United Nations ECE/CTCS/WP.6/2024/11

Economic and Social Council

Distr.: General
3 June 2024

Original: English

Economic Commission for Europe

Steering Committee on Trade Capacity and Standards

Working Party on Regulatory Cooperation and Standardization Policies (WP.6)

Thirty-fourth session
Geneva, 26(pm)–28 August 2024
Item 6 of the provisional agenda
Ad Hoc Team of Specialists on Standardization and Regulatory Techniques

Overarching common regulatory arrangement for the regulatory compliance of products and/or services with embedded artificial intelligence or other digital technologies

Submitted by the WP.6 Chair*

Summary:

International harmonization and interoperability of regulations of products with embedded artificial intelligence (AI) or other digital technologies is a challenge for regulators, but essential in order to achieve regulatory objectives, while avoiding unnecessary technical barriers to trade and multiplication of conformity testing.

Mandate:

The Working Party on Regulatory Cooperation and Standardization Policies (WP.6) Programme of work for 2024 foresees "to promote further horizontal guidance across its subgroups with respect to regulatory challenges related to digitalization. This includes topics like cybersecurity, privacy, artificial intelligence and data-based products." (ECE/CTCS/2023/14, paragraph 7).

Proposed decision:

"Member States took note of the *Overarching common regulatory arrangement for the regulatory compliance of products and/or services with embedded artificial intelligence or other digital technologies* (ECE/CTCS/WP.6/2024/11) and the *Declaration for technical regulation of products with embedded artificial intelligence* (ECE/CTCS/WP.6/2024/12). Member States encourage their relevant agencies and experts to study these two documents and fine tune them over the next six months with a view to adopting and launching the *Declaration* in summer 2025."

* This document is submitted under the responsibility of the WP.6 Chair. This document has also not been edited by a professional editor.

1. Products and/or services making use of embedded artificial intelligence (AI) or other digital technologies are used widely, yet without universally agreed or accepted definitions and under varying regulatory frameworks. Relevant standards are under development aiming at international harmonization. Functional safety and cybersecurity of such products are essential elements for economic competitiveness and product regulations. However, breakthroughs and innovations are moving considerably faster than standards and regulations for industry and wider sustainability.

2. Product regulation is often managed in a siloed approach, whereas embedded technologies are more of a horizontal nature. This may result, in some economies, in a disparity of regulations at the national level where requirements on AI or other embedded technologies might be interpreted or tested differently.

3. A common regulatory arrangement (CRA) as outlined in the Economic Commission for Europe (ECE) WP.6 *Recommendation L on the International Model for Transnational Regulatory Cooperation based on Good Regulatory Practice*[1] provides a voluntary framework for regulatory cooperation that facilitates market access through the use of best practice leading to greater harmonization and the establishment of sectoral and/or product-specific arrangements between interested United Nations Member States.

4. This CRA builds upon the paper developed within WP.6 for the November 2023 Annual Session, "*The regulatory compliance of products with embedded artificial intelligence or other digital technologies*" (ECE/CTCS/WP.6/2023/9),[2] and has taken into consideration the "*Interim Report: Governing AI for Humanity*" December 2023 developed by the United Nations Advisory Board on Artificial Intelligence.[3] In this perspective, it supports the use of AI as a potential means to achieve all 17 Sustainable Development Goals and sustainable development in its three dimensions – environmental, economic and social.

5. Due to the technologies' continuous evolution, it will likely be necessary to periodically review this CRA and update as necessary. References to this document are encouraged to clearly indicate the version number.

I. Scope

6. This CRA provides an overall approach to products and/or services with embedded AI systems or other digital technologies, as a basis for:

- Setting legitimate regulatory objectives
- Identifying and assessing risks
- Identifying relevant international standards for the development of regulations
- Establishing mutually recognizable conformity assessment procedures
- Establishing market surveillance and other enforcement mechanisms

7. It can be used at a national level to promote convergence of national technical regulations among agencies currently in place or yet to be put in place.

8. This CRA provides an overall approach to products and/or services with embedded AI or other digital technologies. It is proposed that use cases be developed to demonstrate how the CRA could be used in different sectors.

A. Terms and Definitions

9. Where not otherwise indicated in the text, the terminology used is based on definitions of the World Trade Organization's (WTO) Agreement on Technical Barriers to Trade (TBT)

[1] See: https://unece.org/DAM/trade/wp6/Recommendations/Recommendation_L_en.pdf
[2] See: https://unece.org/sites/default/files/2023-10/ECE_CTCS_WP6_2023_09_E.pdf
[3] See: https://www.un.org/en/ai-advisory-body

annex 1,[4] or WP.6 recommendations. Terminology related to products and/or services covered by this CRA include:

- Artificial intelligence (AI) system: engineered system that generates outputs such as content, predictions, recommendations or decisions for a given set of human-defined objectives (as for example, covered by the International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) 22989 on AI concepts and terminology).[5]
- Generative artificial intelligence (GenAI): an AI system which can produce a variety of data such as images, videos, 3-D models or audio files. These systems may be embedded into a device or be made available as a Software as a Service (SaaS).
- Product and/or services with embedded AI or other digital technologies: a product and/or service with an embedded, upgradeable (remotely, offline or by other means) AI system or with integrated, upgradeable software or with a combination of both, that operates with varying levels of autonomy and directs its operation and can make decisions influencing physical or virtual environments in a way that is generally intended to further human-defined objectives.

10. This CRA does not intend to cover autonomous wheeled vehicles[6] or autonomous weapons.[7] Both of these are directly covered by other national or international arrangements. However, the guidance in this document may provide useful for these product categories as well.

II. Product and/or services requirements

A. Regulatory objectives and assessing the level of risk

11. Setting regulatory objectives should be based on the acknowledgement that zero risk is not achievable. Determining the tolerable level of risk and risk appetite should be performed as described in ECE WP.6 *Recommendation R on Managing Risk in Regulatory Frameworks*.[8]

12. Certain products with embedded AI or other digital technologies may have an intrinsically high risk, for example, where there is a potential to have direct negative impact on health and safety or fundamental rights of people. Limited risk (or medium risk) is that which could impact safety issues. Low risk AI is that which does not use personal data and/or influence human beings. Governments will need to assess risks and choose the appropriate conformity assessment methods accordingly.

13. In situations, where the risk of an error associated with an AI system is high, human decision making shall be included wherever possible. AI systems should not be able to override human control.

14. Certain sectors have an intrinsically higher severity, where the stakes for life and health are particularly high. Certain medical equipment in hospitals, for example, have sophisticated diagnostic systems using embedded AI. Even if the medical equipment could generate algorithmic decision-making, given liability issues and the potential risk for patients, it is suggested that human decision-making be included wherever possible. Some

[4] See: https://www.wto.org/english/docs_e/legal_e/17-tbt.pdf

[5] See: https://www.iso.org/standard/74296.html

[6] Self-driving cars and autonomous wheeled vehicles are covered under separate United Nations committees. In contrast to autonomous wheeled vehicles, aerial and submarine vehicles as well as robots are within the scope of this CRA. See: https://unece.org/wp29-introduction

[7] The deployment of autonomous weapons and defence products and/or services with embedded AI system and other digital technologies falls within national defence and national security strategies and hence are out of scope of this CRA.

[8] See: https://unece.org/fileadmin/DAM/trade/wp6/Recommendations/Recommendation_R_en.pdf

industrial machinery where humans work alongside programmed robots piloted by AI could also merit obligatory human oversight and intervention.

15. Technology embedded within products is often difficult to assess or to know the actual content (e.g., the method, logic upon which outcomes are reached is unknown). This is true of products with embedded AI or other digital technologies, perhaps even more so as the system itself may react in an unexpected way because of the information it is learning from other sources. This unknown parameter of technology is usually balanced by a series of robust testing of the system within various parameters; this may not be sufficient to discover every unknown of the system, so a certain level of residual risk will remain. Regulators and distributors of such products need to ensure that these residual risks are tolerable and disclosed. For example, further guidance has been developed in the National Institute of Standards and Technology (NIST) - Artificial Intelligence Risk Management Framework (AI RMF1.0).[9]

B. Regulatory objectives and societal impact

16. The embedded system has been conceived in a way to mitigate bias in the AI system itself and within the AI-aided decision making. This includes human cognitive bias, data bias and bias introduced by engineered decisions. These are outlined in the *ISO/IEC TR 24027 on Bias in AI systems and AI aided decision making*.[10]

17. The embedded system has been conceived in a way that will not result in the loss of individual freedom, responsibility or of human autonomy.

18. The embedded system will not negatively impact individual mental wellbeing or wider societal impacts. This includes safeguarding children's vulnerabilities and children's rights on education, home, media and gamification/play. These are outlined in the United Nations *Convention on the Rights of the Child (UNCRC)*,[11] further elaborated in the *Convention's General Comment No 25* (2021) *on children's rights in relation to the digital environment*.[12]

19. The embedded system will not further widen the digital divide as outlined in the United Nations resolution on "*Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*" (A/78/L.49).[13] Products with embedded AI or other digital technologies should be fully functional in emerging economies. They should further aim to not create barriers for emerging economies to trade in these products or to enter into such markets.

C. Regulatory objectives and digital considerations

20. The embedded system should be designed in such a way as to ensure trustworthiness. This would include safeguards against AI-specific security threats, AI-specific privacy threats, unpredictability, opaqueness and challenges in the implementation and use of AI systems. These are outlined in the *ISO/IEC TR24028 on Overview of trustworthiness in artificial intelligence*.[14] Specific concerns on loss or unauthorized access to data is developed for example in the *European Union (EU) General Data Protection Regulation* (GDPR),[15] the *EU AI Act*[16] and other data regulations.

21. The embedded systems should foresee a robust system to protect against cyber-attacks. Data drift, concept drift, reward hacking algorithms and safe exploration should be

[9] See: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf
[10] See: https://www.iso.org/standard/77607.html
[11] See: https://www.unicef.org.uk/what-we-do/un-convention-child-rights/
[12] See: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation
[13] See: http://www.undocs.org/A/78/L.49
[14] See: https://www.iso.org/standard/77608.html
[15] See: https://eur-lex.europa.eu/eli/reg/2016/679/oj
[16] See: https://artificialintelligenceact.eu/

safeguarded within these systems. This is outlined in *ISO/IEC TR5469 on functional safety and AI systems*.[17] For instance, see also *the NIST Adversarial Machine Learning, Taxonomy and Terminology of Attacks and Mitigations (NIST AI 100-2e2023)*[18] and the *NIST Cybersecurity Framework (CSF) 2.0 (NIST CSWP.29)*.[19]

22. Likewise, measures should be taken to ensure that the embedded technology cannot be used for illicit activities (such as unauthorized or illegal control or monitoring, slander, or libel).

III. Reference to international standards

23. Beyond the standards listed in the previous section, there are a number of standards which can assist in designing and regulating products or services with embedded AI or other digital technologies.

24. As highlighted in WP.6 *Recommendation D on Reference to Standards* and further outlined in the WTO *TBT Agreement*, article 2.4: "Where technical regulations are required and relevant international standards exist or their completion is imminent, [WTO] Members shall use them, or the relevant parts of them, as a basis for their technical regulations except when such international standards or relevant parts would be an ineffective or inappropriate means for the fulfilment of the legitimate objectives pursued, for instance because of fundamental climatic or geographical factors or fundamental technological problems."

25. The following international standards may be applied in relation to products and/or services with embedded AI or other digital technologies.

- *ISO/IEC 42001 series of standards on AI management systems*[20]
- *ISO/IEC 23894:2023 series of standards on AI – Guidance on Risk Management*[21]
- *ISO/IEC TR 22100-5 series of standards on the implications of AI machine learning*[22]
- *IEC 62443 series of standards on industrial automation and control systems (IACS)*[23]
- *IEEE 7001-2021 series of standards for transparency of autonomous systems*[24]
- *Organisation for Economic Co-operation and Development (OECD) Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)*[25]
- *United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation on the Ethics of Artificial Intelligence (SHS/BIO/PI/2021/1)*[26]
- *World Health Organization (WHO) Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*[27]

26. Standards that address specific consumer protection and inclusion also need to be taken into consideration.

[17] See: https://www.iso.org/standard/81283.html
[18] See: https://csrc.nist.gov/pubs/ai/100/2/e2023/final
[19] See: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf
[20] See: https://www.iso.org/standard/81230.html
[21] See: https://www.iso.org/standard/77304.html
[22] See: https://www.iso.org/standard/80778.html
[23] See: https://www.iec.ch/blog/understanding-iec-62443
[24] See: https://ieeexplore.ieee.org/document/9726144
[25] See: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjgxvXX3ICFAxVyVKQEHXRyBlUQFnoECBMQAQ&url=https%3A%2F%2Flegalinstruments.oecd.org%2Fapi%2Fprint%3Fids%3D648%26lang%3Den&usg=AOvVaw3bU62HpvCxeAcd6gxRGeJ6&opi=89978449
[26] See: https://unesdoc.unesco.org/ark:/48223/pf0000381137
[27] See: https://www.who.int/publications/i/item/9789240084759

IV. Conformity assessment

27. Current regulatory practices for conformity assessment follow sector-specific mandates while compliance of products and/or services with embedded AI or other digital technologies is of a horizonal nature, requiring new expertise. Horizontal regulatory collaboration as a multi-disciplinary approach is needed to both identify and address risks, vulnerabilities and cyberthreats and increase operational resilience. This necessitates the development of horizontal regulatory capabilities beyond sectoral mandates, that are cutting across mandated procedural silos and are both supportive of the dynamic nature of digital innovation and conducive of enforcement strategies demanded by a digital market.

28. Beyond checking the conformity of the product which is under the supervision of a specific agencies, a product with embedded AI or other digital technology will need to be verified against the AI-specific technical aspects of the product. As noted above, the risk is that each national agency may take a different approach to AI in relation to the products it traditionally oversees. This may result in multiple compliance rules on AI within a single economy. It is suggested to avoid such an approach.

29. Stringency of conformity assessment procedures should be proportionate to risks of non-compliance of products.

30. Functionality of an AI system cannot be described in a standard in such a way that compliance with standards will be sufficient to consider a compliant product safe. Even if it were the case, an economic operator or a conformity assessment body cannot look inside these systems to test them. Also, their functioning can be random, they can behave differently in similar conditions.

31. Regulatory frameworks for AI systems should establish requirements for AI system provider/other stakeholders to mitigate risks of a system during its development and require the residual risk of an AI system to be tolerable.

32. While demonstrating that an AI system has been developed in conditions that are supposed to mitigate the risks and can be achieved by the means of usual conformity assessment procedures, acceptability of residual risk requires a framework for testing and assessment of conformity of AI systems.

33. Such a framework should include at least the following processes:[28]

- Identifying all possible hazards and risk events that could materialize during the functioning of an AI system and cause harm
- Building a list of situations/scenarios that a system can face
- Identifying which hazards can occur in each scenario
- Evaluating potential severity of hazards in scenarios and their frequencies
- Selecting scenarios for testing based on the level of risk: ensuring coverage of the most probable and most dangerous scenarios
- Performing simulation/test and evaluating the residual risk

34. As with the examples above under product regulation, there could be considered three levels of risk: high, limited (or medium) and low. For those products of low risk, governments might consider no specific conformity assessment process, or at most a supplier’s declaration of conformity. For those of limited risk, governments might consider a supplier’s declaration of conformity. And for those of high risk, governments might consider independent third-party conformity assessment.

[28] See: "Key to ensuring continuous compliance: assessing the residual risks of AI systems/products with embedded software", Valentin Nikonov, 23-24 November 2023 WP.6 conference "How to target continuous compliance"

A. Supplier's declaration of conformity

35. A supplier's declaration of conformity may be considered for products or services with embedded AI or other digital technologies of limited risk, and eventually also for those of low risk.

36. Such a declaration should outline that the supplier recognizes the importance of the principles in the product requirements section and that the product or service complies with the relevant international standards. For example, the supplier declaration could usefully reference the *ISO/IEC TR5469 on functional safety and AI systems*[29] and *the Organisation for Economic Co-operation and Development (OECD) Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)*.[30]

B. Third-party conformity assessment

37. A third-party conformity assessment would likely be preferred for products or services with embedded AI or other digital technologies that are considered of high risk. For the AI-related technical aspects of the product or service, such an assessment should outline conformity with the principles in the product requirements section and that the product or service complies with the relevant international standards. The referenced standards may vary depending on the type of product. For medical equipment, it could be the *World Health Organization (WHO) Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*,[31] for example.

38. By adopting this CRA, government agencies would signify their acceptance of third-party conformity assessment on the AI-related aspects of the product or service based on the principle of this CRA. This could be within or beyond pre-existing mutual recognition agreements. In both cases, there would need to be reference to the principles within this CRA for the AI-related aspects.

V. Market surveillance

39. One of the major defining aspects and key regulatory challenges of products or services with embedded AI or other digital technologies is that they may be linked to a remote server that will provide regular updates. The challenge is therefore how to ensure continuous compliance of these products once they have been put onto the market, which may not allow physical follow-up, inspection or verification of changes in product/service properties.

40. Market surveillance authorities will need to integrate methods for continuous compliance into their workflows. This will include regular mandatory independent audits to assure compliance to binary (compliant/non-compliant) government-approved criteria of products or services already on the market and to test their conformity to the principles and standards initially required for entry onto the market based upon the principles within this CRA). These audits will be of particular importance for products or services that had been identified as high risk.

41. Products or services that no longer comply to the principles and standards required of such products to enter the market should be promptly called back and taken off the market. In case of critical non-conformity, an international alert should be put in place to inform other economies.

---

[29] See: https://www.iso.org/standard/81283.html

[30] See: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjgxvXX3ICFAxVyVKQEHXRyBlUQFnoECBMQAQ&url=https%3A%2F%2Flegalinstruments.oecd.org%2Fapi%2Fprint%3Fids%3D648%26lang%3Den&usg=AOvVaw3bU62HpvCxeAcd6gxRGeJ6&opi=89978449

[31] See: https://www.who.int/publications/i/item/9789240084759