

Introduction to AirBox

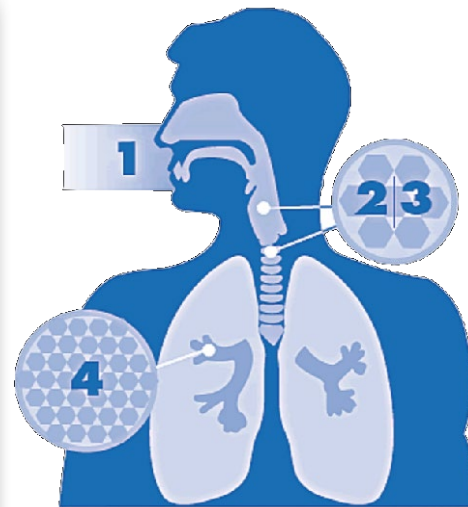
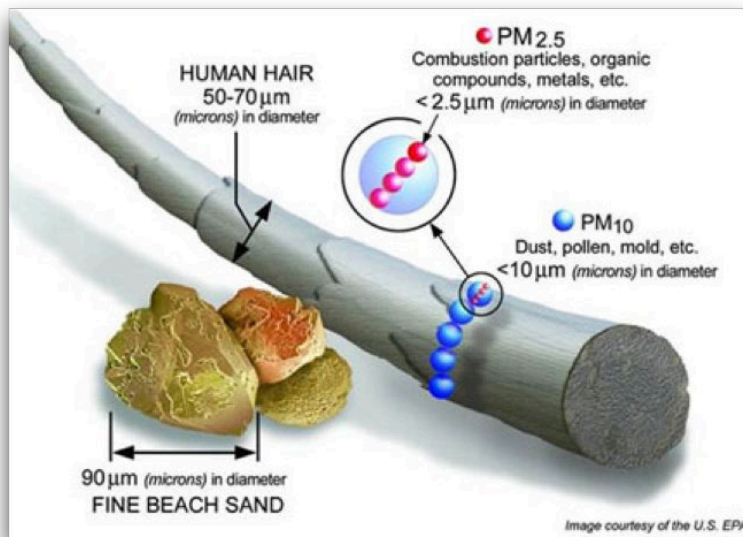
Ling-Jyh Chen

2023/7/19



3 million deaths
attributable to ambient air pollution

4.3 million deaths
attributable to household air pollution

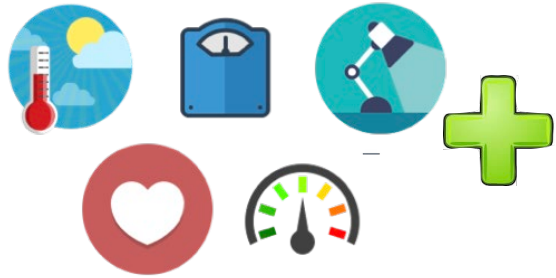


現在 & 這裡的
空氣好不好？

微型空氣品質感測計畫

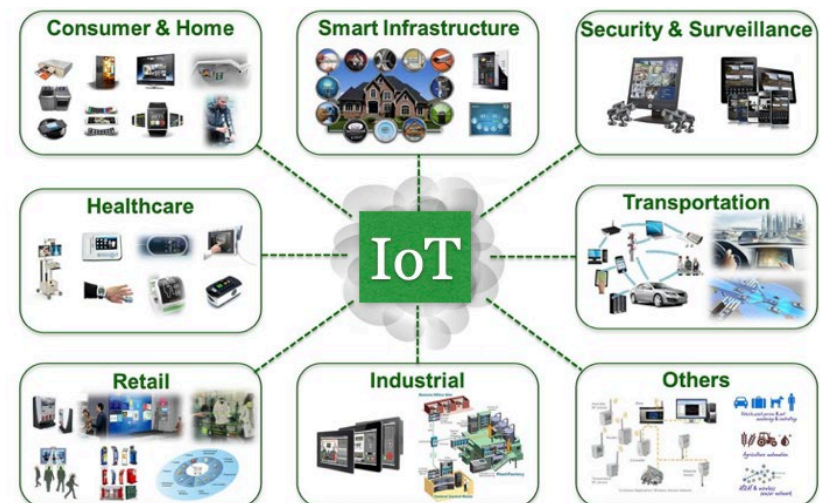
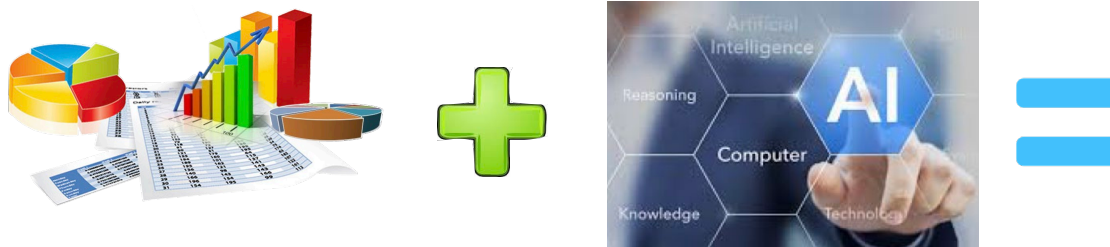
微型感測器

● 微型感測器 ⇒ 大數據 + 物聯網



微型感測器

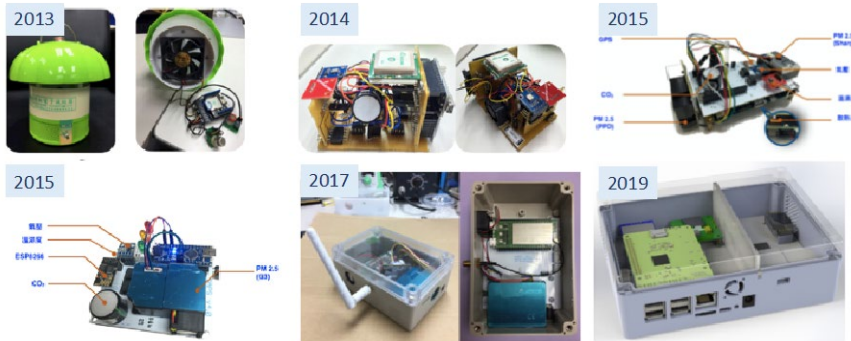
- 微型感測器 ⇒ 大數據、人工智慧、物聯網



Vivante and the Vivante logo are trademarks of Vivante Corporation. All other product, image or service names in this presentation are the property of their respective owners. ©2013 Vivante Corporation

Enable

Investigate low-cost sensors for air quality monitoring

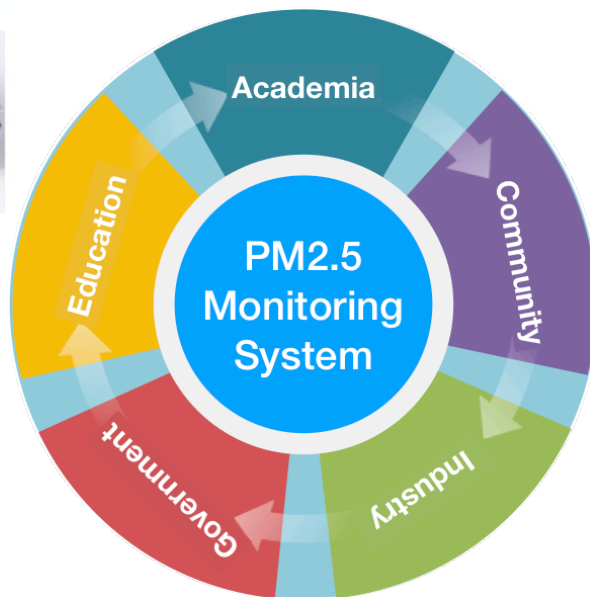


Engage

Encourage people to participate, and contribute

Empower

Reduce the technical boundary by open hardware/software/data



Energize

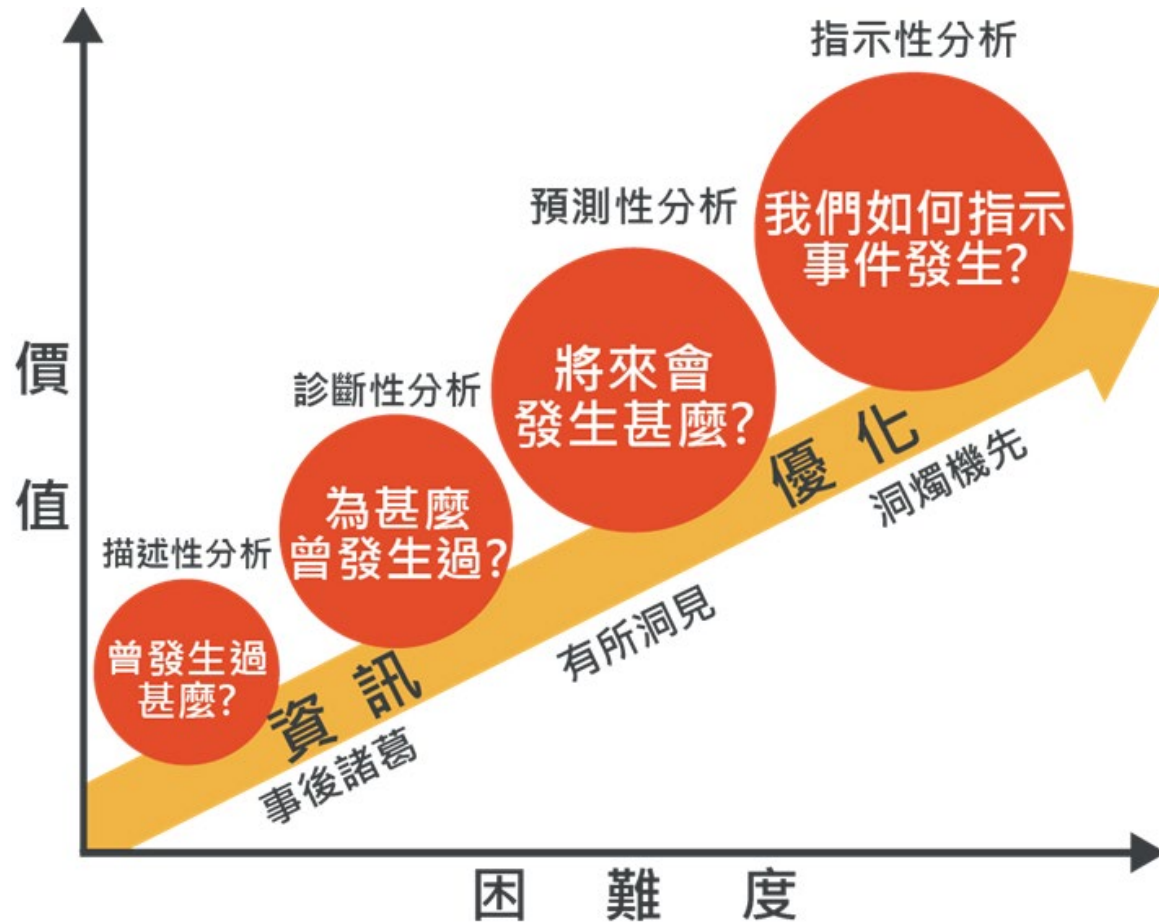
Promote environmental awareness, partnership, and collaboration



59 Countries 20,000+ Devices



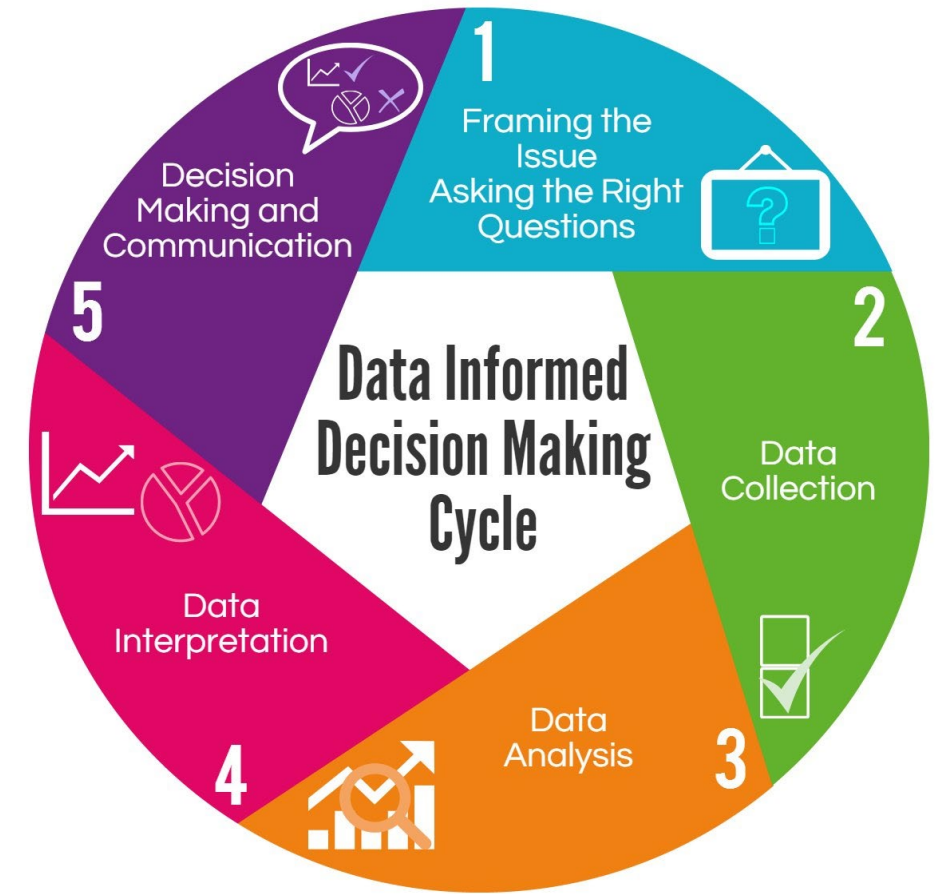
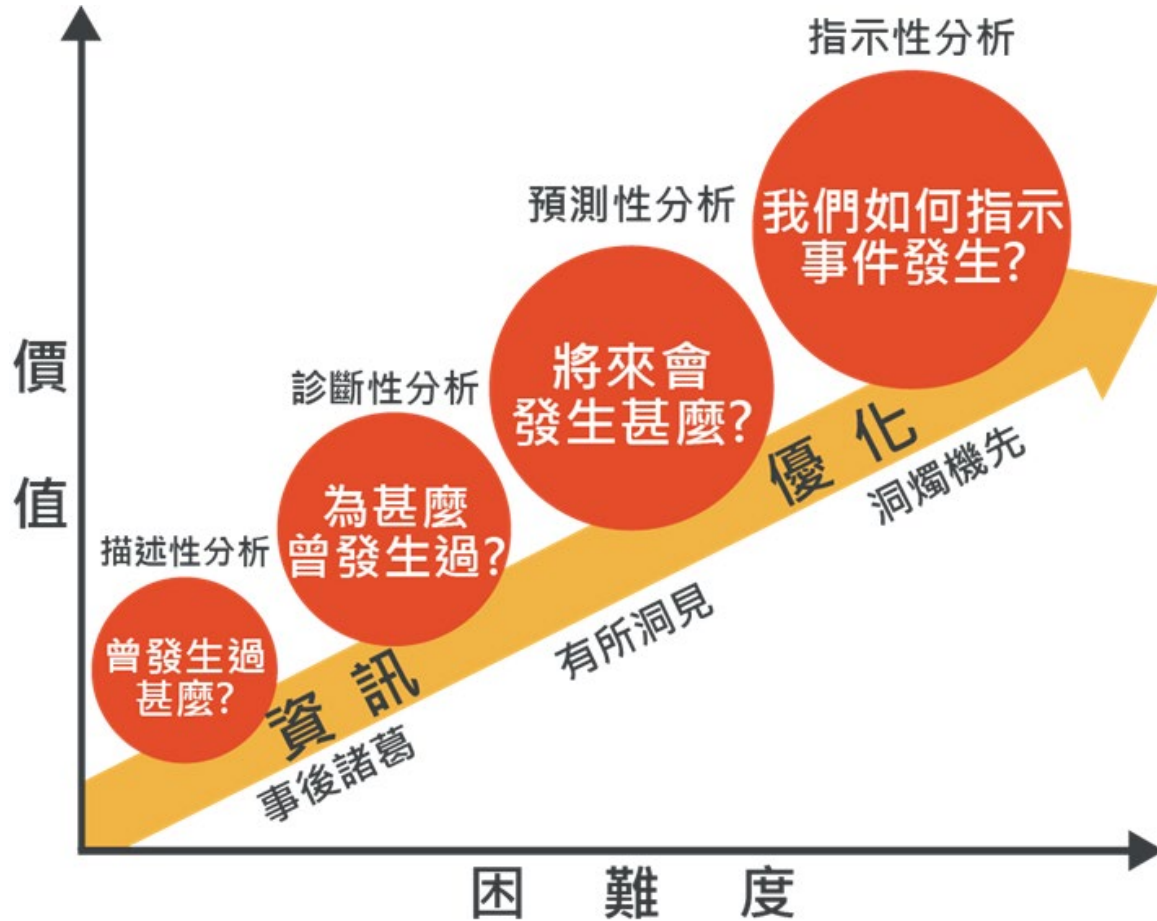

從 微型感測 到 大數據



<https://medium.com/marketingdatascience/資料科學進化論-五種分析方式-types-of-business-analytics-97fe78938769>



從 微型感測 到 大數據 到 資料驅動決策

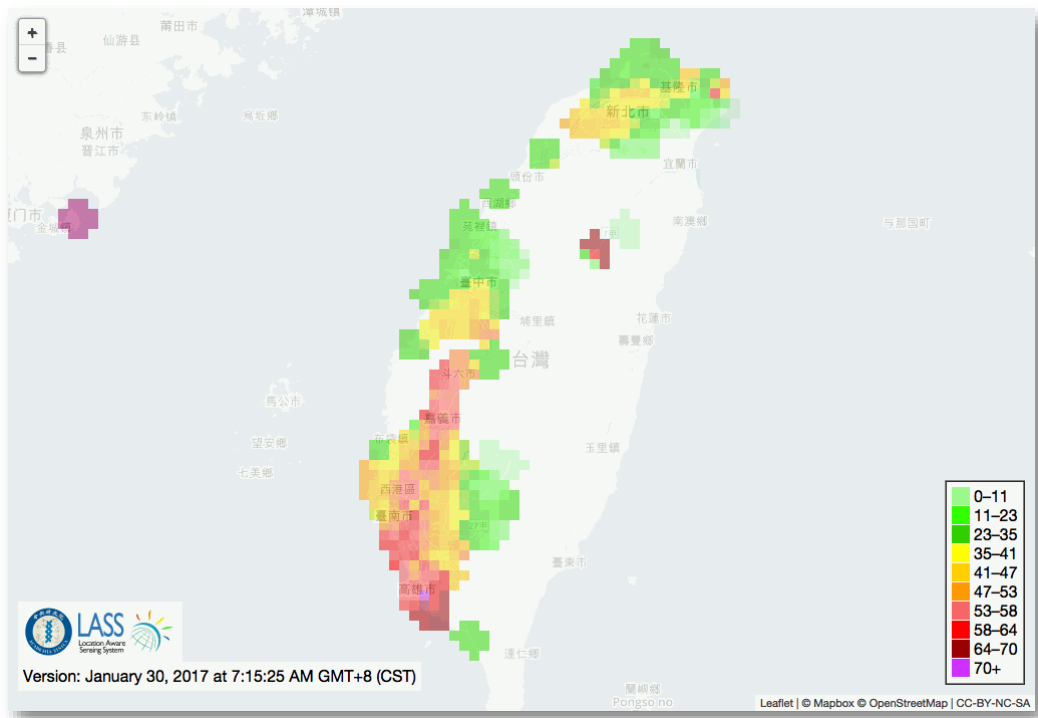


<https://medium.com/marketingdatascience/資料科學進化論-五種分析方式-types-of-business-analytics-97fe78938769>

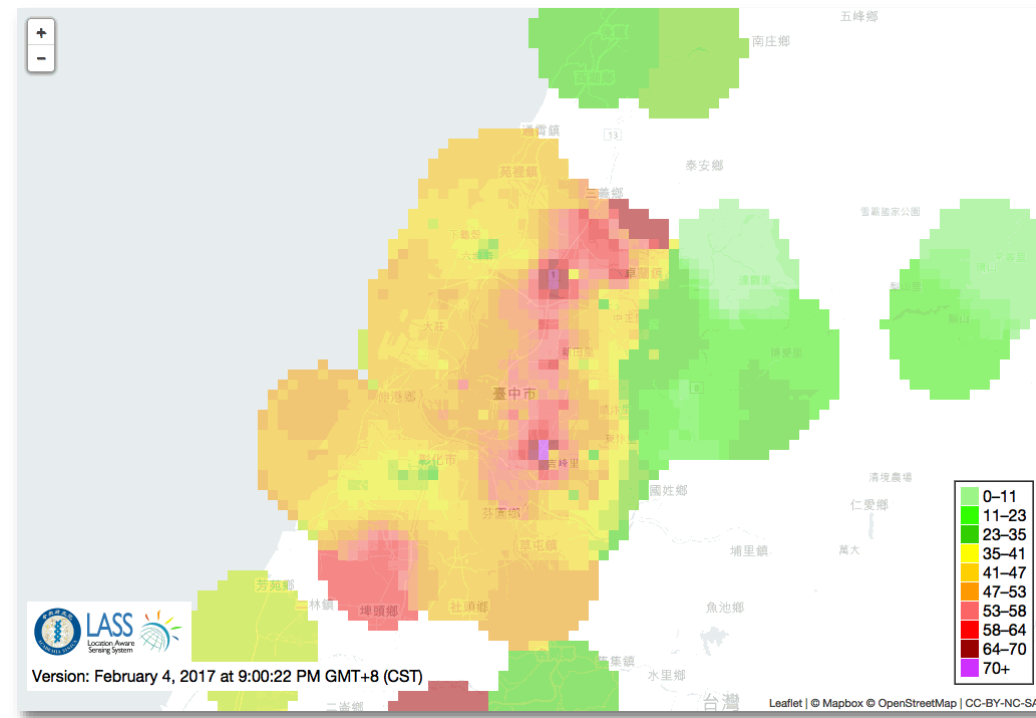
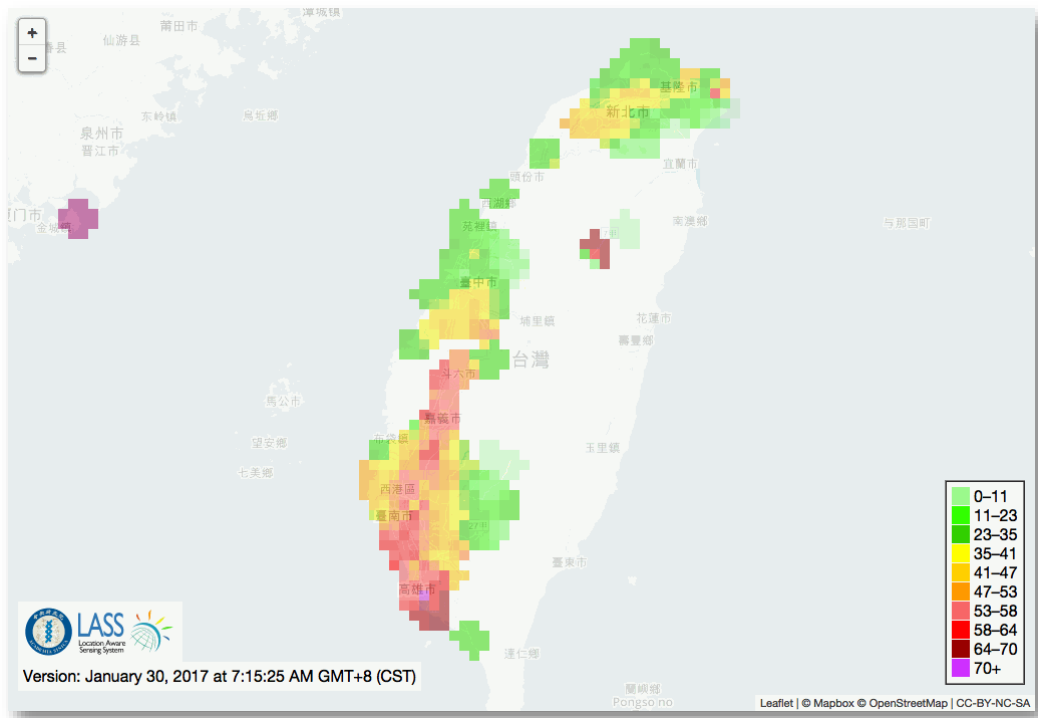
http://www.rosebt.com/uploads/8/1/8/1/8181762/6779091_orig.jpg



描述性分析 Descriptive Analysis

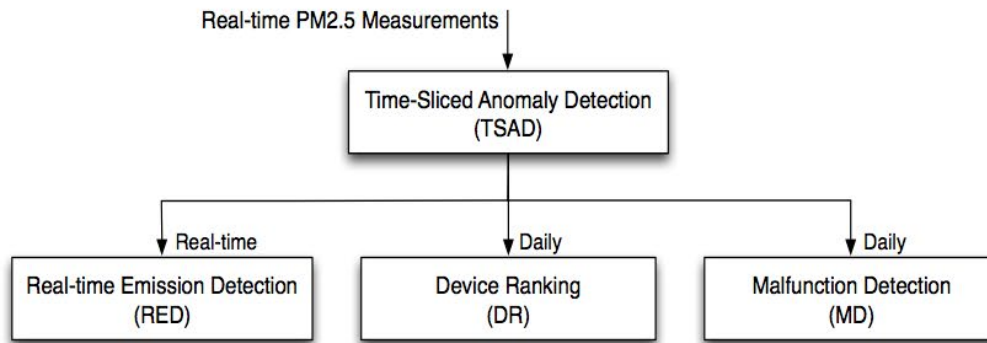


描述性分析 Descriptive Analysis

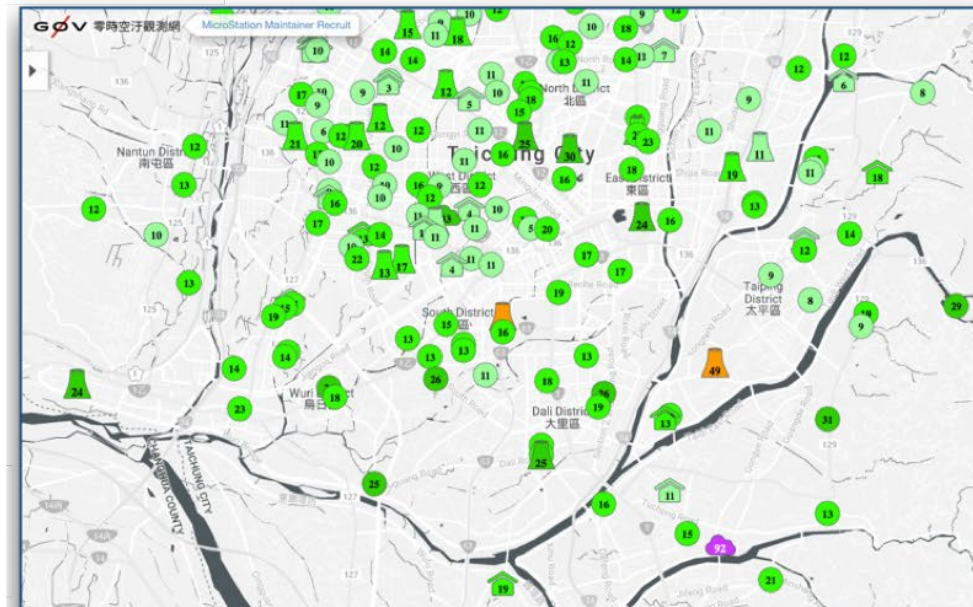


描述性分析 Descriptive Analysis

診斷性分析 Diagnostic Analysis



| | Site | Site ID | Ranking | Survey | Pollution ongoing | Attribute no comment | Attribute indoor | Attribute close to emission | App Version |
|-----|---------|--------------|---------|-------------------|-------------------|----------------------|------------------|-----------------------------|-------------|
| 46 | 台中市翕子國小 | 74DA3895C304 | ★★★★★ | N/A | none | | | maybe (0.352) | N/A |
| 232 | 台中市七星國小 | 74DA3895C384 | ★★★★★ | N/A | none | | | | 0.35.0 |
| 231 | 台中市三光國小 | 74DA3895C420 | ★★★★★ | N/A | none | | likely (0.614) | | 0.35.0 |
| 230 | 台中市三和國小 | 74DA3895C522 | N/A | N/A | none | | | | N/A |
| 229 | 台中市三田國小 | 74DA3895C4EE | ★★★★★ | N/A | none | | | | 0.35.0 |
| 228 | 台中市上安國小 | 74DA3895C276 | ★★★★★ | done (2016-12-03) | none | | | | 0.35.0 |
| 227 | 台中市上楓國小 | 74DA3895C51A | ★★★★★ | N/A | none | | | | 0.35.0 |

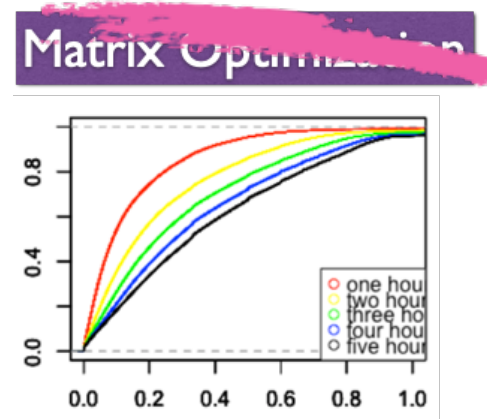


描述性分析
Descriptive Analysis

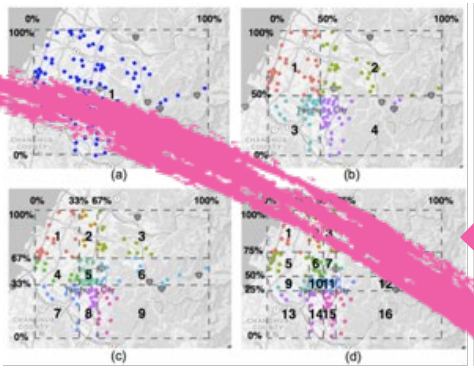
診斷性分析
Diagnostic Analysis

預測性分析
Predictive Analysis

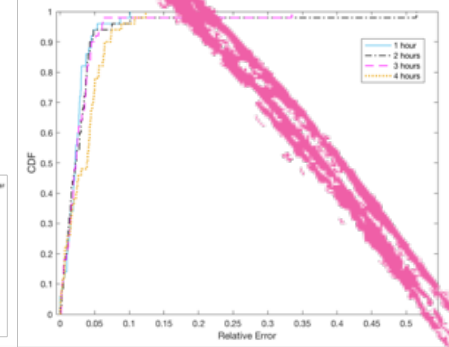
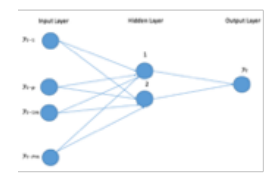
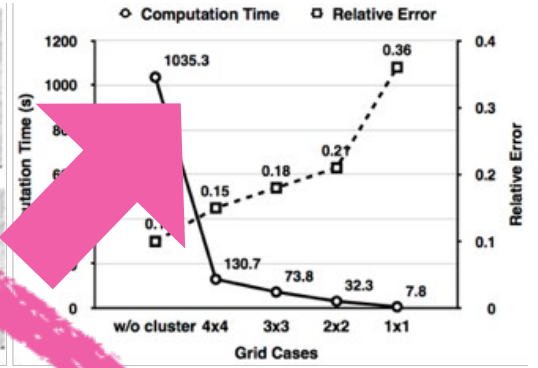
Scalability



$$\min(\lambda * \|M\|_F^2 + \sum_{i=1}^t \|N_{t+i+1} - M * N_{t+i}\|_F^2)$$



Cluster-based Hybrid



Hybrid (NNAR + ARIMA)

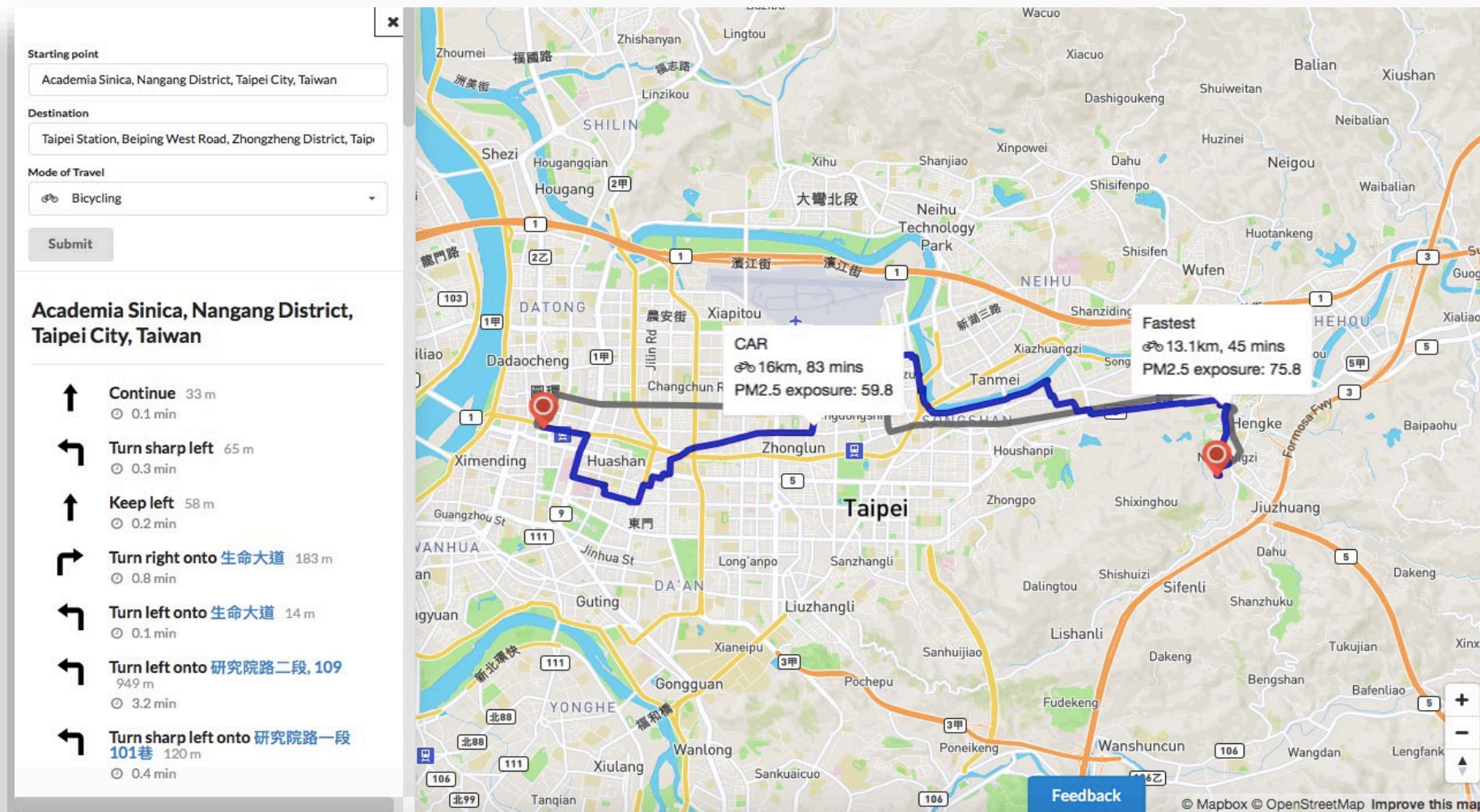
Accuracy

描述性分析
Descriptive Analysis

診斷性分析
Diagnostic Analysis

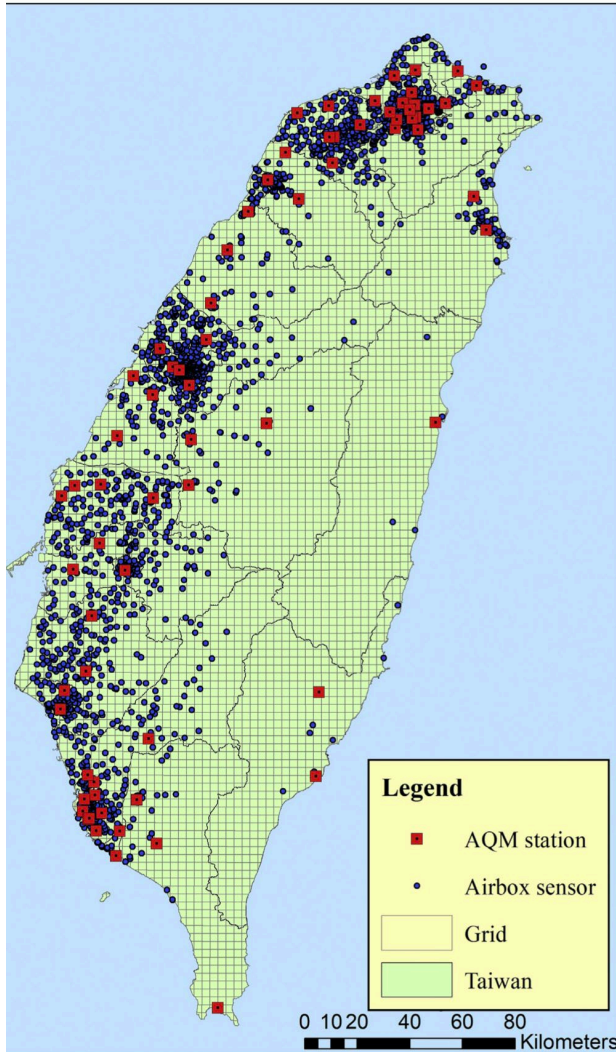
預測性分析
Predictive Analysis

指示性分析
Prescriptive Analysis



Project Results: Fine-grained Ground-level Estimation

Environmental Pollution '20



Data source:

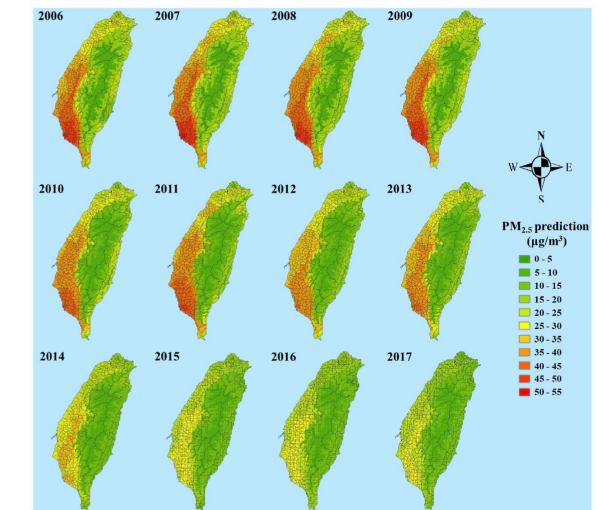
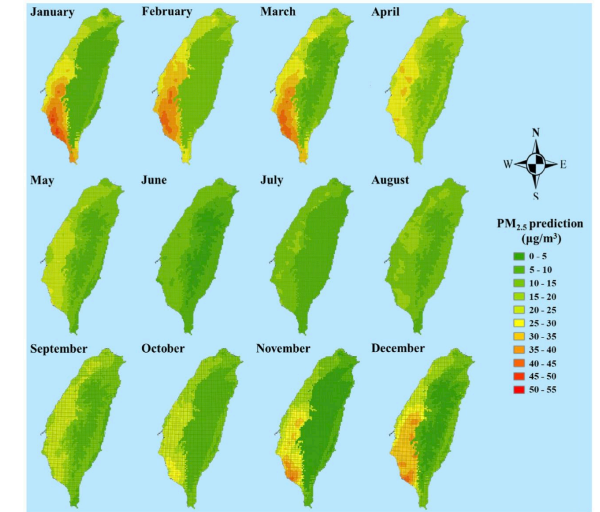
- **E**PA station PM2.5
- **A**irBox PM2.5
- **M**eteorological data: temperature, relative humidity, wind speed, rainfall
- **S**patial predictors: land use, traffic, emission sources, population, elevation

State-of-the-art: consider **E / M / S**

$$R^2 = 0.66 \sim 0.87; \text{RMSE} = 5.0 \sim 12.9 \mu\text{g}/\text{m}^3$$

Our work: consider **E / A / M / S**

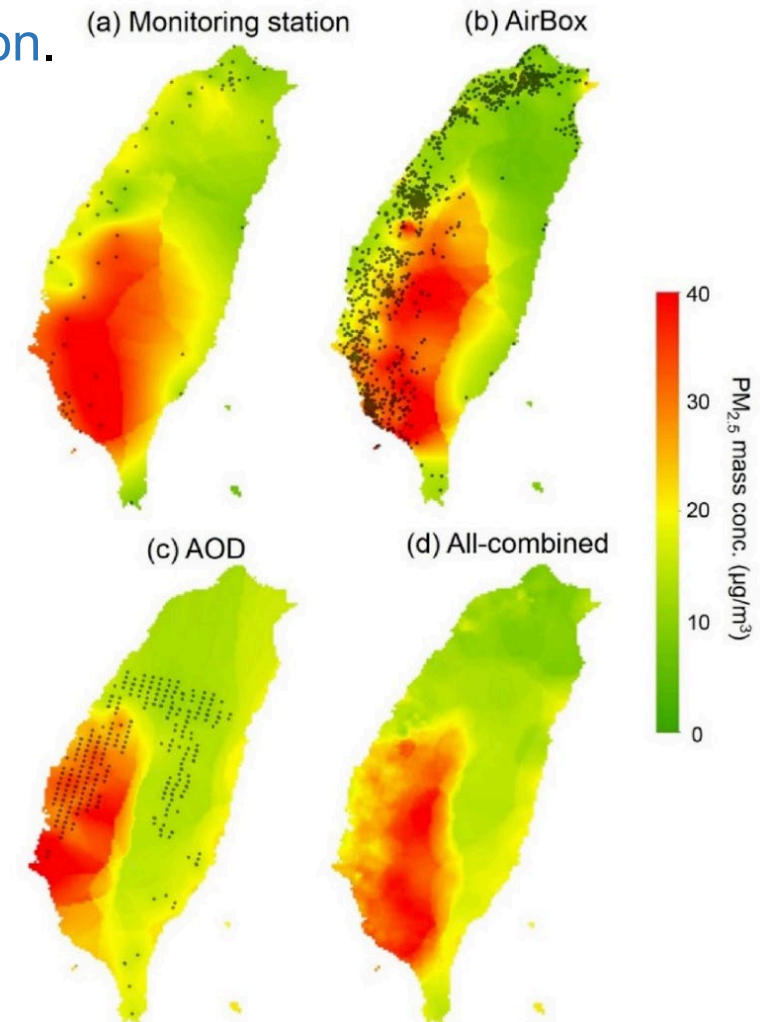
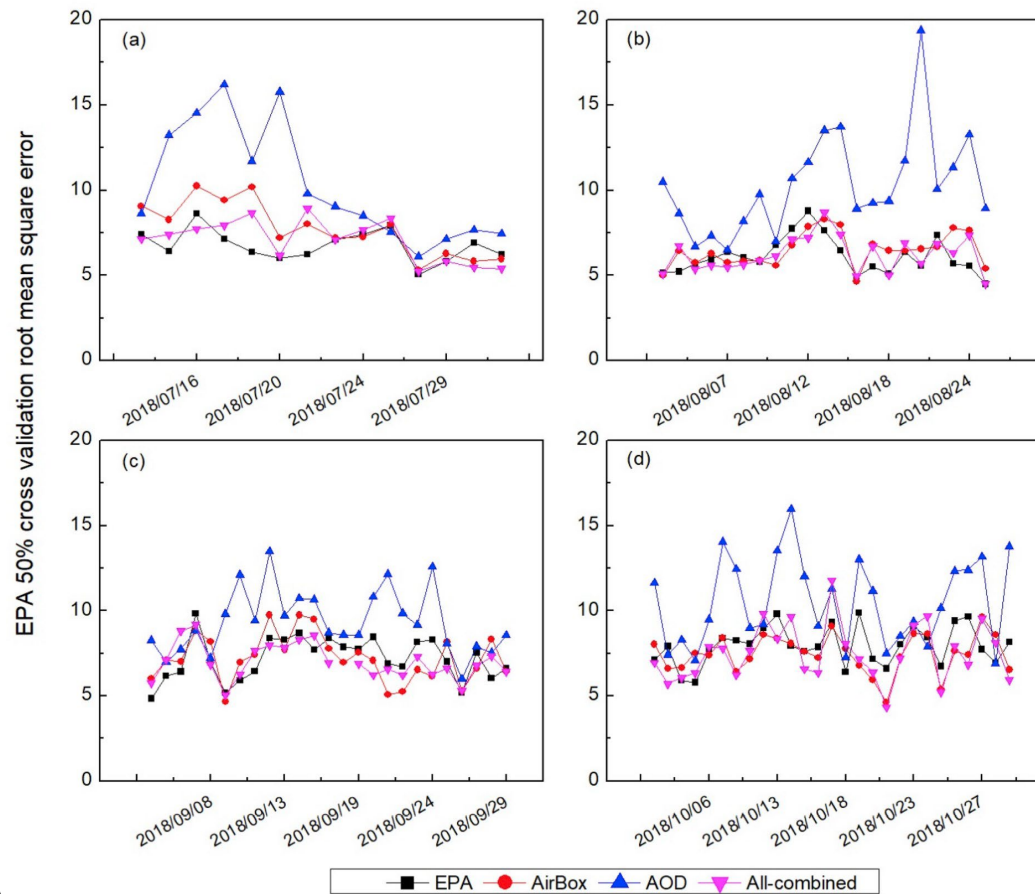
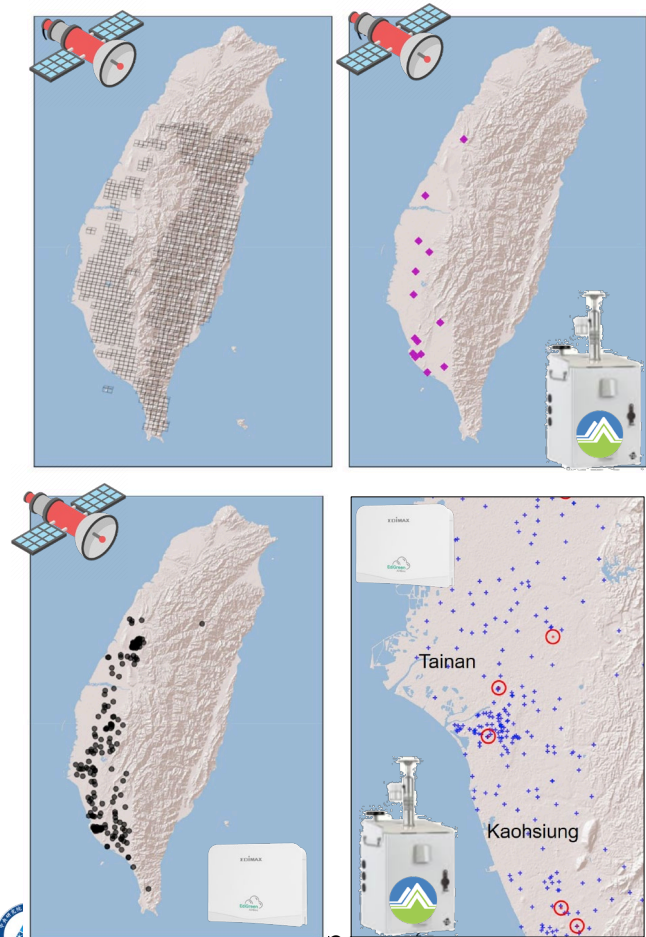
$$R^2 = 0.87; \text{RMSE} = 4.1 \mu\text{g}/\text{m}^3$$



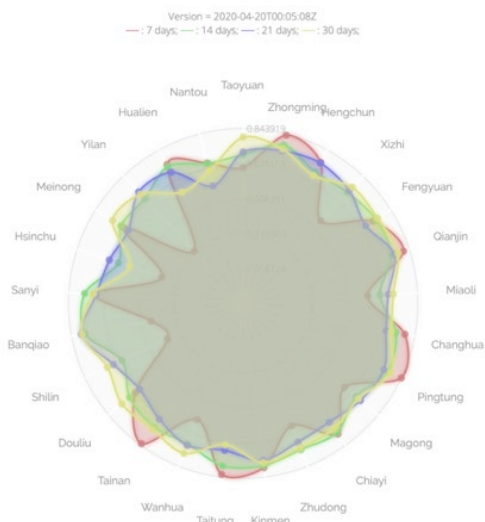
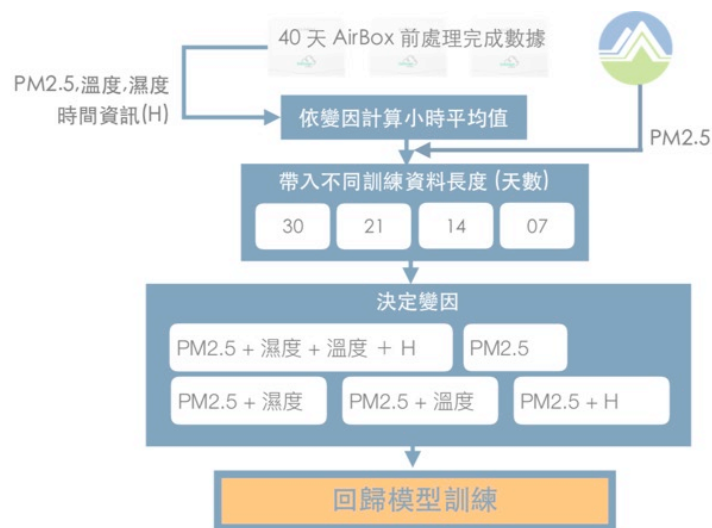
Project Results: Multi-Source Data Fusion

Atmospheric Environment '20

- PM measurement from low-cost sensors, monitoring stations, and remote sensing.
- Characterize different datasets through collocation and calibration.
- Regional scale pollution mapping and pollution events identification.



用開放促進公共化：從微型感測到聯合校正



PM2.5 Open Data Portal AirBox_Status_Report Latest_API Visualization IDW_Animation

Calibration Models Status Report

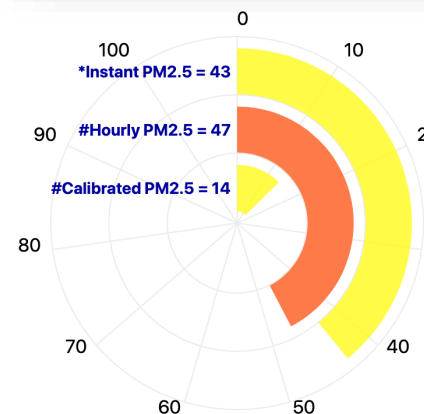
DFC PM2.5 Calibration Open

Data Length
7, 14, 21, 31 days

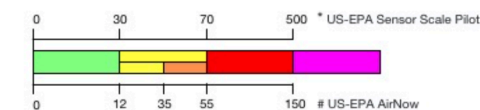
Feature
[P]-PM2.5, [T]-Temperature, [R]-Relative humidity, [H]-Hour at which data is sensed

Method
[RR]-Ridge Regression, [LassoR]-Lasso Regression, [LinearR]-Linear Regression, [BR]-Bayesian Ridge, [RFRR]-Random Forest Regression, [SVR]-Support Vector Regression, [GAM]-Generalized Additive Model

| Date | SiteName | Site GPS | | Data Length | Model Info | | Model Performance | | | | Download Link |
|----------|----------|----------|-----------|-------------|------------|---------|-------------------|------------------------|-------------------------|-----------|---------------|
| | | Latitude | Longitude | | Feature | Method | Validation MAE | Validation Pearson's R | Training R ² | Train MAE | |
| 20200606 | Nantou | 23.913 | 120.685 | 7 | PHR | LassoR | -1.638 | 0.258 | 0.132 | 1.521 | Download |
| 20200606 | Chiayi | 23.463 | 120.441 | 7 | PH | RR | -1.462 | 0.169 | 0.062 | 1.419 | Download |
| 20200606 | Puzi | 23.467 | 120.248 | 7 | PH | LinearR | -1.679 | 0.402 | 0.204 | 1.576 | Download |
| 20200606 | Yilan | 24.748 | 121.746 | 21 | PTR | RR | -1.759 | 0.522 | 0.538 | 1.719 | Download |
| 20200606 | Hengchun | 21.958 | 120.789 | 14 | P | LinearR | -1.448 | 0.367 | 0.156 | 1.407 | Download |
| 20200606 | Pingtung | 22.673 | 120.488 | 7 | PR | BR | -2.013 | 0.448 | 0.3 | 1.934 | Download |
| 20200606 | Changhua | 24.066 | 120.542 | 7 | PT | LinearR | -1.419 | 0.449 | 0.34 | 1.328 | Download |
| 20200606 | Banqiao | 25.013 | 121.459 | 7 | PHR | BR | -2.146 | 0.626 | 0.472 | 1.942 | Download |
| 20200606 | Xizhi | 25.067 | 121.642 | 21 | PR | LinearR | -2.437 | 0.603 | 0.559 | 2.351 | Download |



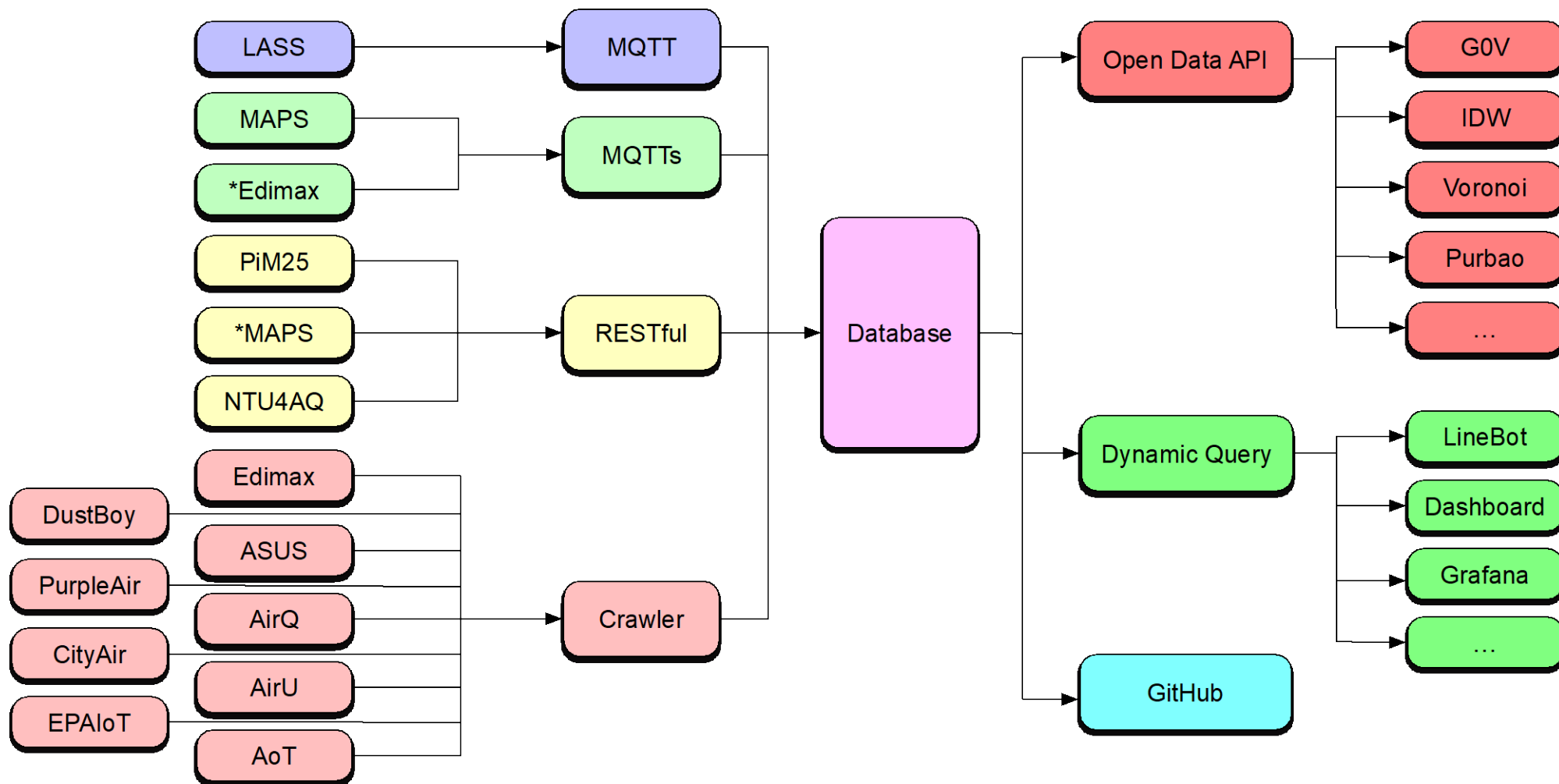
Suggestion:
If medium readings continue (for an hour or more), use the EPA Air Quality Index (AQI) to plan outdoor activities.



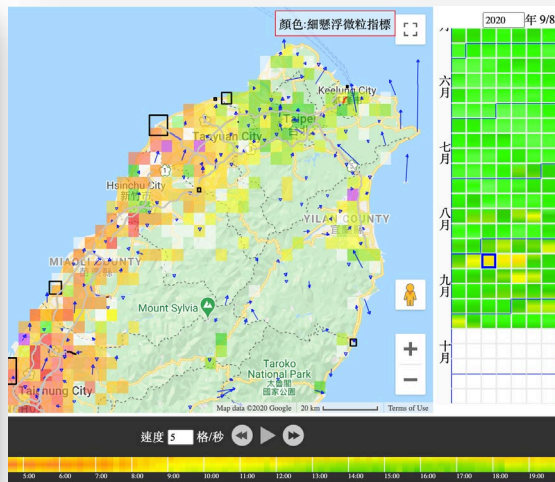
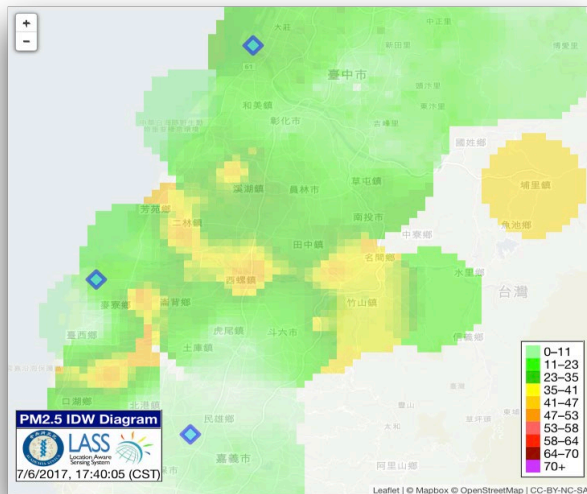
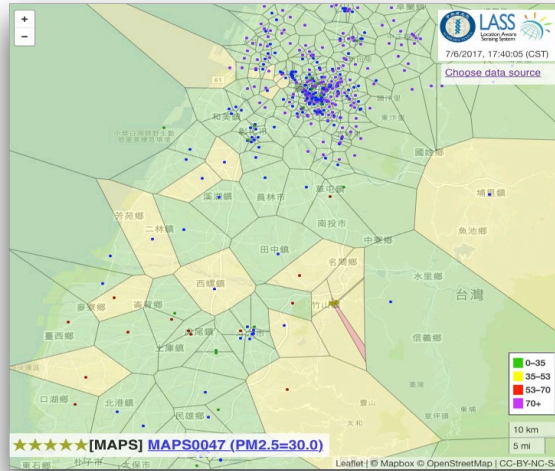
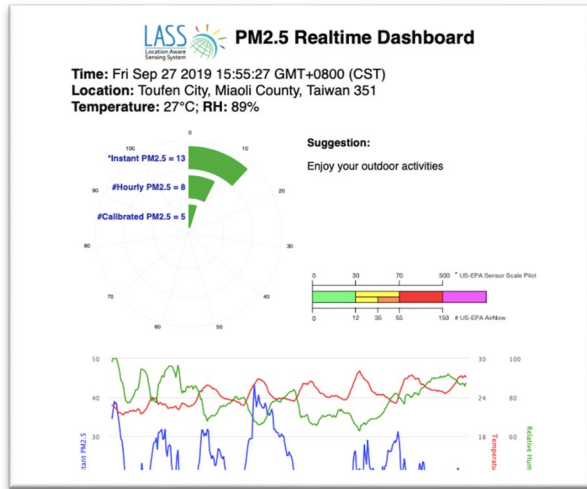
Calibration Model: BRR/7/18.4



用開放促進公共化：從空氣盒子到空氣大聯盟



用開放促進公共化：讓每個想參與的人都被照顧到



SMARTBEAR
SwaggerHub

PM25 Open Data API

1.0.0 OAS3

This Open Data API provides data access to the LASS PM2.5 project, which is a participatory sensing project for large-scale PM2.5 monitoring the PM2.5 sensing project and empowers participants to make low-cost PM2.5 sensing devices on their own. Moreover, it enables PM2.5 monitoring enriches environmental data analysis by making all measurement data freely available for everyone. By 2020, the project has deployed more than 100 countries, and the network is still fast growing. For more details, please visit [PM2.5 Open Data Portal](#).

Servers

<https://pm25.lass-net.org/API-1.0.0/>

Powered by Location Aware Sensing System (LASS) and IIS-NRL, Academia Sinica

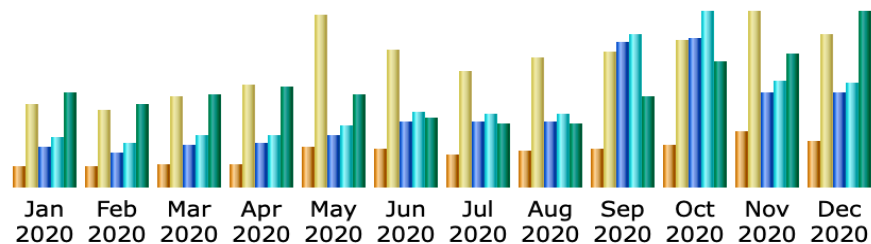
Device

- GET /device/{device_id}/latest/
- GET /device/{device_id}/history/
- GET /device/{device_id}/date/{yyyy-mm-dd}/
- GET /device/nearest/{lat}/{lon}/{lon}/

https://app.swaggerhub.com/apis-docs/I2875/PM25_Open_Data/



用開放促進公共化：用資料驅動自我省視

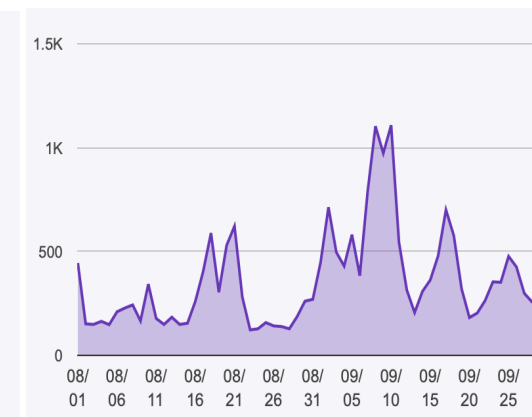
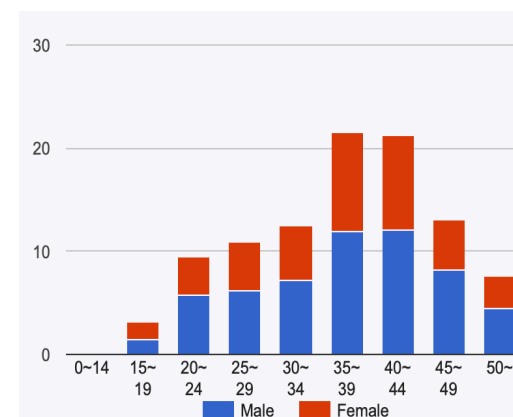
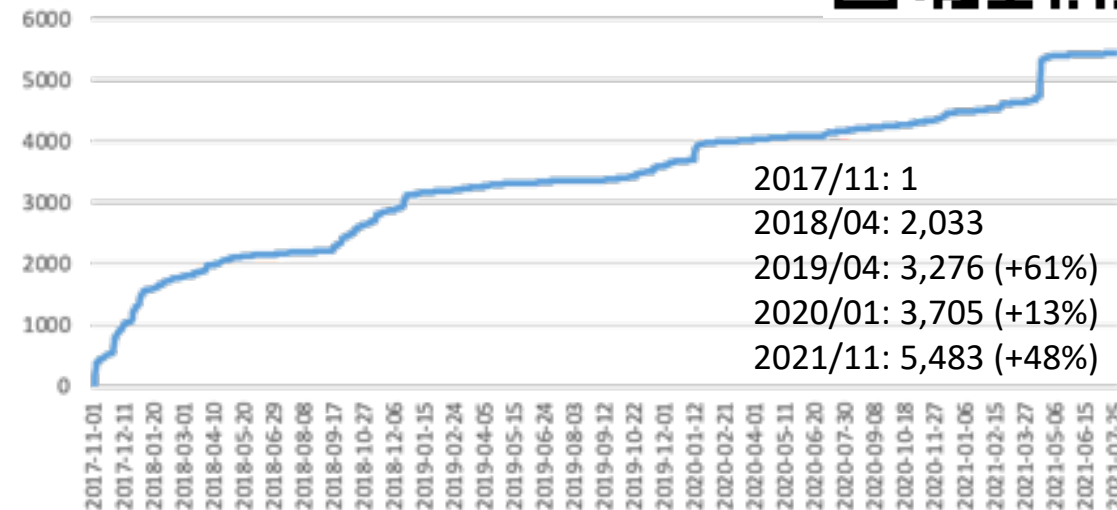


| Month | Unique visitors | Number of visits | Pages | Hits | Bandwidth |
|----------|-----------------|------------------|------------|------------|-----------|
| Jan 2020 | 7,732 | 31,600 | 3,782,432 | 4,853,520 | 245.46 GB |
| Feb 2020 | 7,453 | 29,415 | 3,224,685 | 4,190,210 | 217.10 GB |
| Mar 2020 | 8,267 | 34,537 | 4,082,964 | 4,976,752 | 241.19 GB |
| Apr 2020 | 8,745 | 39,091 | 4,195,639 | 4,936,887 | 260.75 GB |
| May 2020 | 15,398 | 65,021 | 5,002,852 | 5,869,380 | 239.80 GB |
| Jun 2020 | 14,294 | 52,019 | 6,328,794 | 7,224,400 | 179.10 GB |
| Jul 2020 | 11,926 | 43,799 | 6,296,439 | 7,041,846 | 164.09 GB |
| Aug 2020 | 13,925 | 49,284 | 6,240,171 | 6,979,881 | 163.79 GB |
| Sep 2020 | 14,435 | 51,660 | 13,878,826 | 14,729,322 | 237.36 GB |
| Oct 2020 | 15,581 | 55,485 | 14,326,799 | 16,812,616 | 326.88 GB |
| Nov 2020 | 20,759 | 66,458 | 9,115,525 | 10,149,143 | 349.24 GB |
| Dec 2020 | 17,475 | 57,716 | 9,054,528 | 10,042,433 | 457.11 GB |
| Total | 155,990 | 576,085 | 85,529,654 | 97,806,390 | 3.01 TB |

- Yearly visitors: 576,085 => 48,007 person/month, 1,600 person/day, i.e., **1 person/minute**
- Yearly hit count: 97,806,390 => **3.14 hits/sec**
- Yearly traffic: 3.01 TB => **8.25 GB/Day**



PM2.5 OPEN DATA



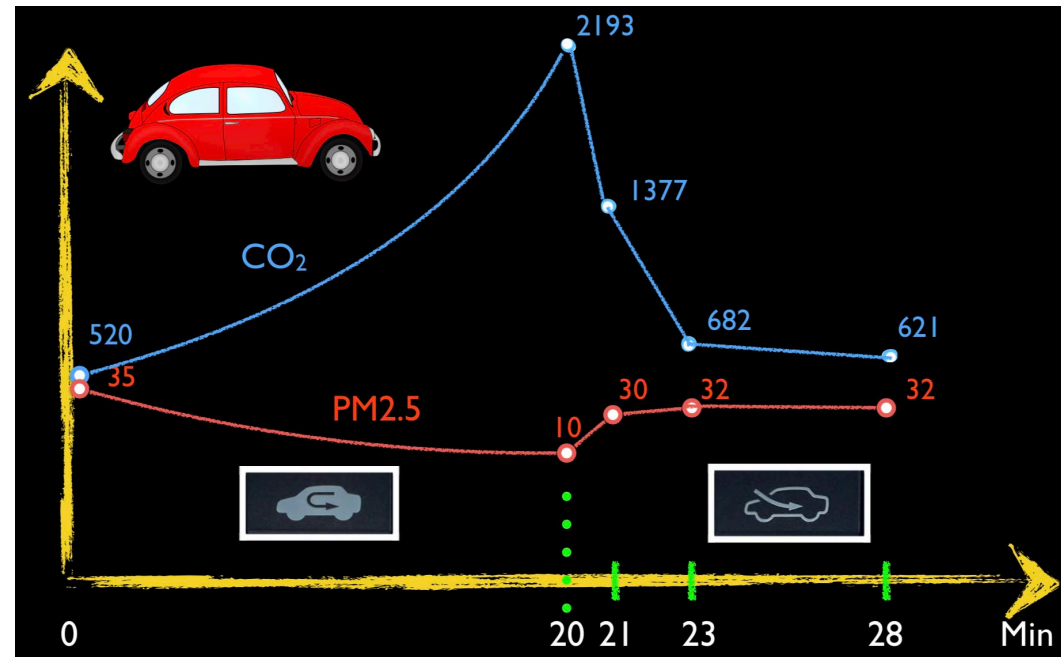
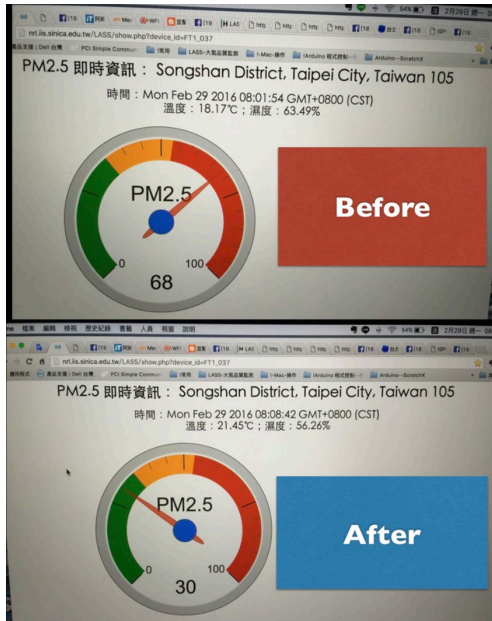
用開放促進公共化：透過社群參與回答每個人的問題



Before



After



用開放促進公共化：用範例演示降低資料應用的門檻

LEARN CIOT

搜尋

- 1. 教學網站簡介 ✓
- 2. 整體課程前言
- 3. 資料取用 ✓
- 4. 時間維度資料分析 ✓
 - 4.1. 時間序列資料處理 ✓
 - 4.2. 時間序列資料預測
 - 4.3. 時間序列屬性分群
- 5. 空間維度資料分析
- 6. 資料應用
- 7. 系統整合應用 ✓

其他連結

- GitHub
- 按照難度等級查詢
- 按照文章標籤查詢
- 按照文章作者查詢

民生公共物聯網資料應用 > 4. 時間維度資料分析 > ...

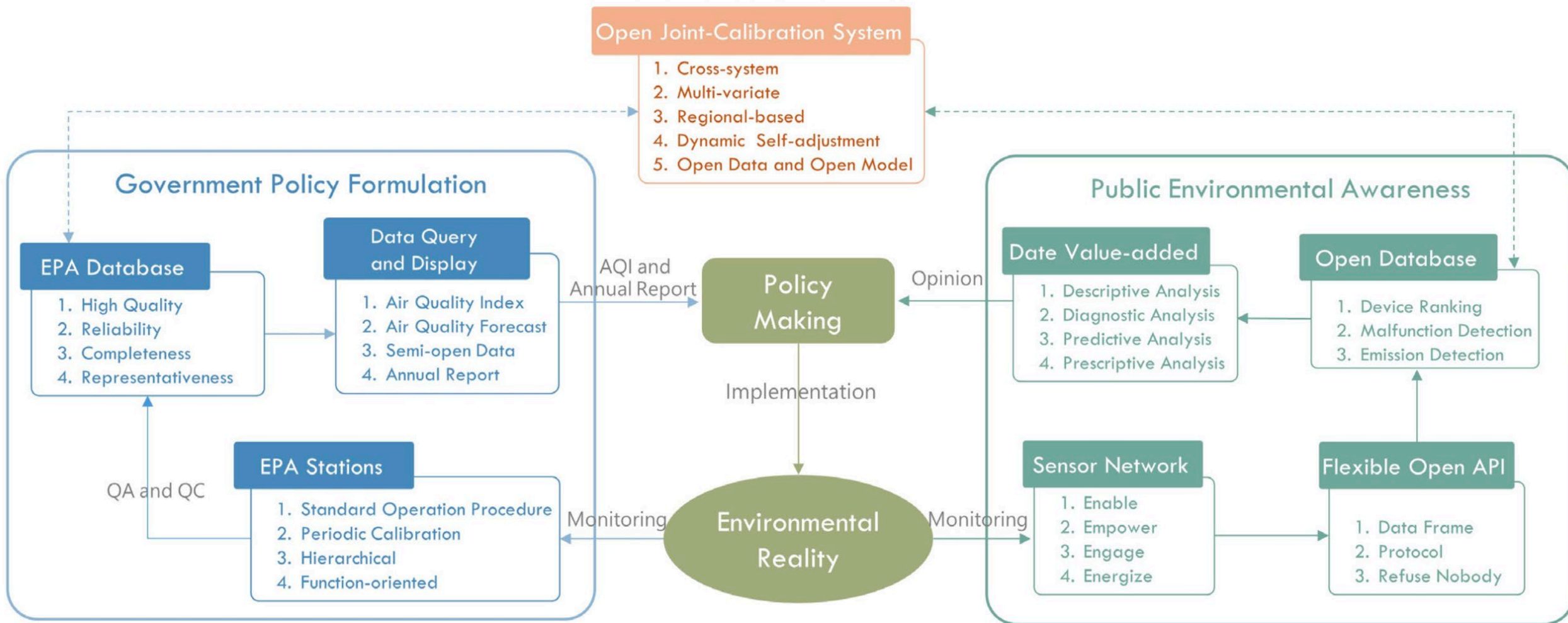
```
# cmap: 設定呈現的顏色色盤  
(https://matplotlib.org/stable/gallery/color/colormap_reference.html)  
# textformat: 設定圖中數字呈現的樣式  
plt = calplot.calplot(data = air_day['PM25'], cmap = 'GnBu', textformat =  
'{:,.0f}',  
figsize = (24, 12), subtitle =  
"PM25 by Month and Year")
```

PM25 by Month and Year

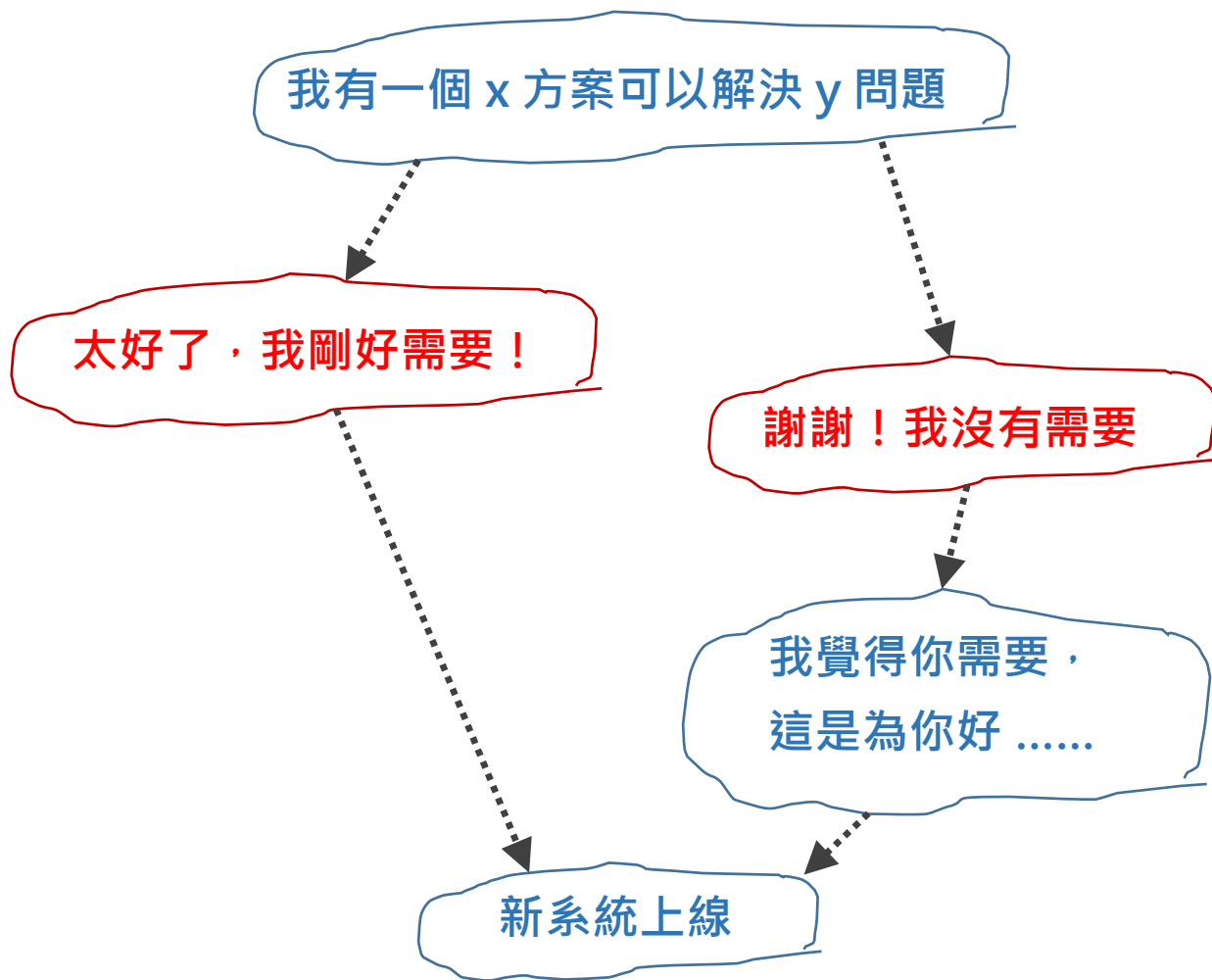
<https://learnciot.github.io/>



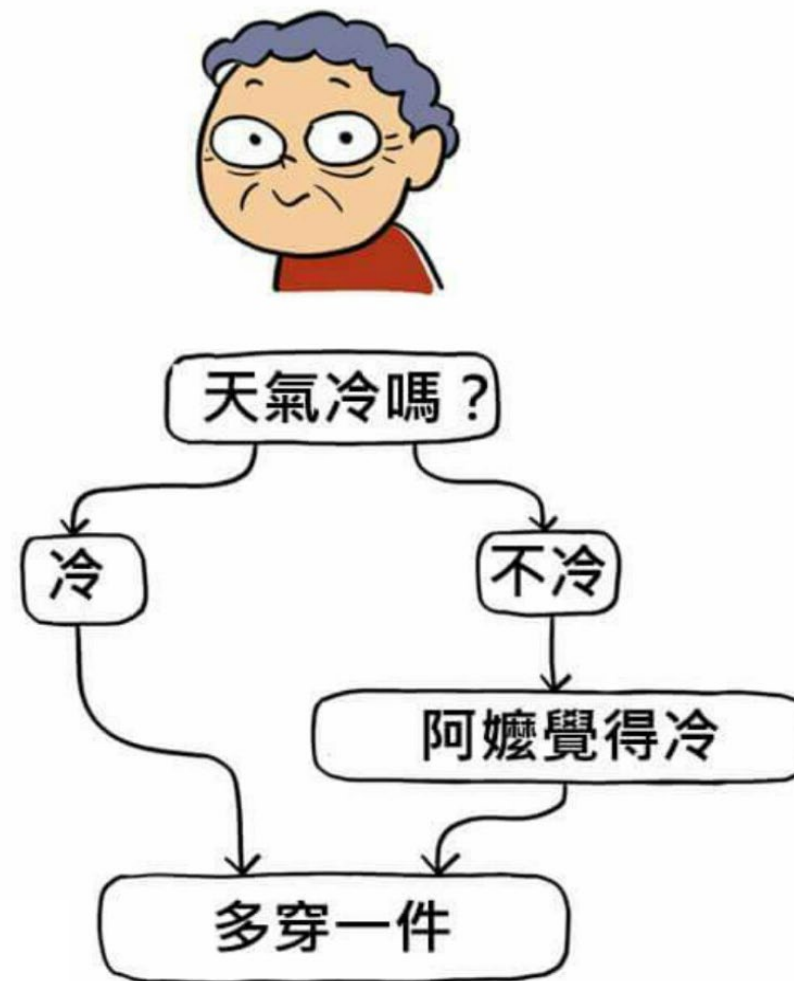
用開放促進公共化：從參與式感測到公私協力



最常遇見/最需要避免的迷思



阿嬤幫孩子選衣服



What's the next?

Device

- Additional sensors
(CO₂, TVOC, sound,
light)
- Green IoT
- IoT security
- Edge computing

What's the next?

Device

- Additional sensors (CO₂, TVOC, sound, light)
- Green IoT
- IoT security
- Edge computing

Data

- Big data fusion
- Cyber-physical-social data fusion
- AI-aided model simulation
- GeoAI and predictive database

What's the next?

Device

- Additional sensors (CO₂, TVOC, sound, light)
- Green IoT
- IoT security
- Edge computing

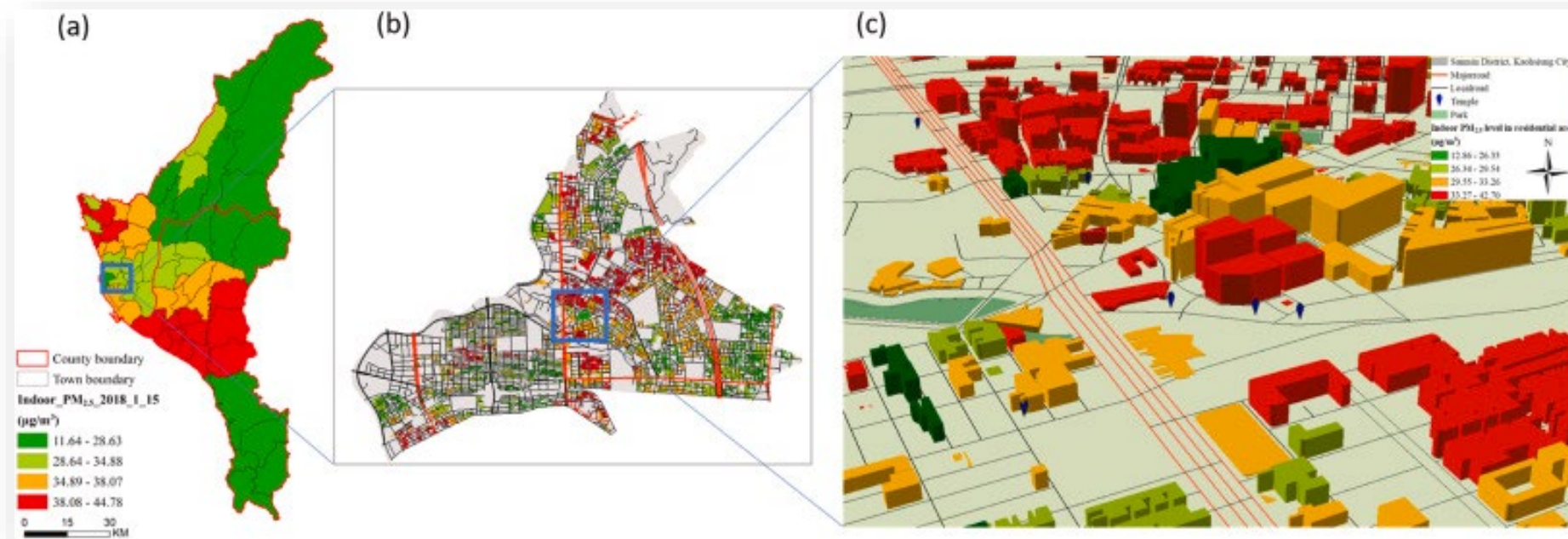
Data

- Big data fusion
- Cyber-physical-social data fusion
- AI-aided model simulation
- GeoAI and predictive database

Application

- Indoor applications
- Cross-disciplinary research
- International collaboration
- Data-informed decision making

室外 -> 室內



Predictive maps of indoor PM_{2.5} concentration variations using LUR integrated XGBoost approach at different scales: (a) county level, (b) township level, and (c) neighborhood level with 3D visualization.

- Yi-Fang Chiang, Ka-Ui Chu, Ling-Jyh Chen, and Yao-Hua Ho. Predicting Risk of COVID-19 Outbreak by using Outdoor Air Pollution Indicators and Population Flow with Queueing Theory. *Atmosphere*, volume 13, issue 10, article number 1727, 2022.
- Pei-Yi Wong, Hsiao-Yun Lee, Ling-Jyh Chen, Yu-Cheng Chen, Nai-Tzu Chen, Shih-Chun Candice Lung, Huey-Jen Su, Chih-Da Wu, Jose Guillermo Cedeno Laurent, Gary Adamkiewicz, and John D. Spengler. An alternative approach for estimating large-area indoor PM_{2.5} concentration – A case study of schools. *Building and Environment*, volume 219, 109249, July 2022.
- Jose Guillermo Cedeño Laurent, Piers MacNaughton, Emily Jones, Anna S. Young, Maya Bliss, Skye Flanigan, Jose Vallarino, Ling-Jyh Chen, Maya Bliss, Borong Lin, Xiaodong Cao, and Joseph G. Allen. Associations between Acute Exposures to PM_{2.5} and Carbon Dioxide Indoors and Cognitive Function in Office Workers: A Multicountry Longitudinal Prospective Observational Study. *Environmental Research Letters*, volume 16, number 9, 094047, Sept. 2021.

空氣盒子 -> 聲音盒子

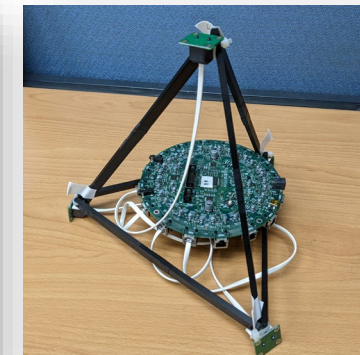
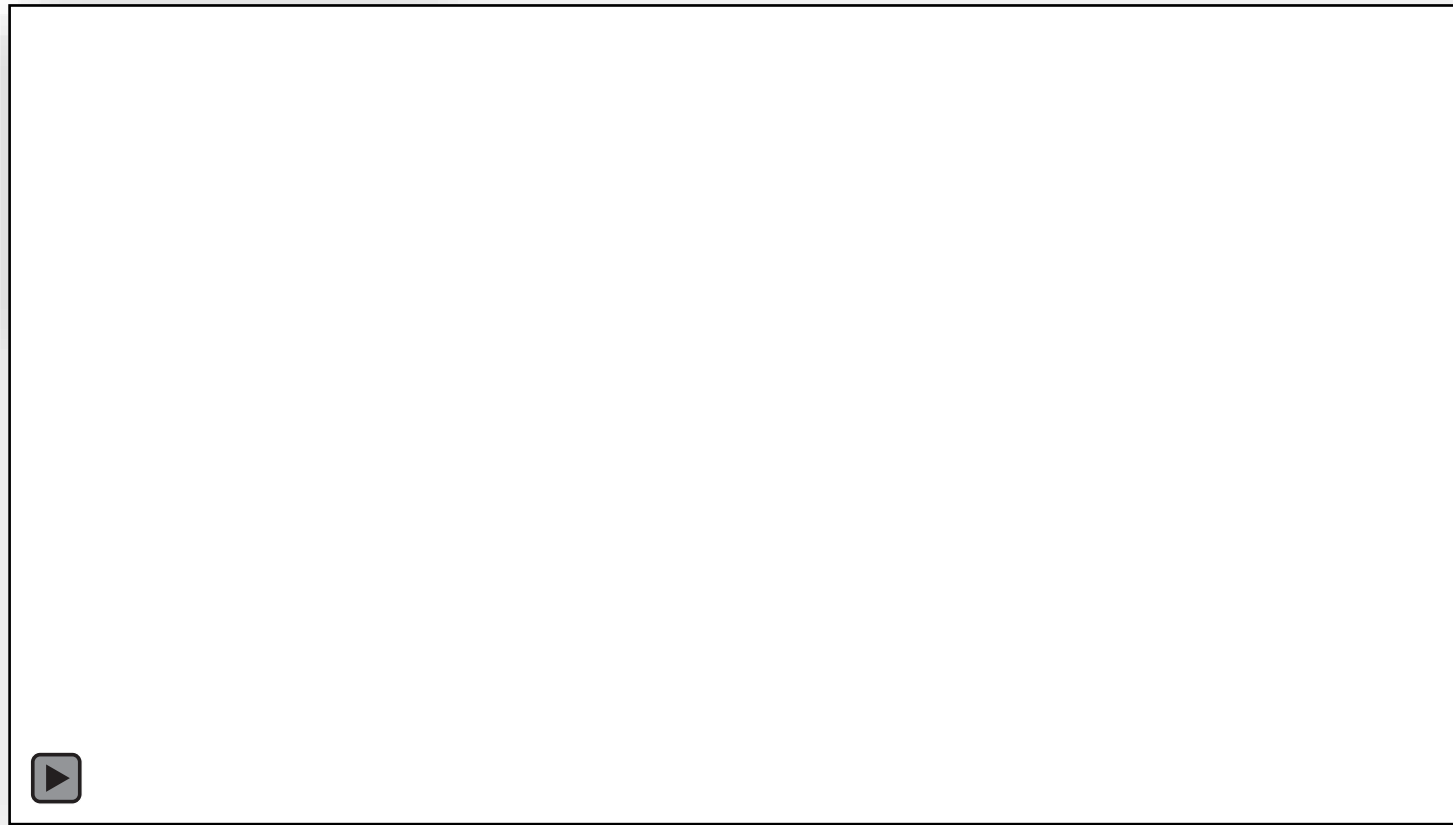
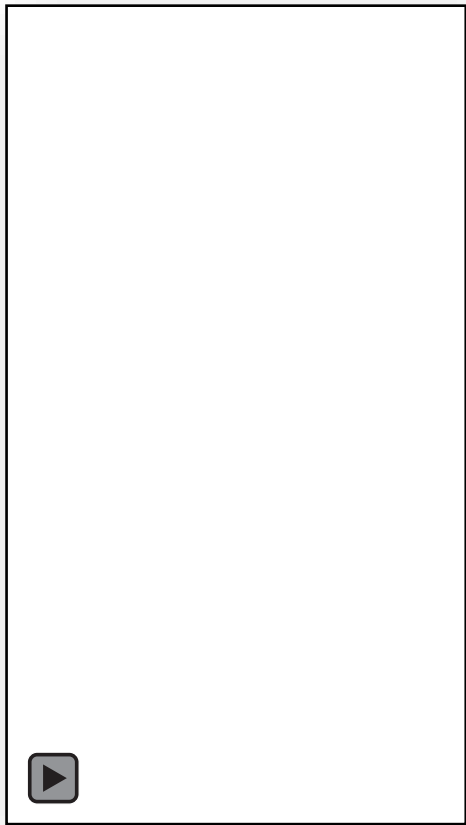
學者聯手打造聲景地圖 揪出北市大安中山區夜間為噪音熱點

2022/9/25 16:56 (9/25 17:18 更新)



- Jia-Hong Tang, Bo-Cheng Lin, Jing-Shiang Hwang, Ling-Jyh Chen, Bing-Sheng Wu, Hong-Lian Jian, Yu-Ting Lee, and Ta-Chien Chan. Dynamic modelling for noise mapping in urban areas. Environmental Impact Assessment Review, volume 97, 106864, November 2022.

聲音盒子 -> 聽音辨位



結語

- 「大、人、物」促進更深層的資料驅動決策與治理
- 「開放」促進溝通、互信、合作、共創
- 「must-have」才是參與式感測的核心
- 智慧環境感測還有許多待解的議題，期待未來一起合作！

Thank You!

