ARTIFICIAL INTELLIGENCE

Law and Regulation

SECOND EDITION

CHARLES KERRIGAN

EE

Published by

Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

Authorised representative in the EU for GPSR queries only: Easy Access System Europe – Mustamäe tee 50, 10621 Tallinn, Estonia, gpsr.requests@easproject.com

A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2025946966

This book is available electronically in the Elgaronline Law subject collection
https://doi.org/10.4337/9781035334353

MIX
Paper | Supporting responsible forestry
FSC® C013604

ISBN 978 1 0353 3434 6 (cased)
ISBN 978 1 0353 3435 3 (eBook)
ISBN 978 1 0353 8024 4 (ePub)

Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY

CONTENTS

EXTENDED CONTENTS

9

ARTIFICIAL INTELLIGENCE AND STANDARDS

Sam De Silva and Barbara Zapisetskaya

A. INTRODUCTION

9.001 Standards have a crucial role to play in the information communications and telecommunications (ICT) sector generally, and in particular in relation to AI. ICT standards are essential in achieving interoperability of new technologies and can bring significant benefits to both industry, consumers and society. The development of ICT standards is integral to AI technologies and systems, in that these standards address inter

alia the capture, storage, retrieval, processing, display, representation, security, privacy and interchange of data and information.[1]

A standard is a document established by consensus and approved by a recognised body. It provides rules, guidelines or characteristics for activities or their results so that they can be repeated. The aim is to achieve the greatest degree of order in a given context.[2] In addition, according to NIST, standards allow technology to work seamlessly and establish trust so that markets can operate smoothly. They: **9.002**

- provide a common language to measure and evaluate performance;
- make interoperability of components made by different companies possible; and
- protect consumers by ensuring safety, durability, and market equity.[3]

The European Union (EU) in particular has acknowledged the critical role of standards in transformative technologies with broad societal implications, such as AI. Over the past few years, the EU and its member countries have been proactive in establishing a unified European strategy for AI governance and standardisation. The EU has consistently utilised the dynamic relationship between regulation and standards as a key component of both its internal market and international commerce, particularly in relation to AI. Therefore, the EU's recognition of this interplay, along with the ongoing efforts to revamp European standardisation in conjunction with the EU AI Act, should be viewed as essential to the future governance of AI, both within the EU and on a global scale. **9.003**

In this chapter we will consider how standards are developed, the two main types of standards and the role of standardisation in the implementation of the AI Act. We provide an overview of three international AI standards which relate to AI risk management, AI quality management and AI governance and explore how two of these standards could be useful in the implementation of the AI Act. **9.004**

B. HOW STANDARDS ARE DEVELOPED AND THE ROLE OF NATIONAL STANDARDS' BODIES

Standards are developed at national, European and international level through designated bodies. Standards are developed at international level by the ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) and at European level by the CEN (European Committee for **9.005**

1 https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf.
2 https://www.gov.uk/guidance/standardisation.
3 https://www.nist.gov/services-resources/standards-and-measurements.

Standardization), CENELEC (European Committee for Electrotechnical Standardization) and ETSI (European Telecommunications Standards Institute), referred to collectively as the European Standardisation Organisations (ESOs).

9.006 Standards are developed by bringing together all interested parties such as manufacturers, consumers, regulators, researchers and experts regarding a particular material, product, process or service. Standardisation is the process through which requirements and recommendations for products, processes and services are developed by experts and agreed upon through a consensus-based mechanism.

9.007 National standards bodies (such as the British Standards Institute (BSI),[4] for the United Kingdom) are usually members of both European and international standards organisations, including the CEN, CENELEC, ISO and IEC, as well as ETSI. To engage with and influence the work of these international bodies, the relevant national body forms 'national mirror committees' composed of national experts. These committees align with their counterparts at ISO, IEC, CEN and CENELEC levels, tracking and participating in international standardisation efforts and recording national positions on work items.

1. Relationship between European and international standards

9.008 The ongoing collaboration between the CEN and ISO, as well as CENELEC and IEC, facilitates a degree of alignment between European and international standards. International standardisation is prioritised whenever feasible to harness the advantages of global standards in fostering international trade and market harmonisation. Due to existing agreements between the ISO and CEN and the IEC and CENELEC, certain internationally developed standards may be adopted as European standards by these entities. It needs to be seen how this alignment will continue to evolve in the context of the AI Act once the European standards to support its implementation are developed. We discuss potential areas of divergence in further detail below.

2. International AI standards development

9.009 In a world that is increasingly interconnected, the development of international AI standards is crucial for fostering a unified global perspective on reliable AI systems and enhancing public confidence in AI technologies. The surge of interest and activity in AI in recent years has underscored the need for a consistent set of international AI standards. To address this global requirement, the ISO and IEC have established a joint standardisation committee for AI, known as ISO/IEC JTC 1/SC 42 (Subcommittee

4 Refer to the section on BSI in the AI Ethics chapter for more details.

42),[5] which is actively engaged in AI and big data. At the time of writing, its working groups are in the process of creating standards that cover various facets of AI.

The Institute of Electrical & Electronics Engineers (IEEE) is also active in AI standardisation, especially regarding autonomous and intelligent systems. At the time of writing, the IEEE is in the process of developing its 7000 Series, which addresses ethical considerations across a wide array of issues related to autonomous and intelligent systems.[6] These include aspects such as transparency, privacy, algorithmic bias, personal data categories, the creation of machine-readable privacy terms for individuals and the ethical ramifications of simulated empathy in AI systems. 9.010

C. HOW STANDARDS WORK

1. Overview

Standards were initially established to ensure safety, quality and interoperability. For businesses aiming to operate on a global scale, international standards are preferred over national or regional ones, as they provide a uniform framework for market operations worldwide. Historically, older standards primarily focused on interoperability as the main societal benefit. However, many modern standards also encompass social, economic and political goals or consequences, thereby setting contemporary norms. 9.011

General standards can encompass a variety of forms, from best practice documents and deployment guidance to specifications for interoperability at different levels, such as physical, network or application. Generally, compliance with these standards is voluntary. 9.012

The other type of standard is a management system standard (MSS). The ISO defines an MSS as the way in which an organisation manages the interrelated parts of its business in order to achieve its objectives. These objectives can relate to a number of different topics, including product or service quality, operational efficiency, environmental performance, health and safety in the workplace and many more. Organisations can get their MSS certified by a third-party auditor. 9.013

2. How do standards work for AI?

Standardisation for AI is still in the early stages. In Europe, organisations such as ETSI and CENELEC have set out ambitious agendas for standardisation, partly 9.014

5 Refer to the section on ISO in the AI Ethics chapter for more details.
6 Refer to the section on IEEE in the AI Ethics chapter for more details.

driven by the AI Act's framework for standards. ETSI has been concentrating on security aspects of AI and machine learning, while CENELEC has been focusing on trustworthiness and ethics. As mentioned above, European-focused standardisation bodies are particularly significant to both the European Commission and the AI Act because they are the only entities capable of developing harmonised standards.

9.015 In addition to security, trust and ethics, the standardisation of AI is poised to significantly influence the cross-sector application of AI. Currently, many AI solutions are sector-specific, such as those in healthcare or intelligent transportation. The ability to utilise large datasets across different sectors is expected to spur even greater innovation in AI. The safe and ethical exchange of data between sectors is anticipated to be a major advancement in AI. Standards that ensure the safe, ethical, efficient and reliable transfer of information will be among the most impactful developments for AI in the years ahead. ETSI is currently focusing on workstreams related to the data supply chain and the availability of training data, while the ISO is engaged in several projects concerning AI and big data.

D. THE ROLE THAT THE AI ACT GIVES TO STANDARDISATION

9.016 As described in detail in Chapter 8, the AI Act categorises three main categories of risk relevant to AI systems that are within scope: (i) unacceptable risk, (ii) high risk and (iii) low risk. The majority of the AI Act establishes the framework that applies to high-risk systems.

9.017 AI systems deemed high-risk are considered to meet the AI Act's requirements if they conform to harmonised standards published in the Official Journal of the European Union (OJEU) (Article 40) or to common specifications set by the European Commission (Article 41). In essence, compliance to established standards or common specifications significantly reduces the compliance obligations for providers of high-risk AI under the AI Act. In certain cases, meeting harmonised standards or common specifications allows for conformity assessment through internal control instead of external evaluation.

E. THE COMMISSION'S STANDARDISATION REQUEST

9.018 The AI Act includes provisions for the use of harmonised standards to comply with the requirements of the AI Act. The Commission's request for standardisation sets the stage for the development of these standards, guiding the creation of technical specifications that align with the EU's regulatory objectives and ensuring that high-risk AI systems meet the necessary safety, quality and ethical standards.

9.019 In December 2022, the European Commission published a draft standardisation request to the CEN and CENELEC in support of the 'key technical areas covered' by the EU AI Act. This request was finally adopted by the European Commission in May 2023 (Standardisation Request) and accepted by the CEN and CENELEC.[7] The CEN and CENELEC are the primary recipients of the Commission Request, with ETSI also mentioned as a potential contributor. Under the Standardisation Request the following new European standards (European Standards) or (Deliverables) that address the following areas are required to be developed, each of which are referred to as Standardisation Requests (SR) numbered from 1 to 10:

- SR1: Risk management system for AI systems: These ESDs will specify the requirements for a risk management system for AI systems. The aim is to establish a continuous iterative process throughout the AI system's lifecycle that prevents or minimises risks to health, safety or fundamental rights. The requirement is for ESDs to be drafted in such a way as to enable usability by relevant operators and market surveillance authorities.
- SR2: Governance and quality of datasets used to build AI systems: These ESDs will include specifications for adequate data governance and data management procedures to be implemented by AI system providers. They will focus on data generation and collection, data preparation operations, addressing biases and ensuring the quality of datasets used to train, validate and test AI systems.
- SR3: Record keeping through logging capabilities by AI systems: These ESDs will define specifications for automatic logging of events by AI systems. The aim is to enable traceability throughout the system's lifecycle, monitor operations and facilitate post-market monitoring by providers.
- SR4: Transparency and information provisions to the users of AI systems: These ESDs will provide specifications for design and development solutions that ensure transparency of AI system operations, enabling users to understand the system's output and use it appropriately. They will also include instructions for use, including system capabilities and limitations, as well as maintenance and care measures.
- SR5: Human oversight of AI systems: These ESDs will specify measures and procedures for human oversight of AI systems. Providers will be required to identify and build these measures into the system before placing it on the market or putting it into service. Users should also be able to implement appropriate oversight measures. These ESDs shall also establish, where appropriate, appropriate oversight measures that are specific to certain AI systems in consideration of their intended purpose. With respect to AI systems intended for remote biometric identification of persons, human oversight measures will need to provide for the possibility that no action or decision is taken by the user on the basis of the identification resulting from the system unless this has been separately verified and confirmed by at least two natural persons.
- SR6: Accuracy specifications for AI systems: These ESDs will outline specifications to ensure an appropriate level of accuracy for AI systems. For the purpose of these ESDs, 'accuracy' shall be understood as referring to the capability of the AI system to perform

[7] https://ec.europa.eu/growth/tools-databases/enorm/mandate/593_en.

the task for which it has been designed. This should not be confused with the narrower definition of statistical accuracy, which is one of several possible metrics for evaluating the performance of AI systems. Providers will be able to declare relevant accuracy metrics and levels, and appropriate tools and metrics will be defined to measure accuracy against defined levels.

- Providers will be able to declare relevant accuracy metrics and levels, and appropriate tools and metrics will be defined to measure accuracy against defined levels.
- SR7: Robustness specifications for AI systems: These ESDs will lay down specifications for the robustness of AI systems, taking into account sources of errors, faults, inconsistencies and interactions with the environment. They will also consider AI systems that continue to learn after being deployed.
- SR8: Cybersecurity specifications for AI systems: These ESDs will provide suitable organisational and technical solutions, to ensure that AI systems are resilient against attempts to alter their use, behaviour or performance or to compromise their security properties by malicious third parties exploiting the systems' vulnerabilities. Organisational and technical solutions shall therefore include, where appropriate, measures to prevent and control cyberattacks trying to manipulate AI-specific assets, such as training datasets (for example, data poisoning) or trained models (for example, adversarial examples) or trying to exploit vulnerabilities in an AI system's digital assets or the underlying ICT infrastructure.
- SR9: Quality management system: These ESDs will be drafted for providers of AI systems to be implemented within their organisation, with particular consideration given to small and medium-sized organisations. Specifications shall be drafted such that the quality management system aspects related to the AI system may be integrated into the overall management system of the provider.
- SR10: Conformity assessment for AI systems: These ESDs will provide procedures and processes for conformity assessment activities related to AI systems and the quality management system of AI providers. They will also include criteria for assessing the competence of individuals involved in conformity assessment activities, considering scenarios where the assessment is carried out by the provider or a professional external third-party organisation.

9.020 According to the Standardisation Request, while it is not always possible to develop ESDs that consider each specific intended purpose, such ESDs shall cover at the minimum a range of technical solutions and options, which the product manufacturer can assess and implement, taking into consideration the intended purpose of their specific system. ESDs shall also possibly include guidance on how such assessment and implementation of solutions and options shall be executed.

9.021 All ten SRs must be completed by 30 April 2025. However, this is a very challenging timeframe given that on average it takes at least 3–4 years to develop a standard.

1. Alignment with the AI Act

Several of the ten areas identified in the Standardisation Request are covered (in part) by international standards that are either published or currently under development. For example, at the time of writing, there are standards addressing risk management (ISO/IEC 23894), data quality and governance (ISO/IEC CD 5259 series), transparency (IEEE 7001, ISO/IEC AWI 12792), human oversight (ISO/IEC 8200, ISO/IEC PWI 18966), accuracy (ISO/IEC TS 4213:2022), robustness (ISO/IEC TR 24029–1, ISO/IEC DIS 24029–2) and cybersecurity (ETSI SAI series). Other standards, although not directly related to the ten areas mentioned in the Standardisation Request, are relevant to some of the requirements of the AI Act, for example, standards on bias (ISO/IEC AWI TS 12791, ISO/IEC TR 24027, P7003), testing (ISO/IEC TR 29119–11) and verification and validation (ISO/IEC TS 17847). **9.022**

We will now focus on two key standards relevant to the Standardisation Request: (1) the standard addressing risk management (ISO/IEC 23894); and (2) the AI management system standard (ISO/IEC 42001). **9.023**

F. RISK MANAGEMENT – ARTICLE 9 AND AI RISK MANAGEMENT STANDARD (ISO/IEC 23894)

Article 9 of the AI Act mandates the creation and maintenance of a risk management system (RMS) for high-risk AI systems throughout their lifecycle. This RMS should be a systematic process that includes identifying and analysing known and foreseeable risks, estimating and evaluating risks from intended use and potential misuse, and considering additional risks from post-market monitoring. After this analysis, appropriate risk management measures must be implemented. **9.024**

The objective of the RMS as required by Article 9 of the AI Act is to ensure that the overall residual risk associated with high-risk AI systems is at an acceptable level. This residual risk should be communicated to users. In implementing risk management measures, priority should be given to eliminating or mitigating risks through thoughtful design and development. For risks that cannot be eliminated, measures should be put in place to control them, and users should be provided with sufficient information and training. **9.025**

Article 9 of the AI Act also stipulates that high-risk AI systems must undergo testing to identify appropriate risk management measures and to verify their consistent performance. The testing procedures should be tailored to the AI system's intended purpose and conducted at appropriate stages during the development process, prior to **9.026**

market release. This ensures that the high-risk AI systems are evaluated for safety and efficacy before they are made available for use.

9.027 Currently, the only existing standard relating to AI risk management is ISO/IEC 23894 (the AI Risk Management Standard). The AI Risk Management Standard builds upon the ISO 31000:2018 (the 'General Risk Management Standard') which provides general guidance on risk assessment. The General Risk Management Standard describes: (i) the underlying principles of risk management (integrated, inclusive, continual improvement, structured and comprehensive risk management); (ii) how risk management frameworks should be integrated into significant activities and functions of an organisation; and (iii) how risk assessment processes and practices help to identify risk and ways to manage risk (as more fully discussed below). Similar to the General Risk Management Standard, the AI Risk Management Standard is not an MSS.

9.028 The AI Risk Management Standard uses the General Risk Management Standard as its base but goes further by suggesting that organisations that develop, deploy or use AI products, systems and services need to manage specific risks relating to this technology and consider the context of AI in an organisation. It is not intended for the specific risk management of products and services using AI for objectives such as safety and security.

9.029 The AI Risk Management Standard states that AI systems can introduce new or emergent risks for an organisation, with positive or negative consequences on objectives, or changes in the likelihood of existing risks. The following principles have been incorporated into the AI Risk Management Standard:

- Inclusivity of stakeholders: As the use of AI systems can result in engagement with multiple stakeholders, organisations should seek dialogue with diverse internal and external groups, both to communicate harms and benefits and to incorporate feedback and awareness in the risk management process. Input from stakeholders will be beneficial for machine learning use cases, and generally for automated decision-making processes and ensuring overall transparency and explainability of AI systems.
- Dynamic risk management: As AI systems are dynamic and require continuous learning, refining and validating, legal and regulatory requirements related to AI need to be frequently updated. Organisations should seek to understand how AI will be integrated with management systems, and how it will impact their environmental footprints, health and safety and legal or corporate responsibilities.
- Best available information: As AI impacts the way individuals interact with and react to technology, it is advisable for organisations to retain information regarding ongoing use of AI systems throughout the entire lifetime of the AI system.
- Human and cultural factors: Human behaviour and culture significantly influence all aspects of risk management at each level and stage. Organisations engaged in the design, development or deployment of AI systems, or any combination of these, should monitor their evolving cultural landscape. Organisations should focus particularly on effects of AI systems or their components on privacy, freedom of expression, fairness, safety, security, employment and environment, and more generally on human rights. Biases in decision-making are overlooked without human interpretation.
- Continual improvement: The identification of previously unknown risks related to the use of AI systems should be considered in the continual improvement process. Organisations engaged in the design, development or deployment of AI systems or system components, or any combination of these, should monitor the AI ecosystem for performance successes, shortcomings and lessons learned, and maintain awareness of new AI research findings and techniques.

1. Gaps with Article 9

9.030 While the AI Risk Management Standard appears to incorporate certain elements of Article 9 and the related Standardisation Request, it unfortunately omits other elements. The AI Risk Management Standard aligns with Article 9 by ensuring that risk management is an iterative process throughout the AI system's lifecycle, thoroughly documented and aimed at mitigating risk to acceptable levels. However, it does not require the communication of residual risks to users or address risks associated with foreseeable misuse. The AI Risk Management Standard also does not specify particular testing procedures for identifying risk management measures, nor does it set metrics or probabilistic thresholds for testing.

9.031 The AI Risk Management Standard diverges from Article 9 of the AI Act in its underlying concept of risk. The AI Risk Management Standard emphasises organisational risks, which are uncertainties that could affect an organisation's goals, whereas the AI Act views risk as potential harm to individuals' health, safety and fundamental rights. This difference limits the AI Risk Management Standard's effectiveness in implementing the AI Act's RMS requirements.

9.032 One of the key limitations is that the AI Risk Management Standard primarily offers recommendations rather than explicit requirements. In the field of standardisation, the distinction between requirements, typically indicated by 'shall', and recommendations, indicated by 'should', is critical because only standards with requirements can be utilised for conformity assessment. Standards that support the AI Act must be suitable for conformity assessment to verify that an AI system complies with the AI Act's requirements. However, since compliance with the AI Risk Management Standard cannot be demonstrated (as it is not an MSS), the AI Risk Management Standard as currently drafted unfortunately does not adequately operationalise the RMS requirements of the AI Act.

G. QUALITY MANAGEMENT – ARTICLE 17 AND AI MANAGEMENT SYSTEM STANDARD (ISO/IEC 42001)

9.033 Article 17 of the AI Act requires a quality management system to be put in place and be fully documented.

9.034 Specifically, compliance with Article 17 requires coverage of how to take into account the high-risk AI ecosystem by defining how to set up: (a) strategy for regulatory compliance; (b) technical specifications to be applied, including standards; (c) communication handling with national authorities; and (d) an accountability framework.

9.035 Article 17 also requires techniques and procedures for: (a) AI design, design verification and development; (b) AI examination, testing and validation; (c) data management; (d) quality control and assurance; (e) reporting of serious incidents and of malfunctioning; (f) record keeping of all relevant documentation and information; (g) conformity assessment; and (h) management of AI modifications.

9.036 Systems for risk management, post-marketing monitoring and resource management also need to be established.

9.037 While a single standard is very unlikely to provide the necessary technical specifications to comply with all requirements of Article 17, a standard that establishes a quality system encompassing all of the above-mentioned aspects is crucial for ensuring compliance with the AI Act. Such a standard would integrate these elements into a systematic quality management system which should facilitate both conformity assessment and post-market quality monitoring.

1. AI Management System Standard (ISO/IEC 42001)

9.038 The AI Management System Standard provides considerable overlap with Article 17 requirement and is a management system standard on the responsible development and use of AI and is developed in relation to a circular process of establishing, implementing, maintaining and continually improving AI systems. It is comparable with other management system standards such as ISO 9000 (Quality) and ISO 27001 (Information Security) and similar to ISO 9000 and ISO 27001. The AI Management System Standard is an MSS.

9.039 In essence, an MSS helps organisations to: (a) improve their performance by specifying repeatable steps that they can implement to achieve their goals and objectives; and (b) create an organisational culture that reflexively engages in a continuous cycle of self-evaluation, correction and improvement of operations and processes through heightened employee awareness and management leadership and commitment.

9.040 The AI Management System Standard focuses on: (a) AI systems that have the potential to change their behaviour through use, presenting a challenge to ensure continuing monitoring and compliance with rules and/or accepted practices; (b) AI systems involved in automatic decision-making (possibly in a non-explainable or transparent way), which require specific management beyond that of a traditional system; and (c) the replacement of human interaction with machine learning, insight and data analysis, which increases the opportunities for applying AI systems while also changing the way those systems are justified, developed or deployed.

9.041 The AI Management System Standard is an auditable and certifiable standard. Audits are a vital part of the management system approach as they enable organisations to check how well their achievements meet their objectives and show conformity to the standard.

2. Gaps with Article 17

9.042 The content in Annex B of the AI Management System Standard, which provides implementation guidelines for AI controls, is highly relevant to Article 17. However, as the AI Management System Standard aims to offer organisations as much flexibility as possible in selecting necessary controls, it may not fully align with the regulatory requirements of the AI Act. This is because the AI Management System Standard allows organisations to determine the relevance of controls and choose which to adopt and does not provide a comprehensive set of justifications for control selection, nor does it establish a connection to regulatory requirements.

9.043 While the AI Management System Standard covers a broad range of Article 17's requirements, it often lacks sufficient technical implementation details. As the AI Management System Standard primarily offers guidance rather than mandatory controls, additional specifications that provide more explicit methodologies and implementation requirements will be required. Such additional specifications could include AI design considerations, quality control and assurance, testing or data management to ensure comprehensive coverage and regulatory compliance.

9.044 Other current gaps with the AI Management System Standard when compared with Article 17 requirements include:

- Post-marketing monitoring system coverage: To enhance compliance with the AI Act, more detailed requirements tailored to the AI Act's concerns and risks are necessary, as AI Management System Standard currently only mandates minimal system and performance monitoring, repairs, updates and support. The existing considerations in the AI Management System Standard should be expanded to include specific mechanisms for monitoring the potential adverse effects of AI system operations on individuals and

society. Additionally, there is a need for measures to identify risks that were not detected during the initial risk assessment and early stages of the AI lifecycle, as well as to recognise risks in the context of continuously learning systems.

- Management of AI modifications: To provide a more detailed framework for managing modifications to the AI system, it is necessary to specify the actions required for continuous assessment and management of risks, particularly when significant changes are made to the AI system. This includes establishing clear requirements for ongoing monitoring and evaluation to ensure the AI system remains compliant with the AI Act throughout its lifecycle.
- Documentation needs: The AI Management System Standard requires the production of multiple documentation items, and there is a need for more clarity and consolidation of these documentation requirements to streamline the application of this standard.

H. INTERNATIONAL STANDARDS AND AI GOVERNANCE

9.045 Although the Standardisation Request does not expressly refer to AI governance in general (as SR2 is limited to governance of datasets), the market had already acknowledged the need for a clear governance framework for AI and accordingly ISO had developed ISO/IEC 38507 (the AI Governance Standard) in 2022. The remainder of this chapter explores the AI Governance Standard.

1. AI Governance Standard: scope

9.046 The AI Governance Standard seeks, in its own words, 'to provide guidance for the governing body of an organization that is using, or is considering the use of, artificial intelligence'.[8] In addition, the AI Governance Standard also provides guidance to:

- internal and external service providers (including consultants);
- executive managers;
- public authorities and policymakers;
- external businesses or technical specialists (including professional bodies, lawyers, accountants or other retail/industrial associations); and
- assessors and auditors.

9.047 The AI Governance Standard is not an MSS.

2. Accountability of governing bodies

9.048 According to the AI Governance Standard, given the potentially diverse applications of AI,[9] which can impact any single area or function of an organisation, responsibility

8 ISO/IEC 38507:2022, p.1.

9 The AI Governance Standard states that reference to 'AI' is intended to be understood to refer to a whole family of technologies and methods, and not to any specific technology, method or application and 'use of AI' is defined in the

and accountability for AI governance should largely sit with an organisation's governing body, as opposed to delegating this to either a lower body (such as managers) or to the AI itself. This is also important from an accountability perspective, as it ensures that accountability for AI governance sits alongside the governance of any other process, measure or tool operated by an organisation. As AI develops and takes on more human tasks and characteristics, governing bodies should not fall into the trap of 'anthropomorphising AI' and giving it human characteristics out of proportion to its actual role and capabilities.[10]

9.049 The AI Governance Standard states that as part of ensuring good and adequate AI accountability, organisations should ensure that they review their procedures and practices to ensure they are suitable for and adequately reflect their uses of AI. On a practical level, this could involve considering:

- Oversight: Does an organisation's use of AI correspond with its risk appetite and are all of the associated procedures (such as measurement, decision assurance and monitoring) fit for purpose? In light of the implications of using AI, the value of its use should also be captured.
- Direction: How does using AI advance the strategy or objectives of the organisation? Are existing organisational statements of values, codes of conduct or ethics appropriate for AI and is the level of resource allocated to AI consummate to its use?
- Reporting: Is there communication and reporting to an organisation's stakeholders on how AI is used?
- Evaluation: Considering how AI has worked (both internally and externally) and other factors such as future opportunities and risks, are the governance processes (and any decisions taken under them) effective and can they be improved going forward?

9.050 As part of the evaluation process, governing bodies need to ensure that they themselves are equipped to effectively govern the use or introduction of AI across their organisation. The AI Governance Standard states that it is advisable to:

- work with stakeholders to build an understanding of AI and address any concerns they may have; or
- review internal processes, timelines and criteria to establish their adequacy in responding to the evolution of AI and its uses. This could involve creating a specialised AI subcommittee or working with other pre-existing committees (such as on strategy) to establish how AI impacts their work.

broadest sense as developing or applying an AI system through any part of its lifecycle to fulfil objectives and create value for the organisation. This includes relationships with any party providing or using such systems.

10 At p.4.

3. Appropriate level of oversight of AI

9.051 The AI Governance Standard acknowledges that the differing uses of AI and its enduring complexity dictate that organisations should take various factors into account when deciding the appropriate level of oversight, including:

- the type of AI used and the purpose of using it;
- the envisaged risks and benefits associated with the AI system;
- the functional layer of the AI ecosystem used;
- the role played by the organisation in the AI value chain (such as a customer, producer or provider) and the stage of implementation of the AI system.[11]

9.052 While the majority of the above would initially be considered by governing bodies at the initial stage of introducing and using AI, it is important to ensure that AI is governed appropriately and effectively throughout its lifecycle.

4. Practical steps organisations can take to alleviate constraints on the use of AI

9.053 The AI Governance Standard emphasises the importance of not only appreciating the benefits of AI, but also understanding the potential risks and constraints it can introduce. To mitigate such issues, organisations can:

(a) Increase oversight of compliance

9.054 Governance oversight within organisations should be based on policies set by the organisation and should identify effective individual and collective accountability in an appropriate chain of responsibility, which is set alongside the context of use of AI. This includes putting policies in place to make sure that AI is used appropriately, that there is sufficient human oversight in place and that any persons using AI are properly trained and know how to raise concerns. Legal requirements or obligations may be determined for using such technologies alongside the risk appetite of the organisation. The AI Governance Standard notes that governing bodies should be aware of new sources of risk posed by AI technologies, including unwanted bias, cyber-threats and a lack of AI expertise.

(b) Address the scope of AI use

9.055 This principle involves considering the formulation of relevant assumptions on data and conducting a prior assessment on the availability, quality, quantity and suitability of data and an examination of potential biases. Formulating a description of the AI system, by way of its algorithms, data and models, would assist in ensuring the AI

11 At p.17.

technology is being deployed for its intended use.[12] The foregoing principle needs to be considered in parallel with the importance of governance of data use, also addressed in the AI Governance Standard, when acknowledging that data is being used for the correct purpose and sensitive data is protected and secured.

(c) Assess and address the impact on stakeholders

9.056 The AI Governance Standard notes that the governing body is responsible outside of the context of AI for shaping and defining the organisation's desired culture, which has an impact on stakeholders connected to the organisation. It also notes the human impact on an organisation's culture and values, which are implicitly embedded in the behaviour of staff, and advocates for human involvement to a degree in the AI process, ensuring that AI systems can be monitored and corrected when needed. Conversely, the AI Governance Standard also highlights that the AI system can itself identify where human decision-making is flawed through bias and discrimination. A 'Cultures and Values Board' or an 'Ethics Review Board' might be set up to supervise the impact of AI systems and make sure they are aligned to an organisation's values and culture.[13]

(d) Determine legal requirements or obligations of using the technology

9.057 As AI and other technologies often used together with it (such as machine learning) develop, they will likely be subject to increasing regulation. Practical examples of this are facial recognition software or automated vehicles. Governing bodies need to ensure that as new legislation or other requirements are introduced, their use of AI remains compliant.[14]

(e) Align the use of AI to the organisation's objectives, culture and values

9.058 The use of AI will likely require an assessment of its impact on an organisation's objectives and strategy, whether it be to drive innovation and development or cut costs. Organisations should also take into account the effect of AI on their culture and values and take appropriate measures to safeguard those, if necessary.[15]

(f) Ensure that problem solving takes due account of context

9.059 Where using AI to solve problems, organisations should ensure that none of the context that would usually be apparent to a human (such as culture, values or behaviour) is lost, lest this lack of background affect the AI's output.[16]

12 At p.10.
13 At p.10.
14 At p.11.
15 At p.11.
16 At p.11.

(g) Examine the additional risk that the use of AI can bring to an organisation

9.060 Organisations should ensure that their use of AI remains within their appetite for risk and that appropriate processes and controls are in place to manage that risk.[17]

I. CONCLUSION

9.061 Given the transformative impact of AI, similar to the advent of computers, AI standards play a pivotal role in promoting its broad acceptance and adoption as a technology that is trustworthy, ethical and safe. These standards facilitate the integration of AI by addressing critical concerns such as safety, fairness, reliability, accountability and transparency, thereby supporting the ethical principles and guidelines that are essential for its widespread adoption. The three AI standards discussed in this chapter in relation to risk management (ISO/IEC 23894), AI management systems (ISO/IEC 42001) and AI governance (ISO/IEC 38507) should contribute to the practical and meaningful application of these ethical principles, fostering the adoption of AI on a global scale.

9.062 Standards are going to play a very important part in the implementation of the AI Act, and it remains to be seen to what extent international standards on AI will play a role in creation of the European standards on AI.

17 At p.11.

10

COMMERCIAL CONTRACTS

Iain Sheridan

A. INTRODUCTION

Pre-contract due diligence requires not only precise drafting of agreed terms but also the foresight to accommodate worst-case scenarios. Commercial contracts across diverse sectors, especially those involving high-value subject matter, require both types of expertise. This need is evidenced by English High Court decisions, which often reveal deficiencies in pre-contract due diligence and imperfect drafting. Additionally, **10.001**