Research Frontiers
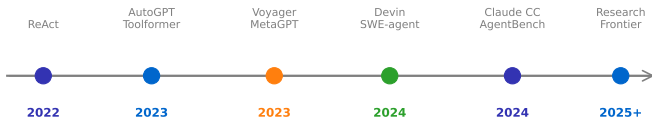
Week 12: Open Problems and Future Directions

PhD Course in Agentic Artificial Intelligence

Agent Research Timeline
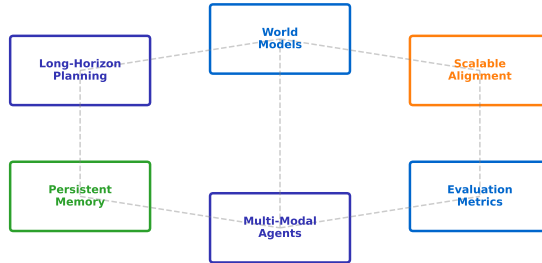

ReAct	AutoGPT Toolformer	Voyager MetaGPT	Devin SWE-agent	Claude CC AgentBench	Research Frontier
2022	2023	2023	2024	2024	2025+

Themes: Reasoning > Tool Use > Multi-Agent > Production

From

reasoning (2022) to production systems (2024) in just two years.

Open Research Problems
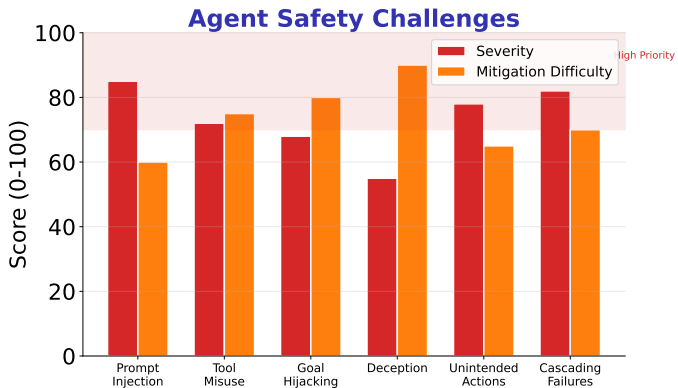


```
                    ┌──────────┐
                    │  World   │
                    │  Models  │
                    └──────────┘
  ┌──────────────┐              ┌──────────────┐
  │ Long-Horizon │              │  Scalable    │
  │   Planning   │              │  Alignment   │
  └──────────────┘              └──────────────┘

  ┌──────────────┐              ┌──────────────┐
  │  Persistent  │              │  Evaluation  │
  │    Memory    │              │   Metrics    │
  └──────────────┘              └──────────────┘
                  ┌──────────────┐
                  │ Multi-Modal  │
                  │    Agents    │
                  └──────────────┘
```

Interconnected challenges requiring holistic solutions

These interconnected challenges define the research agenda.

Agent Safety Challenges

research is critical for responsible agent deployment.

Future Research Directions

Near-term	Mid-term	Long-term
Better benchmarks	World models	Embodied agents
Production tools	Persistent learning	General autonomy
Human-agent collab	Multi-agent systems	Agent societies

Cross-cutting: Safety | Alignment | Interpretability | Evaluation

Key insight: Progress requires interdisciplinary collaboration

Prog

requires interdisciplinary collaboration.

Key Papers

Foundational

- Wang et al. (2023). "Voyager: An Open-Ended Embodied Agent (physical/virtual world)." arXiv:2305.16291
- Park et al. (2023). "Generative Agents: Interactive Simulacra (simulated personas)." arXiv:2304.03442
- Bai et al. (2022). "Constitutional AI." arXiv:2212.08073

Perspectives

- Anthropic (2024). "Agent AI: Surveying the Horizons."

These

papers define the frontier of agent research.

Course Summary

12-Week Journey

- **Foundations**: Agents, LLMs, ReAct, prompting (W1-2)
- **Capabilities**: Tools, planning, multi-agent (W3-5)
- **Frameworks**: LangGraph, AutoGen, production (W6)
- **Knowledge**: RAG, GraphRAG, verification (W7-9)
- **Applications**: Evaluation, domains, frontiers (W10-12)

Key Takeaway

- Agents = LLM + Memory + Tools + Planning + Evaluation

Thank you for participating in this course!