

AI research has become misaligned with what people want from AI

Emma Pierson
UC Berkeley



Outline

- What do people want from AI?
- What is CS research on AI pursuing?
- What should we make of this?

I. What do people want from AI?

Polls are awesome, and underused by computer scientists

- Representative polls from high-quality pollsters are great for questions about AI attitudes, impacts, and use
- CS people:
 - Like fancy methods for data extraction
 - Underestimate the importance of representative data

Samples matter for AI attitudes, a *lot*

- Anthropic survey: 66% of US respondents positive on AI
- Pew survey: 10% positive, 50% negative
- What (primarily) explains huge gap?
Anthropic surveyed *Claude users*, Pew *general public*
- Public is also **60 percentage points** more negative than AI experts

Huang et al. What 81,000 people want from AI. Anthropic Blog, 2026.

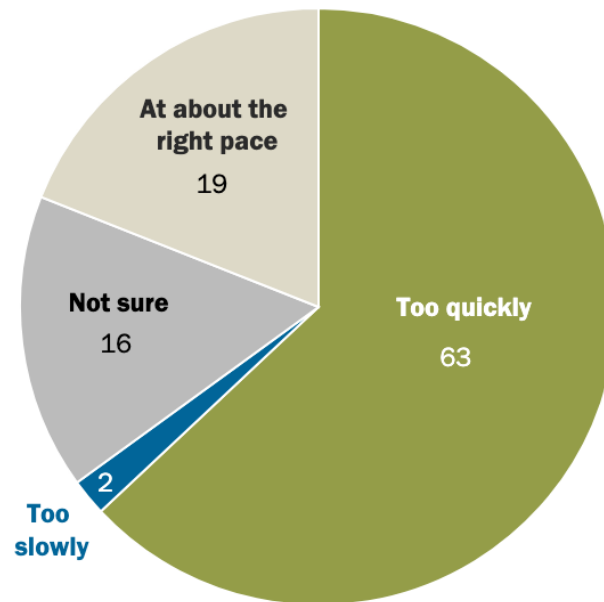
How people around the world view AI. Pew Research Center, 2025.

How the US Public and AI Experts View Artificial Intelligence. Pew Research Center, 2025.

People want AI to slow down

Roughly two-thirds of Americans say AI is advancing too quickly

% of U.S. adults who say they think AI is advancing ...

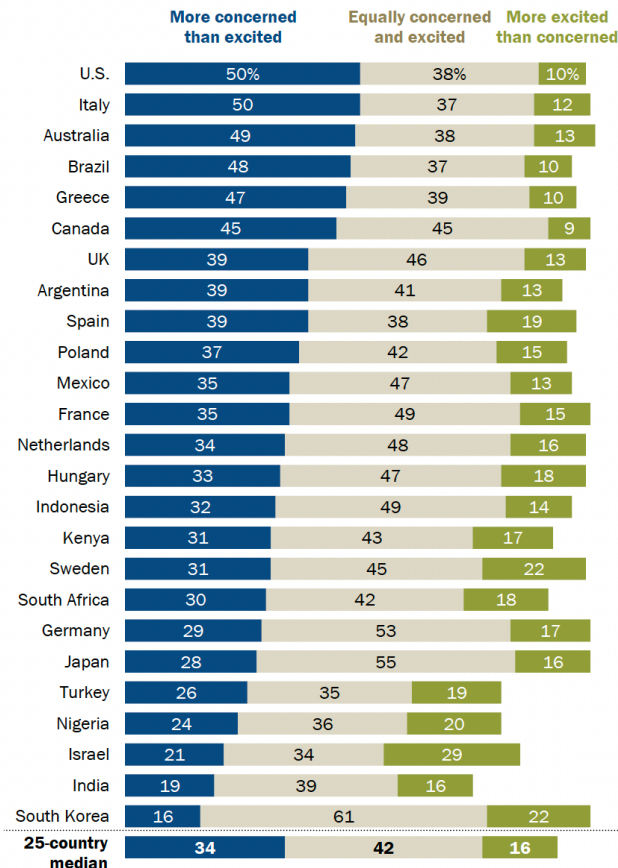


Americans and AI 2026: Chatbots, Smart Devices, and Views on Impact.
Pew Research Center, 2026.

They are more concerned than excited about AI in general

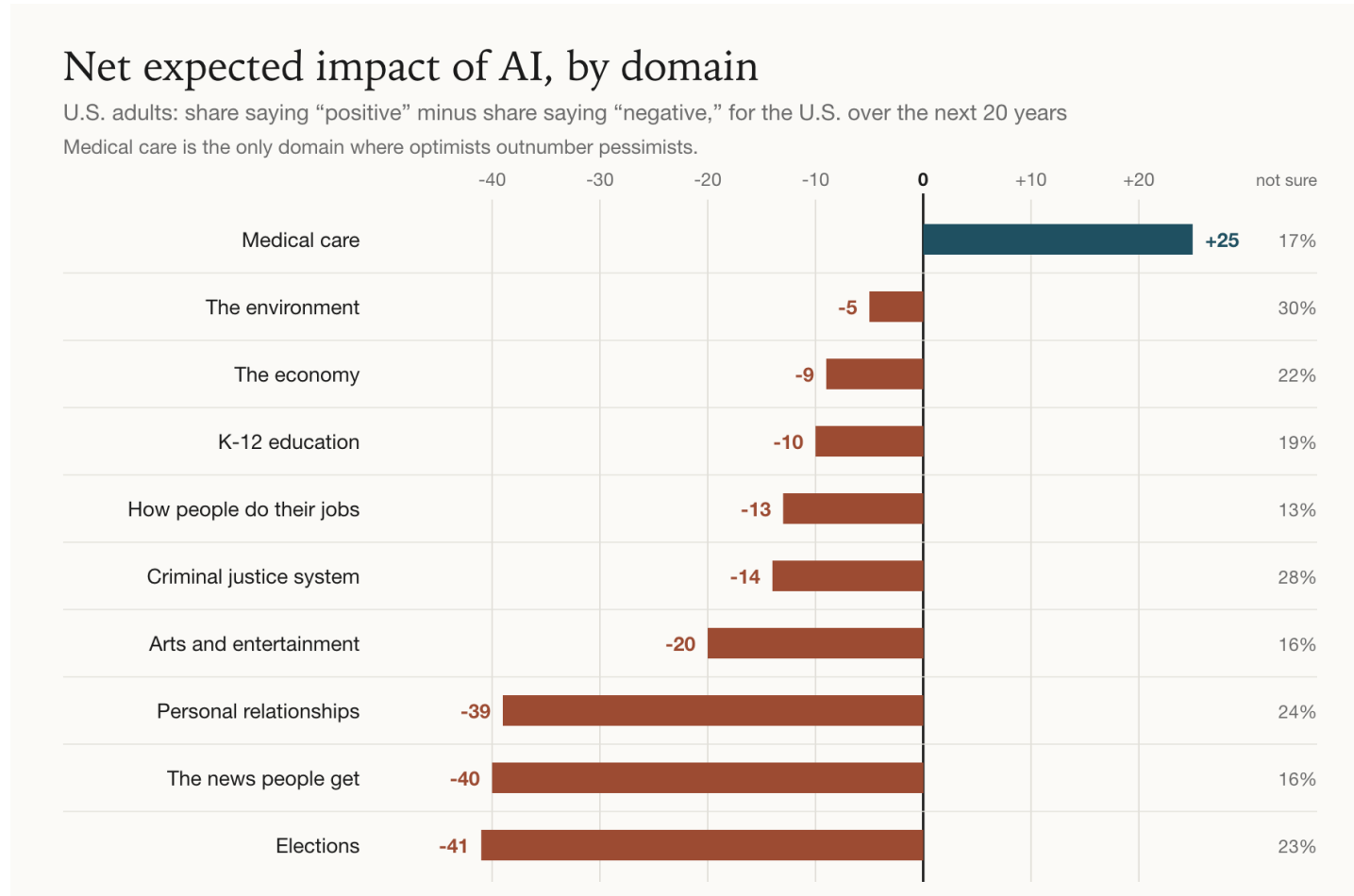
How do people around the world feel about the rise of AI in daily life?

% who say the increased use of artificial intelligence in daily life makes them feel ...



How people around the world view AI. Pew Research Center, 2025.

While people are not excited about AI in general, they are more positive about specific applications



American Trends Panel, Wave 152. Pew Research Center, 2024.

Overall

- People are unenthusiastic about racing forward on generalist AI
- They are more excited about specific AI use cases that might concretely improve their lives
- (Have focused on polls, but interviews by journalists also corroborate this)

II. What is CS research on AI pursuing?

Rarely, what people actually want from AI

- Only a small minority of CS research studies specific AI use cases that might concretely improve people's lives

Some numbers

- I downloaded arXiv papers from 2025 - 2026 (main website where people post CS research), filtered for AI papers, and classified whether each paper solved a concrete, real-world problem that:
 - Did not itself result from AI (e.g., AI misinformation)
 - Most people would agree is uncontroversially good (e.g., not surveillance or military uses)
- Preliminary work done quickly for this talk; take numbers with a grain of salt

Some numbers

- Emma count: 21% of papers passed this test
- Claude count: 17%
- Note: this is subjective, and arguably generous
 - I classified all health papers as “solving real problems that might improve people’s lives” but, having worked in this area, I think a very small fraction are really improving people’s lives

Overall

- Only a small fraction of CS research focuses on what polling suggests people want from AI: using it to concretely improve their lives

III. Is this bad?

To be clear

- I am not arguing that AI research is useless unless it concretely improves people's lives
 - Much of my work doesn't!
 - General AI evals, theory, safety, etc are obviously important

But I do think the overall misalignment in priorities is concerning

- The public pays my salary and supports a great deal of research
- If most of our research effort is devoted to a generalist technology they fear and want to slow down, and not to producing use cases they like, that seems *prima facie* concerning
- Let's consider some counterarguments

Counterargument 1: “Public opinion should not determine research goals”

- Maybe not, if the public is simply indifferent or uninformed
 - Real analysis, metaphysics
- But with AI the public is actively *angry and fearful*, and for understandable reasons

Counterargument 2: “We need to improve AI generally to benefit specific use cases”

- But general improvements in AI also have a lot of downsides, which is why the public is afraid of them
- Also, I think we can already do way more good with the AI we currently have
 - It surpasses our wildest dreams of what was possible five years ago!
 - Helen Toner: “I just think that there’s a lot of juice we could get out of that with models available today if we put the time and the legwork in.”

Counterargument 3:

“Practical impact is the job of industry”

- Industry is often poorly aligned with social good
- Academia's comparative advantage at the moment may well be pursuing socially beneficial AI use cases:
 - We're badly outgunned on general methods work
 - We don't need to make money
 - We have ready access to interdisciplinary collaborators

Counterargument 4: “Other fields, not CS, should translate AI into practical impact”

- But CS people are essential technical collaborators on these teams
- Example (one of many): our collaboration with a county ambulance service
 - They want to understand why there are racial disparities in pain treatment using their clinical notes
 - We bring: the methods (imported from the AI safety folks); the engineers; the hardware

Counterargument 5: “The core purpose of CS is to develop and evaluate general methods”

- This ethos runs deep in our field
 - *Real* computer science
- Historically, this “harder better faster stronger” attitude has often been well-aligned with the public
 - People *do* want faster/more powerful computers, internet, cell phones
 - But we have become misaligned.
People do not want faster/more powerful AI.
- Perhaps we should reconsider our ethos? Times are changing fast, and our field has always changed fast to meet them.

Summary

- What do people want from AI?
Things that concretely improve their lives.
- What is most CS research on AI pursuing?
Not that.
- Is this bad?
In my opinion, yes.

Some tech CEOs are open to this...

- “I think by far the most accurate criticism of AI companies including Anthropic is that we haven’t yet delivered on our big promises to benefit the world. That is totally on us...”

Please reach out if

- You agree with this take
(maybe we can write something)
- You disagree with this take
(maybe you can make me less
heartsick about my field)
- First pass at this argument;
can certainly be refined
- emmaperson@berkeley.edu