



Today's outline

- What's in CS540
- Course logistics
- Software
- Probability



Part I: Course overview

What is artificial intelligence?

Language models

ChatGPT 5

Give me a very short introduction to AI, to begin a lecture to college students.

Here's a concise way you could open your lecture:

"Artificial Intelligence, or AI, is the science of creating machines that can perform tasks requiring human-like intelligence—things like recognizing patterns, learning from data, making decisions, and even generating new ideas. At its core, AI is about building systems that can adapt and improve, not just follow fixed instructions. It powers technologies we use every day—from recommendation systems on streaming platforms to voice assistants and self-driving cars—and continues to shape how we live, work, and think about the future."

Would you like me to make it more **engaging and conversational** (to capture attention at the very start), or more **formal and academic** (to set a scholarly tone)?

+

Ask anything

ChatGPT can make mistakes. Check important info.

DeepMind and OpenAI achieve IMO Gold. What does it all mean?

What we know, what we would like to know, and what it may take years to know

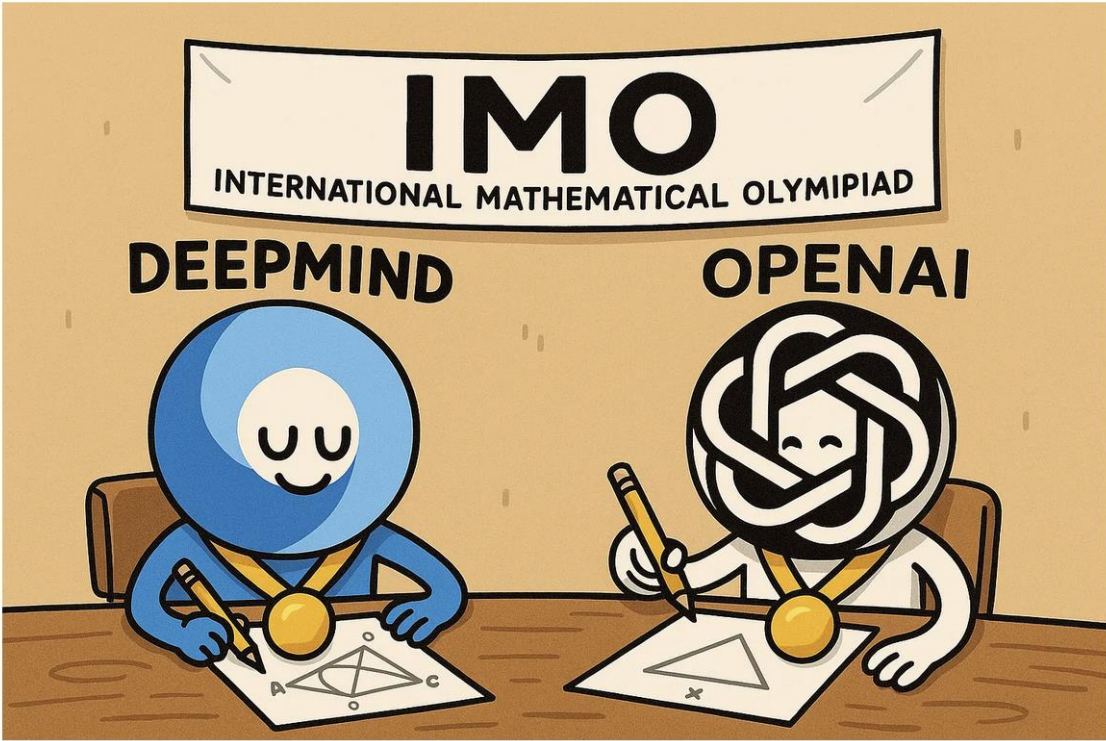
ERNEST DAVIS AND GARY MARCUS
JUL 22, 2025

175

65

24

Share



Drawing (and typo) by ChatGPT

ADVENTURES IN AUGMENTATION

Google CEO says over 25% of new Google code is generated by AI

We've always used tools to build new tools, and developers are using AI to continue that tradition.

BENJ EDWARDS – OCT 30, 2024 11:50 AM | 145

Conversation on 8/7/25

<https://garymarcus.substack.com/p/deepmind-and-openai-achieve-imo-gold>
<https://arstechnica.com/ai/2024/10/google-ceo-says-over-25-of-new-google-code-is-generated-by-ai/>

Autonomous driving



<https://www.sfchronicle.com/sf/article/cruise-waymo-driverless-cars-in-s-f-18282902.php>

Image and Video Generation



Bioengineering

The Nobel Prize in Chemistry 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Chemistry 2024 with one half to

David Baker

University of Washington, Seattle, WA, USA
Howard Hughes Medical Institute, USA.

“for computational protein design”

Demis Hassabis

Google DeepMind, London, UK

“for protein structure prediction”

John Jumper

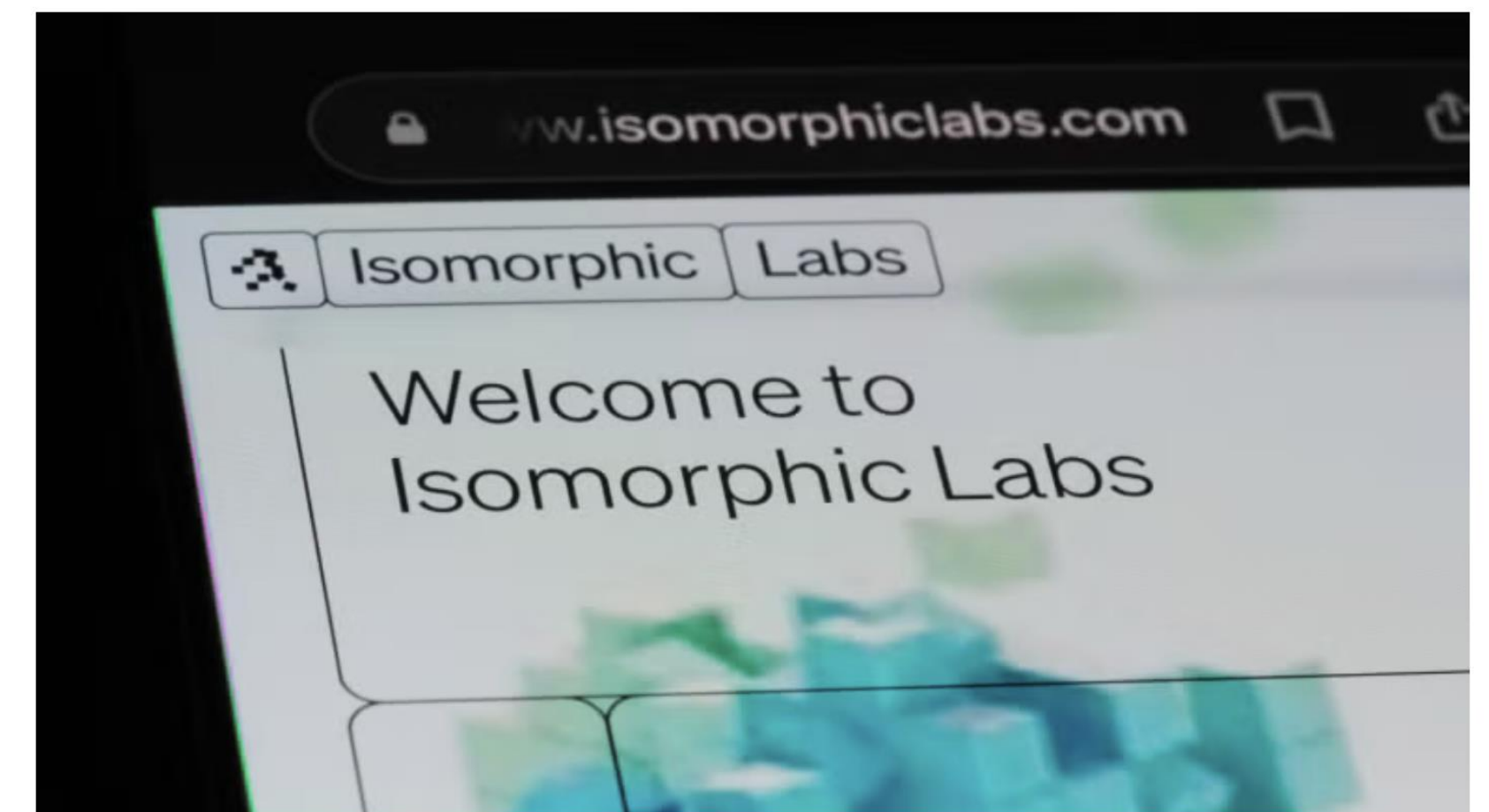
Google DeepMind, London, UK

Isomorphic Labs prepares to launch trials for AI-designed drugs

The Google DeepMind spinout raised \$600m in its first financing round in March 2025.

Abigail Beaney | July 7, 2025

Share



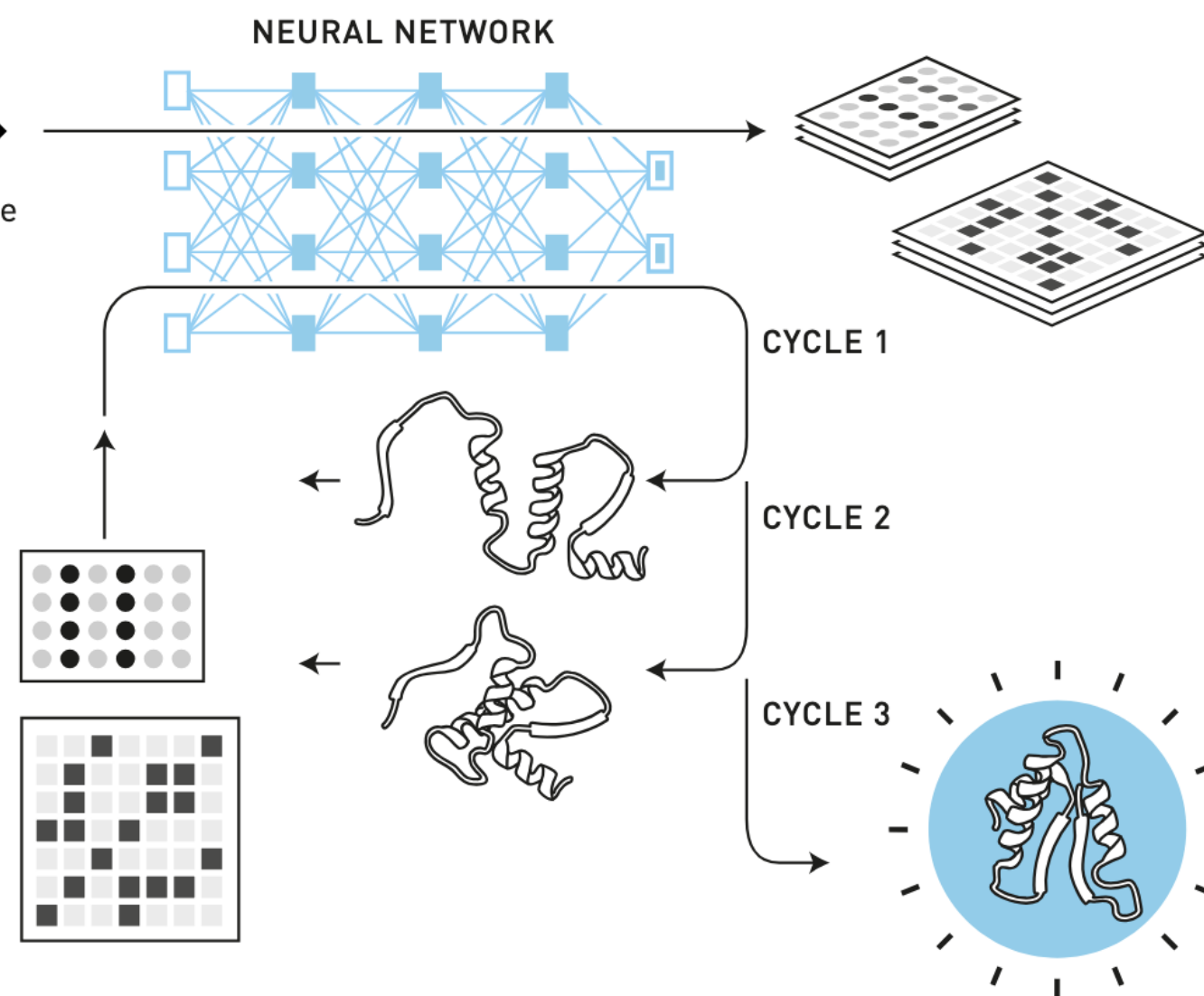
Isomorphic Lab's AlphaFold 3 system, developed with Google DeepMind, won its developers a Nobel Prize. Image credit: Shutterstock / Robert Way

3. AI ANALYSIS

Using an iterative process, AlphaFold2 refines the sequence analysis and distance map. The AI model uses neural networks called transformers, which have a great capacity to identify important elements to focus on. Data about other protein structures – if they were found in step 1 – is also utilised.

4. HYPOTHETICAL STRUCTURE

AlphaFold2 puts together a puzzle of all the amino acids and tests pathways to produce a hypothetical protein structure. This is re-run through step 3. After three cycles, AlphaFold2 arrives at a particular structure. The AI model calculates the probability that different parts of this structure correspond to reality.



<https://www.nobelprize.org/prizes/chemistry/2024/press-release/>

<https://www.clinicaltrialsarena.com/news/isomorphic-labs-prepares-trials-ai-designed-drugs/>

Playing Games

AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol

DeepMind's artificial intelligence astonishes fans to defeat human opponent and offers evidence computer software has mastered a major challenge



📷 The world's top Go player, Lee Sedol, lost the final game of the Google DeepMind challenge match. Photograph: Yonhap/Reuters

Hold 'Em or Fold 'Em? This A.I. Bluffs With the Best

Pluribus, a poker-playing algorithm, can beat the world's top human players, proving that machines, too, can master our mind games.

FUTURE PERFECT

AI triumphs against the world's top pro team in strategy game Dota 2

It's the first time an AI has beat a world champion e-sports team.

by Kelsey Piper

Apr 13, 2019, 10:30 PM UTC



<https://www.vox.com/2019/4/13/18309418/open-ai-dota-triumph-og>

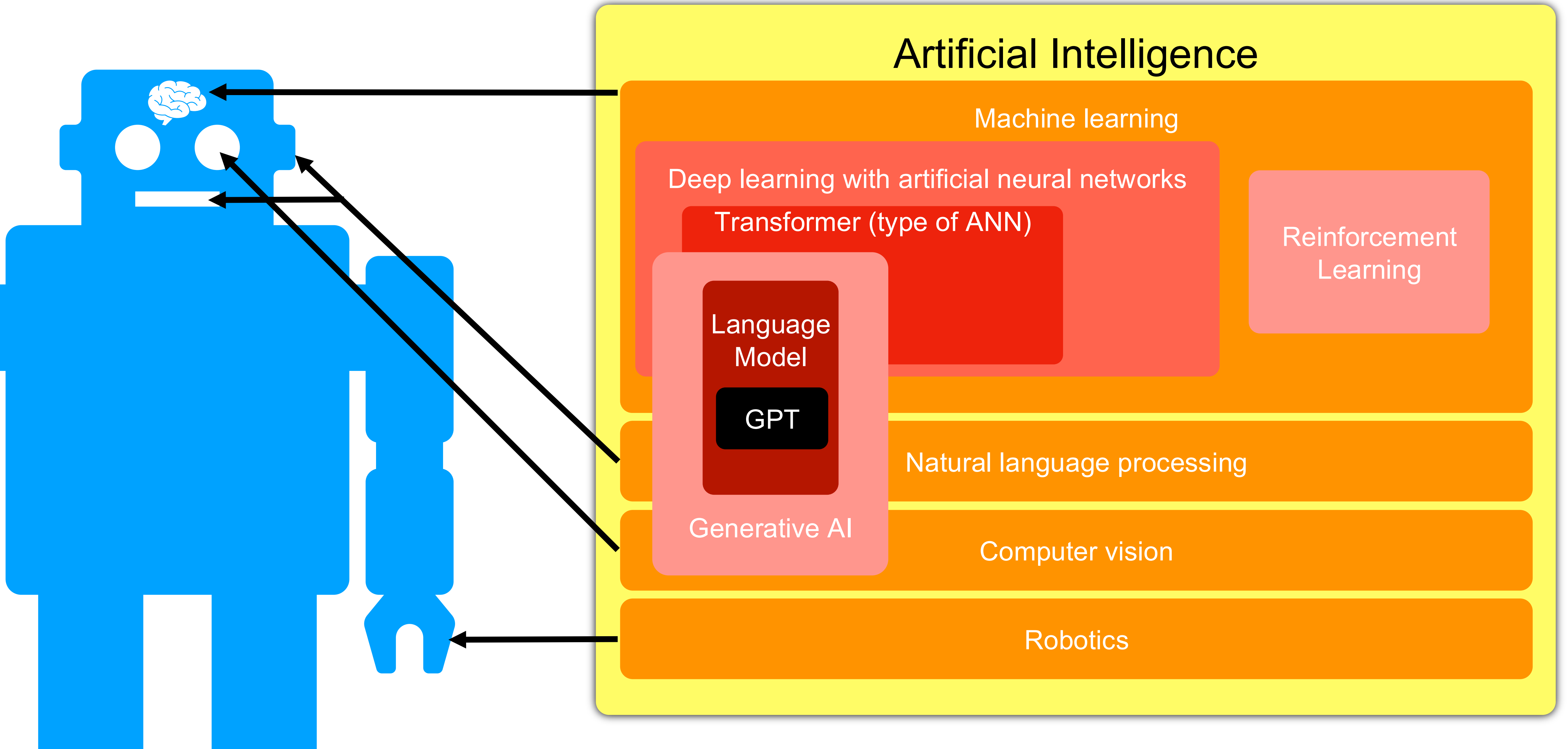
<https://www.nytimes.com/2019/07/11/science/poker-robot-ai-artificial-intelligence.html>

<https://www.theguardian.com/technology/2016/mar/15/googles-alphago-seals-4-1-victory-over-grandmaster-lee-sedol>

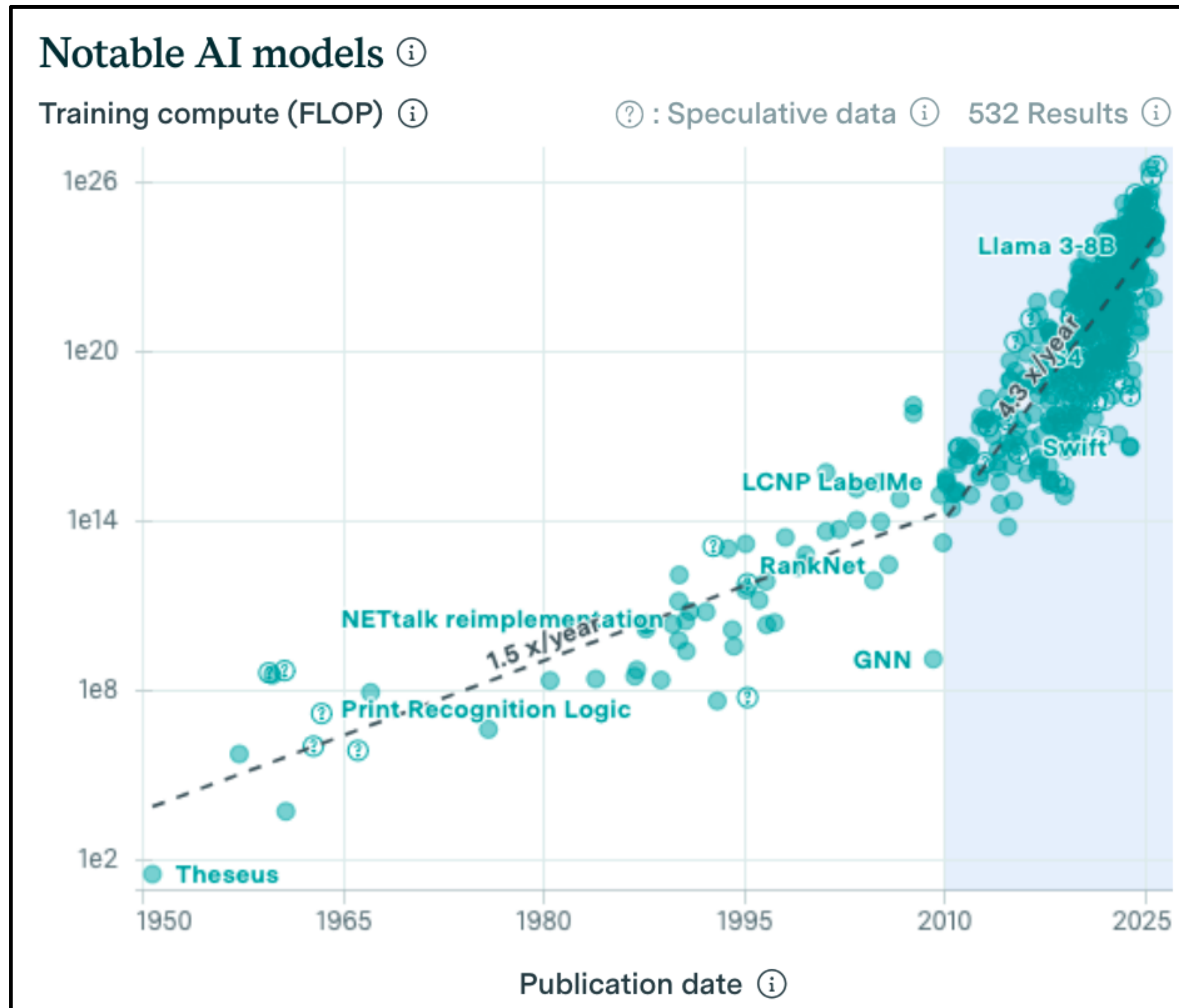
So what exactly is AI?

GPT, large language model, generative AI, deep learning...

Artificial Intelligence Is a Broad Topic



Artificial Intelligence Is Rapidly Changing



Standard textbook on AI is Russell & Norvig's
Artificial Intelligence: A Modern Approach

- First edition in 1995
- Fourth edition in 2020
- Third edition from 2010 devotes just 10 pages to neural networks!

Artificial Intelligence Is Not Magic

It relies on **fundamental** techniques in:

- Algorithms
- Mathematics
- Logic
- Probability and Statistics
- Optimization

What Will You Learn in CS540?

- Foundational mathematical tools
 - Linear algebra, probability, statistics
- Core techniques in natural language processing
 - Language modeling, n-gram models, transformers
- Basics of machine learning
 - Classification, regression, clustering
- Neural networks and deep learning
 - Multilayer perceptron, convolutions, ResNets
- Search and reinforcement learning
 - Game theory, Q-learning, A*
- Ethics in real-world applications

What Will You Learn in CS540?

- Foundational mathematical tools
 - Linear algebra, probability, statistics
- Core techniques in natural language processing
 - Language modeling, n-gram models, transformers
- Basics of machine learning

TL;DR Lots of useful stuff, theory and practice in AI

- Neural networks and deep learning
 - Multilayer perceptron, convolutions, ResNets
- Search and reinforcement learning
 - Game theory, Q-learning, A*
- Ethics in real-world applications

What Will You *Not* Learn in CS540?

This course is *Introduction to AI*. That means a lot of the shiniest stuff is out of scope.

✗ Generative AI & Foundation Models

No in-depth coverage of diffusion models, image generators, or foundation-model training.

✗ Building and operating LLMs

We will discuss the foundations of language modeling and transformers, but not how to build or train them, how to write prompts, or build apps for commercial LLM APIs.

✗ AI Systems

No distributed training, massive datasets, GPU clusters, or the engineering required to train state of the art models.

Our goal: Understand the foundations.

AI techniques change every few months. The math behind them doesn't.

What courses should I take to learn about ChatGPT/LLMs/<insert favorite buzz word>?

Foundations

CS540: Introduction to AI

CS760: Machine Learning

Specific topics

CS762: Deep Learning

CS766: Computer vision

CS769: Natural language processing

CS839: Reinforcement learning

CS839: Foundation models



Part II: Course Logistics

Where to find everything?

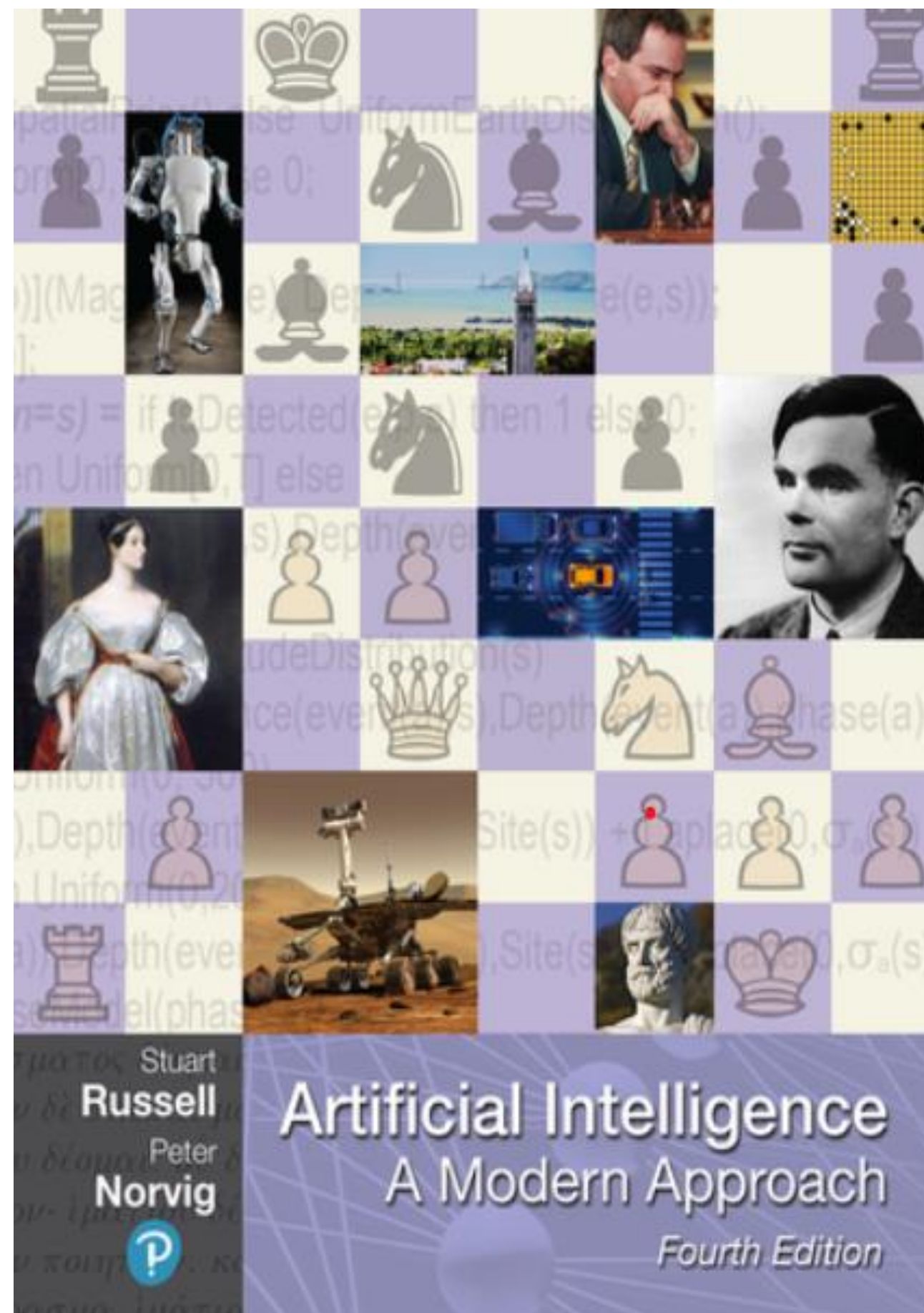
- **Canvas** - [Pointer to everything](#)
 - grades, other private materials that should not be shared
- **Course website** - public materials
 - <https://pages.cs.wisc.edu/~khodak/cs540fall2026>
 - Slides, schedule, policies
- **Piazza**
 - <https://piazza.com/wisc/fall2026/cs540sections12/home>
 - Access code on Canvas
 - Discussion, questions, announcements

Textbook

Artificial Intelligence: A Modern Approach (4th edition).

Stuart Russell and Peter Norvig. Pearson, 2020. ISBN 978-0134610993.

(textbook is optional, but may be useful)



Instruction Team

(See course webpage)

2 Principal Instructors: Kirthi Kandasamy (Section 1), Misha Khodak (Section 2)

Merged across 2 sections:

- 5 Teaching Assistants (TAs): hold office hours, grade your homework
- Grader: grade your homework
- 3 Peer Mentors: hold office hours

Contacting the instructors/TAs

This is a large class, so please use your judgment before emailing the instructors/TAs.

- Please read course webpage and Piazza before you ask us.
- If you choose to contact us
 - OHs > Public Piazza post > Private Piazza post > Email.
 - If you think others will benefit from the response, ask as a public Piazza post. We will prioritize public posts when responding.

Lectures

Section 1: Tue/Thu 11am to 12.15pm CHEM S413

Section 2: Tue/Thu 1pm to 2.15pm Mosse 2650

- Please attend your assigned lecture.
- Lectures will be on slides. Complete slides will be made available prior to class, but are not a substitute for attending lectures.

Office Hours

- Available on the course website
- Instructor OHs
 - Kirthi Kandasamy: Tue/Thu 12.30 - 1.30pm in MH5506 (Office may change)
 - Misha Khodak: Tue/Wed 11-12 in MH5512
- TAs' and PMs' office hours *generally* in **Huddle Room B2580**.
 - See webpage for times.
- All office hours are **merged** across sections, you can go to anyone
- Use TA and Peer Mentor hours for detailed-level questions (e.g. coding related) and use instructor office hours for conceptual questions.

Grading scheme

- **Homework Assignments: 30%**
- **Midterm Exam: 30%**
- **Final Exam: 40%**

Homeworks

- Ten assignments, submit via **Gradescope by 11:59 pm on due date**
- No deadline extensions (plan to submit well ahead of the deadline).
But, *two* lowest homework scores dropped
- Collaboration with other students: You are allowed (and encouraged) to collaborate with other students. But you must write up your own solution.
- Regrade requests: Use Google Form (will be announced) for regrade request.

Homeworks – Policy on AI tools

- You are allowed to use AI tools, however, if you use them to generate entire solutions, you will be very unprepared for your midterm and final.
- Recommended way to use AI tools.
 - Always attempt the problem on your own first. After completing a solution, you may use an LLM to check your work.
 - If you are stuck after substantial effort (at least several hours), you may use an LLM for high-level guidance (e.g., how to approach the problem).
 - If you are still stuck (after a day or two), you may generate a solution using an LLM, but you must understand it fully, verify its correctness, and write the solution in your own words. **Use this as a last resort.**

Exams

- **Midterm:** Will likely fall between 10/19 and 10/30.
 - Will inform you once we receive confirmation from campus.
- **Final:** 12/12/2026, Saturday (12:25PM - 2:25PM)
- **Make-up exams**
 - Only for documented emergencies and with explicit permission from the instructor.
 - Otherwise, up to a 33% penalty.

Integrity



Academic Integrity

Cheating and plagiarism will be dealt with in accordance with University procedures ([UW–Madison Academic Misconduct Rules and Procedures](#)).

Collaboration with other students on homework is encouraged: you may work through the assignments together, compare answers, and debug each other's code. However, each student must write up and submit their own solution, and all examinations must be completed individually.

Use of AI tools and other resources on homework is allowed: you may use artificial intelligence tools (such as ChatGPT, Claude, Gemini, or Cursor) as well as books or online resources on homework without restriction. Note that the exams are closed-book and worth 70% of your grade, so the homework is your main opportunity to learn the material.

Declare your collaborators and your use of AI or other resources at the top of each homework solution: "I discussed Q2–Q3 with X and Y; I used Claude to debug the code in Q4; I found a hint for Q1 on StackOverflow." There is no penalty for anything you declare here. Failing to declare it is an academic integrity violation.

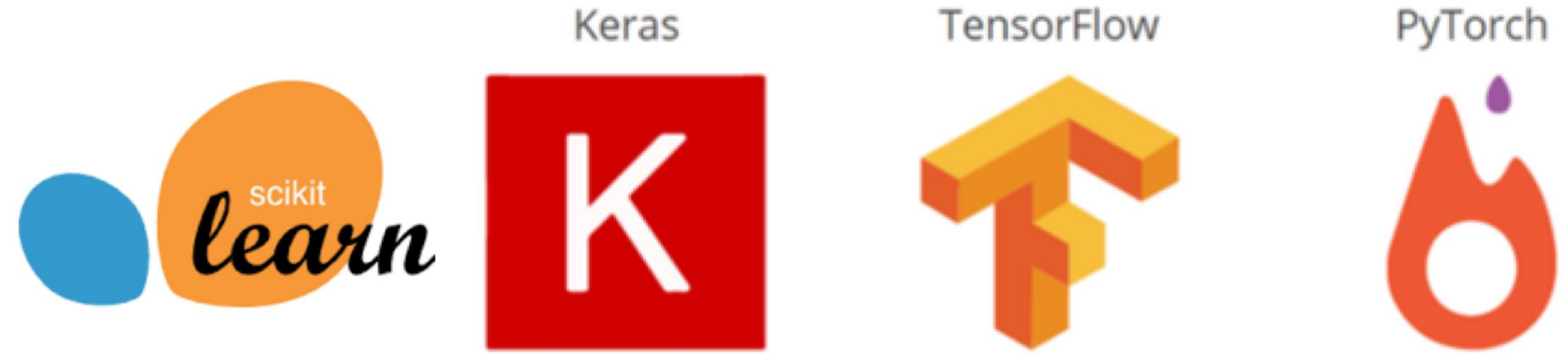
Student responsibilities and expectations

- This will be a challenging course, designed for advanced undergraduate and early graduate students with mathematical maturity.
- You are expected to have a background in probability, statistics, and linear algebra.
 - We will devote the first few lectures to review these topics, but it is your responsibility to be prepared for the course.
 - You can still complete this course successfully without this background. **But be prepared to work hard.**
- You should also have a background in programming (we will use Python for our programming assignments).



Part III: Software

Tools



- **Python**

- Main tool for machine learning & data science
- Conda package manager (for simplicity)

- **Jupyter**

- Sometimes easier to keep track of your experiments
- Obviously, you should put longer code into modules



Part IV: Probability

What is probability?

- Language to express **uncertainty**



In AI/ML Context

- Quantify predictions

$$[p(\text{lion}), p(\text{tiger})] = [0.98, 0.02]$$



$$[p(\text{lion}), p(\text{tiger})] = [0.01, 0.99]$$



$$[p(\text{lion}), p(\text{tiger})] = [0.43, 0.57]$$

* If we know for sure the photo must contain either a lion or a tiger

Model Data Generation

- Model complex distributions



StyleGAN2 (Kerras et al '20)

Basics: Outcomes & Events

- **Outcomes:** possible results of an **experiment**

$$\Omega = \underbrace{\{1, 2, 3, 4, 5, 6\}}_{\text{outcomes}}$$

- **Events:** subsets of outcomes we're interested in

$$\underbrace{\emptyset, \{1\}, \{2\}, \dots, \{1, 2\}, \dots, \Omega}_{\text{events}}$$

- Always include $\emptyset, \Omega$



Basics: Probability Distribution

- We have outcomes and events
- Assign **probabilities**: for each event E , $P(E) \in [0,1]$
- Back to our example

$$\underbrace{\emptyset, \{1\}, \{2\}, \dots, \{1, 2\}, \dots, \Omega}_{\text{events}}$$

$$P(\{1\}) = \frac{1}{6}, \quad P(\{1,2,3\}) = \frac{3}{6} = \frac{1}{2}$$



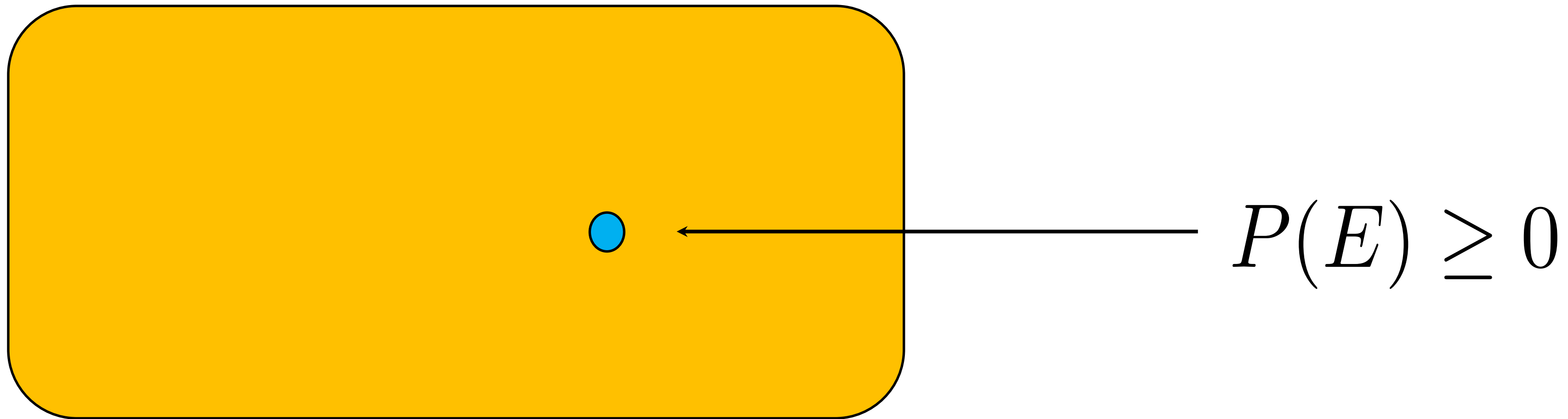
Basics: **Axioms**

- Rules for probability:
 - For all events E , $P(E) \geq 0$
 - Always, $P(\emptyset) = 0, P(\Omega) = 1$
 - For disjoint events, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$
- Easy to derive other laws. Ex: non-disjoint events

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Visualizing the Axioms: I

- Axiom 1: for all events E , $P(E) \geq 0$



Visualizing the Axioms: II

- Axiom 2: $P(\emptyset) = 0, P(\Omega) = 1$

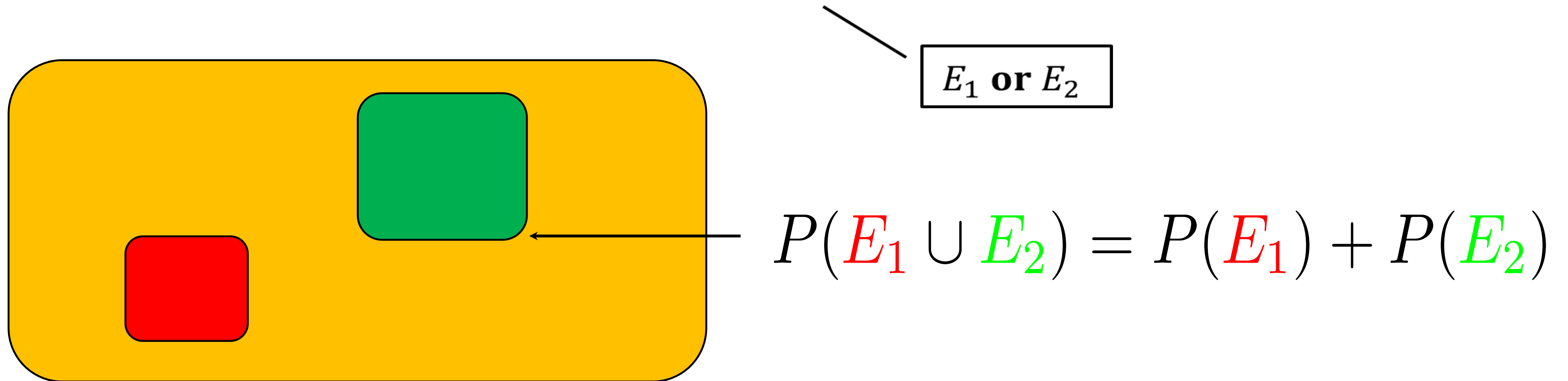


$$P(\Omega) = 1$$



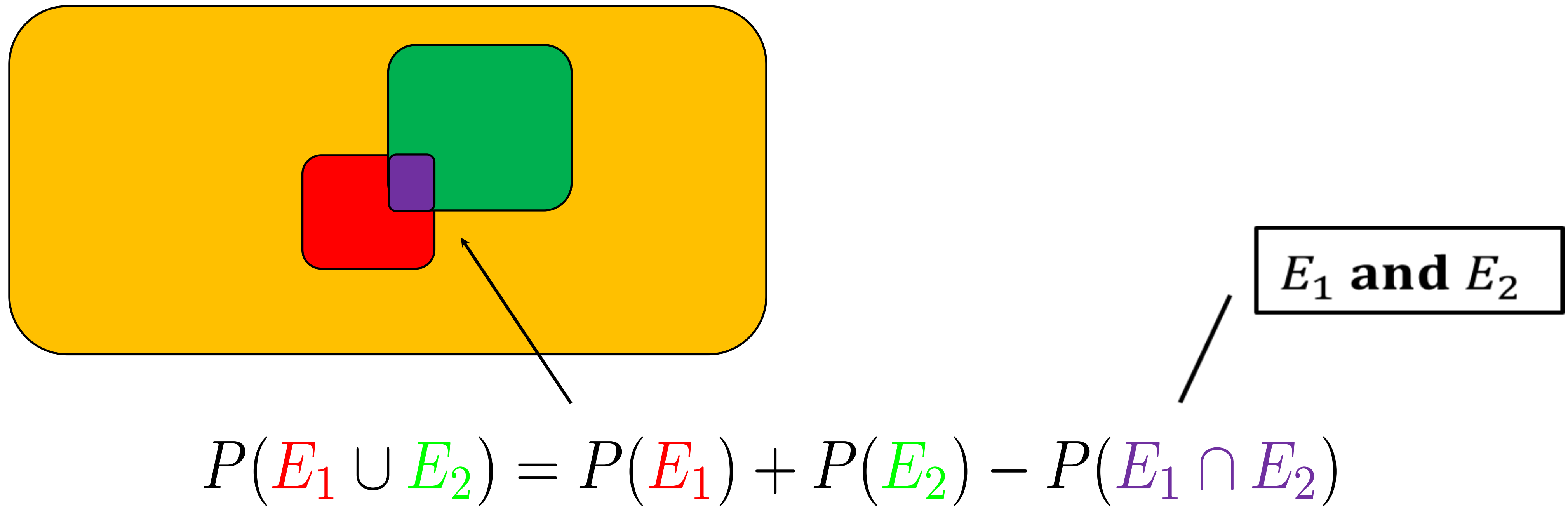
Visualizing the Axioms: III

- Axiom 3: disjoint $P(E_1 \cup E_2) = P(E_1) + P(E_2)$



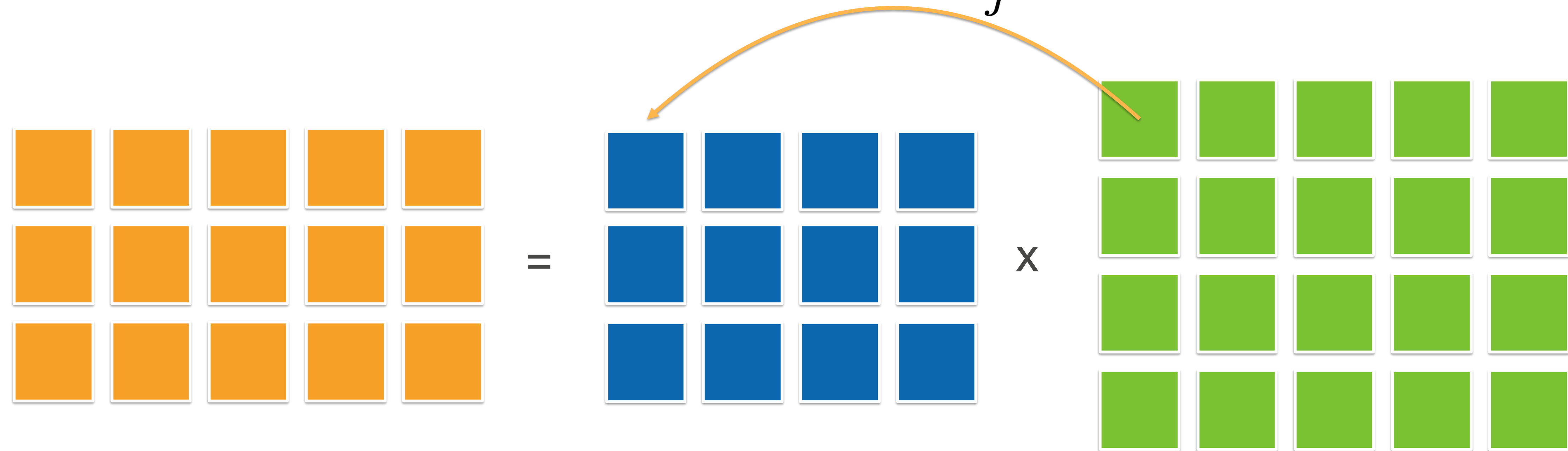
Visualizing the Axioms

- Also, other laws:



Coming up: Probability and Linear Algebra Review

$C = AB$ where $C_{ik} = \sum_j A_{ij} B_{jk}$





Recap

- What's in CS540
- Course logistics
- Software
- Probability



Thanks!