INTERNATIONAL STANDARD

ISO/IEC 23894

First edition
2023-02

Information technology — Artificial intelligence — Guidance on risk management

Technologies de l'information — Intelligence artificielle — Recommandations relatives au management du risque

Reference number
ISO/IEC 23894:2023(E)

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 Principles of AI risk management ... **1**
5 Framework ... **5**
5.1 General ... 5
5.2 Leadership and commitment ... 5
5.3 Integration ... 6
5.4 Design ... 6
5.4.1 Understanding the organization and its context ... 6
5.4.2 Articulating risk management commitment ... 8
5.4.3 Assigning organizational roles, authorities, responsibilities and accountabilities ... 8
5.4.4 Allocating resources ... 8
5.4.5 Establishing communication and consultation ... 8
5.5 Implementation ... 9
5.6 Evaluation ... 9
5.7 Improvement ... 9
5.7.1 Adapting ... 9
5.7.2 Continually improving ... 9
6 Risk management process ... **9**
6.1 General ... 9
6.2 Communication and consultation ... 9
6.3 Scope, context and criteria ... 9
6.3.1 General ... 9
6.3.2 Defining the scope ... 10
6.3.3 External and internal context ... 10
6.3.4 Defining risk criteria ... 10
6.4 Risk assessment ... 11
6.4.1 General ... 11
6.4.2 Risk identification ... 11
6.4.3 Risk analysis ... 14
6.4.4 Risk evaluation ... 15
6.5 Risk treatment ... 15
6.5.1 General ... 15
6.5.2 Selection of risk treatment options ... 15
6.5.3 Preparing and implementing risk treatment plans ... 16
6.6 Monitoring and review ... 16
6.7 Recording and reporting ... 16
Annex A (informative) **Objectives** ... **18**
Annex B (informative) **Risk sources** ... **21**
Annex C (informative) **Risk management and AI system life cycle** ... **24**
Bibliography ... **26**

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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC 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) or the IEC list of patent declarations received (see https://patents.iec.ch).

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.

Introduction

The purpose of risk management is the creation and protection of value. It improves performance, encourages innovation and supports the achievement of objectives.

This document is intended to be used in connection with ISO 31000:2018. Whenever this document extends the guidance given in ISO 31000:2018, an appropriate reference to the clauses of ISO 31000:2018 is made followed by AI-specific guidance, if applicable. To make the relationship between this document and ISO 31000:2018 more explicit, the clause structure of ISO 31000:2018 is mirrored in this document and amended by sub-clauses if needed.

This document is divided into three main parts:

Clause 4: Principles – This clause describes the underlying principles of risk management. The use of AI requires specific considerations with regard to some of these principles as described in ISO 31000:2018, Clause 4.

Clause 5: Framework – The purpose of the risk management framework is to assist the organization in integrating risk management into significant activities and functions. Aspects specific to the development, provisioning or offering, or use of AI systems are described in ISO 31000:2018, Clause 5.

Clause 6: Processes – Risk management processes involve the systematic application of policies, procedures and practices to the activities of communicating and consulting, establishing the context, and assessing, treating, monitoring, reviewing, recording and reporting risk. A specialization of such processes to AI is described in ISO 31000:2018, Clause 6.

Common AI-related objectives and risk sources are provided in Annex A and Annex B. Annex C provides an example mapping between the risk management processes and an AI system life cycle.

Information technology — Artificial intelligence — Guidance on risk management

1 Scope

This document provides guidance on how organizations that develop, produce, deploy or use products, systems and services that utilize artificial intelligence (AI) can manage risk specifically related to AI. The guidance also aims to assist organizations to integrate risk management into their AI-related activities and functions. It moreover describes processes for the effective implementation and integration of AI risk management.

The application of this guidance can be customized to any organization and its context.

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 31000:2018, *Risk management — Guidelines*

ISO Guide 73:2009, *Risk management — Vocabulary*

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

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 31000:2018, ISO/IEC 22989:2022 and ISO Guide 73:2009 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/

4 Principles of AI risk management

Risk management should address the needs of the organization using an integrated, structured and comprehensive approach. Guiding principles allow an organization to identify priorities and make decisions on how to manage the effects of uncertainty on its objectives. These principles apply to all organizational levels and objectives, whether strategic or operational.

Systems and processes usually deploy a combination of various technologies and functionalities in various environments, for specific use cases. Risk management should take into account the whole system, with all its technologies and functionalities, and its impact on the environment and stakeholders.

AI systems can introduce new or emergent risks for an organization, with positive or negative consequences on objectives, or changes in the likelihood of existing risks. They also can necessitate