The Cyclic Journey of AI: From Agents to Agents—and Beyond

YUAN HE | NOV 2024

My Career Path in AI

2016

BSc (Hons) AI & Maths

University of Edinburgh

2020

PhD (DPhil) in Computer Science

University of Oxford

2024

Postdoc (Research Associate)

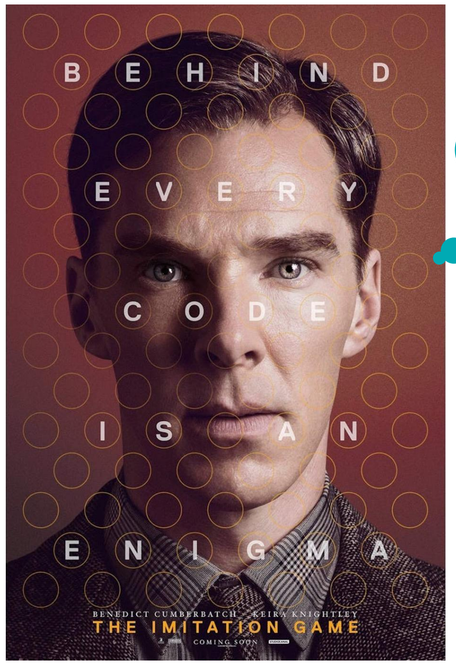
University of Oxford

2025

Applied Scientist

Amazon Rufus

My First Encounters with AI



Encoding & Decoding

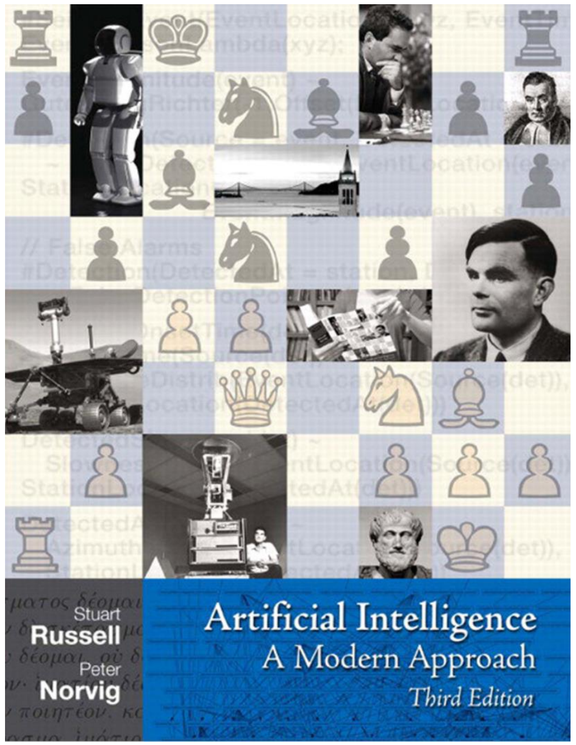
A biographical film about Alan Turing, the father of modern AI

A “Perfect” World?

A classic film exploring the implications of AI on human existence

Taxonomy of AI

- Natural Language Processing
- Knowledge Representation
- Automated Reasoning
- Machine Learning
- Computer Vision
- Robotics

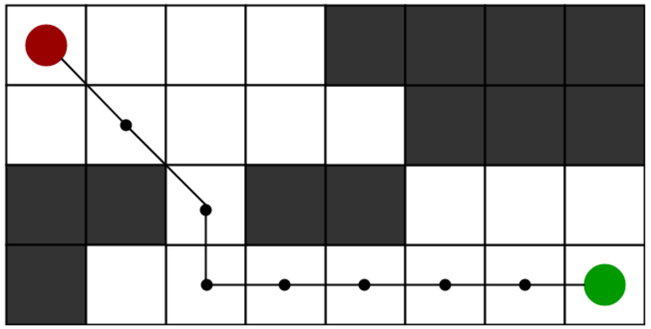


This taxonomy is derived from Russell and Norvig's book: *Artificial Intelligence: A Modern Approach*.

Agents in AI (Last Century)

- **Rational Agent**: An abstraction of agents that can *perceive the environment* and *make decisions* that lead to *best outcomes*.

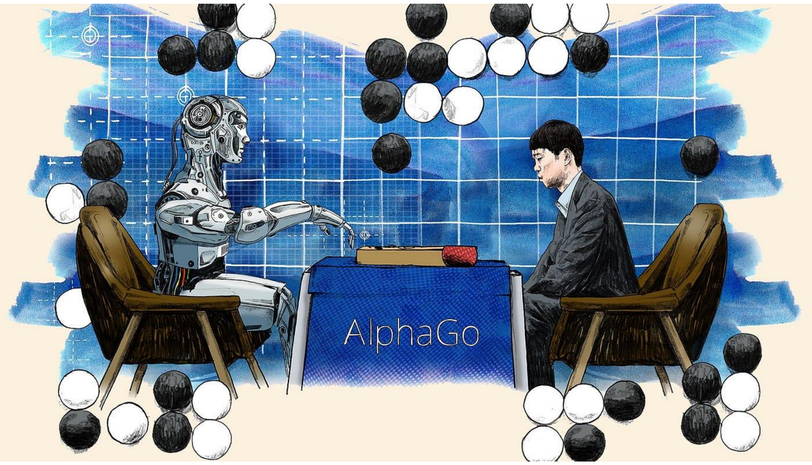
- **Symbolic Agent**: An implementation of agents that rely on expert-curated *rules* and *automated reasoning*.



A* star algorithm for best-route searching and planning commonly used in symbolic agents.
Image: https://www.geeksforgeeks.org/a-search-algorithm/

Agents in AI (Now)

- **Learning Agent**: An implementation of agents that can *learn and capture patterns* from real-world *data*.



AlphaGo (2016) is a learning agent that leverages reinforcement learning (RL) to master the game of Go.
Image: AlphaGo Beats The World's Best Go Player from Medium.

ChatGPT (2022) is a learning agent trained on extensive language data to generate human-like responses.
Image: Brand Guidelines from OpenAI.

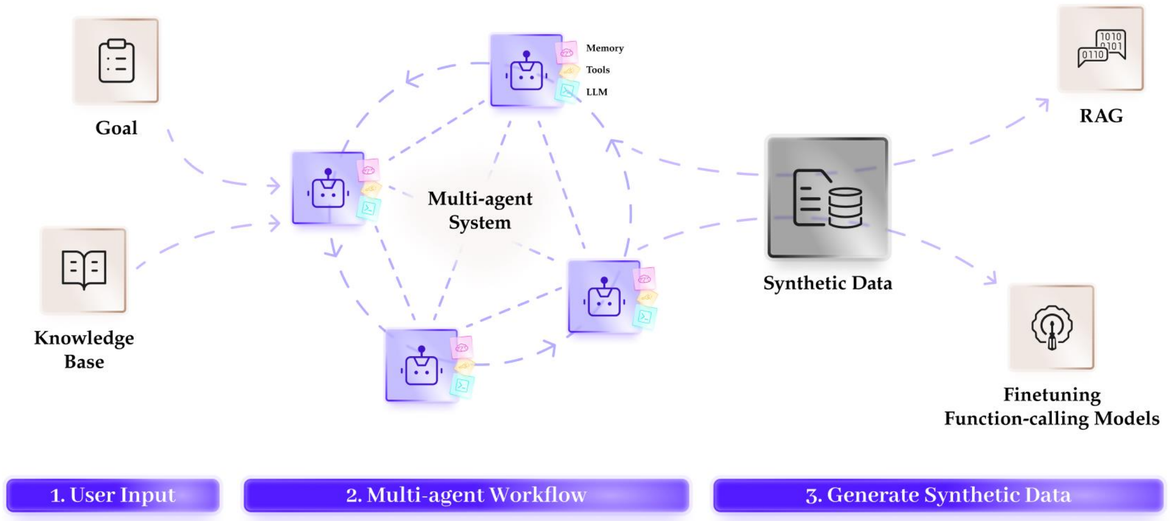
Already Achieve AGI?

- **Q**: Do Large Language Models (LLMs) like ChatGPT become *omni-potent* and nearly solve everything?
- **A**: Nope. LLMs are good at *common sense knowledge*, but they are often wrong when it comes to *domain knowledge* and *computations*.

ChatGPT-4 doesn't know 9.11 is smaller than 9.9.
Image: A dialogue with ChatGPT

Already Achieve AGI?

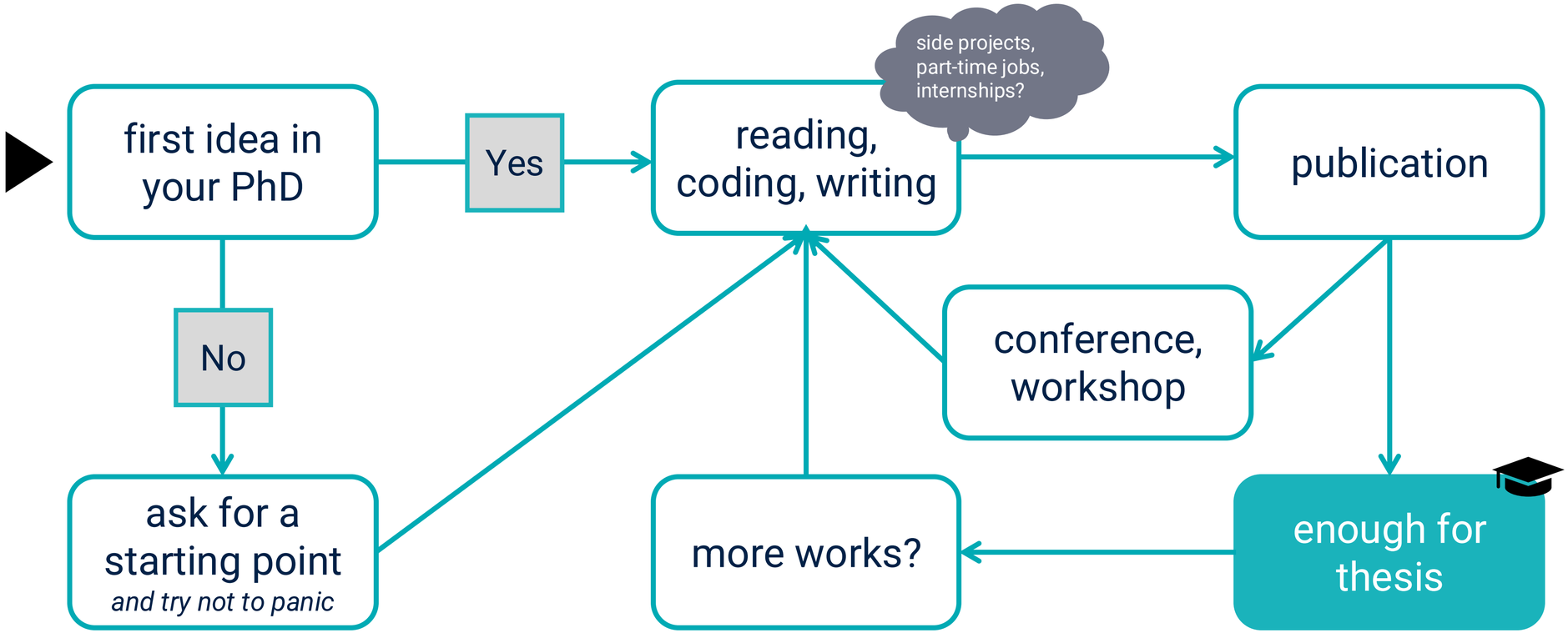
- **Q**: Are old symbolic AI agents completely *obsolete*?
- **A**: Nope. They have new roles in modern *(multi-)agent workflow*; often used to structure *knowledge bases* or support *reasoning* tasks



CAMEL-AI is a well-known multi-agent framework. Image: https://www.camel-ai.org/

My Suggestions for PhDs

An (Arguably) Standard PhD Flow



This is one path—but everyone's journey is unique!

Build Depth in Your Research

- **Avoid frequently changing** your research topic to chase trends.
- **Don’t give up too soon**—evaluate thoroughly and debug carefully.
- Focus on producing **solid, impactful work in one area**, even if it’s not the most popular.
- Interviewers value **deep knowledge of your work**; paper count isn’t everything.

Develop a Solid Foundation

For example, if you work on large language models:

- Can you clearly **define language models**? Can you do the **maths** (e.g., probabilities, linear algebra)?
- Do you **understand key architectures** like transformers, RNNs, and Word2Vec? (not just prompt engineering ☺)
- Do you know the **history of language models**, from early N-gram models to modern transformers, and other **essential NLP concepts**?

Tip: If not, revisit the basics to strengthen your expertise.

Articulate Your Works with Confidence

- Dedicate **~10% of your time** to attending conferences and seminars to improve your **presentation skills** and gain valuable **feedback from other researchers**.

- Present your work with **confidence** and embrace **negative feedback**—it's much better to face an **onsite rebuttal** than to read a "strong reject" rating online on OpenReview!

Rating: 1: strong reject

Confidence: 4: You are confident in your assessment

THANKS!

Scan to visit my homepage if you want to know more and reach out.

UNIVERSITY OF OXFORD

DEPARTMENT OF COMPUTER SCIENCE