



Revista de Administração Contemporânea

ISSN: 1415-6555

ISSN: 1982-7849

Associação Nacional de Pós-Graduação e Pesquisa em Administração

Biondi, Gabrielle Marques Castelo Branco; Cernev, Adrian Kemmer
Nuveo: Ética Digital e Inteligência Artificial para Desafios do Mundo Real
Revista de Administração Contemporânea, vol. 27, no. 3, e220063, 2023
Associação Nacional de Pós-Graduação e Pesquisa em Administração

DOI: <https://doi.org/10.1590/1982-7849rac2023220063.por>

Available in: <https://www.redalyc.org/articulo.oa?id=84075233006>

- How to cite
- Complete issue
- More information about this article
- Journal's webpage in redalyc.org

UABEM
redalyc.org

Scientific Information System Redalyc
Network of Scientific Journals from Latin America and the Caribbean, Spain and Portugal

Project academic non-profit, developed under the open access initiative

Case for Teaching

Nuueo: Digital Ethics and Artificial Intelligence for Real World Challenges



Nuueo: Ética Digital e Inteligência Artificial para Desafios do Mundo Real

Discipline: Information Technology, ESG
Subject: Digital Ethics, Artificial Intelligence
Industry: Information Technology, Public Security
Geography: Brazil

Gabrielle Marques Castelo Branco Biondi¹ 
Adrian Kemmer Cernev^{*2} 

THE OPPORTUNITY OF A NEW BUSINESS

September 2019. While the arrival of spring heralded the beginning of a new cycle for nature, José Flávio Pereira — founder and CEO of Nuueo — wondered if the time had come to start a new cycle for his startup, expanding the offer of its image recognition solutions to something more than process automation.

Nuueo had been acquiring important clients with its proprietary artificial intelligence (AI) technology for analyzing structured and unstructured data in documents and images. José Flávio identified an opportunity to expand its computer vision solution for a purpose that did not involve process automation: facial recognition to locate criminals.

He had read recent research that showed that the use of image recognition technology for surveillance/security would be one of the drivers of the global facial recognition market and estimated that the growth of the Brazilian market would follow the forecasts for Latin America, revolving around 16% a year.

He had also mapped the growing interest of government agencies in this type of tool for public safety, including a forecast of government incentives to encourage the use of this technology¹. One of them, in particular, would soon undertake contracting image recognition technology for public safety.

* Corresponding Author.

1. Fundação Getúlio Vargas, Escola de Administração de Empresas de São Paulo, Mestrado Profissional em Gestão para Competitividade - TI, São Paulo, SP, Brazil.
2. Fundação Getúlio Vargas, Escola de Administração de Empresas de São Paulo, Departamento de Tecnologia e Ciência de Dados, São Paulo, SP, Brazil.

Cite as: Biondi, G. M. C. B., & Cernev, A. K. (2023). Nuueo: Digital ethics and artificial intelligence for real world challenges. *Revista de Administração Contemporânea*, 27(3), e220063. <https://doi.org/10.1590/1982-7849rac2023220063.en>

Published as Early Access: February 15, 2023.
Assigned to this issue: April 25, 2023.

of invited reviewers until the decision:

	1	2	3	4	5	6	7	8	9
1 st round									
2 nd round									

JEL Code: M13, M14, M15.

Editor-in-chief: Marcelo de Souza Bispo (Universidade Federal da Paraíba, PPGA, Brazil) 
Associate Editor: Paula Castro Pires de Souza Chimentil (Universidade Federal do Rio de Janeiro, COPPEAD, Brazil) 
Reviewers: Bruno Fernandes (Universidade Federal do Rio de Janeiro, COPPEAD, Brazil) 
Marco Aurelio de Souza Rodrigues (Escola Superior de Propaganda e Marketing - ESPM, Brazil) 
One reviewer did not authorize the disclosure of his/her identity.

Peer Review Report: The disclosure of the Peer Review Report was not authorized by its reviewers.

Received: March 02, 2022
Last version received: January 11, 2023
Accepted: January 18, 2023

Note: This text is translated from the original Portuguese version, which can be accessed [here](#).

José Flávio knew that this opportunity could allow Nuveo to increase its revenue and quickly consolidate the company in the Brazilian AI market. With the entry of more financial resources and the necessary technological improvement, he would have the potential to expand the portfolio of solutions and customer base, taking his company to a new level.

However, he also knew that, despite the expectation of a considerable increase in revenue for a company that still faced financial limitations and the real possibility of expanding and consolidating its business, the opportunity also brought risks since the degree of accuracy and the biases contained in facial recognition solutions made them susceptible to failure. In addition, these tools were the subject of growing discussions in academic circles and society, under the allegation of being a possible threat to privacy and democracy. Would it be the best time to invest in this new business?

NUVEO

Solving real problems has always been Nuveo's goal. Founded in 2016 with this mission, the startup develops systems and algorithms that use AI to offer solutions mainly aimed at process automation, focusing on improving production capacity, increasing efficiency, and reducing costs.

The idea for the original product emerged when José Flávio was volunteering at a nonprofit organization. While studying for the MBA, he started to work at the GESC Institute² and provide social consultancy for civil society organizations. The NGO *Lar das Crianças* (children's home) was one of them.

The nonprofit was facing a large volume of repetitive and manual work to obtain financial resources from credits from the Nota Fiscal Paulista³. These resources were often not received by the institution in their entirety because there was not enough time to register the credit request since it should be done by the 20th of the month following the document issuance date. The development and implementation of an artificial intelligence solution, simulating human reasoning, allowed the nonprofit to automate its repetitive processes and obtain the maximum credit.

With an entrepreneurial spirit, training, and background in business management, José Flávio soon realized the potential that his solution had for this problem and several others related to bureaucratic and manual processes faced by companies in general. Thus, Nuveo was born to solve real-world challenges with the help of information technology.

Its solutions are developed from API (application programming interface) platforms based on a scalable and elastic cloud architecture, which uses, among others, its patented computer vision algorithms — Ultra OCR and Smart Vision — for image capture and subsequent data processing and analysis and integration with its customers' systems. This process is fully automated, with low impact in terms of customizations and infrastructure, and enables the company to gain customers in different sectors, such as finance, retail, health, and energy.

The startup has tech professionals distributed between its headquarters in São Paulo/SP and its branch in Campina Grande/PB — dedicated exclusively to R&D — and seeks to maintain a close relationship with universities and research centers in the country, especially the Technology Park of Paraíba.

In its three years of existence, Nuveo had structured a team of around 15 employees and gradually increased its revenues, with good acceptance of its product and validation of its business model. However, identifying the demand for its computer vision solution for public safety represented a possibility of exponential growth for the company — the goal of any startup. The company was created with the founder's own resources and did not have financial support from investors.

Given the number of employees and its annual gross revenue, the startup qualified as a small business⁴. José Flávio expected that offering this new solution could at least triple his company's revenue next year and become a showcase for other purposes, further increasing the range of opportunities for the company since the interest of government agencies in this type of tool was not restricted to use in the security area. In addition, he considered that offering a national product at lower costs was an important competitive advantage for Nuveo since most of the solutions already implemented for public management purposes in the country used technology from foreign companies.

Although this seemed to be an excellent opportunity for the company to grow, consolidate in the market, and contribute to tackling a real problem in society — the reduction of crime —, José Flávio could not stop thinking about the possible setbacks that this implementation could bring. When researching the potential market for this solution, José Flávio also encountered ethical issues. And this kept hammering in his head...

Would Nuveo really be contributing to solving a social problem by offering this type of solution? Would it not be better to give the solution more time to mature, preventing it from causing problems for society?

DECIDING THE FUTURE

José Flávio had just learned that a government agency would hire facial recognition technology for public safety. Obtaining this contract would significantly increase the company's revenue and could leverage several other opportunities, but what about his conscience and his company's reputation if a false positive recognition led to serious consequences?

Faced with his doubts and needing to make a decision, José Flávio set up a meeting with three professionals of his utmost confidence: Maria, who worked in the commercial area; Pedro, with vast knowledge in finance; and Arthur, a technology specialist. Upon hearing about the imminent new business opportunity for Nuveo, everyone was excited about the company's great growth potential.

After the initial euphoria, José Flávio began to express his concerns: the current state of technology and algorithmic and data biases allowed false positives and, therefore, an innocent person could be accused of a crime he/she did not commit. A brief internet search found real cases in several countries. Furthermore, data and algorithm biases often occurred with people who already belonged to groups that suffer from social inequalities — such as women, LGBTQIA+⁵ people, and people with non-white skin and, in this way, could lead to discrimination in these more vulnerable groups, further aggravating the already existing asymmetry in society.

Maria listened attentively, reflecting on José Flávio's considerations. She thought for a while and then argued: "False positives really can happen, Flávio. And it's true that this is very bad, but these are isolated cases... On the other hand, think about how many criminals could be arrested with this solution! How many citizens will be safer! The tool can lead to mistakes, but it will be right much more than wrong! The benefits far outweigh the problems."

"Mary is right!" said Peter. He added: "In Brazil, where violence rates are worrisome, solutions developed to collaborate with increased public safety will probably be welcomed also by the population, however controversial this topic may be..."

"Hmm... It's true, Pedro... Who doesn't want a safer city? It is also worth remembering that a failure of the tool would lead to the incorrect identification of a person. Only that. The methodology used to approach the suspect and make their arrest is not Nuveo's responsibility," Arthur added.

"Don't you think?" José Flávio asked. And he continued: "That digital tools can profoundly change our lives for the better, I have no doubt. And that is the purpose of Nuveo! Being able to contribute to the reduction of crime would be fantastic! But what consequences would possible

errors or misuse have for people, for the company, and even for the whole society?"

"Flávio..." continued Arthur, "we know that the accuracy rate of these solutions has been improving rapidly and this could be a great opportunity for the company to evolve technologically! You have a top team and will certainly deliver a tool with a high level of success that will be constantly improved. This can be worked on if there is a concern about algorithmic and data biases. I am sure it is possible to implement mitigation actions in this regard!"

"Of course, the company can act on this! There are international initiatives to guide the development of systems that employ AI in a fairer way for all. But how long will it take before we feel safe to offer our solution? How will we know we've made an appropriate tool? That we will not harm society? We also need to consider that investing time and money in adapting our solution to this new context may impact the deadlines of ongoing projects. In addition, other factors interfere with results in public places: ambient lighting, positioning of cameras..." commented José Flávio.

He continued: "The thing is, we're not dealing with numbers or anything else. We are dealing with lives! And I wouldn't sleep easy if I knew that a tool offered by my company negatively affected someone's life! Also, I care about the company's reputation. If we lose our credibility or society understands that this type of tool is morally offensive, Nuveo could suffer irreversible damage."

"You're right, Flávio..." reflected Pedro, adding: "More than ever, investors and customers are keeping an eye on ESG⁶-related indicators..."

José Flávio added: "That's right... Debates about the possible consequences of this technology for public management are gaining strength worldwide, leading some cities even to ban the use of these tools. This is not limited to the biases that can lead to discrimination. Discussions involve the loss of individual privacy and freedom of expression and even speak of the violation of democracy..."

"Guys, I understand your point of view. However, Nuveo is a startup!" Maria pointed out. "Taking risks is part of the game! Have you ever thought about what it would be like if people and companies gave up on innovations for fear that they could go wrong? How would the world evolve?"

She concluded: "Is it prudent to neglect these market demands and miss the timing of offering this type of solution while other companies take the lead? As they say: the best is the enemy of good... Despite the risks, wouldn't this be the best time to act?"

The meeting ended up bringing more questions than decisions. José Flávio understood that this opportunity would be perfect for consolidating his company in the national AI

market and felt that wasting this chance could pave the way for competitors and jeopardize his company's growth. However, he was also concerned about the impacts the implementation of a facial recognition solution for public safety could bring to Nuveo and society, especially when concerns related to environmental, social, and corporate governance factors intensified among the population and investors.

It was necessary to move on, and José Flávio found himself faced with some possibilities...

One option was to immediately offer a facial recognition solution for public safety agencies, taking advantage of the moment of great market demand to expand the business and consolidate the company as a reference in AI. After all, as he well knew, a startup needs to grow fast!

Another option was to continue with the idea of offering facial recognition solutions but delay the launch. This way, there would be time to invest in improving the solution, only then making the product available to the market for surveillance/security purposes. He could also assess the best way to interact with the public sector. Would participating in partnership programs between governments and startups, carrying out proofs of concept, and a pilot project of the solution be a good alternative? However, this would imply not sending the proposal, giving up the current opportunity, and then waiting for a future opportunity to arise, hoping that no competitor has already dominated this market.

The final option would be to shelve the opportunity to operate in the public safety market and continue with the portfolio of AI solutions for process automation purposes, using this technology safely without additional risks to the business.

If José Flávio was divided before, now the decision seemed to be even more complex... He called his company's directors to define the course of the business. There was no more time to decide whether to submit a proposal for this opportunity. His deadline had run out! It was time for a final decision!

Was this the right time for Nuveo to invest in this new business? Which path should José Flávio take?

NOTES

1. Ordinance No. 793 (2019), October 24th, 2019. It regulates the financial incentive of the actions of the axis of combating violent crime, within the national policy of public security and social defense and the single public security system, with resources from the national public security fund, provided by Law No. 13,756. Ministry of Justice and Public Security. <https://dspace.mj.gov.br/handle/1/1380>
2. GESC Institute: Founded in 2004 by the FIA Business School of the University of São Paulo (USP), the institute brings together several specialized professionals who provide consultancy for civil society organizations (CSOs).
3. Nota Fiscal Paulista (São Paulo Invoice): Program of the Government of the State of São Paulo that aims to encourage fiscal control in the state. When requesting an invoice at the time of purchase, participating consumers can receive back part of the ICMS (State VAT) collected by the establishment or choose to direct this credit to a nonprofit organization.
4. According to Sebrae, a small business has between 10 and 49 employees and annual revenues between BRL 360,000 and BRL 4.8 million (retrieved on October 28, 2022, from <https://www.sebrae.com.br/sites/PortalSebrae/ufs/ac/artigos/epp-entenda-o-que-e-uma-empresa-de-pequeno-porte,305fd6ab067d9710VgnVCM100000d701210aRCRD>).
5. LGBTQIA+: Acronym for: lesbian, gay, bisexual, transgender, queer or questioning, intersex, asexual. The plus symbol (+) includes other sexual orientations and gender identities.
6. ESG: Acronym for environmental, social, and governance, referring to the company's performance in a more holistic way, with environmental care, concern for generating positive social impacts, and the adoption of corporate governance practices.

ANNEX 1

1.1 AI, biometrics, and facial recognition growth prediction

Table A1. Growth rate prediction for AI, biometric, and facial recognition.

Market	Annual growth rate prediction	Dimensions
Global AI market	38.1%	The global AI market reached USD 93.5 billion in 2021 and presents a predicted growth rate of 38.1% per year between 2022 and 2030.
Global biometric market	17.7%	The global biometric market is expected to grow at a rate of 17.67% between 2019 and 2026, reaching USD 62.3 billion in 2026.
Global facial recognition market	18%	The specific global market segment of facial recognition reached more than USD 3 billion in 2019. It has the potential to reach a total revenue of more than USD 12 billion in 2026.
Biometric market in LatAm	17.8%	The revenues of the biometric market in LatAm can grow from USD 2.63 billion in 2021 to USD 7.09 in 2027.
Facial recognition market in LatAm	16.4%	The growth rate estimate for the facial recognition market in LatAm is 16.35% per year for the period from 2019 to 2028.

Note. Source: Elaborated by the authors based on the reports: Grand View Research (retrieved on October 28, 2022, from <https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market>), Maximize Marketing Research (retrieved on October 28, 2022, from <https://www.maximizemarketresearch.com/market-report/global-biometrics-technology-market/44408/>), Global Marketing Insights (retrieved on October 28, 2022, from <https://www.gminsights.com/industry-analysis/facial-recognition-market>), IMARC (retrieved on October 28, 2022, from <https://www.imarcgroup.com/latin-america-biometrics-market>) and MarketResearch.com (retrieved on October 28, 2022, from <https://www.marketresearch.com/Inkwood-Research-v4104/Latin-America-Facial-Recognition-Forecast-13507333/>).

1.2 Use of facial recognition for public management in Brazil

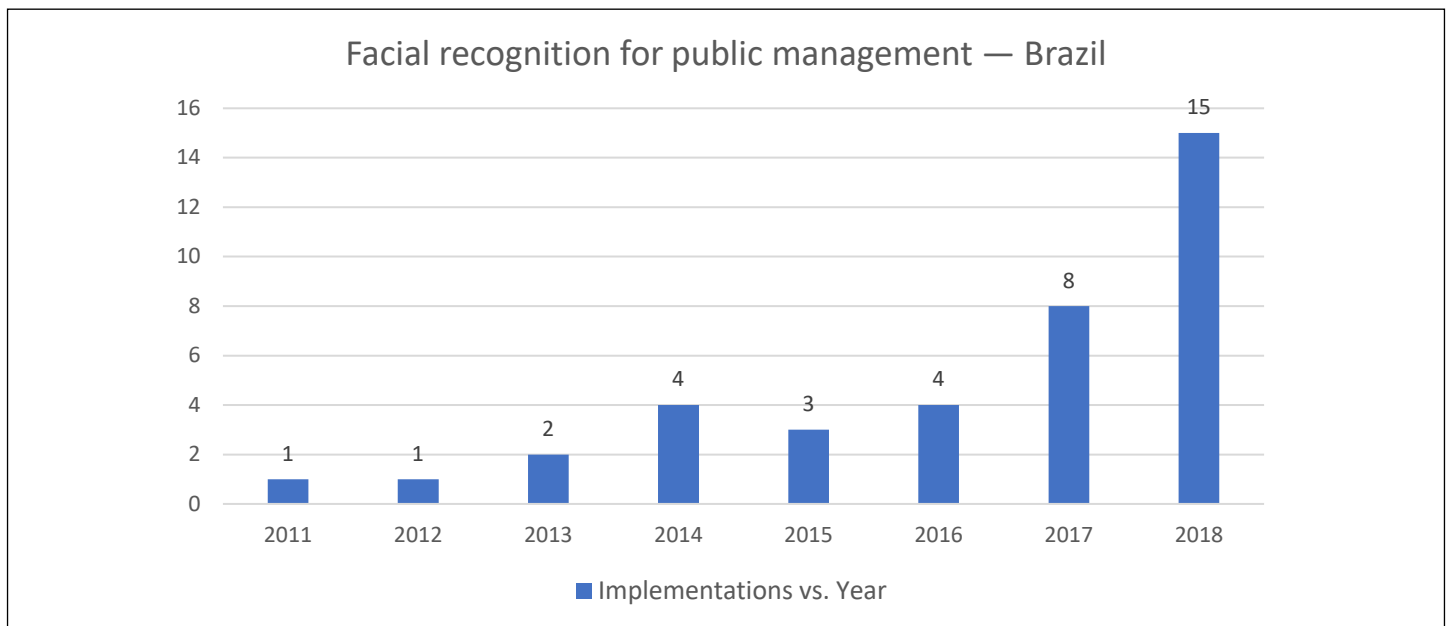


Figure A1. Use of facial recognition tools for public management in Brazil.

Source: Elaborated by the authors based on data from Igarapé Institute (retrieved on September 11, 2021, from <https://igarape.org.br/infografico-reconhecimento-facial-no-brasil/>).

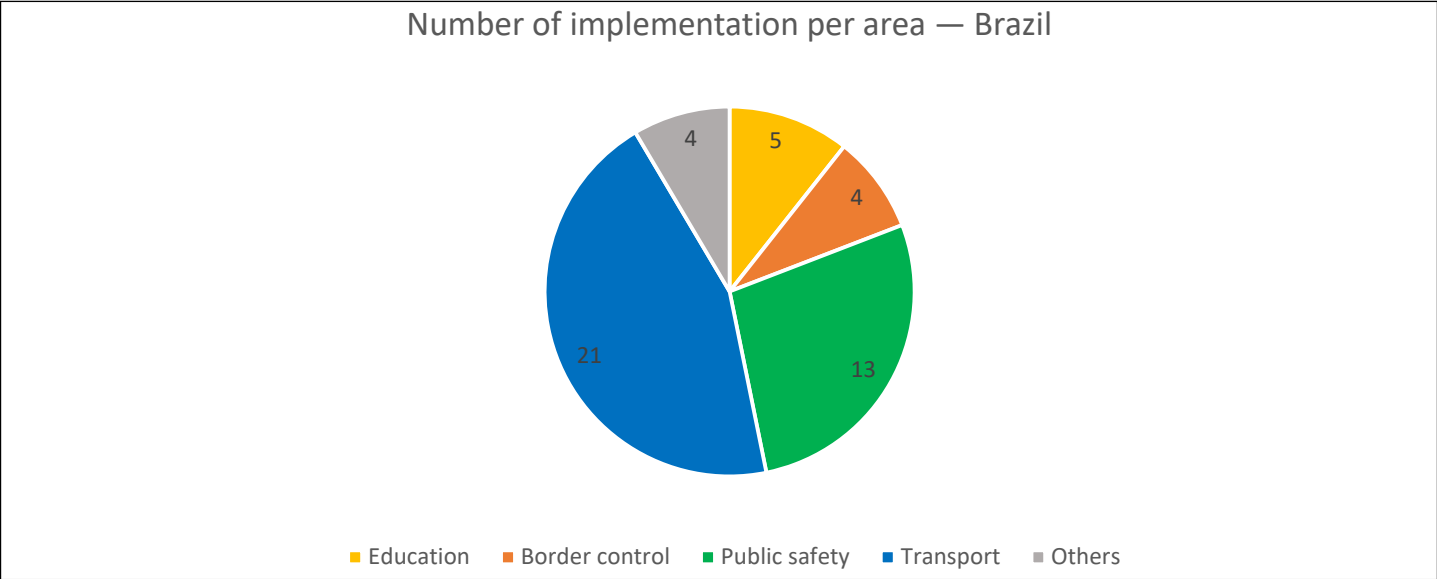


Figure A2. Implementation of facial recognition tools for public management in Brazil between 2011 and March 2019.
Source: Elaborated by the authors based on data from Igarapé Institute (retrieved on September 11, 2021, from <https://igarape.org.br/infografico-reconhecimento-facial-no-brasil/>).

1.3 Accuracy of facial recognition software

Data from the National Institute of Standards and Technology — NIST (retrieved on November 3, 2022, from <https://doi.org/10.6028/NIST.IR.8238>) point out that the accuracy of facial recognition software has been evolving, and error rates are below 0.02% when good quality images are used to find matches from a database with 12 million individuals.

However, according to NIST, facial recognition from video surveillance cameras is much more challenging, as it involves several conditions, such as ambient lighting, camera resolution and positioning, blurred or obstructed faces, speed of movement, face positioning, and aging. Tests performed have shown that the identification error rate can vary from less than 1% to more than 40%, depending on the algorithm used, and the false negative identification rate with images obtained from professional cameras installed on ceilings can range from 20% to over 90%.

In addition, according to the institute, the accuracy rate is influenced by factors such as race, gender, and age group:

Table A2. Effects of ethnicity, age, and sex in the accuracy of the facial recognition software.

Evaluation	Results
Sex	There are two to five times more false positives for women than for men
Ethnicity	Higher rates of false positives were observed among Africans and Asians, followed by Central Americans People from Eastern Europe presented lower rates of false positives Algorithms developed in Chine presented a low rate of errors for Asian, sometimes lower than for Caucasian Algorithms developed in the US presented a higher rate of false positives for Native Americans
Age	Higher rates of false positives for elderly and children

Note. Source: Elaborated by the authors based on research from NIST in 2018 (retrieved on October 16, 2021, from <https://doi.org/10.6028/NIST.IR.8280>).

ANNEX 2

Glossary of technical terms

- . **AI: artificial intelligence.** They are computational systems based on human behavior to solve problems and/or make decisions.
- . **API: application programming interface.** This technology allows communication between different computer systems.
- . **Cloud.** It is a technological architecture that makes remote computational resources available on demand, from data storage to computational capacity.
- . **Deep learning.** It is a branch of machine learning that seeks to teach computer systems to act and interpret data more naturally, usually using neural networks.
- . **Machine learning.** It is a system that can modify its own behavior autonomously, based on recursive training, with minimal human interference.
- . **Smart Vision.** It is a patented computer vision system used for intelligent monitoring.
- . **Ultra OCR: Ultra Optical Character Recognition.** It is a patented system used for optical character recognition.

Teaching Notes

■ ABSTRACT

This teaching case presents the dilemma faced by Nuveo in its quest to grow sustainably and consolidate itself in the Brazilian artificial intelligence market. When the opportunity comes to offer his image recognition technology for public safety, the startup's founder finds himself up against ethical issues. By telling Nuveo's story, this teaching case allows the identification of principles and recommendations for ethical development and use of AI systems, enabling discussions about ethical challenges related to digital transformation and its impacts on individuals, companies, and society. It is intended for use at undergraduate and graduate courses in business administration, public administration, and information technology and can be applied in disciplines that address digital ethics, ESG (environmental, social, and governance), and artificial intelligence.

Keywords: digital ethics; artificial intelligence; ESG; information technology.

SYNOPSIS

This teaching case presents the story of Nuveo, a startup specialized in process automation that seeks to grow and consolidate in the Brazilian artificial intelligence (AI) market. Since its creation, the company has developed and patented dozens of algorithms. Its main patented software are Ultra OCR and Smart Vision, computer vision algorithms that allow capturing images for further data processing and analysis and integration with customers' systems.

Realizing the growing interest of Brazilian government institutions in computer vision solutions for facial recognition, especially for surveillance and public safety, and the expected growth of this type of solution in the global market, José Flávio — founder and CEO of the startup — sees an opportunity to expand the business and exponentially increase revenues.

However, when researching more about this opportunity, he realized the risk of unforeseen and/or unwanted consequences such as discrimination, loss of privacy, and violation of democracy. He also observed that discussions about these solutions' development, implementation, and ethical use were intensifying worldwide.

José Flávio must decide whether his company should send a proposal for facial recognition of criminals to improve public safety. He is faced with three possible options: (a) go ahead and offer his solution and take advantage of the

■ RESUMO

Este caso de ensino apresenta o dilema enfrentado pela Nuveo em sua busca por crescer sustentavelmente e consolidar-se no mercado de inteligência artificial do país. Ao identificar a possibilidade de ofertar o uso da sua tecnologia de reconhecimento de imagens para a segurança pública, o fundador da *startup* depara-se com questões éticas envolvidas com essa oportunidade. Ao contar a história da Nuveo, este caso permite a identificação de princípios e recomendações para o desenvolvimento e uso ético de sistemas de IA, propiciando ainda discussões acerca dos desafios éticos relacionados à transformação digital e seus impactos para os indivíduos, empresas e sociedade. Este caso de ensino foi elaborado para ser aplicado em cursos de graduação e pós-graduação em administração de empresas, administração pública e tecnologia da informação, em disciplinas que abordem ética digital, ESG (*environmental, social, and governance*) e inteligência artificial.

Palavras-chave: ética digital; inteligência artificial; ESG; tecnologia da informação.

identified opportunity; (b) invest in improving his solution and then offer it to the public administration; (c) forget this opportunity and continue with the portfolio of AI solutions for process automation.

The case provides the student with the context of the situation and other information that helps in decision-making. It was written based on the guidelines provided by [Alberton and Silva \(2018\)](#) and [Faria and Figueiredo \(2013\)](#).

POSSIBILITIES OF APPLICATION

This teaching case aims to incorporate digital ethics into undergraduate and graduate curricula, especially in information technology courses that address digital ethics, ESG (environmental, social, and governance), and AI (artificial intelligence).

LEARNING OBJECTIVES

At the end of the discussions, students are expected to:

- Understand some ethical challenges related to digital transformation and its consequences for individuals, companies, and society;
- Understand that ethical principles must be observed from the moment of designing AI solutions to their

implementation and use, as well as when obtaining, processing, and using personal data;

- Identify risk mitigation measures so that the design and implementation of systems that employ AI are carried out responsibly;
- Understand that investing in digital ethics brings positive social impact and business value.

PREVIOUS PREPARATION

Students should be previously prepared to participate effectively in the discussions. The recommended and additional materials below may be approached in a class before discussing the case, helping to level the participants' knowledge of the topics.

RECOMMENDED MATERIALS

According to the lesson plan suggested below, it is recommended that participants read the following materials in advance:

- Document providing an overview of facial recognition technology: [Centre for Data Ethics and Innovation \(2020\)](#);
- Article that presents ethical principles for AI, summarizing six important international initiatives: [Floridi and Cowls \(2019\)](#);
- Framework developed by AI HLEG for the ethical use of AI: [European Commission \(2019\)](#).

ADDITIONAL MATERIALS

Table 1. Additional materials for participants and instructors.

Material	Indication
Gartner report on the importance of responsible use of facial recognition tools and providing recommendations: Buytendijk et al. (2020) . NIST article on a demographic study in facial recognition algorithms: National Institute of Standards and Technology [NIST] (2019) .	Expand and level knowledge about facial recognition technology and support case discussion questions
Article by IBM that addresses the difference between AI, machine learning, and deep learning: Ceron (2019) . Medium article that brings an AI map: Corea (2018) .	Align knowledge about the scope of AI and the difference in approaches
Video on ethical challenges, dealing with the importance of not delegating subjective decision-making exclusively to machines: Tufekci (2016) .	Contribute to reflections on the difference between human and artificial intelligence and on the importance of establishing ethical principles to develop and use AI
Article on the regulation of digital environments: Crespo (2022) . IEEE guide to the ethical use of AI (Institute of Electrical and Electronics Engineers, 2017).	Present challenges for the regulation of digital solutions
Document from the Brazilian government that aims to guide research, the development of solutions, and the use of AI: Ministério da Ciência, Tecnologia e Inovação [MCTI] (2021) .	Present initiatives to promote the development and responsible use of AI
Gartner report that presents digital ethics as a competitive factor: Hamer et al. (2020) . Article by Forbes that presents ethics in AI as a competitive factor: McKendrick (2019) .	Presenting digital ethics as a competitive factor for companies

Note. Source: Elaborated by the authors.

PREPARATION QUESTIONS

- What is a facial recognition system, and how does it work?
- Describe this technology's risks (besides false positives) and benefits.
- What is digital ethics?

QUESTIONS TO DISCUSS THE CASE

1. What problem is José Flávio facing?

This question aims to bring the participant to the heart of the discussion. The central problem faced by the protagonist is related to the provision of AI for public safety. He knew that, although it was an opportunity to

expand and consolidate the business, the change in the context of using his solution — from process automation to public safety — involved ethical aspects that were still under discussion in academia and society and would bring challenges that, if not addressed properly, could negatively impact people and the company.

To discuss the problem faced by the protagonist, it is worth establishing some definitions and concepts presented in the following paragraphs.

According to the original definition, AI is the ability of machines and algorithms to simulate human behavior (McCarthy et al., 1955). However, it is important to keep in mind that intelligence and autonomy are capabilities restricted to people. Although machines and systems can perform tasks that were once exclusive to people (often performing them better than humans), this does not mean that they can think alike, act rationally, or have moral values (European Commission, 2018; Floridi & Cows, 2019).

This difference is crucial to understanding that the abstraction of these concepts and attributing these characteristics to machines and systems is the basis for the emergence of several ethical dilemmas (Institute of Electrical and Electronics Engineers, 2017).

It is necessary to consider that machine learning takes place from the data with which they are fed and trained and that it is also from such data that the results obtained are constantly improved.

Machine learning algorithms use statistical techniques to find patterns in data and make predictions according to how they were programmed. Deep learning algorithms are even more sophisticated, adapting when exposed to new contexts and data patterns (Ceron, 2019). Thus, the data quality, the technique employed, and the context to which these algorithms are submitted directly influence the results achieved and the consequences generated.

The use of data can also lead to ethical issues. They usually involve (a) the use or reuse of data in contexts, times, and for purposes different from the original consented purpose (Herschel & Miori, 2017); (b) the lack of transparency, veracity, and clarity of the terms of consent (Weinhardt, 2020), aggravated by the general public's lack of knowledge of the power of inference of algorithms based on the data obtained (Hinds et al., 2020); and (c) the lack of a single understanding of what constitutes public data, visibly differentiating it from private data (Weinhardt, 2020).

Another relevant point to be considered is that new technologies emerge much more rapidly than the

time required to establish and/or adapt legal instruments. In addition, digital solutions can transcend national territories, increasing the complexity of law enforcement (Crespo, 2022).

Once the protagonist's dilemma is understood, it would be appropriate to ask the participants: What options did he have at that moment?

2. What are the pros and cons of the options evaluated by José Flavio?

Having identified the three options presented in the case, the instructor must ask about the positive and negative points of each one.

For this question, students are expected to make a critical assessment of the situation experienced by José Flávio, weighing the pros and cons of each option presented in the case. Participants should be aware of points such as the maturity and accuracy of facial recognition solutions, regulations, impacts on individuals, society, Nuveo's reputation, and competitiveness.

Option 1 may allow the company to enter a market that is still developing in Brazil, increase its revenues and its customer base, and contribute to locating people wanted by the police. However, this option disregards the factors related to the maturity and accuracy of facial recognition solutions and the ethical issues that have been intensely debated worldwide.

There are several ethical challenges that may arise with the development and adoption of systems that employ AI, causing impacts at different levels, including social, individual, and environmental issues. Some of them are (Bird et al., 2020; Muller, 2020):

- Privacy and surveillance: loss of personal privacy and abuse of mass monitoring;
- Behavior manipulation: use of information to influence people's opinion, manipulating their choices, and modifying their behavior;
- Obscurity in the decision criteria: lack of knowledge about which criteria the algorithms effectively used in decision-making processes;
- Biases in data and algorithms: dissemination of biases contained in the data used for training AI solutions or reflected in the development of algorithms for these tools;
- Accountability: absence of clear mechanisms to identify those responsible for possible damages caused by an AI solution.

In addition, this option also involves reputational risks that can affect the company’s competitiveness.

Option 2 allows the company to implement actions to mitigate ethical issues to offer a solution that obtains the maximum benefit from the technology while reducing the risks for individuals, the company, and society. In addition, investment in actions aimed at digital ethics can give the company a competitive advantage. However, this option requires investment in human and financial resources and

takes time to launch, which can cause the company to miss the timing for entering this market.

Option 3 allows the company to continue working with known solutions and risks. However, it does not allow the company to take advantage of the identified opportunity.

It is recommended that the instructor record the considerations on a class board using a table like the one below.

Table 2. Comparative analysis of options.

Option	Pros	Cons
1. Offer the solution immediately	Entry into a market with high global growth projections Possibility of expanding the customer portfolio Possibility of increasing revenue Availability of a tool aligned with the National Plan for Public Security and Social Defense Contribution to locating people wanted by the police	Degree of maturity and accuracy of facial recognition solutions Risks of discrimination against minority groups caused by algorithmic and data biases, loss of privacy, misuse of images, power asymmetry, among others Reputation risks
2. Delay the solution’s launch	Possibility of implementing risk mitigation actions related to ethical issues before offering the solution Positive social and reputational impacts	Necessary investment (time, money, training) to implement actions to mitigate ethical challenges Risk of missing the timing to become relevant in the market
3. Shelve the opportunity	Continuing to work with already-known solutions and risks	Not entering a market with high global growth projections

Note. Source: Elaborated by the authors.

3. If you were José Flávio, which option would you choose? Why?

After evaluating each option’s pros and cons, the instructor can ask the students which one they would adopt and ask them to explain the reasons.

The instructor can energize the discussion by adding elements that allow students to relate normative ethical theories to their reflections. They can be challenged to think more concretely considering the context, using questions such as: Would you choose the same option if you belonged to a minority group, more susceptible to biases? Is the option you chose more beneficial than problematic to the greatest possible number of stakeholders in the long term?

In the cases of options 1 and 2 (launch immediately or delay the launch), the instructor may ask what measures the company could adopt to mitigate the risks related to the ethical issues involved and what form of relationship with the public sector could be adopted.

The instructor could add aspects regarding the relationship between digital ethics and competitiveness if this issue did not emerge in the previous discussions. The

instructor may ask the participants again to see if there has been a change of opinion.

4. What principles can Nuveo follow for developing and adopting ethical AI systems?

International organizations, academic institutions, civil society, and businesses have joined forces and are committed to defining and establishing principles based on human rights and proposing ethical frameworks to be adopted for AI development and ethical use. Although there are several initiatives in this regard, there is still no universal framework or guide (Floridi & Cowls, 2019).

According to a report by the IDB — Inter-American Development Bank (Gómez Mont et al., 2020), two initiatives stood out: a) the OECD Principles (which Brazil adhered to and used to support the elaboration of the Brazilian Artificial Intelligence Strategy) and b) the guide produced by the European Union AI Expert Group. Another highlight of the IDB’s report is the guide developed by the Global Alliance established by the IEEE — Institute of Electrical and Electronic Engineers, which, in addition to providing guidelines for each situation analyzed within the

context of the ethical design of ‘autonomous and intelligent’ systems, has also been a reference for creating standards (IEEE P7000 series) for the ethical use of AI.

The article by [Floridi and Cows \(2019\)](#) is the result of the authors’ effort to consolidate relevant works in defining ethical principles for developing and adopting AI tools. In general terms, participants are expected to identify: (a) beneficence: development and adoption of AI tools that are beneficial to humanity, promoting well-being, preserving dignity, and being sustainable for the planet; (b) non-maleficence: development and adoption of AI tools that do not cause harm, preventing misuse, preventing breaches of privacy, and with safe limits of operation; (c) autonomy: balance of decision-making power attributed to AI systems; (d) justice: development and adoption of AI tools that promote prosperity and diversity, do not threaten solidarity, and do not cause discrimination and inequalities; and (e) explicability: promotion of transparency — developing systems that are understandable and interpretable — and accountability.

At this point, it would be possible to expand the discussions, asking participants: In addition to observing ethical principles, what recommendations would you give to developers of AI solutions? What does the recommended bibliography suggest?

5. Based on the principles discussed in the previous question, what concrete actions can be adopted by Nuveo in developing and implementing its AI solutions, aiming at an ethical performance?

In general, recommendations for the ethical development and use of AI involve: (a) technical aspects — such as adapting the architecture of systems, explanation methods, the concept of ‘X-by-design,’ and the definition of indicators of the systems’ technical reliability and robustness; (b) aspects of organizational structure and culture — such as accountability based on policies, codes of conduct, and governance frameworks, the promotion of diversity and inclusive design teams, and education and awareness to foster an ethical mindset; (c) educational aspects — that promote society’s digital literacy and that make it possible to understand AI’s potential and impacts; and (d) aspects related to legislation — aimed at protecting and guaranteeing human and social rights and encouraging the development of this technology ([European Commission, 2019](#); [IEEE, 2017](#); [Ministério da Ciência, Tecnologia e Inovação \[MCTI\], 2021](#)).

As education and legislation are outside Nuveo’s scope of action, participants are expected to address technical

aspects and organizational structure and culture, such as those described in Figure 2.

At the end of the class, it is recommended that the instructor summarize the main topics presented and encourage participants to report their conclusions.

TEACHING PLAN

This lesson plan was designed for a 1-hour 35-minute session. It can be adjusted according to the instructor’s interests and needs.

Warm up (10 minutes): The instructor can start the class by presenting a concrete ethical challenge regarding digital transformation, increasing the dynamism of the discussions. For example, the instructor can inform the students that their grade for class participation will be assigned by a tool that identifies the degree of attention and interest based on their facial expressions and check their reactions.

Introduction (20 minutes): The preparation questions aim to level the knowledge of the participants about the subjects approached in this teaching case. If the questions have not been presented in advance, the instructor can ask the students to discuss them in small groups. Alternatively, one or more participants can be encouraged to summarize these issues at the beginning of the class.

Ethics have been studied for more than 2,000 years. It is the branch of philosophy that deals with reflection on human conduct and proposes theories to establish moral principles ([Herschel & Miori, 2017](#)). Digital ethics also addresses reflection on human conduct, but with a focus on establishing moral principles that guide the design, implementation, and proper use of digital tools to minimize or avoid harmful effects they may cause to people and society. It is worth mentioning that the ethical use of AI is part of one of the pillars of the Brazilian Artificial Intelligence Strategy — EBIA ([MCTI, 2021](#)).

The instructor can check participants’ understanding of facial recognition technology by asking about its characteristics and the factors that impact its results. Wrap-up questions could include: What benefits can facial recognition solutions bring? What risks may be involved with using this technology for public safety purposes? The contributions brought by the participants can be summarized and written on the class board, as in the following example (Figure 1).

Then, the instructor should start a new discussion about the dilemma faced by the protagonist and the options presented in the case.

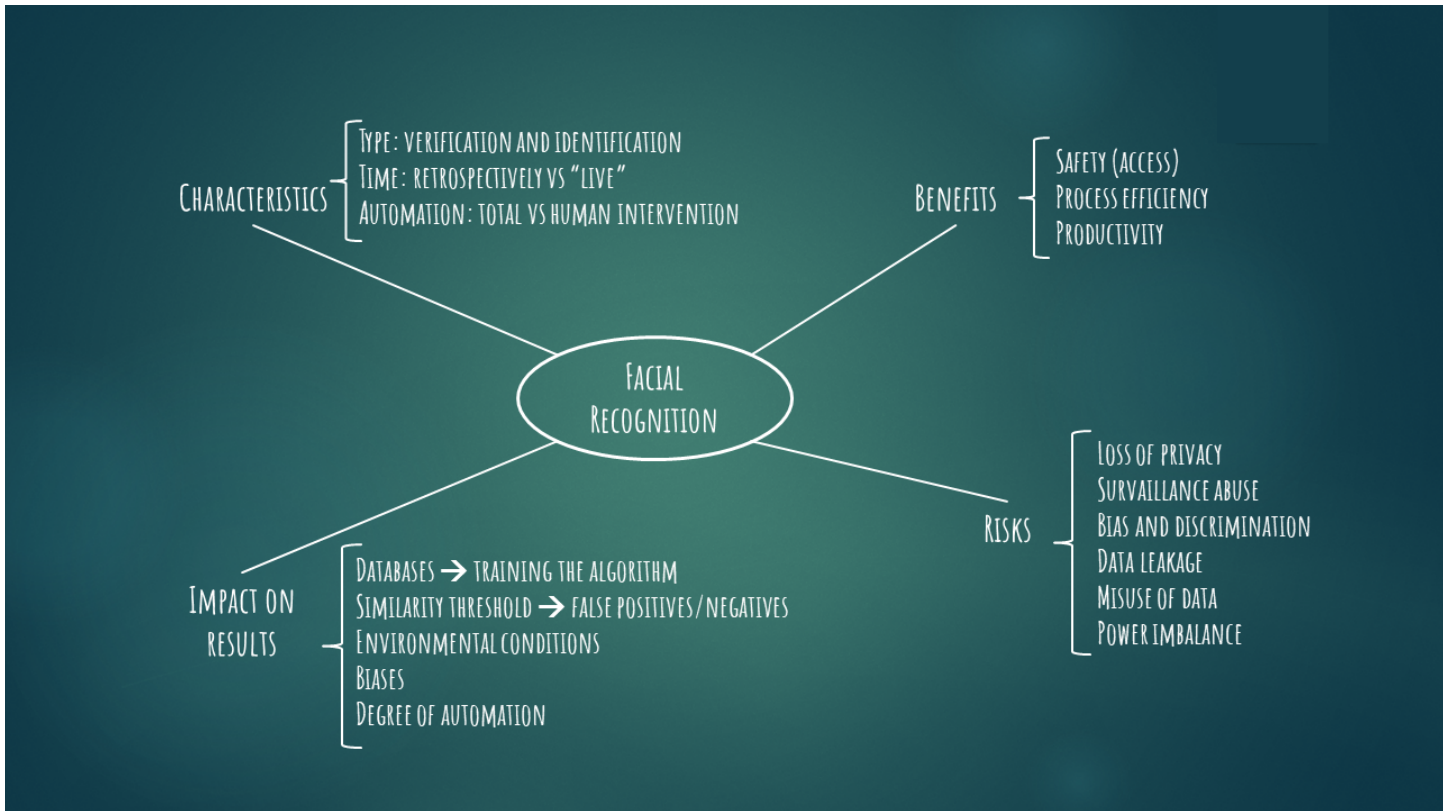


Figure 1. Facial recognition.

Source: Elaborated by the authors based on the [Centre for Data Ethics and Innovation \(2020\)](#).

Analysis of the options (30 minutes): At this stage of the discussion, the instructor may propose a quick poll: If you were the protagonist, would you choose option 1, 2, or 3? Then it will be possible to record contrasting opinions, allowing to collaboratively build a comparative table demonstrating each option's pros and cons. The instructor can also ask the participants if any other options could be envisaged and what would be the positive and negative points of this additional option.

Encouraging reflections in the light of normative ethical theories is recommended, leading participants to analyze the options considering moral values. The discussion can be based on several theories, such as teleological (focusing on the action result and assessing whether an act is ethical or not based on consequences), deontological (based on respect for other people's rights, considering ethical actions that do not violate these rights), and ethical virtue (puts virtues in the middle ground between excess and deficit) (Pollach, 2005; Vial, 2019).

When applicable, the instructor can also bring to the discussion some challenges of interaction with the public sector, such as the issues of infrastructure, governance, and

legal aspects. The discussion may approach possible forms of partnership between startups and governments.

Principles and recommendations related to digital ethics (20 minutes): After discussing the previous issues, the instructor may address the ethical challenges related to AI systems and some principles related to digital ethics. To make this step more interactive, the instructor can ask a participant to present and explain a principle, choosing other participants to complete the list. These reflections can comprehensively address possible ways of mitigating the ethical risks involved with the presented options, involving technical aspects, organizational structure, and culture.

The instructor can highlight the participants' contributions regarding mitigation measures on the class board, as in the following example (Figure 2).

Digital ethics and competitiveness (10 minutes): The instructor could add aspects regarding the relationship between digital ethics and competitiveness if this issue did not emerge in the previous discussions.

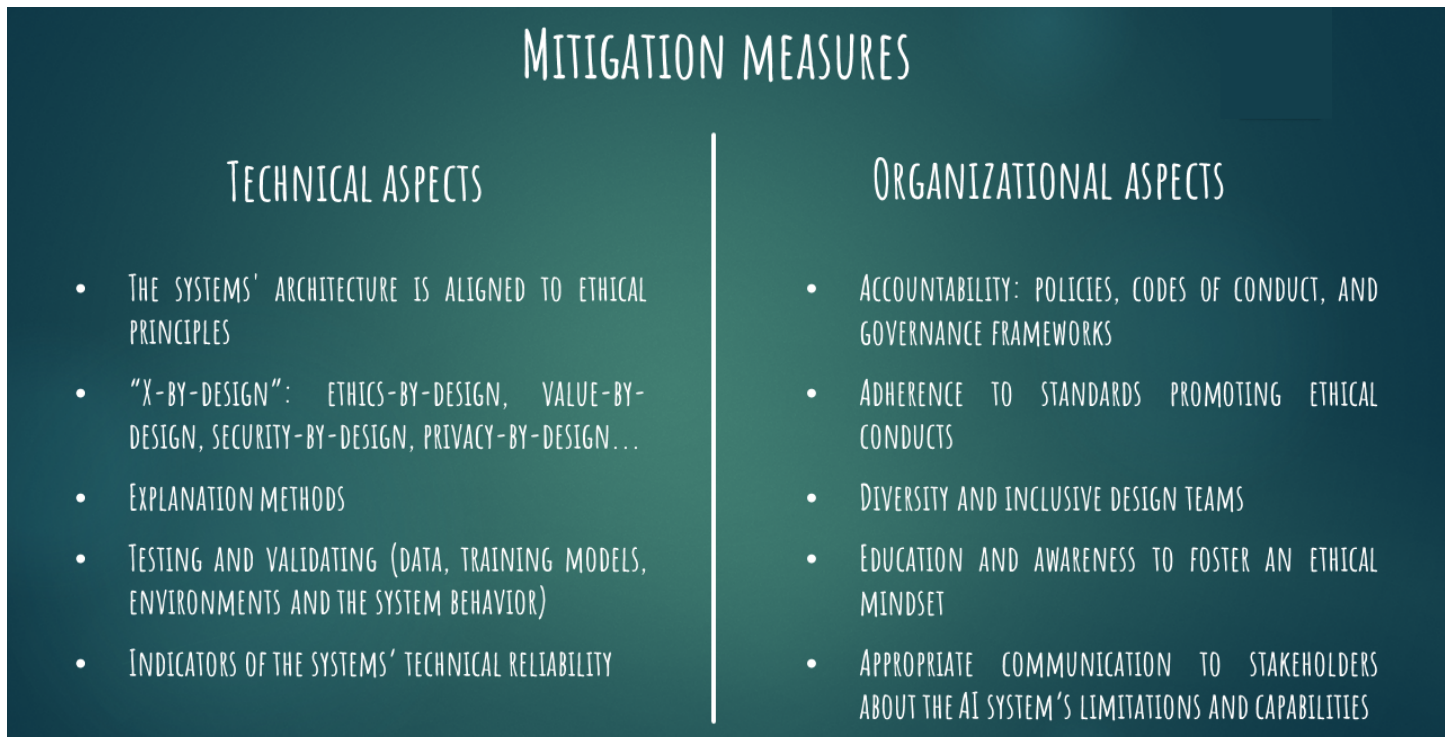


Figure 2. Mitigation measures.

Source: Elaborated by the authors based on the [European Commission \(2019\)](#).

It should be noted that society increasingly demands that companies not only aim at a profit but also consider issues related to ESG in their performance indicators. On the one hand, initiatives that are not in line with current ethical precepts and are considered morally offensive can affect the reputation of companies — and, consequently, their ability to sustain organizational performance (Vial, 2019). On the other hand, investing in digital ethics may be a source of competitive advantage (Hamer et al., 2020; Jones et al., 2018) by increasing trust in the company (Albinson et al., 2019).

A brief round of discussion can be conducted, and the instructor can check whether these new elements led to a change of heart regarding the choice of option.

Wrap-up (5 minutes): The instructor can end the activity by presenting a summary of the discussions and ask the participants (as homework) to prepare a mental map summarizing the topic addressed and their conclusions. This may contribute as a formative evaluation.

CASE DEVELOPMENTS

After analyzing the possible options at the time, José Flávio chose not to expand the use of his computer

vision tool for surveillance/public security purposes, as he understood that the stage of AI solutions and the difficulty of eliminating biases could, even involuntarily, intensify inequalities, intolerances, and social problems. He does not rule out that the company will work with this business in the future but decided to wait until the debates on the subject are more mature in academic circles and society.

Thus, to consolidate and expand the company in the AI market, Nuveo continued to focus on process automation and established strategic partnerships to embed its technology in devices close to the data-generating source, such as IoT devices (internet of things).

DATA SOURCES

This teaching case is based on a true story and was elaborated from data collected in semi-structured interviews with the founder and CEO of Nuveo, José Flávio Pereira. Due to the need for social distancing imposed by the COVID-19 pandemic, the interviews conducted in the second semester of 2021 were carried out using videoconferencing tools. To provide more fluency to the text, the opinions of experts consulted by José Flávio were compiled and transformed into dialogues

idealized by the authors, using fictional characters (Maria, Pedro, and Arthur).

Secondary data were collected from information available on videos and news articles found on the

internet, the company's website and social media, and José Flávio's social media account. Public information and reports about the context of the case can be consulted in the references.

Table 3. Secondary data.

Consulted material	Sources
Company's website	https://www.nuveo.ai/
Company's social media	https://www.linkedin.com/company/nuveo/
Interview with the CTO of Nuveo about the company and AI	https://www.youtube.com/watch?v=x25s0Qwm1q4
Video of the company's presentation at the Data Science Challenge @ ITA Edition 2021	https://www.youtube.com/watch?v=BFAuwRjIMNs
Presentation by the CTO about the company and Nuveo's solutions at the 2nd Computing Week at USP/Rib. Preto — 2020	https://www.youtube.com/watch?v=lnTnM3TKQgA
Article about the company	https://www.techtudo.com.br/noticias/2018/10/tecnologia-automatiza-processos-burocraticos.ghtml
Video of the company's CEO/founder about the organization	https://www.youtube.com/watch?v=i-TIRfRYZAY
Distrito report on AI startups	https://materiais.distrito.me/dataminer-report-inteligencia-artificial
Professional social media account of the company's CEO/founder	https://www.linkedin.com/in/joseflaviopereira/

Note. Source: Elaborated by the authors.

REFERENCES

- Alberton, A., & Silva, A. B. da. (2018). Como escrever um bom caso para ensino? Reflexões sobre o método. *Revista de Administração Contemporânea*, 22(5), 745–761. <https://doi.org/10.1590/1982-7849rac2018180212>
- Albinson, N., Balaji, S., & Chu, Y. (2019). Building digital trust: Technology can lead the way. *Deloitte Insights*. <https://www2.deloitte.com/lu/en/pages/innovation/articles/building-long-term-trust-in-digital-technology.html>
- Bird, E., Fox-Skelly, J., Jenner, N., Larbey, R., Weitkamp, E., & Winfield, A. (2020). *The ethics of artificial intelligence: Issues and initiatives*. European Parliament, Directorate-General for Parliamentary Research Services. <https://doi.org/10.2861/6644>
- Buytendijk, F., Willemsen, B., Jones, L. C., & Hare, J. (2020). Digital ethics: Use facial recognition technology responsibly. *Gartner*. <https://www.gartner.com/en/documents/3990243>
- Centre for Data Ethics and Innovation. (2020). *Snapshot series: Facial recognition technology report*. <https://www.gov.uk/government/publications/cdei-publishes-briefing-paper-on-facial-recognition-technology>
- Ceron, R. (2019, December 5). AI, machine learning and deep learning: What's the difference? *IBM*. <https://www.ibm.com/blogs/systems/ai-machine-learning-and-deep-learning-whats-the-difference/>
- Corea, F. (2018, August 29). AI knowledge map: How to classify AI technologies. *Medium*. <https://francesco-ai.medium.com/ai-knowledge-map-how-to-classify-ai-technologies-6c073b969020>
- Crespo, M. (2022, January 16). *Metaverso: Teremos leis para reger o novo mundo?* LinkedIn. <https://www.linkedin.com/pulse/metaverso-teremos-leis-para-reger-o-novo-mundo-crespo-ph-d-ccep-i/>
- European Commission, Directorate-General for Communications Networks, Content and Technology. (2019). *Ethics guidelines for trustworthy AI*. <https://doi.org/10.2759/346720>
- European Commission, Directorate-General for Research and Innovation, European Group on Ethics in Science and New Technologies. (2018, March 9). *Statement on artificial intelligence, robotics and 'autonomous' systems*. <https://doi.org/10.2777/786515>
- Faria, M., & Figueiredo, K. F. (2013). Casos de ensino no Brasil: Análise bibliométrica e orientações para autores. *Revista de Administração Contemporânea*, 17(2), 176–197. <https://doi.org/10.1590/s1415-65552013000200004>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>

- Gómez Mont, C., Pozo, C. M. D., Pinto, C. M., & Alcocer, A. V. M. D. C. (2020). A inteligência artificial a serviço do bem social na América Latina e no Caribe: Panorama da região e retrato de doze países. *Banco Interamericano de Desenvolvimento*. <https://doi.org/10.18235/0002393>
- Hamer, P. D., Gove, K., Willemssen, B., Struckman, C., & Buytendijk, F. (2020). Digital ethics: From compliance duty to competitive differentiator. *Gartner*. <https://www.gartner.com/en/documents/3989522>
- Herschel, R., & Miori, V. M. (2017). Ethics & big data. *Technology in Society*, 49, 31–36. <https://doi.org/10.1016/j.techsoc.2017.03.003>
- Hinds, J., Williams, E. J., & Joinson, A. N. (2020). “It wouldn’t happen to me”: Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, 143, 102498. <https://doi.org/10.1016/j.ijhcs.2020.102498>
- Institute of Electrical and Electronics Engineers (2017). Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems (Version 2). *IEEE*. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf
- Jones, L. C., Buytendijk, F., & Hare, J. (2018). Data ethics enables business value. *Gartner*. <https://www.gartner.com/en/documents/3894128>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. Dartmouth College.
- McKendrick, J. (2019, July 7). Ethical artificial intelligence becomes a supreme competitive advantage. *Forbes*. <https://www.forbes.com/sites/joemckendrick/2019/07/07/ethical-artificial-intelligence-becomes-a-supreme-competitive-advantage/?sh=3967356f1a8f>
- Ministério da Ciência, Tecnologia e Inovação (2021). *Estratégia brasileira de inteligência artificial*. <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>
- Muller, V. C. (2020). Ethics of artificial intelligence and robotics. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/ethics-ai/>
- National Institute of Standards and Technology. (2019, December 19). *NIST study evaluates effects of race, age, sex on face recognition software*. <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software>
- Pollach, I. (2005). A typology of communicative strategies in online privacy policies: Ethics, power and informed consent. *Journal of Business Ethics*, 62(3), 221–235. <https://doi.org/10.1007/s10551-005-7898-3>
- Tufekci, Z. (2016). Machine intelligence makes human morals more important. *TEDSummit*. https://www.ted.com/talks/zeynep_tufekci_machine_intelligence_makes_human_morals_more_important
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, 28(2), 118–144. <https://doi.org/10.1016/j.jsis.2019.01.003>
- Weinhardt, M. (2020). Ethical issues in the use of big data for social research. *Historical Social Research*, 45(3), 342–368. <https://doi.org/10.12759/hsr.45.2020.3.342-368>

Authorship

Gabrielle Marques Castelo Branco Biondi

Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, Mestrado Profissional em Gestão para Competitividade - TI

Av. 9 de julho, n. 2029, Bela Vista, CEP 01313-902, São Paulo, SP, Brazil

E-mail: gbranco@yahoo.com

 <https://orcid.org/0000-0002-2713-8590>

Adrian Kemmer Cernev*

Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, Departamento de Tecnologia e Ciência de Dados

Av. 9 de julho, n. 2029, Bela Vista, CEP 01313-902, São Paulo, SP, Brazil

E-mail: adrian.cernev@fgv.br

 <https://orcid.org/0000-0002-6300-3976>

* Corresponding Author

Funding

The authors state that there was no financial support for the research in this article.

Copyrights

RAC owns the copyright to this content.

Plagiarism Check

RAC maintains the practice of submitting all documents approved for publication to the plagiarism check, using specific tools, e.g.: iThenticate.

Authors' Contributions

1st author: conceptualization (equal); data curation (lead); investigation (lead); methodology (supporting); project administration (lead); visualization (lead); writing – original draft (lead); writing – review & editing (equal).

2nd author: conceptualization (equal); investigation (supporting); methodology (lead); project administration (supporting); supervision (lead); validation (lead); visualization (supporting); writing – original draft (supporting); writing – review & editing (equal).

Conflict of Interests

The authors have stated that there is no conflict of interest.

Peer Review Method

This content was evaluated using the double-blind peer review process. The disclosure of the reviewers' information on the first page, as well as the Peer Review Report, is made only after concluding the evaluation process, and with the voluntary consent of the respective reviewers and authors.

Data Availability

RAC encourages data sharing but, in compliance with ethical principles, it does not demand the disclosure of any means of identifying research subjects, preserving the privacy of research subjects. The practice of open data is to enable the reproducibility of results, and to ensure the unrestricted transparency of the results of the published research, without requiring the identity of research subjects.

RAC is a member of, and subscribes to the principles of the Committee on Publication Ethics (COPE) for scholarly publication

