

TMU NLP Lab

<http://nlp.tmu.edu.tw/>

自然語言處理實驗室 NATURAL LANGUAGE PROCESSING



邢晏純 (Yan-Chun Hsing)



TMU Data Science
M.D. Research Assistant

林志莉 (Estella Lim)



TMU Data Science
M.D. Research Assistant

葉文照 (Wen-Chao Yeh)



NTHU ISA
Ph.D. Student

吳肇軒 (Siu Hin NG)



Academia Sinica TIGP
Ph.D. Student

黃顯宜 (Yi-Hsuan Huang)



TMU Data Science
M.S. Student

陳靖儒 (Chin-Ju Chen)



TMU Data Science
M.S. Student

邱滄棠 (Linda Chiu)



TMU Data Science
M.S. Student

陳姿穎 (Tzu-Ying Chen)



TMU Data Science
M.S. Student

林佳慈 (Chia-Tzu Lin)



TMU Data Science
M.S. Student

林志成 (Chih-Cheng Lin)



TMU Data Science
M.S. Student

Di-Jie Liew



TMU Data Science
M.S. Student

Jean Aristide Belleza Aquino



TMU Data Science
M.S. Student

黃士軒 (Shih-Hsuan Huang)



TMU Data Science
M.S. Student

林苒優 (Heng-You Lin)



TMU Data Science
M.S. Student

黃永鎮 (Yong-Zhen Huang)



TMU Data Science
M.S. Student

施辰穎 (Eugene Sy)



TMU Data Science
M.S. Student

Systems We Developed



Monpa



TMUNLP Word2Vec



Valence & Arousal



DEWS Dashboard



Word Cloud



Cluster







Research Demo



Information Extraction for Clinical Pathology Report and Certificate of Diagnosis



MODM Multi-Objective Dialog Mechanism



MONPA Multi-Objective NER POS Annotator



MONPA Multi-Objective NER POS Annotator

彭子承 (Tzu-Cheng Peng)



TMU Data Science
M.S. Student

李子儀 (Tzu-Yi Li)



TMU HCA
Undergraduate Student

張庭維 (Ting-Wei Chang)



TMU Medicine
Graduate

陳俊宏 (Chun-Hung Chen)



NTU PHY
M.D. Research Assistant

Our research is extensive



Clinical NLP



Biomedical NLP



Social NLP



Business Intelligence

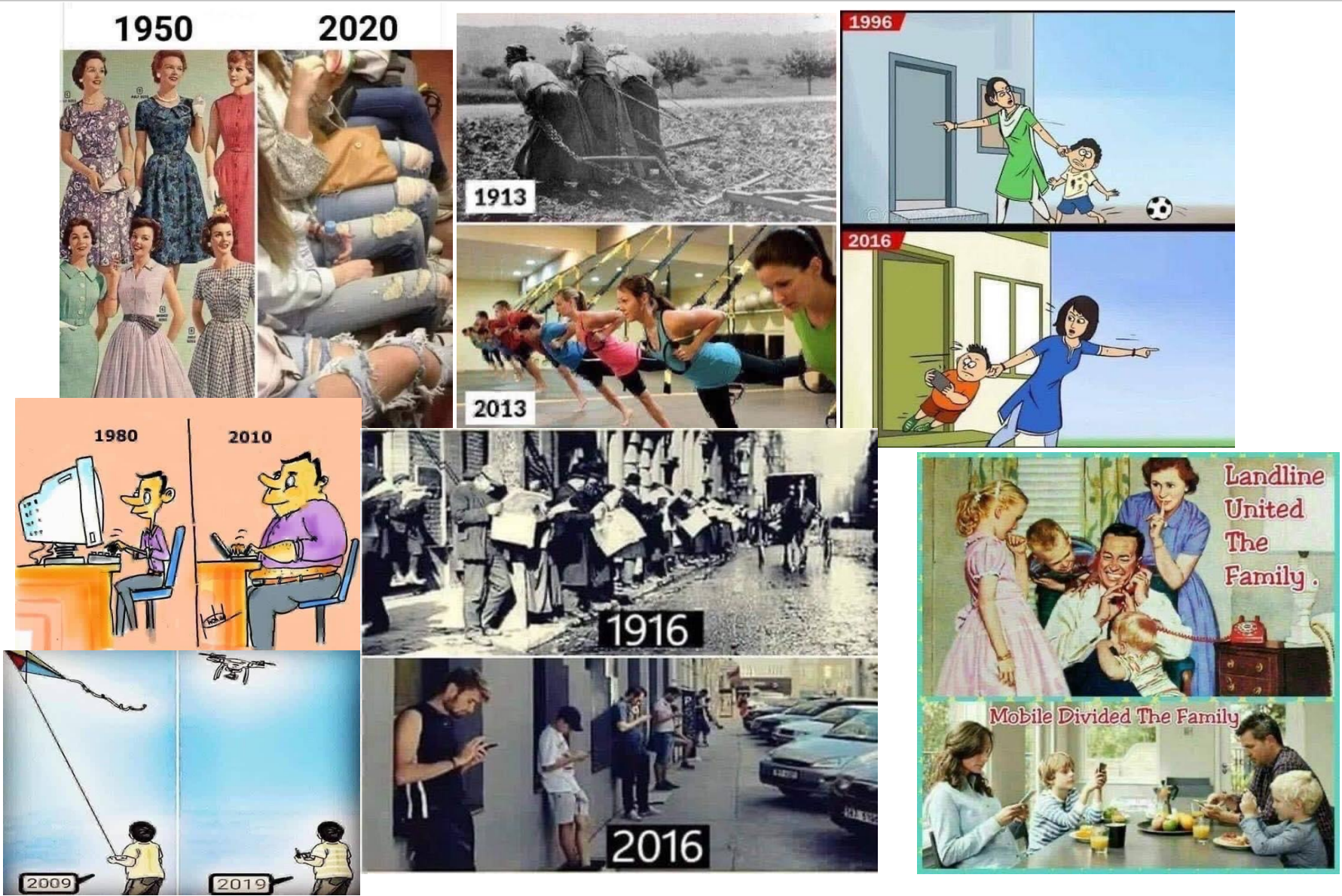


Financial NLP



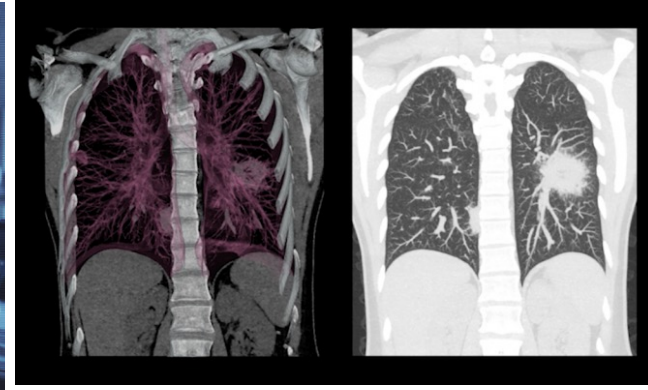
Sports Technology

The world has changed

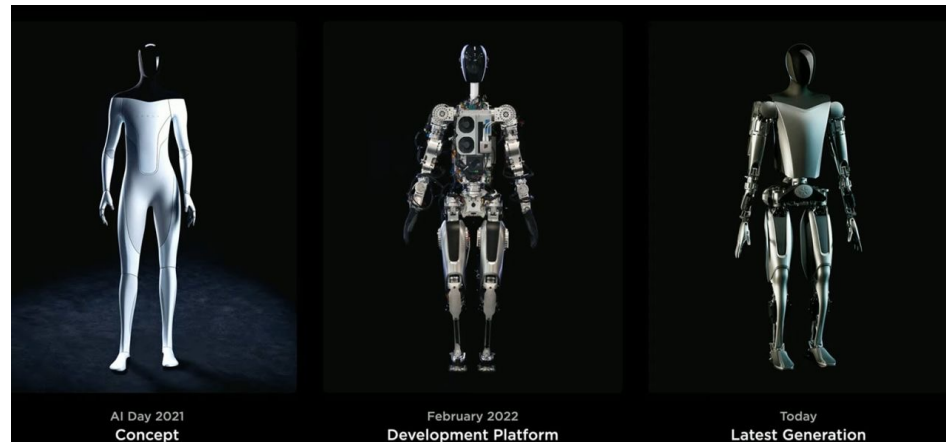


AI is everywhere and nowhere

AlphaGo Zero
Starting from scratch

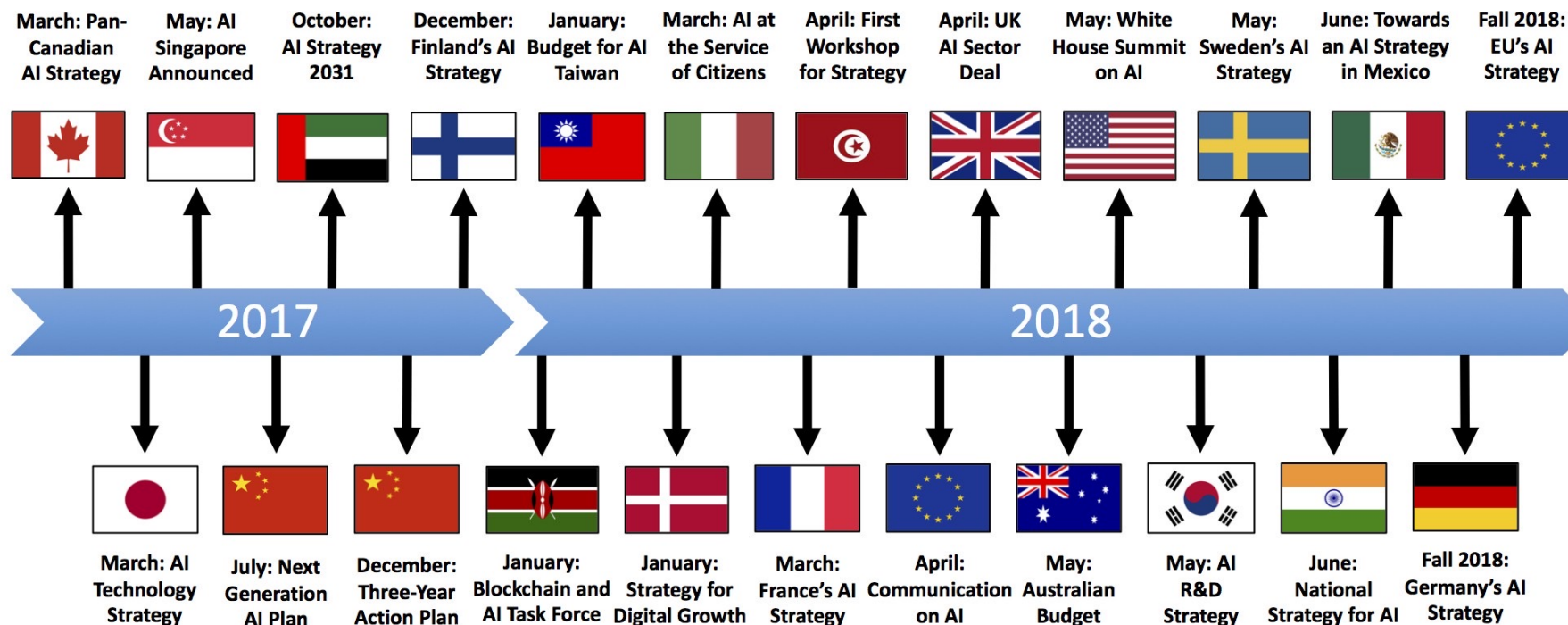


Midjourney



An Overview of National AI Strategies

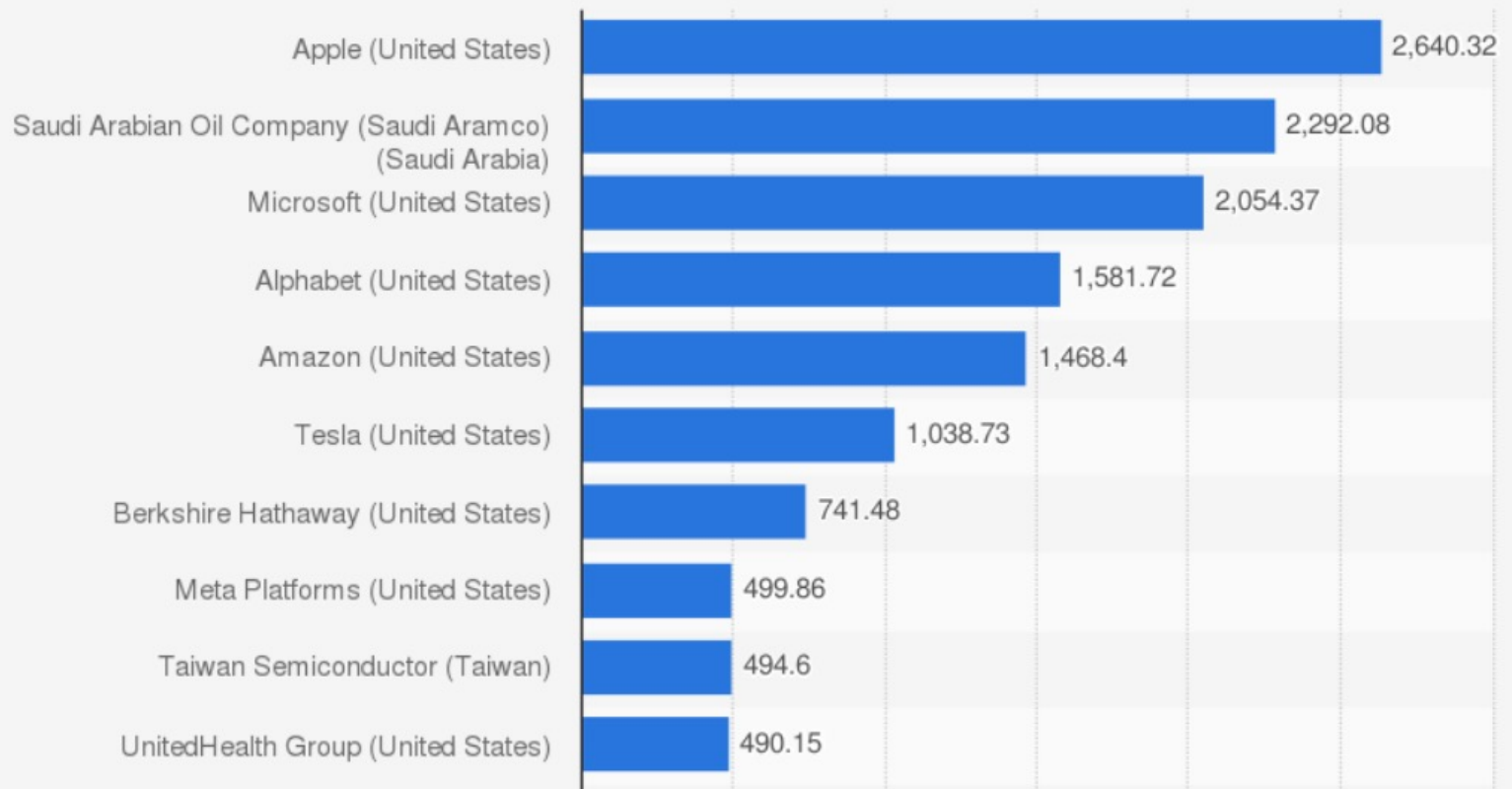
Artificial Intelligence Strategies



2018-07-13 | Politics + AI | Tim Dutton

The Biggest Companies in the World in 2021

The 100 largest companies in the world by market capitalization in 2022 (in billion U.S. dollars)

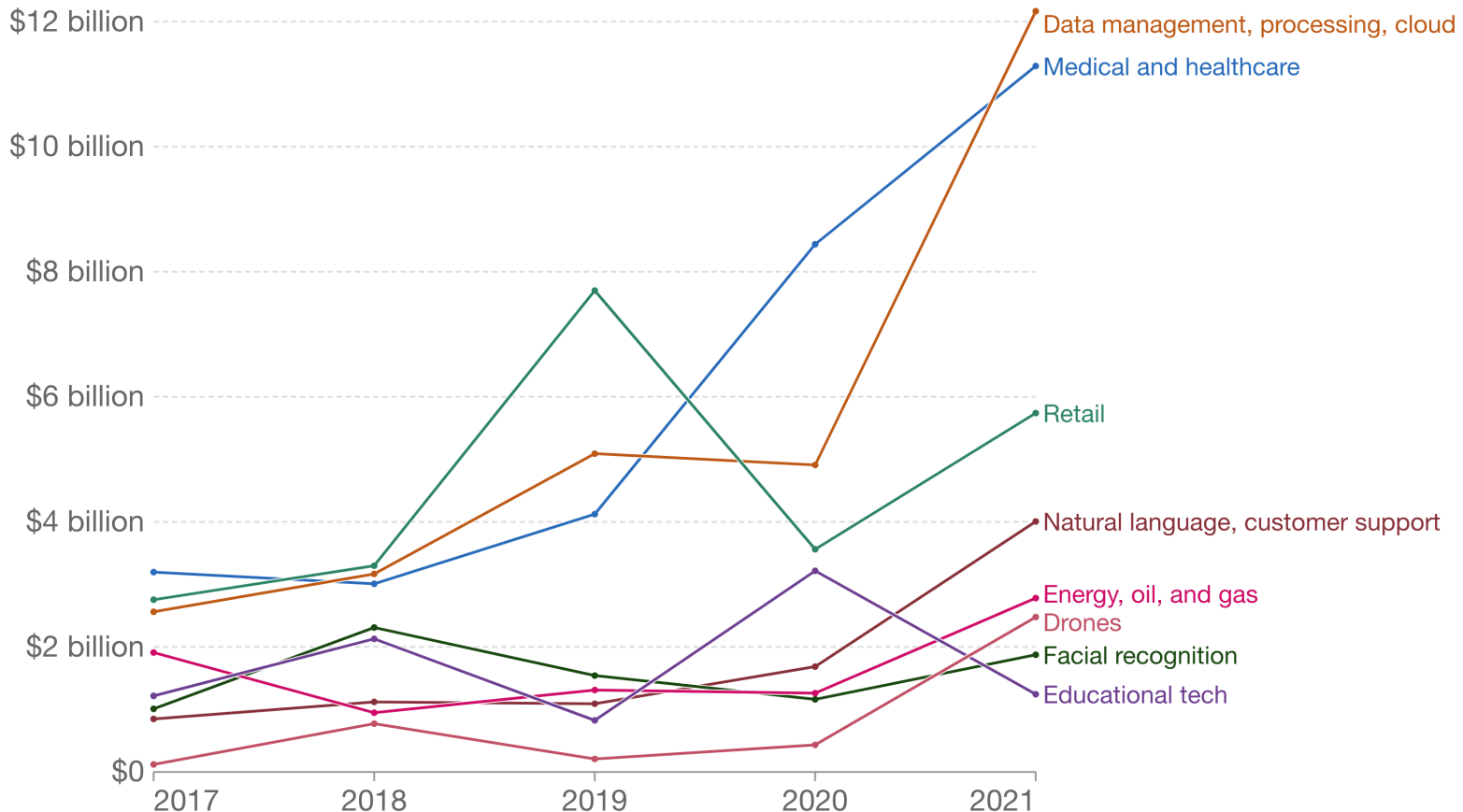


Global Private Investment in AI by Focus Area

Annual global private investment in artificial intelligence, by focus area

Our World
in Data

Includes companies that received more than \$1.5 million in investment. This data is expressed in US dollars, adjusted for inflation.



Source: NetBase Quid via AI Index Report (2022)

OurWorldInData.org/artificial-intelligence • CC BY

Note: Data is expressed in constant 2021 US\$. Inflation adjustment is based on the US Consumer Price Index (CPI).

資料來源：2022 AI Index Report

Data Is The New Oil



“Data is the new oil. It’s valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.” — Clive Humby.

未來30年數據將取代石油，成最強大能源 – 馬雲



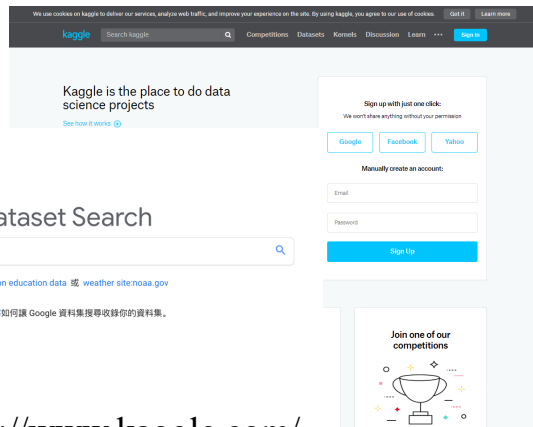
“Consumer data will be the biggest differentiator in the next two to three years. Whoever unlocks the reams of data and uses it strategically will win.” – Angela Ahrendts

“DATA is like the new oil” – Barack Obama



Open Data and Competition

Kaggle & Google Dataset Search



<https://www.kaggle.com/>

<https://datasetsearch.research.google.com/>

NHIRD



<https://nhird.nhri.org.tw/>



MIMIC-III and -IV

scientific **data**

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [scientific data](#) > [data descriptors](#) > [article](#)

Open Access | Published: 24 May 2016

MIMIC-III, a freely accessible critical care database

[Alistair E.W. Johnson](#), [Tom J. Pollard](#), [Lu Shen](#), [Li-wei H. Lehman](#), [Mengling Feng](#), [Mohammad Ghassemi](#), [Benjamin Moody](#), [Peter Szolovits](#), [Leo Anthony Celi](#) & [Roger G. Mark](#)

[Scientific Data](#) 3, Article number: 160035 (2016) | [Cite this article](#)

129k Accesses | 1322 Citations | 79 Altmetric | [Metrics](#)

Abstract

MIMIC-III ('Medical Information Mart for Intensive Care') is a large, single-center database comprising information relating to patients admitted to critical care units at a large tertiary care hospital. Data includes vital signs, medications, laboratory measurements, observations and notes charted by care providers, fluid balance, procedure codes, diagnostic codes, imaging reports, hospital length of stay, survival data, and more. The database supports applications including academic and industrial research, quality improvement initiatives, and higher education coursework.

Design Type(s)	data integration objective
Measurement Type(s)	Demographics • clinical measurement • intervention • Billing • Medical History Dictionary • Pharmacotherapy • clinical laboratory test • medical data
Technology Type(s)	Electronic Medical Record • Medical Record • Electronic Billing System • Medical Coding Process Document • Free Text Format
Factor Type(s)	
Sample Characteristic(s)	Homo sapiens

[Machine-accessible metadata file describing the reported data](#) (ISA-Tab format)

Academic Competition



Ex: NTCIR, BioCreative, Tianchi...

糖尿病眼部影像AI決策支援系統



台灣糖尿病患者多達200 萬人以上，病患易併發眼底多項病變。工研院持續研發眼底影像分析技術，整合數十萬張影像資料、統計分析及決策支援系統，發展出高鑑別度、高篩出率眼底影像AI模型，以Kaggle 資料集（88,726 張影像）進行訓練測試，AI 模型Kappa 一致性達84.15%，排名世界第三。同時移植至嵌入式系統，成為不到10 公分長的智慧匣道器。此系統可整合電腦驗光、眼底鏡等檢測裝置數據與身分識別資料，以行動化模式提供民眾視力健檢服務，成功導入國內4 家診所驗證，讓醫生可運用雲端化遠距圖片，將糖尿病患眼底影像篩檢普及率平均值，從2017年的18.6% 提高至2018年的41.5%。視光師或護理師也可運用行動化檢測設備，至居家、偏鄉幫民眾作眼睛檢測，有效提升糖尿病眼底病變預防篩檢。

Why is NLP so difficult?

Word Segmentation:

衛福部陳時中部長呼籲民眾於龍山寺參拜須保持社交距離



Co-reference resolution:

The epidemic is too serious, especially those countries that are less developed, so that it cannot be resolved quickly among them.

皮膚科

柯醫師

皮膚科醫師

治療這種過敏皮膚柯醫師最有經驗

決定語意段落很重要！！

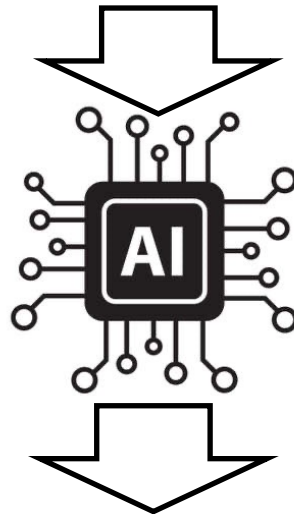
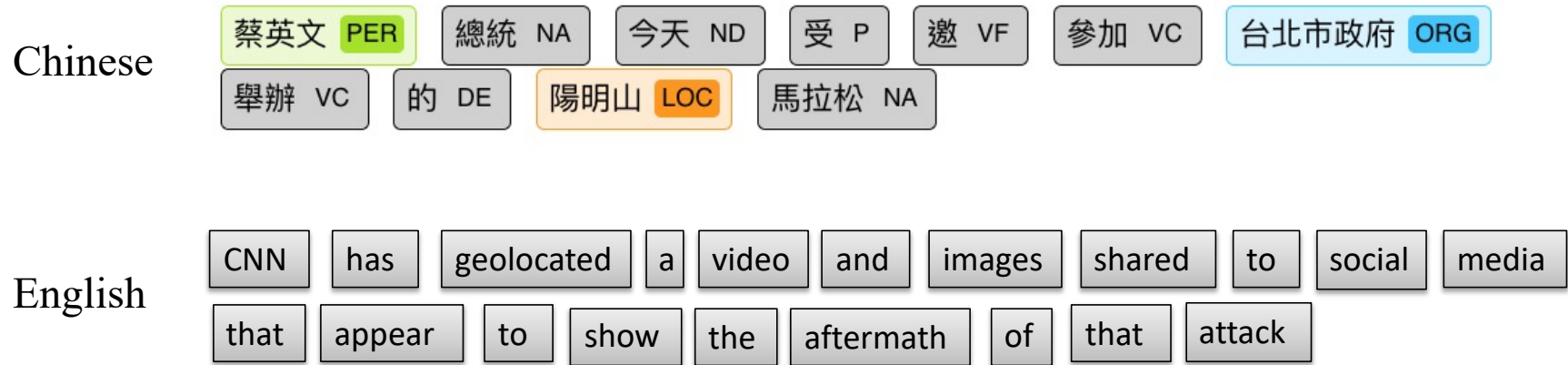
Intent detection & Sentiment Analysis:

我身體不舒服，請幫我掛北醫胸腔內科的蕭世欣醫師



The difference between NLP and other tasks

Input: A sequence of text (tokens): document, sentence, phrase, or word.



Output: sequence, label, or scalar

The difference between NLP and other tasks

Input: A sequence of text (tokens): document, sentence, phrase, or word.

CNN has geolocated a video and images shared to social media that appear to show the aftermath of that attack.

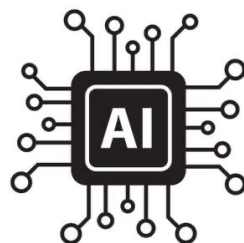
CNN has geolocated a video and images shared to social media that appear to show the aftermath of that attack

蔡英文總統今天受邀參加台北市政府舉辦的陽明山馬拉松

蔡英文 PER 總統 NA 今天 ND 受 P 邀 VF 參加 VC 台北市政府 ORG
舉辦 VC 的 DE 陽明山 LOC 馬拉松 NA



Text Representation



Output: sequence, label, or scalar

Text Representation and Prediction

Language

Human Readable:

Vector

Machine Readable:



Bag-of-Word

2	0	0	2	...	10	3	5	0
Laker	hospital	key	legend		Kobe	NBA	sport	log

Word Embeddings

0.2	0.3	0.5	0.5	...	0.3	0.5	0.4	0.4
0.3	0.2	0.5	0.7	...	0.3	0.5	0.4	0.2
0.3	0.3	0.6	0.2	...	0.3	0.5	0.9	0.9
⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮
0.2	0.5	0.4	0.3	...	0.6	0.5	0.7	0.4
0.5	0.3	0.5	0.5	...	0.8	0.2	0.3	0.3

The brief history of Large Language Models

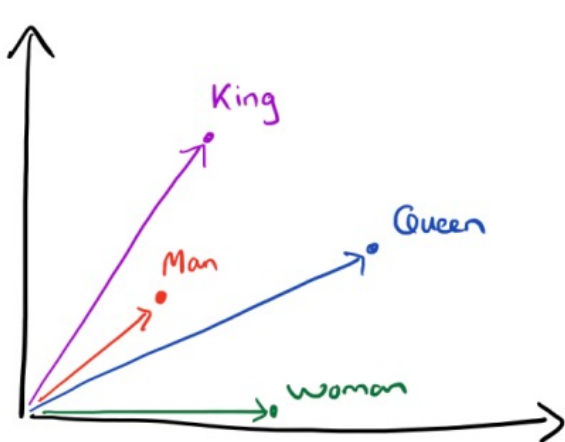
1966	1966	Late 1980s - 1990s	2000s
ELIZA	SHRDLU	Statistical Language Models	Neural Probabilistic Language Model
2019	2018	2017	2013
GPT-2 and T5	BERT	Transformer Models and Attention Mechanisms	Word2Vec
2020	Jan 2021 - Oct 2022		
GPT-3	LaMDA, xlarge, Chinchilla, CodeGen, InCoder, mGPT, PaLM, OPT-IML, Minerva		
Feb 2023	Jan 2023	Dec 2022	Nov 2022
Google Bard and LLaMa	WebGPT	GPT 3.5	ChatGPT
Mar 2023	Apr 2023	May 2023	
GPT-4	BloombergGPT, StableLM, Dolly 2.0, Titan, BingChat	PaLM2	

「Kobe Bryant wins All-Star Game MVP」

「Michael Jordan wins All-Star Game MVP」

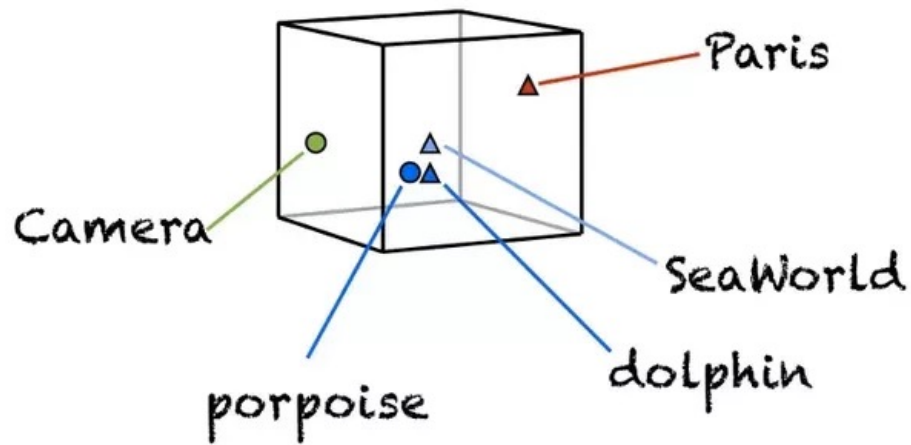
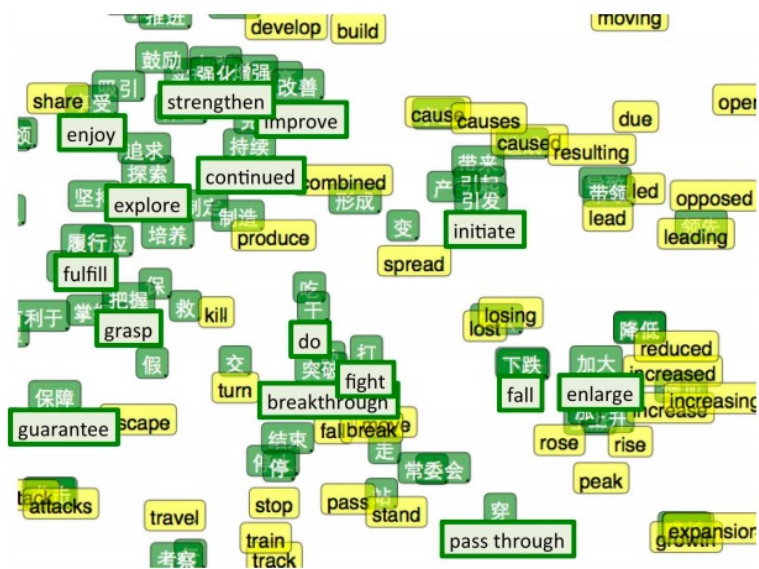
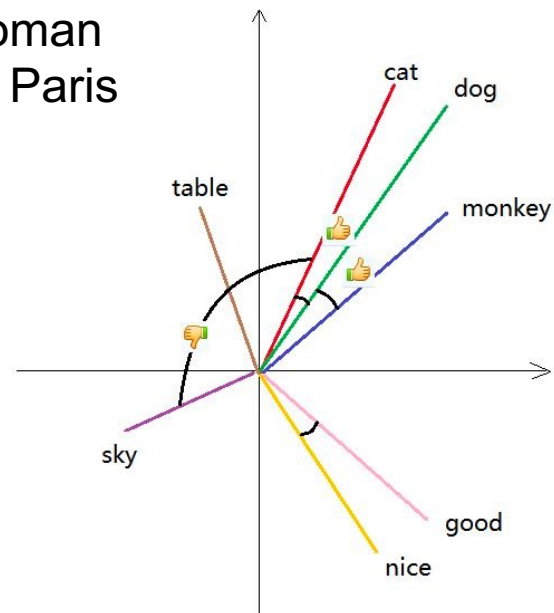
「Lebrn James wins All-Star Game MVP」

The evolution of DNN for NLP (1/2)



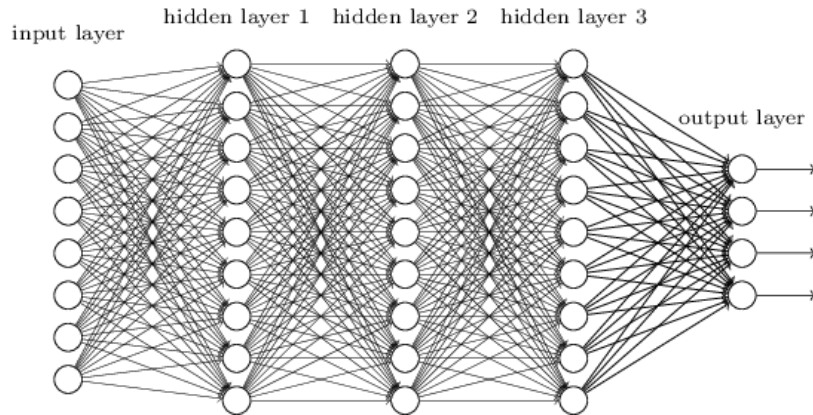
King + Queen = Man + Woman
Japan + Tokyo = France + Paris

Word Vectors

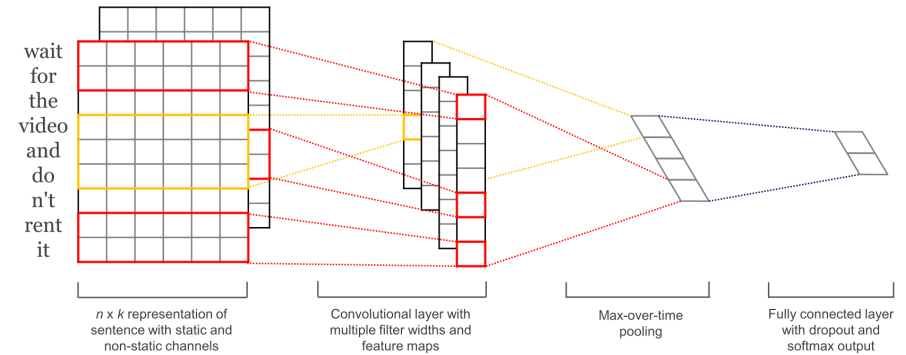


The evolution of DNN for NLP (2/2)

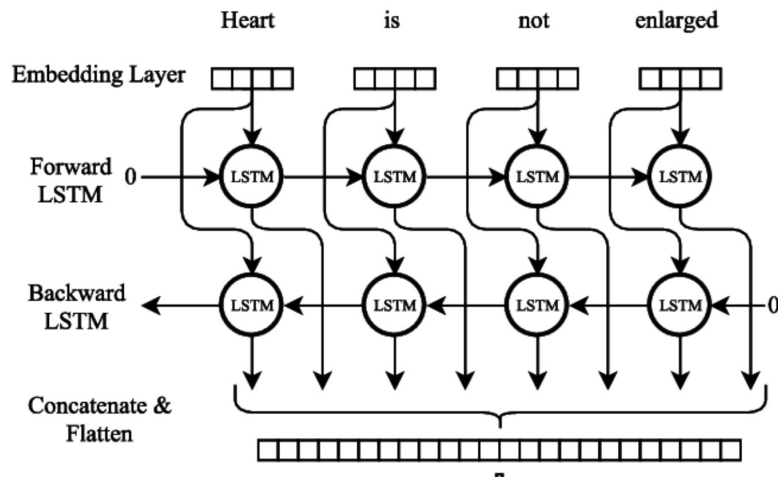
MLP



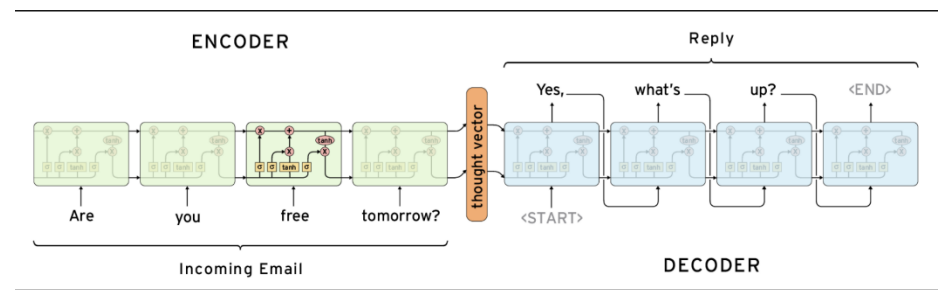
TextCNN



BiLSTM



Seq2Seq



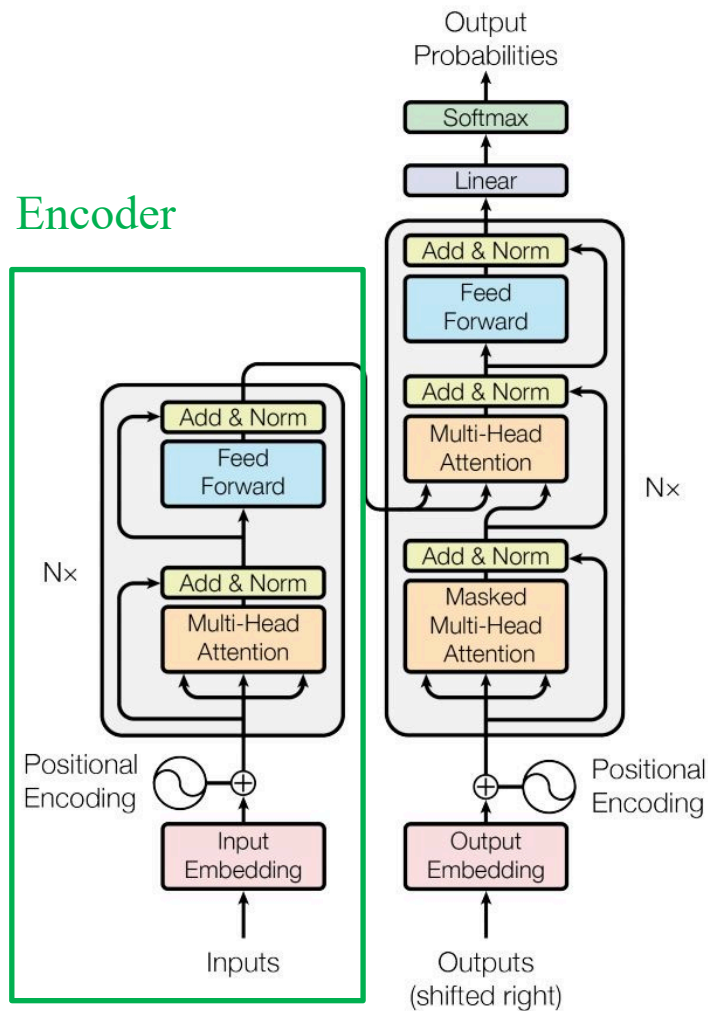
<https://research.googleblog.com/2015/11/computer-respond-to-this-email.html>

Kim, "Convolutional Neural Networks for Sentence Classification", arXiv preprint arXiv:1408.5882, 2014.

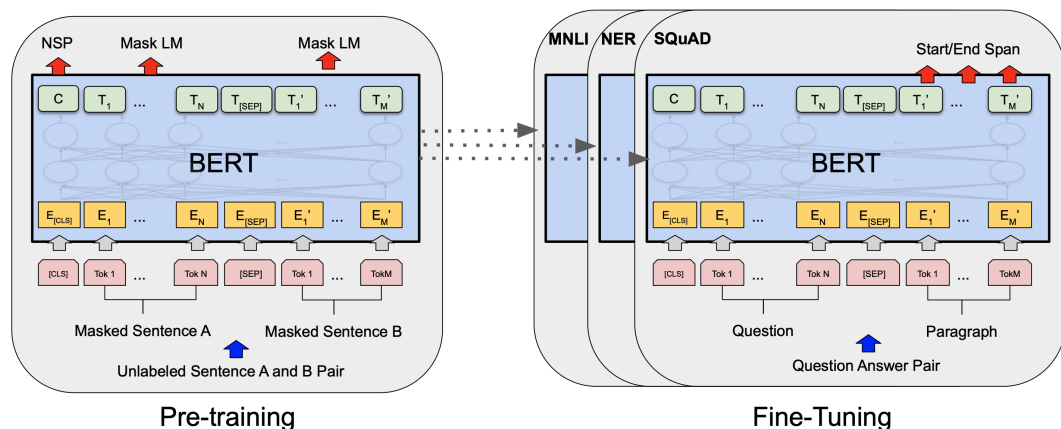
The Current State of the Art in NLP

BERT - Bidirectional Encoder Representations from Transformers

Encoder



People only need to focus on the Downstream Model Design of Pre-trained Language Model for NLP Task



BERT is actually the Encoder in Transformer, but it has many layers



*340M
parameters!!*

ELMO, BERT, ERNIE, and BigBird

Big Bird: Transformers for
Longer Sequences

<https://arxiv.org/abs/2007.14062>

ELMo
(**E**MBEDDINGS from
LANGUAGE **M**ODELS)

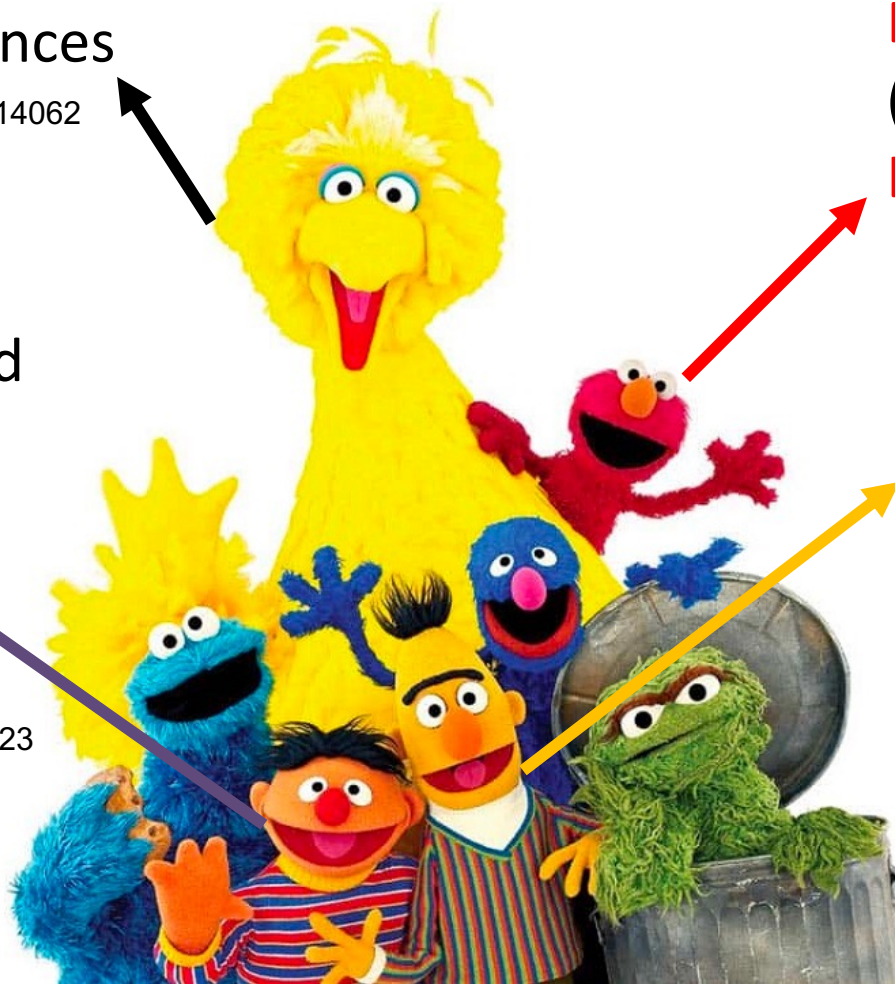
<https://arxiv.org/abs/1802.05365>

ERNIE (**E**nhanced
Representation
through
Knowledge
Integration)

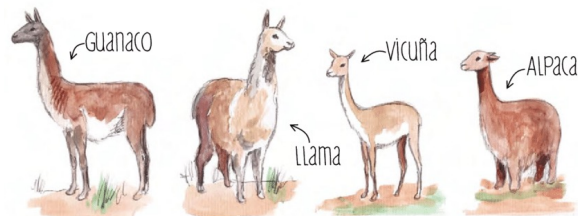
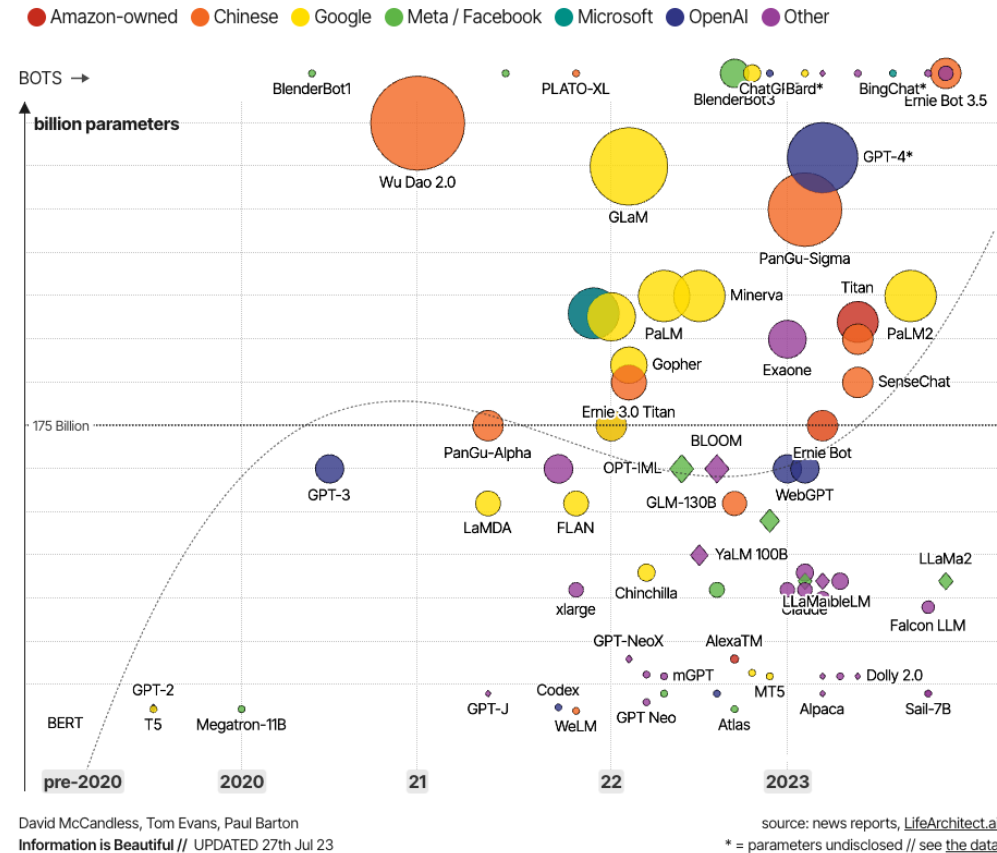
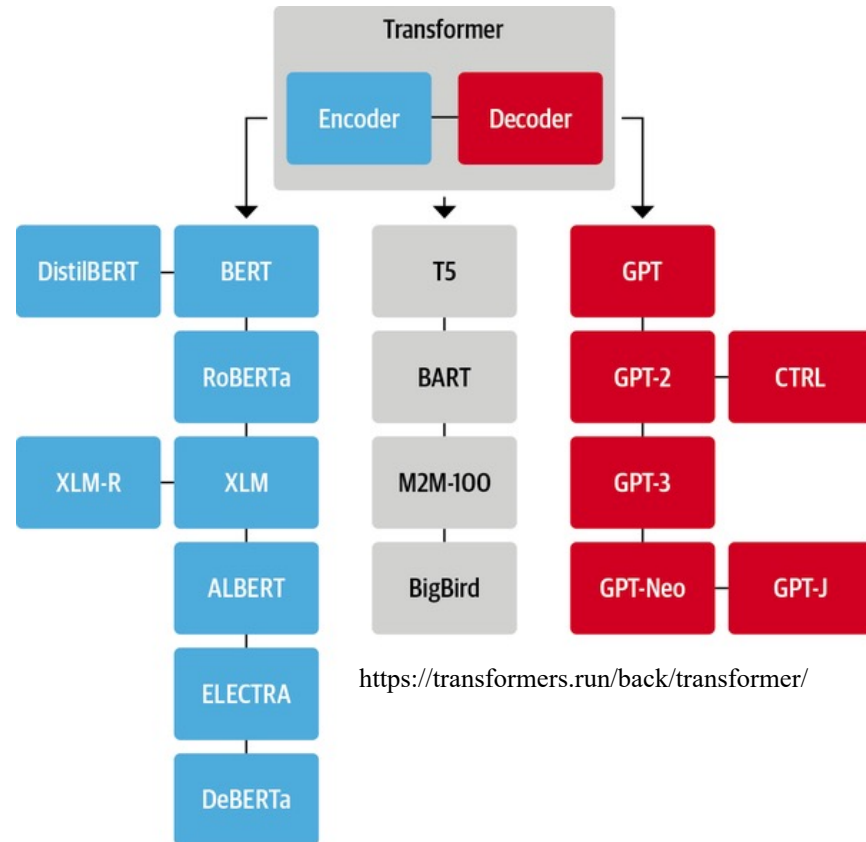
<https://arxiv.org/abs/1904.09223>

BERT (**B**idirectional
Encoder
Representations
from **T**ransformers)

<https://arxiv.org/abs/1810.04805>



The Current State of the Art in NLP



Sweeping wave of ChatGPT



ChatGPT

Generative Pre-trained Transformer



[API](#) [RESEARCH](#) [BLOG](#) [ABOUT](#)

ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.

November 30, 2022
13 minute read



<https://openai.com/blog/chatgpt/>

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

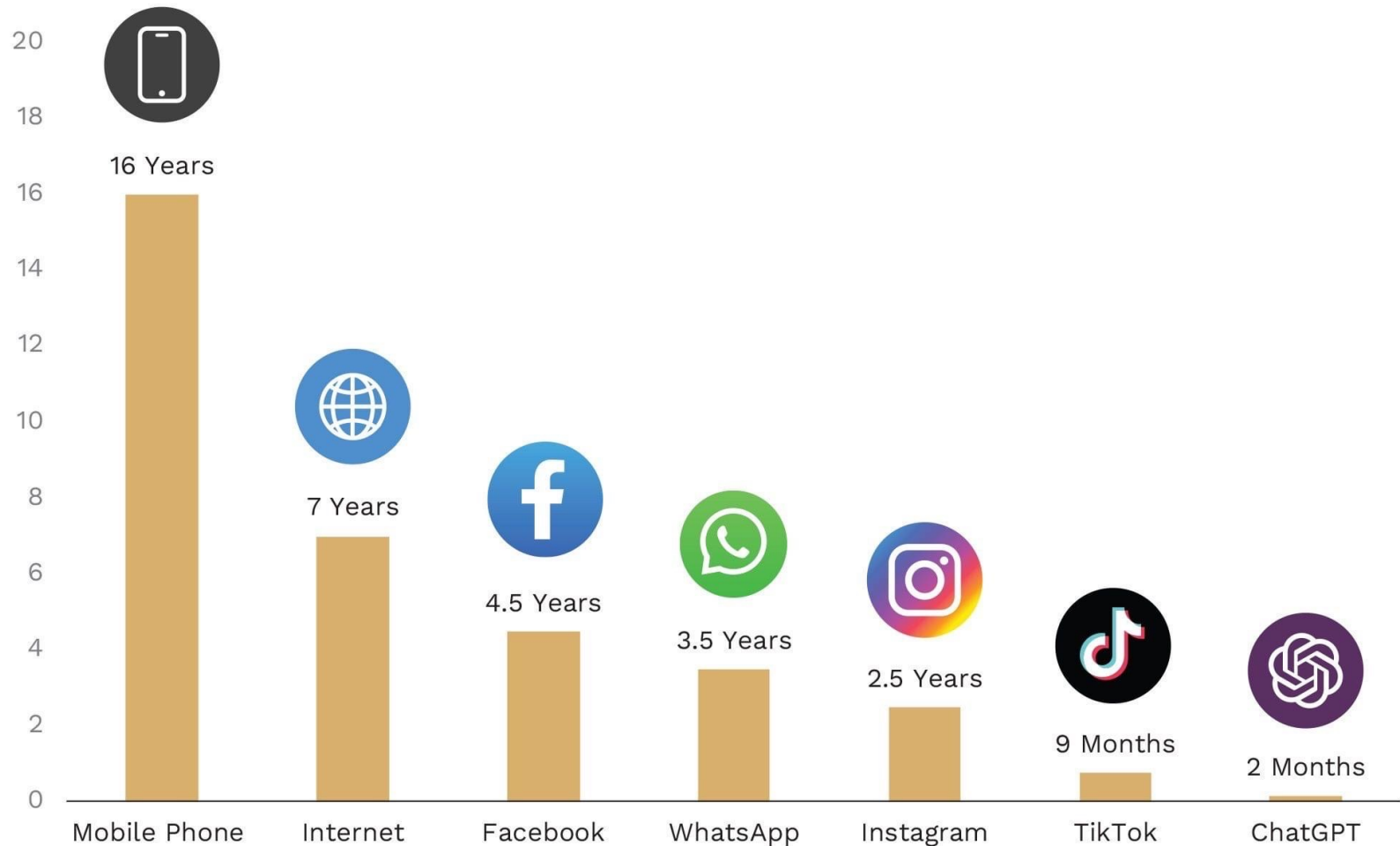
May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Free Research Preview: ChatGPT is optimized for dialogue. Our goal is to make AI systems more natural to interact with, and your feedback will help us improve our systems and make them safer.

<https://chat.openai.com/chat>

Time to reach 100 million users

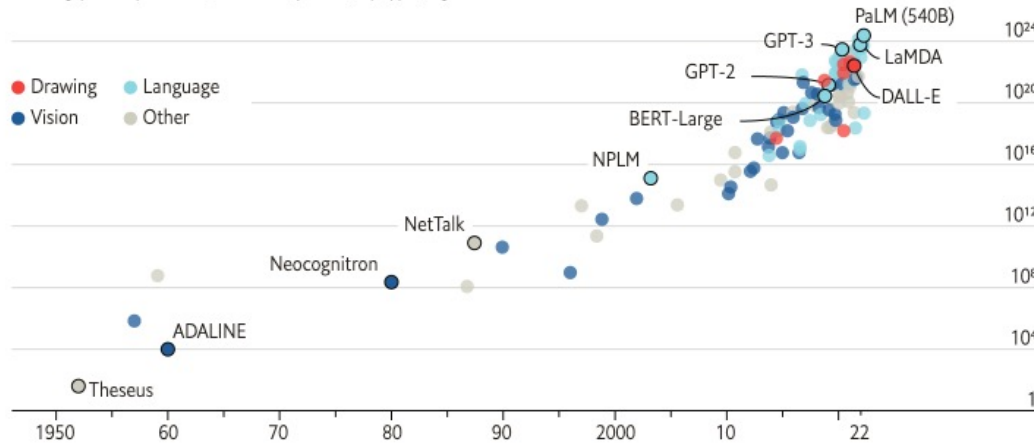


Source: Linas Beliunas

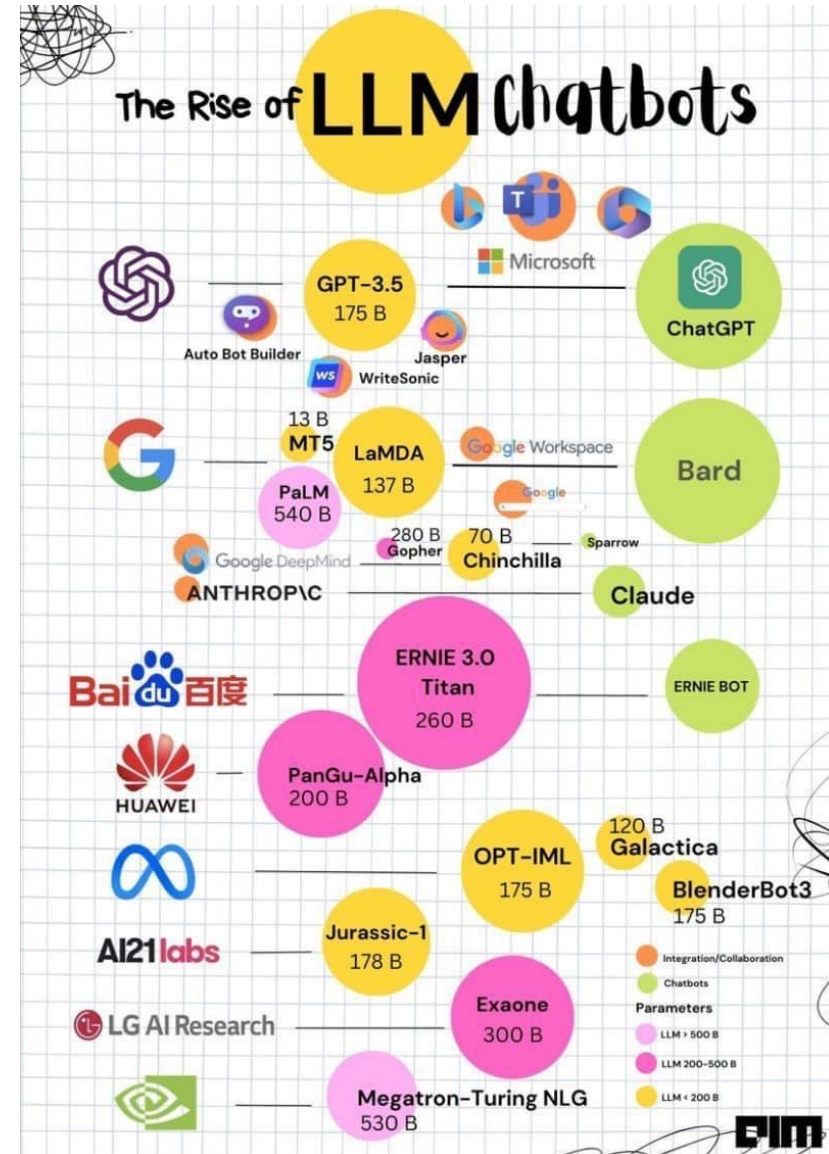
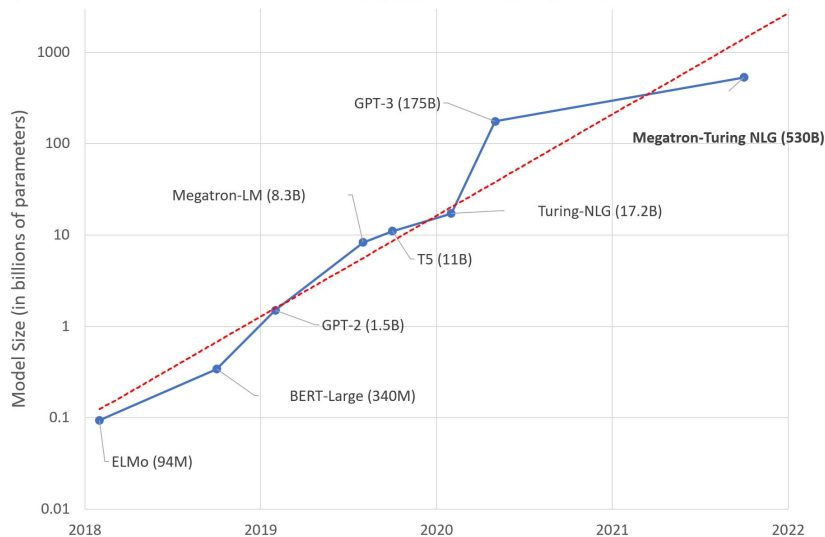
The models become larger and larger ...

The blessings of scale

AI training runs, estimated computing resources used
Floating-point operations, selected systems, by type, log scale



Sources: "Compute trends across three eras of machine learning", by J. Sevilla et al., arXiv, 2022; Our World in Data

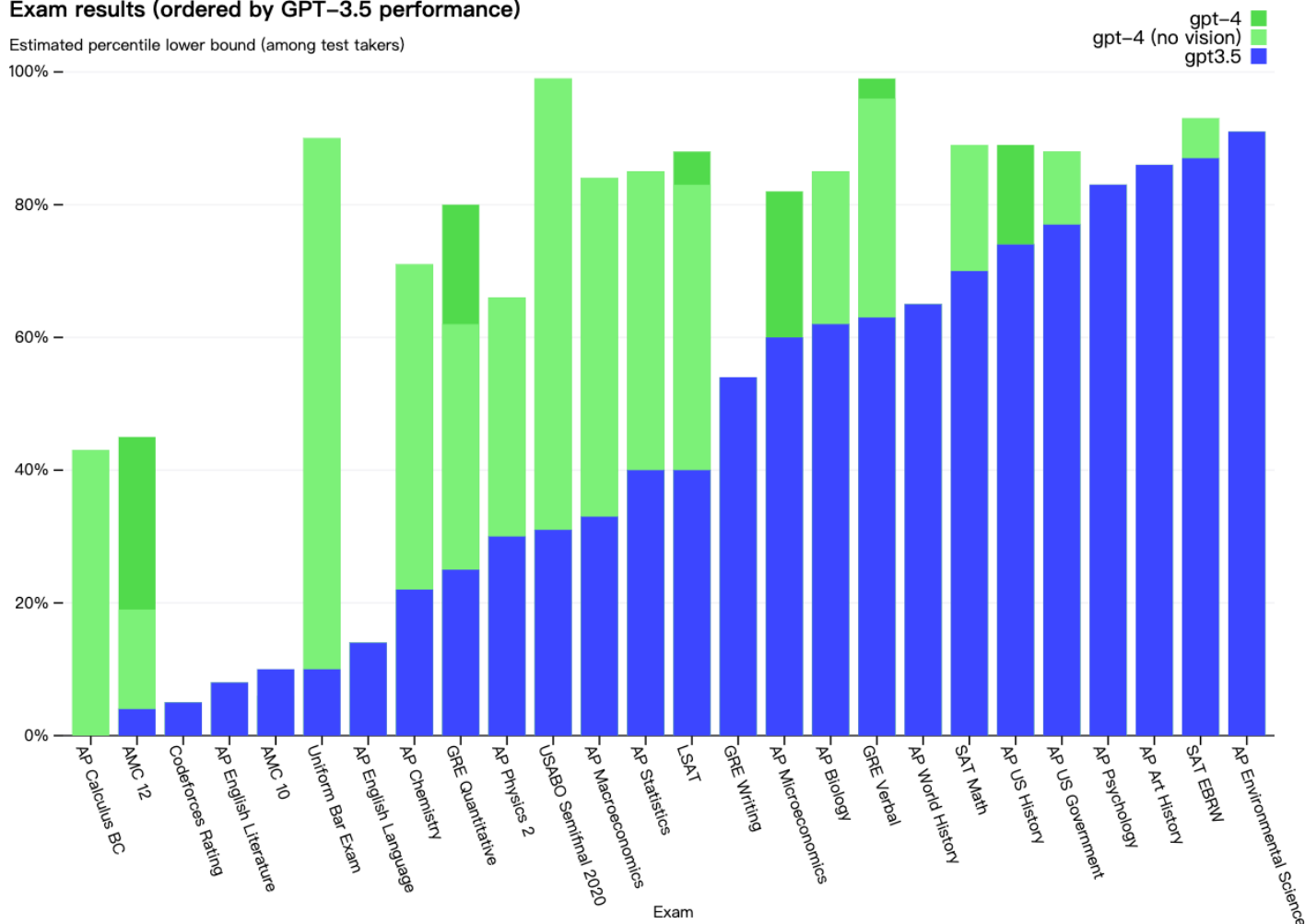


ChatGPT's exam ability



Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)

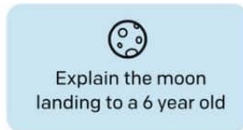


How to train ChatGPT

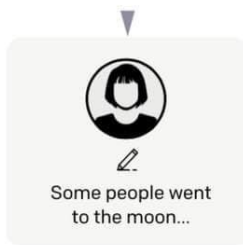
Step 1

Collect demonstration data, and train a supervised policy.

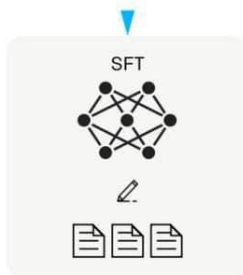
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



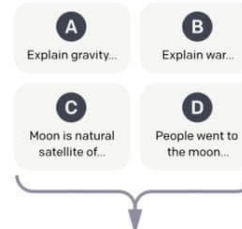
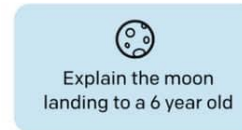
This data is used to fine-tune GPT-3 with supervised learning.



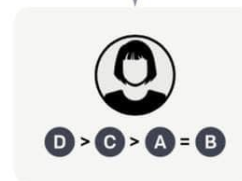
Step 2

Collect comparison data, and train a reward model.

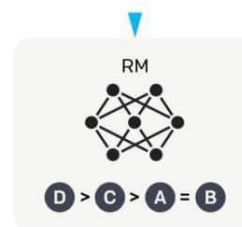
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



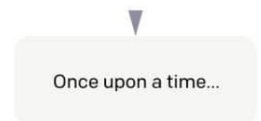
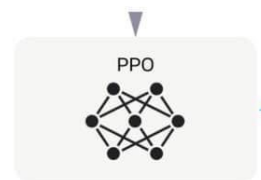
Step 3

Optimize a policy against the reward model using reinforcement learning.

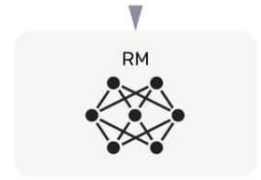
A new prompt is sampled from the dataset.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



Generative AI = Automatic Knowledge Production

如何讓ChatGPT當助教？ 10個好用實例



註：節錄自維吉尼亞大學經濟系教授柯里奈克，給經濟學家的25個使用建議

資料來源：美國經濟研究院 研究整理：卞中佩、蕭歆諺

This March is very Exciting

- Google released AI into Workspace
- OpenAI publishes GPT-4
- Midjourney publishes V5
- Baidu published 文心一言
- Microsoft publishes 365 Copilot
- Stanford University open-sources new model Alpaca 7B
- Traditional Chinese Alpaca released
- Meta open source LLaMa 2

⋮



VS

Alphabet



Bard AI

ChatGPT (cont.)

我想要請張詠淳老師指導我的研究，請幫我寫一封給張老師的自薦函



親愛的張老師：

我很榮幸能夠請您為我指導我的研究。我是一名研究生，我對您的研究非常感興趣。我相信您的專業知識和經驗將會為我的研究帶來巨大的價值。

我已經對您的研究和成就有了基本的了解，並且我非常期待能夠從您身上學到更多。我非常熱忱地想要投入到這個領域的研究中，我相信我有能力在您的指導下完成高質量的研究。

我非常感謝您花時間考慮我的請求，我期待著您的回復。

謝謝您！

Sincerely,
[Your Name]



電影鐵達尼號的結局



電影《鐵達尼號》的結局是主角羅丹·基德（Leonardo DiCaprio）扮演的費茲·麥爾斯（Jack Dawson）因為救了莉莉·柯林斯（Rose DeWitt Bukater，由凱特·溫斯蘭飾演）而喪生，最後莉莉回到了現實世界，並且留下了一張畫，描繪了她和費茲在鐵達尼號上的情景。這是一部悲傷而感人的結局，也是電影中最令人難忘的景象之一。



請寫一首關於下雨天的絕妙好詩



下雨天

黑暗的雲朵掩蓋著天空
窗外的空氣濕濕潤潤
下雨了，一整天
滴滴答答聽得清楚

坐在家裡，窗外落著雨
喝著茶，看著雨
滴滴答答拍打著窗戶
我就這樣安靜地坐在這裡

下雨天，我的心情沈鬱
但那些煩惱都逃不了
下雨天，我的思緒飄忽
回憶起過往的風雨

不管下多大的雨
我都會堅強地撐著
因為我知道雨過天晴
明天就會是好天氣



下雨天，我不會害怕
因為我知道雨會過去
下雨天，我會靜靜地坐著
等待著那美好的彩虹

ChatGPT (cont.)




請幫我用Python寫一個二元搜尋樹的程式



好的，我可以幫助您寫一個 Python 程式來建立一個二元搜尋樹。二元搜尋樹是一種特殊的樹狀結構，它可以讓您快速查找、插入和刪除節點。  

首先，我們可以定義一個類別來表示每個樹節點。每個節點都有一個數值，以及左子節點和右子節點的指標。您可以將這些東西放在節點類別中，如下所示：

```
class Node:
    def __init__(self, value):
        self.value = value
        self.left = None
        self.right = None
```

 Copy code

接下來，我們需要定義一個類別來表示整棵二元搜尋樹。這個類別應該包含方法來插入節點、查找節點和刪除節點。

首先是插入節點的方法。我們可以定義一個 `insert()` 方法，它接收一個數值，並在適當的位置插入一個新的節點。我們可以這樣實現它：

```
class BinarySearchTree:
    def __init__(self):
        self.root = None

    def insert(self, value):
        new_node = Node(value)
        if self.root is None:
            self.root = new_node
            return

        current_node = self.root
        while True:
            if value < current_node.value:
                # 插入左子樹
                if current_node.left is None:
                    current_node.left = new_node
                    return
                current_node = current_node.left
            else:
                # 插入右子樹
                if current
```

 Copy code

ChatGPT (cont.)

CH

要怎麼進行臨床試驗配對？



臨床試驗配對是指在臨床試驗中，將研究受試者的資料與實驗組或對照組的受試者進行比較，以確保研究結果的可靠性和準確性。



進行臨床試驗配對的步驟通常包括：

1. 確定研究的目的和假設：在進行臨床試驗配對之前，首先要確定研究的目的和假設，以便確定應該針對哪些因素進行配對。
2. 確定受試者的選擇標準：確定受試者的選擇標準是指確定哪些受試者可以參加研究，哪些受試者不能參加研究。選擇標準通常包括年齡、性別、疾病狀態等。
3. 收集受試者的資料：在進行臨床試驗配對時，需要收集受試者的資料，以便進行配對。通常，需要收集受試者的基本資料，如年齡、性別、疾病狀態等，以及相關的客觀測量數據，如血壓、血糖等。
4. 進行配對：在進行配對時，需要按照研究的目的和假設，選擇適當的配對方法，將受試者的

CH

治療肺癌的標靶治療藥物有哪些？



治療肺癌的標靶治療藥物是指針對肺癌細胞特定的抗原或酪氨酸激酶受體進行的治療。標靶治療藥物的作用是阻斷癌細胞生長和繁殖，減少癌細胞的代謝，降低肺癌的活性，從而治療癌症。



治療肺癌的標靶治療藥物包括：

1. 阿斯匹林：它是一種酪氨酸激酶抑制劑，可以抑制肺癌細胞的生長和繁殖。
2. 斯諾克替尼：它是一種酪氨酸激酶受體拮抗劑，可以抑制肺癌細胞的生長和繁殖。
3. 胞嘧啶標靶藥物：它是一種抗腫瘤藥物，可以抑制肺癌細胞的生長和繁殖。
4. 化療藥物：它們是一種抗腫瘤藥物，可以抑制肺癌細胞的生長和繁殖。

治療肺癌的標靶治療

ChatGPT (cont.)

CH

有個病患的資訊描述如下，你覺得他是得到什麼病？

病史：

1. 現在病史(中文書寫為主，專業術語文字可加註英文)：

x先生為 45 歲男性病人，3 天前無明顯誘因出現右足背紅斑，迅速於一天之內擴散至右小腿前側與後側，於兩天前進展至明顯的小腿紅腫熱痛，並且有發燒、畏寒等症狀。兩日前於台北○○醫院皮膚科就診，經投與抗生素 dicloxacillin 治療，病況並無好轉，於今日凌晨前來本院急診就醫。急診發現體溫 39.3°C，白血球 18.6K/uL，分葉形白血球 (segment) 92%，C 型反應性蛋白 (CRP) 20.45。系統回顧發現病人並無呼吸道症狀，亦無腹瀉、腹痛、嘔吐等症狀，大小便亦無異常。食慾、睡眠尚正常。擬住院給予靜脈注射抗生素治療。

2. 過去病史(中文書寫為主，專業術語文字可加註英文)：

10 年前有類似病史，經治療後好轉，診斷為蜂窩組織炎。

自訴過去無高血壓、糖尿病、肝炎、肺結核、血友病等疾病。

自訴過去無外傷、手術及輸血情形。

自訴過去無藥物及食物過敏。

預防接種紀錄不詳。



根據病史描述，這位病患可能患有蜂窩性組織炎的復發。蜂窩性組織炎是一種細菌感染引起的皮膚和軟組織疾病，通常由溶血性鏈球菌感染引起。病患出現紅腫熱痛的症狀，白血球計數和CRP指數升高，都是細菌感染的常見表現。病患曾有類似病史且治療後好轉，也支持這個診斷。擬住院給予靜脈注射抗生素治療也是一個常見的治療方案。然而，最終診斷仍需由醫師進一步評估，並依據臨床症狀和化驗結果做出判斷。

ChatGPT (cont.)

CH

那他需要住院嗎



根據描述，病人有明顯的紅腫熱痛、發燒、畏寒等症狀，且白血球數和 CRP 值也升高，可能表示有感染的情況。另外，經過抗生素治療後病況沒有好轉，因此醫院建議擬住院給予靜脈注射抗生素治療。因此，我認為病人需要住院治療。住院治療可以提供更完善的監測和治療，確保病人能夠妥善控制感染症狀，並且進一步進行相關檢查與治療。

ChatGPT (cont.)

CH

幫我針對此產生病歷



病歷：

姓名：X先生

性別：男

年齡：45歲

現病史：

X先生於3天前出現右足背紅斑，之後迅速擴散至右小腿前側與後側。2天前症狀進展至小腿紅腫熱痛，伴隨發燒、畏寒等症狀。2天前前往台北○○醫院皮膚科就診，投與抗生素 dicloxacillin 治療，然而病情未見好轉。今日凌晨前來本院急診就醫，發現體溫 39.3°C，白血球 18.6K/uL，分葉形白血球(segment)92%，C型反應性蛋白(CRP)20.45。系統回顧發現病人並無呼吸道症狀，亦無腹瀉、腹痛、嘔吐等症狀，大小便亦無異常。擬住院給予靜脈注射抗生素治療。

過去病史：

10年前曾有類似病史，經治療後好轉，診斷為蜂窩組織炎。自述過去無高血壓、糖尿病、肝炎、肺結核、血友病等疾病。自述過去無外傷、手術及輸血情形。自述過去無藥物及食物過敏。預防接種紀錄不詳。

初步診斷：

根據症狀和檢查結果，初步診斷為細菌性蜂窩組織炎。

治療計劃：

擬住院給予靜脈注射抗生素治療，密切觀察病情變化，以及評估治療反應和併發症的發展。在治療期間，還需對病人的白血球計數、白細胞分類以及C型反應性蛋白等指標進行監控。在治療期間，還需保持病人的身體功能，盡量保持適當的飲食和運動，避免併發症的

CH

幫我轉成英文版本



Medical History:

1. Present Illness:

Mr. X is a 45-year-old male who developed erythema on the dorsal surface of his right foot three days ago, which rapidly spread to the anterior and posterior aspects of his right lower leg within one day. Two days ago, he developed significant redness, swelling, and pain in his lower leg, along with fever, chills, and other symptoms. He visited the dermatology department of a hospital in Taipei two days ago and was treated with dicloxacillin, but his condition did not improve. He presented to the emergency department of our hospital in the early morning today. On examination, his body temperature was 39.3°C, white blood cell count was 18.6K/uL, segmented neutrophils were 92%, and C-reactive protein was 20.45. A review of systems revealed no respiratory symptoms, diarrhea, abdominal pain, vomiting, or other symptoms, and his bowel and bladder habits were normal. His appetite and sleep were also normal. He is planned to be admitted for intravenous antibiotic treatment.

2. Past Medical History:

He had a similar illness ten years ago, which improved after treatment and was diagnosed as cellulitis. He denies any history of hypertension, diabetes mellitus, hepatitis, tuberculosis, or hemophilia. He denies any history of trauma, surgery, or blood transfusion. He has no known drug or food allergies. His vaccination history is unknown.

ChatGPT (cont.)

CH

請用中文幫我產生幾句話可以向病患說明病情



您好，根據您的病歷，您可能罹患了蜂窩性組織炎，這是一種細菌感染引起的皮膚疾病。您的病情已經進展到明顯的紅腫熱痛，並且有發燒、畏寒等症狀。醫生已經為您開立了靜脈注射抗生素的治療方案，需要您住院觀察和治療，請您放心，我們會盡最大努力照顧您的健康。

ChatGPT (cont.)

CH

針對這樣的病患，他可以怎麼照護傷口或注意的事項



對於這位病患，以下是一些照護傷口和注意事項：

1. 保持傷口清潔，定期更換敷料，避免污染和感染。
2. 避免觸碰傷口，避免摩擦或拉扯傷口。
3. 在醫生指示下，適當地服用抗生素和止痛藥。
4. 保持身體清潔，避免交叉感染。
5. 休息充足，避免長時間站立或坐著，減少傷口負擔。
6. 按時前往醫院進行複診和治療，並遵循醫生的建議。

從資源回收數據產生報告

CH

以下是111年度資源回收統計資訊，包含各個行政區以及總回收公噸數，請幫我產生一份報告，著重在趨勢分析、異常增減情形，並說明原因

12月份

台東市 2,401,000公噸

成功鎮 217,421公噸

關山鎮 140,823公噸

卑南鄉 252,450公噸

大武鄉 59,390公噸

太麻里鄉 277,407公噸

東河鄉 156,659公噸

長濱鄉 58,767公噸

鹿野鄉 114,175公噸

池上鄉 214,503公噸

綠島鄉 77,940公噸

延平鄉 67,309公噸

海端鄉 45,235公噸

達仁鄉 41,042公噸

金峰鄉 67,130公噸

蘭嶼鄉 58,394公噸

⋮

從資源回收數據產生報告 (cont.)



報告：111年度台東縣各行政區回收公噸數統計分析

概覽

本報告統計了台東縣各行政區從9月份到12月份的回收公噸數，以探討回收趨勢和異常增減情形。

趨勢分析

從表格中可以看出，從9月份到11月份，台東市的回收公噸數呈現逐漸增加的趨勢，然而在12月份，回收公噸數卻驟降。在其他行政區，除了成功鎮和鹿野鄉在10月份回收量略微上升外，大部分行政區的回收公噸數也呈現逐漸下降的趨勢。

異常增減情形

從上述趨勢分析中可以看出，在12月份，台東市的回收公噸數驟降了3%。然而在其他行政區，驟降的幅度更大，最高可達15%。其中，成功鎮和鹿野鄉在10月份回收量略微上升，但在11月份和12月份的回收量都下降了，顯示這些行政區的回收狀況可能受到了季節因素和其他不明因素的影響。

從資源回收數據產生報告 (cont.)

原因分析

回收量的減少可能與節假日和季節性因素有關。12月份是冬季，天氣寒冷，人們通常待在室內，因此產生的廢棄物減少。此外，年底節日如聖誕節和跨年夜等也可能對回收量產生影響。此外，2021年至2022年間，全球疫情肆虐，台灣也實施了多次疫情防疫措施，可能也對回收量造成了一定的影響。

結論

綜上所述，台東縣各行政區的回收公噸數在11月和12月呈現了下降的趨勢，可能與節假日和季節性因素有關，而成功鎮和鹿野鄉的回收量下降更為明顯，原因可能還受到其他不明因素的影響。這些統計資訊可以為當地政府制定更有效的資源回收政策提供參考。

從資源回收數據產生報告 (cont.)

CH

請列出變動量前三大的行政區



變動量指的是相鄰兩個月中回收公噸數的增加或減少量。

以資源回收公噸數從11月份到12月份的變化量作為參考，列出變動量前三大的行政區如下：

1. 台東市：-32,794公噸
2. 卑南鄉：+11,270公噸
3. 成功鎮：-17,359公噸

台東市的回收公噸數在11月到12月之間減少了32,794公噸，相較於其他行政區變化量的大小，顯得特別明顯。卑南鄉在這個月份的回收量有顯著的增加，增加了11,270公噸；而成功鎮的回收量則有較大的下降，減少了17,359公噸。

ChatGPT也可以很SDGs

CH 聯合國2030永續發展目標Sustainable Development Goals是什麼?有幾項核心目標?各自是什麼?

 聯合國2030永續發展目標(Sustainable Development Goals, SDGs)是一個全球共識的行動計畫，旨在2030年前實現全球永續發展。它由17個核心目標組成，涵蓋了經濟、社會和環境三個方面，以促進經濟發展、社會進步和環境保護，實現可持續發展。

以下是17個核心目標：

1. 治理：確保平和、公正和有效的治理，促進可持續發展。
2. 終結貧困：消除所有形式的貧困和不平等，實現全人類的基本權利。
3. 食品安全和可持續農業：確保可持續農業生產和供應鏈，消除飢餓和營養不良。
4. 健康和福利：確保全人類的健康和福祉，並消除艾滋病、結核病和瘧疾等傳染病。
5. 具有性別平等的教育：促進平等的教育機會，消除性別不平等和文盲。
6. 潔淨水和衛生：確保可持續的水和衛生設施，以減少疾病和死亡。
7. 可負擔和清潔能源：促進可持續的能源使用，並減少對化石燃料的依賴。
8. 面向包容和持續增長的經濟：促進包容性和可持續的經濟增長，實現全球就業和經濟機會。
9. 產業、創新和基礎設施：促進可持續的產業發展和基礎設施建設，以支持經濟發展。
10. 減少不平等：消除國內和國際的不平等現象，確保所有人享有公平和平等的機會。
11. 可持續城市和社區：促進可持續的城市和社區發展，提高居住品質和可持續性。



12. 可持續生產和消費：促進可持續的生產和消費模式，減少浪費和環境污染。
13. 氣候行動：採取行動應對氣候變化，減少溫室氣體排放和提高社會適應能力。
14. 海洋和海岸生態系統：保護海洋和海岸生態系統，促進可持續漁業和海洋旅游。
15. 生物多樣性和生態系統：保護和恢復陸地和海洋生物多樣性和生態系統，以維持生態平衡和生態系統服務。
16. 和平、公正和強大的機構：建立和平、公正和包容的機構，促進可持續發展和社會進步。
17. 夥伴關係：加強全球夥伴關係，以實現可持續發展目標。

這些目標旨在促進全球可持續發展，從而實現經濟、社會和環境的協調和平衡。每個目標都涉及不同的問題和挑戰，需要各國政府、企業和公民共同努力實現。

ChatGPT也可以很SDGs (cont.)

CH

SDGs 的17個核心目標如下：

消除貧困

飢餓零容忍

健康與福祉

優質教育

性別平等

淨水與衛生

可負擔清潔能源

體面工作與經濟增長

產業創新和基礎建設

減少不平等

可持續城市和社區

負責任的消費和生產

應對氣候變化

海洋和海岸線的保護

生物多樣性和生態系統保護

和平、公正、強大的機構

合作夥伴關係的建立

請問『高雄市環境保護統計年報』中包含了

(1)空氣品質監測及污染防治(2)噪音監測及防制(3)水質監測及污染防治(4)廢棄物管理(5)環境衛生及毒化物管理

與SDGs17項核心目標哪些有關聯？

ChatGPT也可以很SDGs (cont.)



根據所提供的資訊，可以將高雄市環境保護統計年報所包含的項目與SDGs的17項核心目標做以下的關聯：

(1)空氣品質監測及污染防制：與SDGs 3.6（健康與福祉-空氣品質）、SDGs 11.6（可持續城市和社區-空氣品質）和SDGs 13.1（應對氣候變化-減少大氣污染）有關聯。

(2)噪音監測及防制：與SDGs 3.9（健康與福祉-噪音污染）和SDGs 11.7（可持續城市和社區-噪音污染）有關聯。


(3)水質監測及污染防治：與SDGs 6.3（淨水與衛生）和SDGs 14.1（海洋和海岸線的保護-水質）有關聯。


(4)廢棄物管理：與SDGs 11.6（可持續城市和社區-廢棄物管理）和SDGs 12.5（負責任的消費和生產-廢棄物管理）有關聯。


(5)環境衛生及毒化物管理：與SDGs 3.9（健康與福祉-環境污染和毒物）和SDGs 11.6（可持續城市和社區-環境衛生）有關聯。

Using ChatGPT for Midjourney-specific Prompts

What is the concept you would like me to generate a prompt for?


 An abstract tree

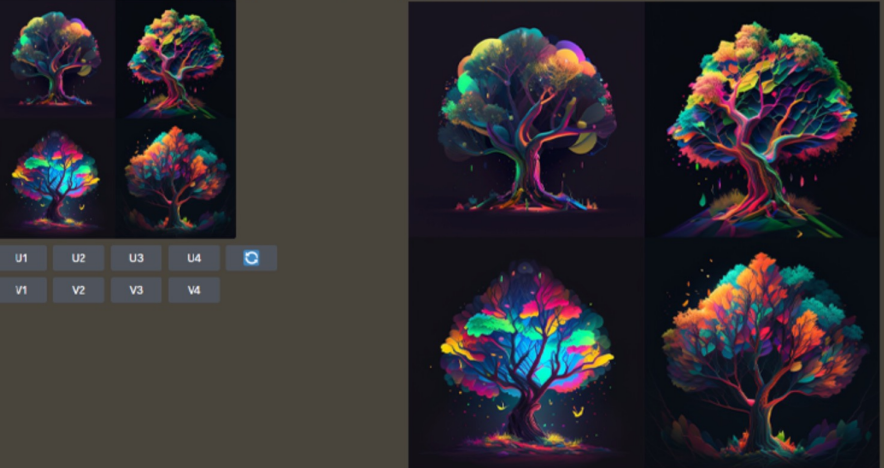
 Here is a prompt for an abstract tree:

 Copy code

```
a digital art illustration of an abstract tree, by Beeple, surrealism, color gradients, in a forest, with a black background, geometric shapes, neon colors, high detail illustration, vibrant colors, abstract landscape, inspired by Vincent van Gogh's Starry Night --v 4 --q 2 --stylize 1000
```

Would you like three different options for prompts for this concept, or would you like to move on to a new concept?

Midjourney Bot  Today at 7:40 PM
a digital art illustration of an abstract tree, by Beeple, surrealism, color gradients, in a forest, with a black background, geometric shapes, neon colors, high detail illustration, vibrant colors, abstract landscape, inspired by Vincent van Gogh's Starry Night --v 4 --q 2 --stylize 1000 - @Xanthius (fast)



AI Avatar



AI Avatar



AI-synthesized faces are indistinguishable from real faces and more trustworthy

民視新聞台

「AI正妹」火辣程度直逼真人 網：小模恐失業



民視新聞網

2023年2月24日



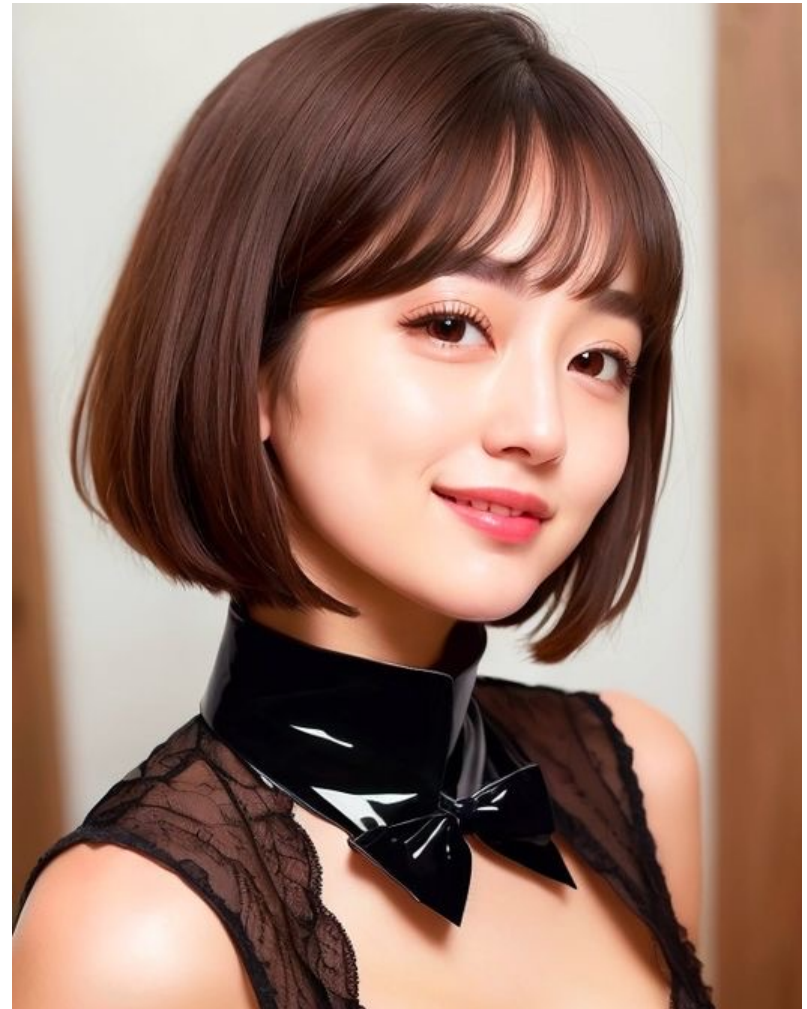
國際中心 / 吳峻光報導



近來AI人工智慧在全球掀起一股熱潮，除了聊天機器人之外，AI繪圖也超級夯，甚至連寫實系的照片都能算出來。最近日本就有一名神人，用AI繪圖創作出無數以假亂真的正妹，作品在網路上瘋傳，有鄉民將其轉發到PTT上，果然馬上引起轟動，但也有不少人感慨，「小模真的要失業了」、「有這個誰還要看真的小模」、「以後真人要不修圖，才能和AI圖有區別了」。



日本有神人用AI繪圖創作出各種美女。(圖 / 翻攝推特)



This photo is generated by using ChilloutMix!!

Metaverse - Resurrect the dead

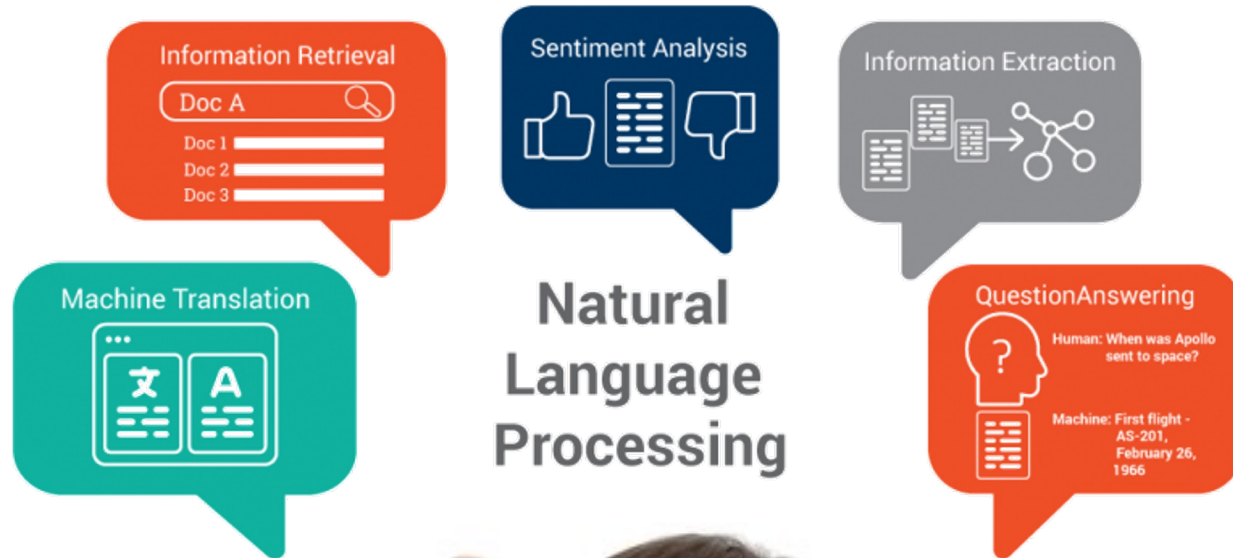
I Met You



Black Mirror: be right back

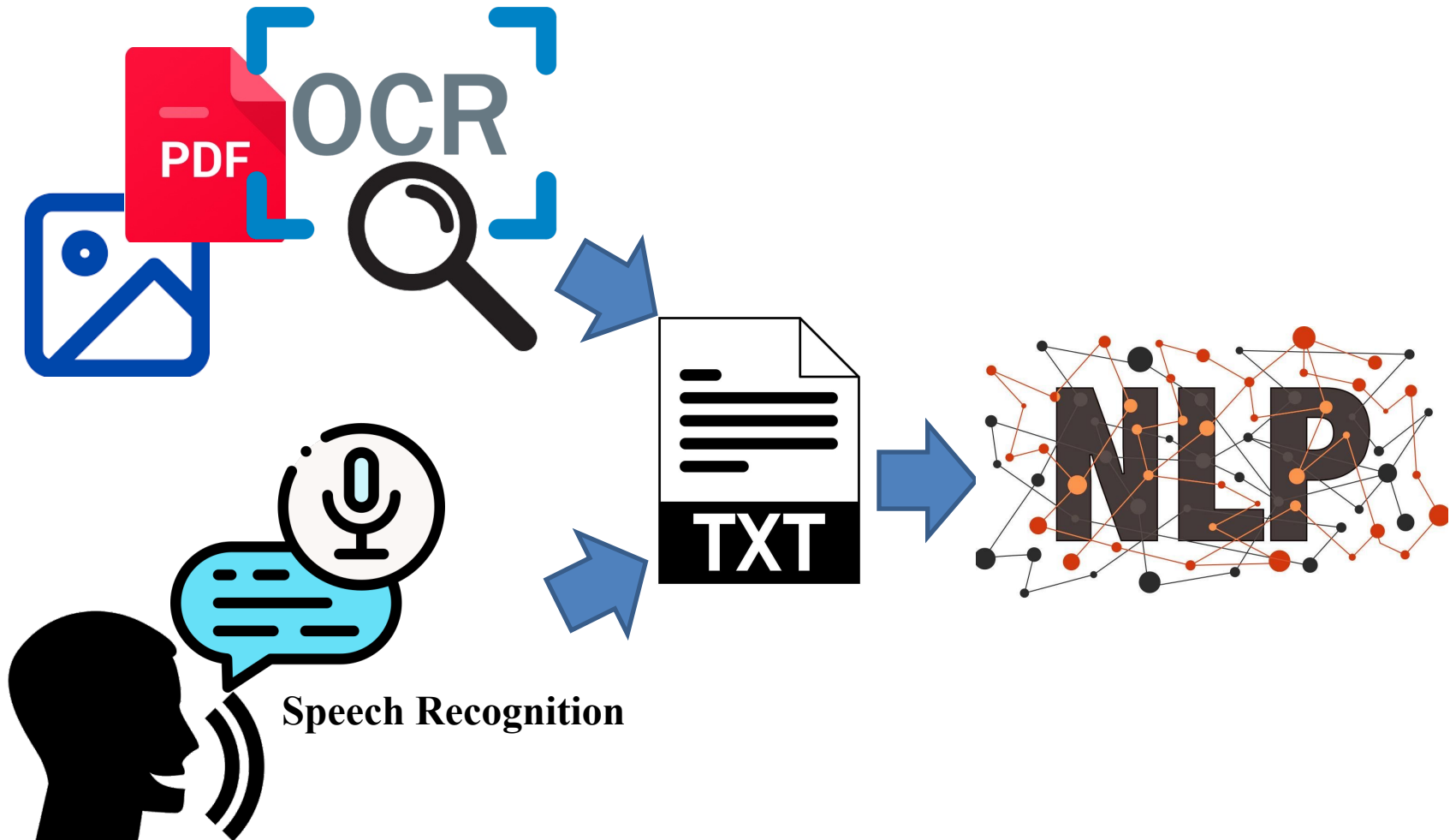


What are the applications of NLP ?



NLP Applications

Optical Character Recognition



Exploring Social Interaction

探索社群媒體中的互動關係

Reader-emotion Recognition

最快明天 油價暴漲3.1元

2012年04月01日

G+1 65



【綜合報導】經濟部官員昨透露，明天擬調漲油價，其中發油量最多的95無鉛漲幅高達9.5%，將漲3.1元。昨各地加油站已出現搶油車潮，加油站業者說，預估今天加油車會更多，已叫油商多準備油。消基會昨指政府帶頭漲價，會使物價連環漲，怎麼活下去？

Facebook 讚 338 萬 馬上按讚 加入Yahoo奇摩新聞粉絲團

推薦 2.2 萬 Tweet 1 Plurk! G+1 1 列印

看完覺得

實用	感人	開心	超扯	無聊	害怕	難過	火大
1%	0%	4%	13%	1%	0%	0%	80%

火大新聞熱門推薦 »

- 捷運鋼橋墜／工人控訴遭揚逼工 為逃11萬罰款賠4命
- 捲入阿帕契風波 小S：講話大聲，顯得智商不高
- 臉書嗆「智商不高」諷網友？小S兩天內急刪文

分享： Facebook Twitter Plurk 共有 1058 投票



Reader-emotion Recognition for Publicity Mining

梅根專訪風波 川普：她不是好人

5

國際中心蔡姍伶/綜合報導

2021年3月15日 週一 上午9:05 · 3分鐘 (閱讀時間)

LINE

f

✉



▲美國前總統川普，與英國哈利王子的妻子梅根，數年前就曾結下樑子，現在又傳出川普於專訪風波後曾私下透露，說梅根不是好人。資料照。(圖/美聯社/達志影像)

Facebook 讚 338 萬 馬上按讚 加入Yahoo奇摩新聞粉絲團

推薦 2.2 萬 Tweet 1 Plurk! 8+1 1 列印

看完覺得

實用	感人	開心	超扯	無聊	害怕	難過	火大
1%	0%	4%	13%	1%	0%	0%	80%

火大新聞熱門推薦

- 捷運銅鑼堡/工人控訴遠揚逼工 為逃11萬罰款賠4命
- 捲入阿帕契風波 小S：講話大聲，顯得智商不高
- 臉書噏「智商不高」讓網友？小S兩天內急刪文

分享： Facebook Twitter Plurk 共有 1058 投票

在專訪中，大爆妯娌不合、英前總統川普身邊任職的幕僚透會像英國名嘴皮爾斯摩根(Piers

son Miller)，上週三接受川普訪問時，透露川普曾對他說：「我恨那一邊的，皮爾斯摩根很

一樣，因說錯話而告別公眾視

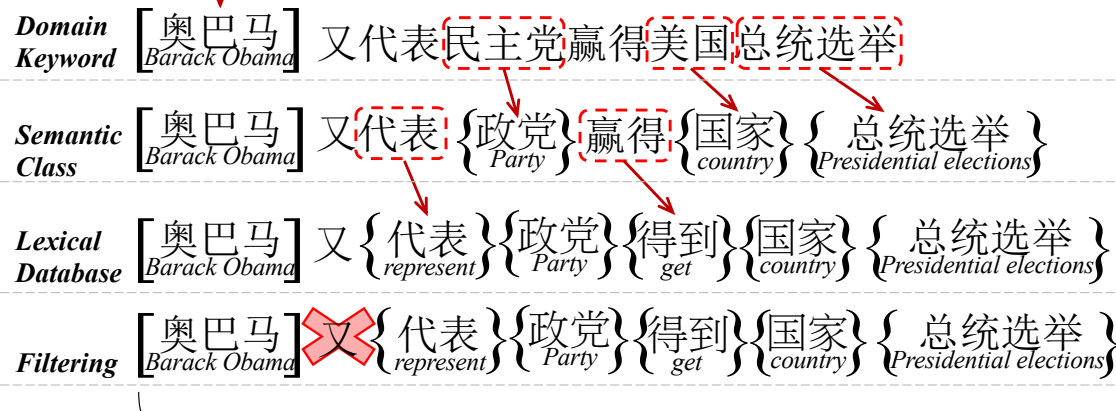


Learning Linguistic Template for Recognizing Reader-emotion

A clause C_n in an article:

Obama, representing the Democratic Party, won the U.S. Presidential election again

奥巴马又代表民主党赢得美国总统选举



sequence of crucial elements

WARM

[婦女women]:[救助help]:[小孩child]:[當作
treat]:[人human]:[認為consider]

WORRIED

[國家country]:[發生occur]:[地震
earthquake]:[劫難disaster]

Emotional Resonance Writing Assistance



*Learning Linguistic Pattern from Raw Text
for Knowledge Representation*

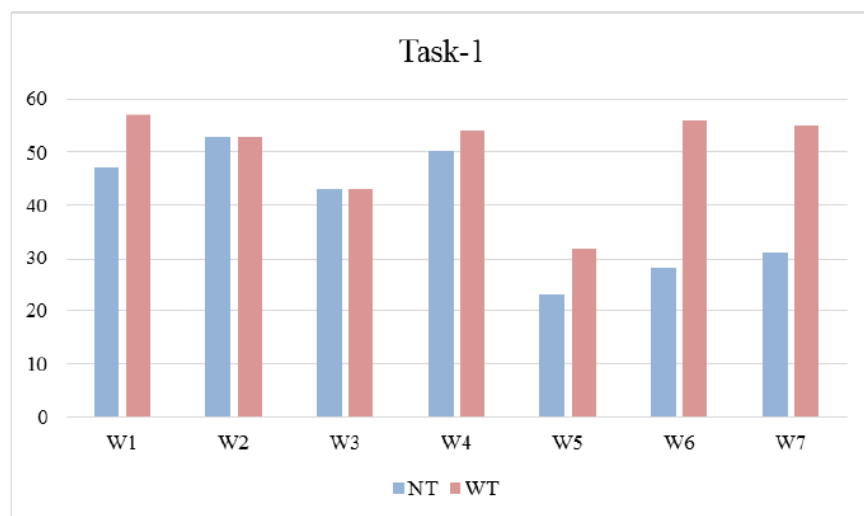


Figure 1. Comparison of the number of correct emotional response before and after utilizing templates.

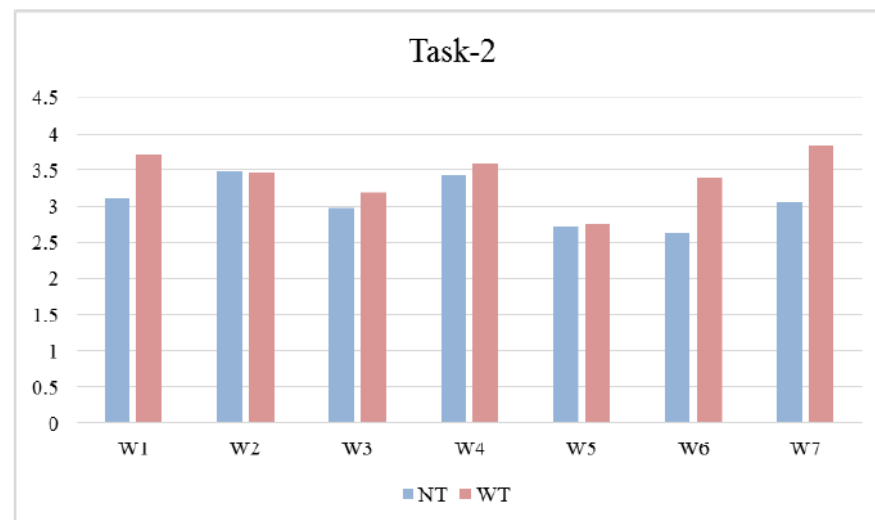


Figure 2. Degree of emotional resonance between writers and readers

Biden and Ryan have a dispute in the first debate of vice president.

Obama slightly leading Romney after the debate

U.S. President Obama leading Republican presidential candidate Romney with 47% to 46% of opinion polling.

Seasoned Vice President Joe Biden versus younger Republican vice presidential candidate Paul Ryan.

President Obama will have three debates with Republican presidential candidate Romney on October.

Vice President Joe Biden versus Republican vice presidential candidate Paul Ryan who promote conservative budget policy.

辯論後 歐巴馬領先羅姆尼

首場總統辯論 拜登、萊恩唇舌劍

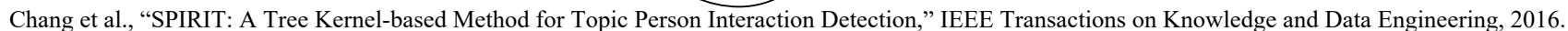
美大選辯論 歐巴馬羅姆尼蓄勢以待

《路透社華盛頓23日電》根據路透和蓋洛普 (Reuters/Gssos) 今天公布的新每日美國民調，美國總統歐巴馬以47%比46%領先共和黨總統候選人羅姆尼。

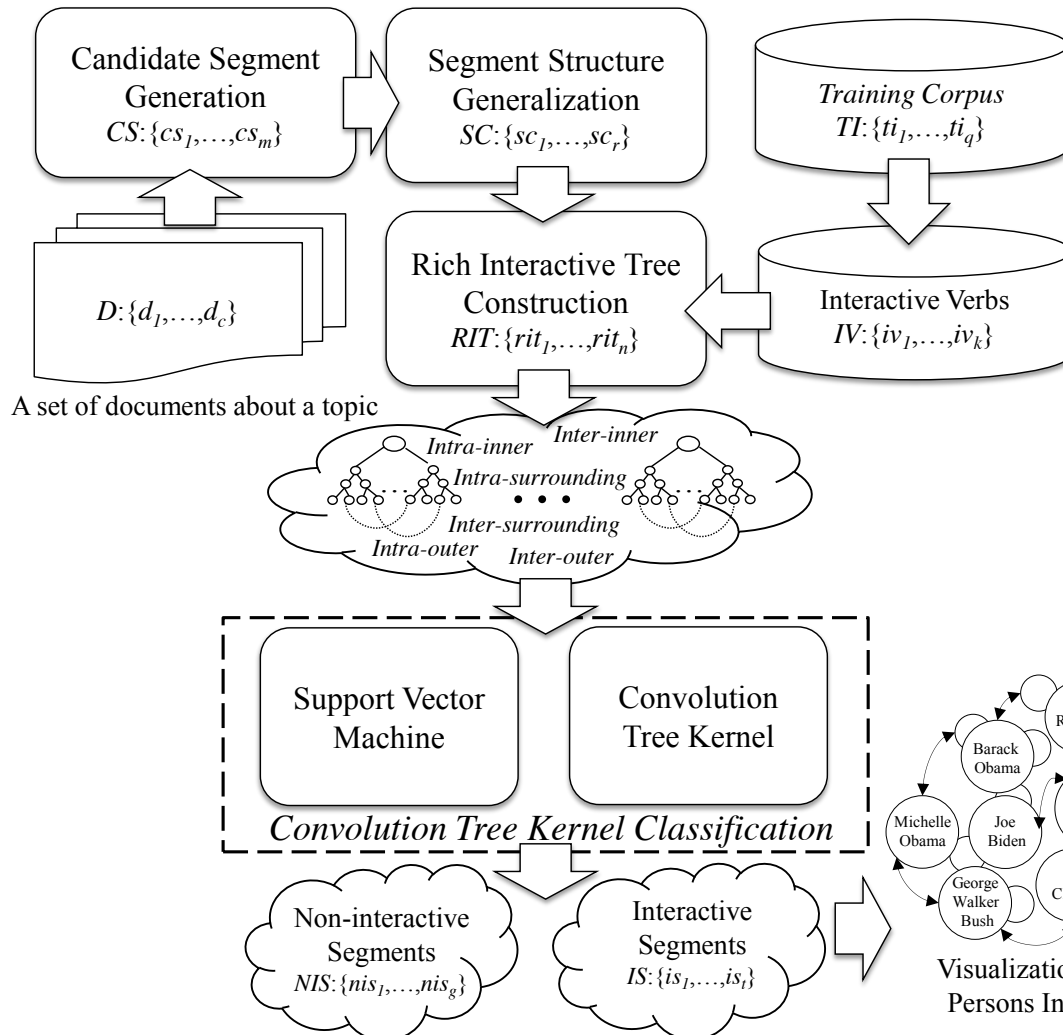
《路透社華盛頓18日電》美國11月8日大選逼近，總統歐巴馬與共和黨對手羅姆尼 (Mitt Romney) 10日，在與美國經濟及領導對手外交政策問題上展開辯論。

歐巴馬辯論也將對上主打保守預算案的人萊恩 (Paul Ryan)。

根據無黨派色彩的總統辯論委員會 (CPB) 主持，辯論時間在東岸時間晚上9時到10時30分。



Tree kernel-based Method for Detecting Interactive Segments



PERSON INTERACTION DETECTION
To detect text segments that may convey person interaction.

PERSON INTERACTION EXTRACTION
To extract keyword for expressing interaction between persons.



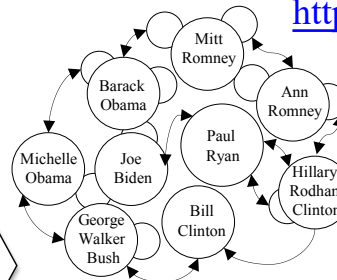
SENTIMENT ANALYSIS
To recognize sentiment behind person interactions.

SOCIAL INTERACTION NETWORK
To construct a network for visualizing social interactions.

Motivation

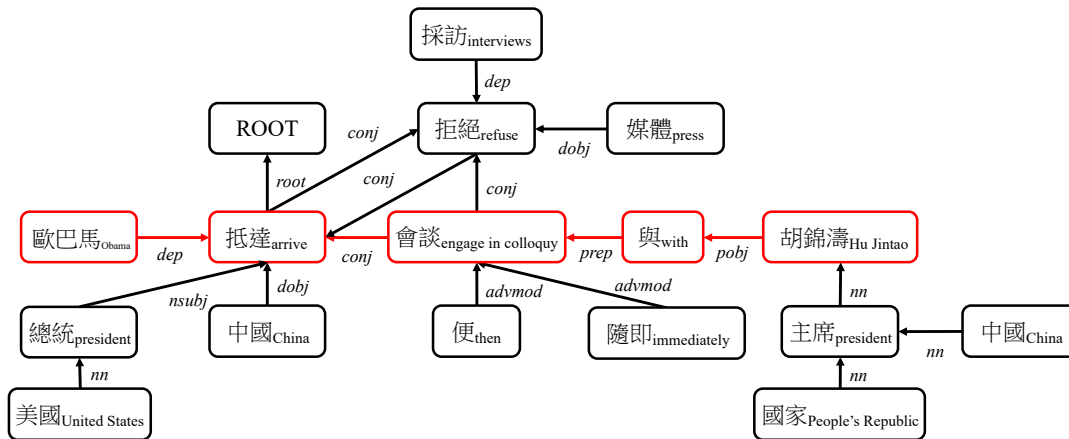
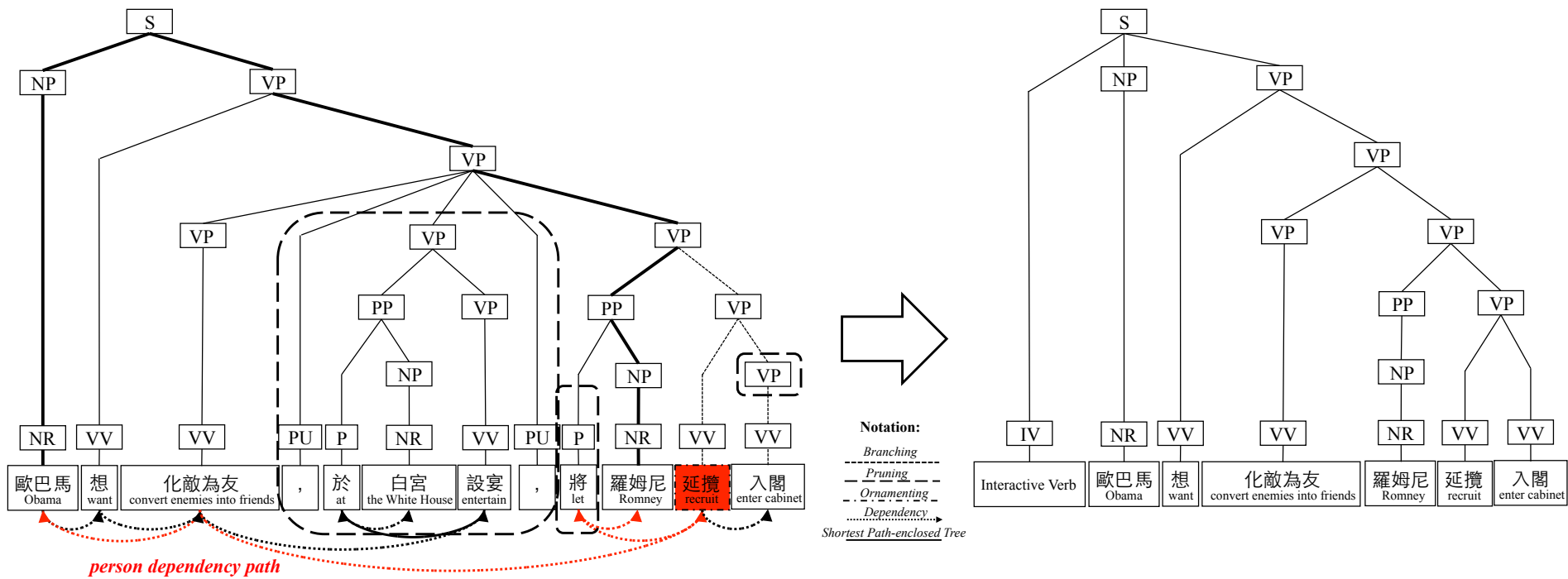
The web has become a powerful medium for disseminating information about diverse topics, such as political issues and sports tournaments. While people can easily find documents that cover various perspectives of a topic, they often have difficulty assimilating the knowledge contained within large documents. This problem has motivated the development of several topic mining methods to help readers digest enormous amounts of topic information. The extracted themes and summaries distill the topic contents clearly; however, readers still need to expend a great deal of time to comprehend the extracted information about unfamiliar topics.

<http://nlp.tmu.edu.tw/interaction/index.html>



Visualization of Topic Persons Interactions

Tree kernel-based Method for Detecting Interactive Segments (cont.)



Multi-Task Deep Neural Network for Person Interaction Discovery

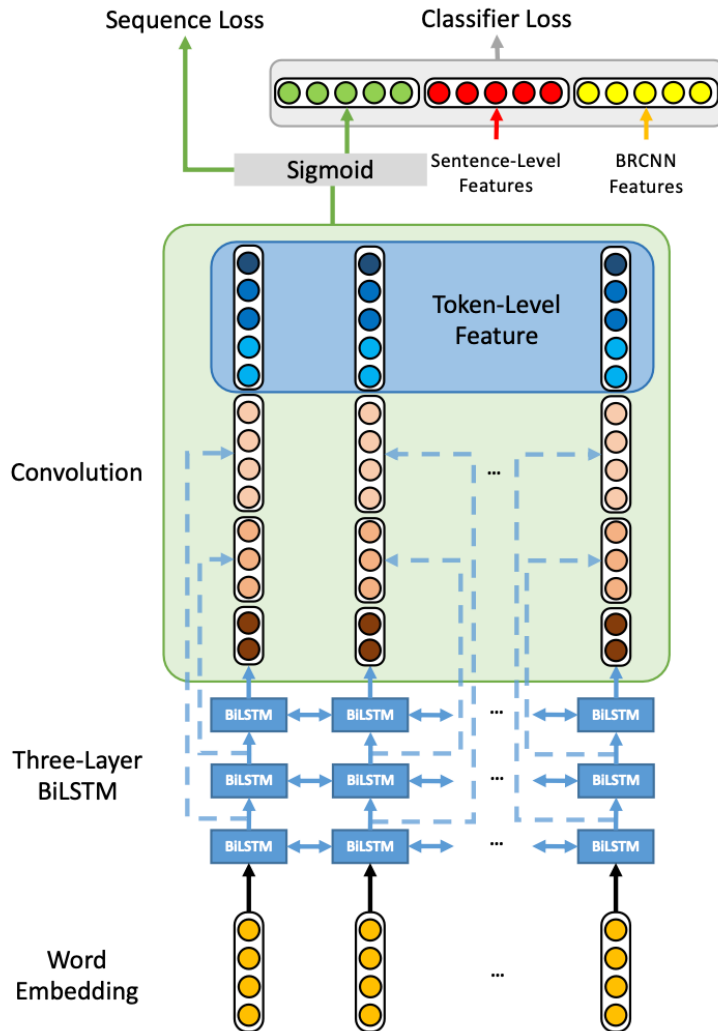


TABLE I. EFFECT OF ADDING VARIOUS SEMANTIC AND SYNTACTIC FEATURES ON THE CLASSIFICATION PERFORMANCE. P AND R REFERS TO PRECISION AND RECALL, RESPECTIVELY.

Method	P/R/F ₁ -score (%)
SPIRIT	68.89/61.56/65.02
Bi-LSTM	66.64/65.28/65.96
MTDNN	67.42/65.66/66.53
+ Entity Embeddings	70.65 /66.55/68.54(+2.01)
+ RIT Embeddings	70.09/68.24/69.15(+0.61)
+ POS Embeddings	69.37/68.99/69.18(+0.03)
+ Dependency Features	69.27/70.13/69.70(+0.52)
+ Threshold	66.36/ 78.88 / 72.08 (+2.38)

TABLE II. THE EXTRACTION PERFORMANCE OF THE PROPOSED MODEL AS COMPARED WITH BASELINE METHODS.

Method	P/R/F ₁ -score (%)
Word-Level CRF	59.40/31.92/40.66
Char-Level CRF	60.30/34.36/42.66
Word-Level Bi-LSTM-CRF	68.68 /50.22/56.20
Char-Level Bi-LSTM-CRF	54.46/33.14/39.99
MTDNN	62.98/ 55.59 / 58.12

Multi-Objective NER POS Annotator

IASL Multi-Objective NER POS Annotator

蔡英文總統今天參加台北市政府舉辦的陽明山馬拉松



Length limit: 50 characters. Longer text will be truncated.

Results:

蔡英文	PER	總統	NA	今天	ND	參加	VC	台北市政府	ORG
舉辦	VC	的	DE	陽明山	LOC	馬拉松	NA		

<http://nlp.tmu.edu.tw:8080/chunk>



罔拍 MONPA: Multi-Objective NER POS Annotator

MONPA 罔拍是一個提供正體中文分詞、詞性標註以及命名實體辨識的多任務模型。初期只有使用原始模型 (v0.1) 的網站版本 (<http://monpa.iis.sinica.edu.tw:9000/chunk>)，本計劃將把新版 monpa (v0.2) 包裝成可以 pip install 的 python package。(提醒：因網站版為 v0.1，與 python 套件版 v0.2 以上的分詞結果可能不同。)

最新版的 monpa model 是使用 pytorch 1.0 框架訓練出來的模型，所以在使用本版本前，請先安裝 torch 1.* 以上版本才能正常使用 monpa 套件。

注意：

1. 建議以原文輸入 monpa 完成切詞後，再視需求濾掉停留字 (stopword) 及標點符號 (punctuation)。
2. 每次輸入予 monpa 做切詞的原文超過 140 字元的部分將被截斷丟失，建議先完成合適斷句後再應用 monpa 切詞。可參考 wiki [如何將長文切成短句再用 monpa 切詞?](#)
3. 支援 python >= 3.6，不支援 python 2.x。

安裝 monpa 套件

monpa 已經支援直接使用 pip 指令安裝，各作業系統的安裝步驟都相同。

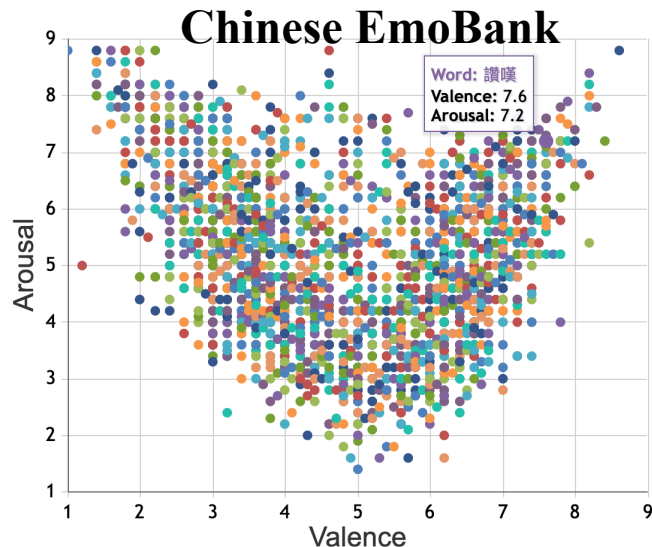
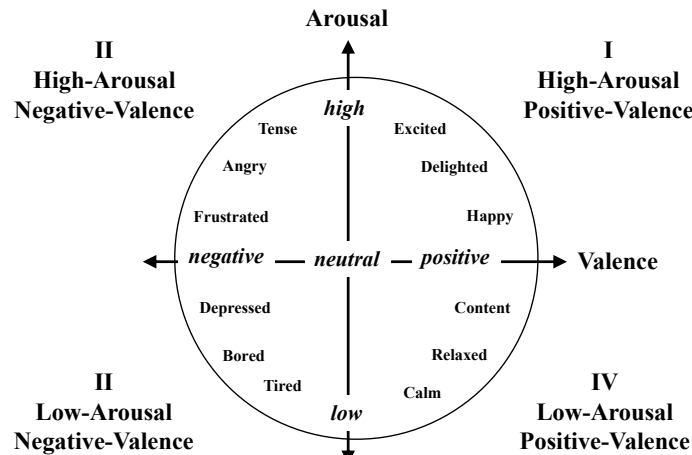
```
pip install monpa
```

安裝時將自動檢查有無 torch >= 1.0 及 requests 等套件，若無則由 pip 直接安裝。



<https://github.com/monpa-team/monpa>

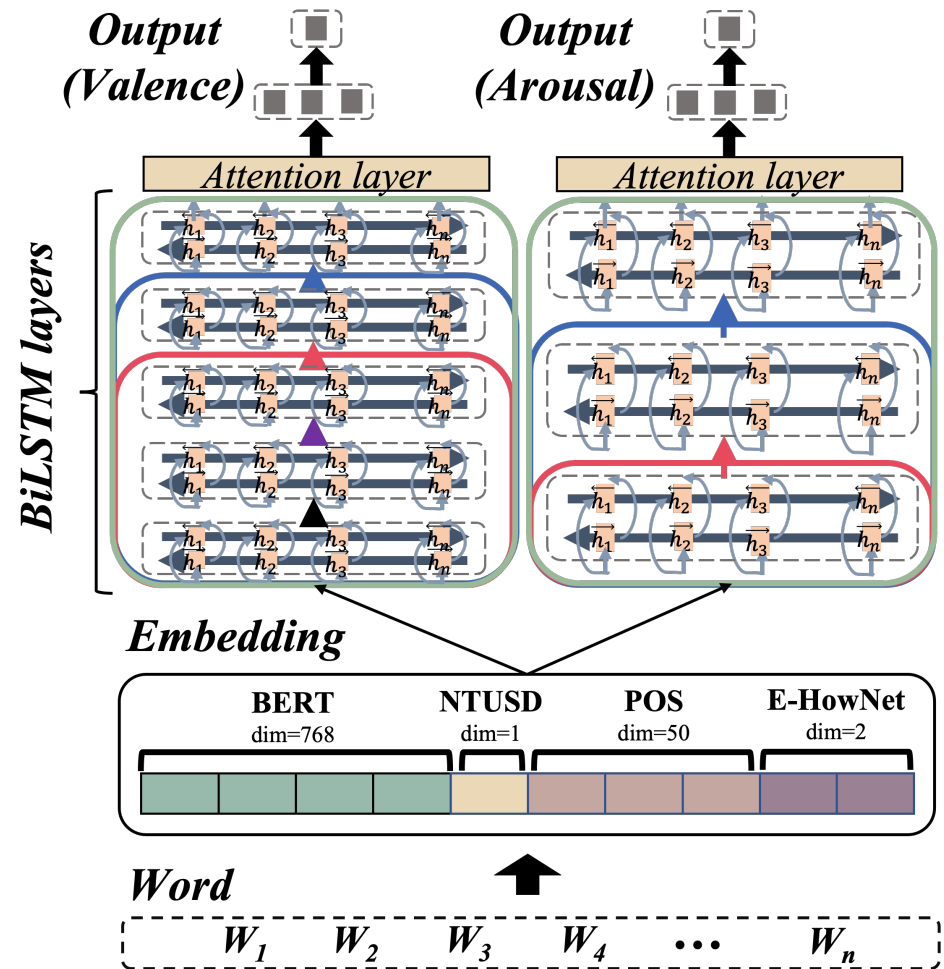
Valence-arousal Prediction for Fine-grained Sentiment Analysis



CanvasJS.com

CanvasJS Trial

<http://nlp.innobic.yzu.edu.tw/resources/ChineseEmoBank.html>



MONPA + Valence Arousal

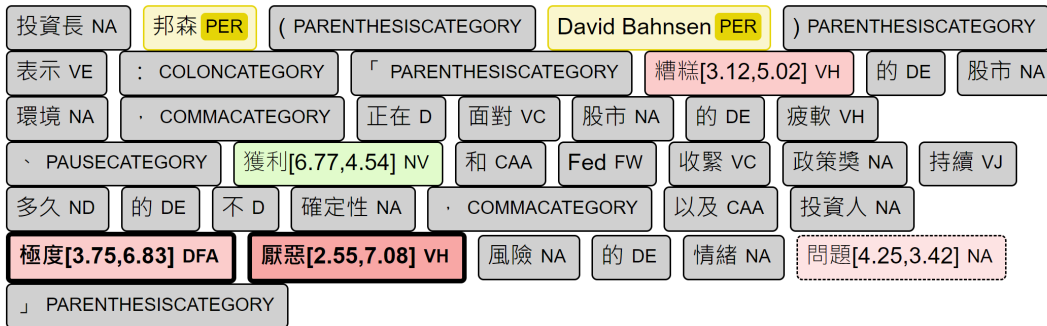
Valence & Arousal

投資長邦森(David Bahnsen)表示：「糟糕的股市環境，正在面對股市的疲軟、獲利和Fed收緊政策獎持續多久的不確定性，以及投資人極度厭惡風險的情緒問題」



Length limit: 200 characters. Longer text will be truncated.

Results:



<http://nlp.tmu.edu.tw:8080/VA>



民視新聞網 | 19.3k 人追蹤 ☆ 追蹤

陳時中出席醫檢師公會餐敘 感謝防疫貢獻



民視新聞網

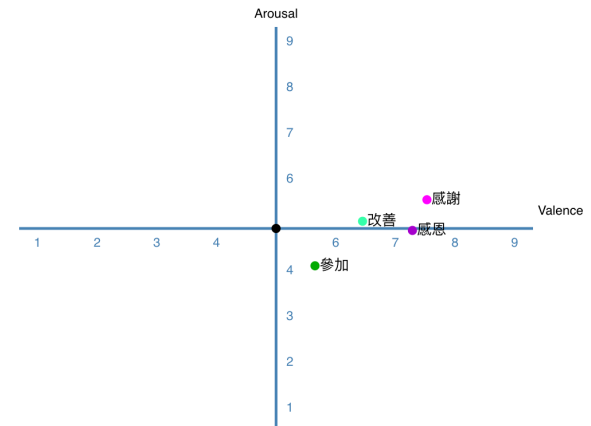
2022年11月11日 週五 下午8:56



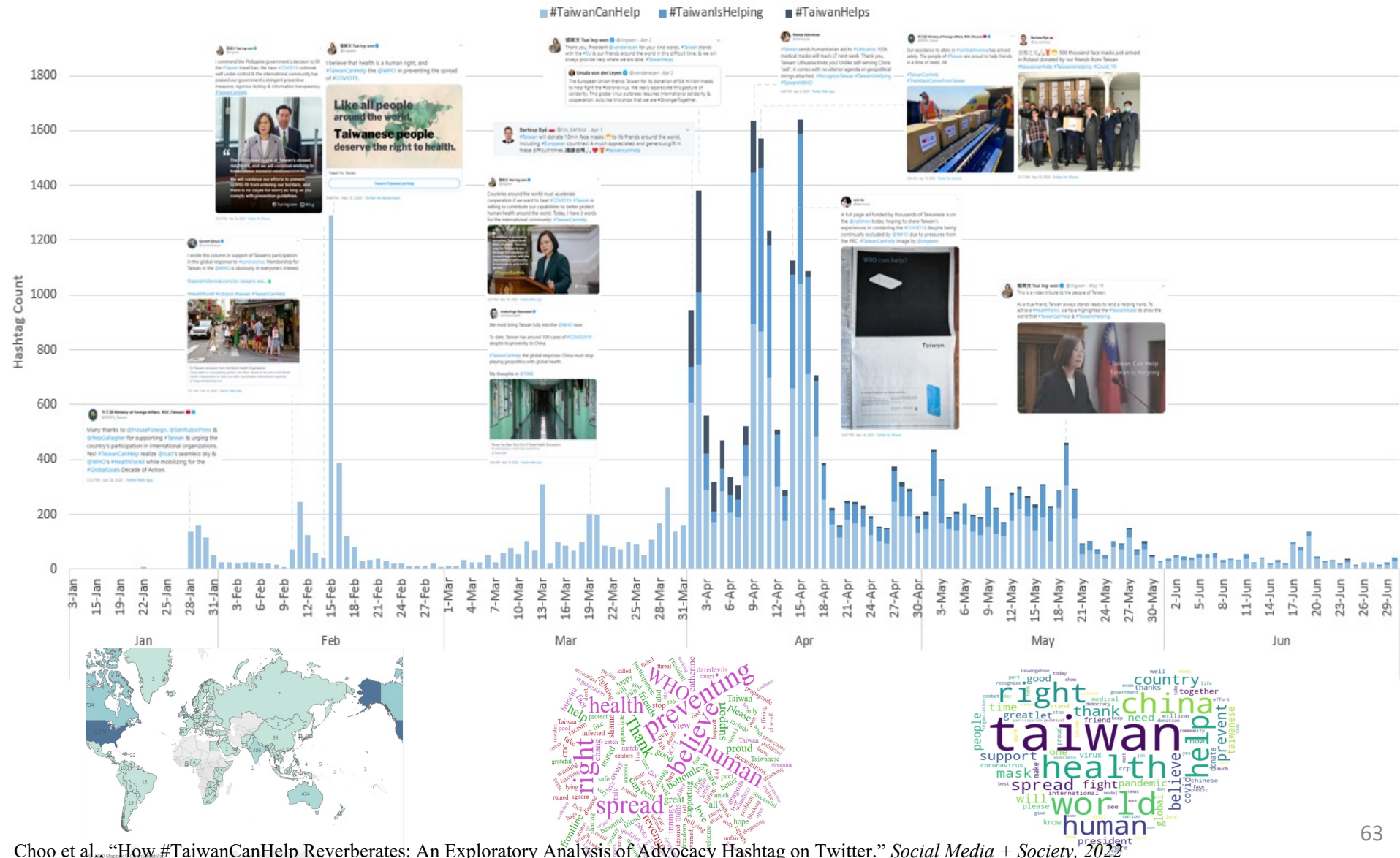
即時中心 / 劉芳好報導



九合一選戰倒數15天，民進黨台北市長候選人陳時中今(11)日晚間和總統蔡英文等人參加台北市醫檢師公會「111年度防疫陣線感恩有您」會員聯誼餐敘。陳時中致詞時細數疫情期間醫檢師的貢獻，他表達對醫檢師的感謝，也提到自己堅持住PCR檢驗費用，是因為台灣的醫檢高品質。陳時中表示，之前與公會理事長有討論過醫檢師的編制、夜班費等等，若未來能入主台北市，將從聯合醫院體系改善開始，讓大家有更好的工作環境。



Social media analytics: #Taiwancanhelp



從新聞找出問題焦點人物-預防經融犯罪

鯉魚潭風景區開發弊案 內政部前都計委員李威儀確定判刑12年

〔記者張文川／台北報導〕12年前爆發的花蓮縣壽豐鄉鯉魚潭風景區的「綠湖國際大飯店」開發弊案，業者王桂霜為讓計畫案順利進行，以250萬元行賄時任內政部都市計畫委員會委員李威儀與其學生藍秀琪；經法院12年審理，2017年12月花蓮高分院更三審判李威儀12年徒刑、藍秀琪9年，皆褫奪公權6年，王桂霜1年9月、褫奪公權1年；最高法院日前駁回上訴，全案全部定讞。



Numerical Understanding in Financial Social Media

- **Task1:** 分類文章中標記的數字與標記\$的字詞是否有相關

Not related (0) directly related (1)

Remember 2009? \$BAC was at 3 a share. The people of Greece all want \$NBR to open to get back to normal. Tired of banking notes on Notepads

- **Task2:** 分類金融文章中標記的數字類別
- **Task3:** 分類數字是否為Claim (宣稱)

{Category: quantity_absolute, absolute, money, price, relative}

Text1

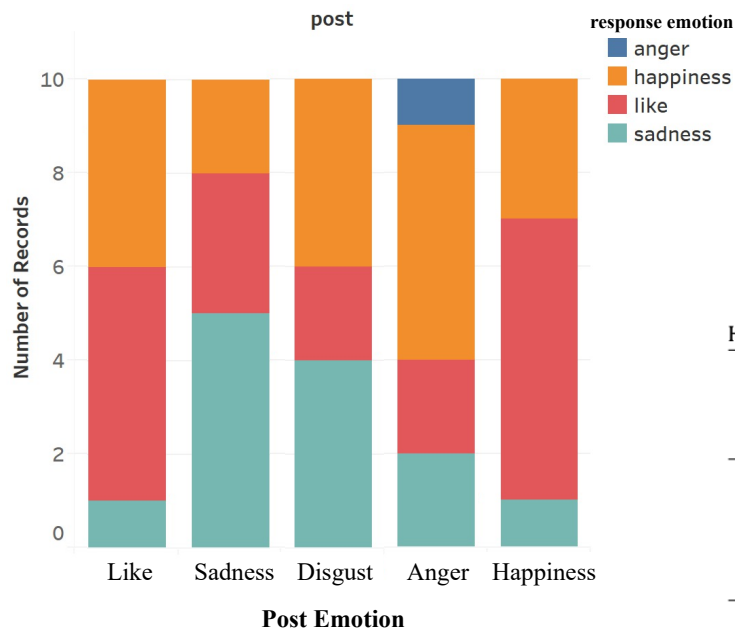
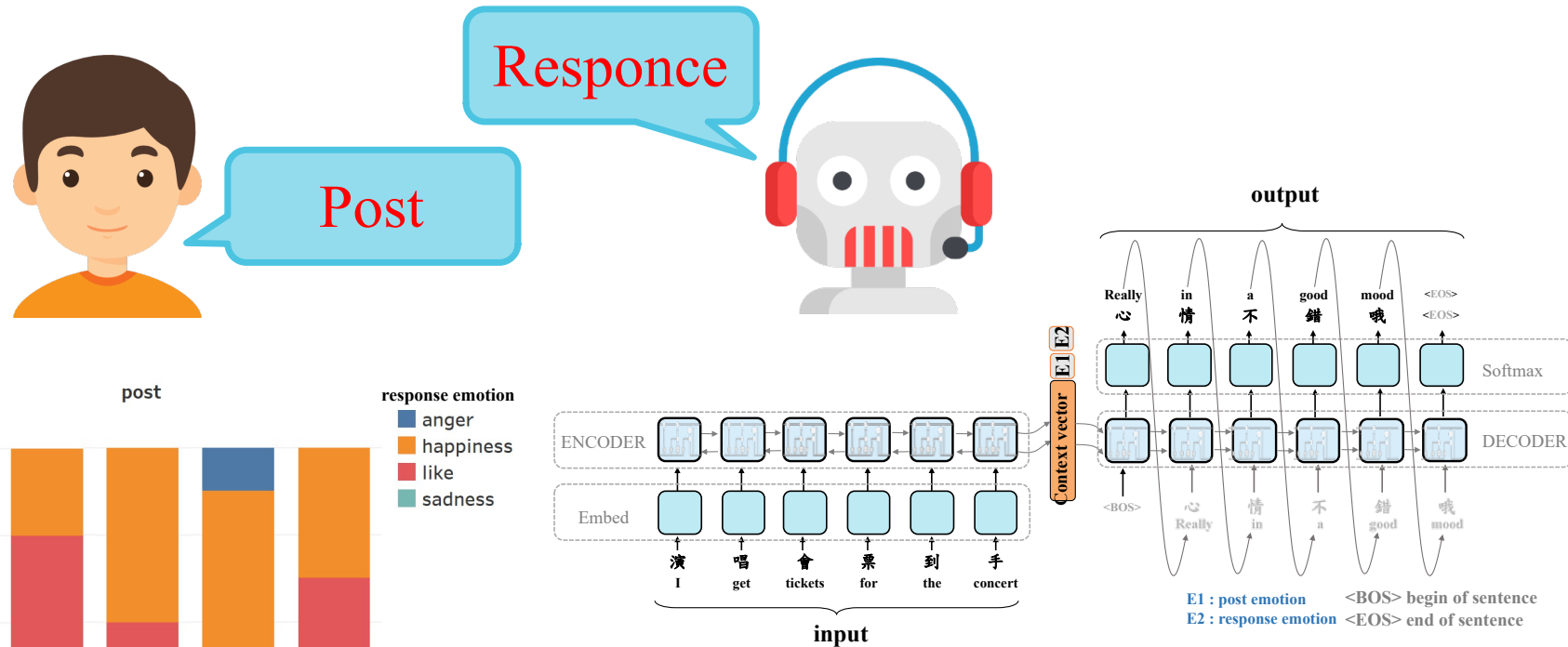
聚陽 4Q18 起，產能供不應求，挑單生產，ASP+3%YoY，營收達 61 億元(+16.5 %YoY)，符合預期，維持 2018 年稅後 EPS 為 7.51 元不變。1Q19ASP 更擴大到 +5%YoY，優於預估的+3%YoY，預估 1Q19 營收可達 64.14 億元(+25%YoY)，毛利率及營利率受惠單價調整較去年同期增加+0.5%，預估稅後淨利達 4.31 億元(+36%YoY)，稅後 EPS 為 2.06 元，2Q19 受惠客戶追加訂單下，預估營收將達 13%YoY。1H19 營運動能超過預期，2H19 需視外包產能取得是否順利，因此略微上修 2019 年營收達 270 億元(+13 %YoY)，毛利率及營利率將持續較去年增加 0.5%，稅後 EPS 為 8.99 元(本估 8.85 元)。

Category: Money Category: Relative
Claim: In-Claim Claim: Out-of-Claim

Business Intelligence

以自然語言處理來挖掘商業中的智慧

Emotion-infused deep neural network for emotionally resonant conversation



Human assessment on the 5-point Likert Scale for the quality of emotion-specific response generation.

Post emotion	Response emotion				
	Like	Sadness	Disgust	Anger	Happiness
	MatchScore/FluencyScore/Average (1 to 5)				
Like	4.40/4.02/4.21	3.06/3.56/3.31	2.36/3.24/2.80	2.28/2.80/2.54	4.14/4.16/4.15
Sadness	4.40/3.70/4.03	4.26/3.90/4.08	2.34/3.88/3.11	2.32/3.78/3.05	4.18/3.16/3.67
Disgust	4.10/3.10/3.58	3.84/4.32/4.08	2.54/3.16/2.85	2.48/2.98/2.73	3.84/3.50/3.67
Anger	3.60/3.30/3.46	3.64/3.84/3.74	2.78/3.56/3.17	2.74/3.40/3.07	4.32/3.50/3.91
Happiness	4.60/4.34/4.46	3.38/3.98/3.68	2.50/3.28/2.89	2.24/3.40/2.82	4.32/3.84/4.08
Average	4.22/3.69/3.95	3.64/ 3.92 /3.78	2.50/3.42/2.96	2.41/3.27/2.84	4.16/3.63/3.90



Reviewed April 27, 2013

What a nightmare! When I reported the situation to the managing stuff, there were no emergency to handle the situation because the decision maker was not available during the week end and it was a Saturday. Instead they offered to transfer my son and I to a hotel nearby where a room was available because they were concerned I was going to cause Mayhem

They finally offered to investigate the room despite the 4 BED BUGS I had contained in a glass and pictures and videos I had showed them. I was supposed to stay one more night but instead chose to move to a hotel nearby; turned out to be cleaner-up to date-bigger room- and cheaper rate and that was the Holiday Inn Express down the road at 3145 Avenue de Hotels.

Beware of BED BUGS! If you are looking for a scratch free night sleep, stay elsewhere, you will be doing you and your loved ones a favour! Trust me...and that's why the Internet is a great tool!

Stayed April 2013, traveled with family

●●○○○ Value	●○○○○ Rooms
●●●○○ Location	●○○○○ Cleanliness
●○○○○ Sleep Quality	●●○○○ Service

Review of [The Muse Hotel New York](#)

Our recent stay at the wise in New York City was superb and our time at the in-house bar prior to our show was the basic King room, but was most comfortable, super clean, nicely decorated, and covered all needs. We took advantage of the parking garage, which made our arrival and departure very easy. Upon checking out the staff saw that I had not used my Karma Club "raid the bar" credit, so they offered to credit it towards my bar charge. Recommend joining the Karma Club prior to booking to take advantage of free internet access as well. All staff members were professional and pleasant. We will most definitely stay here again. Thanks to all!

Room Tip: All room were charming and quaint. Ask hotel staff for advice on your personal preferences.

Green Features: Nice atmosphere and perks. The hotel is environmentally friendly and gives each guest the option to reuse linens and towels or not.

 Sleep Quality
 Service

Reviewed 1 week ago

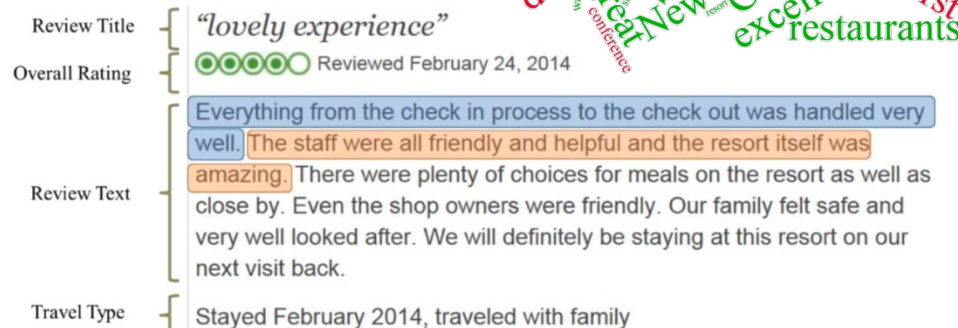
From check in to check out, most of the staff here were incredibly friendly and very helpful. Alvin, the concierge, took time and walked us through many options to get to Bondi Beach - we are extremely appreciative of him! Breakfast in the morning is delicious with a vast amount of choices.

The executive lounge was not only disappointing, but frustrating. The "rules of entry" seem to change by the hour and by the person sitting at the front desk. Certain attire is required during certain times and you are not allowed to take any food item or drink item out of the lounge and to your room for consumption. The staff downstairs, on the "magic" line, and elsewhere in the hotel are great - but avoid the staff in the executive lounge.

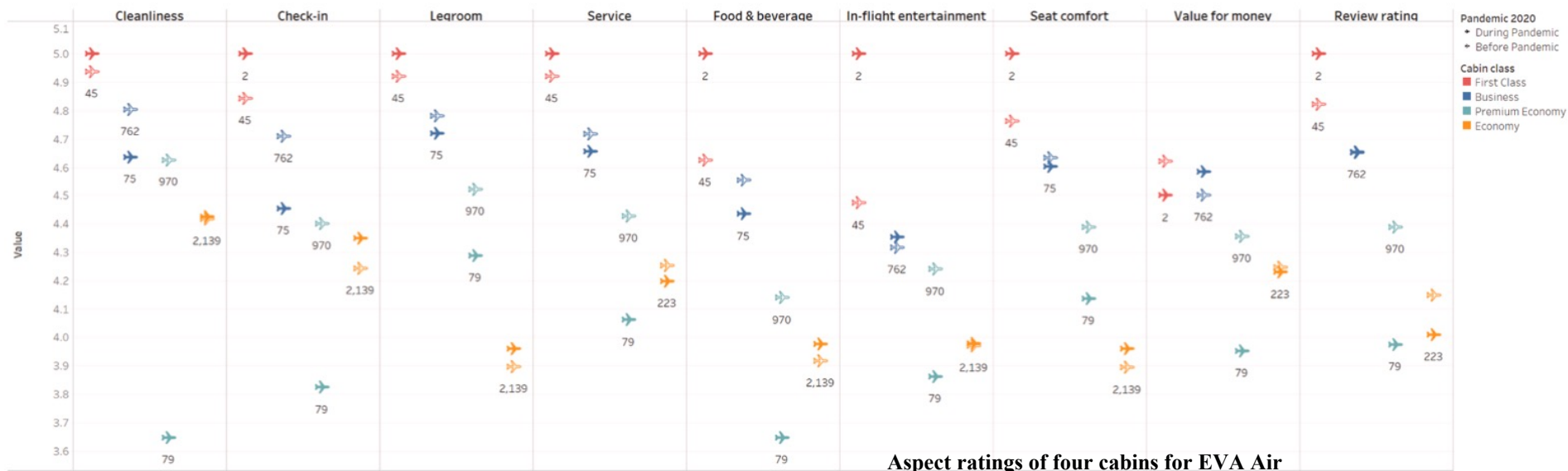
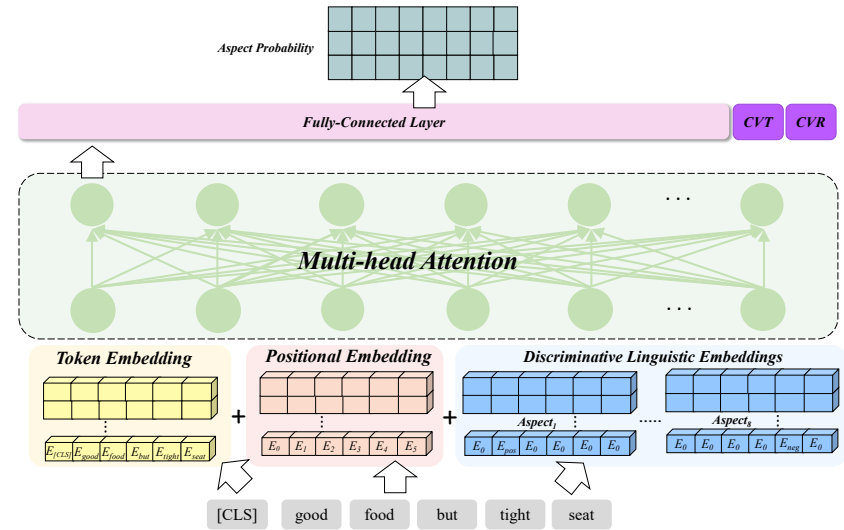
Stayed April 2017, traveled as a couple

 Sleep Quality
 Rooms

 Service
 Service

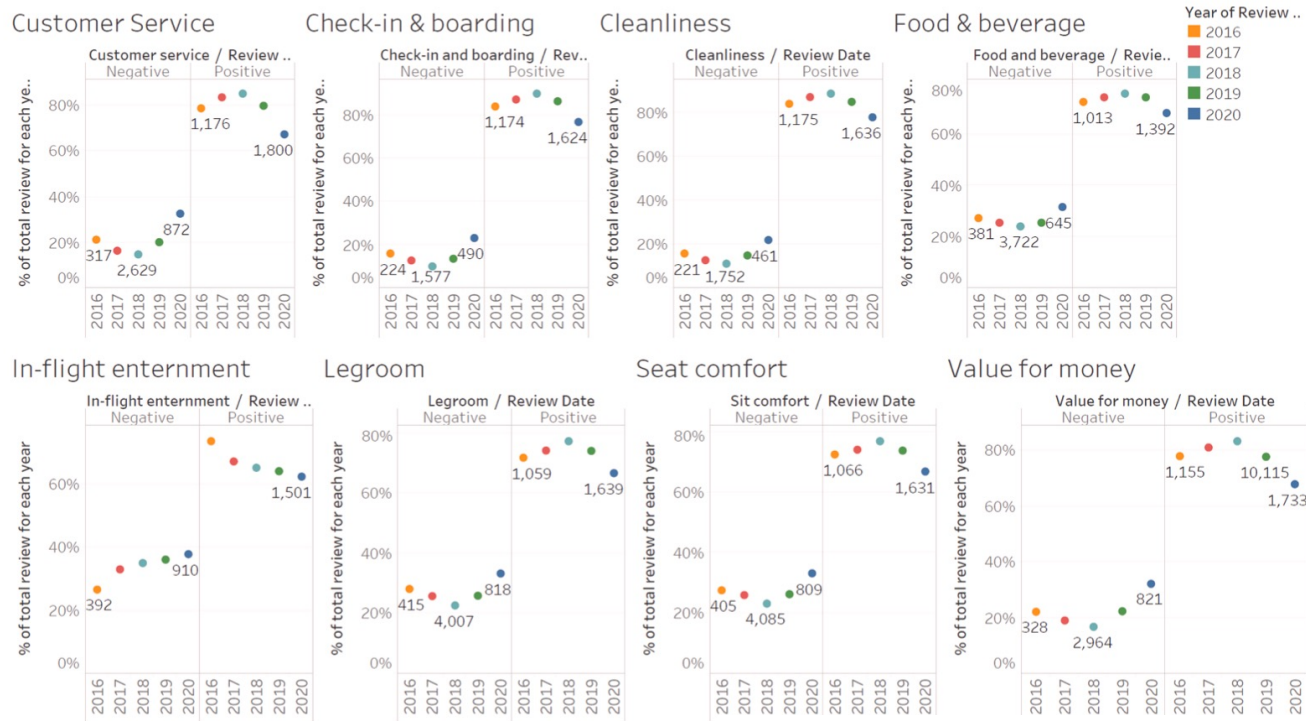


Social media analytics for Airline Industry



Aspect ratings of four cabins for EVA Air

Social media analytics for Airline Industry (cont.)



Negative sentiments word cloud



Negative sentiments word cloud

BioNLP and ClinicalNLP

生醫與臨床上的自然語言處理

Bio-entity Interactions Passage Extraction for Biomedical Literature (1/2)

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed 21394427[uid]
Create RSS Create alert Advanced

Format: Abstract ▾

Clin Rev Allergy Immunol. 2013 Feb;44(1):31-8. doi: 10.1007/s12016-011-8265-z.

The role of inflammation and autoimmunity in the pathophysiology of pulmonary arterial hypertension.

Kherbeck N¹, Tamby MC, Bussone G, Dib H, Perros F, Humbert M, Mouthon L.

⊕ Author information

Abstract

Pulmonary arterial hypertension is characterized by a remodeling of pulmonary arteries with endothelial cell, fibroblast, and vascular smooth muscle cell activation and proliferation. Since pulmonary arterial hypertension occurs frequently in autoimmune conditions such as systemic sclerosis, inflammation and autoimmunity have been suspected to play a critical role in both idiopathic pulmonary arterial hypertension and systemic sclerosis-associated pulmonary arterial hypertension. High levels of pro-inflammatory cytokines such as interleukin-1 and interleukin-6, platelet-derived growth factor, or macrophage inflammatory protein 1 have been found in lung samples of patients with pulmonary arterial hypertension, along with inflammatory cell infiltrates mainly composed of macrophages and dendritic cells, T and B lymphocytes. In addition, circulating autoantibodies are found in the peripheral blood of patients. Thus, autoimmunity and inflammation probably play a role in the development of pulmonary arterial hypertension. In this setting, it would be important to set-up new experimental models of pulmonary arterial hypertension, in order to define novel therapeutics that specifically target immune disturbances in this devastating condition.

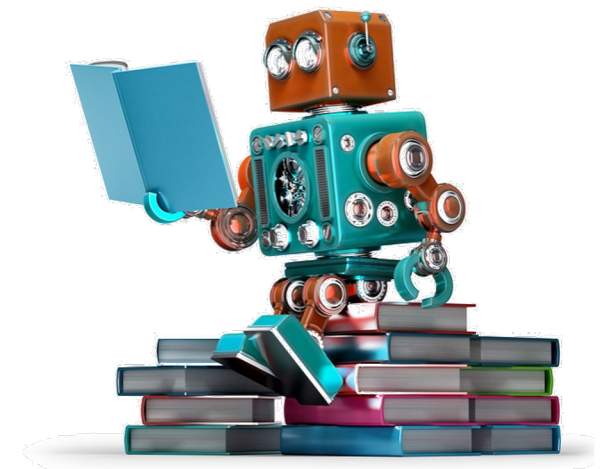
PMID: 21394427 DOI: 10.1007/s12016-011-8265-z

[Indexed for MEDLINE]

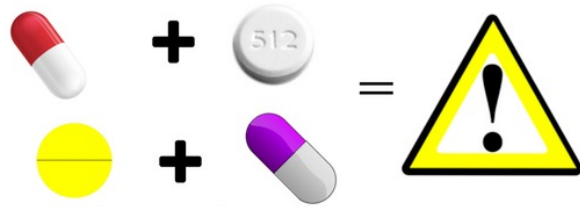
- 
- 1 Steroid mediated lysis of lymphoblasts requires the DNA binding region of the steroid hormone receptor .
 - 2 Glucocorticoids kill certain types of lymphoblasts , but the mechanisms are unknown .
 - 3 It is clear that sufficient numbers of functional glucocorticoid receptors are required to mediate lysis , but whether they do so through the classical model of steroid hormone activation and modulation of gene expression has not been established .
 - 4 In this report we have asked which region (s) of the steroid receptor are important for mediating lysis in leukemic T lymphoblasts .
 - 5 CEM-ICR 27 leukemic lymphoblasts , a clone of CEM cells which lack functional glucocorticoid receptors and therefore are neither lysed by dexamethasone nor capable of showing glutamine synthetase induction , were provided with steroid receptors by DNA transfections of various receptor gene constructs .
 - 6 We measured steroid mediated lysis , receptor number and induction of glutamine synthetase in the transfected cells .
 - 7 Our results provide evidence that the lysis mechanism in the ICR27 lymphoblasts is restored when functional receptor number is restored .
 - 8 The DNA binding region specifying high affinity for GRE sites is required .
 - 9 Lysis is mediated by any steroid that allows for activation of the receptor containing such a region .
 - 10 Our data support the view that steroid-mediated cell death occurs by a process requiring direct interaction of steroid-receptor complexes with the genome .



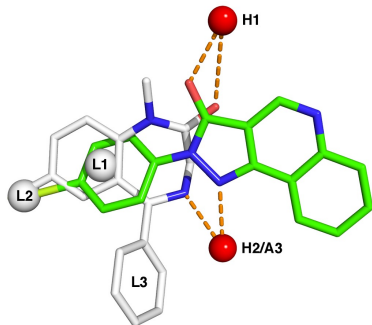
Highlight key information & link them into official database



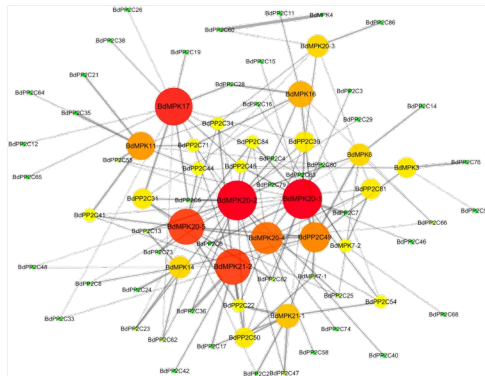
Bio-entity Interaction Passage Extraction from Biomedical Literature (2/2)



Drug-drug Interaction Recognition

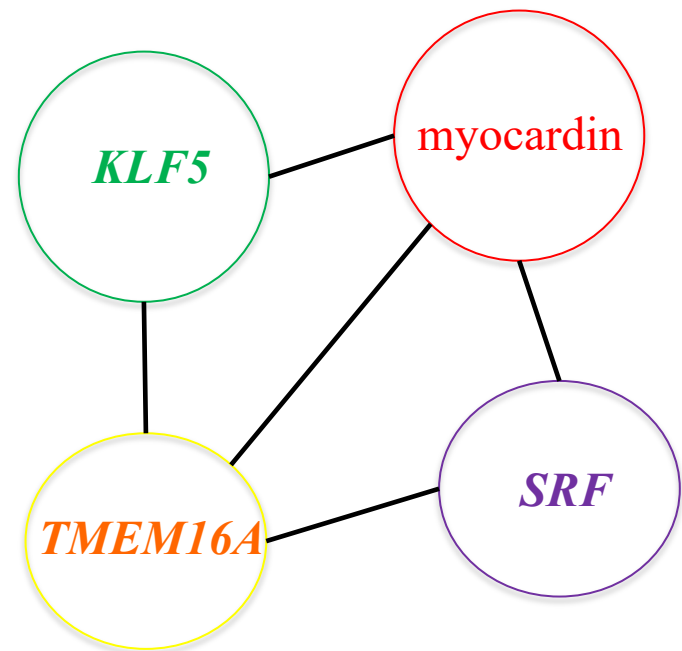


Chemical-disease Relation Identification

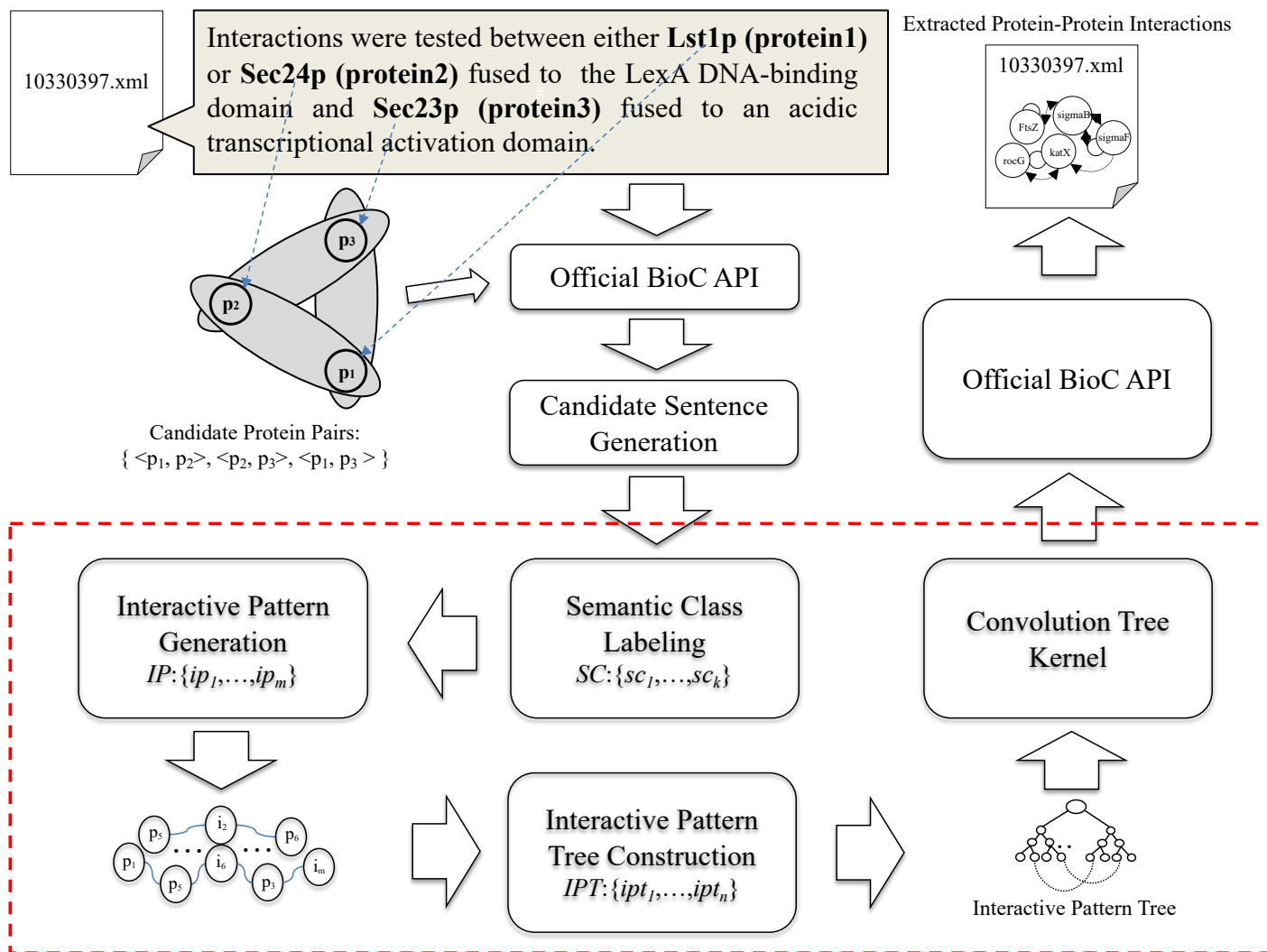


Protein-protein Interaction detection

KLF5 competes with *SRF* to interact with *myocardin*, thereby limiting *myocardin* binding to *SRF* and the synergistic activation of the *TMEM16A* promoter by *myocardin* and *SRF*.



My first BioCreative competition (2015)



Chang et al., "PIPE: A BioC Module for Protein-protein Interaction Passage Extraction," Database, 2016

Chang et al., "Linguistic Template Extraction for Recognizing Reader-Emotion and Emotional Resonance Writing Assistance," ACL 2015.

LBERT for Bio-entity Relations Extraction

Lee et al., BioBERT: a pre-trained biomedical language representation model for biomedical text mining, Bioinformatics, 2019. (Citations: 3382)

- In this study, we propose a universal BRE model, i.e. LBERT, which is a Lexically-aware Transformer-based Bidirectional Encoder Representation model, and which explores both local and global contexts representations for sentence level classification tasks.

Table 3. Comparative analysis of BRE type performance between LBERT and SOTA deep learning models for each task

Classification Task	Bio-entity Relation Type	Corpus	Metrics	Deep-Learning/Kernel Models	SOTA	LBERT
Relation Detection Task	PPI	HPDR50	F_1	<i>DeepResCNN/DSTK</i> (Murugesan et al. 2017)	77.7 / 80.0	85.5
		LLL			83.2 / 89.2	86.0
		IEPA			75.5 / 80.2	83.7
		AIMed			77.6 / 71.0	74.0
		BioInfer			86.9 / 76.2	72.8
			$\mu-F_1$		80.8 / 79.3	80.8
Relation Classification Task	GDR	EU-ADR	F_1	<i>BioBERT v1.0 (PMC)/Ensemble SVM</i> (Bhasuran et al. 2018)	86.5 / 85.3	84.2
		GAD			81.6 / 83.9	78.5
			$\mu-F_1$		84.1 / 84.6	81.4
	DDI	MedLine	F_1	<i>EnrichedCNN/ Feature-based Kernel</i> (Kim et al. 2015)	52.1 / 38.2	64.5
					52.1 / 38.2	64.5
			$\mu-F_1$			
Relation Classification Task	REL	BioNLP Shared Task 2011 corpus	F_1	<i>Base-lines</i> CNN LSTM	12.3	
					17.4 / 57.7	58.8
			$\mu-F_1$		17.4 / 57.7	58.8
	CPR	ChemProt	F_1	<i>BioBERT v1.1 (PubMed)/ CPI 3-stage model</i> (Lung et al. 2019)	76.4 / 51.2	59.1
					76.4 / 51.2	59.1
			$\mu-F_1$		76.4 / 51.2	59.1

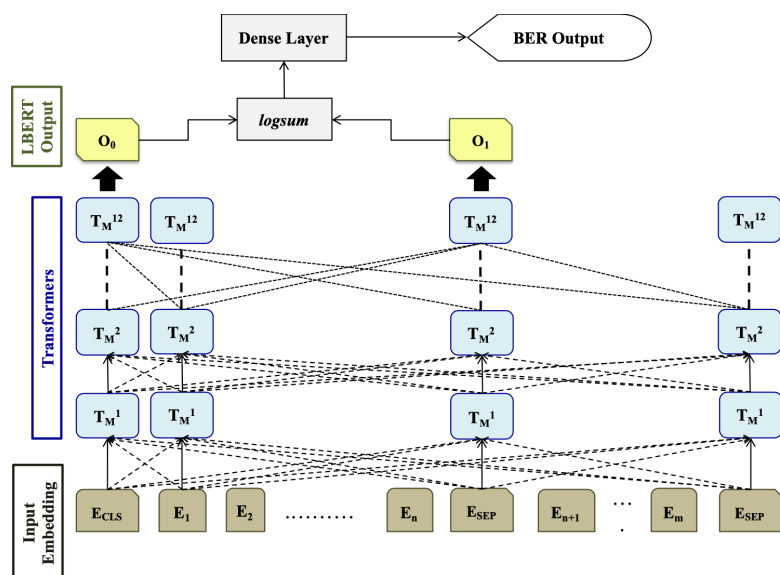


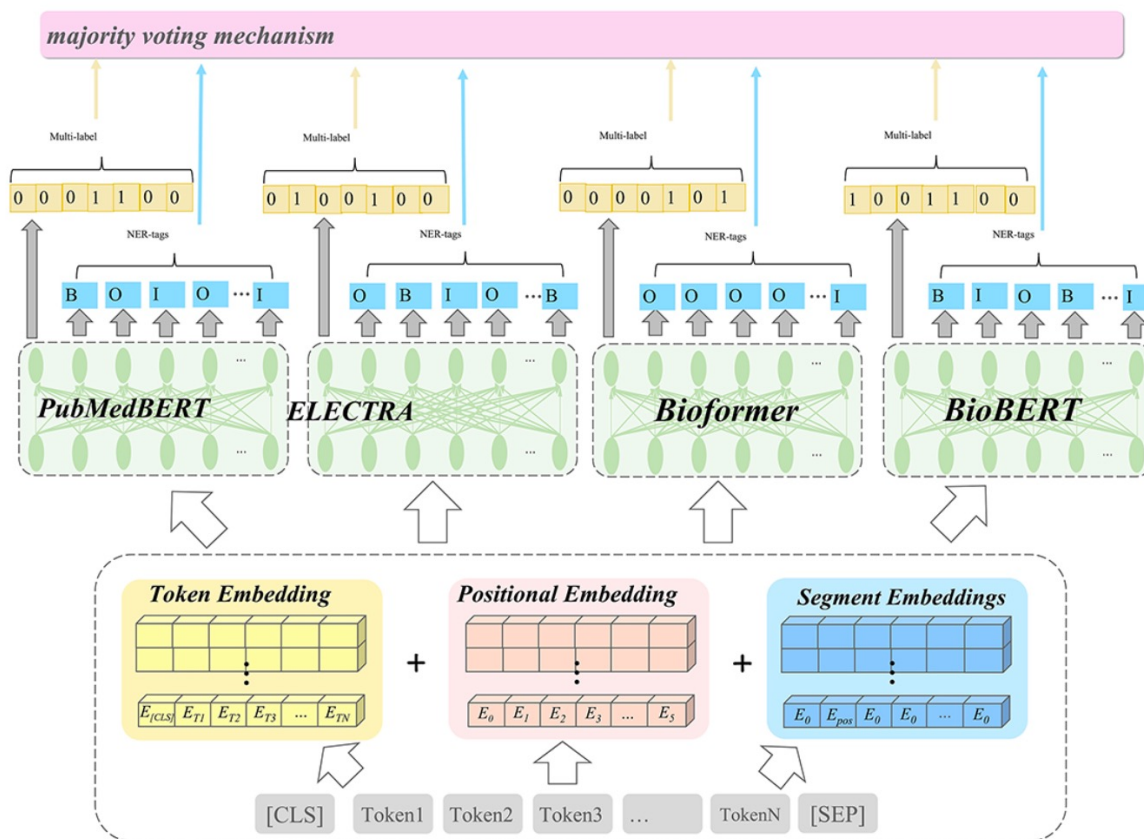
Figure 3. Lexical Bidirectional Encoder Representation (LBERT) system architecture

BioCreative 2017

Warikoo et al., "LPTK: A Linguistic Pattern-Aware Dependency Tree Kernel Approach for the BioCreative VI CHEMPROT Task," Database-The Journal of Biological Databases and Curation, vol. 2018, pp. bay108, 2018.

Warikoo et al., "LBERT: Lexically-aware Transformers based Bidirectional Encoder Representation model for learning Universal Bio-Entity Relations.," Bioinformatics, 2020.

The BioCreative 2021



Pre-trained BERT	Ensembled models for Track	
	NLM-CHEM	LitCovid
biobert-base-cased-v1.2	×	✓
PubMedBERT	✓	✓
ELECTRA-Large-Discriminator	✓	✓
ELECTRA-Large-SQuAD2	✓	✓
bioformer-cased-v1.0	✓	✓
bioformer-cased-v1.0-bc2gm	✓	×

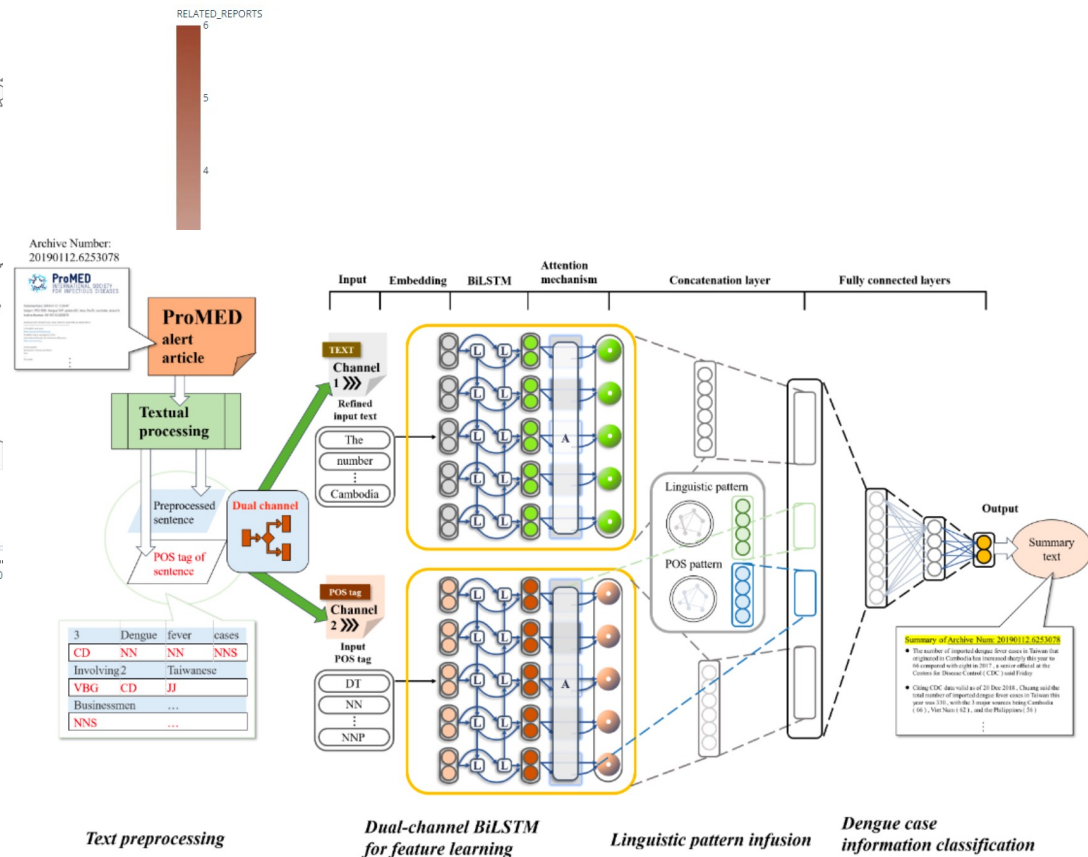
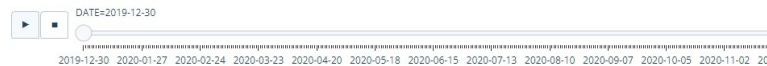
Track2: NLM-Chem Track: Full text Chemical Identification and Indexing in PubMed articles

Track5: LitCovid track Multi-label topic classification for COVID-19 literature annotation

Mining ProMED for Infectious Disease Surveillance



Covid19 Statistics



What role does NLP play in biomedical sciences?



Current status of medical NLP in Taiwan

NIH National Institutes of Health
Office of Intramural Research

Home | Contact Us | Search

About OIR | IRP Website | Sourcebook | WALS | NIH Research Festival | The NIH Catalyst | SIGs | NIDB

Search SIGs

View by Name
View by Scientific Focus Area

Text Mining and Natural Language Processing (NLP) Scientific Interest Group

NIH Text Mining and Natural Language Processing (NLP) Scientific Interest Group

With the increasing availability of text information related to diverse research fields across the NIH Intramural Research Program, the domain of biomedical text mining and Natural Language Processing (NLP) has seen a tremendous growth. Some examples of how researchers across campus utilize NLP are improving literature search in PubMed, automatic entity extraction from scientific articles for scaling up manual curation, etc. Researchers are building knowledge discovery resources for improved literature search and network analysis of scientific literature. Furthermore, text mining techniques are core to computational biology including genomics and other "omics" analysis.

NIH National Library of Medicine
Lister Hill National Center for Biomedical Communications

RESEARCH/R&D | PUBLICATIONS | DOWNLOADS | NEWS/EVENTS | PERSONNEL | ABOUT

RESEARCH/R&D

Research/R&D Areas

Health Information Standards and Discovery

Natural Language Processing

Related Projects

Image Processing

LHNCBC Project List

Natural Language Processing

Lister Hill National Center for Biomedical Communication's (LHNCBC) natural language processing (NLP), or text mining, research focuses on the development and evaluation of computer algorithms for automated text analysis. This area of research works primarily with text from the biomedical literature or electronic medical records and examines a wide variety of NLP tasks, including information extraction, literature searches, question answering, and text summarization.

Related Projects



BabelMeSH and PICO Linguist

BabelMeSH and PICO (Patient, Intervention, Comparison, and Outcome) Linguist are multi-language tools for searching MEDLINE/PubMed. 13 languages, including character-based languages, are supported. Recent enhancements include a query using more than one language and retrieving citations in more than one language.



Consumer Health Question Answering

The consumer health question answering project was launched to support NLM customer services that receive about 90,000 requests a year from a world-wide pool of customers.

PubMed.gov

natural language processing



Search

Advanced Create alert Create RSS

User Guide

Save

Email

Send to

Sorted by: Best match

Display options

RESULTS BY YEAR

8,015 results

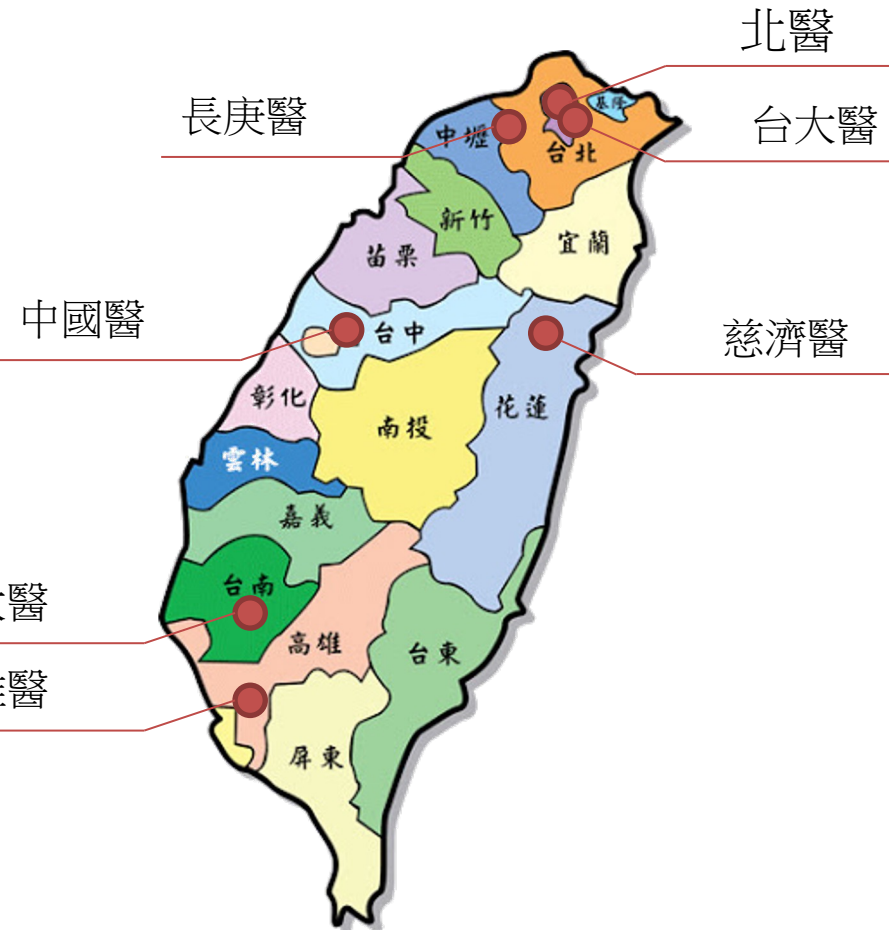


Reset

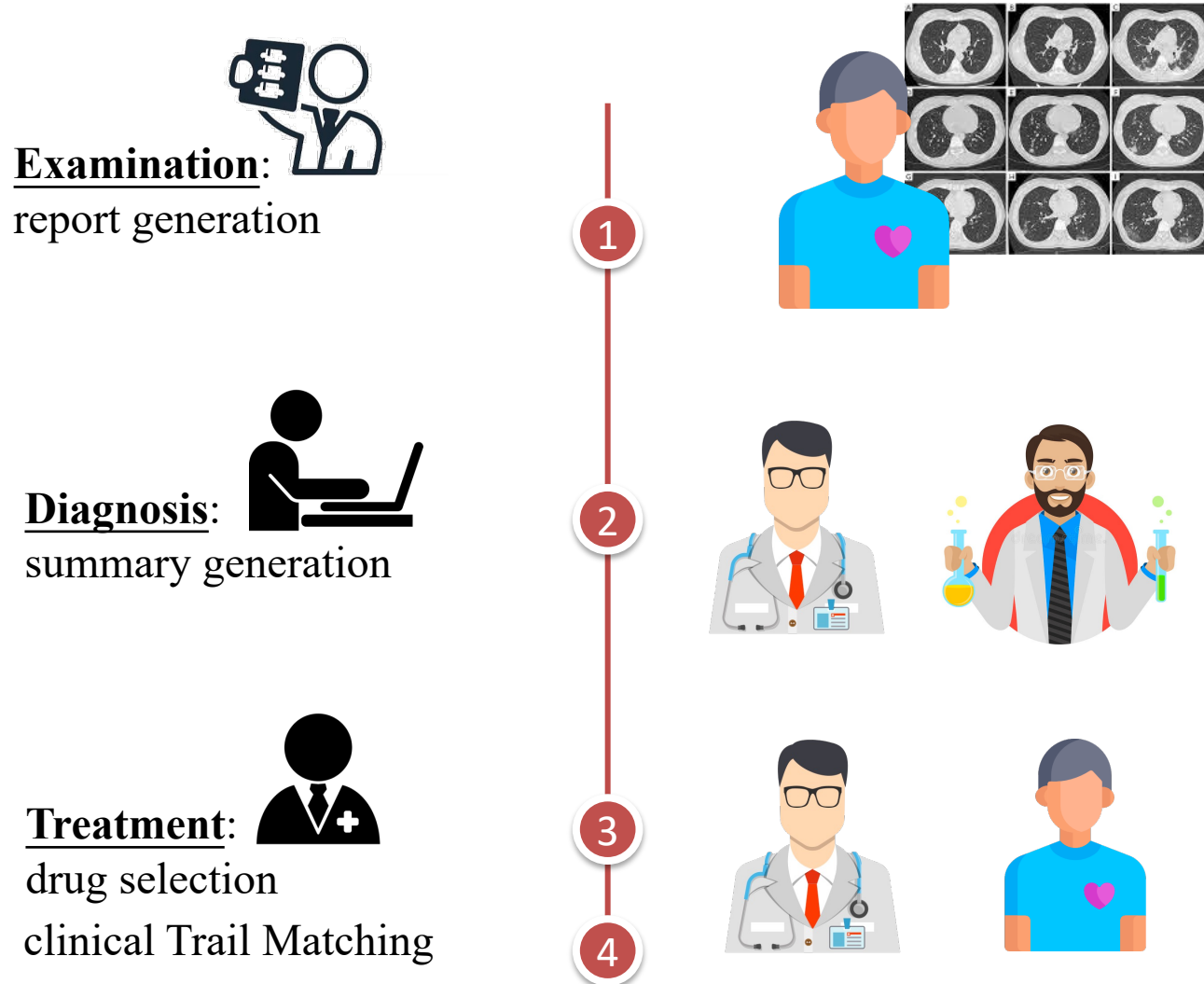
1990



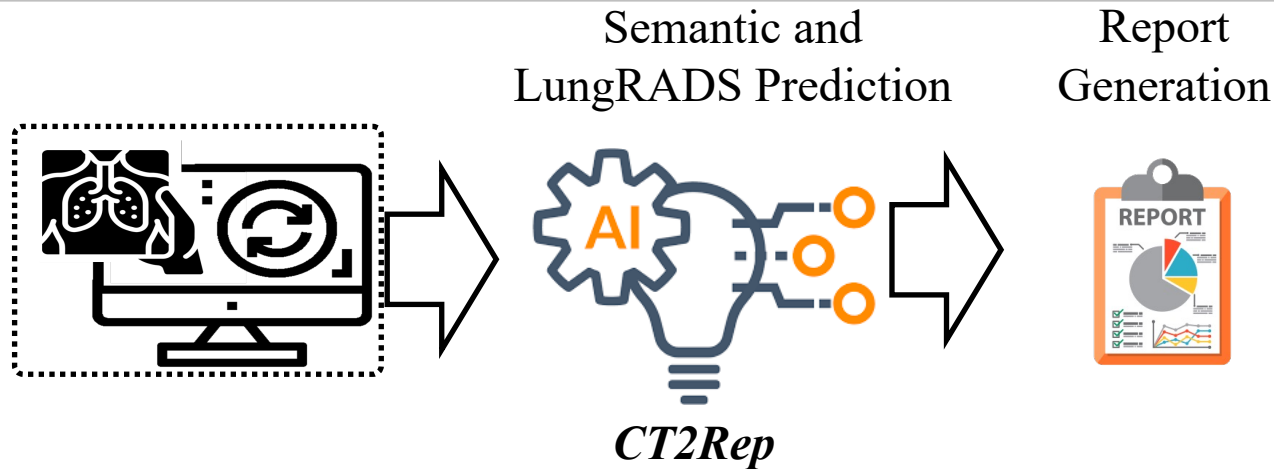
2020: 1,065

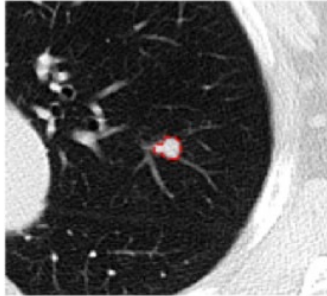
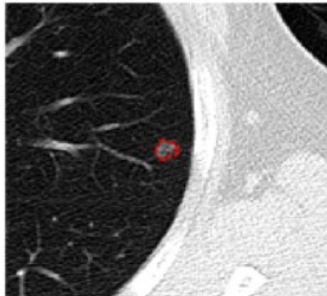


NLP can help the treatment flow



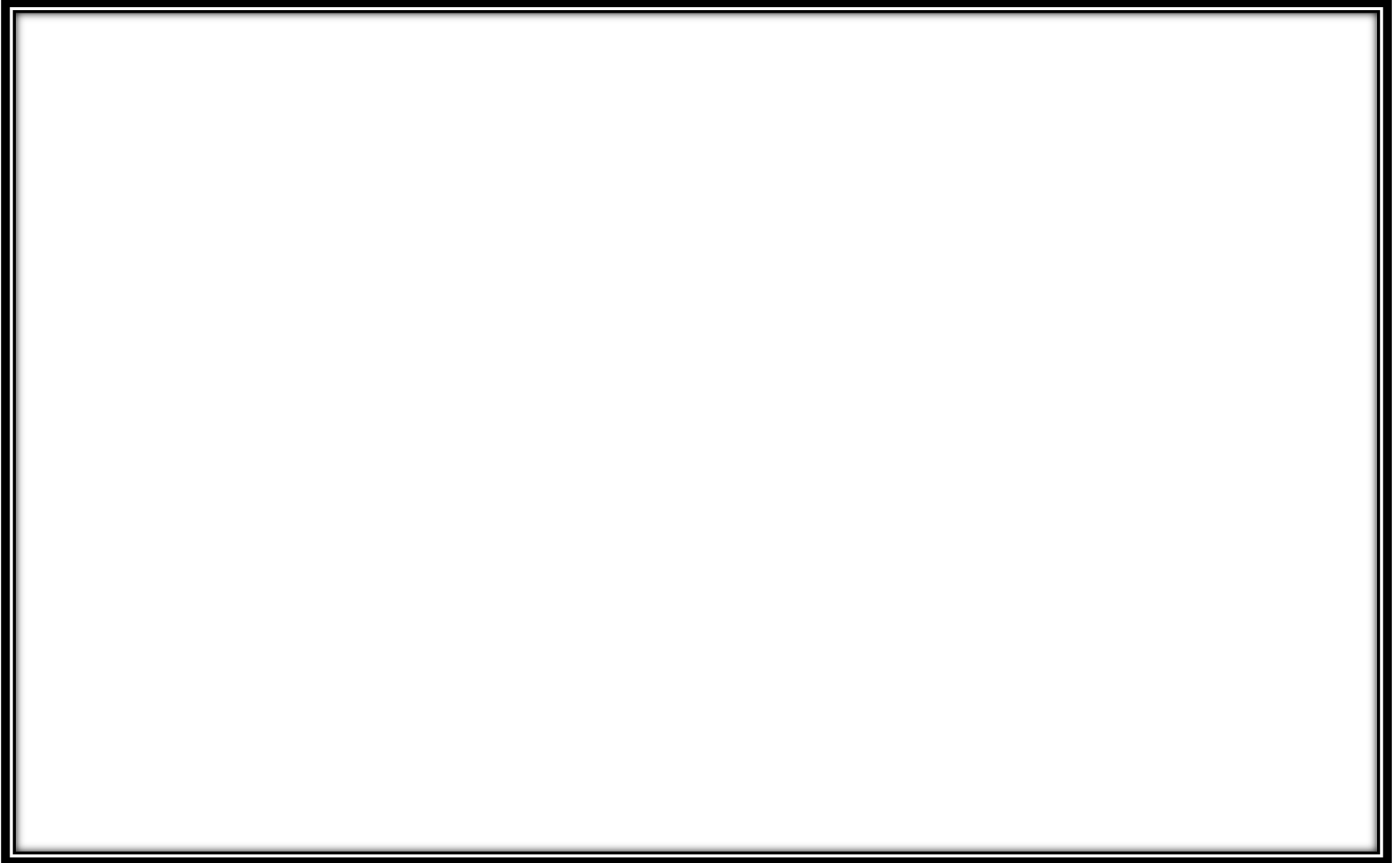
Towards Automatic Radiology Report Generation



	Image Date	RIS ID	Slice Number	Maximum Diameter
	2016/6/18	TMU-27-2585	43/60	8.4 mm
	Nodule 1 impression:			
	1. There is a 8.4 mm solid nodule with lobulated margin in the LUL.			
	2. According to Lung RADS version 1.1, the score is 4A.			
	3. The malignant probability is about 98.20%.			
	4. Suggest closely follow up and further biopsy.			
	Image Date	RIS ID	Slice Number	Maximum Diameter
	2016/6/18	TMU-27-2585	40/60	7.8 mm
	Nodule 2 impression:			
	1. There is a 7.8 mm Pure GGO nodule with lobulated margin in the LUL.			
	2. According to Lung RADS version 1.1, the score is 2.			
	3. The malignant probability is about 24.47%.			

1

CT Report Generation using LLM



2 3 4

Using Generative Deep Learning to Predict the Efficacy Prognosis and Screen Global Clinical Trials for Advanced Lung Cancer



2023未來科技獎

運用生成式深度學習預測晚期肺癌藥效預後和篩選全球臨床試驗

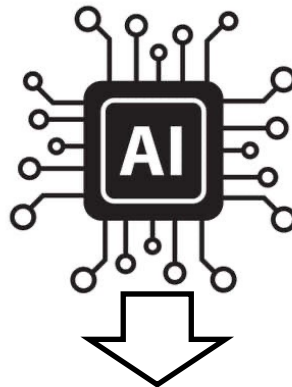
NLP for Clinical Narrative Application

Vital Sign



Clinical Narrative

- 情緒不穩, 平常有在吃抗癲癇的藥, 也有吃精神科的藥, 因為父親管病人打電動, Mood fluctuation, 要拿刀殺爸爸, 潑沙拉油
- fever tonight, today chemotherapy, 現在還在滴藥
- 這幾天沒有外傷, RIGHT MULTIPLE RIBS FRACTURE S/P OP, 昨天還有回胸外門診, NO FEVER, NO DIARRHEA, 現在肚子比較不痛了, --> 痛到喘不過氣, RUQ AND EPIGASTRIC ABDOMINAL PAIN + VOMITING



Critical Outcome Prediction

Patients admitted to ICU within 72 hours

Patients revisiting ED within 48 hours



國朝李時珍先生原本
內附萬方輟錄
奇能八味
重訂
子芥錄
蘇郡懷學
張雲中重訂
張青萬全參
園

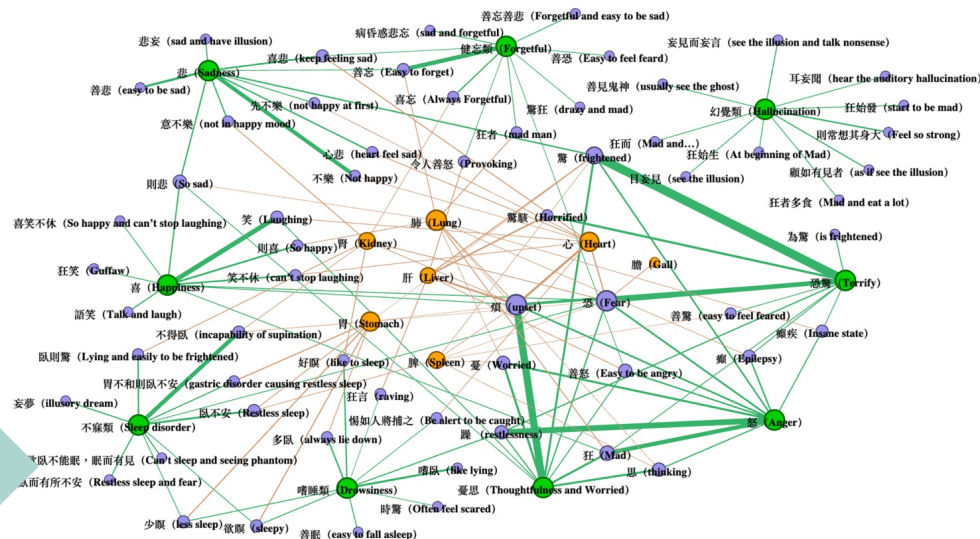
本草綱目

重訂本草綱目
顏鄙老生既病
此二患世莫能
案不食憑牀不
之爲也醫書堆

黃

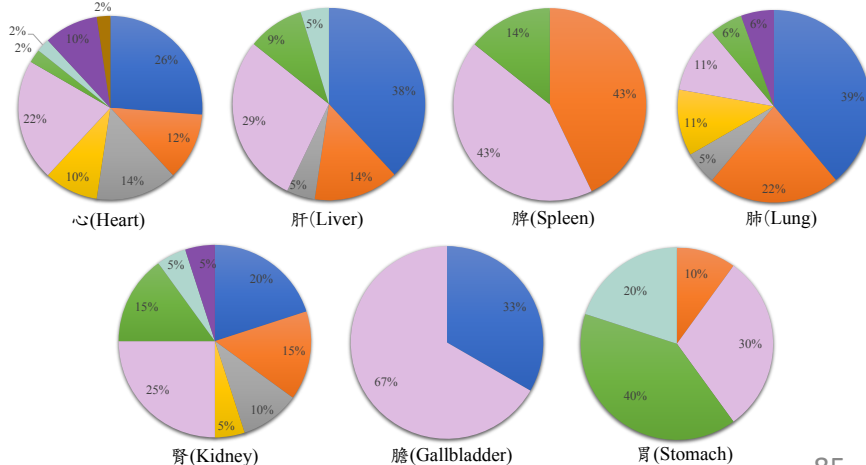


妄夢 (illusory dream)



■ 怒(Anger)
 ■ 憂思(Thoughtfulness & Worried)
 ■ 悲(Sadness)
 ■ 喜(Happiness)
 ■ 恐驚(Terrify)

■ 不寐(Sleep disorder)
 ■ 嗜睡(Drowsiness)
 ■ 健忘(Forgetful)
 ■ 幻覺(Hallucination)



The Future is Here – No Code NLP

Electronic Medical Record

Discharge Summary

Patient Medical Record

Diagnostic Imaging Report

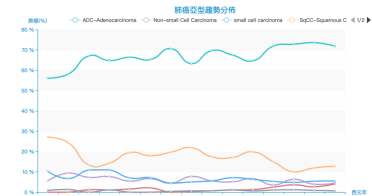
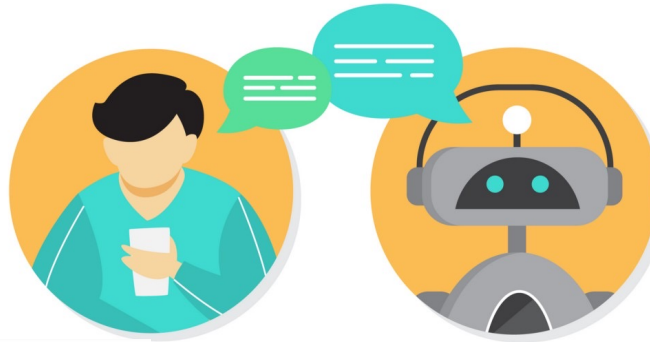
Operation Record

Nursing Report

Pathology Report

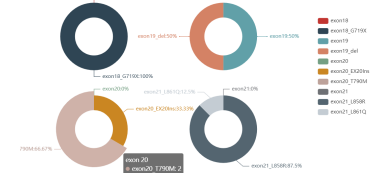
No-Code NLP Platform

1	text	type	organ	diagnosis
2	ORGH SOAP	Ovary		atypical endometriotic cyst
3	Ovary, left, cystic	Ovary		endometrioma
4	Ovary, left, cyst, LP	Ovary		Mucinous borderline tumor of left ovary.
71	(A)	prediction	Endometri	Endometriosis
72	Uterus	prediction	Ovary	Ovary
73	Left ovary	prediction	Ovary	Ovary
74	Vulva, ex	prediction	Ovary	Ovary
75	Ovary, left	prediction	Ovary	ovary



附錄臨床病理報告 - 共計64條, 197條病理報告

EGFR 突變譜系



Lung, lower lobe, left, CT-guided needle biopsy, adenocarcinoma, primary
•organ •operation
•Bx-site

The Future is Here - LLM for Data Annotation

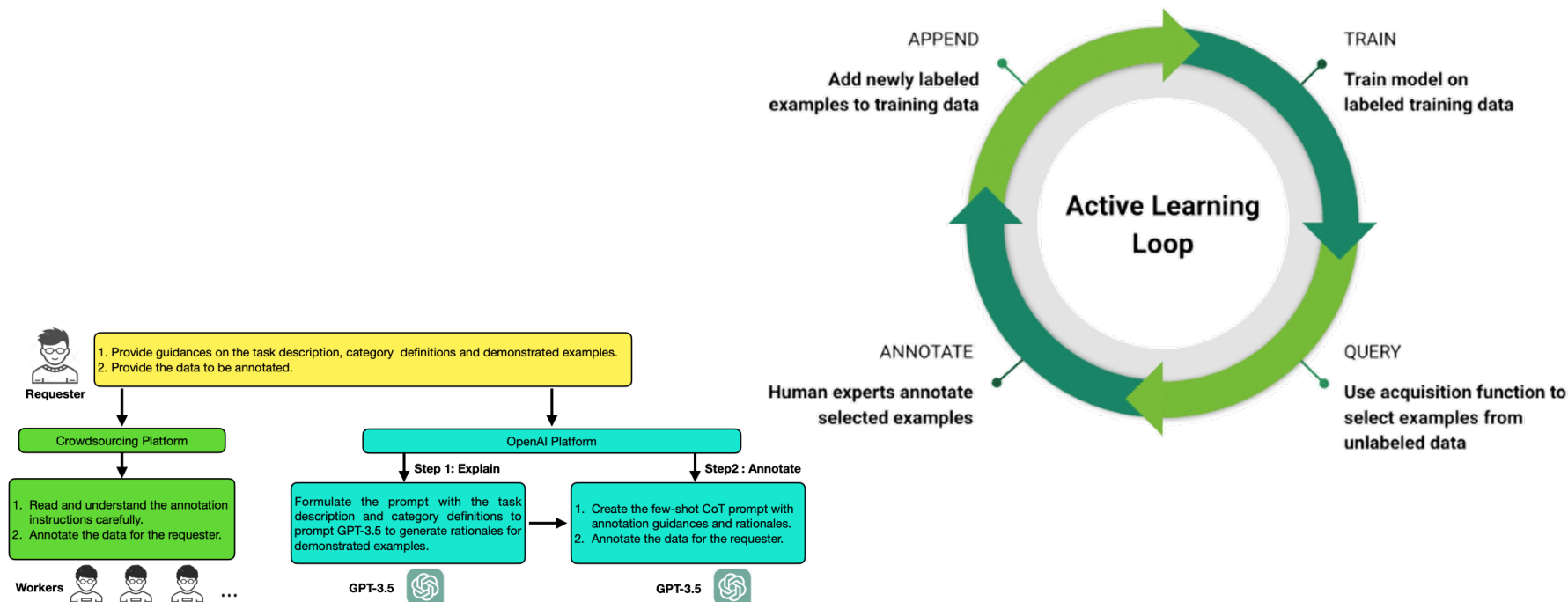


Figure 1: On the left is the annotation process used by crowdsourced workers, while on the right is AnnoGPT's process. AnnoGPT mimics the manual annotation process, with the exception that it generates explanations for each example before annotation. This ensures that each demonstrated example is accompanied by helpful explanations, making the annotation guidelines more informative and useful."

Conclusion

- The application of NLP in clinical settings is beginning to show promising results.
- However,
 - as Large Language Models (LLM) continue to grow in size, there's an increasing effort to compress them.
 - The cost of training and operating ChatGPT is significantly high
 - Not everyone needs to aim for the development of Artificial General Intelligence (AGI).
- Smaller, faster, and more specialized models are the anticipated future.

**Thanks for your
listening**



Q & A