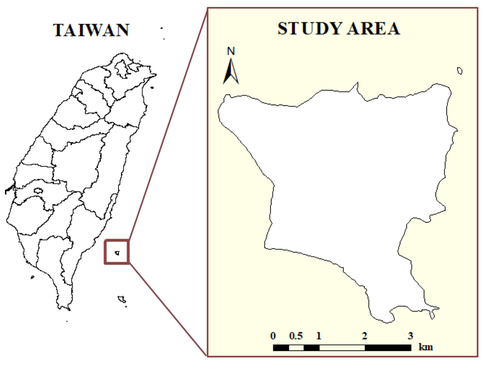
Unveiling Local Environmental Concerns Through Text Mining and DPSIR: A Case Study Using LLM-Based Analysis

Tsai-Chu Chen[1], Cheng-Wei Hung[2], Mei-Hua Yuan[3]

[1]International Degree Program in Climate Change and Sustainable Development, National Taiwan University, Taipei, Taiwan
[2]Graduate Institute of Environmental Engineering , National Taiwan University, Taipei, Taiwan
[3]Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan

ABSRACT

This study uses Lyudao as a case to explore how Large Language Models (LLMs) can enhance qualitative analysis of social–ecological systems. We developed and applied the *GreenMine* system, which integrates LLMs to conduct theme analysis within the DPSIR (Drivers–Pressures–States–Impacts–Responses) framework on in-depth interview transcripts. *GreenMine* operates in a human-in-the-loop manner for semantic relationship identification and includes a repeatable evaluation mechanism, presenting results as visualized DPSIR graphs. Three contextualized themes were designed—"Community Resilience," "Sustainable Tourism," and "Waste Management"—with DPSIR indicators defined based on literature.



METHODS

Iterative Evaluation

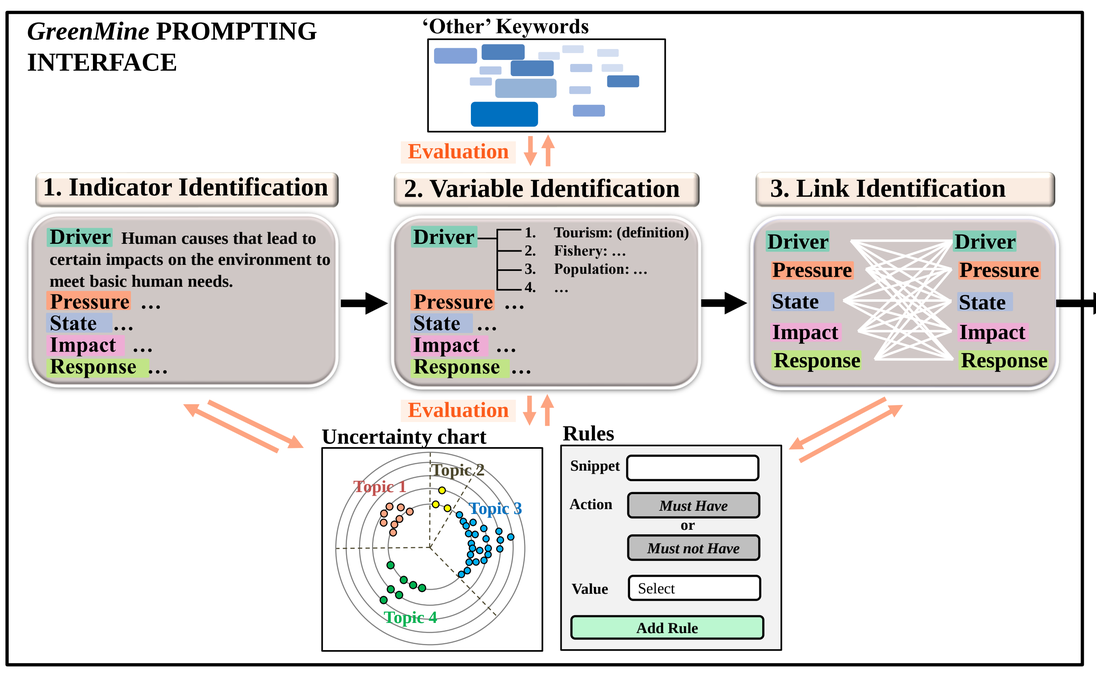
Issue Identification
- **Data collection**
  - News headlines
  - Local council records
- **Identify key issues using** NVIVO

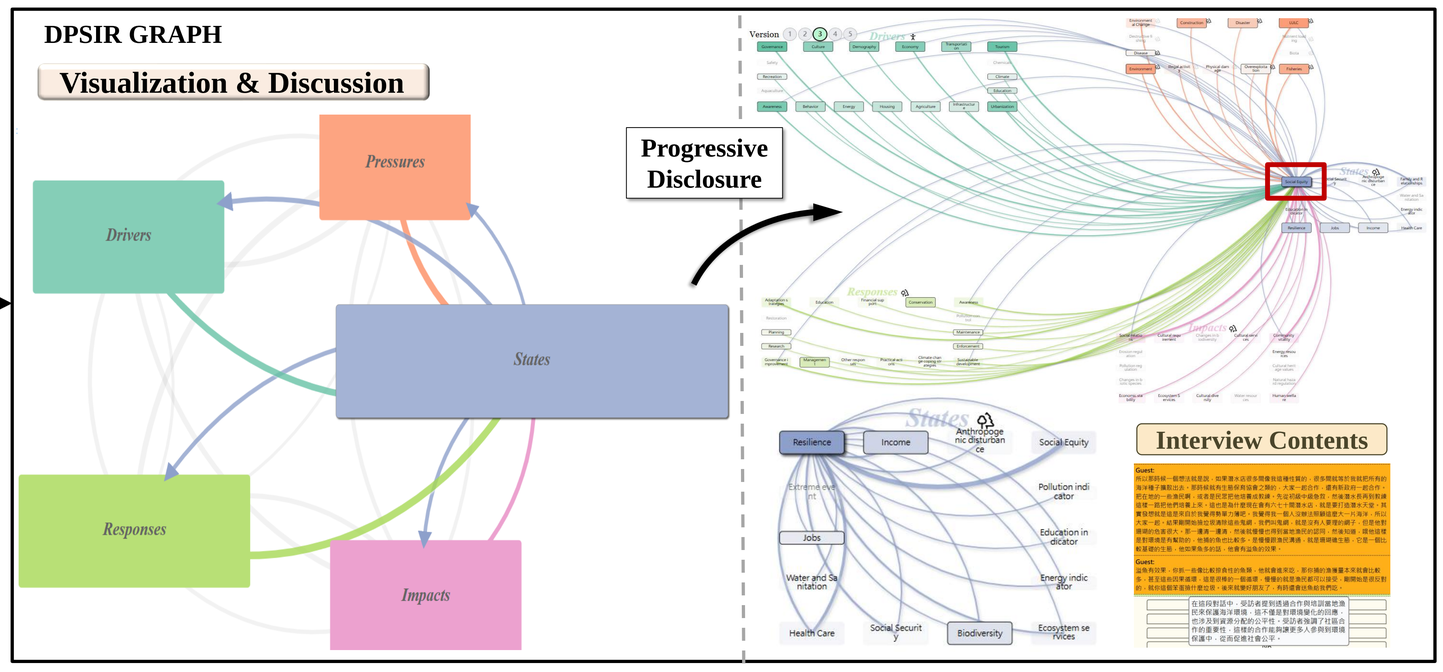
Interview & Data Preprocessing
- **Interviews transcription at Lyudao**
- **Content segmentation**
  - The corpus was segmented into 598 snippets.

Theme Design
- **Three Themes (& EEA)**
  - Distinct indicator definitions
- **Shared standardized variable list**
  - Enables consistency & refinement

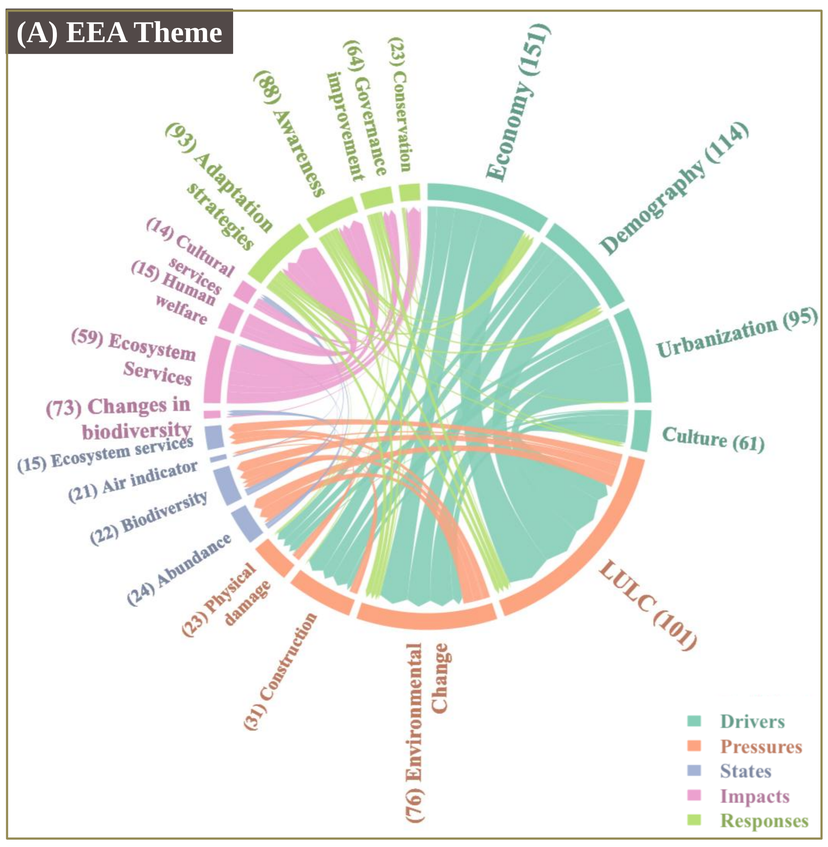
Prompt Execution
- ***GreenMine* System**
  - Indicator Identification
  - Variable Identification
  - Link Identification

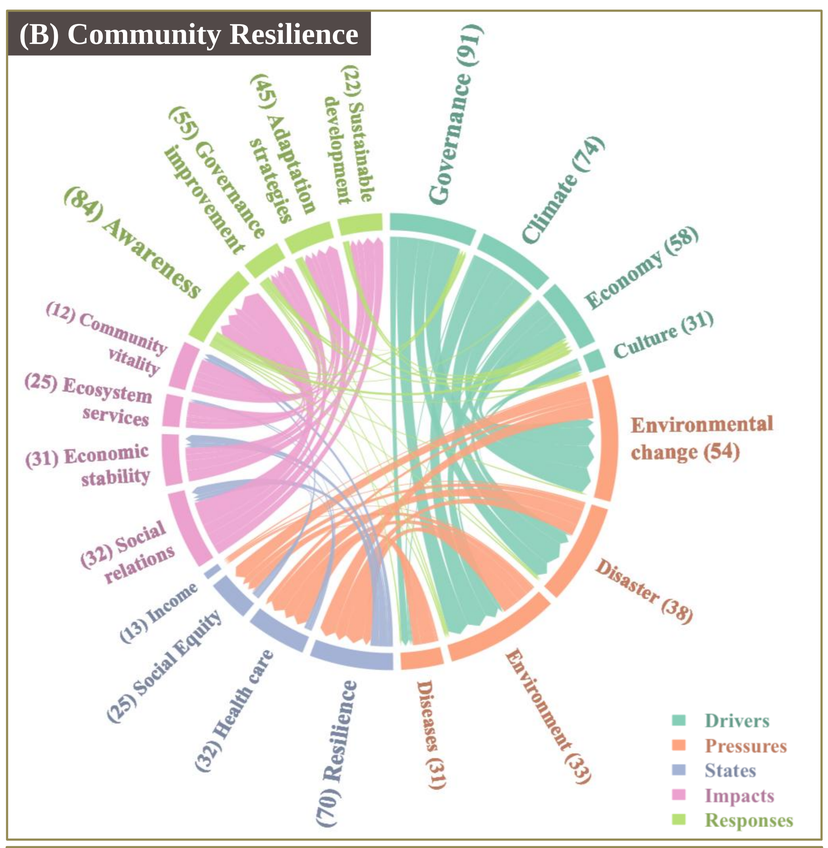
Visualization & Discussion
- **One DPSIR graph for each theme**





RESULTS & DISCUSSION





1. Methodological Features

Automated DPSIR analysis with LLMs, completed in just 24 minutes with visualization.

2. DPSIR Structures by Theme

(A) EEA:
Economy/Population → LULC/Environmental change → Biodiversity loss → Ecosystem decline → Adaptation/Awareness

(B) Community Resilience:
Climate/Economy → Hazards → Resilience → Social stability → Adaptation/Awareness

(C) Sustainable Tourism:
Tourism/Economy → LULC → Anthropogenic disturbance → Welfare loss → Management/Sustainability

(D) Waste Management:
Governance/Awareness → Improper activity → Landscape and ecological damage→ Welfare/Tourism loss → Management/Action

3. Key Observations

Positive Impacts: Beyond negative effects, Impacts also show beneficial influences on society and ecosystems.

Policy Implications of Responses: At the local level, DPSIR captures community responses.

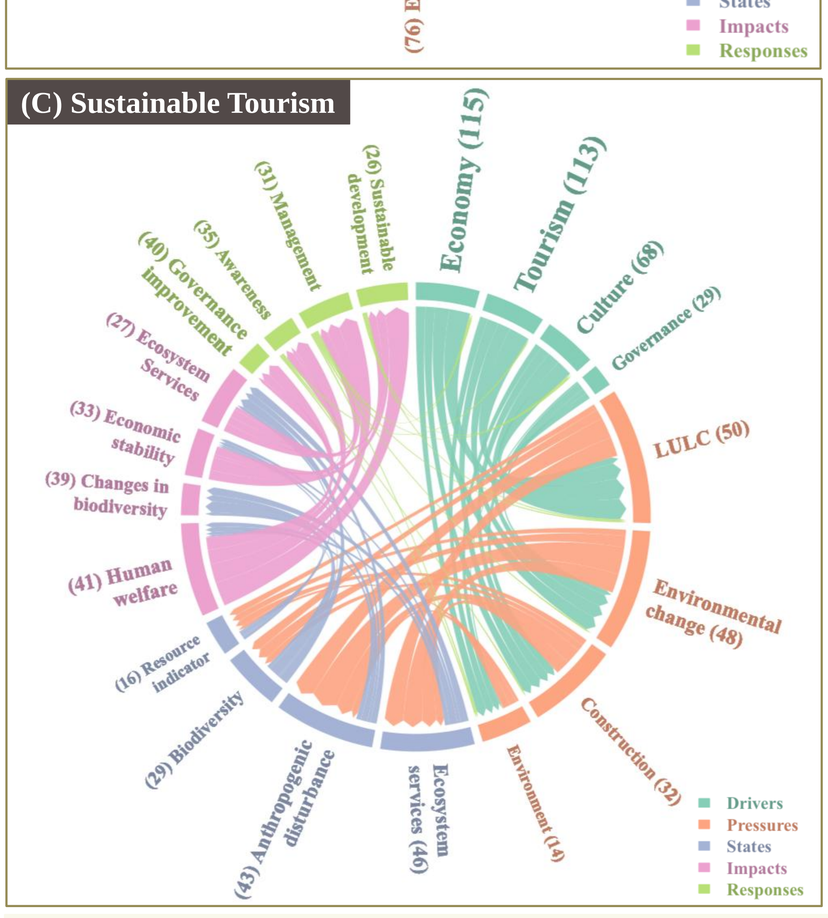
4. Theoretical and Policy Insights

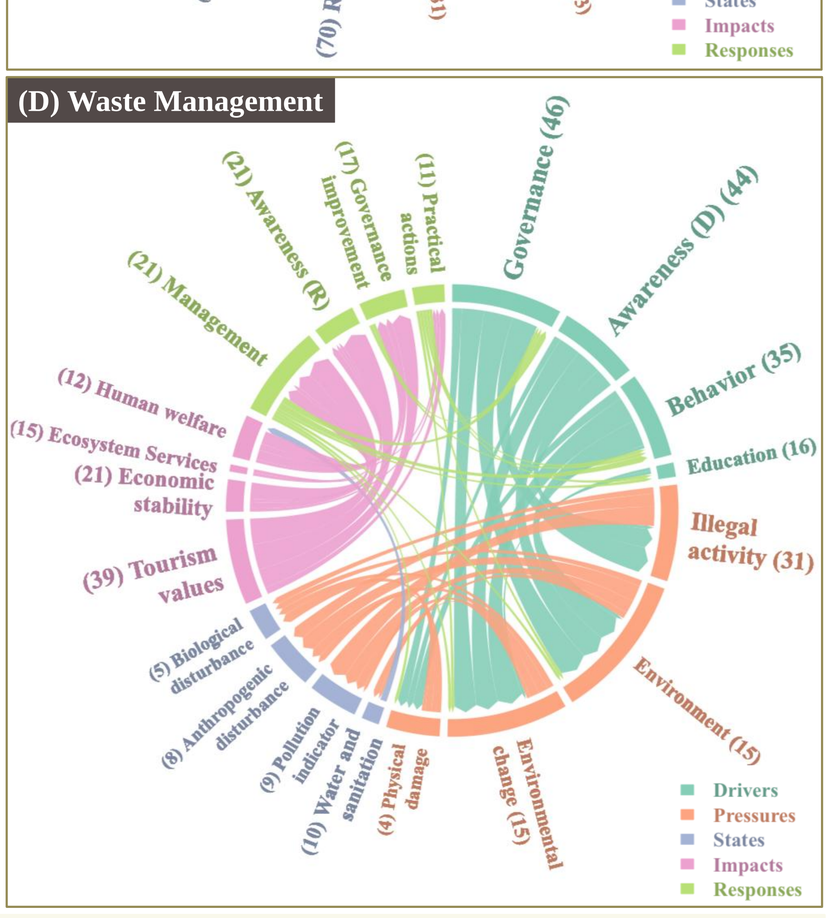
The results reveal bidirectional and non-linear causal interactions across DPSIR indicators.

The use of LLMs helps reduce analysis bias caused by researchers' subjective perspectives, enhancing consistency and objectivity.

Traditional DPSIR responses focus mainly on governance and legislation. This study, through large-scale qualitative analysis and stakeholder inclusion, broadens the scope of "Responses" to include:

- Changing rules and incentive structures to shape behavior
- Investing in state variables (e.g., ecological restoration)
- Shifting values and preferences





▲ **DPSIR Chord Diagrams:** Variables were annotated from snippets using EEA and thematic DPSIR definitions. The top four variables in each category are shown. Font size reflects variable frequency; chord number and thickness represent causal relationship frequency.

CONCLUSION

The results show that LLMs can increase the scale and efficiency of data processing, improve the consistency of DPSIR analysis, and reveal bidirectional and non-linear causal relationships. This approach supports a deeper understanding of complex social–ecological systems and provides context-specific evidence for local sustainability governance.