

## *POPGIS*

# Pollution Observation Platform for Monitoring PM<sub>2.5</sub> using Satellite Data

Project “Improving air pollution monitoring and management of Vietnam with satellite PM<sub>2.5</sub> observation”



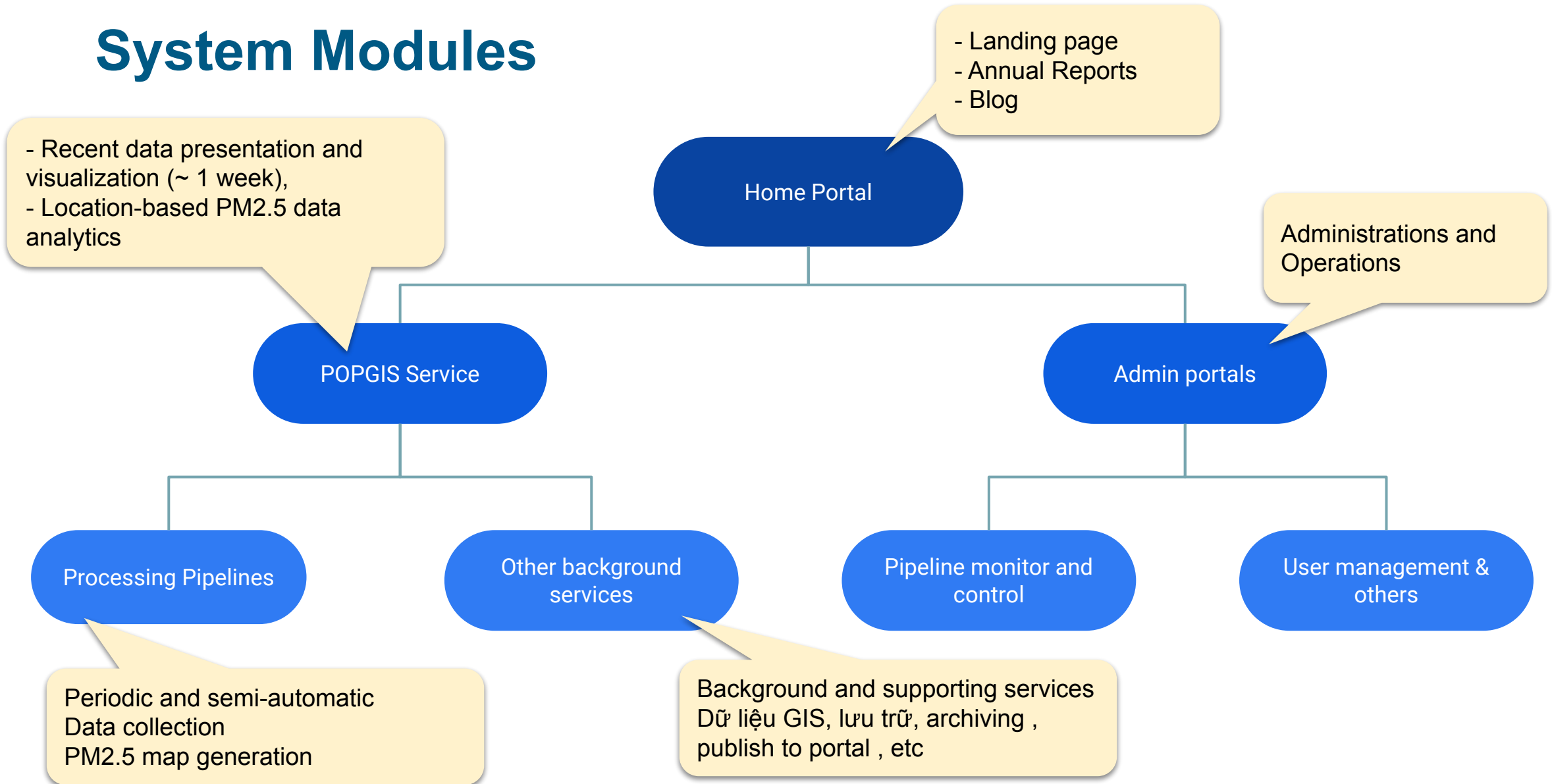
## Content

- ❖ Overview on POPGIS
- ❖ POPGIS in details
  - Architectures, Modules
  - How is it built
- ❖ A short tour on POPGIS
- ❖ Survey on POPGIS

## POPGIS Overview: Objectives

- I. Bring observation data and analytic results to everybody
- II. Knowledge dissemination on facts related to PM<sub>2.5</sub>
  - A. Facts and trendings on PM<sub>2.5</sub> related pollution
  - B. How PM<sub>2.5</sub> affects our health
  - C. Positively influencing our community
- III. Research result verification

## System Modules



POP - Pollution Observation Platform

POP GIS Annual Reports FAQ LaserPulse Admin Area Search this site...

## POP - A Pollution Observation Platform

POP là một sản phẩm thuộc khuôn khổ chương trình [LaserPulse](#)

[Source code](#)

Conducted by ...



Trường Đại học Công nghệ, ĐHQG HN (VNU UET)



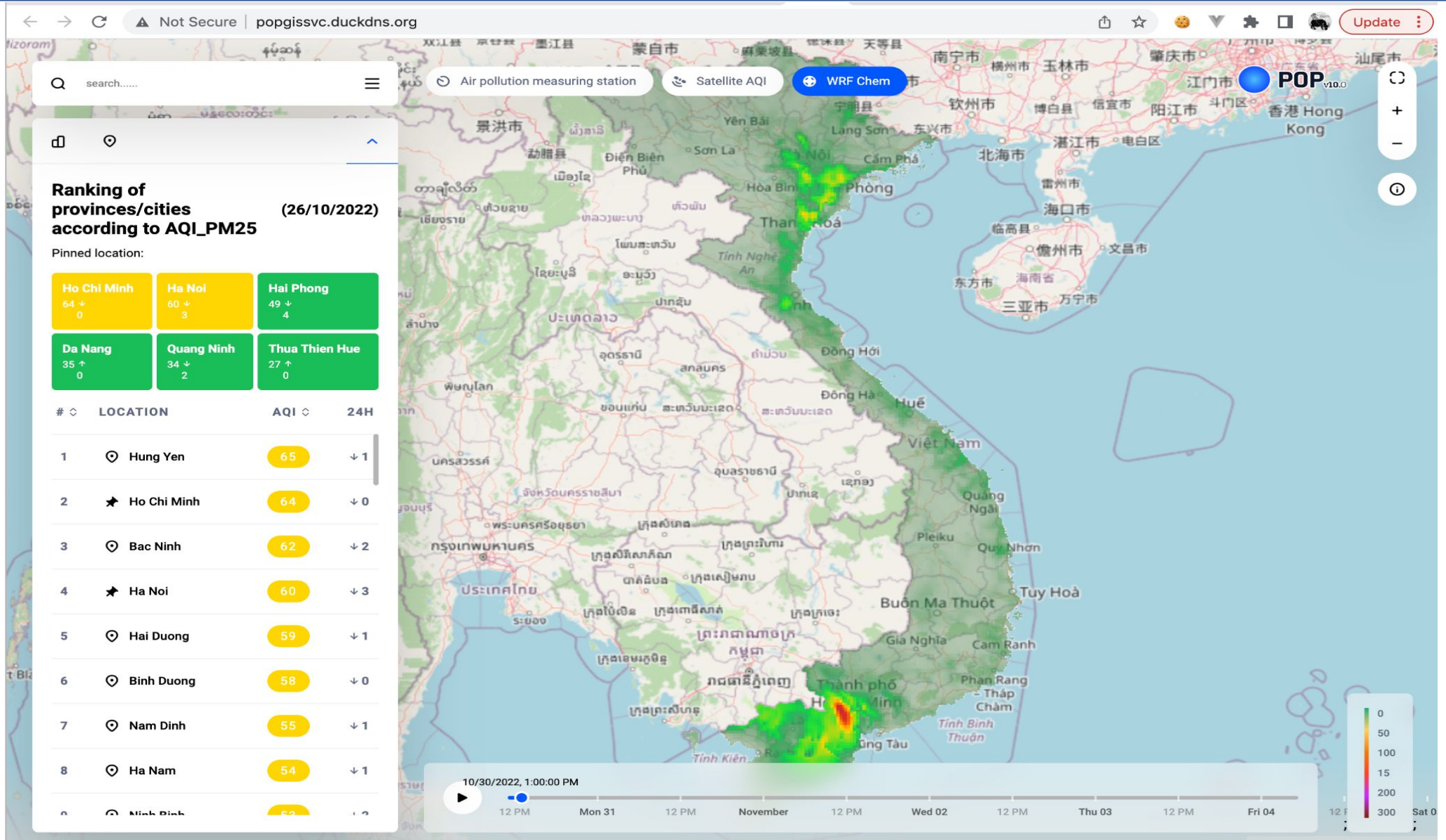
Trung tâm sống và học tập vì Môi trường và Cộng đồng.  
(L&L)



Khoa Khoa học Địa lý, trường Đại học Maryland  
(UMD, USA)




## POPgis Service



## User Management

← → ↻ ⚠ Dangerous | popgisadm.duckdns.org/list\_users 🔑 📄 ☆ 🌟 ⚙️ 🏠 🧑 Update

 POP 🇻🇳 ▼ 🧑 Tung Hoang ▼

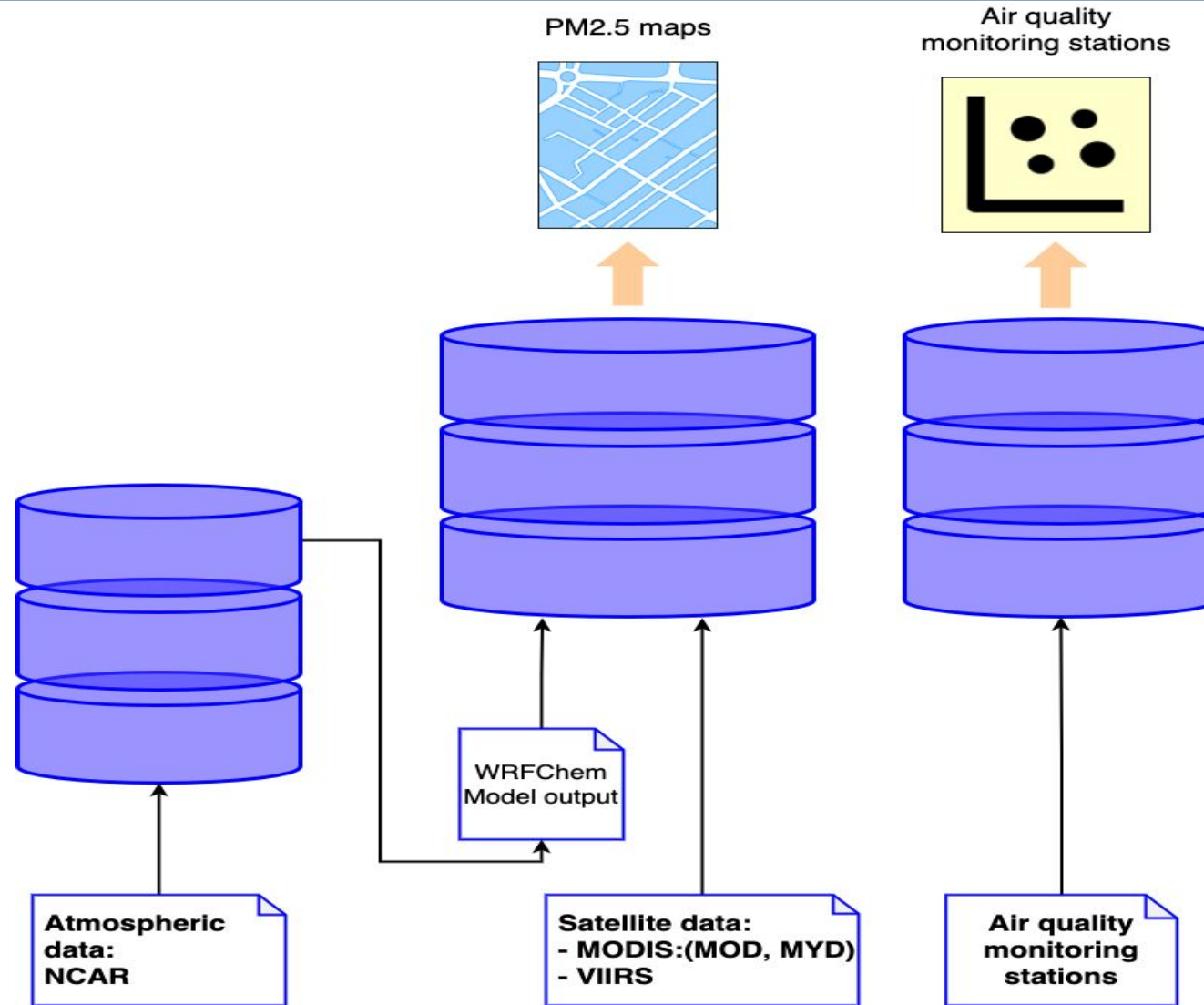
DANH SÁCH NGƯỜI DÙNG - 🗄

[+ Tạo mới tài khoản](#)

STT	Tên người dùng	Tên đầy đủ	Hoạt động mới nhất lúc	Quyền	Trạng thái
1	Tung Hoang		03/11/2022 13:13:59	Administrator	Hoạt động
2	tuannh	Nguyễn Hoàng Tuấn	24/10/2022 23:14:55	Administrator	Hoạt động
3	thanhntn		14/07/2022 21:21:19	Administrator	Hoạt động
4	hapv		30/06/2022 23:43:29	Member	Hoạt động
5	hieupdt		04/05/2022 16:34:03	Member	Hoạt động
6	Lecongphan19010023		14/04/2022 21:46:00	Member	Hoạt động
7	thang		25/03/2022 16:18:30	Member	Hoạt động
8	vuthang226		22/03/2022 20:37:59	Member	Hoạt động
9	nhatminhnt1502		18/03/2022 18:23:50	Member	Hoạt động
10	Khánh Duy		18/03/2022 18:23:49	Member	Hoạt động
11	truongnx		17/03/2022 16:13:03	Member	Hoạt động

POP v0.9.5.23

## Pipelines Processing



Processing pipelines

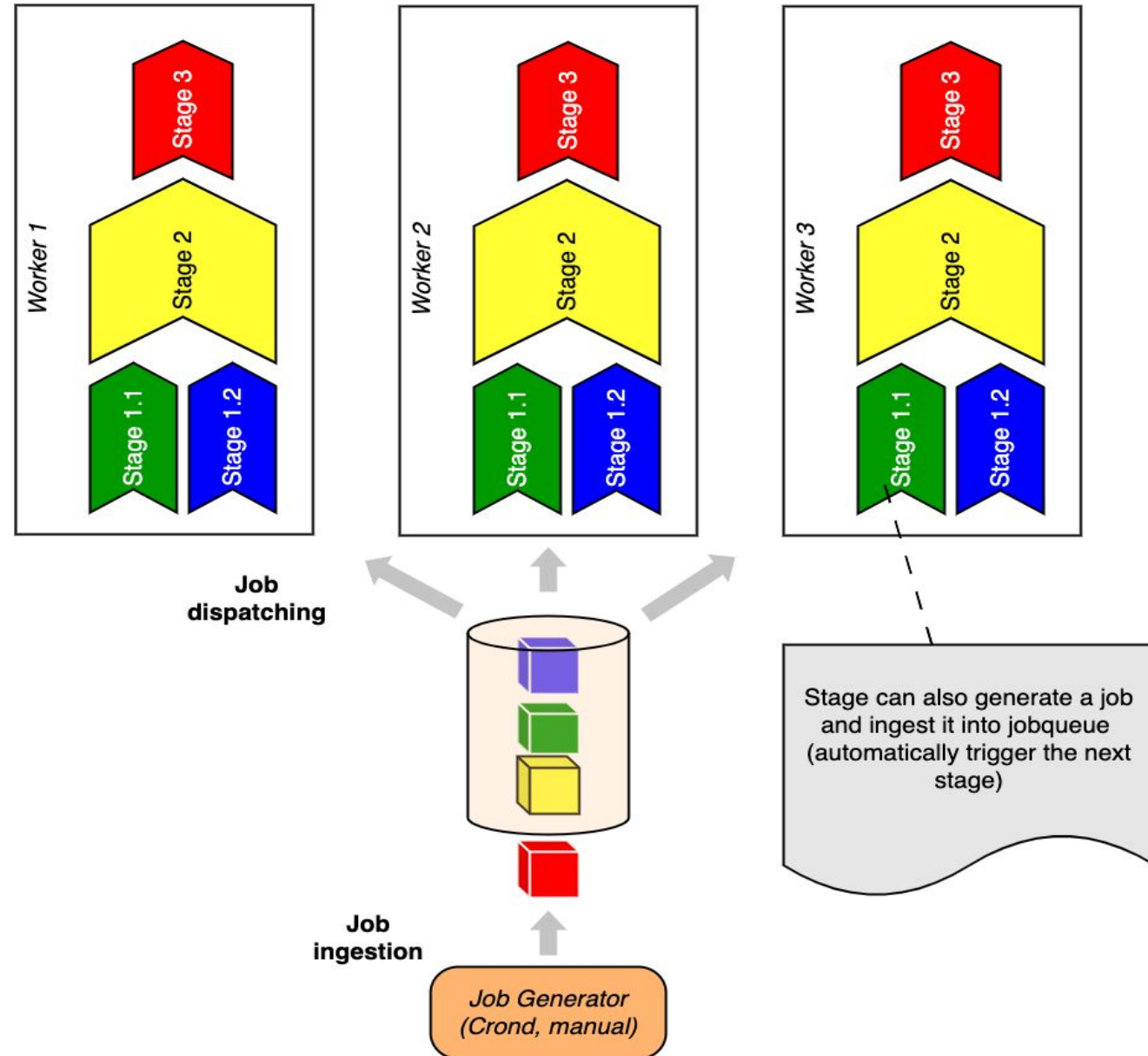


## Processing Pipelines - Challenges

- ❖ Difficult to control:
  - Internal dependency
  - External dependencies (irregular and error-prone data sources)
- ❖ Costly in operation and maintenance: some HUGE jobs
  - Takes 8hrs on 64 cores server (with parallel processing)
  - Generate about 100GB per day

## How pipelines are built?

- ❖ Each stage in a pipeline is developed as independent application with clearly defined input and output
  - Loosely coupled
  - Various technologies can be used: OpenMPI, WRF, HDF5, NetCDF, Scipy, Sci-learn
- ❖ Dependencies between stages are identified based on input data
  - Output of a stage can be input of the next stage
- ❖ A job queue play a crucial role in stage coordination



## Pipelines: Scheduling & Controls

The screenshot displays a web interface for monitoring pipeline jobs. At the top, there are five server status indicators: server5.cluster1(45419) (green, None), server2.cluster1(43586) (red, wrfchem\_pipeline.tasks.runWRF('20230126')), server4.cluster1(31573) (green, None), server3.cluster1(17336) (green, None), and server6.cluster1(25276) (green, None). The main content is divided into two sections for different dates.

**2023-01-26 00:00:00**

downloadNDVI [not run]	NDVIPreprocess [not run]	downloadData [not run]	ImgPreprocess [not run]						
downloadData [failed]	ImgPreprocess [not run]					WRFChemProcessing [not run]	PM25predict [not run]	UploadFile [not run]	UploadToDB [not run]
GetFNL [success]	runWPS [success]	runREAL [success]	runEMISS [success]	runWRF [running]	WRFpostprocessing [not run]			UploadFile [not run]	UploadToDB [not run]

**2023-01-25 00:00:00**

downloadNDVI [success]	NDVIPreprocess [success]	downloadData [not run]	ImgPreprocess [not run]						
downloadData [success]	ImgPreprocess [success]					WRFChemProcessing [success]	PM25predict [success]	UploadFile [success]	UploadToDB [success]
GetFNL	runWPS	runREAL	runEMISS	runWRF	WRFpostprocessing			UploadFile	UploadToDB

## Pipelines - Some important notes

- ❖ Job failures may happen occasionally
  - Data missing from source
  - hardware/network failures
  - Data acquisition interface changes
  - Software bugs
- ❖ If a job failed, ***rerun only the failed job and its dependents***
- ❖ Stage processing should be idempotent ← very important
  - Job rerunning should never affect related stages
  - Help in saving job rerunning

## Pipelines - Some important notes (Cont.)

- ❖ Role of human intervention:
  - Execution checking
  - Rerun job in necessary
  - Resource monitoring and data cleanup

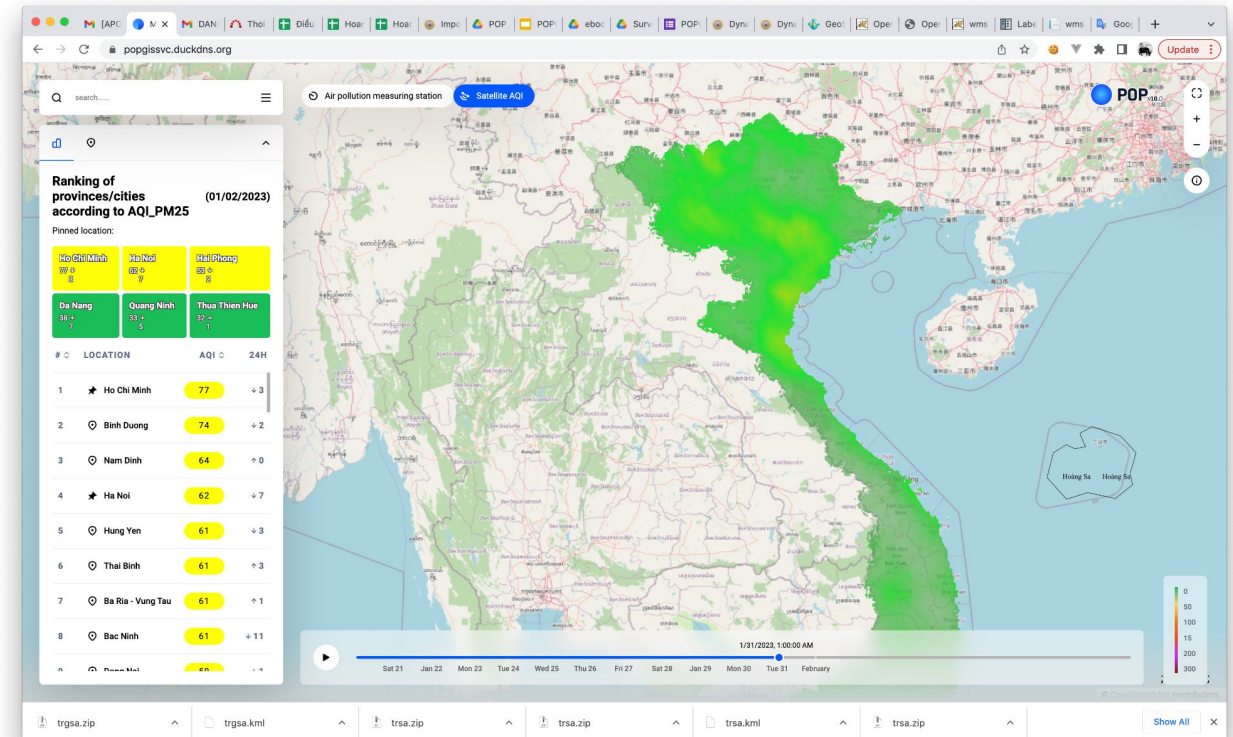
## A short tour on POPGIS

- ❖ POPGIS sites:  
<https://popgis.duckdns.org>  
<https://popgissvc.duckdns.org>



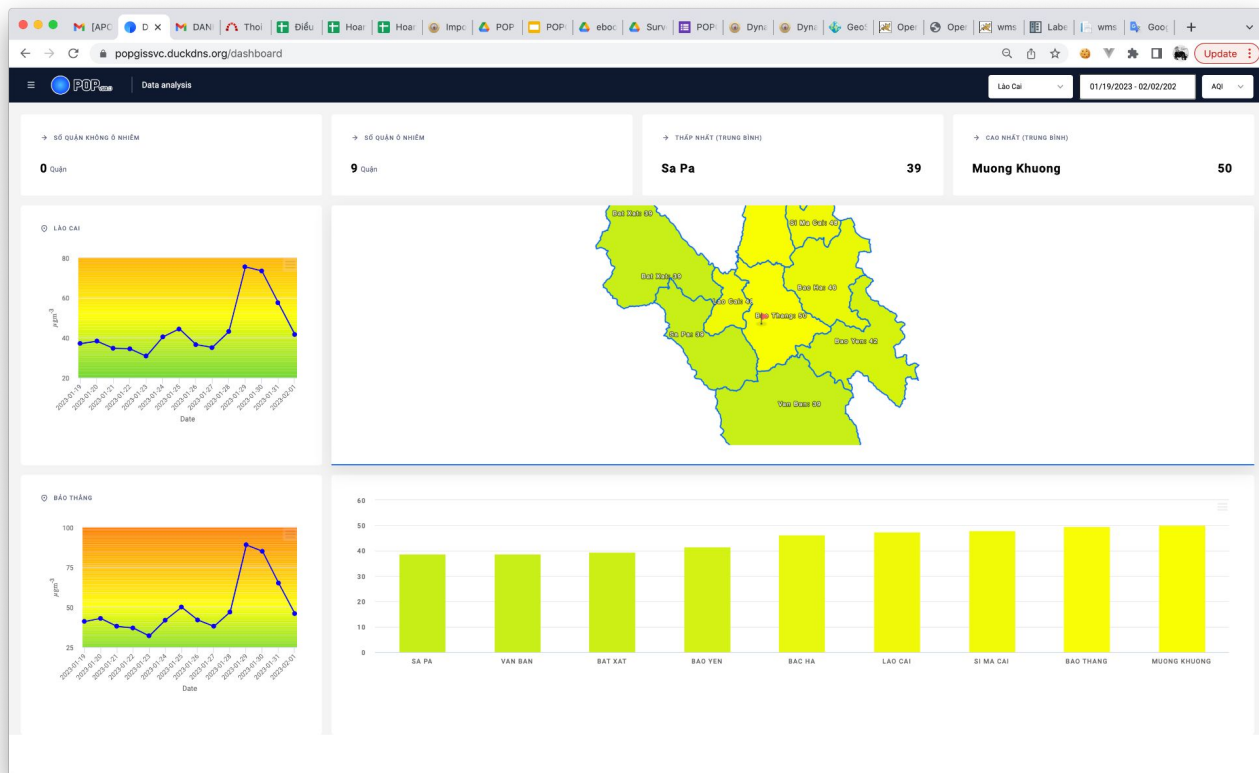
## POPGIS feature summary

- ❖ AQI ranking of cities/provinces (using yesterday data)
- ❖ Viewing and interacting with PM2.5 maps on Vietnam territory for **N** days
- ❖ Pick a location and observe PM2.5/AQI trending for **N** days



## POPGIS feature summary (cont.)

- ❖ Detail data analysis on every city/province
  - For any time range
  - Important metrics:
    - top polluted district,
    - least polluted district
  - Detail AQI/PM2.5 for a selected district



## Finally,

- ❖ We need comments on
  - Our PM2.5/AQI data and presentation sufficiently useful? Any important PM2.5/AQI related data missing?
  - How long should we keep PM2.5 maps visible ? currently it is 10 days
  - Convenient notification channels? Email? Messaging (Telegram/slack)
  - Performance issues ?
- ❖ Please visit POPGIS, experience it and take our survey!
  - <https://forms.gle/SfRRXYcJjZ6F5gqp8>

# Thank you !

