

L4 -Regulatory Mapping- Anthropic

<https://assets.anthropic.com/m/785e231869ea8b3b/original/clause-3-7-sonnet-system-card.pdf>

<https://www.anthropic.com/news/anthropic-achieves-iso-42001-certification-for-responsible-ai>

<https://www.anthropic.com/news/eu-code-practice>

Anthropic achieves ISO 42001 certification for responsible AI

Jan 13, 2025

We are excited to announce that Anthropic has achieved accredited certification under the new [ISO/IEC 42001:2023 standard](#) for our AI management system. ISO 42001 is the first international standard outlining requirements for AI governance and helps ensure AI systems are developed and used responsibly.

Achieving this certification underscores our commitment to AI safety and responsible development. It provides independent validation that we have implemented a comprehensive framework to identify, assess and mitigate potential risks associated with our AI systems. Key elements include:

- Policies and processes to ensure AI systems are designed, developed, and deployed in an ethical, secure and accountable manner;
- Rigorous testing and monitoring to verify systems behave as intended and potential negative impacts are proactively addressed;
- Transparency measures to provide appropriate information to users and stakeholders;
- Established roles, responsibilities and oversight to maintain responsible practices.

We're proud to be **one of the first frontier AI labs** to achieve this certification and hope it provides further assurance to our partners and the public of our strong commitment to AI safety.

The [ISO 42001 certification](#) builds on our existing work to develop AI safely and responsibly, including the release and recent update of our public [Responsible Scaling Policy](#) governance framework, deployment of [Constitutional AI](#) to help models operate in alignment with human values, and active [research](#) into AI safety and robustness. Schellman Compliance, LLC, an ISO certification body accredited by the ANSI National Accreditation Board, issued the certification.

Continuing our commitment to responsible AI leadership

Beyond ISO 42001, we have made a number of [additional voluntary commitments](#) demonstrating our dedication to responsible AI development. These voluntary frameworks cover key areas such as risk assessment and mitigation, security and privacy safeguards, and public awareness efforts.

As AI systems become increasingly powerful, safety and security are at the forefront of our ongoing research and innovation. The ISO 42001 certification is an exciting milestone in that journey.

Information about our ISO certification is now available on our [Trust Center](#). This includes the issued certificate, an overview of the standard, what it entails, and how it reinforces our commitments to responsible AI development.

Anthropic to sign the EU Code of Practice

Jul 21, 2025

After review, Anthropic intends to sign the European Union's General-Purpose AI Code of Practice. We believe the Code advances the principles of transparency, safety and accountability—values that have long been championed by Anthropic for frontier AI development. If thoughtfully implemented, the EU AI Act and Code will enable Europe to harness the most significant technology of our time to power innovation and competitiveness.

A recent analysis found that AI has the potential to add [more than a trillion euros per year](#) to the EU economy by the mid-2030s. The Code, working alongside Europe's [AI Continent Action Plan](#), demonstrates how flexible safety standards can both preserve innovation and enable broader AI deployment. This approach highlights the opportunities and imperatives required for Europe to remain competitive in this transformational technology. With transparent risk assessment processes in place, we can accelerate work to address Europe's most pressing challenges: advancing scientific research, improving public services, and enhancing industrial competitiveness.

We're already seeing signs of what's possible, from Novo Nordisk [accelerating breakthrough drug discovery](#), to Legora [transforming legal work](#), to the European Parliament [expanding access to decades of archives](#) to citizens. Ensuring these benefits materialize with minimal downside requires public visibility into AI safety and transparency practices while preserving private sector agility to deliver AI's transformative potential.

Building on our commitment to transparency

As outlined [previously](#), Anthropic believes the frontier AI industry needs robust transparency frameworks that hold companies accountable for documenting how they identify, assess, and mitigate risks. The EU Code establishes this baseline through mandatory Safety and Security Frameworks that build upon Anthropic's own [Responsible Scaling Policy](#) and will describe important processes for assessing and mitigating systemic risks. This includes

assessment of catastrophic risks—particularly those from Chemical, Biological, Radiological, and Nuclear (CBRN) weapons.

Maintaining flexibility

AI moves fast and changes constantly, which means the best policies are those that can be flexible and adapt alongside the technology.

Over the nearly two years since we first published our Responsible Scaling Policy, we've refined it several times based on practical insights from implementation. For example, our most recent update clarified which actors are in-scope for the [ASL-3 Security Standard](#); this determination was based on a deeper understanding of the relevant threat models and model capabilities.

As an industry, we're still developing best practices for assessing the systemic risks identified in the Code. Different risks require different methodologies. Third-party organizations like the [Frontier Model Forum](#) play a critical role, establishing common safety practices and evaluation standards that evolve with the technology. These groups bridge industry and government, translating technical insights into actionable policy.

We're committed to working with the EU AI Office and safety organizations to ensure the Code remains both robust and responsive to emerging technologies. This collaborative approach—combining regulatory frameworks with flexibility—will be essential for Europe to harness AI's benefits while competing effectively on the global stage.