ARTIFICIAL INTELLIGENCE SECURITY THROUGH INPUT DATA QUALITY STEPS

ORIGINAL ARTICLE

GADELHA, Isaque Araujo[1]

GADELHA, Isaque Araujo. **Artificial Intelligence Security through Input Data Quality Steps.** Revista Científica Multidisciplinar Núcleo do Conhecimento. Year 08, Ed. 12, Vol. 02, pp. 24-38. December 2023. ISSN: 2448-0959, Acess link: https://www.nucleodoconhecimento.com.br/computer-science/input-data, DOI: 10.32749/nucleodoconhecimento.com.br/computer-science/input-data

ABSTRACT

In times of digital transformation, data generation constitutes a new type of asset, as it requires new forms of organizational learning, giving rise to new business models. The focus of all this transformation is Artificial Intelligence (AI), a subfield of Computer Science responsible for creating computational resources with capabilities similar to human reasoning for automated problem-solving. Elements of mathematics and engineering are used to reproduce aspects of human intelligence, as well as insights from other areas such as philosophy, mathematics, economics, neuroscience, psychology, computer engineering, control theory, cybernetics, and linguistics. Hence the name AI. Machines learn to speak, write, interpret data, and solve problems through AI, a tool that is now essential for Industry 4.0, as a transition from industrial society to the knowledge and digital economy. In turn, Generative AI uses multimodal tools to work with elements such as spoken language, images, sounds, body movements, etc. All this technology requires innovations in companies through the adoption of frameworks that enable the selection and storage of qualified data. This original article was developed based on extensive bibliographic research and scientific materials. Its objective is to demonstrate the relevance of data analysis as a way to apply compliance practices to AI. As a result of this research, there was a recognized need for greater attention to input data quality processes for generative AI training.

Keywords: Artificial Intelligence, Industry 4.0, Data Governance, Frameworks, Data Quality.

1. INTRODUCTION

With the intensity and speed of electronic interactions that permeate the information systems of organizations, these systems need to possess technological capacity to analyze and select this dense volume of data, grounded in continuous improvement and necessary innovations. This allows managers to maintain market competitiveness, driven by the new social capital, which is information (Molina and Santos, 2019).

Honório (2022, p.15) raises the issue of the importance of electronic data generation, demonstrating the speed of the changes society has been undergoing: "[...] the telephone took 75 years to reach 50 million people; radio, 38 years to reach the same number of people; television, 13 years; the internet, 4 years; the iPhone only 3 years; Instagram, 2 years; Angry Birds, 35 days, and Pokémon Go, 15 days." According to the author, wealth, once represented by physical assets, now also resides in the product of knowledge, as data is the new asset of organizations.

These examples contribute to demonstrating that innovations in business processes and actions to expand the competitive advantage of organizations require solutions offered by the "knowledge economy, digital economy, Industry 4.0, artificial intelligence, robotization, and clean technologies" (Lima, 2020, p.4).

Regarding the intense volume of data obtained through Artificial Intelligence, the security of this data and its proper storage require cultural changes within companies. The implementation of resources such as frameworks is necessary, allowing the obtained data to be transformed into manageable assets (Lima, 2020).

Generative AI, in turn, is a system that can capture valuable information that may go unnoticed by humans and offer real solutions (Moura, 2023). A recent example reported by the mainstream media was the case of an American boy who, having already seen 17 different doctors and specialists without satisfactory results for his

condition, had his mother search for the child's symptoms on ChatGPT. Comparing exams and symptoms, the AI provided a real diagnosis, identifying it as spina bifida (O Globo, 2023).

Instances like this demonstrate that the technological development of AI contributes to efficiency and productivity in solving situations of varying degrees of difficulty, whose effectiveness is enhanced by the proper analysis of data that is leading humanity down increasingly unexpected paths.

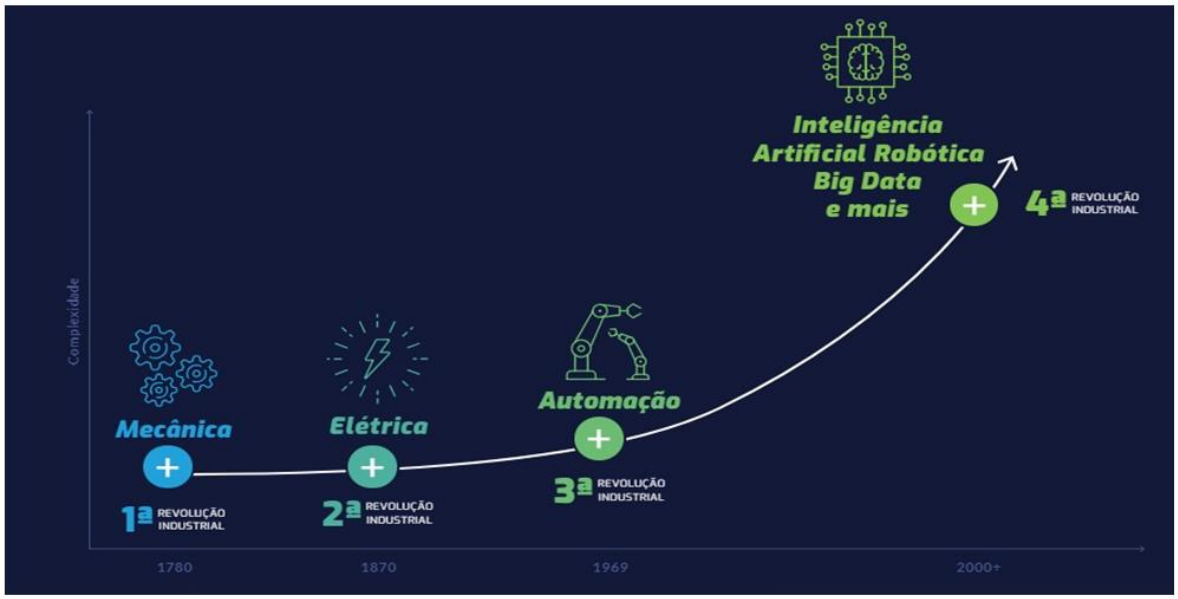
This original article was developed based on extensive bibliographic research and scientific data. Its objective is to demonstrate the relevance of data analysis as a way to apply compliance practices to AI. As a result of this research, the need for greater attention to input data quality processes for generative AI training was identified.

2. THE PHASES OF THE INDUSTRIAL REVOLUTION

Four major leaps in human and social development have occurred to date, detailed by Santos (2019) and Lopes (2019): the **1st Industrial Revolution** (1760 to 1840) brought about the invention of steam engines and mechanical power; the **2nd Industrial Revolution** (1850 to 1945) witnessed the advent of the electrical grid, chemical, petroleum, and steel industries; transformations in transportation (automobiles, trucks, and airplanes) and communication (telephone) also occurred; the **3rd Industrial Revolution** (1950 – 2010) saw the birth of the internet and Information and Communication Technologies (ICTs); the emergence of robotics, genetic engineering, biotechnology, cell phones, solar, wind, and nuclear energy; and the replacement of analog mechanics with digital ones; the **4th Industrial Revolution** (2011 - present), known as Industry 4.0, is a model that combines various technologies: physical, digital, and biological. It is a business model aimed at generating knowledge and productivity.

Lima (2020) explains that there are four evolutionary stages of humanity since the 1940s, when **three sectoral axes** existed in the global economy: the **primary**, focusing on extractive activities; the **secondary**, related to manufacturing production, and the **tertiary**, with the growth of service provision. Different technological revolutions can be visualized in Figure 1.

Figure 1 – The 4 Phases of the Industrial Revolution



Source: Lopes (2019, p. 26).

During World War II, the British government enlisted experts in physics, mathematics, linguistics, physiology, biology, and electronic engineering to decipher German codes, resulting in the theoretical model developed by the British mathematician Alan Turing. Turing created the valve-based calculator Colossus, decrypting messages from the German army (LOPES, 2019).

It was in 1956 that the term Artificial Intelligence (AI) was first used by the computer scientist John McCarthy (BUARQUE, 2023). Post-war, Turing continued his research on machine intelligence, developing Automated Decryption and describing studies on Computational Logic and machine learning techniques. He also worked

on projects for intelligent robots that roam the field, learning from their experiences to aid in computerized agriculture (MUGGLETON, 2014, cited in LOPES, 2019, p.20).

3. ARTIFICIAL INTELLIGENCE (AI)

"Artificial Intelligence (AI) is the science and engineering of making intelligent machines capable of achieving goals in the world" (McCarthy, 2007, p.2, cited in Lopes, 2019, p.22). It emerged from the efforts of specialists to decode German communications during World War II. The birth of Computer Science can be attributed to the English mathematician Alan Turing, who conceived possible machine learning techniques (Lopes, 2019).

At that time, the United States had developed a private system as a strategy to protect its communications during the war, a system that would later expand globally, known as the World Wide Web. Individual microcomputers emerged, transforming the way humanity lives, works, and interacts, turning it into the Information Society. The convergence of innovations in institutional, technological, organizational, economic, political, and social aspects marked what is known as the Information Age (Puc-Rio, 2010).

John McCarthy coined the term Artificial Intelligence in 1956, conducting studies on "the mathematical nature of the thought process, including the theory of Turing machines, the speed of computers, the rel

ationship of a brain model to its environment, and the use of languages by machines" (Lopes, 2019, p.22).

AI involves the efficient use of information, allowing the recognition of market strategies and enabling managers to anticipate decision-making by solving problems satisfactorily. There is a consensus within organizations to build foundations on

innovations and continuous improvements, recognizing that information is value in today's context (Molina and Santos, 2019).

In addition to storing and manipulating data, AI acquires, represents, and manipulates knowledge in its processes due to its ability to deduce new knowledge from existing knowledge using existing algorithms. Buarque (2023, p.2) refers to four categorized concepts for Artificial Intelligence: "1. Systems that think like humans; 2. Systems that act like humans; 3. Systems that think rationally; 4. Systems that act rationally."

This analogy to human abilities makes the AI system unique based on some distinct factors: communication as if it were an entity, internal self-awareness, external knowledge, goal-oriented behavior, and creativity to adopt a specific alternative action if the initial action fails. In this sense, Buarque (2023, p.2) reports the definition given by Patrick Winston about AI: "...the study of computation that enables it to perceive, reason, and act."

4. DIGITAL REVOLUTION OR INDUSTRY 4.0

Honório (2022, p.9) states that Industry 4.0 is characterized by the "miniaturization, cost reduction, and evolution of electronic sensors, artificial intelligence, and massive data generation." Industrial digitization comprises important pillars, as described by Auletta *et al.* (2023, p.1), including "Artificial Intelligence (AI), data analysis, machine learning, cloud computing, and the Internet of Things (IoT)."

It is in this context that companies begin to utilize the resources and tools of AI because "the use of Artificial Intelligence to automate the mapping of the analysis unit's processes, aiming to optimize and innovate management and achieve cost reduction for the company" (Lacerda, 2022, p.7).

Among the benefits of Industry 4.0 are time and cost optimization due to the speed and accuracy of information produced by technological tools; flexibility through the

creation of systems; and product integration and development, known as digital manufacturing, characterized by the rapid, cost-effective production of high-quality products (Lopes, 2019). Despite various nomenclatures for technological advancements, all experts agree that "the generation and diffusion of information and knowledge are sources of value and power in this third millennium of the 21st century" (Puc-Rio, 2010, p.16).

The Digital Revolution or the 4th Industrial Revolution consists of the "[...] fusion of technologies and interaction between physical, digital, and biological domains" (Auletta *et al.*, 2023, p.3).

Venturelli (2018, cited in Auletta *et al.*, 2023, p.3) explains that the Automation Pyramid is a layered structure containing various interfaces in a vertical model, which, when limited to the local environment, lacks flexibility but has a significant influence on decision-making.

All these resources lead to the issue of Information Management or Data Governance, which is a "multifunctional framework for managing data as an organizational asset, focusing on data quality in a dual sense, in addition to critical aspects of security, privacy, and ethics" (Honório, 2022, p.9).

Despite the need and importance of careful handling of obtained data, Barbieri (2019, cited in Honório, 2022, p.17) explains that due to the complexity and the way data are managed, they have been neglected by the management of different companies, which can lead to losses of around 15 to 25% of revenue.

These considerations regarding security, privacy, and ethics can align with Information Management practices – or Knowledge Governance – represented by various formal mechanisms generated by Corporate Governance. These mechanisms require quality management for economic optimization. In other words, it is through systematic optimization and valuation systems of data that one can "co-create knowledge assets" (Honório, 2022, p.17).

5. GENERATIVE ARTIFICIAL INTELLIGENCE

Generative Artificial Intelligence is based on computational heuristic algorithms, utilizing advanced machine learning techniques and deep learning-based neural networks, which are fueled by generative neural networks" (Moura, 2023, p.2). These systems create hypotheses based on standardized data, learning autonomously and recognizing patterns across various layers of processing.

Generative AI has numerous potential applications, as it employs multimodal tools that work with language, images, sounds, and body movements. It can develop texts (text-to-text) through an input command (prompt), producing a response as output. This is done through conversation robots or chatbots. There are models for text-to-image, text-to-3D, text-to-task, and text-to-video (Duque-Pereira and Moura, 2023, p.3).

Machine learning is one of the approaches of AI that allows training and learning from a database, making modifications to improve specific tasks. It can analyze situations and time-series data using mathematical models or conduct analyses based on market trends, such as changes in consumer behavior. Learning techniques include neural networks and decision trees, capable of analyzing patterns and predicting future demands (Sanches and França, 2023).

The enhancement of decision-making guided by Artificial Intelligence involves (Laudon, 2011, p. 338, cited in Lopes, 2019, p.28-29):

- **Case-based reasoning:** using a database where both effective and disastrous solutions are stored, these are solutions that, when processed, will assess which one is more suitable for each case;
- **Fuzzy logic:** it is a type of technology based on rules that represent imprecision and that will create rules with approximate values for the most suitable solution;

- **Artificial neural networks:** devices that mimic the processing patterns of the human brain, following patterns in complex relationships, constructing models, and revising any errors based on large amounts of collected data;
- **Genetic algorithms:** ideal for solving problems that require optimal solutions, resembling neural networks;
- **Intelligent agents:** are back-end software, without direct human intervention, performing specific, repetitive, and predictable activities, used in business processes;
- **Expert systems:** use facts, knowledge, and reasoning techniques to solve problems that require special human skills.

In the present day, 'the volume, speed, and variety of data production rise to the dimension of analysis such as quality, legal aspects, ethics in treatment, and the ability to transform into direct values for the organization, within the concept of monetization and digital transformation' (Barbieri, 2019, cited in Auletta *et al.*, 2023, p.2).

As a consequence, to have a business strategic view that allows maintaining competitiveness, it is crucial for Information Management (IM) to demonstrate to managers that it is necessary to stop considering the volume of collected data as collateral items and implement '...a data governance framework, transforming them into business inputs.' In this sense, Redman (2016, cited in Auletta *et al.*, 2023, p.2) supports this concept, explaining that data storage is as important as knowing how to analyze and create models for this storage, preventing poor data from causing losses to large organizations.

6. IMPORTANCE OF DATA QUALITY AND AI

The quality of data is a crucial factor for organizations to deploy reliable Machine Learning (ML) models, as it is the quality of this data that will lead to optimal ML

performance (Rangineni, 2023). However, literature indicates numerous benefits brought by AI, as well as risks or disadvantages (Wach *et al*., 2023).

The stages that enable a thorough analysis of the ML pipeline phases are: 'data collection, preprocessing, model training, and validation' (Rangineni, 2023, p.16).

Quality criteria for data include accuracy, consistency, integrity, relevance, and ethical considerations, while challenges in the absence of quality include data noise, incompleteness, and biases. All these factors are the focus of expert research for the development of ML tools and the creation of reliable models (Rangineni, 2023).

Regarding data quality and its management, Budac (2022, p.1) explains that, in addition to existing ML models, systems aiming for 'data accuracy, integrity, and consistency' are necessary. The concern is to avoid the use of incomplete, erroneous, or inappropriate data, which leads to inadequate ML training and, consequently, undesirable results.

There are six dimensions to classify data quality, developed by experts and described by Budac (2022, p. 2-5):

- **Consistent representation:** each entity in the real world has only one way to be represented. For example, for the entity 'city,' New York will not be represented as NYC or NY, but only as New York;
- **Completeness:** When values are missing in the real context, for example, a medical sensor to monitor blood pressure, and the sensor fails for some time interval, impacting the result of that measurement;
- **Feature Accuracy:** Accuracy may decrease as real-world data contains errors or inaccuracies in its values, as data can come from various sources. Thus, the greater the deviation from the real value, the lower the accuracy of this data;

- **Target Accuracy:** There is a target feature for each dataset, meaning there is a class/label in classification tasks or a numerical value in regression tasks. For example, 'a fierce dog might be labeled as a 'wolf.'";
- **Uniqueness:** Data redundancy is another factor of great relevance, and the deduplication feature is used to avoid excessive adjustments;
- **Target Class Balance:** 'It is the balanced dataset to achieve satisfactory performance.' It refers to the k-Means algorithm because it 'recognizes clusters of approximately uniform sizes even if it is not the case in the input data".

In turn, the threats of data use that can harm ML models, which can be summarized in 7 main groups, as described by Wach *et al*. (2023, p.7):

1. lack of any regulation in the AI market and urgent need for regulation;
2. poor quality, lack of quality control, misinformation, deepfake content, algorithmic bias;
3. job loss, driven by automation;
4. breach of personal data, social surveillance, and privacy violation;
5. social manipulation, weakening of ethics and goodwill;
6. increased socio-economic inequalities;
7. technological stress of AI.

Rangineni (2023) argues that public and private companies are already aware of the importance of data quality in reducing social risks, lowering costs, and facilitating the efficient use of ML technologies. In this regard, Honório (2022, p.18) states that it is not enough to generate knowledge; it is also necessary to 'create ways to govern it.'

In terms of protecting against the violation of human rights, especially for children and young people aged 9 to 17, Buarque (2023) refers to the manifesto published by the Declaration of Toronto, advocating against any form of distinction,

discrimination, exclusion, restriction, or preference based on factors such as race, color, sex, language, religion, political opinion, or any other.

Wach *et al.* (2023) argue that it is essential to adopt certain practical and legal measures: regulation of the AI/AGI market; ensuring the security and protection of user and organizational data; promoting fair competition; protecting intellectual rights; and addressing privacy and geopolitical risks.

7. CONCLUSION

It is increasingly true that electronic data represents the most valuable asset in the modern world, a concept that extends to Generative AI. However, we must pay utmost attention to the data used for training AI models, given the growing volume of current data and the potential for the evolution of these models to progress rapidly. However, the challenge of managing this data, which is the 'fuel' for these models, lies in keeping up with this growth. The benefits brought by Generative AI are clear, the subject of the present article. However, in today's world, we have also come to realize the risks associated with this new technology, especially regarding unwanted content or content that may harm humanity.

The use of low-quality data, meaning data that may contain real-world problems such as racism and/or prejudice, or even some form of discriminatory or political bias, can cause these learning models to generate responses containing these problems.

The use of data quality steps prior to consumption by learning models is necessary not only to address the integrity and reliability of the data but also for the identification and treatment of ethical and legal issues through data analyses developed specifically for this purpose, thereby enabling the identification of data sets with the potential to generate undesirable results with the use of Generative AI.

Data is the starting point for everything, events presented from various perspectives. To ensure a safe evolution of Artificial Intelligence in the face of the evolution we are

experiencing, care must be taken to ensure that these models do not replicate problems by consuming low-quality data, which involves preprocessing and analyzing input data, a key factor in achieving this goal.

ACKNOWLEDGMENTS

I thank God and my wife Graciane Gadelha for all the support.

REFERENCES

AULETTA, Guilhermo Bellido *et al*. Governança de Dados e a Indústria 4.0. **Revista Científica Senai-SP – Tecnologia, Inovação & Educação**, São Paulo, SP, vol. 1, n. 2, p.01-12, 2023. Disponível em: https://periodicos.sp.senai.br/index.php/rcsenaisp/article/view/23. Acesso em: 16 out. 2023.

BUARQUE, Gabriela. Artificial intelligence and algorithmic discrimination: a reflection on risk and vulnerability in childhood. **Brazilian Journal of Law Technology and Innovation,** vol. 2, n. 2, p. 63-84, set. 2023. DOI:10.59224/bjlti.v1i2.63-86. Disponível em: https://www.researchgate.net/publication/373887331_Artificial_intelligence_and_algorithmic_discrimination_a_reflection_on_risk_and_vulnerability_in_childhood. Acesso em: 18 out. 2023.

BUDACH, Lukas *et al.* **The Effects of Data Quality on Machine Learning Performance**. arXiv, 2022. Disponível em: https://arxiv.org/pdf/2207.14529.pdf. Acesso em: 16 out. 2023.

DUQUE-PEREIRA, Ives da Silva; MOURA, Sergio Arruda de. Compreendendo a Inteligência Artificial Generativa na perspectiva da língua. **SciELOPreprints**, 2023. https://doi.org/10.1590/SciELOPreprints.7077. Disponível em: https://preprints.scielo.org/index.php/scielo/preprint/view/7077/13284. Acesso em: 31 out 2023.

HONORIO, Roseli. **Modelo Conceitual de Governança de Dados como suporte à Governança do Conhecimento Organizacional.** 2022. Dissertação [Mestrado em Engenharia e Gestão do Conhecimento] apresentada ao Programa de Pós-Graduação da Universidade Federal de Santa Catarina. Florianópolis, 2022. Disponível em: https://repositorio.ufsc.br/bitstream/handle/123456789/243667/PEGC0739-D.pdf?sequence=-1&isAllowed=y. Acesso em: 16 out. 2023.

LACERDA, Lidia Correa de. **O uso da inteligência artificial na gestão da inovação tecnológica:** automatização do processo de mapeamento de dados. 2022. Trabalho de Conclusão de Curso [Graduação em Engenharia de Produção] apresentado à Escola de Engenharia de Petrópolis da Universidade Federal Fluminense. Petrópolis, 2022. Disponível em: https://app.homologacao.uff.br/riuff/handle/1/25224. Acesso em: 16 out. 2023.

LIMA, Jean Santos. **A Vantagem Competitiva das Nações no Limiar da 4ª Revolução Industrial:** A Importância da Economia do Conhecimento, da Sinergia entre Indústria e Serviços, e da Política Internacional. 2020. Tese [Doutorado em Relações Internacionais] apresentada ao curso de Pós-graduação em Relações Internacionais do Instituto de Relações Internacionais da Universidade de Brasília (UnB). Brasília, 2020. Disponível em: http://icts.unb.br/jspui/bitstream/10482/39098/1/2020_JeanSantosLima.pdf. Acesso em: 19 out. 2023.

LOPES, Roberta da Silva. **Inteligência Artificial na Contabilidade em Organizações Públicas:** Potencialidades e Desafios. Dissertação [Mestrado em Controle de Gestão] apresentada à Universidade do Estado do Rio de Janeiro. Rio de Janeiro, 2019. Disponível em: https://www.bdtd.uerj.br:8443/handle/1/8054. Acesso em: 16 out. 2023.

MOLINA, Leticia Gorri; SANTOS, Juliana Cardoso dos. **Gestão da Informação e a 4a Revolução Industrial**. AtoZ: novas práticas em informação e conhecimento, vol. 8, n.2, p 39-48, jul./dez. 2019. DOI: 10.5380/atoz.v8i2.65784. Acesso em: 19 out. 2023.

MOURA, Mariana Vasques. **A Inteligência Artificial Generativa como autora de invenções patenteáveis:** um estudo analítico do "Caso Dabus". O Globo, 2023. Disponível em: https://repositorio.animaeducacao.com.br/bitstream/ANIMA/36309/1/2023.06.04%20-%20TCC%20Mariana%20Moura%20-%20VF%20-%20PDF-A.pdf. Acesso em: 31 out 2023.

O GLOBO. **Criança é diagnosticada com doença rara pelo ChatGPT após passar por 17 médicos**. O Globo, 2023. Disponível em: https://oglobo.globo.com/saude/medicina/noticia/2023/09/13/crianca-e-diagnosticada-com-doenca-rara-pelo-chatgpt-apos-passar-por-17-medicos-entenda.ghtml. Acesso em: 31 out. 2023.

PUC RIO. **Conhecimento Tecnológico e Informação:** a Era da Sociedade Informacional. Puc-Rio, 2010. Disponível em: https://www.maxwell.vrac.puc-rio.br/16712/16712_3.PDF. Acesso em: 19 out. 2023.

RANGINENI, Sandeep. An Analysis of Data Quality Requirements for Machine Learning Development Pipelines Frameworks. **International Journal of Computer Trends and Technology**, v.71, n. 8, ago. 2023, p.16-27. DOI:10.14445/22312803/IJCTT-V71I8P103. Disponível em: https://www.researchgate.net/publication/373821198_An_Analysis_of_Data_Quality_Requirements_for_Machine_Learning_Development_Pipelines_Frameworks. Acesso em: 16 out. 2023.

SANCHES, Felipe Norato; FRANÇA, Celso Ap. de. O uso de algoritmos de classificação para determinar estoques de segurança. **Repositório Institucional UFSCar**, 2023. Disponível em: https://repositorio.ufscar.br/handle/ufscar/18405. Acesso em: 16 out. 2023.

SANTOS, Leon. **Conheça as quatro Revoluções Industriais que moldaram a trajetória do mundo**. CFA, 2019. Disponível em: https://cfa.org.br/as-outras-revolucoes-industriais/. Acesso em 19 out. 2023.

WACH, Krzysztof *et al*. The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. **Entrepreneurial Business Magazine**, vol.11, n. 2, p. 7-30, Jun. 2023. DOI:10.15678/EBER.2023.110201. Disponível em: https://www.researchgate.net/publication/371987305_The_dark_side_of_generative_artificial_intelligence_A_critical_analysis_of_controversies_and_risks_of_ChatGPT. Acesso em: 16 out. 2023.

Submitted: November 14, 2023.

Approved: November 27, 2023.

[1] Postgraduate Specialization in Project and Process Management; Bachelor's Degree in Information Systems. ORCID: 0009-0005-4593-7897.