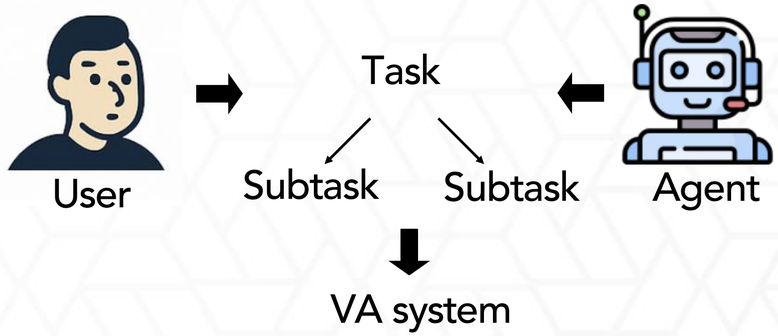
FDUVIS Lab

VIS2025

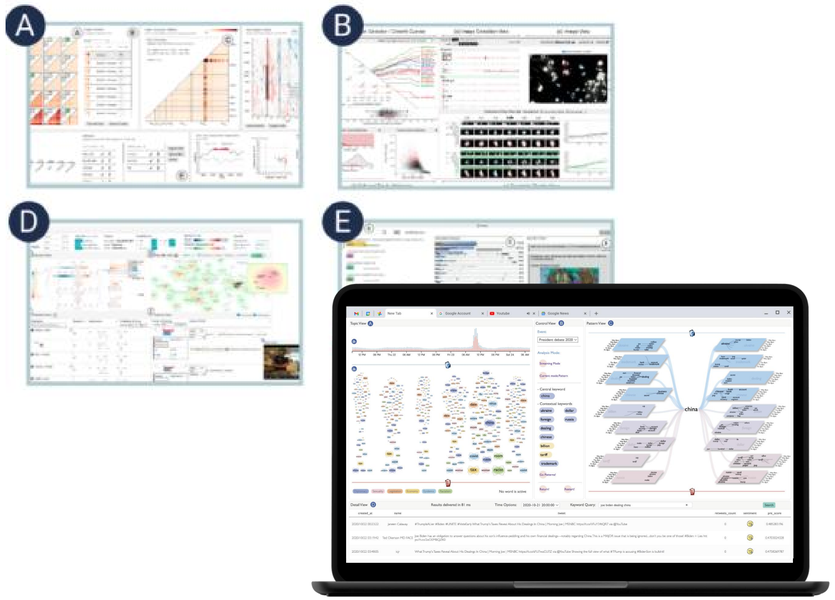


LightVA: Lightweight Visual Analytics With LLM Agent-Based Task Planning and Execution

Yuheng Zhao[1], Junjie Wang, Linbing Xiang, Xiaowen Zhang, Zifei Guo, Cagatay Turkay[3], Yu Zhang[2], Siming Chen[1]

[1] Fudan University [2] University of Oxford [3] University of Warwick

Motivation



Visual Analytics System

Large exploration space

Fixed Space

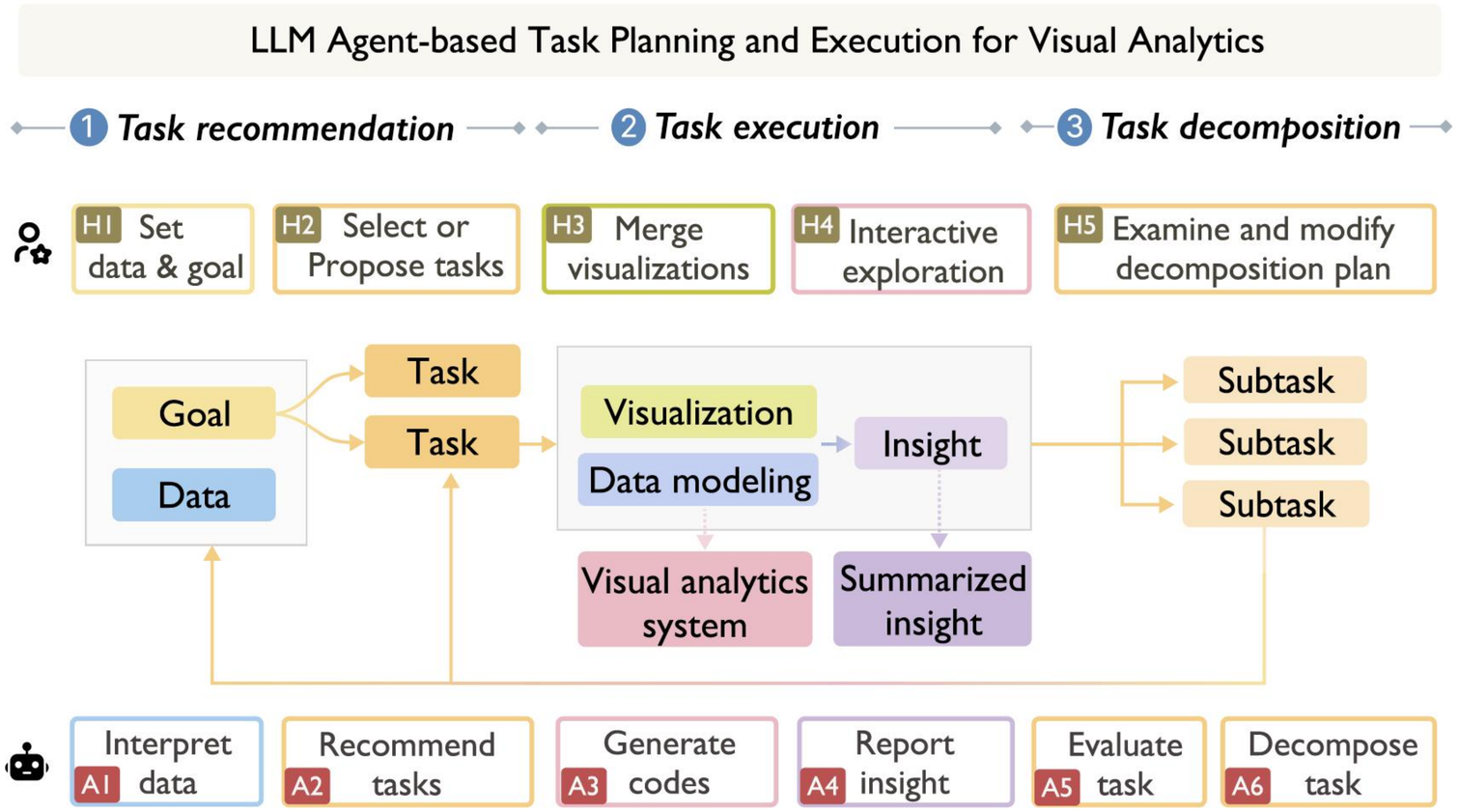
Task planning

Analyst

Have to plan and decompose tasks

Some new tasks/ideas cannot be solved promptly

Framework



Task decomposition: The goal–task–subtask progressive structure that helps users navigate a large analytical space.

Malleable interface: generate visualization & insights based on user needs.

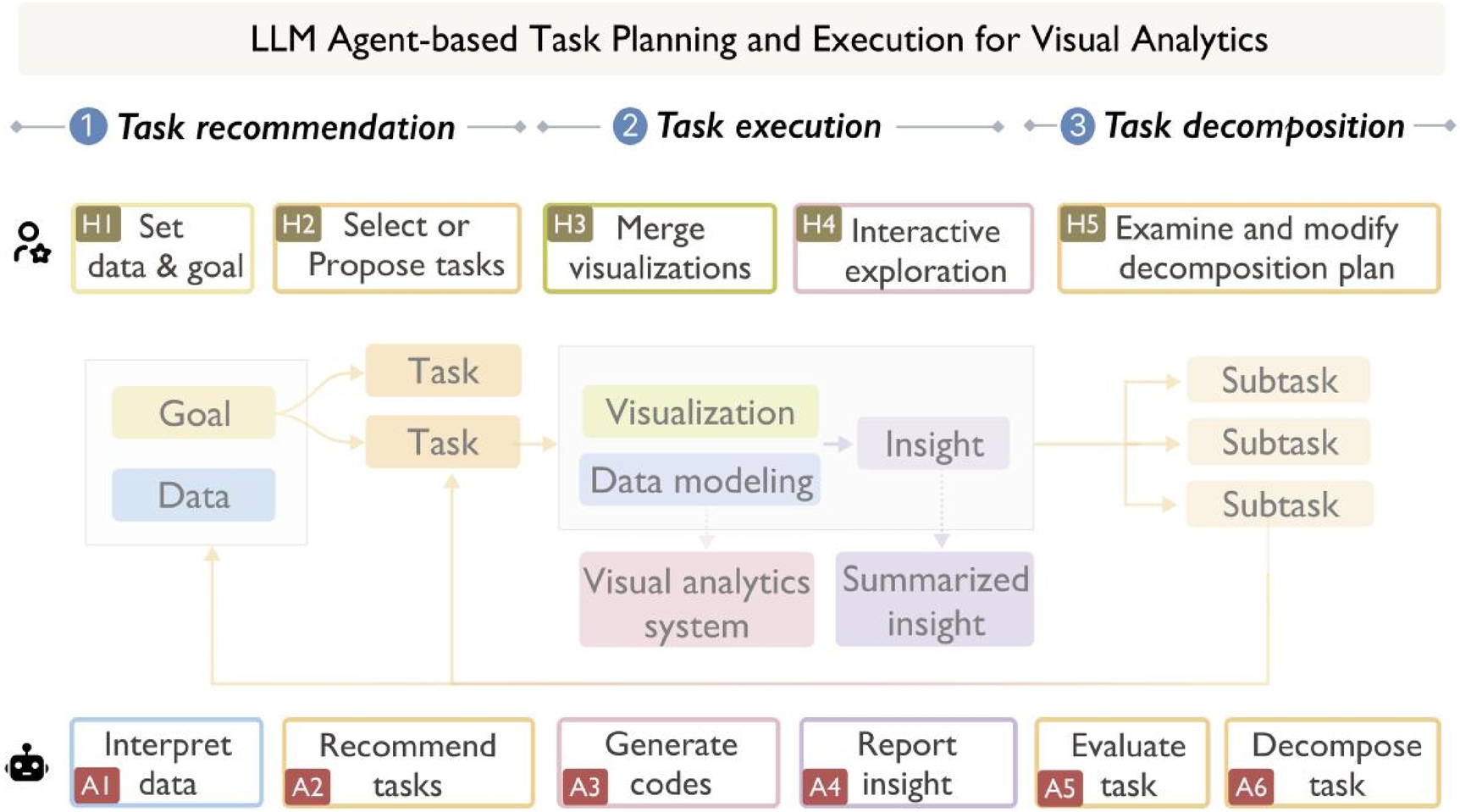
Framework

Human roles:

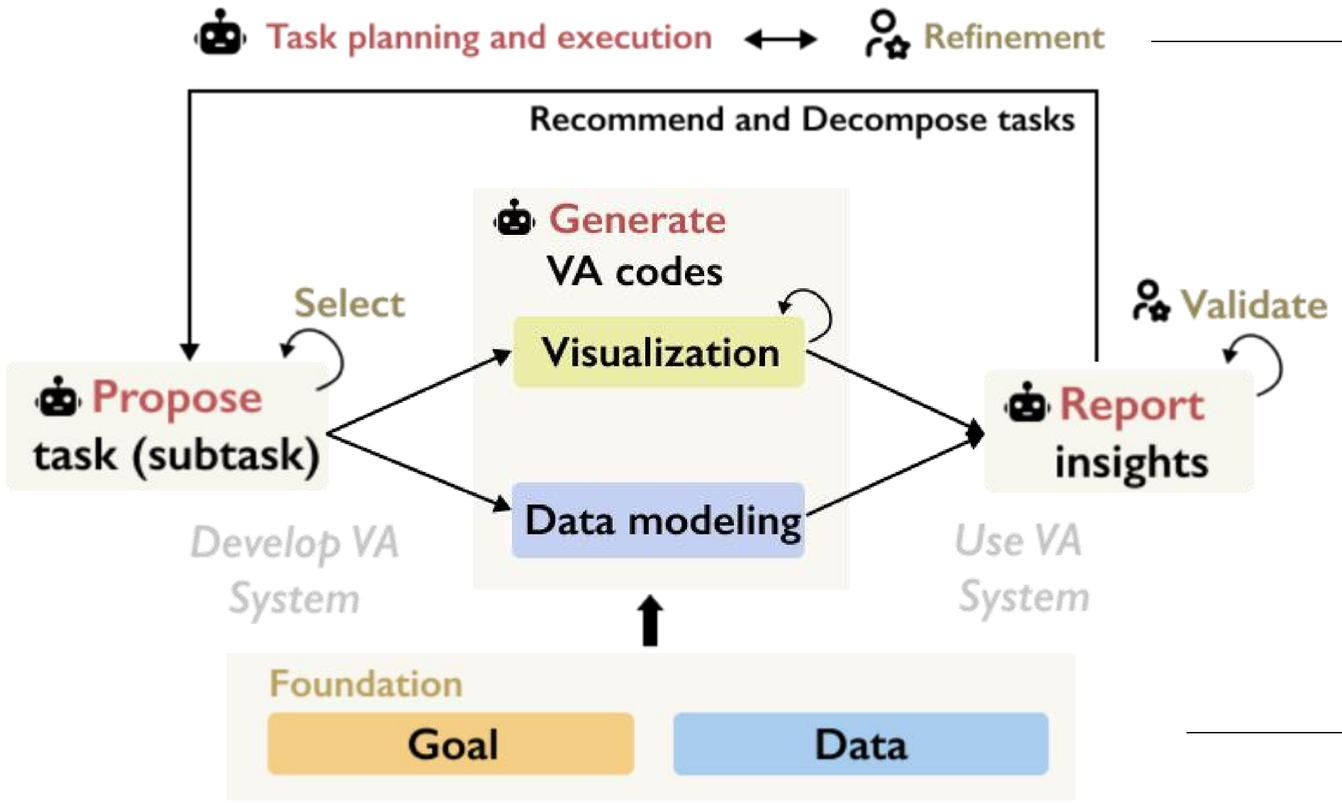
- **Goal setting**
- **Visual perception**
- **Validation**

Agent roles

- Task planning
- Data modeling
- Explanation



Framework



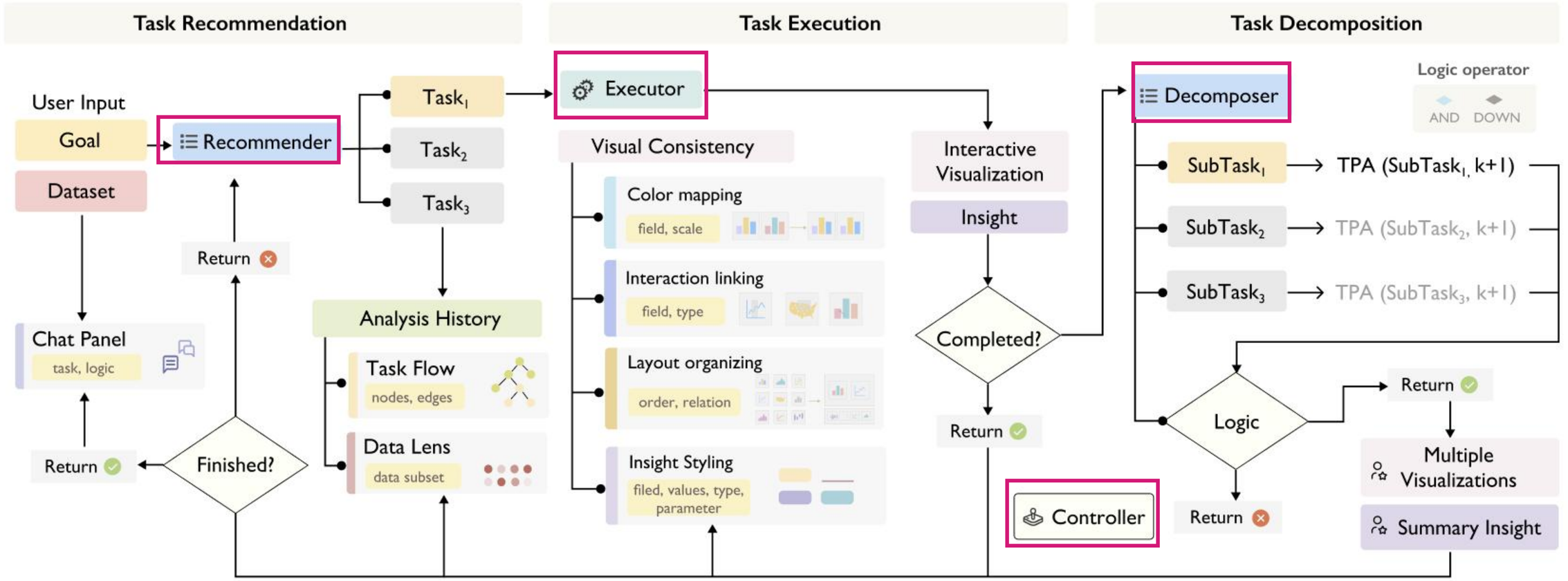
Collaboration mode:
Autonomous planning +
human refinement

Method: Task dynamic planning

Output: Visualization, insight

Input: Goal, Dataset

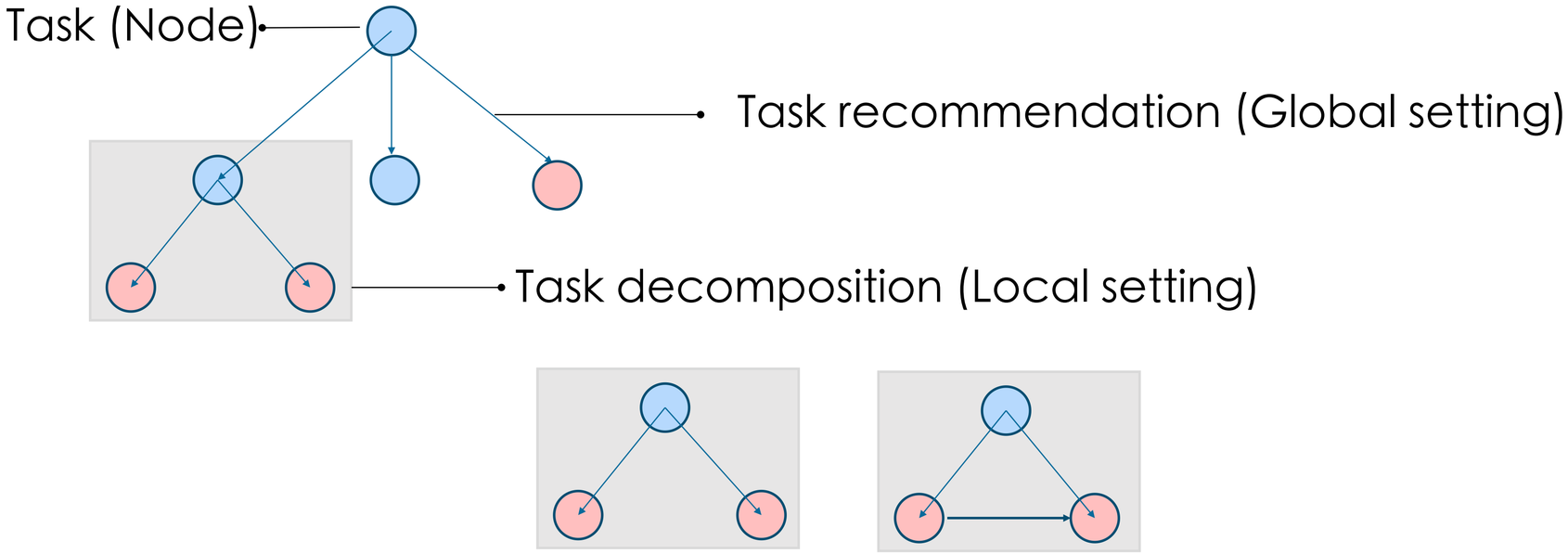
Workflow: Agent-Based Task Planning



Modeling

The definition of the task planning process

> We define the task flow as a directed graph, where nodes represent tasks, edges represent the relationships between tasks, and tasks are executed with the logic.

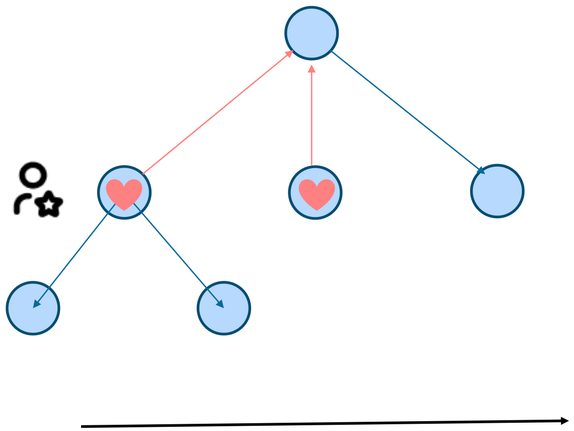


Indicates a newly proposed task

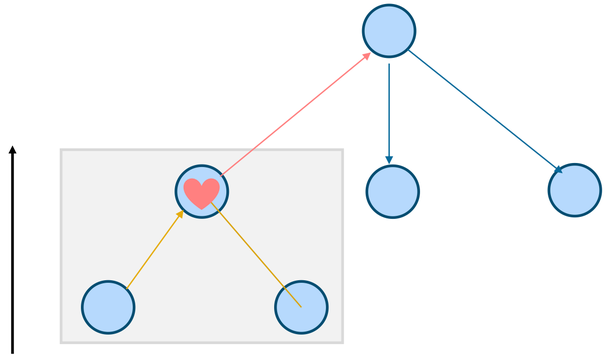
Modeling

The definition of goal completion

> We define *task progress* for visual analysis of the context, meaning:
> (1) The user first needs to confirm whether the task should be executed; if confirmed
> (2) The completion of this task depends on whether its subtasks are completed.



(1) When all tasks that the user expects to be completed are done, the goal is achieved.



(2) Subtasks are executed according to the decomposition plan; when all subtasks are completed, the main task is considered complete.

The circle with a heart indicates a task the user expects to execute.

Modeling

Task recommendation

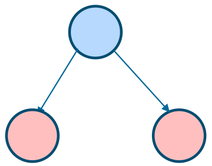
Goal: “Tell me about fuel efficiency”

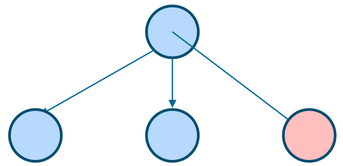
Task 1: Identify the key factors affecting fuel efficiency, such as engine size, weight, and aerodynamics.

Task 1: Identify the key factors affecting fuel efficiency, such as engine size, weight, and aerodynamics.

Initial stage

- Initially determine the *data variables* related to the goal
- Recommend different *task types* according to the goal and data





Based on historical analysis records

- Propose *new tasks* according to the current context and previous tasks
- Remind users of *unexplored tasks* that need to be revisited

The circles with a red outline represent newly proposed tasks.

Modeling

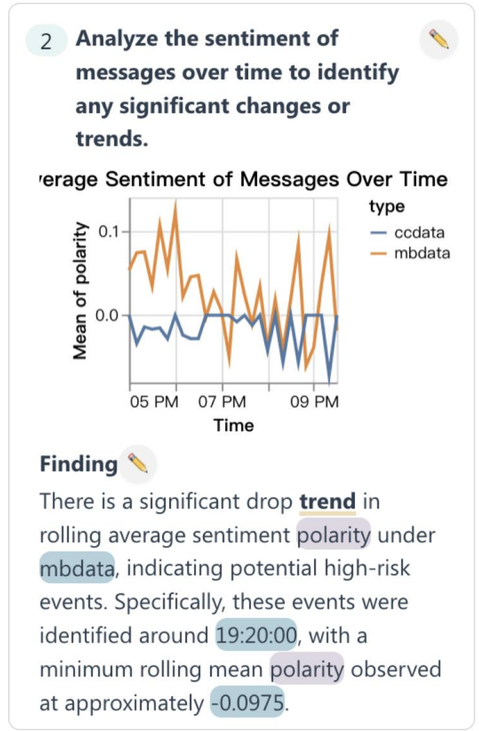
Task execution

Step1. Code generation

- Data transformation/modeling
- Vega-lite generation

```python
import altair as alt
import pandas as pd
def plot(data: pd.DataFrame):
    # Data preprocessing
        <codes>
    # Chart generation
    chart = alt.Chart().mark_bar().encode()
    return chart
```

```
{
"text": "There is a correlation between A and B, indicating xx generally exhibit lower B for xx in C.",
"type": "'Correlation",
"parameters": "negative",
"data_variables": ['A', 'B'],
"data_values": {
'A': '2000 to 5000lbs',
'B': '10 to 35MPG',
'C': 'xx'
}
}
```



Step2. Insight generation

- Compete Insight specification

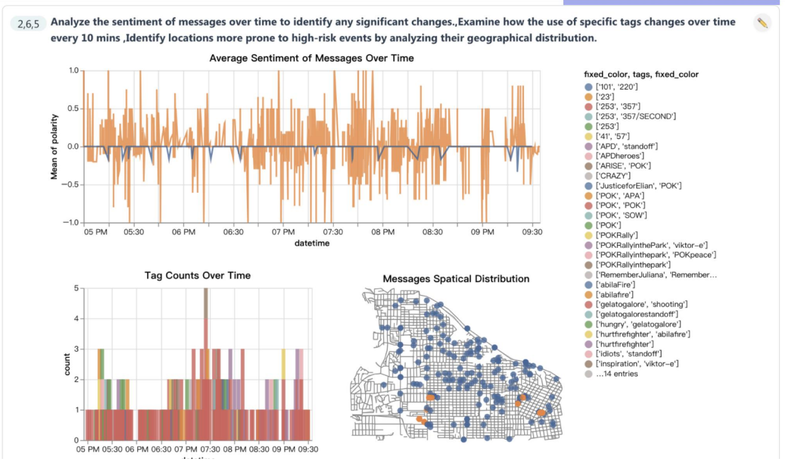
Step3. Multi-view generation

- Cognitive consistency constraint
  - Colors, layout [1]
- Interaction coordination constraints
  - Interaction Type (brush | filter | click)

Cognitive consistency

Interactive coordination

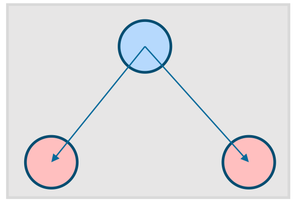
Multi-linked Views



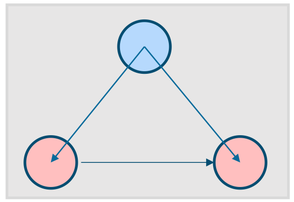
[1] Qu, Zening, and Jessica Hullman. "Keeping multiple views consistent: Constraints, validations, and exceptions in visualization authoring." *IEEE transactions on visualization and computer graphics* 24, no. 1 (2017): 468-477.

Modeling

Task Decomposition



and
(parallel)



down
(sequential)

Goal: “Tell me about fuel efficiency”

Task 1: Identify the key factors affecting fuel efficiency, such as engine size, weight, and aerodynamics.

Subtask1: Analyze the relationship between vehicle weight and mpg.

Subtask2: Analyze the correlation coefficient between vehicle weight and mpg.

Subtask3: Perform a linear regression analysis to quantify the relationship.

1. Task Completion Check

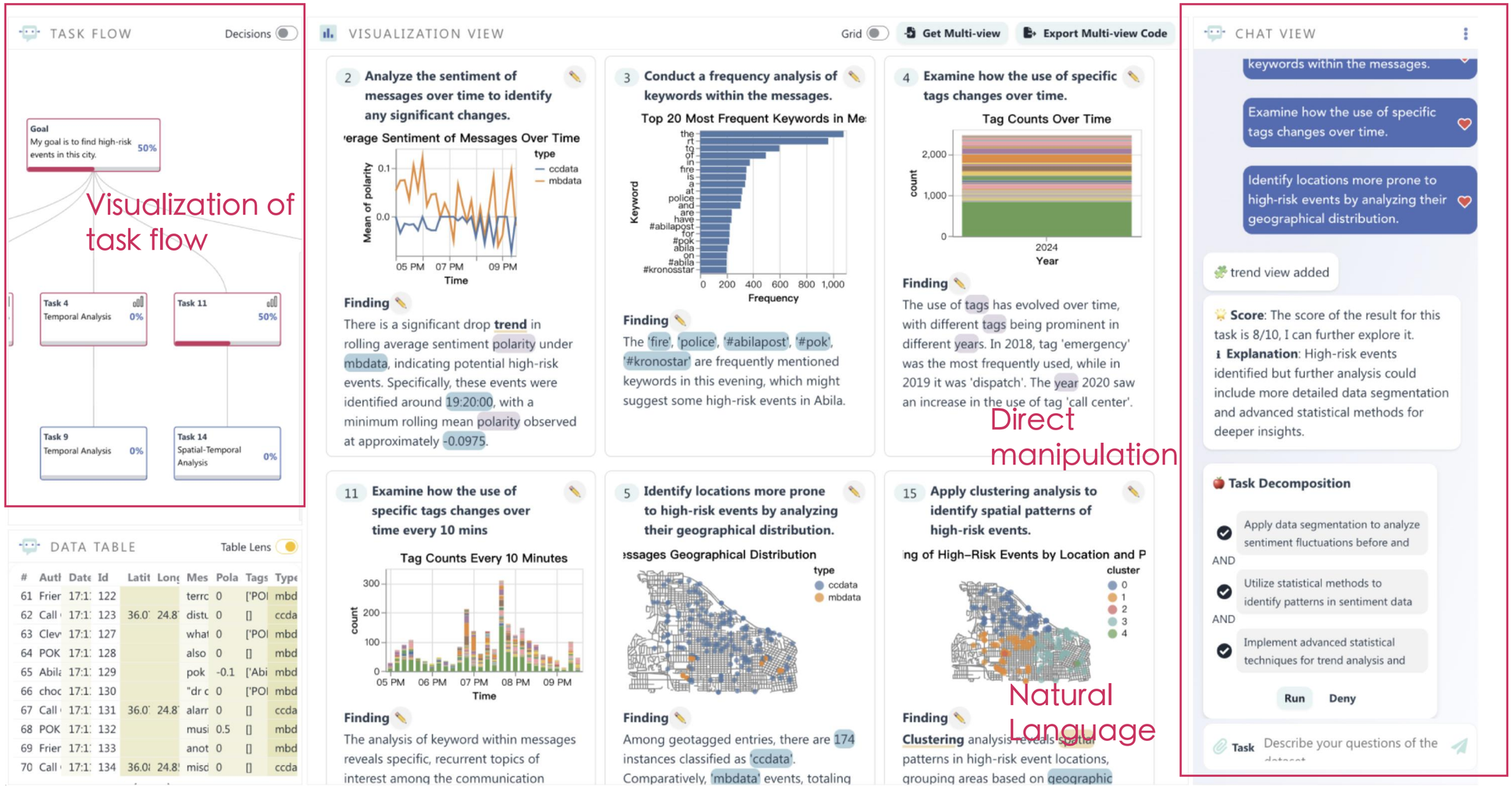
- Data coverage: Covers all relevant subsets
- Insight complexity: Need advanced analysis (regression, clustering, etc.)
- Visualization: Try better ways to present insights

2. If incomplete → Decomposition Plan

- And: Multiple agents run in parallel
- Down: Single agent runs sequentially, using previous results as context

Logic:= “(Subtask1 AND Subtask2) DOWN Subtask3”

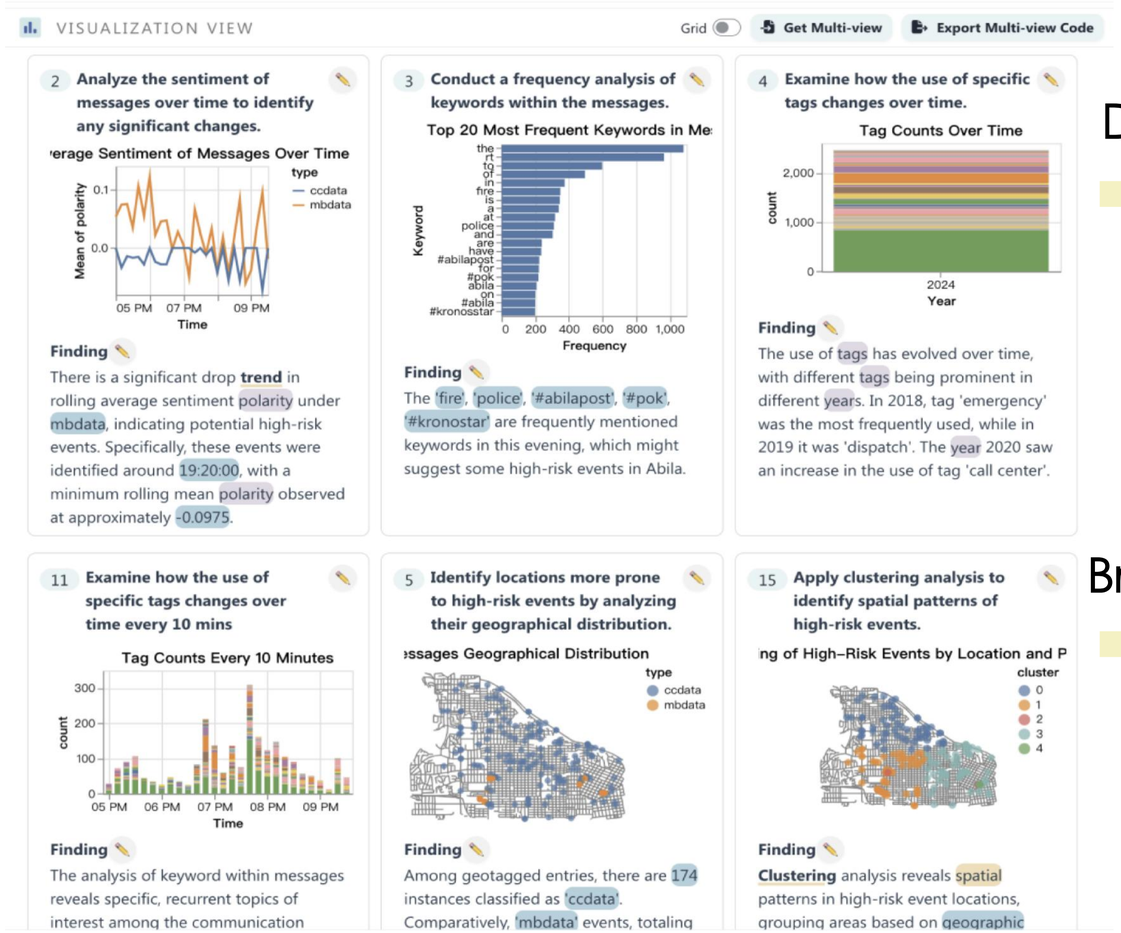
Interface



Users steer the agent’s reasoning via visualization, direct manipulation, and natural language.

Interface

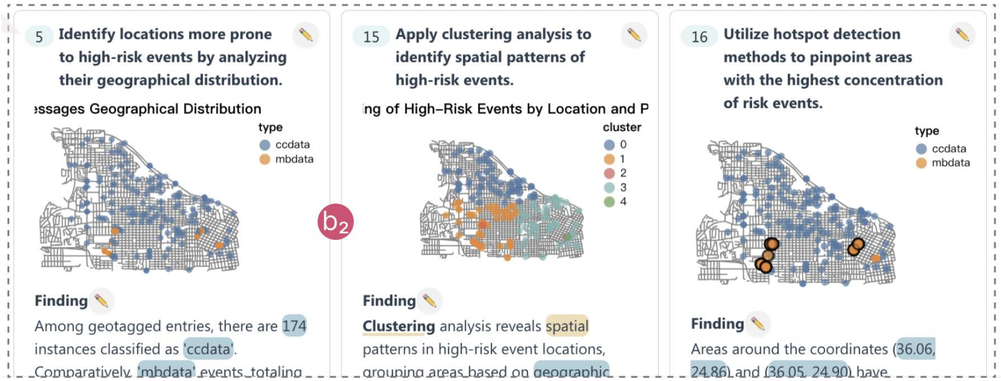
Insight gallery



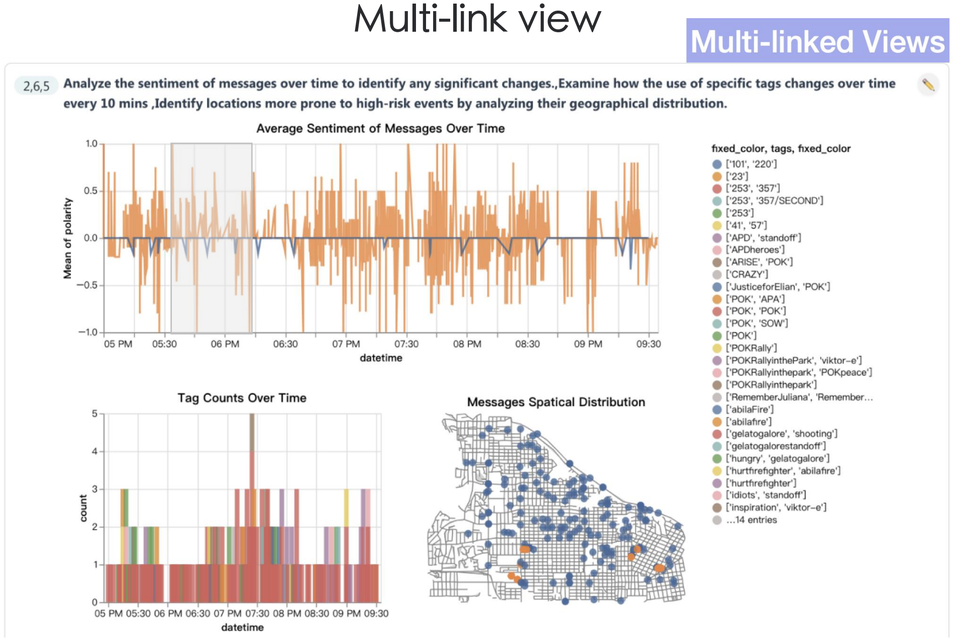
Depth

Breadth

Spatial analysis -> Cluster analysis + hot spot detection analysis



Multi-link view



A Case Study Demo

Usage Scenario · VAST Challenge 2021 · Mini-Challenge 3

- **Background**: On January 23, 2014, multiple events unfolded in Abila.

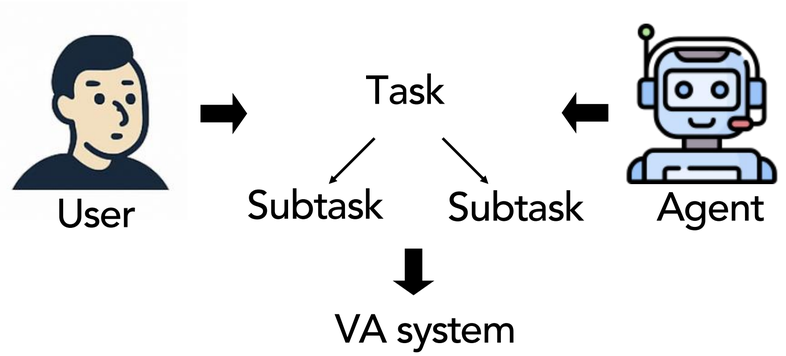
- **Data**: Social media messages with sentiment scores and locations

To evaluate our framework

Discussion

Findings from the expert studies

- LightVA shortens the process of system development and analysis that originally took about 30 days to just 2-3 hours, expanding the depth and breadth of exploration.
- Compared with end-to-end agent-based data analysis, progressive task co-planning help enhance user engagement.



Open questions

Towards mixed-initiative collaboration, not only tool, but also partners:

- Visualization (perception & interaction) can serve as medium a that benefits customization, validation, and explanation of planning process.
- Building the task graph from different domain VA as knowledge base, more adaptive task planning
- Rethinking evaluation — beyond task accuracy toward collaborative performance (insight discovery ratio)

FDUVIS Lab

Thank you!

yuhengzhao@fudan.edu.cn
simingchen@fudan.edu.cn

More details in our paper

Project Website: https://zyh1222.github.io/LightVA/

VIS2025