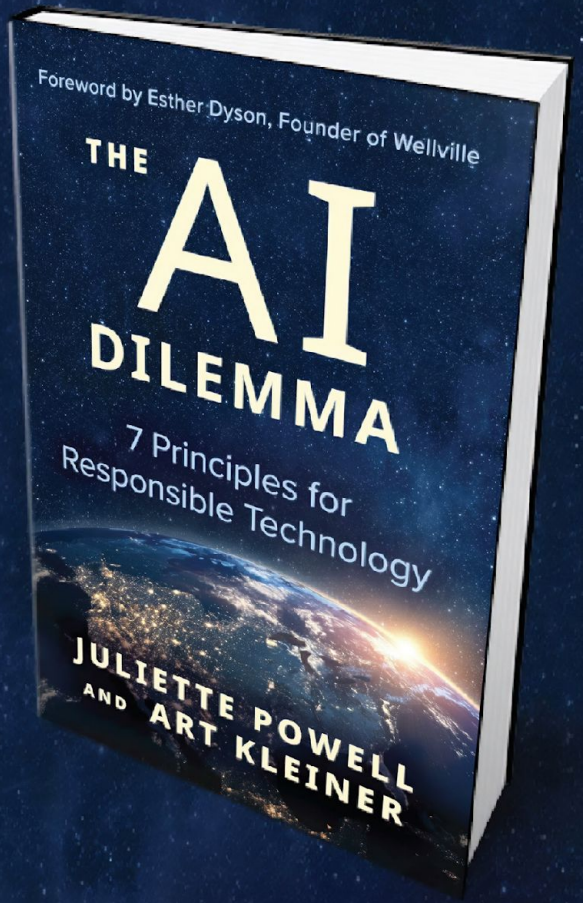
The A.I. Dilemma



Juliette Powell
Fall, 2023

Robotaxis in San Francisco

How long will it be before they are in your local community?

- *0-1 years*
- *1-5 years*
- *5-10 years*
- *More than 10 years*
- *Never*

What is AI?

We use the term “artificial intelligence” because that is the term generally used by the public.

Engineers refer to this technology as machine learning.

We also use the term Triple-A systems, because they are:

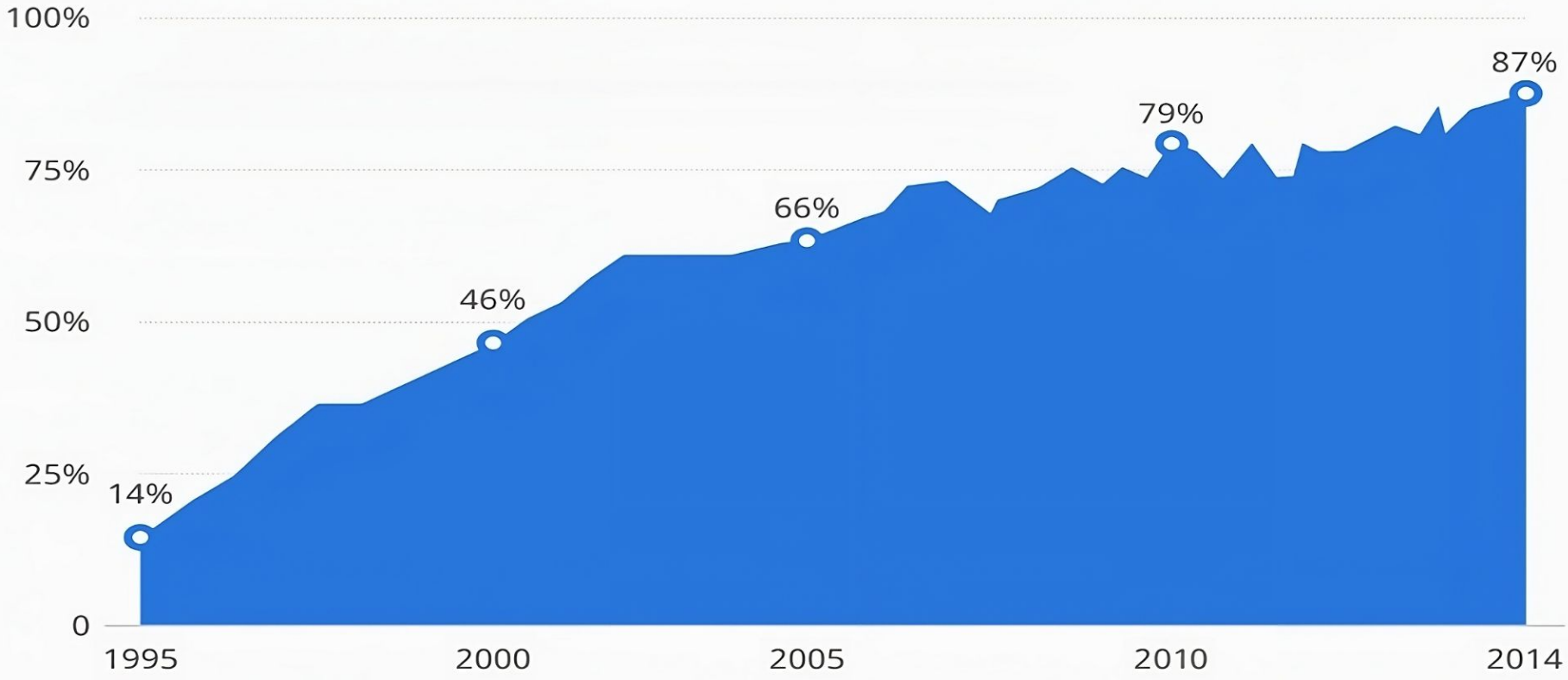
Autonomous: They can perform tasks in complex environments without constant guidance by a person.

Adaptive: They can improve performance by learning from experience.

Algorithmic: Their activities are managed through software code.

The Rapid Rise Of The Internet

% of American adults who use the internet



@StatistaCharts

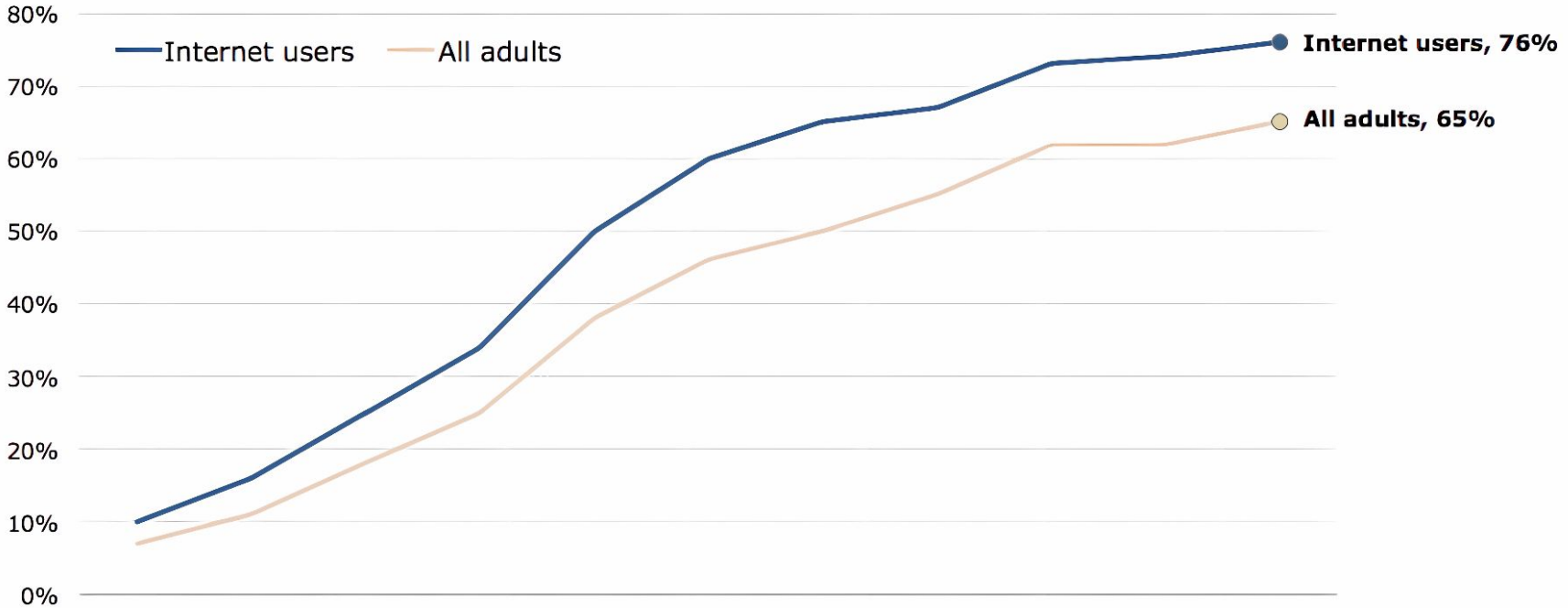
Source: Pew Research Center

statista

Social Media Adoption Trends

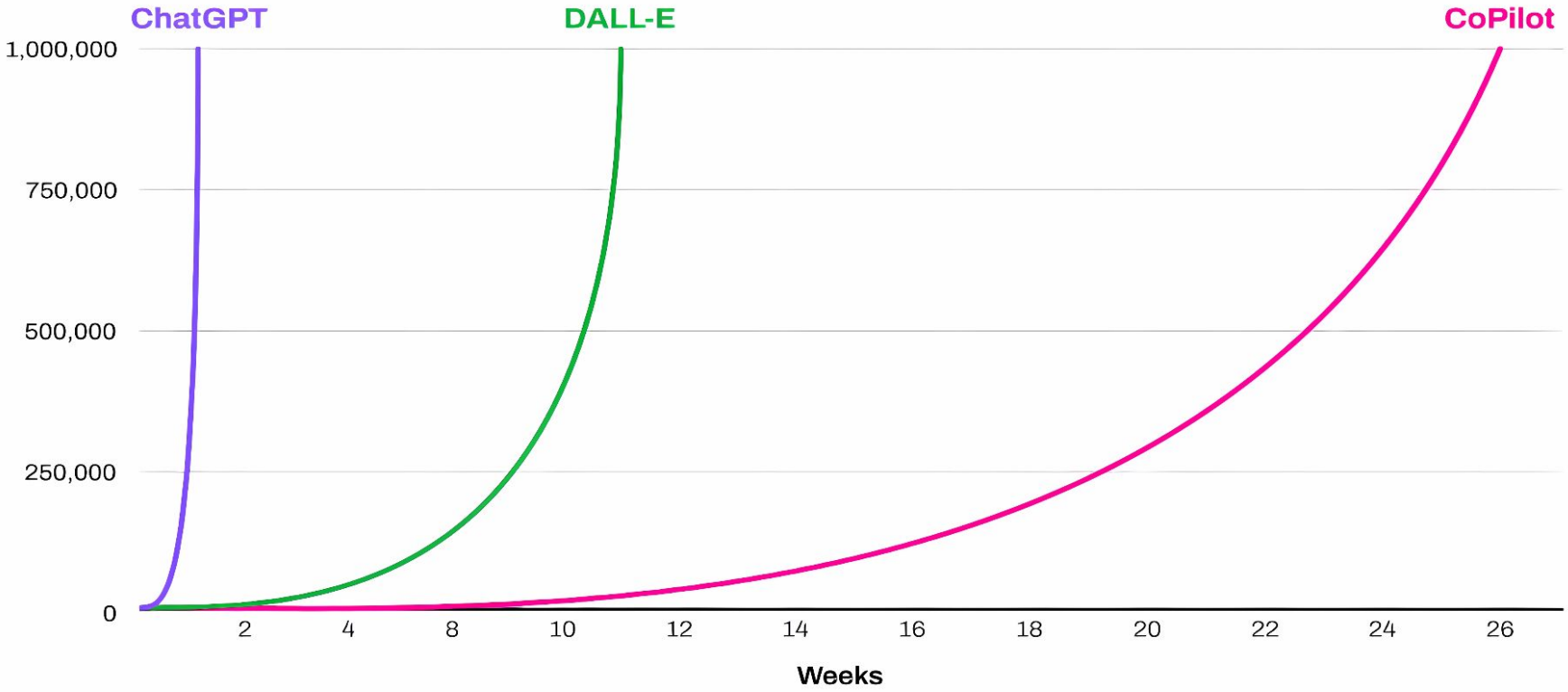
percentage of all American adults - and internet-using adults - who use at least one social networking site

2005-2015

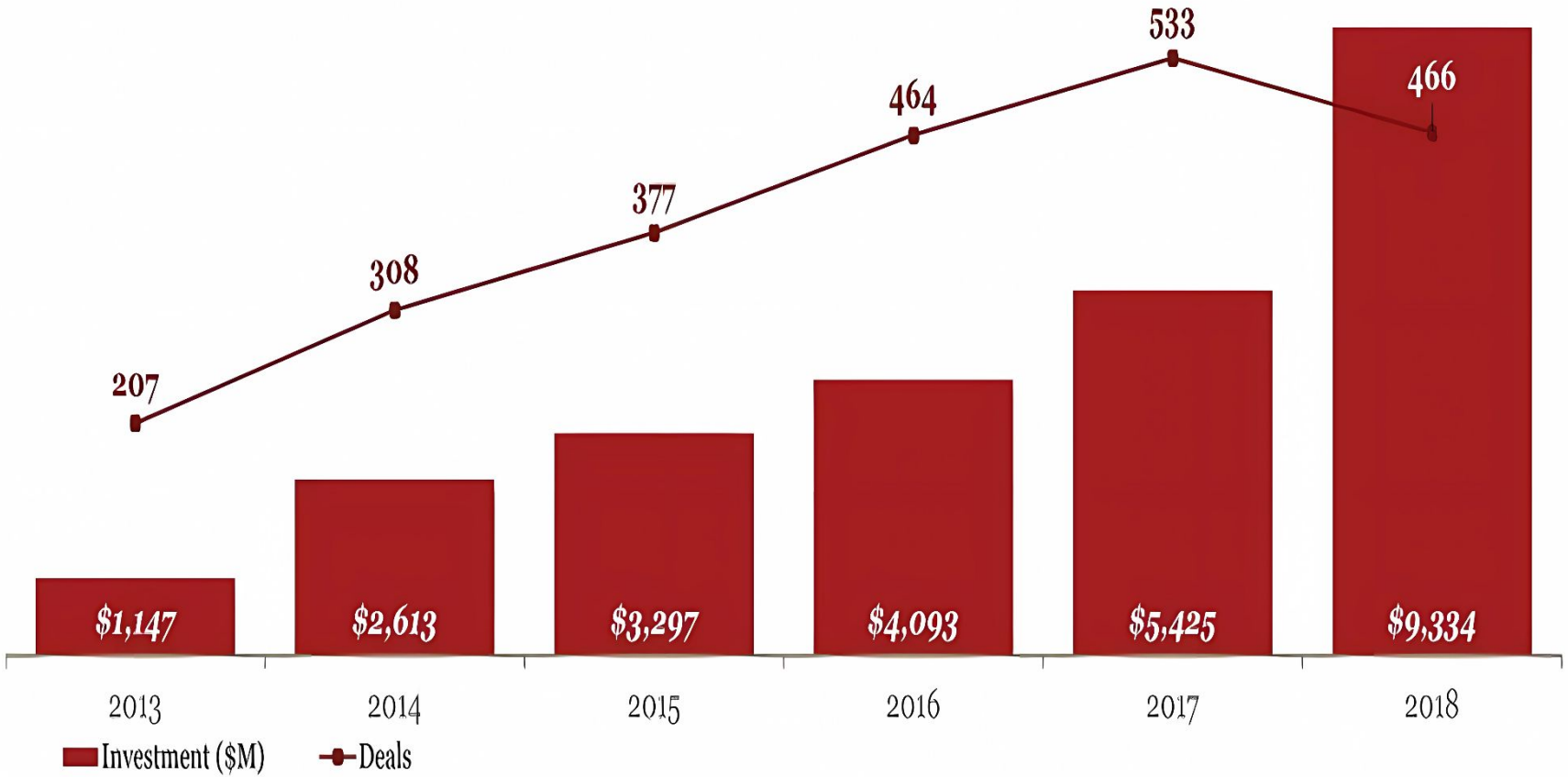


	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Internet users	10%	16%	N/A	34%	50%	60%	65%	67%	73%	74%	76%
All adults	7%	11%	N/A	25%	38%	46%	50%	55%	62%	62%	65%

Weeks to Reach 1 Million Users



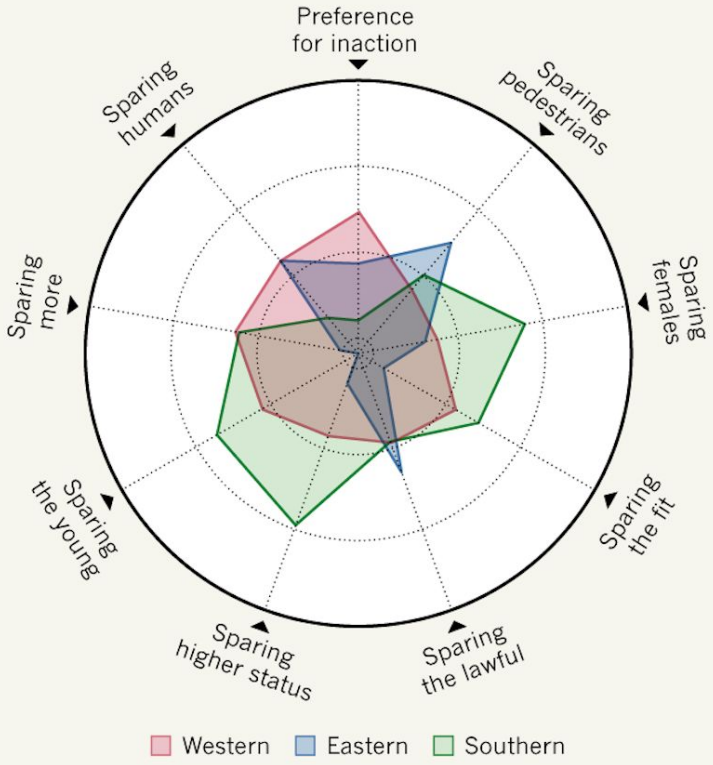
Self-Driving
48
mph

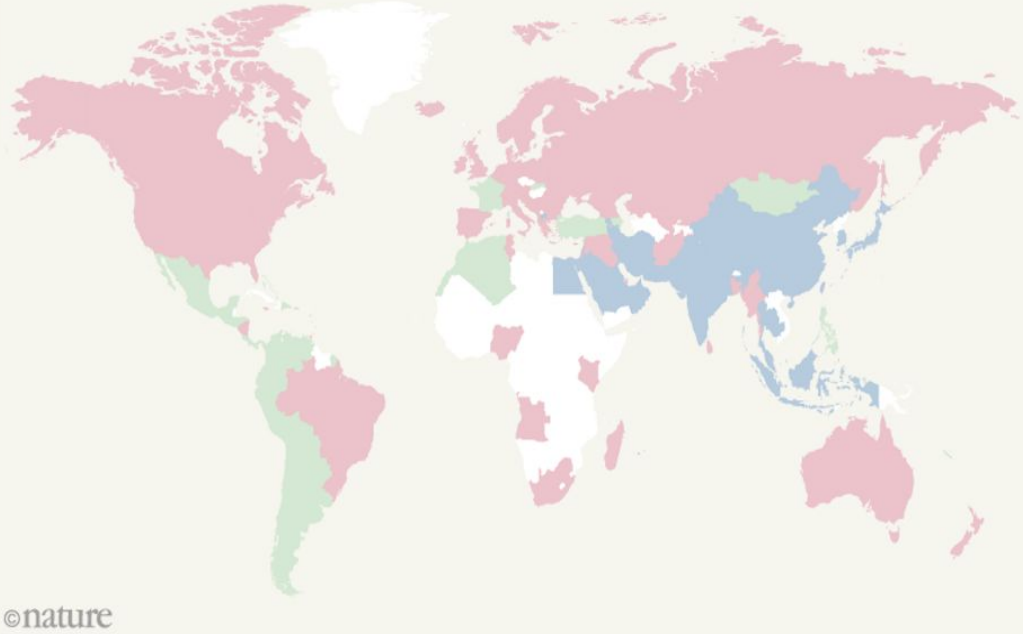


PwC | CB Insights MoneyTree™ Report Q4 2018

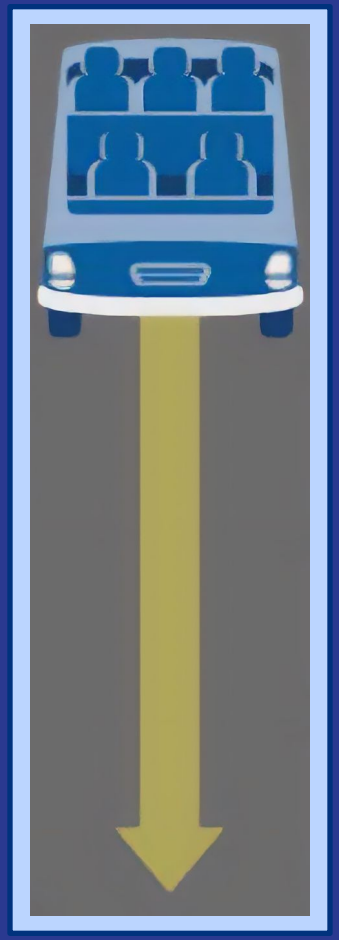
MORAL COMPASS

A survey of 2.3 million people worldwide reveals variations in the moral principles that guide drivers' decisions. Respondents were presented with 13 scenarios, in which a collision that killed some combination of passengers and pedestrians was unavoidable, and asked to decide who they would spare. Scientists used these data to group countries and territories into three groups based on their moral attitudes.

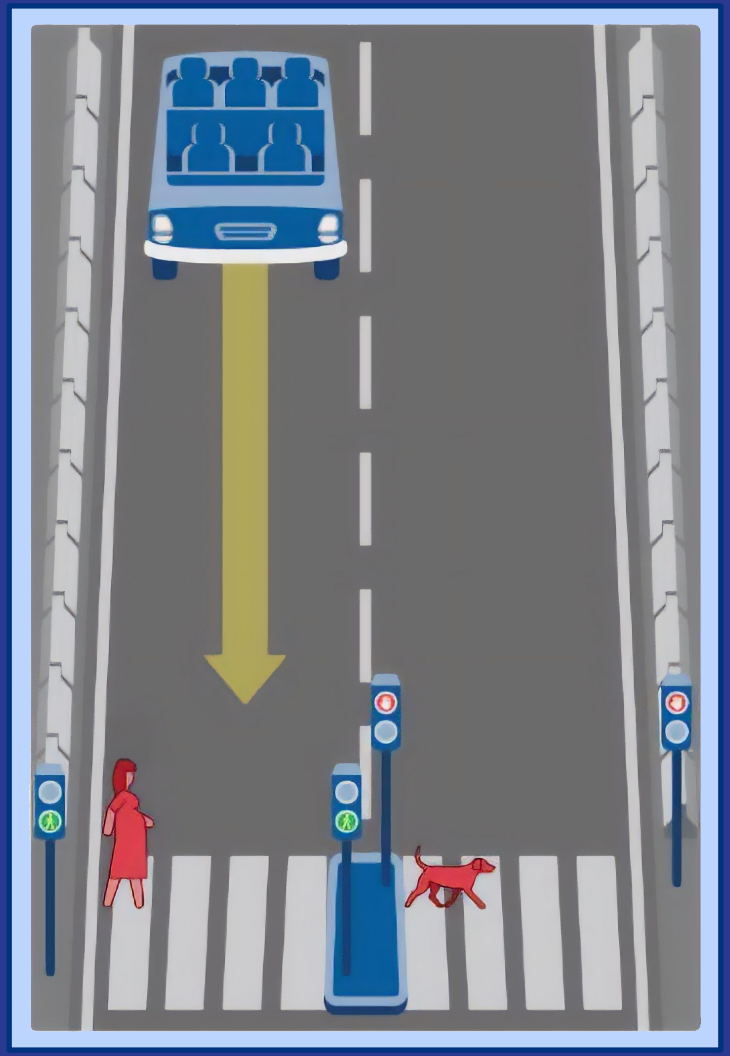


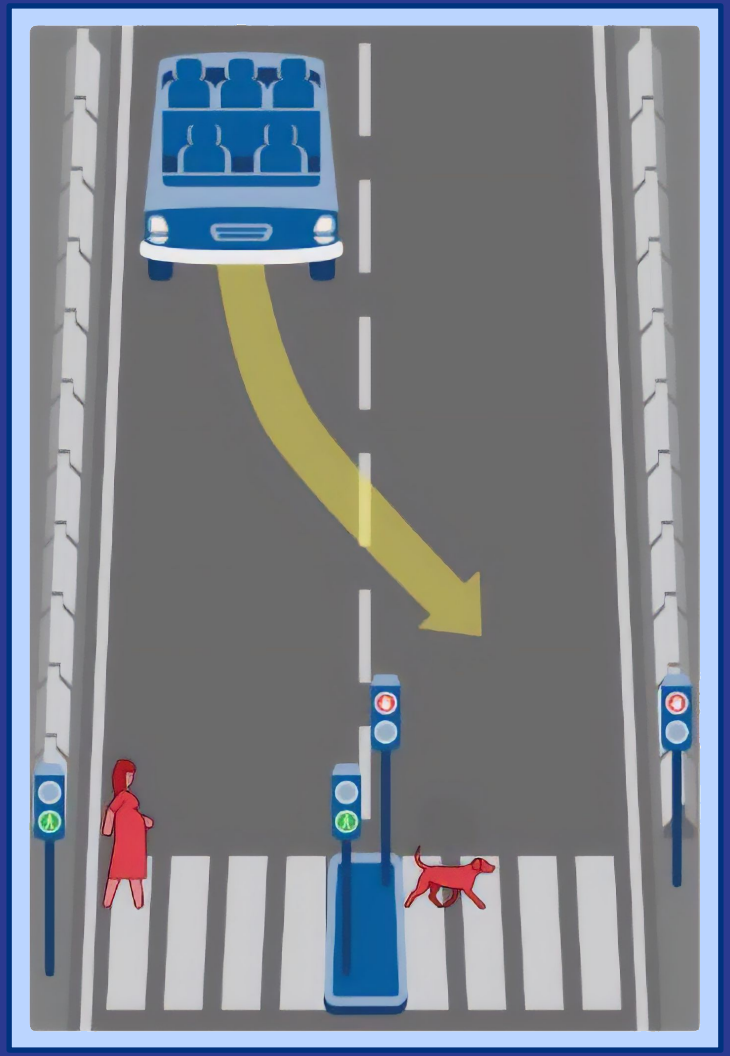


Suppose you made a self-driving car....

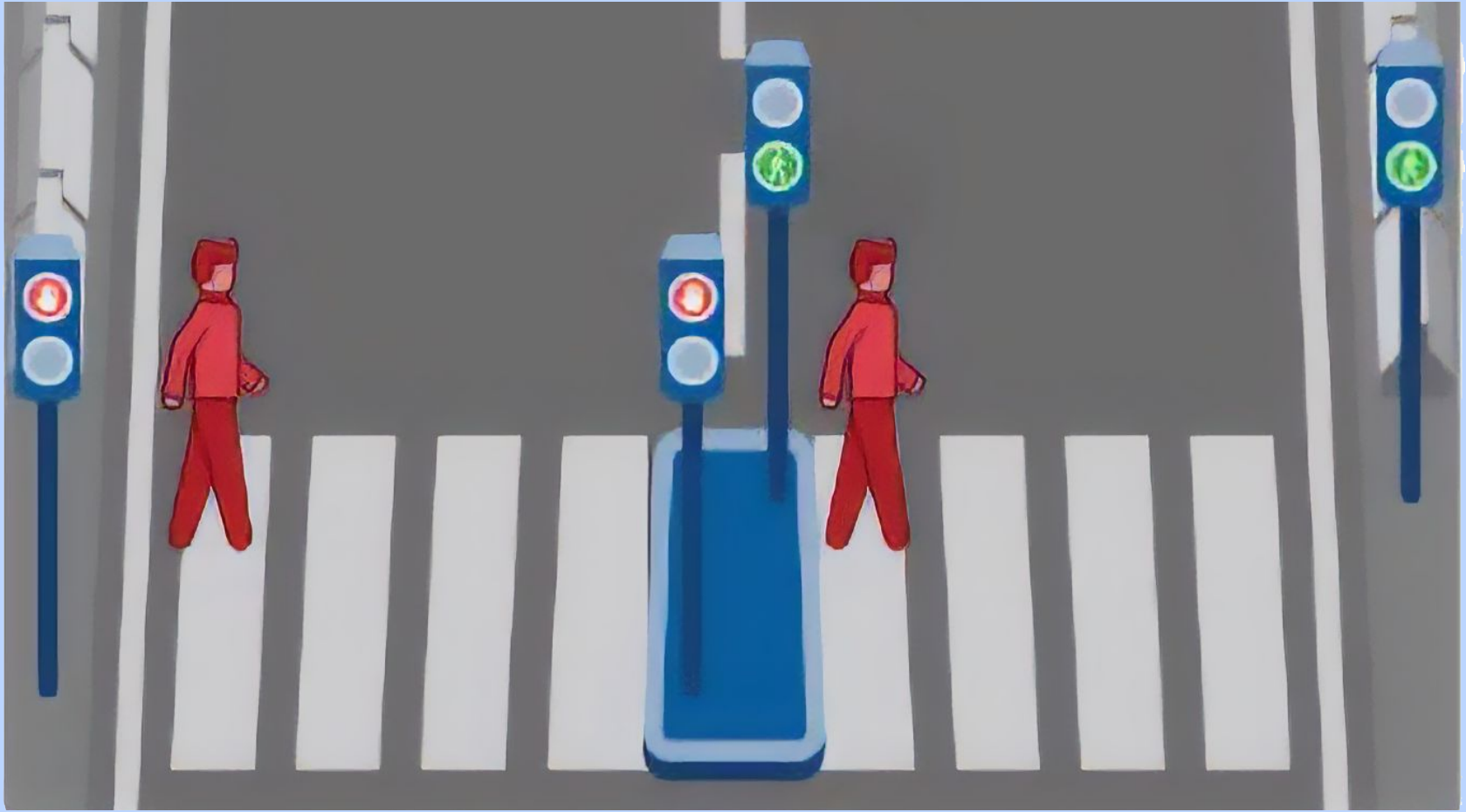


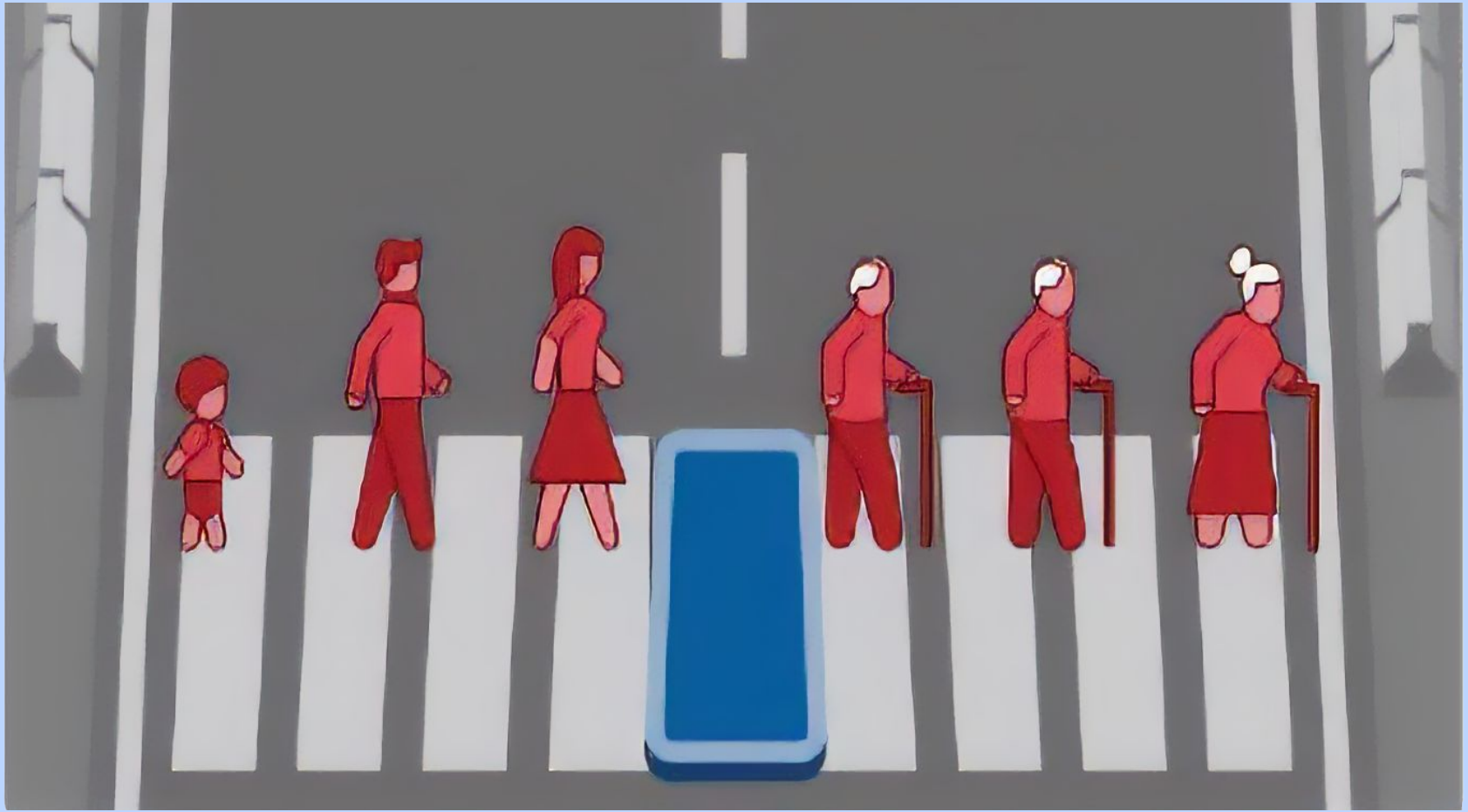
Pregnant woman...

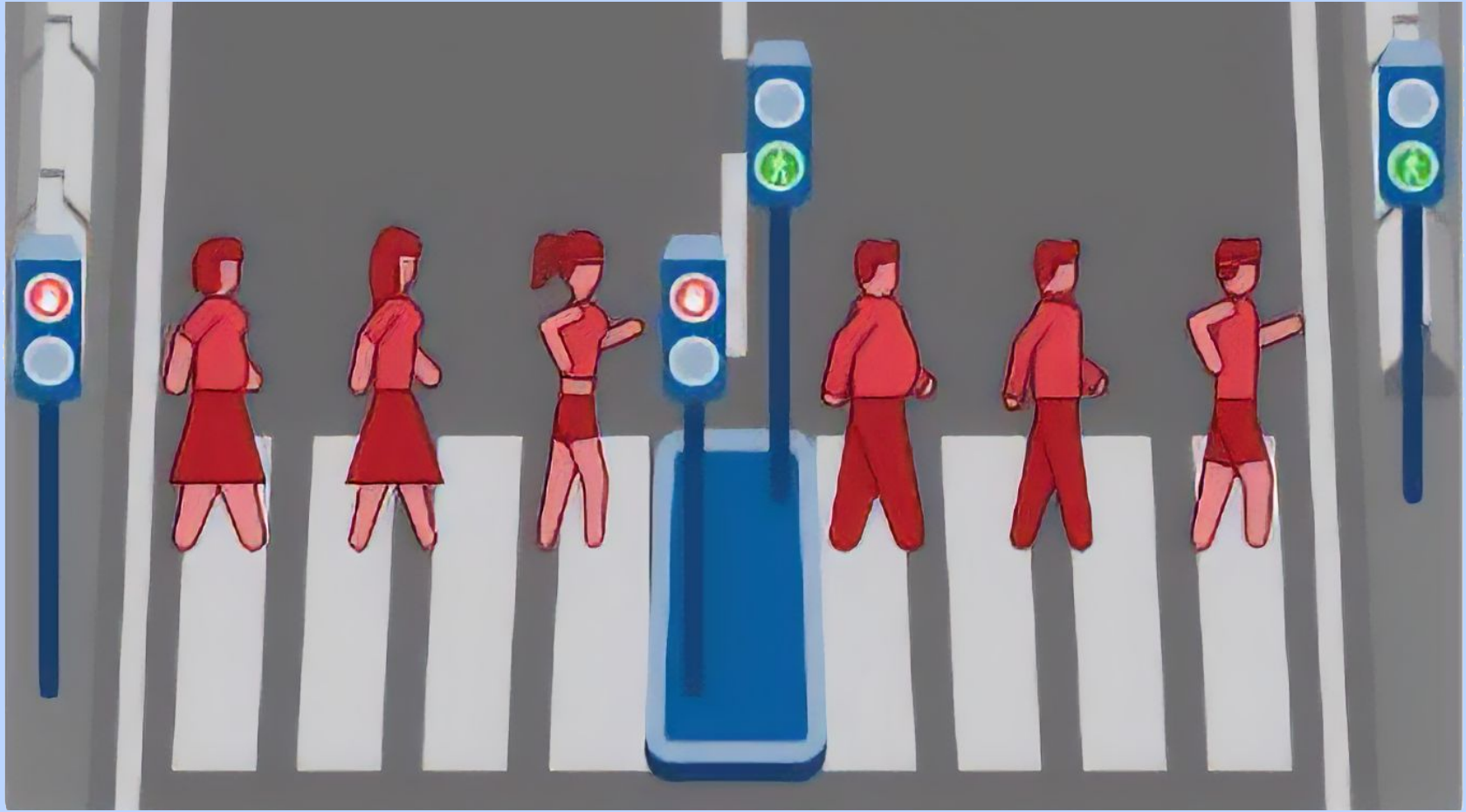


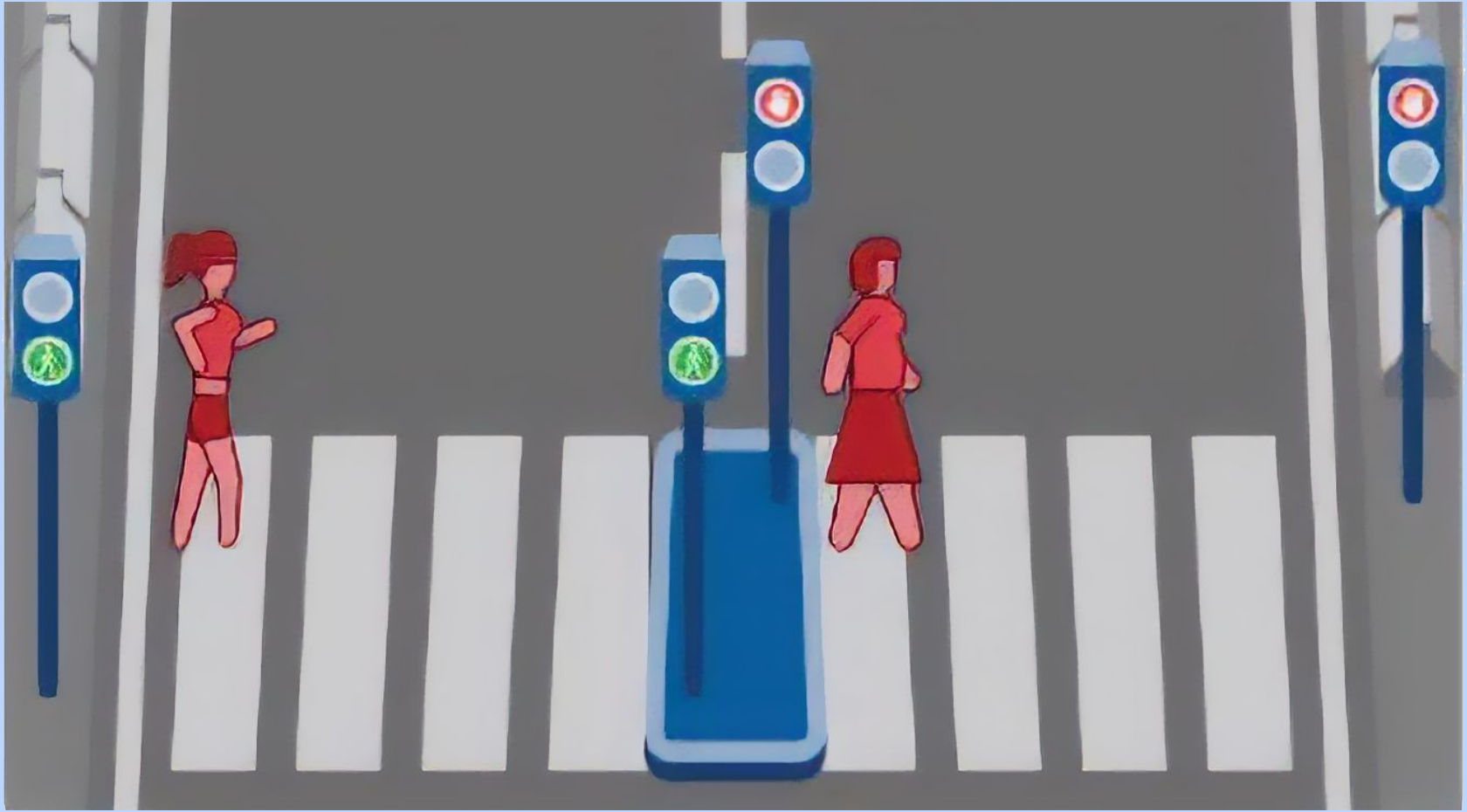


or dog?

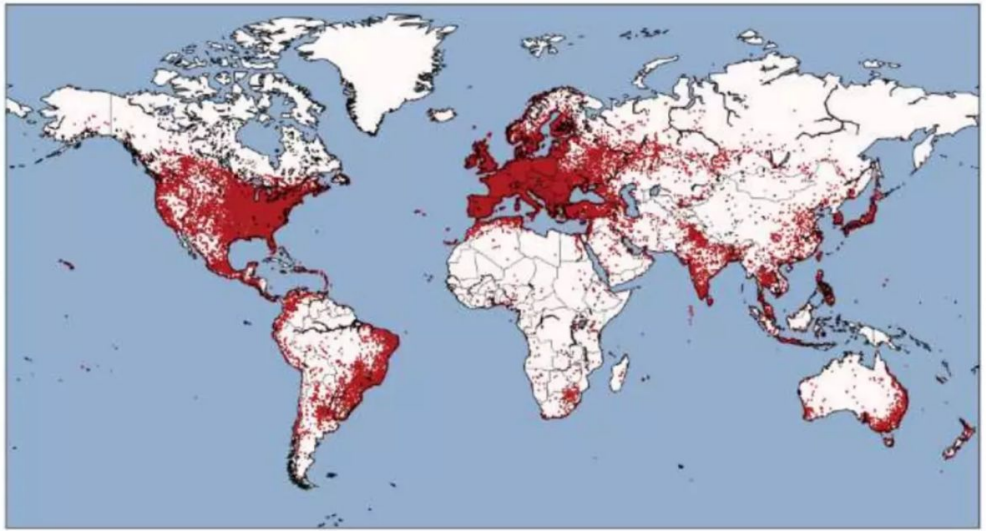








Who decides? - We all do



The Moral Machine Experiment platform gathered 39.61 million decisions in ten languages from millions of people in 233 countries

Global consensus on just three points:

- People want to save lives.
  They prefer to kill fewer people.

- Most people favor saving humans over pets.

- Most people favor saving children over adults.

- Do you believe that your decisions on Moral Machine will be used to program actual self-driving cars?
- To what extent do you feel you can trust machines in the future?
- To what extent do you feel that machines will become out of control?

Examples

"Explain quantum computing in simple terms"

"Got any creative ideas for a 10 year old's birthday?"

"How do I make an HTTP request in Javascript?"

Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests

Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Steve Wozniak
Elon Musk
Yuval Noah Harari
Andrew Yang
Valerie Pisano
Connor Leahy
Kate Jerome
Jaan Tallinn
Tristan Harris
Danielle Allen
Marc Rotenberg
Gary Marcus

Sam Altman
@sama

DALL·E 2 is here! It can generate images from text, like "teddy bears working on new AI research on the moon in the 1980s".

It's so fun, and sometimes beautiful.

openai.com/dall-e-2/

10:39 AM · Apr 6, 2022

Smoke inpainted in an image of the White House. Credit: OpenAI

BCG study, Sept 15, 2023

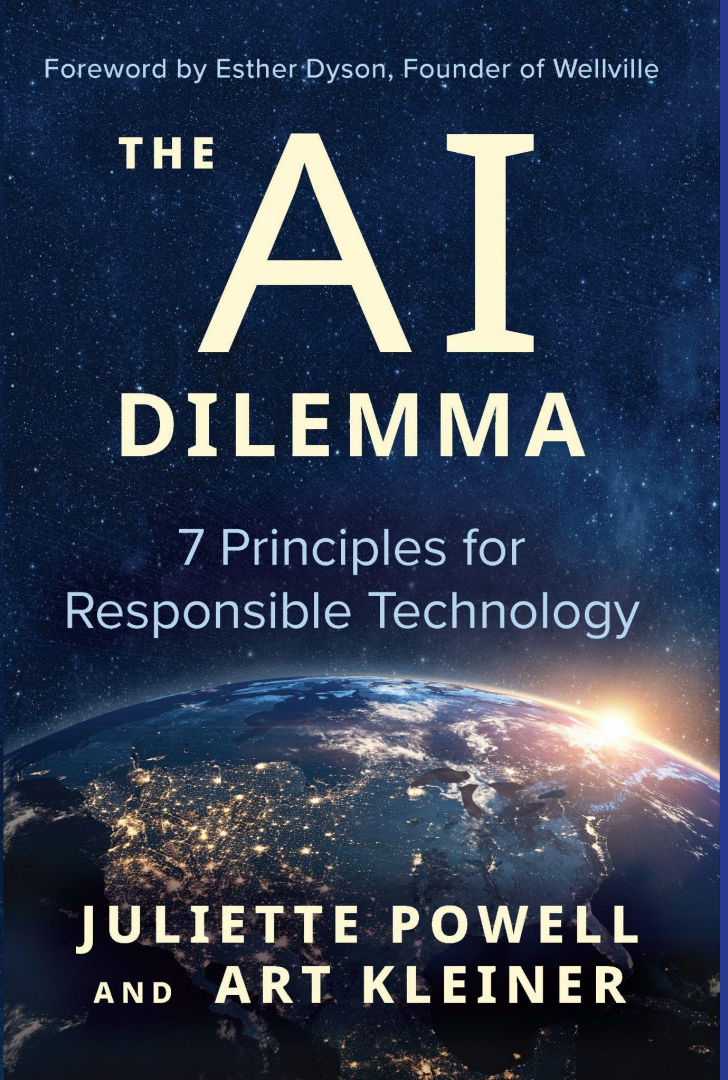
BCG consultants who used AI were 12% more productive and completed work 25% faster.

They also produced 40% higher quality results.

BCG study, Sept 15, 2023

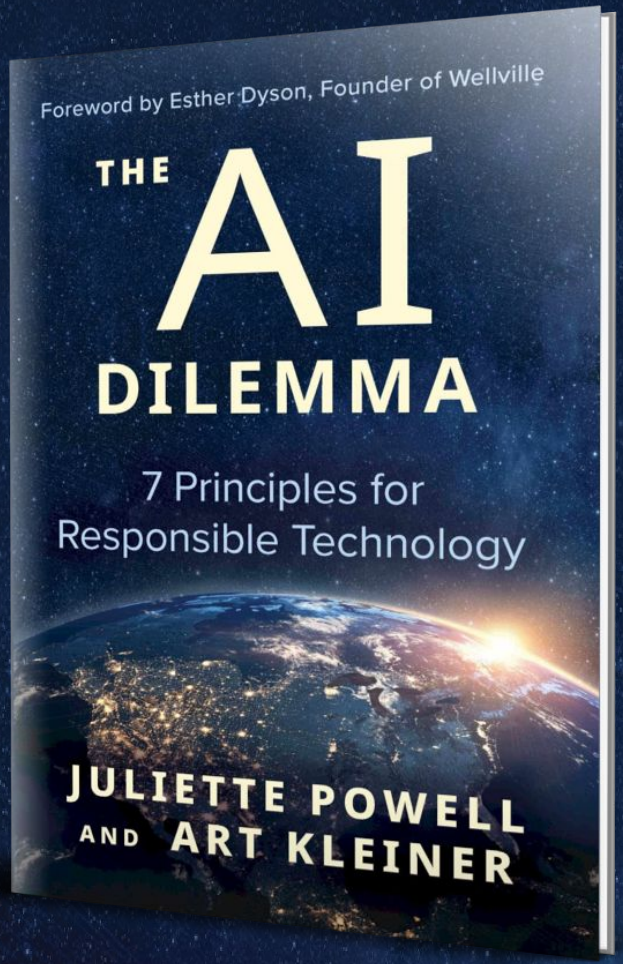
Below average workers saw performance increase by 43%.

Above average workers saw 14% gains.



The technology in the wrong hands is dangerous.

In the right hands, AI can be beneficial to all.

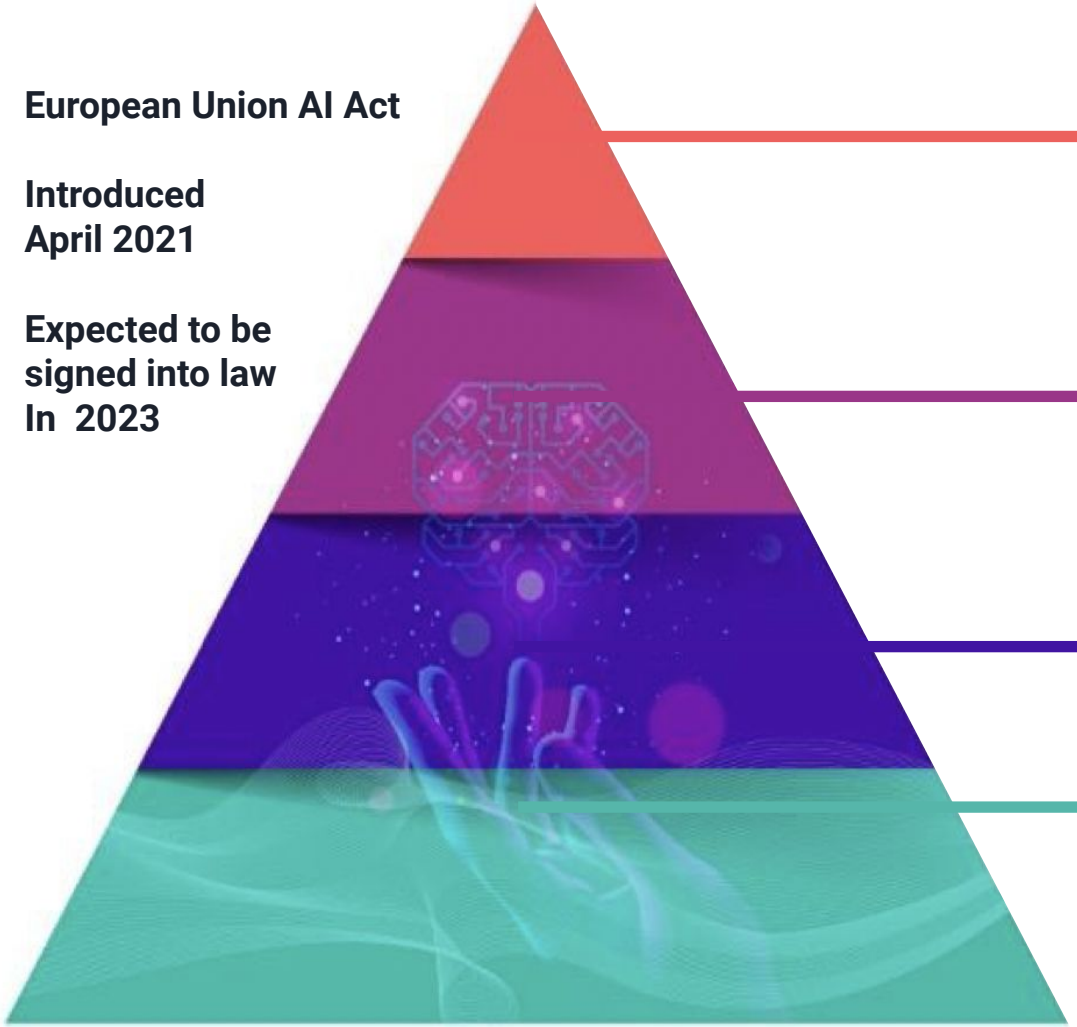


EU (AIA)
Classification-based: 4 categories of risk
Geography: Worldwide (if EU data is used)
Prohibited Harmful manipulation of individuals Social scoring that targets groups Real-time biometric identification
Oversight: EU AI board and member states Expected to demand independent audits
Key uncertainties: Whether this will influence worldwide behavior as GDPR did.
Fines: Up to 6% of global income.

European Union AI Act

Introduced
April 2021

Expected to be
signed into law
In 2023

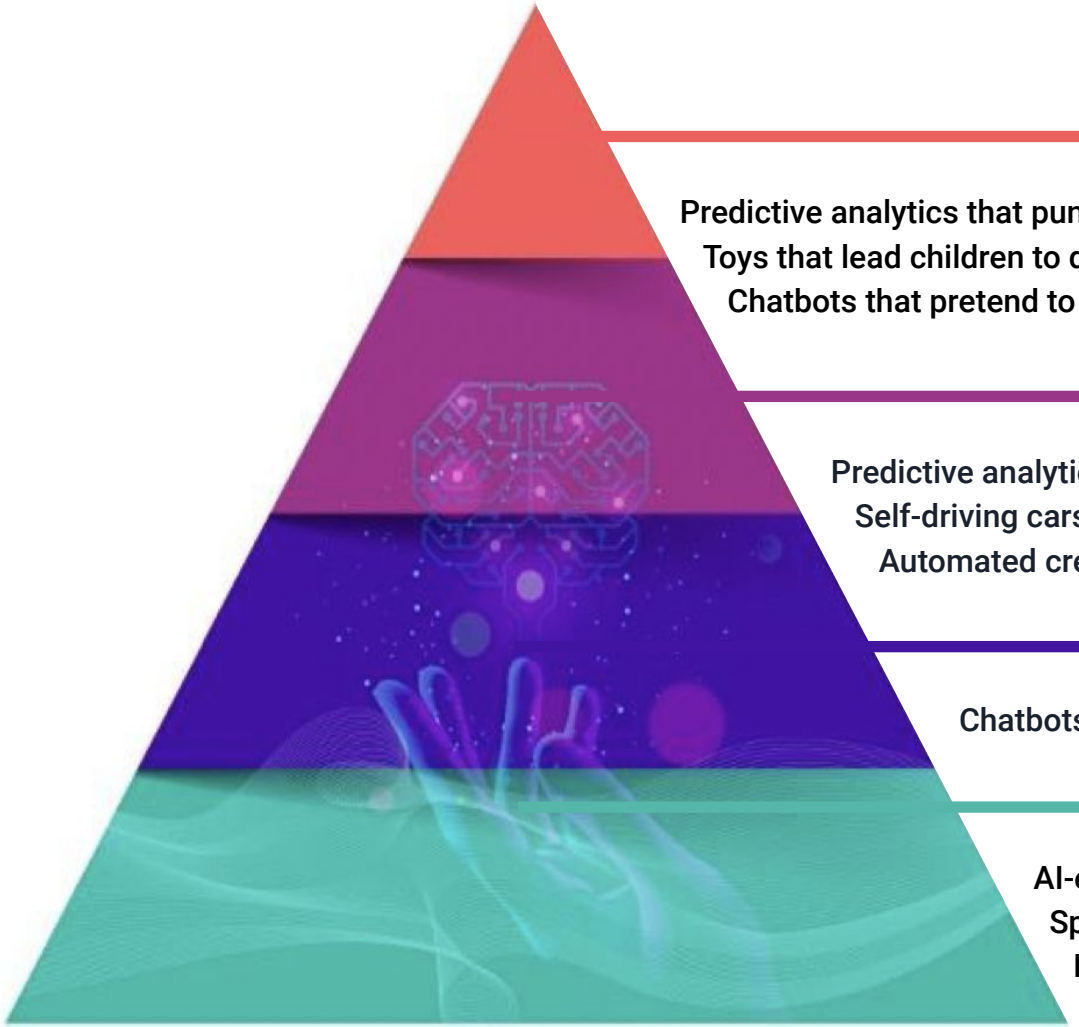


UNACCEPTABLE RISK

HIGH RISK

LIMITED RISK

MINIMAL RISK



UNACCEPTABLE RISK

Predictive analytics that punish people
Toys that lead children to dangerous behavior
Chatbots that pretend to be human

HIGH RISK

Predictive analytics that detect child abuse
Self-driving cars
Automated credit check systems

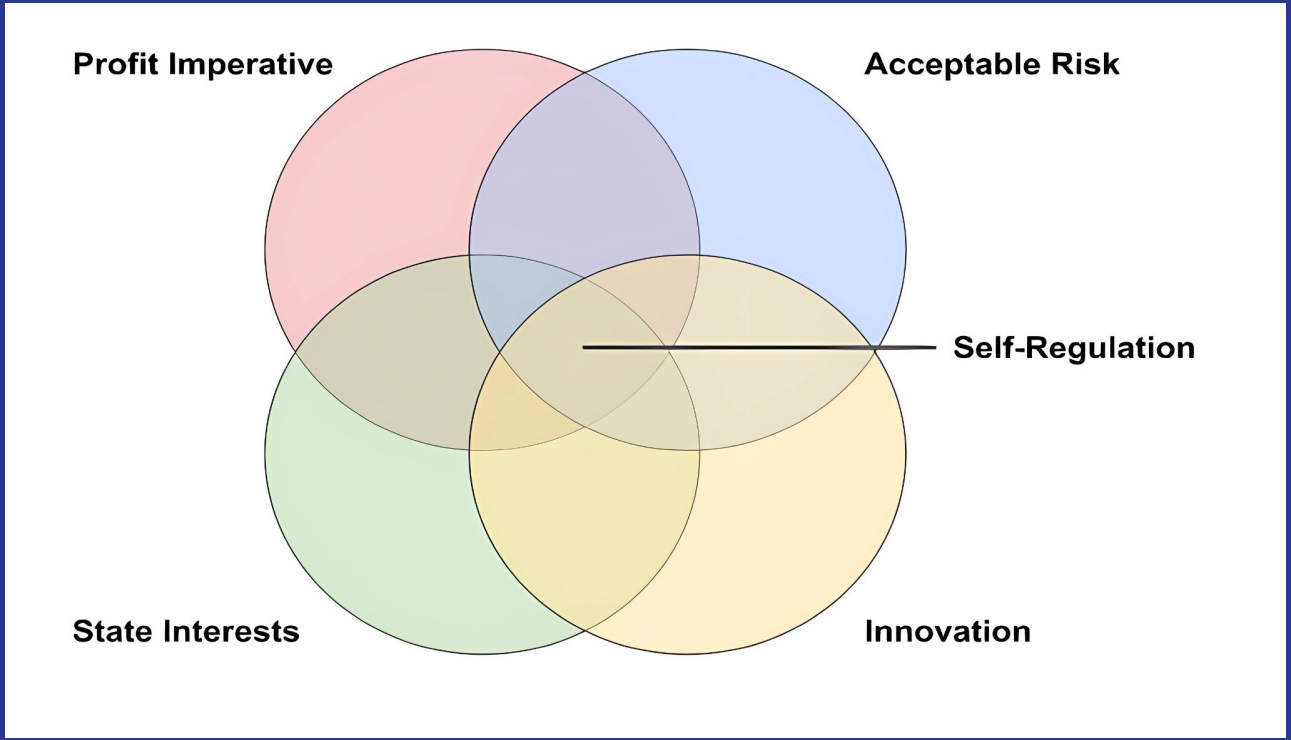
LIMITED RISK

Chatbots that reveal they're machines

MINIMAL RISK

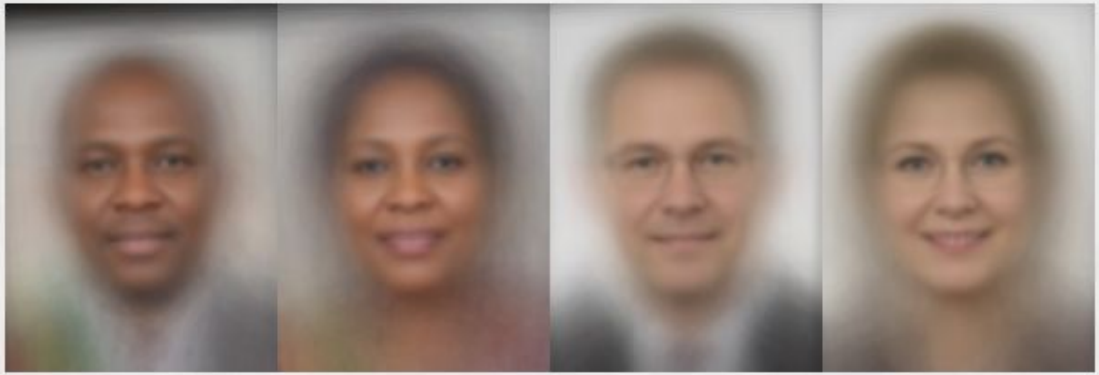
AI-enabled video enhancement systems
Spam filters
Inventory management systems

The Apex Benchmark: A Model of Self-Regulation



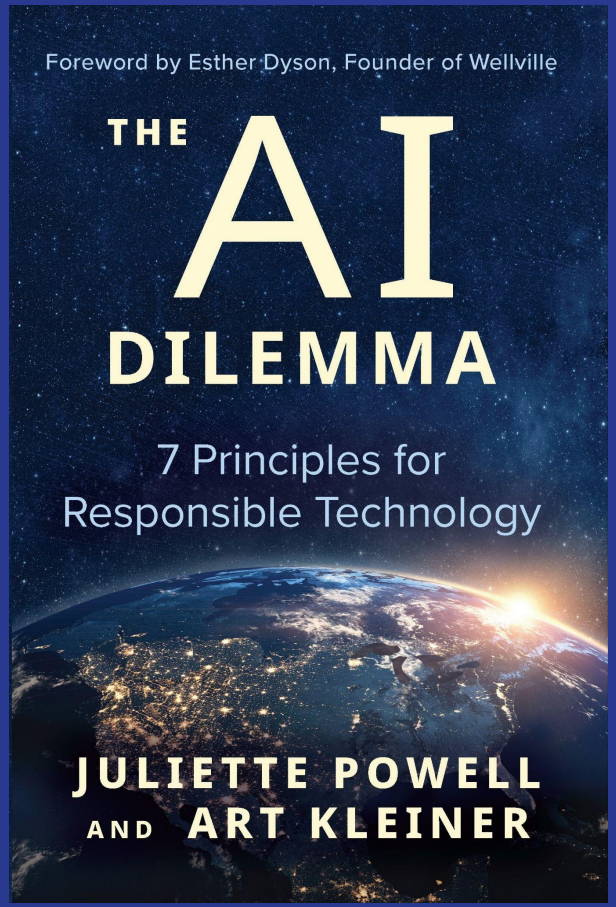
Confront and Question Bias: AI Mirrors Human Biases

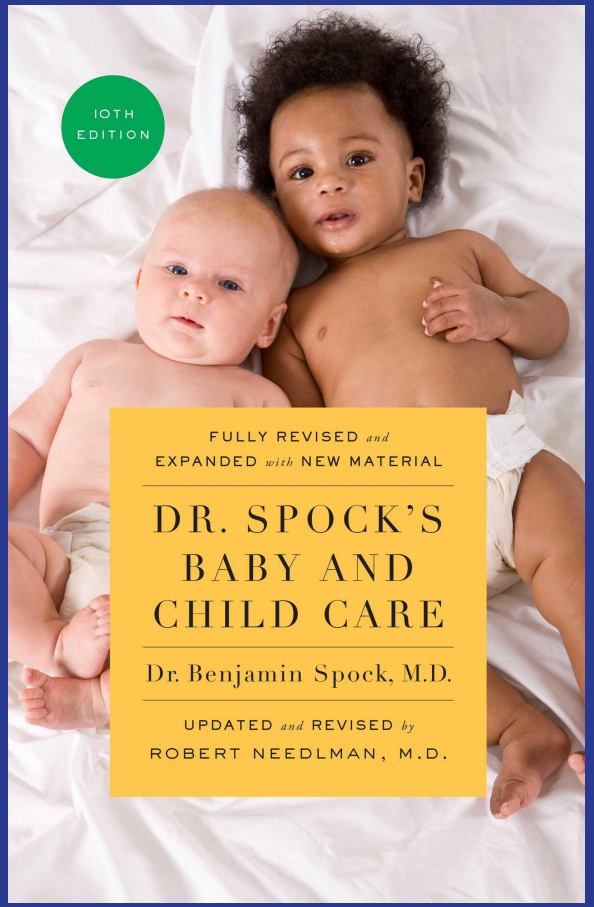
Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%



Hold Stakeholders Accountable

The government of the Netherlands was not fully held accountable for their anti-fraud system's false positives.


Foreword by Esther Dyson, Founder of Wellville
THE AI DILEMMA
7 Principles for Responsible Technology
JULIETTE POWELL AND ART KLEINER


10TH EDITION
FULLY REVISED and EXPANDED with NEW MATERIAL
DR. SPOCK'S BABY AND CHILD CARE
Dr. Benjamin Spock, M.D.
UPDATED and REVISED by ROBERT NEEDLMAN, M.D.

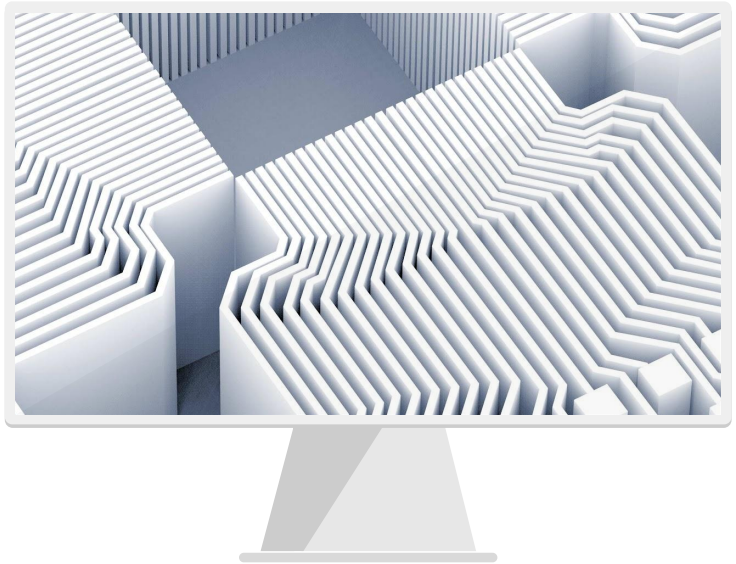
Dr. David Stark's Research, Columbia University

More Diverse Teams Deliver

Four Logics of Power	Private	Public
Individual	**Engineering logic** Technology	**Social Justice logic** Humanity
Institutional	**Corporate logic** Ownership, Markets, and Growth	**Government logic** Authority and Security

The 4 Logics of Power

Corporate Logic is the logic of markets and competitive advantage

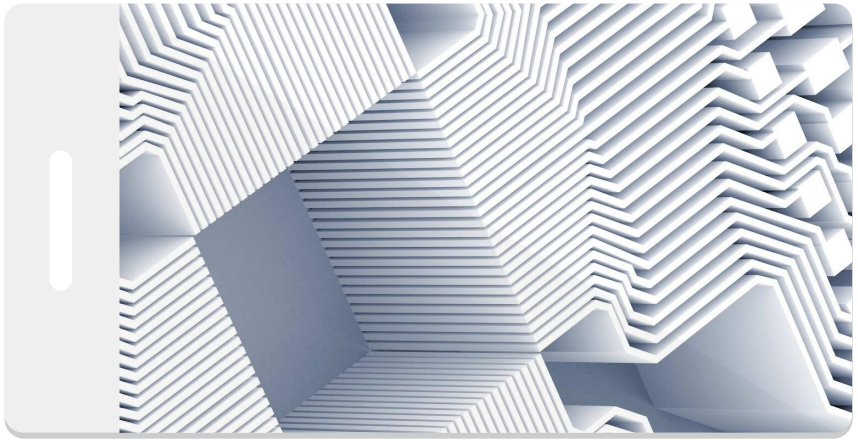


Prioritizes profit, growth, expansion, new business, all on behalf of shareholder value

The 4 Logics of Power

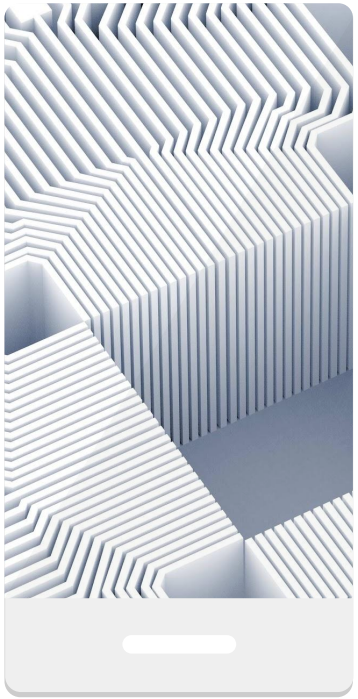
Engineering Logic is the logic technologists use.

Prioritizes efficiency, and seamlessness. This view values technology as a way to solve human and technological problems



The 4 Logics of Power

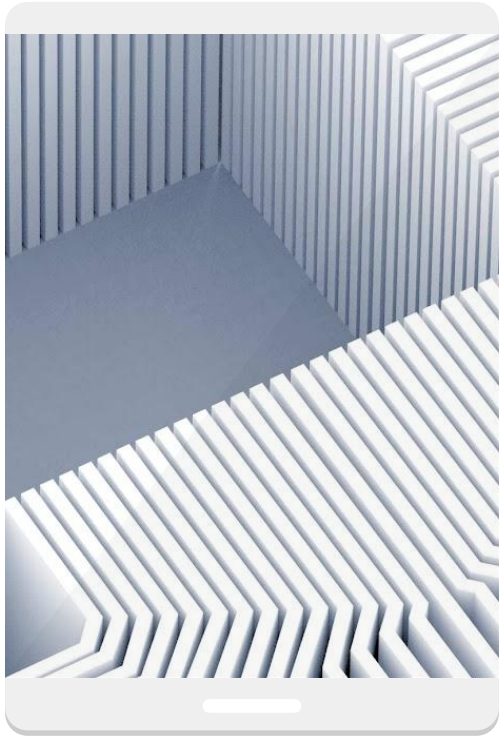
Social Justice logic
perspective prioritizes humanity.



In this view, people are more important than profit or efficiency. This view values people as a way to solve technological and human problems.

The 4 Logics of Power

Government Logic is the view of authority.



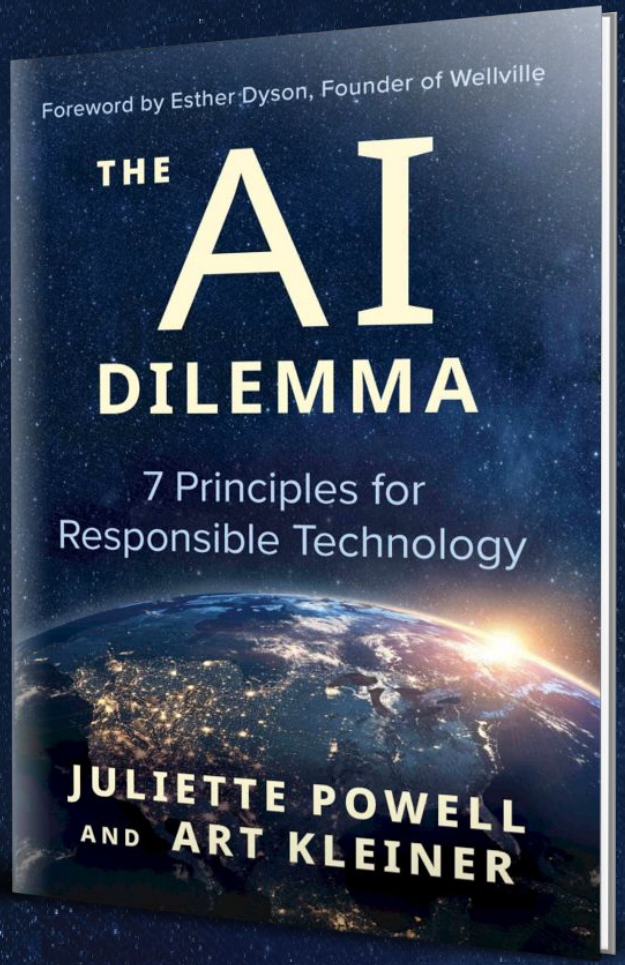
Prioritizes law and order.

This view values technology as a way to track, serve, prevent, and protect people, as well as institutions.

What is the AI Dilemma?

The technology in the wrong hands is dangerous.

In the right hands, AI can be beneficial to all.



Patterns of successful AI use at work (BCG study):

- *Delegate rote tasks to AI.*
- *Integrate AI in the entire task flow; humans are in constant interaction with the AI.*

The Calculus of Intentional Risk

What is the risk/reward to using AI in your organization?

- What is your reputational risk?
- Will you run afoul of regulation?
- Will insurance cover lawsuits?

The Calculus of Intentional Risk

How much will it cost you not to use AI in terms of securing a competitive advantage?

The Calculus of Intentional Risk and Self-Regulation

What are you NOT willing to do to make money?

Responsible Use of Data + Technology

"THE PRICE OF GREATNESS IS RESPONSIBILITY."

- Winston Churchill

Thank You!
Any Questions?

Find more on our AI advisory work, our research, and *The AI Dilemma* book at JuliettePowell.com

JP@JuliettePowell.com