Name(s) ______________________ Period ________ Date ____________

AI Research Areas

Instructions

This document contains several AI ethics research areas, along with sample articles and videos for each. For each area, we showcase a technology or industry using AI. We present some use cases for that technology as well as cases of bias, ethical concerns, or current issues that AI practitioners are working on! This is not a complete list, so use it to help you get started before exploring beyond!

Facial Recognition

Facial recognition software is a technology used to analyze faces. It can be used to identify a specific person, recognize someone's face, or simply detect if there is a face present in an image. It works by surfacing patterns and characteristics from a large training set of images – determining which features are most important (like width between eyes, thickness of lips, etc). It's used in lots of different ways, from unlocking your phone to verifying your ID. However, it has been shown that facial recognition technology has bias towards some races/genders over others.

Use Cases

🤳phone unlocking
📹theft detection
✈️airport security
💅beauty filters
🌏immigration check
🖼️image tagging

Examples of How This Technology Has Impacted Society

- ➔ Pros and Cons of Facial Recognition Technology
- ➔ Pros and Cons of Facial Recognition for Law Enforcement Video
- ➔ Video Analysis to Prevent Traffic Accidents

Ethical Considerations

- ➔ Gender Shades: Algorithmic Bias in Detecting a Face
- ➔ Wrongfully Accused by an Algorithm
- ➔ Facial Recognition at the Airport: Privacy and Racial Bias Concerns
- ➔ Software that monitors students taking exams perpetuates inequality
- ➔ Facial Recognition Mistakenly Matches 28 Members of Congress to Mugshots
- ➔ Digital Beauty Filters and Colorism

✨Spotlight Person: Joy Buolomwini

Social Media Recommender Systems

On social media, recommender systems play a large role in what we see. The algorithms determine how content gets ordered, what ads are served to us, which content is recommended, and which friends are suggested. Typically, these systems do this through a combination of methods that take into account a user's past behavior, user demographics, and what similar users tend to like. From the news you see to the memes you share, recommender systems play a role. However, there is increased concern around how misinformation spreads online, and whether or not social media encourages us into narrow `echo chambers'. There are also concerns about how social media content impacts mental health and body image.

Use Cases

📱organizing newsfeed
🙋suggesting friends
🛍️showing ads
📰presenting news
💡recommending content

Examples of How This Technology Has Impacted Society

- ➔ 10 Positive Effects of Social Media
- ➔ Social Media Activism for Marginalized Groups
- ➔ Benefits of Recommendation Systems for Business

Ethical Considerations

- ➔ When Algorithms Decide Whose Voices Are Heard
- ➔ Echo Chambers and Filter Bubbles
- ➔ Body Image and Social Media
- ➔ Diet Culture and TikTok
- ➔ Youtube Rabbit Holes and Misinformation
- ➔ What TikTok Does to Your Mental Health

✨Spotlight Person: Casey Fiesler

Applicant Tracking Systems

Large companies and universities receive thousands to even millions of applications a year. Automated tools can help with such a large influx of applications. Also known as Applicant Tracking Systems (ATS), these AI tools are a type of software that manages the recruiting and hiring process, including job postings and job applications. Some of these systems include automatic resume sorting or applicant selection. They typically work by learning on successful past candidates as well as some criteria, to surface the characteristics of a desirable candidate and search for those. However, there is evidence that these tools have the potential to discriminate.

Use Cases

🔍background checks 📄sift through resumes 🎓 college admissions

Examples of How This Technology Has Impacted Society

- ➔ Helping Kids Get Into College
- ➔ Creating diverse, inclusive & transparent workplaces with AI
- ➔ Diversifying Hiring with AI

Ethical Considerations

- ➔ Objective or Biased?
- ➔ The Truth About Algorithms | Cathy O'Neil
- ➔ Hiring Tool Shows Bias Against Women
- ➔ Hiring Tool Biased Against Protected Classes
- ➔ AI and Bias in University Admissions
- ➔ Hiring Software and Disability

✨Spotlight Person: Cathy O'Neil

Insurance, Loan, and Financial Aid Approval Algorithms

AI can help us determine appropriate financial decisions, especially when caseloads and needs are high. These systems tend to work similarly to Applicant Tracking Systems, relying on a variety of different algorithms to learn from past data. For example, an AI system could look at the history of who was awarded loans and who paid them back. It can then use this data to predict who is most likely to pay back their loan on time, and then can be used to decide who is awarded a loan. However, similar to ATS, these systems can learn stereotypes or reinforce systemic biases by repeating cultural norms. This can result in systemic barring of certain populations from accessing care, education, opportunity, and financial independence.

Use Cases

🏦loan approval 💸financial aid allocation ⚕️ insurance calculation

Examples of How This Technology Has Impacted Society

- ➔ Benefits of AI for Insurance
- ➔ AI Can Make Bank Loans More Fair
- ➔ AI Helping Students Apply for Financial Aid
- ➔ Enrollment Algorithms Pros and Cons

Ethical Considerations

- ➔ Understanding Algorithmic Bias and Financial Access (video)
- ➔ How AI is Biased in Loan Decision Making
- ➔ Is AI Entrenching Bias in Healthcare?
- ➔ AI and Enrollment Crisis (Financial Aid Algorithms)

✨Spotlight Person: Ruha Benjamin

Medical Software

AI has numerous applications for medicine, from detecting diseases to helping develop new drugs. Hospitals are including AI in a variety of ways – computer vision that looks at radiology scans, or algorithms trained on patient data that help determine health insurance coverage. IBM even launched a project to use AI for cancer diagnosis. However, there have been high-profile failures for many medical AI projects – from IBM's cancer diagnostics tool recommending fatal drug combinations to computer vision algorithms failing to detect skin cancer for darker skin tones.

Use Cases

🩻 radiology 🏥insurance coverage 💊drug research 🚑ER triage

Examples of How This Technology Has Impacted Society

- Empowering Healthcare Providers in Kenya With Technology
- 5 Uses of Medical AI in 5 Minutes
- Using AI to Prevent Blindness
- Benefits of AI in Healthcare
- 5 Ways AI is Transforming Healthcare
- Personalization For Healthcare with AI

Ethical Considerations

- Bias Discovered in Widely Used Hospital Algorithm
- Bias in AI Skin Cancer Detection
- Human Bias Influences AI for Healthcare
- Bias in AI Models for Medical Applications
- Racial Bias in Health Care
- Where Did IBM go wrong with Watson Health?

✨ Spotlight Person: Ziad Obermeyer

Voice Recognition and AI Assistants

Over the past decade, AI voice assistants have increasingly become part of our daily lives. From “Hey Siri?” to “Alexa!”, voice recognition technology has advanced to being able to process speech in real time. These technologies not only add convenience, but also boost accessibility for those who face barriers with accessing screens, text entry, or performing screen gestures. However, there are a few ethical considerations to take into account when it comes to voice recognition and assistants – some people critique how most voice assistants are marketed as female, and are concerned about the privacy risks of a technology that is always listening. Further, these voice assistants tend to make more errors or misunderstandings for speakers with accents or African American Vernacular English (AAVE).

Use Cases

♿Accessibility 🏠Smart Home 🗨Translation 🤹Play

Examples of How This Technology Has Impacted Society

- ➔ Learn to Read with the Power of Voice
- ➔ Project Relate: Bridging Communication Gaps
- ➔ How Voice Assistants Improve Accessibility

Ethical Considerations

- ➔ Privacy Risks of Voice Assistants
- ➔ Voice Assistants Don’t Recognize Voices Equally
- ➔ Voice Assistants All Sound the Same

✨Spotlight Person: Michael Running Wolf

Generative AI

A lot of what we refer to as “AI” is specifically generative AI: a type of artificial intelligence technology that can produce various types of content, including text, imagery, audio and synthetic data. Examples of this include ChatGPT for language, and image generation software like Midjourney or Dall-E for images. These technologies are trained on massive amounts of data, either text or images, and use patterns in the data to generate new content. However, this has implications for the credibility of the content we see – generative AI can make news stories that never happened, even generate video of politicians saying things they never said. Generative AI is also prone to bias, reinforcing stereotypes about race, gender, disability, and more.

Use Cases

🤖Chatbots 🎨Art 🎶Music 🔍Information Search
📰News 🎮Games 🧑🏽‍💻Personalized Content 👩‍🏫Education

Examples of How This Technology Has Impacted Society

- ➔ How to Harness Generative AI for Social Good
- ➔ ChatGPT for Therapy

Ethical Considerations

- ➔ Personalized Disinformation Made by AI
- ➔ Generative AI Images Reinforce Stereotypes
- ➔ Joy Buolomwini Talks About Generative AI
- ➔ Is AI Stealing from Artists?
- ➔ What are Deepfakes?
- ➔ Bias in ChatGPT

✨Spotlight Person: Emily Bender

Risk Scores for Vulnerable Populations

AI can be used to help us make sense of risk. For example, given previous diagnoses, age, and some other data, can we determine how likely someone is to need hospital treatment this year? AI would be able to give us a likelihood of that risk. We also use it to determine the risks of criminal activity, child neglect, security threats, or even detecting the possibility of a student having a learning disability. This is done using a variety of algorithms that use past data to surface patterns and probabilities around risk. However, some of these systems are more invasive and incorrect than they are useful. They perpetuate bias, mistakenly accuse people, and can prevent people from getting opportunity and access.

Use Cases

🚸Child Welfare
📚Detecting Learning Disabilities
⚖️ Criminal Justice
🤝Social Scoring
🔐Assessing Security Risk
🚑Health Risk Scoring

Examples of How This Technology Has Impacted Society

- ➔ Reducing the Risk of Houselessness with AI
- ➔ Can AI Help Homelessness?
- ➔ AI to Support Better Outcomes in Juvenile Justice

Ethical Considerations

- ➔ Machine Bias: Arrest and Risk Scores
- ➔ How They Analyzed the Machine Bias
- ➔ “Social Scoring”
- ➔ AI Bias in Child Welfare
- ➔ AI Assumes Disabled People are Neglectful Parents
- ➔ It Matters How AI Communicates Risk
- ➔ Disability Exclusion
- ➔ AI Decisions Aren’t ‘Objective’

✨Spotlight Person: Renée Cummings

Environment

AI is being successfully applied to help our environment! From forecasting natural disasters to protecting oceans and biodiversity, AI can help us with environment and sustainability challenges. However, creating AI models actually has an environmental cost of its own. Something you may not realize is that anything we do with AI relies on physical computers, which require energy to run. While AI can help us with sustainability, we also need to be aware of the carbon footprint of training large AI algorithms, which may increase greenhouse gasses and cause the global temperature to rise.

Use Cases

🌊Forecasting Natural Disasters

📈Predicting Climate Change

🗑️ Detecting Pollution

🌳Monitoring Biodiversity Decline

🦺Helping People Stay Safe in Weather Events

🐟 Supporting Wildlife

Examples of How This Technology Has Impacted Society

- ➔ Forecasting Floods with AI
- ➔ Helping People Adapt to Heat Waves
- ➔ Tackling Environmental Challenges With AI
- ➔ Protecting Bio Diversity with AI
- ➔ 4 Ways AI Can Help Tackle Climate Change Video

Ethical Considerations

- ➔ The Carbon Footprint of Training AI
- ➔ Environmental Costs of AI Video
- ➔ Pros and Cons of AI For the Environment
- ➔ Using Water to Train AI Models during a Drought

✨Spotlight Person: