

SLOVENSKI STANDARD **oSIST prEN ISO/IEC 12792:2024**

01-julij-2024

Informacijska tehnologija - Umetna inteligenca - Taksonomija preglednosti sistemov UI (ISO/IEC DIS 12792:2024)

Information technology - Artificial intelligence - Transparency taxonomy of AI systems (ISO/IEC DIS 12792:2024)

Informationstechnologie - Künstliche Intelligenz - Transparenz-Taxonomie von KI□Systemen (ISO/IEC DIS 12792:2024)

Technologies de l'information - Intelligence artificielle - Taxonomie pour la transparence des systèmes d'IA (ISO/IEC DIS 12792:2024)

Ta slovenski standard je istoveten z: prEN ISO/IEC 12792

ICS:

35.020 Informacijska tehnika in

Information technology (IT) in

tehnologija na splošno general

oSIST prEN ISO/IEC 12792:2024 en,fr,de oSIST prEN ISO/IEC 12792:2024

iTeh Standards (https://standards.iteh.ai) Document Preview

oSIST prEN ISO/IEC 12792:2024

https://standards.iteh.ai/catalog/standards/sist/9c97ebf2-845a-4c0f-9a14-487241728307/osist-pren-iso-iec-12792-2024



DRAFT International Standard

ISO/IEC DIS 12792

Information technology — Artificial intelligence — Transparency taxonomy of AI systems

ICS: 01.040.35; 35.020

ISO/IEC JTC 1/SC 42

Secretariat: ANSI

Voting begins on: **2024-04-23**

Voting terminates on: **2024-07-16**

iTeh Standards
(https://standards.iteh.ai)
Document Preview

oSIST prEN ISO/IEC 1279 2:2024

https://standards.iteh.ai/catalog/standards/sist/9c97ebf2-845a-4c0f-9a 4-487241728307/osist-pren-iso-iec-12792-2024

This document is circulated as received from the committee secretariat.

ISO/CEN PARALLEL PROCESSING

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENTS AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

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.

iTeh Standards (https://standards.iteh.ai) Document Preview

oSIST prEN ISO/IEC 12792:2024

https://standards.iteh.ai/catalog/standards/sist/9c97ebf2-845a-4c0f-9a14-487241728307/osist-pren-iso-iec-12792-2026



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office CP 401 • Ch. de Blandonnet 8 CH-1214 Vernier, Geneva Phone: +41 22 749 01 11 Email: copyright@iso.org Website: www.iso.org

Published in Switzerland

Contents			Page
Fore	eword		v
		on	
_			
1	-)e	
2	Norr	mative references	1
3	Tern	ns and definitions	1
4	Sym	bols and abbreviated terms	4
5	Over	4	
	5.1	rview	
	5.2	Organization and usage of the taxonomy	
	5.3	Concept of transparency	
6	Stak	7	
	6.1	seholders' needs and objectives General	7
	6.2	Selected stakeholder roles in transparency	
7	Cont	text-level taxonomy	10
	7.1	General	
	7.2	Societal context	10
		7.2.1 General	10
		7.2.2 Labour practices	
		7.2.3 Consumer needs	
	7.3	Environmental context	14
8	System-level taxonomy		
	8.1	General LLEN ST2M02110S	
	8.2	Basic information	17
	8.3	Organizational processes 8.3.1 General	18
		8.3.1 General 8.3.2 Governance	18 10
		8.3.3 Management system	
		8.3.4 Risk management	
		8.3.5 Quality management	
	8.4	Applicability Applicability	19
		8.4.1 General	
		8.4.2 Intended purposes	
		8.4.3 Capabilities	
		8.4.4 Functional limitations 8.4.5 Recommended uses	
		8.4.6 Precluded uses	
	8.5	Overview of technical characteristics	
	0.0	8.5.1 General	
		8.5.2 Expected inputs and outputs	20
		8.5.3 Production data	
		8.5.4 Logging and storing	
		8.5.5 System decomposition	
		8.5.6 Application programming interface	
		8.5.7 Human factors 8.5.8 Deployment methods	
		8.5.9 Configuration management	
	8.6	Access to internal elements	
	8.7	Quality and performance	
		8.7.1 General	
		8.7.2 Verification and validation processes	
		8.7.3 Runtime measurements	
		8.7.4 Comparison with alternative systems	24

9	Model-level taxonomy		25
	9.1	General	25
	9.2	Basic information	25
	9.3	Usage	
		9.3.1 Processing performed by the model	
		9.3.2 Dependence on other models	
		9.3.3 Coherence with AI system's intended purposes	26
	9.4	Technical characteristics	
		9.4.1 Type of technology used	
		9.4.2 Features extracted from input data	
		9.4.3 Algorithm used for processing	
		9.4.4 Procedure for building the model	
		9.4.5 Hyperparameters	
		9.4.6 Input and output formats	
		9.4.7 Compute hardware	
		9.4.8 Computational costs	
		9.4.9 Models in evolutive systems	
	9.5	Data used	
	9.6	Functional correctness	30
10	Data	set-level taxonomy	30
	10.1	General	30
	10.2	Basic information	30
	10.3	Data provenance	
	10.4	Data properties	
	10.5	Dataset domain and purposes	34
	10.6	Data biases and limitations	
	10.7	Societal considerations	34
	10.8	Data preparation performed	35
	10.9	Data preparation performed Dataset maintenance	36
Anne	x A (in	formative) Examples of AI system transparency	37
Anne	x B (in	formative) (Information disclosures for language data	38
Bibli	ograpł	Document Preview	39

SIST prEN ISO/IEC 12792:2024

https://standards.iteh.ai/catalog/standards/sist/9c97ebf2-845a-4c0f-9a14-487241728307/osist-pren-iso-iec-12792-202

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 42, *Artificial intelligence*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found atwww.iso.org/members.html.

Document Preview

oSIST prEN ISO/IEC 12792:2024

https://standards.iteh.ai/catalog/standards/sist/9c97ebf2-845a-4c0f-9a14-487241728307/osist-pren-iso-jec-12792-202

Introduction

The objectives of this document include:

- improving trustworthiness, accountability and communication among different AI stakeholders, including partners in a supply chain, customers, users, society and regulators, by establishing a consistent terminology around transparency of AI systems;
- providing AI stakeholders with information about different elements of transparency with their relevance and possible limitations to different use cases and target audience;
- serving as a basis for developing technology-specific, industry-specific or region-specific standards for transparency of AI systems.

Transparency for AI systems is the property of a system that stakeholders receive relevant information about the system. This can include information on items such as system features, limitations, data, system design and design choices (see ISO/IEC 22989:2022, 5.15.8).

Increased transparency provides information for relevant stakeholders to better understand how an AI system is developed and used. For example, this allows an AI customer (such as an AI user) of the AI system to determine if it is appropriate for their situation, and supports an AI auditor in assessing if the system complies with conformity requirements.

A standardized transparency taxonomy of AI systems helps people with different backgrounds to better understand each other by using the same terminology. This in turn supports an improved understanding of the AI systems, and provides a foundation for developing interoperable and coherent transparency related standards.

This document is structured as follows:

- <u>Clause 5</u> provides an overview of this document and describes the concept of transparency of AI systems;
- <u>Clause 6</u> discusses how transparency needs can vary depending on the AI system context and on the stakeholders involved;
- <u>Clause 7</u> discusses transparency items that describe the context of the AI system;
- <u>Clause 8</u> pertains to documenting how the AI system interacts with its environment;
- <u>Clause 9</u> focuses on documenting the internal functioning of the system;
- <u>Clause 10</u> offers guidance on the documentation of datasets as stand-alone items.

AI systems often affect society and the environment (see <u>Clause 7</u>). However, it is also true that society and the environment can affect the performance of an AI system. This can include various aspects, such as:

- environmental conditions that affect the functioning and longevity of an AI system;
- introduction or reinforcement of unwanted bias;
- organizational practices that can cause poor outcomes;
- effects such as concept and data drift;
- formation of unwanted feedback loops (e.g. negative reinforcement of discriminatory patterns), which can be especially problematic in the case of continuous learning.

While these are important items for consideration, a detailed analysis is not provided in this document.

Information technology — Artificial intelligence — Transparency taxonomy of AI systems

1 Scope

This document defines a taxonomy of information elements to assist AI stakeholders with identifying and addressing the needs for transparency of AI systems. The document describes the semantics of the information elements and their relevance to the various objectives of different stakeholders.

This document uses a horizontal approach and is applicable to any kind of organization and application involving an AI system.

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.

ISO/IEC 22989:2022, Information technology — Artificial intelligence — Artificial intelligence concepts and terminology

3 Terms and definitions

terms and definitions

The Terms and definitions clause is a mandatory element of the text.

For rules on the drafting of the Terms and definitions, refer to the ISO/IEC Directives, Part 2:2018, Clause 16.

To insert a new terminological entry, go to the Structure tab and click on Insert Term entry.

For the purposes of this document, the *following terms and definitions given in ISO/IEC 22989:2022, as well as the following* apply.

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

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

3.1

attestation

issue of a statement, based on a decision, that fulfilment of *specified requirements* (3.13) has been demonstrated

Note 1 to entry: The resulting statement is intended to convey the assurance that the specified requirements have been fulfilled. Such an assurance does not, of itself, provide contractual or other legal guarantees.

Note 2 to entry: First-party attestation and third-party attestation are distinguished by the terms declaration, certification and accreditation, but there is no corresponding term applicable to second-party attestation.

[SOURCE: ISO/IEC 17000:2020(en), 7.3; modified: removed "referred to in this document as a "statement of conformity" from Note 1.]

3.2

carbon footprint of a product

CFP

carbon footprint

sum of greenhouse gas emissions and greenhouse gas removals in a product system, expressed as carbon dioxide equivalents and based on a life cycle assessment using the single impact category of climate change

[SOURCE: ISO 14050:2020(en), 3.11.1; modified: added the admitted term "carbon footprint"]

3.3

confidential data

data to which only a limited number of persons have access and which are meant for restricted use

[SOURCE: ISO 5127:2017, 3.1.10.18; modified: removed Note 1 to entry]

3.4

conformity assessment body

body that performs conformity assessment activities, excluding accreditation

[SOURCE: ISO/IEC 17000:2020(en), 4.6]

3.5

data protection

implementation of administrative, technical, or physical measures to guard against the unauthorized access to data

[SOURCE: ISO/IEC 2382:2015(en), 2126371; modified: removed all Notes to entry.]

3.6

digital fingerprint

bit sequence generated from a digital document using an algorithm that uniquely identifies the original document

Note 1 to entry: Any digital document modification will produce a different fingerprint.

[SOURCE: ISO 14641:2018, 3.15]

3.7

evolutive system

AI system whose behaviour can change without an explicit action from its AI developer. st-pren-iso-jec-12792-2024

Note 1 to entry: Examples of evolutive systems include AI systems with a storage component, AI systems that leverage user-side actions and feedback, and AI systems involving continuous learning.

Note 2 to entry: Change of behaviour does not imply a change of goal, but can result in a better or worse fulfillment of the targeted goal.

3.8

feature

measurable property of an object or event with respect to a set of characteristics

Note 1 to entry: Features play a role in training and prediction.

Features provide a machine-readable way to describe the relevant objects. As the algorithm will not go back to the objects or events themselves, feature representations are designed to contain all useful information.

[SOURCE: ISO/IEC 23053:2022, 3.3.3; modified: removed <machine learning> domain tag]

3.9

life cycle impact assessment

LCIA

environmental impact assessment

phase of life cycle assessment aimed at understanding and evaluating the magnitude and significance of the potential environmental impacts for a product system throughout the life cycle of the product

[SOURCE: ISO 14050:2020, 3.6.5; modified: added admitted term "environmental impact assessment"]

3.10

object of conformity assessment

object

entity to which specified requirements (3.13) apply

EXAMPLE Product, process, service, system, installation, project, data, design, material, claim, person, body or organization, or any combination thereof.

Note 1 to entry: The term "body" is used in this definition to refer to conformity assessment bodies and accreditation bodies. The term "organization" is used in its general meaning and may include bodies according to the context. The more specific ISO/IEC Guide 2 definition of an organization as a body based on membership is not applicable to the field of conformity assessment.

[SOURCE: ISO/IEC 17000:2020(en), 4.2; modified "in this document" to "in this definition" in Note 1 to entry]

3.11

power usage effectiveness

PUE

ratio of the data centre total energy consumption to information technology equipment energy consumption, calculated, measured or assessed across the same period

Note 1 to entry: Sometimes the inverse value of PUE, referred to as Data Centre Infrastructure Efficiency (DCiE), is used.

[SOURCE: ISO/IEC 30134-2:2016(en), 3.1.3]

3.12

scope of attestation

range or characteristics of *objects of conformity assessment* (3.10) covered by *attestation* (3.1)

[SOURCE: ISO/IEC 17000:2020(en), 7.6] SIST prEN ISO/IEC 12792:2024

3.13

specified requirement

need or expectation that is stated

Note 1 to entry: Specified requirements can be stated in normative documents such as regulations, standards and technical specifications.

Note 2 to entry: Specified requirements can be detailed or general.

[SOURCE: ISO/IEC 17000:2020(en), 5.1]

3.14

taxonomy

systematic classification of items into generic groups based on factors possibly common to several of the items

[SOURCE: ISO/IEC 14224:2016(en), 3.92]

3.15

water footprint

metric(s) that quantifies the potential environmental impacts related to water

[SOURCE: ISO 14050:2020(en), 3.10.1]

4 Symbols and abbreviated terms

AI artificial intelligence

ML machine learning

API application programming interface

URL universal resource locator

REST representational state transfer

JSON javascript object notation

GUI graphical user interface

5 Overview

5.1 General

<u>Clause 5</u> provides an overview of this document and describes the concept of transparency of AI systems, in particular in relation to other trustworthiness characteristics.

5.2 Organization and usage of the taxonomy

This document describes a multi-faceted taxonomy of transparency for AI systems. <u>Clause 6</u> discusses how transparency needs can vary, depending on the AI system context and the role of different AI stakeholders. The remainder of this document covers multiple levels and viewpoints on the system. It is organized in four parts:

- <u>Clause 7</u> describes transparency taxonomy elements (referred to simply as "taxonomy elements") informing on the context of the AI system (e.g. transparency needs in the case of labour stakeholders).
- Clause 8 pertains to documenting how the AI system interacts with its environment (e.g. its recommended and precluded uses, or API documentation).
- <u>Clause 9</u> focuses on the internal functioning of the system, at the level of model components and algorithmic processes.
- <u>Clause 10</u> offers guidance on the documentation of datasets as stand-alone items, considering that the
 AI system's life cycle involves those datasets (thereby also worth transparency considerations), but also
 that datasets can be reused as is across several AI systems for different use cases.

For a given AI system, all four parts of the taxonomy (<u>Clauses 7</u> to <u>10</u>) can be considered, as they provide complementary information. An illustration of this document's structure can be found in <u>Figure 1</u>.