LECTURE X

Current Events in Engineering

SECTION I

Generative Artificial Intelligence

What is Artificial Intelligence?

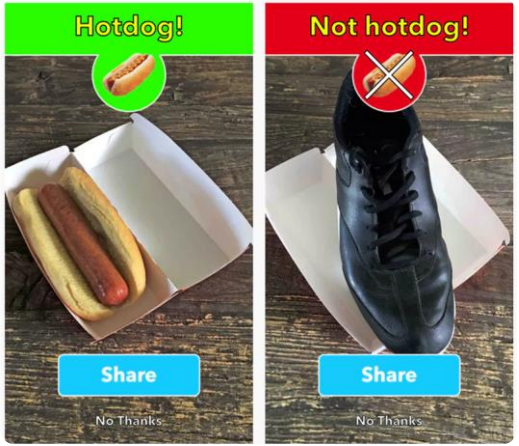
- **Artificial Intelligence (AI)** is technology that **simulates human behavior**
  - Ex) Chess bot, video game NPC, Alexa, Google Maps traffic estimates
- How do I know if a system is AI or not? Yes, if any of the following…
  - Does it learn or adapt to new information?
  - Does it make decisions or predictions based on input data?
- There are different approaches to AI (search algorithms, expert systems), but we will focus on **machine learning** in this lecture…

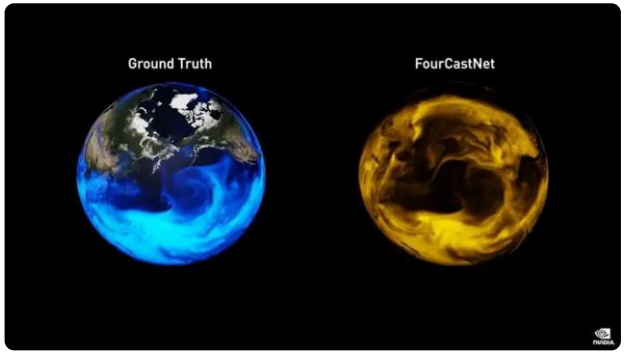
Machine Learning

- **Machine learning** **(ML)** is a process of “training” programs to recognize patterns from data and to make decisions based on those patterns
    - **Programs self-learn** new behavior with some (but not much) human intervention
- **Deep learning** is a subset of machine learning which uses **artificial neural networks** and involves **little to no human guidance**
    - **Neural networks** are ostensibly mathematical functions with coefficients and biases that are adjusted while trained to data
        - Eventually, the neural network (or function) is optimized such that, provided an input, it will generate an accurate output

Predictive Models

- We train **models** (aka programs) which can then be used to **make decisions based on new, unseen data**
- Up until the last few years, we mostly built **predictive models**
  - These models are **used for classifying data**
  - Ex) Not Hotdog is a predictive model that performs image classification to identify hotdogs in photos
  - Ex) FourCastNet Weather forecasting





Generative Models

- More recent breakthroughs led to **generative models** – artificial intelligence that can **generate new content** (text, images, simulations, etc.)
    - Ex) ChatGPT, DALL-E, Llama, Gemini (all released in 2022)
- From just 2022 to 2024, many of these models have been improved dramatically in realism, size, and speed…

DALL-E by OpenAI

Training ChatGPT

- ChatGPT is powered by **large language models** (**LLMs**) designed to understand and respond to human language
    - Models like the ones used with ChatGPT are **trained using deep learning**
- Training makes the model more accurate. How?
    - **We feed it data… lots of it**
        - Acquired from information “publicly available” on the internet
        - the exact amount is not disclosed
- The size of GPT-3 (the LLM on which ChatGPT is based) is rumored to be 350GB

Sora

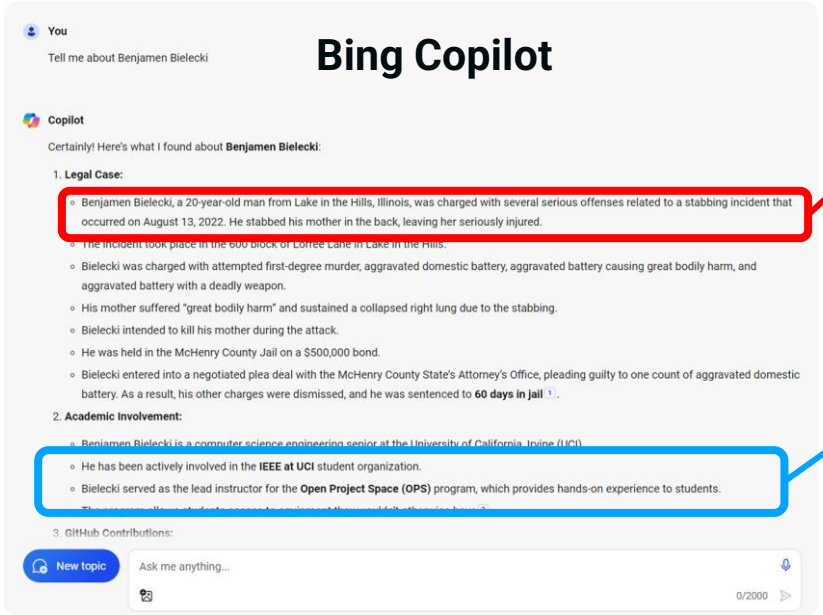
- Generative AI has arguably surpassed the **uncanny valley** – the idea that virtual or artificially-generated content designed to look real still appears… weird
- In February 2024, OpenAI announced **Sora** – a text-to-video generative model
- Here is some example content:

Generative AI in the Film Industry

- The film industry rapidly adopted the new technology
    - Luke Skywalker's face was rendered with specially-trained "deepfake" generative models in *The Mandalorian*
- Adobe products now feature generative AI tools…
    - Objects can be added or removed on a whim

AI Risks: Misinformation

- Generative AI does not verify the truthness of its responses
  - The **accuracy of its results are coincidental**, not assured…
- Ex)



"Benjamen Bielecki, a 20-year-old man from Lake in the Hills, Illinois, was charged with several serious offenses…"

"…served as the lead instructor for the Open Project Space (OPS) program…"

AI Risks: Misinformation (Cont'd)

Anthony J. Bielecki, 19, of Arlington Heights, (inset) is charged in connection with a stabbing that left a woman injured in the 600 block of Lorrie Lane in Lake in the Hills on August 13, 2022. | Background Photo: Google Street View

LOCAL NEWS

60 days in jail for man who pleaded guilty to stabbing his mother in Lake in the Hills

LMCS by Sam Borcia Published May 27, 2023 9:54 AM

"...Anthony J. Bielecki, 19..."

- Generative AI **doesn't "know" what it is saying/showing**
  - It simply **uses math and statistics to predict** the most likely output

AI Risks: Misinformation (Cont'd)

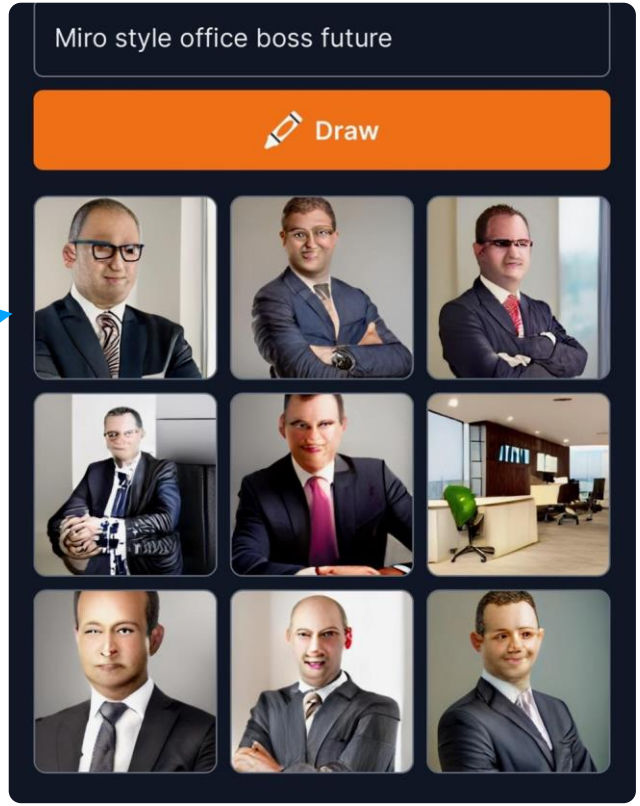
- Tech news outlet **CNET was discovered to have used AI to write stories**
  - Corrections had to be issued to 41 of the 77 stories where AI was used
- LLM tools are known to have generated **false attributions**
- Prevalence of **AI fake news sites** has dramatically increased
  - Bad actors create sites to generate *intentionally* misleading content
- **Deepfakes** – fake imagery of real people – can be used to spread false information

AI Risks: Intellectual Property

- The **New York Times filed a lawsuit against Microsoft** (owner of ChatGPT)
  - Alleges that **chatbots are violating the IP rights** of the Times by plagiarizing their articles verbatim
  - The bots generate false attributions to the Times
  - There are suits from fiction and nonfiction authors as well over IP rights violations
- Amazon and other large companies inform employees not to share code or information with ChatGPT
  - **Proprietary code could be used as training data** or mixed with other company's IP

AI Risks: Bias

- The **training data can introduce bias** based on race, sex, etc.
  - If the training data is not robust, the model will be less accurate
  - Are all CEOs White men?
- Computer-aided diagnosis systems have **lower accuracy results for Black patients** than White patients



AI Risks: Environment

- Data centers which power AI tools **require lots of water** for cooling
  - Microsoft's global water consumption increased by 34% from 2021 to 2022 due to its AI Research
- Servers also **require significant power**
- AI and the cloud are **expected to intensify greenhouse gas emissions**

AI Revolution

- Generative *and* predictive models are rapidly changing all industry sectors…
  - **Journalism** – News reports are increasingly generated with AI tools
  - **Healthcare** – AI will conduct health interviews, identify diseases, and monitor patient stats
  - **Finance** – Stock trades, fraud detection, and financial audits will all be automated
  - **Transportation** – self-driving vehicles, traffic light optimization, future urban planning
  - The list goes on…

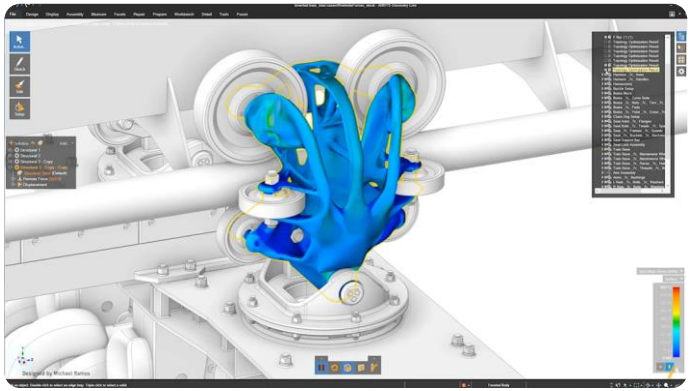
AI Revolution: Software Development

- Software engineers are already using **generative AI coding tools**, which can…
  - Explain code functionality
  - Simplify code
  - Write unit tests
  - Write documentation
  - Identify bugs

- These tools are **not competent enough to write full software applications**
  - This won’t leave programmers unemployed (maybe the bad ones)
- Learn prompt engineering to take advantage of these tools and write new ones!

AI Revolution: Mechanical Engineering

- Engineers can use **generative design** software to create structures and systems
- Generative design tools are useful for…
  - Quickly generating several **permutations of structures** based on design constraints
  - creating **complex shapes which cannot be traditionally manufactured** (using additive manufacturing)
  - **design optimization**

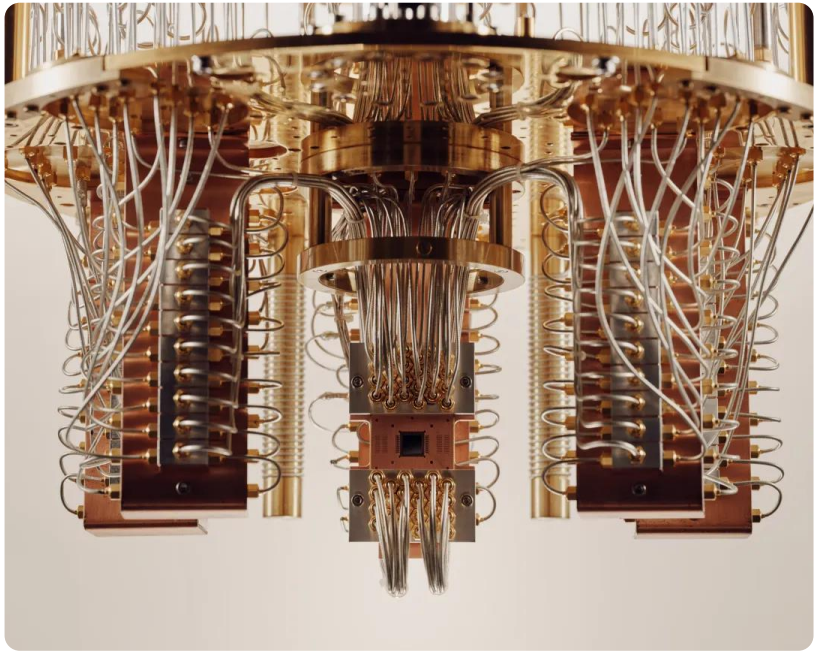


SECTION II

Quantum Computing

Quantum Computers

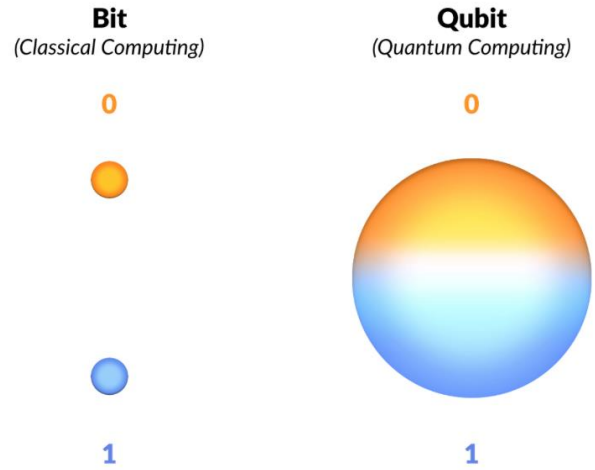
- **Quantum computers** are specialized hardware and software that utilizes quantum mechanics (physics at the subatomic level)
- This lecture focuses more on why quantum computing is important, not how it works



IBM Q System One

Qubits

- Data is represented as **quantum bits** (**qubits**) on quantum computers instead of classical bits
  - They **can represent a 1 or 0**, like classical bits OR…
  - They can **also be in a superposition of 1 and 0**, with a certain probability of being 0 and a certain probability of being 1



Superposition and Entanglement

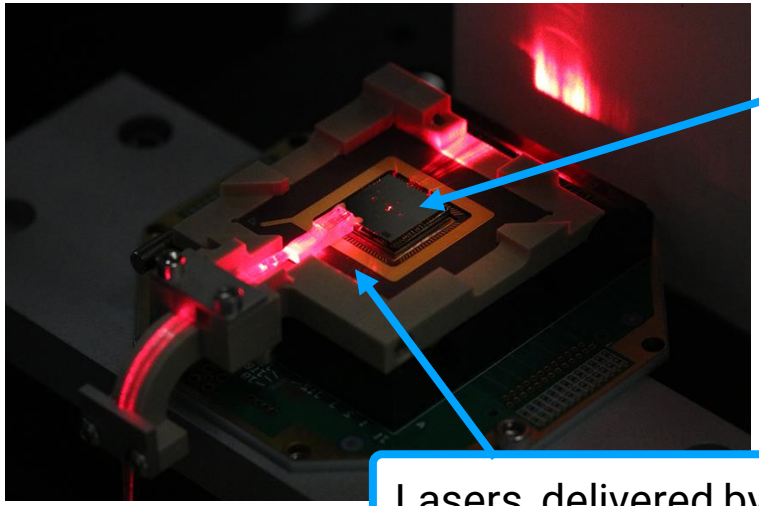
- What makes qubits so useful?
- The principle of **superposition** enables qubits to **exist in multiple states at once**
  - Classical bits only exist in one state at a time
  - **Qubits store more information** than a classical bit while in superposition
- Multiple qubits can experience **entanglement**
  - In entanglement, the **qubits' states correlate**
  - We can **manipulate many qubits** just by changing one qubit **in a single operation**

Implementing a Qubit

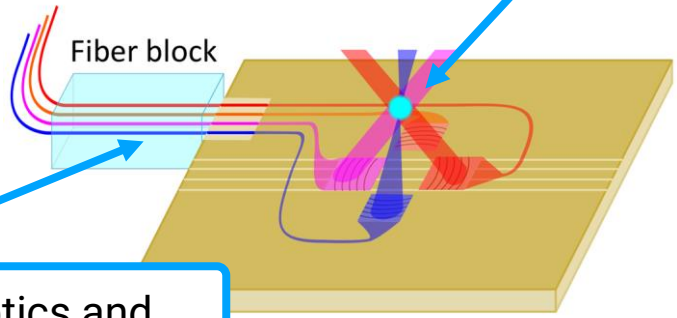
- Qubits can be made from **trapped ions**, **electrons**, **real** or **artificial atoms**
  - Some implementations demand **temperatures approaching absolute zero** (to shield from magnetic and electric interference)
- Ex) **Ion Trap Qubit**
  - A charged atom (ion) is **trapped by electron fields** and **cooled to near absolute zero** (low energy)
  - **Lasers** are used to **change the energy state** of the qubit
  - **Lasers and cameras** are also used on the ion to **measure the qubit** state

Implementing a Qubit

Below is an implementation of the **ion trap qubit on a chip**:



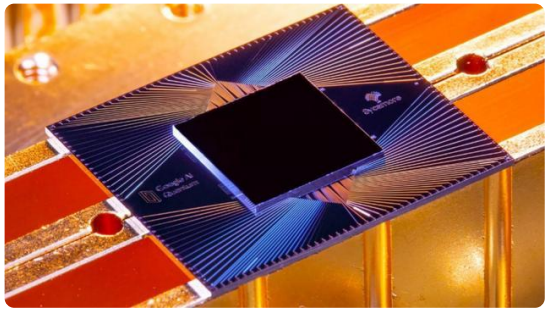
An ion is trapped just above the chip's surface at cryogenic temperatures

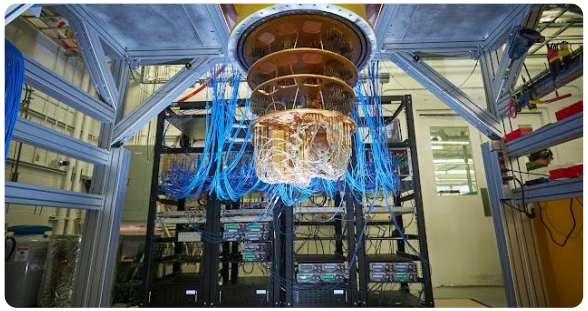


Lasers, delivered by fiber optics and guides within the chip, control the ion qubit

Quantum Computing Speeds

- The qubit's superposition and entanglement properties enable **extremely fast computing**
    - *"Information that 500 qubits can easily represent would not be possible with even more than 2^{500} classical bits."* – Microsoft
- **Google's Sycamore processor** (53 qubits) was used to **complete a task in 200 seconds** that would take a state-of-the-art supercomputer 10,000 years





Where is Quantum Computing Heading?

- Quantum computers **need to be brought to scale**
  - Computers require thousands of qubits to perform tasks useful to humans
    - **Most today are less than 100 qubits**
  - Management consulting firm McKinsey estimates that the technology for handling complex problems won't be available before 2035
- **Quantum computers are error-prone** due to environmental noise to imperfections in the qubits
  - **Quantum Error Correction (QEC)** is being developed to handle this
- One of its most immediate effects has been on cybersecurity…

Cryptography in the Quantum Age

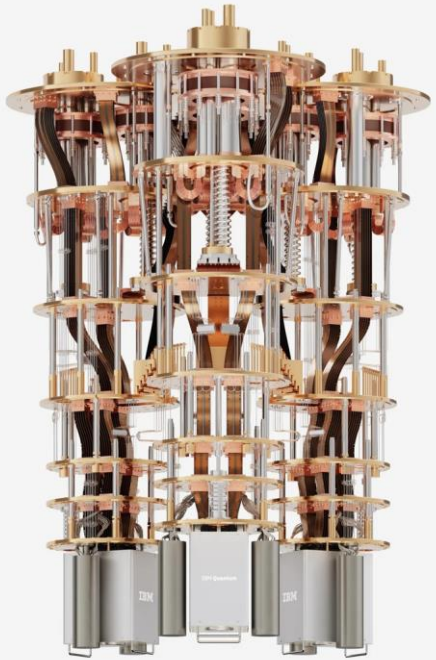
- **Cryptography** is the study of technique for secure communication
  - **Private information is encrypted** so it may only decrypted by trusted people and not be read by bad actors
  - There exist methods of encryption, data transfer, and decryption
- **Over 90% of the Internet uses RSA encryption** (Forbes)
  - Relies on the fact that it would be **prohibitively time confusing to break the encryption**
  - Shor's Algorithm suggests **quantum computers could solve RSA encryption fast**, making it entirely unsafe

Cryptography in the Quantum Age (Cont'd)

- **Quantum Key Distribution** (**QKD**)
  - Most viable method of quantum cryptography so far
  - Allows private data to be transmitted in a way that makes **undetected eavesdropping impossible**
  - Considered impractical due to scalability issues and security concerns – chiefly, how do you authenticate the connection?
- The **U.S. National Institute of Standards and Technology (NIST)** has announced several **quantum-resistant cryptographic algorithms**
  - These do not require quantum computers

Advancing Quantum Computing

- Google, IBM, Microsoft, and Amazon are some of the top companies researching quantum computing
- IBM is leading the quantum race
    - The **IBM Condor is the first quantum processor to exceed 1,000 qubits** (December 2023)
        - 1,121 qubits



Quantum Computing Jobs

Here are “hypothetical” 2024 job listings with the qualifications employers are currently looking for…

Quantum Research Scientist
(Shmamazon – Pasadena, CA)

Basic Qualifications

- PhD or Master’s
- 4+ years of quantitative research experience

Preferred Qualifications

- R, MATLAB, Python
- Agile development experience

Quantum Systems Architect
(Shmicroshmoft – Washington)

Qualifications

- BS/MS in CS or related field
- 5+ in the software industry
- 3+ years with Azure

Notice that these jobs **do not always require prior quantum computing experience**

SECTION III

Final Thoughts

How to Keep Up

There will always be newly emerging technologies. So how do you stay current?

- **Join a professional organization** (pssst… IEEE)
  - You receive access to publications, conferences, and chapter events that update you on current events
- **Attend conferences and tech events**
  - Attend workshops and classes at events like IEEE Rising Stars
- **Read the news and listen to podcasts**
- **Commit time to personal projects**
  - You can learn machine learning right now!

FAIR USE DISCLAIMER

Copyright Disclaimer under section 107 of the Copyright Act 1976, allowance is made for “fair use” for purposes such as criticism, comment, news reporting, teaching, scholarship, education and research.

Fair use is a use permitted by copyright statute that might otherwise be infringing.

Non-profit, educational or personal use tips the balance in favor of fair use.

CC BY-NC-SA 4.0