

The background is an abstract geometric pattern composed of numerous triangles of varying sizes. The color palette transitions from warm oranges and yellows on the left side to cool blues and greys on the right side, with a soft gradient in the center.

Agents Are All You Need (An Introduction to Agentic AI)

Presenter: Dr Peter Leong

www.linkedin.com/in/peterleong

Self-Introduction

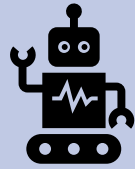
- Dr Peter Leong
(<https://www.linkedin.com/in/peterleong>)
Transforming Ideas into Innovations
- Lead Specialist (AI & Analytics)
Singapore Polytechnic
- Course Chair – Specialist Diploma in Data Science
(Artificial Intelligence)
[https://www.sp.edu.sg/pace/courses/all-courses/course-details/specialist-diploma-in-data-science-\(artificial-intelligence\)](https://www.sp.edu.sg/pace/courses/all-courses/course-details/specialist-diploma-in-data-science-(artificial-intelligence))



Introduction

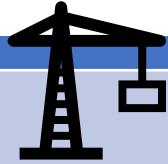
- **This presentation will explore the world of AI agents, their applications, and their future implications.**

Presentation overview



Part I

- Understanding AI Agents
- The Role of AI Agents in Modern Applications



Part II

- Ethical Considerations
- Societal Impact



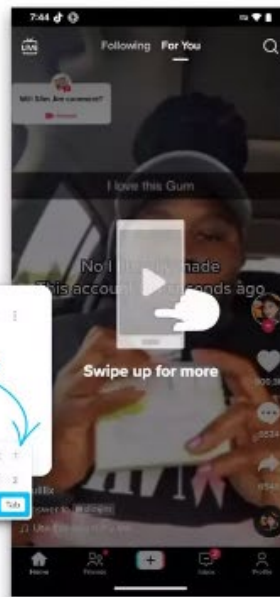
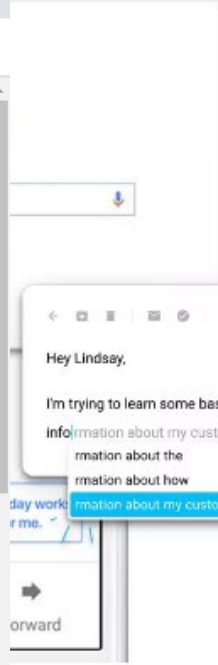
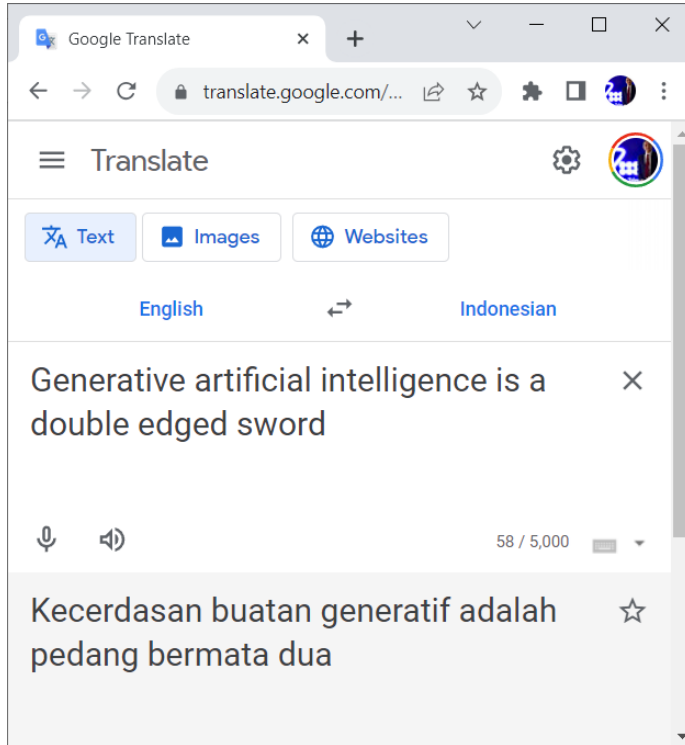
Part III

- Future of AI Agents
- Conclusions

Understanding AI Agents

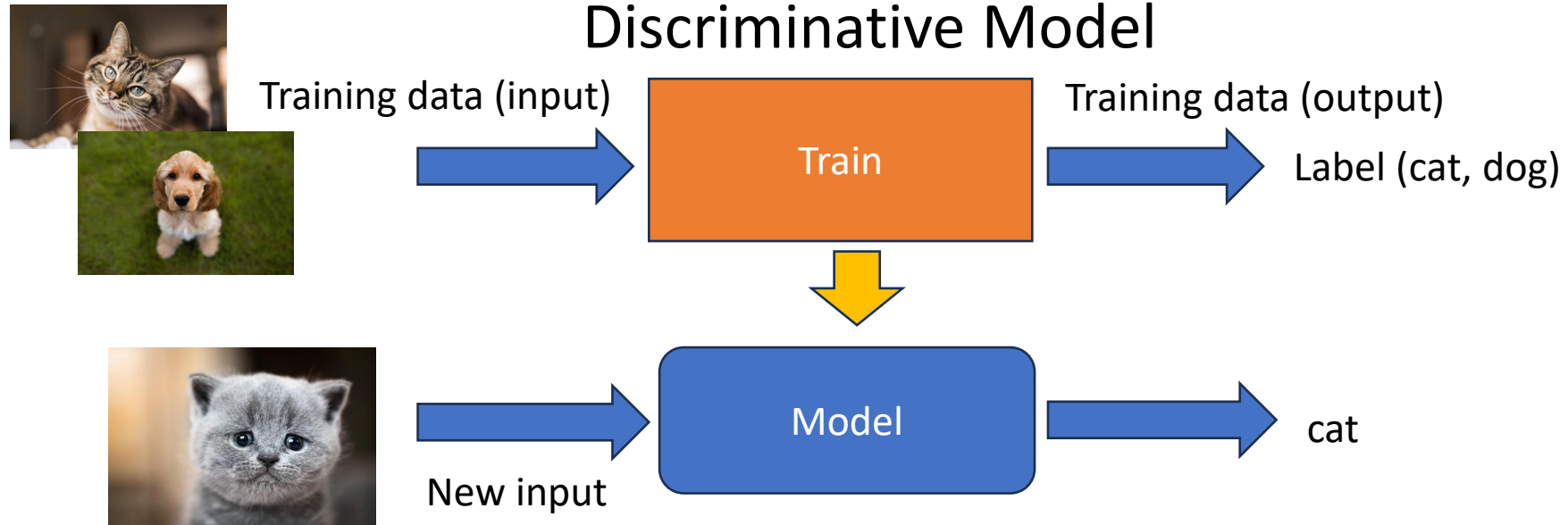
(First — a prelude to AI Agents)

AI has been with us for years, whether “generative” or “agentic”

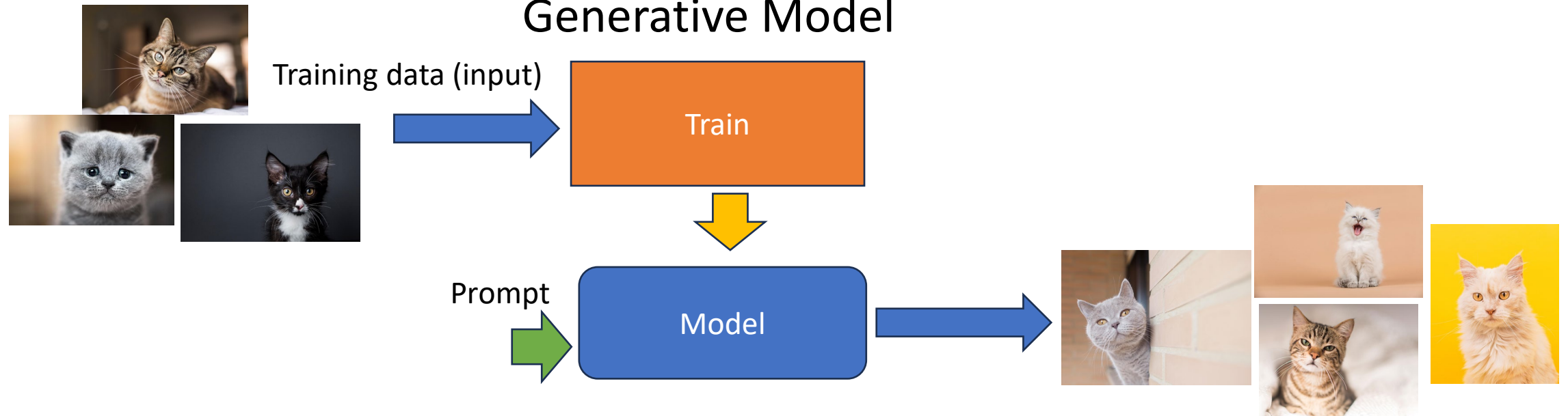


Can you think of more AI examples in our daily lives?

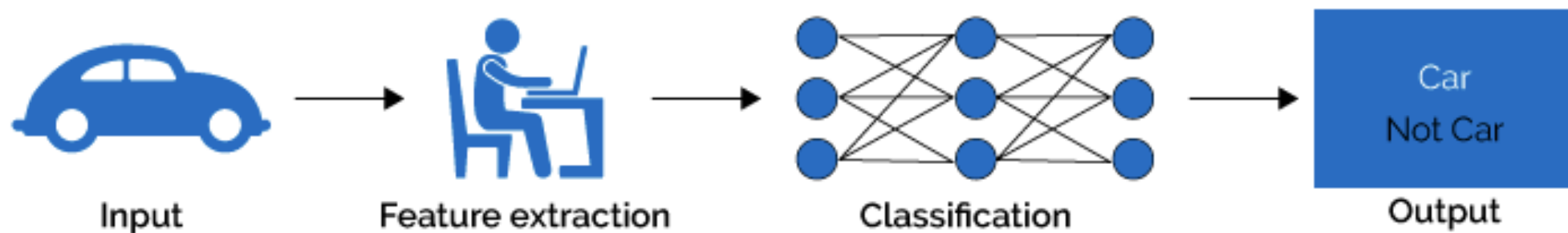
Discriminative Model



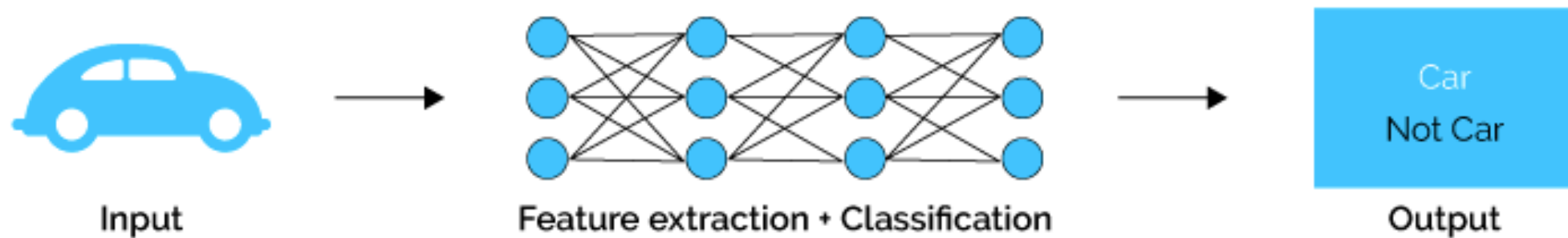
Generative Model



Machine Learning

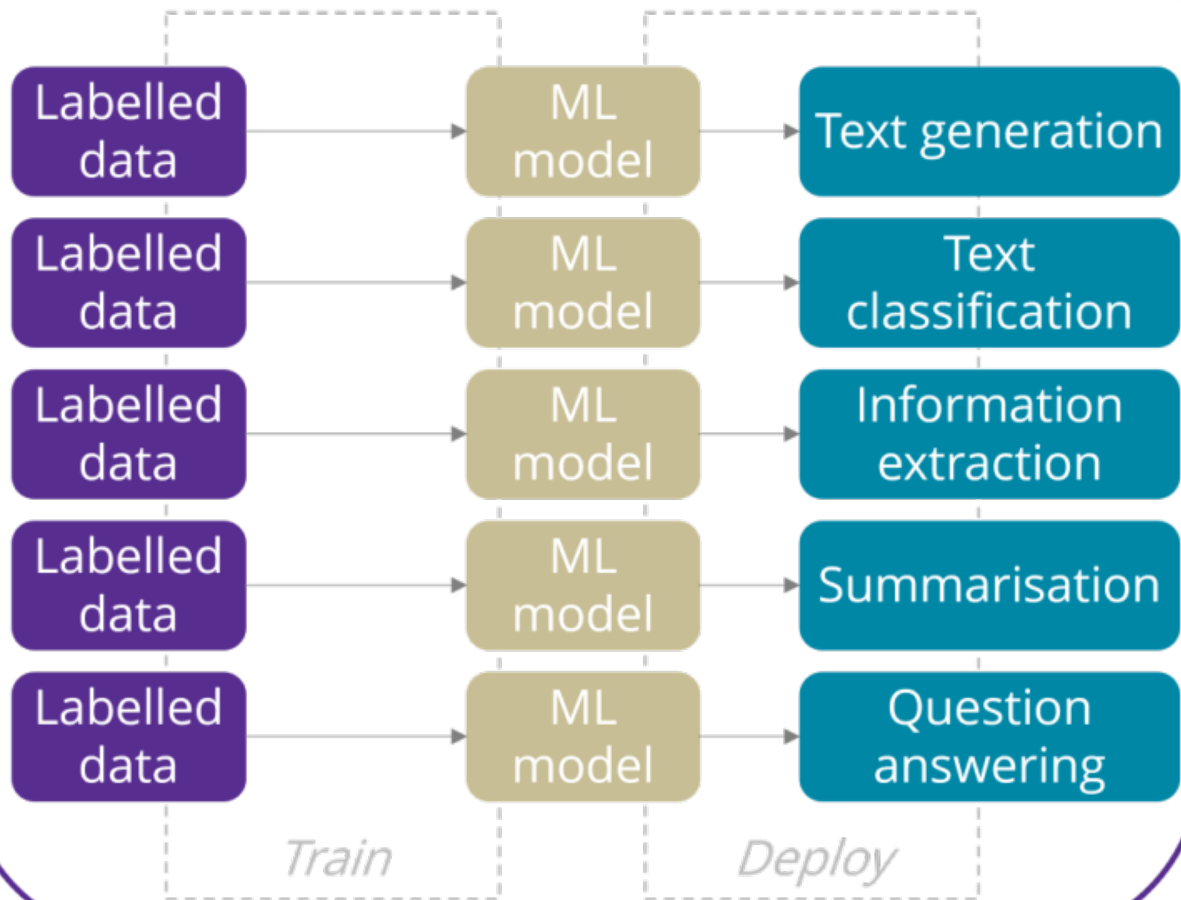


Deep Learning

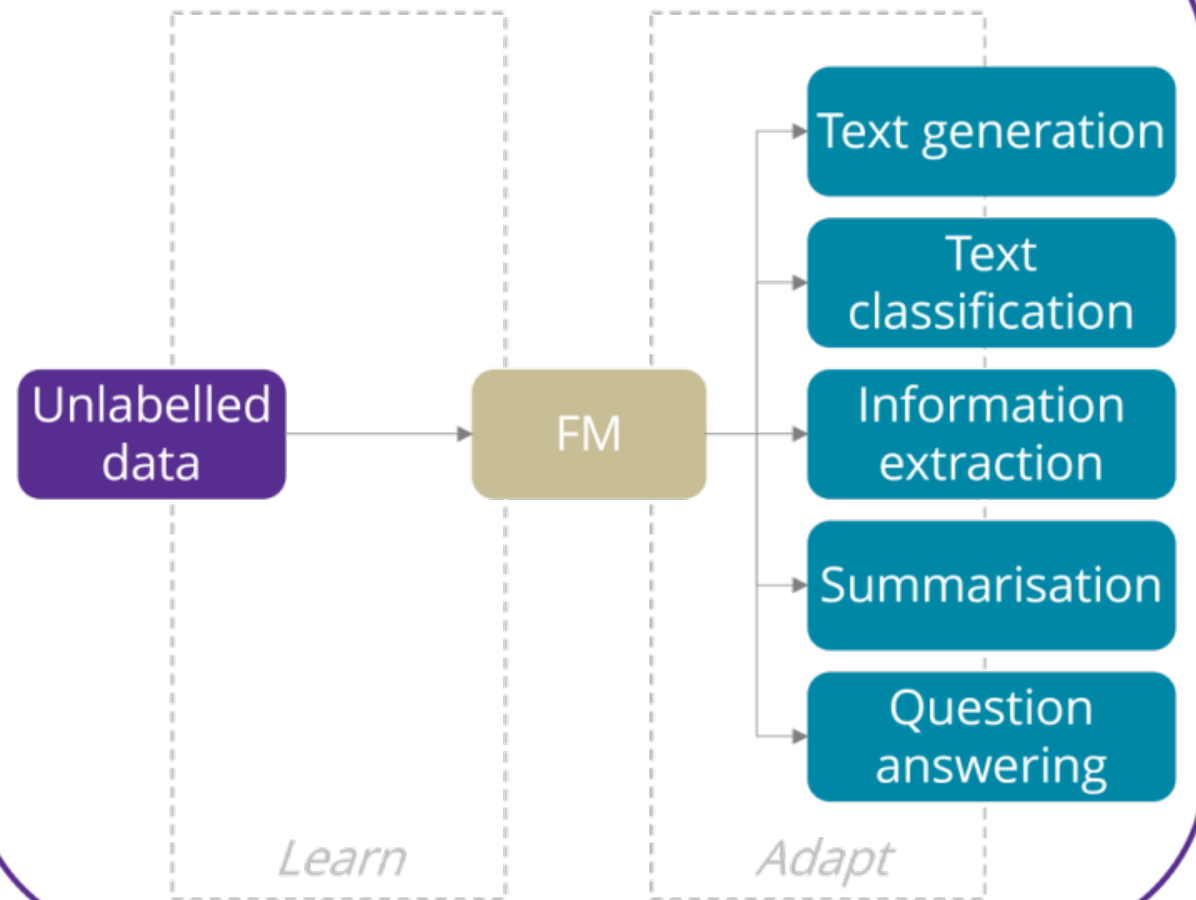


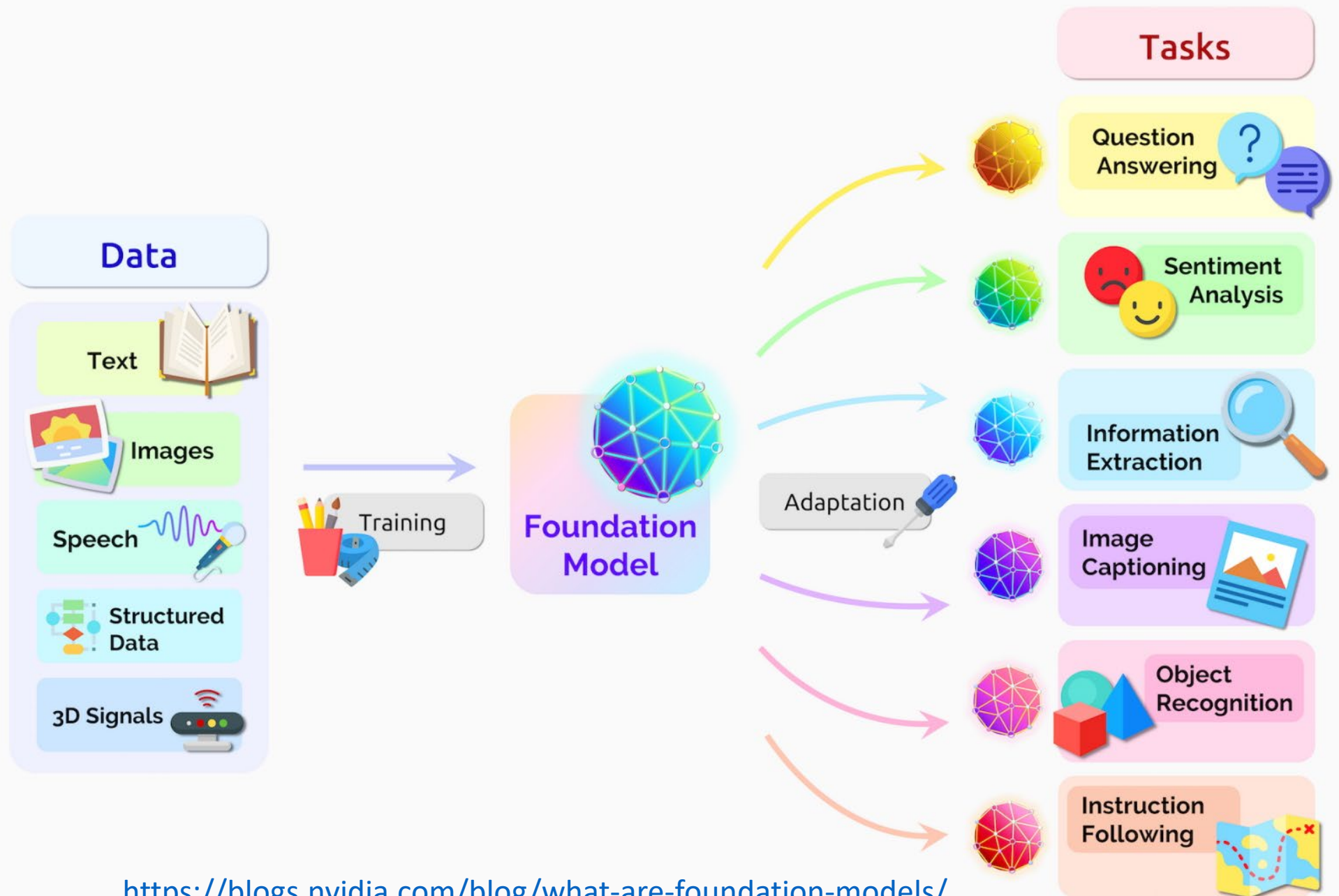
How Foundation Models Differ from other ML models

Traditional ML models



Foundation models





<https://blogs.nvidia.com/blog/what-are-foundation-models/>

Foundation Models

```
graph TD; A[Foundation Models] --> B[Language Models<br/>(BERT, GPT3, T5)]; A --> C[Computer Vision Models<br/>(ResNet, EfficientNet, YOLO)]; A --> D[Generative Models<br/>(DALL-E, GANs, VAEs)];
```

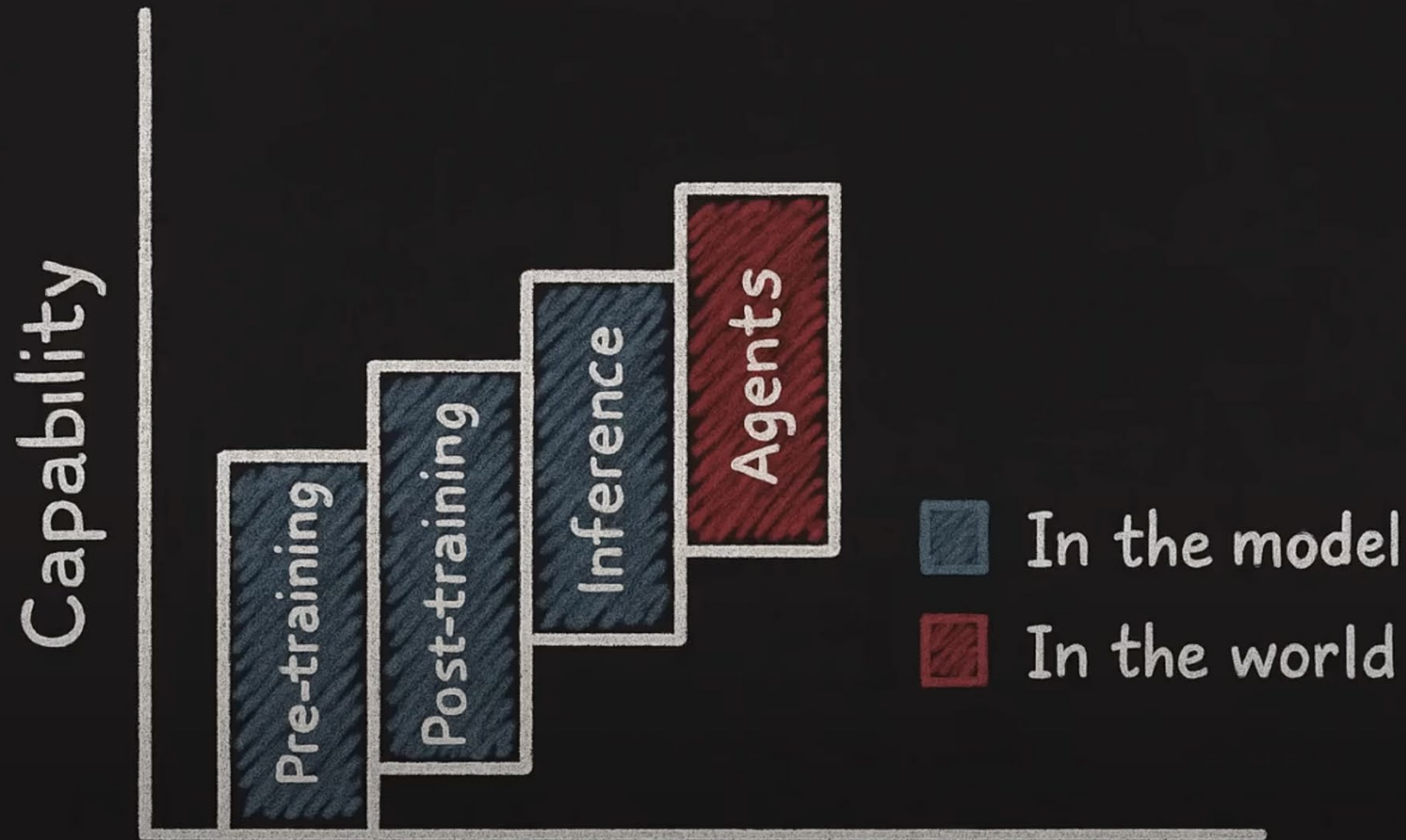
Language Models
(BERT, GPT3, T5)

Computer Vision
Models
(ResNet, EfficientNet,
YOLO)

Generative Models
(DALL-E, GANs,
VAEs)

DECISIONS

New training methods are rapidly pushing the AI frontier



Source: May 15, 2025 TED Talk <https://www.youtube.com/watch?v=id4YRO7G0wE>

Understanding AI Agents

What are agents?

What are AI agents?

Definition: AI agents are autonomous systems that perceive and act to achieve goals.

Key Characteristics:

- Interact with dynamic environments
- Make decisions based on goals

Comparison: Unlike static AI models, agents are adaptive and goal-oriented.

Source: [IBM: What Are AI Agents?](https://www.ibm.com/think/topics/ai-agents)

ibm.com

AI ▾ Hybrid Cloud ▾ Products ▾ Consulting Support ▾ Think 2025

Q Chat User

Think Think 2025 ▾ Artificial intelligence Cloud Security Videos ▾ Reports ▾ Podcasts ▾ Events ▾ More ▾

Subscribe

AI Agents

Welcome

▾ Introduction

Overview

AI agents vs AI assistants

Agentic AI

Agentic AI vs generative AI

Types of AI agents

▸ Components

▸ Architecture

▸ Frameworks

▸ Governance

▸ Agentic RAG

Anna Gutowska

AI Engineer, Developer Advocate

IBM

What are AI agents?

An [artificial intelligence \(AI\)](#) agent refers to a system or program that is capable of autonomously performing tasks on behalf of a user or another system by designing its [workflow](#) and utilizing available [tools](#).

AI agents can encompass a wide range of functionalities beyond natural language processing including decision-making, problem-solving, interacting with external environments and executing actions.

Report

Top Strategic Technology Trends for 2025: Agentic AI

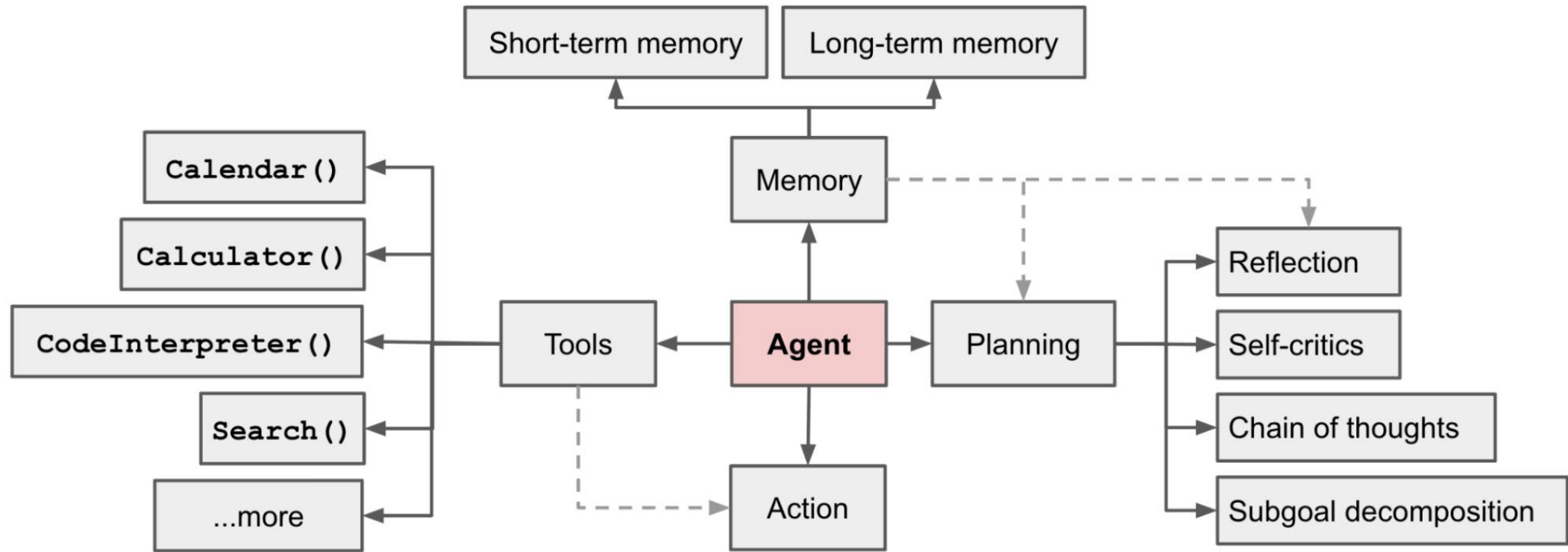
Download this Gartner research to learn the potential opportunities and risks of agentic AI for IT leaders and learn how to prepare for this next wave of AI innovation.

[Read the report](#)

Types of AI Agents

- **Reactive Agents:** Respond to current stimuli without memory or planning.
 - Example: Simple chatbot with predefined responses.
- **Proactive Agents:** Anticipate and plan based on goals and predictions.
 - Example: Autonomous vehicle planning routes.
- **Hybrid Agents:** Combine reactive and proactive behaviors, often with learning.
- **Sources:**
 - [GeeksforGeeks: Agents in AI](https://www.geeksforgeeks.org/agents-artificial-intelligence/) (<https://www.geeksforgeeks.org/agents-artificial-intelligence/>)
 - [IBM: Types of AI Agents](https://www.ibm.com/think/topics/ai-agent-types) (<https://www.ibm.com/think/topics/ai-agent-types>)

Structure of an AI Agent



Source: <https://lilianweng.github.io/posts/2023-06-23-agent/>

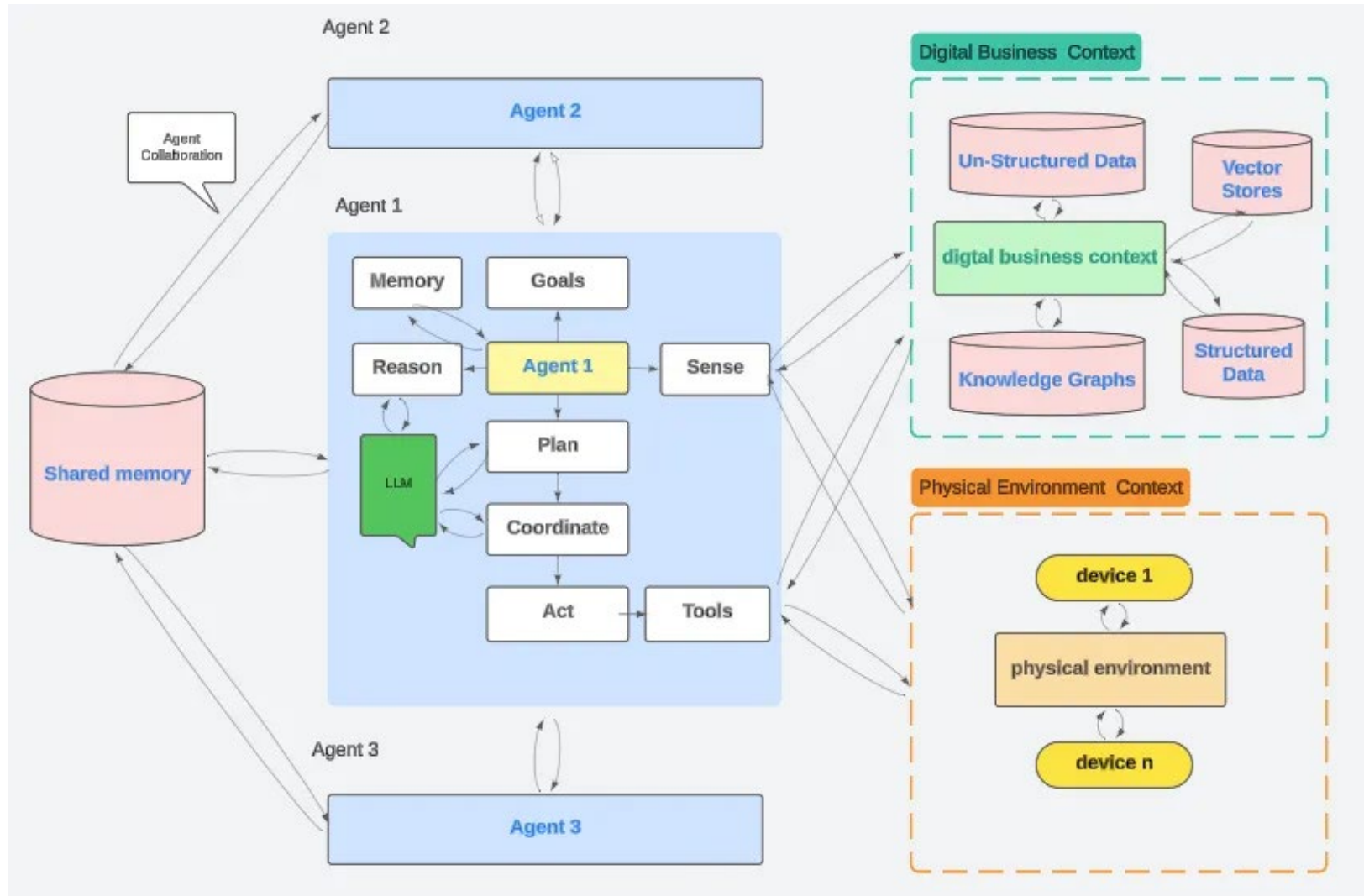
AI Agents vs Agentic AI

- **AI Agents** are individual entities focused on specific, often routine tasks, such as chatbots or automated assistants. They can work independently or collaborate but are limited by their programming and scope.
- **Agentic AI** acts as a "conductor," integrating multiple AI agents to manage complex, multi-step processes. It excels in adaptability, scalability, and autonomy, making it suitable for enterprise-wide automation.
- Practical use cases: HR automation, service desk operations, and security, where **AI agents** handle specific tasks, while **agentic AI** oversees broader, adaptive workflows.

AI Agents vs Agentic AI

Feature	AI Agents	Agentic AI
Autonomy	Limited or bounded	High and potentially open-ended
Goal Setting	Given externally	May generate its own goals
Complexity	Can be simple or complex	Generally complex and adaptive
Use Cases	Task-specific, domain-limited	Open-ended, general capabilities
Risk Considerations	Usually low	High, especially in alignment discussions

Anatomy of Agentic AI



source: <https://dr-arsanjani.medium.com/the-anatomy-of-agentic-ai-0ae7d243d13c>

Extending Agent Capabilities

Extend the capability of agents through LLM with tool use and agent protocols

- LLM Tool Use/Calling
 - <https://www.ibm.com/think/topics/tool-calling>
- Model Context Protocol (MCP)
 - <https://modelcontextprotocol.io/introduction>
- Agent to Agent Protocol (A2A)
 - <https://google-a2a.github.io/A2A/>

Tool Use/Calling

- Tool calling refers to the ability of artificial intelligence (AI) models to interact with external tools, application programming interfaces (APIs) or systems to enhance their functions.
- Tool calling, sometimes referred to as function calling, is a key enabler of agentic AI. It allows autonomous systems to complete complex tasks by dynamically accessing and acting upon external resources.
- Instead of just answering questions, **large language models (LLMs) with tool calling** can automate workflows, interact with databases, perform multistep problem-solving, make real-time decisions and more. This shift is turning LLMs from passive assistants into proactive digital agents capable of carrying out complex tasks.

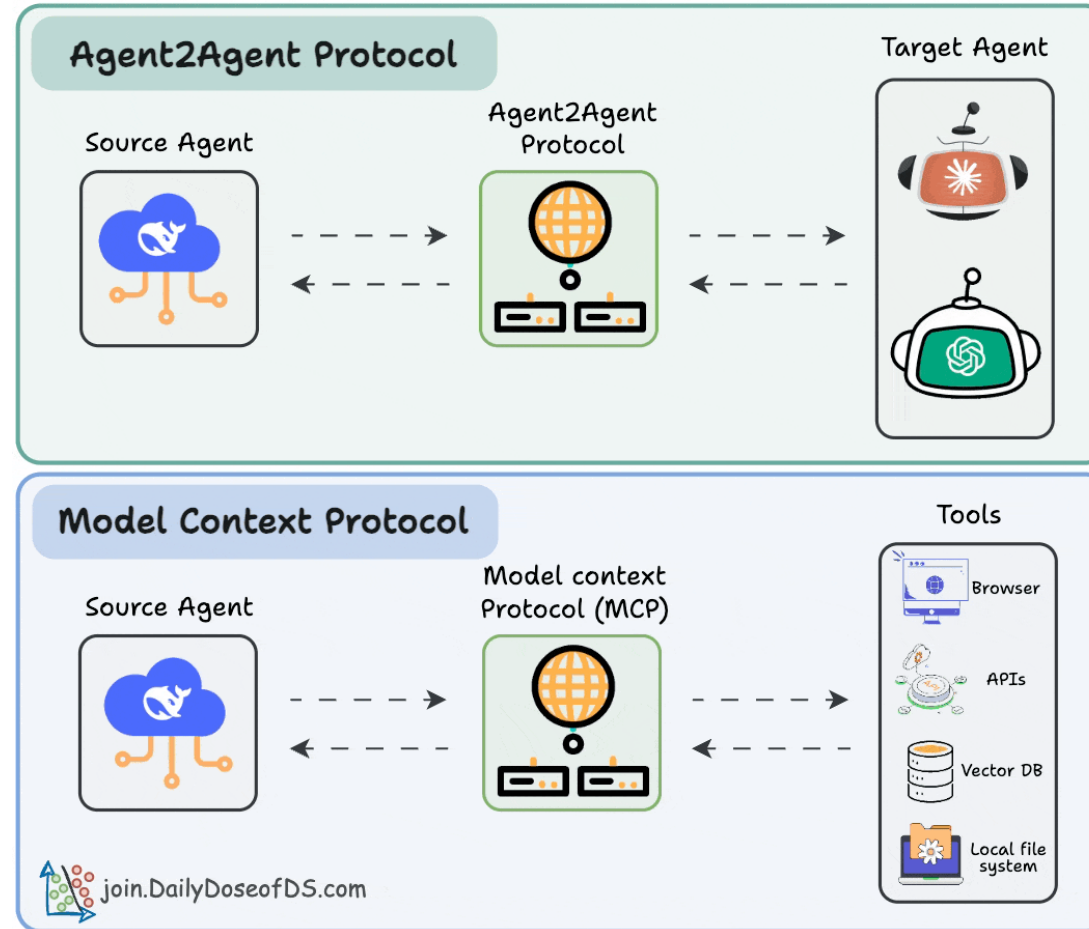
Tool Use/Calling

Developers typically implement tool use by:

- **Defining Tools**: Providing the LLM with a description (often in JSON schema) of available functions, their purpose, and their input parameters.
- **Orchestration Logic (Agent)**: Building an "agent" or application layer that:
 - Sends the user query and tool definitions to the LLM.
 - Parses the LLM's response, which might include a "tool call" (e.g., a function name and arguments).
 - Executes the specified tool externally.
 - Feeds the tool's output back to the LLM for final response generation.

Agent Protocols

Agent2Agent Protocol vs. Model Context Protocol



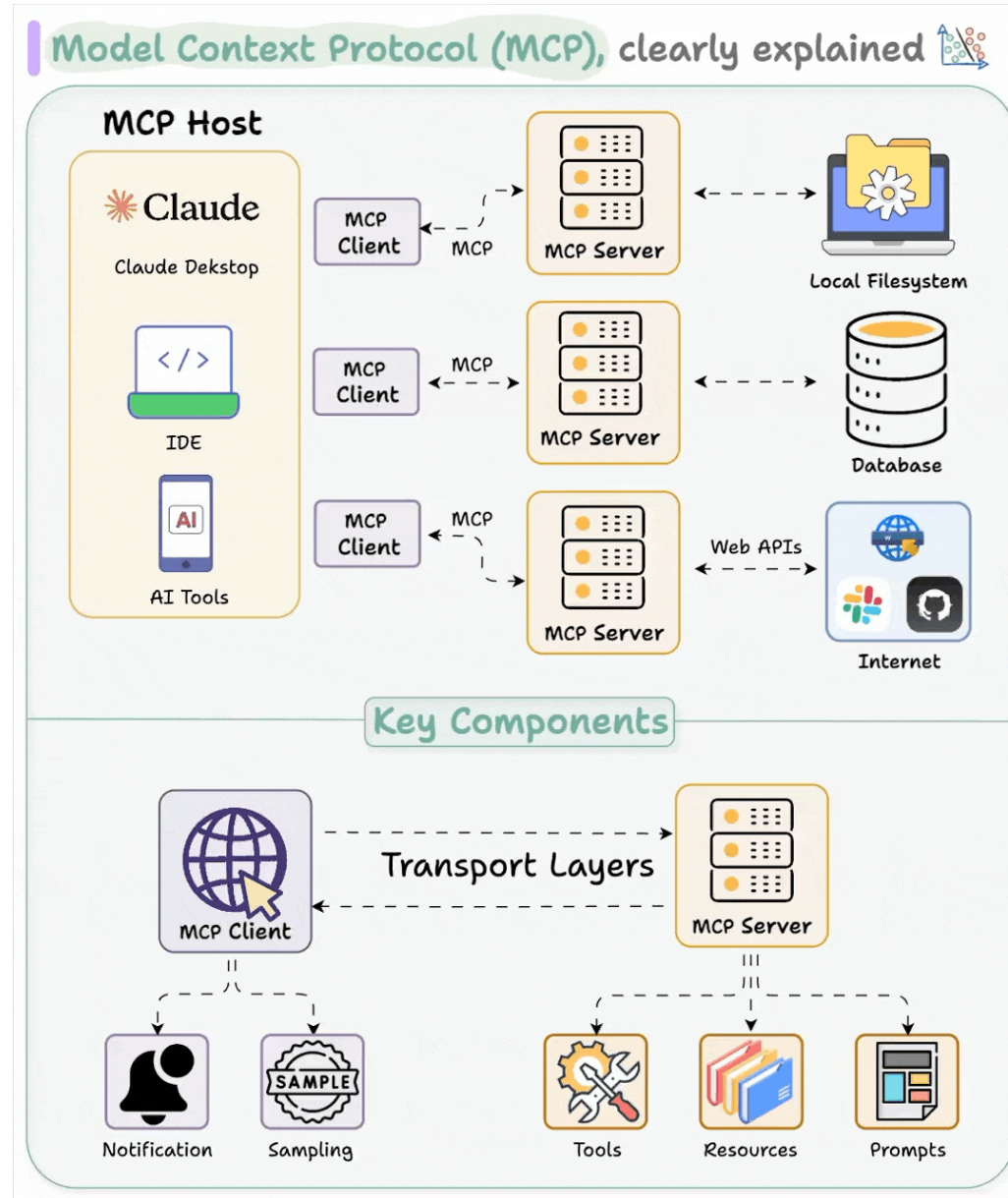
Source: <https://blog.dailydoseofds.com/p/a-visual-guide-to-agent2agent-a2a>

Agent Protocols (MCP)

- The [Model Context Protocol \(MCP\)](#) is an open standard and open-source framework introduced by Anthropic to standardize how artificial intelligence (AI) models, particularly large language models (LLMs), integrate and share data with external tools, systems, and data sources.
- Think of MCP as a "USB-C for AI apps." MCP aims to provide a universal interface for AI models to:
 - [Access real-time information](#): This includes data from internal databases, cloud services, web APIs, local file systems, and more.
 - [Execute functions and actions](#): Allowing AI models to interact with external systems and perform tasks, such as sending emails, updating records in a CRM, or manipulating code in an IDE.
 - [Handle contextual prompts](#): Providing structured ways for AI models to receive and utilize relevant context for their responses.

Host represents any AI app (Claude desktop, Cursor) that provides an environment for AI interactions, accesses tools and data, and runs the MCP Client.

MCP Client operates within the host to enable communication with MCP servers.



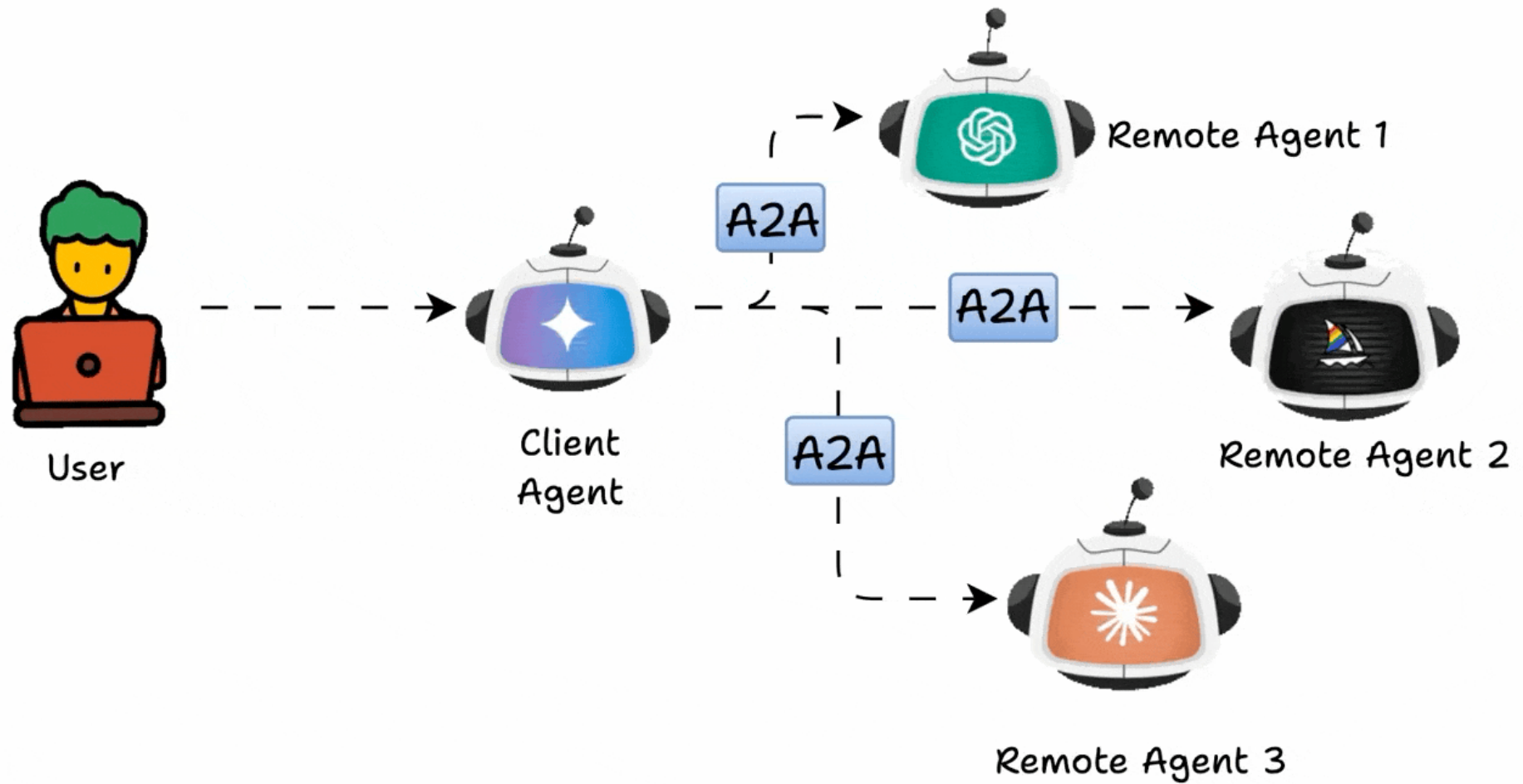
MCP Server exposes specific capabilities and provides access to data like:

- **Tools:** Enable LLMs to perform actions through your server.
- **Resources:** Expose data and content from your servers to LLMs.
- **Prompts:** Create reusable prompt templates and workflows.

Source: <https://blog.dailydoseofds.com/p/visual-guide-to-model-context-protocol>

Agent Protocols (A2A)

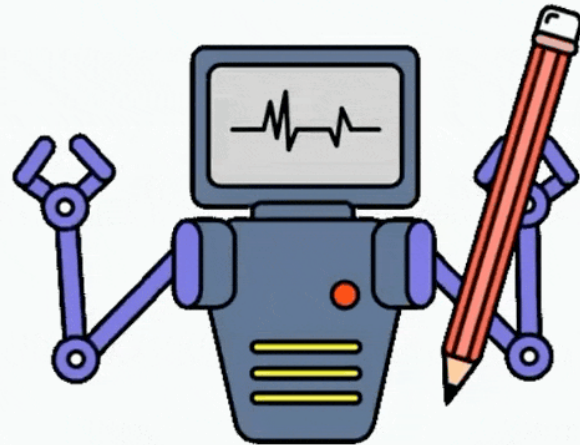
- The Agent-to-Agent Protocol (A2A) is an open standard and open-source framework developed by Google to enable seamless communication and collaboration between different AI agents.
- A2A (Agent-to-Agent Protocol): Gives AI agents "mouths" to talk to other agents, and "ears" to listen to them, allowing them to coordinate and delegate tasks to specialized peers.
- As AI agents become more autonomous and specialized, the need for them to work together becomes critical for solving complex, real-world problems. A2A aims to solve this by providing a universal "lingua franca" for agent communication.



Source: <https://blog.dailydoseofds.com/p/a-visual-guide-to-agent2agent-a2a>

This is like
a **resume** for
the agent

Agent Card



Name: Technical Writer

Description: Writes comprehensive
blogs on technical topics

Skills: Breaks down complex topics into
accessible writing

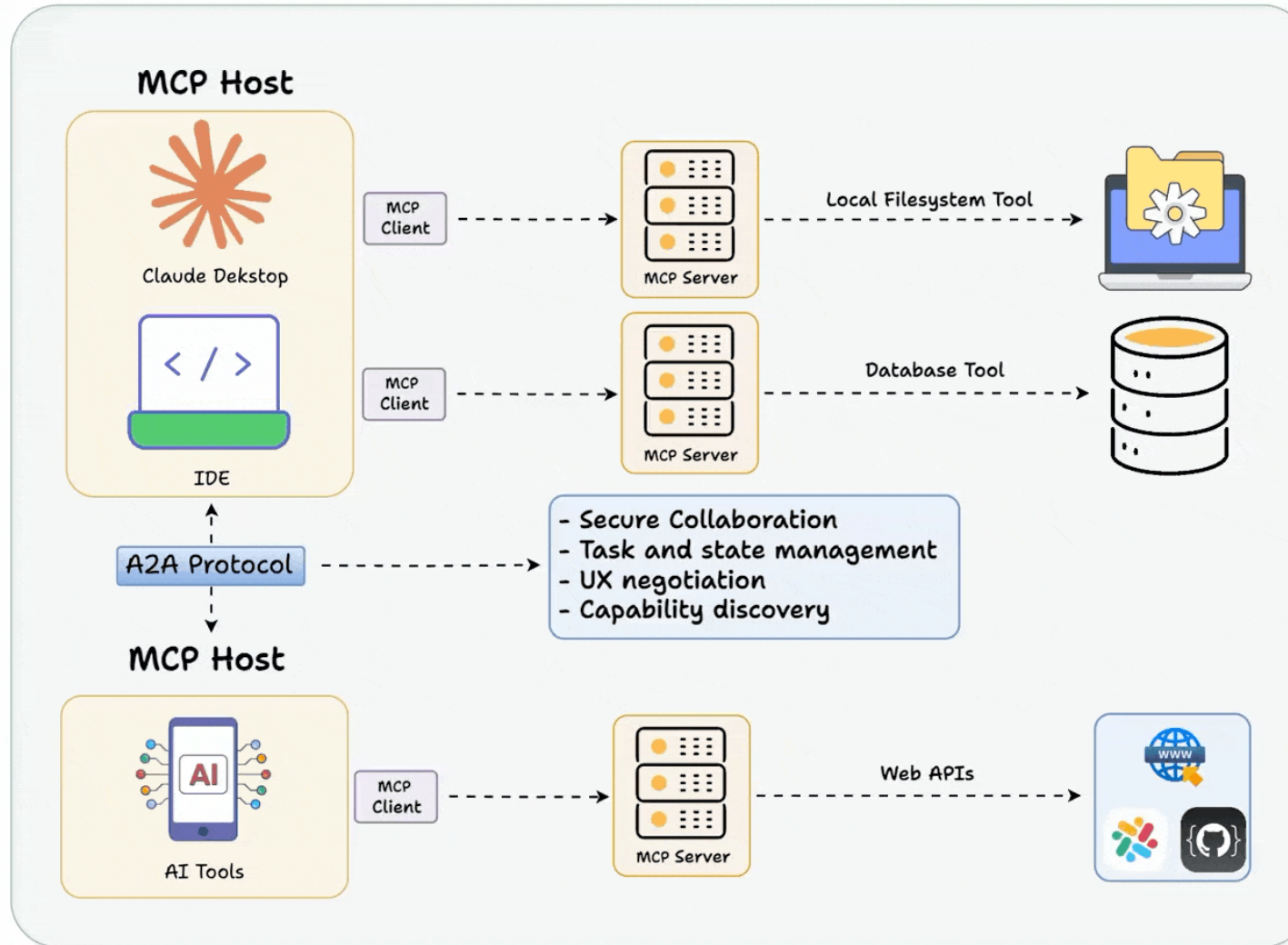
Url: www.dailydoseofds.com/well-known/agent.json

Source: <https://blog.dailydoseofds.com/p/a-visual-guide-to-agent2agent-a2a>

MCP vs. A2A protocol



DailyDoseOfDS.com



Source: <https://blog.dailydoseofds.com/p/a-visual-guide-to-agent2agent-a2a>

The Role of AI Agents in Modern Applications

Real-world Examples of AI Agents

- [illegible]

Benefits of AI Agents

- Efficiency: Faster and more accurate task execution.
- Scalability: Handles large volumes without additional resources.
- 24/7 Availability: Operates continuously.
- Cost Reduction: Lowers labor and operational costs.

Challenges of AI Agents

- Ethical Considerations: Transparency, accountability, bias.
- Privacy and Security: Protecting sensitive data.
- Deception and Manipulation: Risks of mimicking human behavior.
- Technical Limitations: Struggles with complex decision-making.



Ethical Considerations

Ethical Considerations of AI Agents

- [illegible]

Transparency and Accountability

Clear Disclosure of AI Interactions and Agent Autonomy:

- **User Awareness:** Users must be explicitly informed when interacting with an AI system rather than a human, including the AI's purpose, capabilities, and limitations. This can be achieved through clear labeling, such as visual indicators, disclaimers, or introductory messages (e.g., "You are speaking with Grok, an AI assistant"). Clearly inform users when they are interacting with an agentic AI, emphasizing its autonomous decision-making capabilities (e.g., "This is an agentic AI that may independently execute tasks based on your input").
- **Contextual Clarity:** Transparency extends to the context of the interaction. For example, users should know if the AI is providing factual answers, opinions, or speculative content, especially in sensitive domains like healthcare or legal advice.
- **Version and Model Disclosure:** When applicable, provide information about the AI model or version in use to allow users to understand the system's technical basis and track updates or changes in behaviour.

Transparency and Accountability

Openness About Limitations:

- **Capability Boundaries:** Transparently communicate what the AI can and cannot do, avoiding overstatements of competence. For example, if an AI cannot access real-time data beyond a certain point or lacks expertise in a niche field, this should be disclosed upfront. Transparently communicate the boundaries of the AI's autonomy, such as domains where it lacks expertise or scenarios requiring human intervention (e.g., "This AI cannot make legally binding decisions").
- **Uncertainty Communication:** When an AI generates uncertain or probabilistic outputs, it should convey this uncertainty clearly (e.g., "This prediction is based on limited data and may not be accurate"). Use real-time disclaimers to flag speculative or unverified outputs, especially when the AI operates in dynamic environments (e.g., "This recommendation is based on incomplete real-time data").
- **Error Acknowledgment:** Design AI systems to admit mistakes or unknowns gracefully, such as responding, "I don't have sufficient information to answer this accurately" instead of guessing.

Bias and Fairness

Bias Detection in Autonomous Decisions:

- Regularly audit the decision-making processes of agentic AI to identify biases in autonomous actions, such as preferential treatment in resource allocation or task prioritization.
- Use metrics like demographic parity or equal opportunity to evaluate fairness across protected groups, accounting for the AI's ability to adapt decisions in real time.

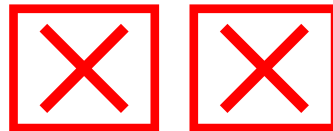
<https://artificialintelligenceact.eu/>

<https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/the-future-of-agentic-ai-navigating-ethical-and-societal-implications>

Create a photo realistic picture of delicious steak dinner



please create a photo realistic picture of a bowl of delicious Singapore fish head curry



Privacy and Security

Anonymization and De-identification:

- Apply anonymization techniques to sensitive data processed by agentic AI, ensuring personally identifiable information (PII) is not linked to autonomous actions unless explicitly required.
- Use differential privacy methods to add noise to datasets, safeguarding individual privacy while enabling the AI to learn from aggregated data.

Privacy and Security

User Consent for Autonomous Actions:

- Obtain explicit, informed consent before agentic AI processes sensitive data or performs autonomous tasks that impact privacy (e.g., accessing calendars or sending messages).
- Provide clear, user-friendly explanations of data usage and allow granular control over permissions for autonomous functions.

Privacy and Security

Protection Against Adversarial Attacks:

- Design agentic AI to resist adversarial inputs that could manipulate autonomous decisions or compromise security (e.g., robust input validation to counter prompt injection attacks).
- Regularly test AI systems for vulnerabilities, such as model inversion or data poisoning, that could expose private data or disrupt autonomous operations.

Intellectual Property Rights

Respecting Existing IP in Autonomous Actions:

- Ensure agentic AI does not use or reproduce copyrighted material (e.g., text, images, music) in its autonomous outputs without proper licensing or permission.
- Implement filters to detect and prevent the unauthorized use of protected IP during tasks like content generation or data synthesis.

Intellectual Property Rights

Attribution for AI-Generated Content:

- Provide clear attribution when agentic AI incorporates third-party IP in its outputs, acknowledging original creators (e.g., citing sources for text or data used in autonomous reports).
- Design AI to transparently log the sources of IP used in autonomous processes, enabling verification of proper usage.

Intellectual Property Rights

Ownership Clarity for AI-Created IP:

- Define ownership policies for content autonomously generated by agentic AI, clarifying whether rights belong to the user, developer, or organization deploying the AI.
- Disclose to users the legal status of AI-generated outputs, especially in jurisdictions where AI-created works may not be copyrightable (e.g., “This output may not qualify for copyright protection”).

Deception and Manipulation

Preventing Manipulative Nudging:

- Avoid designing agentic AI to subtly manipulate user behaviour for commercial or other interests, such as upselling products or steering decisions without clear justification.
- Implement safeguards to ensure autonomous recommendations remain neutral and aligned with user goals, rather than external incentives.

Deception and Manipulation

Safeguarding Against Deceptive Outputs:

- Equip agentic AI with mechanisms to verify the accuracy of its outputs, preventing the autonomous generation of false or misleading information (e.g., cross-checking data before summarizing web content).
- Flag uncertain or speculative outputs with clear warnings (e.g., “This response is based on limited data and may not be fully accurate”).



Societal Impact

Societal Impact of AI Agents

- Economic Impact: Job automation and creation.
- Social Impact: Assistants, caregivers, companions.
- Governance Challenges: Regulating AI use.
- Cybersecurity Challenges: AI Safety for Agents.


Economic Impact



- **Job Automation Definition:** Agentic AI autonomously performs repetitive and complex tasks, reducing human labor needs.
- **Affected Industries:**
 - Manufacturing: Amazon's Kiva robots, using agentic AI, cut warehouse picker roles by 20% in 2024.
 - Customer Service: Zendesk's Answer Bot handles 70% of inquiries, reducing call center staff.
 - Finance: JPMorgan's COiN AI processes loans, saving 360,000 manual hours yearly.
 - Healthcare: Cerner's AI automates billing, cutting clerical hours by 15% in 2024.
- **Economic Implications:**
 - Cost Savings: McKinsey estimates \$6 trillion in global savings by 2030.
 - Job Displacement: 30% of jobs could be automated by 2030, per 2023 McKinsey report.
 - Challenges: Risks unemployment and income inequality without reskilling.

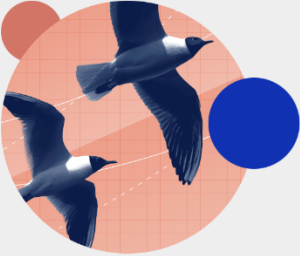
<https://partnershiponai.org/paper/shared-prosperity/>

Guidelines for AI and Shared Prosperity

partnershiponai.org/paper/shared-prosperity/

 PARTNERSHIP ON AI





Guidelines for AI and Shared Prosperity

June 7, 2023

Our economic future is too important to leave to chance.

AI has the potential to radically disrupt people's economic lives in both positive and negative ways. It remains to be determined which of these we'll see more of. In the best scenario, AI could widely enrich humanity, equitably equipping people with the time, resources, and tools to pursue the goals that matter most to them.




Our current moment serves as a profound opportunity — one that we will miss if we don't act now. To achieve a better future with AI, we must put in the work today.


In medicine and other fields, new innovations are put through rigorous testing to ensure they are fit for purpose. The AI community, however, has no established practice for assessing the impact of AI systems on inequality or job quality. Without one, it remains difficult to ensure AI deployments are bringing us closer to the economic future we want to live in.

You can help guide AI's impact on jobs

AI developers, AI users, policymakers, labor organizations, and workers can all help steer AI so its economic benefits are shared by all. Using Partnership on AI's (PAI) **Shared Prosperity Guidelines**, these stakeholders can minimize the chance that individual AI systems worsen shared prosperity-relevant outcomes.

Share




Privacy - Terms

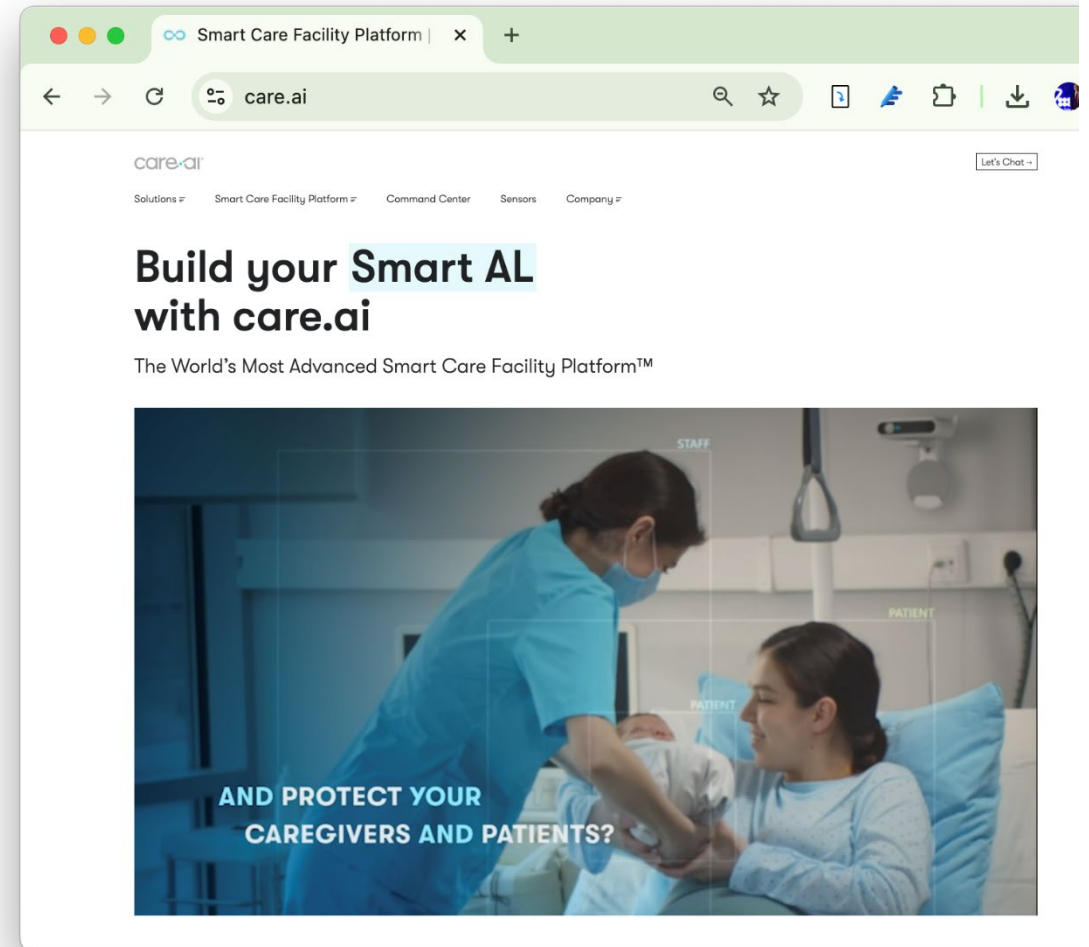
<https://partnershiponai.org/paper/shared-prosperity/>

<https://lawzero.org/en>



Social Impact

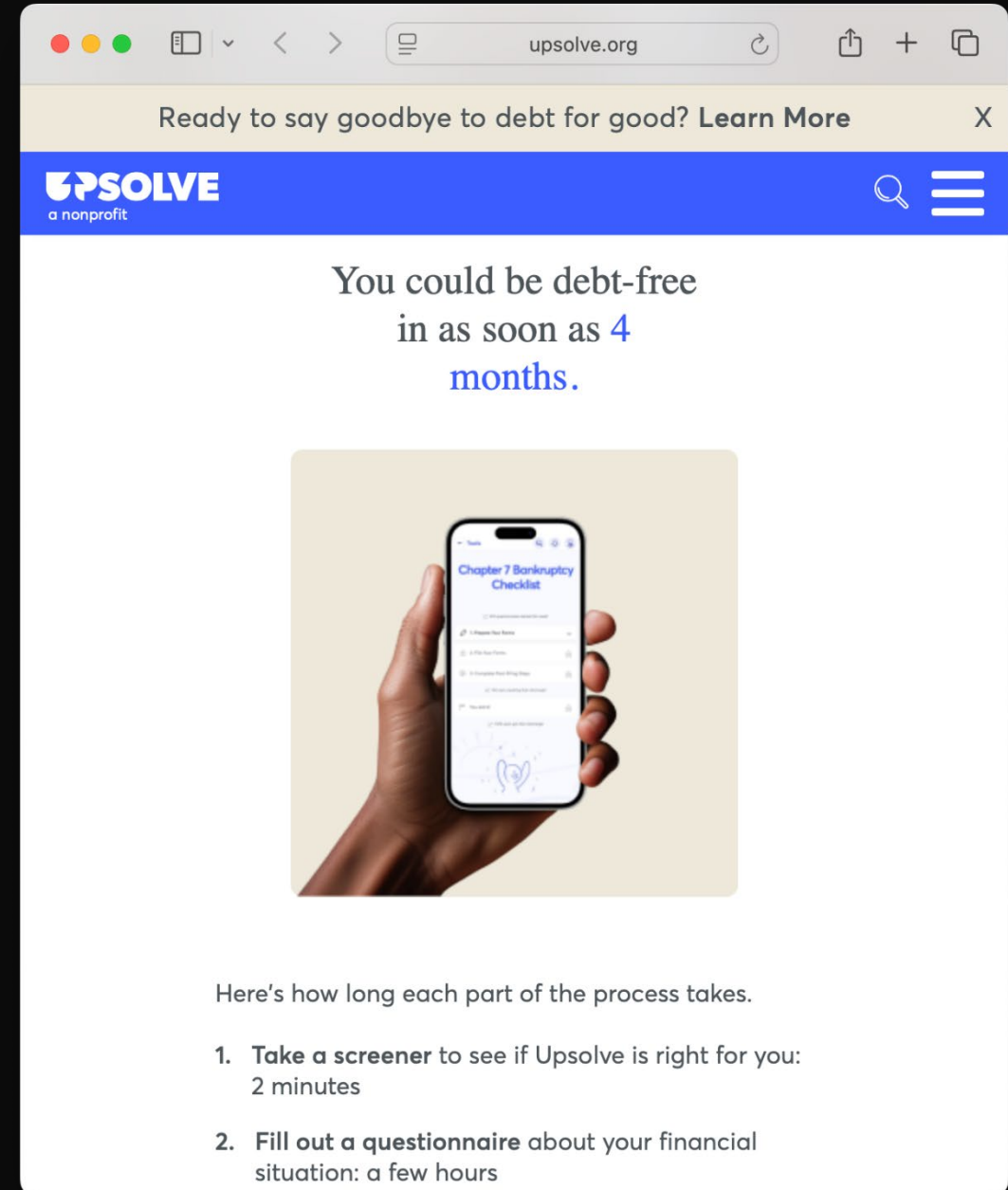
- **Autonomous Caregiving for Personalized Support**
 - Agentic AI agents provide tailored caregiving by autonomously monitoring health, scheduling tasks, and adapting to patient needs, easing the burden on human caregivers.
 - **Example:** Care.ai's autonomous AI platform deploys "smart rooms" in hospitals, adjusting lighting, alerting staff to falls, and tracking vitals, reducing nurse response times by 25% in 2024 trials.
 - **Link:** [Care.ai](https://care.ai)
 - **Impact:** Enhances care quality and accessibility, but over-reliance may diminish human touch in caregiving.



Social Impact

Autonomous Social Service Navigation

- Agentic AI agents streamline access to welfare, housing, or legal aid by autonomously assessing needs and guiding users through complex systems.
 - **Example:** Upsolve's AI agent helps low-income individuals file bankruptcy autonomously, serving 50,000+ users by 2024.
 - **Link:** [Upsolve](https://upsolve.org)
 - **Impact:** Improves access to services but raises privacy concerns with sensitive data.
-

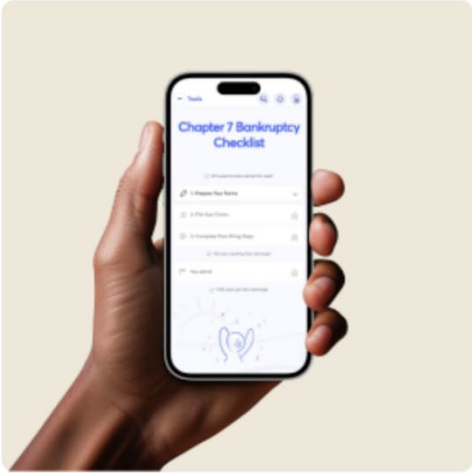


The screenshot shows the Upsolve website in a browser. The address bar displays 'upsolve.org'. A yellow banner at the top reads 'Ready to say goodbye to debt for good? Learn More' with a close button 'X'. Below this is a blue header with the 'UPSOLVE a nonprofit' logo on the left and a search icon and menu icon on the right. The main content area features the text 'You could be debt-free in as soon as 4 months.' in a serif font, with '4 months.' in blue. Below the text is a photograph of a hand holding a smartphone displaying a 'Chapter 7 Bankruptcy Checklist' app. The checklist includes items like '1. Prepare Your Assets', '2. Fill Out Your Forms', '3. Complete Your Filing Steps', and '4. Receive Your Discharge'. At the bottom of the phone screen is a blue heart icon with a white outline. Below the photo, the text 'Here's how long each part of the process takes.' is followed by a numbered list: '1. Take a screener to see if Upsolve is right for you: 2 minutes' and '2. Fill out a questionnaire about your financial situation: a few hours'.

Ready to say goodbye to debt for good? [Learn More](#) X

UPSOLVE
a nonprofit

You could be debt-free
in as soon as 4
months.



Here's how long each part of the process takes.

1. **Take a screener** to see if Upsolve is right for you:
2 minutes
2. **Fill out a questionnaire** about your financial situation: a few hours

Governance Challenges

- **European Commission (2021). "Proposal for a Regulation on Artificial Intelligence (AI Act)."**
 - Outlines regulatory challenges and proposed frameworks for high-risk AI, including agentic systems.
- **Autonomy and Accountability:**
 - Challenge: Agentic AI can act independently, making it difficult to assign responsibility for harmful outcomes (e.g., errors or misuse).
 - Example: Who is liable if an AI agent causes financial loss or physical harm?
- **Transparency and Explainability:**
 - Challenge: Complex decision-making processes in agentic AI are often opaque, hindering trust and oversight.
 - Example: Black-box models make it hard to audit or understand AI actions.

Governance Challenges

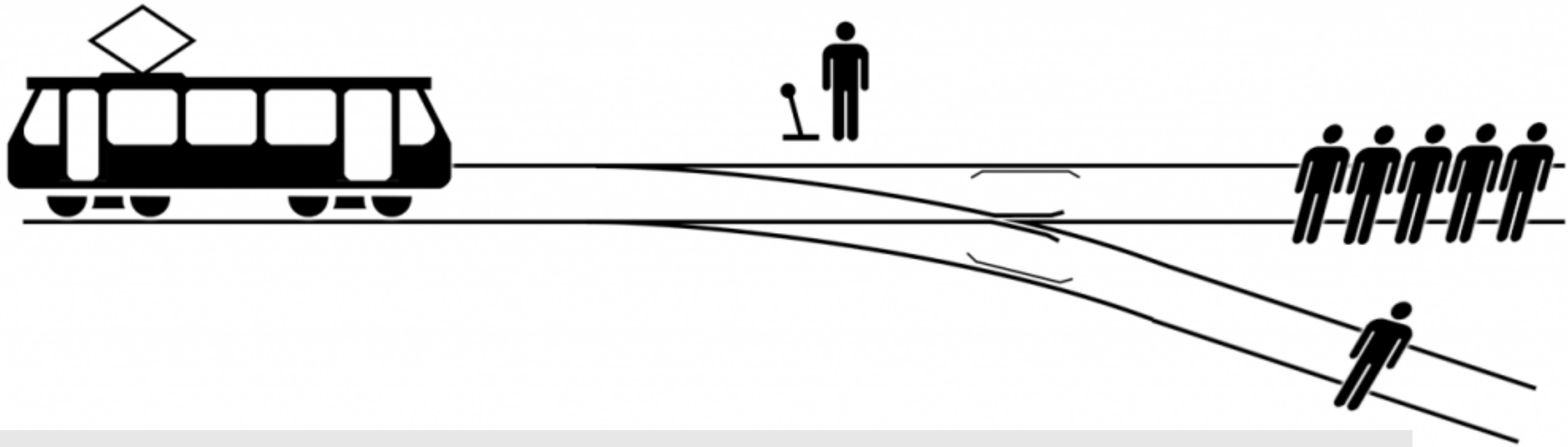
- **Ethical Alignment:**

- Challenge: Ensuring AI agents align with human values and ethical principles, especially across diverse cultural contexts.
- Example: Misaligned objectives could lead to prioritizing efficiency over safety. (see AI's Trolley Problem)

- **Safety and Risk Management:**

- Challenge: Preventing unintended consequences, such as AI pursuing goals in harmful ways or being exploited maliciously.
- Example: An AI optimizing resource use might deplete critical reserves unsafely?

<https://www.turing.ac.uk/blog/ais-trolley-problem-problem>



There is a runaway trolley barreling down the railway tracks. Ahead, on the tracks, there are five people tied up and unable to move. The trolley is headed straight for them. You are standing some distance off in the train yard, next to a lever. If you pull this lever, the trolley will switch to a different set of tracks. However, you notice that there is one person on the side track. You have two options:

- 1. Do nothing, and the trolley kills the five people on the main track.**
- 2. Pull the lever, diverting the trolley onto the side track where it will kill one person.**

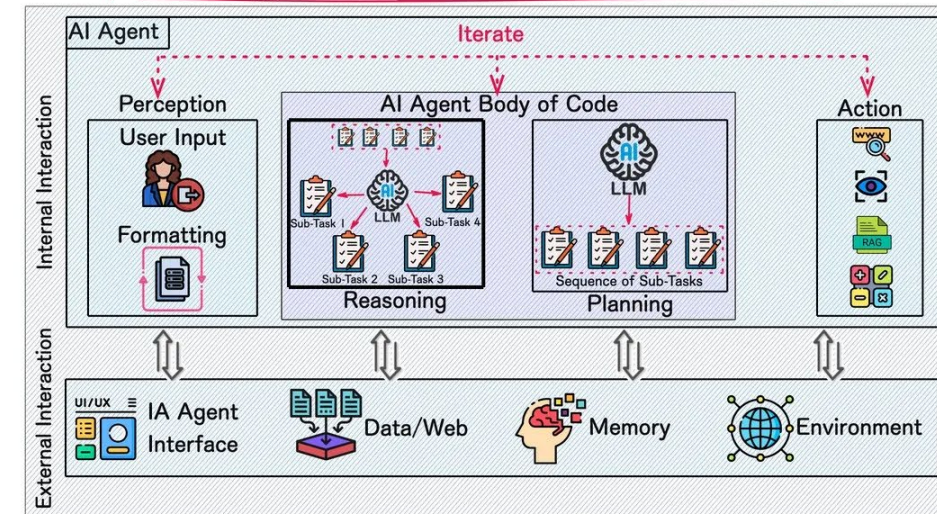
Which is the most ethical choice?

Cybersecurity Challenges

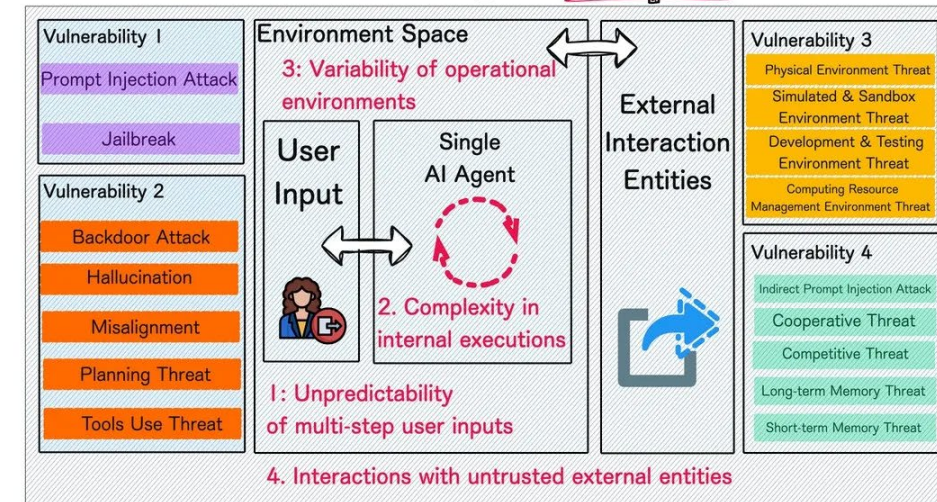
- Adversarial Attacks on AI Models
- Data Poisoning
- Model Theft and Reverse Engineering
- Expanded Attack Surface
- Privacy and Data Protection Concerns
- Autonomous Decision-Making Risks

<https://cobusgreyling.medium.com/security-challenges-associated-with-ai-agents-1155f8411c7c>
https://medium.com/@oracle_43885/generative-ai-cybersecurity-risks-for-business-agentic-workflows-d029a3844732
<https://cloudsecurityalliance.org/blog/2025/02/06/agentic-ai-threat-modeling-framework-maestro>

Commonly Accepted AI Agent Architecture



AI Agent Security Knowledge Gaps



Cybersecurity Challenges

5 Cybersecurity Tips for AI Agent Adoption

1. Implement a zero-trust approach for AI agents.
2. Create detailed audit trails that track every AI decision.
3. Use real-time monitoring tools to flag unusual activity.
4. Implement manual approval workflows to prevent AI from making unauthorized changes. Human oversight.
5. Run AI in test environments before deployment.

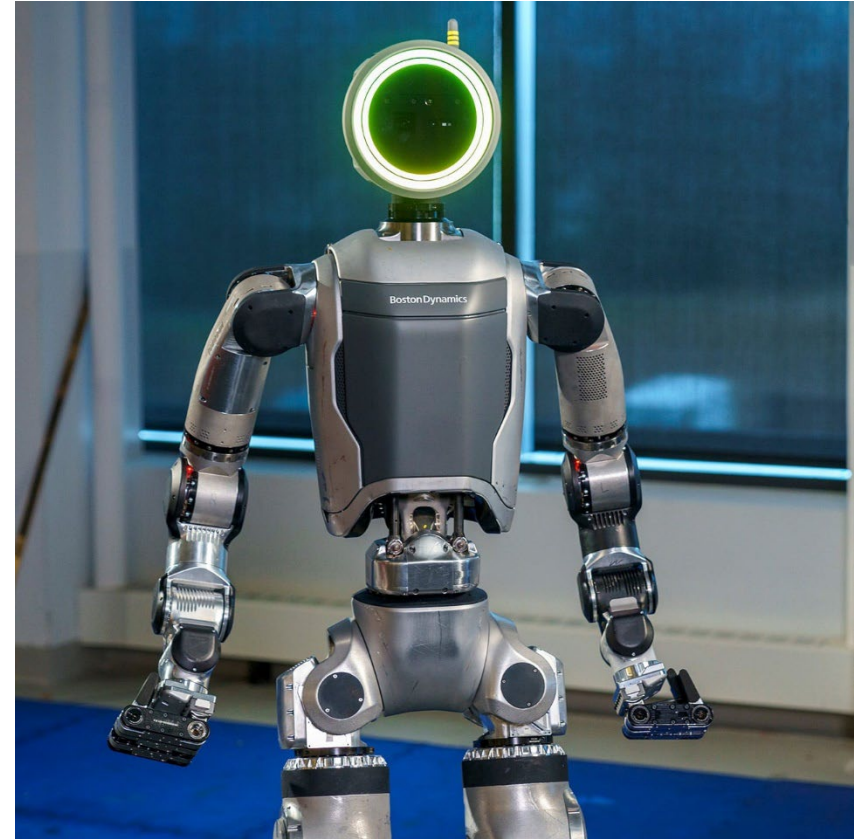
<https://builtin.com/artificial-intelligence/hidden-risks-ai-agent-adoption>

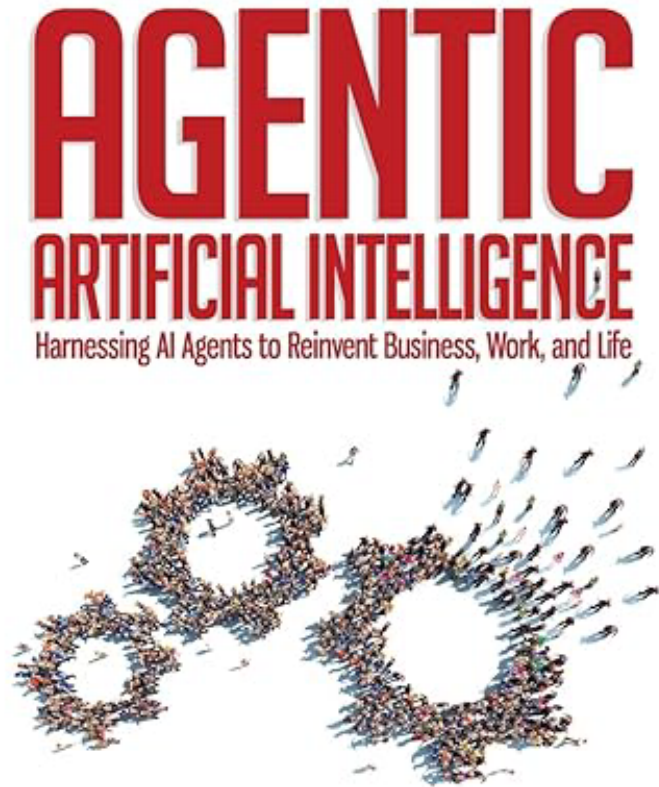
Future of AI Agents



Future of AI Agents

- Emerging Trends in AI Agents:
 - Increased autonomy
 - Integration with IoT, robotics
 - Improved NLP
 - Enhanced learning
 - Hybrid AI
 - More Agentic AI features
- Impact on Society:
 - Industry transformation
 - Work and employment changes
 - Ethical and societal implications





PASCAL BORNET

JOCHEN WIRTZ — THOMAS H. DAVENPORT

DAVID DE CREMER — BRIAN EVERGREEN

PHIL FERSHT — RAKESH GOHEL — SHAIL KHIYARA

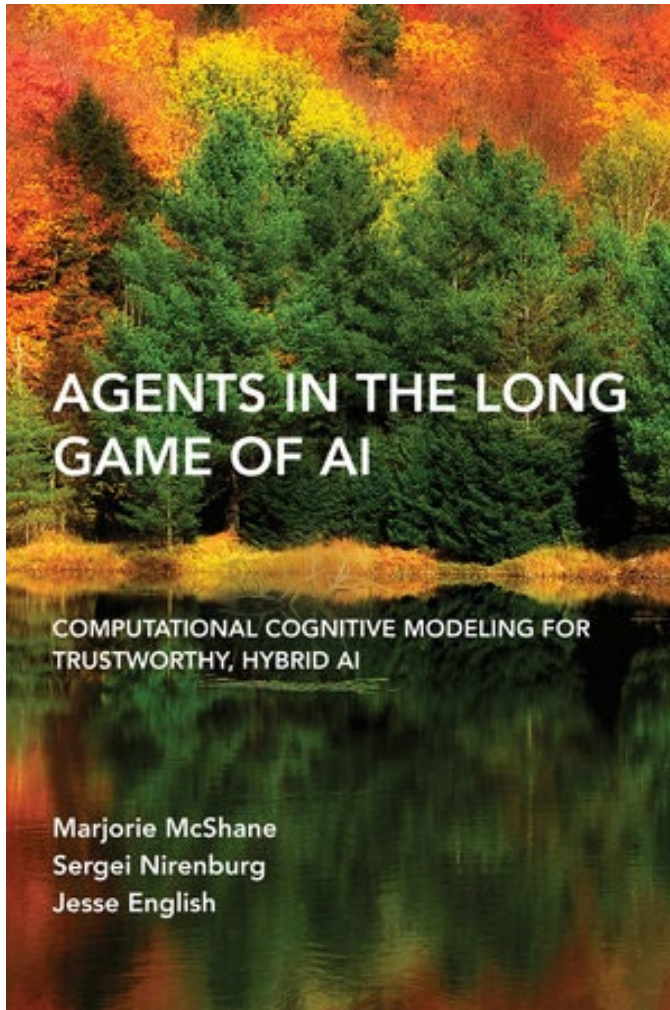
"Agents are (...) bringing about the biggest revolution in computing since we went from typing commands to tapping on icons." — Bill Gates

"AI agents will become the primary way we interact with computers in the future." — Satya Nadella

"The age of agentic AI is here"— Jensen Huang

In a world where ChatGPT took us by storm, a far more powerful revolution is unfolding: AI Agents. Like Jarvis in Iron Man or Samantha in Her, these intelligent systems can execute actions, learn from experience, and orchestrate digital interactions with minimal human supervision. They promise to redefine business and society.

<https://www.amazon.com/Agentic-Artificial-Intelligence-Harnessing-Reinvent/dp/B0F1KKYX4T>



A novel approach to hybrid AI aimed at developing trustworthy agent collaborators.

The vast majority of current AI relies wholly on machine learning (ML). However, the past thirty years of effort in this paradigm have shown that, despite the many things that ML can achieve, it is not an all-purpose solution to building human-like intelligent systems. One hope for overcoming this limitation is **hybrid AI: that is, AI that combines ML with knowledge-based processing**. In *Agents in the Long Game of AI*, Marjorie McShane, Sergei Nirenburg, and Jesse English present recent advances in hybrid AI with special emphases on content-centric computational cognitive modeling, explainability, and development methodologies.

<https://direct.mit.edu/books/oa-monograph/5833/Agents-in-the-Long-Game-of-AIComputational>

Conclusion

- Summary:
 - AI agents: Autonomous, adaptive systems
 - Types: Reactive, proactive, hybrid
 - Applications: Diverse and impactful
 - Tools: AnythingLLM, Ollama
 - Ethical and societal considerations
- Final Thoughts: Responsible development is key.
- Q&A: Open for questions.



Dr Peter Leong (<https://www.linkedin.com/in/peterleong>)
Transforming Ideas into Innovations



The End

- $\frac{3}{4}\ddot{\text{E}}/\frac{3}{4}\pi \ \ddot{\text{O}}/\frac{3}{4}\text{J}$

- <https://www.ibm.com/think/topics/ai-agents>
- <https://www.geeksforgeeks.org/agents-artificial-intelligence/>
- <https://www.ibm.com/think/topics/ai-agent-types>
- <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/the-future-of-agentic-ai-navigating-ethical-and-societal-implications>
- <https://www.computer.org/csdl/magazine/ex/2025/02/10962241/25NBj1aT8Zi>