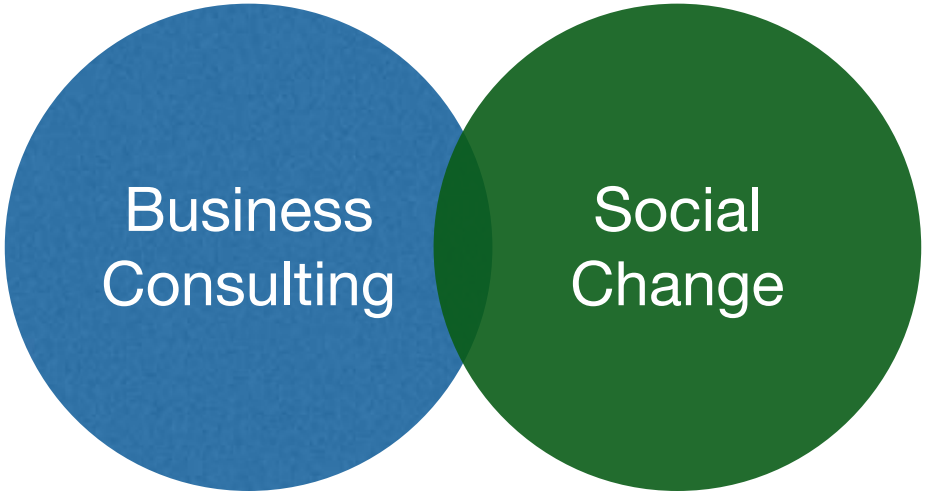
Data for Social Good

Chia-Kai Liu (DSP, Inc.)
ck@dsp.im | @agentCKL

2018.12.21

DSP

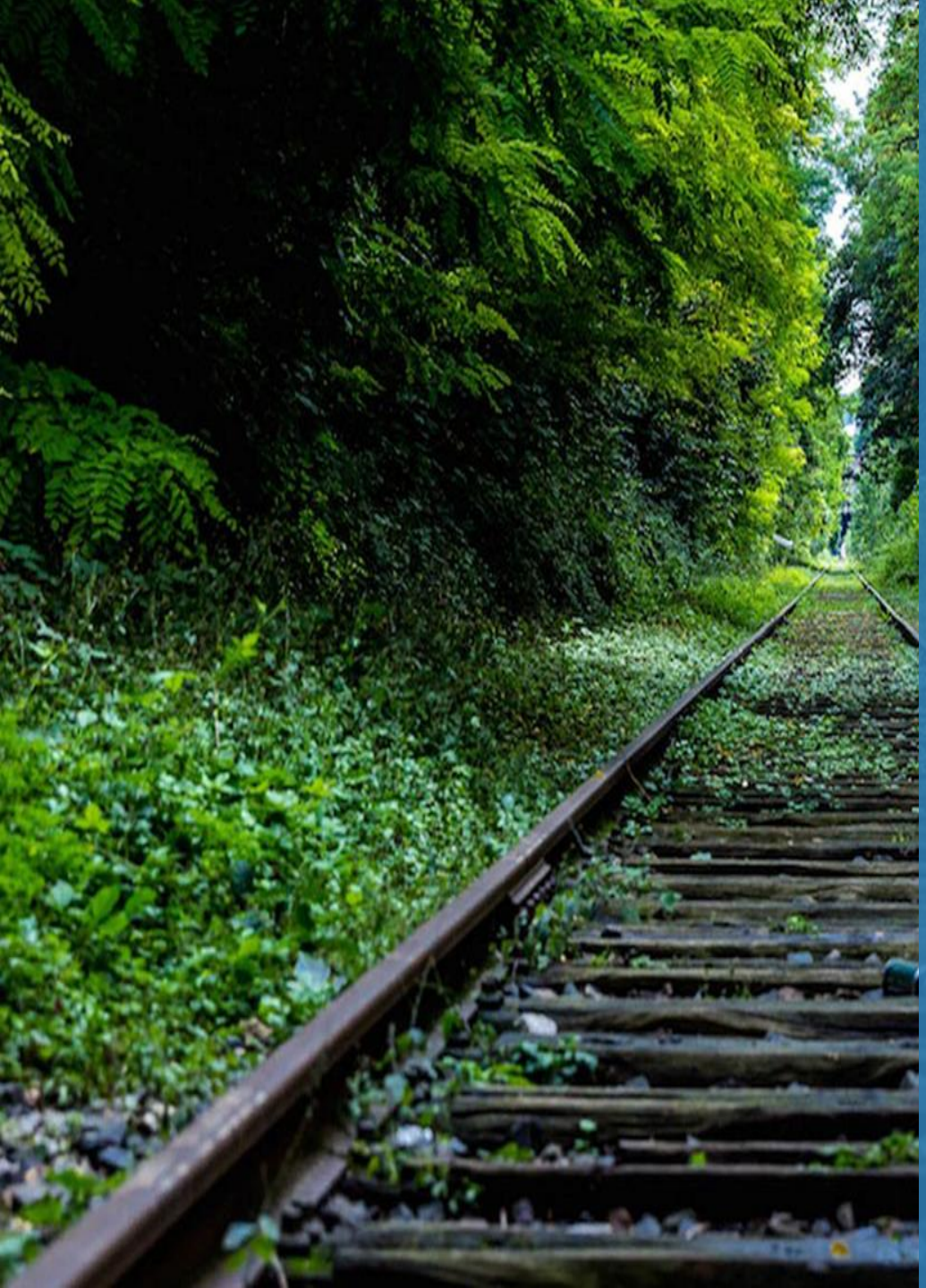


Use Data Science to Solve Big Challenges

https://dsp.im/

Our Clients


台灣化學纖維股份有限公司
FORMOSA CHEMICALS & FIBRE CORPORATION
台灣電力公司
中華電信
Chunghwa Telecom
統一企業
大潤發
RT-MART
高醫
福懋
TREND MICRO
BenQ
SYSTEX 精誠資訊
富邦人壽
全球人壽
天下雜誌
聯合報系
United Daily News Group
BUSINESS WEEKLY GROUP
商周集團
風傳媒
THE STORM MEDIA
經理人
MANAGER today
痞客邦 PIXNET
今周刊
Leo Burnett
國家發展委員會
NATIONAL DEVELOPMENT COUNCIL
中華民國經濟部
Ministry of Economic Affairs, R.O.C.
行政院農業委員會
COUNCIL OF AGRICULTURE, EXECUTIVE YUAN
嘉義市政府
Chiayi City Government
臺北市政府社會局
Department of Social Welfare, Taipei City Government
新北市政府社會局
Social Welfare Department, New Taipei City Government
工業技術研究院
Industrial Technology Research Institute
中央研究院
ACADEMIA SINICA
財團法人資訊工業策進會
INSTITUTE FOR INFORMATION INDUSTRY
TAITRA



Topics

1. D4SG introduction
2. Case Studies
3. Lessons learned
4. Next steps

Data for Social Good (D4SG)

Data for Social Good

A Data Revolution for the Social & Public Sectors

http://d4sg.org/

Founding Partners

DSP

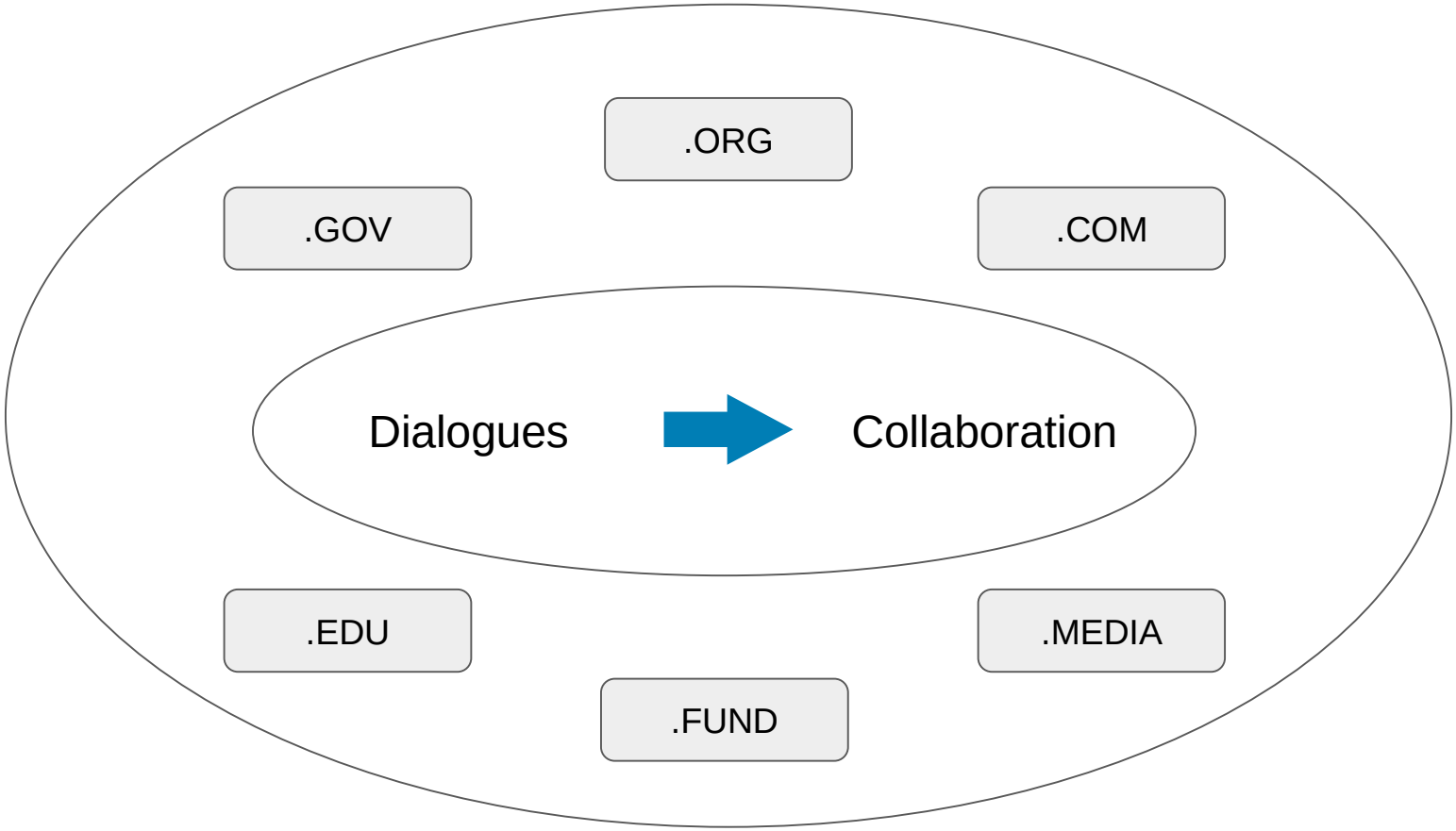
Frontier Foundation

Social Computing & Big Data Project

Objectives

- Foster stakeholder dialogues
- Develop data-driven, evidence-based policies
- Improve operational efficiency
- Stimulate innovation in public services

Multi-Stakeholder Partnership

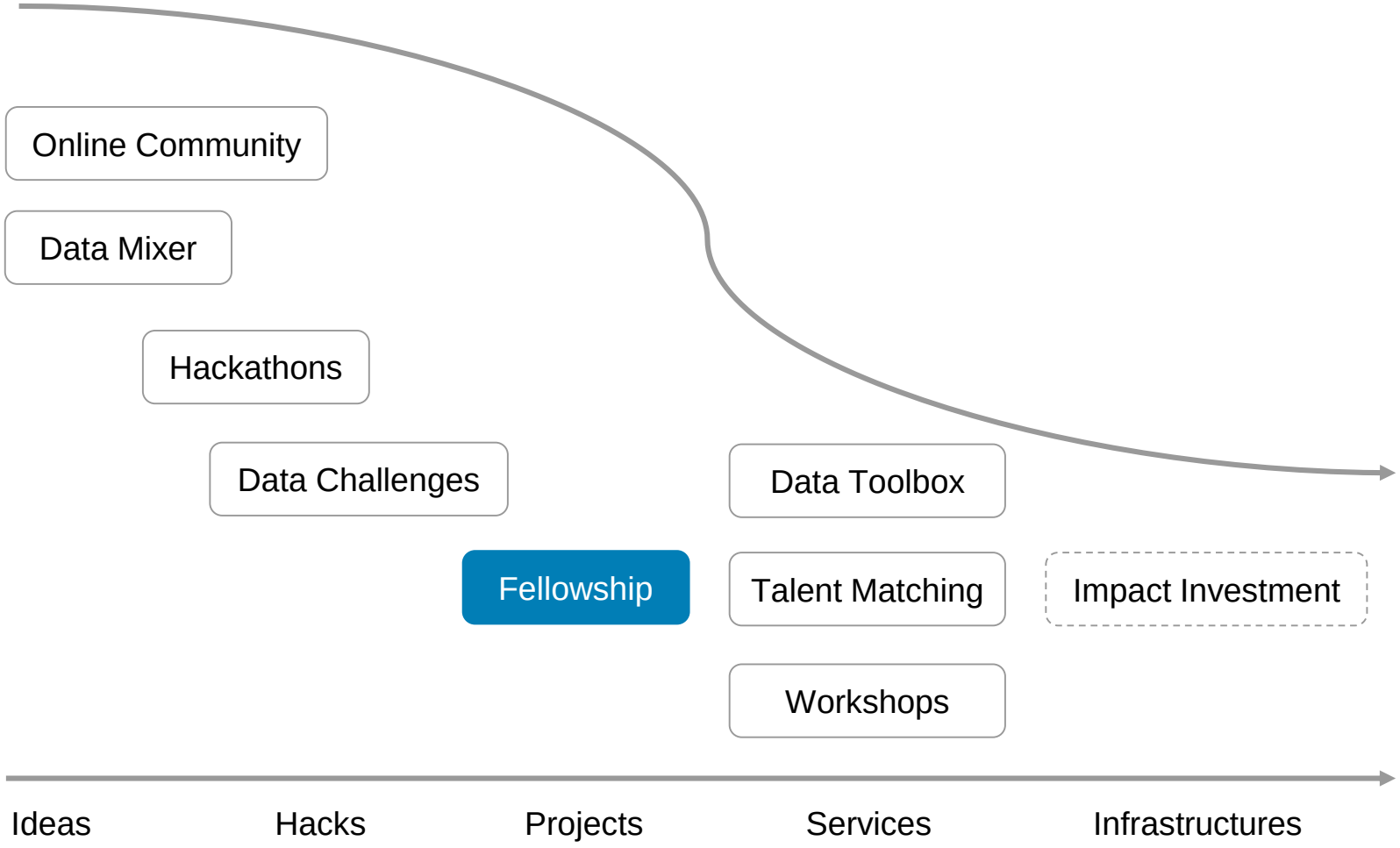


Operating Model

Seek Collaboration

Provide & Learn

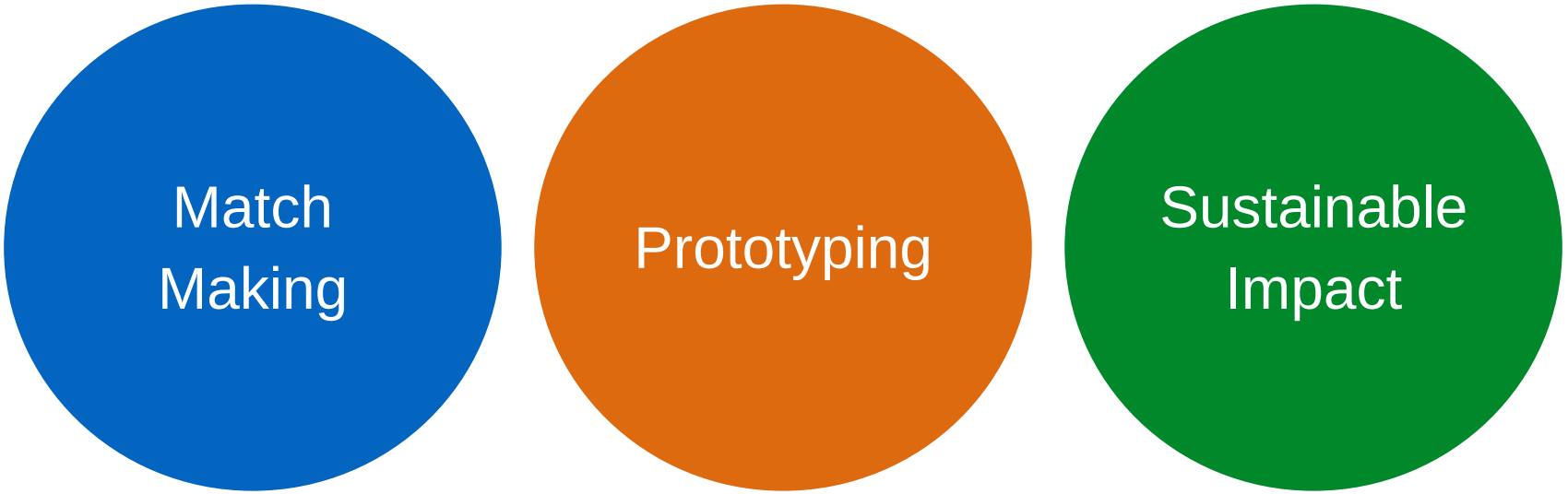
Approaches


Online Community
Data Mixer
Hackathons
Data Challenges
Fellowship
Data Toolbox
Talent Matching
Impact Investment
Workshops
Ideas
Hacks
Projects
Services
Infrastructures

Project Selection Criteria

- High social impact potential
- Already have data
- Dedicated liaison
- Executive support
- Willingness to adopt if the results are good
- Scalability

From Idea to Impact



Sustainable Development Goals

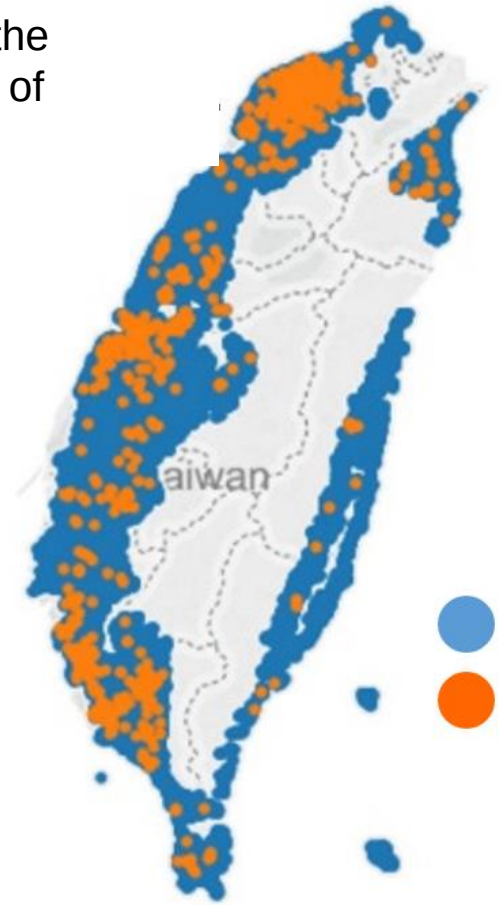
Case Studies

Farmland Pollution Management

http://d4sg.org/farmland-pollution/

Heavy Metal Pollution on Farmland

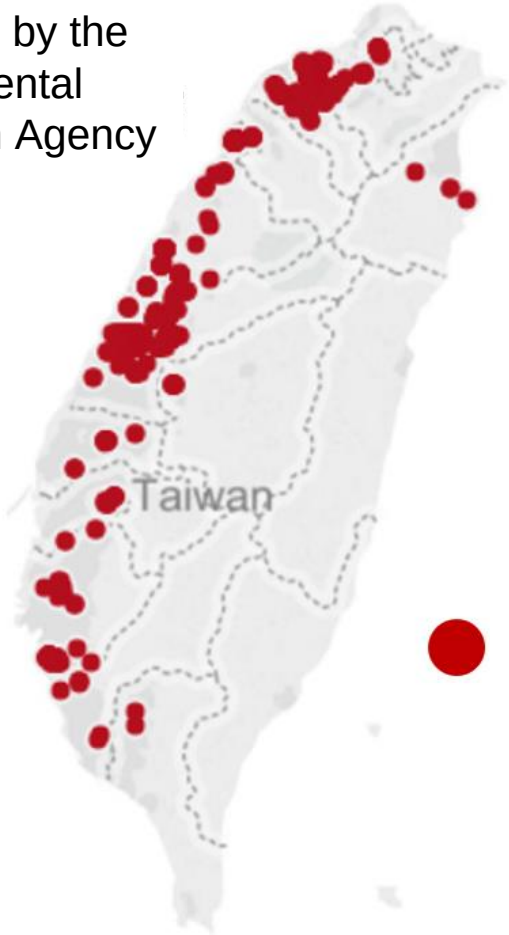
Tested by the Committee of Agriculture



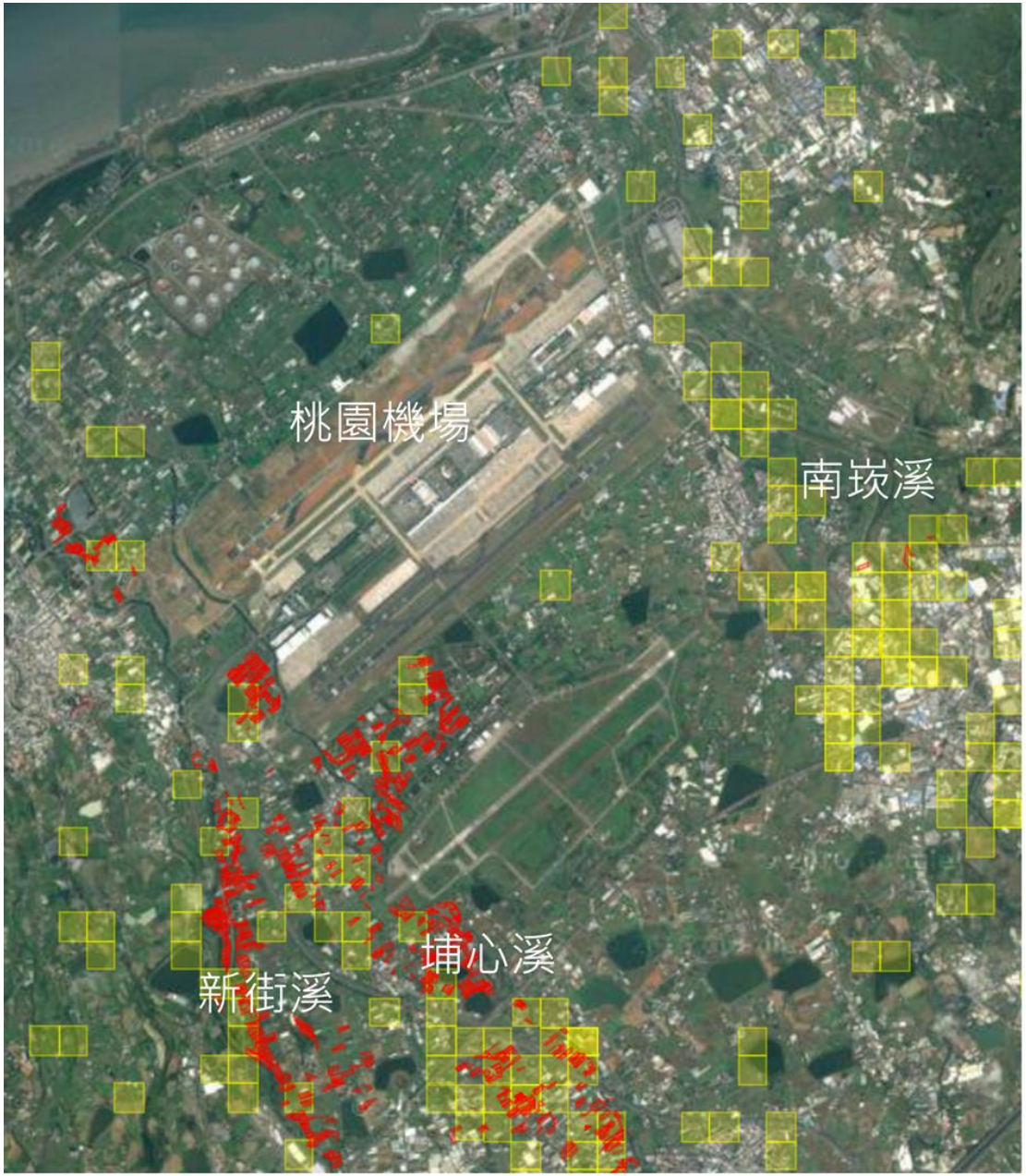
Not toxic

Toxic

Monitored by the Environmental Protection Agency

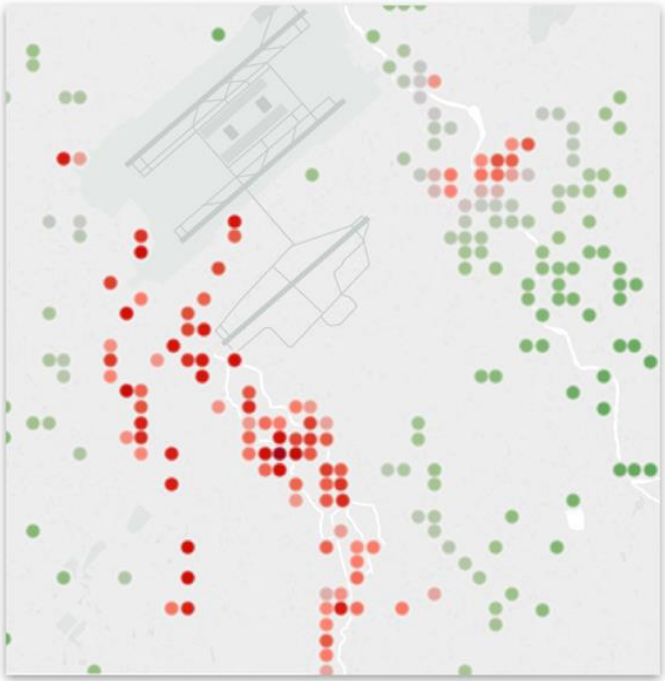


Monitored

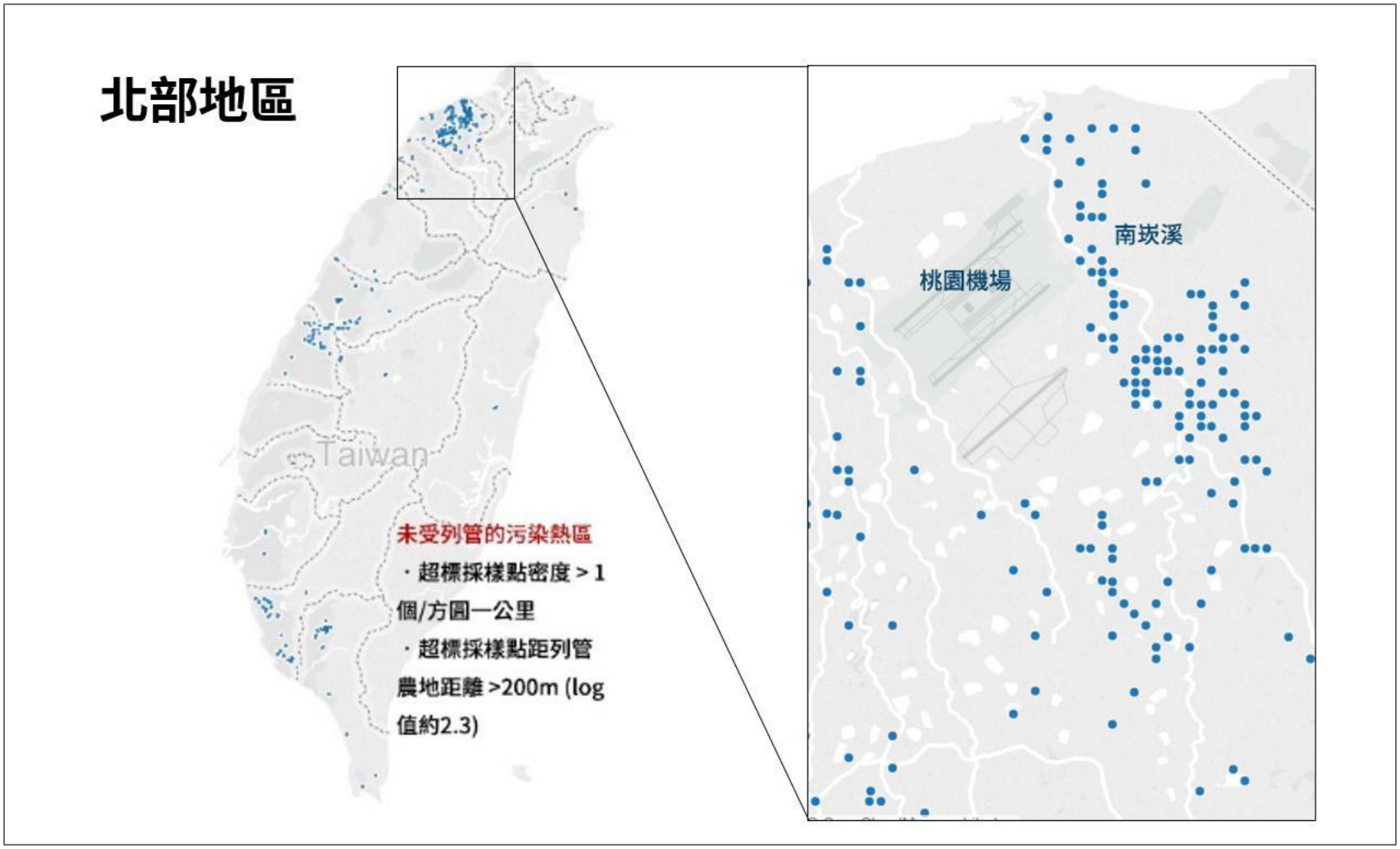


Yellow/Green: Toxic
Red: Monitored

重金屬超標與列管農地關係



Unmonitored Toxic Locations

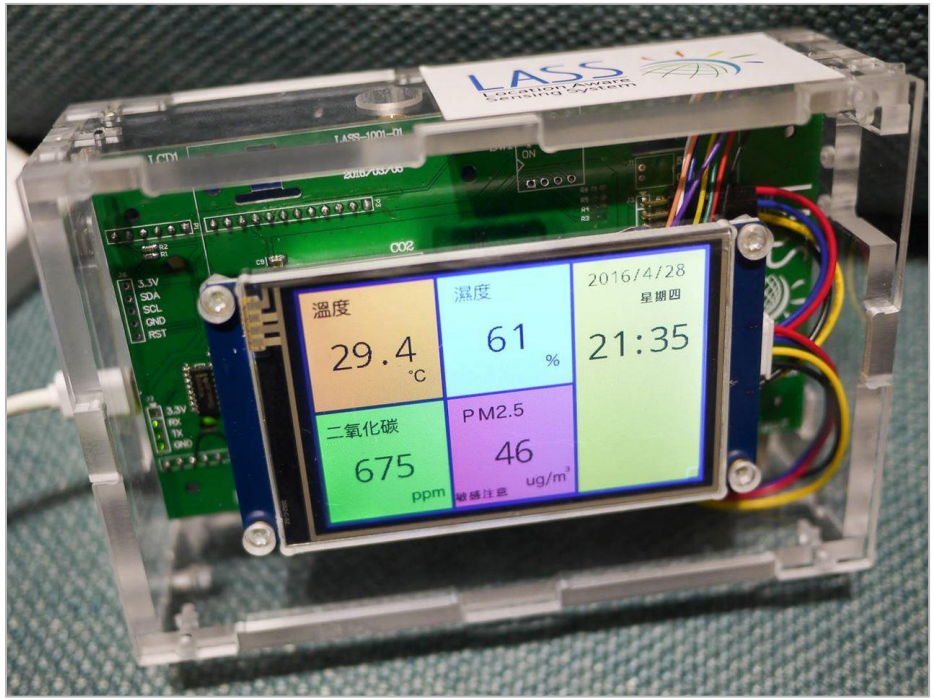


https://youtu.be/OZyusema_H0

Air Pollution Tracking

“Airbox” for PM2.5 Monitoring

LASS
Location Aware
Sensing System



https://pm25.lass-net.org/

AirBox Open Data APIs

https://sites.google.com/site/pm25opendata/open-data

- the latest samples of LASS devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-lass.json
- the latest samples of AirBox devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-airbox.json
- the latest samples of MAPS devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-maps.json
- the latest samples of ProbeCube devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-probecube.json
- the latest samples of Indie devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-indie.json
- the latest samples of Webduino devices
  - update frequency: every 5 min
  - https://data.lass-net.org/data/last-all-webduino.json
- the latest samples of TW-EPA stations
  - update frequency: every hour
  - https://data.lass-net.org/data/last-all-epa.json

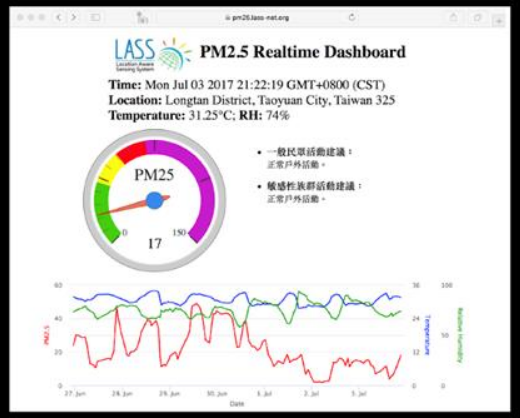
https://scidm.nchc.org.tw/organization/environment-quality

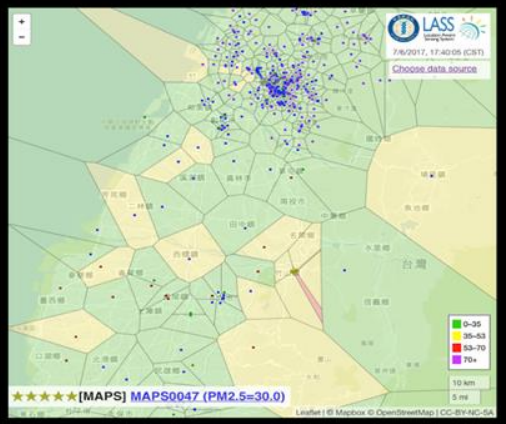
- Data analysis results
  - ranking results of all LASS/AirBox devices by IIS-NRL algorithm
    - update frequency: every 5 minute
    - https://data.lass-net.org/data/device_ranking.json
  - indoor LASS/AirBox devices detected by IIS-NRL algorithm
    - update frequency: every 5 minute
    - https://data.lass-net.org/data/device_indoor.json
  - Malfunctioned LASS/AirBox devices detected by IIS-NRL algorithm
    - update frequency: every 5 minute
    - https://data.lass-net.org/data/device_malfunction_daily.json
  - Local pollutions detected by IIS-NRL algorithm
    - update frequency: every 5 minute
    - https://data.lass-net.org/data/device_pollution.json
- AirBox status report
  - update frequency: every 5 minutes
  - https://pm25.lass-net.org/AirBox/

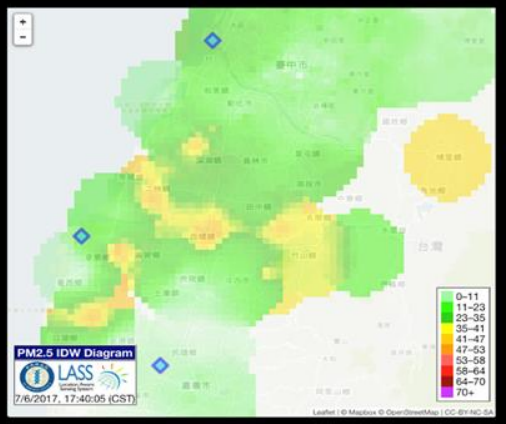
- the latest sample of one particular device
  - update frequency: upon request
  - https://data.lass-net.org/data/last.php?device_id=XXX
- the last 1-week samples of one particular device
  - update frequency: upon request
  - https://data.lass-net.org/data/history.php?device_id=XXX
  - other options (can be used at the same time):
    - format: CSV or JSON (default)
      - https://data.lass-net.org/data/history.php?device_id=XXX&format=CSV
    - date: in YYYY-MM-DD format
      - https://data.lass-net.org/data/history.php?device_id=XXX&date=YYYY-MM-DD
- the latest sample of the nearest device
  - update frequency: upon request
  - http://nrl.iis.sinica.edu.tw/LASS/nearest.php
- the explanation about the field names used in the open data
  - https://lass.hackpad.com/LASS-Data-specification-1dYpwINtH8R

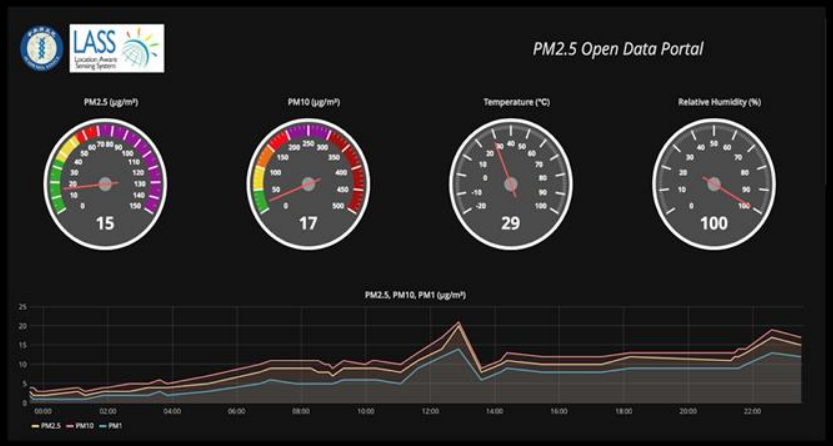
https://pm25.lass-net.org/

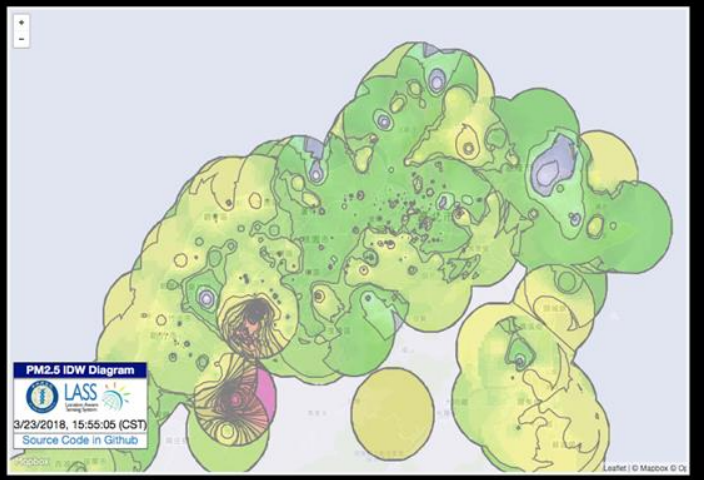
AirBox Visualization











https://pm25.lass-net.org/

https://pm25.lass-net.org/

The key of this project: OPEN

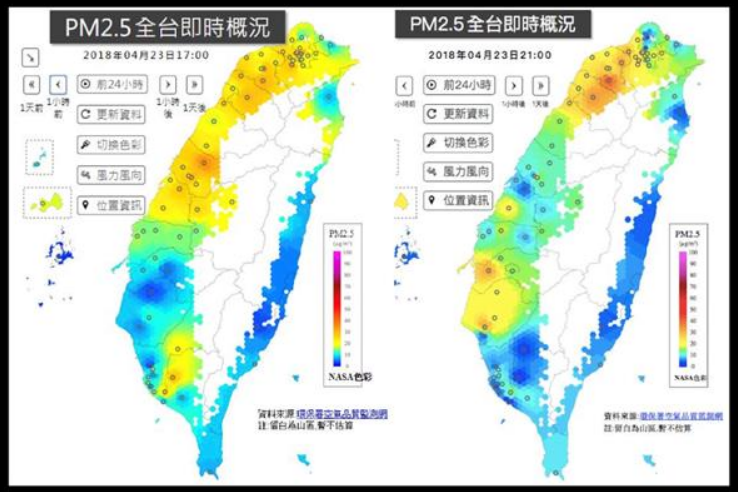
App. → Air Quality Monitoring

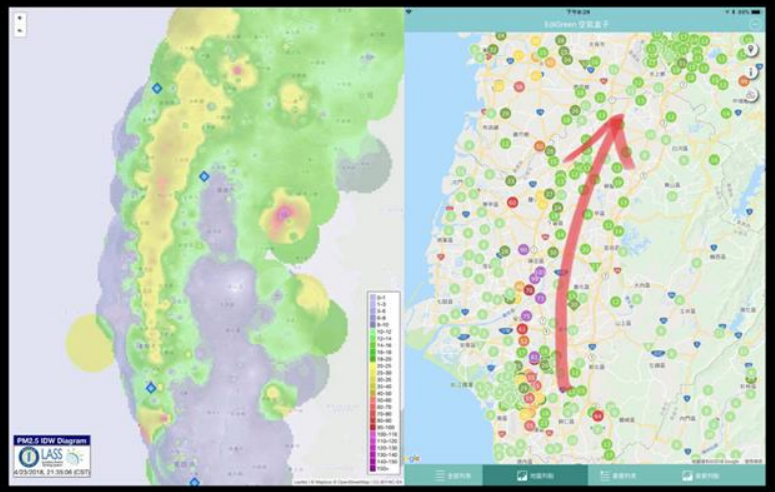
Spirit → Open HW | Open SW | Open Data

Method → Maker | Participation | Environment

Internet of Things + Citizen Science

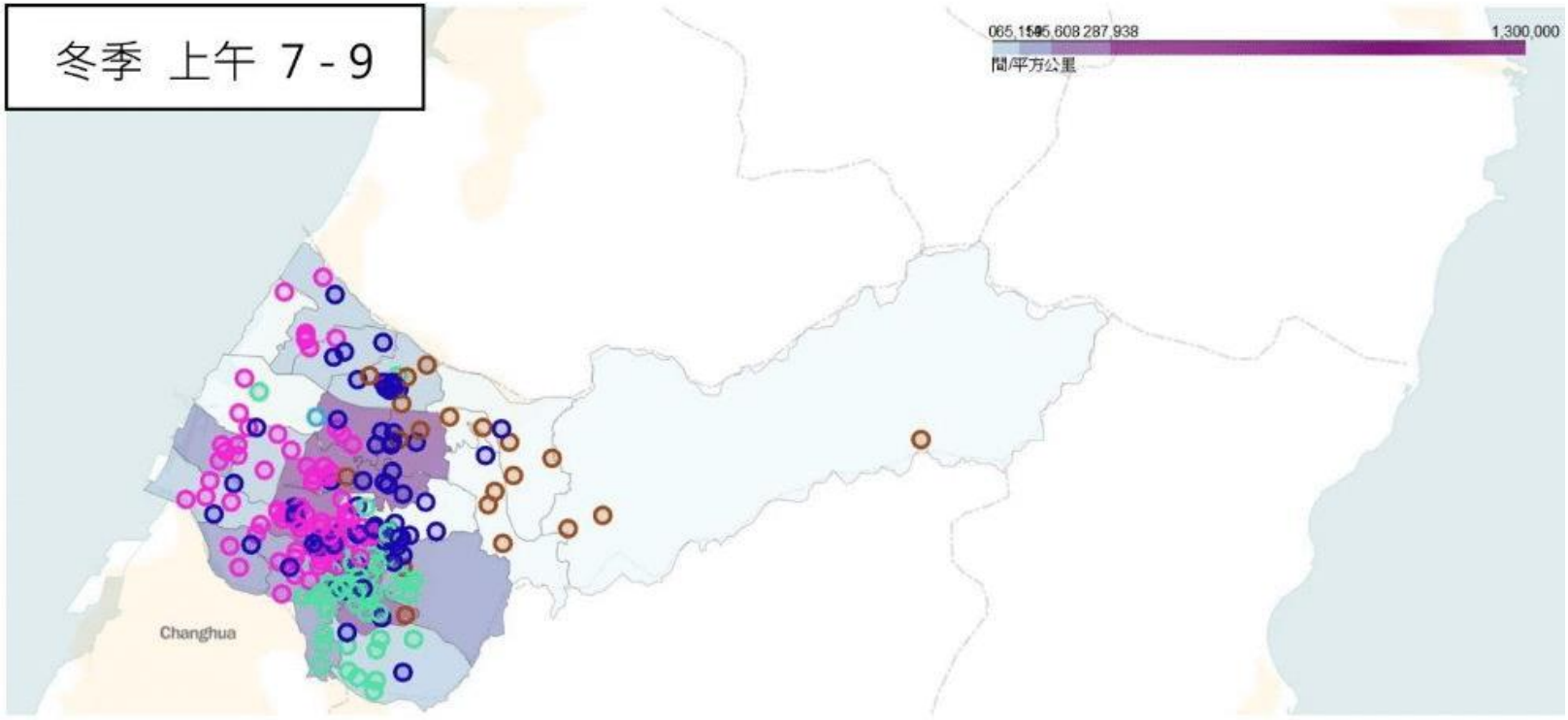
A plastics factory on fire in N. Kaohsiung on 2018/4/23…



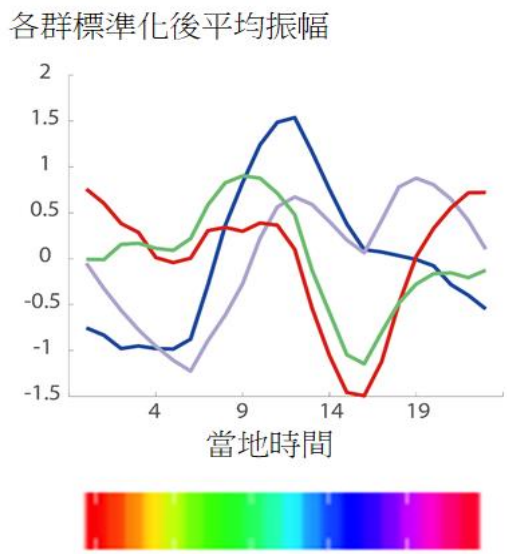


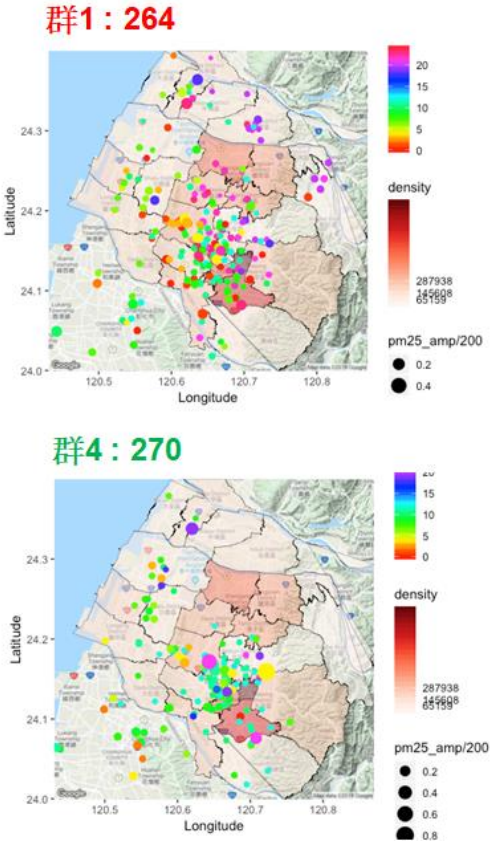
https://pm25.lass-net.org/

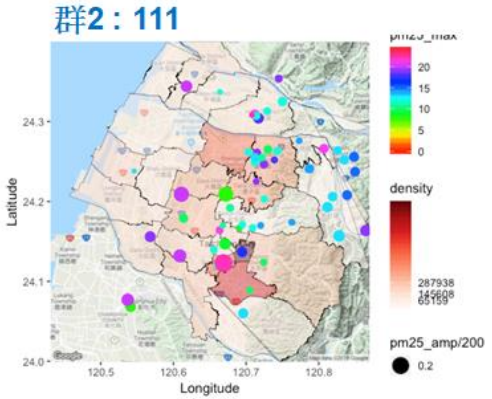
Seasonal Clusters (Taichung City)

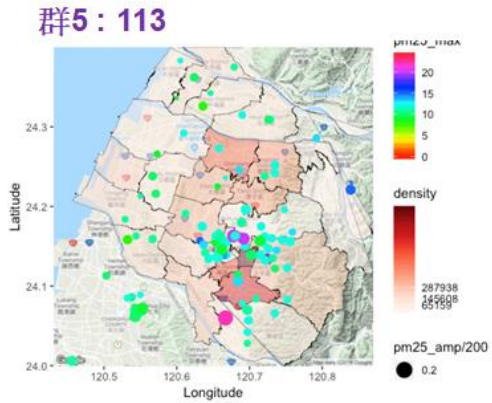


Intraday Clusters (Taichung City)

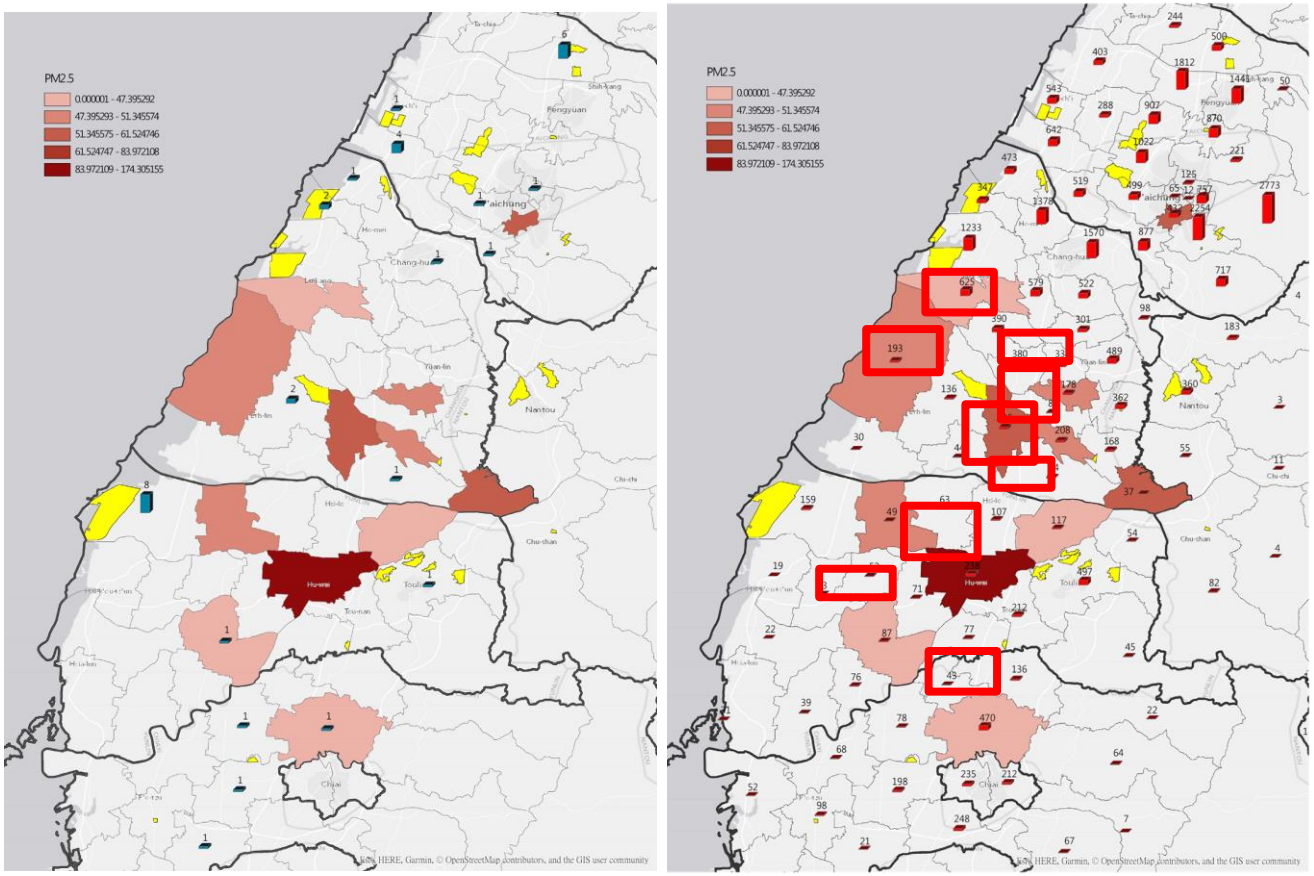








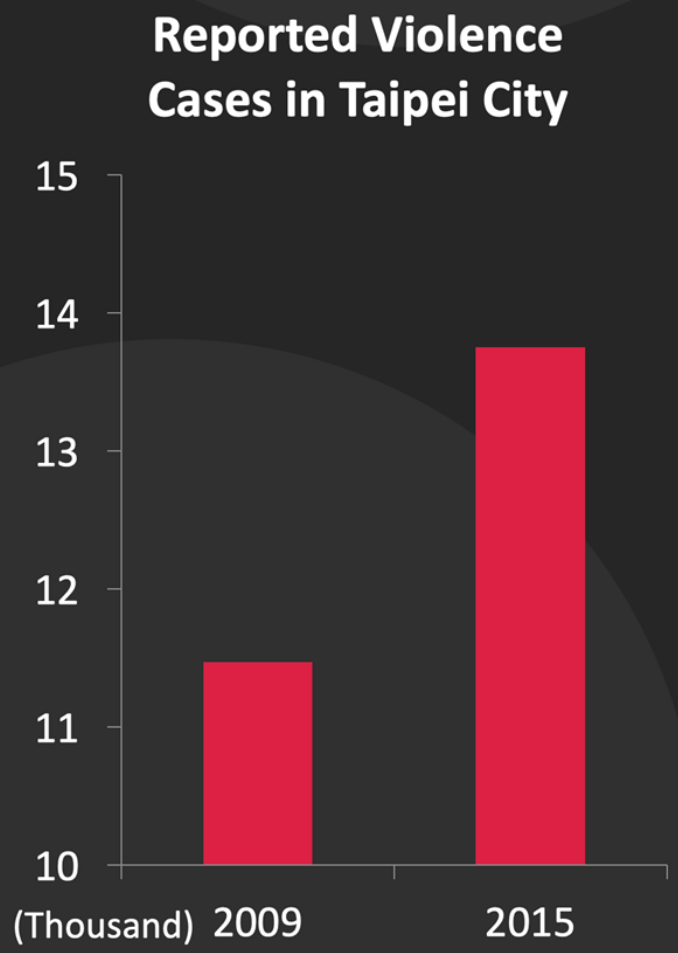
Airbox Location Recommendation for Monitoring Factory Air Pollution



Risk Prevention of Domestic Violence

https://pm25.lass-net.org/

Current Situation of the Domestic Violence (DV) in Taiwan



13,000 people need victim protection and support per year

1 social worker handles **116** cases

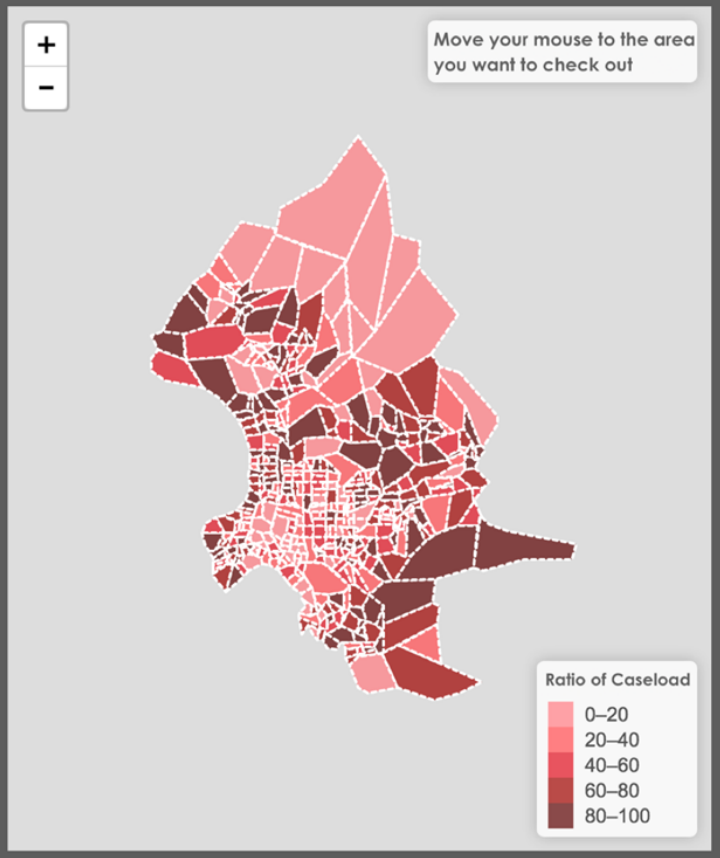
52% are Intimate Partner Violence cases

33% **are Repeat Victimization**

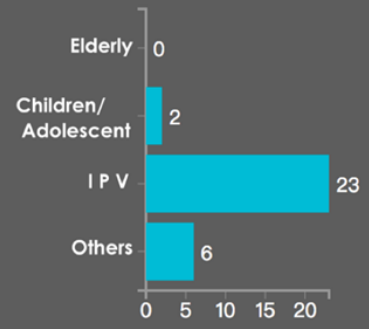
Community Level:
Interactive DV Prevention Risk Map

Taipei DV Prevention - Victim Risk Map & Dashboard

MapType | EventType | District | Village



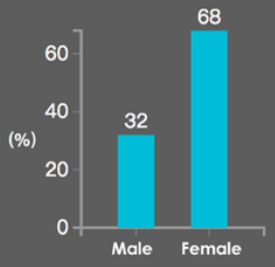
Caseload of the NQ Dist. CJ Vil. 31



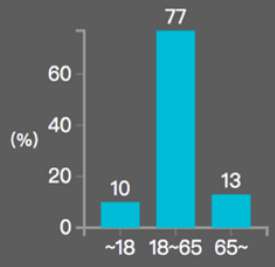
Number of DV victims from low- and middle-income families in Taipei city by village

Number of DV victims with disabilities or mental illness in Taipei city by village

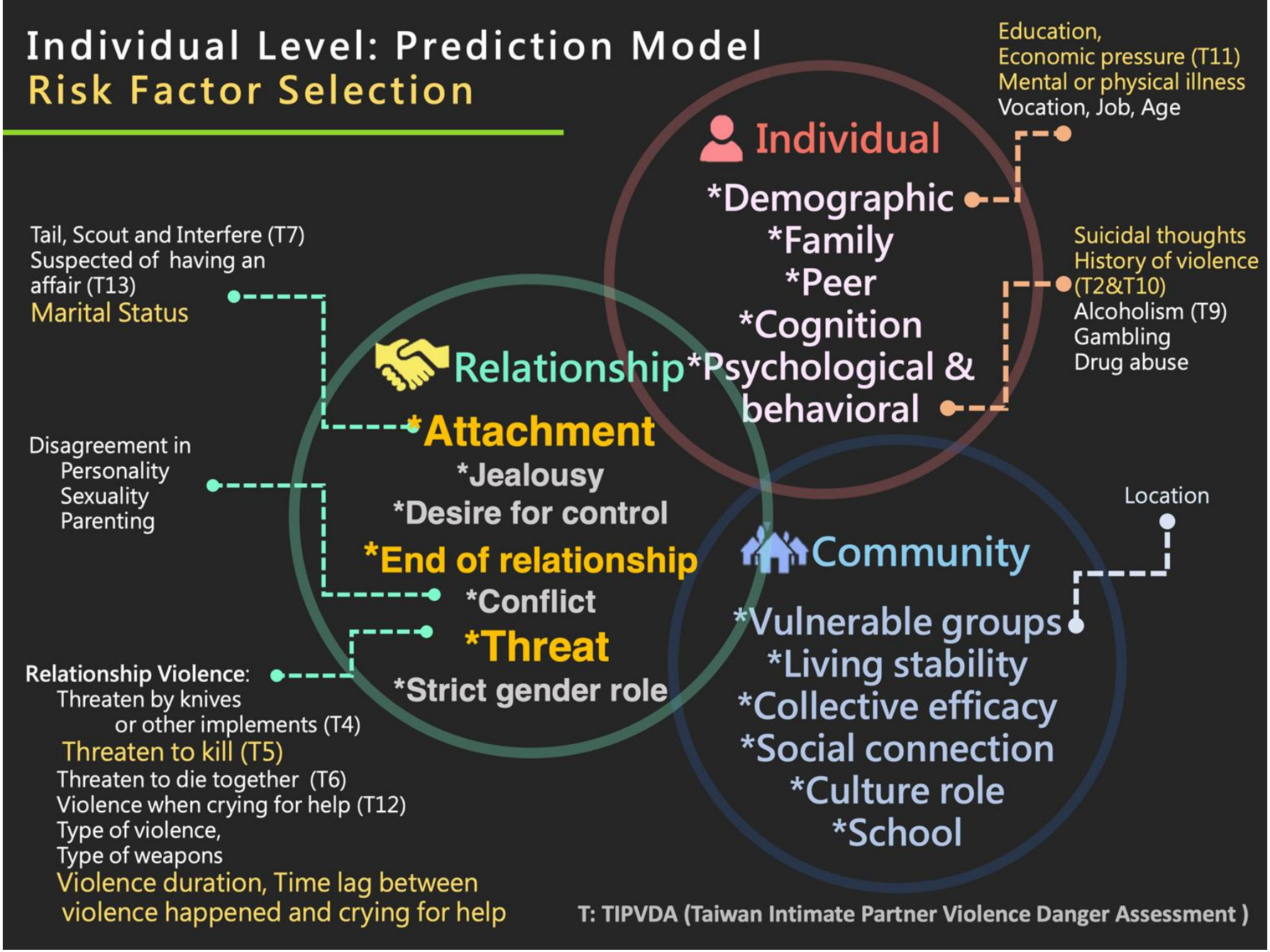
Victim Gender Ratio



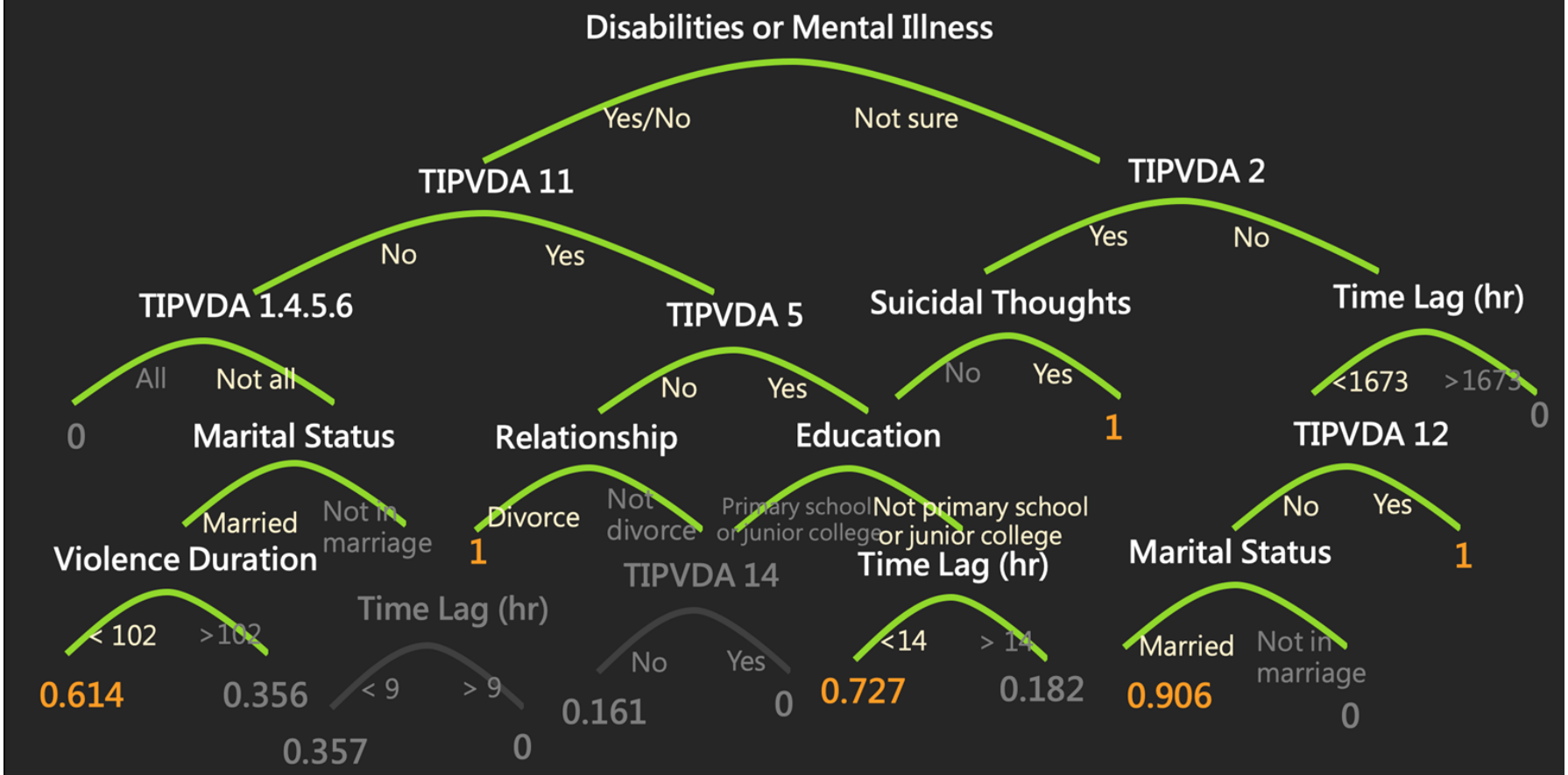
Victim Age Ratio



Individual Level: Prediction Model
Risk Factor Selection



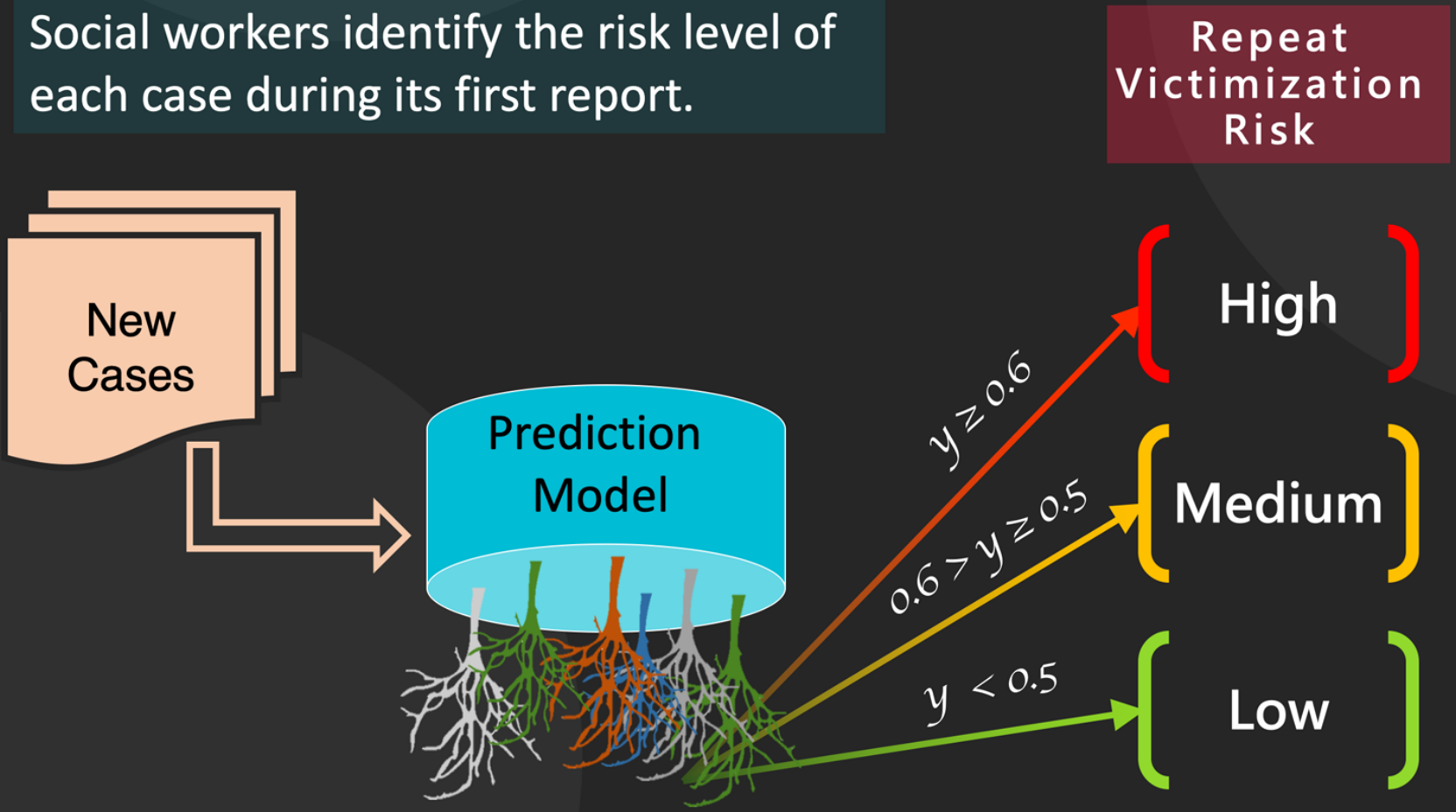
Individual Level:
Repeat Victimization Risk Prediction Model



T: TIPVDA (Taiwan Intimate Partner Violence Danger Assessment)

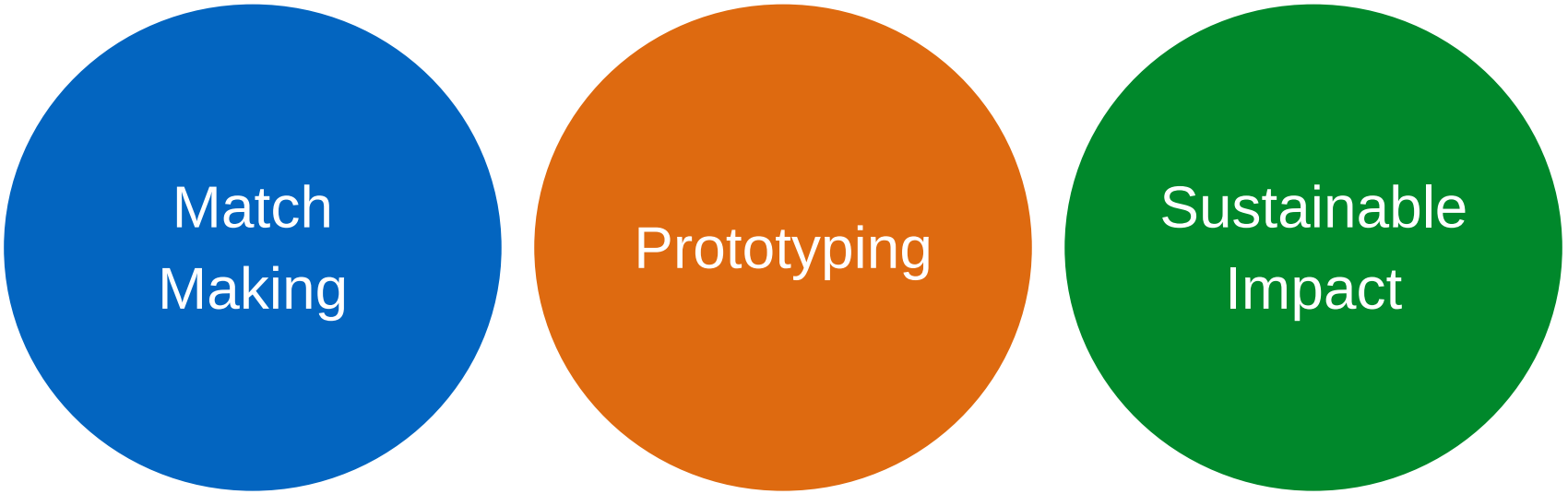
Individual Level:
Repeat Victimization Risk Prediction Model

Social workers identify the risk level of each case during its first report.



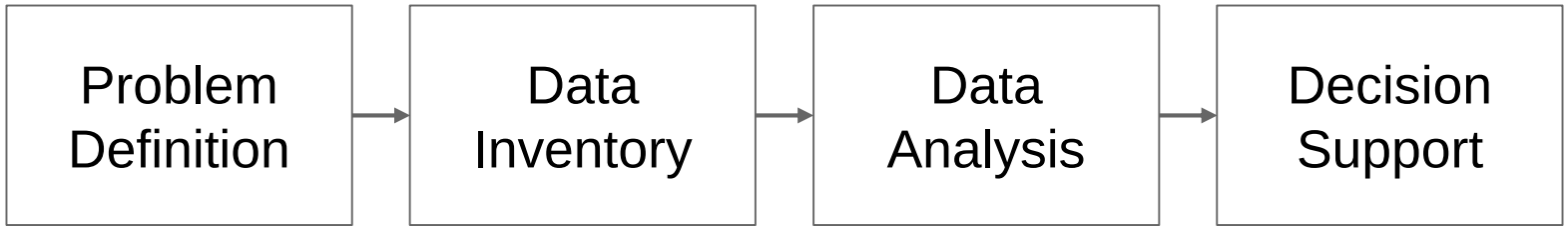
Lessons Learned

From Idea to Impact



Lessons Learned

- What is your goal?
- What are the proper measures for current status, trends or project/policy performance?
- Can the measures be found in, or calculated from your data sources?
- How to consolidate various data sources?
- Can you interpret the analysis results?
- What do you plan to do with the findings?



Next Steps

Next Steps

- Do you want to start a D4SG project?
- How can we cooperate?
  - Workshops or bootcamps
  - D4SG team projects
  - Learning data science with R