INTERNATIONAL STANDARD

ISO/IEC 38507

First edition
2022-04

Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations

Technologies de l'Information — Gouvernance des technologies de l'information — Implications de gouvernance de l'utilisation par des organisations de l'intelligence artificielle

Reference number
ISO/IEC 38507:2022(E)

© ISO/IEC 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

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

© ISO/IEC 2022 – All rights reserved

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**
3.1 Terms related to AI ... 2
3.2 Terms related to governance ... 2

4 Governance implications of the organizational use of AI ... **2**
4.1 General ... 2
4.2 Maintaining governance when introducing AI ... 3
4.3 Maintaining accountability when introducing AI ... 4

5 Overview of AI and AI systems ... **6**
5.1 General ... 6
5.2 How AI systems differ from other information technologies ... 6
5.2.1 Decision automation ... 6
5.2.2 Data-driven problem-solving ... 7
5.2.3 Adaptive systems ... 7
5.3 AI ecosystem ... 8
5.4 Benefits of the use of AI ... 9
5.5 Constraints on the use of AI ... 10

6 Policies to address use of AI ... **11**
6.1 General ... 11
6.2 Governance oversight of AI ... 12
6.3 Governance of decision-making ... 13
6.4 Governance of data use ... 14
6.5 Culture and values ... 15
6.6 Compliance ... 16
6.6.1 Compliance obligations ... 16
6.6.2 Compliance management ... 17
6.7 Risk ... 17
6.7.1 Risk appetite and management ... 17
6.7.2 Risk management ... 18
6.7.3 Objectives ... 19
6.7.4 Sources of risk ... 20
6.7.5 Controls ... 21

Annex A (normative) **Governance and organizational decision-making** ... **23**

Bibliography ... **27**

© ISO/IEC 2022 – 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).

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 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 jointly by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittees SC40, *IT service management and IT governance* and 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 2022 – All rights reserved

Introduction

The objective of this document is to provide guidance for the governing body of an organization that is using, or is considering the use of, artificial intelligence (AI).

This document provides guidance on the role of a governing body with regard to the use of AI within their organization and encourages organizations to use appropriate standards to underpin their governance of the use of AI.

This document addresses the nature and mechanisms of AI to the extent necessary to understand the governance implications of their use: what are the additional opportunities, risks and responsibilities that the use of AI brings? The emphasis is on governance (which is done by humans) of the organization's use of AI and not on the technologies making up any AI system. However, such governance requires an understanding of the implications of the technologies.

Artificial intelligence (AI)

AI embraces a family of technologies that bring together computing power, scalability, networking, connected devices and interfaces, together with vast amounts of data. Reference to 'AI' in this document is intended to be understood to refer to a whole family of technologies and methods, and not to any specific technology, method or application. For AI concepts and terminology, see ISO/IEC 22989:—[1)].

Use of AI

"Use of AI" is defined in this document in the broadest sense as developing or applying an AI system through any part of its life cycle to fulfil objectives and create value for the organization. This includes relationships with any party providing or using such systems.

Governance implications of the use of AI

The scope of this document is concerned with the implications for an organization of the use of AI. As with any powerful tool, the use of AI brings new risks and responsibilities that should be addressed by organizations that use it. AI is not inherently 'good' or 'evil', 'fair' or 'biased', 'ethical' or 'unethical' although its use can be or can seem to be so.

The organization's purpose, ethics and other guidelines are reflected, either formally or informally, in its policies. This document examines both governance and organizational policies and their application and provides guidance to adapt these for the use of AI. The operational aspects of the policies are implemented through management. This document refers to other standards for details on related topics including social responsibility, trustworthiness (such as risk management, management of bias, and quality) and compliance management.

1) Under preparation. Stage at the time of publication: ISO/IEC FDIS 22989:2022.

© ISO/IEC 2022 – All rights reserved

Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations

1 Scope

This document provides guidance for members of the governing body of an organization to enable and govern the use of Artificial Intelligence (AI), in order to ensure its effective, efficient and acceptable use within the organization.

This document also provides guidance to a wider community, including:

— executive managers;

— external businesses or technical specialists, such as legal or accounting specialists, retail or industrial associations, or professional bodies;

— public authorities and policymakers;

— internal and external service providers (including consultants);

— assessors and auditors.

This document is applicable to the governance of current and future uses of AI as well as the implications of such use for the organization itself.

This document is applicable to any organization, including public and private companies, government entities and not-for-profit organizations. This document is applicable to an organization of any size irrespective of their dependence on data or information technologies.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitute 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:—[2], *Information technology — Artificial intelligence —Artificial intelligence concepts and terminology*

ISO/IEC 38500:2015, *Information technology — Governance of IT for the organization*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22989, ISO/IEC 38500 and the following apply.

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

— IEC Electropedia: available at https://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

2) Under preparation. Stage at the time of publication: ISO/IEC FDIS 22989:2022.

© ISO/IEC 2022 – All rights reserved

3.1 Terms related to AI

3.1.1
use of AI
developing or applying an AI system through any part of its life cycle to fulfil an organization's objectives

Note 1 to entry: This term is scoped to any action or activity related to AI that can have governance implications.

3.2 Terms related to governance

3.2.1
oversight
monitoring of the implementation of organizational and governance policies and management of associated tasks, services and products set by the organization, in order to adapt to changes in internal or external circumstances

Note 1 to entry: Effective oversight needs general understanding of a situation. Oversight is one of the 'principles of governance' covered in depth in ISO 37000:2021, 6.4.

3.2.2
risk
effect of uncertainty on objectives

Note 1 to entry: An effect is a deviation from the expected. It can be positive, negative or both, and can address, create or result in opportunities and threats.

Note 2 to entry: Objectives can have different aspects and categories and can be applied at different levels.

Note 3 to entry: Risk is usually expressed in terms of risk sources, potential events, their consequences and their likelihood.

[SOURCE: ISO 31000:2018, 3.1]

3.2.3
risk appetite
amount and type of *risk* (3.2.2) that an organization is willing to pursue or retain

[SOURCE: ISO Guide 73:2009, 3.7.1.2]

3.2.4
compliance obligations
requirements that an organization mandatorily has to comply with as well as those that an organization voluntarily chooses to comply with

[SOURCE: ISO 37301:2021, 3.25]

3.2.5
compliance
meeting all the organization's *compliance obligations* (3.2.4)

[SOURCE: ISO 37301:2021, 3.26]

4 Governance implications of the organizational use of AI

4.1 General

The governance of organizations is enabled by the application of principles that help the organization fulfil its organizational purpose and, in doing so, generate value for the organization and its stakeholders. According to ISO 31000:2018, 5.3 governance guides the course of the organization, its external and internal relationships, and the rules, processes and practices needed to achieve its

© ISO/IEC 2022 – All rights reserved

purpose. Management structures translate governance direction into the strategy and associated objectives required to achieve desired levels of sustainable performance and long-term viability.

An overview of the concepts of governance and organizational decision-making (and in particular, references to existing standards) that shall be followed, is given in Annex A.

The governing body's responsibility to set goals in traditional contexts extends to both financial objectives and non-financial outcomes including culture, values and ethical outcomes. Organizational and governance policies are generally created and enforced through a combination of controls, business plans, strategies, position descriptions, professional discipline accepted practice, regulation, training, key performance indicators and a variety of executive communications.

The governing body remains accountable for all activities of an organization. This accountability cannot be delegated.

The governing body of an organization has an ongoing responsibility to consider the implications on the organization of any new tool, technique or technology being introduced.

The members of the governing body should assure themselves and be able to demonstrate to stakeholders that their policies (together with the implementation of those policies) are sufficient for the organization, its products and interactions, and the human resources, processes and technology the organization uses. In this respect, the responsibility for and resulting from the introduction of AI is not new. However, AI has the potential to enable new organizational objectives, and to fulfil or extend existing ones, and do so more effectively and more efficiently.

4.2 Maintaining governance when introducing AI

The governing body sets the purpose of the organization and approves the strategies necessary to achieve that purpose. However, it is possible that existing governance is no longer fit-for-purpose when AI is being used within that organization. The specific choice of tools, e.g. AI systems, should be a management decision, made in light of and in line with guidance from the governing body. In order to establish such guidance, the governing body should inform itself about AI in general terms because its use can bring:

— significant benefit to the organization strategically;

— significant risk to the organization, with the potential for harm to its stakeholders;

— additional obligations to the organization.

The governing body should assess its intended use of AI as part of its risk appetite. Risk can change rapidly. New insights and a proactive approach provide an organization with the means to respond to risk. The organization should therefore demonstrate willingness to modify or abort projects, if deemed necessary. For further guidance see ISO/IEC 38506.

New implications arise from the use of AI, including but not limited to:

— increased reliance on technology and systems for the acquisition of data and assurance of its quality;

— transparency and explainability of AI systems (including insight into the objectives, assumptions and rules included in them) when partly or fully automated systems are used for addressing tasks and problems that were previously performed by humans (e.g. credit scoring) together with adequate processes to modify and update those algorithms;

— the possibility that existing direction and controls are not appropriate to ensure required outcomes (and mitigate the risk of undesirable consequences) or can even be compromised. This is due to the differences in assumptions that can be made when delegating to a human, as opposed to when making use of, or acquiring support from, AI.

EXAMPLE 1 An instruction to "defer credit repayment until after the holidays" is sufficiently clear in context to another human operator but insufficiently precise for an AI system to execute correctly.

© ISO/IEC 2022 – All rights reserved

— competitive pressure due to the sales and operations of an organization not using AI;

— accepting the use of AI systems without awareness or consideration of potential bias, error or harm, or of the implications of embedding AI within existing complex systems;

— the growing disparity between the speed of change in automated learning systems and the corresponding human controls of compliance;

— the impact of AI on the workforce, including concerns about discrimination, harm to the fundamental rights of workers, redundancy due to automation or de- and re-skilling, and the possible loss of organizational knowledge, but also leveraging AI to increase human creativity, increased quality of work by delegating repetitive, trivial or dangerous tasks to an AI system;

— the impact on commercial operations and to brand reputation.

The use of AI can also reduce or eliminate certain existing risks and the governing body should review and adjust its risk assessment accordingly.

EXAMPLE 2 An AI system can reduce the risk of error when deployed to complement humans engaged in repetitive tasks, or where humans are required to continuously monitor systems looking for rare anomalies (e.g. security guards).

4.3 Maintaining accountability when introducing AI

Members of the governing body are responsible for oversight and outcomes of the organization as well as for the systems and practices that enable such assurances to be made. They are accountable for the decisions made throughout the organization, including those that are made through the use of AI and for the adequacy of governance and controls where AI is being deployed. They are thus accountable for the use of AI considered acceptable by the organization.

The governing body should take responsibility for the use of AI, rather than attributing responsibility to the AI system itself. Members of the governing body are responsible for informing themselves about the possibilities and risks raised by using AI systems. Members of the governing body should be conscious of the risk of anthropomorphising AI, a phenomenon by which human characteristics (e.g. thinking, emoting, judging, moralizing) are unduly attributed to AI systems, out of proportion, or in a manner inappropriate, to that which is necessary in order to understand the role played by the use of AI.

Members of the governing body can be held to account for the mis-actions of the organization in cases where inadequate diligence, care, guidance, training, oversight and enforcement within the organization allow issues to arise. Such accountability can be ensured by the governing body itself or imposed by stakeholders or through other means. Members of the governing body can face a penalty, removal from office, or legal redress.

The governing body therefore should ensure that its practices are fit-for-purpose for the specific uses to which AI is being applied within the organization. This can include review and, where necessary, enhancement of:

— **Direction**: through policy, strategy, allocation of resources, codes of ethics, statements of values, purpose or other instruments relating to the use of AI in the organization;

— **Oversight**: through an evaluation of AI, an assessment of its value to the organization and the organization's risk appetite, and assurance of implementation, monitoring, measurement, decision assurance and other mechanisms relating to the use of AI in the organization;

— **Evaluation**: considering different elements, e.g. the internal and external factors relating to the organization, current and future threats and opportunities, outcomes achieved, effectiveness and efficiency of the governance mechanisms in place, and judgements about decisions and options taken.

© ISO/IEC 2022 – All rights reserved

— **Reporting**: to demonstrate to stakeholders that the use of AI is being effectively governed by those accountable (compare this with the tasks of 'evaluate', 'direct' and 'monitor' in ISO/IEC 38500:2015, 4.2).

The governing body should also ensure that it has sufficient capabilities to deal with the implications of the use of AI. Actions to address this can include:

— improving AI-related skills among its members;

— increasing the frequency of review of the organization's use of IT and AI in particular;

— examining and updating the criteria used to monitor both the internal and external environment;

— ensuring that staff interests and concerns (e.g. workplace safety, staff training, quality of work) are represented;

— strengthening oversight by establishing or enhancing subcommittees dealing with strategy, risk, assessment or audit, and ethics.

The governing body's accountability should be established across all aspects of intended or actual use of AI and in a manner that is sufficient to ensure the intended outcomes, notably:

— when considering the potential impacts of the use of AI;

— when crafting business strategies that incorporate the use of AI;

— at purchase, implementation, configuration, deployment, testing and other project phases throughout an AI system's life cycle;

— changes in the environments to which the AI is exposed, the learning and actions, decisions and outputs of the AI system, as well as its impacts on stakeholders;

— that appropriate security controls are in place to protect the organization, its stakeholders and its data;

— at decommissioning, including the knowledge and data that are contained in the AI system.

Alongside issues associated with AI itself, there are other issues associated with newly introduced technologies that can affect the organization and its stakeholders, including:

— misunderstanding the nature of the technologies;

— making inappropriate governance decisions;

— omitting appropriate governance oversight of AI;

— failing to include AI in the scope of existing governance;

— applying the technologies inappropriately or ubiquitously without context-specific awareness, appropriate planning, policy or training;

— failing to protect and secure information and assets against automated attacks that use AI to identify vulnerabilities;

— failing to address the implications of emerging relationships between humans and AI systems.

© ISO/IEC 2022 – All rights reserved

5 Overview of AI and AI systems

5.1 General

AI systems come in a range of forms and warrant different degrees of oversight by the governing body. As such, the governing body should understand what the "use of AI" entails and at what stage in its use the governing body should be involved either directly or through appropriate governance mechanisms.

AI systems build on existing IT capabilities including networking, Internet of things devices, e.g. sensors and actuators, big data and cloud computing.

Most of the recent advances in the field of AI technologies relate to the domain of machine learning (ML). ML is an AI technique that gives computers the ability to "learn" without being explicitly programmed. Data are key: they can represent, e.g. text, numbers, pictures, symbols, formulae, graphs, images, speech, sound or videos. A model of an existing data set is created and applied to new data to solve a particular problem, predict an outcome or to categorize new input data.

The nature of AI systems based on ML, including the objective of their use, the choice of algorithms, data driven approach, training methodologies and probability-based outputs, is such that there is potential for additional risk to, and opportunities for, the organization (see also 6.7).

AI systems can automate decision-making by analysing data to provide a potentially probabilistic outcome and, in many cases, acting on that outcome. AI systems can change the nature of products, processes and relationships as well as how an organization operates. This can have material impacts across most industries.

As with any powerful tool that offers benefits, the potential for harm also exists. Therefore, the use of AI should be included in the organization's risk assessment.

The use of AI can result in new obligations for the organization. These can be legal requirements or as a consequence of the adoption of voluntary codes of practice, whether directly within an AI system's automation of decision-making processes or indirectly through its use of data or other resources or processes. The potential for an AI system to cross the boundary between presenting options for action and executing the action itself, without a human involved, should be a major consideration for the governing body.

AI can be distinguished from other technologies by the sheer volume and complexity of data gathered from various sources that can be too complex for humans to handle or adequately process, and specifically, from the perspective of governance implications, by:

— the capability for decision automation;

— the use of data analysis, insights and learning rather than explicit human coded logic to solve problems;

— the capability to adapt as the AI system's environment changes, in ways that are not explicitly coded and necessarily known in advance.

These three elements have wide ranging implications for the organization and its governance.

5.2 How AI systems differ from other information technologies

5.2.1 Decision automation

AI systems generally create output based on historical and current data chosen for the tasks for which the AI system is designed. In modern AI systems, in particular those based on ML, the resultant prediction is usually represented as a probability. For example:

— there is a 97 % probability that this part does not meet the quality requirements;

© ISO/IEC 2022 – All rights reserved

— there is a 92 % probability that taking this route will be the fastest;

— there is a 98 % probability that the face in front the of camera belongs to “Barry Jones”.

Based on these probabilities, different courses of action can be taken by the AI system (actions which can change over time due to retraining). In some cases, automated decisions can be made in fractions of a second.

The speed at which such decisions can be made is a key element of AI systems, making them powerful tools for organizations to use. The scope of automated decision-making can involve a large amount and variety of input data being taken into account.

5.2.2 Data-driven problem-solving

AI systems, in particular those based on ML, typically examine large amounts of data, and identify and refine patterns found in that data. This capability enables such systems to address certain types of problems that are too difficult or too time consuming to be solved by humans on their own or by using classical programming techniques.

Rather than a human driving each logical step and coding towards solving problems, data drive the process. For example:

— recognizing a face is something humans do very well but describing exactly how that is achieved is very difficult. AI systems achieve this by examining thousands of images of faces, learning patterns in those data and applying those patterns to new images of faces;

— the process of finding a cancer cell amongst millions of healthy cells is very difficult but ML techniques for image recognition have proven to be applicable for cancer diagnosis.

Increasing use of ML, together (in the case that a system evolves due to retraining and continuous learning) with ever-changing data, dramatically speeds up problem solving and brings enormous potential as well as governance challenges.

5.2.3 Adaptive systems

Some AI systems undergo retraining during their operational phase or perform ongoing training. Input data perceived during operation are used by the AI system to improve and optimize its internal models. Hence, over time such AI systems can generate different outputs from the same input. Such adaptation can evolve as a reaction to external stimuli such as changes in the environment or feedback reaching the AI system. As an example, consider a recommendation system that monitors a user’s reaction to its recommendations and adapts to increase the acceptance rate. This is ‘continual learning’ (ISO/IEC 22989:—, 3.1.10).

Adaptation can also result from self-improvement, e.g. in a chess-playing AI system, which is able to learn solely from playing against itself.

Such adaptive systems also have implications for governance:

— Capabilities are potentially required of an AI system to provide insight into its functionality, e.g. in an assessment or audit situation.

— As an AI system’s functionality changes over time, assurance that the system is still operating within acceptable boundaries can require further assessments or other control mechanisms.

— The adaptability of the AI system can increase the organization’s agility and provide it with a strategic advantage.

— AI systems can inadvertently circumvent existing governance controls. Management should ensure that AI systems explicitly comply with existing governance controls.

© ISO/IEC 2022 – All rights reserved

5.3 AI ecosystem

Understanding the many components of an AI system and how these can relate to the organization requires a layered approach to the AI ecosystem.

As Figure 1 shows, the AI ecosystem is broad and includes a spectrum of different technologies. A more detailed version of this figure is available in ISO/IEC 22989:—, Figure 6, where further details on ML elements and computational resources are described. If AI techniques are used, some of the other components in this ecosystem become a functional part of the overall AI system and should be treated as such from a governance perspective.

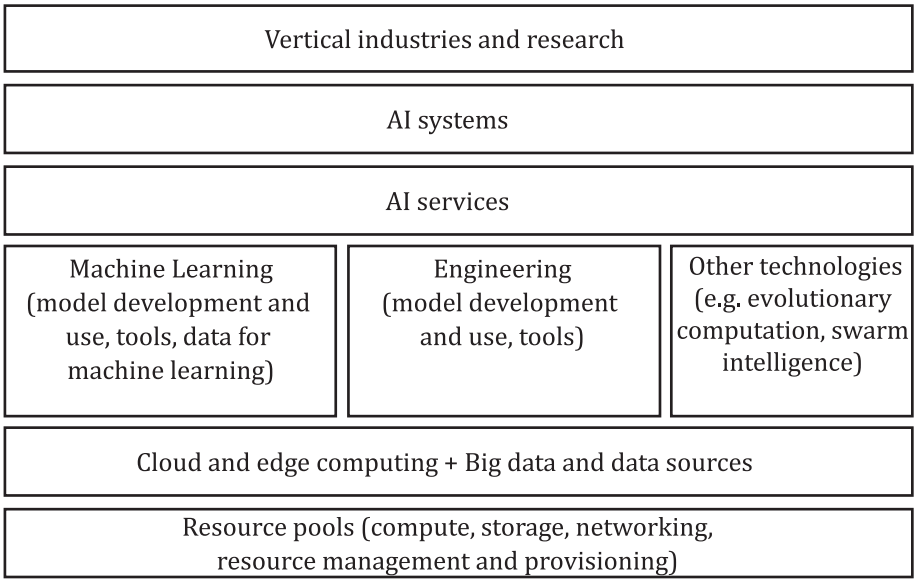

Figure 1 — Simplified version of an AI ecosystem from ISO/IEC 22989:—, Figure 6

This layered representation of an AI ecosystem can help delineate the level and the scope of accountability based on the organization's role in the ecosystem as well as the appropriate level of accountability of the governing body.

The complex nature of AI ecosystems means that the scope for oversight will vary greatly based on several factors including:

— the intended purpose of the AI system;

— the type of AI used;

— the functional layer of the AI ecosystem used;

— the potential benefit the AI system will deliver;

— the new risks that can accompany the AI system;

— the stage of implementation of the AI system;

— the role played by the organization in the AI value chain (e.g. role as AI provider, producer, customer). See ISO/IEC 22989: —, 5.17).

These factors are influenced by the organization's purpose and objectives, and the decision to use an AI system in preference to other information technologies. These factors are mostly considered at the beginning of the development and deployment of an AI system. However, there are further organizational considerations that should be factored during the entire life cycle.

© ISO/IEC 2022 – All rights reserved

Figure 2 shows the AI system life cycle from inception to retirement. Once an organization has decided on its use of AI, this can assist governing bodies with identifying both the points or "gates" at which key governance questions can arise and need to be addressed by the governing body.

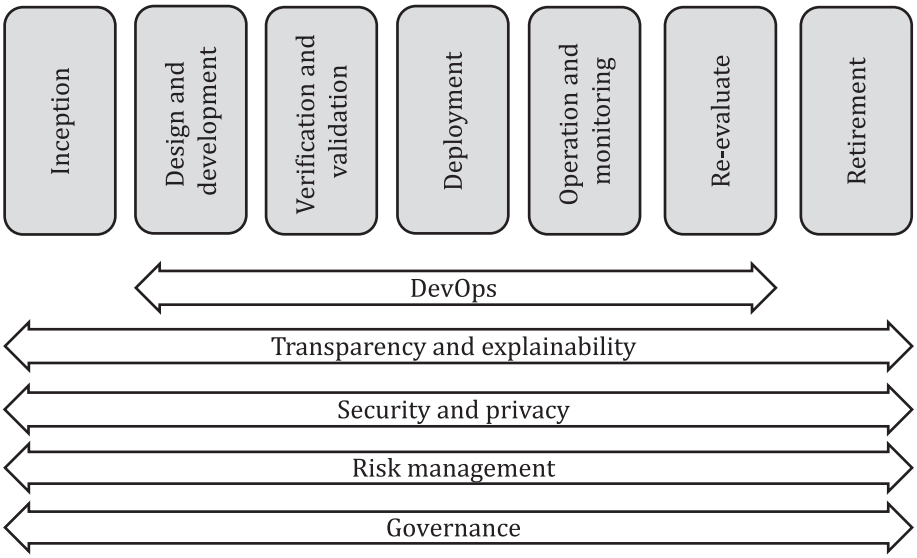

Figure 2 — AI system life cycle model from ISO/IEC 22989:—, Figure 3

With these AI system life cycle considerations, the organization can begin to understand how to evolve their existing governance mechanisms. The life cycle can help determine where additional governance is required given an AI system's dependence on data. The governance of data is addressed in the ISO/IEC 38505 series. The life cycle can also help the governing body identify at each stage where the organization can be exposed to additional risks and then take decisions on how to manage such risks. In 5.4 and 5.5, additional governance considerations arising from the use of AI are described further.

5.4 Benefits of the use of AI

A governing body can consider deploying AI in order to pursue specific opportunities that the organization has identified, including the organization's future growth or better achieving the organization's purpose and objectives. In such cases, the governing body needs to weigh those opportunities against risk and other implications of use. It should ensure that oversight mechanisms exist and are used to review whether the deployment remains consistent with the organization's strategic intent, purpose and values. AI can be used to fulfil objectives and create value for the organization. This can result in

— a reduction in costs of operations,

— new products and services that will accelerate commercial disruption of existing businesses, or

— the transformation of organizations such as government entities or not-for-profit organizations.

An organization is potentially already using AI in its daily operations.

Examples of uses of AI during the daily operation of an organization include:

— the use of a satellite-based navigation system that finds the most efficient route for trucking;

— most internet searches use AI to answer a query, find a similar picture or a document;

— productivity applications use AI to suggest a better design for a slide in a presentation or an improvement to grammar in a document.

© ISO/IEC 2022 – All rights reserved

Further applications can be found in various domains (e.g. agriculture, defence, education, healthcare, manufacturing, transportation) and use various deployment models (e.g. cloud services or on premises-systems, cyber-physical systems). An extensive collection of use cases is included in ISO/IEC TR 24030, covering these and other domains and deployment models.

An AI system can, for example:

— change the nature of products, processes and services;

— automate feedback collection, comparisons, translations, reviews or analyses of data, leaving value-added activities to humans (e.g. conclusions to be drawn or decisions to be made);

— increase customer service, productivity and product quality;

— change the relationship with customers and suppliers (and thus also between employees, customers and supply chains) using intelligent agents such as chatbots to collect relevant information, triage and route requests, help with product, service or billing questions or resolve problems;

— increase the efficiency of manufacturing and agriculture by moving from “automated assistance” to “fully automated”;

— identify the most likely intent of the speaker based upon samples of spoken language;

— predict the spread of diseases based upon historical data of similar outbreaks of diseases;

— suggest a possible instance of cancer based on examination of cells;

— recommend granting or refusing a loan.

Such use of AI can bring benefits to the organization but can also present new or more acute challenges, e.g. additional risk or bias, issues which are addressed in more detail in Clause 6.

Equipped with AI, an organization can reshape its strategy and apply models suitable for its sector. It can change how it creates value for its stakeholders and how it captures this value. Using AI also creates opportunities for the workforce, e.g. greater creativity mobilized through the use of AI, increased competitivity and different job opportunities that eliminate repetitive, trivial or unsafe tasks. In this way, AI technologies can bring profound transformation to organizations. The governing body should learn about the opportunities that AI presents for the organization and promote its adoption when appropriate.

By using AI to handle the process burden of such tasks, an organization can focus efforts and resources on tasks in which the value added comes from human efforts that cannot be automated or replicated by AI.

5.5 Constraints on the use of AI

While the organization can easily identify the benefits of the use of AI, the governing body should understand the constraints and obligations that such use places on the organization. To adequately govern these constraints and obligations, the organization should take some of the following actions:

— **Increase oversight of compliance**. Without the appropriate oversight, the use of AI can automate processes, produce outcomes that undergo frequent change, can be difficult to explain or conflict with organizational policies (see 6.6).

— **Address the scope of use**. Ensure that the scope of automation is overseen by the governing body and implemented by appropriately authorized and skilled people (see 6.3). The governing body should ensure that the requisite authority, responsibility and accountability are maintained and that the consequences of such automation are examined and understood before implementation.

— **Assess and address the impact on stakeholders**. While some decisions will negatively impact stakeholders (e.g. refusing a bank loan to a customer), the organization should ensure that such impacts are not exacerbated by the use of AI. Existing risk and impact assessments, together with

© ISO/IEC 2022 – All rights reserved

mitigation processes, e.g. those preserving legal rights or ensuring legal certainty, should remain effective (see 6.7).

— **Determine legal requirements or obligations of using such technology**. It is increasingly likely that the use of AI and associated solutions, e.g. facial recognition or automated vehicle movement, will be challenged and can be the subject of new legal requirements. The governing body should show how such obligations are met and that they are in line with requirements and suitably explained if required.

— **Align the use of AI to the objectives of the organization**. The innovative use of new technologies is critical to the viability and health of many organizations and, in those cases, governance will encourage such innovation. Not every project will be strategically important (e.g. some will only reduce costs), but overall, the use of AI should assist the organization in reaching its objectives.

— **Align the use of AI to the organization's culture and values.** Decisions proposed by an AI system should take into account organizational policies, expectations (including impact of use) and ethics.

— **Ensure that problem solving takes due account of context**. An organization needs to ensure that contextual elements, essential to understanding behaviour, values and culture, are not missing, or omitted from the data that it is using to solve problems.

— **Examine the additional risk that the use of AI can bring to the organization**. By examining the organization's purpose and objectives and considering the possible sources of risk identified if it decides to use AI, together with measures to treat these risks, the organization should remain within the bounds of its risk appetite (see 6.7).

Beside these broad constraints on the use of AI by the organization (and any others that are applicable to the organization or its stakeholders), AI systems themselves will have technology constraints. The governing body should also seek assurance from management that such constraints are adequately managed.

6 Policies to address use of AI

6.1 General

AI systems use data to build models and make predictions as well as potentially automate decisions and actions. As with any use of human resources, processes or technology, the governing body remains accountable for these decisions and actions. Some use of AI can inadvertently result in policy non-compliance or result in outcomes that existing policies did not foresee.

The specification and authorization of directions and controls that are appropriate for an organization not using AI, are potentially not appropriate for an organization that uses AI.

Continued effective governance requires reviewing and potentially reinforcing existing governance mechanisms and controls to ensure that they are robust, explicit and appropriate to cover the additional governance considerations (as well as direction to management) that AI systems bring to the organization and its stakeholders.

The governing body and managers should further involve stakeholders that can be impacted by the use of AI systems, such as personnel and their representatives, throughout the entire process of implementing AI systems at organisations, via information, consultation and participation procedures.

The guidance in this clause is intended to help the governing body understand and revise policies it should consider in order to reduce the occurrence of avoidable and unintended consequences of the organization's use of AI. It is not and cannot be comprehensive.

According to ISO/IEC TR 38502:2017, 4.1.3, managers are responsible for ensuring the achievement of the objectives of the organization within the strategies and policies established by the governing body. The task of governing is accomplished in close cooperation between the governing body and managers as shown in Figure 3.

© ISO/IEC 2022 – All rights reserved

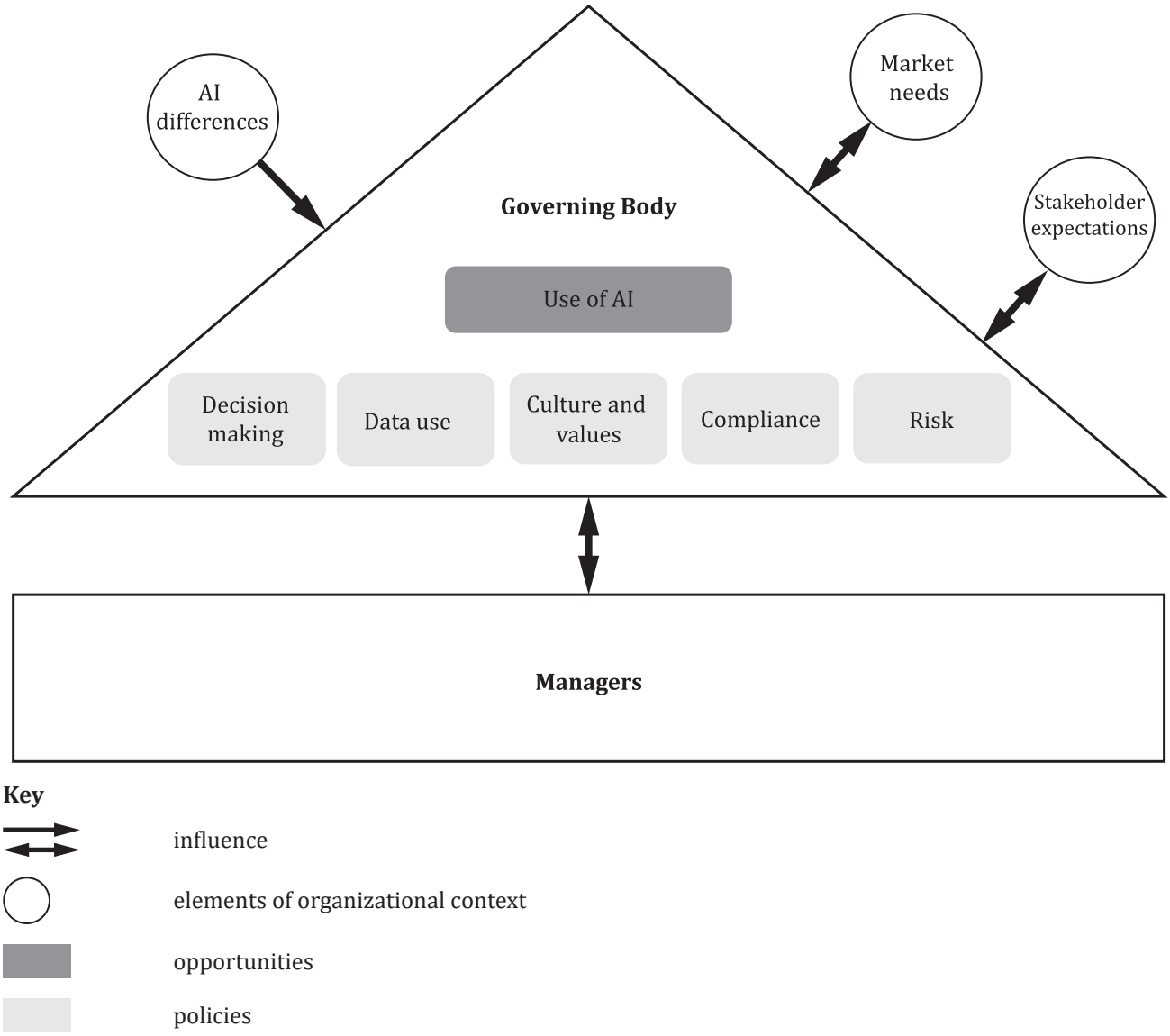

Figure 3 — Governance implications of the use of AI

Figure 3 is based upon and modifies the "Model for Governance of IT" in ISO/IEC 38500:2015, Figure 1 (reproduced in Figure A.1).

Regardless of the use of any technology, the governing body should set and oversee the achievement of outcomes that are aligned to its principles. Such principles can arise internally or be suggested or imposed by external organizations (e.g. in[18],[19],[20] and[21]).

The speed of technological innovation and ever-changing legal requirements should encourage the organization to actively maintain a set of principles for its use of AI and ensure that they remain appropriate for the organization's use of AI.

6.2 Governance oversight of AI

The governing body should ensure that oversight arrangements for AI are established and are appropriate to the risks associated with the organization's use of AI.

Governance oversight, based on policies set by an organization, should identify the individual and the collective accountability in a chain of responsibility. Good governance oversight should be based on the

© ISO/IEC 2022 – All rights reserved

general understanding of the usage (and context of use) of a system. See "External and internal context" in both ISO/IEC 23894:—, 6.3.3[3)] and ISO 31000:2018, 6.3.3.

As part of its governance oversight of AI, a governing body should ensure that:

— policies are in place to ensure the appropriate use of AI;

— responsibilities, chain of responsibility, accountability, authority and potential delegation of authority are clearly defined and agreed both within the organization and, where applicable, between different parties in any value chain;

— adequate human oversight is in place while using AI;

— any human using AI or responsible for the use of AI:

  — has appropriate understanding of the AI system being used;

  — is properly informed and trained, including knowing how and with whom to raise any concern;

  — has the authority to make decisions (or knows to whom to request that a decision be made) and knows to whom to report back;

  — has sufficient control over the AI system, including the possibility to intervene when necessary.

6.3 Governance of decision-making

Governance of decision-making is part of the overall governance of the organization.

Authority and responsibility are delegated to people throughout an organization in order to spread the burden of work done and decisions made. Wherever such delegated authority or responsibility resides, accountability remains with the governing body for all work done and decisions made (see 4.3). Some of those decisions and some of that work can be increasingly made and carried out by AI systems but the governing board remains accountable in the same manner as for human decisions and work.

Revising decision-making policies regarding the use of AI is intended to ensure that a human or group of humans is and remains clearly accountable for the decisions to which they have delegated authority and responsibility. That is regardless of whether the decision was partly or completely automated.

The governing body should monitor the types of decision and output generated by automated systems and direct management to ensure that such systems are configured to operate within acceptable bounds by implementing appropriate controls. Such controls should provide the governing body with appropriate visibility of the conformance of decision-making to organizational policies, together with any exceptions thereto.

The governing body should seek assurance that active oversight of such controls is delegated to an appropriately resourced member of staff who has the authority to make or instigate responses to issues identified. The use of automated decision-making, delivered by an AI system, does not alter the accountability of the governing body (nor the responsibilities of any delegated authority) for such decisions.

Important factors relating to decision-making within the organization include:

— **Alignment to objectives.** Decisions should align to the organizational objectives while keeping within the allocated resources, defined risk and other controls imposed by the organization.

— **Level of responsibility.** Ensuring that the level of decision-making matches the authority granted and responsibility associated to the decision is a critical element of good governance. Defining the scope and impact of possible decisions and matching those to the levels of responsibility is necessary to empower staff to act appropriately and thus make the whole organization more agile. Human accountability for the actions and decisions performed by an AI system should be clearly defined

3) Under preparation. Stage at the time of publication: ISO/IEC DIS 23894:2022.

© ISO/IEC 2022 – All rights reserved

and assigned to staff with the appropriate authority and tools to execute the tasks assigned them (e.g. visibility, oversight, quality control) and to take sufficient corrective action if problems are identified.

— **Decision-making capability.** Decision-makers should be adequately skilled and trained for the decisions for which they are responsible. Controls should be implemented to ensure AI systems are adequate to the task they have been set. See ISO/IEC TR 24028.

— **Decision-making process.** Stakeholders are showing increasing demand for transparency of decision-making. This transparency includes reporting requirements, strategic direction and interactions with all internal and external users. Along with this increased transparency comes an increase in the expected level of engagement and care in decision-making. Therefore, the processes of decision-making (together with actions required to mitigate the consequences of poorly made decisions) are an important governance mechanism and will become increasingly important as the use of AI increases.

— **Decision-making oversight**. The governing body should ensure that there is adequate oversight, that controls are implemented to ensure effective decision-making capabilities and that there is appropriate visibility of both conformity of decision-making to organizational policies and any exceptions. For conformity exceptions in decision making, the need for additional transparency and accountability should be determined. Appropriate means should be given to all stakeholders to identify and report non-compliant behaviour, or decision-making outcomes generally (whether or not they include AI systems) and be given meaningful, timely and adequate response.

6.4 Governance of data use

Some AI systems rely on data to build and train a model and therefore governance of data use is critical to the responsible use of AI. The governing body should ensure at an early stage that existing governance and management are adequate for the purpose for which that data are being used and that sensitive data are protected and secured (see ISO/IEC 38505-1). This includes, for example, documenting how the organization complies with its obligations including in regard to procurement and sale, privacy and security, retention and disposal of data as well as overseeing its own policies regarding data management (see also 6.7.4).

Additional governance considerations in the use of data for AI systems include:

— the relevant design choices made for the AI system;

— data collection;

— relevant data preparation processing operations, e.g. annotation, labelling, cleaning, enrichment and aggregation;

— the formulation of relevant assumptions, notably with respect to the information that the data are supposed to measure and represent;

— a prior assessment of the availability, quality, quantity and suitability of the data;

— examination in view of possible biases;

— the identification of any possible data gaps or shortcomings, and how those gaps and shortcomings can be addressed.

One major difference between AI systems (particularly those involving ML) and other IT systems is that they rely on data inference to produce a result with a level of confidence. A second major difference is the existence and use of training data which, depending on its quality, can improve or degrade the quality of any result. This is different than software development that relies on humans conceiving and programming the specific logical steps that are required to produce a result. Subclauses 5.2.2 and 5.2.3 outline how such systems differ. As data quality is critical to the performance of AI systems, the governing body should seek assurance from management that data are of requisite quality for its intended use in AI systems.

© ISO/IEC 2022 – All rights reserved

Depending on the application, using historical data to train the AI system can result in:

— repeating earlier mistakes;

— making 'unconscionable' decisions in relation to outcomes not recorded in the data;

— making decisions in relation to novel combinations of circumstances or data that are difficult to understand by a human, e.g. the linking of different datasets that are not obviously related;

— surveillance and behavioural information not necessarily voluntarily disclosed;

These results are best understood as mistakes or errors, introduced by the use of incomplete, erroneous or inappropriate data. This is sometimes (and often incorrectly) referred to as "AI bias". However, bias is introduced or reinforced by humans, not by AI systems themselves (see also 6.7.4).

Existing governance of the use of data and data management practices should be reviewed where data are used in AI systems. In addition, for shared AI systems such as for industry analysis, additional governance policy and management controls can be required.

Other technology assets (e.g. algorithms, neural networks and natural language processing systems) can also require similar enhancements to existing organizational and governance policies.

The organization should consider applying the provisions of ISO/IEC 38505-1, relating to the use of data, to these aspects of governance of the use of AI:

— **Value:** the quality and quantity of the data, its timeliness, the context and cost of its use.

— **Risk:** data security, misuse and privacy, as well as competitive risks involved in not leveraging the use of data.

— **Constraints:** legal requirements, contractual obligations, copyright or commercial interests.

See also A.3 on governance of data use.

Data and their use by organizations is an increasingly important issue for all organizations and their stakeholders. In accordance with the principles, models and data-specific aspects of governance outlined in ISO/IEC 38505-1, governing bodies should take actions that ensure the effective governance of, and investment in, the organization's use of data, and treat risks involved in that use. This includes ensuring that the correct data are being used for the correct purpose.

6.5 Culture and values

The governing body is responsible for defining the organization's desired culture and values with respect to stakeholders, markets and regulators, and changes in societal expectations of the organization's operations and impact. Governing bodies should be aware of emerging guidance and international norms of behaviour such as those reflected in the Universal Declaration of Human Rights,[22] the Johannesburg Declaration on Sustainable Development[23] and other instruments. Further details can be found in ISO 26000:2010, 3.3.2.

Any decision or action of the organization should align with its culture and values. However, much of an organization's culture and values are implicitly embedded in the behaviour of its staff and processes. An AI system has no equivalent of human understanding of context, of common sense, morality or knowledge to guide its output. Instead, an AI system relies on models, algorithms, or training data to achieve comparable results.

For these reasons, the governing body should be explicit about its culture and values and have the appropriate governance mechanisms and policies to ensure such AI system behaviours can be monitored and corrected when needed. In some cases, the scope and impact of the AI system should be constrained, and augmented by human action, so that governance policies can be ensured. How that is achieved will vary according to circumstances. Equally, an AI system can help identify where human decision-making is flawed (e.g. by inappropriate bias or discrimination, or by poor reasoning

© ISO/IEC 2022 – All rights reserved

or deduction) and thus provide insights into an organization that the governing body can subsequently address. This raises future challenges for the governance of organizations.

Depending on the nature and purpose of the AI system intended to be used, it is possible that the organization will need to explicitly oversee these systems to ensure they comply to the culture and values of the organization. Governance mechanisms can take the form of a "Culture and Values Board" or an "Ethics Review Board". Oversight can consist of a review being required before a sensitive or high-risk AI systems project is approved, or certain escalation criteria are met. Additional organizational policies in this area can require the use of technical controls within an AI system in order to assist with compliance to those policies. For more information on ethics and societal concerns relative to the use of AI, see ISO/IEC TR 24368[4].

6.6 Compliance

6.6.1 Compliance obligations

To ensure that the organization meets its compliance obligations, the organization establishes an ongoing compliance process. To be effective, the compliance process is shaped by organizational leadership and embedded in the culture of the organization. A compliance management system, such as that described in ISO 37301, addresses these requirements and helps the organization to demonstrate its commitment to meeting its obligations and comply with the expectations of its stakeholders.

The governing body should seek assurances that management configures and maintains any AI system used by the organization to meet the organization's compliance obligations and avoid breaches of compliance. Examples of breaches include pricing mechanisms that violate anti-trust legal requirements or the use of data for training that violates civil rights or is discriminatory.

Compliance is concerned with the behaviour of operational staff. AI can reduce the effect of human behaviour in certain situations, thereby reducing the risk of non-compliance that is a result of human behaviour. As a consequence, the human factor is now elevated or moved towards those that design, develop and implement AI systems. Such a shift poses a whole new challenge in terms of awareness, training and monitoring.

When examining the role of AI in the context of the organization and its compliance obligations, the governing body should address these items:

— stakeholder expectations regarding the use and impact of AI (ensure that they align with existing policies and compliance obligations);

— the cultural implications of the use of AI within the organization, such as the increased sophistication of an AI system that can lead staff to assume that it is error-free. The organization should establish appropriate processes to validate the output of an AI system;

— the extent of the use of AI by the organization;

— the impact of AI on the organization's compliance requirements;

— changes to the organization's risk appetite arising from the use of AI;

— the extent to which the existing oversight processes, structures and controls address the specific aspects of AI;

— the impact of the use of AI on the accountability structures within the organization;

— the existing compliance management system should be assessed against the implications posed by the use of AI (where for example, designs and solutions can be difficult to explain and be subject to frequent changes), which can include, e.g. extending the compliance policy, or expanding the periodic risk assessment;

4) Under preparation. Stage at the time of publication: ISO/IEC DTR 24368.

© ISO/IEC 2022 – All rights reserved

— whether individuals are affected by the use or creation of personal data by an AI system that is constrained, e.g. by law or by organizational policy.

6.6.2 Compliance management

The use of a compliance management system standard allows the organization to overlay its specific management requirements, e.g. security, quality and privacy, on their management system in an integrated way. This makes it easier to add additional compliance measures to the management system resulting from the use of AI.

At the compliance management level, particular attention should be paid to:

— extending compliance processes to account for the speed, scope or sophistication of the AI system (e.g. to increase the level or frequency of monitoring);

— requests for human re-evaluation of decisions made by the AI system;

— additional controls to ensure the AI systems remain within required compliance conditions;

— ensuring that the use of data for model building or training complies with policy;

— the use of AI to monitor other AI systems and the extra monitoring or alerting that can be required.

The organization should assess the impact of any planned use of AI using the compliance management system. Possible uses of AI within the management system should be part of this evaluation.

6.7 Risk

6.7.1 Risk appetite and management

Governance of an organization, however it is performed, includes defining the organization's purpose and objectives as well as the strategy for how these are to be achieved. In considering its strategy, the governing body of the organization decides the risk appetite, in pursuit of its objectives. To keep within this risk appetite, and to assist the governing body in making decisions regarding the risk appetite, the organization establishes a risk management process.

Risk management involves assembling relevant information for an organization to make decisions and address risk. While the governing body will have defined the overall risk appetite and organizational objectives, it delegates the decision-making process of identifying, assessing and treating risk to management within the organization. The differing roles of governance and management in addressing risk are discussed in:

— this document, which describes additional governance considerations for the organization regarding the development, purchase or use of an AI system. Such considerations include new opportunities, potential changes to the risk appetite as well as new governance policies to ensure the responsible use of an AI system by the organization;

— ISO/IEC 23894, which describes the management processes that should be performed within the organization to address the additional risks to the organization through the introduction of an AI system.

Treating risk involves both the governing body, which determines and is accountable for the risk appetite, and management, which works to maintain risk within the bounds of the agreed risk appetite.

A review of current risk management processes should particularly examine whether the risks involved in decision-making, data use, culture and values, and compliance are well understood and managed. In this way, the context of the additional risks that AI systems bring to the organization can be clarified. Once those additional risks are identified, the governing body is in a better position to:

— ensure that risks are adequately considered when setting the organization's objectives;

© ISO/IEC 2022 – All rights reserved

— understand the risks facing the organization in pursuit of its objectives;

— ensure that systems to manage such risks are implemented and operating effectively;

— ensure that such risks are within the risk appetite of the organization;

— ensure that information about such risks and their management is effectively communicated (using, e.g. ISO 31000:2018, 5.2).

6.7.2 Risk management

Risk management is integral to all organizational activities. Although AI systems can deliver benefit to the organization, the organization's objectives related to good governance of decision-making, to use of data, and to the organization's desired culture and values should be revised to take account of possible impacts of the use of AI.

The organization strives to protect its principles and values, its identity and reputation, its stakeholders, its market, its environment and furthermore to protect its freedom of action, in order to succeed.

In order to address the risks posed by AI, the governing body should put in place:

— appropriate internal rules and policies;

— appropriate specific sub-organizations, processes and tools designed to guarantee or enforce values, principles and internal controls that are foundational to good governance.

In addition, the governing body should ensure that the organization's contractors and sub-contractors abide by the same codes of practice and policies.

These measures provide the means to any stakeholder to identify and report non-compliant AI system related behaviour and be given meaningful and adequate responses. They are particularly important where complex supply chains are involved and where reputational, financial or other risk can accrue to a lead contractor or procurer. The organization should acquire a clear understanding of the implications of the use of AI within contractual relationships where the risk appetites of different organizations are in play.

Figure 4 shows examples of:

— **objectives** that the organization is pursuing (such as the protection of its reputation);

— additional **sources of risk** that the use of AI can bring to the organization;

— some technical and organizational **controls** that can be used to treat the risks.

© ISO/IEC 2022 – All rights reserved

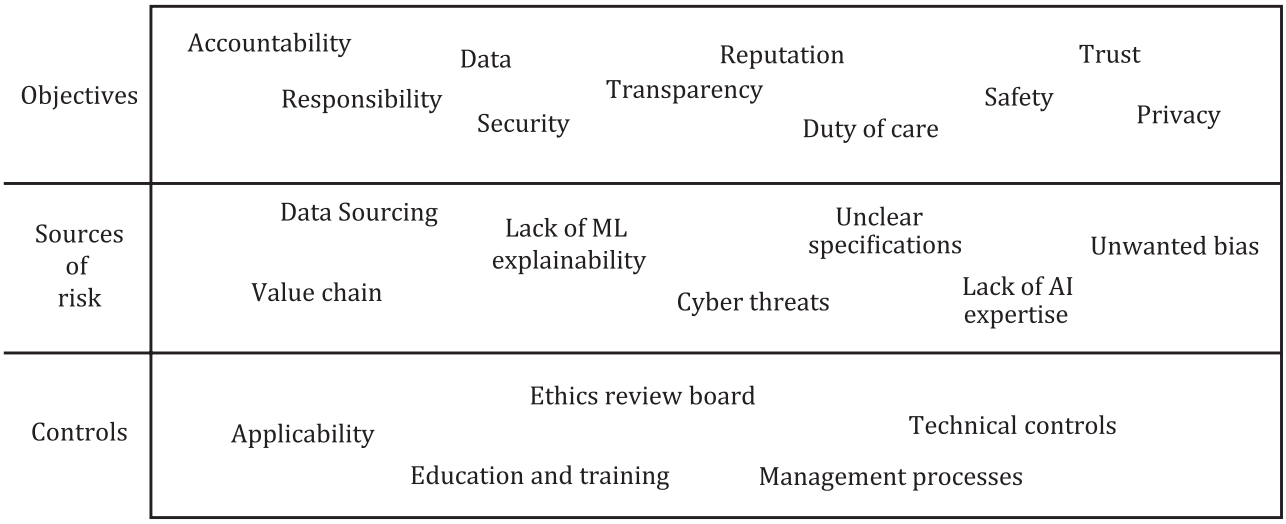

Figure 4 — Risk management and the use of AI

6.7.3 Objectives

The objectives that the organization wants to protect through risk management processes include not only the protection of assets such as data, information systems, application code, algorithms and equipment, but also its duty of care, culture and values, its reputation and its strategy.

Figure 4 shows examples of objectives that the organization is pursuing:

— **Accountability** and **Responsibility**. The governing body should maintain its accountability as well as oversight of the organization's responsibilities both internally and externally for the use of AI.

— **Reputation** and **Trust.** The perspective of risk considered here is to the organization itself. However, the organization does not exist in isolation and therefore consideration shall be given to its stakeholders and the environment in which it operates. Consequences and likelihood of risks affecting stakeholders such as customers and suppliers should be key considerations in terms of risk to the organization.

— **Duty of care.** The organization has a responsibility and possibly a legal duty of care to its stakeholders, both internal and external. This can involve an obligation (with regulatory or legal consequences, depending on jurisdiction) to ensure the well-being of all stakeholders and the protection of their rights, e.g. access to significant financial, health or housing services, and of human rights (including freedom of movement or privacy).

  In situations where there is a significant risk to the duty of care, the governing body should require additional organizational controls to effectively treat such risks and ensure they do not exceed the risk appetite of the organization (see 6.7.5 for examples).

— **Safety**. Where the use of an AI system carries a significant risk of physical or emotional harm, the organization should be especially alert to the nature and consequences of that harm. Where necessary, the organization should put in place appropriate systems for the ongoing management of safety as well as considering how the use of AI can reduce the exposure of humans to dangerous activities.

— **Security** and **Privacy**. The organization should ensure the security of its operations, especially where confidentiality is necessary and the privacy of individuals is important. These objectives should not be altered by the use of AI, particularly given the ability of AI systems to infer new information from patterns in data.

— **Data**. Data are an important resource for the organization and their protection and integrity should be an organizational objective.

© ISO/IEC 2022 – All rights reserved

— **Transparency**. Transparency of decision-making by the organization is likely to be an objective that the governing body wishes to maintain. Stakeholders expect to understand at least some of the important inputs and variables that explain how decisions are made by the organization. This is regardless of how much automation is used in reaching that decision.

Some of these objectives are treated in depth in ISO/IEC TR 24368 and ISO/IEC TR 24028, as well as in ISO 31050[5)].

6.7.4 Sources of risk

The organization should expect additional sources of risk depending on the scope and nature of the domain to which an AI system is applied and the type of AI system deployed. Some uses of AI will be contained and controlled, bringing little, or no additional risk to an organization. Other uses will carry significant exposures that have not previously been present in the organization.

Furthermore, as stated in ISO/IEC 23894, less mature technologies such as those used in the development and application of AI systems can impose risks that are unknown to the organization or are hard to assess. On the other hand, for mature technologies a larger variety of experience data can be available, making risks easier to identify and to assess. A set of "readiness levels" characterizing the maturity of an AI system can be of use in assessing and monitoring risk.

The direct mapping of assets, values and objectives to sources of risk or risks controls is not possible because the effect of any of these items can affect any other item.

The example sources of risk shown in Figure 4 are those that relate to the use of AI, though these sources can also apply to many other technologies or processes. The organization should expect additional sources of risk depending on the scope and nature of the AI systems used. The examples shown include:

— **Data sourcing.** Because data are used to build and train the model in an ML system, the quality and appropriateness of the data are critical and should be aligned with the intended use and objectives of the system. Manipulated data can permit an adversarial attack resulting in model poisoning and misclassifications.

— **Unclear specifications.** In traditional software development, the nature of the problem and its solution are tightly bound because the people involved use similar approaches for both. However, an AI system can arrive at a solution in a very different manner long after the system is initially designed and developed, so the accuracy, clarity and scope of problem specification, system requirements, designed system goals, and incorporated behaviour boundaries can all play significant roles.

— **Value chain.** The supply and distribution of AI systems can include risks as can the use of AI systems by those outside the organization. Accountability is to be clearly defined and agreed between different parties in the value chain.

— **Unwanted bias.** The algorithm, training and test data, and ML model used in AI systems often aim to reflect and predict real world situations. By their nature, these systems are only a small sample or perspective and can produce outcomes that reflect the sample rather than the real world. This can lead to unwanted bias being amplified. See ISO/IEC TR 24027:—[6)].

— **Lack of ML explainability.** Risk arises from the complexity of AI systems. This can make it very difficult to explain how an AI system reached a particular conclusion. This differs from traditional IT systems where a human defines an algorithmic procedure for determining an answer. A human can be questioned, double-checked, assessed by their standing and reputation and be held accountable and any code tested to ensure that it produced the responses predicted.

— **Lack of AI expertise.** The use of AI requires different skills from traditional software development, although those skills are also needed. Additional skills include an understanding of data analytics

5) Under preparation. Stage at the time of publication: ISO/CD 31050.

6) Under preparation. Stage at the time of publication: ISO/IEC DTR 24027:2022.

© ISO/IEC 2022 – All rights reserved

and statistics, modelling and algorithm design and testing, as well as human skills such as ethics and empathy.

— **Cyber threats.** The implementation of some AI systems can be vulnerable to specific and difficult to recognize cyber threats, which leave few or no visible traces behind.

Other sources of risk relate to use of AI and its indirect impact on humans. For example:

— **Hollowed out skills.** The increased use of AI in routine decision-making roles means that humans are decreasingly exposed to experience over time. A reduction in staff experience ('agent atrophy') will remove the opportunity to strengthen human skills, decision-making abilities and expertise. This hollowing out of human skill development is a risk to organizations and society in general.

— **Contractual issues.** The governing body should ensure that applicable contracts, laws and best practices remain valid when AI systems are used. In many AI applications, systems learn from the data and practice of their host organization, in an initial, and potentially ongoing, basis. This uniquely makes AI systems reflective of the organization they serve in a way that can affect the applicability of contracts, laws and best practices.

— **Environmental issues**. The governing body should consider the risk of carbon emissions resulting from energy consumption for AI training and data processing. It should also consider the risk of pollution and resource depletion due to accelerated obsolescence of hardware in favour of newer AI-capable cloud and edge devices.

— **Personal autonomy**. The use of AI increasingly can determine the range of choices individuals encounter in terms of news, entertainment, interactions, products and services. Though individually small, the increasing frequency and ubiquity of these determinations means the governing body should consider the impact of the use of AI on the autonomy of humans, as individuals and as communities.

— **Missed opportunities**. Governing bodies are sometimes too conservative to take advantage of new opportunities. If the governing body does not take up opportunities that the use of AI offers, the organization can face increased competition from organizations that do take up such opportunities.

6.7.5 Controls

Controls aim to maintain or modify risks so that they remain within the range of the risk appetite of the organization. These controls can be organizational controls such as those effected through management structures, review boards or management processes or they can be technical controls that are implemented through measures such as database access restrictions, software code or data filtering.

The controls shown in Figure 4 are just a small selection of those that can be related to AI systems.

Figure 4 shows the following examples of risk controls:

— **Applicability.** A description of the AI system, including for example, its algorithms, data and models, should be transparent enough to ensure its applicability to the intended use.

— **Ethics review board**. For applications that the organization has determined high risk, high stakeholder impact or other thresholds, a common risk control is the use of an ethics review board that can ensure alignment with the culture and values of the organization.

— **Management processes**. The design and implementation of management processes is a common practice to address requirements such as quality, security, privacy and compliance. Similar processes can be constructed for AI system-related requirements.

— **Technical controls**. Technical controls embedded in the software can implement actions to assist treating some risks. See, for example, ISO/IEC TR 24029-1.

© ISO/IEC 2022 – All rights reserved

— **Education and training.** Everyone involved in all stages of the use of AI should receive adequate training to ensure that they acquire and deploy the requisite skills.

© ISO/IEC 2022 – All rights reserved

Annex A
(normative)

Governance and organizational decision-making

A.1 Overview of relevant governance standards

A.1.1 General

ISO 37000 describes the governance of organizations as the system by which an organization is directed, overseen, and accountable for achieving its defined purpose.

At its foundation this includes:

— setting the purpose, mission, vision, organizational ethos, organizational values, and culture to give the organization direction;

— steering the strategy and balancing resources appropriately to achieve that purpose;

— exercising oversight of the organization's performance, ensuring compliance and viability;

— engaging with and accounting to stakeholders.

In the domain of information technology, "governance" is defined in ISO/IEC 38500:2015 as the "system of directing and controlling" an organization through three main tasks of evaluating, directing and monitoring the use of IT.

A.1.2 Direction

ISO/IEC 38500:2015 provides guidance to governing bodies in assigning responsibility for strategies and policies that set the direction for investments in IT and what IT should achieve. They should encourage a culture of good governance of IT in their organization and direct the submission of proposals for approval to address identified needs (see Figure A.1).

© ISO/IEC 2022 – All rights reserved

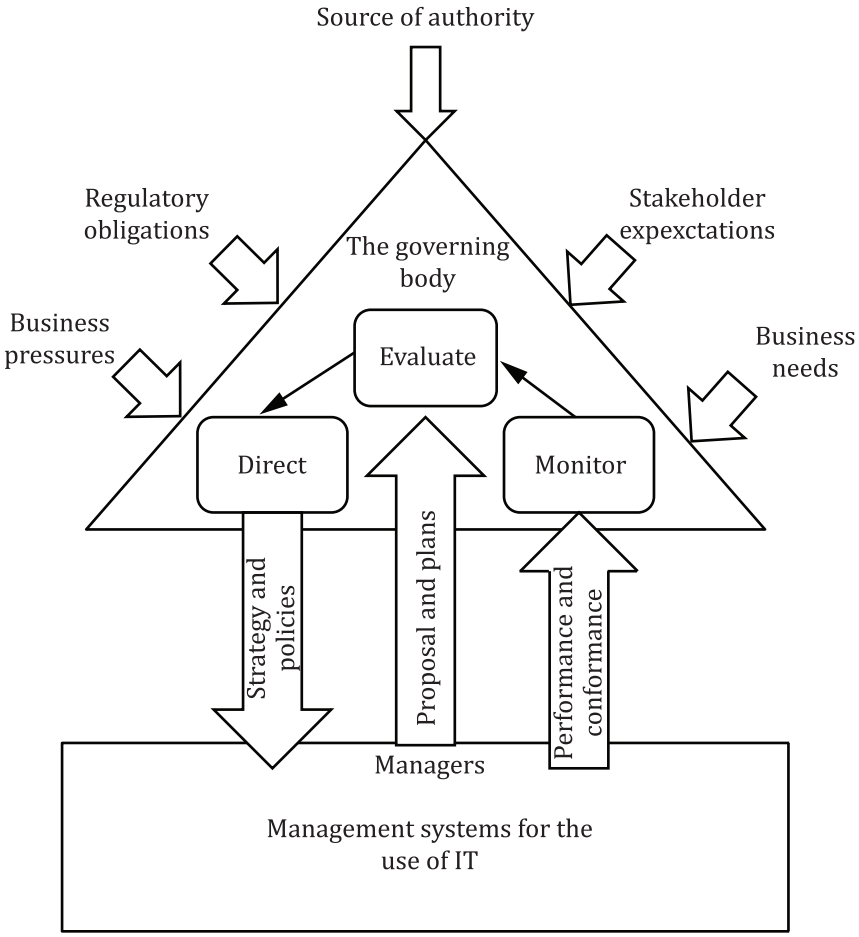

Figure A.1 — Model for governance of IT (from ISO/IEC 38500:2015)

A.1.3 Accountability

The governing body of an organization is responsible for the decisions, actions and inactions of the whole organization and is accountable to the stakeholders of the organization. This accountability cannot be delegated.

The organization uses people, processes and technologies to achieve its purpose and, regardless of the resources utilized, the governing body remains accountable for the outcomes of the organization. For example, the governing body can be responsible and be held accountable for staff members' misuse of IT equipment where it is deemed that there was insufficient guidance, training, oversight and enforcement within the organization to prevent such misuse from occurring. The governing body can be similarly at risk as a result of insufficient governance of the use of AI.

The "acceptable use of AI" is determined by the stakeholders of the organization in the various contexts in which they operate and that "acceptable use" can vary across those contexts. It is through the process of stakeholder engagement that the organization can understand what "acceptable use" means and translate those requirements into organizational policies. Such policies will then reflect not only the applicable legal requirements, but also societal expectations.

Stakeholder engagement, which is a key principle of the governance of organizations, requires that the organization stays reasonably engaged with stakeholders to ensure the organization is creating benefit over time and doing so in a manner that is acceptable to those stakeholders. The relationship can require legal or moral obligations as well as alignment with the norms or expected behaviours of the society, economy and environment in which the organization operates.

© ISO/IEC 2022 – All rights reserved

A.1.4 Oversight

ISO 37000:2021, 6.4.1 states that the governing body should oversee the organization's performance to ensure that it meets the governing body's intentions for, and expectations of, the organization, its ethical behaviour and its compliance obligations.

Organizational policies reflect the applicable legal and societal obligations. They will also include policies that align and drive the organization to achieve its objectives. In order to remain accountable to its stakeholders, the governing body uses compliance to organizational policies as a means to retain effective control of the organization.

The governing body is also responsible for determining the nature and extent of the risks that the organization is prepared to face in achieving its goals. Once this risk appetite is determined, it oversees the appropriate management of that risk and any mitigation that is necessary to ensure that risk appetite is not exceeded.

Good oversight also ensures that the governing body provides adequate organizational capability to perform its tasks, protect the organization's and stakeholders' interests and highlight new opportunities.

A.2 Governing body guidance over management decisions

The governing body sets the purpose and other parameters for the organization and is accountable for the whole organization. The governing body decides what parameters to set and these will vary between organizations. Typically, such parameters can include (either formally or informally):

— Strategic decisions including markets to serve and products and services to provide.

— Stakeholder engagement to set investments, expected behaviours, culture and values.

— Risk appetite to outline the nature and extent of risks that the organization is willing to take on in the pursuit of its goals.

The role of management is to efficiently and effectively deliver the objectives of the organization within these parameters.

The differentiation and clarity of the governance and management roles are critical to the success of the organization. The roles involve different responsibilities and require different skill sets. Their separation gives them well-defined freedom for innovation and entrepreneurship. The interface between the roles brings the necessary controls in place to safeguard the organization. For example, if management sees new opportunities or risks that would be outside the strategy or risk appetite, they would propose a change to the relevant parameters. The governing body would then decide if a change in those parameters is in the best interest of the organization and its stakeholders.

Although each of these roles is distinct, multiple roles can be performed by the same person, provided that there is no conflict of interest between the roles and that the person is aware of the discrete responsibility that each role entails. For example, an executive manager can also be a member of the governing body. However, it is important that the roles remain distinct and that the person remains cognisant of the role that is being performed at any time.

A.3 Governance of data use

ISO/IEC 38505-1:2017, Clause 5 highlights the following data-specific aspects of governance:

— **Value.** Data are the raw material for useful information. Some data are potentially not useful, while other data are extremely valuable to the organization. However, this value is not known until it is used by the organization and therefore all data are of interest to the governing body that is ultimately accountable for it. The term "value" in this case also includes the quality and quantity

© ISO/IEC 2022 – All rights reserved

of the data, its timeliness, the context (which are, in themselves, data) and the cost of its storage, maintenance, use and disposal.

— **Risk.** Different classes of data and types of processing bring different levels of risk and the governing body should understand the risks of data and how to direct managers to manage these risks. The risks not only manifest themselves in data breaches, but also in the misuse of data and any reputational risks that arise therefrom.

— **Constraints.** Most data come with constraints on their use. Some of these are imposed externally on the organization through legislation, regulation or contractual obligations and include issues of, e.g. privacy, copyright, commercial interests. Other considerations include ethical or societal obligations, or organizational policies that restrict the use of the data. Strategies and policies are required to account for these constraints in any use of the data by the organization.

© ISO/IEC 2022 – All rights reserved

Bibliography

[1] ISO Guide 73:2009, *Risk management — Vocabulary*

[2] ISO/IEC 23894: —, *Information technology — Artificial intelligence — Risk management*[7)]

[3] ISO/IEC TR 24027, *Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making*

[4] ISO/IEC TR 24028, *Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence*

[5] ISO/IEC TR 24029-1, *Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview*

[6] ISO/IEC TR 24030, *Information technology — Artificial intelligence (AI) — Use cases*

[7] ISO/IEC TR 24368, *Information technology — Artificial intelligence — Overview of ethical and societal concerns*[8)]

[8] ISO 26000:2010, *Guidance on social responsibility*

[9] ISO 31000:2018, *Risk management — Guidelines*

[10] ISO 31050, *Guidance for managing emerging risks to enhance resilience*[9)]

[11] ISO 37000:2021, *Governance of organizations — Guidance*

[12] ISO 37301:2021, *Compliance management systems — Requirements with guidance for use*

[13] ISO/IEC TR 38502:2017, *Information technology — Governance of IT — Framework and model*

[14] ISO/IEC 38505-1:2017, *Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data*

[15] ISO/IEC TR 38505-2, *Information technology — Governance of IT — Governance of data — Part 2: Implications of ISO/IEC 38505-1 for data management*

[16] ISO/IEC TS 38505-3, *Information technology — Governance of IT — Governance of data — Part 3: Data Classification*

[17] ISO/IEC 38506, *Information technology — Governance of IT — Application of ISO/IEC 38500 to the governance of IT enabled investments*

[18] Secretariat J.C. *Social Principles of Human-Centric AI*, March 2019 [viewed January 2021]. Available from: https://www8.cao.go.jp/cstp/english/humancentricai.pdf

[19] OECD *The OECD Principles on AI*, May 2019 [viewed January 2021]. Available from https://www.oecd.org/going-digital/ai/principles/

[20] European Commission. *Ethics Guidelines for Trustworthy Artificial Intelligence (AI)*, April 2019 viewed January 2021]. Available from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

7) Under preparation. Stage at the time of publication: ISO/IEC DIS 23894:2022.

8) Under preparation. Stage at the time of publication: ISO/IEC DTR 24368.

9) Under preparation. Stage at the time of publication: ISO/CD 31050.

© ISO/IEC 2022 – All rights reserved

[21] Government of Singapore. *Model Artificial Intelligence Governance Framework, Second Edition*, January 2020 [viewed January 2021]. Available from https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf

[22] United Nations *Universal Declaration of Human Rights* [viewed October 2021]. Available from https://www.un.org/en/about-us/universal-declaration-of-human-rights

[23] United Nations *Johannesburg Declaration on Sustainable Development, World Summit on Sustainable Development*, 2002 [viewed October 2021]. Available. at http://www.un-documents.net/jburgdec.htm

© ISO/IEC 2022 – All rights reserved

ICS 35.020

Price based on 28 pages

© ISO/IEC 2022 – All rights reserved