

EMBARGOED UNTIL July 30 2025

Contact: Karl Koch, karl.koch@aiwi.org

New Campaign Calls on AI Companies to Publish Their Whistleblowing Policies
OpenAI Is the Only Frontier Company That Currently Makes Their Policy Public

WASHINGTON, DC - As part of the [National Whistleblower Day Even](#) on Capitol Hill, a new campaign will be launched to call on AI companies to publish their whistleblowing policies and thereby open them up to public review. Backed by over 30 global organizations and experts in AI governance and whistleblower protection, including Stuart Russell, [Lawrence Lessig](#), [Transparency International](#), and the [National Whistleblower Center](#), the campaign aims to ensure these policies truly protect those who speak up.

AI carries great potential, but also great risks. Many of these risks will only be visible to company insiders, who have both access to internal information and the expert knowledge needed to understand it. Insiders should be able to safely raise their concerns, internally and externally – a call echoed in the “[Right to Warn](#)” letter by current and former AI insiders, and in the recently released Future of Life Institute’s [AI Safety Index](#). The first step, and the campaign’s focus, is for frontier AI companies to publish their policies, a minimum standard to help protect insiders and the public. The second step for these companies is to provide further transparency around the effectiveness of these policies.

No frontier AI companies, including Anthropic, Google DeepMind, Meta, xAI and Mistral, have published their whistleblowing policies. An exception is OpenAI, who published their policy following revelations around their highly restrictive non-disparagement clauses. The public should not wait for further scandals, and the campaign calls on these companies to enhance transparency today.

“Everyone wins when companies commit to sharing their whistleblower policies,” says Karl Koch, founder of [The AI Whistleblower Initiative](#) (formerly known as OAISIS), a nonprofit leading the campaign. “It helps insiders and the public to understand what processes and protections are in place, and where there are gaps. Public feedback can further help improve policies and protections - which also benefits companies through increased employee satisfaction and innovation, as evidenced by [shareholder action calling on Google to improve its whistleblowing system](#). Commitment to transparency can help address [declining public trust in AI companies](#).”

“Companies from a range of industries, including some tech companies, make their whistleblower policies public. AI companies should show the same leadership,” say Naomi Colvin, Program Director at [Blueprint for Free Speech](#), another campaign member.

Making whistleblowing policies public – and subjecting them to public scrutiny – is an important first step toward ensuring they are strong enough to support insiders in reporting problems. Following that, whistleblowing policies must be backed by strong implementation, an organizational culture that encourages the reporting of safety concerns, leadership committed to taking whistleblower reports seriously, and transparency about how whistleblowing systems function in practice.

###