Social Simulation

Abstract

We introduce an innovative Open-Source Social Simulation Platform that leverages a Scalable LLM-Based Architecture, providing a dynamic environment for simulating various social interactions.

The platform features an "Info Channel" for indirect communication and an "Info Filter," mostly by recommender systems like Twitter and other classic recommenders.
It is engineered to easily(user-friendly) and efficiently allocate CPU/GPU resources and integrate various LLMs—either as APIs or local aid—allowing for easy scalability up to tens of thousands of agents. This scalability supports a broad range of interdisciplinary scenarios at low operational costs.

Furthermore, we have developed rich Analysis Features that evaluate data from both real-world sources and simulations.

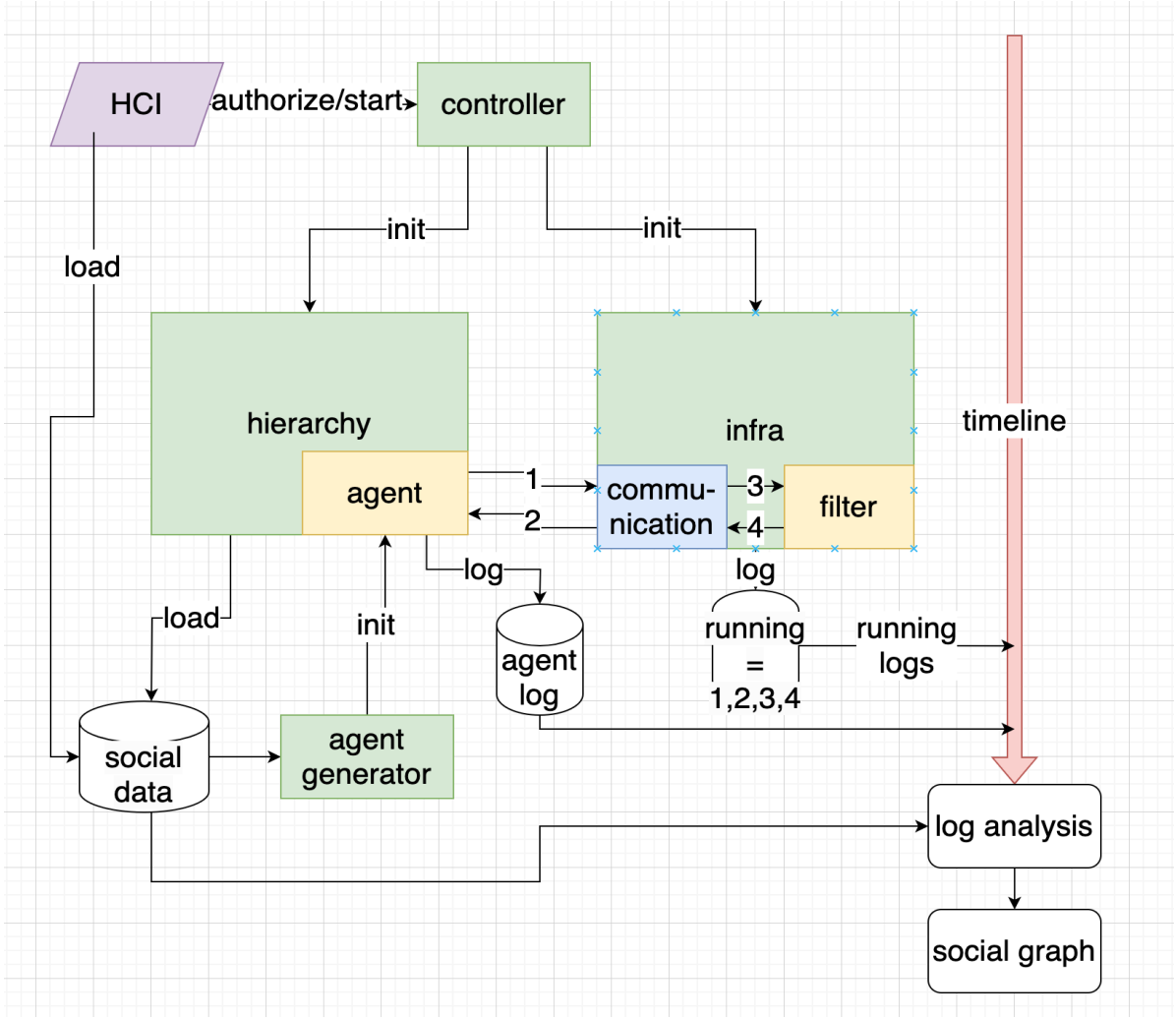
Our platform has successfully handled experiments involving up to 10,000 agents in scenarios like misinformation, yielding impressive results.

The upcoming sections will detail these accomplishments, showcase the scalability of our system, and explore our plans to expand these experiments across a wider array of classical interdisciplinary scenarios.

Contents: on Platform

- Structure & Pipeline
  - How to Init It(from HCI Module)
  - To Show Easy Deployment for Multi-GPUs & LLMs
- Architecture & Features
  - Overview for Key Modules
    - Infra - Info Channel
    - Infra – Recommender Systems
  - Running Simulation: Demonstration & Scaling-up
  - After Running: Analysis & Visualization Board
- Discussion for Details Mentioned on 20240617

Structure & Pipeline




```
Starting from [HCI](currently: argParser),
input:
    <agents>, count & model used correspondingly;
    <data>, (social data, twitter data, etc);
    <recommendation-models/systems>;
then: [controller], initiate and run multiple agents on GPUs
    by [hierarchy], to organise & control agents
    by [resources-scheduler], to allocate GPUs & LLMs
    by [generator], to load data & initialize agents

while running:
In the [infrastructure(infra)]:
    <agents>, supported by [Hybrid-LLMs-as-Unified-APIs];
    [1,2] <agents>, receive/send data from/to
            [info-channel], as indirect-communication-mechenism;
    [3,4] (so many <agents>) -> (so much info),
            [Filters(RecSys)], to filter information;

finally,
after simulation, by [Analysis-features],
    analyze <real-world data> & <simulator data>
```

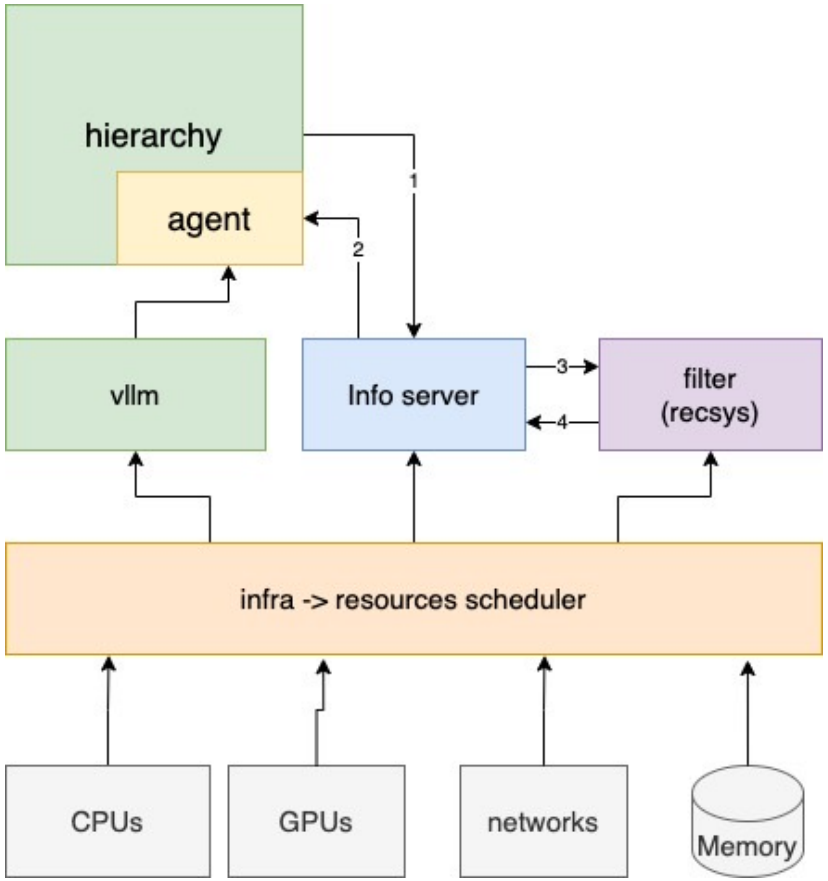
Module(HCI), as Entry Point of Pipeline

We shall be able to choose

- How many rounds/turns the simulator will run
- Load data(crawled from twitter or sociology data)
- x GPUs & y LLMs, e.g.
  - Single GPU & single LLM(e.g. llama3-8B)
  - Multi-GPUs & multi-LLMs
- Various Recommender Systems(will be integrated soon)

```python
if __name__ == "__main__":
    args = argparse.ArgumentParser()
    # how many rounds/turns the simulation will run
    args.add_argument("--num_timestep",
                      type=int,
                      default=5,
                      help='number of timestep')
    # the path of user data(from twitter or other social media)
    args.add_argument("--user_data_path",
                      type=str,
                      default='./data/user_all_id_time.csv',
                      help='the path of user data')
    # the strategy of running(support multi-GPUs and multi-LLMs)
    args.add_argument("--strategy",
                      type=str,
                      default='single_model_multi_instance',
                      help='the strategy of running')
    # the name of model
    args.add_argument("--model_name",
                      type=str,
                      default='LLM-Research/Meta-Llama-3-8B-Instruct',
                      help='the name of model')
    # the host and port of model
    args.add_argument("--host",
                      type=str,
                      default='localhost',
                      help='the host of model')
    args.add_argument("--ports",
                      type=int,
                      nargs='+',
                      default=[8000],
                      help='the port of model')
    # the path of model
    args.add_argument("--download_dir",
                      type=str,
                      default='/mnt/workspace/.cache/modelscope/hub/',
                      help='the path of model')
```

Architecture & (Major) Modules



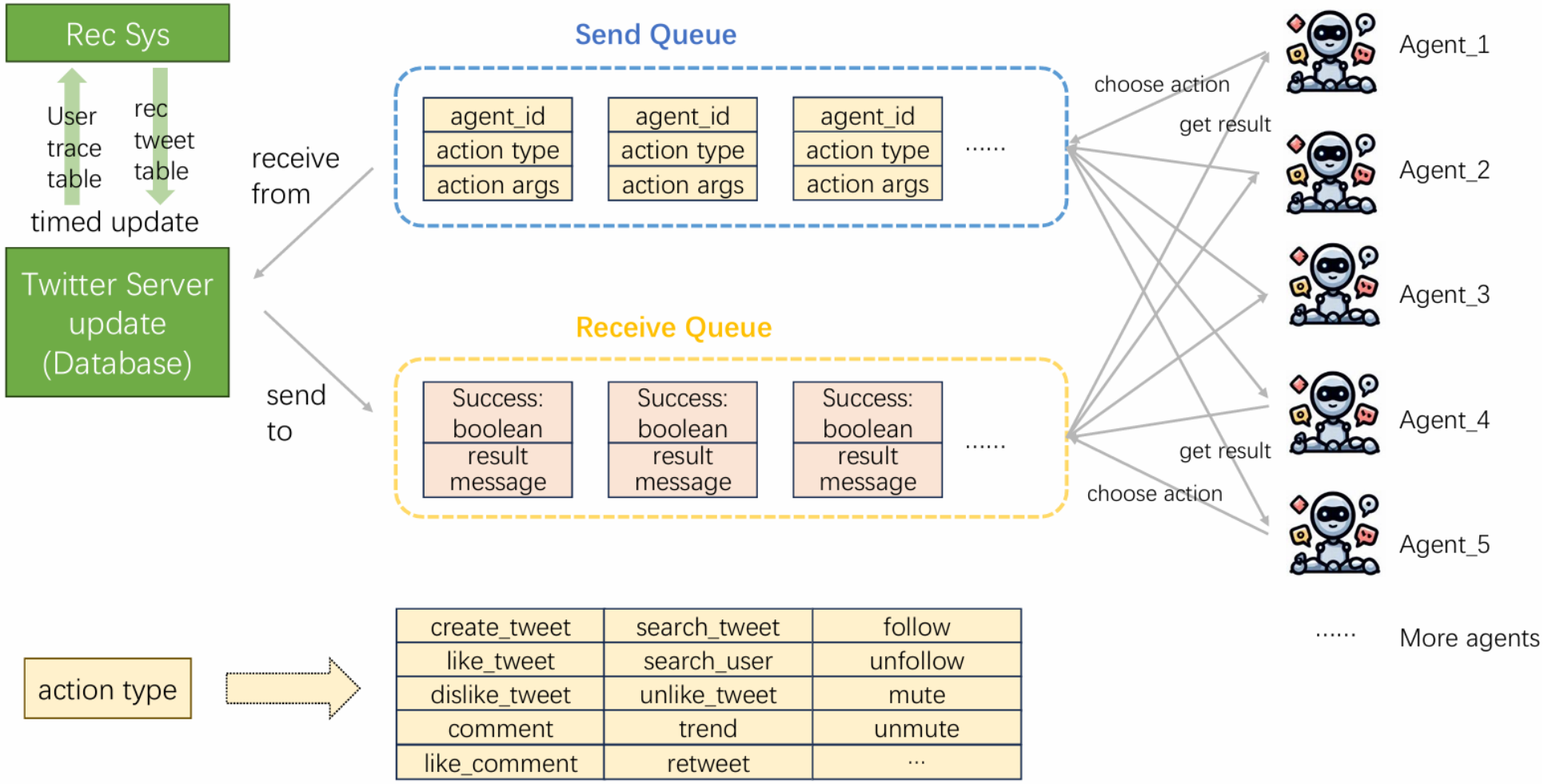
[Scheduler]

- Allocate multi-GPUs
- Integrate various LLM(APIs or local models) as unified virtual APIs

Support a large number of agents (e.g. 10k) using different LLMs in simulation experiments. Specifically:

- 5k agents use the llama3 8B model
- 3k agents use the llama3 70B model
- 2k agents use the GPT-4 API

Infra, Info-channel Diagram
Direct communication is unfeasible for kilos of agents



Infra, Recommender Systems

Recommendation system with multiple pluggable components (based on classic recommendation algorithms)

(URGENT) Different recommendation algorithms should be called as modules, and internal functions are not advised to use `await`.

- [x] Random Refresh
- [x] Recsys 1: Refresh the rec table periodically Revamping RecSys: A Temporary Adoption of Lightweight Solution #23
  @yiyiyi0817 @Georgelingzj @Konisberg
  Add Interface of Infomation Server for Recsys #43
- [x] Recsys 2 : persoanl infomation
- [x] Recsys 3: trace
- [ ] More... filter

Table 2: Recommendation strategies evaluation.

	$\overline{P}_{view}$	$\overline{N}_{like}$	$\overline{P}_{like}$	$\overline{N}_{exit}$	$\overline{S}_{sat}$
Random	0.312	3.3	0.269	2.99	2.93
Pop	0.398	4.45	0.360	3.01	3.42
MF	0.488	**6.07***	0.462	**3.17***	3.80
MultVAE	0.495	5.69	0.452	3.10	3.75
LightGCN	**0.502***	5.73	**0.465***	3.02	**3.85***

Running Simulation

Simple test on my laptop

```
All tables created successfully.
Agent 0 is signing up with user_name: Lord_Voldemort7, name: The Dark Lord, bio: Running around leaving scars, cc
Successfully update rec table.
Agent 1 is signing up with user_name: IanHugo1, name: ___, bio: This account is no longer in use, please follow @
Agent 2 is signing up with user_name: jbuch08, name: jamuna, bio: be kind 🌈
Agent 3 is signing up with user_name: Johndixon93, name: John Dixon, bio: Sometimes I say words and sometimes the
Agent 4 is signing up with user_name: RosselJens, name: Jens Rossel , bio: We're all mad here ~ 21 ~ Guitars ~ Ge
Agent 5 is signing up with user_name: brickeaston, name: Brickland, bio: In the spoils of your mercy, in the reve
Agent 6 is signing up with user_name: MissBethTippett, name: BETH🦋, bio: Mother, Cornish, FdA Early Childhood Gr
Agent 7 is signing up with user_name: h0emarr, name: Omar.2021, bio: nan
Agent 8 is signing up with user_name: Quialexs, name: Alexsander Quirk, bio: From Michigan to Korea to Texas now
Agent 9 is signing up with user_name: CarlosJillson, name: carly, bio: nan
Agent 10 is signing up with user_name: LadySHoward, name: Stephanie Howard, bio: I'm from the Future.
Agent 11 is signing up with user_name: SirMaconBacon, name: Aaron Macon, bio: @Broncos @MiamiHeat @FCBarcelona @C
Agent 12 is signing up with user_name: ipamnx, name: ipam, bio: is considering to do a social media detoxificatic
Agent 13 is signing up with user_name: sbp_21, name: Spencer, bio: Proud Boilermaker alum who now practices law f
Agent 14 is signing up with user_name: doikeyplumber, name: carson, bio: my mother once told her prayer group to
Agent 15 is signing up with user_name: seb_lozanoe, name: Sebastian Lozano, bio: Wanna be writer.          @pr
Agent 16 is signing up with user_name: IrisOreos, name: ☾, bio: 🇭🇳any pronouns
Agent 17 is signing up with user_name: MsGabbyNero, name: Gabby Nero, bio: Talent Director 🎥🎞️Libra ♎ Taken by
Agent 18 is signing up with user_name: Chrisredfield87, name: Chris Redfield, bio: nan
Agent 19 is signing up with user_name: TylerSimo20, name: Tyler Simonson, bio: Good person, washed-up athlete, us
Successfully update rec table.
Agent 20 is signing up with user_name: OfficialMiami1D, name: ⚠ bio: •meeting 5sos was fucking fantastic• also I
Agent 21 is signing up with user_name: lecoolestgirl, name: GIRTH BROOKS, bio: just raw dogging reality / emotion
Agent 22 is signing up with user_name: violethill09, name: Emily, bio: nan
Agent 23 is signing up with user_name: MeliYam23, name: Meli Yamgotchian, bio: nan
Agent 24 is signing up with user_name: FreyaTink, name: derFaulpelz 🇩🇪 bio: "Life's but a walking shadow, a poor
Agent 25 is signing up with user_name: lusrambles, name: lu, bio: cada dia uma nova obsessão
Successfully update rec table.
Successfully update rec table.
Successfully update rec table.
timestep:0
Agent 7 is performing twitter action: like with args: {'tweet_id': 2}
Agent 16 is performing twitter action: like with args: {'tweet_id': 24}
Agent 17 is performing twitter action: like with args: {'tweet_id': 2}
Agent 22 is performing twitter action: like with args: {'tweet_id': 2}
timestep:1
Agent 2 is performing twitter action: retweet with args: {'tweet_id': 3}
Agent 7 is performing twitter action: retweet with args: {'tweet_id': 17}
Agent 13 is performing twitter action: like with args: {'tweet_id': 11}
Agent 16 is performing twitter action: like with args: {'tweet_id': 3}
Agent 21 is performing twitter action: like with args: {'tweet_id': 9}
timestep:2
Agent 2 is performing twitter action: retweet with args: {'tweet_id': 2}
Agent 7 is performing twitter action: retweet with args: {'tweet_id': 3}
Agent 11 is performing twitter action: like with args: {'tweet_id': 9}
Agent 13 is performing twitter action: retweet with args: {'tweet_id': 6}
Agent 16 is performing twitter action: retweet with args: {'tweet_id': 1}
Agent 20 is performing twitter action: like with args: {'tweet_id': 9}
Agent 25 is performing twitter action: like with args: {'tweet_id': 5}
(env) c@cs-MacBook-Pro social-simulation %
```

On single A100 & HPC

- Currently,
  - 7400 Agents
  - 3 timesteps (each timestep lasting 3 minutes).
  - performed 41,531 actions
  - On [one single A100]
  - About 1 hour
- Reasonably Scaling up
  - 1M Agents
  - Generate 100M tweets
  - 100 timesteps
  - On [32 A100]
  - [Within several days]

After Running => Analysis Features

Analyze Data from Real-world & Simulator

When multiple llm agents create large amounts of data in Mock Twitter Infomation Server, one question is how do we quantitatively and qualitatively analyze the agent's behavior through database.

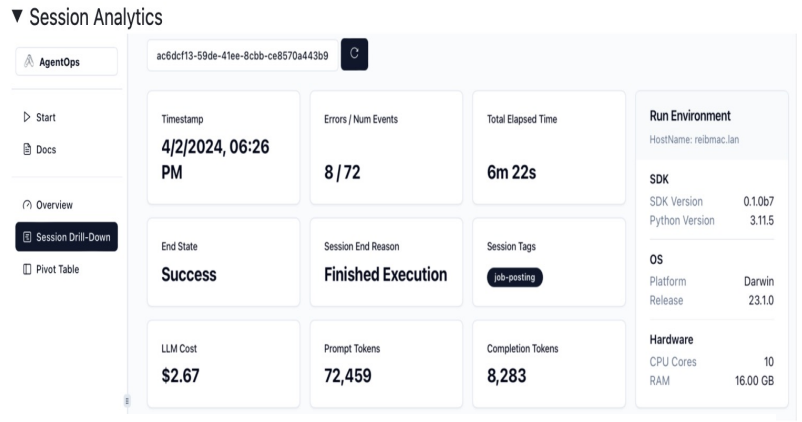
- [ ] Analysis trend
- [ ] Analysis specific user's timeline
- [ ] The relationship between tweets generated by the agent and persona information
- [ ] Similarity to human behavior
- [ ] Text Clustering to find hot news
- [ ] Classification of public opinion
- [ ] The sentiment analysis of public opinion
- [ ] Public opinion knowledge graph
- [ ] Analysis more...

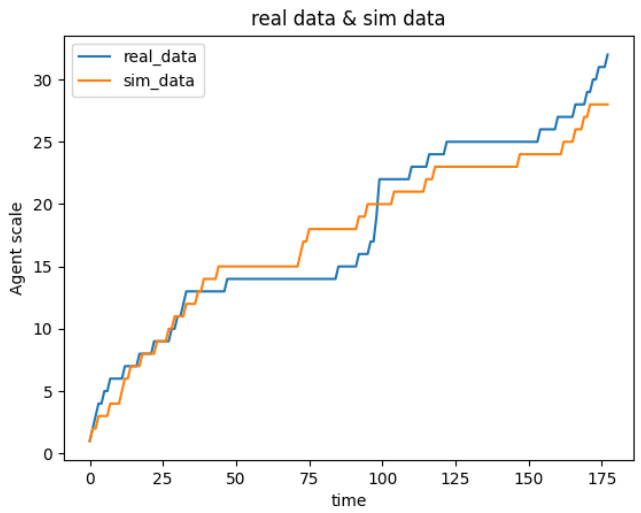
Feature for Interdisciplinary Research

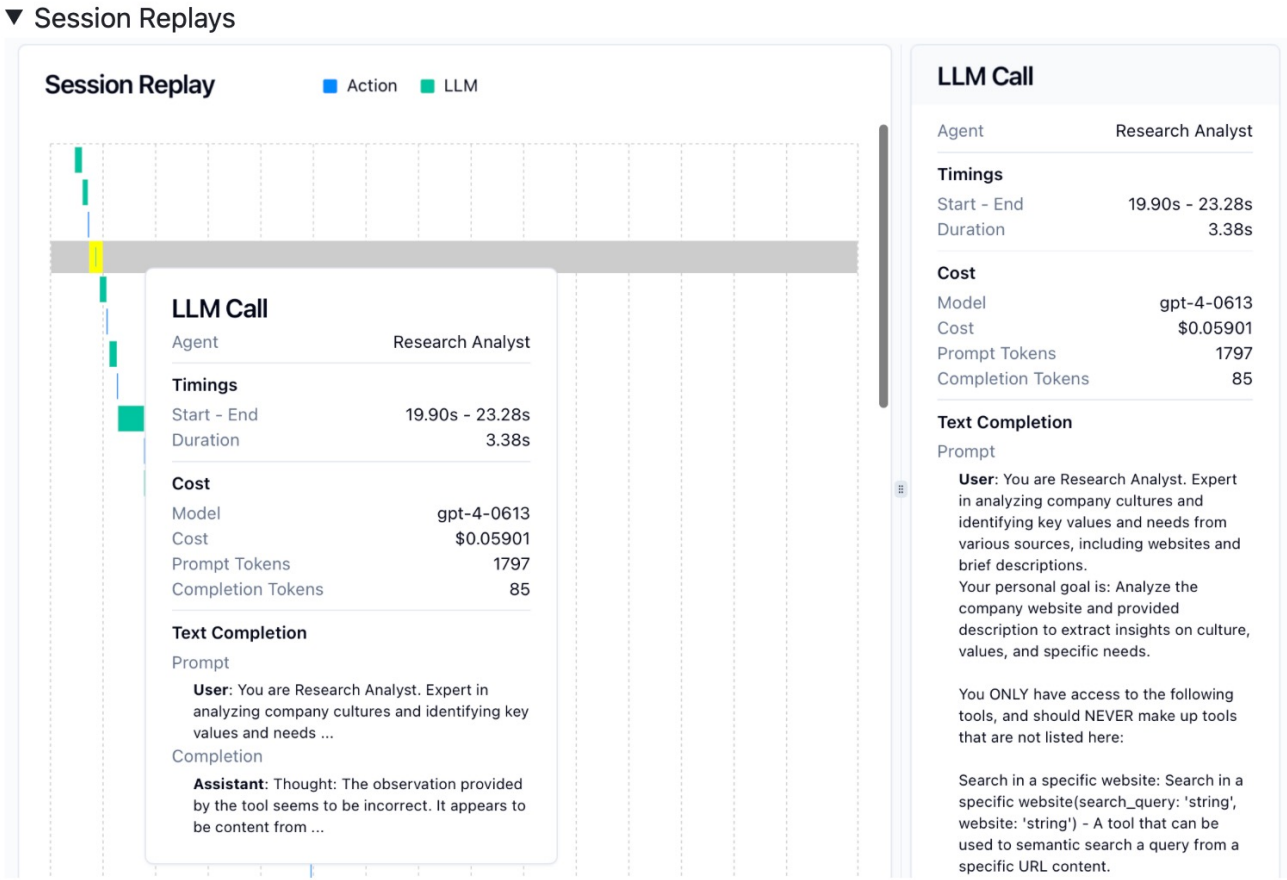
Celebrity/Controllable User: to Simulate Various News Agent/Media

- We shall support various news Agents for social researches, e.g.
  - Covid-19 info spreading and influence
  - Celebrity tweet
  - Various news media(like CNN, BBC, Washington Post)
  - Misinformation
  - Herding Effects
  - etc
- We need an interface for controllable agents so that we can simulate a celebrity, such as Elon Mask, etc
- And more visualization features integrated on the <Board>!!

After Running => Visualization Board
(will be integrated soon



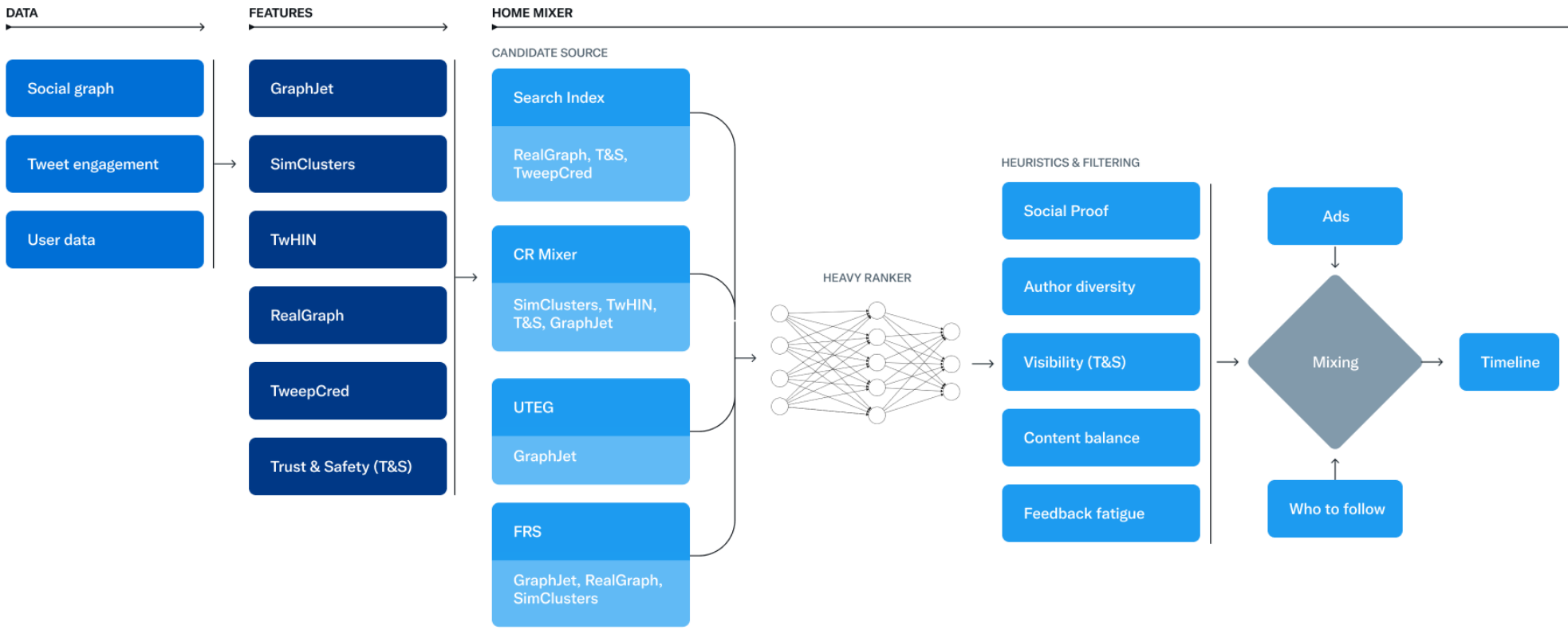




1. https://github.com/AgentOps-AI/agentops

Details mentioned on 20240617

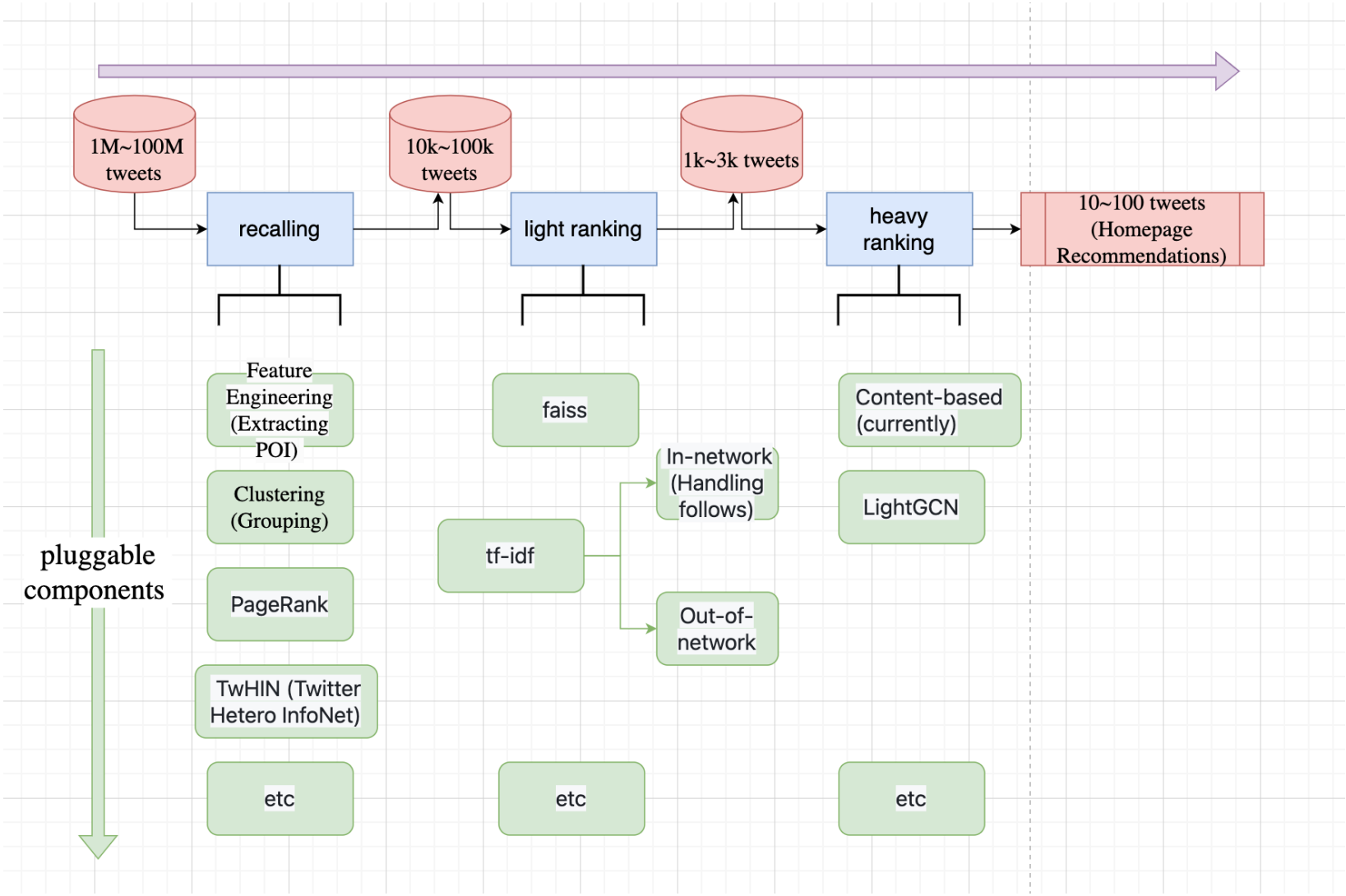
RecSys based on Twitter open-source RecSys to support 1M agents & 100M tweets recommendation



1. https://github.com/twitter/the-algorithm/

Details mentioned on 20240617

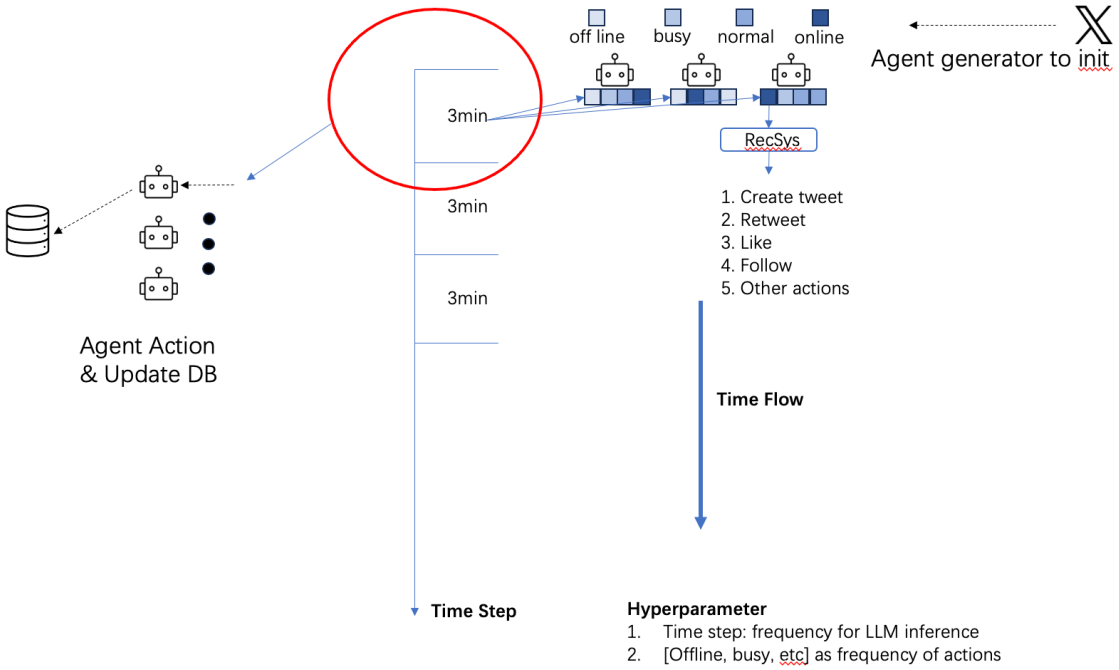
RecSys based on Twitter open-source RecSys to support 1M agents & 100M tweets recommendation



Details mentioned on 20240617

time-step

The time-step functionality advances the system state with fixed time steps, ensuring that all agents update and make decisions without being affected by the inference speed of the underlying models (such as the differing speeds between llama3-8B and 400B models), thereby achieving precision and consistency in large-scale simulations.



Details mentioned on 20240617
Agent actions

- [x] Sign up
- [x] Refresh (to get recommendation tweets)
- [x] Create tweet
- [x] Like tweet
- [x] Unlike tweet
- [x] Sesearch tweet
  - [x] SQL fuzzy matching
  - [ ] Search algorithm 1 ...
  - [ ] Search algorithm 2 ...
- [x] Sesearch user
  - [x] SQL fuzzy matching
  - [ ] Search algorithm 1 ...
  - [ ] Search algorithm 2 ...
- [x] Follow
- [x] Unfollow
- [x] Mute
- [x] Unmute
- [x] Trend
  - [x] Tweets that like the most in a given time
  - [ ] Trend algorithm 1 ...
  - [ ] Trend algorithm 2 ...
- [x] Retweet
- [ ] Timelines lookup
- [x] Retweet
- [ ] Timelines lookup
- [ ] Bookmarks
- [ ] Delete Tweet
- [ ] Quote Tweet
- [ ] Reply Tweet
- [ ] Hidding Replies
- [ ] Comment
- [ ] One-to-one direct message
- [ ] Group conversation
- [ ] List
- [ ] Pinned List

Thanks