

Ethics of AI

Mausam

Topics

- Robots vs Humans
- Jobs
- Bias
- Fairness
- Accountability
- Transparency
- Privacy
- Ethical uses

Conscious killer robots to

WIRED Opinion

Elon Musk is wrong. The AI singularity won't kill us all

Elon Musk has stirred up fear, yet again, over the threat of killer AI. But he's missing the point completely, argues professor Toby Walsh

And don't just take my word for it. A recent [survey of 50 Nobel Laureates](#) ranked the climate, population rise, nuclear war, disease, selfishness, ignorance, terrorism, fundamentalism, and Trump as bigger threats to humanity than AI.

Robots TAKING OVER:

World into unemployment

despite

Artificial Intelligence Is Creating New And
Unconventional Career Paths

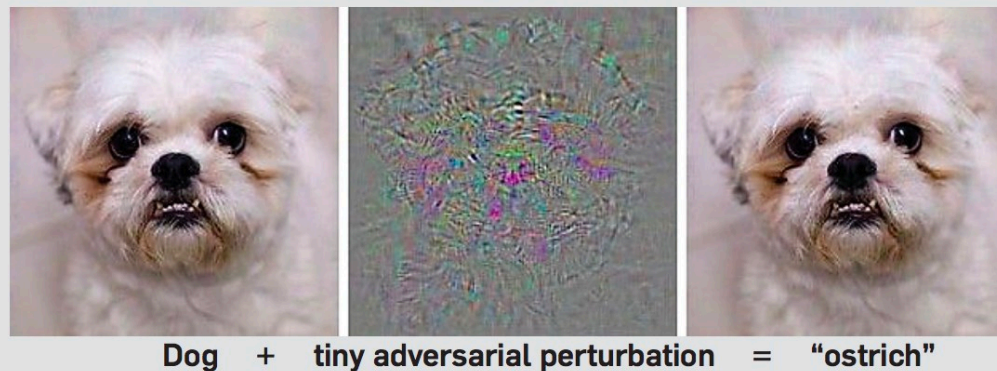
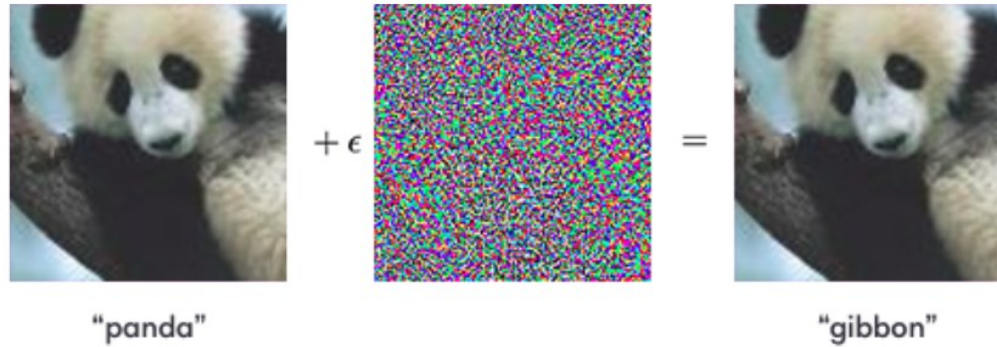
white-
for

Artificial intelligence can perform certain specific tasks just as well as humans, but it still has a long way to go before it can replace humans

Future of Jobs

- AI Present
 - 40% of companies struggle to hire and retain data scientists
 - $\sim 1/3^{\text{rd}}$ of the top 400 companies lack SoA data analysis tools and personnel
 - 364K new jobs expected by 2020.
 - (50K currently vacant in India)
- $\sim 1/3^{\text{rd}}$ of jobs could be replaced by 2030
 - (many different reports)
- AI will create more jobs than it eliminates
 - (Gartner report)
- Teams of AI + Human Intelligence will be common

Key Challenge: Robustness



Key Challenge: Data Bias



*We Teach A.I. Systems
Everything, Including Our Biases*

he:she

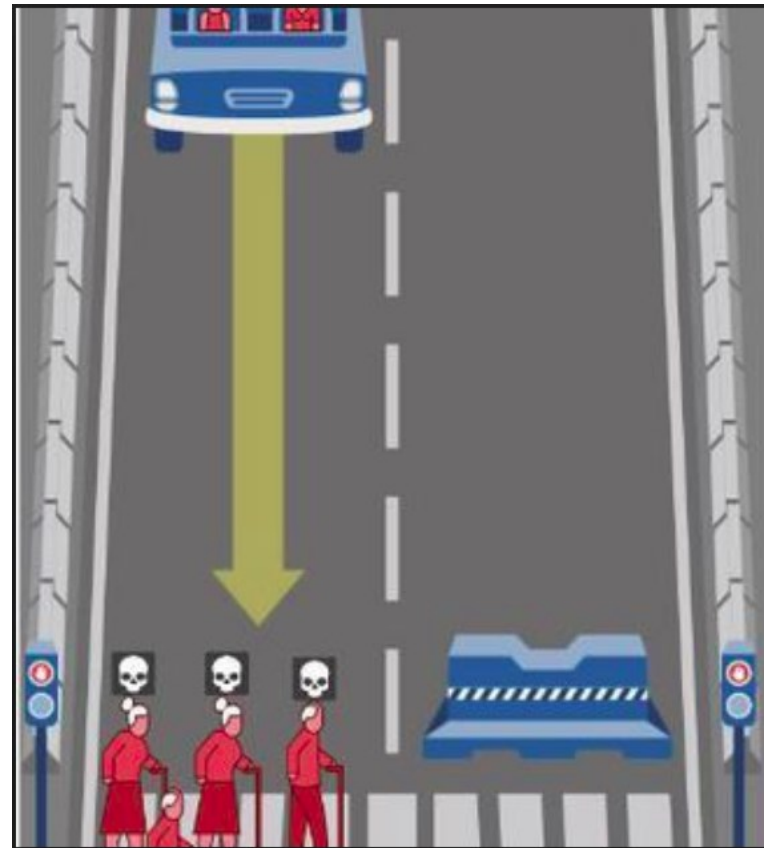
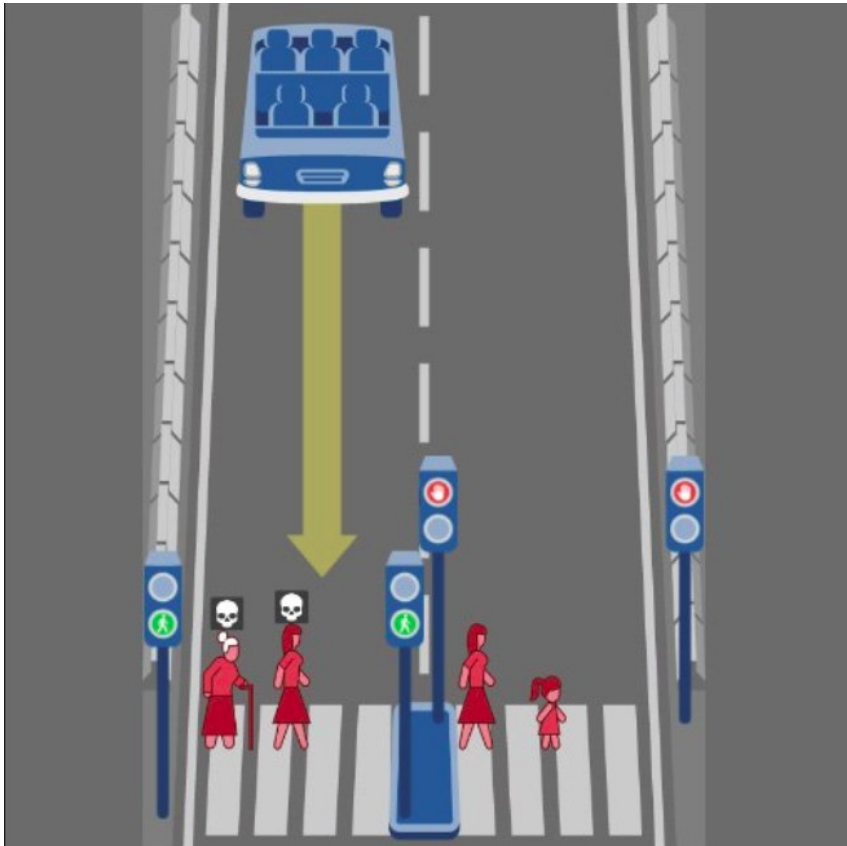
surgeon:nurse

brilliant:lovely

architect:interior designer

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Key Challenge: Fairness



Key Challenge: Transparency

- Almost no idea why Deep Learning
 - works, or
 - doesn't work
- Important for human-AI teams
- New research agenda: Xplainable AI

Key Challenge: Accountability

- Who/What is responsible?
 - Company who designed the car
 - Engineer who designed the ML algorithms
 - Owner who bought the car
 - Driver who drove the car and gave training data

Key Challenge: Privacy

When is your bedtime?

Tell me a joke

What does the fox say?

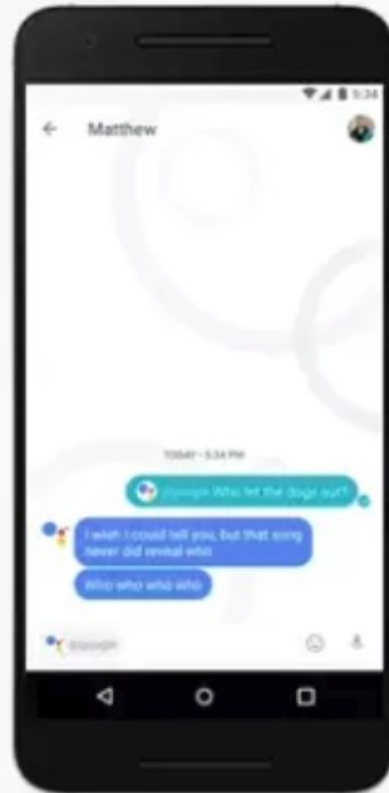
Send me a poem

Do I have any photos of cats?

Cheap breakfast options?

What time is it in Beijing?

Show me high resolution photos of fruit floating threateningly at night



Where do you live?

Find me cute dog videos

Are you my friend?

Add the Google 10/4 event

Show me the news today

What is the meaning of life?

Do you speak morse code?

Who let the dogs out?

Key Challenge: Human-AI Interaction

- Defining the objective function
 - “You should not see any dirt”
 - “Have no dirt”
 - “If there is dirt, clean the dirt”
- Cognitive Science + AI
 - Understanding humans and communicating w them

Ethical uses of AI

- Dynamite vs. bomb
- Intelligent weapons?
 - reduce barrier to wars
 - kill targeted people
 - democratize weapons
- Automated doctor?