

Welcome!

---

CS 232:  
Artificial Intelligence

Fall 2023

---

Prof. Carolyn Anderson  
Wellesley College

# My Work in AI

## AI-Assisted Programming

*(How Can Machines Help Us Program?)*

**Knowledge Transfer from High-Resource to Low-Resource Programming Languages for Code LLMs**

F Cassano, J Gouwar, F Lucchetti, C Schlesinger, CJ Anderson, ...  
arXiv preprint arXiv:2308.09895

**StarCoder: may the source be with you!**

R Li, LB Allal, Y Zi, N Muennighoff, D Kocetkov, C Mou, M Marone, C Akiki, ...  
arXiv preprint arXiv:2305.06161

**StudentEval: A Benchmark of Student-Written Prompts for Large Language Models of Code**

HML Babe, S Nguyen, Y Zi, A Guha, MQ Feldman, CJ Anderson  
arXiv preprint arXiv:2306.04556

**SantaCoder: don't reach for the stars!**

LB Allal, R Li, D Kocetkov, C Mou, C Akiki, CM Ferrandis, N Muennighoff, ...  
arXiv preprint arXiv:2301.03988

**MultiPL-E: a scalable and polyglot approach to benchmarking neural code generation**

F Cassano, J Gouwar, D Nguyen, S Nguyen, L Phipps-Costin, D Pinckney, ...  
IEEE Transactions on Software Engineering

## Natural Language Processing

*(How Do Machines Understand Language?)*

**Solving and Generating NPR Sunday Puzzles with Large Language Models**

J Zhao, CJ Anderson  
arXiv preprint arXiv:2306.12255

**Do All Minority Languages Look the Same to GPT-3? Linguistic (Mis)information in a Large Language Model**

S Nguyen, CJ Anderson  
Proceedings of the Society for Computation in Linguistics 6 (1), 400-402

**ProSPer: Probing human and neural network language model understanding of spatial perspective**

T Masis, C Anderson  
Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting ...

**Guess who's coming (and who's going): Bringing perspective to the rational speech acts framework**

CJ Anderson, BW Dillon  
Proceedings of the Society for Computation in Linguistics 2 (1), 185-194

**Tell me everything you know: a conversation update system for the rational speech acts framework**

CJ Anderson  
Proceedings of the Society for Computation in Linguistics 2021, 244-253

eliminating  
bias

ChatGPT - issues

- how does it work?

How does image  
generation work?

# What are you interested in learning about AI?

technical  
applications  
of AI in  
industry

computer vision

ethics  
identifying  
human

AI v.  
products  
decision-making

critical thinking

learning from the environment  
explainability?

growth

accuracy  
widely agreed upon

informed decisions

# What is intelligence?

data retention  
analyze data

predicting  
applying information  
to new experiences

generate new things  
fill-in-the-blank  
understand vague  
or general  
instructions

# Types of AI

# Types of AI

---

- ◆ Narrow / weak AI: human-like performance on a single task
- ◆ General / strong AI: AI that can do everything humans can do

# Task-based AI

# Task-based AI

---

Our goal is to write programs that can solve tasks. This is sort of the goal of all computer science.

In AI, though, the tasks we focus on are ones that seem to require human intelligence. This is a moving standard- what seems impossible for a computer to solve one day may eventually become very easy.



# AI Tasks

---

Almost all AI tasks can be grouped into one of three main categories:

- ◆ Search
- ◆ Classification
- ◆ Generation

# Learning

# How do people learn?

---

- ◆ Some tasks have a **critical acquisition period**:
  - Language acquisition
  - Vision
  - Music training
- ◆ Others can be consciously acquired

# Example: Child Language Acquisition

## Example 1



**cj and ember manning liked**



**Gareth Roberts** @garicgymro · 45m

Just overheard from two of my kids:

Osian (5;1): Look how I caught Mickey!

Eirwen (8;2): Do you mean caught?

Osian: ... yeah.

Eirwen: But you can keep saying caught!

Osian: Look how I caught him!

## Example 2



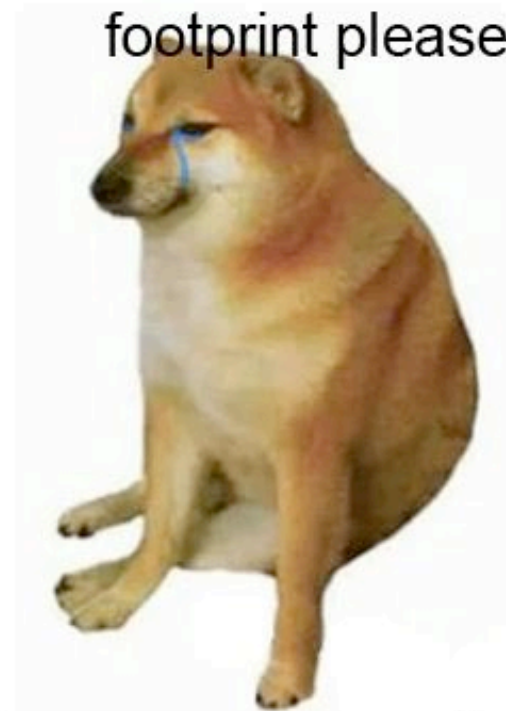
# How do people learn?

---



human infants

with fast mapping, i  
can learn the meaning  
of a word in 3  
exposures in my  
human brain powered  
by food



large language models

1 trillion  
parameters and  
a carbon  
footprint please

# How does AI learn?

---

- ◆ Symbolic AI
  - Program rules for the model
- ◆ Machine learning
  - ◆ Supervised learning
    - Give model input/output pairs to learn from
  - ◆ Reinforcement learning
    - Give model a reward function
  - ◆ Unsupervised learning
    - Model tries to **separate** data

# Practicalities

# Staff

---



Carolyn Anderson (she)  
Professor



Lepei Zhao  
Tutor



Lyra Kalajian  
Tutor



Jess Yao  
Grader



# Help Hours

---

- ◆ Tutor hours:

- TBA

- ◆ My help hours:

- Monday 4-5:30
- Thursday 4:30-5:30
- By appointment

Come to my help hours to ...

- ◆ Get help with CS232
- ◆ Talk about AI

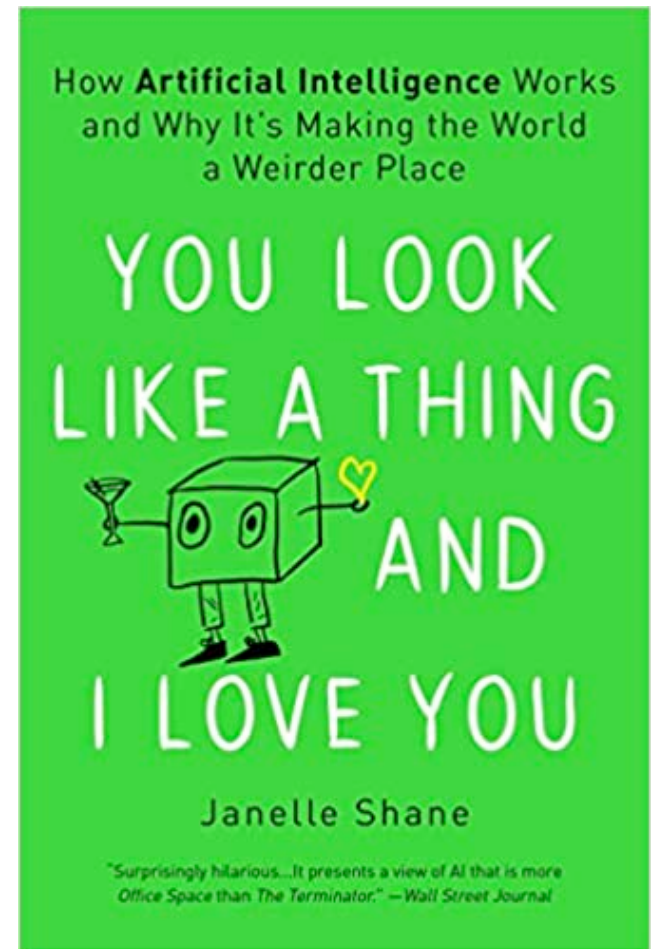
# Readings

---

Most readings will be from Janelle Shane's book *You Look Like A Thing And I Love You* (abbrev. YLLATAILY).

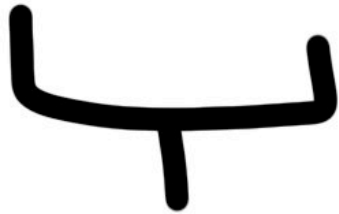
All readings are listed on the schedule. Some recommended readings are also posted there—many from Russell & Norvig's *AI: A Modern Approach*.

Please finish each week's required reading **before our Tuesday class.**



# Homework will be in Python

I recommend setting up a Python 3.8 virtual environment.



This will be a fun programming language to learn



wait this is a snake

photo credit: [Kat Maddox](#)

# Assignments

---

- ◆ Assignments are due on **Mondays at 10 PM**
- ◆ Homework submission will be through Gradescope
- ◆ Expect an assignment **every week**
- ◆ Get help early!

# Late Policy

---

You have **5 late days** for the semester, which you can use all at once, or spread across assignments. ~~I will not accept late work beyond these days.~~

5% penalty by day  
at the end of all late days

Important: I will not answer questions on late work during help hours. are used

If you have a prolonged illness or unexpected circumstance, let me know and **we'll work together** to make a **custom plan**.

# Collaboration policy

---

In this class, you can talk at a high-level with other students about assignments, but **you cannot show them your code.**

If you discuss a homework problem with another student, **please note this** on your assignment when you submit it.

**You may not use ChatGPT, Bard, Codex or any other AI system** unless explicitly stated in the homework assignment.

# Midterm and Final Project

---

We will have an in-class midterm on October 17th

There is no final exam. Instead, you will work on a **final project**. We will have presentations on the last day of class.

# Feedback and Questions

---

You can submit anonymous feedback or anonymous questions through the **Anonymous Question Form**.

Questions submitted using the form will be answered in the **Q&A document**. Check it regularly for help with assignments!

If you are submitting feedback about the course rather than a question for the Q&A document, just say that in the form.



FALL 2023 WELLESLEY COLLEGE THEATRE MAINSTAGE PRODUCTION

# R.U.R.

by Karel Čapek  
Adapted by  
Marta Rainer



*Auditions*

WEDNESDAY 9/6 | 6-8:30PM

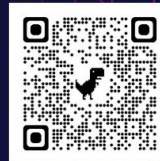
THURSDAY 9/7 | 6-8:30PM

RUTH NAGEL JONES THEATRE

SEEKING 10 ENSEMBLE MEMBERS: WE WILL NEED YOU, BRAVE HUMANS OF WELLESLEY,  
TO PERFORM - AND POTENTIALLY SAVE HUMANITY. OPEN TO ALL!

MAKE THEATRE, COLLABORATE AND BOND WITH YOUR PEERS AND EARN 1.0 ACADEMIC  
CREDIT - WITH THE SUPPORT OF A TEAM OF INDUSTRY PROFESSIONALS!

THST 345 REHEARSES M/T/W 6:30-9:30  
& TECH & PERFORMANCES DEC 7-10, 2023



# Next Class: Agents

---

## **PRE-RECORDED VIDEO**

(I'll be at a conference)

I will post the video on the Google group;  
please watch by next Tuesday.