UCI

# Thinking in DAGs

## How Humans and Agents Collaborate in Data Science

texera

**Speaker: Chen Li**
University of California, Irvine

# A Data Science Task

**movies**

| id | title | year | rating |
|---|---|---|---|
| 1 | Inception | 2010 | 8.8 |
| 2 | Avatar | 2009 | 7.8 |
| 3 | Dark Knight | 2008 | 9 |
| 4 | Titanic | 1997 | 7.2 |
| 5 | Pulp Fiction | 1994 | 8.9 |
| 6 | Parasite | 2019 | 8.6 |

.....

**budgets**

| id | title | budget_millions |
|---|---|---|
| 1 | Inception | 160 |
| 2 | Avatar | 237 |
| 3 | Dark Knight | 185 |
| 4 | Titanic | 200 |
| 5 | Pulp Fiction | 8.5 |
| 6 | Parasite | 11 |

.....

**Question:** Which highly-rated movies have below-average budgets?

# AI Agents using Scripts (Python) as the Abstraction

In [1]: **Script 1**

```
movies  = pd.read_csv('movies.csv')
budgets = pd.read_csv('budgets.csv')
print(len(movies), len(budgets))
```

In [2]: **Script 2**

```
merged = movies.merge(budgets, on='id')
avg = merged['budget_millions'].mean()
print(avg)
```

In [3]: **Script 3**

```
hit = merged[(merged['rating'] >= 8.0) &
             (merged['budget_millions'] < avg)]
print(hit[['title','rating','budget_millions']])
```

**User task**

Load both files and preview them.

▸ `generate` **`script1`**

Join on id and compute the average budget.

▸ `generate` **`script2`**

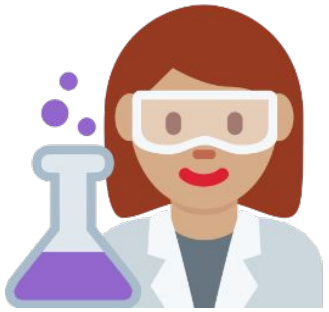
Keep movies rated ≥ 8.0 with below-average budget.

▸ `generate` **`script3`**

Send message to ScriptAgent…

# Scripts are Hard to Understand by End Users

Scripts (SQL, Python, etc.)

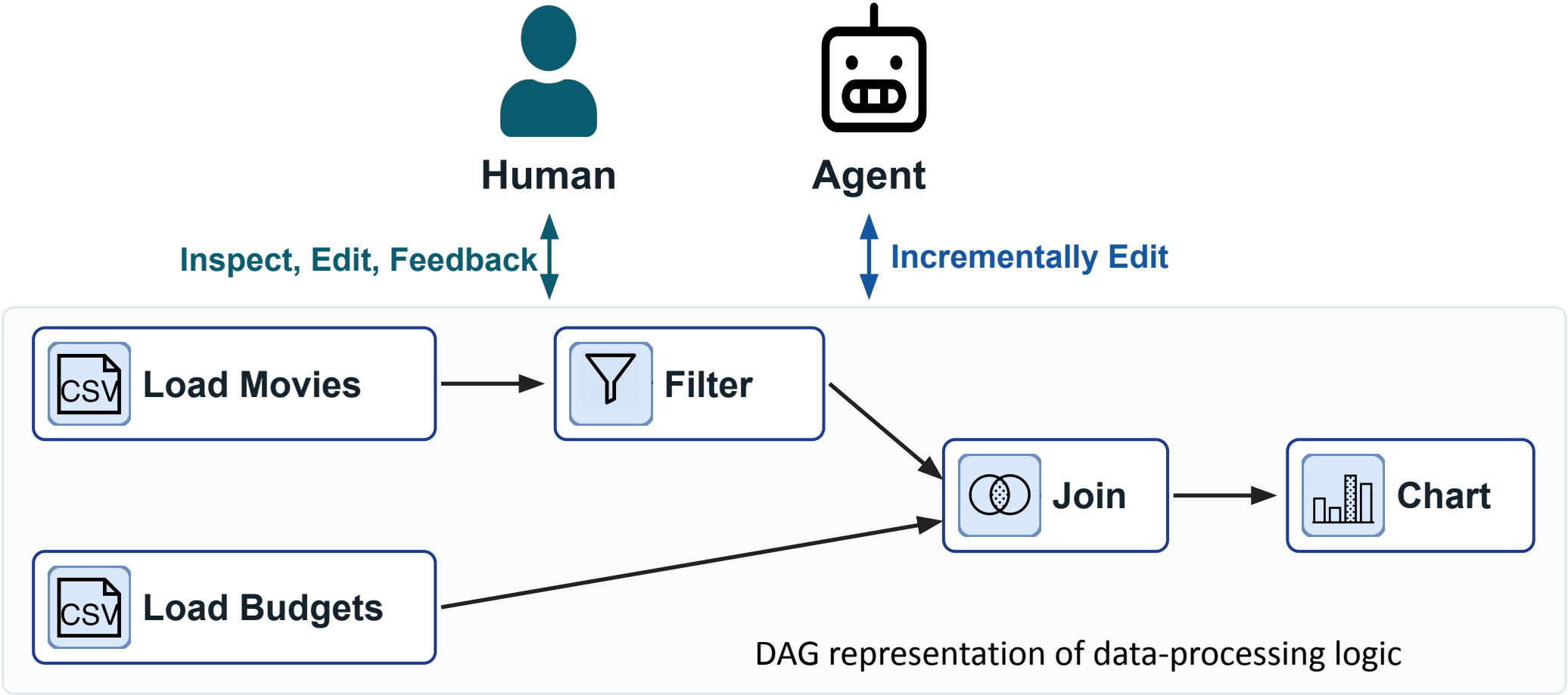


Domain Scientists

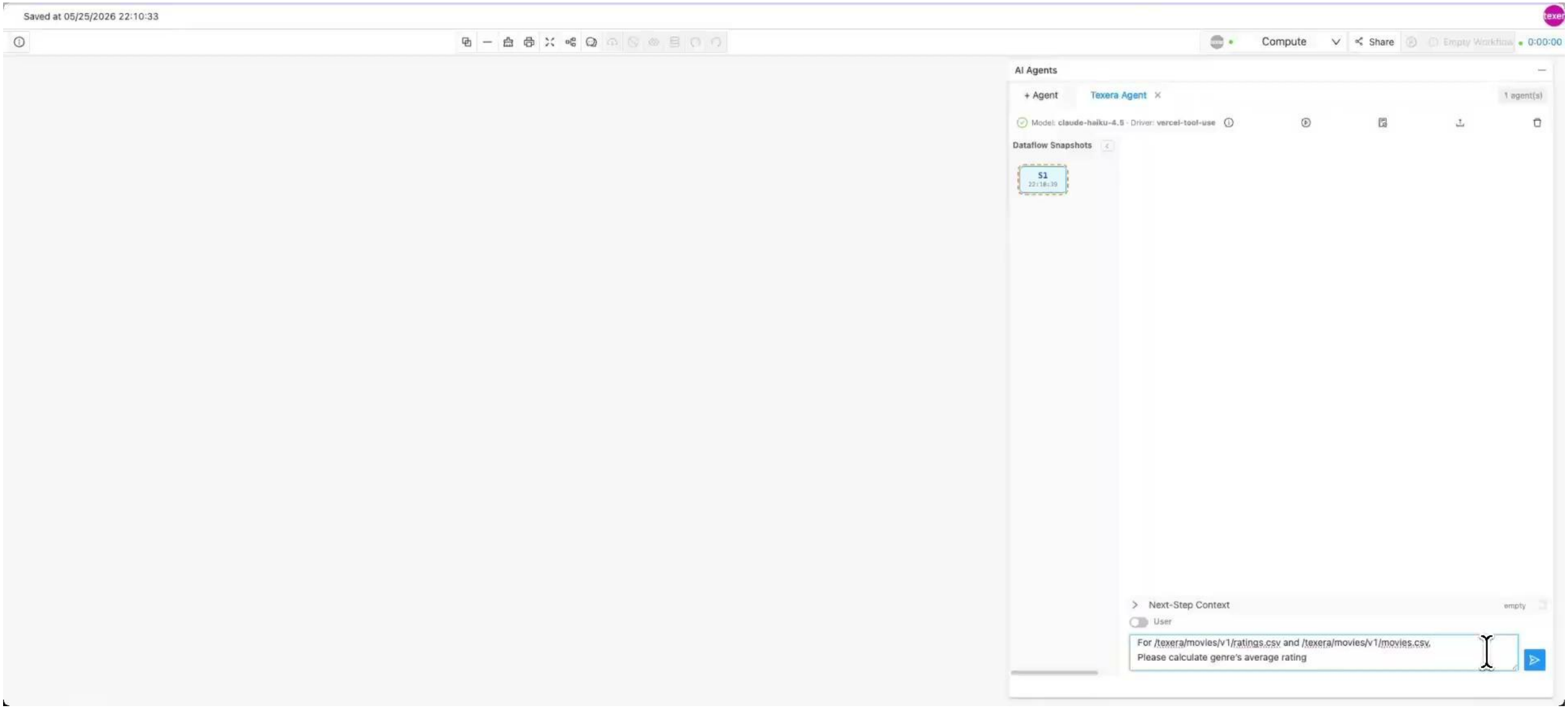
K-12 Students

Programmers

# Our Belief: a DAG is a better Abstraction for Human-Agent Collaboration



# Demo of BobFlow: Human-Agent Collaborations using Dataflows



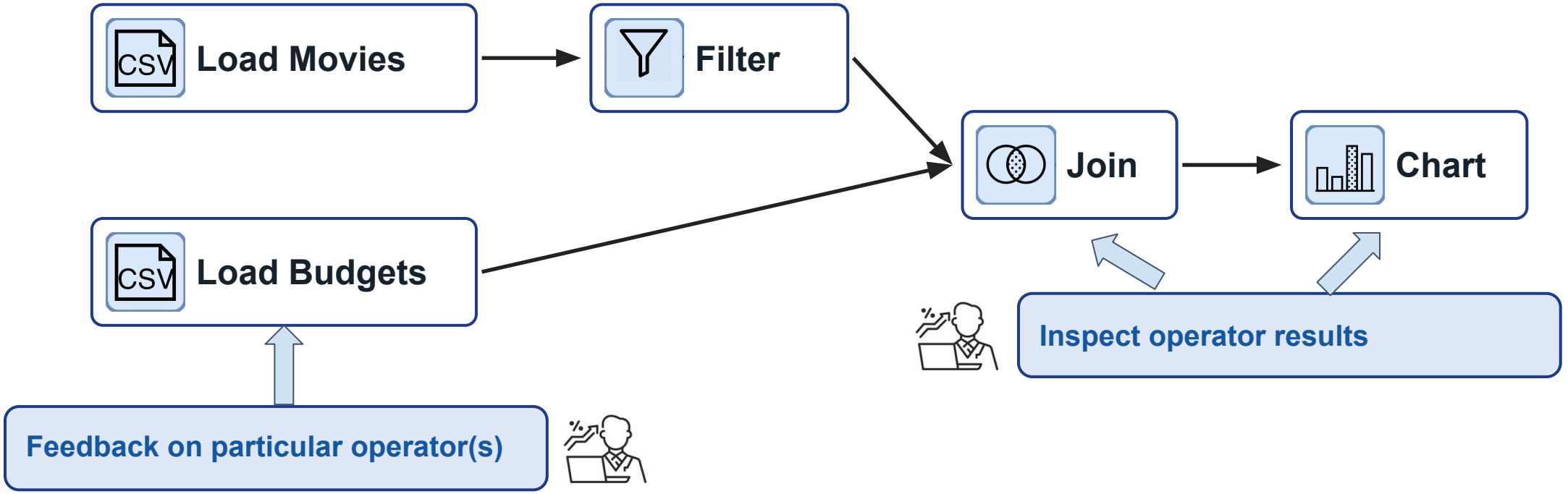
# Three key questions to answer

1. Benefits to **humans**?

2. Benefits to **agents**?

3. Requirements for the **runtime engine**?

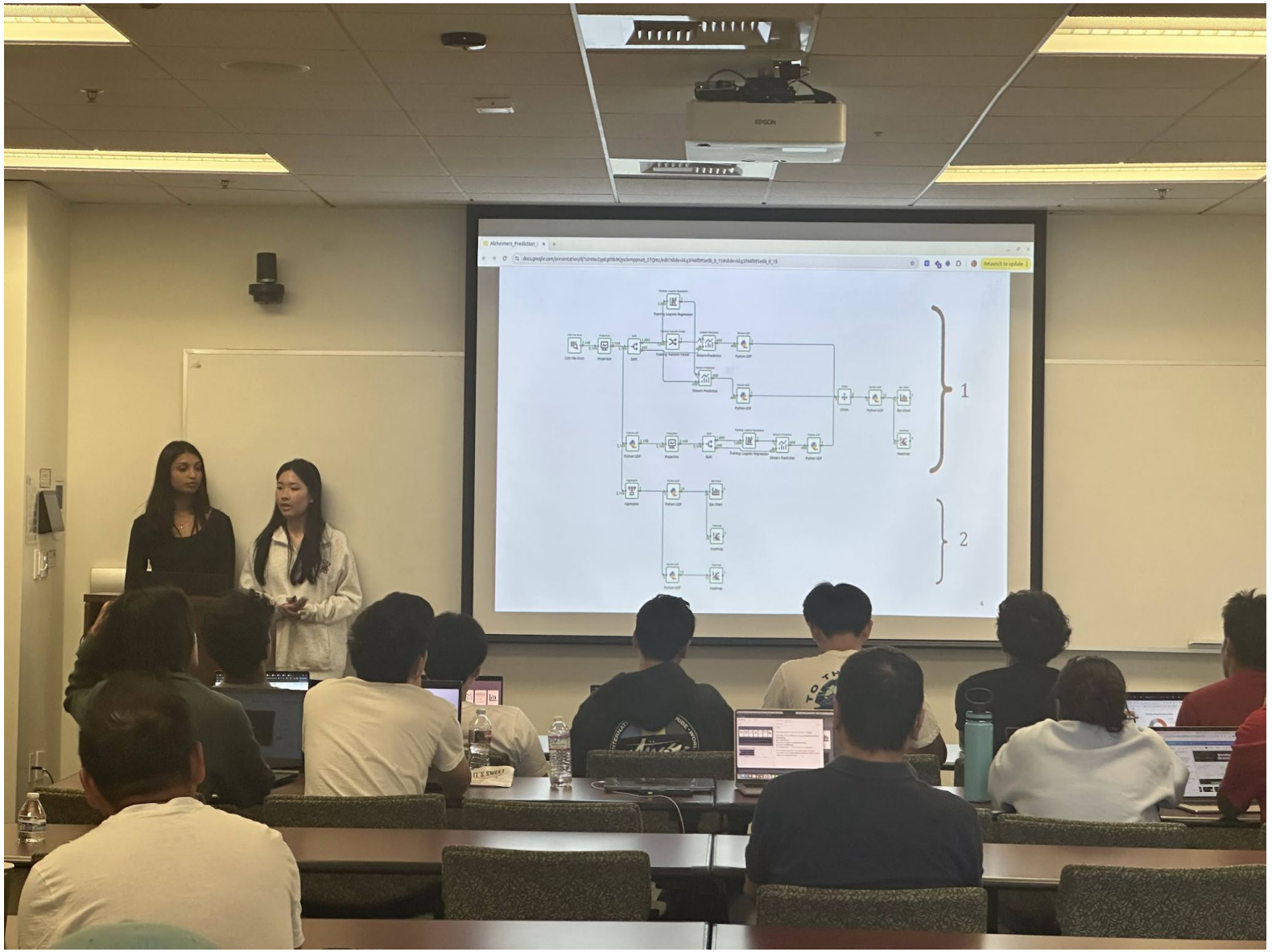
# Benefits to Humans




**Intuitive**

**Accurate**

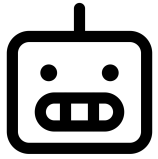
# Teaching High-Schoolers using DAGs



# Three key questions to answer

1. Benefits to **humans** ? ✓
2. Benefits to **agents** ?
3. Requirements for the **runtime engine**?

# Benefits to **Agents**?

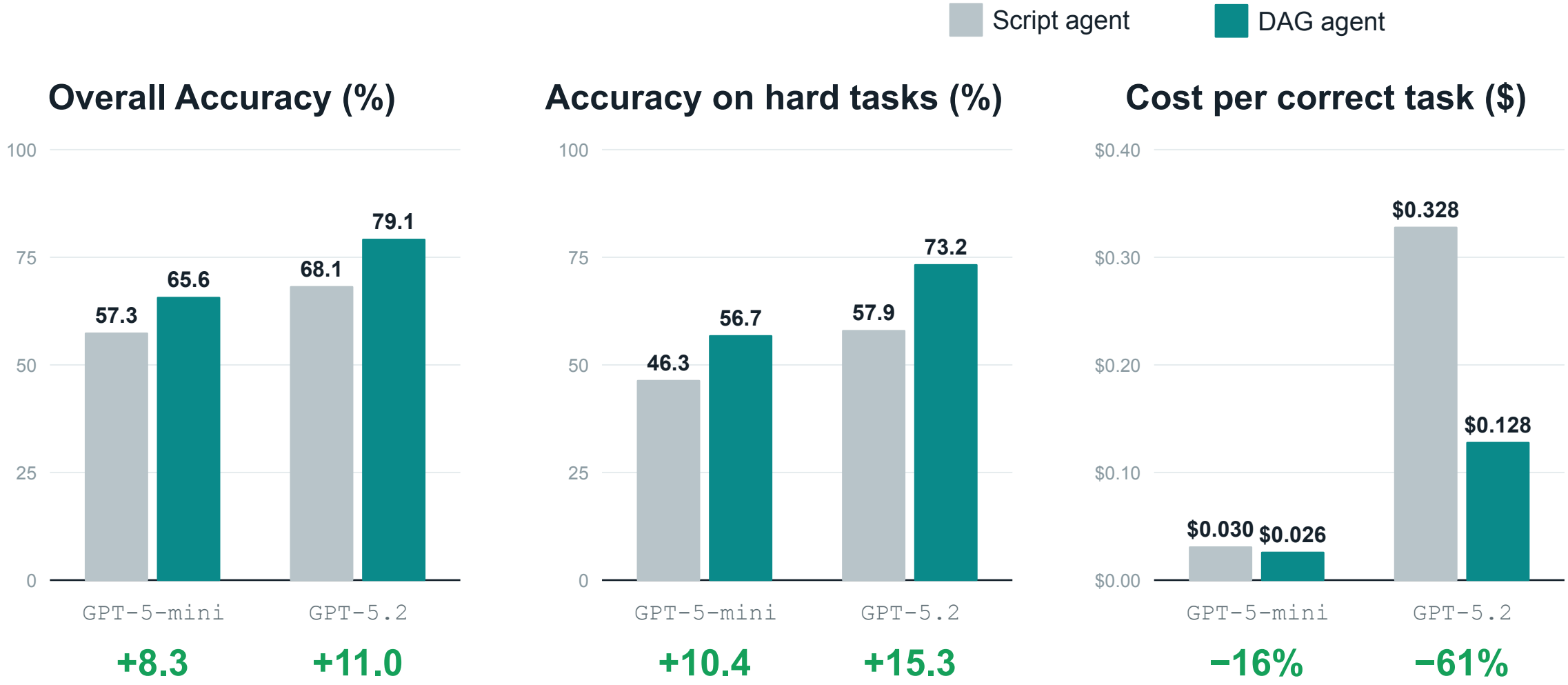


Increase the accuracy?

Decrease the cost?

# Preliminary Results

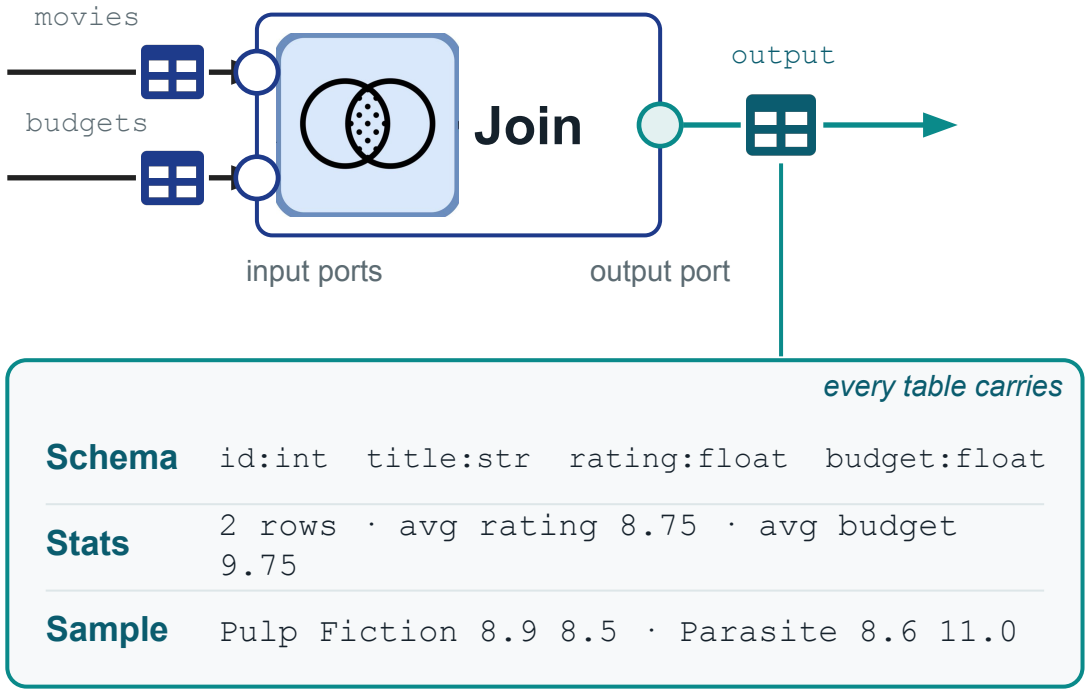
Benchmarked over 104 tasks in KramaBench.

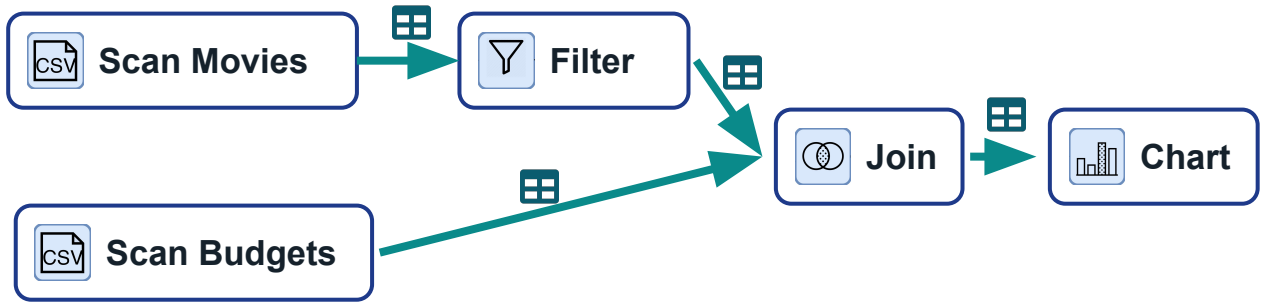



The Higher The Better

The Lower The Better

# Observation 1: DAG-based Contexts Improve Accuracy

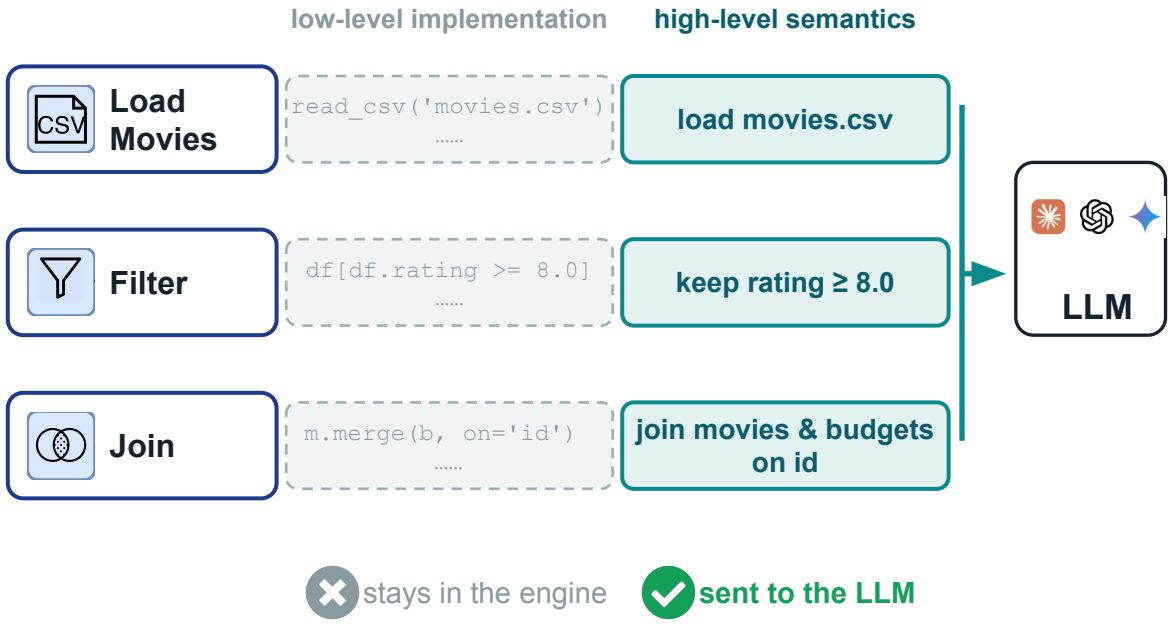


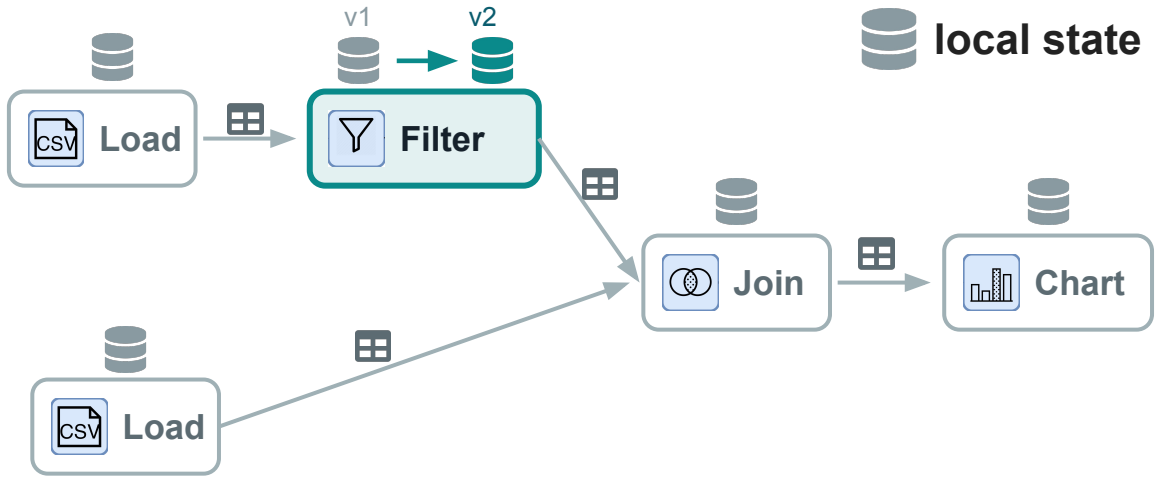




1 Data explicitly defined

2 Lineage explicitly defined

# Observation 2: DAG-based Contexts Decrease Cost



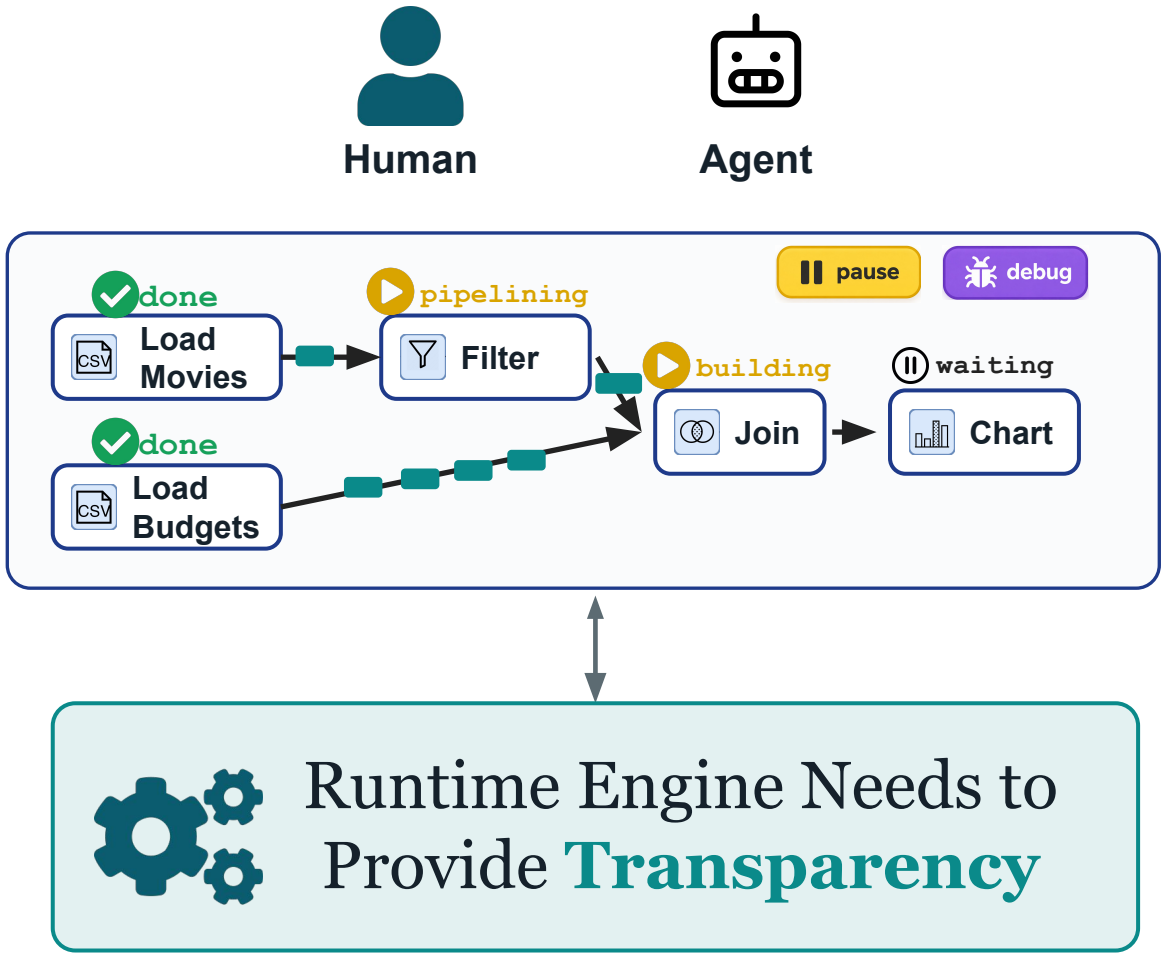

1 Dense input information via **high-level operator semantics**

2 Minimal output tokens via **shared-nothing execution paradigm**

# Three key questions to answer

1. Benefits to **humans** ? ✓
2. Benefits to **agents** ? ✓
3. Requirements for the **runtime engine**?

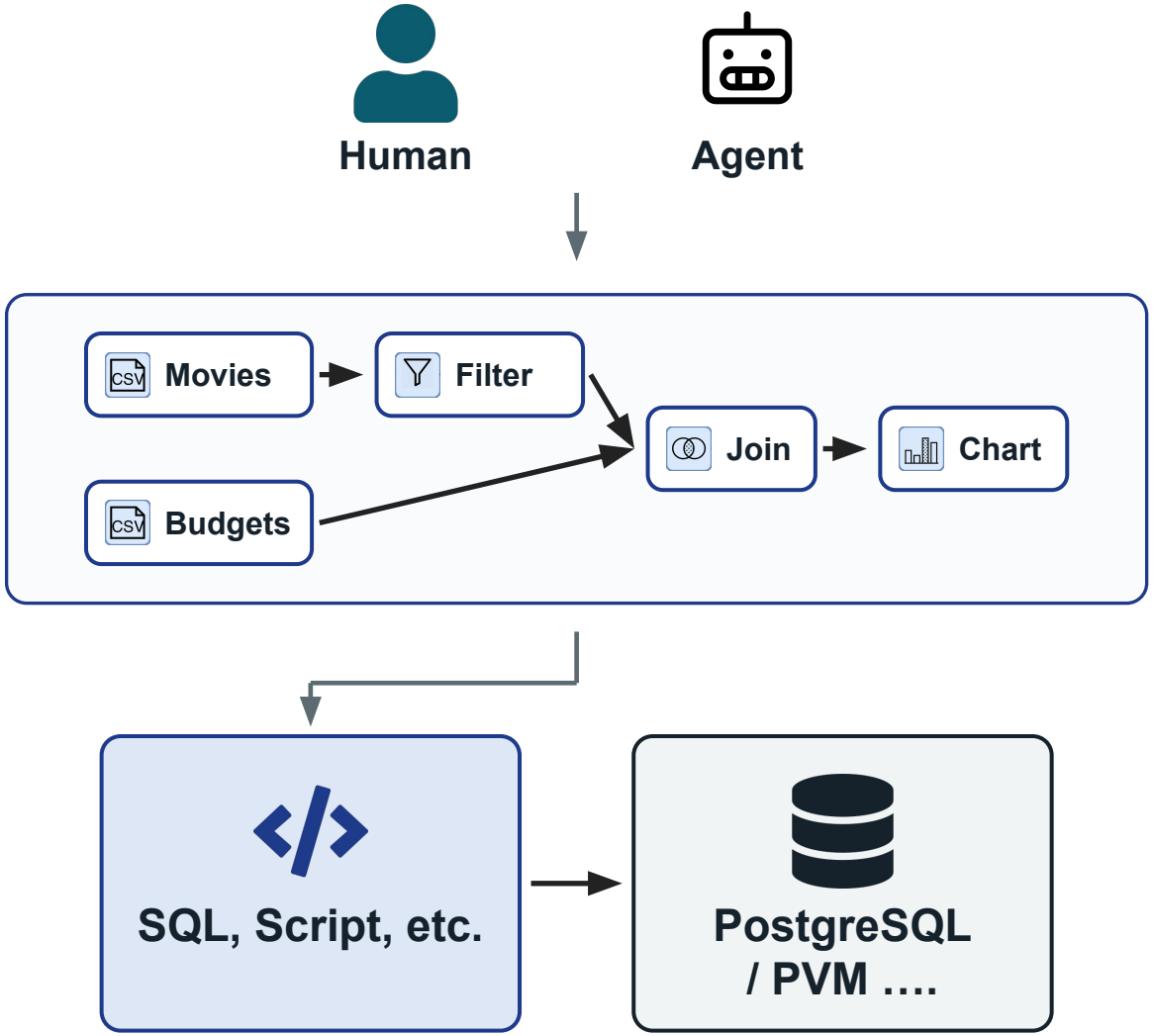
# Humans Need to Collaborate With Agents



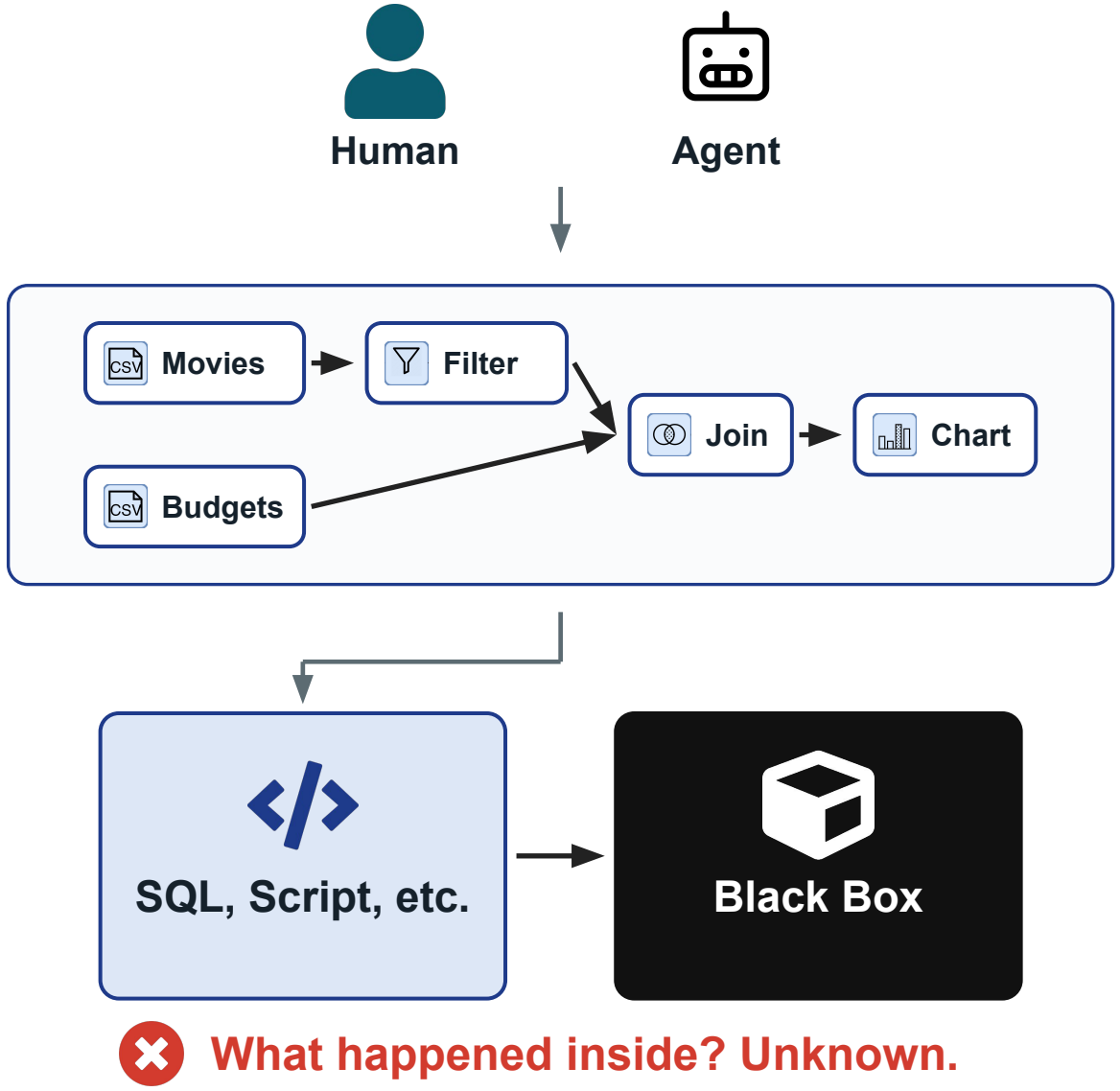

To collaborate, humans need:

- Runtime State
- Intermediate Results
- Execution Control

# Option: translate DAG as SQL, Script(Python), etc.

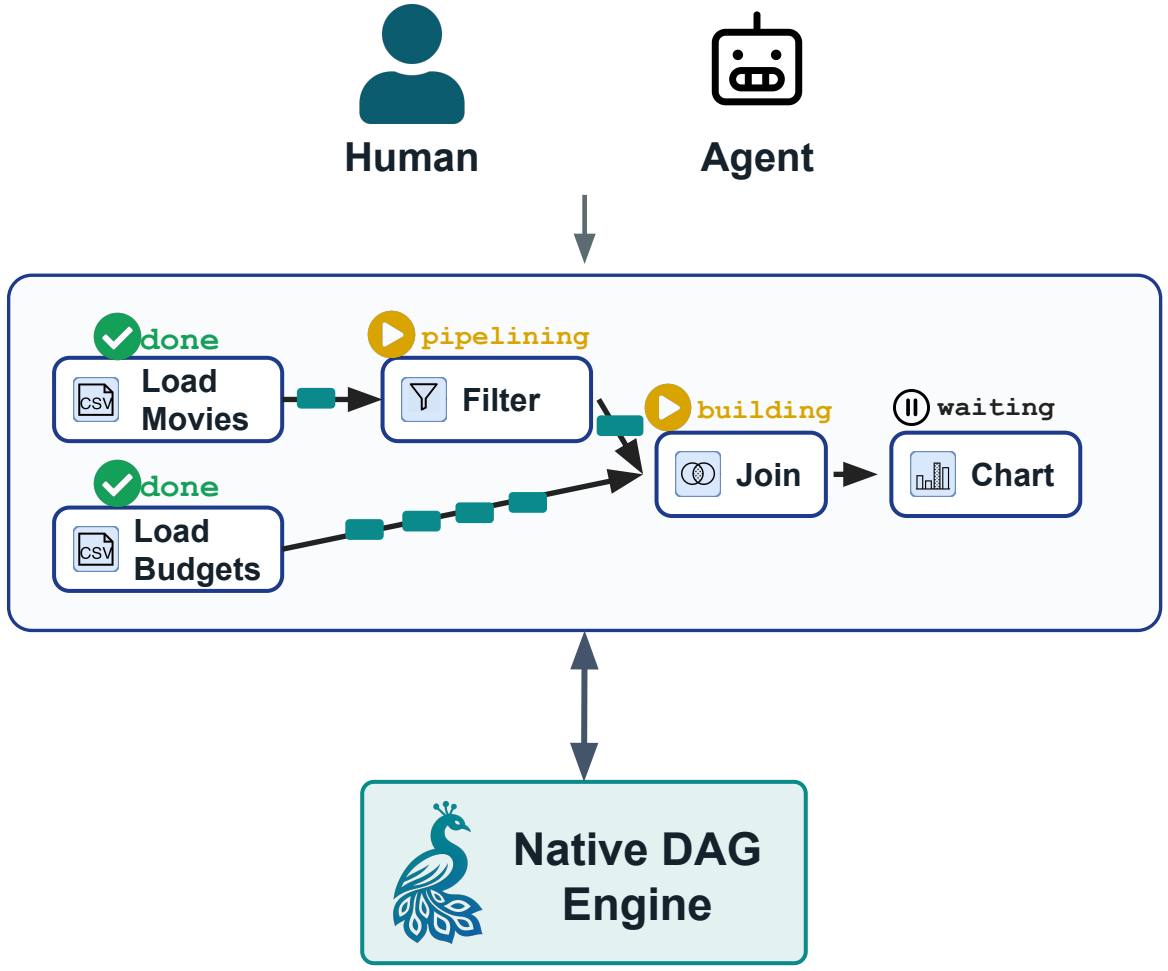


# Option: translate DAG as SQL, Script(Python), etc.




What happened inside? Unknown.

# A Native DAG Engine: what you see is what you get!



# Four requirements for a native DAG engine

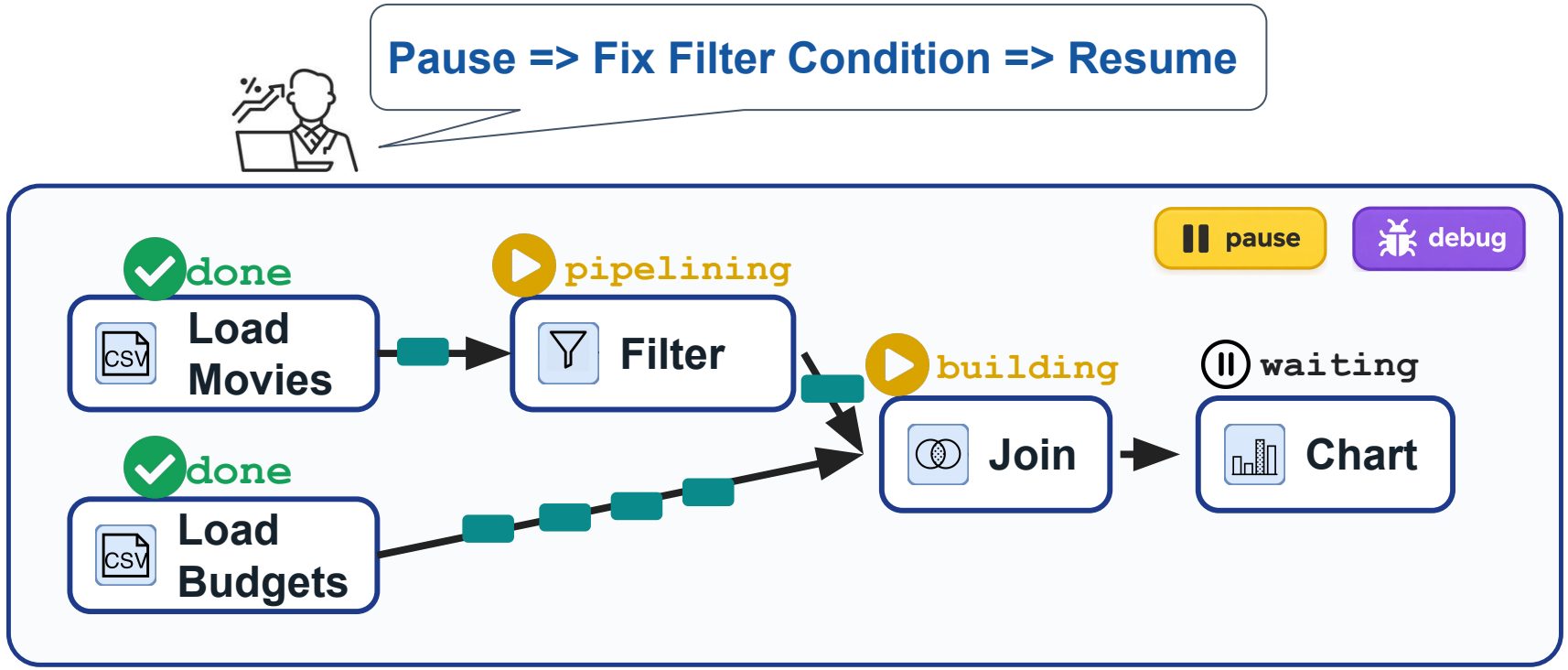
1. **User interactions** during execution

2. **Expressiveness** of operators

3. **Incremental** execution of agent-built DAGs

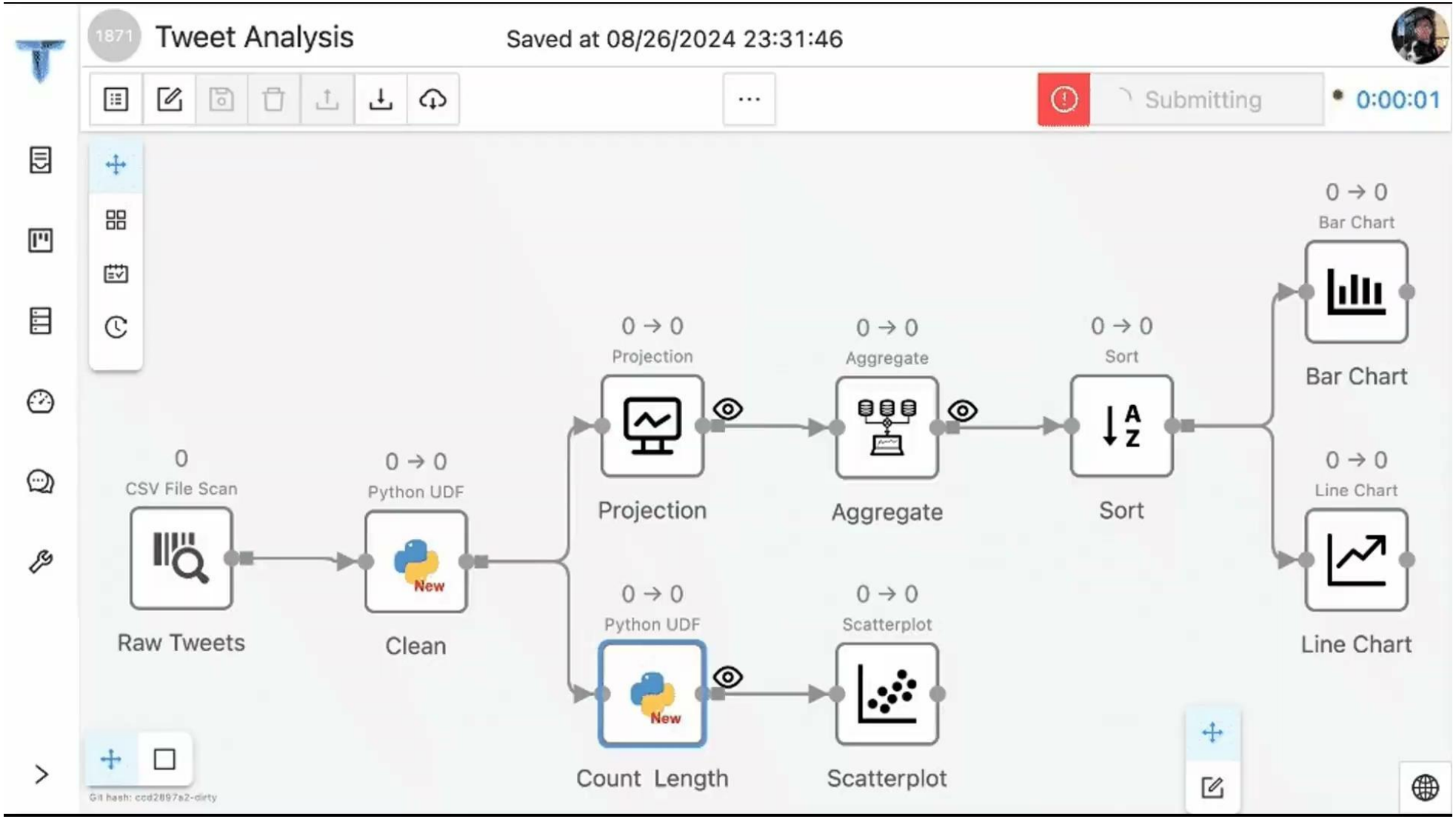
4. **Time travel** on conversation history

# Requirement 1: User Interactions during Execution



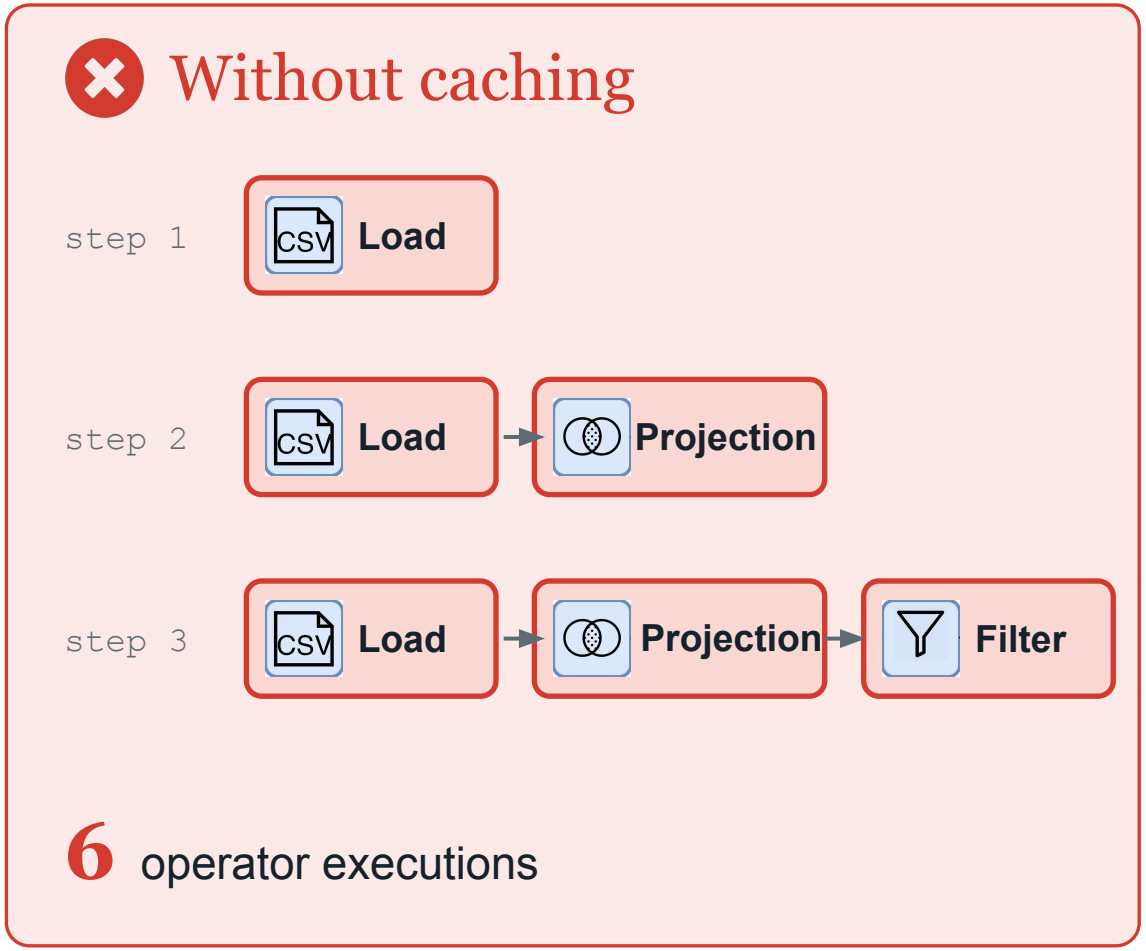
# Requirement 1: User Interactions during Execution

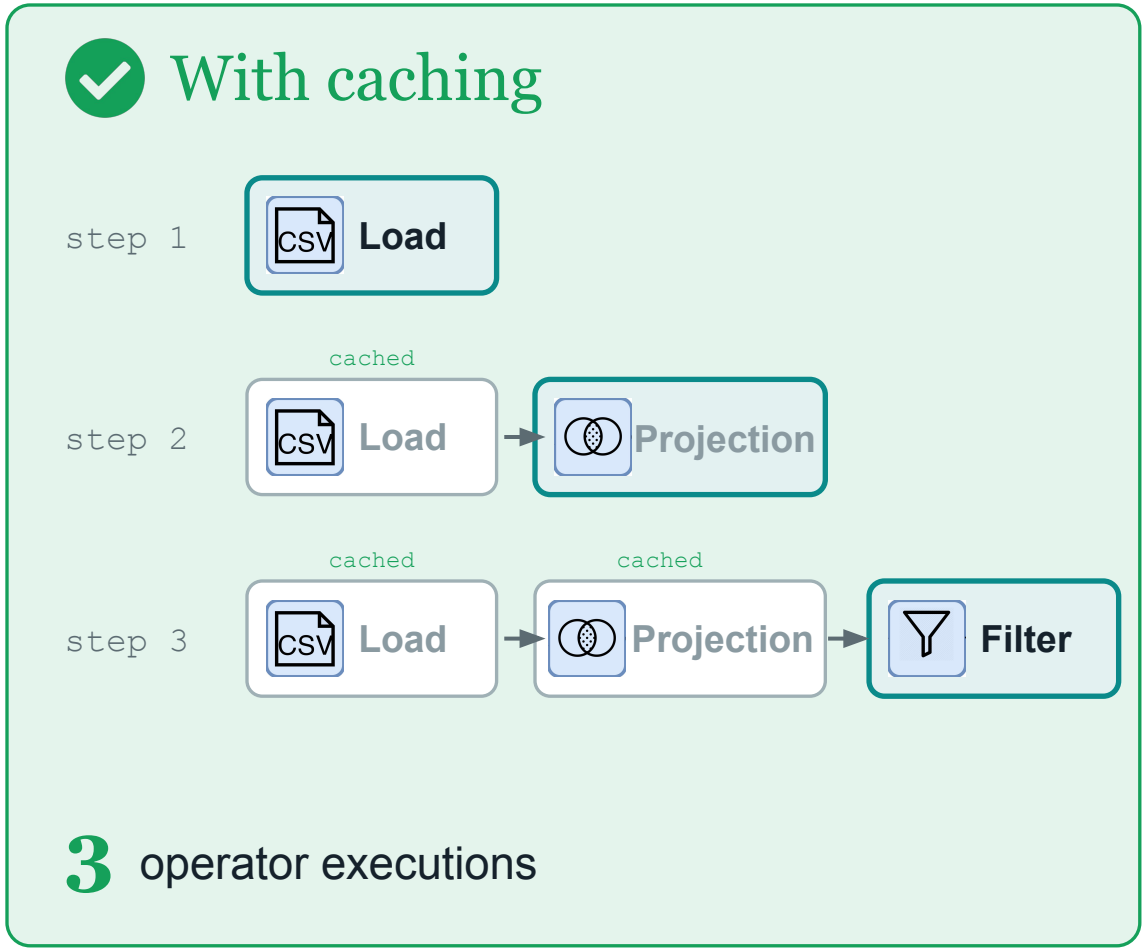
Humans should be able to pause, inspect, edit, and give feedback **while the DAG runs**.



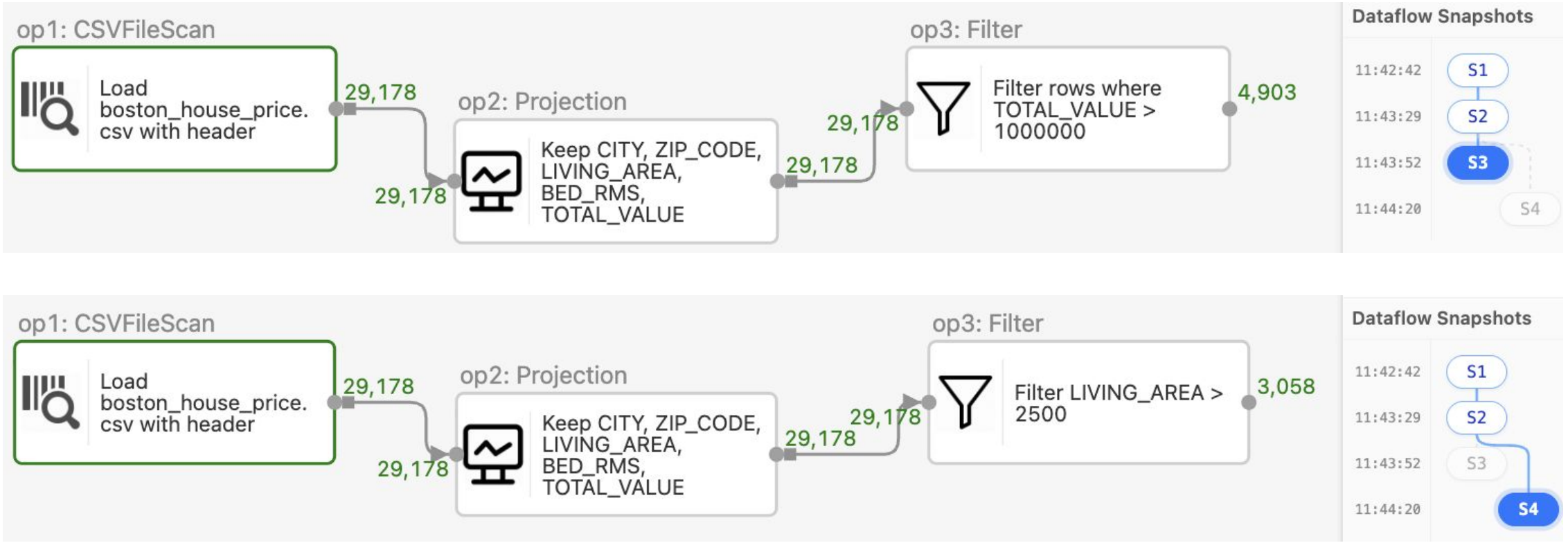
# Requirement 2: Caching for Incremental Execution

Each step adds one operator. Without caching, everything upstream runs again.



# Requirement 3: Time Travel on the Conversation History



# Requirement 4: Expressive Operators, e.g. loops

## One flattened operator

**SingleOp**

```
for f in files:
    df = load(f); df = df[df.rating >= 8];
plot(df)
```

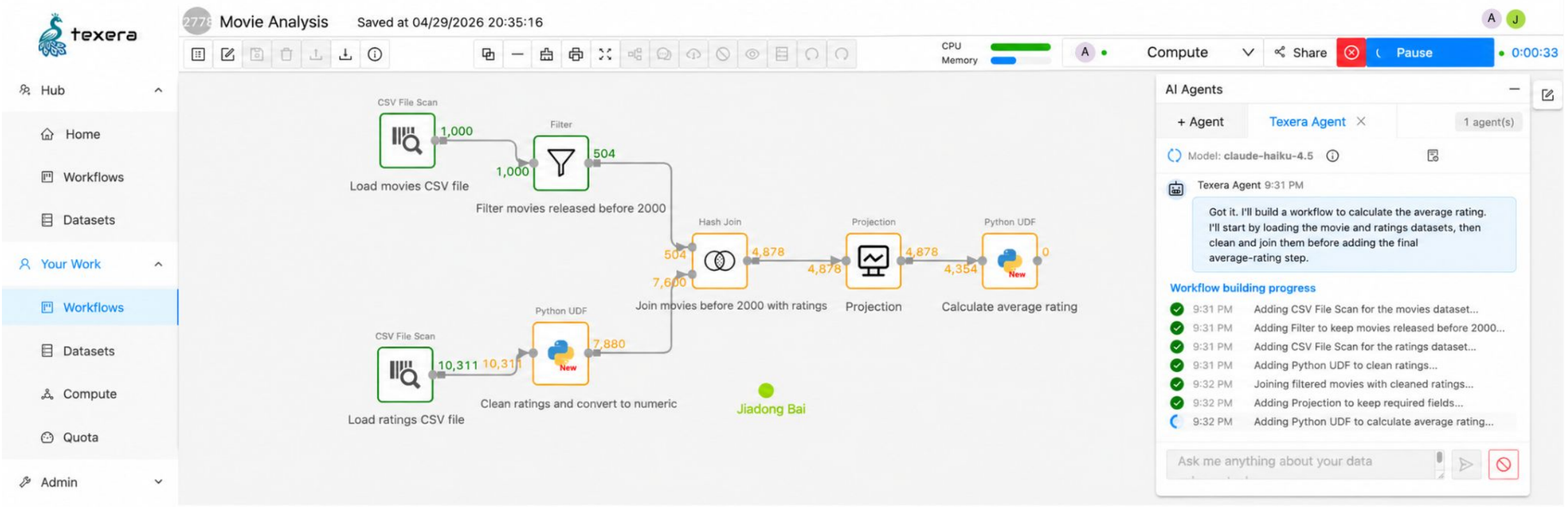
*all logic hidden inside code*

## ForLoop operator with a sub-DAG

**ForLoop** once per iteration

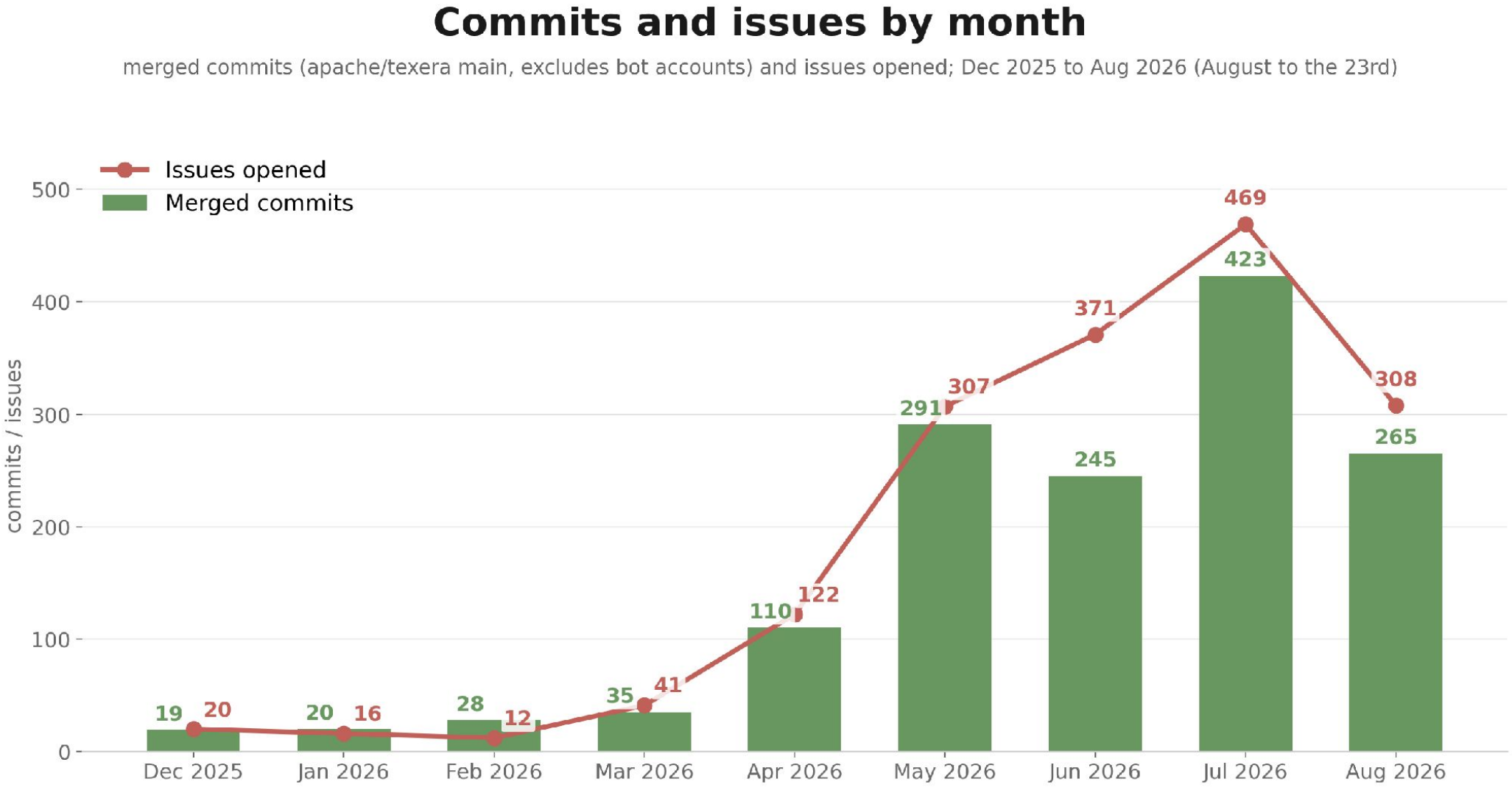
*sub-DAG: every step is visible*

# Apache Texera: Human-AI Collaborative Data Science Using Visual Workflows




- Native **DAG engine** with a web UI
- Operators in **ETL, UDF, ML, Visualization**
- Focus: **Human–AI collaboration**

# Apache Texera: Active Open-Source Community



# Summary

1. Benefits to **humans**?

   Visual, intuitive, precise; inspect and give feedback anywhere

2. Benefits to **agents**?

   Higher accuracy, lower cost

3. Requirements for the **runtime engine**?

   A native DAG engine for human-agent collaboration

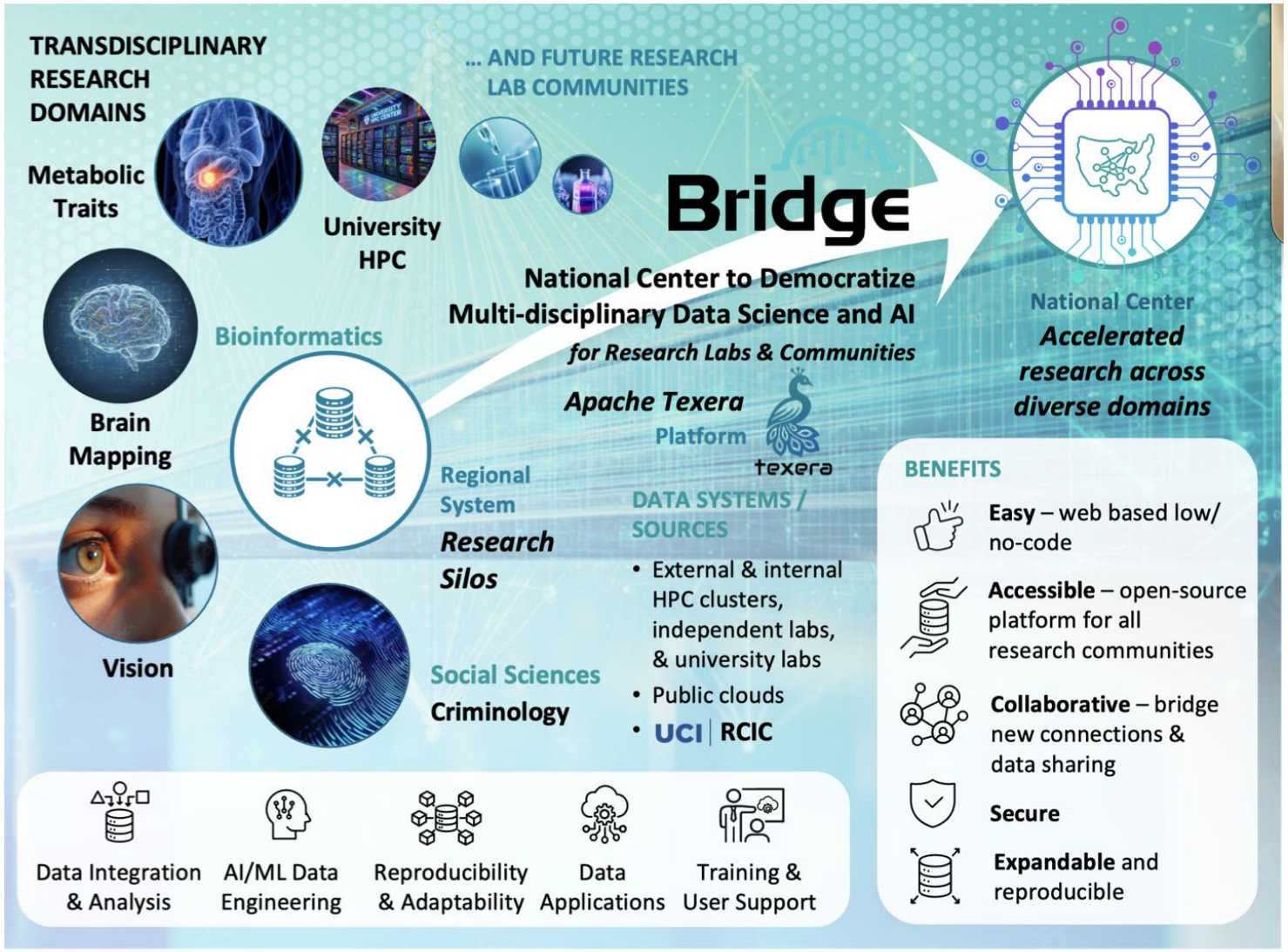
# NSF Bridge Infrastructure: Call for Collaborations!

**$9M**

NSF, 3 years

**from 7/12/2026**

**Looking for partners**



# Thinking in DAGs

## How Humans and Agents Collaborate in Data Science

**Speaker: Chen Li**
University of California, Irvine

Thank you!