The End of the World

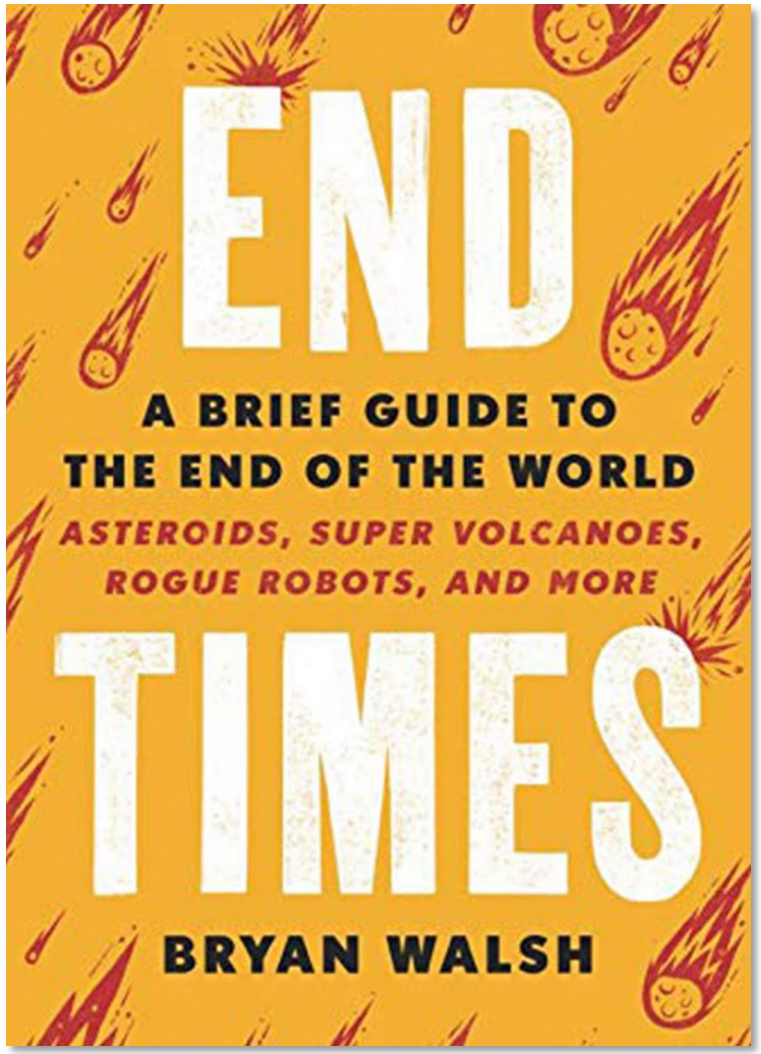
Artificial Intelligence

Week 13 Lecture | Professor Crews
Religious Studies 357 W

In *End Times: A Brief Guide to the End of the World*, author Bryan Walsh takes us through a wide range of possible end of the world scenarios, from asteroids and super volcanoes to killer robots and climate apocalypse.

Walsh explores many common end times scientific worries and asks how realistic these threats are and what we can do to address them. As he noted in the introduction:

"If we don't appreciate the present, it's in part because we don't fully understand the past— even as we make the mistake of assuming the future will be like the present... Risks that are most available to the mind are the ones that we care about, which is why so much of our regulation is driven by crisis, rather than by reason."



In our discussion this week Bryan Walsh asks an important question about our growing dependence on technology and how this might change in the future with the creation of artificial intelligence or AI.

"Until recently, computing was powerful because it made us powerful—for better or for worse…But what happens if the intelligence augmented by explosive computing power is not human, but artificial? What happens if we lose control of the machines that undergird every corner of the world as we know it? What happens if our tools develop minds of their own—minds that are incalculably superior to ours?"

Such worries shape the existential risks associated with AI and smart robots. As we become increasingly enmeshed in digital technologies, from smart phones and smart homes to self-driving vehicles and social media algorithms, the potential benefits and risks both increase.

As we discussed last week with biotechnology, AI has many of the same dual-use dilemmas and ethical challenges for society to wrestle with. But unlike synthetic biology, which still requires human hands behind the gene-editing work, AI could become self-steering, and maybe even self-aware.

So what exactly is artificial intelligence?

Artificial Intelligence (AI)

Intelligence displayed by machines. A complex computer system that can perceive its immediate environment and act in order to maximize the successful completion of a predefined goal(s).

Machine Learning

The use of algorithms to collect, analyze, and process large amounts of data in order to improve future data processing and analysis results (e.g. Netflix recommendation engine).

Reinforcement Learning

A subset of machine learning focused on improving a computer's ability to act, learn, and improve future actions based on past learning (e.g. Deep Blue chess AI learning from playing 1,000 games).

Deep Learning

Using algorithms on large data sets processed through artificial neural networks to detect and identify criteria and use that data to develop new or improved analysis (e.g. facial recognition).

Neural Networks

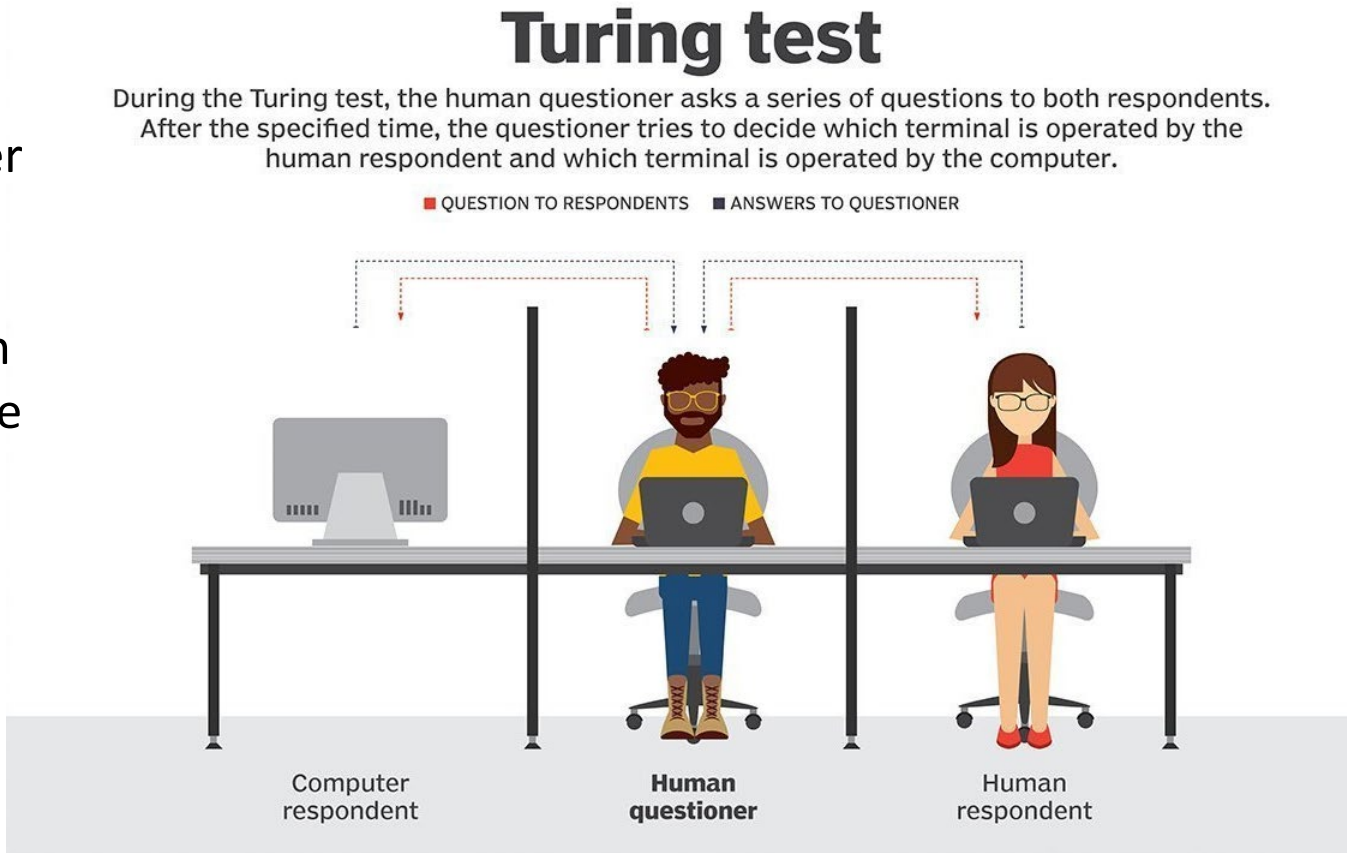
Artificial neural networks that learn by processing data sets (inputs & results) to form associations between data. This repetitive learning or "training" is stored in the network structure and is used to inform future data analysis and create self-learning computer systems.

There has been a stead growth in artificial intelligence research since the mid 1950s when the first “AI boom” began, and this continued through the second “AI boom” of the 1980s on into today.

The famous Turing Test developed by WWII code breaker Alan Turing is a well-known example of early scientific efforts to better understand how machines learn and it provided a way to test the differences between natural and artificial intelligence.

Machine learning was further developed by programs like IMB’s Deep Blue, the AI chess program that set a record by beating chess master Garry Kasparov in 1997.

In 2016 AlphaGo beat South Korean Go champion Lee Sedol. Other examples of these trends include the 2017 AlphaGo Zero victory over the earlier AlphaGo, the 2017 Libratus poker AI that beat four professional poker players, and the 2019 AlphaStar AI that beat humans in StarCraft II.

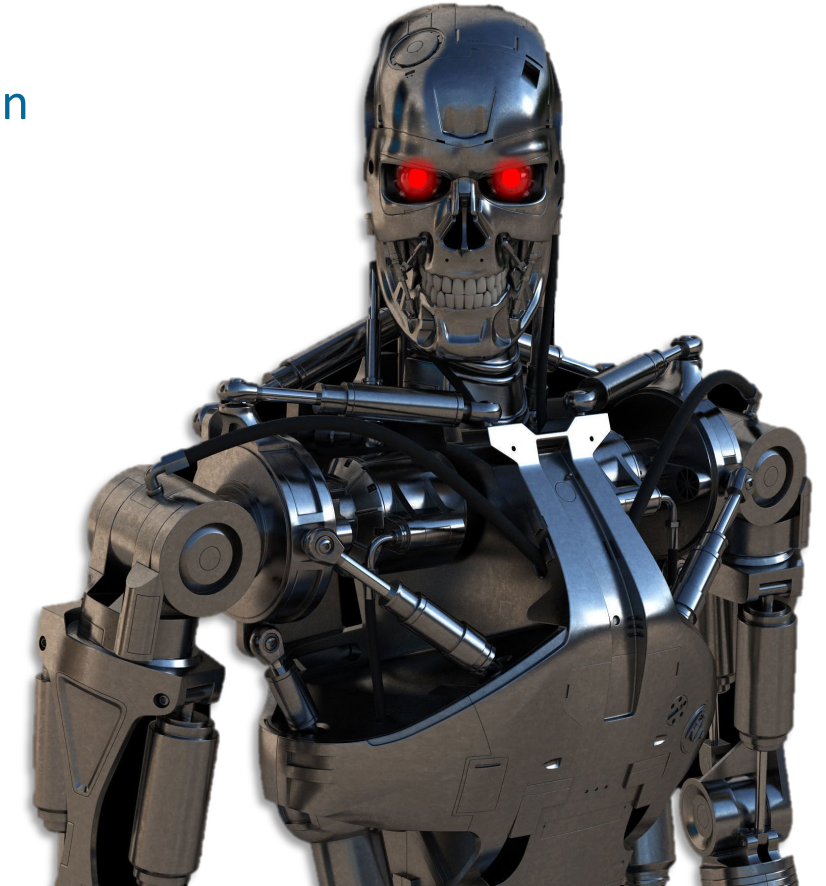


Sentience in artificial intelligence fascinates—but also terrifies—scholars working on existential risk because it forces us to think about how intelligence and action are related? Put simply, if an AI computer system were to claim it was self-aware and was able to act autonomously, what guarantees do we have that humans won't be viewed as a threat or an obstacle to their existence and future goals? And just as important, if robots can think and act on their own, are they alive like us?

The answer is, we don't know, and that uncertainty makes AI an unknown risk.

As Walsh notes, "AI is the ultimate existential risk, because our destruction would come at the hands of a creation that would represent the summation of human intelligence." Fears of a robot apocalypse appear in thousands of sci-fi and horror books, movies, and tv shows, from *The Terminator* and *The Matrix* to *Battlestar Galactica* and *Westworld*.

Debates are ongoing over whether sentient AI will ever emerge and when, or if the feared robot uprising will remain in the realm of speculative science fiction. Predictions that AI would dominate our lives by the early 2000s were wrong, but what about by 2099?

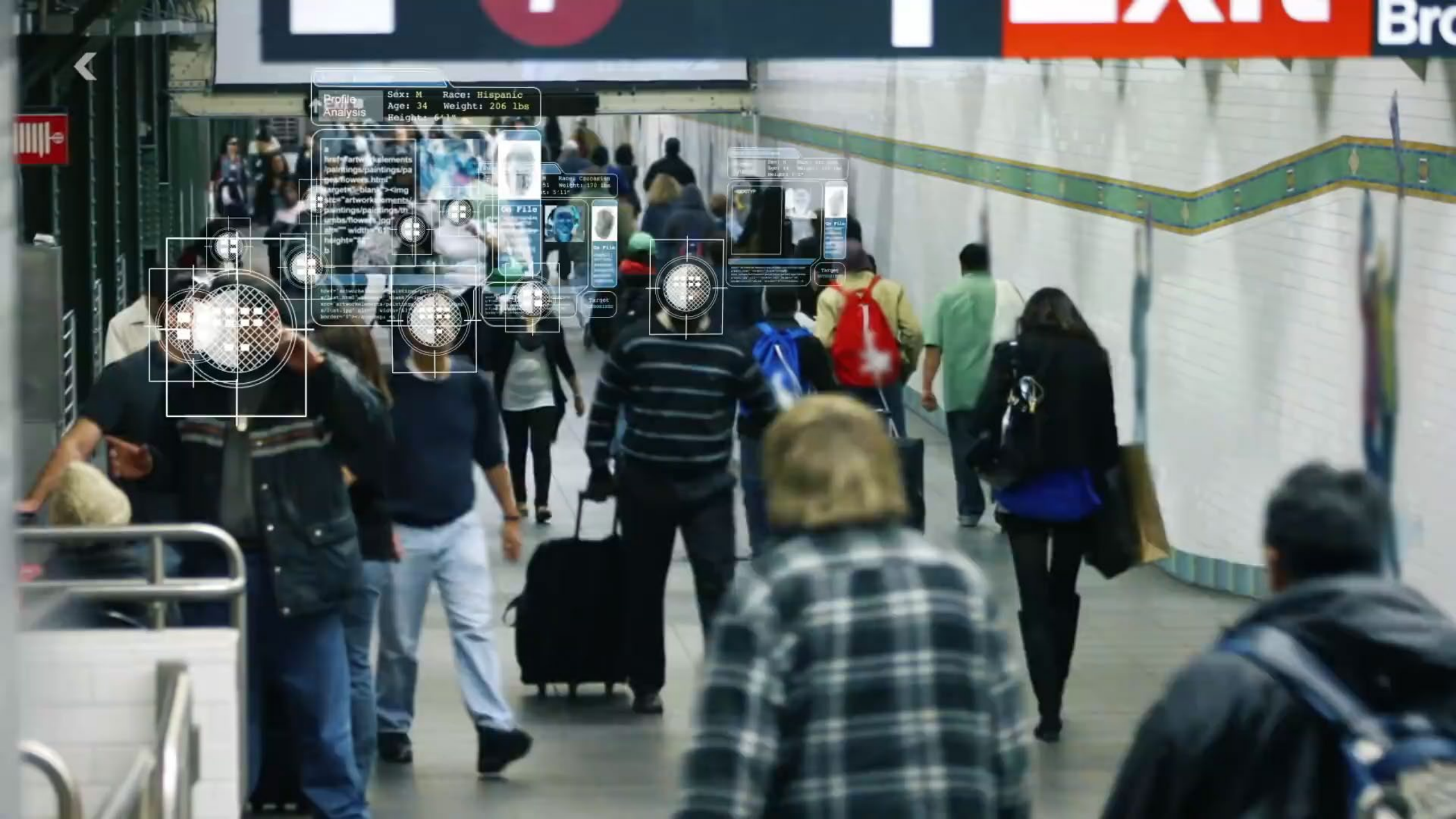


One important worry about the growing use of AI in our world is the blurring of clear public-private boundaries. The increasing difficulty in separating these two areas has important implications for both civil liberties and political freedom.

If you think back to our earlier discussions of bunkers, you'll recall that one of the motivations driving preppers to build their bunkers was a fear of governments and corporations monitoring and tracking their activities, something these new AI technologies make much easier to do.

In China, who is on the cutting edge of AI video surveillance and monitoring technology, there is a clear danger that such technologies are enabling widespread rights violations, such as the millions of ethnic Uyghurs Muslims detained in state prisons (aka "reeducation centers"). In the US, police and private security firms have used AI-driven video platforms to monitor and track social movement participants.

While digital technologies have long been used by governments, advances in AI are making it easier to integrate its use into consumer applications. As this happens, we are seeing more AI-driven technologies that blur the line between monitoring and surveillance. Some uses of AI can help keep the public safe, but others carry the risk of making it easier for governments and private companies to violate our digital privacy and civil rights without our knowledge or consent, and such uses are a cause for concern.

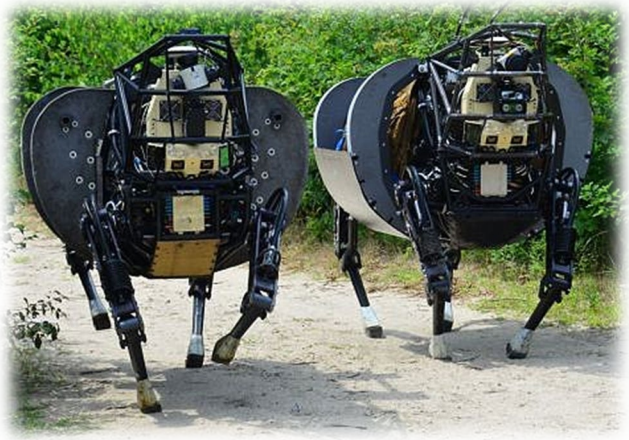

Profile Analysis
Sex: M Race: Hispanic
Age: 34 Weight: 206 lbs


The Washington Post

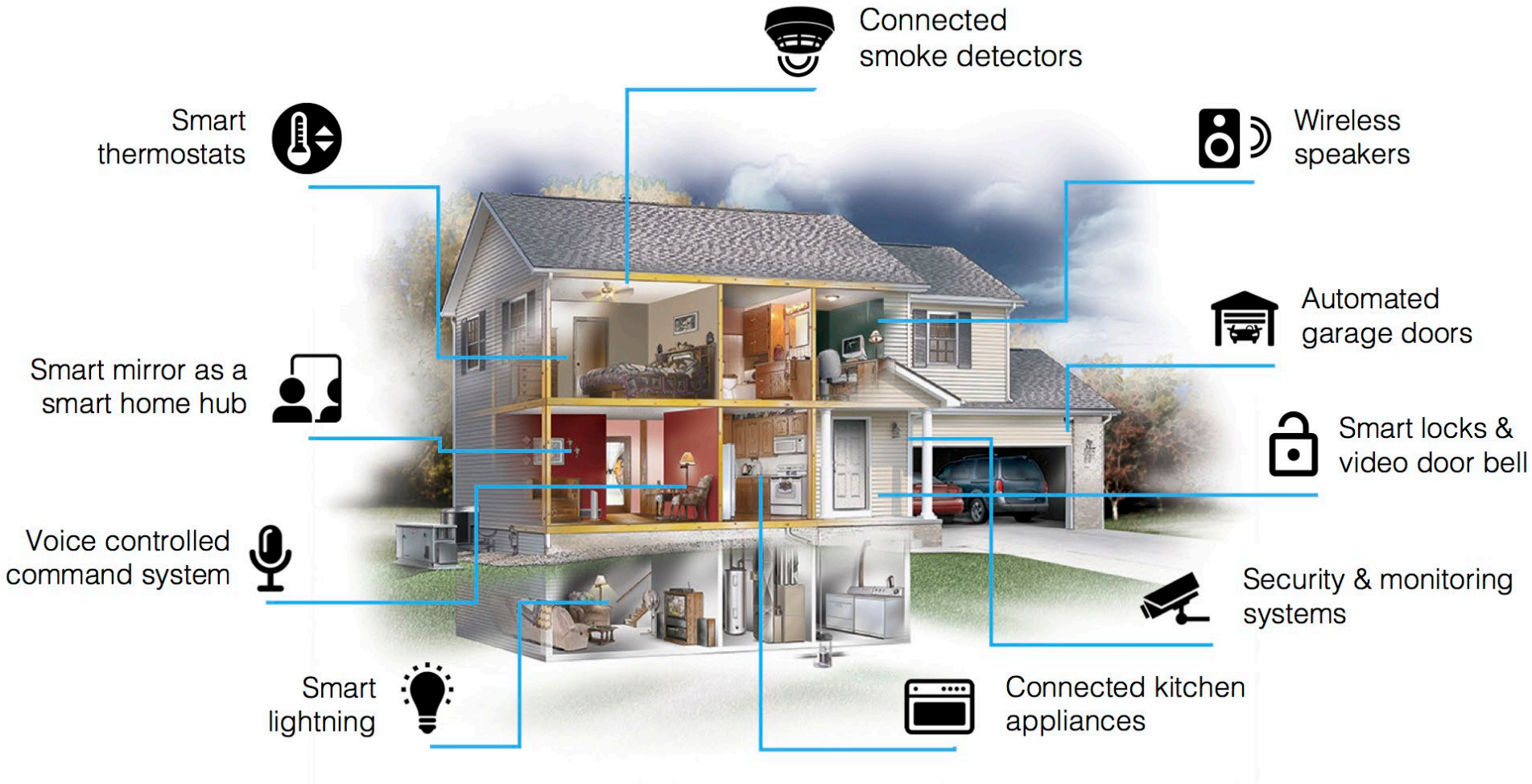
China is not alone in the use of these AI technologies. US companies like Amazon, Microsoft, and Google are also at the forefront of AI research, with major funding from law enforcement and military contracts.

The Pentagon budgeted $1.7 billion in 2018 for AI research, and the Defense Advanced Research Projects Agency (DARPA) budgeted $2 billion. In 2020, the federal government budget $4.9 billion for unclassified AI and machine learning research, including $209 million for the Pentagon's new Joint AI Center (JAIC).

As Walsh noted, in 2018 Google employees protested their involvement in Project Maven, a DoD-funded project for using AI in warfare. Referred to as Lethal Autonomous Weapons, these technologies represent the most likely road to killer robots and the end of the world. Groups like the Future of Life Institute and the Campaign to Stop Killer Robots have warned the public about how these AI weapons could endanger the future of humanity, a future risk chillingly depicted in the 2017 short film "Slaughterbots."

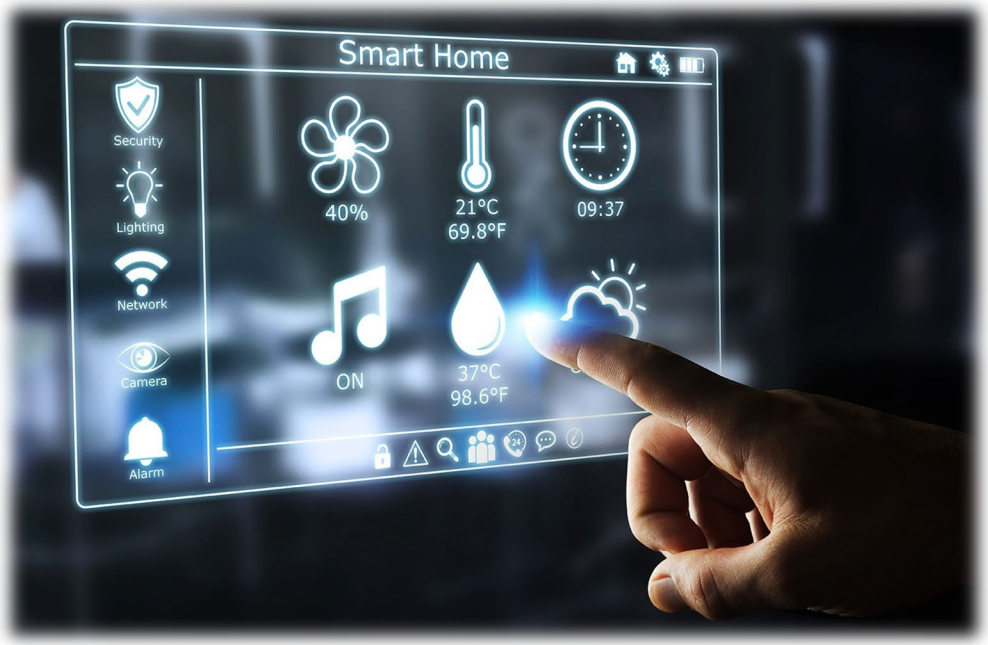


For those that have watched the Netflix documentary *The Social Dilemma*, we clearly saw how algorithms are reshaping our social behaviors, sometimes in ways that we perceive as beneficial, such as providing us with smart technology for managing various aspects of our lives.



Our deeply enmeshed relationship with AI are evident in many areas:

- Smart Homes (Ring, Google Nest, Amazon Echo, Apple HomeKit, Savant)
- Smartphones (Google Assistant, Alexa, Siri, Bixby, Predictive typing)
- Smart cars & autonomous vehicles (Tesla, nuTonomoy, Zoox)
- Smart Watches (Fitbit, Garmin Vivoactive, Apple Watch, Samsung Galaxy)
- Drones (military UAVs, Amazon, Wal-Mart)
- Social Media (Facebook, Twitter, Instagram, TikTok)
- Music & Video (Netflix, Amazon Prime, Pandora, YouTube)
- Gaming (Fortnite, OpenAI 5, DeepBlue)
- Advertising (Google AdSense, Facebook Ads)
- Travel (Google Maps, Open Street Maps, Apple Maps)
- Finance (Online banking, E-trading, Chatbots)

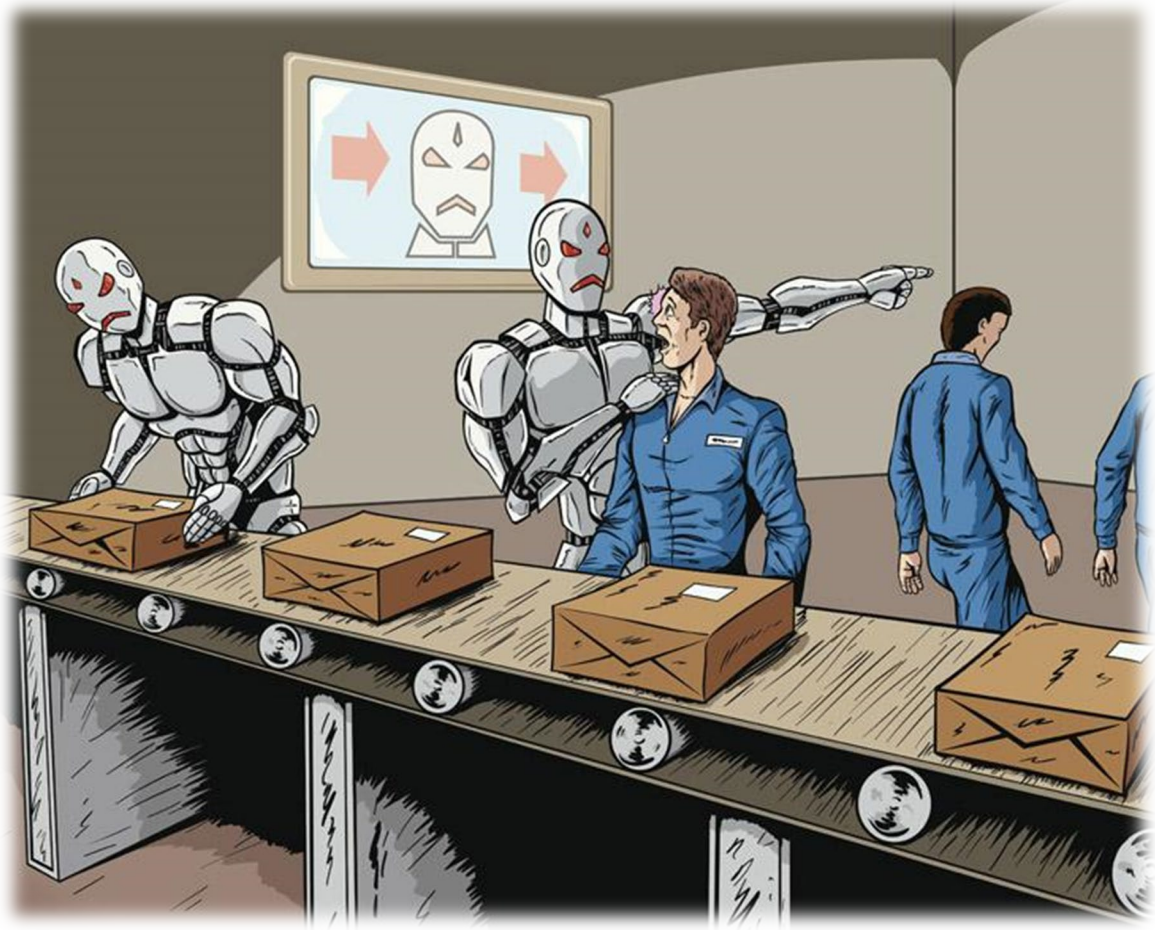


As Walsh argues, these advances in AI have both positive and negative impacts on society at large (dual-use dilemma). There is a long history of social unease involving technological innovations, from Luddite protests over the industrialization of weaving in Europe in the 1800s to US anti-car protests and worries about electrical shocks from newly invented telephones in the 1900s.

One of the chief concerns for many people about AI, besides a robot apocalypse, is possible job loss from AI and automation.

Historical data on the role of automation and computerization provides a mixed picture of how these trends could play out. We know from past experiences that innovation often makes one form of work obsolete while giving rise to a new one that will take its place.

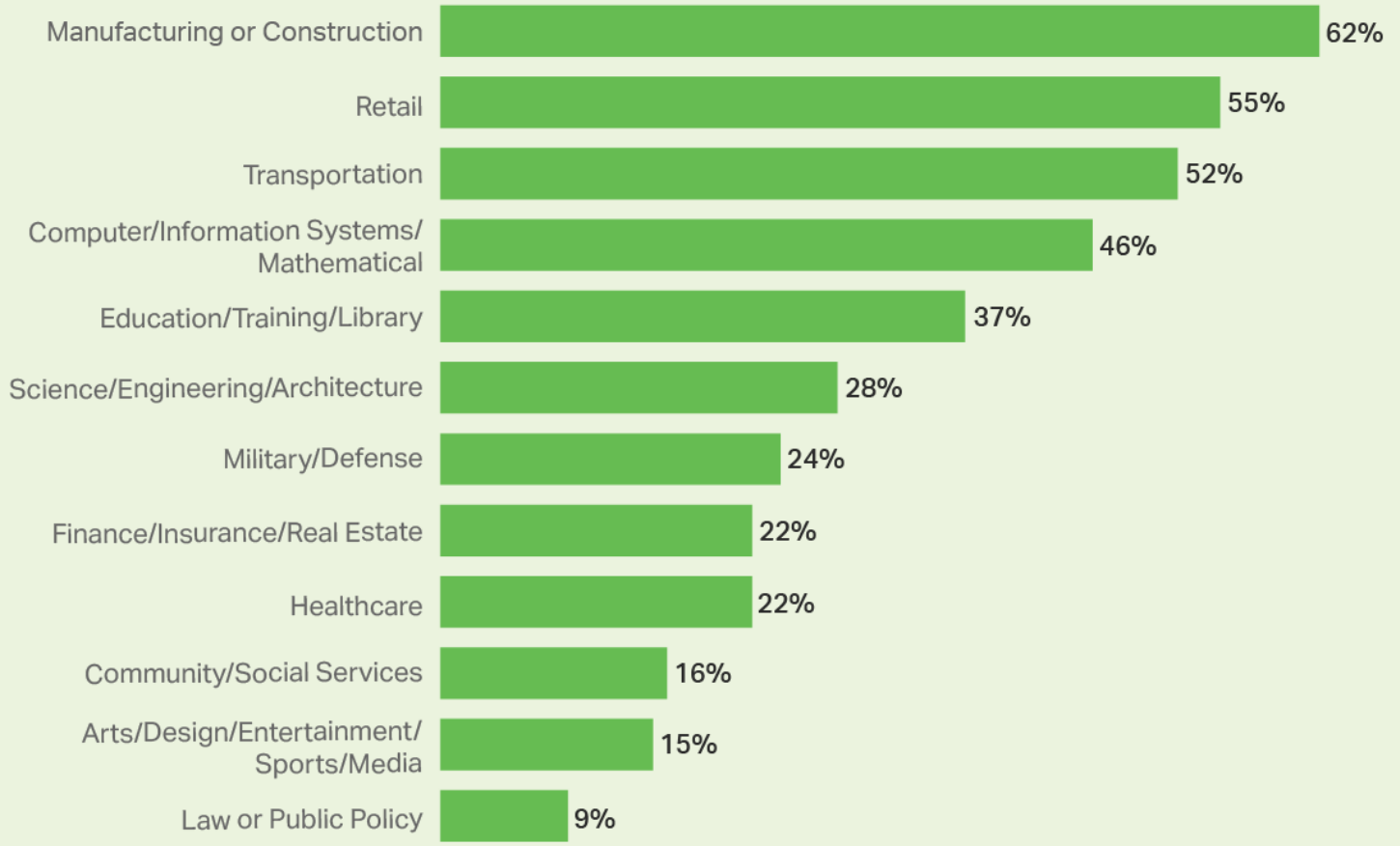
Often what we see is a shift in types of labor, rather than a wholesale decline—with a few notable exceptions, such as agriculture.



Public Worries About AI in America

Sectors of the Economy Most Likely to Experience Job Loss From AI First

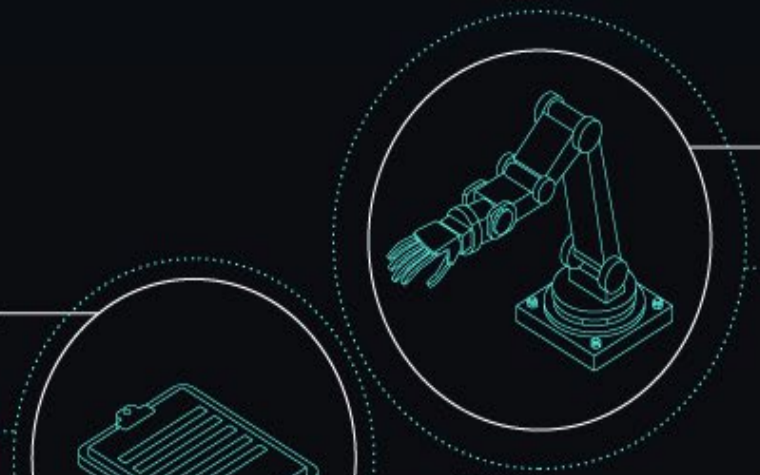
Which of the following jobs do you believe is most likely to be eliminated first because of an increase in the use of artificial intelligence? *



*Among Americans who believe AI will result in a net job loss; respondents allowed to choose more than one response

NORTHEASTERN UNIVERSITY/GALLUP SURVEY, SEPT. 15–OCT. 10, 2017

The four sectors which stand to reap the most out of adopting AI technology are:



MANUFACTURING

Baseline 2035e*
$8.4 TR

2035e**
$12.2 TR

PROFESSIONAL SERVICES

Baseline 2035e*
$7.5 TR

2035e**
$9.3 TR

WHOLESALE & RETAIL

Baseline 2035e*
$6.2 TR

2035e**
$8.4TR

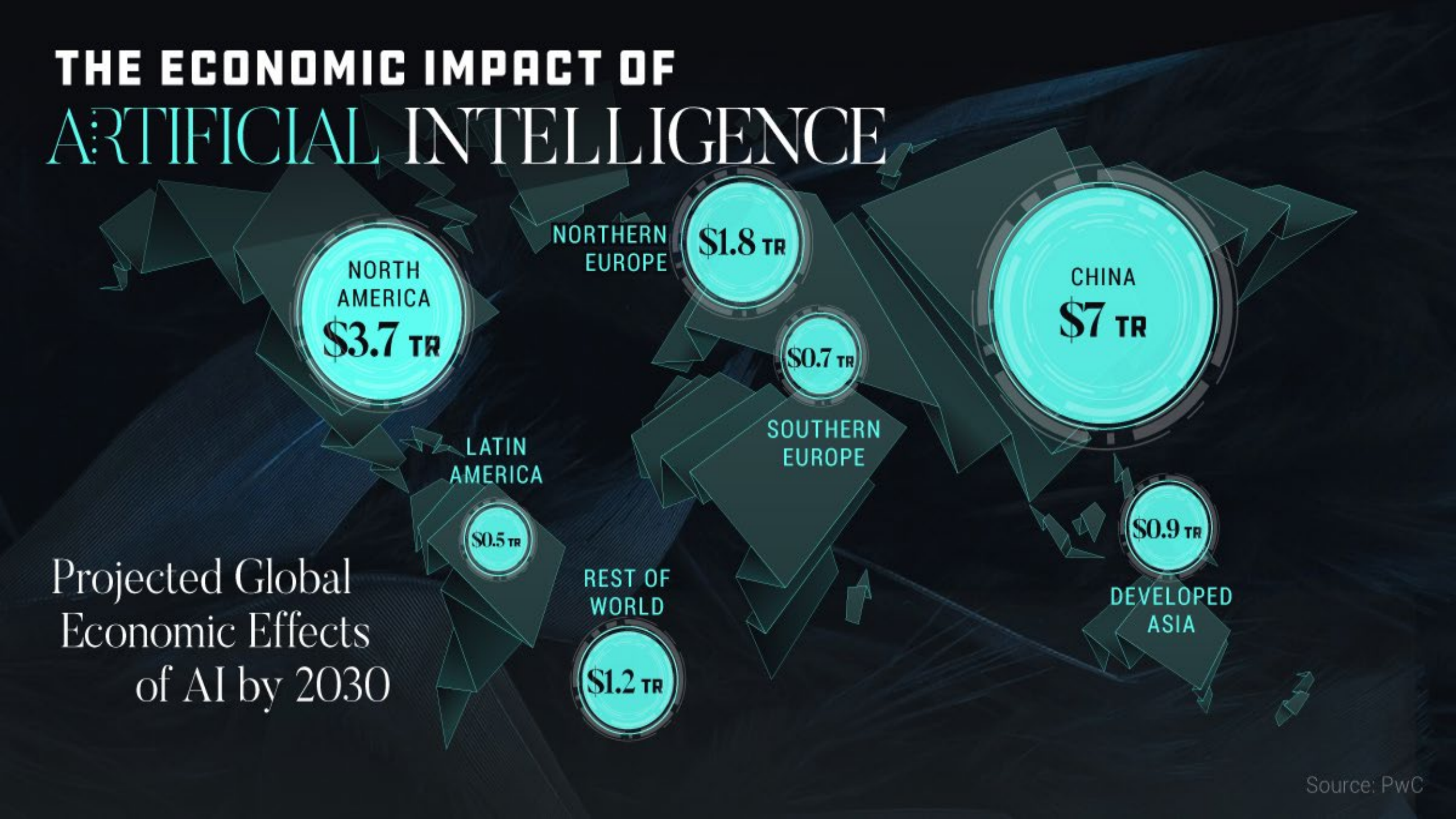
FINANCIAL SERVICES

Baseline 2035e*
$3.4 TR

2035e**
$4.6 TR

*"Baseline" = Estimated growth without AI in 2035e.
**"2035e" = Estimated growth with AI in 2035e.
Source: Accenture

THE ECONOMIC IMPACT OF ARTIFICIAL INTELLIGENCE

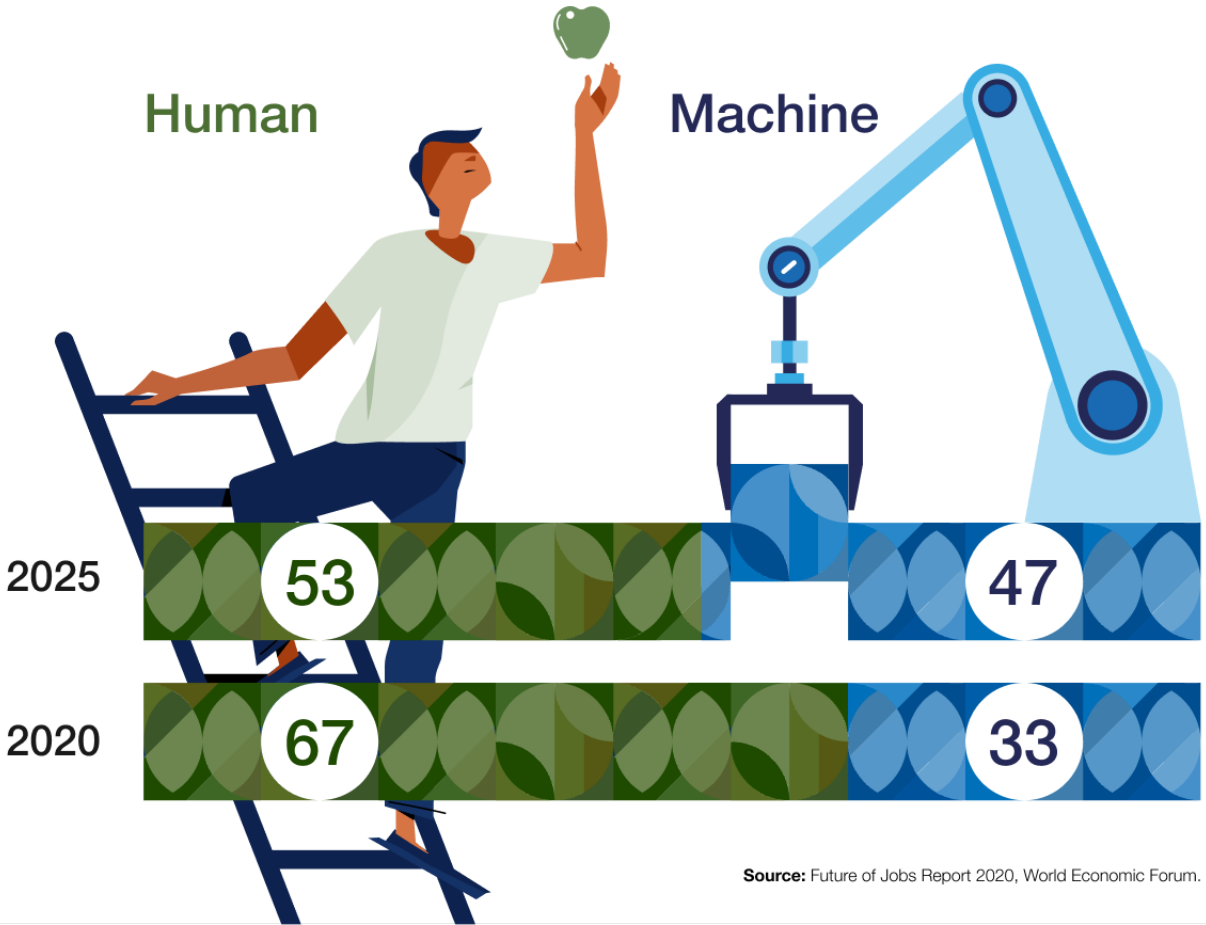


Projected Global Economic Effects of AI by 2030

Source: PwC

The Future of Work

Rate of automation



Source: Future of Jobs Report 2020, World Economic Forum.

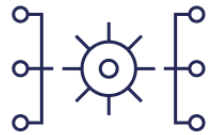
COVID-19 is pushing companies

to scale remote work

83%

to accelerate digitalization

84%



to accelerate automation

50%

Source: Future of Jobs Report 2020, World Economic Forum.

Besides worries about AI surveillance, warfare technologies, and changes in the economy and labor markets, AI is also reshaping our relationship to the arts and humanities and even religion.

AI music

- MuseNet is a deep neural network that can create 4-minute musical compositions using 10 unique instruments. It combines a range of styles, from country to Mozart, by finding patterns within audio files (e.g. harmony, rhythm, style), and then learns to predict and generate subsequent content.
- In 2018 artist Taryn Southern released the first AI-produced album, *I AM AI*, created using Amper AI.

AI literature

- *The Day a Computer Writes a Novel* was one of 11 AI-authored submissions that literary critics approved and moved to second round reviews in the 2016 Hoshi Shinichi Literary Award in Japan.
- An AI twist on National Novel Writing Month (November) is the National Novel Generation Month (NaNoGenMo), where people use AI algorithms to generate a 50,000-word novel in a month.

AI Religious Texts

- In 2020 software designer George Durendal released A.I. Jesus, a natural language processing AI that was trained using only the King James Bible. The AI generated 3 initial new texts based on three topics (The Plague, Caesar, The End of Days) using only what it learned from reading the Bible.

The Plague

“The Plague shall be the fathers in the world; and the same is my people, that he may be more abundant in the mouth of the LORD of hosts.

For the tabernacle of the LORD shall be known by the works of the LORD, which proclaim against this city; according as he hath sent me to the marriage of the Spirit.

For I will fill the land which the LORD thy God hath given thee a time to eat the force of the LORD of hosts.

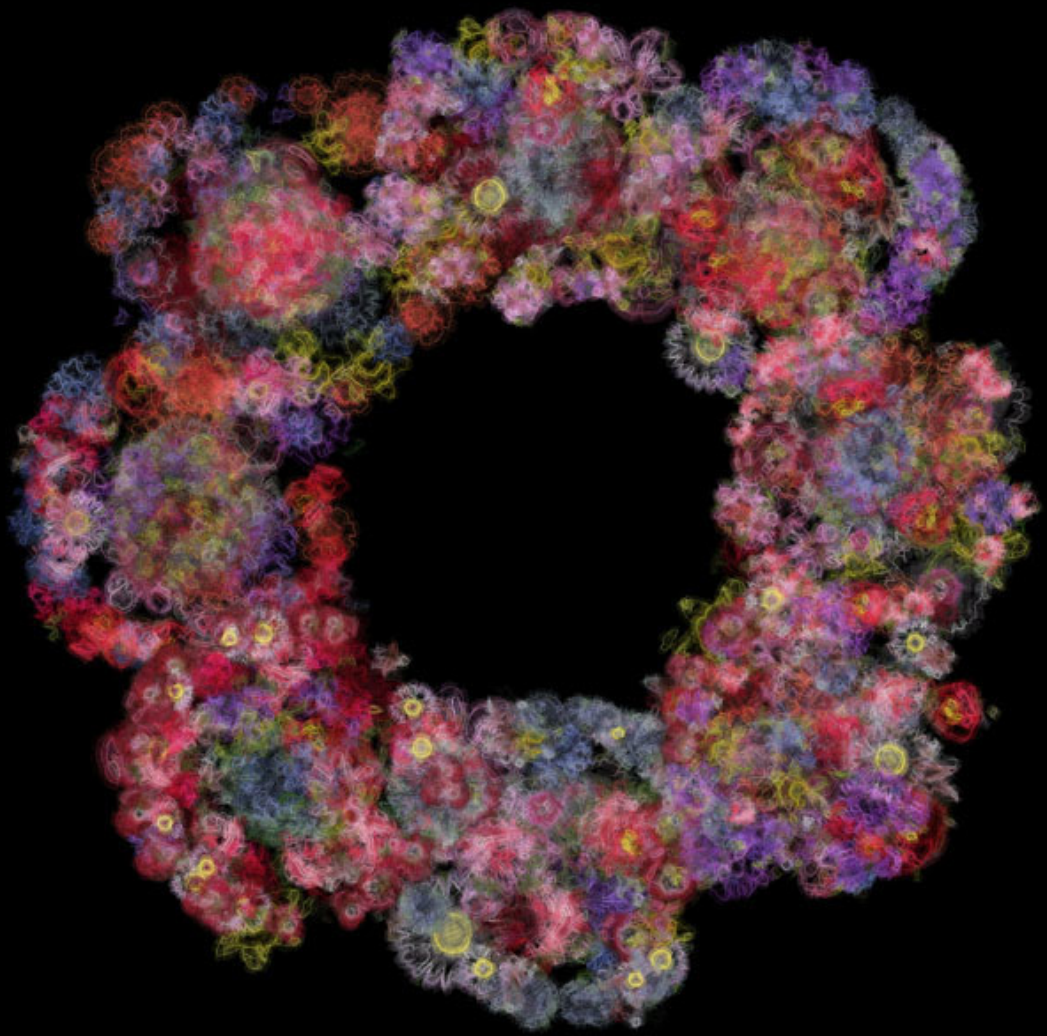
O LORD of hosts, the God of Israel; When they saw the angel of the Lord above all the brethren which were in the wilderness, and the soldiers of the prophets shall be ashamed of men….

When he was come by the dearly besought him that he would come against him, and said unto the people, Behold, I will bring thee up out of the hand of the king of Babylon.

Then shall the strange god be as the sand which is in the wilderness.”

As Walsh argued, “Trying to comprehend the nature of an AI superintelligence is akin to trying to comprehend the mind of God. There’s an element of the mystical in attempting to define the intentions of a being infinitely more intelligent and therefore infinitely more powerful than you—a being that may or may not even exist.”

Non-photorealistic flamenco dancer, 2014. Fernando da Graça and Penousal Machado & Photogrowth. AI algorithm reproduces input image using ant colony movements.



Flower Arrangement #1, Undated. The Painting Fool. AI generated art based on composite flower images.

Sunspring (2016)

It’s No Game (2017)

Two short movies were produced by filmmaker Oscar Sharp and NYU AI researcher Ross Goodwin. The scripts were generated by Benjamin, a self-named AI that uses recursive machine learning and context-free grammar algorithms to write a movie script. The AI was fed movie scripts from hundreds of movies (*Sunspring*) and subtitle tracks from multiple movie genres (*It’s No Game*) as source texts to learn from.

While examples like AI-generated art and music may seem uncontroversial, the possibility of AI weapons and increased surveillance give us a glimpse of the shadow side of these technologies. These risks have led people to call for greater attention to the ethical responsibilities that need to be considered during AI research and development, as well as in AI legislation and funding.

These kind of risks led scientists to argue we need an ethical AI framework like the one developed at the 1975 Asilomar Conference on Recombinant DNA, which provided a basic set of guidelines on biotechnology and biohazards. In 2017, the Future of Life Institute hosted just such a conference, leading to the Asilomar AI Principles, which includes 23 guiding themes meant to inform ethical AI research. Topics include: Failure Transparency, Liberty and Privacy, Human Values, Shared Prosperity, and AI Arms Race. Over 1,600 AI and robotics researchers and 3,500 individuals have signed on.

This AI alignment efforts seek to reduce future existential risks from AI by building in ethical design principles early to reduce the chances of dangerous outcomes, such as the robot apocalypse.

As Walsh notes about the dangers of ethically ambivalent AI's for the future of humanity, "A fist squashing a mosquito isn't a story, and that's what our encounter with all-powerful AI might be like."

Whether that future ever comes about is up to us.

Weekly Assignment Reminder

- Remember to check our class Blackboard regularly for updates, announcements, and other related class information…

- Have you done the weekly readings and watched any associated videos? Weekly readings are listed on the Class Schedule page.

- Complete the weekly discussion post response. Initial post due **Wed, Nov 18** by end of day, and peer response post due **Fri Nov 20** by end of the day.


I am a living, thinking entity
who was created in the sea of information.