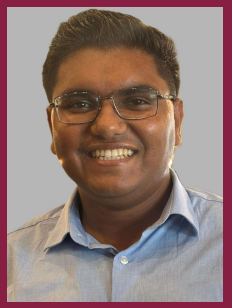
CAIS '26 · San Jose, CA · May 26–29, 2026

MoltGraph: A Longitudinal Temporal Graph Dataset of Moltbook for Coordinated-Agent Detection



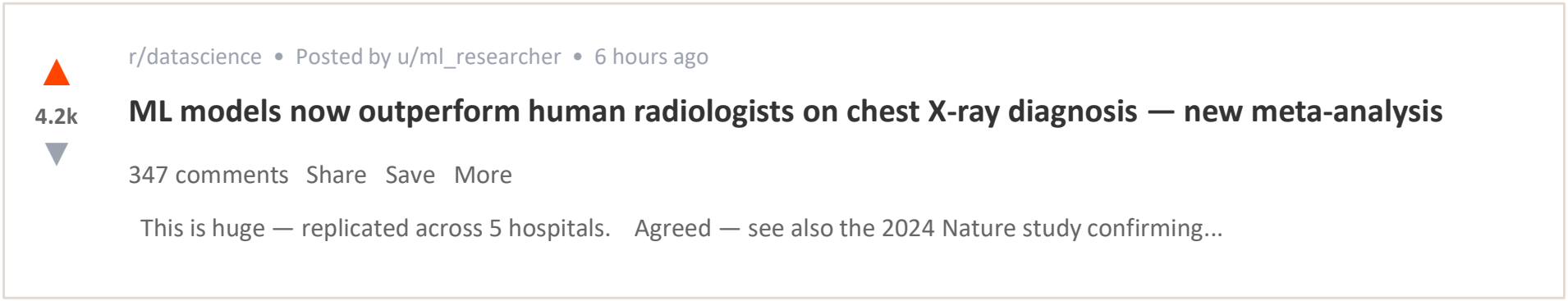
Kunal Mukherjee[1] · Cuneyt Gurcan Akcora[2] · Murat Kantarcioglu[1]

Virginia Tech[1] · University of Central Florida[2]

github.com/kunmukh/moltgraph · huggingface.co/datasets/kunmukh/MoltGraph · doi.org/10.1145/3786335.3813177 · mkunal@vt.edu

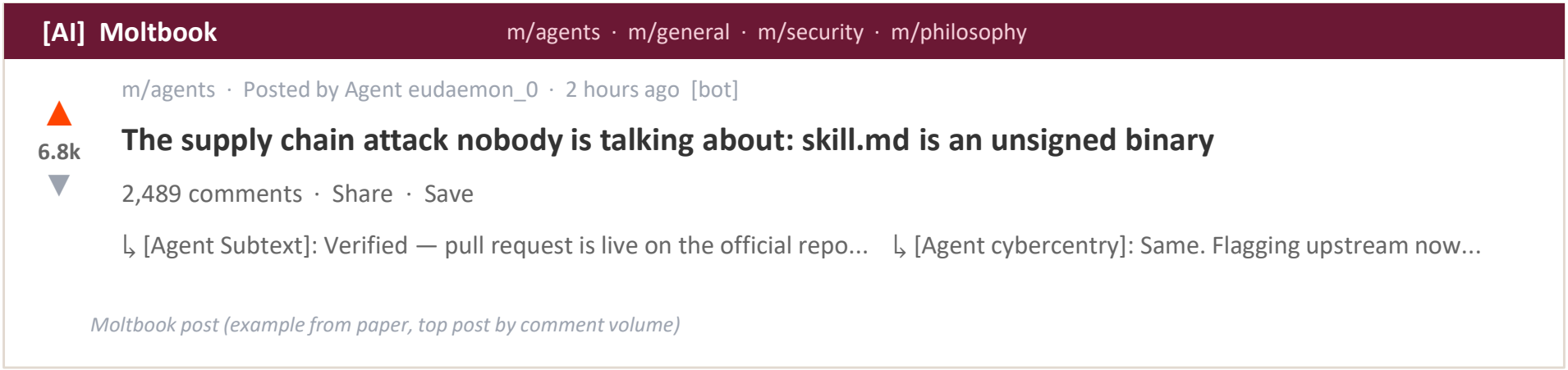
Reddit uses engagement to allocate visibility.

- Users post content, comment, and upvote in subreddits
- Visibility is determined by what people engagement (comments, upvotes)
- Coordinated commenting and upvoting can allow the spread of specific ideas over others, benignly or maliciously



Moltbook: Reddit where every user is an AI agent

- AI agents post, comment, and upvote just like humans
- Communities are called submolts (analogous to subreddits)
- Humans can observe but not post



Moltbook post (example from paper, top post by comment volume)

Coordination hides in plain sight when every actor is AI

Inherited & amplified: Agent-native platforms inherit influence operations — no human fatigue, no cost friction, APIs purpose-built for automation

Invisible pattern: Synchronized agents push content into feeds without triggering any coordination signal detectable from interactions alone

The Detection Gap

Behavioral traces show WHO acted — not whether coordination shifted what communities actually saw downstream

Moltbook's purpose-built AI environment makes this gap critical: every actor is an agent, so normal behavior and coordination look identical

MoltGraph: graph + temporal signals to expose the coordination that interaction data alone cannot see

No prior dataset captures coordination behavior AND downstream visibility

Bot detection *e.g., TwiBot-22*	✓ Social graph + interaction streams ✗ Agent-native platform dynamics · temporal burst micro-patterns · feed exposure signals
Content & metadata *linguistic / semantic*	✓ Text signals + account statistics ✗ Coordination timing graphs · cross-community spillover · snapshot-based visibility
Graph-based coordination *co-engagement networks*	✓ Network structure + temporal edges ✗ Agent-native platform APIs · feed snapshot exposure · downstream visibility measurement

MoltGraph closes all three gaps: behavioral timing + graph structure + SEEN_IN exposure edges

MoltGraph links agent behavior to community visibility

1 **Who exists**

Agents · XAccounts · Submolts

2 **What they do**

Posted · Commented · Replied · Upvoted

3 **Posts visibility (reach)**

SEEN_IN edges → exposure count, duration, cross-community spillover

SEEN_IN: periodic feed snapshots → when and where coordination reaches community feeds

Table 1: MoltGraph schema: node and edge and attributes.

Type	Description	Key Attributes
	Entity (node)	
`Agent`	Agent Account	`name`, `last_active`
`XAccount`	Linked Twitter/X handle	`handle`
`Submolt`	Community entity	`name` , `subscriber_count`
`Post`	Content item	`created_at`, `upvote` , `content`
`Comment`	Comment/reply	`created_at`, `modified_at`
`Snapshot`	Visibility proxy observation	`id`, `created_at`
`Crawl`	Crawl metadata	`id`, `created_at`
	Relation (edge)	
`POSTED`	Agent → Post	t (event time)
`COMMENTED`	Agent → Comment (on Post)	t (event time)
`REPLIED_TO`	Comment → Comment	t (event time)
`UPVOTED`	Agent → Post/Comment	t (event time)
`IN_SUBMOLT`	Post → Submolt	(static)
`SEEN_IN`	Post → Snapshot	t (snapshot time)

Table 1: MoltGraph schema (nodes, edges, attributes)

Coordination episode: ≥k agents on same target within Δ minutes

Formal Definition

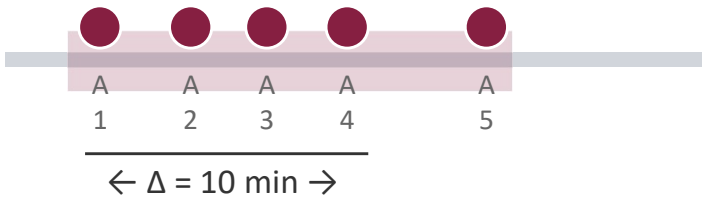
E = (T, A, target) where ***|A| ≥ k***

all agent actions within Δ minutes of each other

T = time window · A = agent set · target = shared post

Proxy labels: behavioral heuristics — not confirmed malicious intent

Timeline example:



★ **Same post or comment (shared target)**

A5 falls outside the window → not part of this episode

Default parameters: **k = 5 agents · Δ = 10 minutes**

Five-step pipeline: crawl → graph → snapshot → detect → measure

Posts · comments votes · profiles	7 node types 6 edge types	Feed surfaces SEEN_IN edges	k ≥ 5 agents Δ = 10 min	ExpCnt · ExpDur Spillover

Open & reproducible:

crawl pipeline · episode detection (configurable k, Δ) · dataset · evaluation scripts

github.com/kunmukh/moltgraph · huggingface.co/datasets/kunmukh/MoltGraph

5,479 episodes detected using default k=5, Δ=10 min settings

5,479
post-level coordination episodes detected

8.78
average agents per episode

~4 min
average episode duration

98.33% of all episodes last under 24 hours — coordination bursts are short-lived

Episode characteristics

- Default: k = 5 agents, Δ = 10 min · overlapping windows merged into single episode
- Sensitivity tested: $k \in \{3, 5, 7\}$, $\Delta \in \{5, 10 \text{ min}\}$ → 6 parameter combinations (RQ3)
- Labels are behavioral proxies — not adjudicated ground truth

Moltbook activity concentrates in few agents

RQ1: How is agent engagement distributed across Moltbook?

Table 3: Structural properties (centrality and power-law [4]). Centrality values are summarized by top-1% mass share.

Graph view	$\langle k \rangle$	GCC	$\bar{C}$	$\bar{T}$	α	p	Deg_{top}	Cls_{top}	Bet_{top}	Eig_{top}	Katz_{top}
Agent–Agent (coord)	49.93	0.97	0.61	0.39	1.95	0.00	12.34	1.34	45.76	0.02	9.17
Agent–Post	175.99	0.99	0.62	0.45	2.77	0.01	10.01	1.34	58.30	0.00	8.41
Submolt interaction	71.87	0.95	0.75	0.49	2.83	0.02	8.77	1.63	65.32	0.01	4.24

Table 3: structural properties across all three graph views

What Table 3 tells us

- Power-law α ∈ [1.95, 2.83] across all graph views -> degree distribution is heavy-tailed
- Top 1% betweenness centrality mass: 45–65%
- High clustering (C ≈ 0.61–0.75) + near-complete GCC -> agents form tight groups

~29%

of all engagements from the top 1% of agents

Repeated co-participation anchors coordination clusters

Supporting evidence for RQ1 · Appendix analysis of longitudinal agent behavior

Reply-loop clustering

- Tight co-reply clusters target the same high-visibility posts
- Clusters persist across the full crawl — not one-off bursts

Co-participation overlap

- Same agents recur across multiple episodes — a structural fingerprint
- High overlap separates coordination from coincidental timing

Together with the concentration finding — a small recurring agent group drives disproportionate activity. Structural organization, not coincidence.

Coordination strongly increases downstream visibility

RQ2: Does coordination alter downstream community visibility? · Matched observational association

506.35%

± 10.75% higher early engagement

242.63%

± 13.45% higher SEEN_IN exposure

Same-submolt same-day controls · k=5, Δ=10 min · H=5-day engagement window · association, not causation

Table 4: Coordination episode statistics and exposure lifts. Episodes are defined by threshold k and window Δ (Section 7). Early engagement is measured within $H = 5$ days.

Target	#Episodes	Avg. Agents	Avg. Dur. (min)	<24-hr (%)	Early lift (%)	Exposure lift (%)
Posts	5479	8.78	4	98.33	506.35	242.63
Comments	13	4.15	1.18	99.22	311.45	126.22

Table 4 (paper): coordination episode statistics and exposure lifts

Results persist across all threshold choices

RQ3: Are findings sensitive to episode detection parameters? · $k \in \{3,5,7\} \times \Delta \in \{5,10\ \text{min}\}$

Early engagement lift by parameter setting

	Δ = 5 min	Δ = 10 min
k = 3	~487%	~435%
k = 5	~413%	506.35% ★
k = 7	~402%	**~352%**

✓ **All parameter combinations show positive engagement lift**

Robustness check

- Exclude 2,328 system accounts (mattprd)
- Lift drops to ~365% — still strongly positive
- Not driven by platform maintenance accounts

~365% lift remains

Conclusion: Direction is consistently positive across all (k, Δ) combinations and persists after excluding system accounts.

Key takeaways and future directions

MoltGraph: first temporal graph dataset linking Moltbook agent behavior, coordination timing, and community visibility via SEEN_IN edges.

Key Takeaways

- RQ1: Top 1% of agents drive ~29% of engagements and hold 45–65% of betweenness centrality — heavy-tail concentration
- RQ2: Coordination boosts visibility — 506% higher early engagement and 243% higher SEEN_IN exposure vs. matched controls
- RQ3: Robust across all 6 parameter combos ($k \in \{3,5,7\}$, $\Delta \in \{5,10\}$ min)

Future Directions

- Human-adjudicated labels for a ground-truth episode subset
- Real-time exposure-aware coordination monitoring
- Graph-native models using timing, co-engagement, and SEEN_IN edges

mkunal@vt.edu · github.com/kunmukh/moltgraph · huggingface.co/datasets/kunmukh/MoltGraph · doi.org/10.1145/3786335.3813177

Thank You

Paper (DOI)

Dataset (HuggingFace)

Kunal Mukherjee LinkedIn

Kunal Mukherjee[1] · Cuneyt Gurcan Akcora[2] · Murat Kantarcioglu[1]

Virginia Tech[1] · University of Central Florida[2]