



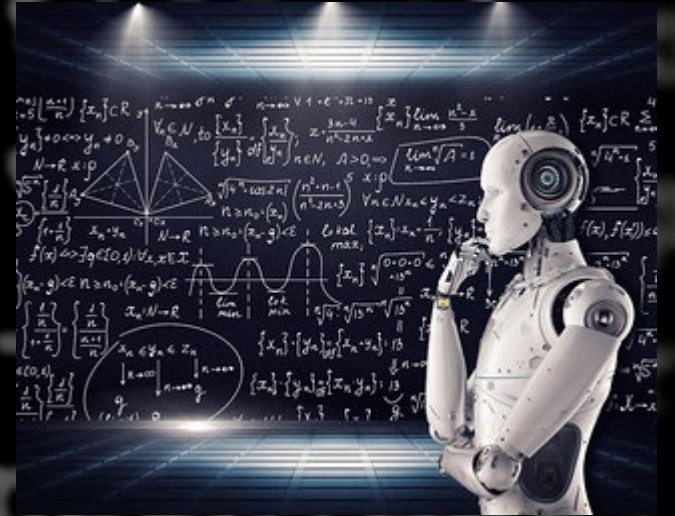
**HD**  
**FILM TRIBUTES**



**F HD**

# AGENDA

- I. Overview
- II. What is AI?
- III. A Brief-ish History of AI
- IV. Uses of AI by Municipal Entities
- V. Legal and Ethical Issues Associated with AI
- VI. Selecting AI Vendors
- VII. Regulation
- VIII. Other Jurisdictions
- IX. Resources
- X. The Future of AI



# I. OVERVIEW

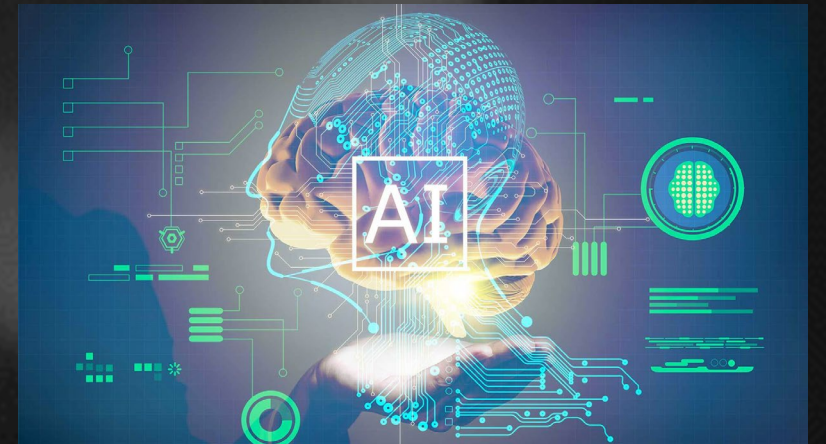
- Ask the right questions, even if you don't have all the answers
- Stay current
  - Technological advances in hardware and software
  - Academic and industry studies
  - Legislation
  - Use resources listed at the end of the paper

# I. OVERVIEW (CONT.)

- AI has great promise, but also potentially great risks
  - Industry introduced AI to the general public in November 2022
  - “Reveal” for public use came without any prior regulation
  - Due to massive investment, market pressures, political considerations, and international security risks and competition, comprehensive regulations by federal government may have limited impact
  - State and cities likely to be at the forefront of regulating high risk uses

## II. WHAT IS AI?

- An intelligent machine
  - What's "intelligent?"
- Machine learning
- Large Language Models (LLM)
- Artificial Generative Intelligence (AGI)



# III. A BRIEF-ISH HISTORY OF AI

- Humans have been trying to create intelligent machines for millennia
- Advances in “software:” mathematical logic; statistics; Bayesian probability theory; game theory; data analysis
  - Limits: Gödel's incompleteness theorem; halting problem; reachability problem; complexity and chaos theory
  - Results
    - Fundamentally unpredictable behavior of AIs
    - Inability to explain outputs: can't run it backwards to recover a unique set of inputs

# III. A BRIEF-ISH HISTORY OF AI (CONT.)

- Advances in “hardware:” Moore’s law for semiconductor chips; computing power; memory storage; neural networks designed like human brains
  - Limits: Speed of light, size of physical dimensions and objects; quantum mechanical effects (quantum tunneling, inherently probabilistic behavior, Heisenberg uncertainty principle)
  - Results: human brains and AIs both have fundamental components that are nanometer-sized



# III. A BRIEF-ISH HISTORY OF AI (CONT.)

- Welcome to the modern era
  - Some notable achievements
    - Games: checkers, chess, Jeopardy, poker, Go
    - SAT, LSAT, Bar Exam
    - Siri, DARPA, autonomous vehicles, robots that can operate first in constrained and then novel environments
    - Discovery of new proteins and materials; used for medical diagnosis and treatments



# III. A BRIEF-ISH HISTORY OF AI (CONT.)

- Welcome to the modern era
  - Some notable recent “concerning situations” created by AI
    - JWST first discovered certain exoplanets, when it didn’t
    - Depicted the Pope as a female
    - Depicted the Founding Fathers as including dark-skinned and indigenous people
    - Sydney: a case study
      - A mirror image of our own minds?
  - But current AIs pass the Turing Test

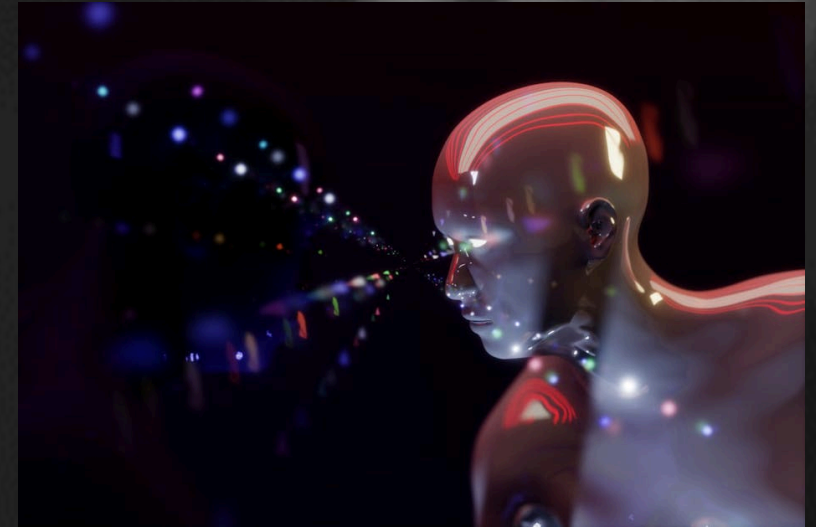
A man with short blonde hair, wearing a dark suit jacket over a light-colored shirt, is seated at a wooden desk. He is looking slightly to his left with a neutral expression. The room is dimly lit, with a prominent brass desk lamp on the desk behind him, casting a warm glow. On the desk, there is a tray with a glass and a bottle. The background is a plain, light-colored wall.

ex machina  
**OFFICIAL CLIP**

# IV. USES OF AI BY MUNICIPAL ENTITIES

## ➤ HR

- Meet and confer; effects bargaining
- Job descriptions: training to use
- Recruitment: watch out for bias
- Negotiations; dueling AIs
- Disability Accommodation: virtual assistants; screening bias
- Language translation: multi-lingual employees; translate public meetings
- EAP: use caution if considering using AI
- Executive Coaching: use humans



# IV. USES OF AI BY MUNICIPAL ENTITIES (CONT.)

- Public Works: potholes; traffic; homelessness prediction and resources
- Emergency services departments: crime investigation and reduction
- Administrative departments: summaries, modeling, risk assessment, legal documents
- Use by employees
  - Policies/regulations to govern use
  - Writing...everything?
  - Virtual assistants

# IV. USES OF AI BY MUNICIPAL ENTITIES (CONT.)

- Use by employees (cont.)
  - Productivity: recent studies
  - Workweek reduction; industry thinks AI will result in cutting jobs and boost productivity
  - Bottom line: define the problem, then decide if using AI to solve it is a good idea
- Use by unions
  - Salary surveys
  - Contracting out
  - Organized activity; NLRB memo

# IV. USES OF AI BY MUNICIPAL ENTITIES (CONT.)

- Use by the public
  - Environmental documents: is an AI comment a comment requiring a response?
  - Social media: can AIs be blocked?
- Use by bad actors
  - “Zoom bombing” taken to a new level
  - Sophisticated cyber scams

# V. LEGAL AND ETHICAL ISSUES ASSOCIATED WITH AI

- Data scraping: whose data, who keeps it, who gets to use it
- Consents, permissions, and copyright: litigation; AIs can't get patents
- Privacy
  - Protecting privacy: is it really protected?
  - Asymmetry of information: vast datasets and unparalleled ability to aggregate and analyze
  - Facial recognition: watch out





# V. LEGAL AND ETHICAL ISSUES ASSOCIATED WITH AI (CONT.)

- Bias
  - Bias tradeoffs in models: privacy vs. bias reduction; chaotic behavior
  - Speech and image biases
  - The “current employee” bias: use of enterprise data sets
  - Recent litigation: *Mobley v. Workday, Inc.*
- Transparency
  - Most models score low
  - Mandate disclosures and watermarks?



# V. LEGAL AND ETHICAL ISSUES ASSOCIATED WITH AI (CONT.)



- Explainability (the “black box problem”)
  - The most intractable problem facing AI today
  - Use caution if using AI to make “consequential” decisions
- Hallucinations:
  - No one knows what’s causing them
  - Plan on human involvement at important stages



# V. LEGAL AND ETHICAL ISSUES ASSOCIATED WITH AI (CONT.)

- Security
  - “Deepfakes:” virtually no barriers to entry
  - What is true? Freedom to parody; confirmation bias
  - Buy good insurance (and make sure your vendors have too)
- Risks of not using AI
  - AI can be better and less biased than humans at some tasks
  - Don’t confuse apparent objectivity and accuracy for consistent and fair results

# VI. SELECTING AI VENDORS

- Find, hire, or become an expert
  - Products and capabilities changing rapidly
  - Operation of system isn't fully transparent and likely can't be made so
  - City should have capability to audit and test system periodically for compliance with legal and other city requirements
- Read and negotiate the fine print in the contract
  - Vendor should be responsible and certify that product complies with all laws including anti-discrimination laws

# VI. SELECTING AI VENDORS (CONT.)



- Read and negotiate the fine print in the contract (cont.)
  - Broad defense, indemnity, and hold harmless; include insurance that covers any possible city risks, including copyright or similar claims relating to data scraping
  - Confirm product tested and passes the “4/5ths” rule from EEOC; mandate periodic testing
  - Prohibit access to confidential or proprietary city data
  - Require vendor to stop using product if disparate impact or other problems arise
  - Ensure city owns data (including “synthetic data”) and product outputs; prohibit data transfer to anyone else
  - See San Jose’s GovAI Coalition template, referenced in the paper



## VII. REGULATION

### ➤ International

➤ EU AI Act: describes AI categories based on risk level and matches regulation accordingly; effective late 2025 or early 2026

### ➤ Prohibited

➤ Biometric ID in public spaces by law enforcement, except when investigating certain serious crimes (terrorism, murder, kidnapping...)

➤ Biometric ID using sensitive characteristics: political, religious, philosophical, sexual orientation, race

➤ Facial recognition, emotion recognition in workplaces, social scoring based on personal characteristics, manipulative/exploitive behavior

# VII. REGULATION (CONT.)

- International (cont.)
  - High risk
    - Significant potential to harm health, safety, fundamental rights, environment, democracy, elections, rule of law
    - Requires impact assessment, mitigation of risks, testing, reporting, security guarantees
  - Lower risk: Generative AI (ChatGPT, etc.), chatbots
    - Transparency, disclosures about data, uses



# VII. REGULATION (CONT.)

## ➤ Federal

### ➤ October 2022: Presidential AI Bill of Rights

- Safety and effectiveness
- Protection against algorithmic discrimination
- Data privacy
- Notice and explanation
- Human alternatives, consideration, and feedback



# VII. REGULATION (CONT.)

- Federal (cont.)
  - July 2023: voluntary commitments from seven leading AI companies
    - Internal/external security testing before public release
    - Share risk information with industry, government, academia
    - Prioritize cybersecurity and protect proprietary components
    - Disclosures: watermarking for AI-generated content
    - Public reporting of capabilities, limitations
    - Prioritize research on risks (bias, discrimination, privacy)
    - Develop systems to address societal challenges

# VII. REGULATION (CONT.)

- Federal (cont.)
  - October 2023: Executive Order
    - Implements many of the points of voluntary agreement
    - Additional provisions to not worsen job quality or cause labor force disruptions
    - Advance equity and civil rights
    - Not disadvantage marginalized groups
  - February 2024: AI Safety Institute Consortium
    - Resource Center with detailed and comprehensive information

# VII. REGULATION (CONT.)

- Federal (cont.)
  - Congressional Legislation: on the increase
    - Use the Brennan Center's AI Legislation Tracker
    - Includes all current AI legislation that
      - Imposes restrictions on high-risk AI
      - Requires evaluations or assessments
      - Imposes transparency, notice, or labeling requirements
      - Creates or designates a regulatory authority to oversee AI
      - Protects consumers with liability measures
      - Directs government to study AI to inform potential regulation



# VII. REGULATION (CONT.)

- Federal (cont.)
  - EEOC
    - May 2023 Technical Assistance Document
      - Not binding but contains very helpful FAQ section
      - Employers using AI to screen job applicants held to “adverse impact” standard
        - “The 4/5 Rule:” selection rate for any race, sex, or ethnic group that’s less than 4/5 of the rate for the group with the highest rate is evidence of adverse impact
        - \$365K settlement: employer screened out candidates over 55





## VII. REGULATION (CONT.)

### ➤ California

- September 2023 Executive Order directing state agencies to adopt proactive approach to AI by July 2024
  - Risk-analysis report
  - Procurement blueprint
  - Beneficial use report, including harms/risks for public employees
  - Deployment and analysis framework, including pilot programs
  - State employee training

# VII. REGULATION (CONT.)

- California (cont.)
  - September 2023 Executive Order directing state agencies to adopt proactive approach to AI by July 2024 (cont.)
    - Establish partnership with UC Berkeley and Stanford for impact analysis on workforce
    - Recommend legislation
    - Ongoing evaluation of AI impacts
  - Department of Technology created “AI Community:” meets quarterly
    - Issued GenAI Guidelines for Public Sector Uses, Procurement, and Training



# VII. REGULATION (CONT.)

- California (cont.)
  - Legislature
    - Five bills introduced as of date paper written; likely to change
    - Most include the same provisions as in EO, mainly addressing high-risk uses and disclosure/transparency when AI is being used
  - Civil Rights Division (CRD)
    - Proposed regulations released February 2023
      - Address AI decisions that have adverse impact on applicants based on protected characteristics; data retention requirements

# VII. REGULATION (CONT.)

- California (cont.)
  - Local policies
    - San Jose “GovAI Coalition”
      - Wide representation, template policies, manuals, vendor contracts and registry
    - San Francisco Generative AI Guidelines
      - Always review and fact check
      - Always disclose use
      - Always incorporate expert humans in decision making
      - Never provide confidential information or generate images that could be mistaken for real people



## VIII. OTHER JURISDICTIONS

- New York City Bias Audit Law; effective January 2023, enforced July 2023
  - Companies operating and hiring in NY that use AI for hiring or promotion must undergo independent bias audit
- Seattle, Washington
  - November 2023: adopted generative AI policy, similar to others discussed

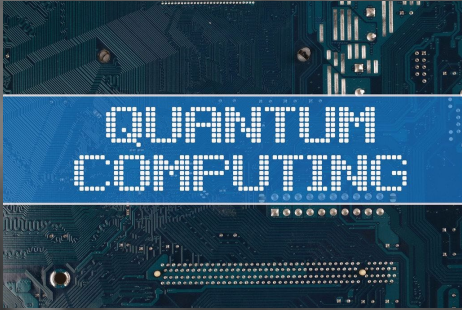


# IX. GENERAL RESOURCES



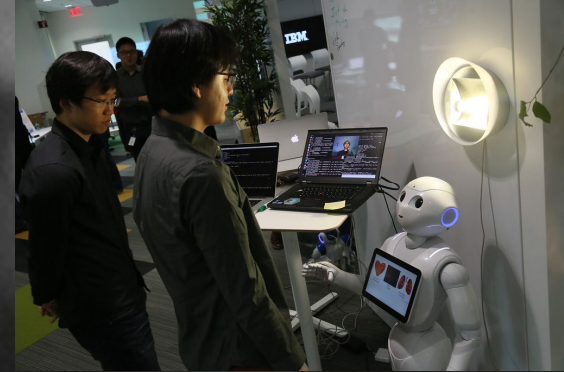
- How to use them
  - Sign up for Stanford and Berkeley (and/or other) newsletters
    - Current research; annual summary in charts/graphs
  - Check Markkula Center for Law and Technology for webinars
  - Use NIST's AI Resource Center for general and detailed information on every aspect of AI
  - Use SJ GovAI Coalition for California-specific templates
  - Use prior CalCities AI papers for advice on AI tools for attorneys

# X. THE FUTURE OF AI: IT'S ALREADY HERE



- Emergent behavior: occurs in natural systems, perhaps also in artificial systems
  - Increasing complexity + probabilistic algorithms = unpredictable and novel AI behavior, a hallmark of human behavior and creativity
- Quantum computing meets AI
  - Fundamentally non-deterministic; could solve currently intractable problems; could also break encryption protocols
- Do AIs have first amendment rights?
  - The corporations that invented them do

# X. THE FUTURE OF AI: IT'S ALREADY HERE (CONT.)



- Robots in the workplace
  - Already used extensively in industry to perform repetitive, dangerous, or undesirable jobs
    - Requires retraining/upskilling current employees if workforce will be retained
  - Interactions with human workers
    - Will workplace rights extend from humans to AI robots?
- Data centers in space: it's cold there

# ARE YOU SMARTER THAN AN AI?

## ➤ Takeaways

- AI is a transformational human achievement that already has, and will continue to change the world
- In the U.S., meaningful regulation will likely occur first at the state and local levels
- AI technology is changing rapidly and being deployed with minimal regulatory oversight
- Cities can take significant first steps to implement AI for beneficial uses and limit use/mitigate risks when using for consequential decisions



NOVA  
WONDERS







**STREAM ON THE PBS APP**