



AI 專案經理 的挑戰

台灣人工智慧學校新竹分校經理人班

彭其捷
foxfirejack@gmail.com

CAMEO
卡米爾股份有限公司

2019/5/11



為什麼選擇
這個主題？



Question

AI 專案經理人 / 管理者，要學
Deep Learning 嗎？要學
Python 嗎？要學演算法嗎？



博客來 google 『人工智慧』，大多是趨勢 / 技術書籍



[AI創世紀：即將來臨的超級人工智慧時代](#)

中文書, 埃米爾·侯賽因 游力秦, 賽鼎, 出版日期: 2018-11-02

優惠價: 79 折, 284 元

[放入購物車](#)

[試閱](#)

「2018人工智慧領域最值得一讀」選書 企業決策者「必讀AI書籍」Top 10入選 Amazon網路書店4.6星讀者口碑好評
波音、思科、福特汽車等跨國企業聯名推薦 終有一天，超級人工智慧必將實現 到那時候，人類有什麼.....

[more](#)



[LINE Bot與人工智慧辨識開發實戰：使用C#與Azure Cognitive Services](#)

中文書, 董大偉, 基峰, 出版日期: 2019-03-11

優惠價: 79 折, 458 元

[放入購物車](#)

[試閱](#)

近年來，各種人工智慧、bot的開發工具與套件紛紛出籠，各大廠紛紛拿出壓箱寶，把手上各種放在實驗室中已久的
人工智慧服務、套件、API端上檯面，讓這個市場突然之間熱鬧了起來。至今筆者的團隊已經建構出了許多LINE Bot
相關應用。同時間.....

[more](#)



[下一個統治世界的企業：人工智慧讓47%以上工作被機器人取代，我如何把威脅變機會？](#)

中文書, 小林雅一 劉錦秀, 大是文化, 出版日期: 2016-08-25

優惠價: 79 折, 253 元

[放入購物車](#)

[試閱](#)

做什麼工作」。人類的最後堡壘絕對不是人工智慧智能，我們還有更厲害的..... 畢竟發明機器人的，還是人。更重要的是，目前相關專業人才屈指可數，你現在追還不遲——。各界推薦 《數位時代》編輯總監 盧諭緯 奇點大學.....

[more](#)

Youtube 也大多在講趨 勢 / 技術議題



筛选器

artificial intelligence management



The Future of Management in an Artificial Intelligence-Based World (Conference Highlights)

IESE Business School • 觀看次數：3815次 • 10 個月前

The Future of Management in an Artificial Intelligence-Based World The Future of Leadership Development Conference Series ...



Artificial Intelligence: New Challenges for Leadership and Management

IESE Business School • 觀看次數：2211次 • 直播時間：1 年前

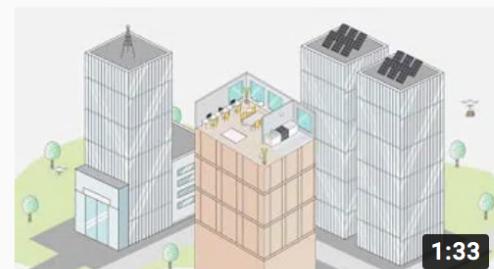
The Future of Management in an Artificial Intelligence-Based World For more info about the conference: <https://bit.ly/2J30TD3> Live ...



Can Management Be Replaced By AI?

Jacob Morgan • 觀看次數：3437次 • 2 年前

There are new advancements in AI every day and people are questioning what types of jobs will be taken over by this new ...

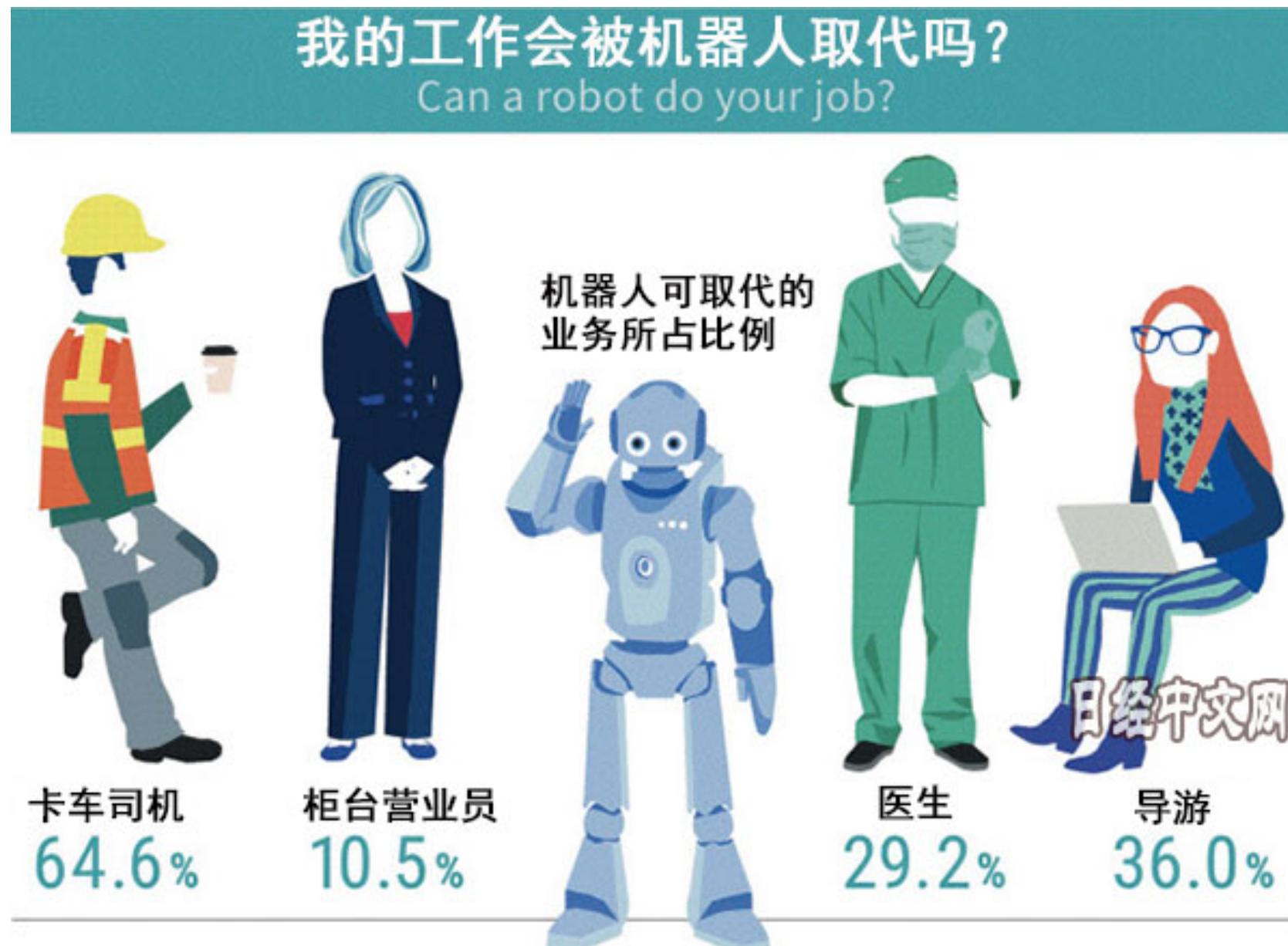


Artificial Intelligence for Facility Management

ACCIONA ✓ 觀看次數：2.1 萬次 • 2 年前

We discover buildings more and more intelligent and sustainable thanks to technology. In this video you can watch curiosities ...

媒體 / 電影：大多著墨於社會與科技面對於未來的想像



市面上的資訊，主要分成三大類

- **意見領袖**：講 AI 的大趨勢與機會
- **媒體與電影**：講 AI 未來可能帶來的風險
- **技術專家與書籍**：講 Python / Deep Learning / 統計學

我覺得，較少給管理者/經理人的實務經驗分享資訊，因此決定了今天的主題



ABOUT ME

關於我

關於我 About Me



國立台灣大學 地理環境資源學系
地理計算科學 研究室
Laboratory for Geospatial
Computational Science
NTU Geography

 財團法人資訊工業策進會
INSTITUTE FOR INFORMATION INDUSTRY

 國立交通大學
National Chiao Tung University

【卡米爾】專案經理

負責領域：資料科學、UI/UX 設計

【台大】地理與環境博班

研究領域：空間分析演算法

【資策會】研發替代役

數據系統設計、系統開發、前端工程師、資料庫管理員

【交大】資訊管理研究所

研究領域：資料探勘與資料庫



相關教學 / 演講 / 經歷

- 清華大學：AI 與體驗設計講師
- 台灣大學：大數據與視覺化講師
- 國泰集團：資料視覺化 UX 設計 - 企業內訓講師
- 科智公司：資料視覺化 - 企業內訓講師
- 2017 資料科學年會：資料視覺化主題 Speaker
- KKDAY：Tableau 視覺化軟體 - 企業內訓講師
- 台灣鐵路局：新興科技與 UX - 企業內訓講師
- 中原大學：前端工程與RWD設計 - 課程講師
- 東吳大學：EMBA 科技與群眾募資課程講師
- 玉山科技論壇：團隊對談講師
- 交通部：跨多單位體驗設計講師
- 資策會：創新樂活服務設計工作坊 - 主導講師
- 環保署：跨單位服務設計講師
- 天地人：資料視覺化課程講師
- YOTTA：設計思考課程講師



CAMEO

卡米爾股份有限公司



INSTITUTE FOR INFORMATION INDUSTRY



Taiching Veterans General Hospital
TCVGH



台灣鐵路



演講時間/地點

台灣資料科學年會 Day 2 11/12 (日) 16:45–17:30 A0

演講摘要

許多人都了解資料視覺化的優點，例如可將資料抽象化、讓資訊降維等等，讓數據以更直提升閱讀者的吸收意願。然而，我們依然可以舉出許多例子，即使是視覺化後的數據，也難以勾起人們理解的意願。『使用者經驗（UX）』設計的最終目標是『Don't Make Me Think』，暢的閱讀過程，讓使用者能夠更輕鬆理解背後的寶貴價值，而這之中除了透過視覺設計，還包括了使用者研究、需求釐清、文案優化等等技巧，如果能夠將使用者經驗設計相關方法應用在資料視覺化上，一次降低數據解讀的門檻，讓更多人了解其價值意涵，也讓更多人被數據所感動。

主持人

Yves Lin / Titansoft

講者簡介



彭其捷 ↗ 個人網站 <https://about.me/divaka>

卡米爾股份有限公司 / 數據專案經理

彭其捷畢業於交大數據分析實驗室，對於網路服務設計有深入研究，擅長於機器學習、資料視覺化、前端工程、服務設計等等類型專案，出版過多本相關技術書籍，並於許多國際會議上發表論文。他專長於打造良善的網路服務體驗，近期多投入資料科學相關主題的研究，著重於如何將複雜的數據知識轉化為易懂的資訊，並強調人機互動的連結，讓數據分析更平民化。



2017 台灣資料科學年會

- 交大資管，資料探勘實驗室
- 現職：數據專案經理
- 經歷：前端工程師
- UI/UX 講者

資料視覺化講者、專欄作者



【資料治理怎麼做】環保稽查也要科學辦案，環保署靠AI才識破不肖工廠排污

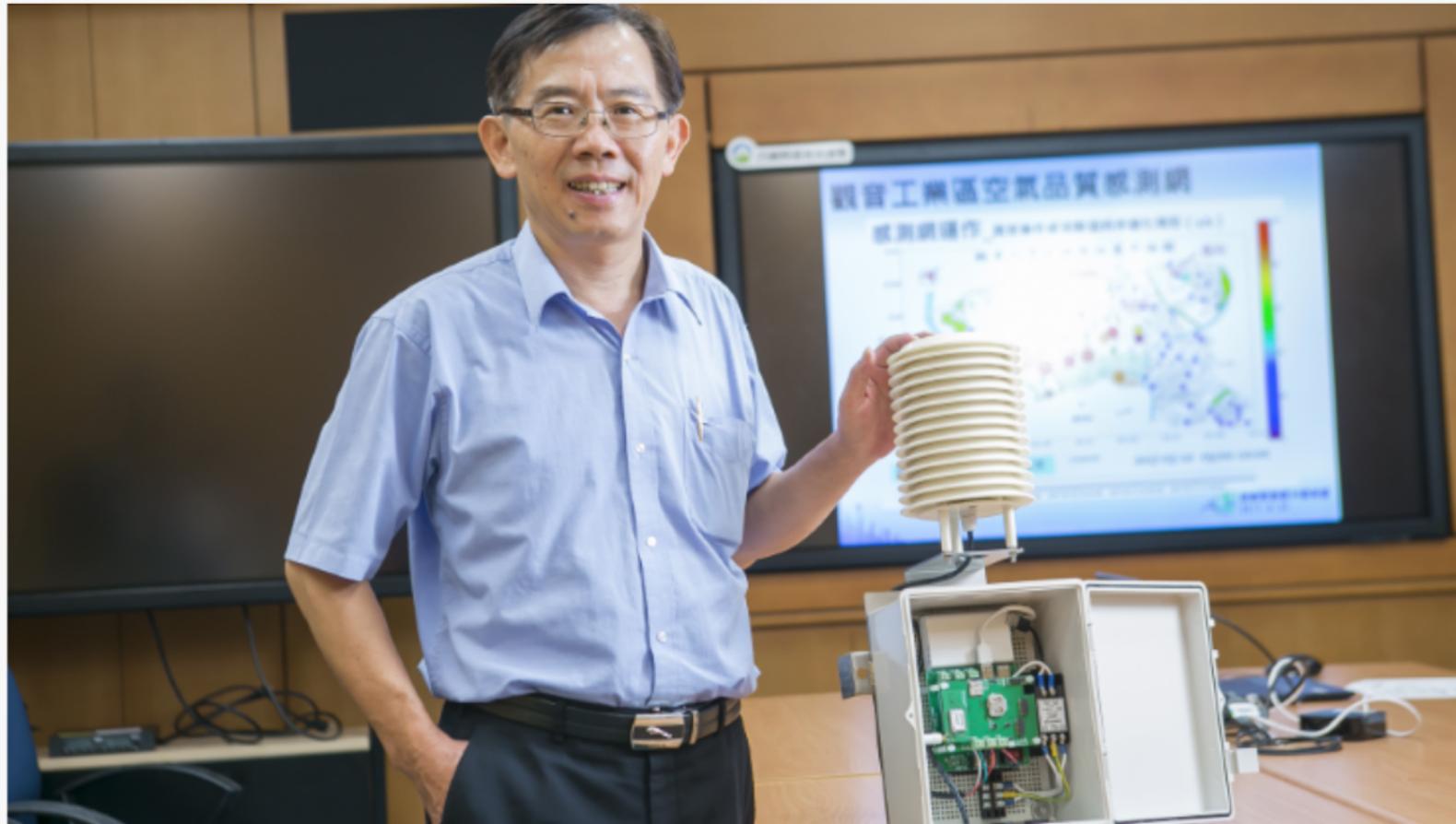
不肖業者常以各種取巧、偽裝來規避稽查員蒐證和抽查，環保署改從大量監測數據中，靠機器學習分析污染排放規律，並鎖定可疑區域，幫助稽查員找出關鍵稽查時機，現場人贓俱獲，逮住不良工廠暗夜偷排廢氣

文/ 余至浩 | 2017-10-09 發表

讚 4.4 萬 按讚加入iThome粉絲團

讚 456 分享

G+



大數據監控空汙 讓台灣的天空更乾淨

2018-11-30

作者: 廣告部企劃製作



A A A



卡米爾專案經理彭其捷。(圖/經濟部工業局提供)



大數據監控空汙 讓台灣的天空更乾淨

用人工智慧偵測空汙 推動城市治理智慧

空氣汙染從綠色、紅色到紫色，代表PM2.5由少到多的程度，這是天氣預報常見的指標。但如何知道住家附近最即時的空汙狀況呢？

未來，一張立體的城市鳥瞰圖，紫紅色的柱狀最高、綠色的最低，這些高高低低、不同顏色的柱體，一看就會讓您知道自己此刻是住在PM150以上紅色警戒區？或者是PM15.5以下的綠色清新區？

**AI分析空氣感測數據
智慧城鄉應用**

空汙是民眾高度關注的議題，經濟部工業局透過智慧城鄉計畫，在全台推動智慧城市

市應用服務，促成智慧生及。其中，在空氣品質監面，特別透過跨部會合作，合公私資源，透過物聯網、數據分析等技術，推動一個明化的智慧稽查系統。

這套系統中所需的AI分析軟體技術，是由成立有18年的卡米爾CAMIO團隊負責。這間集合大數據分析、機器學習、UI/UX設計、視覺化商業智能分析、策略專家的公司，技術核心由三位交通大學資訊博士設計。

示，目前全國13個縣市已經建置了200個空氣品質微型感測器，需要有人協助分析感測器所偵測到的數據，故2017年

科技看非凡

非凡新聞 HD

AI 立體成像助整合
物聯網感測通報資料



台中市

19-23

美元兌日圓

113.470

19:20:56

卡米爾專案經理 彭其捷

比較高污染的一個狀況

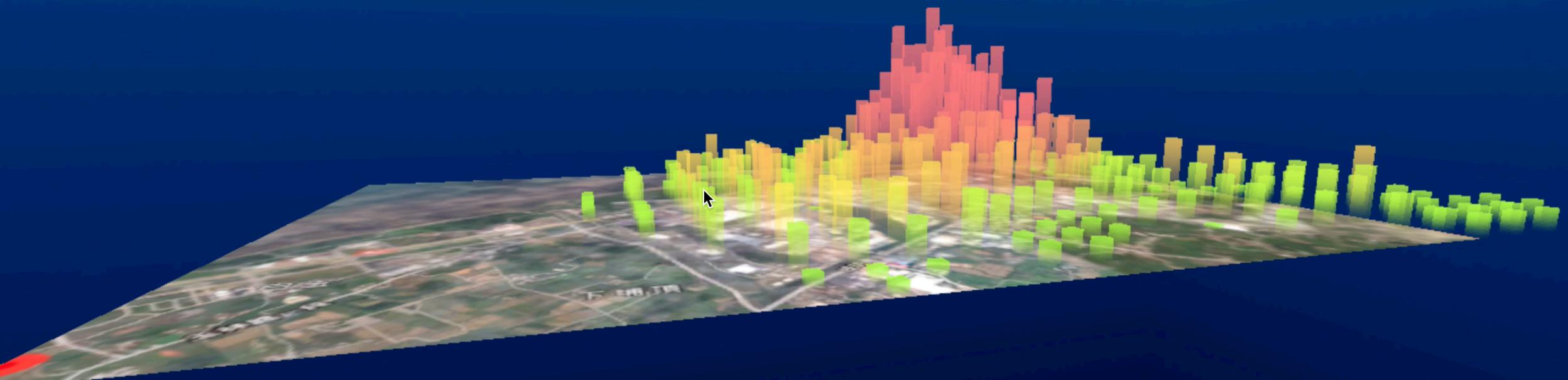
臺指選12

臺指選擇權 9800買 182.00

▼15.00 178



人工智慧環保稽查

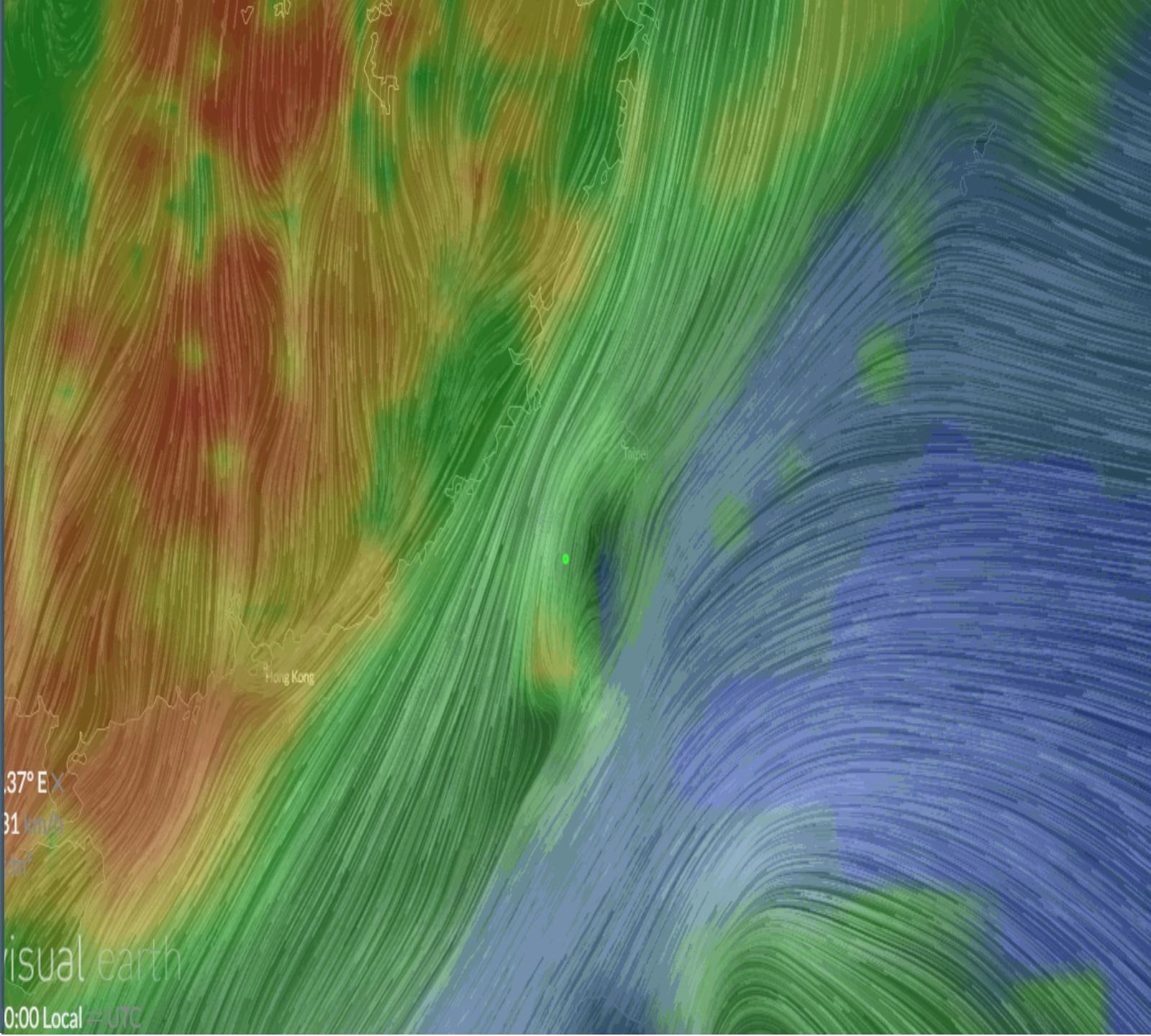




AI 專案

機器學習

空氣污染預測



全文檢索



熱門搜尋：環境即時通APP



關於環保署

行政公開資訊

環保統計

便民服務

環保法規

相關連結



「環境即時通」推出即時、適地之空品預報

輪播各圖片說明與連結

空氣品質(AQI)預報

(未來24小時)

發布時間：2018/05/02 16:30



空氣品質現況



愛環境資訊網



水質監測



紫外線現況



Question

有人出門之前，會先查看
空氣品質狀況嗎？



空污敏感族群，會想要知道極短時之空氣品質變化



- ❖過敏族群
- ❖氣喘患者
- ❖皮膚過敏
- ❖父母族群
- ❖運動族群
- ❖單車族群

環保署提供之空氣品質預測 (2019/05/11 10:30 提供之預測)

全國各空品區空氣品質指標 (AQI) 預報

圖形

日期	05/11		05/12			05/13	
空品區	AQI 指標	指標 污染物	AQI 指標	指標 污染物	AQI 指標	指標 污染物	
北部	60	細懸浮微粒 (60)	70	細懸浮微粒 (70)	100	臭氧八小時 (100)	
竹苗	80	細懸浮微粒 (80)	85	臭氧八小時 (85)	90	臭氧八小時 (90)	
中部	125	臭氧八小時 (125) 細懸浮微粒 (105)	130	臭氧八小時 (130) 細懸浮微粒 (105)	125	臭氧八小時 (125)	
雲嘉南	155	細懸浮微粒 (155) 臭氧八小時 (105)	130	臭氧八小時 (130) 細懸浮微粒 (105)	140	細懸浮微粒 (140)	
高屏	140	細懸浮微粒 (140) 臭氧八小時 (105)	125	臭氧八小時 (125) 細懸浮微粒 (105)	130	細懸浮微粒 (130)	
宜蘭	50		50		50		
花東	40		40		40		
馬祖	110	臭氧八小時 (110)					
金門	105	細懸浮微粒 (105)					
澎湖	65	臭氧八小時 (65)					
分類	良好	普通	對敏感族群不健康		對所有族群不健康		非常不健康
指標等級	0~50	51~100	101~150		151~200		201~300
	301~500						

備註：AQI預報為該預報日有效時間內(0至24時)，出現之可能最大即時AQI值及當時之指標污染物。但污染物濃度可能受其日夜變化影響，即時之指標污染物將有所變化。



Taipei City
上午10:49 CST



毛毛雨

↑ 23° ↓ 20°

22°

© 由eric802222-雞雞毛... flickr

預報

上午11時 下午12時 下午1時 下午2時 下午3時 下午4時 下午

50	50	50	50	50	50	50
23°	23°	23°	22°	22°	21°	21

星期五 23° 20°

星期六 24° 18°

星期日 20° 19°

星期一 25° 21°

，，

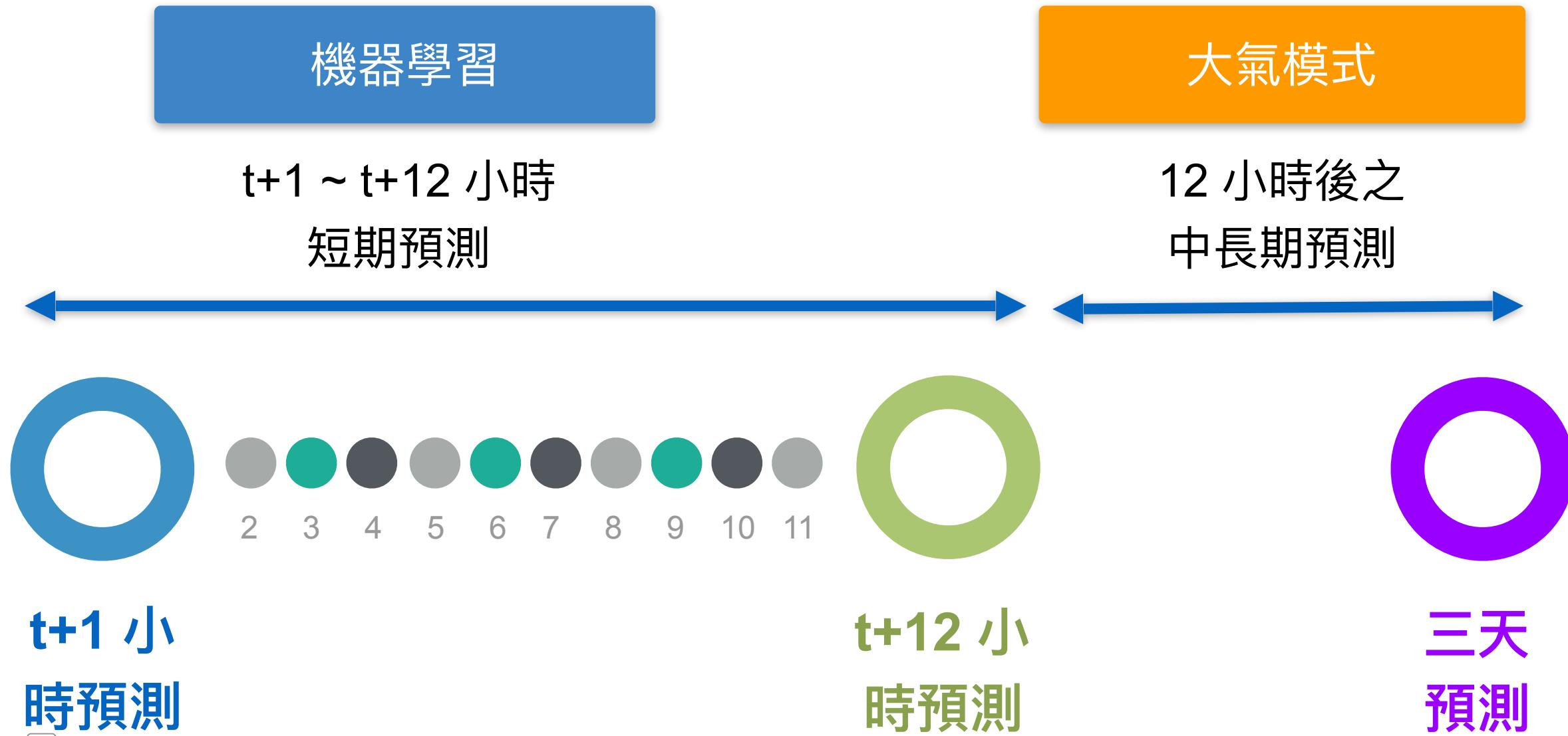
氣象領域，已經可以提供 hourly 的預測結果，空氣品質是否也可以呢？





環境即時通

空氣品質 (1) 短期預測 搭配 (2) 中長期預測





環保署 國家級測站資料，單一測站每天每小時的資料

累積 10 年，總共超過 600 萬 筆資料進行運算

日期\時間	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2017/03/01	10	12	13	14	11	6	9	13	12	15	13	19	10	14	17	17	17	19	26	38	39	45	51	61
2017/03/02	63	47	31	30	30	38	35	36	32	32	34	29	24	28	27	27	17	24	22	28	27	36	30	23
2017/03/03	18	19	21	15	15	12	11	8	10	5	7	8	12	11	5	3	11	17	21	14	15	23	27	37
2017/03/04	30	30	20	23	27	21	22	26	33	30	23	25	23	21	12	12	12	26	26	32	34	29	26	25
2017/03/05	43	45	45	39	41	41	32	39	40	53	42	39	36	39	39	34	40	36	45	38	38	30	28	
2017/03/06	27	21	12	3	3	3	3	5	14	17	20	23	20	18	#	21	28	40	47	60	71	75	72	60
2017/03/07	47	34	21	17	20	21	20	24	20	22	15	13	18	14	17	13	20	21	22	22	33	33	35	31
2017/03/08	41	38	47	39	39	34	41	47	42	48	47	49	43	38	38	29	31	26	24	23	30	33	33	30
2017/03/09	32	38	36	36	30	33	32	30	28	26	30	25	25	20	23	19	22	16	12	11	11	21	14	12
2017/03/10	11	12	12	9	9	9	10	10	14	15	19	12	17	12	*	12	13	14	8	12	9	6	8	4
2017/03/11	6	2	3	7	5	*	11	18	7	11	9	6	10	4	11	12	13	9	10	10	8	ND	4	8
2017/03/12	8	7	5	7	5	10	12	11	16	23	27	21	14	12	12	23	31	35	39	39	38	30	24	33
2017/03/13	39	51	48	39	30	31	34	42	33	30	12	13	12	14	12	11	12	7	7	6	5	ND	ND	4
2017/03/14	5	#	#	NA	x	NA																		
2017/03/15	NA	#	#	33	24	16	12	16	17	19	16	10	12											
2017/03/16	12	12	10	12	10	9	8	5	10	12	12	7	3	4	11	12	7	8	6	12	11	10	7	8
2017/03/17	6	3	4	6	11	9	9	7	7	9	13	10	13	10	12	NA	9	8	7	10	14	14	10	7
2017/03/18	9	10	8	4	6	5	15	12	12	13	17	26	20	21	20	21	30	25	28	24	18	16	9	11
2017/03/19	12	8	7	4	11	12	7	3	3	4	5	6	5	7	9	11	9	8	3	4	8	12	10	5
2017/03/20	11	16	20	20	19	19	7	2	2	22	22	19	19	22	22	22	22	22	22	22	19	19	19	22

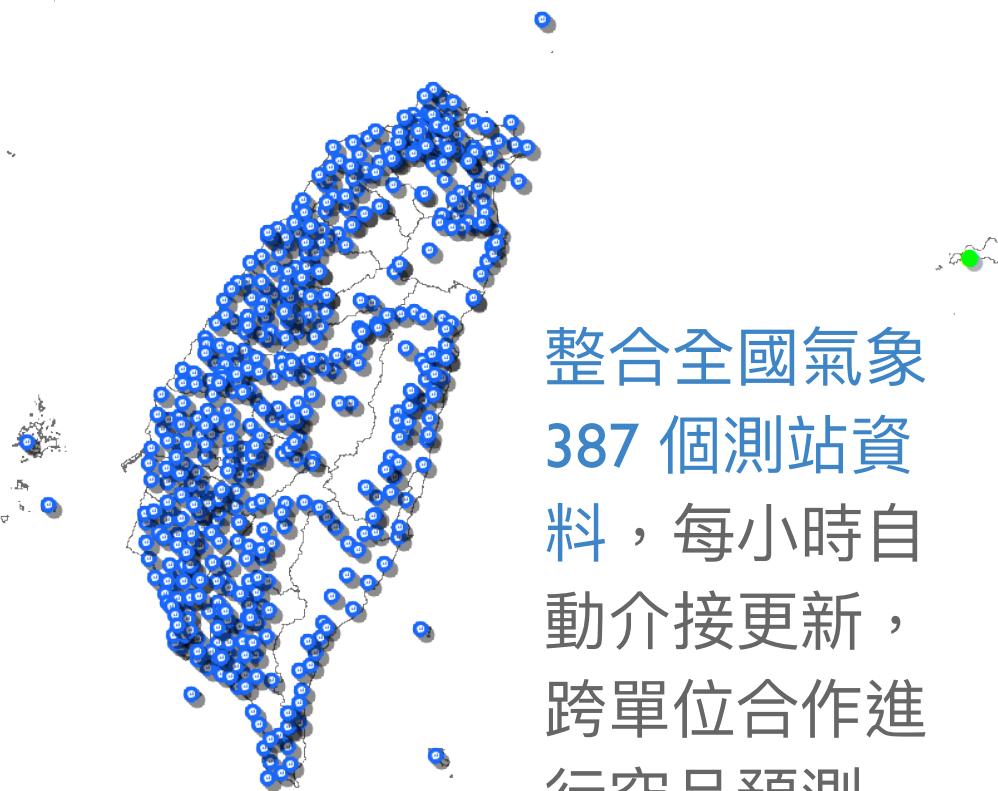
整合氣象局『氣象資料』與環保署『空品數據』



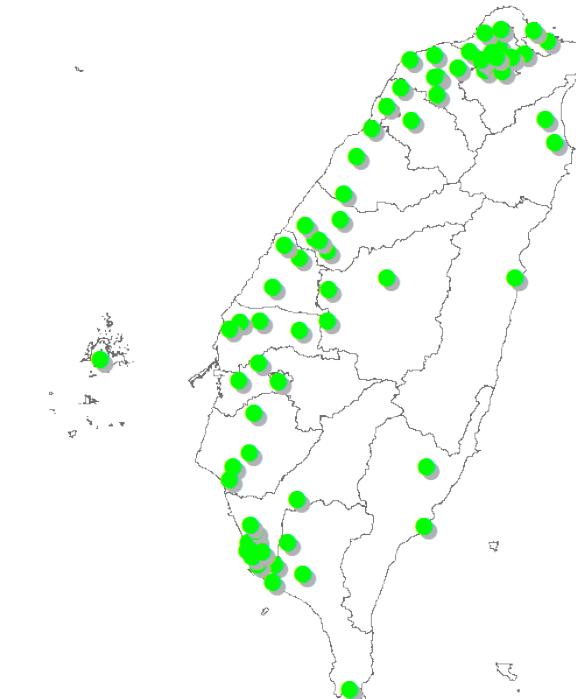
交通部中央氣象局
Central Weather Bureau



行政院環境保護署
Environmental Protection Administration
Executive Yuan, R.O.C. (Taiwan)

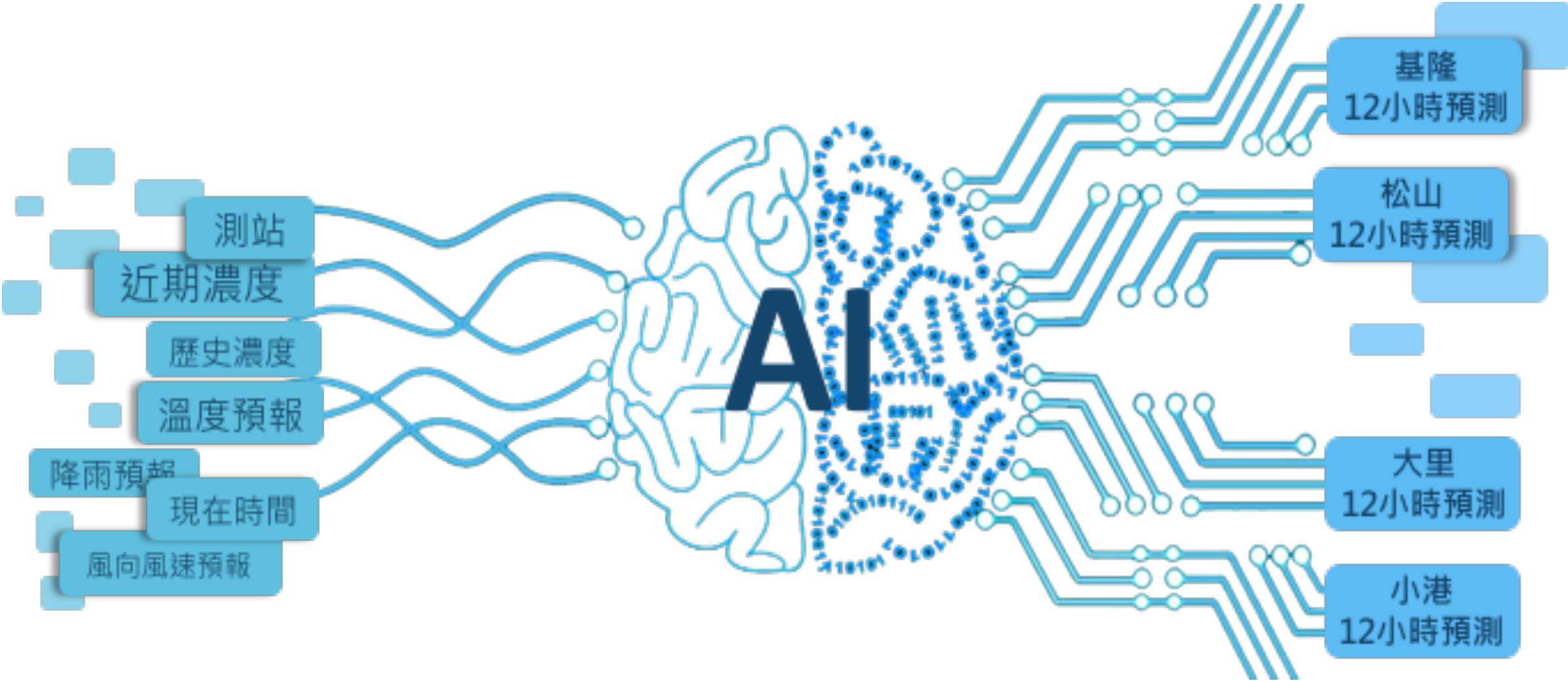


整合全國氣象
387 個測站資
料，每小時自
動介接更新，
跨單位合作進
行空品預測



10年76個國家
級測站資料，
每日每小時的
資料，訓練資
料超過 600 萬
筆

透過 AI 類神經網路技術：預測空氣品質



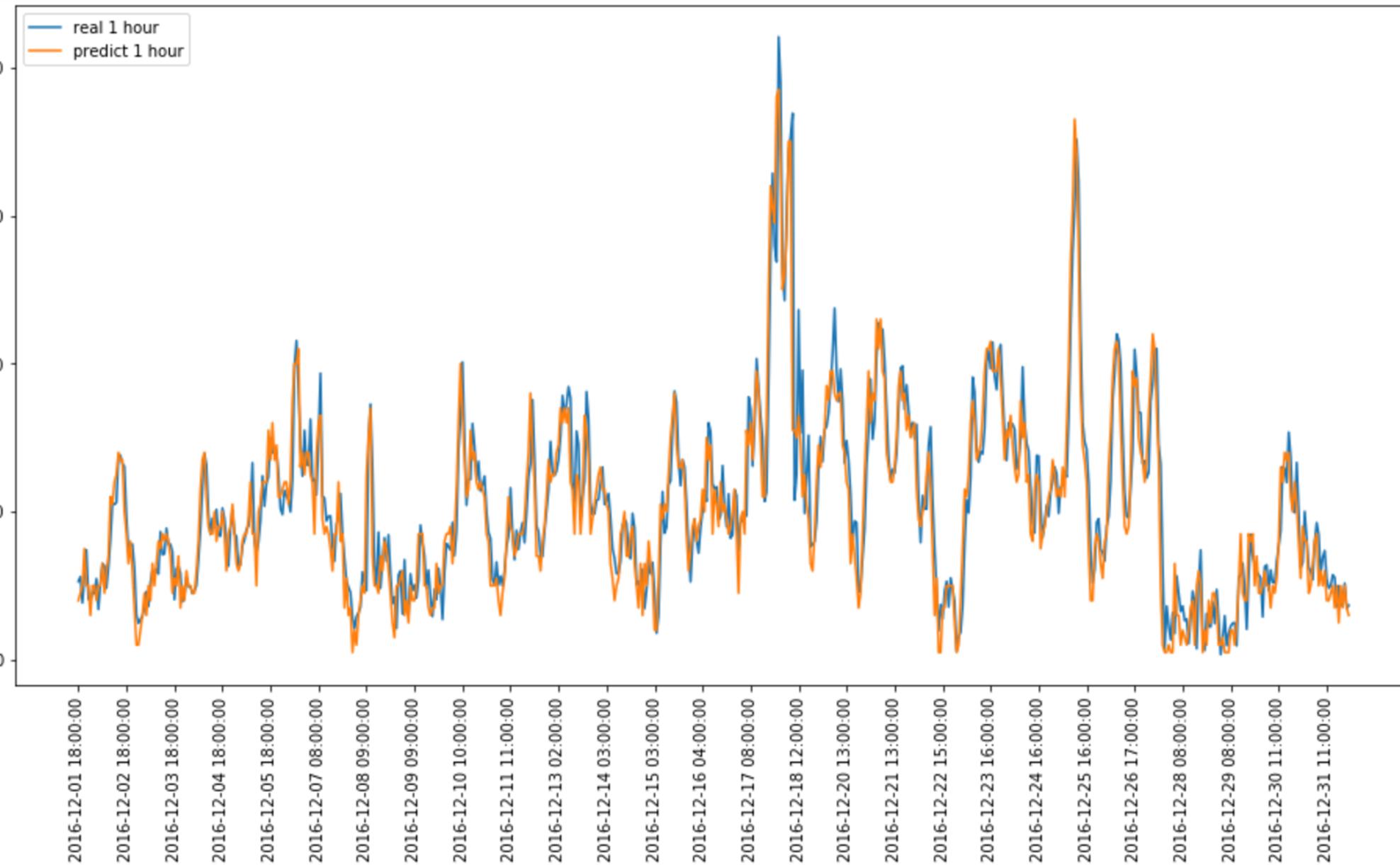


2018年，極短時空氣品質預測服務 上線記者會

29



實驗：空品『預測』與『真實值』之貼合程度



	AQI 誤差
t+1	4.2
t+2	5.5
t+3	7.2
t+4	8.9
t+5	10.3
t+6	11.4
t+7	12.4
t+8	13.4
t+9	14.3
t+10	15.2
t+11	16
t+12	16.9



AI 專案執行過程中：五大挑戰

AI 專案經理的五大挑戰





#1 技術深化

AI Knowledge / Algorithm



Question

經理人 / 管理者 / PM
需要了解技術議題嗎？

A photograph of a man and a woman in professional attire talking outdoors. The man, on the left, has a beard and glasses, wearing a white shirt and red tie. The woman, on the right, has curly hair and is wearing a grey blazer over a white top with a necklace. They are both holding coffee cups. The background shows a brick building and some greenery.

[情境] 工程師：我想使用 LSTM 方法優化看看演算法

Which One Is Better ?



心動指數
10%



我想使用多項
式回歸方法優
化看看，可以
嗎？



請問
LSTM 是
什麼？

心動指數
50%



我想使用多項
式回歸方法優
化看看，可以
嗎？



你覺得合適的
話，我們就來
試試看吧！

心動指數
100%



我想使用多項
式回歸方法優
化看看，可以
嗎？

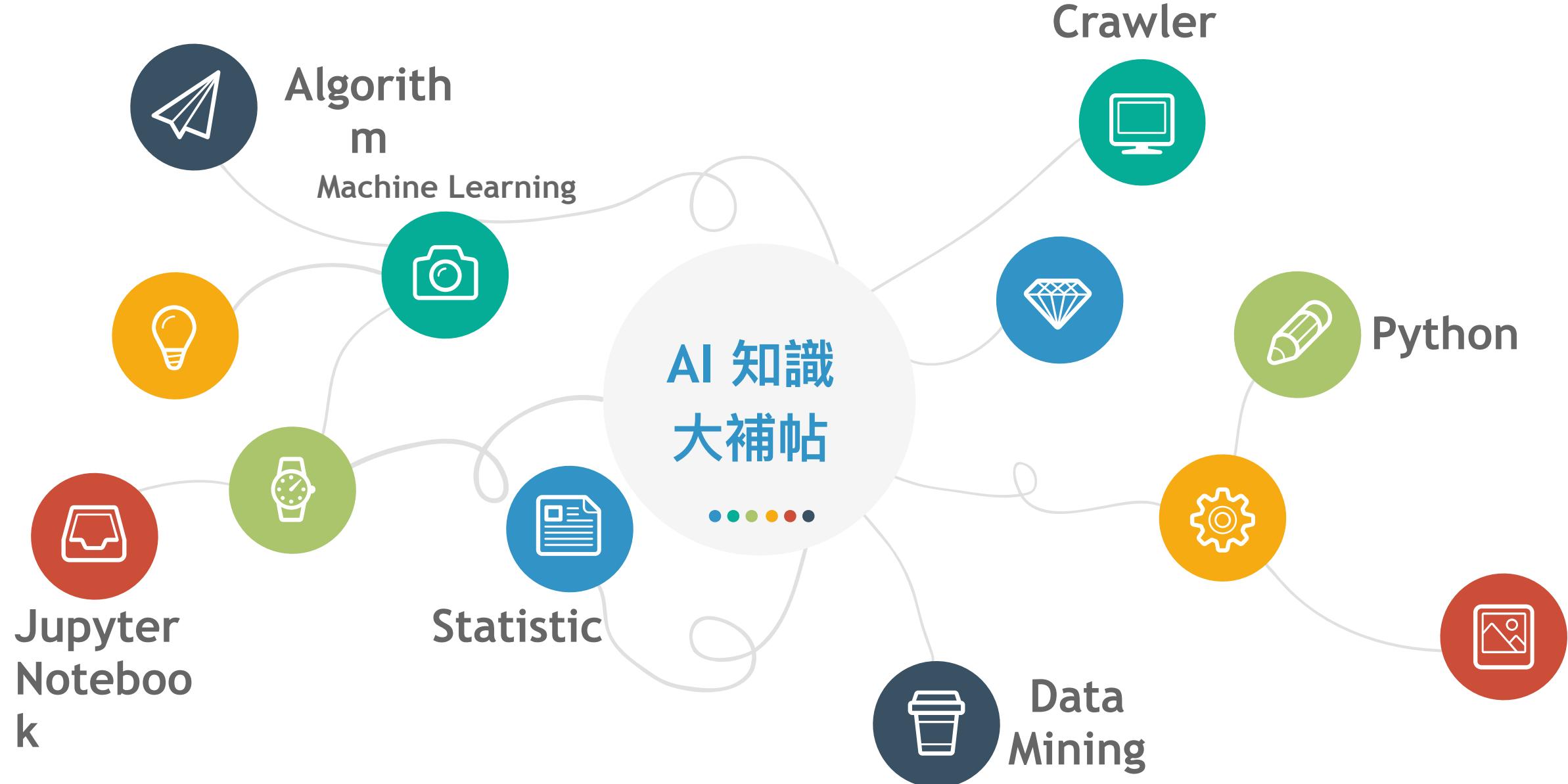


太棒了！我記得那
是一個加入時間先
後次序的演算法，
會納入時間序列的
概念！



Question

技術很重要
但對於管理者來說
應該要了解多深呢？



不必然需要了解，但可以有大概的概念，就是加分



<https://software.intel.com/zh-cn/articles/hands-on-ai-part-3-the-anatomy-of-an-ai-team>

數據科學家

數據科學家通常要求掌握以下技能：

- 算法設計與分析
- 線性代數
- 概率論
- 數理統計
- 腳本語言（主要是Python*）
- 數據建模語言（主要是Python、R*、MATLAB*、Mathematica*）
- 數據管理語言（SQL 及其衍生品，比如Pig* 和Hive*）
- 機器學習（分類、回歸、聚類）
- 關係數據庫（PostgreSQL*、MySQL*、Oracle*）
- 版本控制（例如Git*）

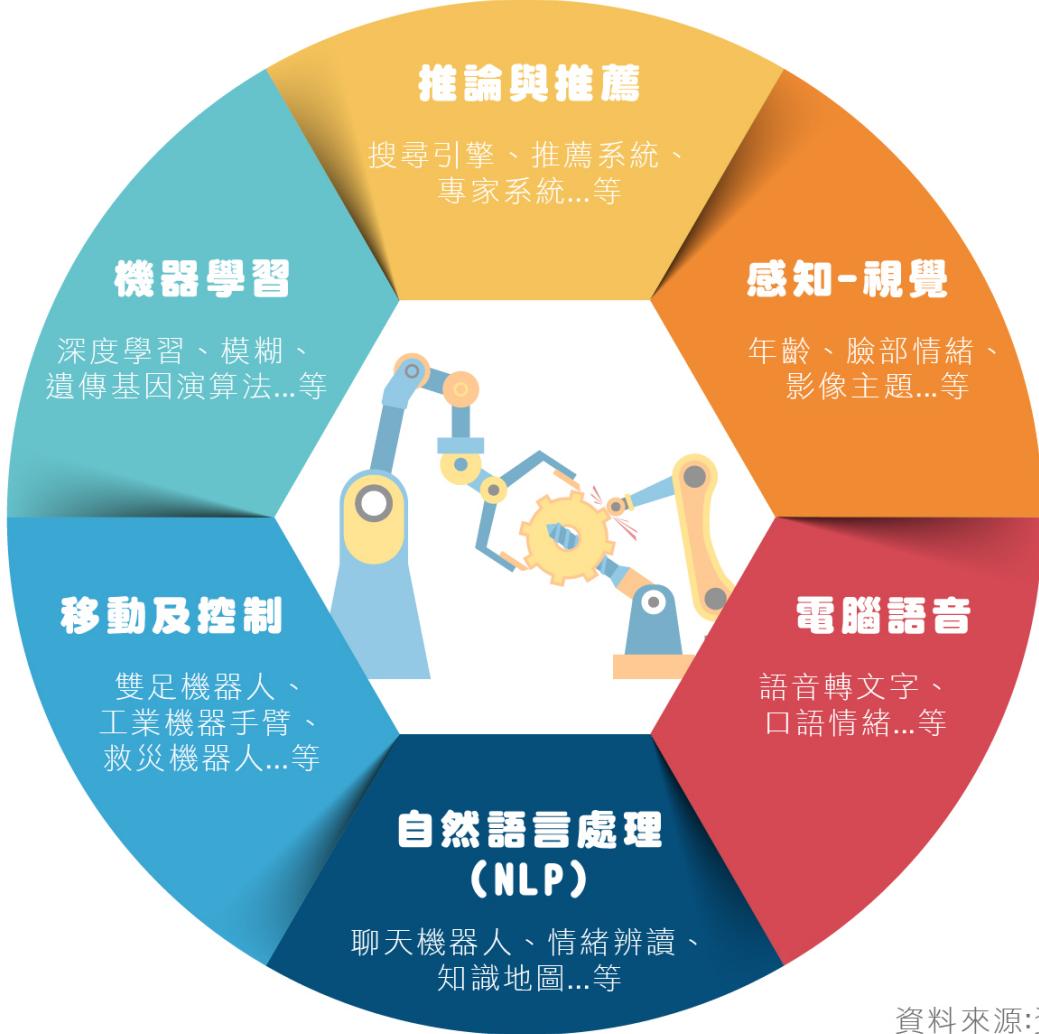
數據工程師

該角色與大數據概念有著直接的關係。數據工程師可操作機器學習算法處理大型數據集。儘管數據工程師要求的技能與數據科學家有所重疊，但兩者之間還是存在一些重要差異。

要求掌握的技能包括：

- 算法設計與分析
- 腳本語言（主要是Python）
- 精通機器學習理論（分類、回歸或聚類）
- 分佈式機器學習框架（Spark*、Storm*、H2O.ai*）
- 數據管理語言（SQL 及其衍生品，比如Pig 和Hive）
- 關係數據庫（PostgreSQL*、MySQL*、Oracle）
- 分佈式系統和NoSQL* 數據庫（Hadoop*、Cassandra*、HBase*、Riak*、Kafka*、Dynamo*、Redis*、MongoDB* 或ElasticSearch*）
- 版本控制（例如Git）
- 雲計算（例如Amazon Web Services*、Microsoft Azure* 或Google Cloud*）
- Linux*

More is Better : 學習專業知識的方法



- 從大分類開始了解起 (不要想說什麼都學)
- 其實 google 就很好用的
- 找時間看看 AI 的書 (技術的書，而不是只有趨勢的書)
- 打破砂鍋問到底
- 看論文 (也可專注看 Abstract / Summary 的說明)

資料來源:資策會MIC, 2017年6月



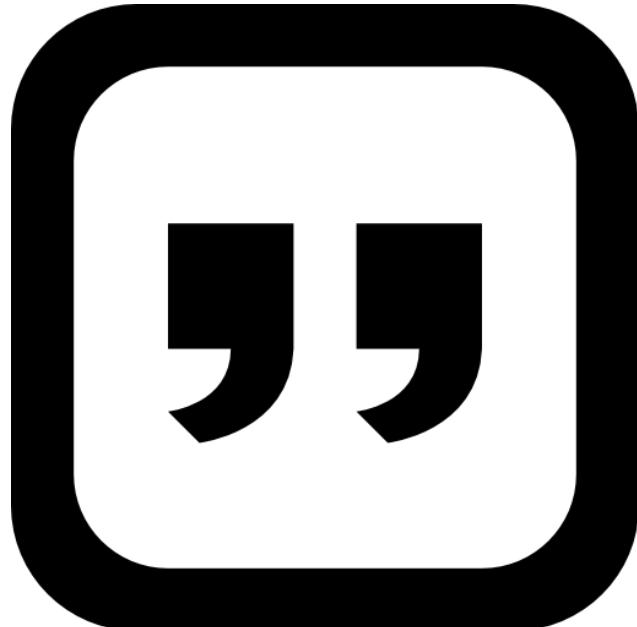
#2 領域知識

Field Knowledge



Question

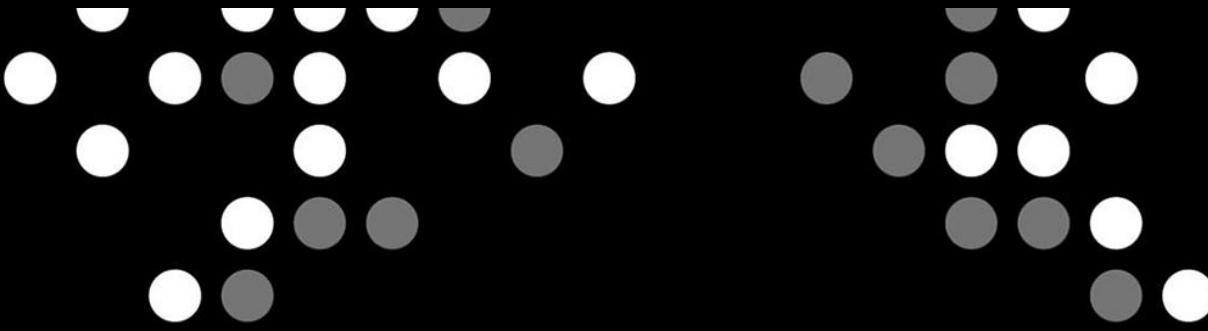
只要擁有高深的 AI 技術，就能設計出良好的各類 AI 嗎？



領域專家

擅長某一特定行業或垂直行業。通常情況下，領域專家提供的信息可用於面向機器學習和數據註釋的特性設計。

☆ 領域專家不一定要懂 AI



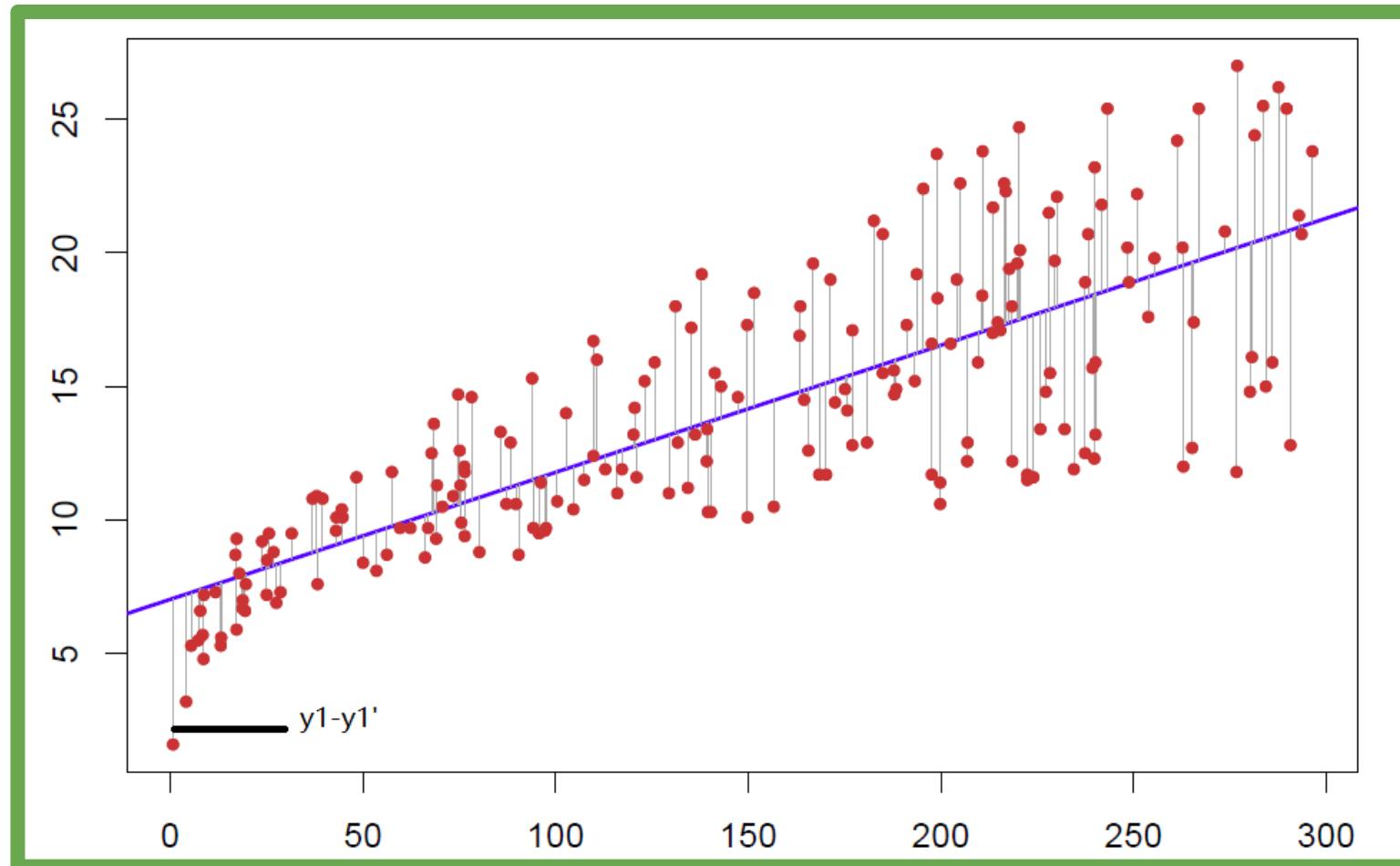
ALPHAGO

A DOCUMENTARY • SPRING 2017



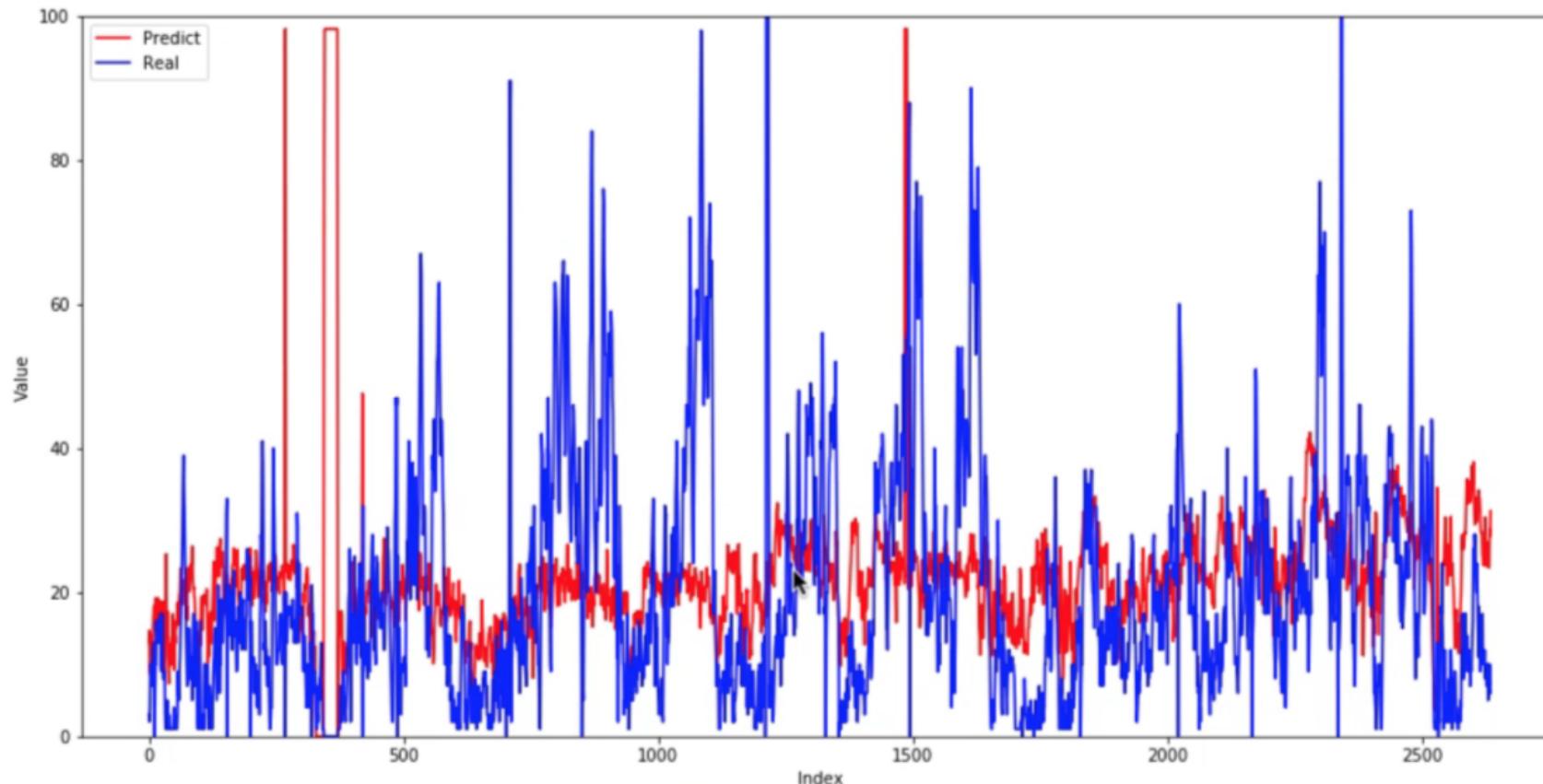
Phase 1：建立 AI baseline

數值預測，透過線性回歸，可作為分數的基礎（Baseline）



持續納入領域知識，可不斷優化運算結果

線性預測 baseline：RMSE 分數都 > 10 (誤差過高)

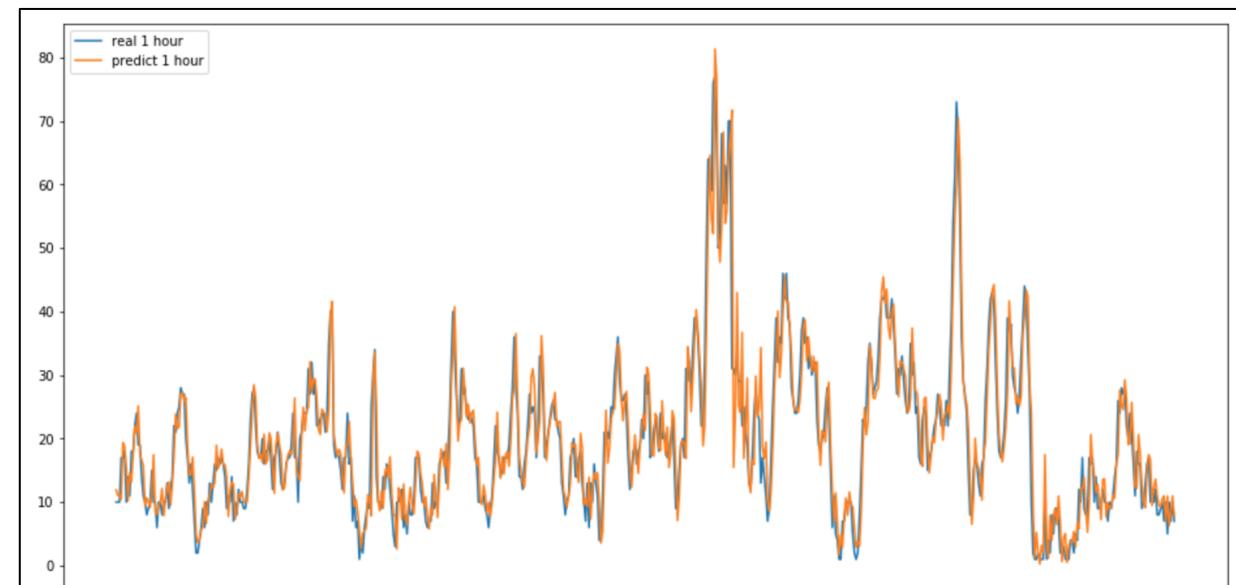
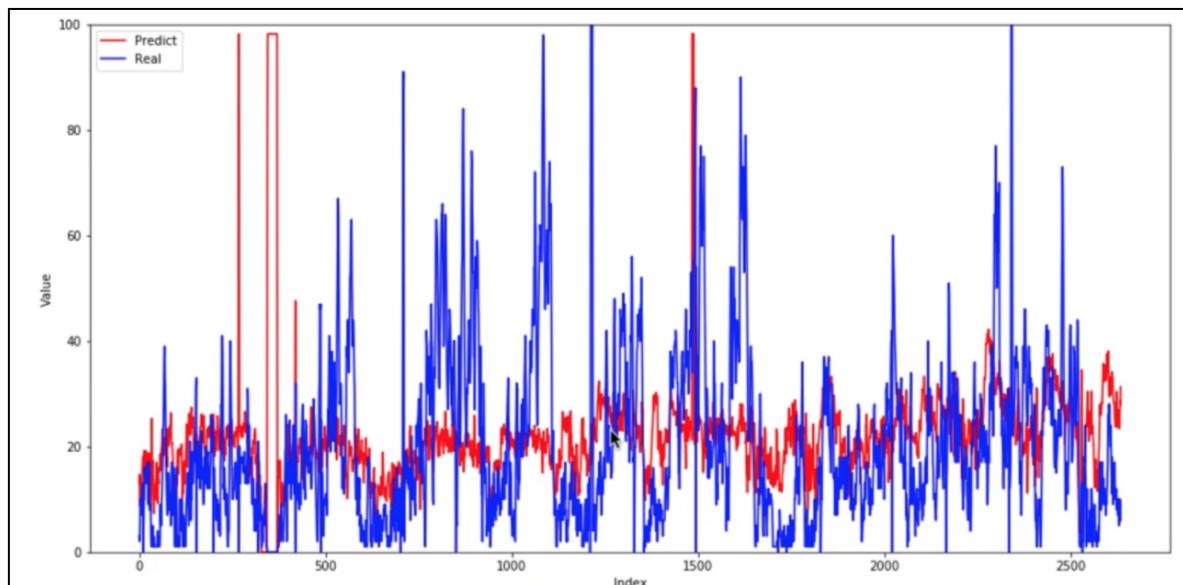
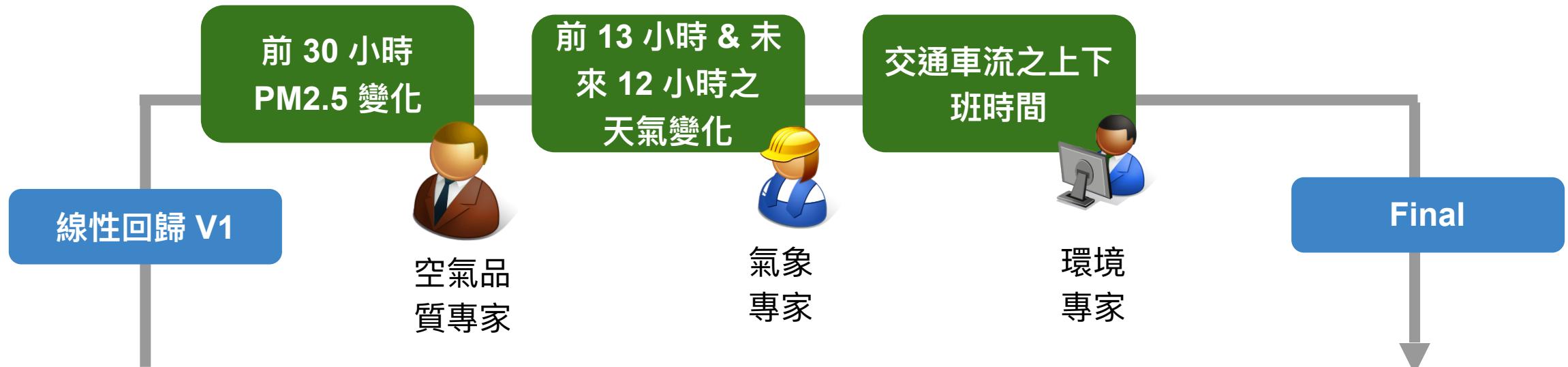


預測不夠準
的話，會造成
民眾恐慌



行政院環境保護署
Environmental Protection Administration
Executive Yuan, R.O.C. (Taiwan)

V1 > Final : 持續透過領域知識優化



跑了超過 400 次的實驗，逐步釐清有效的特徵

	M	N	O	P	Q
1	10次的平均	最好的分數	分數改變	訓練時間	備註
2	4.746117	4.670003	baseline	20.3	使用刪除有缺漏值的資料 PM2.5需連續五筆資料才可以被訓練與預估
3	4.695131	4.657201	變好	29	是目前最好的組合，且透過model selection驗證，使用MLP預設的參數表現最好
4	4.780128	4.674715	變差	25.2	加入月份，反而訓練的不佳
5	5.026653	4.682835	變差	18	單純加入PM10輔助，反而影響平均的表現
6	4.859515	4.632743	變好	24.4	再次驗證加入'hour'特徵，表現較好
7	4.91092	4.65723	變差	27.9	單純加入'month'反而使預測準度下降
8	5.440511	5.41658	沒感覺	6.7	GradientBoostingRegressor，表現普普
9	5.445867	5.405747	沒感覺	19.7	GBR沒辦法有效學習到'hour'特徵
10	5.452919	5.42444	變差	23.8	表現依然與MLP有些差距
11	5.025852	5.025852	沒感覺	0.3	表現稍微比GBR好，但RMSE仍比MLP差0.33
12	5.49617	5.391339	baseline	7.8	冬天RMSE最高，(5579, train) (1401, test)
13	5.314498	5.248047	baseline	8.2	春天表現普通，(6234, train) (1406, test)
14	4.096829	3.941363	變好	10.1	夏天的訓練資料與測試資料比較其他季節少，但預測效果卻好，(4845, train) (1200, test)
15	4.897542	4.722512	baseline	11.7	秋天預測效果也普通，但是缺漏的資料較少，(6746, train) (1492, test)

訪問領域專家：來取得數據特徵（特徵工程）

根據：空污歷史數據，納入前 30 小時狀態進行建模

前 30 小時、前 29 小時、前 28 小時

... 前 3 小時、前 2 小時、前 1 小時

可將白天晚上週期影響納入考量

根據：天氣歷史數據，納入 13 小時單位進行建模

風速：前 13 小時、前 12 小時 、前 1 小時

風向：前 13 小時、前 12 小時 、前 1 小時

溫度：前 13 小時、前 12 小時 、前 1 小時

降雨：前 13 小時、前 12 小時 、前 1 小時

考慮連續的歷史氣象特徵變化

納入：未來12小時天氣

納入：未來 12 小時的天氣狀況，比對過去的歷史數據，進行污染的預設



另外一種領域專家：『標記資料庫』

海沃德 - Google地圖 × G imagenet - Google 搜尋 × ImageNet Tree View × About WordNet - Wor × (109) ViTBAT - Video T × kai

image-net.org/explore?wnid=n02512053

Fish

Any of various mostly cold-blooded aquatic vertebrates usually having scales and breathing through gills; "the shark is a large fish"; "in the living room there was a tank of colorful fish"

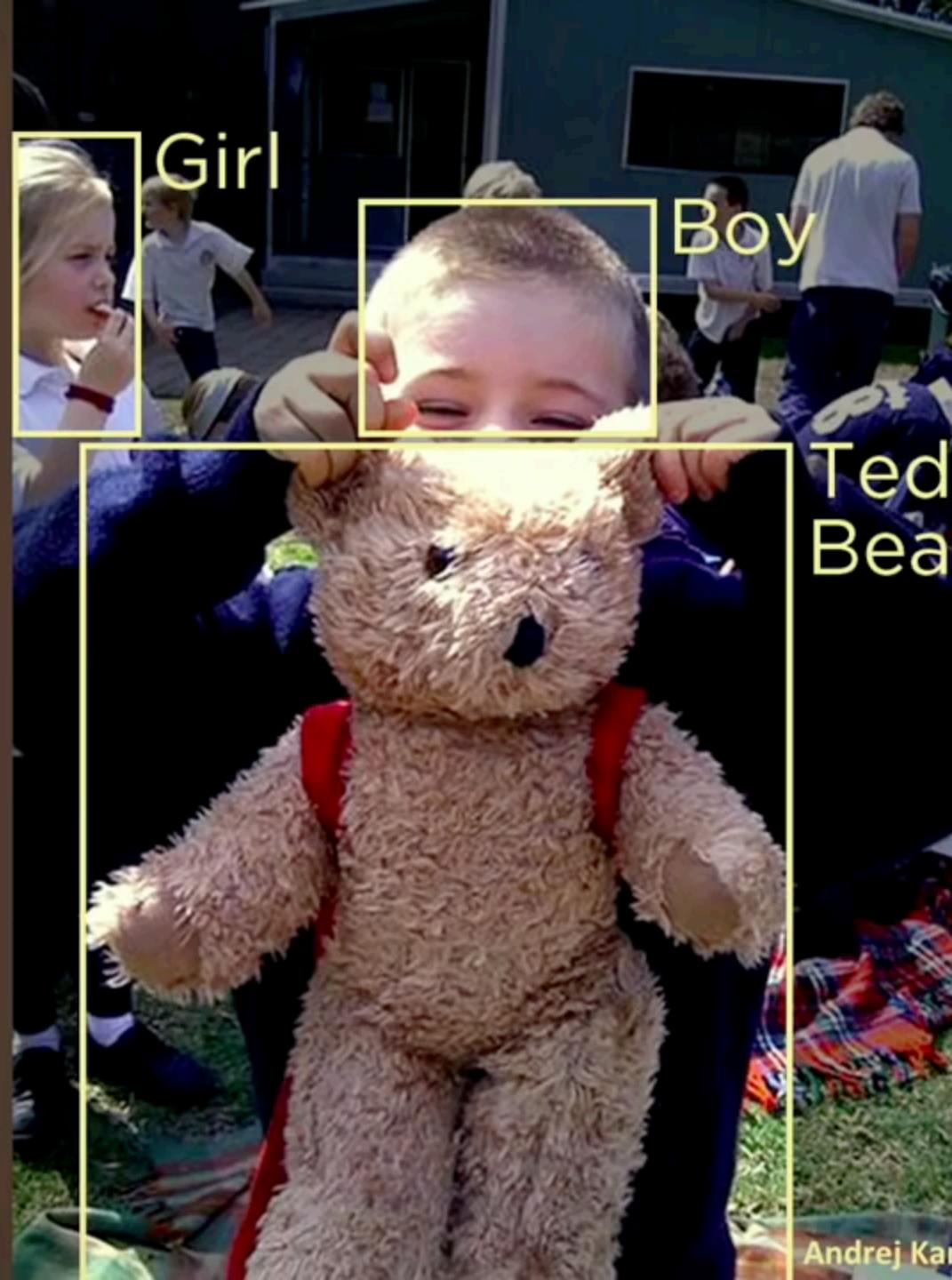
1307 pictures 91.33% Popularity Percentile Wordnet IDs

- plant, flora, plant life (4486)
- geological formation, formation (175)
- natural object (1112)
- sport, athletics (176)
- artifact, artefact (10504)
- fungus (308)
- person, individual, someone, somet
- animal, animate being, beast, brute, invertebrate (766)
- homeotherm, homolotherm, hom
- work animal (4)
- darter (0)
- survivor (0)
- range animal (0)
- creepy-crawly (0)
- domestic animal, domesticated a
- molter, moulter (0)
- varmint, varment (0)
- mutant (0)
- critter (0)
- game (47)
- young, offspring (45)
- poikilotherm, ectotherm (0)

Treemap Visualization Images of the Synset Downloads

IMAGENET

接近 1500 萬
張標記圖片



<https://www.youtube.com/watch?v=40riCqvRoMs>

Andrej Karpathy, Armand Joulin, & Li Fei-Fei, NIPS, 2014

Vertical

All

Filter

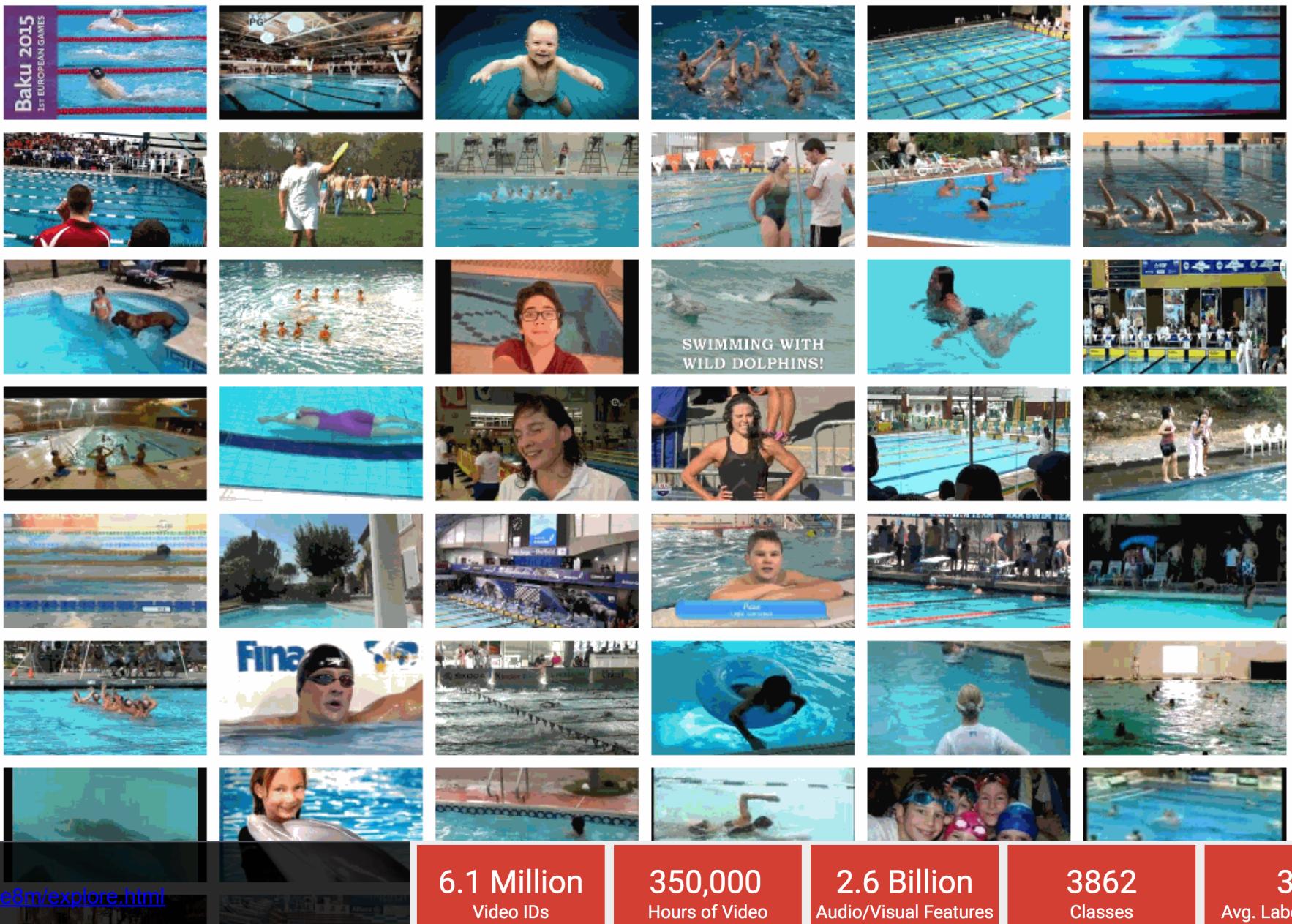
human

Entities

Human swimming (4291)

Human hair color (503)

Humanoid robot (139)



9

Your query "semiconduct" matches 0 synsets.

Are you looking for [semi conduct](#), [semi-conduct](#)?

© 2016 Stanford Vision Lab, Stanford University, Princeton University support@image-net.org Copyright infringement

YouTube | 8M

Vertical

All

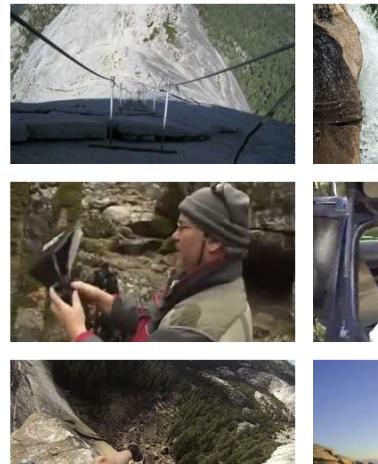
Filter

semi

Entities

Semi-trailer (192)

Yosemite National Park (165)



**許多領域的 AI，必須遵從以下
步驟逐步展開**

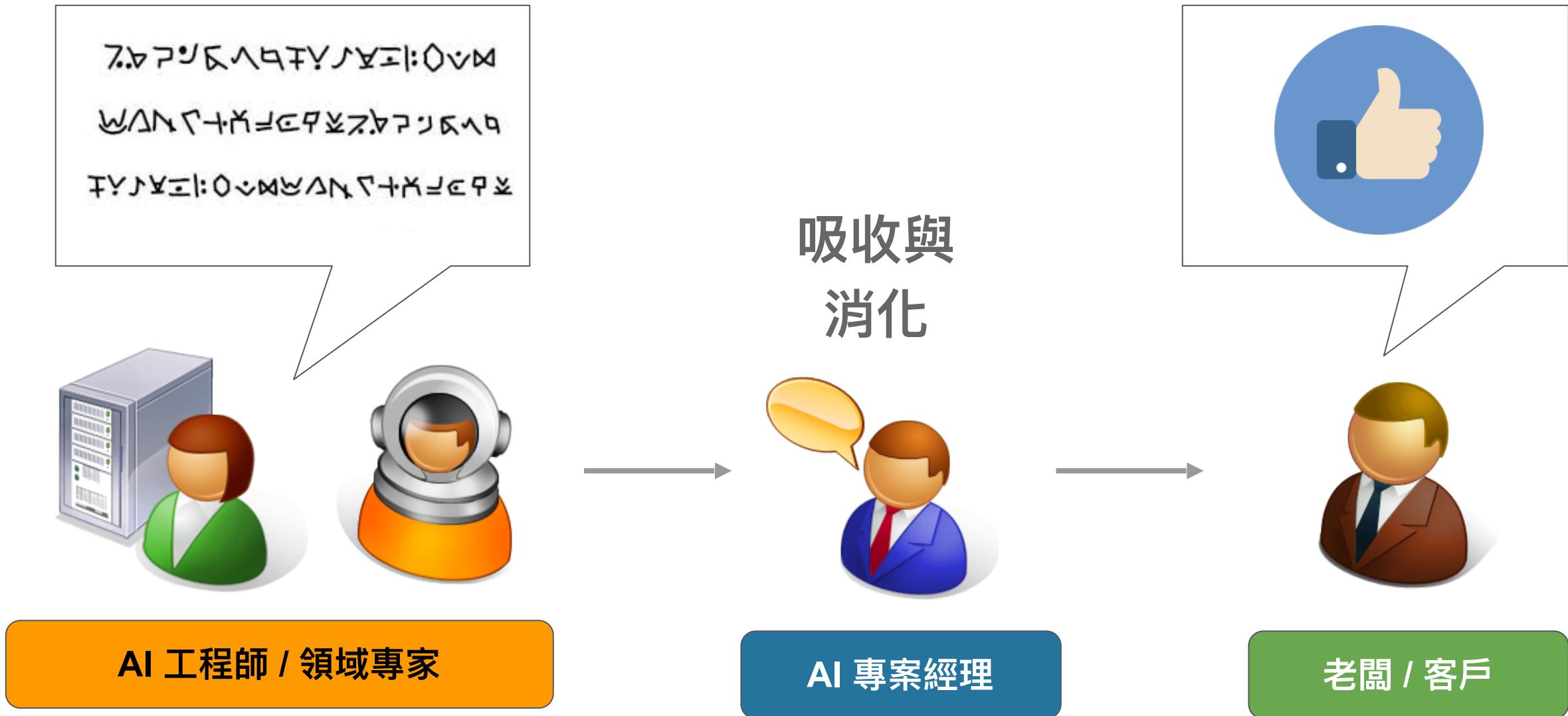
- 收集資料開始
- 彙整領域專家經驗
- 進行資料標記



#3 紀錄能力

Taking Note > Learning

專案進行過程，會有許多專業術語





Question

然而，資訊量通常很龐大又
精深，怎樣有效紀錄呢？

Record Everything / Record Detail



將討論內容可視化，有利後續的討論



**Realtime
Board**

<https://miro.com/>

#1 跑Graphlab 的四個 regression 方法	#2 之後再跑回歸 並加入更多的特徵 可以從看起來可疑的 feature 著手加入 ex. 晴天、雨天、秋天	# 3 如何找出最佳的 正規化組合
在 T+1 情境 scikitlearn + 神經網路最佳	#1 將 coefficient 匯出 excel 格式	# 2 實驗：Time Window 的切法 維持固定數量



Me

氣象領
域專家

閱讀工程師研發手稿 & 程式碼，輔助閱讀 / 邏輯確認

jupyter spectrogram (autosaved)

File Edit View Insert Cell Kernel Help Python 3

CellToolbar

Simple spectral analysis

An illustration of the [Discrete Fourier Transform](#)

$$X_k = \sum_{n=0}^{N-1} x_n \exp \frac{-2\pi i}{N} kn \quad k = 0, \dots, N-1$$

```
In [2]: from scipy.io import wavfile  
rate, x = wavfile.read('test_mono.wav')
```

And we can easily view its spectral structure using matplotlib's builtin specgram routine:

```
In [5]: fig, (ax1, ax2) = plt.subplots(1,2,figsize(16,5))  
ax1.plot(x); ax1.set_title('Raw audio signal')  
ax2.specgram(x); ax2.set_title('Spectrogram');
```

The Jupyter Notebook interface shows two plots. On the left, a line plot titled "Raw audio signal" displays a fluctuating blue waveform. On the right, a spectrogram titled "Spectrogram" is shown as a color heatmap where horizontal axes represent time and vertical axes represent frequency, with a color gradient from red at the bottom to blue at the top.





#4 視覺能力

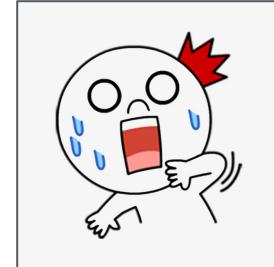
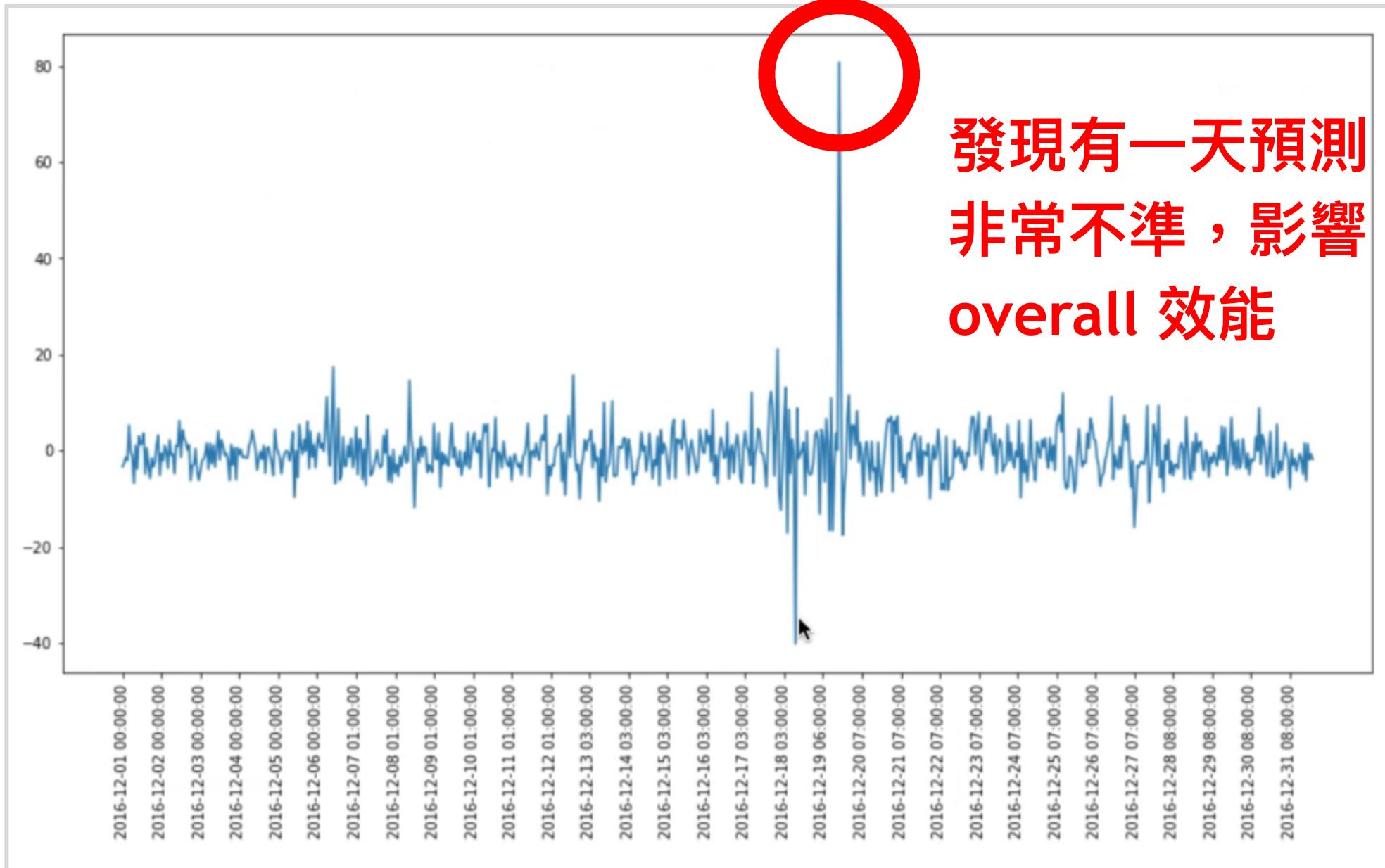
Visual Skill



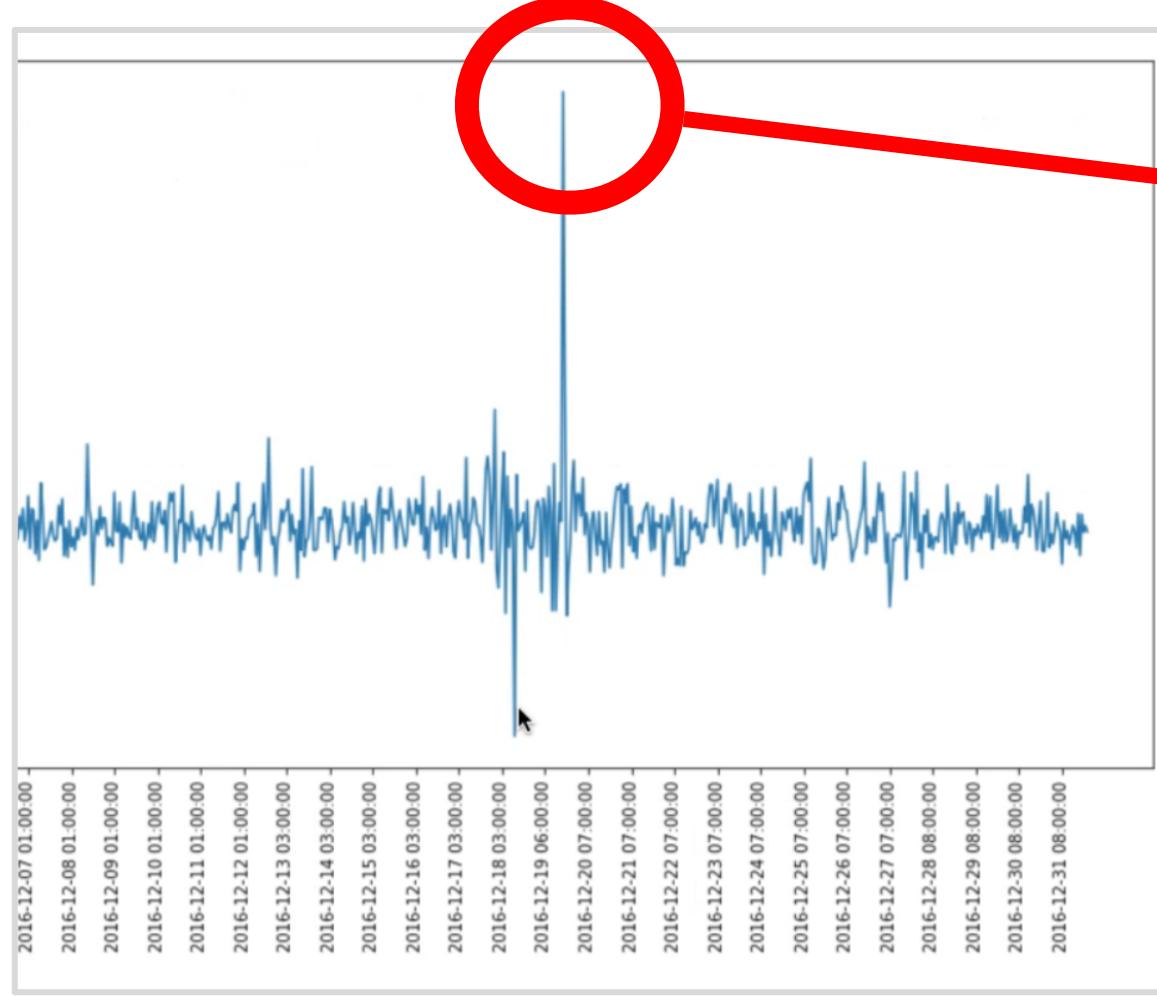
Question

AI 不是通常都是運算結果，
會跟視覺有關係嗎？

將 AI 產出視覺：可幫助我們洞察數據狀態



將 AI 產出視覺：可幫助我們洞察額外的數據狀態



- 2016/6/30受稻草燃燒影響空氣品質
- $\text{PM}_{2.5}$ 占 PM_{10} 約80%

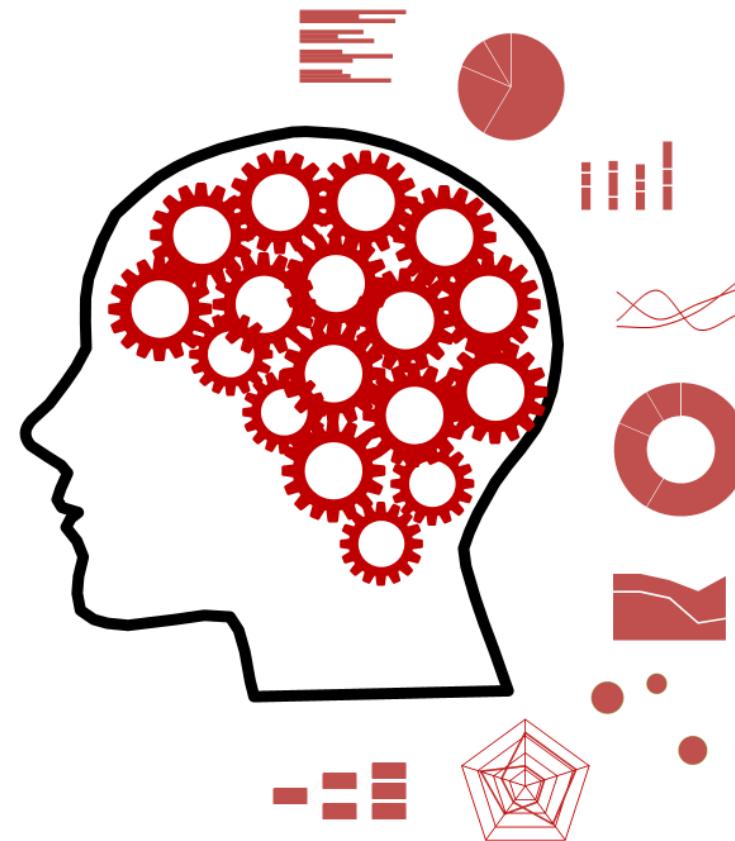
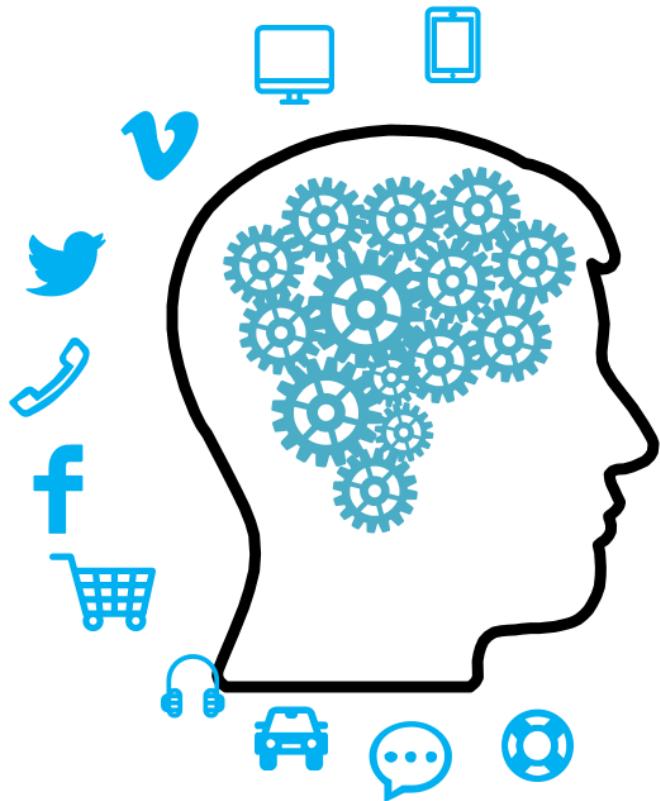


燒稻草

火災

視覺：可幫助更清楚呈現 AI 效果

【資料】
【分數】
【數據】



清楚
視覺





專案經理，如果能自己產
出 AI 結果的視覺化，能
夠對 AI 解析更有主導性

通常工程師的產出：是序列化的數值 (1/2)

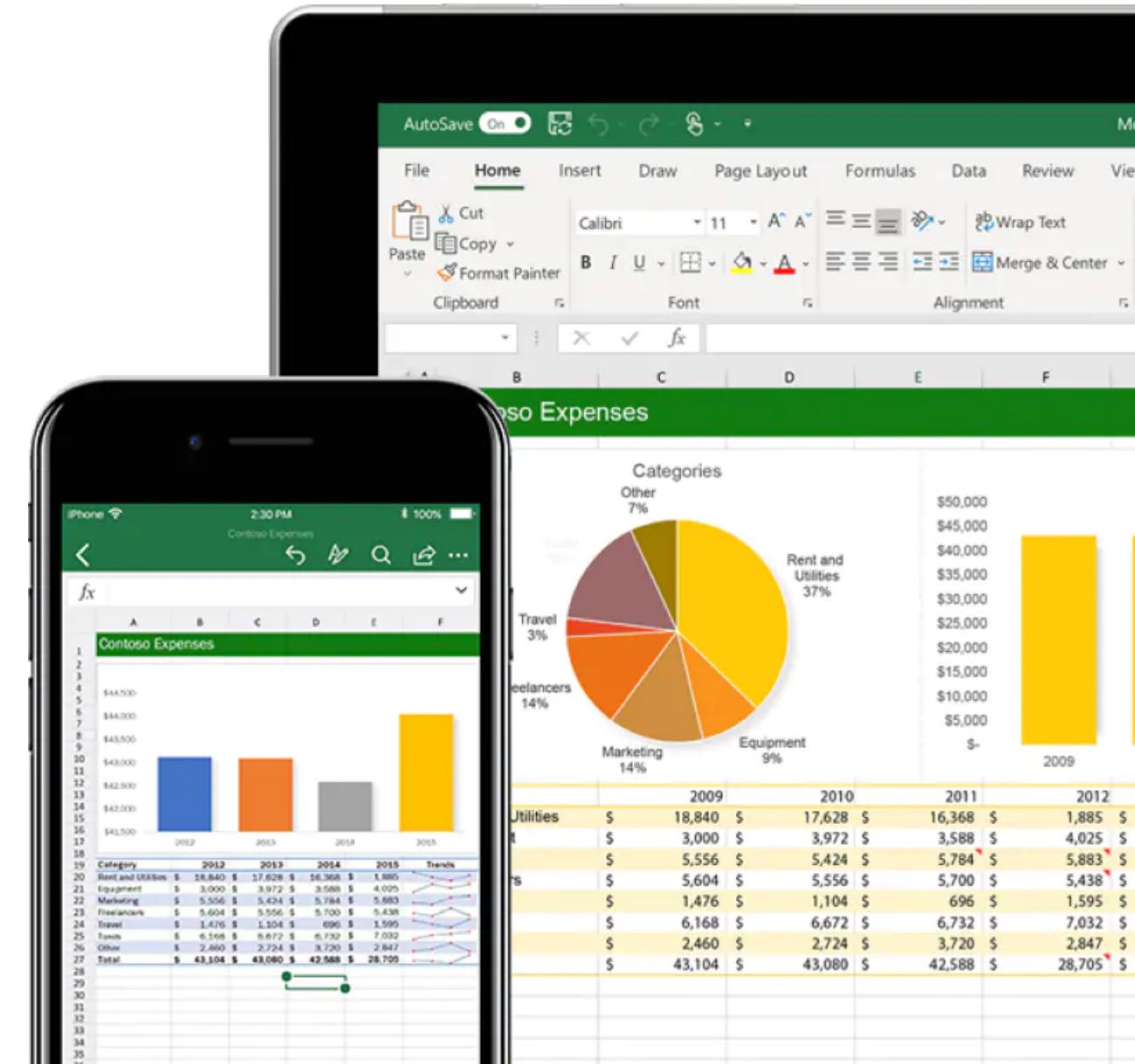
datetime	station	aqi_12hr_4hr_avg_1	aqi_12hr_4hr_avg_2	aqi_12hr_4hr_avg_3	aqi_12hr_4hr_avg_4	aqi_12hr_4hr_avg_5	aqi_12hr_4hr_avg_6
2018-01-01 00:00:00	三義	85	80	77	74	73	72
2018-01-01 00:00:00	中山	66	61	58	57	58	59
2018-01-01 00:00:00	林口	80	72	67	64	62	61
2018-01-01 00:00:00	桃園	83	74	69	66	64	63
2018-01-01 00:00:00	汐止	84	78	74	71	68	69
2018-01-01 00:00:00	沙鹿	87	79	73	69	69	70
2018-01-01 00:00:00	淡水	71	67	65	64	63	64
2018-01-01 00:00:00	湖口	85	80	76	74	72	70
2018-01-01 00:00:00	竹東	82	73	69	67	67	68
2018-01-01 00:00:00	線西	117	109	101	95	92	90
2018-01-01 00:00:00	臺西	124	112	105	99	95	92
2018-01-01 00:00:00	苗栗	97	89	86	84	83	84
2018-01-01 00:00:00	菜寮	68	65	64	63	61	60
2018-01-01 00:00:00	萬華	74	70	64	60	58	57
2018-01-01 00:00:00	萬里	90	84	79	76	75	75
2018-01-01 00:00:00	西屯	105	93	84	77	72	72
2018-01-01 00:00:00	觀音	107	100	95	93	93	93
2018-01-01 00:00:00	豐原	73	66	62	56	53	53
2018-01-01 00:00:00	頭份	92	86	82	81	82	81

通常工程師的產出：是序列化的數值 (2/2)

空氣品質：預測 17:00 之後的數值

```
{"station": "\u53e4\u4ead", "datetime": "2018-09-20 16:00:00", "pred_1": 39.72, "pred_2": 42.9, "pred_3": 40.14, "pred_4": 39.45, "pred_5": 39.42, "pred_6": 35.49, "pred_7": 36.38, "pred_8": 34.85, "pred_9": 31.99, "pred_10": 30.96, "pred_11": 27.22, "pred_12": 28.12, "pred_13": 25.57, "aqi_12hr_4hr_avg_1": 112.0, "aqi_12hr_4hr_avg_2": 112.0, "aqi_12hr_4hr_avg_3": 103.0, "aqi_12hr_4hr_avg_4": 107.0, "aqi_12hr_4hr_avg_5": 110.0, "aqi_12hr_4hr_avg_6": 110.0, "aqi_12hr_4hr_avg_7": 110.0, "aqi_12hr_4hr_avg_8": 109.0, "aqi_12hr_4hr_avg_9": 105.0, "aqi_12hr_4hr_avg_10": 102.0, "aqi_12hr_4hr_avg_11": 96.0, "aqi_12hr_4hr_avg_12": 93.0, "aqi_12hr_4hr_avg_13": 90.0, "error_bound_t4": 8.54762, "error_bound_t7": 11.756, "error_bound_t8": 12.9405, "error_bound_t12": 16.0833, "error_bound_t6": 10.7738, "error_bound_t9": 13.6905, "error_bound_t5": 9.625, "error_bound_t13": 16.6964, "error_bound_t10": 14.5774, "error_bound_t1": 0.434524, "error_bound_t3": 6.36905, "error_bound_t2": 1.125, "error_bound_t11": 15.3095, "pollute_6": "PM2.5", "pollute_10": "PM2.5", "pollute_12": "PM2.5", "pollute_11": "PM2.5", "pollute_3": "PM2.5", "pollute_4": "PM2.5", "pollute_1": "O3", "pollute_7": "PM2.5", "pollute_9": "PM2.5", "pollute_2": "O3", "pollute_5": "PM2.5", "pollute_8": "PM2.5", "pollute_13": "PM2.5"}"
```

透過 Excel / Tableau 將數據產出視覺化





tableau

Connect

To a File

Excel

Text File

Access

Statistical File

Other files

To a Server

Tableau Server

Microsoft SQL Server

|Search

Tableau Server

Actian Matrix

Actian Vector

Amazon Aurora

Amazon EMR

Amazon Redshift

Aster Database

Cisco Information Server

Cloudera Hadoop

DataStax Enterprise

內部系統：透過 Tableau 每月持續觀察預測誤差

**T+1 小時
AQI 誤差
2/15-2/20**

空品-真實值

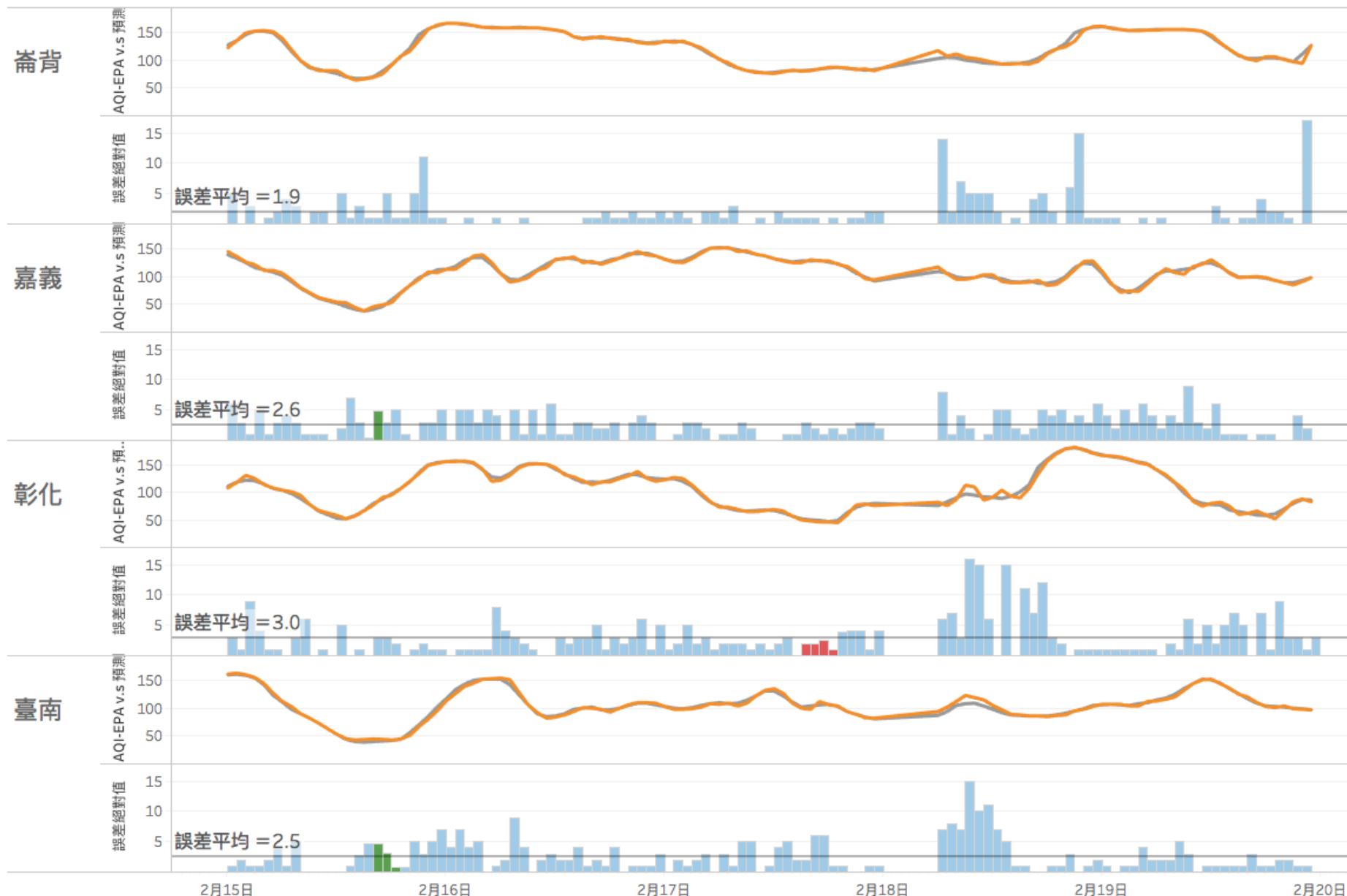
空品-預測值

AQI 指標物

O3

PM2.5

PM10





#5 沟通能力

Communication

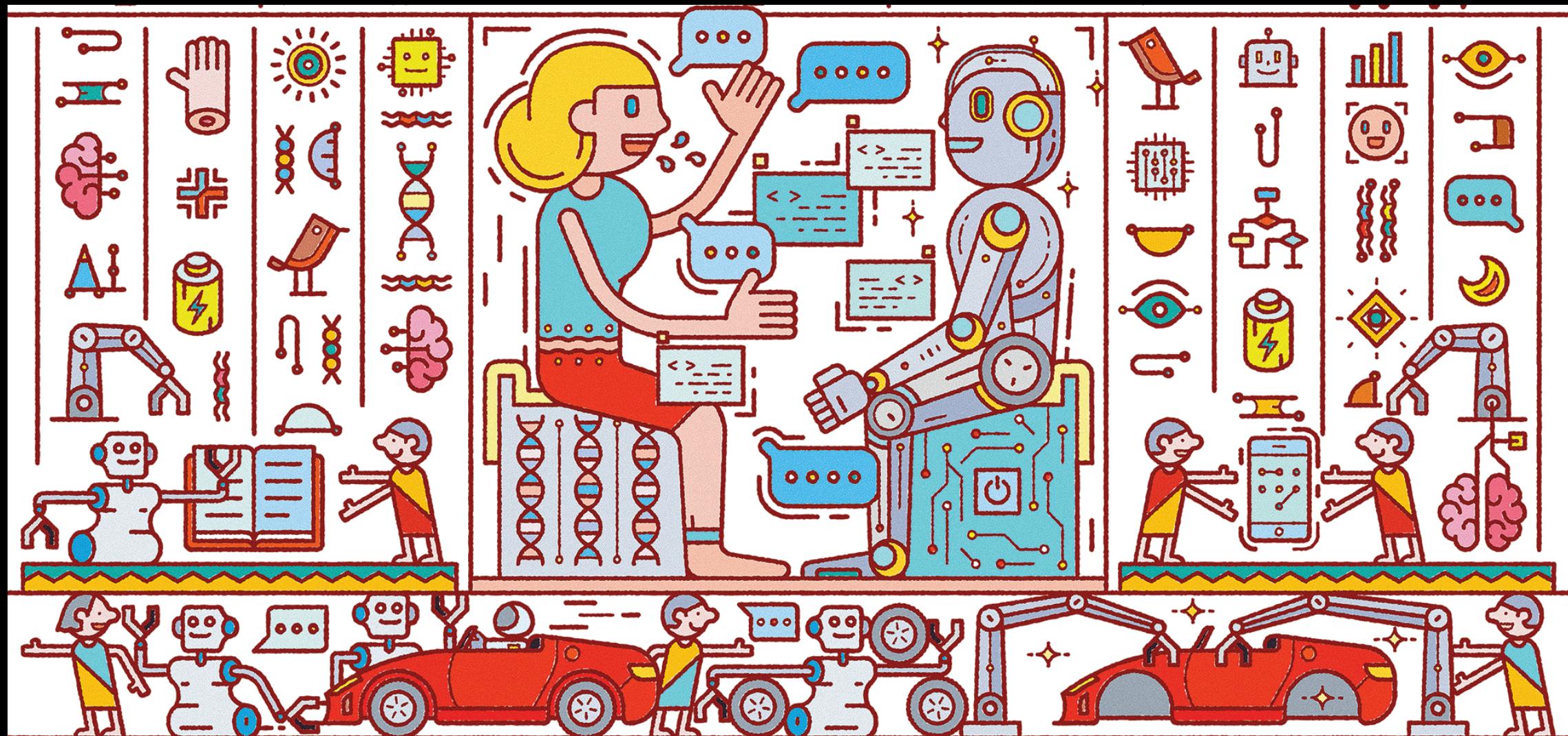
領域術語

演算法術語

商業術語

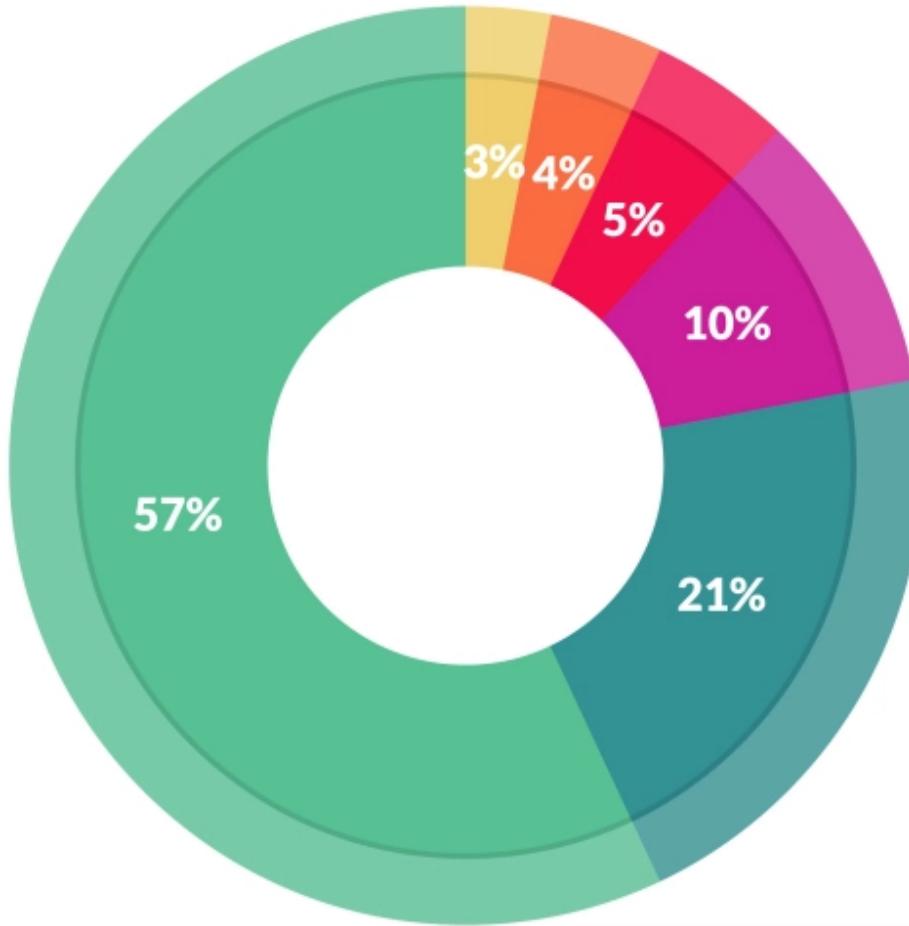
資料術語

使用者術語



圖來源：<https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces>

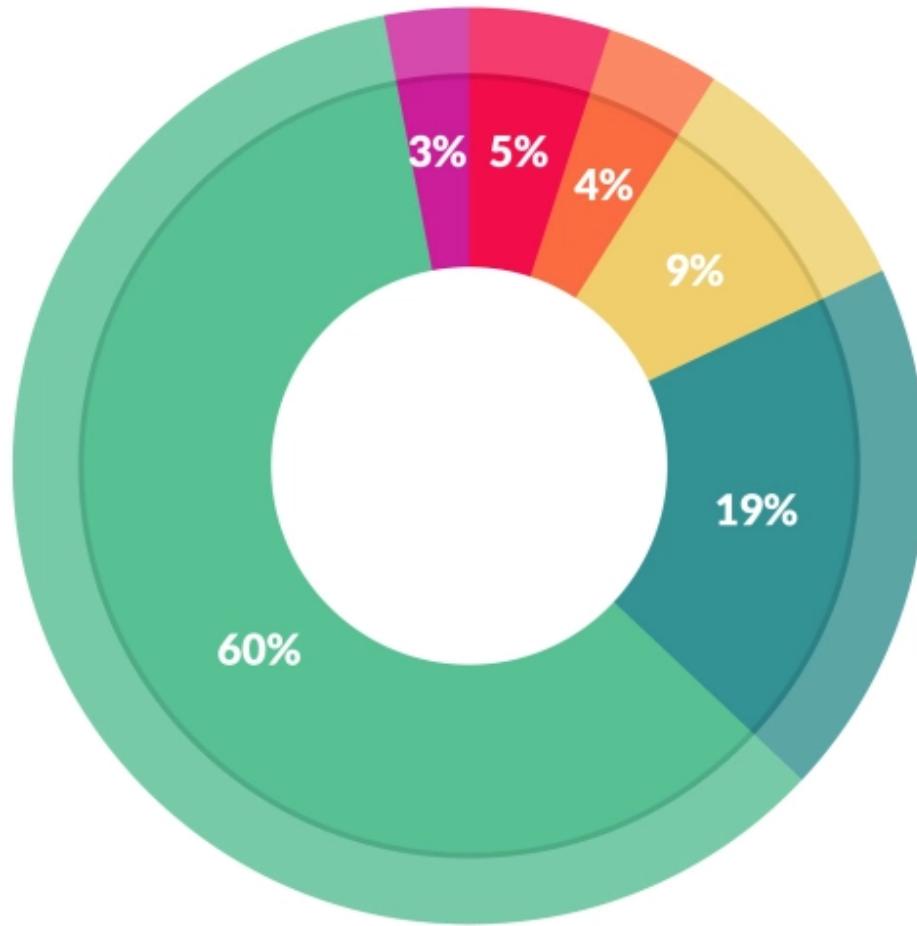
資料科學家/工程師 最不喜歡 的事情：清洗資料



What's the least enjoyable part of data science?

- Building training sets: 10%
- Cleaning and organizing data: 57%
- Collecting data sets: 21%
- Mining data for patterns: 3%
- Refining algorithms: 4%
- Other: 5%

資料科學家/工程師 最花時間 的事情：清洗資料



What data scientists spend the most time doing

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19% -----
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%



KEEP
CALM
AND
LOVE AN
ENGINEER

專案經理，需要持續透過
溝通推動團隊鬥志



清洗資料



清洗資料



清洗資料



設計AI演算法



訓練AI演算法



觀察結果



清洗資料



清洗資料



清洗資料



結論





AI 專案經理的五大挑戰





”

people need the ability to make sense of information, to tell the difference between what is important and what is unimportant



THE END