Technical Report

ISO/IEC TR 42106

First edition
2026-07

Information technology — Artificial intelligence (AI) — Overview of differentiated benchmarking of AI system quality characteristics

Technologies de l'information — Intelligence artificielle (IA) — Aperçu de l'analyse comparative différenciée des caractéristiques de qualité des systèmes d'IA

Reference number
ISO/IEC TR 42106:2026(en)

© ISO/IEC 2026



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

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

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**

4 Overview of relevant benchmarking methods ... **2**
4.1 Review of benchmarking definitions ... 2
4.2 Types of benchmarking ... 3
4.3 Metrics, measures and criteria ... 3

5 Benchmarking AI systems ... **4**
5.1 Benchmarking AI system quality ... 4
5.2 Context of use ... 5
5.3 Complex adaptive systems ... 5
5.4 Limitations in benchmarking AI systems ... 6

6 Frameworks for differentiated benchmarking ... **7**
6.1 AI management frameworks ... 7
6.2 Classification-based frameworks ... 7
6.3 Levels of specification ... 9

7 Feasibility analysis ... **11**
7.1 Case Example 1: On-job training recommendation system ... 11
7.2 Case Example 2: User intent recognition ... 12
7.3 Case Example 3: Generation of clinical pathways ... 13

Annex A (informative) **Sample levels of specification** ... **15**

Annex B (informative) **Descriptions of measures** ... **16**

Bibliography ... **18**

© ISO/IEC 2026 – All rights reserved

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

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 document 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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint 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 at www.iso.org/members.html and www.iec.ch/national-committees.

© ISO/IEC 2026 – All rights reserved

Introduction

Artificial intelligence (AI) systems are diverse in nature and heterogeneous in terms of their potential effect on consumers and third-parties. For example, a company's use of disposable income estimates using a linear regression model to serve ads targeted at different socio-economic profiles reflects a system with low complexity and potential impact. Conversely, a bank's use of disposable income estimates using a large neural network model to make housing loan decisions reflects a system with both high complexity and high impact.

Benchmarking is often used to compare quality characteristics of software systems against reference or target values. Given the diverse nature and heterogeneous impact of AI systems, such reference or target values can differ widely across systems and across deployments of similar systems across contexts of use. To accommodate this diversity and heterogeneity, differentiated benchmarking can target different quality characteristics at different reference values. This document reviews AI management frameworks for their ability to offer guidance for differentiated benchmarking of AI system quality characteristics.

By evaluating frameworks for specifying differing levels of benchmarks and benchmarking, commensurate with expected social impact of AI systems, this document identifies gaps in current frameworks that, when filled, can yield guidance for differentiated benchmarking of AI systems, which can help rationalize the standardization implementation effort of AI providers, while maintaining system trustworthiness for AI customers and AI partners.

© ISO/IEC 2026 – All rights reserved

Information technology — Artificial intelligence (AI) — Overview of differentiated benchmarking of AI system quality characteristics

1 Scope

This document provides an overview of conceptual frameworks for graded benchmarking of artificial intelligence (AI) system quality characteristics. The aim is to examine the feasibility of using differentiated benchmarking of quality characteristics based on the complexity and context of use of 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, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*

ISO/IEC 23053, *Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)*

ISO/IEC 25059, *Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for AI systems*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989, ISO/IEC 23053, ISO/IEC 25059 and the following 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
benchmark
reference point against which comparisons can be made

Note 1 to entry: For AI system benchmarking, an AI system quality characteristic (ISO/IEC 25059 [1]) is the object of comparison.

[SOURCE: ISO/IEC 29155-1:2017 [2], 3.2, modified —Note 1 to entry has been replaced.]

3.2
benchmarking
activity of comparing objects of interest to each other or against a *benchmark* (3.1) to evaluate (a) characteristic(s)

Note 1 to entry: For AI system benchmarking, the object of interest is an AI system quality characteristic (ISO/IEC 25059 [1]).

[SOURCE: ISO/IEC 29155-1:2017 [2], 3.3, modified — Note 1 to entry has been replaced.]

© ISO/IEC 2026 – All rights reserved