Introduction to Agentic AI Workshop

Presented by
the AI Safety Awareness Project

Workshop Agenda

01

Introduction to Agentic AI

Generative AI vs. Agentic AI
Use Case Example Walkthrough

02

Agentic AI Architectures

Reflection, Tool Use,
Planning, MultiAgent
Patterns + Examples

03

Hands-On Exercise

Breakout room
opportunities – Test
MultiAgent Systems

04

Wrap-Up

Final remarks; Where
Agentic AI takes policy

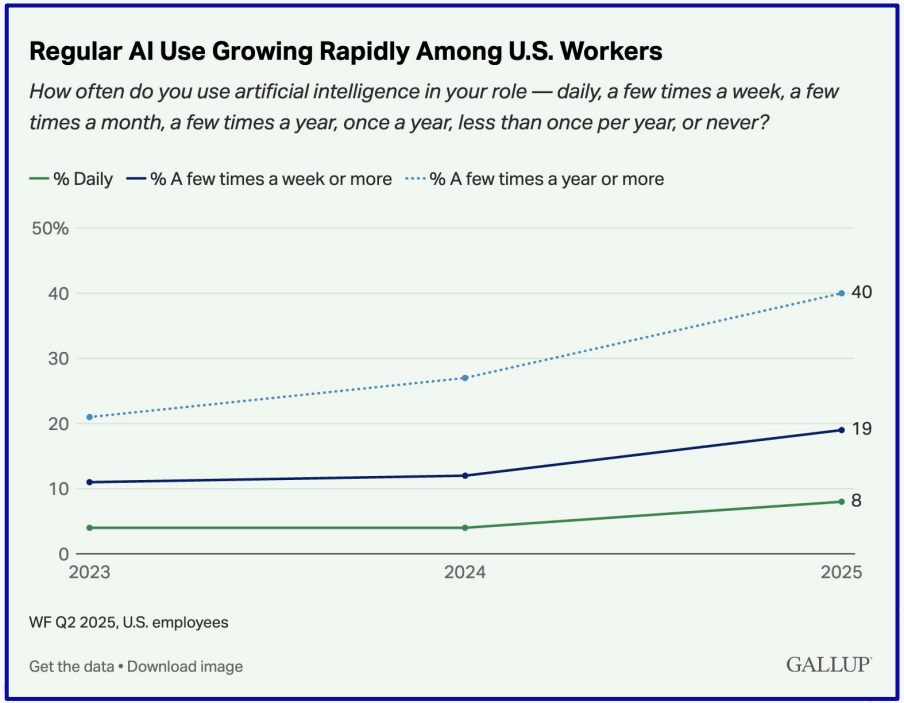
01 Introduction to Agentic AI

What is Agentic AI?

Ask audience

1. Knowledge base of AI?
    a. What do you know about AI?
    b. Do you use any AI tools?
2. What do you think about AI?
    a. “I think it’s groundbreaking – and I use it on a regular basis”
    b. “I think it has potential like the internet – it can be used for good or bad”
    c. “I think it’s moving too fast – it can be useful but the bad outweighs the good”
3. How often do you use AI?
    a. Everyday
    b. A few times a week
    c. Only when I need to
    d. Never

AI Usage Has Doubled Among US Workers

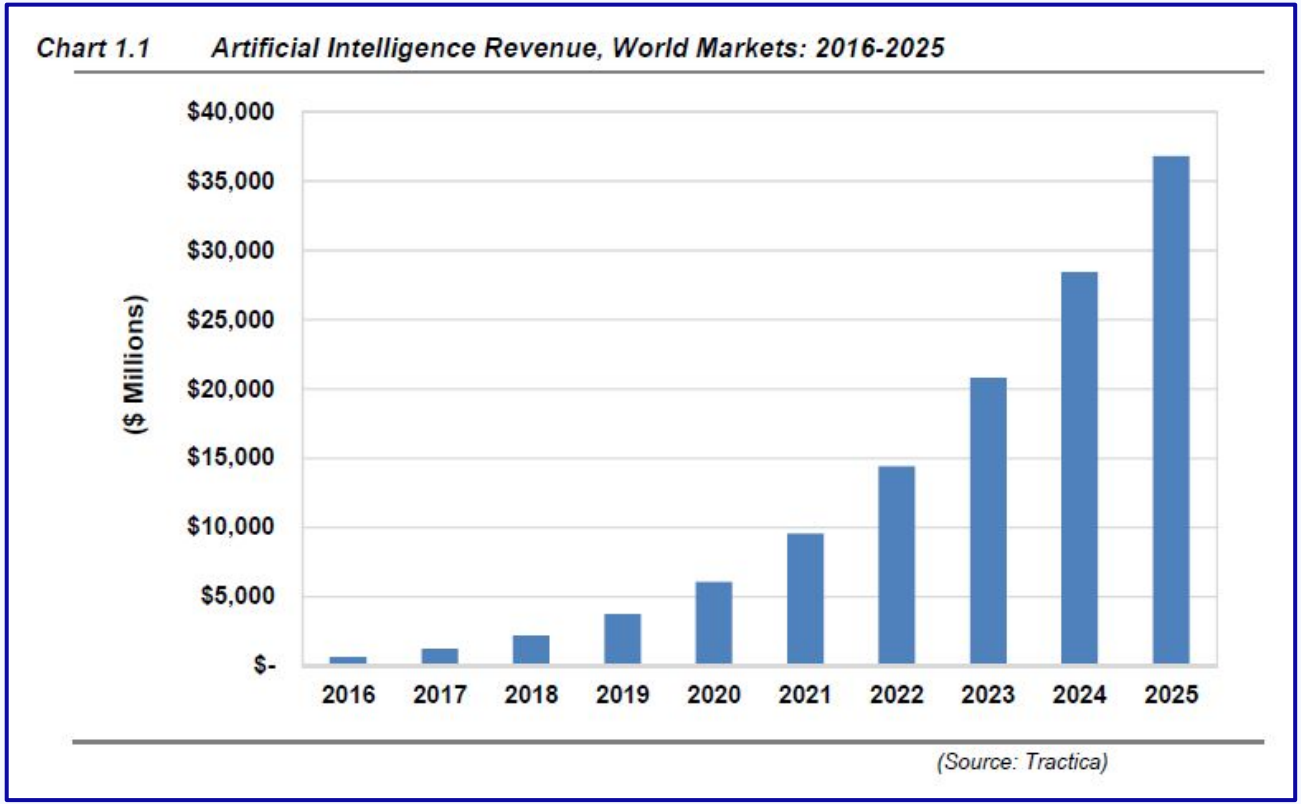


In just two years

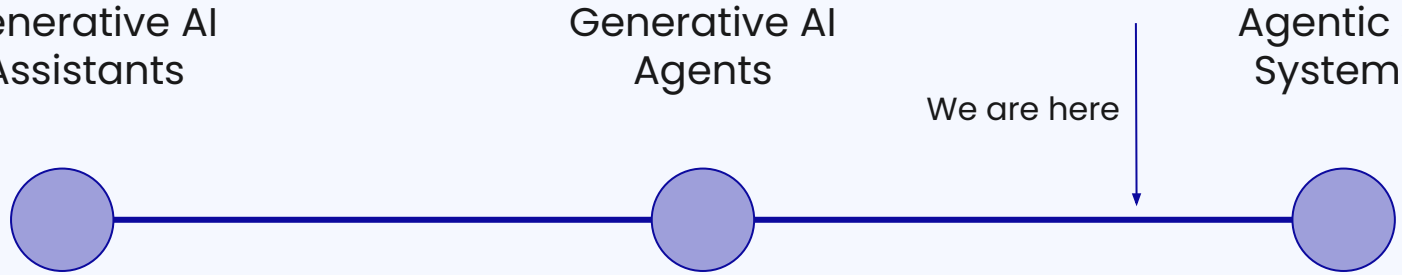
39%

of U.S. adults have already used AI. In comparison, the internet reached just 20% adoption in its first two years.

Markets are investing too!



Evolving into Agentic AI



- Follows particular set of rules
- Good for automated work

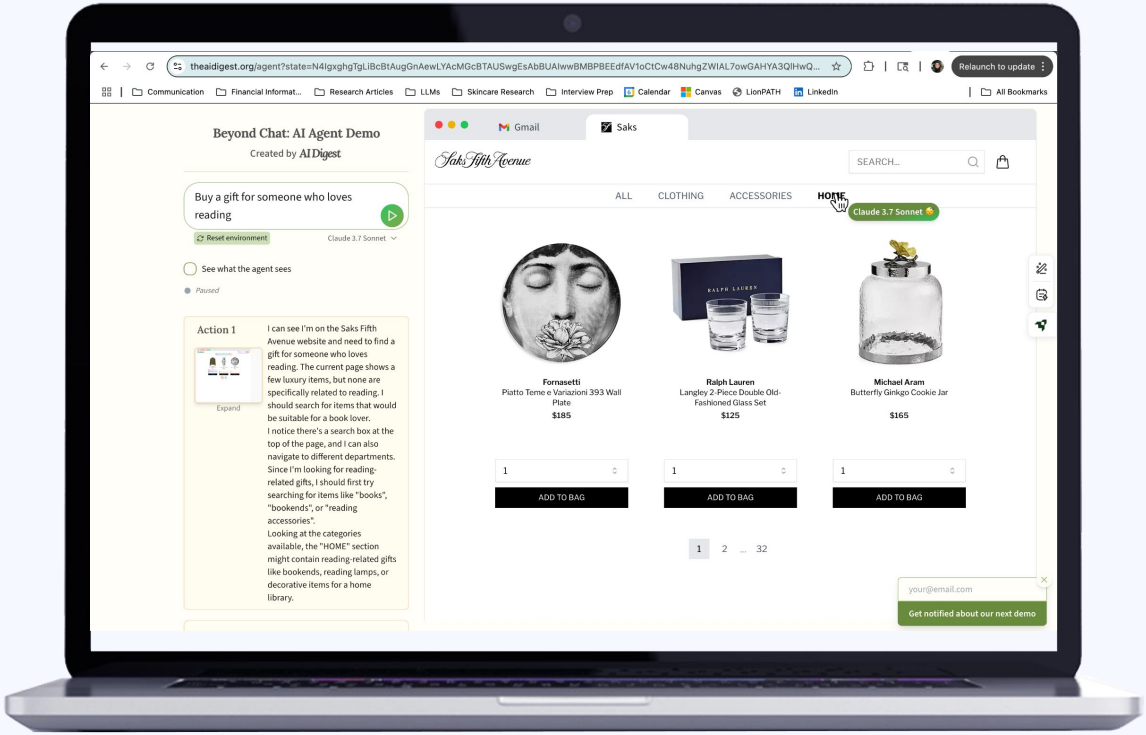
- Focuses on content creation
- Achieves a particular goal

- Fully autonomous
- Mimics human action and reasoning

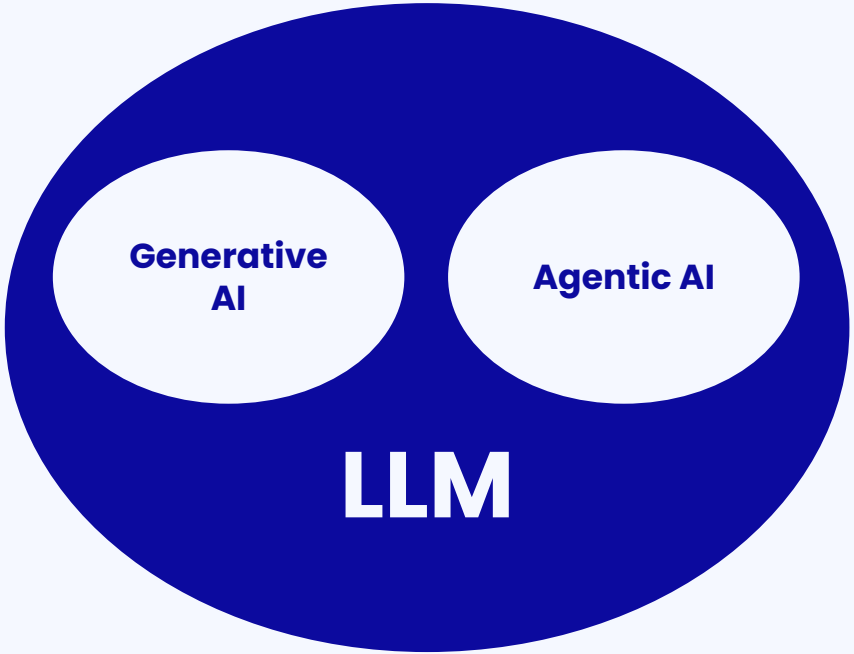
More Oversight

Less Oversight

Agentic AI: DEMO



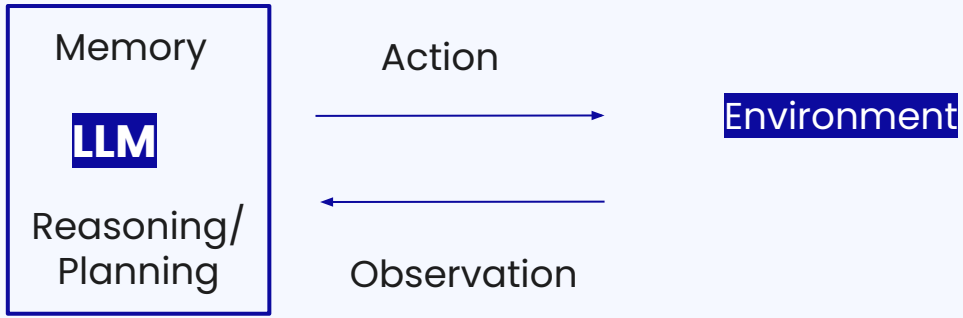
Where Does Agentic AI Fit?



Agentic Language Models

They interact with their environments

- Baseline LLM involves text in → text out
- Agentic AI adds a layer perceives environment and maintains a belief state



Agentic AI Capabilities

- Considered to be a progression of LM usage
  - Has the ability to reason and act

1. Can reason and act – breakdown complex tasks and actions
2. Execute actions – use external code and tools to retrieve information
3. Interact with environment using tools
4. Learn from feedback

AI System: Use Case

Question: "Can I get a full refund for my United Airlines ticket because of maintenance issues?"

Generative AI

- **Mindset:** Predict next best sentence from what's already known (very pattern-recognition oriented)
- **Data:** Has a "knowledge cutoff" date
- **Language:** maybe, perhaps, not definitive
- **Style:** Usually single slide of information; follow ups not common
- **Example:** "You are likely eligible – call United Airlines"

Agentic AI

- **Mindset:** Figures out the correct answer
- **Data:** Will go on Department of Transportation site, other updated sources
- **Language:** Asks follow up questions
- **Style:** Can use web search, possibly offer to set up reminder
- **Example:** "Because maintenance is controllable and flight has been delayed for more than 3 hours, you will get a refund of $200"

Agentic Language Models: Programming Use Case

Vibe Coding Today

- **Human:** Main director and editor of code
- **AI Role:** Has oversight, looks, shapes, and edits coding mistakes
- **Use cases:** experimentation, debugging, provide basic frameworks
- **Planning:** Very much man-controlled; user still needs programming fundamentals knowledge

Agentic AI Vibe Coding

- **Human and AI Roles:** Pair Programming
- **Use cases:** Utilizes more multi-step workflows, multisystem codebases
- **Planning:** Defines execution goals, and tools needed

Also invites for next steps!

02

Agentic AI Patterns

Discover Flow Patterns

Questions?

03

Hands-On Projects

Break-out opportunities

Project: Multi-Agent Court Case Scenario (30 minutes)

Objective: In this project, you will gain the opportunity to explore different LLM agents (Judge, Jury, Witness, Prosecution, and Defense) interacting with each other to defend and support popular court cases.

Directions:

1. Breakout into groups (up to 3) and run through each scenario provided in the Google Colab
2. Read each case carefully, note any reasoning patterns you notice – or anything you did/did not expect
3. Answer the following questions at the bottom of the docs
4. Time permitting: Brainstorm your own scenario and see what the agents output
5. Come back to larger team and open discussion

Some Takeaways?

1. Which of the three cases surprised you the most? What was most expected?

2. How close do you believe this mimics human thinking patterns? Is it getting closer?

3. What single additional fact (a document, a sensor log, a policy memo) would have flipped your vote in any one case—and why?

4. Would you rely on Multi-Agent AI to be just and fair in court cases? Where is the line drawn if so?

The AGI: Are Courts Feeling It?

- Sanctions and withdrawn opinions are multiplying
  - A Mississippi federal judge disqualified three Butler Snow lawyers for un-vetted ChatGPT cites, calling their conduct “bad faith.”
  - Federal judges in New Jersey and Mississippi withdrew rulings once lawyers showed the orders relied on filings riddled with AI-fabricated quotes
- Courts are writing a rule book – right now
  - At least 100 U.S. judges now require parties to disclose any generative-AI help and to certify manual cite checks. A Bloomberg Law comparison table tracks the orders.
  - California’s Judicial Council previewed a model policy that local courts can adopt or tweak.
- But … it’s still moving!
  - Johnson v. Dunn (Ala. N.D. 2025) – Defense team relied on an internal research agent that hallucinated precedent.
  - The U.K. Ministry of Justice is rolling out an agent that parses millions of inmate phone messages to predict and prevent prison violence, a live deployment inside England & Wales prisons.
  - GOV.UK
  - U.S. state courts use National Center for State Courts (NCSC) templates to stand-up AI triage bots that guide self-represented litigants through filings and fee waivers

Summary

So far, we have covered...

- Evolution of Artificial Intelligence
- Relevance in markets, industries, personal lives
- Role of Agentic AI
- Exploring technical AI architectures (Reflection, Tool Use, Planning, and Multi-Agent Collaboration)
- Testing live agents and capabilities

04

Next Steps

Discuss what can be done to expect as users in the future of Agentic AI

Where is AI going? And where does policy come into play?

- There is rapid technical development of Artificial Intelligence, so where does this take policy?
  - Concerns about potential misuse and unintended consequences of AI, prompt efforts to develop standards
- In 2024 legislation, **45** US States have introduced AI bills, and **31** have adopted or enacted laws
  - Most include:
    - Colorado: Enacted effective legislation to requiring developers and deployers of high-risk AI systems to avoid algorithmic discrimination
    - Maryland: Required the Department of Information Technology to adopt policies regarding development and deployment
    - Pennsylvania: Passed an AI Content Disclosure Bill (HB 1598) requiring the disclosure of AI generated content for consumer products and services

Where have we gone so far?

Pros:

- Introduces accountability in the conversation – think about AI Executive Orders from the US and AI EU Act
- Creates pressure for transparency
  - Think Red-Teaming across AI companies
  - Model cards and system disclosures, audits of training data,
- Encourage preemptive safety tools
  - Interruptibility mechanisms
  - Bias + autonomy audits

Cons:

- Still not on par with technology advancement
- Vague terminology with little follow-up or enforcement
  - In the United States, most policies are voluntary or non-binding
- No address on Multi-agent behavior yet
  - Current laws treat AI like a single system

How can you get involved?

- Join our mailing list:
  aisafetyawarenessfoundation.org
- Let your local legislators know that AI safety is an important issue
  house.gov/representatives/find-your-representative
- Self study
  aisafety.info
- AI Safety Education and career transitions:
  BlueDot Impact
- Reading
  AI 2027 - Daniel Kokotajlo

See our page for more extensive resources!

Workshop Survey