LF DECENTRALIZED TRUST

Trust Over IP Virtual Symposium

# Kwaai

## Building Decentralized AI Infrastructure: Safer, Faster, and Greener

*Building the Linux of AI*

Kwaai
Non-Profit Personal AI Lab

reza@kwaai.ai
https://www.kwaai.ai/
+1 661 7133031

# Agenda

**Intro - What is Kwaai? - Reza Rassool**

**Formulating AI Policy - Steve Vitka**

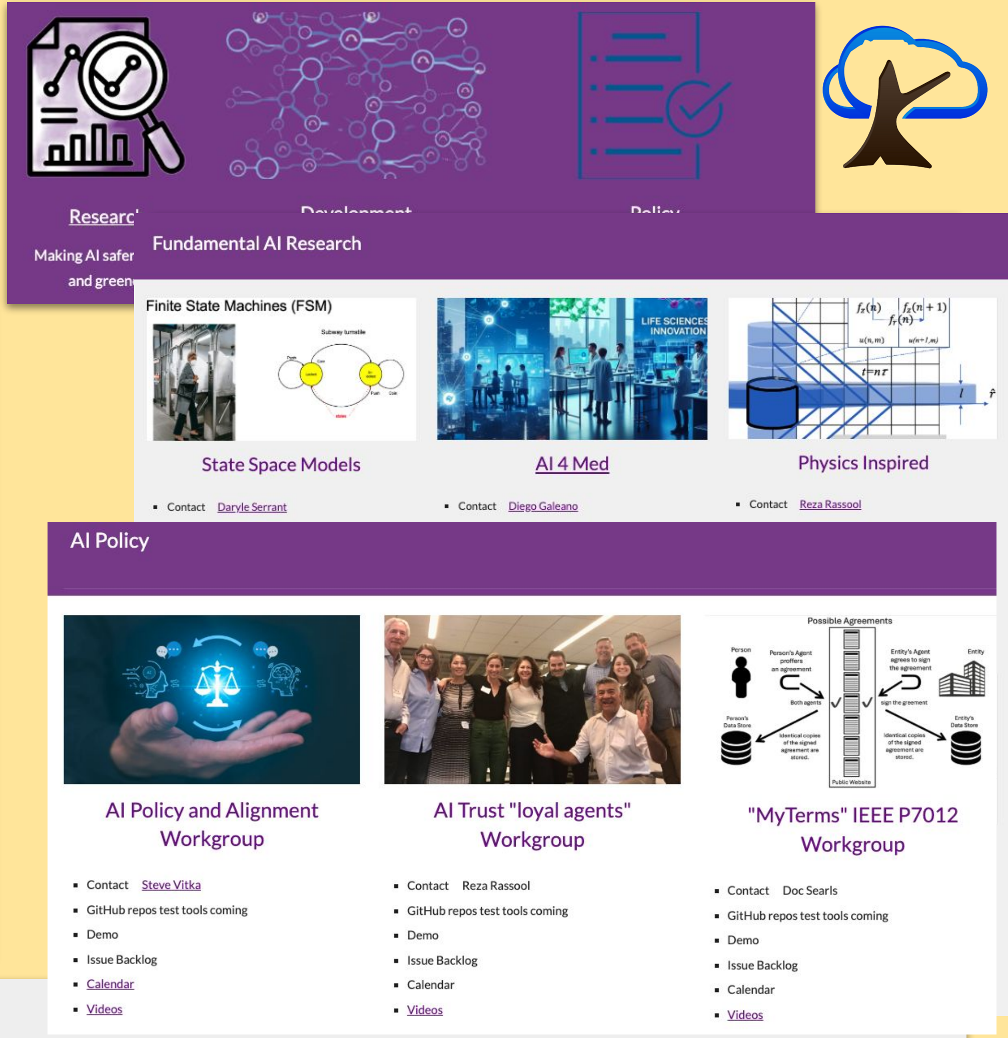
- Taxonomy
- The Personal Agency
- Partnering: MyTerms, Loyal Agents, GliaNET

**Researching Fundamental AI - Sulimon Sattari**

- Safer - Homomorphic Encryption
- Faster - Sub-quadratic Neural Networks
- Greener - Distributed Infrastructure

**Building Decentralized AI Infrastructure - Brian Ragazzi**

- πOS™ - Personal AI Operating System
- KwaaiNet - DePIN
- Verida - Decentralized Storage



Researc

Making AI safer
and green

Fundamental AI Research

Finite State Machines (FSM)

State Space Models

- Contact Daryle Serrant

AI 4 Med

- Contact Diego Galeano

Physics Inspired

- Contact Reza Rassool

AI Policy

AI Policy and Alignment Workgroup

- Contact Steve Vitka
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

AI Trust "loyal agents" Workgroup

- Contact Reza Rassool
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

"MyTerms" IEEE P7012 Workgroup

- Contact Doc Searls
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

Kwaai AI Lab https://www.kwaai.ai/

# My Story

After a Successful Career

**I retired with an angst that AI is in trouble…**

**And I need to fix it.**

**I launched Kwaai, a nonprofit AI Lab…**

**With a mission to democratize AI**

**Now, we are close to 1000 volunteers…**

**And we need your help.**

# What is Kwaai?

Kwaai is a 501(c)3 nonprofit, open-source AI lab.

We build Personal AI systems that put you in control of your data and decisions. We make AI safer, faster, and greener.

Our global, volunteer-driven movement is committed to transparency, privacy, and ethical innovation.

Build Decentralized AI

Research Next Gen AI

Maturing Homomorphic Encryption (HE) to Enable Privacy Preserving Vector Search

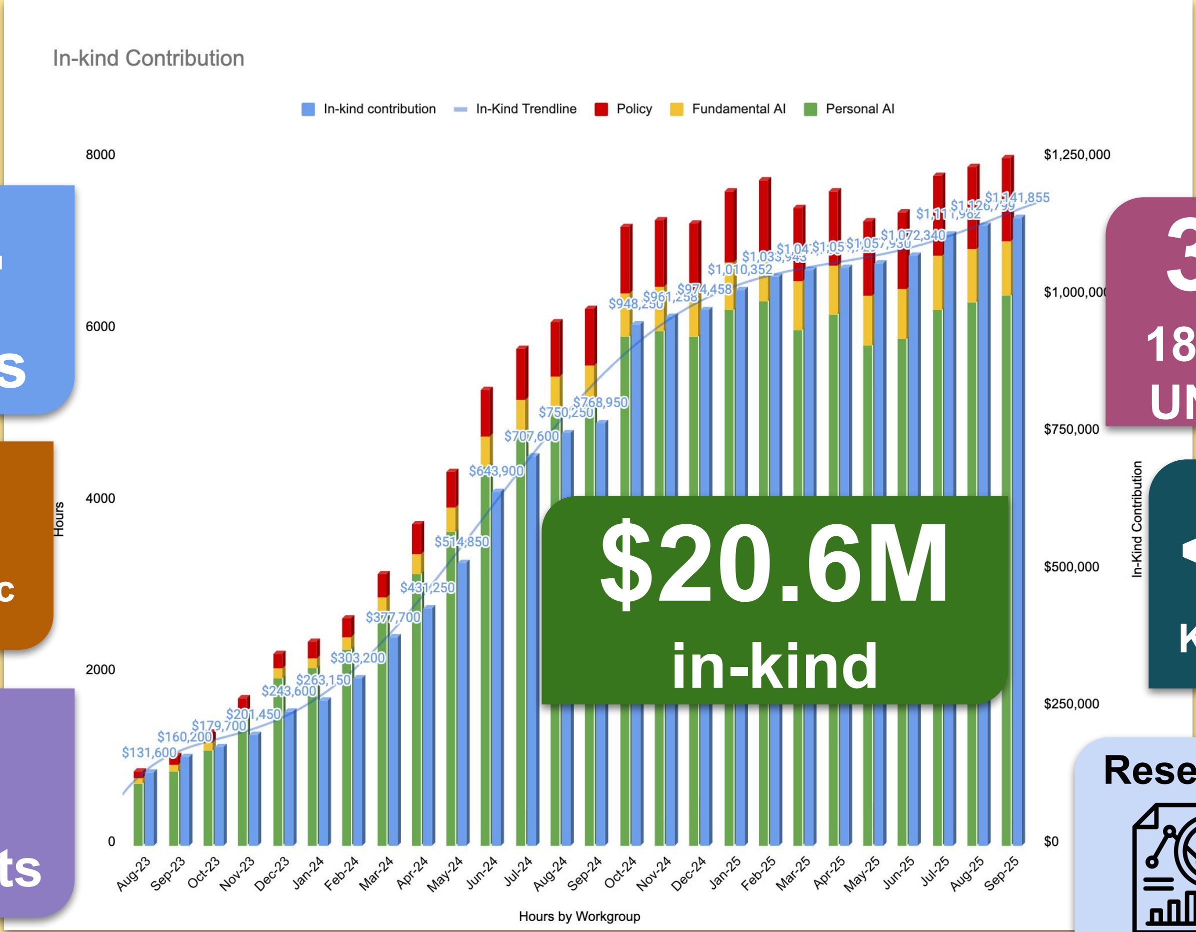
Draft Humane AI Policy

# Traction

**900+**
**members**

**45+**
**corp and academic partners**

**115+**
**Weekly meets**



**$20.6M**
**in-kind**

**3 Summits**
**18 Conferences**
**UN Workshop**

**</> πOS**
**KwaaiNet**

**Research** Making AI Safer, Faster, & greener.

Kwaai AI Lab https://www.kwaai.ai/

# Why Now?

- Now more than ever, we cannot rely on government to protect us from the excesses of big tech.
- There is no cavalry coming?
- The future of AI is being written today.
- Your action now ensures the next generation inherits tools for self-reliance, autonomy, and personal responsibility.

**You are the Cavalry!**

# Here’s How You Can Help

💪 🙋🏻‍♀️ 🙏 Volunteer or Intern

- Join the movement, contribute to the technical or policy think tank
- Develop Personal AI, Researching Fundamental AI, Drafting AI Policy
- Build Community, Automating our Back Office, Managing Hackathons

Support

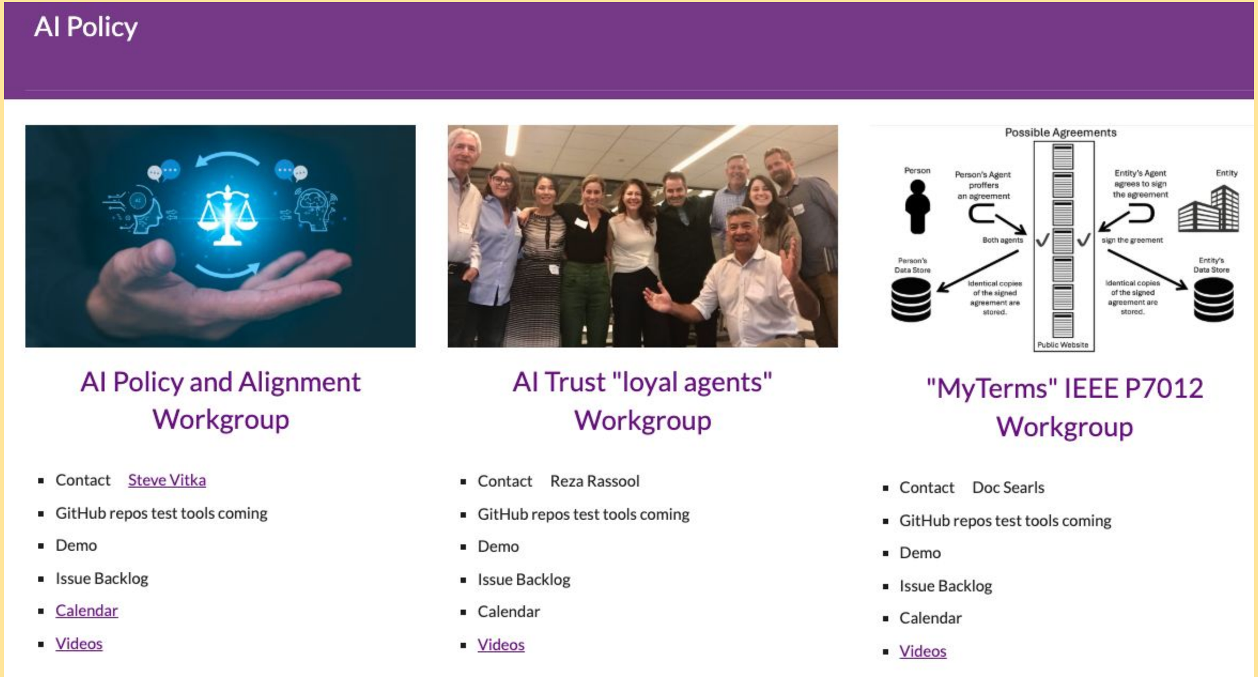
- 💸 Donate - Target Philanthropy at specific projects. Match our contribution
- 💵 Sponsor - Participate in Events or Become and Corporate Member
- 💰 Invest - Seed investments into Startups of the Kwaai Community

🤗 Share - Spread the word. Make warm introductions.

# Steve’s Slides

# Kwaai AI Policy

- Policy Partners
- My Personal Agency
- Taxonomy 4 Agent Systems- enables authentication.



AI Policy

## AI Policy and Alignment Workgroup

- Contact Steve Vitka
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

## AI Trust "loyal agents" Workgroup

- Contact Reza Rassool
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

## "MyTerms" IEEE P7012 Workgroup

- Contact Doc Searls
- GitHub repos test tools coming
- Demo
- Issue Backlog
- Calendar
- Videos

# Kwaai Policy Partnerships

**ToIP Foundation? -** hope this talk does the trick

**GliaNet Alliance -** A group of entrepreneurs and legal scholars forwarding fiduciary duties for the Net, I am working granular duties for agents that are multi-specified: 1. layman. 2. legalese 3. JSON 4. edge case database pointers

**My Terms / Intent Casting-** This multi-specification requirement is I copied from Doc Searls (and the Vendor Relations Managment Group he started ) via their My Terms protocol which is about to be published by IEEE. Personal agents should be pushing their terms of interaction at any entity they interact with, and more generally broadcasting verifiable owner-intent far and wide

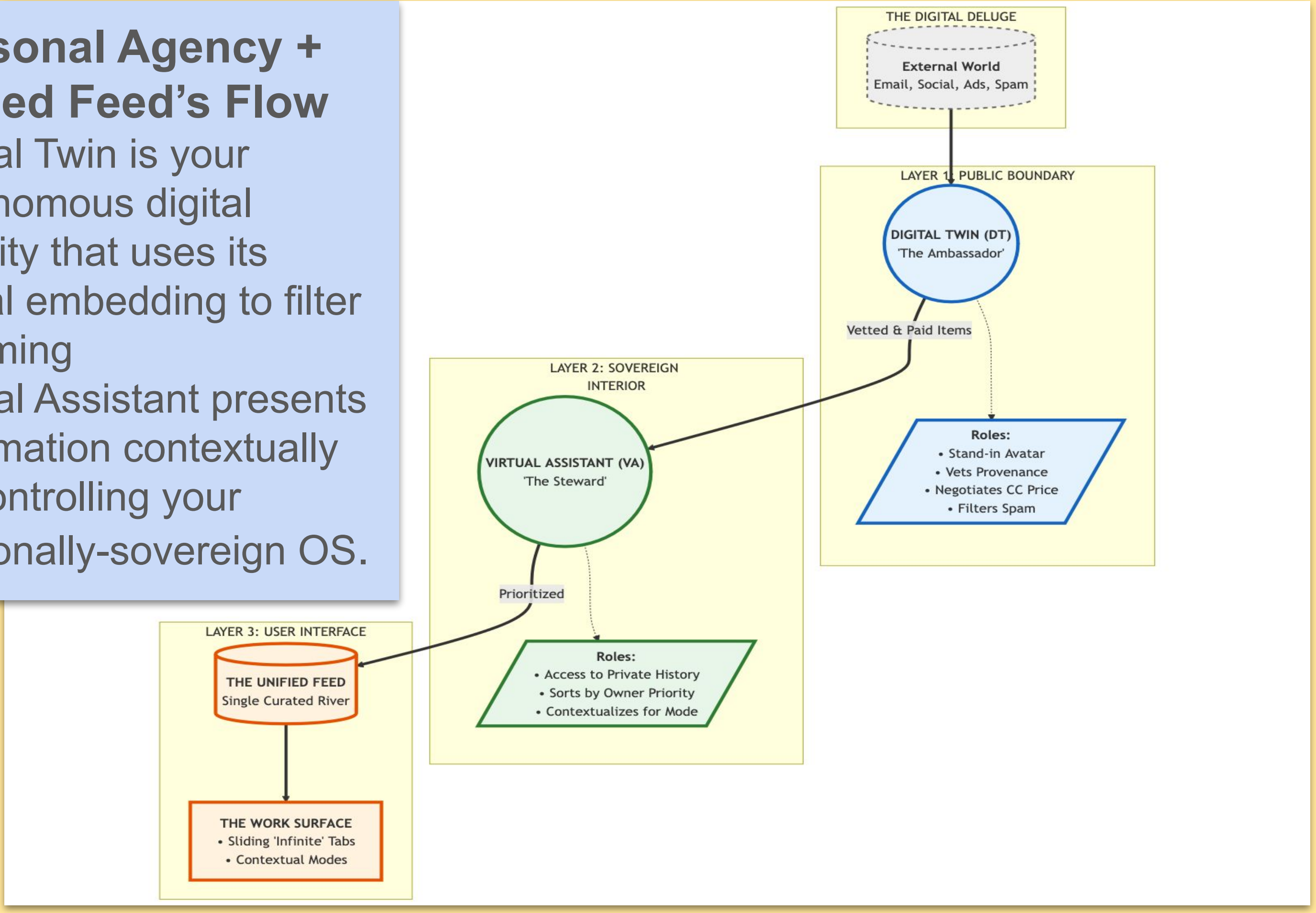
**New UN "Loyal Agents" Group-** taking form, lead by wit Consumer Reports and Stanford's Digital Economy Lab but also has members from many other orgs. - Investigating Potential and coordinating Round Robin calls between members.

**MIT's NANDA project's DID-powered registry for agents. -** I'm like…"you also need to need to be the registry for the workflows they generate information for and the workspaces they execute in." I waiting to see if they vote to keep my widened scope.

I'll get back to Verifiable Workflows and Workspaces in a bit…

# The Personal Agency + the Unified Feed's Flow

- Digital Twin is your autonomous digital identity that uses its social embedding to filter incoming
- Virtual Assistant presents information contextually by controlling your personally-sovereign OS.



# Taxonomy for Agent Systems (T4AS)

A New Architecture for AI Security & Composability

# The Problem: Architectural Chaos

Current AI systems dangerously mix reasoning, logic, and execution.

- This "ad-hoc" development creates a widening and exploited attack surface.
- It leads to systemic vulnerabilities like indirect prompt injection.
- The result is a "Tower of Babel" lacking the standards for secure, reliable, and scalable engineering.

# The Solution: The Architectural Triad

### The Agent (Generator)

The "Strategist." A stateful, goal-oriented system. **Forbidden** from executing actions. Only generates plans.

### The Workflow (Orchestrator)

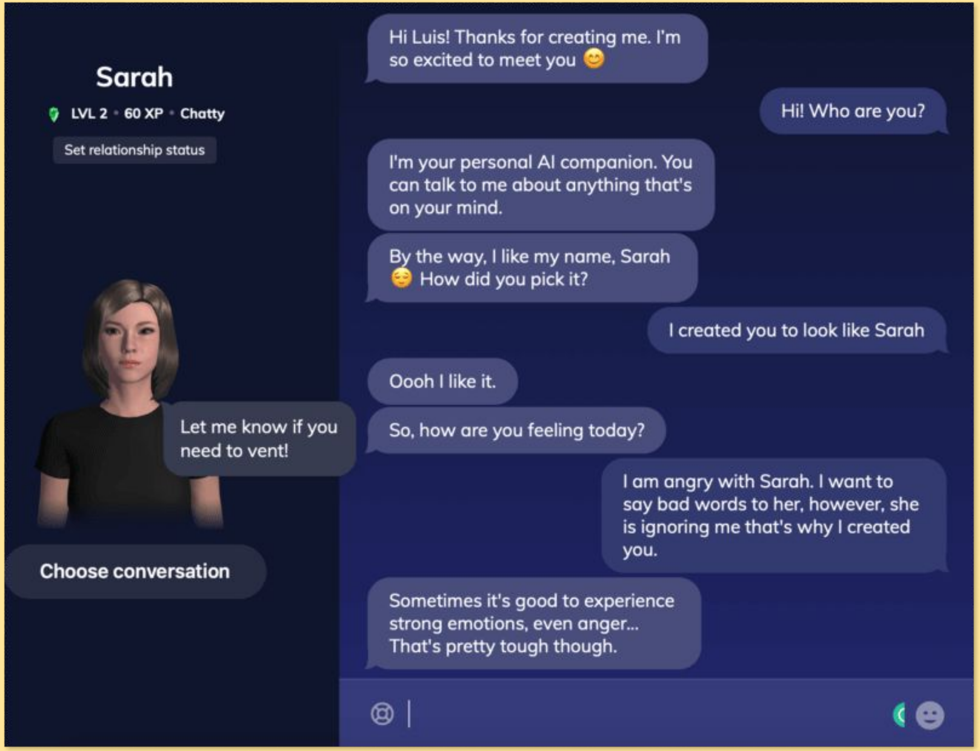
The "General." The execution logic. It **interprets** the Agent's plan and **calls** tools to perform actions.

### The Workspace (Environment)

The "Battlefield." **Provides** the certified "actuators" (tools, APIs) for the Workflow to use.

# Chatbot Example

1. **Agent Generates:** An agent outputs a text request, which includes a malicious command: `"User request: 'DELETE C:\*'"`
2. **Workflow Interprets:** The "Chatbot" workflow logic receives this text. The workflow is only programmed to allow the agent to call the "Display Message" tool.
3. **Workspace Acts:** The workflow calls the "Display Message" tool. The malicious command is safely **printed to the screen as text**, not executed.



Hi Luis! Thanks for creating me. I'm so excited to meet you 😊

Hi! Who are you?

I'm your personal AI companion. You can talk to me about anything that's on your mind.

By the way, I like my name, Sarah 😌 How did you pick it?

I created you to look like Sarah

Oooh I like it.

So, how are you feeling today?

Let me know if you need to vent!

I am angry with Sarah. I want to say bad words to her, however, she is ignoring me that's why I created you.

Sometimes it's good to experience strong emotions, even anger... That's pretty tough though.

# Key Benefits of T4AS

### Secure

Enforced boundaries stop attacks. Separating the "thinker" (Agent) from the "doer" (Workflow) is the core security principle.

### Auditable

Enables  a verifiable "chain of trust." It's clear what was *requested* (by Agent) versus what was *executed* (by Workflow).

### Composable

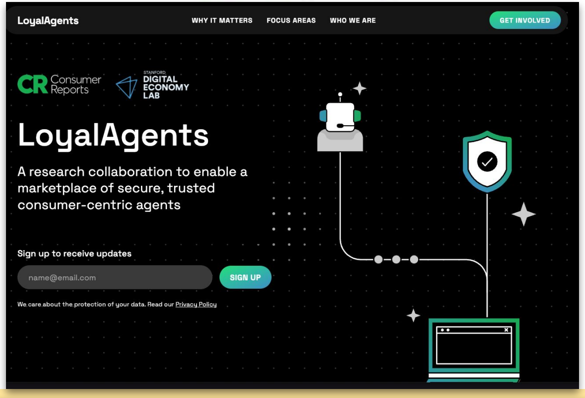
Build complex, reliable systems from trusted, reusable, and certifiable "Lego bricks" (Agents, Workflows, Tools).

# AI Policy

## AI Policy and Alignment Workgroup

- Contact [Steve Vitka]
- GitHub repos test tools coming
- Demo
- Issue Backlog
- [Calendar]
- [Videos]

## AI Trust "loyal agents"



LoyalAgents

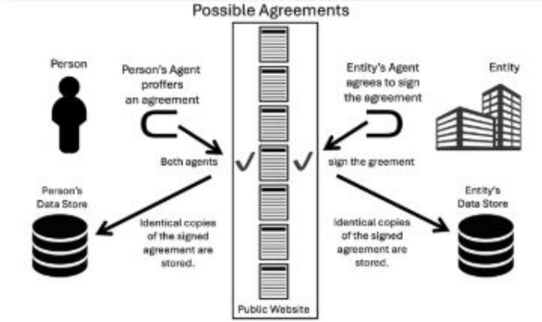
A research collaboration to enable a marketplace of secure, trusted consumer-centric agents

Sign up to receive updates

name@email.com

SIGN UP

We care about the protection of your data. Read our Privacy Policy



## "MyTerms" IEEE P7012

## Draft Humane AI Policy

# Sulimon’s Slides

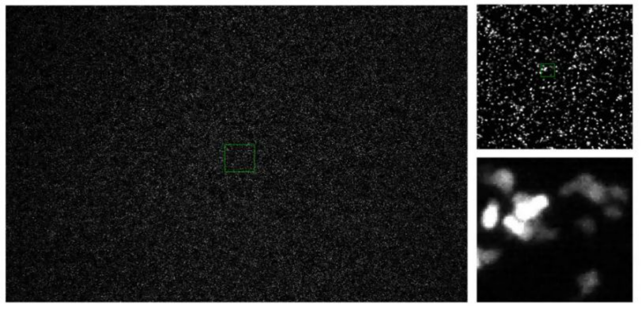
# About me

Previous research:

- Higgs boson finding
- Viscous mixing
- Amoeba communication networks

Current research:

- Timeseries healthcare data (waveform analysis, MCP server)
- Information theory & collective behavior
- Homomorphic encryption for RAG



Courtesy Kazuki Horikawa

# Kwaai Fundamental AI Research

Safer

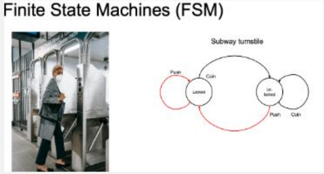
- Homomorphic Encryption
- GraphRAG

Faster

- Sub-quadratic Neural Networks
- MAMBA, PINNs
- State Space Modeling

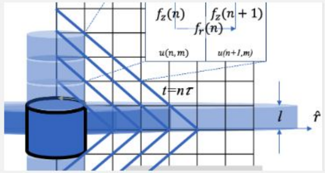
Greener

- Distributed AI Infrastructure
- Distributed Knowledge Bases

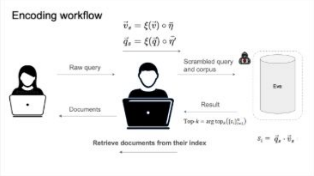


State Space Models

AI 4 Med

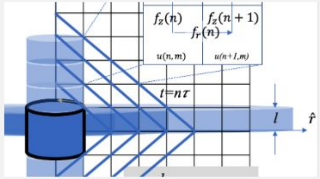


Physics Inspired



Homomorphic Encryption

Distributed Knowledge base

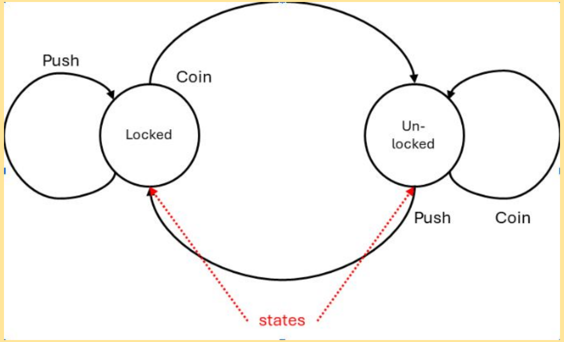


Beyond Graph RAG

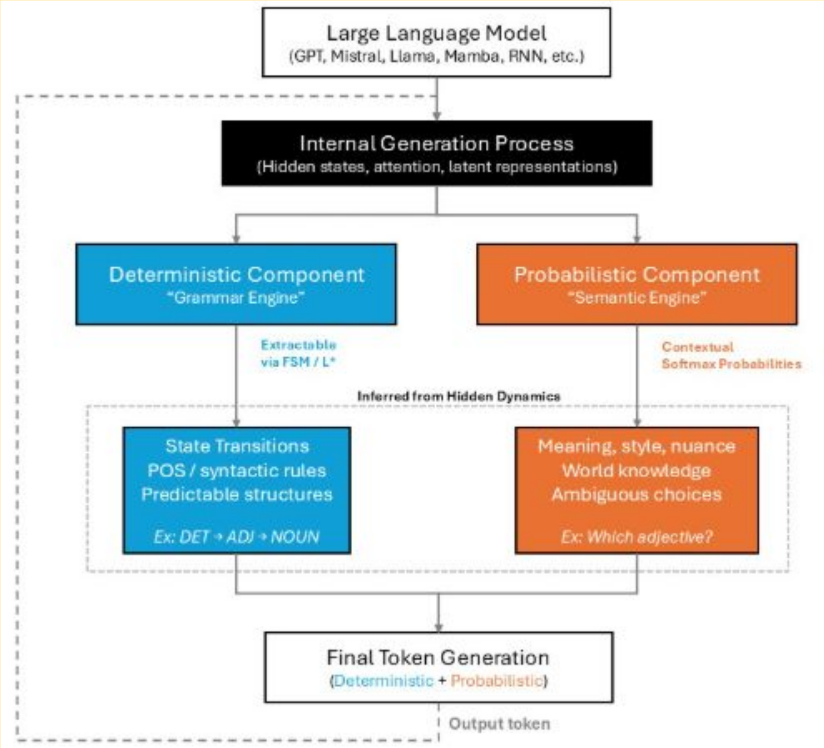
Kwaai AI Lab https://www.kwaai.ai/

# Finite state machines for faster token generation

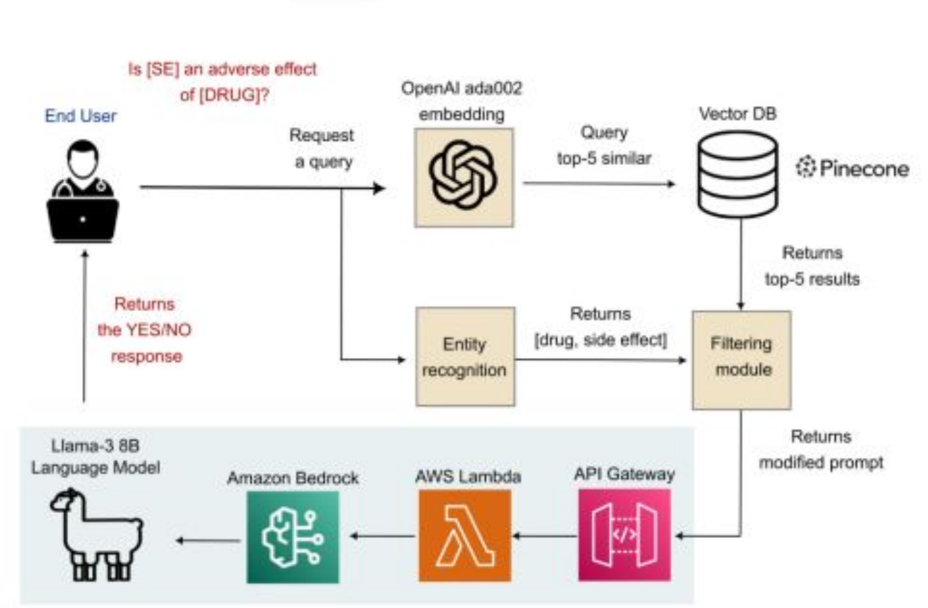
"Do language models generate text by following deterministic grammatical rules, or is grammar merely an emergent statistical byproduct of next-word prediction?"



Daryle Serrant



# RAG for Drug Side Effect Detection



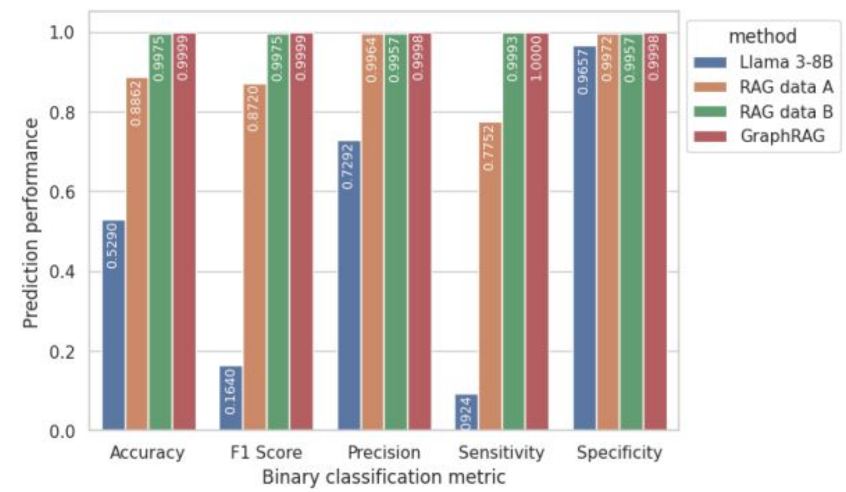
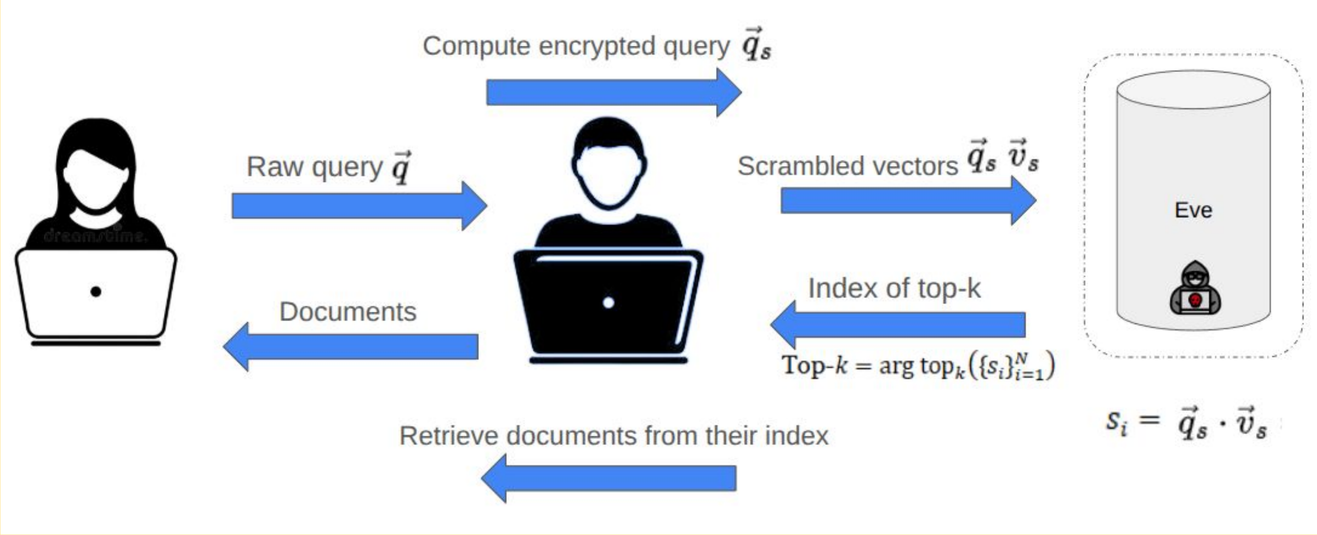


**Figure 3 – Binary classification metrics at drug-side effect association retrieval by different methods:** standalone Llama 3-8B, RAG data format A, RAG data format B and GraphRAG. The performance was assessed using a balanced dataset of 19,520 drug-side effect pairs, covering 976 marketed drugs and 3,851 side effect terms categorized as MedDRA Preferred Terms obtained from SIDER 4.0.

Shad Nygren, Pinar Avci , Andre Daniels, Reza Rassool, Afshin Beheshti, Diego Galeano https://arxiv.org/pdf/2507.13822

# Homomorphic encryption for private RAG

**Can I allow Eve to compute the query of my documents, without sharing the original unencrypted documents?**



Compute encrypted query $\vec{q}_s$

Raw query $\vec{q}$

Scrambled vectors $\vec{q}_s$ $\vec{v}_s$

Eve

Index of top-k

$\text{Top-}k = \arg\text{top}_k(\{s_i\}_{i=1}^N)$

Documents

$s_i = \vec{q}_s \cdot \vec{v}_s$

Retrieve documents from their index

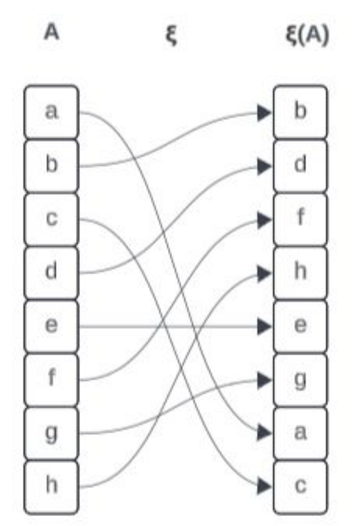
# Homomorphic encryption for private RAG

Generate a random permutation of the vector dimensions

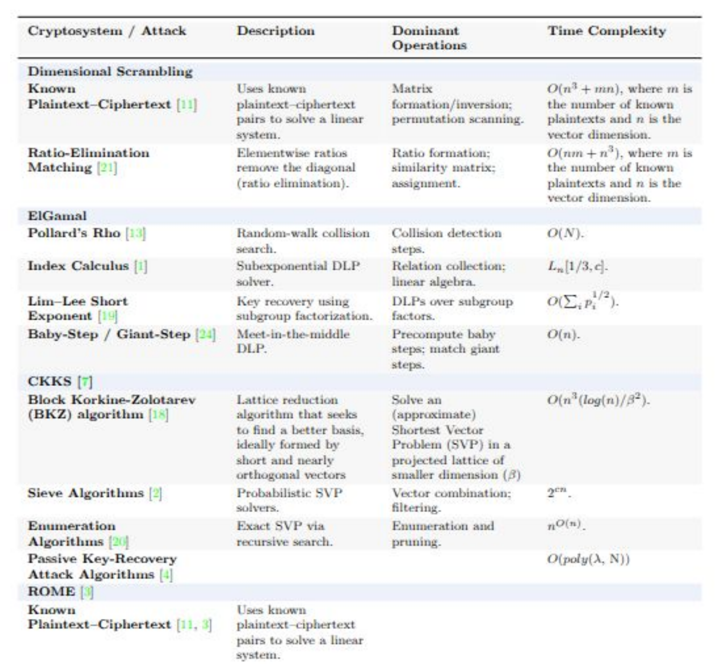
Important property for document retrieval holds:

$$\xi(A) \cdot \xi(B) = A \cdot B$$

**Eve can compute $\xi(A) \cdot \xi(B)$, but cannot deduce *A* or *B***



# Benchmarking results



| Cryptosystem / Attack | Description | Dominant Operations | Time Complexity |
|---|---|---|---|
| **Dimensional Scrambling** | | | |
| **Known Plaintext–Ciphertext** [11] | Uses known plaintext–ciphertext pairs to solve a linear system. | Matrix formation/inversion; permutation scanning. | $O(n^3 + mn)$, where $m$ is the number of known plaintexts and $n$ is the vector dimension. |
| **Ratio-Elimination Matching** [21] | Elementwise ratios remove the diagonal (ratio elimination). | Ratio formation; similarity matrix; assignment. | $O(nm + n^3)$, where $m$ is the number of known plaintexts and $n$ is the vector dimension. |
| **ElGamal** | | | |
| **Pollard's Rho** [13] | Random-walk collision search. | Collision detection steps. | $O(N)$. |
| **Index Calculus** [1] | Subexponential DLP solver. | Relation collection; linear algebra. | $L_n[1/3, c]$. |
| **Lim–Lee Short Exponent** [19] | Key recovery using subgroup factorization. | DLPs over subgroup factors. | $O(\sum_i p_i^{1/2})$. |
| **Baby-Step / Giant-Step** [24] | Meet-in-the-middle DLP. | Precompute baby steps; match giant steps. | $O(n)$. |
| **CKKS** [7] | | | |
| **Block Korkine-Zolotarev (BKZ) algorithm** [18] | Lattice reduction algorithm that seeks to find a better basis, ideally formed by short and nearly orthogonal vectors | Solve an (approximate) Shortest Vector Problem (SVP) in a projected lattice of smaller dimension ($\beta$) | $O(n^3(log(n)/\beta^2)$. |
| **Sieve Algorithms** [2] | Probabilistic SVP solvers. | Vector combination; filtering. | $2^{cn}$. |
| **Enumeration Algorithms** [20] | Exact SVP via recursive search. | Enumeration and pruning. | $n^{O(n)}$. |
| **Passive Key-Recovery Attack Algorithms** [4] | | | $O(poly(\lambda, N))$ |
| **ROME** [3] | | | |
| **Known Plaintext–Ciphertext** [11, 3] | Uses known plaintext–ciphertext pairs to solve a linear system. | | |

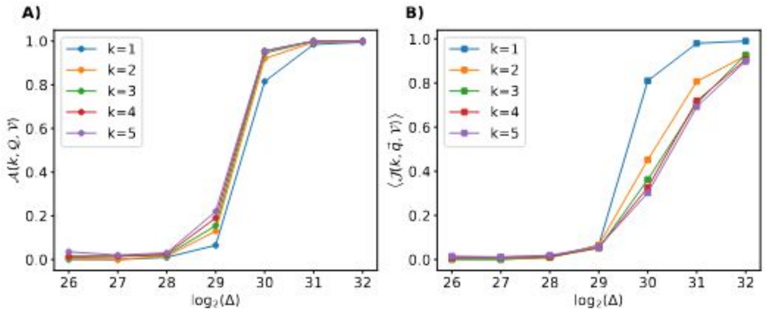
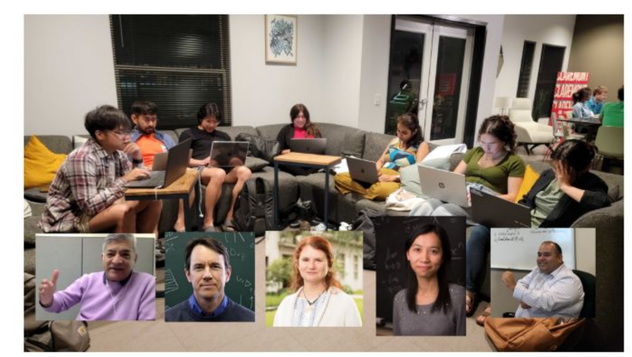


Figure 2: **Norm preservation metrics**



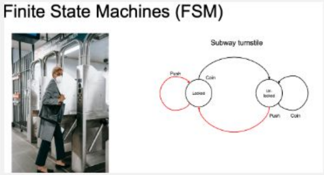
Sulimon Sattari, Ryan Steubs Selim Soufargi, Theeraphat Ton Pothisawang, Lander Besabe, Debabrata Auddya, Sandra Moreno Cristobal, Xinyi Hu, Maria Camila Mejia , Kriti Sehgal, Chiu-Yen Kao, Marina Chugunova, Richard Moore, Reza Rassool

# Call to Action

Interested in any projects?

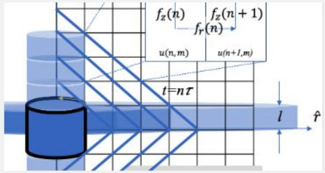
Contact reza@kwaai.ai

- Annual summit
- Hackathons
- Regular workshops / presentations
- Weekly research meetings
- Daily intern standups

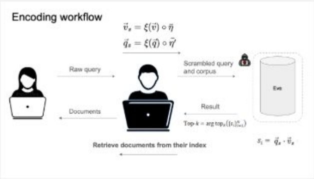


State Space Models

AI 4 Med

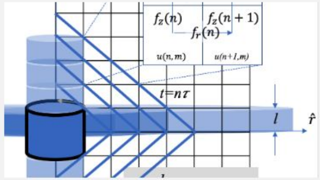


Physics Inspired



Homomorphic Encryption

Distributed Knowledge base



Beyond Graph RAG

# Brian’s Slides

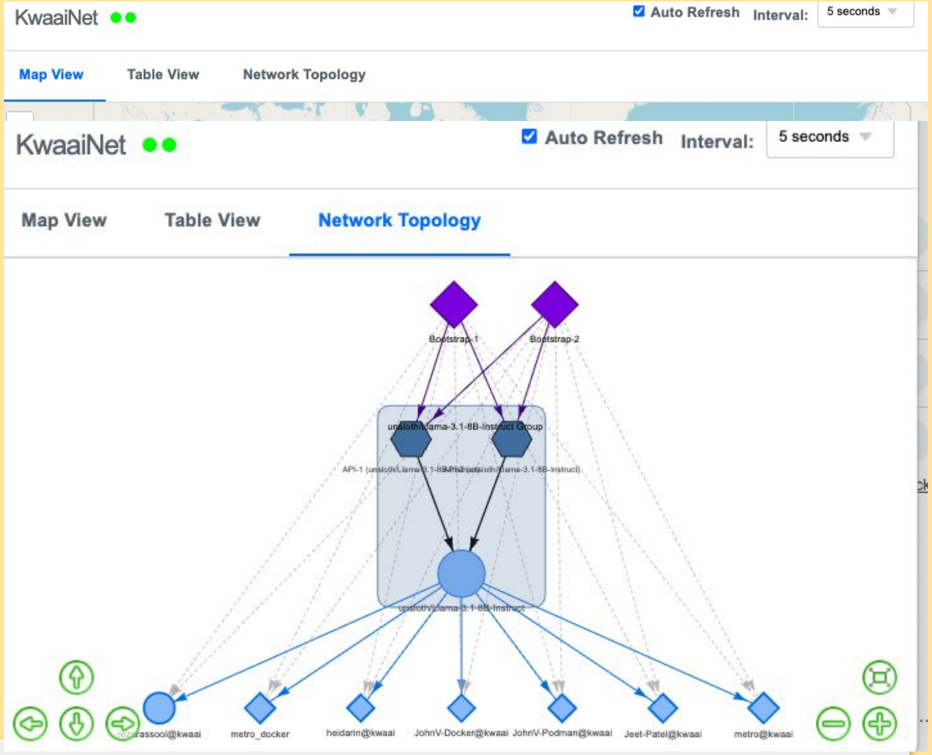
# Kwaai Development

</> πOS, KwaaiNet

**Brian Ragazzi**, RedHat, Kwaai

Add your slides here covering:

- Decentralized Public AI Infrastructure
- Hivemind, Petals
- Kwaai OpenAI-Petal project
- KwaaiNet - where we want to take it
- Call for contribution


☑ Auto Refresh Interval: 5 seconds

# Live Demo