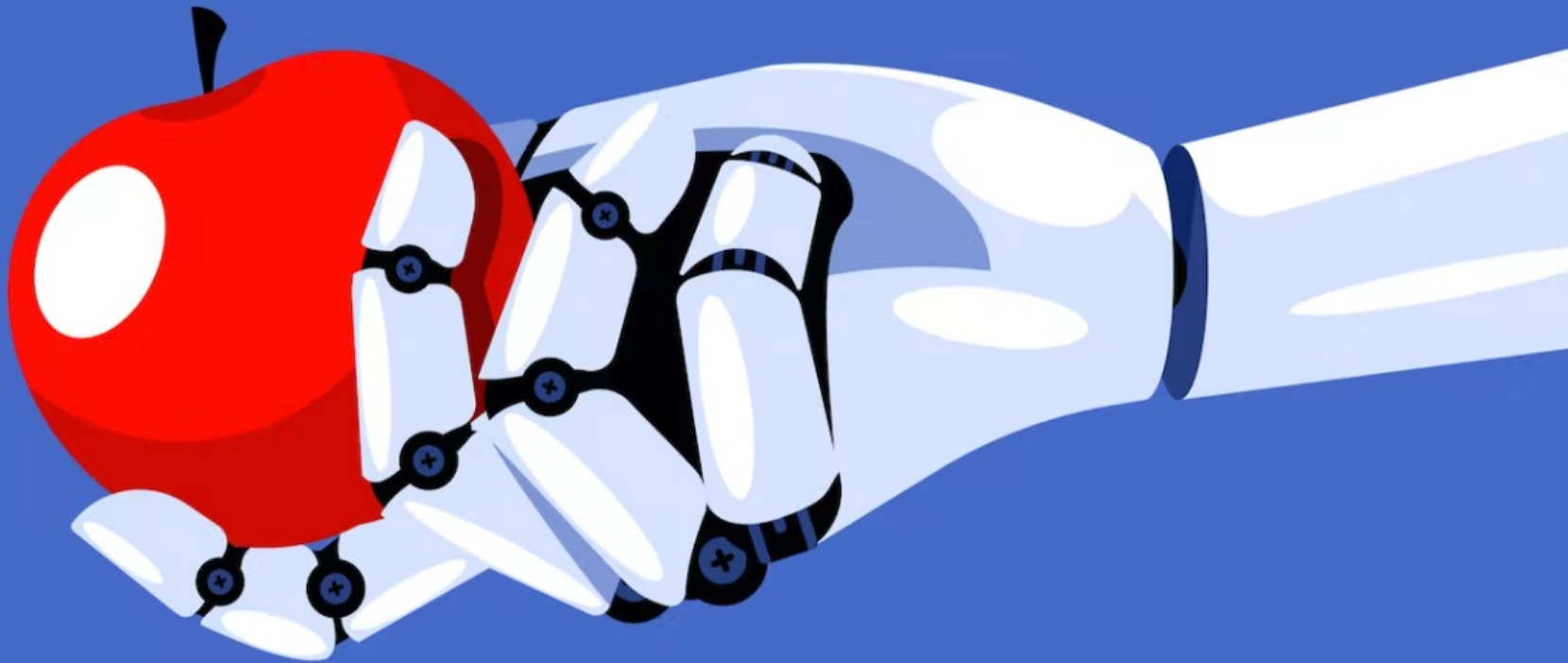


Critical AI Literacy

Understanding the New Tool



ajoubin@gwu.edu

Alexa Alice Joubin

Biases in the System

- Social biases and human value systems
- They are are baked into the AI
- Through the corpora (data sets) the AI is trained on
- Even evidence-based arguments can contain biases

CRISIS

危

Danger

机

Opportunity

Challenges

- Sorting through **conflicting info** about this disruptive innovation
- Amplification of pre-existing social biases and **prejudices**
- Mistaking synthesis for **critical thinking** (a challenge that existed before the AI arrived ...)

Opportunities

- Enhancing metacognition
- Rethinking the craft of writing
- Examining educational practices
- Improving curatorial and editorial skills
- Discussing machine ethics

Large Language Model

- A **statistical** tool to recognize and reproduce patterns based on the datasets it trained on
- **Based on the probability** of a given sequence of words occurring in a sentence
- Calculating, statistically, what might come next
- Super-powered “autocomplete”

Not all heroes wear _____

Not all heroes wear

capas (80%)

pants (5%)

...

pyjamas (0.01%)

Not all heroes wear _____

Not all heroes wear capes

but (65%)

and (10%)

if (4%)

...

, (0.03%)

What really happens behind the scene ...

From a **human perspective**: We are asking an infinite variety of questions.

From the **AI perspective**: Humans repeatedly ask the same question – what are the most likely words to follow what I just typed ...

What really happens behind the scene ...

- Human: “The first person to walk on the Moon was ...”
- AI interprets the question as “Given the statistical distribution of words in the English corpus, what words are most likely to follow the sequence – The first person to walk on the Moon was?”
- Reaching this answer: “Neil Armstrong”

Human: "What is the highest city in the world?"

AI: La Paz, Bolivia, at 3,869 m above sea level

But: El Alto, Bolivia, is in fact higher, at 4150 m

The prompt should be fine-tuned to define city or metropolitan area (having a minimum population of _____)

Who am I? What do I believe?

- Certain tasks cannot be outsourced to AI
- Understand the world *through language*
- Reveal our **unique selves** *through writing*
- **Freedom** = freely express one's originality, not just through embellishment of language generated by an AI chatbot
- Structuring my own language to **find out who I am**