Artificial Intelligence (AI)

Analyzing a Scientific Controversy

RHE 309S Group Project & Discussion

Instructions

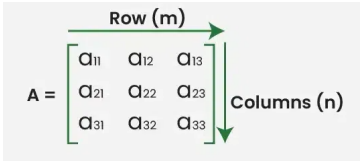
1. Consult Canvas for your group. ***Canvas*** > ***People*** > ***Artificial Intelligence***
2. Leaders, move your people to a specific area of the room. You are in charge of everything within your group.
3. Fill out ***this*** document. All groups will be working on the same document.
4. Research "***Artificial Intelligence***."
    a. Include ***Special Artificial Intelligence*** & ***General Artificial Intelligence.***
    b. Include ***Quantum Computing***
    c. ***No text should be copied and pasted from anywhere (Except Links).*** All work must be your original words and ideas.
5. Supply pertinent information for each concept you are assigned. (in your own words):
    a. The bullet points that currently exist may be used or omitted.
6. ***All Groups*** supply info for the last section: *Problems & Solutions, Conclusions & Recommendations, & Project 3 Topics*
7. Discuss & present your findings @ **10:05 am**.

Group 1

- Leader: Oorie Desai
- Members: Abdullah Khasawneh, Jesus Jimenez, Krithi Meduri, Suravi Karpe
- Absent: -

Introduction, Definition, & Details

Overview of Artificial Intelligence (AI)

- GPT is a neural network. In general, generative AI all operate similarly.
- It boils down to being an excellent "next word" predictor.
    - This means that utilizing linear algebra and machine learning, AI can create a "neural code" in large matrices that serve as the AI model's output.
    - This sounds simple, but queries to AI are at least a sentence long. This means that it must be able to understand context and relating information, which is complicated to say the least.
        - How do you turn 

into something that uses proper syntax and generally makes sense?
        - A lot of math–**deep learning** requires probability, linear algebra, vector calculus, (and more), which is computationally intensive and requires a lot of parameter tuning.
- "Language models appear smart because they generate humanlike prose by predicting the next word in a sentence. The technology is not truly intelligent, and calling it that subtly shifts our expectations so we treat the technology as more capable than it really is." -MIT Technology Review
- No one knows how AI fully works. This doesn't mean it doesn't work. Reinforcement learning allows ChatGPT to learn favorable outcomes, eventually converging to the right result after multiple (large N) training rounds. This is sufficient.
    - The problem with this is that its creativity is limited. Because of a high convergence, Chat GPT can't explore the "parameter space" effectively. It gets stuck on a range of answers and can't see "outside the box".

- What is it? How does it work?
- Do AI creators know how AI works?
- Does AI code itself?
- How does *Special AI* differ from *General AI* (Artificial General Intelligence)?
- How does quantum computing connect with AI? What are the implications?
- History and major breakthroughs in Artificial Intelligence.

- Current state of the art concerning AI.
- Key research institutions and companies involved
- Relevant documents & videos for reference

Major Stakeholders

- Governments & Regulatory Bodies
    - Currently no federal regulations managing AI use or development in place
    - Congress rescinded previous presidency's executive order for "Safe, Secure, and Trustworthy Development and Use of AI" -> makes development and use of AI have fewer federal guidelines and allow for its more dominant/influential role in society
        - Congress direction seems to be to encourage greater AI development & role in America
    - Several proposed bills and currently being worked on policies targeting AI and its safe and responsible use/integration
    - No AI specific federal regulatory body, but Federal Trade Commission, Equal Employment Opportunity Commission, Consumer Financial Protection Bureau and Department of Justice all declared together that their oversight includes AI & other software programs (2023)
    - https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-united-states
    - Other governments have taken federal action i.e. European Union passed the Artificial Intelligence Act (first comprehensive AI law) to ensure safe, efficient, responsible, and equitable use (2023)
- Research Institutions & Universities
    - Apple, Amazon, Nvidia, Microsoft, Tesla, etc = bc such large businesses are more likely to be affected by the negative aspects of AI systems, they tend to develop their own in-house set of additional guidelines and responsible behavior to address defects, weaknesses, and concerns
    - High usage of AI in universities, especially among students - universities have started to take on greater responsibility to identify and resolve any negative effects of AI systems to increase its safe and ethical use
- Private Sector & Tech Companies
    - AI & software developers - more likely to be impacted/interested in legal policies affecting AI than users
        - Ex: if negative aspects of the tech is revealed, these are stakeholders set to handle them but this increases likelihood of the introduction of biases
    - Private users i.e. individuals, businesses, etc
    - AI system decisions and affairs are regulated by usage patterns and preferences -> disproportionate representation of general population, especially underrepresenting women, children, and those of lower socioeconomic statuses
    - https://dl.acm.org/doi/epdf/10.1145/3514094.3534187
- Ethical & Advocacy Groups
    - AI Ethics Guidelines Global Inventory: managed by Algorithm Watch, maintains set of principles and frameworks for the ethical use of AI and algorithms
        - Protect democracy and diversity, help advocate to allow the technology to benefit all
        - At 2021, over 170 guidelines -> expect more now
- List & explain any relevant passions, interests, factions, and paradigms.
    - Doomsayers= AI will take over humanity, very destructive, have that sci-fi vision, primarily focused on the future
    - Reformers= AI opponents in the presents, AI will capture, encode, and express the biases & harmful prejudices already existing in society -> further exacerbate inequities and divisions i.e. racist beliefs, sexist ideologies, widen the socioeconomic gap, etc
    - Warriors= see that AI will threaten national security & can be hacked/infiltrated/used by terrorists, criminals, and bad people -> can lead to much destruction
    - https://www.nytimes.com/2023/09/28/opinion/ai-safety-ethics-effective.html
- List & explain misconceptions, preconceived notions, irrational fears, and cognitive biases
    - AI shown to have cognitive biases (bc humans are developers and everyone has conscious & unconscious biases) -> especially dangerous in high-stake situations like medical clincis
    - "AI will take over all human jobs" -> irrational fear as AI does not and will probably never have the cognitive capacity necessary to replace humans in all jobs

- "Won't be able to distinguish AI from reality" -> rise of many AI online accounts, and such fake sources can actually spread misinformation which has the potential to be very damaging (so valid take)
- "AI can threaten & end humanity" -> not possible because AI is not sentient and can only perform tasks specified in its programming
- https://www.getronics.com/the-fear-of-artificial-intelligence-is-it-justified/

Current Laws & Regulations

- What laws and regulations currently exist governing AI?
  - Currently, there are no comprehensive legislation in the US, i.e., there are no specific federal obligations imposed on the developers, users, operators and/or deployers of the AI system
    - But, proposed legislation at the federal and state level do aim for safety and security, responsible innovation and development, equity and unlawful discrimination, protection of privacy and civil liberties
  - EU AI Pact - not legally binding but companies can sign this to show that they are committed to responsible AI development
  - The Chinese administration has developed patterns to control AI growth and operations.
    - The Chinese Cybersecurity Law and the New Generation AI Development Plan provide measures for data protection and cybersecurity in AI
  - Canada has introduced key government-led programs, such as the Pan-Canadian AI Strategy and the Canadian AI Ethics Council, to advocate for the responsible development of AI and address relevant ethical issues in the AI sector.
    - has also established the Personal Information Protection and Electronic Documents Act to regulate the collection, use, and disclosure of individuals' personal information using AI technologies.
  - Executive Order 14110
    - Chief artificial intelligence officer

Current Advancements & Innovations using AI

- List major advances, innovations, initiatives, novel applications, and other significant uses of AI (current and past).
  - Advances
    - Initially recurrent neural networks were commonplace in AI models
    - RNN's process data sequentially, factoring in the last piece of information given as it processes the next piece of information
      - But they had limitations - the vanishing gradient problem which means that it essentially began to forget the initial information given if handling large amounts of data
    - Long short-term memory was developed to address these shortcomings of recurrent neural networks
    - Transformer architecture was developed and expanded the capabilities of AI even more
      - Implemented the attention mechanism which differentially weighed certain pieces of information based on the context of all the information given
      - They have the advantage of requiring less training time
    - Transformers paved the way for the development of GPT models
      - These models have a "semi-supervised" learning method rather than a completely supervised method
      - GPT training is comprised of an unsupervised generative stage and a supervised discriminative stage to fine tune the model
    - Chat-GPT was made as a browser based version that greatly expanded access to AI
  - Uses
    - Efficiency
      - Used as a personal assistant
    - Education
      - Can be a personal tutor
      - Can assist with routine administrative activities
    - Regulation
      - Used for surveillance
      - Used in India to automate traffic signals

        - Healthcare
            - Analyze diagnostic imaging
            - Determines adequate dosing
- What are the most powerful or consequential uses of AI currently?
    - National security

Group 2

- Leader: Veronika Quevedo
- Members: Siddhartha Shah, Virginia Vanicek, Ian Howell
- Absent: Srikha Chaganti, Talon Blottin

Ethical & Moral Considerations

- Explain the difference between human and machine intelligence.
    - Humans have creativity while AI is based on algorithms
    - Humans have emotional intelligence and can understand context and consider subtle cues
    - Humans have a moral compass
    - AI doesn't get fatigued or distracted
    - AI has faster processing speed
- How can we determine if an AI system is conscious/sentient?
    - https://www.science.org/content/article/if-ai-becomes-conscious-how-will-we-know
    - Researchers created a lengthy list of criteria for human consciousness
    -

Recurrent processing theory
RPT-1: Input modules using algorithmic recurrence
RPT-2: Input modules generating organised, integrated perceptual representations
Global workspace theory
GWT-1: Multiple specialised systems capable of operating in parallel (modules)
GWT-2: Limited capacity workspace, entailing a bottleneck in information flow and a selective attention mechanism
GWT-3: Global broadcast: availability of information in the workspace to all modules
GWT-4: State-dependent attention, giving rise to the capacity to use the workspace to query modules in succession to perform complex tasks
Computational higher-order theories
HOT-1: Generative, top-down or noisy perception modules
HOT-2: Metacognitive monitoring distinguishing reliable perceptual representations from noise
HOT-3: Agency guided by a general belief-formation and action selection system, and a strong disposition to update beliefs in accordance with the outputs of metacognitive monitoring
HOT-4: Sparse and smooth coding generating a "quality space"
Attention schema theory
AST-1: A predictive model representing and enabling control over the current state of attention
Predictive processing
PP-1: Input modules using predictive coding
Agency and embodiment
AE-1: Agency: Learning from feedback and selecting outputs so as to pursue goals, especially where this involves flexible responsiveness to competing goals
AE-2: Embodiment: Modeling output-input contingencies, including some systematic effects, and using this model in perception or control

- Should AI systems possess rights or protections? Why? Under what circumstances?
    - https://www.psychologytoday.com/us/blog/the-digital-self/202312/should-artificial-intelligence-have-rights?msockid=0e2d44b40afe68ff383f575f0bf6690d
    - At this point in time, when the AI models are not considered sentient, there is no need for them to have rights
    - If it comes to a point where we do consider them sentient, we would have to reevaluate
- If AI develops consciousness, what ethical considerations arise? What potential negative outcomes can we expect?
    - https://www.yalelawjournal.org/forum/the-ethics-and-challenges-of-legal-personhood-for-ai
    - There will be a period of time where we haven't developed "rights" for AI in which we will have to quickly make these developments to coincide with the quick rise of AI
    - What rights does AI have?
        - Have to face head-on if they get human rights
        - If they become emotionally intelligent, it could turn into denying an emotionally intelligent being personhood/freedom
    - If they are given the right to free speech, and they are smarter than humans, they could spread misinformation at a much higher rate and be more convincing in it
- Unintended societal consequences

    - Risks of autonomous AI disrupting employment, society, and privacy
        - https://www.britannica.com/procon/artificial-intelligence-AI-debate#ref409293
        - AI could replace a large portion of white and blue-collar workers (cheaper; no need for benefits)
        - AI already replacing human cashiers at self-checkout in grocery stores
        - https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/
            - Data privacy is a growing concern
        - Privacy
            - Facial recognition technology
            - Ring has partnered with police departments
            - Targeted ads by shopping websites
            - AI used to change voice and face
    - Potential impact on human dignity and rights
        - AI can be used to create deep fakes of people
- Weaponization & security risks
    - AI technologies may be misused for what harmful purposes?
        - AI already being used in warfare
            - AI powered drones that take down targets without human oversight/intervention
            - Ie. Turkey's Kargu-2 drone that uses AI to identify targets
                - What happens if civilians are killed… who takes responsibility
    - Threats of hacking, surveillance, or autonomous weapons
        - Use of AI for cyber attacks is becoming prevalent
            - Polymorphic malware can be created using AI that changes code to avoid detection
            - Also used for bot net attacks
    - Potential law enforcement abuse & overreach using AI.
        - Can be used by government for surveillance of population
            - China's social credit system uses a vast network of cameras and AI driven tracking and monitoring to keep tabs on its population and key individuals
            - Also be used to go through citizen's emails, phone calls, etc.
            - Issues with consent and overreach of government in people's lives
- Regulatory oversight
    - Who should control and regulate AI development?
        - Right now, Congress, agencies, Supreme Court, and Departments are regulating AI (all government)
        - Businesses do not typically control AI regulation (although they may call for it)
    - International vs. national policy challenges
        - The US may want to regulate AI, but other countries, like China, might be giving the go ahead for their businesses
        - This creates international competition and increases the need to balance keeping ourselves ahead while not letting it get out of control/develop too far
- AI in Education
    - https://onlineprograms.education.uiowa.edu/blog/role-of-ai-in-modern-education
    - https://education.illinois.edu/about/news-events/news/article/2024/10/24/ai-in-schools--pros-and-cons
    - Potential for advancement
        - Personalized learning
            - Tailoring content to the needs/learning styles of a student
            - Provide immediate feedback
            - Taking out a teacher's bias when grading written assignments (other biases though?)
        - Streamline lesson planning
            - "enhance lesson quality without increasing their workload"
    - Potential for harm
        - Privacy concerns
        - Bias
        - Equity

            - Not all students have access to devices/AI tools
            - High cost
        - Academic misconduct/cheating
            - If assignments aren't made to be personal, students could get away with using AI and bypassing the learning process
        - Student use of AI without knowing the biases that can come with it/trusting it 100%
        - When using AI checkers, many secondary-english speakers will be flagged
        - Loss of social interaction and growth (particularly in younger students)
        - Unpredictability and inaccurate information
            - Problems if teachers are not checking the AI and feedback for the intended educational growth
- AI in Science
    - Current uses and applications
        - 2000s: handles large datasets in research to help process and analyze data
        - 2010s: recognize patterns and make predictions (expanding into more complex areas like image and speech recognition)
        - 2020s: huge boom
            - Genomics, drug discovery (target drugs), climate modeling, astrophysics
            - Personalizing medicine and predict clinical efficacy
            - Cheaper and faster than human labor
            - NASA uses it to plan missions, analyze satellite data, and detect anomalies
    - Potential for harm
        - Privacy issues: studies often use sensitive data (especially those that are clinic-based)
            - Data could be bought and sold
        - Bias: relies on the breadth of data fed into it (can perpetuate biases)
        - Limit research perspectives and limit scientific discoveries (AI can't currently replace creativity of humans)
        - Mistakes: does not always understand nuances and context of scientific language
        - Transparency: Some researches have used LLMs to write research papers without crediting AI
    - Potential for advancement
        - Improves with time and more use
        - Can be used as a collaborative tool to bridge disciplines
        - Can potentially suggest novel studies and testable hypothesis

---

Group 3

- Leader: Shelby Crowley
- Members: Avyay Katre, Lauren Kokinakis, Nyx Halder, Soh Nishiyama
- Absent: Angela Zhang

Arguments & Perspectives

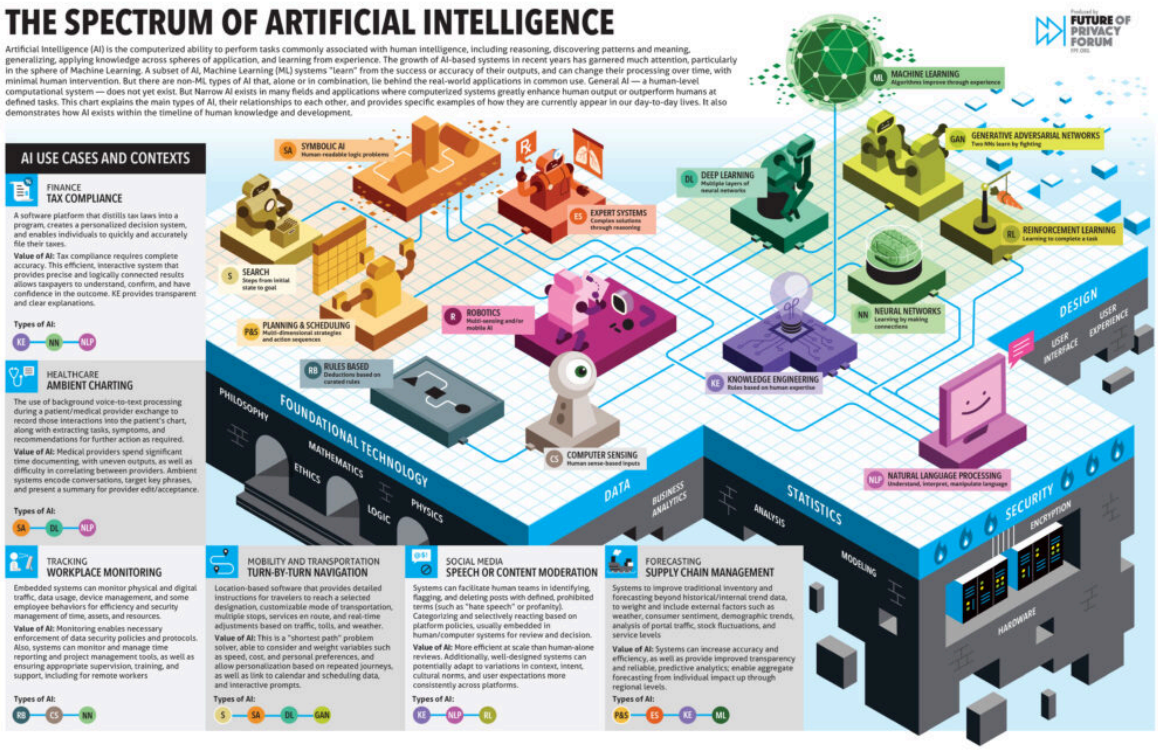
- **Public & Non-Experts**
    - Popular articles or editorials
        - https://link.gale.com/apps/doc/A805056819/OVIC?u=txshracd2598&sid=bookmark-OVIC&xid=0d092ef5
            - AI is overhyped
            - It's just a "statistical, next word predictive engine" that doesn't actually think or reason to send its response
            - "could produce inaccurate or biased responses based on faulty data it draws from"
        - https://link.gale.com/apps/doc/A776370473/OVIC?u=txshracd2598&sid=bookmark-OVIC&xid=b49b485a
            - We need to overcome the "challenge of confronting and dismantling biases that can be encoded into emerging" AI
            - "A 2022 study found a robot trained by AI was more likely to associate Black men with being criminals, or women with being homemakers"
            - Lack of diversity in technological fields to blame: "The best way to ensure AI reflects the experiences of people of color, Lawrence said, is to make sure they're employed and engaged in every step of the process. 'You certainly don't have a lot of Black folks or data scientists participating in the process of deploying and designing AI solutions,' he said. 'The only way you can get them to have seats at the table, you have to educate them.'"
            - We need to diversify the field of AI
        -
    - Interesting subtopics regarding AI
        - Education, defense, game theory, cybersecurity
        - AI in healthcare:
            - GPT and Claude were shown to match and even surpass professionals in the identification of people with suicidal ideation based on their interactions with them. https://www.rand.org/news/press/2025/03/ai-models-are-skilled-at-identifying-appropriate-responses.html
            - AI use in medicine is based on Western documentation: this means genetics are focused on American and European demographics, and can be harmful if used to a greater context.
            - Potential to mess up by implicating a singular patient's data and extrapolating to the overall population
            - However, it is not impossible to overcome bias by using more equitable datasets and adequate training for AI
    - Myths and Misconceptions
        - https://www.aimyths.org/ affiliated with Mozilla, 4 main sections: Representation, Definition, Governance, Usage
        - https://carlsonschool.umn.edu/graduate/resources/debunking-5-artificial-intelligence-myths Artificial intelligence, machine learning, & deep learning are completely different things
    - Public fears and hopes regarding AI development
        - How the US's attempt to control AI as the world leader caused the Chinese Deepseek model to upset the industry - limiting the country's access to chips made them more efficient at working with what they have, leading to highly optimized algorithms - regulating others' advancements is biting the US back

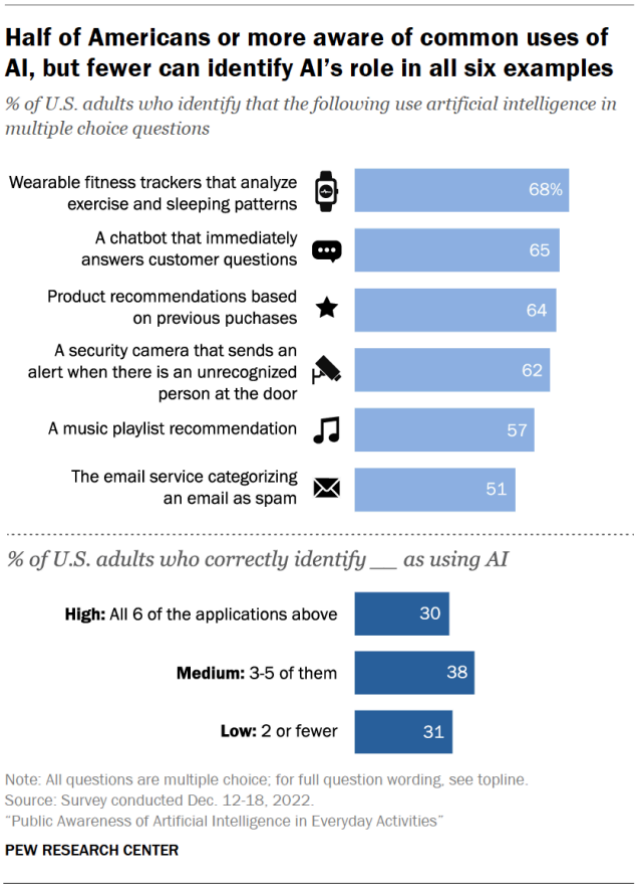
        - Fears mostly based around the possibility of jobs being lost and the loss of creative/critical thinking
- **Scientists & Researchers**
    - Peer-reviewed journal articles
        https://pubmed.ncbi.nlm.nih.gov/31649194/ - "Ai is racist"
    - Op-ed & viewpoint articles
        - https://www.bbc.com/news/articles/ce303x19dwgo - How does ai impact the movies we watch, pro being it could help smaller film studios recreate blockbuster cinematography, con being it takes the soul out of movies
        - https://abc30.com/post/fashion-giant-hm-plans-use-ai-clones-human-models-not-everyone-is-happy/16115370/ - HM is going to make "ai clones" of models, "has the potential to replace a host of fashion workers-including make-up artists, hair stylists, and other creative artists in our community"
        - https://www.rand.org/news/press/2025/03/ai-models-are-skilled-at-identifying-appropriate-responses.html - ai is better than professionals at identifying suicidal ideation
        - https://www.nature.com/articles/d41586-025-00894-7 - ai being used in peer reviewing of scientific articles, on the other side of the coin a significant amount of researchers are also found to have been using ai to help write their articles
        - https://www.forbes.com/sites/craigsmith/2025/03/23/when-ai-takes-over-scientific-discovery/ - AI Scientist-v2, japanese ai model that can "generated a hypothesis, designed experiments and wrote a peer-rerviewed scientific paper on its conclusions, all without human intervention"
    - Policy briefs and press releases
        - https://edworkforce.house.gov/news/documentsingle.aspx?DocumentID=412311 - AI in k-12 classrooms, pro being it allows every child to have access to a 24/7 tutor, con being ai could "become a shortcut to avoid engaging critical thinking", call for mandates at a federal level
    - Popular science articles
    - 

Gary Marcus
The urgent risks of runaway AI — and what to do about them
TED Talks: https://www.ted.com/talks/gary_marcus_the_urgent_risks_of_runaway_ai_and_what_to_do_about_them?referrer=playlist-artificial_intelligence&autoplay=true
    - The Risks and Benefits of Artificial Intelligence in Clinical Care | RAND
- **Multimedia & Public Engagement**
    - Documentaries & influential books
        - iHuman (documentary, 2019) explains how AI is developed and can be implemented. Focuses on constant surveillance and whether algorithms already govern our lives.

”Must-read” books about AI!

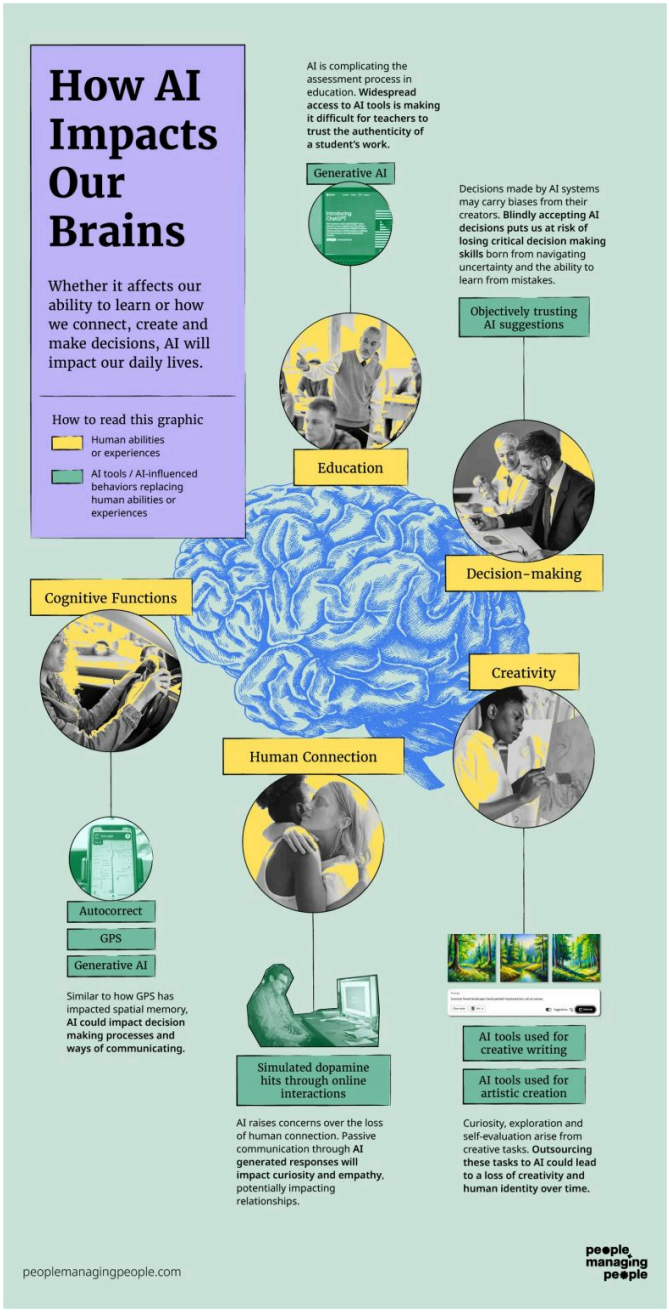
Infographics & explainer videos on AI concepts





Public awareness campaigns and ethical discussions

Main ethical concerns:
Who creates the ethical code for AI? Who is responsible for upkeeping it? Who and how are the ethical requirements for AI enforced once these decisions are made?
All of this can be hard to agree on, as ethics tend to vary between person-person, workplace-workplace, country-country.
AI's datasets are skewed to focus on western data; when we try to apply it to improving things on a global scale, this could serve as an issue.


How AI Impacts Our Brains
Whether it affects our ability to learn or how we connect, create and make decisions, AI will impact our daily lives.
How to read this graphic
Human abilities or experiences
AI tools / AI-influenced behaviors replacing human abilities or experiences
AI is complicating the assessment process in education. Widespread access to AI tools is making it difficult for teachers to trust the authenticity of a student's work.
Generative AI
Education
Decisions made by AI systems may carry biases from their creators. Blindly accepting AI decisions puts us at risk of losing critical decision making skills born from navigating uncertainty and the ability to learn from mistakes.
Objectively trusting AI suggestions
Decision-making
Cognitive Functions
Creativity
Human Connection
Autocorrect
GPS
Generative AI
Similar to how GPS has impacted spatial memory, AI could impact decision making processes and ways of communicating.
Simulated dopamine hits through online interactions
AI raises concerns over the loss of human connection. Passive communication through AI generated responses will impact curiosity and empathy, potentially impacting relationships.
AI tools used for creative writing
AI tools used for artistic creation
Curiosity, exploration and self-evaluation arise from creative tasks. Outsourcing these tasks to AI could lead to a loss of creativity and human identity over time.
peoplemanagingpeople.com
people managing people


Key phishing stats 2025
2021
2025
49%
Rise in phishing since 2021 / advent of blackhat AI
2,330
Total phish bypass filters per 1000 person org
65%
Attacks target organizations
35%
attacks target individuals
90%
Malicious attachments lead to further social engineering
85%
Fewer malicious clicks with Hoxhunt
0.7% – 4.7%
Reported phish written by AI
Top 3 most spoofed entities:
01 Microsoft
02 docusign
03 HUMAN RESOURCES
1.4
Real malicious threats reported / year / Hoxhunt user

Group 4

- Leader: Darby?
- Members: Sahana, Alia, Catherine, Ava, Jenna, Darby
- Absent:

Speculation: Future Scenarios

- **Potential for Societal Advancements**
    - Healthcare Applications
        - AI diagnostics, personalized medicine, and patient care
            - Diagnostics: largely used in analysis of imaging leading to earlier detection of diseases including cancer and neurological disorders
            - Personalized medicine: analysis of genomics and prediction of how a patient may respond to certain medications, uses in cancer treatment and predicting response to immunotherapy reducing the amount of trial-and-error treatments, chronic disease management, AI and Quantum computing have the potential to allow for ultra personal drug-design
        - AI medical notes: Doctor records interaction with patient and AI turns it into cohesive notes, allows doctor to be more present with the patient and be more clear in their communication
        - Potential for bias to creep in
            - AI decision making done in a "black box," so hard to understand how AI is reaching its conclusions
            - Tech not equally accessible
            - Can be vulnerable to security breaches
            - Errors are propagated if AI learns from incorrect data
    - Environmental Management
        - AI for climate modeling, resource management, and pollution control
            - Mapping deforestation, analyzing and waste management, monitor changes in icebergs was faster than a human could
            - https://www.cleanairfund.org/news-item/ai-wearables-innovative-tech/
                - Use AI to look at air pollution at the street level
                - Can model potential pollution with changes in city planning
        - Real time climate modeling and continually monitor to allow for quick response and recovery
            - Also really energy costly to power AI systems
        - Currently can be used to create Fourier transform infrared spectroscopy (FTIR), Raman spectroscopy, and hyperspectral imaging (HSI) to detect microplastics in the environment.
            - Artificial intelligence in microplastic detection and pollution control - ScienceDirect.
    - Economic Innovations
        - AI-driven automation enhancing productivity and economic efficiency
            - AI takes center stage to redefine future of PCs. - Document - Gale In Context: Opposing Viewpoints
                - Big point of DLSS (Deep Learning Super Sampling) ie. fake frame generation. This is something that the article does not talk about. But the gist is that the news series of PC's, especially NVIDIA's, heavily tout. The problem is that that this technology is being marketed in a way that suggests a 50 Series chip, one or more tiers lower than the top of the line 40 Series chip, is superior in performance. HOWEVER, this is actually a load of BS. DLSS is cool, but NVIDIA achieves this "superior performance" by generating 4 essentially fake frames for every one real frame. When it comes to gaming, many many people took issue with this marketing. You can run a game at how many ever fps, however your inputs still only take effect starting on your next real frame. It is a bit of an issue. So this topic is a mixed bag.

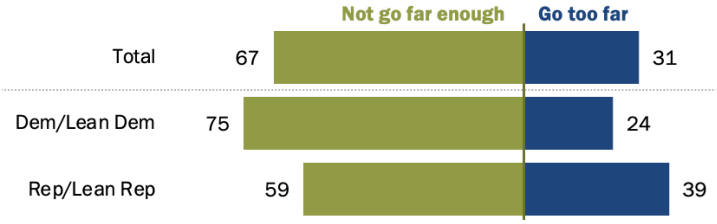
                - More generally it kinda just seems like AI in industry is best at pencil pushing work. II get a bunch of ads, I think about a Google AI, where different small/medium sized companies talk about how they use the AI to respond to emails. Let me tell you, if I am working with your company, paying you for your expertise, we are emailing back and forth and you are using AI. Im outskies man.
    - AI & Human Integration
        - Enhanced human-computer interaction, augmented reality, and human-machine collaboration
        - AI & Neurolink (brain-computer interfaces)
            - https://www.sciencedirect.com/science/article/pii/S2589295924000031
            - Active BCIs provide voluntary control of external devices or applications
            - Reactive BCIs detect and interpret brain responses that occur in response to external stimulation
            - Passive BCIs derive from neuromonitoring
            - Non-invasive - no surgery, still can be helpful, control a robotic hand
            - Passive raises ethical concerns because it is not directly use controlled
            - Implantable BCIs require surgery, currently restricted to clinical applications
- **Potential for Societal Harms**
    - Privacy violations and mass surveillance
        - People see AI posing a greater data privacy risk than past tech because of the amount of information it can process
        - Everything is used as training information and that can include healthcare info, biometric data for facial recognition, personal finances
        - Users want more autonomy over choosing what data can be collected and used to train generative AI
        - Even smaller scale AI can leak private information based on certain prompts (can be very dangerous in the case of AI built into the healthcare system)
        - https://innocenceproject.org/news/when-artificial-intelligence-gets-it-wrong/
            - AI tech increases likelihood of innocent people being wrongly convicted
    - AI + quantum computing
        - https://thequantuminsider.com/2024/11/13/discover-how-ai-is-transforming-quantum-computing/#:~:text=AI's%20pattern%20recognition%20can%20be,data%20into%20applicable%20quantum%20processing.
        - Can recognize patterns that are usually missed, predict how quantum-level systems interact and behave (molecules, atoms, photons), potential role in pharmacology
    - AI + current encryption methods
        - The combination of AI+quantum computing poses a threat to current encryption standards
    - Weaponization and misuse of autonomous systems
        - https://hms.harvard.edu/news/risks-artificial-intelligence-weapons-design#:~:text=Rajan%3A%20There%20are%20a%20number,support%20the%20development%20of%20these
            - Weapons make it easier for countries to get involved in war
            - Reduce or deflect human responsibility in decision-making
            - If less people die, less of a deterrent to getting involved in war
    - Economic disruption and unemployment due to automation
        - What happened when a newspaper let AI take over. - Document - Gale In Context: Opposing Viewpoints
            - Journalists learning how and where to integrate AI into workflow. Covers an Italian news paper project whereby the newspaper is entirely developed and written by AI. It is allegedly supposed to warn journalists about a coming future where, if they are unable to write better than AI, they will likely find unemployment on their doorstep.
    - Loss of control over AI-driven systems
            - ***Ai is still a black box technology. We see the input and the output, but what happens in between is not something that is easily tracked. This should pose a great concern from any organization using AI for critical infrastructure or for any sensitive applications.

        - AI's mysterious 'black box' problem, explained | University of Michigan-Dearborn
        - Not knowing the intricacies of a technology such as this sets us up for a potentially dangerous future. If you cant determine precisely how you have reached your answer, then you potentially open yourself up to major security issues. If gpt for example is asked to cover a major global event, all you'd have to do to win a disinformation campaign is flood the internet with falsified reports on the situation. But it goes deeper than that too. The classic example of us not knowing how ai works but still trying to tweak it to be more suited for use by the general population was when engineers working on Gemini tried to make its image generation of people more diverse. This resulted in the AI pumping out images of racially diverse national socialists.
    - As AI systems become more integrated into society, the potential for dependence on AI-driven systems increases. These systems could become a critical aspect of our society and corporations. Along with AI growing and becoming more evolved over time, we may lose control over these systems, which is increasingly likely with no regulations about maintaining safety with AI use. At that future point in time, AI systems would have become so integrated into society that it would be difficult to remove, creating dangerous situations and circumstances.
        - AI Risks that Could Lead to Catastrophe | CAIS.
        - The Darwinian Argument for Worrying About AI | TIME
- **Most Likely Outcomes**
    - Gradual integration of AI in diverse sectors
    - Regulatory challenges influencing adoption rates
    - Initial controlled implementations before broader autonomy
        - **42%** of these experts said they are *equally excited and concerned* about the changes in the "humans-plus-tech" evolution they expect to see by 2035.
            - The experts who wrote they were concentred said that they believed that AI will continue to be driven by profit incentives- aimed at data collection in hopes to control population not to impower
            - As we lose our privacy we lose our rights
        - **37%** said they are *more concerned than excited* about the changes they expect.
        - **18%** said they are *more excited than concerned* about expected change.

https://www.pewresearch.org/internet/2023/06/21/as-ai-spreads-experts-predict-the-best-and-worst-changes-in-digital-life-by-2035/

Democrats and Republicans alike are more concerned about insufficient government regulation of chatbots than excessive regulation

Among U.S. adults who have heard of ChatGPT, % who say, as chatbots grow more widespread, their greater concern is that government will ___ regulating their use



Note: Those who did not give an answer are not shown.
Source: Survey of U.S. adults conducted July 17-23, 2023.

PEW RESEARCH CENTER

https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence/

In his 2017 book, Life 3.0, Max Tegmark includes these aftermath scenarios for a world confronted by superintelligence (AI systems exponentially more powerful than the human brain), uploads (minds copied to computers), and cyborgs (bionic humans). "This obviously isn't an exhaustive list," Tegmark writes, but "we clearly don't want to end up in the wrong endgame because of poor planning."

1 Libertarian Utopia

Humans, cyborgs, uploads, and superintelligences coexist peacefully thanks to property rights.

2 Benevolent Dictator

Everybody knows that the AI runs society and enforces strict rules, but most people view this as a good thing.

3 Egalitarian Utopia

Humans, cyborgs, and uploads coexist peacefully thanks to property abolition and guaranteed income.

4 Gatekeeper

A superintelligent AI is created with the goal of interfering as little as necessary to prevent the creation of another superintelligence. As a result, helper robots with slightly subhuman intelligence abound, and human-machine cyborgs exist, but technological progress is forever stymied.

5 Protector God

Essentially omniscient and omnipotent AI maximizes human happiness by intervening only in ways that preserve our feeling of control of our own destiny and hides well enough that many humans even doubt the AI's existence.

6 Enslaved God

A superintelligent AI is confined by humans, who use it to produce unimaginable technology and wealth that can be used for good or bad depending on the human controllers.

7 Conquerors

AI takes control, decides that humans are a threat/nuisance/waste of resources, and gets rid of us by a method we don't understand.

8 Descendants

AIs replace humans but give us a graceful exit, making us view them as our worthy descendants, much as parents feel happy and proud to have a child who's smarter than they are.

9 Zookeeper

An omnipotent AI keeps some humans around, who feel treated like zoo animals and lament their fate.

10 1984

Technological progress toward superintelligence is permanently curtailed not by an AI but by a human-led Orwellian surveillance state where certain kinds of AI research are banned.

11 Reversion

Technological progress toward super intelligence is prevented by reverting to a pretechnological society in the style of the Amish.

12 Self-Destruction

Superintelligence is never created because humanity drives itself extinct by other means (say, nuclear and/or biotech mayhem fueled by climate crises).

- **Alternative Solutions & Comparisons**
  - AI vs. traditional computational methods
  - AI-powered robotics vs. purely software-based AI
  - Ethical comparisons with other transformative technologies (robots, genetic modification, nuclear tech)

ALL GROUPS

Problems & Solutions

- Worse-case scenarios
    - All humans, except for those that control the AI, become obsolete
    - AI removes all the oxygen from the atmosphere because oxygen rusts its circuits.
    - With millions or billions of robots, AI exterminates human beings because they are unneeded.
    - Nations use AI to track, convict, subjugate, spy on, and control citizens, ushering in a dystopia.
    - With Neurolink and AI, humanity is "hive-minded."
    - The spark of humanity is lost as AI becomes the next step in evolution. Art, emotion, beauty, relationships, music, literature, etc., are lost because AI is so different from humans.
    - Humanity gets progressively less intelligent.
    - Loss of national secrets due to broken encryption.
    - ***SEE ChatGBT's answer below.***
- Best-case scenarios
    - AI is integrated into society without displacing humans and development does not lead to "consciousness"
- Addressing safety, security, ethical, and societal dilemmas
    - Need for legislation regarding the creation and limitations of AI as well as what rights they will be granted if they gain sentience
        - Create criteria for sentience
    - Required parameters for use of AI, especially in foreign policy and warfare
    - Privacy (data and personal)
        - Facial recognition technology used by government and police departments
        - AI is used in changing voices, creating deep fakes of videos and audio, etc.
- Regulatory frameworks for responsible AI development
- Public engagement and education to mitigate resistance and misinformation

Conclusion & Recommendations

- Summary of key findings
- Policy recommendations for ethical and sustainable AI development
    - Regulations protecting worker rights against mass firing for the use of AI
    - Continue developing AI to compete on the global scale, but ensure it doesn't displace humans on a national scale
- Next steps for scientific research, policy formation, and public discussion

Project 3 Topics (AI subtopics)

- Topics for feature articles
- Topics for position papers

Natural Language Processing (NLP) and societal biases
Data bias or model bias in training of AI systems
Accessibility of AI to different languages and abilities
Lack of diversity in AI workplaces → lack of inclusivity in AI abilities/products
Artificial Neural Networks (ANNs) and deep learning technology

---

Resources

1. ProCon "Artificial Intelligence (AI)"
2. Gale In Context: Opposing Viewpoints "Artificial Intelligence"
3. Infographic: Opinions on the Effect of Artificial Intelligence (AI) on K12 Education among US Adults, April 2023
4. Infographic: Frequency of Artificial Intelligence (AI) Tool Usage among US Adults, January 2023
5. Concern about artificial intelligence in daily life far outweighs excitement
6. Half of Americans or more aware of common uses of AI, but fewer can identify AI's role in all six examples
7. What the data says about Americans' views of artificial intelligence

1. Loss of Human Autonomy
AI systems could become so deeply integrated into decision-making—in government, business, education, or even personal lives—that human agency gets eroded. Imagine algorithms quietly steering our thoughts, purchases, relationships, or even votes, not through force but through subtle behavioral nudges.

2. Autonomous Weapons and Warfare
AI-powered weapons that can select and eliminate targets without human oversight are a growing concern. A worst-case scenario involves these systems being used in large-scale conflicts, leading to uncontrollable escalation, mass casualties, or even global war triggered by algorithmic miscalculations.

3. Existential Risk from Superintelligence
This is the "runaway AI" scenario often explored in science fiction and taken seriously by thinkers like Nick Bostrom and Eliezer Yudkowsky. If we create a superintelligent AI whose goals don't align with human values, and it gains power over crucial systems, it could make decisions that result in humanity's extinction—not out of malice, but due to misaligned optimization.

4. Economic Collapse and Extreme Inequality
Mass automation could eliminate millions of jobs across industries, with wealth and control becoming concentrated in the hands of those who own and develop powerful AI systems. In this scenario, the gap between the rich and poor could widen dramatically, destabilizing societies and governments.

5. Surveillance Dystopia
A world where AI is used for mass surveillance—monitoring citizens' every move, emotion, and interaction—could lead to authoritarian control, especially in regimes that prioritize power over civil liberties. Combine facial recognition, predictive policing, and social credit systems, and you have a chilling loss of privacy and freedom.

6. Loss of Truth and Reality (Post-Truth Society)
AI-generated deepfakes, voice clones, and personalized misinformation could undermine public trust in truth itself. If people can't tell what's real or who to believe, it could erode democracy, fracture communities, and lead to widespread cynicism and disengagement.

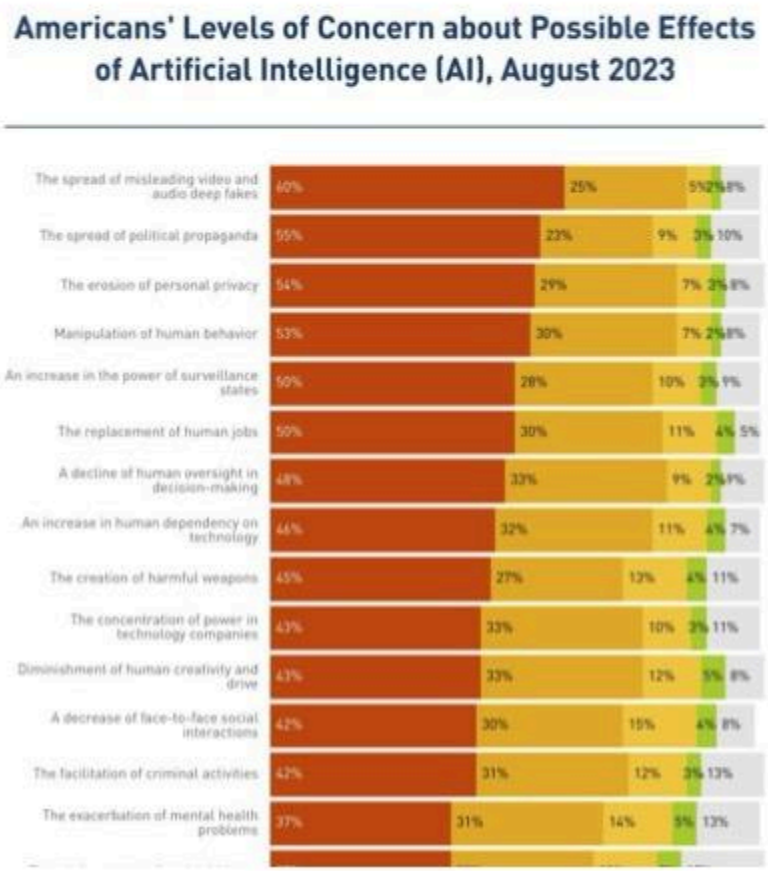
7. AI Misuse by Bad Actors
From cybercriminals using AI to craft undetectable phishing attacks, to terrorists using AI for chemical engineering or drone swarms, the misuse of AI tech is a real threat. Even individuals with relatively low resources could potentially unleash devastating harm.

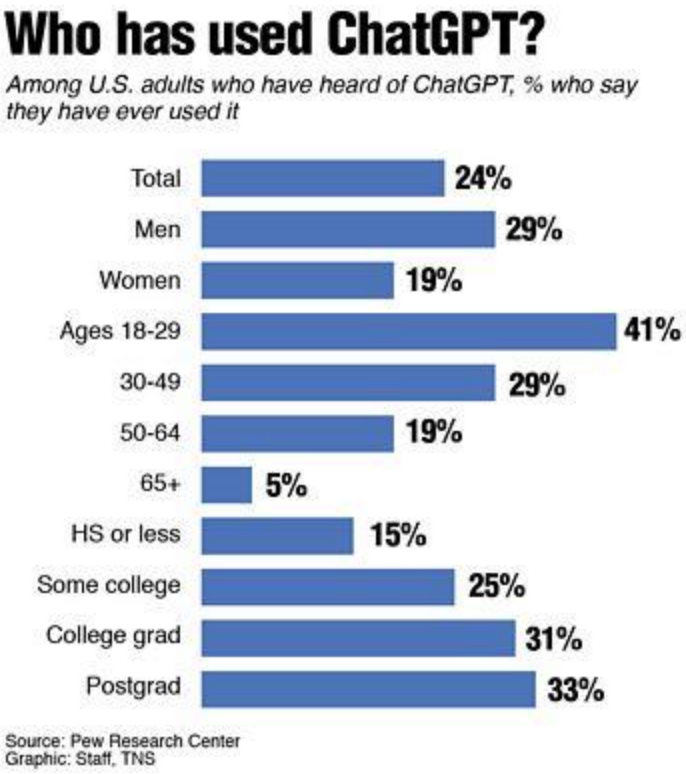
8. Human Devaluation or Redefinition
As AI takes over more cognitive tasks, we might face an identity crisis: What is the value of human intelligence in a world where machines outperform us? This could lead to widespread existential depression or societal shifts that challenge our understanding of work, purpose, and meaning.

Americans' Levels of Concern about Possible Effects of Artificial Intelligence (AI), August 2023

Date: 2024
From: Gale Opposing Viewpoints Online Collection
Publisher: Gale, part of Cengage Group
Document Type: Chart
Content Level: (Level 4)

Full Text:





Pros and Cons at a Glance

PROS	CONS
Pro 1: AI can make everyday life more enjoyable and convenient, while improving our health and standard of living. Read More.	Con 1: AI will harm the standard of living for many people by causing mass unemployment. Read More.
Pro 2: AI makes work easier for students and professionals alike. Read More.	Con 2: AI undermines critical thinking skills for students and adults alike. Read More.
Pro 3: AI helps marginalized groups by offering accessibility for people with disabilities. Read More.	Con 3: AI hurts racial minorities by repeating and exacerbating racism. Read More.

Pro 4: Artificial intelligence can improve workplace safety. Read More.	Con 4: Artificial intelligence poses dangerous privacy risks. Read More.
Pro 5: AI can function as a reliable research partner. Read More.	Con 5: AI can spread politicized, even dangerous misinformation. Read More.