



SLOVENSKI STANDARD oSIST prEN 18229-3:2026

01-september-2026

Okvir zaupanja v umetno inteligenco - 3. del: Človeški nadzor

AI trustworthiness framework - Part 3: Human oversight

Framework für Vertrauenswürdigkeit von Künstlicher Intelligenz - Teil 3: Transparenz und menschliche Aufsicht

Sample Document

Ta slovenski standard je istoveten z: prEN 18229-3

get full document from standards.iteh.ai

ICS:

01.040.35	Informacijska tehnologija. (Slovarji)	Information technology (Vocabularies)
35.020	Informacijska tehnika in tehnologija na splošno	Information technology (IT) in general

oSIST prEN 18229-3:2026

en,fr,de

Sample Document

get full document from standards.iteh.ai

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

DRAFT
prEN 18229-3

July 2026

ICS 35.240.01

English version

AI trustworthiness framework - Part 3: Human oversight

Framework für Vertrauenswürdigkeit von Künstlicher
Intelligenz - Teil 3: Transparenz und menschliche
Aufsicht

This draft European Standard is submitted to CEN members for enquiry. It has been drawn up by the Technical Committee CEN/CLC/JTC 21.

If this draft becomes a European Standard, CEN and CENELEC members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

This draft European Standard was established by CEN and CENELEC in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN and CENELEC member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

CEN and CENELEC members are the national standards bodies and national electrotechnical committees of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Republic of North Macedonia, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Türkiye and United Kingdom.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Warning : This document is not a European Standard. It is distributed for review and comments. It is subject to change without notice and shall not be referred to as a European Standard.



CEN-CENELEC Management Centre:
Rue de la Science 23, B-1040 Brussels

Contents	Page
European foreword	4
Introduction	5
1 Scope.....	7
2 Normative references.....	7
3 Terms and definitions.....	7
3.1 Terms relating to the EU AI Act.....	7
3.2 Terms relating to AI systems.....	8
3.3 Terms related to risk management and quality management.....	10
3.4 Terms related to human oversight.....	11
4 Abbreviations.....	13
5 Human oversight.....	13
5.1 Objectives	13
5.2 Risk-management integration	13
5.2.1 Identification of risk scenarios	13
5.2.2 Reaction timeframe.....	14
5.2.3 Selection and documentation of measures.....	14
5.2.4 Technical infeasibility of real-time human intervention	14
5.2.5 Overload and conflict	15
5.2.6 Review of oversight assignments.....	15
5.3 Human oversight measure categories.....	15
5.3.1 General.....	15
5.3.2 Retrospective review measures	15
5.3.3 Alert-triggered intervention measures	17
5.3.4 Continuous monitoring measures	18
5.4 Specification of deployer-implemented oversight measures.....	20
5.4.1 Conditions for deployer-implemented measures and documentation.....	20
5.4.2 Limitations on residual risk.....	21
5.4.3 Specifications for deployer-implemented notifications	21
5.5 Automation bias and human factors.....	21
5.5.1 Awareness of automation bias	21
5.5.2 Testing for automation bias.....	21
5.6 Interpretation of outputs and interpretation tools.....	22
5.6.1 Output interpretation support for designated persons.....	22
5.6.2 Training material for designated persons.....	22
5.7 Functional requirements for designated person intervention functions	22
5.7.1 General.....	22
5.7.2 Disregard	23
5.7.3 Override	23
5.7.4 Reverse	23
5.7.5 Safe state transition	23
5.8 RBI systems - human oversight requirements.....	24
5.8.1 Technical prevention of unreviewed identification outputs	24
5.8.2 Review process and interface requirements	25
5.8.3 Manual reviewer independence mechanisms.....	26

5.8.4	Documentation requirements	26
5.8.5	Manual reviewer competency requirements	26
6	Documentation	27
6.1	Instructions for use.....	27
6.2	Technical documentation	27
Annex ZA (informative)	Relationship between this European Standard and the essential requirements of Regulation (EU) 2024/1689 aimed to be covered	28
Bibliography		29

Sample Document

get full document from standards.iteh.ai

prEN 18229-3:2026 (E)**European foreword**

This document (prEN 18229-3:2026) has been prepared by Technical Committee JTC 21 “Artificial Intelligence”, the secretariat of which is held by DS.

This document is currently submitted to the CEN Enquiry.

This document has been prepared under a Standardization request addressed to CEN-CENELEC by the European Commission. The Standing Committee of the EFTA States subsequently approves these requests for its Member States.

For relationship with EU Legislation, see informative Annex ZA, which is an integral part of this document.

Sample Document

get full document from standards.iteh.ai

Introduction

0.1 General

This document specifies requirements for human oversight of AI systems, in support of compliance with Regulation (EU) 2024/1689 (the EU AI Act). Annex ZA shows how the document supports Article 14 of the Regulation (EU) 2024/1689.

This document is part of the AI trustworthiness framework series: prEN 18229-1:— addresses record-keeping (Article 12), prEN 18229-2:— transparency and the provision of information to deployers (Article 13), prEN 18229-4:— accuracy and prEN 18229-5:— robustness (Article 15 except cybersecurity).

The risks that human oversight addresses are identified, estimated and evaluated through the risk management process specified in prEN 18228:— for each risk scenario identified through the risk management process, the provider determines a reaction timeframe, and that timeframe drives the selection of the human oversight measure categories and of the intervention functions made available to designated persons. The requirements of this document address the provider of the AI system; obligations that depend on the deployment context are conveyed to the deployer through the instructions for use specified in prEN 18229-2:—. The information addressed in this document is the subset necessary to enable and support human oversight; the general transparency information provided to deployers is specified in prEN 18229-2:— in support of the Regulation (EU) 2024/1689. Record keeping supporting oversight relies on prEN 18229-1:—. The human oversight measure categories specified in 5.3 address the information conditions of oversight (monitoring and the interpretation of outputs); the intervention functions specified in 5.7 address the control conditions (disregard, override, reverse and the transition to the safe state).

The evidence generated in designing and verifying those measures is recorded in the technical documentation of this document (Clause 6), while the risk management file (prEN 18228:—, 4.6) remains the home of the risk determinations on which those measures are calibrated. The design, verification and modification of the human oversight measures, and the monitoring of their effectiveness after the AI system is placed on the market, are governed by the provider's quality management system in accordance with EN 18286:2026 (Article 17); the risk management process of prEN 18228:— (Article 9) calibrates the measures through the reaction timeframe; this document specifies the measures themselves.

The allocation of responsibilities in this document is intended to reflect the distinction between provider and deployer obligations set out in Regulation (EU) 2024/1689, under Article 14. The requirements upon the provider in this document relate to the design and development of the AI system, including the design of human oversight measures and the provision of information to deployers. Where provisions in this document refer to outcomes that depend on deployer implementation, such provisions can be understood as requiring the provider to design and provide for those outcomes within the scope of the provider's control at the time of placing on the market, and based on the assumptions documented for the intended purpose and use.

Subclause 5.1 establishes the objectives of human oversight and identifies the parties to whom they are addressed.

Subclause 5.2 specifies the procedure by which the provider integrates human oversight into the risk management process: the identification of risk scenarios, the determination of the reaction timeframe, the selection and documentation of measures, the treatment of technical infeasibility, and the review of oversight assignments.

Subclause 5.3 specifies the requirements for each human oversight measure category: retrospective review measures (5.3.2), alert-triggered intervention measures (5.3.3) and continuous monitoring measures (5.3.4).

prEN 18229-3:2026 (E)

Subclause 5.4 specifies the deployer-implemented human oversight measures and the information on those measures to be provided in the instructions for use in accordance with prEN 18229-2:—, 6.1.

Subclause 5.5 specifies the requirements and recommendations addressing automation bias and human factors.

Subclause 5.6 specifies the requirements enabling designated persons to interpret the outputs of the AI system, in connection with the transparency mechanisms selected in accordance with prEN 18229-2:—, 5.2.

Subclause 5.7 specifies the functional requirements for the designated person intervention functions, including the stop procedure implemented through the safe state.

Subclause 5.8 specifies the human oversight requirements for remote biometric identification systems.

Clause 6 specifies the recording of evidence produced under Clause 5 in the instructions for use and the technical documentation.

0.2 Verbal forms

In this document, the following verbal forms are used:

- “shall” indicates a requirement;
- “should” indicates a recommendation;
- “may” indicates a permission;
- “can” indicates a possibility or a capability.

Sample Document

get full document from standards.iteh.ai

1 Scope

This document provides terminology, concepts, requirements, and guidance for human oversight of AI systems.

It is primarily intended for organizations placing on the market or putting into service AI systems and is not specific to any particular sector.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

prEN 18228:—,¹ *AI Risk Management*

prEN 18229-1:—,² *AI trustworthiness framework – Part 1: Logging*

prEN 18229-2:—,³ *AI trustworthiness framework – Part 2: Transparency*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp/>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1 Terms relating to the EU AI Act

3.1.1

provider

natural or legal person, public authority, agency or other body that develops an *AI system* (3.2.1) or a general-purpose AI model or that has an *AI system* (3.2.1) or a general-purpose AI model developed and places it on the market or puts the *AI system* (3.2.1) into service under its own name or trademark, whether for payment or free of charge

[SOURCE: EU AI Act 2024/1689, (Article 3(3)), modified – removed “a”]

3.1.2

deployer

natural or legal person, public authority, agency or other body using an *AI system* (3.2.1) under its authority except where the *AI system* (3.2.1) is used in the course of a personal non-professional activity

[SOURCE: EU AI Act 2024/1689, (Article 3(4)), modified – removed “a”]

¹ Under preparation. Stage at time of publication, Enquiry

² Under preparation. Stage at time of publication, Enquiry

³ Under preparation. Stage at time of publication, Working Draft