

From Data Layers to Decision Questions: A Question-Oriented Interface Model for an Air-Quality Digital Twin in Thailand

A UX Case Study of a PM2.5 Digital Twin for Thai Provincial Officers

R. Sukhahuta¹, S. Thipcharoen², T. Wongsas³, D. Sukhahuta⁴, K. Saksri⁵

¹Chiang Mai University · ²The Far Eastern University · ³RMUTL-Nan ·

⁴Maejo University · ⁵Aware Corporation Limited

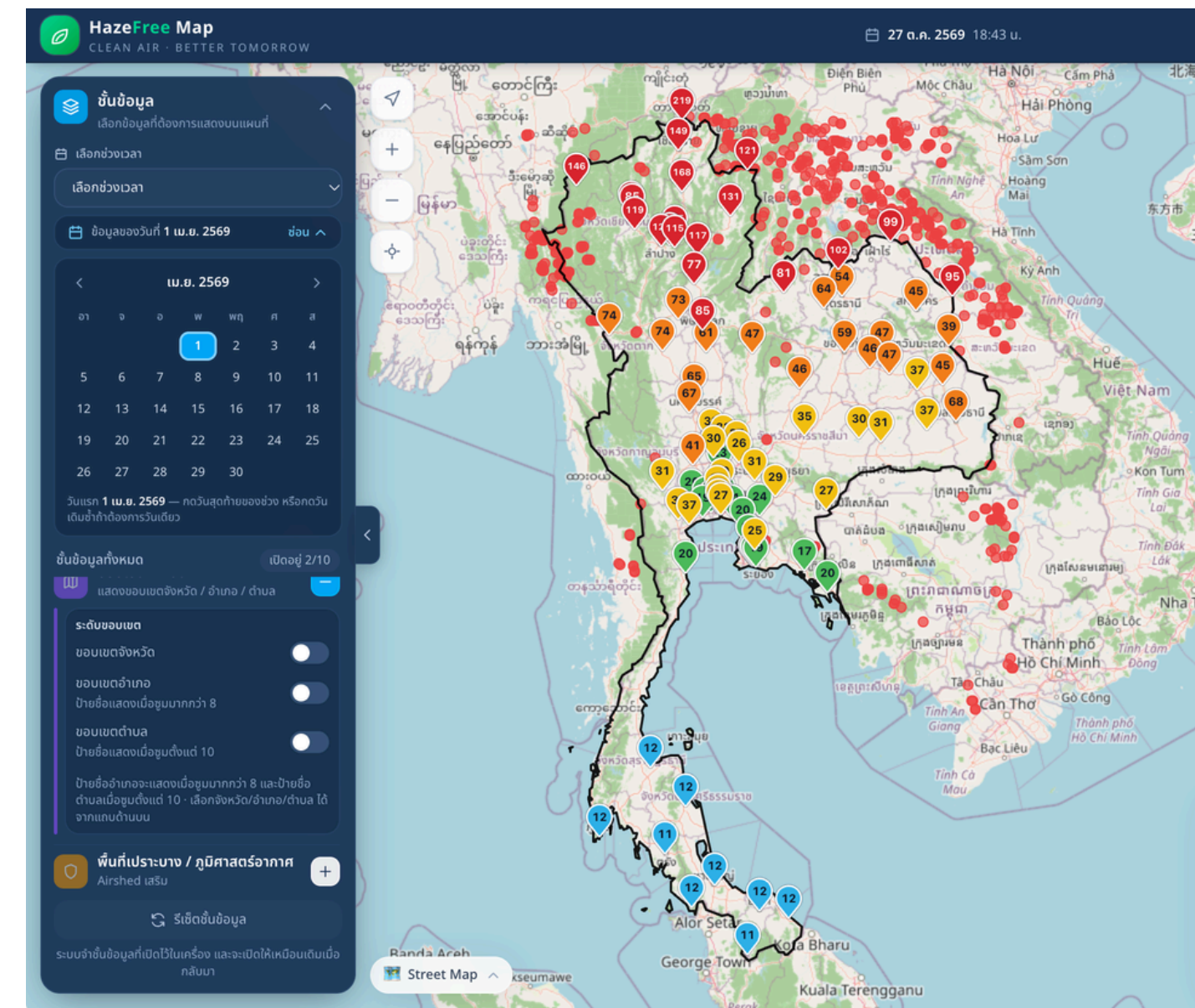
The 9th International Conference on Culture Technology 2026
Shanghai University of Engineering Science, Shanghai, China

Supported by NRCT grant N25B(PM)690041



The Problem: Data Mirrors vs. Decision Needs

- Motivating context: Northern Thailand has a severe PM2.5 episode every January-April.
- Officers must issue warnings, coordinate agencies and choose control measures—while reconciling multiple sources and uncertain forecasts.
- Most environmental dashboards are organized by data layer (stations, satellite, weather). This structure forces time-poor, accountable officials to manually assemble the data into a decision, rather than guiding them through a decision process.
- Machine learning provides a distribution of possible outcomes, but dashboards overwhelmingly display a single number. This invites over-confidence, leading officials to act on forecasts with unwarranted precision.



Most environmental platforms are organized by data layer (stations | satellite | weather | terrain | model). The user must assemble the decision.

Research Gap & Contributions



Gap: Interface-level studies of the tools Thai agencies actually use appear absent from the literature.

1

Question-oriented information architecture — each screen named after the decision question it answers, ordered by situation-awareness theory (Endsley, 1995).

2

Eight interface patterns for communicating machine forecasts to non-expert officials, derived by coding every panel in the implemented interface.

3

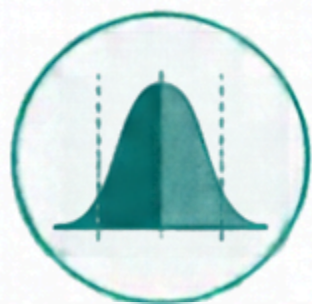
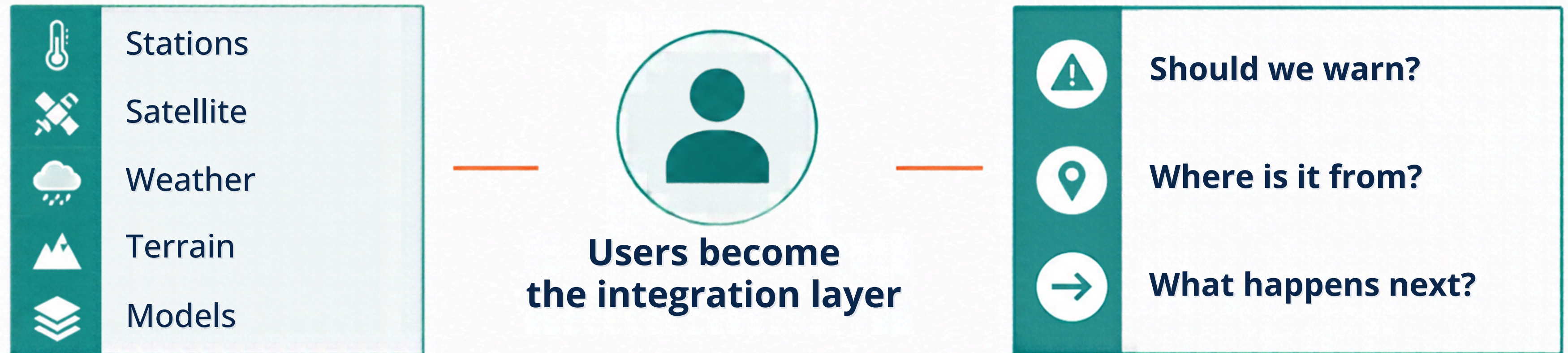
Expert usability inspection by 30 UX/UI evaluators, with redesign priorities.

The Evaluation Based on Three Decisions

Indicator	Role in Forest Fire Situations	Tools / Satellites Used for Measurement	Limitations / Caveats
Number of Hotspots <i>(Thermal anomalies)</i>	"Current fire indicator" Used to identify active high-heat spots or real-time burning areas.	Thermal IR wavelength satellites, e.g., NASA FIRMS (MODIS, VIIRS).	May be false positives from global thermal anomalies, hot asphalt, or industrial flares.
PM2.5 <i>(Fine particulate matter)</i>	"Atmospheric impact" Smoke and fine particles released from burning vegetation, spreading according to wind patterns.	Ground monitoring stations, Low-cost sensors, Satellite Aerosol Optical Depth (AOD).	Dust/smoke may disperse far from the fire source depending on atmospheric pressure and wind direction.
Burn Scar <i>(Burned area)</i>	"Physical impact on land" Evidence of damage to forest or agricultural areas after a fire.	Optical satellite and Radar, e.g., Sentinel-2, Landsat (using NBR, NBRT indices).	Must wait for cloud cover to clear or the fire to extinguish before clear boundaries of the damaged area are visible.

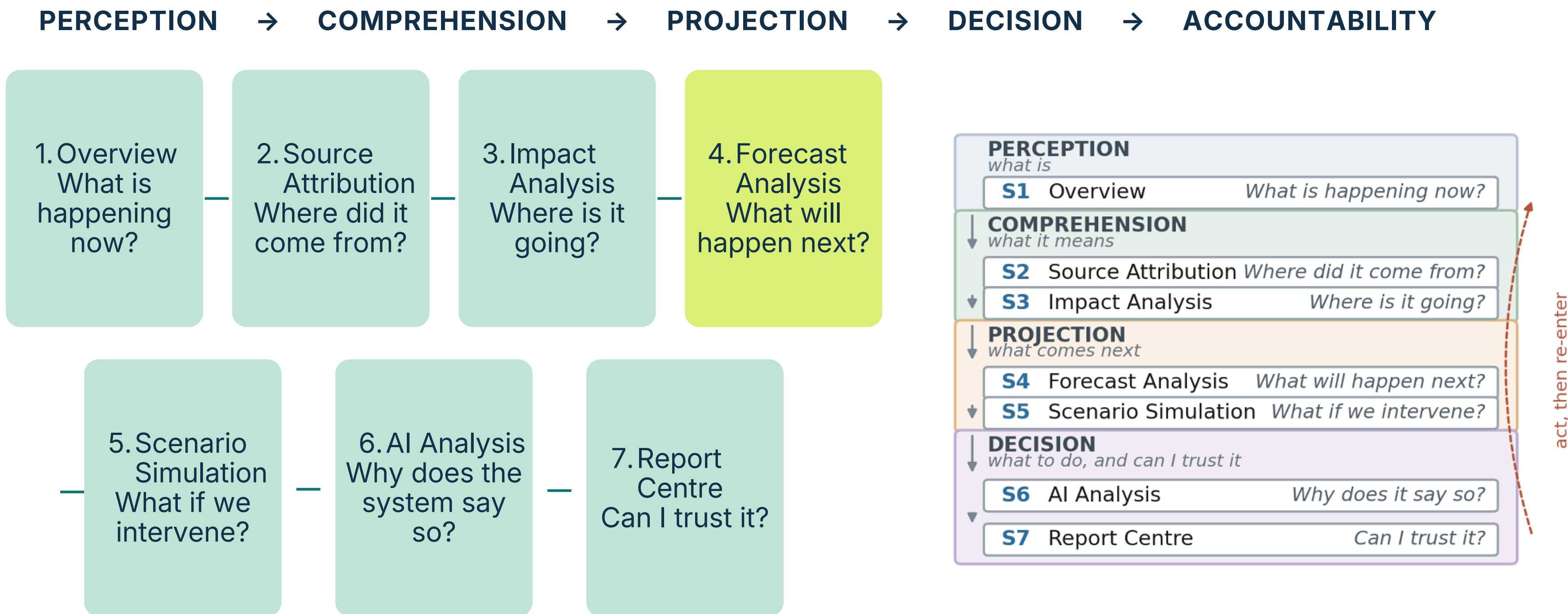
The Decision Gap

Dashboards mirror data layers.
Decisions do not.



Probabilistic forecasts shown as single numbers create false certainty—the same communication failure exposed by the classic 'cone of uncertainty'

Seven analysis steps for a decision pathway



Shared layout: left rail = state + confidence · centre = map + timeline · right rail = answers + hand-off

The Concept of Framework Structured Methodology Designed

P1

Decision-question
framing

- Clearly and sharply defining the core question or decision that needs to be made before diving into analysis or conclusions.

P2

Plain-language gloss

- Translating technical jargon, complex statistical outputs, or dense bureaucratic/academic phrasing into accessible, everyday language.

P3

Confidence disclosure

- Explicitly stating the level of certainty, reliability, or risk associated with a recommendation, forecast, or finding

P4

Provenance tags

- Visual labels that explicitly distinguish raw **measured** data (e.g., from a physical ground station) from **derived** data (e.g., estimated by the AI model, interpolated between stations, or forecasted)

P5

Scalar decision
proxies

- Visually distinguishing raw measured data from derived, modeled, or estimated values so users know the source of the number.

P6

Co-located
explanations

- Compressing complex model outputs (like a transport matrix) into one simple, bounded score (e.g., 0–100) to support a quick choice.

P7

Shared temporal
grammar

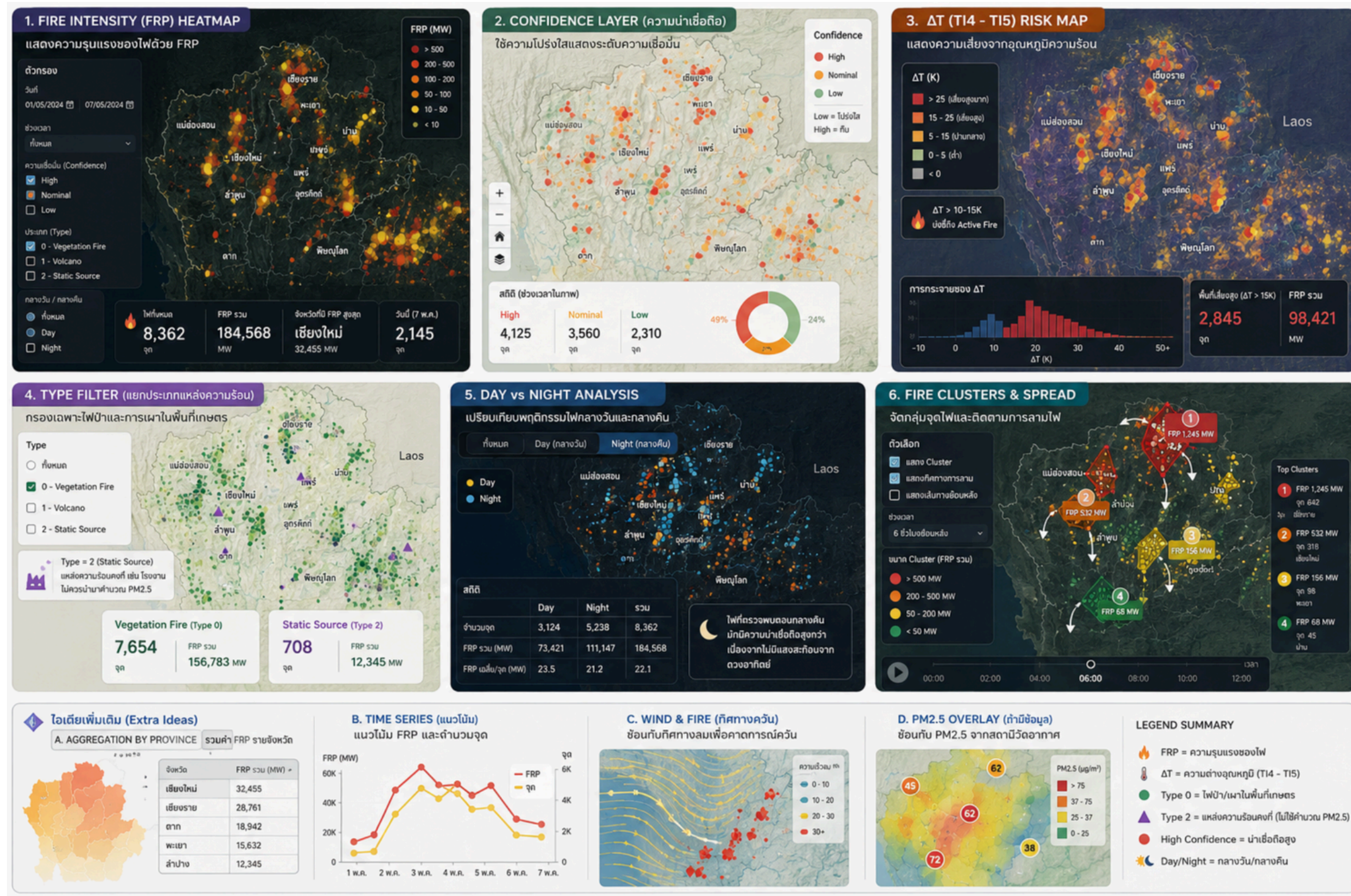
- Placing the reason beside the value (why the system says so) rather than hiding it in a separate menu.

P8

Threshold anchoring

- Tying visual encodings (like color bands) to national regulatory standards rather than to arbitrary data quantiles.

Our First Prototype of UX/UI Designs



- S1: Overview – National map, regional means, and active warnings to answer "What is happening now?"
- S2: Source Attribution – Ranked upwind contributors and smoke paths to answer "Where did it come from?"
- S3: Impact Analysis – Downwind provinces, expected increases, and arrival times to answer "Where is it going?"
- S4: Forecast Analysis – Peak values, timing, and exceedance risk by lead time to answer "What will happen next?"
- S5: Scenario Simulation – Before-and-after comparisons of control measures to answer "What if we intervene?"
- S6: AI Analysis – Driver rankings and narrative cause accounts to answer "Why does the system say so?"
- S7: Report Centre – Past accuracy records, event detection, and export tools to answer "Can I trust it?"
- S8: Settings – System configuration and user preferences (The supporting utility screen).

The Seven Decision Questions

#	Screen	Question	What it principally presents
S1	Overview	What is happening now?	National map, regional means, active warnings, 7-day outlook
S2	Source Attribution	Where did it come from?	Ranked upwind contributors ($\mu\text{g}/\text{m}^3$ and %), smoke paths
S3	Impact Analysis	Where is it going?	Downwind provinces, expected increase, time to arrival
S4	Forecast Analysis	What will happen next?	Peak value and time; exceedance risk by lead time
S5	Scenario Simulation	What if we intervene?	Baseline vs. measure; per-measure contribution; benefit estimate
S6	AI Analysis	Why does the system say so?	Driver ranking, narrative cause account, prioritized actions
S7	Report Centre	Can I trust it?	Past accuracy by lead time/province, event detection, export

Evaluating the UX/UI: Coding the Interface Patterns

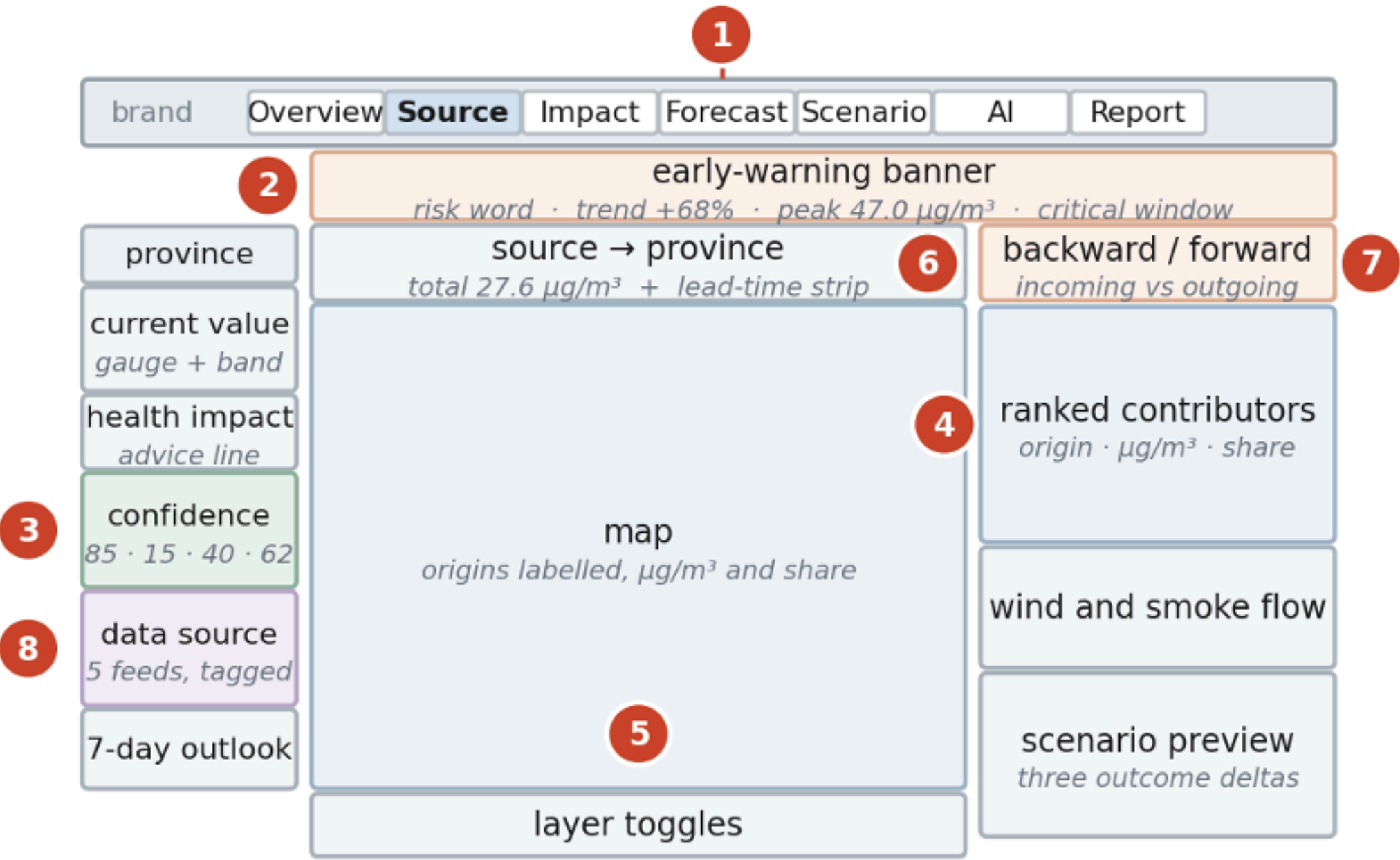
To evaluate whether our question-oriented design actually worked, we couldn't just rely on opinion. We needed a systematic, empirical method. We used an inductive qualitative content analysis

- The Unit of Analysis:
 - Seven screen dashboard is broken into 97 distinct panels (regions or UI elements like maps, charts, or banners).
- The Coding Process:
 - We did this in three passes.
 - i. Described what every panel was doing.
 - ii. Grouped similar functions into candidate patterns.
 - iii. Compared these against established design guidelines. We only kept a pattern if it appeared on at least two different screens.
- The Results (Table & Heatmap)
 - 8 distinct interface patterns (P1 to P8). The table shows the definition of each and how many screens/panels they appear on. The heatmap (Fig. 7) shows the distribution: 91 out of 97 panels use at least one pattern, and 6 of the 8 patterns appear on every single screen.

S1 Overview	2	3	1	3	–	1	4	5
S2 Source Attribution	4	5	1	1	–	–	3	5
S3 Impact Analysis	5	4	2	1	3	1	4	1
S4 Forecast Analysis	3	4	4	4	1	1	4	3
S5 Scenario Simulation	8	1	4	3	1	–	3	1
S6 AI Analysis	4	3	6	2	1	5	3	1
S7 Report Centre	1	6	2	7	–	1	1	1
	P1	P2	P3	P4	P5	P6	P7	P8

cell = panels on that screen carrying the pattern; column totals 6–27

Conceptual Framework for Designing the Wireframe



1. Situation Awareness

The human progression of Perceiving elements in the environment, Comprehending what they mean, and Projecting future states.

2. Cognitive Load Theory (UX/HCI)

Keeping the layout frame (left rail, map, right rail) identical across all screens, we eliminate extraneous load (energy spent finding controls), allowing the user to devote all mental resources to the actual decision.

3. Decision-Support Architecture

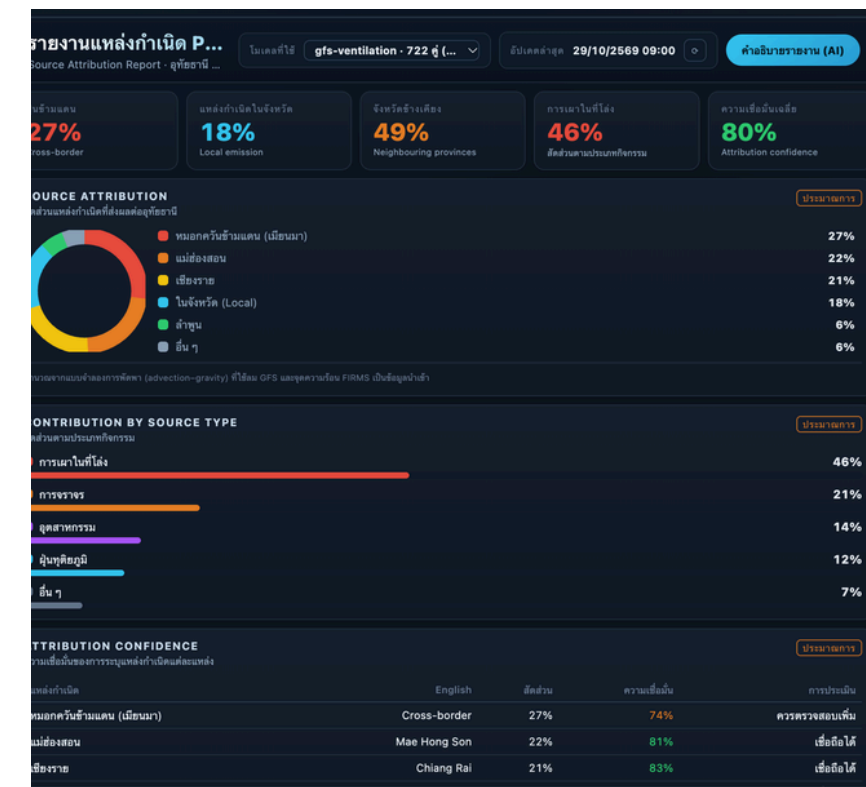
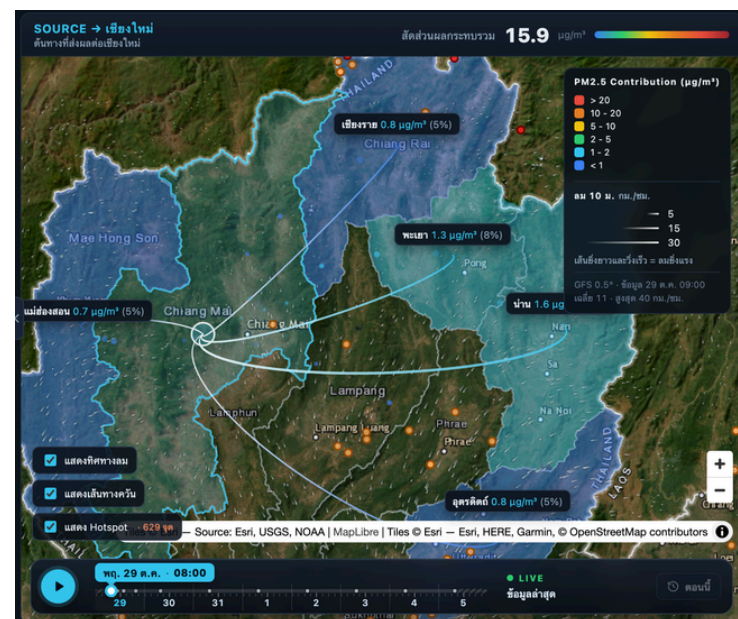
A system designed to assist—rather than replace—human judgment by organizing data around a specific choice.

Evaluator Feedback: From Silos to Systemic Fixes

Topic	Problem Found by Evaluators	Design Fix Implemented
1. False Precision	Values shown to 0.1 µg/m³ (e.g., 28.1), but the model's own Mean Absolute Error is 2.1 µg/m³.	Round to whole units and display the error band beside the value.
2. The "Zero" Trap	Missing data (e.g., mixing-layer height) is rendered as "0.0", making an absent value look like a real measurement.	Create a distinct "No Data" visual token; never render missing data as a number.
3. Misleading Metrics	A 40% weighted physics share sits in the exact same visual dial as an 86% data-quality score, implying poor performance.	Rename share labels and visually separate component weights from confidence scores.
4. Unit Errors	An upper-casing transform destroyed the "µ" prefix in map legends, labeling micrograms as milligrams (wrong by 3 orders of magnitude).	Lock unit strings from text transformations; render units from a single secure token.
5. Unaccounted Math	Source screen attributes only 85% of pollution and leaves the rest blank; Analysis screen renormalizes it to 99%.	Attribute data on one unified system-wide basis and show the unattributed remainder explicitly.
6. Cross-Screen Conflicts	"Population at Risk" reads 0 in a header tile, but a map callout on the same screen reads ~195,000.	Derive all labels from one shared definition; inspect the UI along the decision path, not just screen-by-screen.

Five Levels of End-to-End System Functionality

- **Monitor** – Observe the current situation (Dashboard, Map)
- **Diagnose** – Analyze sources and causes (Source Contribution, AI Analysis)
- **Predict** – Forecast 7 days, provide early warnings, and assess cross-provincial impacts
- **Prescribe** – Simulate measures (reduce burning by 30/50/100%, royal rainmaking, traffic, etc.) and recommend decisions
- **Report** – Generate 7 types of reports, verify model accuracy, and export/schedule automatic delivery"



Conclusion

- In this study, we design User Interface Design for PM2.5 digital-twin prototype for Thailand consisting of 174 monitoring stations, satellite fire detections, weather forecasts, terrain data → hourly forecasts up to 7 days, source attribution, early warning, and what-if simulation.
- The interface was divided into seven decision questions (not data layers) shifts integration work from the user to the system.
- We are dividing eight interface patterns into 97 panels for presenting forecasts to non-expert officials; 91 of 97 panels use at least one, and six patterns appear on every screen.
- A 30-evaluator expert inspection produced 14 findings (6 major) — most about uncertain values shown with unwarranted precision, in the wrong unit, or inconsistently across screens.
- Users may not tell you what the UI should look like, but their operational needs tell you what the dashboard must include — the questions it must answer, the confidence it must disclose, and the limits it must make visible.

**Thank you for
your attention**

Questions?

ICCT 2026 · 29-31 October 2026

Supported by NRCT grant N25B(PM)690041

