Artificial Intelligence

ECE 57000, Fall 2025
Smith Hall 108

Website: https://www.davidinouye.com/course/ece57000-fall-2025/

David I. Inouye

The Big Questions

- What is Artificial Intelligence (AI)? (Past & Present)
- What are the limits of AI? (Present & Near-Term)
- What is the future of AI? (Long-Term & Existential)

What is Artificial Intelligence?

(Past & Present)

What Is Artificial Intelligence?

- **Merriam-Webster Dictionary**

> a branch of computer science dealing with the simulation of intelligent behavior in computers

- **Oxford Dictionary**

> the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

AI as a Moving Target

In reality, the definition of AI shifts as technology advances. What was once AI is now just software.

- **Yesterday's AI (1st Wave - Handcrafted Knowledge):**
  - Examples: Chess solvers, TurboTax
  - **Good at:** Logical reasoning from rules
  - **Bad at:** Perceiving and learning from data
- **Today's AI (2nd Wave - Statistical Learning):**
  - Examples: Facial recognition, personalized recommendations
  - **Good at:** Perceiving patterns and learning from data
  - **Bad at:** Reasoning and robustness

> AI is anything that humans can do that computers cannot *yet* do.

A Third Wave: The Age of LLMs

- Today's AI frontier is increasingly defined by **Large Language Models (LLM)—e.g., ChatGPT, Gemini, Claude, Grok.**
- This "3rd Wave" builds on the foundation of the second wave.
  - The key difference is **unprecedented scale** in data and computation.
  - Because LLMs are trained on billions of human-written sentences, **they reflect human intelligence**—they do not create intelligence out of nothing.
- **Course Goal:** To teach the fundamentals of the second wave that when scaled up led to the third wave of modern advances.

What are the limits of AI?

(Present & Near-Term)

Pairs Interaction Format

- Find a nearby buddy (ideally pairs but okay with a group of 3).
- **Format of Interaction:**
  - The first person answers the question.
  - When I say **“Switch,”** the second person answers.
  - When I say **“Stop,”** return your attention to the front.
- **Rule:** If I ask for responses, you can only share what your **partner** said.
- Listen well—your partner’s answer might become your own voice!

- **First questions:**
  - What do you think AI will be able to do in the next 5-10 years?
  - What do you think AI will NOT be able to do in the next 5-10 years?

Framing the Limits: Two Mental Models

To understand AI's limitations and potential, I suggest two mental models.

1. **AI as a Tool**
2. **AI as a Book**

These models help us frame appropriate applications and anticipate risks.

Mental Model 1 — Tool 🛠️

> ## Getting Things Done (More) Efficiently

- **Concept:** AI as a tool for organizing and executing tasks (more) efficiently.
- **Historical viewpoint:**
  - Engines/machines transformed physical work.
  - LLMs are transforming knowledge work.

How should we use AI? Wise and Foolish Use of AI

Three farmers were each given a strong ox and a new plow by their master.

- The first said, ‘Oxen are too messy and unpredictable—I’ll do it by hand.’ He worked hard, but his field remained small and yielded little.
- The second said, ‘This ox can do the work for me,’ and let it run without guidance. It wandered, trampled crops, and the field became a ruin.
- But the third farmer yoked the ox, guided it with care, tilled straight rows, and prayed for rain. In the end, his field bore a rich harvest.

Question: Which of them was wise in the eyes of the master? Which farmer do you relate to most right now—and why?

Mental Model 2 — Book 📖

> **Personalized & Dynamic Knowledge**

- **Concept:** AI as a personalized, adaptive, and dynamic book.
- **Historical viewpoint:**
  - Oral cultures → Written texts → Printing press → Internet → Google Search → ChatGPT
  - Represents a new layer of interactive knowledge distillation
- **The Storyteller Twist:** This “book” is also a creative storyteller.
  - It can produce persuasive words, but they may not be true.
  - Responses are often plausible and confident but could be false.

A Note on Trust: Books vs. AI

The “Book” model raises a critical question: **When should you trust it?**

- We have heuristics for traditional media. We might trust a peer-reviewed textbook more than a website, which we trust more than an anonymous chat message.
- However, we have poor mental models for AI. An LLM can generate content that *sounds* as authoritative as a textbook but may be completely fabricated.
- Developing discernment for AI-generated content is an essential new skill.

My Hypothesis 1: The Value Judgement Limit

A fundamental limit of AI is its inability to make genuine **value judgements**.

- Value judgements are tied to **morality, lived experience, and an embodied existence**. AI lacks the physical and emotional context that shapes the human moral landscape.
- An AI can be given an objective (e.g., “maximize profit”), and it may perform the task with superhuman skill.
- However, it cannot independently determine if that objective is **right, just, or meaningful**. That requires a value system it doesn’t possess.

My Hypothesis 2: The Diminishing Returns Limit

- The exponential growth of AI will not last forever.
- We will encounter **diminishing returns**, particularly when the costs outweigh the benefits.
- As AI models become larger, the computational cost and data required to achieve smaller gains will grow exponentially.

What is the future of AI?

(Long-Term & Existential)

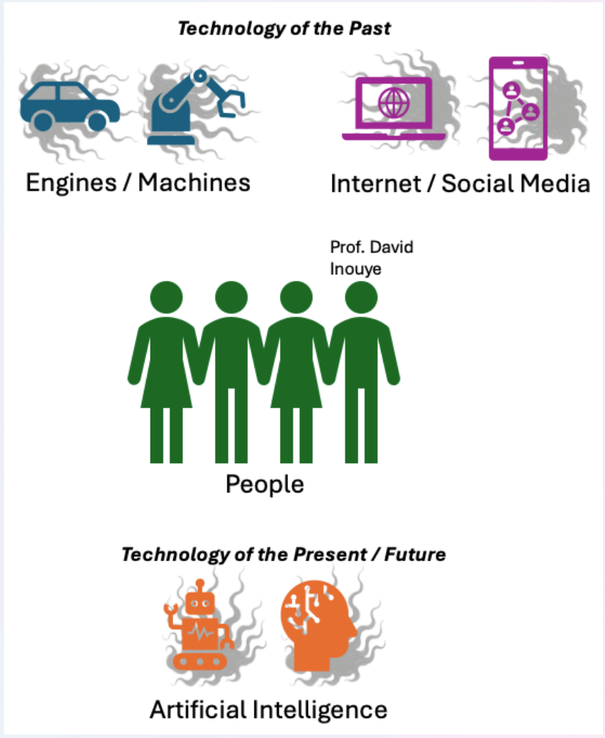
Discussion: Will AI take over or destroy the human society?

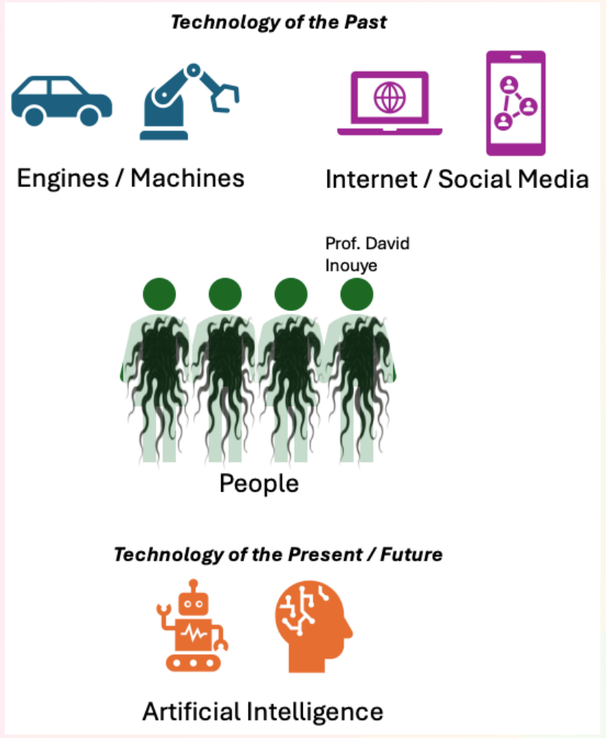
Let’s discuss with your partner.

- Will AI take over the world?
- What are the biggest dangers of AI?

Misplacing the Threat

Questions like “Will AI destroy the world?” often carry an implicit assumption.





The Common Fear: An external, superintelligent AI will emerge, and its goals will conflict with ours, leading to our demise.

Questioning the Assumption: This view supposes that evil resides *within the technology itself*, it’s external.

My personal view: The key problem of evil resides not within technology but within people including myself!

My Hypothesis: The Real Danger - Evil Within

> A good master gave each of his servants a sharp knife. One servant used it to carve furniture from wood he had gathered. He worked with care. The other servant used the knife to threaten his neighbor and steal furniture. When the master returns, will he judge the knife—or the ones who used it?

- AI is not the source of good or evil; it is a powerful **amplifier** of the intentions of its creators and users.
- The true existential threat is not the technology, but the **human heart** that wields it.
- The ultimate question is not what AI will do, but what **we** will choose to do with it.

Logistics and Syllabus

Agenda

1. Announcements
    - Emergency preparedness
2. Syllabus
3. Logistics
4. (time-permitting) Defining AI

Emergency Preparedness

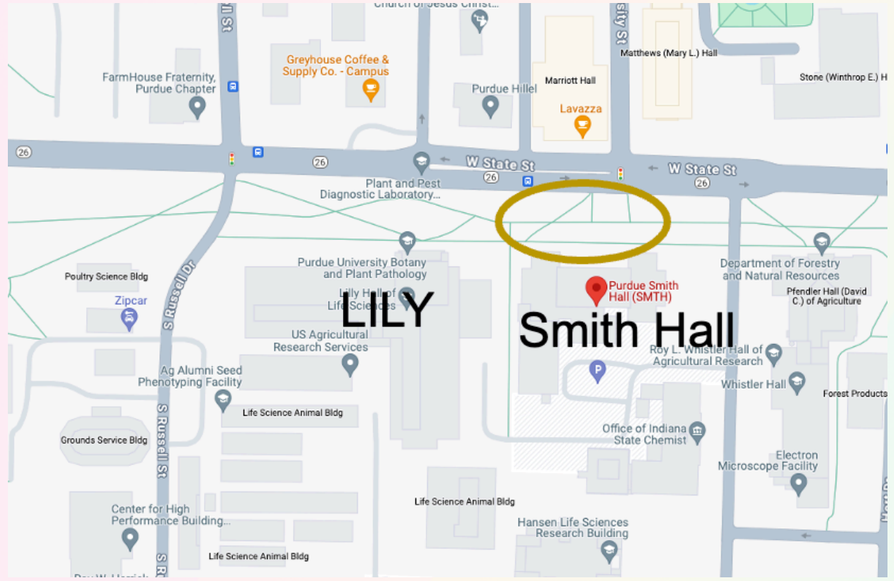
As we begin this semester, I want to take a few minutes and discuss emergency preparedness. Purdue University is a very safe campus and there is a low probability that a serious incident will occur here at Purdue. However, just as we receive a “safety briefing” each time we get on an aircraft, we want to emphasize our emergency procedures for evacuation and shelter-in-place incidents. Our preparedness will be critical IF an unexpected event occurs!

Emergency preparedness is your personal responsibility. Purdue University is actively preparing for natural disasters or human-caused incidents with the ultimate goal of maintaining a safe and secure campus.

Emergency Preparedness

Let's review the following procedure:

- For any emergency text or call 911.
- There are more than 300 Emergency Telephones (aka blue lights) throughout campus that connect directly to the Purdue Police Department (PUPD).
- If you feel threatened or need help, push the button and you will be connected right away.
- If we hear a fire alarm, we will immediately evacuate the building and proceed to the grassy area **north of Smith Hall**. Do not use the elevator.
- If we are notified of a Shelter in Place requirement for a **tornado warning, active shooter, or hazardous waste**, we will stop classroom or research activities and shelter in **this classroom**.



Map of Emergency Location

Syllabus!

All significant updates to the syllabus or course schedule will be posted on **Piazza**. See **course website** for syllabus and schedule.

Other Related Classes

- ECE 57000: Artificial Intelligence (both semesters, project-based)
- ECE 50024: Machine Learning I (fall - online, spring - in-person) by Prof. Stanley Chan and Prof. Qi Quo
- ECE 59500: Intro. to Data Mining by Prof. Jing Gao
- ECE 59500: Reinforcement Learning by Prof. Mahsa Ghasemi
- ECE 60146: Deep Learning (spring semester) by Prof. Kak and Prof. Bouman
- ECE 69500: Machine Learning in Bioinformatics and Healthcare by Prof. Joy Wang
- ECE 69500: Optimization in Deep Learning by Prof. Abolfazl Hashemi

More Technical Introduction to AI

Developing a Wise Mental Model for AI

A wise mental model is critical for effective knowledge work in the age of LLMs.

- **The New Skillset:** Knowledge workers must be able to make quick **value judgements** and rapidly assess the **value and limits** of AI tools.
- **Avoiding Extremes:**
  - Avoid **overhype** (the idea that AI can solve all problems).
  - Avoid **unfounded resistance** (e.g., overregulation because of a misunderstood existential threat).
- **Your Goal in this Class:**
  - Learn how to use AI to **enhance your learning and understanding**.
  - Avoid using it as a crutch that makes you think less or become complacent.

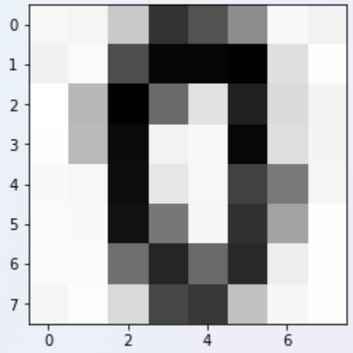
To build this model, it helps to see how AI “perceives” the world.

Computers Don’t “Think” Like We Do (Matrix)

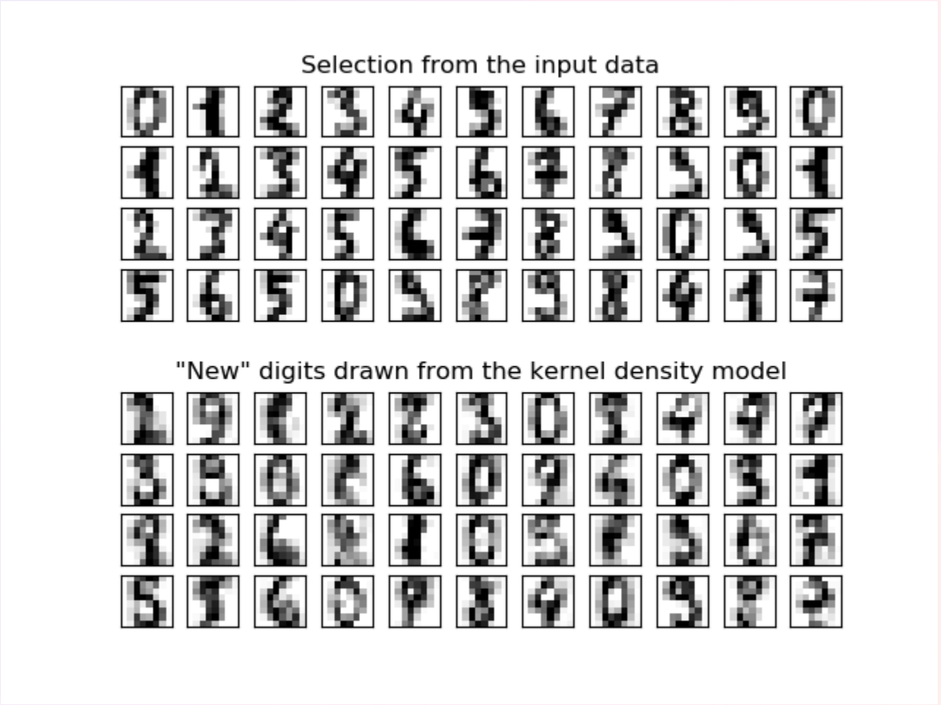
$$\begin{bmatrix} 5.49e-04 & 7.15e-04 & 5.60e-03 & 1.35e-02 & 9.42e-03 & 1.65e-03 & 4.38e-04 & 8.92e-04 \\ 9.64e-04 & 3.83e-04 & 1.38e-02 & 1.55e-02 & 1.06e-02 & 1.59e-02 & 5.07e-03 & 8.71e-05 \\ 2.02e-05 & 3.83e-03 & 1.58e-02 & 2.87e-03 & 9.79e-04 & 1.18e-02 & 8.46e-03 & 7.81e-04 \\ 1.18e-04 & 4.64e-03 & 1.21e-02 & 9.45e-04 & 5.22e-04 & 8.41e-03 & 8.26e-03 & 7.74e-04 \\ 4.56e-04 & 5.57e-03 & 8.02e-03 & 6.18e-04 & 6.12e-04 & 9.62e-03 & 8.94e-03 & 6.82e-04 \\ 3.60e-04 & 4.44e-03 & 1.17e-02 & 6.02e-05 & 1.67e-03 & 1.27e-02 & 7.21e-03 & 1.29e-04 \\ 3.15e-04 & 2.36e-03 & 1.46e-02 & 5.44e-03 & 1.10e-02 & 1.21e-02 & 2.09e-04 & 1.61e-04 \\ 6.53e-04 & 2.53e-04 & 6.47e-03 & 1.32e-02 & 1.02e-02 & 1.10e-04 & 6.56e-04 & 1.38e-04 \end{bmatrix}$$

A matrix of numbers as the computer “sees”. Do you know what this matrix represents?

Computers Don’t “Think” Like We Do (Image)



The same matrix of numbers displayed as an image.



Other examples from this dataset.

Nor Is Human Imitation Necessarily the Goal

Seeing how AI “perceives” the world helps us build better mental models. The goal isn’t to perfectly imitate human intelligence, but to understand the underlying principles of intelligence itself.

Imitation

Understanding + Engineering
(Underlying principles)

A More Pragmatic Definition

> AI is that which appears in academic conferences on AI.

The Modern AI Landscape

The field of AI is broad and rapidly evolving. Key areas include:

Core Disciplines

- **Machine Learning:** Algorithms that learn patterns from data.
- **Deep Learning:** A subset of ML using neural networks with many layers.
- **Reinforcement Learning:** Agents that learn by interacting with an environment.

Major Application Areas

- **Natural Language Processing (NLP):** Understanding and generating human language.
- **Computer Vision:** Interpreting and understanding information from images and videos.

Key Emerging Topics

- **Generative AI:** Creating new content (e.g., LLMs, diffusion models).
- **AI Safety & Alignment:** Ensuring AI systems are safe and beneficial.
- **Responsible AI:** A focus on fairness, explainability, and ethics.
- **Embodied AI & Robotics:** AI agents that can interact with the physical world.

This Course Will Only Cover a Small Set of Topics

1. Introduction to artificial intelligence
2. Machine learning basics (e.g., optimization, gradient descent, regularization)
3. Deep learning basics (e.g., MLPs, CNNs)
4. Natural language processing basics (e.g., RNNs, Transformers)
5. Dimensionality reduction (e.g., PCA, VAE)
6. Generative models (e.g., VAE, Diffusion models)
7. Markov decision processes (e.g., basics of reinforcement learning)
8. Special topics

Questions?