



# TOWARDS A HIROSHIMA ARTIFICIAL INTELLIGENCE PROCESS CODE OF CONDUCT REPORTING FRAMEWORK: FINDINGS FROM THE PILOT

*REPORT OF THE G7 ITALIAN PRESIDENCY*

## Background: Hiroshima AI Process Code of Conduct Draft Reporting Framework

One of the priorities of the G7 Italian Presidency in 2024 is to facilitate the adoption and implementation of the Hiroshima AI Process International Code of Conduct for Organisations Developing Advanced AI Systems, welcomed as an outcome of the G7 Hiroshima AI Process. At the request of the G7 membership and in line with the commitments of the [Trento Declaration](#) (March 15, 2024), the G7 Digital and Tech Working Group under the Italian Presidency called on the OECD to identify and develop appropriate mechanisms to monitor the voluntary implementation of the Code of Conduct by organisations choosing to implement it. G7 Leaders in their [Apulia Communiqué](#) reaffirmed their commitment to developing a reporting framework in cooperation with the OECD.

To initiate this process, the OECD supported the development of a draft reporting framework, with input from experts from a broad array of organisations including businesses, academia, civil society and research institutes across G7 countries. To refine and assess the effectiveness of the draft reporting framework, the OECD conducted a pilot, in which a diverse range of organisations provided detailed feedback. The findings of the pilot were discussed at the [G7 Technology and Digital Ministerial Meeting](#) on 15 October 2024 in Cernobbio:

---

*“We extend our gratitude to all the participating organisations for their invaluable support in enhancing the reporting framework during this period which has been a truly multi-stakeholder effort, engaging companies, academia, civil society, international organisations, and governments in a collaborative exercise. In this regard, we note [the Italian Presidency’s overview of the OECD pilot of the Hiroshima AI Process reporting framework](#). We continue to work to develop the reporting framework with the aim to advance it by the end of the year, in collaboration with the OECD and the participating organisations.”*

---

This report, developed by the OECD Secretariat, summarises key feedback and highlights strengths and areas for improvement in the context of developing the operational “1.0” version of the reporting framework. It was prepared for publication by the government of Italy, as the G7 Presidency in 2024. The positions expressed herein do not necessarily reflect the official views of OECD Member countries, of G7 members, or of the Italian G7 Presidency.

Following discussions within the membership, it was decided that submissions during the pilot phase would not be disclosed. This approach encouraged open and candid input, allowing participants to share detailed insights and challenges, without concerns of disclosure of sensitive information, and fostered trust and constructive engagement. While responses to the pilot were confidential, responses to the operational version of the framework are expected to be made public and should not include commercially confidential or sensitive information.

## **Overview of pilot phase of the reporting framework**

The pilot phase of the reporting framework, conducted from 19 July to 6 September 2024, garnered responses from 20 organisations across 10 countries<sup>1</sup>. Organisations responding to the pilot of the reporting framework ranged from large technology companies to start-ups and included developers of AI systems, global technology companies, research institutes, academic institutions, and consulting firms involved in AI compliance and auditing.

The rich and well-rounded feedback is summarized in the following sections.

---

<sup>1</sup> Canada, Denmark, Germany, Israel, Japan, the Netherlands, Spain, Switzerland, the United Kingdom, and the United States

## Summary of feedback from the pilot phase of the reporting framework

Findings from the pilot phase include: (1) Participants found that the reporting framework had a comprehensive coverage of AI governance topics, was well-aligned with the Hiroshima AI Code of Conduct and recognized its potential to become an international mechanism for reporting and promoting consistent practices across governance frameworks for advanced AI systems. (2) Participants identified areas for improvement, including on streamlining the reporting framework and improving guidance, enabling different disclosure levels (e.g. full disclosure or aggregate results only), and ensuring alignment between the reporting framework and other international AI governance frameworks.

## Strengths of the draft reporting framework

Responses from the pilot identified the following strengths of the draft reporting framework:

- **Comprehensive coverage of AI governance topics:** One of the most frequently cited strengths of the reporting framework was its comprehensive coverage of AI governance topics, ensuring that all critical areas of AI development, risk management, and compliance are addressed.
- **Direct alignment with the Hiroshima Code of Conduct:** Another notable strength, mentioned by several organisations, is the framework's alignment with the Hiroshima Code of Conduct, as it reflects a set of internationally recognised actions to promote safe, secure and trustworthy AI. It provides organisations with clear expectations and a trusted foundation for compliance.
- **Potential to become an international mechanism for AI reporting, promoting alignment across AI governance frameworks:** Additionally, several respondents highlighted the framework's potential to become an international mechanism for AI reporting, particularly in its ability to promote alignment with other AI governance reporting frameworks.

## Areas for improvement of the draft reporting framework

Responses from the pilot identified the following opportunities for improvement, particularly in the structure and content of the reporting process.

- **Consolidate repetitive questions:** Fourteen organisations suggested reducing redundancy by consolidating repetitive questions to help streamline the reporting process, make it more efficient and lessen the time burden for reporting organisations.
- **Enhance formatting options:** Thirteen organisations asked for more flexible formatting options, including the ability to use bullet points or hyperlinks to improve the clarity of responses.
- **Clarify survey instruction:** Twelve organisations requested clearer instructions to guide them through the framework, specifically asking for the inclusion of word limits and examples to help ensure that responses be concise, focused, and consistent across organisations.

- **Explain key terms:** Relatedly, nine respondents requested explanations of key terms identified as ambiguous, such as "advanced AI systems," "unreasonable risks," and "significant incidents."
- **Improve alignment with other voluntary reporting mechanisms:** Six organisations suggested improving alignment with other voluntary reporting mechanisms, such as the Frontier AI Safety Commitments and the White House AI Voluntary Commitments, to make filling out the report easier.
- **Clarify the use and sharing of responses:** Six respondents expressed a need for further information regarding how the input they provide would be used and shared. Of these, four noted concerns about public disclosure of confidential or commercially sensitive information.

## Conclusion

The feedback obtained from the pilot phase will be considered by the G7 in their revision of the reporting framework. The G7 will publish a revised version of the framework and continue to engage with diverse stakeholders to support implementation and update the framework as needed over time.