

ARTIFICIAL INTELLIGENCE (AI) MANAGEMENT SYSTEM, KEY ELEMENTS

Biljana Markovic^{1, a} and Miljan Savic^{2, b}

¹ University in East Sarajevo, Faculty of Mechanical Engineering, East Sarajevo, RS, Bosnia and Herzegovina

² Institute for Standardization of Bosnia and Herzegovina, East Sarajevo, RS, Bosnia and Herzegovina

^abiljana.markovic@ues.rs.ba, ^bmiljan.savic@isbih.gov.ba

Abstract Artificial Intelligence (AI) has become an integral part of modern technologies, but its development brings a range of challenges related to security, ethics, and privacy. This paper explores the role of Artificial Intelligence management systems (AI) in ensuring the responsible application of these technologies. Special attention is given to the ISO/IEC 42001:2023 standard, the first global standard for managing AI systems, which provides a framework for risk management and achieving transparency and reliability in the application of AI. The standard enables organizations to establish clear guidelines for the development, implementation, and oversight of AI systems in accordance with ethical principles. This paper analyzes key aspects of this standard in detail, including risk assessment, documentation management, impact assessment on individuals and society, as well as the obligations of organizations regarding transparency and accountability.

Keywords: Artificial intelligence AI, application; management system; ISO/IEC 42001; Risk assessment; AI policy.

1. INTRODUCTION

The development of AI technology brings many advantages but also numerous challenges, especially in terms of risk management, transparency, and ethics. Artificial intelligence is now integrated into various industries, from healthcare to finance, necessitating the introduction of globally recognized standards for its management. AI enables the automation of processes, pattern recognition, decision-making, and data management in ways that were previously unimaginable. However, with the increasing adoption of AI technologies, there is a growing need for clear rules and standards to ensure that these technologies are used ethically and responsibly. ISO/IEC 42001:2023 is designed to enable organizations to implement, utilize, and monitor AI systems in a responsible and ethical manner by establishing clear rules and guidelines [1].

The ISO 42001 standard was published on December 18, 2023, to guide the organizations that design, develop, and deploy AI systems. The prime concerns are implementing the standards and maintaining transparency and accountability. In addition, it mitigates risk and ensures safety. ISO 42001 is a voluntary standard and is not legally binding. However, given its significance and emerging recognition, it is highly likely to become the benchmark for AI management systems in the future. Organizations should anticipate possible regulatory developments and consider proactively adopting ISO 42001.

2. APPLICATION OF AI

The application of Artificial Intelligence (AI) encompasses a wide range of sectors, bringing changes in the way organizations operate and function, addressing social problems, and how people interact with technology. Key sectors for the application of artificial intelligence include:

- *Healthcare:* Artificial intelligence has enormous potential in medicine, enabling better diagnostics, personalized treatments, and optimization of processes in healthcare institutions. Faster and more accurate diagnostics, tailored treatments for patients, and reduced errors in medicine are significant benefits.
- *Financial Sector:* AI is used in banking, insurance, and other financial sectors for data analysis, risk assessment, and automation of business processes. This includes faster processing of large volumes of data, improved risk assessment, and real-time fraud detection.
- *Automotive Industry:* Applications of artificial intelligence in the automotive industry are focused on establishing processes for autonomous vehicles. This leads to a reduction in human error while driving, increased road safety, and traffic optimization. Legal regulations and responsibilities in the event of accidents, security against hacking attacks on autonomous vehicles, and technical challenges in decision-making in complex traffic situations are also significant considerations.
- *Manufacturing and Logistics:* AI is used to optimize supply chains, forecast demand, and automate production lines through robots that can operate independently and make decisions. This results in reduced production costs, faster and more precise product delivery, and decreased production downtime.
- *Education:* AI brings innovations to the education sector by providing teachers and students with tools for personalized learning and improving educational processes. This includes personalized learning, faster and more efficient assessment, and support for teachers in monitoring student progress.
- *Smart Cities and Infrastructure:* AI is utilized to enhance infrastructure and manage resources in urban areas, making cities more efficient and sustainable. This involves more efficient resource use, reduced pollution, improved mobility, and decreased traffic congestion.
- *Media and Entertainment:* AI transforms the way content is created and distributed in the media industry. It allows for better personalization of the user experience, faster production and analysis of content, and prediction of media trends.

The development of AI technologies has enabled the creation of a wide range of products used across the aforementioned sectors and applications. These products often combine sophisticated algorithms and machine learning to provide intelligent functions that help users solve problems, automate processes, and improve productivity. Key products based on AI technologies include:

- *Virtual Assistants:* Virtual assistants use AI to enable user interaction through voice commands or text queries. They can perform tasks such as searching for information, managing calendars, setting reminders, controlling smart devices, and even making decisions.
- *AI in Autonomous Vehicles:* AI systems that enable autonomous vehicles use advanced sensors, cameras, and machine learning algorithms to detect and analyze the environment, identify objects, make decisions, and operate vehicles without drivers.

- *AI Tools for Data Analysis and Business Intelligence:* These products use AI to analyze large volumes of data to uncover patterns, predict trends, and generate useful insights for businesses. They often involve machine learning and deep learning to support decision-making and process optimization.
- *AI Tools for Image and Facial Recognition:* AI systems that use machine learning to recognize and analyze images or video content. These products are used in security applications, social networks, and data analysis from surveillance systems.
- *Chatbots and Virtual Agents:* AI chatbots use natural language processing (NLP) to communicate with users, answer their questions, provide support, and perform tasks automatically. These products are widely used in customer support and sales.
- *Generative AI Tools:* Generative AI systems use deep learning models to create new content—text, images, music, or even video—based on input data received from users.
- *AI for Natural Language Processing (NLP):* Tools that use AI for the analysis, processing, and generation of natural language. These systems help in processing large volumes of textual data, understanding semantics and context, as well as automatic translation and key information extraction.
- *AI Recommendation Systems:* AI algorithms are used to recommend content based on user preferences. These products assist in personalizing the user experience through recommendations for music, video content, products, and services.

3. AI MANAGEMENT SYSTEM ACCORDING TO THE REQUIREMENTS OF ISO/IEC 42001:2023

According to the ISO/IEC 42001:2023 standard, the AI management system is a set of interconnected elements that help organizations establish policies, objectives, and processes to achieve these goals in a responsible manner [1]. Key elements include an artificial intelligence policy, processes for identifying and treating risks, processes for assessing the impact on society and individuals, as well as processes for the continuous improvement of the system [3]. The following components are essential for effectively managing an artificial intelligence system (Figure 1). Some of the key requirements covered in the this standard include:

- *Leadership:* Top management should demonstrate leadership and commitment to the AI management system (AIMS) and establish policies and objectives that are consistent with the organization's strategic direction.
- *Planning:* Identify and assess risks and opportunities associated with AI and develop a plan to address them.
- *Support:* Provide resources and support for the AIMS, including training, awareness, and communication.
- *Operation:* Establish processes and procedures for the development, deployment, and maintenance of AI systems.
- *Performance evaluation:* Monitor, measure, analyze, and evaluate the performance of AI systems and take corrective actions when necessary.
- *Continual improvement:* Continually improve the AIMS, and ensure that it remains relevant and effective.

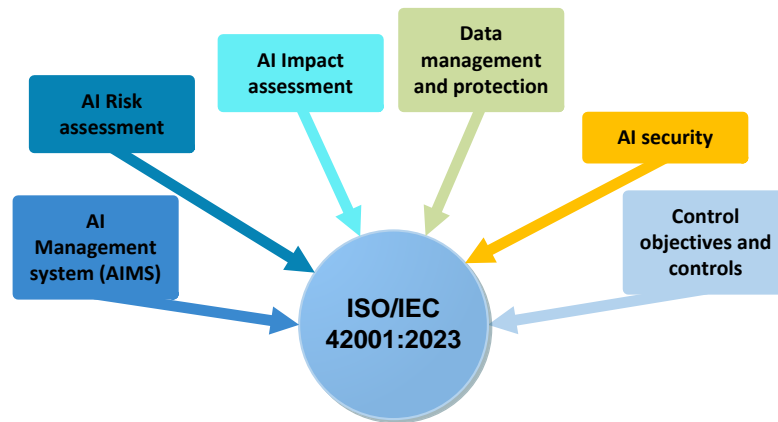


Figure 1. Key components of an AI management system.

3.1. AI Management System

The Artificial Intelligence (AI) management system refers to a structured approach to implementing, controlling, monitoring, and maintaining AI technologies within organizations [1]. Given the growing role of artificial intelligence across various industries, an effective management system is becoming increasingly important for achieving ethical, responsible, and transparent application of AI. This system enables organizations to ensure that AI technologies are used in a safe, legal, and reliable manner [3].

Top management within the organization shall establish an artificial intelligence policy that encompasses ethical principles, accountability, and a commitment to continuous improvement [1]. This policy should align with the organization's business strategies and legislative requirements.

Top management plays a crucial role in the implementation of the AI management system [1]. They shall ensure the provision of resources necessary for the system's operation, set AI objectives, and ensure that the AI policy is applied across all business processes.

3.2. AI Risk Assessment

According to ISO/IEC 42001:2023, risk assessment is a key component of managing artificial intelligence. The risk assessment process includes:

- Identifying risks that may prevent the organization from achieving its established AI objectives;
- Analyzing the potential consequences of identified risks on the organization, individuals, and society;
- Evaluating the likelihood of these risks materializing and prioritizing their treatment.

Risks are identified based on the context of the AI system application, which involves defining external and internal factors affecting the organization.

Based on the results of the risk assessment, the organization should implement a risk treatment process that includes selecting appropriate controls to minimize or eliminate the risks. Organizations can implement control measures or develop their own controls [1]. A key aspect of risk management is documenting all relevant steps and outcomes.

3.3. AI Impact Assessment

The impact assessment of artificial intelligence (AI) systems is a critical process within the management of AI technologies. This process helps organizations identify, analyze, and understand the potential consequences that the application of AI may have on people, society, businesses, and the environment [2]. Through this assessment, organizations can recognize risks and opportunities and make decisions that ensure responsible and ethical use of AI. The impact assessment process includes:

- Identifying possible scenarios in which an AI system may negatively impact individuals or groups;
- Analyzing the likelihood of these scenarios occurring and assessing their consequences;
- Providing recommendations for mitigating risks through technical and organizational measures (sections 6.1.4 and A.5.2).

The results of the impact assessment shall be documented and stored in accordance with the organization's data retention policies. Organizations are required to conduct impact assessments at regular intervals or when significant changes occur in the AI systems [2].

3.4. Data Management and Protection

Data management and protection are critical aspects of developing and implementing artificial intelligence (AI) systems. AI systems heavily rely on the data used for their development and training, operations, and continuous learning. Effective data management and protection within the AI management system ensure that data is properly collected, processed, protected, and used, while also safeguarding data integrity.

Within the ISO/IEC 42001:2023 standard, data management and protection play a crucial role in maintaining user trust, ensuring compliance with laws, and minimizing risks related to data privacy and security [1].

Data management in artificial intelligence systems refers to the processes of collecting, storing, processing, analyzing, and retaining data that are critical for the operation of AI technologies [1]. This data may include personal user information, business data, or publicly available information.

Data collection is the first step in managing AI systems. For the AI system to be effective, the data shall be of high quality, relevant, and diverse. When collecting data, organizations shall ensure that:

- Data is collected in accordance with privacy laws, such as the General Data Protection Regulation (GDPR);

- There is a clear legal basis for data collection, either through user consent or other legitimate interests of the organization;
- Data is collected in a transparent manner, informing users of the purpose of its use.

Data storage in AI systems shall be secure, providing adequate protection against unauthorized access, data breaches, and cyber attacks. The organization shall:

- Implement security measures such as data encryption, access controls, and activity monitoring;
- Establish data retention policies, defining how long data will be kept and when it will be deleted or archived.

Data protection is essential to avoid risks related to data breaches, unauthorized access, and infringement of user privacy. AI systems often utilize large amounts of sensitive data, making data protection crucial.

3.5. AI Security

The security of artificial intelligence (AI) is of paramount importance, given the application of AI systems in various sectors, including finance, healthcare, education, autonomous vehicles, and more. AI security refers to the technical, organizational, and legal measures that ensure AI systems operate safely, reliably, and in compliance with regulations, regardless of the complexity of the tasks they perform or the volume of data they process.

Since AI technology can be susceptible to manipulation, errors, or malicious attacks, key challenges in AI security include data protection, model integrity, prevention of manipulation, and ensuring that AI systems do not make harmful or erroneous decisions.

The European Union (EU) has recognized the enormous potential of artificial intelligence (AI) and its role in technological advancement, but it is also aware of the challenges and risks that AI presents. As a result, the EU has launched initiatives to regulate AI systems to ensure responsible, safe, and ethical development of this technology. The EU is particularly focused on protecting fundamental human rights, safety, transparency, and accountability in the application of AI technologies [4].

The most significant legal framework that the EU is developing to regulate artificial intelligence is the EU Artificial Intelligence Act (AI Act), proposed in 2021, which aims to become the first global legal framework regulating AI technologies [4].

We hope that the process of implementing AI technology will also be legally regulated in Bosnia and Herzegovina in the near future.

3.6. Control Objectives and Controls

Control Objectives and Controls are key elements for effective management of Artificial Intelligence (AI) systems. They ensure that organizations establish clear guidelines and procedures to achieve

desired outcomes and minimize potential risks. Control objectives and controls enable consistent, responsible, and safe use of AI technologies [1].

Objectives are important because they define the direction an organization aims to follow when using AI technologies. They guide the development of AI systems in line with business strategies, ethical principles, and legal requirements. Objectives also help monitor the performance of AI systems, allowing organizations to assess whether the system is achieving the expected results.

Based on the results of the risk assessment, organizations define a risk treatment for high-level risks and implement all necessary controls. These controls are measures the organization takes to ensure that objectives are met. By implementing controls, data protection is ensured, and data is used in a secure and legal manner. Transparency and accountability are also guaranteed, so that AI systems operate in accordance with the highest ethical standards. Figure 2 shows the basic controls that an organization can implement, and if additional requirements arise, the organization may introduce additional controls.

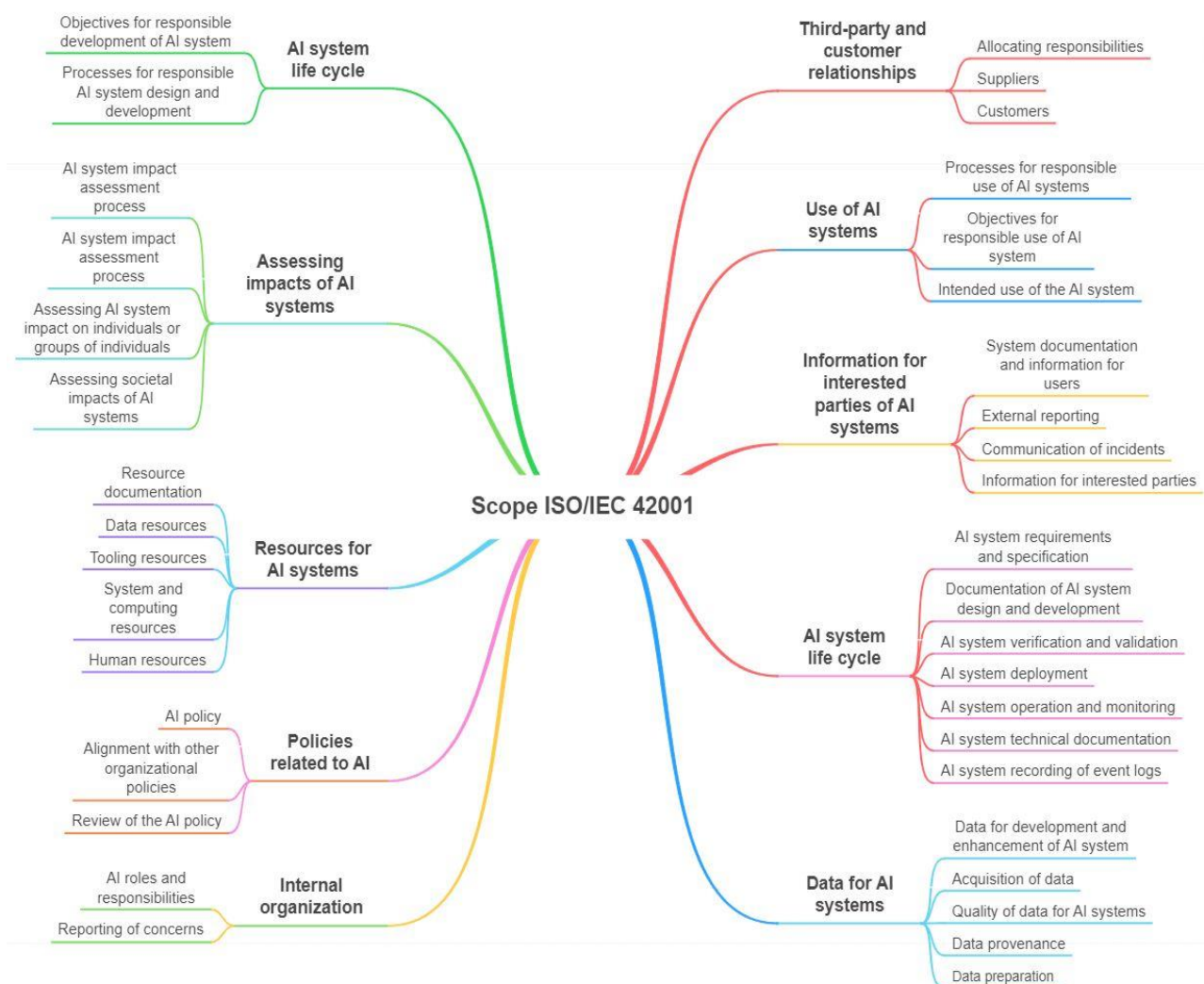


Figure 2. Basic controls.

3.7. Potential Challenges-dilemmas Versus Benefits

AI can analyze huge amounts of data in a short amount of time, enabling informed decisions that can improve business processes. Personalization is another key advantage: AI allows products and services to be tailored to the specific needs and preferences of users, thereby improving the user experience.

However, AI is not without its challenges. One of the main disadvantages is the potential for job losses, especially those that are routine and do not require special skills. Also, using large amounts of data to train AI models may raise concerns about user privacy. Ethical concerns are also important – the ethical implications of AI decisions need to be carefully considered, particularly in sectors such as healthcare and criminal justice. Finally, developing and implementing AI solutions can be expensive, which can be a barrier for small businesses.

In the end, the benefits of implementing ISO/IEC 42001 can be: the use of AI is inevitable, the standard enables control within a structured framework, supports compliance with legal and regulatory standards, enables the management of risks that are specific to AI, raises the visibility, recognition and reputation of the organization, encourages innovation among employees, and, above all, ensures responsible and ethically correct use of AI.

4. CONTINUOUS IMPROVEMENT OF AI MANAGEMENT SYSTEMS

Artificial intelligence management is directly related to risk management, as well as other processes related to the establishment of Industry 4.0. Therefore, before applying the requirements of the ISO/IEC 42001 standard, it is necessary to study the requirements of related standards such as ISO/IEC 22989, which establishes terminology for AI and describes concepts in the field of AI; ISO/IEC 23053, which establishes an AI and machine learning (ML) framework to describe a generic AI system using ML technology, and ISO/IEC 23894, which provides guidance on AI risk management for organizations.

In addition, ISO/IEC 42001 is a management system standard (MSS). Implementing this standard means establishing policies and procedures for the sound management of an organization related to AI, using a Plan-Do-Check-Act methodology, and as such provides value to any business or entity. This cycle allows organizations to:

- Plan AI objectives and processes to achieve them;
- Implement planned activities and control measures;
- Regularly check and measure the results of AI systems to determine whether they are achieving the desired goals;
- Conduct necessary corrective actions to improve the management system.

Continuous improvement is a fundamental principle of the management system, as AI technology is constantly evolving, requiring flexibility and adaptation of organizational management systems [1].

5. CONCLUSION

ISO/IEC 42001:2023 provides a comprehensive framework for managing artificial intelligence, enabling organizations to address the complex challenges that AI brings. A management system based on risk assessment, accountability, and transparency allows organizations to responsibly use AI technologies, reducing potential risks to society and individuals. As AI continues to develop, managing these technologies becomes increasingly important, and the ISO/IEC 42001 standard provides key guidelines for ethical and safe application. Focusing on quality, safety, and ethics can bring long-term benefits and contribute to the development of sustainable AI technologies.

AI-based products are becoming increasingly prevalent in everyday life and industries. The ISO/IEC 42001 standard is aligned with governmental initiatives such as the U.S. Risk Management Framework and the EU AI Act, reflecting the growing global emphasis on responsible and trustworthy artificial intelligence. The EU Artificial Intelligence Act is a legislative framework that lays the foundation for responsible, transparent, and safe application of AI in Europe [4]. This act provides a clear legal framework for all market players, from small businesses to large corporations, guiding the development of AI technologies in line with fundamental rights and freedoms of EU citizens. [4] The act also ensures that innovations in AI can continue with minimized risks to users and society as a whole.

In Bosnia and Herzegovina, there is a need to align the legal framework that regulates the development, application, and safety of AI technologies. Since Bosnia and Herzegovina is on the path to joining the European Union, one of the requirements in the coming period will be the alignment of the national legal framework with the EU Artificial Intelligence Act.

Organizations can be certified according to the requirements of the ISO/IEC 42001 standard. Certification under ISO/IEC 42001 represents an important step for organizations that wish to establish and demonstrate an effective AI management system. This process helps achieve high standards of quality and safety, as well as protect the basic rights and freedoms of users. Through proper preparation, implementation, and monitoring, organizations can ensure that their AI systems operate in accordance with best practices and ethical principles.

In Bosnia and Herzegovina, AI is slowly coming onto the scene, mainly through IT companies and academic institutions. Although it is in the initial stages, the application of AI in BiH has a huge potential for the improvement of various sectors such as industry in general, health, agriculture, production and public administration. By using AI technologies in the right way, the IT community of BiH can achieve a significant increase in efficiency, create new jobs and improve the quality of services.

Certification bodies in Bosnia and Herzegovina mostly come from the countries of the region and do not have adequate supervision by accreditation bodies, and market competition for issuing certificates leads to low quality of the certification process. Based on our own experience, and based on our participation and organization in a large number of audits by certification bodies, there are auditors in BiH who are not fully competent to check compliance with the requirements of international ISO standards. Therefore, it is bold to say, but very certain, as a result, in the future there will be certified

management systems in institutions and companies, which have a certificate, but do not adequately meet the requirements of the standard, especially the ISO/IEC 42001 standard.

This can pose a great danger to the control of the use of AI for positive purposes, for which there is a well-founded fear, already today.

References

- [1] ISO/IEC, 2023, Information technology — Artificial intelligence — Management system, ISO/IEC 42001:2023, International Organization for Standardization, Geneva.
- [2] ISO/IEC, under development, Information technology — Artificial intelligence — AI system impact assessment, ISO/IEC DIS 42005, International Organization for Standardization, Geneva.
- [3] International Organization for Standardization, 2024, Artificial intelligence, <https://www.iso.org/sectors/it-technologies/ai/>, [Accessed 5 September 2024].
- [4] The AI Act Explorer, 2024, Artificial Intelligence Act (Regulation (EU) 2024/1689), Official Journal version of 13 June 2024, Available: <https://artificialintelligenceact.eu/ai-act-explorer/> [Accessed 5 September 2024].