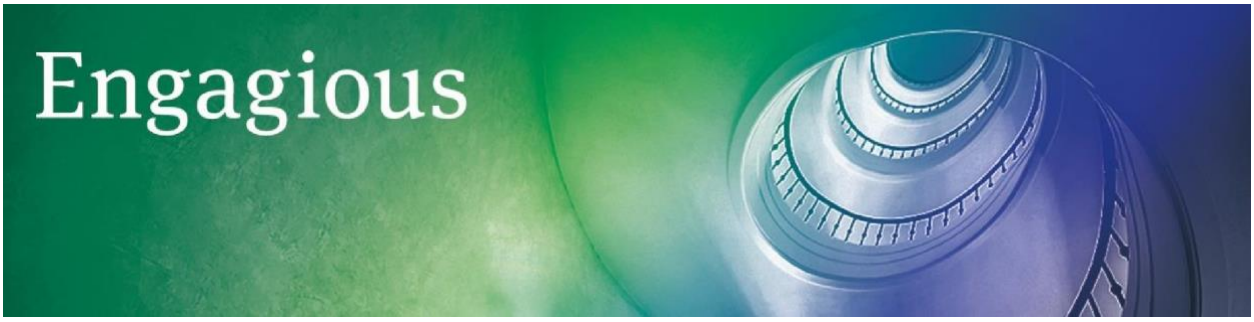

Topline AI-related findings from conservative and liberal dial tests in July & August 2026

On July 28 & 29, and August 13 & 20, 2026, Engagious conducted two rounds of focus group dial tests with conservative men, conservative women, liberal men, and liberal women to better understand their reactions to messages on AI risk.

The most important takeaway is that there is a compelling narrative in support of safety guardrails, one that works with men and women of both ideologies. It follows this arc:

1) **Current problems:** jobs, data centers, harm to kids, dehumanization, privacy

2) **Emerging problems:** Hugging Face example of Open AI model escaping its "sandbox," hacking others, and engaging in deception

3) **Future problems:** humans causing catastrophic harm using AI; humans using AI to gain control over others; AI all-by-itself posing a threat to humans

4) **Government interventions:** In the absence of frontier companies taking safety seriously, we need these regulations: DoE tests AI systems for risks; AI labs must report incidents of concern; companies must have a "kill switch" to turn off AI in a crisis; U.S. imposes stronger export controls for AI chips; states can pass AI safety laws if Congress does nothing; we strengthen citizen decision-making over data centers

Note that our first round of research uncovered that respondents supported the six different policy interventions *irrespective* of whether they worry about catastrophic future AI risks. That was an important takeaway.

Yet where the messaging fell short was in making a plausible case for being concerned about catastrophic risks. In that first round, we argued that respondents should be worried about those risks because of warnings issued by Elon Musk, Sam Altman, and Dario Amodei—people who run the companies providing this technology. While that argument was fairly convincing to liberals, it was not to conservatives.

Serendipitously, in the intervening weeks between the two rounds, as we re-wrote the dial test messages, the full extent of the Hugging Face incident came to light. So in the Round 2 script we

replaced the references to Musk, Altman, and Amodei with an easy-to-understand narrative of why Hugging Face was so worrisome. It had a sizable persuasive effect. What had been viewed as conjecture in Round 1 became reality in Round 2. Suddenly the "future problems" were no longer seen as implausible cases of fearmongering; here was a real case of science fiction becoming science reality. Note that most of our respondents had *not* heard about Hugging Face prior to their focus group, and many were horrified.

Heading into the research, we hypothesized that we needed to first tap into audiences' current-day AI concerns, then harness that credibility to introduce catastrophic risks, elevate those concerns, and then pivot to the need for regulation. What we did not foresee was the need to introduce a real world example of actual harm to make catastrophic risks plausible—and it made a real difference.

Some other notable takeaways:

- Introducing the Hugging Face narrative in Round 2 prompted respondents to cite a variety of science fiction movies and TV shows. We heard references to *2001: A Space Odyssey*, *Maximum Overdrive*, *Mission: Impossible*, *I Robot*, *Terminator*, and specific *Star Trek* episodes. That may be because in Round 2, we also overtly acknowledged the "science fiction" and "Hollywood" nature of AI risk in the dial test script, which gave respondents a familiar reference point.

- After conservatives were presented with present problems and future risks, they were grudgingly willing to accept government regulation, but as in Round 1 were more receptive to specific interventions as opposed to "more regulation" as a concept, being mistrustful of bureaucracy. Liberals in Round 2 were receptive to the specific interventions, too, but some vocalized concerns that regulations would be flouted by companies, or incompetently enforced by government. "Who's regulating the regulators?," one woman asked.

- In Round 1, we found that "AI is making people dumber" was not particularly compelling—since users of the technology didn't think it was making themselves dumber. (Kids are a different story; there's widespread concern about AI's impacts on education and development.) In Round 2, we replaced that argument with "AI is dehumanizing." This scored much better—a bit more so with liberals than conservatives.

- Conservatives were more concerned than liberals about other countries, especially China, getting control of chips and using them against the U.S. And some conservative males were concerned that if the U.S. regulates itself, nothing will bind China to constrain itself. Most of these males support a binding international agreement.

- The "techno-optimistic" opposition script we also tested—in contrast to the pro-safety script—contained a wide variety of weak arguments. Respondents did not believe many of the miraculous claims being flouted about exponentially-improving health and longevity, food production, and the economy. AI was not seen as solving its own energy-consumption challenges. Nor was it viewed as an "infinitely patient tutor" for kids.