

Introduction

AIE 1901 - AI Exploration - LLM for Optimization

Jie Wang, 2025/09/01

Instructor

- Prof. Jie Wang
- Office: TA409a
- Office Hour: 3pm - 4pm, every Wednesday, TA409a
- Class Venue and Time: Tuesday 1:30pm - 3:20pm, TA206 or TC501 (Let's use TA206 for the first three weeks)
- Email: jwang@cuhk.edu.cn

Teaching Assistant

- Mr. Zhaoliang Yuan
- Office Hour: To be determined
- Email: 225080014@link.cuhk.edu.cn
- **Teaching Schedule: Jie Wang and Zhaoliang Yuan will teach in alternating weeks.**

Introduction to myself

- Assistant Professor at School of Artificial Intelligence (home department) and School of Data Science, The Chinese University of Hong Kong, Shenzhen
- Contact: jwang@cuhk.edu.cn
- Webpage: <https://mypage.cuhk.edu.cn/academics/jwang/>
- Got Bachelor in Pure Mathematics (2016-2020) at CUHK SZ
- Got Ph.D. in Industrial Engineering (2020-2025) at Georgia Tech
- Research focus: optimization and statistics for machine learning
- You can call me Prof. Wang or Dr. Wang

Selected Learning Objectives

- Understanding core principles and technologies behind Large Language Models (LLMs)
- Exploring the tools and infrastructure necessary for developing and deploying LLMs
- Understand basic optimization modeling for practical applications
- Apply LLM agents to solve real-world problems.
- Address the ethical considerations of using LLM technology
- Cultivating collaboration, communication, and leadership skills

Course Organization

This course equips you with technical skills essential for Generative AI and Operations Research professionals.

- **Lectures:** Slides posted on Blackboard.
- **Assignments:** Four individual assignments to master core concepts.
- **Project:** A culminating group project to apply LLM techniques for optimization modeling and solving.

Readings

- **Essential Instructional Materials:** Lecture slides, tutorial materials, and selected articles will be distributed throughout the semester.
- **Recommended Readings:**
 - Building LLM Apps: Create Intelligent Apps and Agents with Large Language Models. by Valentina Alto
 - Tutorials by Nvidia: <https://developer.nvidia.com/blog/introduction-to-lm-agents/>
 - Hands-On Large Language Models by Jay Alammar, Maarten Grootendorst, published by O'Reilly Media, Inc
 - An article ORLM: A Customizable Framework in Training Large Models for Automated Optimization Modeling in <https://arxiv.org/abs/2405.17743>

Grading Scheme

Assignments (Individual): 40%

Midterm Group Project Presentation: 20%

Final Group Project Presentation: 40%

Assignment

- Homework Assignments: Several exercises will be assigned during the semester. They are individual assignments and must be completed INDEPENDENTLY to receive full credits. Copied work among several students will result in a zero grade for all parties involved.
- Late Assignments: Late submissions will result in 10% reduction per day late unless approved by the instructor prior to the due date.
- Integrity: We assume that you have complete integrity in all your class efforts. Violations of the university's rules will be taken extremely seriously, and they will be addressed promptly according to established university procedures.

Project Teams

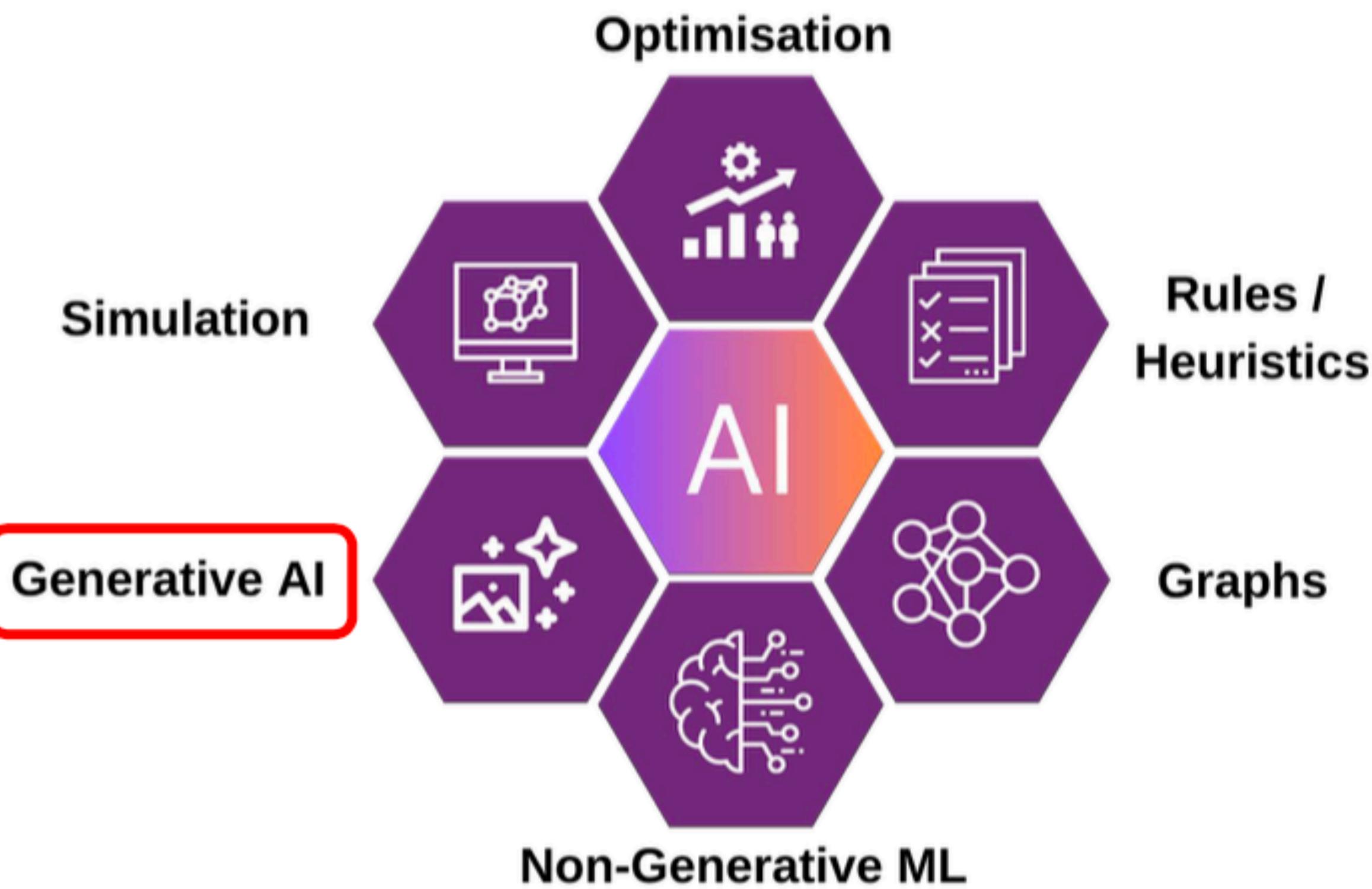
- No more than four members per team
- A team leader should be elected by the members
- Each team will meet the instructor / TA to discuss each report
- Work should be divided clearly among team members
- Should keep a work log that records the contributions made by each team member
- Late project reports will result in 10 % reduction per day late

Project Deliverables

- **Midterm Presentation (20%):** Project pitch & initial progress
- **Final Presentation (40%):** Demonstrate your working solution
- Presentation/Report guidelines will be discussed in class and posted online
- The instructors / TA will schedule a meeting with each group before each presentations to answer questions and ensure smooth project execution

Impacts of Large Language Model

Generative AI is a New Member of AI



Since “Artificial Intelligence” was coined in 1950s, many techniques of AI have been developed, Generative AI is the most recent and the most popular AI technique

AI Will Revolutionize Higher Education

Benefits of AI Higher Education



Access to Vast Amounts of Information

AI-powered search engines and recommendation systems offer students diverse, relevant resources..



Personalized Learning Experience

AI in higher education tailors learning paths based on students' past performance, learning styles, and difficulties.



Enhanced Learning Efficiency

AI-driven tutoring systems adapt content to individual learning styles, boosting confidence and motivation.



Personalized Support

AI chatbots serve as virtual tutors, providing instant, tailored assistance to students.

An AI School is Born in the USA



Arizona School's Curriculum Will Be Taught by AI, No Teachers

By Todd Feathers Published December 19, 2024

- The Arizona State Board for Charter Schools on Monday approved an application from Unbound Academy to open a fully online school serving grades four through eight.
- Under a 2hr Learning model, students spend just two hours a day using personalized learning programs from companies like IXL and Khan Academy

LLMs Enables More Personalized Learning

How can LLMs be best utilized in education?



Personalized Learning

Adapts to individual learning styles and paces, provides tailored learning experiences.

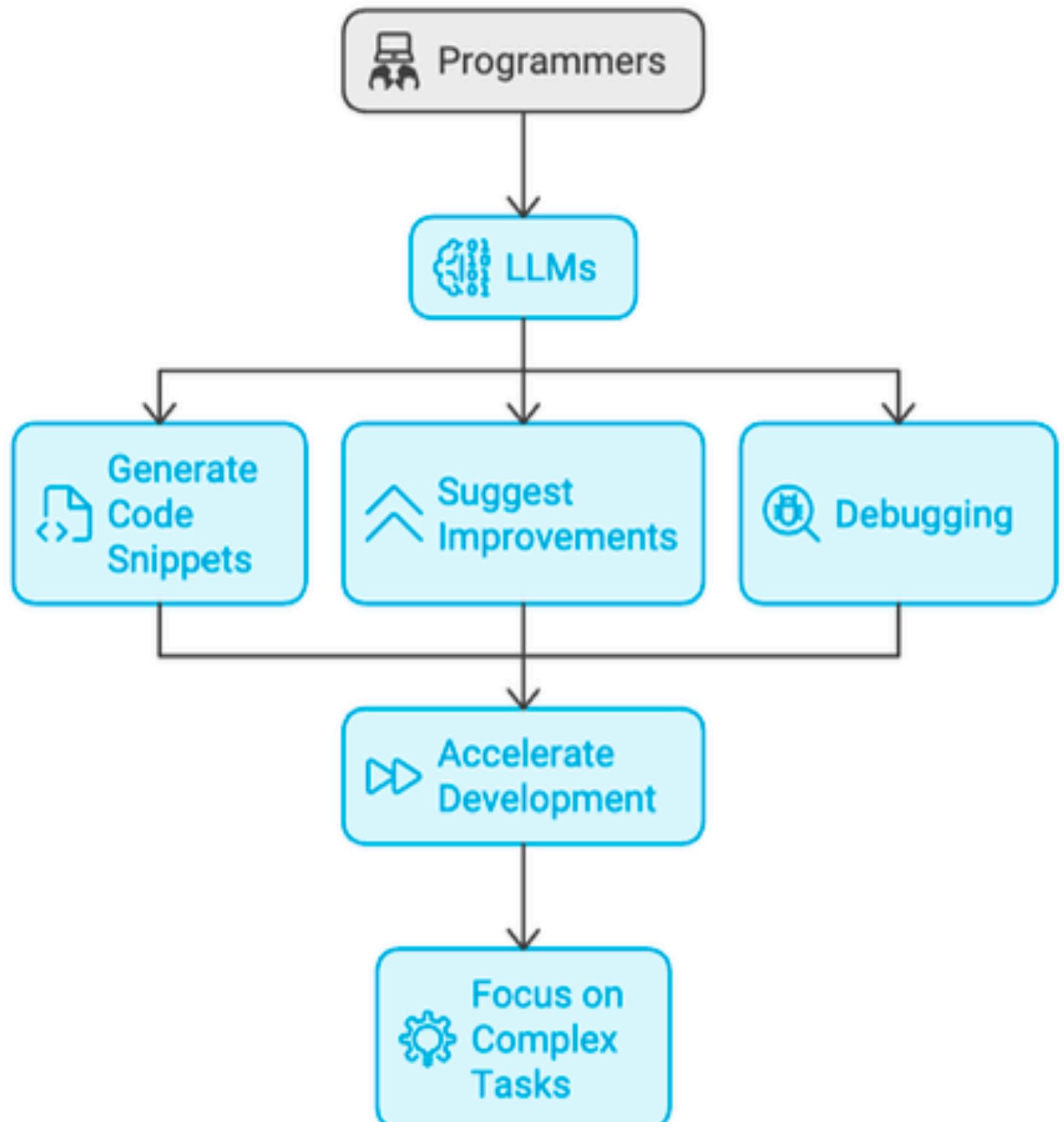


Traditional Learning

Standardized curriculum, uniform pace, teacher-led instruction.

Impact of LLMs on Programming

For Programmers, LLMs can:

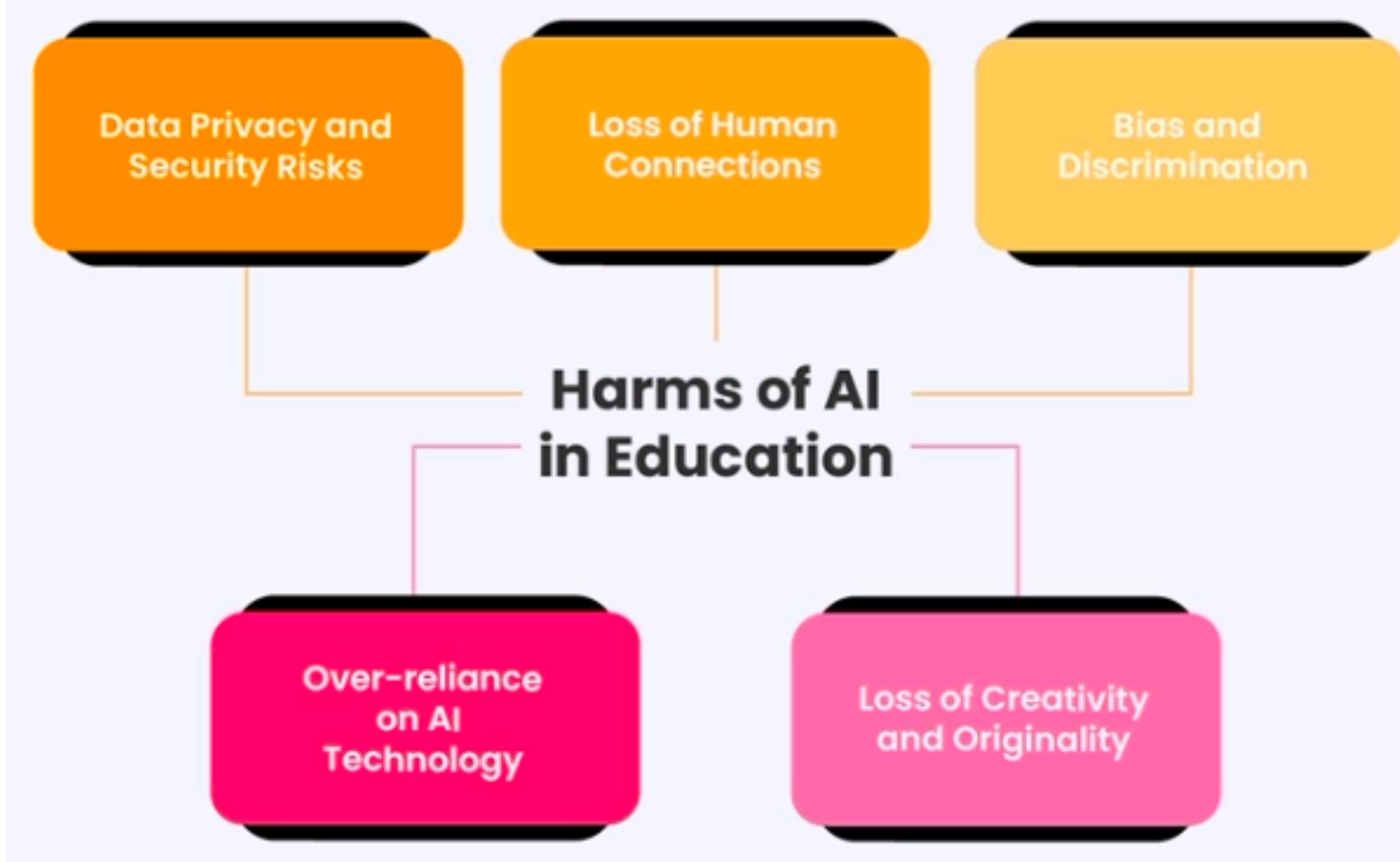


- Generate code snippets and suggest improvements
- Accelerate debugging
- Free up developers to focus on more complex tasks

GenAI Wisdom 2 (GW2)

- Coding is becoming more automatic.
- Programming with natural language is becoming reality.

Potential Negative LLM Impacts



GenAI Wisdom 3 (GW3)

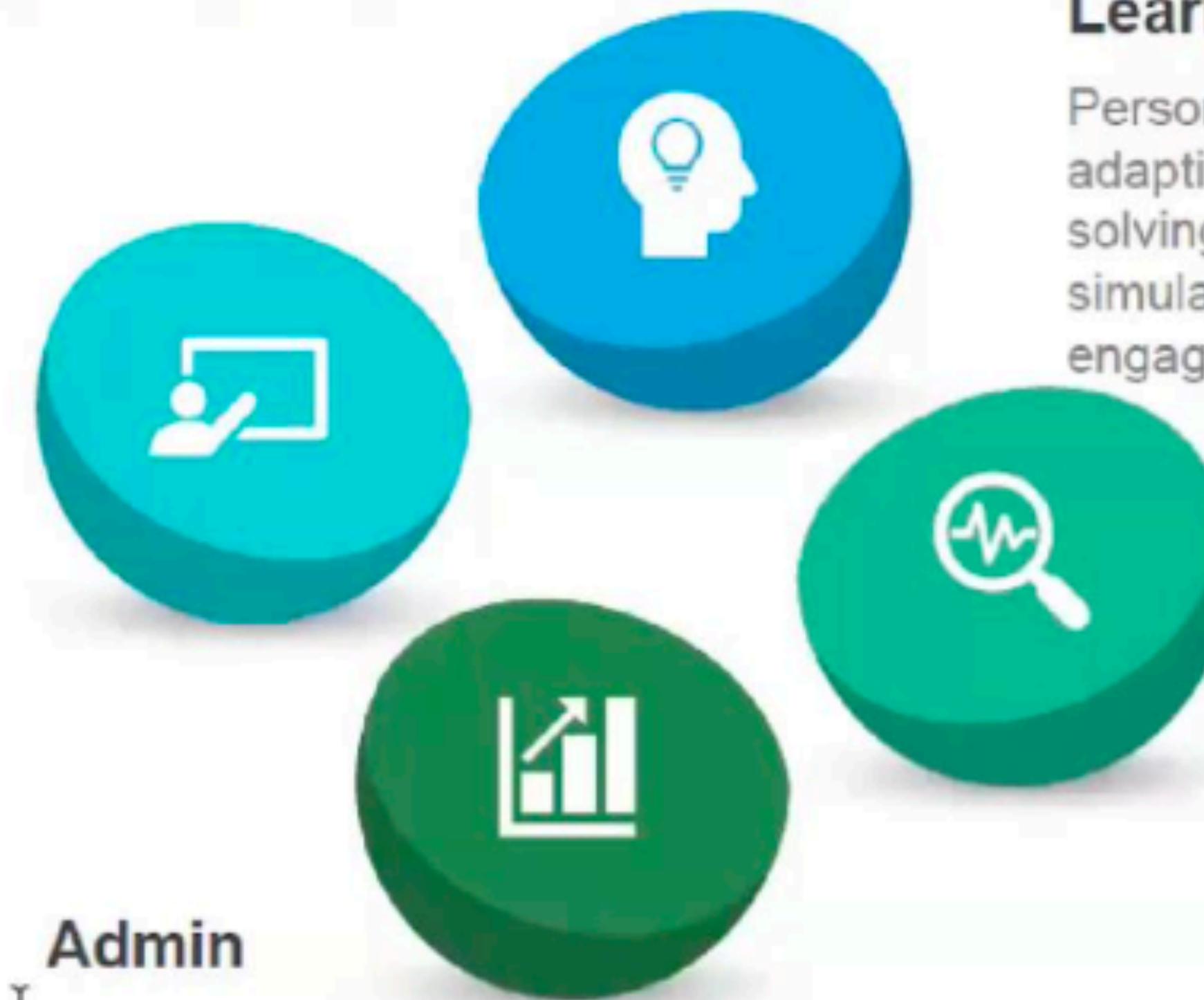
- GenAI adoption can have negative impacts.
- GenAI adoption must be managed strategically.

Four Roles of GenAI in Education

GEN AI IN HIGHER EDUCATION

Teaching

Create syllabi, exams, essays, feedback, and presentations for their courses, as well as generate images, audio, video, and text for educational content



Admin

Enrollment management, student retention, financial aid, marketing campaigns, and academic advising that optimize their efficiency and effectiveness.

Learning

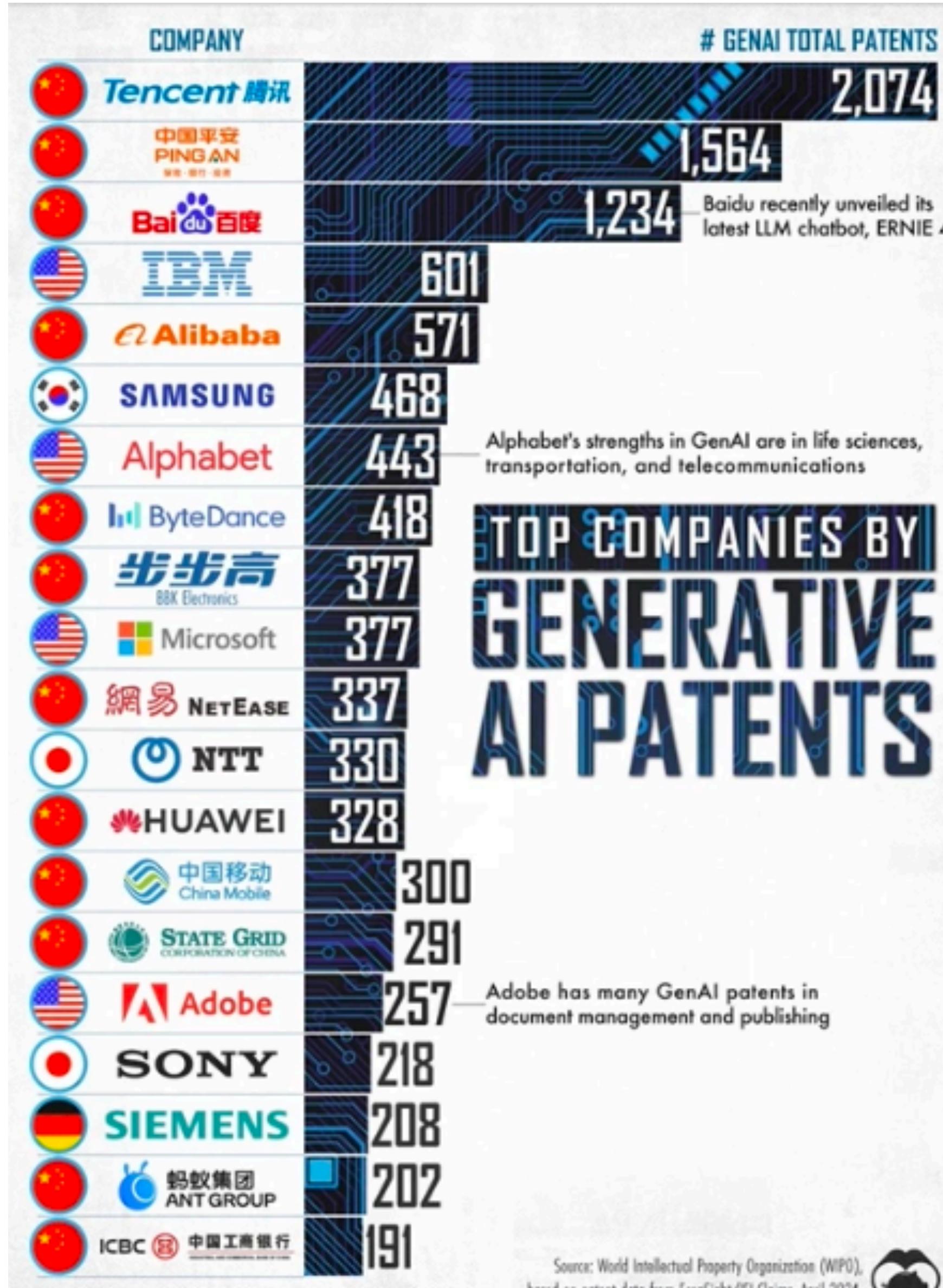
Personalized learning paths, adaptive assessments, problem-solving hints, and interactive simulations that enhance their engagement and understanding

Research

Data analysis, literature review, hypothesis generation, research paper writing, and peer review that improve their productivity and quality

Opportunities and Challenges in Large Language Model

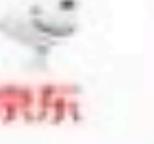
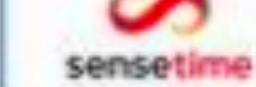
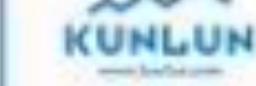
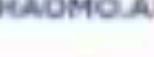
China Leads in LLM Patents



- **Invention:** LLMs were invented in the US
- **Innovation & Patents:** China is catching up and leads in LLM patents
- **Opportunity:** China will create more LLM applications
- More jobs will involve LLM development
- **Our Focus:** This course is focused on using LLM for Optimization

Many LLMs in China

Chinese Tech Firms Vie for Dominance in Large AI Models

Company	Large Model	Company	Large Model
Baidu	 Wenxin Yiyan	NetEase	 Fuxi
Alibaba	 Tongyi Qianwen, M6	JD.com	 ChatJD
Tencent	 Hunyuan	CloudWalk	 Hangye Jingling
Huawei	 Pangu	Langboat	 MChat
SenseTime	 SenseNova, SenseChat	Inspur	 Yuan 1.0
Kunlun Tech	 Tiangong 3.5	Haomo.AI	 DriveGPT
Zhihu	 Zhihaitu AI	Data Grand	 Cao Zhi
Iflytek	 Iflytek Spark	Mobvoi	 Xulie Houzi
360	 360 Zhinao		

Note: Information accurate as of May 2

Source: Public information compiled by Caixin

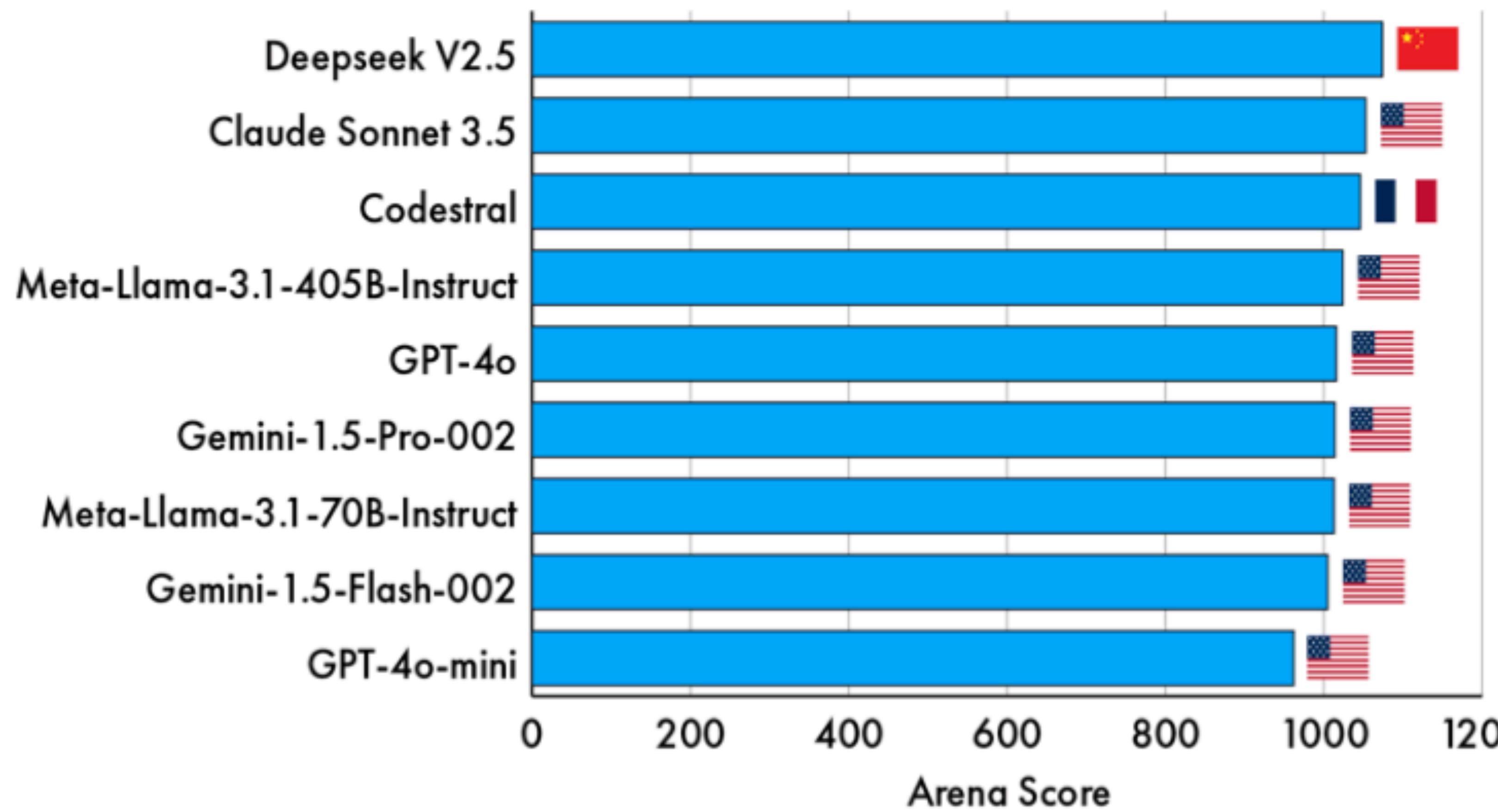
Every company tries to jump into the LLM game, and many jobs will be available as more LLM applications are created

Caixin

The Best Coding LLMs

DECEMBER 8, 2024

An initial leaderboard by Copilot Arena, which pits different AI models against each other on coding tasks, ranked DeepSeek as its top performer.



Copilot Arena is an open-source code AI coding assistant that provides paired autocomplete completions from different LLMs.

Alibaba is All In LLMs

阿里所有产品未来将接入大模型全面升级

DECEMBER 8, 2024

Integrating Alibaba's Offerings with Tongyi Qianwen



LLMs Across Industries

All industries:
LLMs can now
automate '60-70% of
employees' time'
[\[McKinsey\]](#)



Agriculture:
Decrease operating
costs by 22%
[\[ARK\]](#)

Consulting: 25.1%
faster, 40% higher
quality
[\[Harvard/BCG\]](#)



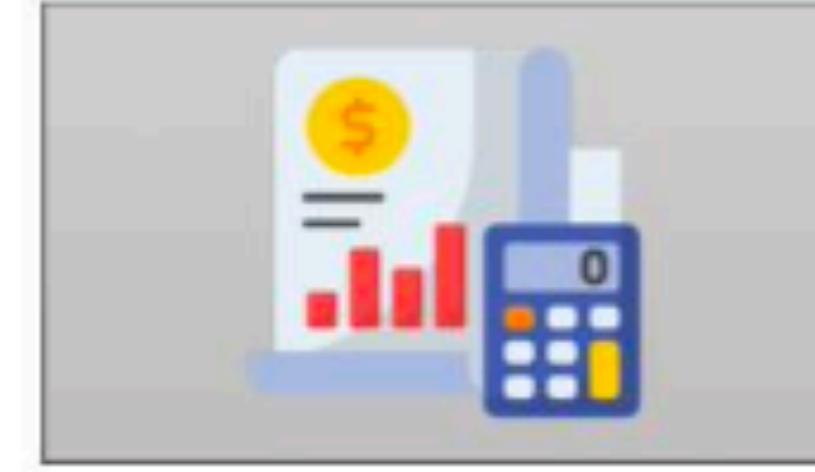
Healthcare:
58.7% better
interpersonal skills
[\[Nature\]](#)



**Software
development:**
55.8% faster
[\[Microsoft/MIT\]](#)



Medicine:
Charts 75% faster,
250% more detail
[\[Carbon Health\]](#)



**Finance, accounting,
auditing:** 'Fully
exposed,' 100% faster
[\[OpenAI/UPenn\]](#)



Therapy/coaching:
AI promises 'more
effective' therapy
[\[Seligman\]](#)

Chemistry:
'GPT-4 represents the
future of chemistry'
[\[White\]](#)



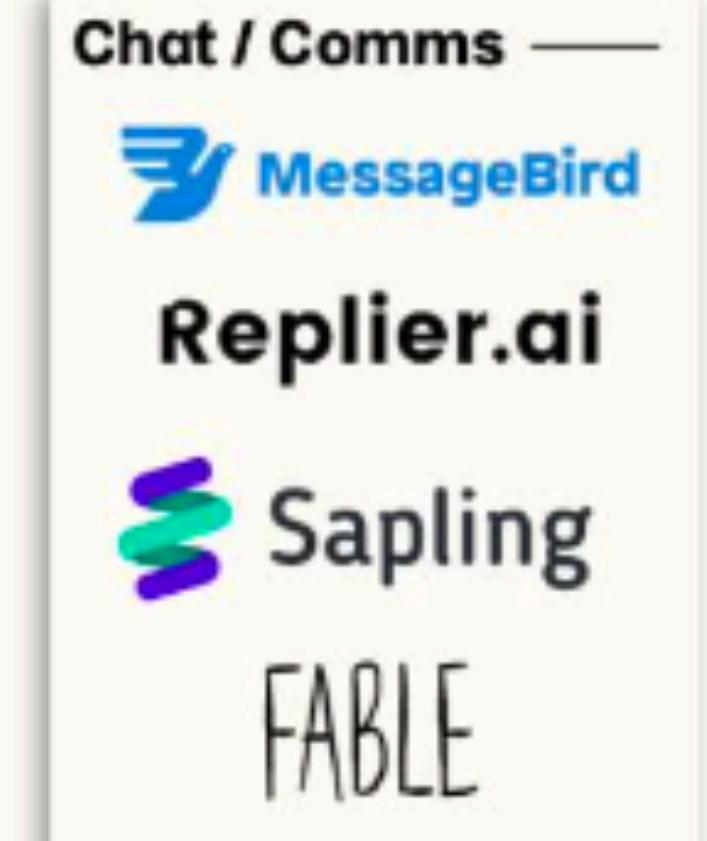
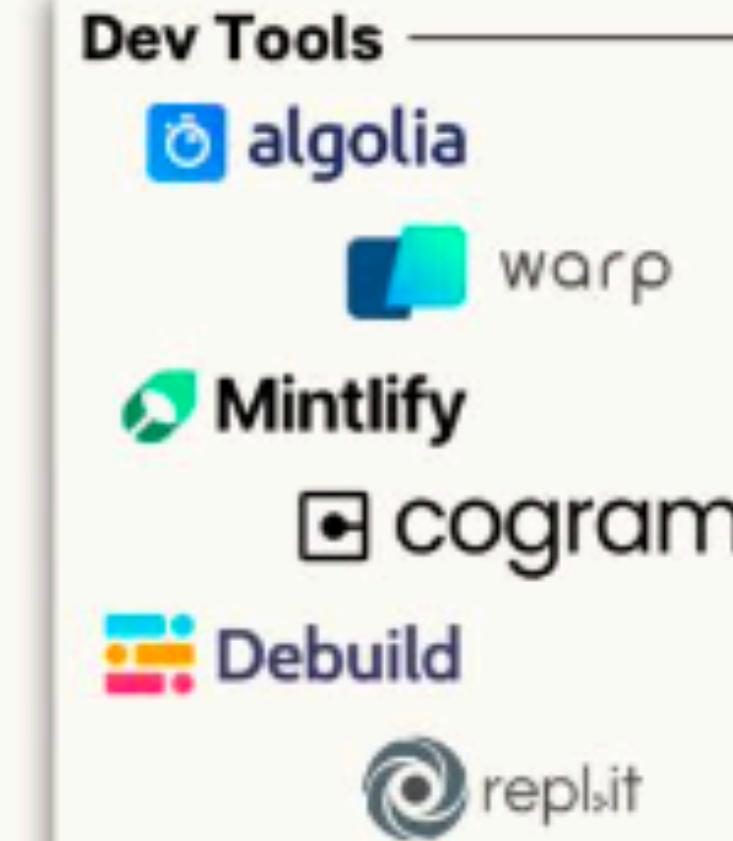
Customer service:
14% more issues
resolved
[\[McKinsey\]](#)



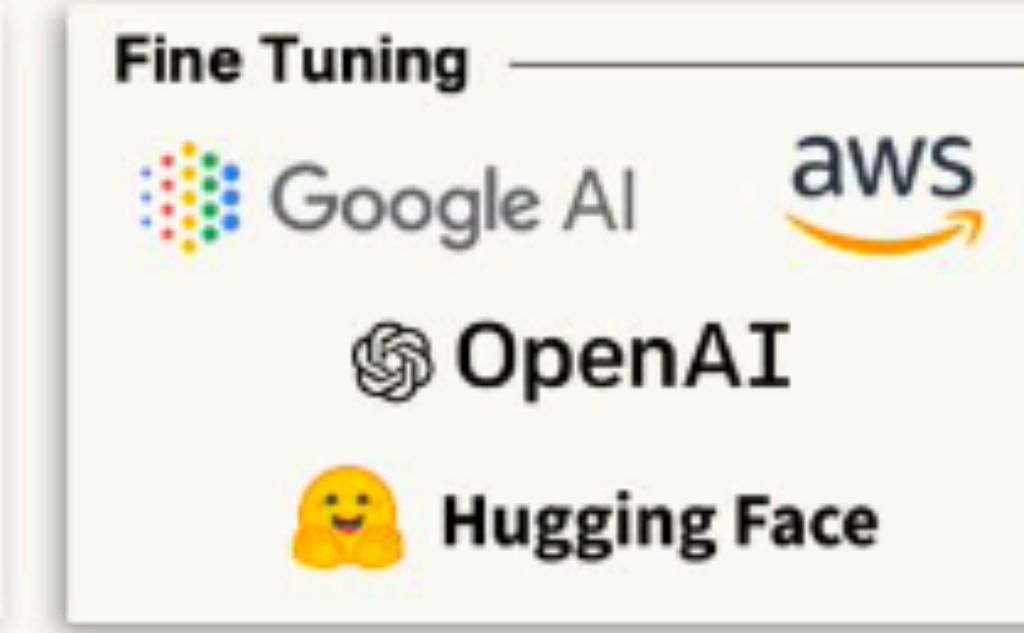
Writers: 'Fully
exposed,' 100%
faster
[\[OpenAI/UPenn\]](#)

Various Types and Uses of LLMs

Large Language Models



Infrastructure Layer



LLM Use Case Categories



DATA TRANSFORMATIONS

A data product that transforms a text input into a text output, e.g. classify, summarize, convert to JSON



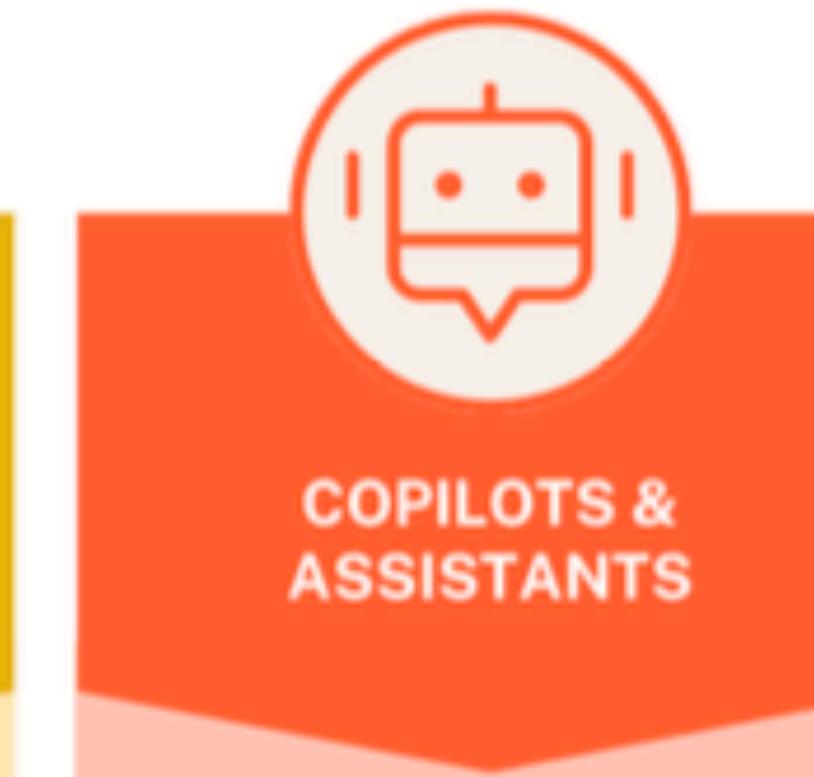
NATURAL LANGUAGE INTERFACES

A language-based interface to data or a tool, e.g. chat-your-documents, sql query



WORKFLOW AUTOMATIONS

Automate predefined workflows using access to data and tools, e.g. write a proposal, book a flight



COPILOTS & ASSISTANTS

A mixture of natural language interfaces and automation capabilities, used in the loop with a human user , e.g. Microsoft Copilot



AUTONOMOUS AGENTS

Automate arbitrary, unseen workflows using data and tools

Less complex



More complex

LLM Agents and Agentic Workflow

Data Analysis

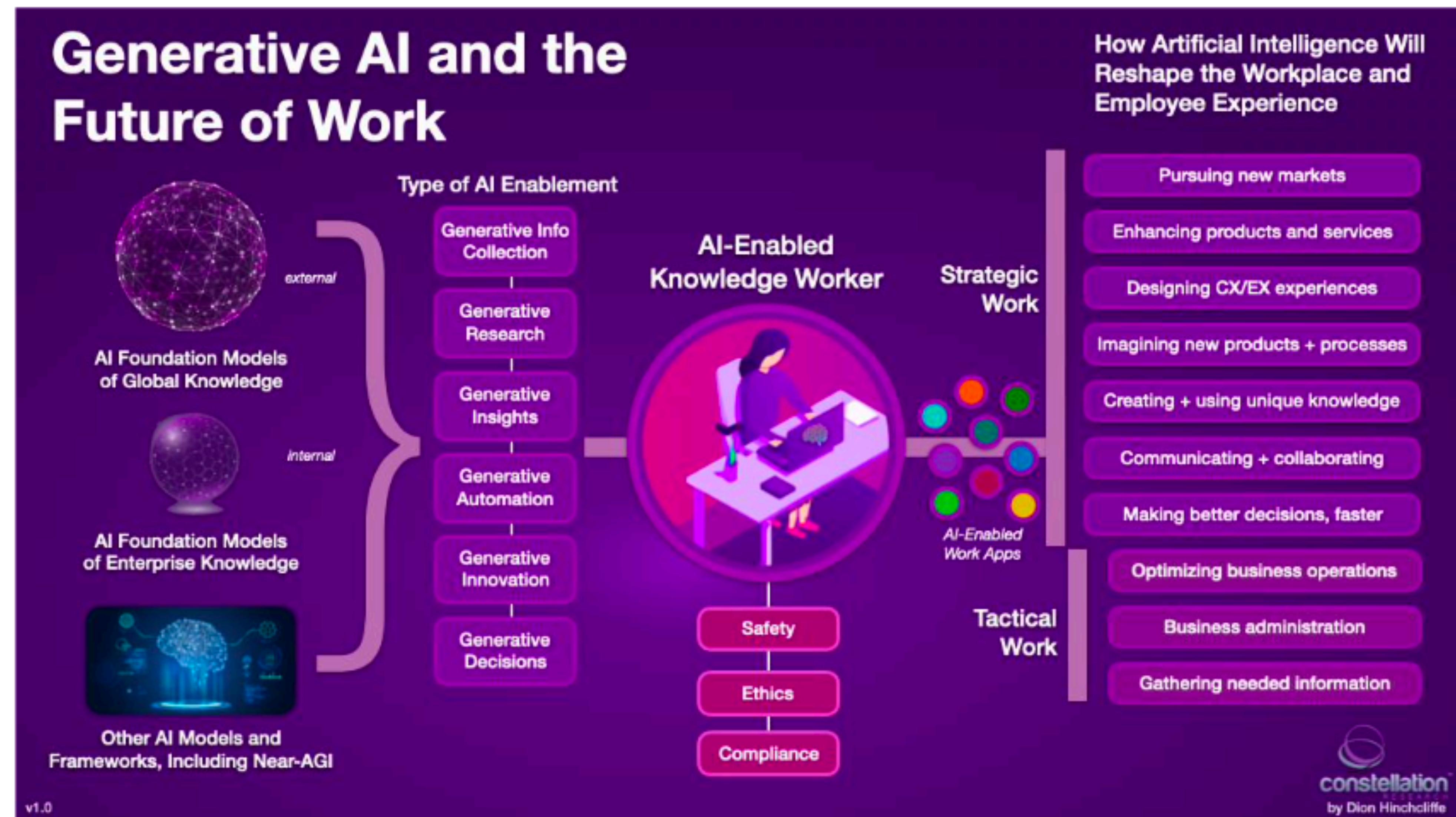


Human-Readable Summaries

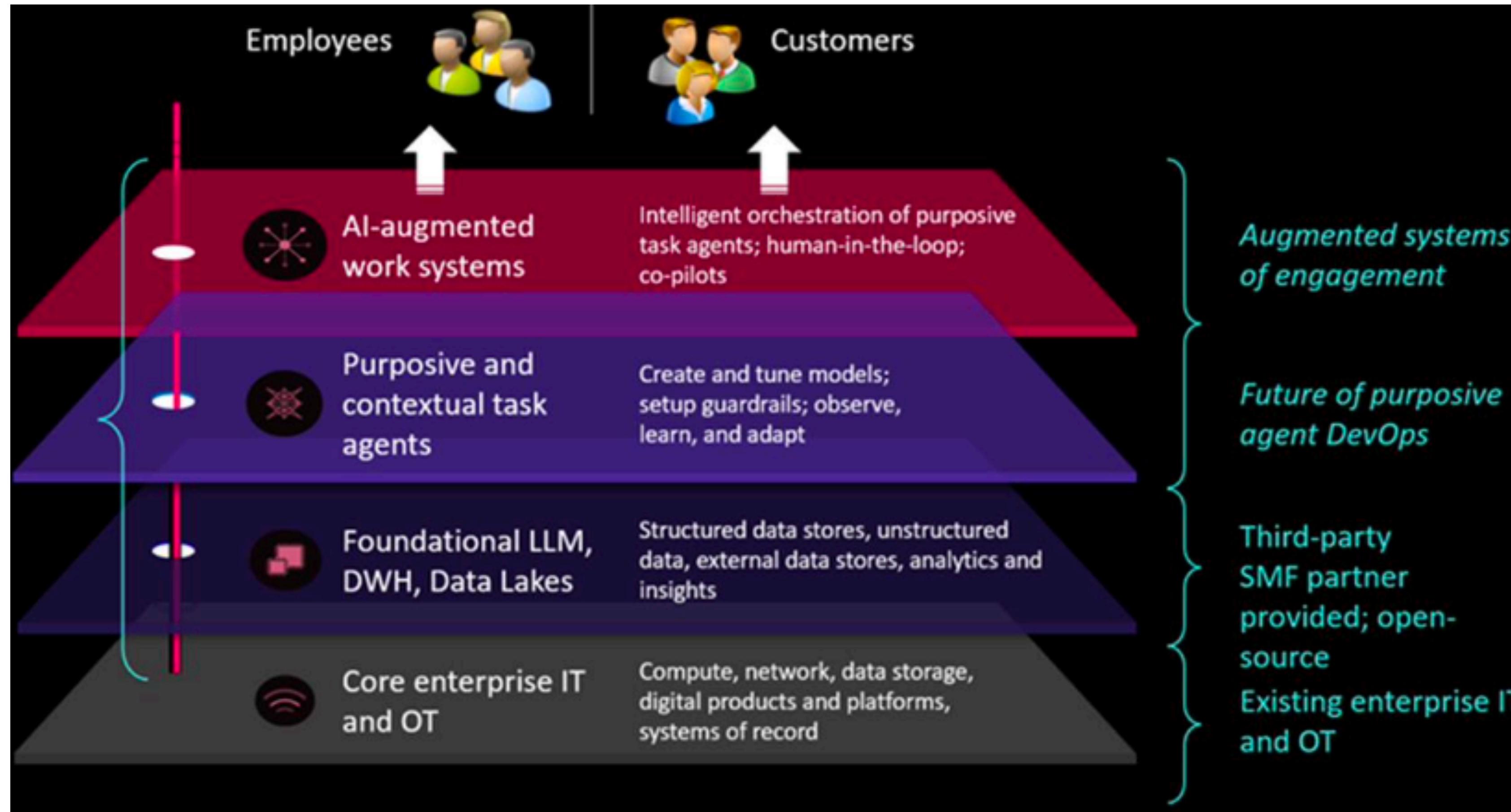
Human-Readable Summaries

Informed Decision-Making

Generative AI Is Reshaping the Workplace



AI Is Part of the Enterprise Architecture



AI-enabled enterprise architecture will support employees and customers.

LLM Job Opportunities for Students

LLM Agents and Agentic Workflow

Data Analysis



Human-Readable Summaries

Human-Readable Summaries

Informed Decision-Making