ISSN 2949-2483

Volume 3 Number 1

JOURNAL OF DIGITAL TECHNOLOGIES AND LAW

2025

ELECTRONIC SCIENTIFIC AND PRACTICAL JOURNAL

JOURNAL OF DIGITAL TECHNOLOGIES AND LAW
Volume 3, No. 1
2025
DOI: 10.21202/2949-2483.2025.1
ELECTRONIC SCIENTIFIC AND PRACTICAL JOURNAL
eISSN 2949-2483
Published since 2023, frequency – 4 issues a year. DOI: 10.21202/2949-2483

Address: 42 Moskovskaya Str., 420111 Kazan, Russian Federation
Tel.: +7 (843) 231-92-90
Fax: +7 (843) 292-61-59
E-mail: lawjournal@ieml.ru
Website: https://www.lawjournal.digital
Telegram: https://t.me/JournalDTL_world
VKontakte: https://vk.com/JournalDTL
Yandex.Dzen: https://dzen.ru/JournalDTL
Odnoklassniki: https://ok.ru/JournalDTL

Founder and publisher of the Journal

Kazan Innovative University named after V. G. Timiryasov. Address: 42 Moskovskaya Str., 420111 Kazan, Republic of Tatarstan, Russian Federation. Tel.: +7 (843) 231-92-90. Fax: +7 (843) 292-61-59. E-mail: info@ieml.ru. Website: https://ieml.ru

Certificate on registering a mass medium: EL no. FS 77-84090 of 21.10.2022, issued by Roskomnadzor.
Territory of distribution: Russian Federation, foreign countries.

Date of signing the issue for publication: 2025, March 25. Hosted on the website https://www.lawjournal.digital: 2025, March 30.

JOURNAL OF DIGITAL TECHNOLOGIES AND LAW
Volume 3, No. 1
2025
DOI: 10.21202/2949-2483.2025.1
ELECTRONIC SCIENTIFIC AND PRACTICAL JOURNAL
eISSN 2949-2483

Content

Ndiyun R. K., Mukonza R. M.
Digital Transformation of Civil Registration System in Cameroon: Innovations in E-governance 7

Tzimas Th.
Evolution of Copyright in the Era of Artificial Intelligence: Analysis of Conflicts of Law and Judicial Precedents 35

Bhatt N.
Crimes in the Age of Artificial Intelligence: a Hybrid Approach to Liability and Security in the Digital Era 65

Ilin I. G.
Constitutional-legal Aspect of Creating Large Language Models: the Problem of Digital Inequality and Linguistic Discrimination 89

Rakhmetov B., Khaizabekov K.
Behavioral Biometrics in the European Union: Legal Challenges and Technological Prospects 108

Hadi M. A., Abdulredha M. N.
A General Information Security Governance Framework for Public and Private Organizations under Laws and Regulations 125

Correia P. M. A. R., Pedro R. L. D., Videira S.
Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection 143

Research article
UDC 34:004:347.6:004.9:004.056
EDN: https://elibrary.ru/eaqyqj
DOI: https://doi.org/10.21202/jdtl.2025.1

Digital Transformation of Civil Registration System in Cameroon: Innovations in e-Governance

Robert Kosho Ndiyun ✉
Tshwane University of Technology, Pretoria, South Africa

Ricky Munyaradzi Mukonza
Tshwane University of Technology, Pretoria, South Africa

Keywords

Cameroon,
civil status act,
data protection,
digital literacy,
digital technologies,
e-governance,
e-government,
law,
legislation,
public service

Abstract

Objective: to study the innovative transformations in the field of e-Governance introduced into Cameroon's civil registration system during the 2024 legislative reforms. The focus is on assessing the impact of these transformations on improving governance efficiency, transparency, accessibility of services for citizens, as well as improving statistical accounting of vital events.

Methods: the work uses general scientific methods of analysis and synthesis, classification, systematic and functional approaches, as well as formal legal and comparative legal methods.

Results: the research shows that measures like introduction of electronic declaration of civil status acts, creation of a national database and transition to electronic certificates can dramatically improve the efficiency and accessibility of services for the population. However, the authors emphasize that the successful implementation of digital innovations requires overcoming significant barriers, such as insufficient technological equipment, limited Internet access, and low digital literacy of citizens. These challenges make it necessary to develop additional regulatory and support mechanisms. Particularly important is the balance between digitalization and ensuring the rights of citizens in the context of electronic registration.

✉ Corresponding author

Scientific novelty: the work provides unique empirical data on digitalization of public services in Cameroon. This is especially important for the countries of the global South, where such transformations are slow and fragmentary. The study makes a significant contribution to the scientific debate by expanding understanding of digital technology adoption models through the lens of expected usefulness and perceived ease of use in developing countries.

Practical significance: recommendations for legislators, government officials and other stakeholders were developed. The authors emphasize the need to adopt a regulatory framework as soon as possible, introduce educational programs for employees and citizens, and ensure access to digital technologies. These measures aim at creating a sustainable infrastructure for an effective transition to electronic systems and improving the quality of public services. The work contributes to the study of public governance digitalization, offering both theoretical concepts and practical solutions that can be adapted for other countries with similar challenges.

For citation

Ndiyun, R. K., & Mukonza, R. M. (2025). Digital Transformation of Civil Registration System in Cameroon: Innovations in e-Governance. *Journal of Digital Technologies and Law, 3*(1), 7–34. https://doi.org/10.21202/jdtl.2025.1

Contents

Introduction
1. Literature review
   1.1. Conceptualising e-Governance
   1.2. E-Governance approaches and domains
   1.3. The significance and relevance of e-Governance
   1.4. The pitfalls of e-Governance
2. Theoretical framework
3. Methodology
4. Results: Overview of electronic innovations in Cameroon's new law
   4.1. Technological Innovations
   4.2. Service Delivery Enhancements
   4.3. Governance and Accountability
   4.4. Data Protection and regulatory frameworks
   4.5. Technological infrastructure in Cameroon
5. Discussions
6. Recommendations
Conclusion
References

Introduction

Information and Communication Technology's importance in boosting governance and service delivery has become central in a progressively interconnected world. As countries struggle for more efficiency and transparency, incorporating electronic governance (e-Governance) mechanisms has developed as a transformative tool in public administration (Bannister & Connolly, 2012; Oliveira et al., 2020). Countries have adopted and implemented these tools to varying degrees. The United Nations (UN) reported that 161 countries offer online platforms for the application of birth certificates, 152 countries provide digital services for marriage certificates, and 151 countries offer electronic application services for death certificates, recording an increase of 3, 1 and 8 % respectively from 2022 data[1]. This increase in the computerisation of vital events declaration and registration reflects the importance of civil status systems to governance. While most countries in the Global North have registered full digitalisation in these services, the reverse holds for the Global South, particularly in Africa where the full digitalisation rate stands at 11 % for birth certificate applications and 7 % for marriage and death certificates[2]. In Cameroon, the public sector is transitioning to electronic service delivery processes[3]. This transformation is contained in the 2016 National ICT Strategic Plan by 2020, and the National Development Strategy by 2030 which highlight the digitalisation of public governance to achieve sustainable development and uplift Cameroon to an emerging economy by 2035 (Sevidzem, 2024).

The constantly developing technologies are parallel with citizens' demands, requiring governments to align public service delivery to ICT. This paper focuses on the electronic innovations introduced in civil status registration systems in Cameroon by Law No. 2024/016 of 23 December 2024. This law resonates with the policy documents mentioned above, which include digitising the civil status and accompanying the digital transition with appropriate institutions and ICT tools within all government institutions (Sevidzem, 2024). Civil status registration – an indispensable tool that reinforces citizenship rights and societal organisation – stands at the lead of these policy innovations. The ability to correctly and efficiently record vital events such as births, marriages, and deaths is fundamental to national identity and social equity.[4] Also, civil

1 United Nations. (2024). E-Government Survey, 2024. UN (2020). https://clck.ru/3Gf8K

2 Ibid.

3 Sindeu, E. (2013). Implementation of e-government in Cameroon. In 7th Annual E-Government Forum, Muyuno, Uganda, 25–27 March (pp. 1–24).

4 World Economic Forum. (2022, October 22). Civil registrations and vital statistics: Here's why they're fundamental to society. https://clck.ru/3Gf8Qt

status is crucial in establishing the legal identity of citizens, guaranteeing the exercise of their rights and facilitating their engagement in the social and political life of the country[5].

The need for improvement in civil status registration in Cameroon had become particularly persistent. The country's civil status registration system has historically encountered multiple barriers, including bureaucratic inefficiencies, limited access to services, and a lack of public awareness and low birth registration rate with over seven million citizens currently without birth certificates[6]. After independence and reunification, Cameroon adopted Law No. 68-LF-2 OF 11 June 1968 to govern civil status systems. This law was replaced by Ordinance No. 81-02 of 29 June 1981, and later modified by Law No. 2011/011 of 6 May 2011. At the operational level, the National Civil Status Bureau (BUNEC) instituted in 2011 was only created in 2013 by a presidential decree on the organization and functioning of the structure[7]. This institution in charge of the supervision, control, regulation and evaluation of the national civil status system only went operational in 2016, and by 2023 did have an official comprehensive record of civil status events in Cameroon[8].

The process of transforming Cameroon's civil status registration was initiated in 2007 with Cameroon's Civil Status Rehabilitation Programme (PRE2C). In 2010, Cameroon adhered to the African Programme on Accelerated Improvement of Civil Registration and Vital Statistics (APAI-CRVS), and later adopted the 2018-2022 strategic plan for civil status rehabilitation. By 2015, Prime Ministerial Order No. 019/CAB/PM of 24 February 2015, a steering committee was created to ensure the progress of PRE2C, chaired by the Minister of Decentralization and Local Development[9]. This was followed by the adoption of a master plan for the computerisation of the national civil registry system for the period 2019-2023. These activities contributed to the 2024 law on civil status registration systems in Cameroon.

Cameroon's civil status system is semi-decentralised, with local civil status centres and BUNEC at the central level. The decentralised centres comprise primary civil status

5 National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

6 Ayang, M. (2024, November 29). Cameroon introduces draft laws on civil registration reform, data protection. Biometric Updates. https://clck.ru/3Gf8ZM

7 National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

8 Ibid.

9 Zewoldi, Y. (2019). Snapshot of Civil Registration and Vital Statistics in Cameroon. Centre of Excellence for Civil Registration and Vital Statistics. https://clck.ru/3Gf8q5

centres at the level of Councils, and secondary civil status centres in villages and neighbourhoods. Cameroon currently counts 375 main civil status centres in Cameroon and 56 in Consular Posts and Diplomatic Missions, and 2455 secondary centres[10]. The registration of vital events at these centres is carried out by Civil Status Registrars – City Mayors and their Deputies, Mayors and their Deputies assisted by Secretaries (Ordinance No. 81-02, article 7).

Notwithstanding these measures, the declaration and registration of births and deaths in Cameroon is undermined by poor knowledge of the importance of civil status certificates, the non-mastery of the legal frame work governing civil status, financial constraints, accessibility to centres by enclaved population, and administrative bottlenecks.[11] The promulgation of the 2024 law marks a fundamental innovation in addressing these issues through implementing electronic systems framed to modernise processes and boost citizen participation. This law mirrors a commitment to transforming public administration and improving the overall quality of governance in Cameroon[12]. By 2018, the birth registration rate in Cameroon stood at 54 %, while death registrations for 2020 were at 9.71 %[13].

While there is increasing body of research on e-Governance and its bearings across various domains, limited scholarship has specifically focused on the electronic innovations of civil status registration in the Cameroonian context introduced by the 2024 law. Existing research on e-Governance in Cameroon has focused on taxation procedures (Djossa-Tchokoté et al., 2024), local public service delivery (Sevidzem, 2024), e-particpation (Xin et al., 2023), general e-Governance implementation[14].

This study seeks to fill the gap by critically investigating the e-Governance innovations introduced by the 2024 law and their potential to reform civil registration operations in Cameroon.

10 National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

11 Ibid.

12 Ayang, M. (2024, November 29). Cameroon introduces draft laws on civil registration reform, data protection. Biometric Updates. https://clck.ru/3Gf8ZM

13 National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

14 Sindeu, E. (2013). Implementation of e-government in Cameroon. In 7th Annual E-Government Forum, Muyuno, Uganda, 25–27 March (pp. 1–24).

The research's central question is: How can e-Governance innovations in civil status registration improve administrative efficiency and citizen involvement in Cameroon? To harness this problem, the study poses the following research questions: (1) What key electronic innovations have been introduced by the 2024 law on civil status registration in Cameroon? (2) How do these innovations impact service delivery and governance in the civil status registration sector? And (3) What challenges and opportunities do stakeholders face in implementing the innovative system?

The paper aims to: (1) explore the specific electronic innovations introduced by the 2024 law, (2) evaluate their implications for service delivery and governance, and (3) identify the barriers stakeholders may encounter during implementation and propose strategies to overcome them.

The research outcomes demonstrate that establishing a National Civil Status Database and the introduction of electronic declarations can significantly enhance the efficiency and accessibility of civil registration services[15]. However, challenges such as technological infrastructure gaps and public resistance to change must be addressed to achieve the full potential of these innovations.

This paper is structured as follows: The following section explores a comprehensive literature review, emphasising main themes and gaps related to the topic, followed by a theoretical framework. The methodology section presents the research design and data collection methods used in the study. The findings section details the results of the analysis, followed by a discussion that contextualises the results within the e-Governance broader landscape. The paper concludes with recommendations for ameliorating implementation and maximising the effectiveness of civil status registration in Cameroon.

1. Literature review

1.1 Conceptualising E-Governance

E-Governance denotes a transformative move that utilises information and communication technology (ICT) to boost the efficiency and effectiveness of public service delivery. It incorporates all aspects of government operations that involve ICT, including service delivery, citizen participation, and data management to achieve accountability, effectiveness and transparency of political processes (Grigalashvili, 2022; Heeks, 2006). This definition aligns with Umbach & Tkalec (2022) who suggest that it entails the application of digital machineries to improve public service delivery to citizens. Furthermore, e-Governance is viewed as a reciprocal interface between

[15] Ayang, M. (2024, November 29). Cameroon introduces draft laws on civil registration reform, data protection. Biometric Updates. https://clck.ru/3Gf8ZM

the government and its internal and exogenous stakeholders employing digital tools.[16] E-Governance transcends mere digitisation of services and aims to promote a more responsive and accountable public service by enhancing real-time relations between the population and public institutions (Heeks, 2006). This revolution is particularly essential in civil registration, where the incorporation of technology can rationalise hitherto traditionally sluggish and bureaucratic operations. Effective e-Governance requires robust technological infrastructure and appropriate planning to establish a more citizen-centric system.

While some scholars (Palvia & Sharma, 2007) contend that there is no difference between e-government and e-Governance, Grigalashvili (2022) argues that the two concepts and processes aim to digitally ameliorate government-user interaction, with the former being a substantial component of the latter. E-Governance aims to enhance the government's effectiveness through improved information flow and policy-making between the government and the population by employing ICT tools. By facilitating the exchange of information to achieve policy goals (Muttoo et al., 2019), e-Governance creates an informed community, reducing government intervention time interval, and enhancing service delivery and public participation in a digital process (Umbach & Tkalec, 2022).

Some countries have successfully applied e-Governance innovations in civil registration. For instance, Estonia's e-residency system has been applauded as a model for digital governance, offering residents and citizens protected access to various government services online, including civil registration (Tammpuu & Masso, 2018; Kattel & Mergel, 2019). Likewise, India's Digital India approach aims to enhance the user-friendliness of civil registration services through mobile applications and online systems (Suthar et al., 2019). These case studies demonstrate that successful e-Governance application often entails robust stakeholder commitment, vigorous technological set-up, and a resolve to constant upgrading. The progress in e-Governance systems highlights a shift towards a participatory management approach that empowers citizens. However, governance innovations should integrate technological advancement and socially inclusive sensitive approaches. The effectiveness of e-Governance innovations significantly relies on contextual features such as available technological facilities, public trust in government, and legal mechanisms (Suthar et al., 2019).

[16] Oyedokun, G., Adeolu-Akande, M., & Oyedokun, D. (2022). Assessing the Status and Challenges of e-Governance and e-Public Services Delivery in Nigeria. BAM 2022 Conference, University of Manchester (pp. 1–15).

1.2. E-Governance approaches and domains

Scholars have proposed various models adopted in e-Governance describing the movement of data and services between the service provider – the government –and the users – citizens (Grigalashvili, 2022; Halachmi, 2004; Prashar & Bawa, 2023). These approaches include e-advocacy, critical flow and comparative analysis (Prashar & Bawa, 2023). The critical flow model contends with the swift flow of vital information to the targeted recipients with the aid of ICT. The comparative analysis model concentrates on finding best practices in e-Governance and applying them as benchmarks to test other management systems (Halachmi, 2004; Grigalashvili, 2022). The e-advocacy model emphasises strengthening the various public domains to influence government policies through their input and feedback (Albert, 2009). The critical flow model's strength lies in its leverage on ICTs to achieve instant information transfer, reducing distance and time. The comparative analysis model on the other hand draws its strength from the unlimited capacity of electronic systems to stock valuable information and retrieve and disseminate it instantly across numerous barriers (Halachmi, 2004). On its part, the e-advocacy model, one of the widely used digital administration approaches, holds its strength in mobilising and pulling human resources and information beyond institutional and physical obstacles and implementing it for specific action.

Researchers have classified e-Governance and e-government into various types or domains, based on digital interactions between the government, citizens, economic sector, employees and other non-profit, political and social institutions (Fang, 2002). These interactions have been grouped into the following domains with a two-way digital interaction: Government-to-Government (G2G); Government-to-Business (G2B); Government-to-Citizens (G2C); Government-to-Employees (G2E); and Government-to-Nonprofit (G2N) (Fang, 2002; Kaisara & Pather, 2011). G2G involves digital cooperation, communication, information and commodities between government departments or agencies to enhance government efficiency and effective coordination. G2B preoccupies with digital communication between governments and businesses, including e-procurements, e-business incorporation and fiscal transactions. G2C focuses on creating electronic public services by the government, for access by the citizens. G2E concerns how governments undertake initiatives that facilitate the internal flow of information among state employees to digitise civil service processing and management systems. Finally, G2N entails information communication and transactions between the government and nonprofit organisations, political parties and other social groups.

1.3. The significance and relevance of e-Governance

Implementing the various e-Governance types discussed above provides multiple merits for the government and other actors. E-Governance enhances citizen engagement in policymaking procedures, achieving improved accountability (Schuppan, 2009). E-Governance eliminates barriers displacement cost, and time by citizens to participate in government processes. These barriers are removed through the effortless availability of public service information and improved mechanisms to liaise with the government (Sharma et al., 2021). By participating in government decision-making through management, citizens become co-initiators of decisions concerning them, in collaboration with the government. Digital governance therefore increases citizen participation across various segments of society such as economic, socio-cultural and geographical domains.

Moreover, e-Governance significantly influences the efficiency and accessibility of public service delivery through information diffusion (Sharma et al., 2021). By digitising procedures, states considerably cut processing times, enhancing user fulfilment and promoting higher participation rates (Halachmi, 2004). E-Governance ameliorates the quality of information communicated with its partners. Studies have revealed that countries applying online civil registration systems witness reduced waiting times and enhanced accuracy in record-keeping, which are essential for effective administration (Suthar et al., 2019). For instance, adopting digital birth and death registration in countries like Ghana has led to a notable increase in registration rates, demonstrating the potential of e-Governance to improve essential services (Suthar et al., 2019). The introduction of the m-Birth computerised infant birth registration system in Ghana resulted in a 15.5 % increment in birth registration between 2014 and 2017[17]. In March 2023, Kenya also introduced an electronic service for registering births and deaths[18], leading to an increase in under-5 birth registration to 76 % as opposed to 65% in 2014[19].

Furthermore, the e-Governance system's effectiveness can be weighed through various lenses, including user fulfilment, transparency, service delivery speed, and compliance rates. Heeks (2003) contends that e-Governance implementation is slow in developing countries due to the delay or non-implementation of projects and programmes, and the adverse outcome of implemented projects. Due to the scarcity of expertise causing digital illiteracy, most states encounter a digital divide, undermining

[17] UNICEF. (2018). Assessment of the m-birth project in Ghana. https://clck.ru/3Gf9Jp

[18] Njoya, S. (2023, February 6). Kenya to start issuing digital death and birth certificates. https://clck.ru/3GfFXf

[19] UNICEF. (2023). Country Office Annual Report 2023: Kenya. https://clck.ru/3Gf9MH

successful e-Governance (Naqvi et al., 2021). This highlights that most countries in the global south experience limited e-Governance implementation due to political, economic and socio-cultural factors. Studies also reveal that countries employing inclusive monitoring and evaluation frameworks like the case of South Africa during COVID-19 (Naqvi et al., 2021), are more furnished to measure the effectiveness of their e-Governance processes (Suri & Sushil, 2017). These tools enable governments to detect improvement sectors and frame data-driven policies that promote service delivery.

1.4. The pitfalls of e-Governance

Notwithstanding the numerous benefits of e-Governance innovations discussed above, the process also has disadvantages. One of these is the vulnerability of the data which transit along e-Governance platforms. Muttoo et al., (2019) suggest that inadequately designed and implemented e-Governance processes can expose government and citizen's data to cyber security threats and illegal third-party access. This presents the need to integrate data safeguard tools in implementing e-Governance technologies. The threats of data loss and unauthorised access deter some countries from switching from paper systems to electronic processes (Munyoka, 2020), likewise, citizens fear possible government misuse of their information captured in electronic systems without their consent (Makwanya, 2022). The likelihood of public information misuse by unauthorised state institutions is evident, while the unconsented dissemination of citizens' personal information in electronic systems can undermine their right to privacy[20].

Furthermore, segments of the population in areas with inadequate internet connectivity – due to lack of facilities and low purchasing power – can become more deprived and sidelined from electronic innovations in service delivery, engraining the digital gap in society (Foster, 2020). E-governance implementation requires electronic tools, and most importantly internet literacy among service users. The absence of these facilities and knowledge may result in social segregation and relegation (Palvia & Sharma, 2007). Moreover, not all public sector employees overseeing e-Governance implementation possess the required technology proficiency[21]. Consequently, rich and technology-literate citizens and those in areas with internet connection will possibly access better public service than their poor and illiterate counterparts, and those in areas with poor or low internet connectivity. These demonstrate the noteworthy setbacks of digitising

[20] Zungu, S. (2024). The use of monitoring and evaluation as an improving e-Governance for enhanced service delivery: Master thesis. University of Johannesburg.

[21] Ibid.

service delivery (Coe et al., 2001). This challenge is visible in Cameroon due to substantial infrastructural disparities across regions, restricting citizen participation in government action through e-Governance to persons living in urban areas and with resources to purchase electronic gadgets and access internet connection.

2. Theoretical framework

The theoretical lens for assessing the e-Governance innovations in civil registration in Cameroon is grounded in two main theories: e-Governance theory and the Technological Acceptance Model (TAM). These theories provide a solid base for understanding how technological innovations can transform public service delivery and facilitate accessibility. The theories offer a comprehensive perspective through which the electronic innovations of the new civil registration law can be measured.

E-Governance theory suggests incorporating information and communication technologies (ICTs) into government processes to improve service delivery, efficiency, transparency, and accessibility (Heeks, 2006). This theory accentuates the transformative strength of ICTs in enhancing public service delivery and promoting greater citizen participation. In the case of Cameroon's new civil registration law, e-Governance theory contends that adopting electronic systems can streamline registration operations, diminish bureaucratic delays, and improve the accuracy and consistency of civil records. By leveraging digital tools, the state can guarantee that citizens have easier access to registration facilities, fostering inclusivity and mitigating barriers to participation (Lubis et al., 2024).

The Technology Acceptance Model (TAM) is a solid theoretical lens for understanding user acceptance of electronic innovations, especially in the case of civil status registration in Cameroon. TAM suggests that perceived usefulness and perceived ease of use are the main factors influencing users' behaviour towards technology adoption (Davis et al., 1989). With electronic civil status registration, the perceived usefulness may be triggered by the efficiency and accessibility provided by digital systems, permitting citizens to declare vital events such as births, marriages, and deaths more expediently. Furthermore, the perceived ease of use is essential, as citizens must find the electronic process spontaneous and straightforward to access. This model potentially identifies challenges to adoption, such as technological literacy and infrastructure barriers, which are particularly pertinent in the Cameroonian context where digital literacy fluctuates significantly among the citizens.[22] The TAM has been applied in various fields of study,

[22] National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

including e-commerce (Araújo & Casais, 2020), e-learning, (Al-Gahtani, 2016) and e-justice (Reiling & Contini, 2022). By emphasising the constructs of expected usefulness and perceived ease of use, this research evaluates how these factors impact user attitudes and eventually their goal to adopt the electronic civil status registration system. This is particularly pertinent in Cameroon, where the government is modernising public services through digital technologies.

3. Methodology

This study is qualitative drawing on existing literature and legal and policy documents. This design allows for a detailed exploration of the new civil registration law in Cameroon and its e-Governance innovations. Qualitative research is appropriate for examining intricate phenomena and understanding the setting in which they ensue (Creswell & Poth, 2016). Data for this study was gathered from various sources to guarantee an in-depth assessment of electronic innovations and new civil registration laws. These sources included: government policy documents which are relevant for understanding the government's view and the official narrative related to e-Governance innovations; academic publications debating e-Governance innovations and civil registration systems to provide rich perspectives into the theoretical foundations and empirical evidence linked to the research; and reports from nongovernmental organisations and international bodies to derive an external viewpoint on e-Governance in civil registration systems.

The data collected from these sources was analysed using thematically. This approach involves categorising, scrutinising, and reporting patterns (themes) within the data. Thematic analysis is a flexible and valuable research instrument that offers a rich and detailed outline of the data (Clarke & Braun, 2017). Main themes, trends, and challenges associated with the automated innovations in the new civil registration law will be highlighted and discussed.

It is important to acknowledge the limitations of this study.

One possible limitation of this study's methodology is the dependence on secondary data sources, which may present biases in the original documents analysed. Additionally, the absence of primary data from actors like government officials, population, and civil society, may narrow the perimeter of the analysis. addressing these methodological shortcomings, the study triangulates data sources by tapping official government data, scholarly publications, and data from other independent bodies and civil society.

4. Results: Overview of electronic innovations in Cameroon's new law

A textual analysis of law No. 2024/016 of 23 December 2024 on establishing the civil status registration system, particularly its provisions relating to e-Governance of civil registration has produced the following innovations.

4.1. Technological Innovations

The 2024 law introduces several important definitions and legal bases essential for understanding the use of e-Governance in civil registration in Cameroon. According to the new law, an electronic certificate is an electronic document secured by the electronic signature of its user, which attests, upon verification, to the authenticity of its content. An approved certification authority issues a qualified electronic certificate. These definitions lay the basis for the protected and authentic electronic recording and archiving of vital events. These innovations guarantee the authenticity of civil status documents, which is essential for upholding the integrity of records. Also, they introduce a supplementary degree of security and reliability, ensuring that the population can trust the documents delivered electronically.

Furthermore, the 2024 law in section 83 introduces a Unique Personal Identification Number (UPIN), allocated at birth, streamlining identification processes within Cameroon's civil status systems. This unique identifier is projected to ease communications between citizens and government agencies, reducing double identifications and bureaucratic ineptitudes. Also, the law dictates the digital recording and archiving of certificates associated with declared vital events, significantly improving data management abilities and permitting more efficient data retrieval when necessary.

4.2. Service Delivery Enhancements

One of the key remarkable aspects of the new law is the provision for electronic declaration of vital events, such as births, marriages, and deaths (Section 9). This innovation from purely traditional paper-based procedures to computerised systems improves citizens' accessibility, expedites the registration process and lessens the administrative load on citizens and government employees. Also, the 2024 law introduces the automation of the declaration of vital events, drawing up, issuance, and archiving of civil status documents, and the production of civil status statistics (Section 80). This automation boosts the efficiency and accuracy of civil registration processes. However, the modalities for electronic declarations will be stipulated by subsequent regulations, possibly ensuring that the process is tailored to users' needs. Unlike before, where the loss of an original copy of a civil status document required a lengthy judicial process for

obtaining a new one, the new law allows for obtaining copies of civil status documents from the NCSD in case of loss (Section 62(3)).

Moreover, the law expands the modes of publication of marriage banns from hardcopy pasting to online on the Council's website and introduces electronic communication of notices of intent to marry (Section 17). This innovation reinforces communication between Civil Status Registrars, improves transparency, and accelerates the marriage registration procedure. Significantly, civil status documents in Cameroon can now be delivered in printed and digital forms, with both formats holding the same legal validity (Section 29). This dual approach guarantees that citizens have flexibility in obtaining their documents, thereby improving overall service delivery. Furthermore, in section 126, the law demonstrates that the full implementation of e-Governance in civil service registration in Cameroon will flow from a pilot phase. Specifically, this provision of the law accords legal validity to "civil status certificates drawn up as part of pilot operations undertaken to computerise the civil status…" Moreover, section 70(2) of the 2024 law highlights the switch to digital administration, stipulating that any civil status centre which becomes fully digitalised shall be exempted from keeping registers in hardcopies. This demonstrates the government's commitment to progressively transforming civil registration in Cameroon through e-Governance.

4.3. Governance and Accountability

Creating a National Civil Status Database is a fundamental feature of the new law, providing a centralised repository for all civil status documents in Cameroon (Section 30). The database will host "all information, data, documents, copies, or forms, in both paper and digital form, related to the declaration of vital events and the drawing up of civil status documents" (Section 77). The enumerated data is kept and managed either in hardcopy or electronically by the civil status management body (Section 78(3)). This database ameliorates data integrity and improves accessibility for both citizens and civil status officials. Consolidating civil status records in a single location is expected to expand the efficiency of civil status management while keeping the data up-to-date and accurate.

The law also introduces electronic objections to marriage solemnisation, permitting individuals to lodge their protests through digital canals (Section 42). The same computerised channel applies to requests for dispensation for publication of banns addressed to the competent State Counsel (Section 18). This modernisation rationalises the procedure, making it more efficient and time-saving for citizens and judicial authorities.

4.4. Data Protection and regulatory frameworks

The law emphasises data protection and security in section 82, stating that "personal civil status data contained in the national civil status database shall be protected according to the laws on personal data protection." This provision resonates with law No. 2024/017 of 23 December 2024 on personal data protection in Cameroon. This innovation echoes the government's pledge to preserve citizens' privacy and security in an increasingly digital world. Furthermore, Section 82 of the law on the civil registration system in Cameroon grants Civil Status Registrars and their Secretaries direct access to civil status data associated with their respective centres. The same access is granted to public administrations connected to the national civil status systems for specific or general consultation. To ensure efficiency and security, the law guarantees that the technical features of and procedure for electronic communication of civil status data must adhere to the law regulating electronic communications in Cameroon (Section 83). Electronic communications in Cameroon are governed by Law No. 2010/013 of 21 December 2010 as amended and supplemented by Law No. 2015/016 of 20 April 2015.

The 2024 law stipulates that most aspects of the digital innovations introduced shall be governed by separate legal instruments and regulations. These include the modalities for issuing civil status documents in electronic form (Section 29(5)); methods of signatories and signature for electronically drawn marriage certificates (Section 41(2)); the conditions for access to data in the National Civil Status Database, and the issuance and certification of civil status documents from the NCSD (Section 79); the conditions for automated processing of civil status documents (Section 80); the features of and conditions for use and assigning of UPIN (Section 81 (3)); and the list administrations and procedure for access to data in the NCSD (Section 82(4)). The anticipated regulatory frameworks will possibly guarantee that all electronic communications conform to legal standards, mitigating the risk of fraud and ensuring the security of personal data. Furthermore, by opting for separate robust policy frameworks that regulate the functionalities of the new e-Governance model and the protection of individual rights, the law establishes a safe and structured milieu for civil registration. This emphasis on governance tools is crucial for establishing public trust in the system, as citizens will have a clearer picture of how their data is stored and accessed.

4.5. Technological infrastructure in Cameroon

Cameroon has recorded important strides in the ICT sector, particularly through adopting the National ICT Strategic Plan 2020 which acknowledges the digital economy as a key development catalyst[23]. The country's national optical fibre backbone of about 12000 km has

[23] Toussi, S. (2019, September 12). Overview of Cameroon's Digital Landscape. CIPESA. https://clck.ru/3Gf9m2

linked 209 of the 360 subdivisions. Cameroon's Ministry of Post and Telecommunications revealed that by 2018, 83 % of Cameroonians had subscribed to mobile phone services, with an internet penetration rate of 35 %[24]. According to ICT Development Index 2024, Cameroon currently ranks 31st out of 47 countries, with a score of 44.2[25].

The national ICT strategic plan 2020 covers 8 key areas, aligning with the Country's emergence vision by 2035. Some public institutions have been created to boost Cameroon's development of digital public service[26]. Based on this institutional framework, some e-Governance initiatives have been undertaken to digitise and manage state employees, public finances, electoral register management, customs operations, and transport documents[27]. Other e-Governance innovations are operational through introducing e-taxes, e-visa, and e-commerce and trade[28]. As part of the Strategic Plan for the Rehabilitation of Civil Registration in Cameroon (2018–2022), a digitalisation plan for civil registration was designed in 2018 as a basis for the 2024 law on civil registration systems. The aim as per the master plan is to establish a platform connecting the civil status system and other segments, such as the national identity institution, Ministries of Justice, Transport and Health for effective service delivery.

5. Discussions

The paper aimed to investigate the electronic innovations introduced in the civil registration system of Cameroon by the new law of 23 December 2024, with emphasis on their implications for service delivery, governance and citizen participation. Employing a qualitative research design, the study examined existing literature, policy and legal documents to assess the effectiveness of the new e-Governance framework in the civil status registration sector. Thematic analysis was adopted to categorise major themes and challenges of the new law. The research revealed several relevant innovations introduced by the 2024 law, including concepts of digital certificates and qualified electronic certificates, the introduction of UPIN, electronic declaration of vital events, the adoption of electronic signatures and the creation of an NCSD. The automation

24 Ministry of Posts and Telecommunication (2018). Posts, Telecommunications, and ICT: Preclous assets of the seven-year mandate. https://clck.ru/3Gf9pe

25 Ecofin agency. (2024, July 16). ICT Development Index 2024: Ranking of African Countries. https://clck.ru/3Gf9tk

26 Telecommunications Regulatory Board (ART); the National Agency for ICT (ANTIC); and the National Centre for the Development of Computer Services (CENADI).

27 Alypova, S. (2024). E-government Development in Cameroon. Centre for African Studies, HSE University. https://clck.ru/3Gf9zu

28 Ibid.

of civil status operations and the publication of marriage banns online constitute other noteworthy novelties. These innovations align Cameroon's civil status registration system with the critical flow model of e-Governance. Also, the innovations fit into the G2G and G2C domains of e-Governance interactions.

The key findings demonstrate significant innovations contained in the new law but identify challenges with implementation linked to technology literacy rate, technological infrastructure and internet coverage in Cameroon which impact service delivery, and governance accountability. With regards to e-government, the United Nations e-government ranking 2024 places Cameroon at 155th out of 193 countries as opposed to 139th position in 2004, and 105th in e-participation with improvement from 84th position in 2004[29]. With an internet coverage of less than 50 %, the country's Network Readiness Index in 2024 positions it at 113th out of 133 countries with recorded improvement in governance[30]. While the demand for internet services constantly increases, the precious service is almost inaccessible in some rural areas.

One of the remarkable results is the creation of a National Civil Status Database to centralise and store civil status documents. This innovation is essential as the NCSD has the potential to improve data integrity record-storage culture and modernise access to valuable information. This finding is relevant in its potential to renovate civil status registration operations from a disjointed system into an interconnected, efficient system that addresses citizens' needs effectively. Automating civil status transactions further increases efficiency by mitigating bureaucratic delays and errors. Furthermore, introducing electronic declaration of vital events – births, marriages and deaths – reflects a remarkable paradigm change in the interaction between citizens and public services. This innovation shortens processing time, and increases accessibility, especially for marginalised segments of the population who may encounter challenges in the traditional declaration processes[31]. These transformations resonate with the TAM, which suggests that citizens accept technology based on its perceived ease of use and usefulness (Reiling & Contini, 2022).

The introduction of electronic certificates demonstrates a milestone towards safeguarding the authenticity and security of civil status documents. By exploiting digital signatures and encoding algorithms, the 2024 law aims to improve the integrity and trustworthiness of civil status documents in Cameroon. This innovation identifies

29 UN E-Government Knowledgebase. https://clck.ru/3GfA9A

30 Portulans Institute. (2024). Network Readiness Index 2024: Cameroon. https://clck.ru/3GfABt

31 Zewoldi, Y. (2019). Snapshot of Civil Registration and Vital Statistics in Cameroon. Centre of Excellence for Civil Registration and Vital Statistics. https://clck.ru/3GfADM

with the e-Governance theory's principle which highlights the transformative capabilities of ICTs in enhancing public service delivery (Heeks, 2006). Moreover, UPIN eases civil status records' archival, management and retrieval. The UPIN serves as a unique identifier for citizens, reorganising the registration operation and minimising duplication risks.

These results align with previous studies on the benefits of e-Governance in civil registration and other domains. For instance, studies from countries like Estonia and Kenya have demonstrated that centralised computerised systems can dramatically augment service delivery, and citizen fulfilment and reduce bureaucratic burdens (Tammpuu & Masso, 2018). Furthermore, Ghana's m-birth model justifies how digital birth registration can remarkably increase the declaration rate of vital events[32]. Likewise, by introducing e-Governance in civil registration processes, the Cameroonian legislator aims to improve the efficiency and accessibility of civil registration services to citizens and other institutions through centralised records, digital certificates and UPIN. The Cameroonian innovation in civil registration can pull vital lessons from these countries' successes, to guarantee the effective materialisation of its innovation through tailored regulatory frameworks. However, while scholars highlight the merits of e-Governance, they also emphasise the relevance of addressing infrastructure barriers, and possible public confrontation to the innovations[33].

The research findings resonate with the e-governance theory. This theory suggests that ICT incorporation into governance operations can improve transparency, accountability, and citizen participation (Heeks, 2006). The results echo this theory as the introduction of UPIN, electronic declarations of vital events, objections to marriage celebrations, electronic signatures for marriage solemnisation, and the electronic publication of marriage banns, all align with the principles of accessibility and transparency. Furthermore, the findings relate to TAM. TAM contends that apparent ease of use and perceived relevance significantly determine citizen's acceptance of technology (Davis et al., 1989). The new law seeks to ameliorate the ease of use of civil registration operations by automating transactions and implementing electronic declarations, signatures and certificates. The effective application of digital innovations in the civil registration system in Cameroon greatly relies both on the functionality of the system and how easily civil status personnel and citizens can access and use the new technologies. As such, if the population view the automated system as relevant and easy to use, they will likely welcome the innovations.

---

32 UNICEF. (2018). Assessment of the m-birth project in Ghana. https://clck.ru/3GfAKK

33 Zungu, S. (2024). The use of monitoring and evaluation as an improving e-Governance for enhanced service delivery. Master thesis. University of Johannesburg.

Despite the remarkable findings on the potential of the 2024 law to reinforce the civil registration system's efficiency and accessibility, alternative analysis should be acknowledged. For instance, the success of the new computerised civil registration system may depend on the government's ability and willingness to ameliorate technological facilities and offer routine training to civil status registrars, secretaries, and other stakeholders involved in the chain. If these foundational factors are not handled appropriately, the projected results may not be as expected. Moreover, the breakthrough of the new law may be shaped by the regulatory mechanisms, the political will and the capacity of the government establishments to enforce the new law. Also, the prevailing digital gap in Cameroon poses a profound menace, especially between urban and rural areas. If the existing inequality in access technology is not addressed, digital innovations could inadvertently aggravate existing social disparities.

The research recognises some limitations. First, the dependence on secondary data may introduce prejudices, as it does not gather direct perspectives from the main actors in enforcing the innovative digital system in the new law. Moreover, the research focuses on a recently adopted legislation whose effective implementation is still pending the adoption of some regulatory frameworks. There is a lack of empirical data on user fulfilment and the effectiveness of the digital system as no research has been conducted on it so far. This restricts the possibility of concluding the practical successes of the digital innovations introduced.

Further research could engage an assessment of the long-term repercussions of digital innovations on the civil registration system in Cameroon. This could involve an empirical study to collect qualitative primary data from the main actors and users of the civil registration system, offering a profound understanding of citizen experience and domains for amelioration. Also, investigations into the intersection of ICTs and social inclusion could be essential in revealing how to guarantee that all citizens have access to, and gain from these innovations without discrimination.

6. Recommendations

To boost the enforcement and increase the effectiveness of digital innovations in the new civil registration law in Cameroon, several measures should be considered: The research findings on the rate of technological infrastructure in Cameroon present a need for improvement to ensure the success of e-Governance innovations. The government should capitalise on advancing and increasing internet connectivity, especially in rural and underserved localities. This also includes allocating ICT kits and other resources to all civil status centres to ease efficient data collection, management and transmission to the centralised repository. This will guarantee citizens' access to computerised civil

registration services, bridging the digital gap, as emphasised by the National Institute of Statistics[34].

Furthermore, the effective application of digital aspects of the new law requires that state employees and civil status staff are well-trained in using digital tools. Inclusive training programmes should be designed to improve their technical skills and acquaint them with the new processes. These skills improvement initiatives should be continuous to keep the staff informed on technological advancement and best practices in the domain. This will minimise resistance to transformation and ensure easy reception of the new systems[35]. Also, regular monitoring and evaluation of the enforcement operation is essential for identifying difficulties and adopting appropriate adjustments. The government should institute a monitoring mechanism to track the new law's evolution and measure its influence on civil registration services. Systematic stakeholder feedback should be integrated to build up the system.

Equally, the new law contains some provisions requiring further regulations and guidelines for enforcement. The government should sequence the rapid design of these regulations to provide clear modalities on the procedures for electronic declarations, issuance of electronic certificates, the generation and issuance of UPIN, and the functioning of the NCSD, among others. This will guarantee reliability and compliance in all civil status centres (Heeks, 2006). The security of personal civil status data is essential for ensuring public trust in the innovative system. The government should integrate robust data protection strategies, including encryption, access restrictions, and routine audits with the NCSD, to secure sensitive data. Also, alignment with the personal data protection law should be strictly guaranteed. Through these mechanisms, the government will establish transparency and accountability in the enforcement process and readiness to address any issues connected to data security and privacy. Achieving these will build public trust and confidence in the system, guaranteeing its success and mitigating resistance.

Moreover, engaging stakeholders and educating the population are basic strategies for successfully implementing e-Governance innovations in the new civil status registration law. The key actors include state agencies, employees, civil society, and community leaders. They should be fully involved in the design of regulations, training policies, and monitoring mechanisms. Their involvement and feedback can offer valued perspectives and assist in addressing possible setbacks. Again, education

[34] National Institute of Statistics. (2023). Rapport sur les statistiques de l'état civil au Cameroun (2018–2022). https://clck.ru/3Gf8Uq

[35] Zewoldi, Y. (2019). Snapshot of Civil Registration and Vital Statistics in Cameroon. Centre of Excellence for Civil Registration and Vital Statistics. https://clck.ru/3GfADM

and sensitisation operations are essential to enlighten citizens about e-Governance innovations in civil status registration operations and their benefits. These campaigns should employ various communication media, including radio, television, social media, posters, community meetings, churches and other social gatherings to attain a wide audience. Clear and accessible information should inform the population about the significance of civil status registration, the associated rights and how to use the computerised system. In tandem with awareness campaigns, the government should adopt programmes that intensify digital literacy, mainly in rural and underserved zones to empower citizens to access digitised systems effectively.

Conclusion

This study has examined the digital innovations enshrined in the 2024 law governing the civil status registration system in Cameroon. The research demonstrates valuable insights into how these innovations can alter civil status registration processes. The introduction of electronic declaration of vital events, digital certificates, a Unique Personal Identification Number (UPIN), and the creation of the NCSB are crucial novelties that improve the efficiency, user-friendliness, and trustworthiness of civil status registration processes. These findings are chiefly remarkable as they highlight the potential of ICTs to transform public administration and strengthen service delivery. One of the main unexpected results of this research is that some key aspects of computerised innovations will only be implemented through modalities and procedures determined by regulations. These regulations have not been adopted and there is a likelihood of delay, which may undermine the full implementation of the digital innovations.

The paper's results are significant for civil status registration and public sector governance in Cameroon. By incorporating digital technologies, the new law modernises civil status registration operations, decreases bureaucratic foot-dragging, and increases the accuracy and trustworthiness of records. This can potentially ameliorate public service delivery efficiency and effectiveness, profiting citizens and government institutions. The study contributes to scholarly discourse by presenting an insightful assessment of digital innovations embedded in the new civil status registration law in Cameroon. While scholars have investigated the application of e-Governance in civil registration in other countries like Kenya, Ghana, South, Africa and Estonia, there exists no research so far on the specific case of Cameroon. This paper fills this gap by assessing the unique challenges and opportunities associated with e-Governance in the new law and its potential influence on civil status registration operations. The outcomes highlight the need for further inquiry into the long-term bearings of e-Governance innovations and the factors influencing their success.

The study's results have crucial implications for the evolution of civil registration and public sector governance in Cameroon. The innovations have the potential to modernise and interconnect civil status registration processes in Cameroon, making them more efficient, accessible, and reliable. For policymakers and practitioners, the paper's recommendation on stakeholder involvement and public awareness creation demonstrates the need for a collaborative approach to enforcing the digital system. This will guarantee that citizens are educated and guided on accessing digital services. The study also highlights the importance of data security and privacy in implementing e-Governance. By ensuring personal civil status data protection, the new law aims to build public trust and confidence in the new system. This is crucial for successfully adopting digital civil registration services and the overall effectiveness of e-Governance initiatives.

References

Albert, I. O. (2009). Whose E-Governance?: A Critique of Online Citizen Engagement In Africa. *International Journal of eBusiness and eGovernment Studies*, *1*(1), 27–40.

Al-Gahtani, S. S. (2016). Empirical investigation of e-learning acceptance and assimilation: A structural equation model. *Applied Computing and Informatics*, *12*(1), 27–50.

Araújo, T., & Casais, B. (2020). Customer acceptance of shopping-assistant chatbots. In *Marketing and Smart Technologies: Proceedings of ICMarkTech 2019* (pp. 278–287). Singapore: Springer.

Bannister, F., & Connolly, R. (2012). Defining e-Governance. *e-Service Journal*, *8*(2), 3–25. https://doi.org/10.2979/eservicej.8.2.3

Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

Coe, A., Paquet, G., & Roy, J. (2001). E-governance and smart communities: a social learning challenge. *Social Science Computer Review*, *19*(1), 80–93.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). Technology acceptance model. *Journal of Management Science*, *35*(8), 982–1003.

Djossa-Tchokoté, I., Teutio, A. O. N., & Nyongo, A. S. A. (2024). E-Governance, Digitized Tax Procedures and SME's Business Process Performance in Cameroon. *iBusiness*, *16*(3), 65–84. https://doi.org/10.4236/ib.2024.163006

Fang, Z. (2002). E-government in digital era: concept, practice, and development. *International Journal of the Computer, the Internet and Management*, *10*(2), 1–22.

Foster, K. (2020). Smarten Up: Paths to bottom-up smart cities and the risks of top-down smart governance. *Smart Cities Paper Series: Smart Governance in South African Cities*.

Grigalashvili, V. (2022). E-government and E-governance: Various or Multifarious Concepts. *International Journal of Scientific and Management Research*. *5*(01).183–196. https://doi.org/10.37502/ijsmr.2022.5111

Halachmi, A. (2004). E-government theory and practice: The evidence from Tennessee (USA). In M. Holzer, M. Zhang, & K. Dong (Eds.), *Frontiers of Public Administration: Proceedings of the Second Sino-U.S. International Conference: Public Administration in the Changing World* (pp. 24–43). United Nations Public Administration Network.

Heeks, R. (2003). E-Government in Africa: Promise and practice. *Information Polity*, *7*(23), 97–114. https://doi.org/10.3233/ip-2002-0008

Heeks, R. (2006). *Implementation and managing e-Governance*. London: Sage publications Ltd.

Kaisara, G., & Pather, S. (2011). The e-Government evaluation challenge: A South African Batho Pele-aligned service quality approach. *Government Information Quarterly*, *28*(2), 211–221. https://doi.org/10.1016/j.giq.2010.07.008

Kattel, R., & Mergel, I. (2019). Estonia's digital transformation: Mission mystique and the hiding hand. In M. E. Compton, & P. Hart (Eds.), *Great Policy Successes* (pp. 143–160). Oxford University Press. https://doi.org/10.1093/oso/9780198843719.003.0008

Lubis, S., Purnomo, E. P., Lado, J. A., & Hung, C. F. (2024). Electronic governance in advancing sustainable development goals through systematic literature review. *Discover Global Society*, *2*(1) 77. https://doi.org/10.1007/s44282-024-00102-3

Makwanya, M. (2022). Digital practice for social work in Zimbabwe: Success, challenges and opportunities. In A. L. Peláez, S. Suh, & S. Zelenev (Eds.), *Digital Transformation and Social Well-Being* (pp. 160–168). Routledge.

Munyoka, W. (2020). Electronic government adoption in voluntary environments–a case study of Zimbabwe. *Information Development*, *36*(3), 414–437. https://doi.org/10.1177/0266666919864713

Muttoo, S. K., Gupta, R., Pal, S. K., & Muttoo, S. K. (2019). *E-Governance in India* (pp. 13–25). Singapore: Springer. https://doi.org/10.1007/978-981-13-8852-1

Naqvi, S. A. M., Alyas, T., Tabassum, N., Namoun, A., & Naqvi, H. H. (2021). Post Pandemic World and Challenges for E-Governance Framework. *International Journal of Advanced Trends in Computer Science and Engineering*, *10*(3), 2630–2636. https://doi.org/10.30534/ijatcse/2021/1571032021

Oliveira, T. A., Oliver, M., & Ramalhinho, H. (2020). Challenges for Connecting Citizens and Smart Cities: ICT, e-Governance and Blockchain. *Sustainability*, *12*(7), 2926. https://doi.org/10.3390/su12072926

Palvia, S. C. J., & Sharma, S. S. (2007). E-government and e-governance: definitions/domain framework and status around the world. In *International Conference on E-governance* (Vol. 5, No. 1, pp. 1–12).

Prashar, K., & Bawa, S. S. (2023). Studying the Effect of Artificial Intelligence on E-Governance. In P. Tyagi, S. Grima, K. Sood, B. Balamurugan, E. Özen, & T. Eleftherios (Eds.), *Smart Analytics, Artificial Intelligence and Sustainable Performance Management in a Global Digitalised Economy (Contemporary Studies in Economic and Financial Analysis, Vol. 110A)* (pp. 87–101). Emerald Publishing Limited, Leeds. https://doi.org/10.1108/S1569-37592023000110A005

Reiling, D., & Contini, F. (2022). E-Justice Platforms: Challenges for Judicial Governance. *International Journal for Court Administration*, *13*(1), 6. https://doi.org/10.36745/ijca.445

Schuppan, T. (2009). E-Government in developing countries: Experiences from sub-Saharan Africa. *Government Information Quarterly*, *26*(1), 118–127. https://doi.org/10.1016/j.giq.2008.01.006

Sevidzem, M. C. (2024). Use of ICT and the Application of E-Governance Strategies in Service Delivery by Local Councils in Cameroon: The Case of Local Councils in the Bamenda Municipality. *PanAfrican Journal of Governance and Development (PJGD)*, *5*(1), 3–27. https://doi.org/10.46404/panjogov.v5i1.5354

Sharma, S., Kumar Kar, A., & Gupta, M. P. (2021). Unpacking Digital Accountability: Ensuring efficient and answerable e-governance service delivery. In *Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance* (pp. 260–269). https://doi.org/10.1145/3494193.3494229

Suri, P. K., & Sushil. (Eds.) (2017). *Strategic Planning and Implementation of E-Governance*, Springer.

Suthar, A. B., Khalifa, A., Yin, S., Wenz, K., Ma Fat, D., Mills, S. L., ... & Mrkic, S. (2019). Evaluation of approaches to strengthen civil registration and vital statistics systems: a systematic review and synthesis of policies in 25 countries. *PLoS Medicine*, *16*(9), e1002929. https://doi.org/10.1371/journal.pmed.1002929

Tammpuu, P., & Masso, A. (2018). 'Welcome to the virtual state': Estonian e-residency and the digitalised state as a commodity. *European Journal of Cultural Studies*, *21*(5), 543–560. https://doi.org/10.1177/1367549417751148

Umbach, G., & Tkalec, I. (2022). Evaluating e-governance through e-government: Practices and challenges of assessing the digitalisation of public governmental services. *Evaluation and program planning*, *93*, 102–118. https://doi.org/10.1016/j.evalprogplan.2022.102118

Xin, G., Esembe, E. E., & Chen, J. (2023). The mixed effects of e-participation on the dynamic of trust in government: Evidence from Cameroon. *Australian Journal of Public Administration*, *82*(1), 69–95. https://doi.org/10.1111/1467-8500.12569

Authors information

Robert Kosho Ndiyun – PhD (Political Studies), Post-Doctoral Research Fellow, Department of Public Affairs, Tshwane University of Technology
Address: Staatsartillerie Rd, Philip Nel Park, 0183 Pretoria, South Africa
E-mail: NdiyunRK@tut.ac.za
ORCID ID: https://orcid.org/0000-0002-6435-0494
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57762430700
WoS Researcher ID: https://www.webofscience.com/wos/author/record/JCD-5501-2023
Google Scholar ID: https://scholar.google.com/citations?user=siNlKWQAAAAJ

Ricky Munyaradzi Mukonza – DTech (Public Management), Associate Professor of Public Management, Department of Public Affairs, Tshwane University of Technology
Address: Staatsartillerie Rd, Philip Nel Park, 0183 Pretoria, South Africa
E-mail: MukonzaRM@tut.ac.za
ORCID ID: https://orcid.org/0000-0001-8121-1501
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56157610200
Google Scholar ID: https://scholar.google.com/citations?user=LXj9QUQAAAAJ

Authors' contributions

The authors have contributed equally into the concept and methodology elaboration, validation, formal analysis, research, selection of sources, text writing and editing, project guidance and management.

Conflict of interest

The authors declares no conflict of interest.

Financial disclosure

The research had no sponsorship.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – January 18, 2025
Date of approval – January 31, 2025
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:347.6:004.9:004.056
EDN: https://elibrary.ru/eaqyqj
DOI: https://doi.org/10.21202/jdtl.2025.1

Цифровая трансформация системы регистрации актов гражданского состояния Камеруна: инновации в электронном управлении

Роберт Кошо Ндиюн ✉

Технологический университет Тшване, Претория, ЮАР

Рикки Муньярадзи Муконза

Технологический университет Тшване, Претория, ЮАР

Ключевые слова

акт гражданского состояния,
государственная услуга,
законодательство,
защита данных,
Камерун,
право,
цифровая грамотность,
цифровые технологии,
электронное правительство,
электронное управление

Аннотация

Цель: исследование инновационных преобразований в сфере электронного управления, внедренных в систему регистрации актов гражданского состояния Камеруна в результате законодательных реформ 2024 года. Основное внимание уделяется оценке влияния этих преобразований на повышение эффективности управления, прозрачности, доступности услуг для граждан, а также на совершенствование статистического учета жизненно важных событий.

Методы: в работе использованы общенаучные методы анализа и синтеза, классификации, системный и функциональный подходы, а также формально-юридический и сравнительно-правовой методы.

Результаты: внедрение электронного декларирования актов гражданского состояния, создание Национальной базы данных и переход на электронные свидетельства способны существенно повысить эффективность и доступность услуг для населения. Однако авторы подчеркивают, что успешная реализация цифровых инноваций требует преодоления значительных барьеров, таких как недостаточная технологическая оснащенность, ограниченный доступ к Интернету и низкий уровень цифровой грамотности среди граждан. Эти вызовы делают необходимой разработку дополнительных механизмов регулирования и поддержки. Особое значение придается балансу между цифровизацией и обеспечением прав граждан в контексте электронной регистрации.

✉ Контактное лицо

Научная новизна: заключается в предоставлении уникальных эмпирических данных о процессе цифровизации государственных услуг в Камеруне, что особенно актуально для стран глобального Юга, где подобные преобразования происходят медленно и фрагментарно. Исследование вносит значительный вклад в научную дискуссию, расширяя понимание моделей внедрения цифровых технологий через призму ожидаемой полезности и воспринимаемой простоты использования в условиях развивающихся стран.

Практическая значимость: состоит в разработке конкретных рекомендаций для законодателей, государственных служащих и других заинтересованных сторон. Авторы подчеркивают необходимость скорейшего принятия нормативной правовой базы, внедрения образовательных программ для сотрудников и граждан, а также обеспечения доступа к цифровым технологиям. Эти меры направлены на создание устойчивой инфраструктуры для эффективного перехода к электронным системам и повышение качества государственных услуг. Работа представляет собой важный вклад в изучение процессов цифровизации государственного управления, предлагая как теоретические выкладки, так и практические решения, которые могут быть адаптированы для других стран с аналогичными вызовами.

Для цитирования

Ндиюн, Р. К., Муконза, Р. М. (2025). Цифровая трансформация системы регистрации актов гражданского состояния Камеруна: инновации в электронном управлении. *Journal of Digital Technologies and Law*, *3*(1), 7–34. https://doi.org/10.21202/jdtl.2025.1

Список литературы

Albert, I. O. (2009). Whose E-Governance?: A Critique of Online Citizen Engagement In Africa. *International Journal of eBusiness and eGovernment Studies*, *1*(1), 27–40.

Al-Gahtani, S. S. (2016). Empirical investigation of e-learning acceptance and assimilation: A structural equation model. *Applied Computing and Informatics*, *12*(1), 27–50.

Araújo, T., & Casais, B. (2020). Customer acceptance of shopping-assistant chatbots. In *Marketing and Smart Technologies: Proceedings of ICMarkTech 2019* (pp. 278–287). Singapore: Springer.

Bannister, F., & Connolly, R. (2012). Defining e-Governance. *e-Service Journal*, *8*(2), 3–25. https://doi.org/10.2979/eservicej.8.2.3

Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

Coe, A., Paquet, G., & Roy, J. (2001). E-governance and smart communities: a social learning challenge. *Social Science Computer Review*, *19*(1), 80–93.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). Technology acceptance model. *Journal of Management Science*, *35*(8), 982–1003.

Djossa-Tchokoté, I., Teutio, A. O. N., & Nyongo, A. S. A. (2024). E-Governance, Digitized Tax Procedures and SME's Business Process Performance in Cameroon. *iBusiness*, *16*(3), 65–84. https://doi.org/10.4236/ib.2024.163006

Fang, Z. (2002). E-government in digital era: concept, practice, and development. *International Journal of the Computer, the Internet and Management*, *10*(2), 1–22.

Foster, K. (2020). Smarten Up: Paths to bottom-up smart cities and the risks of top-down smart governance. *Smart Cities Paper Series: Smart Governance in South African Cities*.

Grigalashvili, V. (2022). E-government and E-governance: Various or Multifarious Concepts. *International Journal of Scientific and Management Research*. *5*(01).183–196. https://doi.org/10.37502/ijsmr.2022.5111

Halachmi, A. (2004). E-government theory and practice: The evidence from Tennessee (USA). In M. Holzer, M. Zhang, & K. Dong (Eds.), *Frontiers of Public Administration: Proceedings of the Second Sino-U.S. International Conference: Public Administration in the Changing World* (pp. 24–43). United Nations Public Administration Network.

Heeks, R. (2003). E-Government in Africa: Promise and practice. *Information Polity*, *7*(23), 97–114. https://doi.org/10.3233/ip-2002-0008

Heeks, R. (2006). *Implementation and managing e-Governance*. London: Sage publications Ltd.

Kaisara, G., & Pather, S. (2011). The e-Government evaluation challenge: A South African Batho Pele-aligned service quality approach. *Government Information Quarterly*, *28*(2), 211–221. https://doi.org/10.1016/j.giq.2010.07.008

Kattel, R., & Mergel, I. (2019). Estonia's digital transformation: Mission mystique and the hiding hand. In M. E. Compton, & P. Hart (Eds.), *Great Policy Successes* (pp. 143–160). Oxford University Press. https://doi.org/10.1093/oso/9780198843719.003.0008

Lubis, S., Purnomo, E. P., Lado, J. A., & Hung, C. F. (2024). Electronic governance in advancing sustainable development goals through systematic literature review. *Discover Global Society*, *2*(1) 77. https://doi.org/10.1007/s44282-024-00102-3

Makwanya, M. (2022). Digital practice for social work in Zimbabwe: Success, challenges and opportunities. In A. L. Peláez, S. Suh, & S. Zelenev (Eds.), *Digital Transformation and Social Well-Being* (pp. 160–168). Routledge.

Munyoka, W. (2020). Electronic government adoption in voluntary environments–a case study of Zimbabwe. *Information Development*, *36*(3), 414–437. https://doi.org/10.1177/0266666919864713

Muttoo, S. K., Gupta, R., Pal, S. K., & Muttoo, S. K. (2019). *E-Governance in India* (pp. 13–25). Singapore: Springer. https://doi.org/10.1007/978-981-13-8852-1

Naqvi, S. A. M., Alyas, T., Tabassum, N., Namoun, A., & Naqvi, H. H. (2021). Post Pandemic World and Challenges for E-Governance Framework. *International Journal of Advanced Trends in Computer Science and Engineering*, *10*(3), 2630–2636. https://doi.org/10.30534/ijatcse/2021/1571032021

Oliveira, T. A., Oliver, M., & Ramalhinho, H. (2020). Challenges for Connecting Citizens and Smart Cities: ICT, e-Governance and Blockchain. *Sustainability*, *12*(7), 2926. https://doi.org/10.3390/su12072926

Palvia, S. C. J., & Sharma, S. S. (2007). E-government and e-governance: definitions/domain framework and status around the world. In *International Conference on E-governance* (Vol. 5, No. 1, pp. 1–12).

Prashar, K., & Bawa, S. S. (2023). Studying the Effect of Artificial Intelligence on E-Governance. In P. Tyagi, S. Grima, K. Sood, B. Balamurugan, E. Özen, & T. Eleftherios (Eds.), *Smart Analytics, Artificial Intelligence and Sustainable Performance Management in a Global Digitalised Economy (Contemporary Studies in Economic and Financial Analysis, Vol. 110A)* (pp. 87–101). Emerald Publishing Limited, Leeds. https://doi.org/10.1108/S1569-37592023000110A005

Reiling, D., & Contini, F. (2022). E-Justice Platforms: Challenges for Judicial Governance. *International Journal for Court Administration*, *13*(1), 6. https://doi.org/10.36745/ijca.445

Schuppan, T. (2009). E-Government in developing countries: Experiences from sub-Saharan Africa. *Government Information Quarterly*, *26*(1), 118–127. https://doi.org/10.1016/j.giq.2008.01.006

Sevidzem, M. C. (2024). Use of ICT and the Application of E-Governance Strategies in Service Delivery by Local Councils in Cameroon: The Case of Local Councils in the Bamenda Municipality. *PanAfrican Journal of Governance and Development (PJGD)*, *5*(1), 3–27. https://doi.org/10.46404/panjogov.v5i1.5354

Sharma, S., Kumar Kar, A., & Gupta, M. P. (2021). Unpacking Digital Accountability: Ensuring efficient and answerable e-governance service delivery. In *Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance* (pp. 260–269). https://doi.org/10.1145/3494193.3494229

Suri, P. K., & Sushil. (Eds.) (2017). *Strategic Planning and Implementation of E-Governance*, Springer.

Suthar, A. B., Khalifa, A., Yin, S., Wenz, K., Ma Fat, D., Mills, S. L., ... & Mrkic, S. (2019). Evaluation of approaches to strengthen civil registration and vital statistics systems: a systematic review and synthesis of policies in 25 countries. *PLoS Medicine*, *16*(9), e1002929. https://doi.org/10.1371/journal.pmed.1002929

Tammpuu, P., & Masso, A. (2018). 'Welcome to the virtual state': Estonian e-residency and the digitalised state as a commodity. *European Journal of Cultural Studies*, *21*(5), 543–560. https://doi.org/10.1177/1367549417751148

Umbach, G., & Tkalec, I. (2022). Evaluating e-governance through e-government: Practices and challenges of assessing the digitalisation of public governmental services. *Evaluation and program planning*, *93*, 102–118. https://doi.org/10.1016/j.evalprogplan.2022.102118

Xin, G., Esembe, E. E., & Chen, J. (2023). The mixed effects of e-participation on the dynamic of trust in government: Evidence from Cameroon. *Australian Journal of Public Administration*, *82*(1), 69–95. https://doi.org/10.1111/1467-8500.12569

Сведения об авторах

Ндиюн Роберт Кошо – PhD в области политологии, пост-док, кафедра по связям с общественностью, Технологический университет Тшване
Адрес: Южно-Африканская Республика, 0183, г. Претория, ул. Штаатсартиллери, парк Филипа Нела
ORCID ID: https://orcid.org/0000-0002-6435-0494
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57762430700
WoS Researcher ID: https://www.webofscience.com/wos/author/record/JCD-5501-2023
Google Scholar ID: https://scholar.google.com/citations?user=siNlKWQAAAAJ

Муконза Рикки Муньярадзи – доктор технологии в области государственного управления, доцент в области государственного управления, кафедра по связям с общественностью, Технологический университет Тшване
Адрес: Южно-Африканская Республика, 0183, г. Претория, ул. Штаатсартиллери, парк Филипа Нела
E-mail: MukonzaRM@tut.ac.za
ORCID ID: https://orcid.org/0000-0001-8121-1501
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56157610200
Google Scholar ID: https://scholar.google.com/citations?user=LXj9QUQAAAAJ

Вклад авторов

Авторы внесли равный вклад в разработку концепции, методологии, валидацию, формальный анализ, проведение исследования, подбор источников, написание и редактирование текста, руководство и управление проектом.

Конфликт интересов

Авторы сообщают об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.07.45 / Право и научно-технический прогресс
Специальность ВАК: 5.1.3 / Частно-правовые (цивилистические) науки

История статьи

Дата поступления – 18 января 2025 г.
Дата одобрения после рецензирования – 31 января 2025 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:347.78:340.13:347.9:004.8
EDN: https://elibrary.ru/ppfjub
DOI: https://doi.org/10.21202/jdtl.2025.2

Evolution of Copyright in the Era of Artificial Intelligence: Analysis of Conflicts of Law and Judicial Precedents

Themistoklis Tzimas

Democritus University of Thrace, Komotini, Greece

Keywords

artificial intelligence,
copyright law,
court,
digital technologies,
intellectual property,
interdisciplinary approach,
law,
legal regulation,
legislation,
technological advancement

Abstract

Objective: a comprehensive critical analysis of the modern legal regulation of artificial intelligence technologies arising at the junction of intellectual property and artificial intelligence norms. Special attention is paid to the study of conflicts between existing European copyright legislation and new technological realities.

Methods: the work uses an interdisciplinary approach, including historical, formal-legal and comparative-legal research methods. The historical method allowed tracing the evolution of legislative and doctrinal approaches to intellectual property regulation in the era of digitalization. The formal-legal method made it possible to conduct a detailed analysis of the legal norms of various states. The comparative-legal method provided an opportunity to compare different approaches to regulating relations in the use of artificial intelligence for creative activities.

Results: the study examines the issues of copyright for works created using artificial intelligence, including complex aspects of determining authorship, as well as the issues of anthropocentrism in modern legislation. The author analyzes judicial precedents, mainly in the context of the European Union legislation, which is actively adapting to new technological challenges. Various approaches are investigated to determine the legal status of works created using artificial intelligence and their impact on traditional intellectual property concepts.

Scientific novelty: the article presents a unique comprehensive assessment of the impact of the AI creative capabilities on the fundamental intellectual

property concepts. The scientific significance lies in the author's original assessment of the impact of artificial intelligence technologies on copyright legislation, based on a detailed analysis of judicial precedents and doctrinal approaches. The author investigate the prospective development of legal regulation in the context of technological progress.

Practical significance: the paper proposes legal and governmental solutions aimed at creating a balanced and effective intellectual property regime in the era of artificial intelligence. Recommendations were developed to improve legislation, taking into account existing judicial precedents and the needs of the digital economy. The research results can be used to develop new regulations and improve the existing legal framework of artificial intelligence regulation.

For citation

Tzimas, Th. (2025). Evolution of Copyright in the Era of Artificial Intelligence: Analysis of Conflicts of Law and Judicial Precedents. *Journal of Digital Technologies and Law, 3*(1), 35–64. https://doi.org/10.21202/jdtl.2025.2

Contents

Introduction
1. AI Ontology and the role of autonomy
2. The fundamentals of intellectual property norms and AI impact. The EU law perspective
3. Judicial precedents assessing the impact of AI on copyright law
Conclusions
References

Introduction

In the rapidly evolving landscape of intellectual property, the integration of Artificial Intelligence (AI) into the creative process has become an undeniable force, reshaping the very nature of original works and introducing a host of complex legal considerations (Greenstein, 2022). Legal systems are struggling to keep up with the transformative effects of AI on intellectual property[1]. At the center of these effects lies AI's unique ontology and more specifically its autonomy, which makes AI capable of producing creative and original work without any or at least any critical human intervention.

[1] Love, J. (2023, August 7). We Need Smart Intellectual Property Laws for Artificial Intelligence, Scientific American. https://clck.ru/3GEWjn ; Ogwuche, Perpetua. (2022, October 16). Artificial Intelligence: The Legal Implications of Intellectual Property Rights for AI-generated Inventions. https://clck.ru/3GEWku

Although this is a very recent legal issue, a number of court decisions from various legal systems worldwide have begun to be produced[2]. This article attempts to explore the intersection between AI-generated works and intellectual property norms with reference to relevant recent court decisions, and with a view to examining how these may affect the direction of European Union (EU) law. The point of reference consists in the role of anthropocentrism as a prerequisite for IP protection.

The implications stemming from AI's burgeoning role in the creation of copyrighted, patented, and owned works are potentially cataclysmic, because of the expanding intelligence of AI and its capacity to emulate human intelligence characteristics which at least resemble with creativity and originality. On such grounds it is reasonable to wonder about who should owe AI-generated inventions in general and in particular under EU law. The next part refers briefly to some elements of AI ontology which are critical in order to comprehend why anthropocentrism is challenged by AI in relation to IP law.

1. AI Ontology and the role of autonomy

AI is defined in the EU AI Act as "a system that is designed to operate with a certain level of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of human-defined objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts"[3]. This definition is used in the present article in order to avoid extensive parathesis of AI definitions as well as because of its comprehensive meaning. The definition in "AI act" refers to a wide range of AI output, from decision-making to content. It is on the basis of this ontology of AI that the question about the regulation of AI-generated work emerges[4].

The concept of AI, fundamentally is built around the quest to create machines which can emulate human intelligence or aspects of it; in other words, AI evolution refers to the quest for a new type of intelligent beings (Gerdes, 2018). The fundamental element of AI is its expanding, intellectual autonomy that provides it with the capacity

2 The article makes reference to the most important among them.

3 The EU Artificial Intelligence Act. https://clck.ru/3GEWoH

4 Prior to this legislative initiative, it was the European Commission's, "Artificial Intelligence for Europe", Communication from the Commission to the European Parliament, which set as goals to "Boost the EU's technological and industrial capacity and AI uptake across the economy… Prepare for socio-economic changes brought about by AI… nsure an appropriate ethical and legal framework"; EC, "Artificial Intelligence for Europe", Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM/2018/237 final, 25 April 2018.

to adapt to novel environments (Omohundro, 2008; Russell & Norvig, 2010). This is what makes AI so unique: the shift from automation, to autonomy. Autonomy means that AI does not constitute the mere outcome of predetermined, software programming but imitates and reproduces human learning procedure and aspects of human intelligence, through machine learning (McCarthy, 2008; Lake et al., 2016). Alan Turing's approach was that computers could imitate children's minds, methodology and evolution (Turing, 1950).

Machine learning means that AI is "taught" how to deliver certain outputs (Bostrom & Ćirkovic, 2008). The goal of machine learning is to achieve in terms of the intelligence of AI, natural-like, evolutionary patterns and therefore to come up with solutions to a wide range of not predetermined, problems, without necessarily having humans in the loop (Bostrom, 2014). Therefore, what AI is doing is to learn and decide, on the basis of the action with the highest expected utility, in light of the basic preferences and goals (Bostrom, 2014).

"Machine learning" takes place on the basis of bigdata harvesting and use so that algorithms can be trained. AI is trained on the basis of our collective, socially produced data and in this sense, AI autonomy is fundamentally – at least partially – a social outcome. It is this procedure that makes AI capable of developing and demonstrating characteristics such as logic[5] – as a tool of analysis[6] – creativity, problem solving, pattern recognition, classification, learning, induction, deduction, analogies building, optimization, surviving in an environment and language processing (Hutter, 2010; Hallevy, 2018), cognitive autonomy, intuition and strategic thinking (Yanisky-Ravid & Liu, 2018; Hallevy, 2018; Suchman & Weber, 2016)[7]. AI does not yet understand all these characteristics as a human does since it does not possess self-reflective intelligence but it already produces in many areas and cases, outputs that in humans prerequisite such intellectual capacities (Laton, 2016; Russell & Norvig, 2010)[8].

Machine learning also explains why as AI evolves its ontology becomes probabilistic, nonlinear, complicated, opaque and therefore unpredictable, raising fundamental uncertainties that have been described as the "black-box effect"; we cannot be certain what the outcome of machine learning and AI actions will be[9]. We know the input and we

5 Thomason, R. Logic and Artificial Intelligence. Stanford Encyclopedia of Philosophy. https://clck.ru/3GEWqZ

6 Ibid.

7 Camett, J. B., & Heinz, E. (2006, Apr 19). John Koza Built an Invention Machine. Popular Science. https://clck.ru/3GEWtB

8 Pyle, D., & San Jose, C. (2015, June). An executive's guide to machine learning. McKinsley Quarterly. https://clck.ru/3GEWuK

9 InFERENCe. (2015, August 13). The Two Kinds of Uncertainty an AI Agent Has to Represent. https://clck.ru/3GEWvd

see the output of machine learning but we are not certain of the in-between of the two (Karppi & Crawford, 2016; Van Asselt & Renn, 2011). The unique advantage of AI – its autonomy and therefore adaptability – comes with a significant risk as well – this is the "black box" (Castelvecchi, 2016).

Of course, there are serious disagreements about what and mainly when AI can achieve breakthroughs leading it to a level of general intelligence[10]. Nevertheless, even present, narrow AI autonomy produces transformative results – among other areas – in relation to original intellectual work, already significantly limiting or diminishing the human presence in the loop (Martinez, 2019).

The debate about whether AI can be creative and original or these are uniquely human characteristics is interdisciplinary and largely unanswered yet (Hashiguchi, 2017b; Hattenbach & Snyder, 2018)[11]. For a certain part, AI – at least the one that we currently have – is solely guessing patterns and therefore it can never be creative and original. From another perspective this is a highly unfair approach, overlooking that even at a limited extent, AI is emulating human mind characteristics, including aspects of creativity. What partially bypasses this ontological debate but also answers it in the area of law is that regardless of whether AI can be considered as ontologically creative or not, the fact is that it produces work which if produced by human authors would be considered creative and therefore protected under copyright norms.

Therefore, algorithms produce work which when produced by humans is protected under intellectual property norms. How must law treat such work? Should it be protected by IP law for the benefit of natural or legal persons or should it be freely accessible? (Xu et al., 2018)[12]. The next section briefly examines the foundations of IP law in general and EU law in particular. Then, the relevant judicial precedents are examined in order to assess the impact of AI on IP law in general and under EU law in particular.

---

[10] For example, and indicatively enough, while "generative AI" is considered by many as a unique scientific breakthrough, by another part of experts is downplayed in the sense that what AI does is to predict sequences on the basis of vast data. While we are not certain about what human intelligence exactly does, it certainly does "more" than the above.

[11] Gottschalk v. Benson, 409 U.S. 63, 67 (1972); Hauser, L. Artificial Intelligence. Internet Encyclopedia of Philosophy. https://clck.ru/3GEX2y

[12] Schwab, K. (2015). The Fourth Industrial Revolution: What It Means and How to Respond. https://clck.ru/3GEX4J ; Xiang, F. (2018). AI Will Spell the End of Capitalism. Available via The Washington Post. https://clck.ru/3GEX6N ; Acemoglu, D., & Restrepo, P. (2017). Robots and Jobs: Evidence from US Labor Markets. MIT Department of Economics Working Paper, 17-04. https://clck.ru/3GEX7H ; Yongjun, Xu et al. (2021, November 28). Artificial intelligence: A powerful paradigm for scientific research. The Innovation, 2, 100179.

2. The fundamentals of intellectual property norms and AI impact. The EU law perspective

In order to address the issue of AI and IP law, first the foundations of intellectual property law must be briefly examined. To begin with, intellectual property "…very broadly, means the legal rights which result from intellectual activity in the industrial, scientific, literary and artistic fields"[13]. Intellectual property law is very simple at its core: it transforms knowledge and its practical applications into economic value (Manderieux, 2010). It is supposed to be able to achieve a balance between competitive interests – private and public – and to regulate access to benefits (Pila & Torremans, 2019).

In the framework of intellectual property, scientific works belong to the copyright branch and inventions to industrial property[14]. As inventions are defined the new solutions to technical problems, whereas scientific discoveries consist of "the recognition of phenomena, properties or laws of the material universe not hitherto recognized and capable of verification"[15]. The fundamental element of intellectual property is the intersection of the creative, nonobvious, original idea or invention, with the practical application of industrial utility. The determination of the fulfillment of each specific criterion constitutes a legal challenge.

The theoretical foundations of intellectual property norms are the labor/desert and the utilitarian/incentive theory (Khoury, 2017). The former emphasizes on the reward of the work of the creator, whereas the second on the motivation to creators to further work on new ideas and new inventories (Fisher, 2001)[16]. Both of them are built on two fundamental assumptions: the first one is that there is a human author behind the protected work and the second one is that this human must be rewarded for her/his work.

The concept of intellectual property protection has been criticized on the basis of the lack of social utility of intellectual property norms as promoters of monopolies and therefore as obstacles to innovativeness (Hemel & Ouellette, 2013; Rai, 1999).

---

13 World Intellectual Property Organization. (2014). WIPO Intellectual Property Handbook. According to the Convention Establishing the World Intellectual Property Organization (WIPO), "intellectual property shall include rights relating to: literary, artistic and scientific works; performances of performing artists, phonograms and broadcasts; inventions in all fields of human endeavor; scientific discoveries; industrial designs; trademarks, service marks and commercial names and designations; protection against unfair competition, and all other rights resulting from intellectual activity in the industrial, scientific, literary or artistic fields"; Convention Establishing the World Intellectual Property Organization (as amended on September 28, 1979) (Authentic text). https://clck.ru/3GEX8a

14 Regarding AI, copyright also could be relevant given that it refers to "computerized systems for the storage and retrieval of information."; WIPO Intellectual Property Handbook, (2014).

15 World Intellectual Property Organization. (2014). WIPO Intellectual Property Handbook; The Geneva Treaty on the International Recording of Scientific Discoveries, Article 1.

16 United Nations The Role of Patents in the Transfer of Technology to Developing Countries. E. 75. II. D. 6, 1975.

Further, intellectual property rights do not emerge as other property rights out of scarcity but create scarcity eventually leading to wealth reducing, at least at a general social level (Krauss, 1989). A strong public domain is the "engine", opening the public to new ideas and inventions, whereas intellectual property protection excludes or restricts access to the protected work, therefore limiting the free flow of ideas and applications (Cohen, 2006; Salzberger, 2006). The "ocean" is the public domain and intellectual property are the "islands", which eventually "collapse" into the former (Khoury, 2017). Therefore, intellectual property rights under all legal systems must always balance with the wider public interest so that they are not abused as rights: public access must not be unfairly restricted in favor of a natural or legal person. This becomes even more critical when positioned in the framework of AI and the gradual expulsion of human from the loop.

Another common element among all legal systems is that they exclude mental activities from intellectual property protection. A variety of legal precedents have clarified this issue by introducing a crucial distinction between mental activities per se and mental activities with an industrial application. The former, mental activities without industrial application are not patent-eligible. The latter, mental activities with industrial applications may be patent-eligible on the basis of the assessment of the relationship between the mental activities and their industrial application.[17]

According to the European Patent Office, there are four basic requirements for intellectual property protection: "there must be an «invention», «susceptible of industrial application», which is «new» involving an «inventive step[18]. Under the EU legal system, it

[17] Indicatively see: (Hashiguchi, 2017a); Elec. Power Group, LLC v. Alstom S.A., 830 F.3d 1350, 1351 (Fed. Cir. 2016), 2351-2359; In re TLI Communications LLC Patent Litigation, 823 F.3d 607-613 (Fed. Cir. 2016); Alice Corp. Pty., 134 S. Ct. at 2354 (citing Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 133; Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1339 (Fed. Cir. 2016); Mayo Collaborative Servs. v. Prometheus Labs., 566 U.S. 66, 77 (2012) ("The question before us is whether the claims do significantly more than simply describe these natural relations. To put the matter more precisely, do the patent claims add enough to their statements of the correlations to allow the processes they describe to qualify as patent-eligible processes that apply natural laws?"); European Patent Office, Convention On The Grant Of European Patents 108; McRO, Inc. v. Bandai Namco Games Am. Inc., 837 F.3d 1299, 1302-1316 (Fed. Cir. 2016); Fitbit Inc. v. AliphCom, No. 16-cv-00118-BLF (N.D. Cal. Mar. 2, 2017) at 10, 22; Decision of the European Patent Office. (2004, Apr. 21). Technical Board of Appeal, Case T 258/03–3.5.1, Reasons for the Decision, 3.3, 3.7, 4.1, 4.3, 4.4, 4.7. https://clck.ru/3HrAYg ; In re Sesame Active System, 15/01962, Cour d'Appel de Paris [Court of Appeal of Paris] (26 fevrier 2016 [Feb. 26, 2016]); In re Dassault Systèmes, 14/06444, Cour d'Appel de Paris [Court of Appeal of Paris] (16 décembre 2016 [Dec. 16, 2016]); (Hashiguchi, 2017b); Decision of the European Patent Office. (1988, Oct. 5). Technical Board of Appeal, Case T 22/85–3.5.1, Reasons for the Decision, 5. https://clck.ru/3HrAYg ; Decision of the European Patent Office. (1995, Jan. 20). Technical Board of Appeal, Case T 0605/93-3.5.1, 5.3, 5.7. Reasons for the Decision, 5.9. https://clck.ru/3HrAYg ; Further, intellectual property norms do not apply to inventions that make use of the laws of nature; S. Ct. 2107, 2116 (2013); Mayo Collaborative Servs. v. Prometheus Labs., Inc., 132 S. Ct. 1289, 1293 (2012)).

[18] European Patent Office. Patentability requirements. https://clck.ru/3HrAEr ; European Patent Office, Convention on the Grant of European Patents 108 (16th ed., 2016, June). https://clck.ru/3HrAKW (compiling the European Patent Convention articles) [hereinafter European Patent Convention].

is the personality of the author that is protected by copyright laws[19] (Kur et al., 2013). The author's personality and the creativity that the latter demonstrates constitute a synthesis leading to the originality of work (Hugenholtz & Quintais, 2021). The level of creativity is assessed ad hoc and on the basis of general guidelines (van Gompel, 2014). What is undoubted is that without author's personality there is no further assessment that needs to be done[20].

The fact that the protection of the personality of the author constitutes the foundation of EU copyright law is apparent in a variety of rules: a creation cannot be modified or distorted without the permission of the author, regardless of any potential transfer of the copyright[21]; it must be associated with the author's name; the disclosure of the creation is prohibited until the author adheres to it the author retains the right to retract the creation[22]. The above elements essentially comprise the EU natural law theoretical approach (Holst, 2006; Adler, 2009).

Under EU law and in particular the influence of the European Court of Justice (ECJ) originality is linked with the "author's own intellectual creation" and therefore implicitly with human authorship (van Eechoud, 2012). It was during a short period of time, between 2009 and 2012, that the ECJ elaborated further on the linkage of originality with authorship through five decisions[23]. In these decisions it was

---

19 Lundstedt, L. (2016). Territoriality in intellectual property law: a comparative study of the interpretation and operation of the territoriality principle in the resolution of transborder intellectual property infringement disputes with respect to international civil jurisdiction, applicable law and the territorial scope of application of substantive intellectual property law in the European Union and United States: Doctoral dissertation. Stockholm University.

20 Parenthetically it must be stressed that the EU does not hold the exclusive authority to legislate on IP norms; (Kur et al., 2013); European Commission, Shaping Europe's digital future, The EU copyright; Directive 2009/24/EC of 23 April 2009 on the legal protection of computer programs, article 1(3)), Directive 96/9 of 11 March 1996 on the legal protection of databases, article 3(1)), Directive 2006/116/EC of 12 December 2006 on the term of protection of copyright and certain related rights, article 6; Case C-277/10 – Martin Luksan v. Petrus van der Let (2012) ECLI:EU:C:2012:65 (Luksan), para. 59, and Case C-310/17 – Levola Hengelo BV v. Smilde Foods BV (2018) ECLI:EU:C:2018:899 (Levola Hengelo), paras. 38–39 legislation. https://clck.ru/3GEXR2 ; The Berne convention among other things simplifies the procedures for the protection of authors' rights, establishes a minimum period of protection and protects the moral rights of authors; Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as last amended, July 24, 1971 European Commission, Commission adopts Action Plan on Intellectual Property to strengthen EU's economic resilience and recovery. https://clck.ru/3GEXRY

21 (Hansmann & Santilli, 1997). There are exceptions to copyright protection in the name of public interest for the promotion of science, education and culture, as well as for data and data mining "by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research". The InfoSoc Directive, Art. 5; Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L 130/92 -DSM Directive.

22 Ibid.

23 Infopaq International v. Danske Dagblades Forening [2009]; Bezpečnostní softwarová asociace v. Ministerstvo kultury [2010]; Football Association Premier League v. QC Leisure and Karen Murphy v. Media Protection Services [2011]; Eva-Maria Painer v. Standard VerlagsGmbH [2011]; Football Dataco v. Yahoo! [2012].

clarified that originality is established in "author's own intellectual creation", which presupposes free and creative choices. This type of choices is absent when the relevant techniques, functions or rules make it imperative for the author to express in only one specific way, therefore leaving no space for free choice. Creativity and free choice are essentially qualitative and not quantitative characteristics. It is not the effort of work that is put in each creation that matters but the level of creativity which goes with freedom of choice[24]. This is the case with both traditional works and the ones based on new technologies[25].

Until the emergence of AI, authorship was obviously human. This dimension was inferred by the fact that creativity, as well as freedom of choice, all presuppose the intellectual capacity, which was self-evidently the realm of human[26]. Since

[24] "[T]he significant labour and skill required for setting up that database cannot as such justify such a protection if they do not express any originality in the selection or arrangement of the data which that database contains" Football Dataco v Yahoo [2012], 53(1).

[25] The determination of the exact level of originality that is required remains with each member state. Directive 98/71/EC of 13October 1998 on the legal protection of designs, article17; Regulation (EC) No. 6/2002 of 12 December 2001 on community designs, article 96; Computer Programs Directive, recital 8: "In respect of the criteria to be applied in determining whether or not a computer program is an original work, no testsas to the qualitative or aesthetic merits of the program should be applied". Ramalho, A. (2019). Originality redux: an analysis of the originality requirement in AI-generated works. AIDA, 9; (Ricketson & Ginsburg, 2005); Case C-5/08 Infopaq International A/S v. Danske Dagblades Forening [2009] ECR I-6569, ECLI:EU:C:2009:465, para. 37; Case C-393/09 Bezpecˇnostn softwarova´ asociace – Svaz softwarove´ ochrany v. Ministerstvo kultury [2010], ECR I-13971, ECLI:EU:C:2010:816, para. 45; Joined Cases C-403/08 and C-429/08, Football Association Premier League v. QC Leisure and Karen Murphy v. Media Protection Services Ltd [2011] ECR I-09083, ECLI:EU:C:2011:631, para. 97; Guide to the Berne Convention (1978), 17–18; (Hutukka, 2023).

[26] The degree of creativity varies among different legal systems with the threshold being higher – such as for example in the US- or lower depending on the legal system and tradition. In the Painer case, which was about a photographic portrait the decision it was determined that the author: "can make free and creative choices in several ways and at various points in its production. […] By making those various choices, the author of a portrait photograph can stamp the work created with his "personal touch". Consequently, as regards a portrait photograph, the freedom available to the author to exercise his creative abilities will not necessarily be minor or even nonexistent". In the Cofemel case the Court argued that "if a subject matter is to be capable of being regarded as original, it is both necessary and sufficient that the subject matter reflects the personality of its author, as an expression of his free and creative choices". Feist Publications v. Rural Telephone Service 499 U.S. 340 (1991), 346; CCH Canadian v. Law Society of Upper Canada [2004] 1 S.C.R. 339; Case C-145/10 – Painer, paras. 90–93; Case C-145/10 – Painer, para. 92; Case C-683/17 – Cofemel, para. 30; Case C-5/08 Infopaq International A/S v. Danske Dagblades Forening [2009] ECR I-6569; ECLI:EU:C:2009:465, para. 37; Case C-393/09 Bezpecˇnostnı´ softwarova´ asociace - Svaz softwarove´ ochrany v. Ministerstvo kultury [2010], ECR I-13971, ECLI:EU:C:2010:816, para. 45; Joined Cases C-403/08 and C-429/08, Football Association Premier League v. QC Leisure and Karen Murphy v. Media Protection Services Ltd. [2011] ECR I-09083, ECLI:EU:C:2011:631, para. 97; Case C-604/10 Football Dataco and Others v. Yahoo! UK Ltd. and Others[2012]; ECLI:EU:C:2012:115, para. 38; Case C-5/08 Infopaq International A/S v. Danske Dagblades Forening [2009] ECR I-6569, ECLI:EU:C:2009:465, para. 45; Case C-393/09 Bezpecˇnostnı´softwarova´ asociace – Svaz softwarove´ ochrany v. Ministerstvo kultury [2010], ECR I-13971, ECLI:EU:C:2010:816, para. 50; Case C-145/10 Eva-Maria Painer v. Standard VerlagsGmbH and Others [2011] ECLI:EU:C:2011:798,paras. 89, 92; Parenthetically, originality describes a condition where the work is not copied and is the result of "skill, judgement and/or labour"; (Bently & Sherman, 2014).

intellectual property protects author's personality, what personality is to be protected if there is no human author?[27] It is in this relationship that AI steps in, which explains why all these self-evident until recently facts are tested under the influence of AI.

The advent of technology, as already mentioned prompted fundamental reconsiderations of the traditional concept of authorship even prior to AI. From the 1988 Green Paper, to the Directive 91/250/EEC of 14 May 1991 on the legal protection of computer programs, the issue of human authorship gradually emerges as a matter of debate (Walter & von Lewinski, 2010). Until AI there was little doubt about the anthropocentrism of authorship. Computers are "automata", not autonomous. AI however, through its variety of applications can produce the type of work that until the AI era could be only the output of human intelligence (Hugenholtz & Quintais, 2021). Even existing, narrow AI (ANI) can by now produce creations in practically all the areas of human creativity and intellectual activity (Senftleben & Buijtelaar, 2020; Gervais, 2019; Senftleben & Buijtelaar, 2020; Butler, 1982).

Given that human presence in the loop – or on the loop – decreases, the linear causality between the (distant) human "mind" behind an AI algorithm and the final invention or work sublimates at least up to a significant extent or even – by now – completely. How are we going to assess it so that we can reach a conclusion about whether copyright protection should be provided or not (Hashiguchi, 2017a)? When is it not fair anymore to reward a human for AI's work (Spector, 2006; Jaszi, 1992; Grinmelmann, 2016)[28]?

---

[27] For example InfoSoc case by the CJEU at the expense of legal entities. In even clearer terms, it was Advocate General Trstenjak in her opinion in the Painer case, the one who directly linked intellectual property norms with human nature: "only human creations are therefore protected…"; Case C-277/10 – Luksan; Case C-572/13 – Hewlett-Packard Belgium SPRL v. Reprobel SCRL (2015); ECLI:EU:C:2015:750 (Reprobel); Opinion AG Trstenjak in Case C-145/10 – Painer, para. 121; Identical is the US Copyright Office approach as well: "works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author" if it determines that a human being did not create the work"; USPTO, Compendium of U.S. Copyright Office Practices § 101 (3rd edn. 2017). https://clck.ru/3GEXSn , Arts. 306 and 313(2).

[28] Sloman, A. (2007). What is Artificial Intelligence?, School of Computer Science The University of Birmingham. https://clck.ru/3Hr9gj ; Burrow-Giles Lithographic Co. v. Sarony, Ill U.S. 53 (1884); Midway Mfg. Co. v. Artic Intern., Inc., 704 F. 2d 1009, 1011 (7d Cir. 1983). Back in 1965, in the US, the Register of Copyrights submitted to the Congress, a report, about computer-generated work, raising the issue of copyright, given that part of such work is generated by computer. The, then established National Commission on New Technological Uses of Copyrighted Works (CONTU) held that computers are no different from cameras or typewriters, with copyright belonging only to the user. U.S. COPYRIGHT OFFICE, SIXTY-EIGHTH ANNUAL REPORT OF THE REGISTER OF COPYRIGHTS 5 (1965). https://clck.ru/3GEo8C ; NAT'L COMM'N ON NEW TECH. USES OF COPYRIGHTED WORKS, FINAL REPORT OF THE NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS 44-45 (1978). https://clck.ru/3GEXrV ; Arsheeya Bajwa. IBM beats profit estimates as AI shift boosts software performance, shares surge. (2025, January 30). Reuters. https://clck.ru/3HK9es ; van den Oord, A., et al. (2016, Sept. 8). WaveNet: A Generative Model for Raw Audio. arXiv. https://clck.ru/3HK9hP ; Mordvintsev, A. et al. (2015, June 17). Inceptionism: Going Deeper into Neural Networks, GOOGLE RES. BLOG. https://clck.ru/3HK9ip ; (Hashiguchi, 2017a).

In this context, a number of court rulings internationally help us further elaborate our approach to the effect of AI on the requirement of an author who must be protected under IP law.

3. Judicial precedents assessing the impact of AI on copyright law

The relevant judicial precedents emerge both from EU member states' and internationally. All of them are useful in order to clarify the intersection of AI and IP protection in general and under EU law.

A US Court, the District Court of Columbia held that an entirely AI-generated artwork, in the production of which there is not any human involvement, is ineligible for copyright protection because of the lack of human authorship. The case was built on the request of the owner of a computer system called "Creativity Machine" to register for copyright protection the visual art which was produced by the AI and then transfer to him the copyright because he was the owner of the system. The US copyright office declined the copyright protection because of the lack of human authorship. The plaintiff invoked the common law "work-for-hire" doctrine in his favor. The Court held that these arguments concern "…to whom a valid copyright should have been registered, and in so doing put the cart before the horse". it further held that "Copyright is designed to adapt with the times. Underlying that adaptability, however, has been a consistent understanding that human creativity is the sine qua non at the core of copyrightability, even as that human creativity is channeled through new tools or into new media"[29].

In another case, the US Copyright Office concluded that AI-generated content must be disclaimed in the registration application to provide the Office with the information relevant to the preparation or identification of the work or to the existence, ownership or duration of the copyright, eventually refusing to register under copyright norms the relevant AI-generated work. In particular the Board found that "… the Work contains more than a de minimis amount of content generated by artificial intelligence ("AI"), and this content must therefore be disclaimed in an application for registration"[30]. The work under assessment was upgraded by the human who claimed copyright protection but the Board found that since he refused to disclaim the material produced by AI and this material exceeded a de minimis amount of AI-generated content, copyright protection could not be provided. The Board in its decision reiterated the findings of the Thaler v. Perlmutter case in which it was held that "human authorship is a bedrock requirement

[29] US District Court For The District Of Columbia, Stephen Thaler v Shira Perlmutter, Case 1:22-cv-01564-BAH Document 24 Filed 08/18/23, at pp. 7, 8.

[30] United States Copyright Office, Second Request for Reconsideration for Refusal to Register Théâtre D'opéra Spatial (SR # 1-11743923581; Correspondence ID: 1-5T5320R), (2023, September 5).

of copyright"[31]. The Board also mentioned another famous case, "Urantia Found. v. Kristen Maaherra", the judge of which had found that "some element of human creativity must have occurred in order for the Book to be copyrightable" and that "it is not creations of divine beings that the copyright laws were intended to protect"[32].

The US copyright office has been consistent in that the fundamental necessity for copyright protection is human authorship. In this framework it issued public guidance according to which the fundamental question consists in ""whether the 'work' is basically one of human authorship, with the computer [or other device] merely being an assisting instrument, or whether the traditional elements of authorship in the work (literary, artistic, or musical expression or elements of selection, arrangement, etc.) were actually conceived and executed not by man but by a machine"[33].

The United States Patent and Trademark Office (USPTO) Inventorship Guidance for AI-Assisted Inventions, which came into effect by February 13, 2024, establish joint inventorship rules between human and AI in AI-assisted inventions. In cases where a human and generative AI are each instrumental in creating an invention they are determined as co-inventors. Given however the fact that AI cannot be protected as inventor the crucial issue becomes whether, for each patent application, "at least one natural person has made a "significant contribution" that satisfies the joint inventorship Pannu factors required of being an inventor. If not, then the invention cannot be patented because there is no inventor to list"[34].

In order therefore to have intellectual property protection in a joint human-AI scheme, the human co-inventor must be able to prove that simultaneously and crucially participated both in the conception and in the industrial application of the invention; not only in one of the two. The USPTO guidelines are particularly helpful in cases such as for example when the human simply owns the AI system, or provides to the latter the problems that it must solve. These are not cases of IP protection. On the contrary, the design, building and training of AI in order to solve a particular problem, as well as a critical participation of the human in the problem-solving procedure of AI may lead to recognition of inventorship[35]. The afore-mentioned guidelines coincided with

31 Ibid.

32 Ibid.

33 Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16,190, 16,192 (Mar. 16, 2023) ("AI Registration Guidance").

34 The Pannu factors consist of the significance of the contribution on the conception and materialization of the invention, the significance of the qualitative contribution and originality/ creativity in the sense of doing more than explaining well- known concepts. Katsulis, A. (2024, May 14). Clarifying AI and inventorship: USPTO's guidance for AI-assisted inventions. Inside Tech Law. https://clck.ru/3GEY7Q

35 Ibid.

the judicial precedent of the Thaler v Vidal case, where the recognition of AI as inventor was declined[36].

Both US courts and authorities have contributed significantly in the relationship between AI and intellectual property: AI-oriented inventions are in principle excluded from IP protection. It is not the actual capacity of AI to invent things that is denied but its potential protection under IP law. In this sense, IP protection is safeguarded only after careful, case-by-case, quantitative and qualitative examination, assessing the role of the human factor and AI work[37].

The Internet Court of Beijing partially distanced itself from the afore-mentioned reasoning in a similar case[38]. It also made a quantitative assessment between human work and AI but found that the work was eligible for copyright protection because it was original. Its originality stemmed "from the numerous positive and negative prompts inserted and the adjustments and amendments made by the human user to select the final image that matched his expectations"[39]. The Court held that in spite of AI involvement, the AI-generated image "… reflected the plaintiff's individual creativity and aesthetic choices made during the creation process". The plaintiff's creative intellectual input included designing, selecting prompts, setting parameters during the image's creation, amending and adjusting the output image several times, until he reached a final image that matched his expectations. In light of the above, the Court found that "AI models lack legal personality and humans remain the creators of works generated using this technology"[40].

The Court in particular has reached a rather problematic conclusion in this regard. On the one hand, it maintains the assessment of the degree of impact of human intelligence on the produced work on a case-by-case basis, but on the other hand it lowered the threshold of necessary human participation considerably. Therefore, while it does not deny the prerequisite of human authorship in principle, it practically relativizes its significance. Further, the fact that AI does not have legal personality should not imply that copyright passes to its human owner automatically. Why the lowering of the threshold of necessary human involvement is problematic becomes profound on the basis of the founding on intellectual property protection on human personality, as among various cases,

36 Thaler v. Vidal, 43 F.4th 1207 (Fed. Cir. 2022).

37 The Canadian Intellectual Property Office (CIPO) followed a similar path, but in a somewhat ambiguous way. On the one hand, it ruled that AI could not be the inventor, and on the other, it provided a loophole by proposing the applicant to submit an application on behalf of the artificial intelligence system and identify oneself as its legal representative.

38 US and Chine pioneer in the evolution of AI. Their Courts' decisions therefore are points of reference.

39 European Union Intellectual Property Office. (2024, May). Recent European Case-Law On The Infringement And Enforcement Of Intellectual Property Rights, at pp. 5–6.

40 Ibid, at p. 5.

the recent one "Lithoss Nv V Vimar S.P.A. And Vecolux BV" reiterated[41]. The reasoning of the Beijing Court seems to falsely confuse automation with autonomy. It also seems to be partially "manipulating" the true ontology of AI in furtherance of not jeopardizing the continuation of intellectual property protection and profit.

Chinese courts nevertheless have also come to decisions which are better aligned with the goals of intellectual property norms and their emergence from human authorship. The Beijing Internet Court held in the "Beijing Film Law Firm v Beijing Baidu Netcom Science & Technology Co Ltd (Film)" case, that creation by natural persons was a prerequisite for protection under the Chinese copyright law. In that case the work produced by AI was not to be protected by copyright law, regardless of its originality (Yong Wan & Hongxuyang Lu, 2021).

Not all states' copyright laws and judicial precedents follow the same logic with the US and the EU or at least not always: the United Kingdom, South Africa, New Zealand, Ireland, and India, all have recognized copyright protection for computer-generated work, even absent any human intervention. According to the High Court of England and Wales (High Court of Justice), composite screen frames generated by a computer program of a coin operated video game are computer-generated works because the software built up composite images by overlaying the digital image of a pool table with images of the balls and cue.[42] "In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken"[43]. While UK courts decrease up to some extent the level of copyright protection in the afore-mentioned cases and the subsequent profitability from the relevant copyright protection, still in that instance had reached a legally and ethically unjustifiable norm of securing profits and ownership for humans who have shown no real authorship. Such approaches constitute an abuse both of the intellectual property norms and rights, as well as of the employer-employee relationship, given as previously mentioned that machines are not employees but owned technology, part of capital (Hristov, 2017).

Nevertheless, in the "Comptroller-General of Patents, Designs and Trade Marks v Emotional Perception AI Ltd" case, the English Court of Appeal overturned the previous argumentation and found that an Artificial Neural Network (ANN) was not patentable. The Court dealt with three questions: "Is an ANN a "computer"? If it is a computer, is an ANN a program for a computer "as such", within the meaning of section 1(2) of the Patents

41 Antwerp Court of Appeal, 2021/AR/1900, LITHOSS NV v VIMAR S.p.A. and VECOLUX BV [13 September 2023].

42 Nova Productions v. Mazooma Games, [2006] EWHC 24 (Ch) (UK).

43 Copyright, Designs, and Patents Act 1988, c.48 §§ 12(7), 79(2), 81(2).

Act 1977 (meaning that it would be excluded from patentability unless the next bullet applies). If the ANN is a program for a computer, did the ANN nevertheless fall outside the computer program exclusion (and so be patentable) because it had a "technical contribution" outside itself?"[44] The Court reached the conclusion that ANN is a computer, functioning as a computer program and in this sense that its recommendations are not technical but a matter of aesthetics. In this sense it "corrected" the precedent of the High Court and ascertained that there can be no IP protection for AI generated work.

The Federal Court of Australia, as well, in the Thaler case of 2022 had concluded that AI cannot be recognized as an inventor[45]. The Court had found that only natural persons can be labelled as "inventors" and that it is necessary to have a legal relationship between the human inventor"… and the person first entitled to the grant, which is a legal impossibility in… case where the purported inventor has no legal identity and therefore cannot give effect to an assignment"[46].

A crucial decision albeit with a problematic syllogism was issued by the German Federal Court of Justice (Bundesgerichtshof – BGH), in the "DABUS" case, according to which only natural persons can be named as inventors under IP protection law. The German Court reached three landmark points in its judgment: first that AI can not be the inventor; second that behind every AI-generated invention there is a certain level of human contribution, which even if not inventive or substantial can lead to the designation of a human as the inventor provided that (s)he is the one with the decisive influence; and third that patent applications must not include contradictory statements. It must be declared that either the human or AI conceived the invention[47]. The decision is crucial but at the same time characteristic of the perplexity of certain part of the judiciary. In fact, it is crucial especially because of this last reason.

The court confused the recognition of AI under intellectual property law as inventor with the actual capacity to conceive inventions. In this sense it confused law with ontology. The true ability of AI to conceive inventions is almost undoubted. This decision ignores this real condition. Instead, the court's attempt in this decision appears to be to ignore AI's ability in question, to banish AI from the true ability of inventing, and then, on the basis of the above false conception, to attribute invention to a human, even if his/her actual participation is entirely secondary or even insignificant. The Court

---

44 Maloshchinskaia, P. (2024, July). Artificial intelligence: English Court of Appeal decides artificial neural network is not patentable. Inside Tech Law. https://clck.ru/3GEY92

45 Commissioner of Patents v Thaler [2022] FCAFC 62 (Thaler FC).

46 O'Brien, J., & Taylor, I. (2022, May 5). Demise of the machines: Full Court of the Federal Court of Australia overturns ruling on AI as a patent 'inventor'. Inside Tech Law. https://clck.ru/3GEYAf

47 Kalhor-Witzel, R. (2024, July). Germany: AI cannot be named as inventor – insights from the Bundesgerichtshof's DABUS decision, Norton Rose Fulbright. https://clck.ru/3GEYCa

limits itself only to the formality of whether only human is referred to as inventor or not. This is an attempt by the court to cling to the traditional approach to IP law, without taking into account the transformations taking place, resulting in an abuse of rights and unfair restriction of public access to the benefits of AI. The problem with the Court decision is not that it does not recognize AI as inventor under IP law but that it insists to do so with a human, regardless of the latter's actual role in the invention. The result of this approach is to limit public access to an invention when there is really no human personality that needs to be protected. In this sense, it conflicts with EU law's natural law-oriented approach to the need to protect the personality of the person who actually invents something original thanks to his creative conception.

Czech Courts produced a landmark decision, consistent with the requirement of human authorship. The case was brought in front of the Municipal Court in Prague and the plaintiff had used DALL-E, an AI program, in order to generate an image according to his request: "create a visual representation of two parties signing a business contract in a formal setting, such as a conference room or a law firm office in Prague. Just show your hands"[48]. The image that was used by the plaintiff on his website, was copied and posted without his authorization by the defendant. The argument of the plaintiff did not doubt that it was AI that created the image but claimed that because of his assignment to the AI he should be copyright protected as the author. The Court rejected his argument by reaching the decision that first AI cannot be recognized as the author and secondly there was no unique creativity in his action. The Prague Court very adamantly held that the AI-generated work did not constitute a «work» because it was not the unique result of the creative conception of a natural person. According to the court: «Copyright is an absolute right belonging to an individual. If the image in question was not created personally by the applicant, but by an artificial intelligence, it cannot, by definition, be a copyrighted work"[49].

All the above jurisprudence, despite its internal contradictions, is particularly illuminating for the effect of AI-oriented or generated inventions on intellectual property norms internationally and in particular under EU Law. The first and fundamental point of reference is that copyright protection is exclusively safeguarded for human authors. It is a timeless and international common foundation of IP law. Characteristic in these regards – before the AI era – is the precedent in the Burrow-Giles Lithographic Co. v. Sarony case, where the Court held that the crucial factor for copyright protection is the originality of the ideas of the author. The definition of the term "author" is according to the court anthropocentric through the characterization of copyright as "the exclusive

48 Czech court finds AI tool cannot be an author of a copyright work. https://goo.su/MQNOed

49 Novagraaf Team. (2024, May 1). AI and copyright: First ruling from a European court, Novagraaf. https://clck.ru/3GEYFv

right of a man to the production of his own genius or intellect"[50]. This reasoning has been followed both in the era of automation, as well as in the era of autonomy, i.e. of AI.

The central question is if a natural person can be protected as the author of the work that has been generated by AI. Different approaches can be found here. The common thread between case-law from different legal systems is that copyright is not automatically recognized in favor of the natural person who owns the AI system. In some legal systems, mainly that of the USA and EU, there is the requirement of a substantial, significant contribution of the natural person to the invention, both in terms of conception and industrial application, in order to be recognized as a co-inventor. In other legal systems, as we observed, it is easier to deliver intellectual property protection even with a minimal participation of the natural person. In any case, however, the weighting of the degree of human participation, whether a higher or lower level of contribution is required, is also a horizontal among different legal systems point of reference.

In order to assess the contribution of these judicial precedents in the clarification of EU law we must keep in mind that copyright protection under EU law is founded on natural law theories. Intellectual property protection exists not for speculative reasons but as a material reward of the personality of the human author and more specifically of her/his creativity. Absent this element, there is no moral and legal foundation for copyright protection in general and especially under EU law (Sobel, 2017). In principle, there is no sense of fairness in protecting AI-oriented creations under intellectual property norms. The attribution of such rights to a natural person for work that does not include her/ his own creativity constitutes an abuse of intellectual property norms since a human or a legal entity will be profiting by restricting the wider public's access to work that was not of any human author.

In order to understand even better how unbalanced and disproportionate the attribution of such rights is, the role of machine learning must be also taken into account, which is making AI at a significant extent a social project, given that machine learning is conducted on the basis of collective, big data, produced by us all.

In addition, intellectual property norms by definition are not designed in order to create a framework of "the goose that lays the golden egg", or to disproportionately restrict public access and interest. Intellectual property is designed to protect each specific human author's creativity leading to original work with industrial application as an "island" in the "ocean" of ecumenically accessible knowledge and applications.

The definition of creativity under EU law also deserves attention in conjunction with the afore-mentioned judicial precedents as well. The CJEU talks about "…creative

[50] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 54, 56, 58, 61 (1884).

abilities [of the author] in the production of the work by making free and creative choices"[51]. The word "free" implies "autonomy". From such a perspective, AI already can at least up to some extent make choices which if done by humans would be considered as free choices and therefore be creative. Once this is the case, there is no human creativity to be rewarded (Ginsburg & Budiardjo, 2019). There is only AI's creativity which is excluded from IP protection however.

An attempt to "come around" the lack of human authorship is the one arguing in favor of re-defining the concept of employer and employee (Hristov, 2017). Under this approach, the relationship between the human owner and the AI is considered as an adjusted relationship between employer and employee. The problem persists however: there is no human personality being the main – even more the sole – author of the copyrighted work and therefore nothing to be protected under IP protection norms. Copyright protection regardless of the lack of human author would simply constitute an abuse of the right. It is very simple: the foundations of copyright protection for AI-generated work are not there since no human author's personality can be found. Copyright constitutes an anthropocentric concept (Zurth, 2021). In a characteristic case brought in front of a US Court about copyright from a photo clicked by a monkey it was very clearly held that only human authorship can be protected under copyright law: "[W]e conclude that this monkey – and all animals, since they are not human – lacks statutory standing under the Copyright Act"[52].

Further, "baptizing" machines as employees is simply arbitrary. All types of machines, until at least they reach the level of general or super intelligence and the equivalent, fully autonomous legal personality is considered as part of capital. The ECJ has also answered to the afore-mentioned approach: copyright protection's prerequisite is the expression of author's "creative abilities in the production of the work by making free and creative choices"[53]. It is not the owner who makes free and creative choices but the one conceiving the idea and producing the original work.

After all, under EU law and according to the European Court of Justice precedents, there is no doubt about the linkage between author's personality and intellectual property; any work is original once it is the "author's own intellectual

51 Case C-469/17 – Funke Medien, para. 19; Case C-145/10 – Painer, paras. 87–88.

52 Naruto v. Slater, 888 F.3d 418, 420 (9th Cir. 2018)").

53 Case C-145/10, Painer v. Standard VerlagsGmbH, 2011 E.C.R. I-12594, 89; see also Case C-604/10, Football Dataco Ltd. v. Yahoo! UK Ltd., ECLI:EU:C:2012:115, 38 (Mar. 1, 2012).

creation"[54]. Even skill and work are secondary to the "personal stamp" element. They cannot justify protection by themselves but constitute elements in the wider scheme of intellectual property protection[55]. The potential protection of AI-oriented work under copyright law constitutes a violation of the "alterum non laedere" principle: public access to AI benefits will be restricted without any fair and legitimate basis for this to happen. Such an approach is both disproportionate and unreasonable (Sganga & Scalzini, 2017; Mizaras, 2012). For intellectual property norms to apply they must be relevant with each case, proportional and fair.[56] This is obviously not the case when one of the fundamental pillars of intellectual property – human author – are absent. All of the above lead us to the same direction: copyright law is anthropocentric (Ginsburg, 2018).

The ineligibility of AI – oriented work for copyright protection however does not completely clarify the issue of the required threshold for human contribution in order to talk about human author or co-inventor. A more difficult question in other words, is whether we can have some qualitative and quantitative criterion about when AI's intervention is so catalytic that human can no more be considered as the author of the work under potential copyright protection.

The threshold of human contribution is as shown above, debated. As already mentioned, EU law in this area is natural law-oriented. Therefore, the EU legal system must be interpreted as leaning to impose the highest possible threshold of human contribution in order to have copyright protection. Such an approach aligns with EU Courts' decision – and the ones of the US. The human author's personality must be directly linked with the invention either in the sense of being the sole inventor or as co-inventor. This means that for example it is not enough to simply present the question to AI but that either AI does some secondary work or solei or at least that the human has re-arranged and combined AI's outputs in a creative way.

---

54 Infopaq International v. Danske Dagblades Forening [2009]; C393/09 Bezpečnostní softwarová asociace v. Ministerstvo kultury [2010] E.C.R. I-13971; C-403/08 and C-429/08 Football Association Premier League and Others v. QC Leisure and Others and Karen Murphy v. Media Protection Services [2011]E.C.R. I-09083; C-145/10 Eva-Maria Painer v. Standard VerlagsGmbH and Others [2011] E.C.R I-12533; C-604/10 Football Dataco v. Yahoo! UK and Others [2012] EU:C:2012:115) Football Dataco v Yahoo [2012], 53 (1): "The significant labour and skill required for setting up that database cannot as such justify such a protection if they do not express any originality in the selection or arrangement of the data which that database contains".

55 Directive 2009/24, of the European Parliament and of the Council of 23 April 2009 on the Legal Protection ofComputer Programs, art. 1, 3, 2009 O.J. (L 111) 16, 18 (EC).

56 Productores de Música de España (Promusicae) v Telefónica de España SAU, Case C-275/06, [2008] ECR I-271, para. 68. Football Association Premier League Ltd and Others v QC Leisure and Others, Case C-403/08, Karen Murphy v Media Protection Services Ltd, Case C-429/08, (2012) EWHC 108.

The concept of abuse of copyright law can further clarify the issue at hand. While the issue of the potential abuse of copyright law rights because of AI-oriented inventions has not been fully developed, still some guidelines can be inferred from EU courts' past decisions in relation to human author and his/her behavior as rightsholder, in cases such as "Deutsche Grammophon", "Coditel" I "Coditel II" and "Metronome Musik"[57]. The fundamental notion in all of the afore-mentioned cases is that of reasonableness, proportionality and appropriateness in "the protection of the moral and economic rights" of the author[58]. When there is a human author reasonableness, proportionality and appropriateness take the form of reasonable renumeration from the commercial use of the creation. When there is no human author, the same concepts should take the form of non-attribution of intellectual property norms due to the ellipsis of the primary prerequisite of copyright protection. When there is some human interference, this should be proven to be creative outside the output of AI. According to the CJEU copyright laws require the "…"indispensable" intervention by the operator (without this intervention, the customers would not be able to enjoy the work)" (Xalabarder, 2016). In AI- oriented work the question must be: is the human contribution in the invention indispensable for its materialization?

On such grounds, a normative framework which will provide us with guidance on a step-by-step basis is required. Just as the EU AI Act was adopted, a law about AI and intellectual property can be designed and put in practice as well. Ahead of such a potential legal development we need to envisage what could be the case for such a development.

An initial prerequisite could consist in the obligation to maintain accessible by the relevant patent authorities "log files" showing every step until the final work is produced and each "actor's" – both human and AI – participation in the final outcome. This is a "sine qua non" so that we can then assess qualitatively and quantitatively each actor's contribution. The log files should be submitted to the relevant authorities and in order for a human to be eligible for copyright protection these files should be accessible for assessment.

On the basis of this first step, the second one could be the reconstruction of the creative process. Some crucial thresholds of creativity must be determined: the selection of the area of interest, meaning in which scientific and industrial area an effort to become creative will take place; the conception of the original idea and of its later versions; the data on the basis of which the training will take place; the repetitive work until the original idea or its versions have been finalized. It is obvious that not all of these steps share the same qualitative value in the conception and the produce of the final creative work.

---

57 Case C-78/70 Deutsche Grammophon Gesellschaft mbH v. Metro-SB-Großma¨rkte GmbH & Co; KG.Deutsche Grammophon v. Metro SB [1971] ECR 487, para. 11; Case C-262/81 Coditel v. Cine´Vog Films II (Coditel II) [1982] ECR 3381; Case C-200/96, Metronome Musik GmbH v. Music Point Hokamp GmbH [1998] ECR I-1953; (Xalabarder, 2016).

58 Case 158/86 Warner Brothers and Another v. Christiansen [1988] ECR 2605, para. 13.

Through the log files the relevant authorities will be able to determine who is the subject of the causal link between the conception of the idea and the industrial application. In order for a work to be copyright-eligible, it should be proven that it is the human intelligence the one behind both the conception of the idea and the work that is needed in order for this idea to become industrial application. If the human fails to prove both the conception of the idea and its transformation to industrial application, then there can be no copyright protection. Secondary contributions by the human, such as improvements in the final work or only partial revisions of it will not be enough to justify copyright protection. Obviously, the relevant administrative and judicial authorities will have to make assessments.

There may be cases of course where regardless of the provision of whatever log files it will be impossible to determine whether it is the human or the artificial intelligence the subject of the causal link, due to a constant, back and forth interaction between the two. In such a case, of close and equally creative collaboration between human and artificial intelligence, again it would be unfair for the human to gain profit since (s)he is not the sole creator. Even if the time frame of copyright protection is reduced, for this period, the human will benefit from something that is not only human work. Therefore, the request must be for solely human authorship and for the burden of proof on the human.

Conclusions

EU law has been facing increasing challenges in terms of dealing with the impact of emerging technologies and especially AI on intellectual property norms (Rosati, 2014). In fact, it is not only the EU legal system that faces such difficulties but legal systems all over the world. AI is raising new challenges in front both of lawmakers and courts – which especially under EU law have played significant role in shaping the intellectual property norms (Favale et al., 2016). It cannot be questioned that AI transforms the intellectual property landscape at an unprecedented degree (Cabay & Lambrecht, 2015).

Intellectual property norms are not absolute. They must be balanced with public interest and competitive rights. Several EU member-states' national legislations contain such provisions, as well as the EU law. While the exact extent of intellectual property protection both at the constitutional level as well as at the level of ordinary laws differs among member-states, the need for balanced approach is not questioned[59]. The CJEU has adopted this position as well. In the "Scarlet Extended" and in "NetLog" cases it held

[59] Geller, P. E. (2009–2010). A German Approach to Fair Use. Test Cases for TRIPS Criteria for Copyright Limitations, in 57 Journal of the Copyright Society of the USA 553, 907; Moscarini, A.(2006). Proprietà privata e tradizioni costituzionali comuni, Milano, 2006, 161 ff.

that while intellectual property rights are protected, "there is (…) nothing whatsoever in the wording of that provision or in the Court's case-law to suggest that that right is inviolable and must for that reason be absolutely protected"[60]. Intellectual property protection must be fair towards the author – rewarding author's personality – and also proportional in relation to public interest which needs the widest possible access to knowledge, creations and industrial applications.

The proposal of this article is that in AI-oriented work, AI can be ontologically creative in spite of the fact that its inventions cannot be protected under copyright law. AI is not mere automata but a unique, distinct actor with capacities of autonomous creation. The prohibition of public access to its work therefore, is unfair, disproportionate and abusive under intellectual property norms, both under EU law and international copyright law. What is needed is in fact a new set of norms and regulations under EU law, a type of AI autonomy standards and metrics that can guide us in terms of when an AI entity is so autonomous that its outputs must be freely accessible by us all. The principle must be ecumenical access to AI-generated work. This must be the new guiding principle in the emerging era of AI-generated work.

References

Adler, A. (2009.). Against moral rights. *California Law Review*, *97*, 263–301.

Bently, L., & Sherman, B. (2014). *Intellectual property law*. New York: Oxford University Press.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

Bostrom, N., & Ćirkovic, M. (2008). Introduction. In N. Bostrom, M., Ćirkovic, & M. Rees (Eds.), *Global Catastrophic Risks*. Oxford: Oxford University Press.

Butler, T. L. (1982). Can a computer be an author – copyright aspects of artificial intelligence. *Hastings Communications and Entertainment Law Journal*, *4*, 707.

Cabay, J., & Lambrecht, M. (2015). Remix prohibited: how rigid EU copyright laws inhibit creativity. *Journal of Intellectual Property Law & Practice*, *10*(5), 359–377. https://doi.org/10.1093/jiplp/jpv015

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, *538*(7623), 22. https://doi.org/10.1038/538020a

Cohen, Ju. E. (2006). Copyright, Commodification, and Culture: Locating the Public Domain. In L. Guibault, & P. B. Hugenholtz (Eds.), *The Future of the Public Domain* (pp. 121–166). Kluwer Law Intl.

Favale, M., Kretschmer, M., & Torremans, P. C. (2016). Is there an EU Copyright Jurisprudence? An Empirical Analysis of the Workings of the European Court of Justice. *The Modern Law Review*, *79*, 31–75. https://doi.org/10.1111/1468-2230.12166

Fisher, W. W. (2001). Theories of Intellectual Property. In S. Munzer (Ed.), *New Essays in the Legal and Political Theory of Property* (pp. 168–199). Cambridge University Press.

Gerdes, A. (2018). An Inclusive Ethical Design Perspective for a Flourishing Future with Artificial Intelligent Systems. *European Journal of Risk Regulation*, *9*(4), 677–689. https://doi.org/10.1017/err.2018.62

Gervais, D. J. (2019). The machine as author. *Iowa Law Review*, *105*, 2053–2106.

Ginsburg, J. C. (2018). People Not Machines: Authorship and What It Means in the Berne Convention. *International Review of Intellectual Property and Competition Law (IIC)*, *49*, 131. https://doi.org/10.1007/S40319-018-0670-X

Ginsburg, J. C., & Budiardjo, L. A. (2019). Authors and machines. *Berkeley Technology Law Journal*, *34*(2), 343. https://doi.org/10.2139/ssrn.3233885

Greenstein, S. (2022). Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*, *30*, 291–323. https://doi.org/10.1007/s10506-021-09294-4

[60] Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM), Case C-70/10, 24 November 2011, para. 43.

Grinmelmann, J. (2016). There's No Such Thing as a Computer-Authored Work – and It's a Good Thing, Too. *Columbia Journal of Law & the Arts*, *39*, 403. https://doi.org/10.31228/osf.io/rk8cm

Hallevy, G. (2018). *Dangerous Robots – Artificial Intelligence vs. Human Intelligence*. http://dx.doi.org/10.2139/ssrn.3121905

Hansmann, H., & Santilli, M. (1997). Authors' and artists' moral rights: a comparative legal and economic analysis. *The Journal of Legal Studies*, *26*(1), 95. https://doi.org/10.1086/467990

Hashiguchi, M. (2017a). The global artificial intelligence revolution challenges patent eligibility laws. *Journal of Business & Technology Law*, *13*(1).

Hashiguchi, M. (2017b). Artificial intelligence and the jurisprudence of patent eligibility in the United States, Europe, and Japan. *Intellectual Property & Technology Law Journal*, *29*(12), 3–15.

Hattenbach, B., & Snyder, G. (2018). Rethinking the mental steps doctrine and other barriers to patentability of artificial intelligence. *Columbia Science and Technology Law Review*, *19*(2), 313–339.

Hemel, D. J., & Ouellette, L. L. (2013). Beyond the Patents–Prizes Debate. *Texas Law Review*, *92*(2), 303. https://doi.org/10.2139/ssrn.2245691

Holst, K. (2006.). A case of bad credit?: The United States and the protection of moral rights in intellectual property law. *Buffalo Intellectual Property Law Journal*, *3*(2), 105.

Hristov, K. (2017). Artificial Intelligence and the Copyright Dilemma. *IDEA*, *57*, 431.

Hugenholtz, P. B., & Quintais, J. P. (2021). Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output? (2021). *International Review of Intellectual Property and Competition Law – IIC*, 52, 1190–1216. https://doi.org/10.1007/s40319-021-01115-0

Hutter, M. (2010). *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*. Springer.

Hutukka, P. (2023). Copyright Law in the European Union, the United States and China. *International Review of Intellectual Property and Competition Law – IIC*, 54, 1044–1080. https://doi.org/10.1007/s40319-023-01357-0

Jaszi, P. (1992). On the Author Effect: Contemporary Copyright and Collective Creativity. *Cardozo Arts & Entertainment Law Journal*, *10*(2), 293–320.

Karppi, T., & Crawford, K. (2016). Social Media, Financial Algorithms and the Hack Crash. *Theory Culture & Society*, *33*(1), 73. https://doi.org/10.1177/0263276415583139

Khoury, A. H. (2017). Intellectual Property Rights For "Hubots": On The Legal Implications Of Human-Like Robots As Innovators And Creators. *Cardozo Arts & Entertainment Law Journal*, *35*(3), 635–668.

Krauss, M. (1989). Property, Monopoly, and Intellectual Rights Non-Posnerian Law and Economics Symposium. *Hamline Law Review*, *12*(2), 305.

Kur, A., Dreier, T., & Luginbuehl, S. (2013). *European intellectual property law: text, cases and materials* (2nd edn.). Edward Elgar Publishing, Cheltenham.

Lake, B. Ullman, T., Tenenbaum, J., & Gershman, S. (2017). Building Machines That Learn and Think Like People. *Behavioral and brain sciences*, *40*, 1–72. https://doi.org/10.1017/S0140525X16001837

Laton, D. (2016). Manhattan_Project.Exe: A Nuclear Option for the Digital Age. *Catholic University Journal of Law & Technology*, *25*(1), 94.

Manderieux, L. (2010). Secured Transactions as a Tool for Better Use of Intellectual Property Rights and of Intellectual Property Licensing (including Patent Licensing). *UNIDROIT Uniform Law Review*, *2010-1*, 447.

Martinez, R. (2019). Artificial Intelligence: Distinguishing Between Types & Definitions. *Nevada Law Journal*, *19*(3), 1015–1041.

McCarthy, J. (2008). The Well-Designed Child. *Artificial Intelligence*, *172*(18). https://doi.org/10.1016/j.artint.2008.10.001

Mizaras, V. (2012). Lithuania, In R. M. Hilty, & S, Ne´risson (Eds), *Balancing copyright – a survey of national approaches* (pp. 623–644). Springer, Berlin.

Omohundro, S. M. (2008). The Basic AI Drives. In Pei Wang et al. (Eds.), *Artificial General Intelligence 2008: Proceedings Of The First Agi Conference* (p. 483).

Pila, J., & Torremans, P. (2019). *European Intellectual Property Law*. Oxford University Press.

Rai, Arti Kaur. (1999). Regulating Scientific Research: Intellectual Property Rights and the Norms of Science. *Northwestern University Law Review*, *94*, 77. https://doi.org/10.2139/ssrn.172032

Ricketson, S., & Ginsburg, J. (2005). *International copyright and neighbouring rights: The Berne Convention and beyond* (2d ed.). New York: Oxford University Press.

Rosati, E. (2014). Copyright in the EU: in search of (in)flexibilities. *Journal of Intellectual Property Law & Practice*, *9*(7), 585–598. https://doi.org/10.1093/jiplp/jpu034

Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3d ed.). Pearson.

Salzberger, E. (2006). Economic Analysis of the Public Domain In L. Guibault, & P. B. Hugenholtz (Eds.), *The Future of the Public Domain* (Ch. III, pp. 27–59). Kluwer Law Intl.

Senftleben, M., & Buijtelaar, L. (2020). Robot creativity: an incentive-based neighboring rights approach. *European Intellectual Property Review*, *42*, 797–806. https://doi.org/10.2139/ssrn.3707741

Sganga, C., & Scalzini, S. (2017). From Abuse of Right to European Copyright Misuse: A New Doctrine for EU Copyright Law. *International Review of Intellectual Property and Competition Law (IIC)*, *48*(4), 405–435. https://doi.org/10.1007/s40319-017-0584-z

Sobel, B. L. W. (2017). Artificial Intelligence's Fair Use Crisis. *Columbia Journal of Law & The Arts*, *41*(1), 45–97. https://doi.org/10.7916/jla.v41i1.2036

Spector, L. (2006). Evolution of artificial intelligence. *Artificial Intelligence*, *170*(18), 1251–1253. https://doi.org/10.1016/j.artint.2006.10.009

Suchman, L., & Weber, J. (2016). Human-Machine Autonomies. In N. Bhuta, S. Beck, R. Geib, H. Yan Liu, & C. Kreb (Eds.), *Autonomous Weapon Systems: Law, Ethics, Policy* (pp. 39, 40). Cambridge: Cambridge University Press.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *49*(236), 433–460. https://doi.org/10.1093/mind/lix.236.433

Van Asselt, M. B. A., & Renn, O. (2011). Risk Governance. *Journal of Risk Research*, *14*(4), 431–449. https://doi.org/10.1080/13669877.2011.553730

van Eechoud, M. (2012). Along the road to uniformity: diverse readings of the Court of Justice Judgments on copyright works. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, *3*(1), 60–80.

van Gompel, S. (2014). Creativity, autonomy and personal touch. A critical appraisal of the CJEU's originality test for copyright. In M. van Eechoud (Ed.), *The work of authorship* (pp. 95–143). Amsterdam: University Press.

Walter, M., & von Lewinski, S. (2010). *European copyright law: a commentary*. New York: Oxford University Press.

Xalabarder, R. (2016). The Role of the CJEU in Harmonizing EU Copyright Law. *International Review of Intellectual Property and Competition Law – IIC*, *47*, 635–639. https://doi.org/10.1007/s40319-016-0509-2

Xu, M, David, J. M., & Kim, S. H. (2018). The Fourth Industrial Revolution: Opportunities and Challenges. *International Journal of Financial Research*, *9*(2), 90–95. https://doi.org/10.5430/ijfr.v9n2p90

Yanisky-Ravid, S., & Liu, X. (2018). When Artificial Intelligence Systems Produce Inventions: The 3A Era and an Alternative Model for Patent Law. *Cardozo Law Review*, *39*, 2215–2263.

Yong, Wan, & Hongxuyang, Lu. (2021). Copyright protection for AI-generated outputs: The experience from China. *Computer Law & Security Review*, *42*. https://doi.org/10.1016/j.clsr.2021.105581

Zurth, P. (2021). Artificial Creativity? A Case against Copyright Protection for AI-Generated Works. *UCLA Journal of Law & Technology*, *25*(2).

Author information

Themistoklis Tzimas – PhD (Public International Law and Political Science), Adjunct Assistant Professor, School of Law, Democritus University of Thrace
Address: PS 66100, University Campus, Komotini, Greece
E-mail: themis.tzimas@gmail.com
ORCID ID: https://orcid.org/0000-0002-0454-8220
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56615756300
Google Scholar ID: https://scholar.google.com/citations?user=XYVSDaIAAAAJ

Conflict of interest

The author declare no conflict of interest.

Financial disclosure

The research had no sponsorship.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – October 20, 2024
Date of approval – November 8, 2024
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:347.78:340.13:347.9:004.8
EDN: https://elibrary.ru/ppfjub
DOI: https://doi.org/10.21202/jdtl.2025.2

Эволюция авторского права в эпоху искусственного интеллекта: анализ правовых коллизий и судебных прецедентов

Фемистоклис Цимас

Университет Фракии имени Демокрита, Комотини, Греция

Ключевые слова

авторское право,
законодательство,
интеллектуальная собственность,
искусственный интеллект,
междисциплинарный подход,
право,
правовое регулирование,
суд,
технологический прогресс,
цифровые технологии

Аннотация

Цель: комплексный критический анализ современных проблем в области правового регулирования технологий искусственного интеллекта, возникающих на стыке норм интеллектуальной собственности и искусственного интеллекта. Особое внимание уделяется исследованию коллизий между существующим европейским законодательством об авторском праве и новыми технологическими реалиями.

Методы: в работе применяется междисциплинарный подход, включающий исторический, формально-юридический и сравнительно-правовой методы исследования. Исторический метод позволил проследить эволюцию законодательных и доктринальных подходов к регулированию интеллектуальной собственности в эпоху цифровизации. Формально-юридический метод дал возможность провести детальный анализ правовых норм различных государств. Сравнительно-правовой метод обеспечил возможность сопоставления различных подходов к регулированию отношений, связанных с использованием искусственного интеллекта в творческой деятельности.

Результаты: в ходе исследования детально рассмотрены вопросы авторского права на произведения, созданные с помощью искусственного интеллекта, включая сложные аспекты определения авторства и проблемы антропоцентризма в современном законодательстве. Проведен анализ судебных прецедентов, преимущественно в контексте законодательства Европейского союза, которое активно адаптируется к новым технологическим вызовам. Исследованы различные подходы к определению правового статуса произведений, созданных с помощью искусственного интеллекта, и их влияние на традиционные концепции интеллектуальной собственности.

Научная новизна: в статье впервые представлена комплексная оценка влияния творческих возможностей искусственного интеллекта на

фундаментальные концепции интеллектуальной собственности. Научная значимость заключается в оригинальной авторской оценке воздействия технологий искусственного интеллекта на законодательство об авторском праве, основанной на детальном анализе судебных прецедентов и доктринальных подходов. Исследованы перспективы развития правового регулирования в условиях технологического прогресса.

Практическая значимость: в работе предложены конкретные правовые и государственные решения, направленные на формирование сбалансированного и эффективного режима интеллектуальной собственности в эпоху искусственного интеллекта. Разработаны рекомендации по совершенствованию законодательства с учетом существующих судебных прецедентов и потребностей цифровой экономики. Результаты исследования могут быть использованы при разработке новых нормативных актов и совершенствовании существующей правовой базы в области регулирования искусственного интеллекта.

Для цитирования

Цимас, Ф. (2025). Эволюция авторского права в эпоху искусственного интеллекта: анализ правовых коллизий и судебных прецедентов. *Journal of Digital Technologies and Law*, *3*(1), 35–64. https://doi.org/10.21202/jdtl.2025.2

Список литературы

Adler, A. (2009.). Against moral rights. *California Law Review*, *97*, 263–301.

Bently, L., & Sherman, B. (2014). *Intellectual property law*. New York: Oxford University Press.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

Bostrom, N., & Ćirkovic, M. (2008). Introduction. In N. Bostrom, M., Ćirkovic, & M. Rees (Eds.), *Global Catastrophic Risks*. Oxford: Oxford University Press.

Butler, T. L. (1982). Can a computer be an author – copyright aspects of artificial intelligence. *Hastings Communications and Entertainment Law Journal*, *4*, 707.

Cabay, J., & Lambrecht, M. (2015). Remix prohibited: how rigid EU copyright laws inhibit creativity. *Journal of Intellectual Property Law & Practice*, *10*(5), 359–377. https://doi.org/10.1093/jiplp/jpv015

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, *538*(7623), 22. https://doi.org/10.1038/538020a

Cohen, Ju. E. (2006). Copyright, Commodification, and Culture: Locating the Public Domain. In L. Guibault, & P. B. Hugenholtz (Eds.), *The Future of the Public Domain* (pp. 121–166). Kluwer Law Intl.

Favale, M., Kretschmer, M., & Torremans, P. C. (2016). Is there an EU Copyright Jurisprudence? An Empirical Analysis of the Workings of the European Court of Justice. *The Modern Law Review*, *79*, 31–75. https://doi.org/10.1111/1468-2230.12166

Fisher, W. W. (2001). Theories of Intellectual Property. In S. Munzer (Ed.), *New Essays in the Legal and Political Theory of Property* (pp. 168–199). Cambridge University Press.

Gerdes, A. (2018). An Inclusive Ethical Design Perspective for a Flourishing Future with Artificial Intelligent Systems. *European Journal of Risk Regulation*, *9*(4), 677–689. https://doi.org/10.1017/err.2018.62

Gervais, D. J. (2019). The machine as author. *Iowa Law Review*, *105*, 2053–2106.

Ginsburg, J. C. (2018). People Not Machines: Authorship and What It Means in the Berne Convention. *International Review of Intellectual Property and Competition Law (IIC)*, *49*, 131. https://doi.org/10.1007/S40319-018-0670-X

Ginsburg, J. C., & Budiardjo, L. A. (2019). Authors and machines. *Berkeley Technology Law Journal*, *34*(2), 343. https://doi.org/10.2139/ssrn.3233885

Greenstein, S. (2022). Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*, *30*, 291–323. https://doi.org/10.1007/s10506-021-09294-4

Grinmelmann, J. (2016). There's No Such Thing as a Computer-Authored Work – and It's a Good Thing, Too. *Columbia Journal of Law & the Arts*, *39*, 403. https://doi.org/10.31228/osf.io/rk8cm

Hallevy, G. (2018). *Dangerous Robots – Artificial Intelligence vs. Human Intelligence*. http://dx.doi.org/10.2139/ssrn.3121905

Hansmann, H., & Santilli, M. (1997). Authors' and artists' moral rights: a comparative legal and economic analysis. *The Journal of Legal Studies*, *26*(1), 95. https://doi.org/10.1086/467990

Hashiguchi, M. (2017a). The global artificial intelligence revolution challenges patent eligibility laws. *Journal of Business & Technology Law*, *13*(1).

Hashiguchi, M. (2017b). Artificial intelligence and the jurisprudence of patent eligibility in the United States, Europe, and Japan. *Intellectual Property & Technology Law Journal*, *29*(12), 3–15.

Hattenbach, B., & Snyder, G. (2018). Rethinking the mental steps doctrine and other barriers to patentability of artificial intelligence. *Columbia Science and Technology Law Review*, *19*(2), 313–339.

Hemel, D. J., & Ouellette, L. L. (2013). Beyond the Patents–Prizes Debate. *Texas Law Review*, *92*(2), 303. https://doi.org/10.2139/ssrn.2245691

Holst, K. (2006.). A case of bad credit?: The United States and the protection of moral rights in intellectual property law. *Buffalo Intellectual Property Law Journal*, *3*(2), 105.

Hristov, K. (2017). Artificial Intelligence and the Copyright Dilemma. *IDEA*, *57*, 431.

Hugenholtz, P. B., & Quintais, J. P. (2021). Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output? (2021). *International Review of Intellectual Property and Competition Law – IIC*, 52, 1190–1216. https://doi.org/10.1007/s40319-021-01115-0

Hutter, M. (2010). *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*. Springer.

Hutukka, P. (2023). Copyright Law in the European Union, the United States and China. *International Review of Intellectual Property and Competition Law – IIC*, 54, 1044–1080. https://doi.org/10.1007/s40319-023-01357-0

Jaszi, P. (1992). On the Author Effect: Contemporary Copyright and Collective Creativity. *Cardozo Arts & Entertainment Law Journal*, *10*(2), 293–320.

Karppi, T., & Crawford, K. (2016). Social Media, Financial Algorithms and the Hack Crash. *Theory Culture & Society*, *33*(1), 73. https://doi.org/10.1177/0263276415583139

Khoury, A. H. (2017). Intellectual Property Rights For "Hubots": On The Legal Implications Of Human-Like Robots As Innovators And Creators. *Cardozo Arts & Entertainment Law Journal*, *35*(3), 635–668.

Krauss, M. (1989). Property, Monopoly, and Intellectual Rights Non-Posnerian Law and Economics Symposium. *Hamline Law Review*, *12*(2), 305.

Kur, A., Dreier, T., & Luginbuehl, S. (2013). *European intellectual property law: text, cases and materials* (2nd edn.). Edward Elgar Publishing, Cheltenham.

Lake, B. Ullman, T., Tenenbaum, J., & Gershman, S. (2017). Building Machines That Learn and Think Like People. *Behavioral and brain sciences*, *40*, 1–72. https://doi.org/10.1017/S0140525X16001837

Laton, D. (2016). Manhattan_Project.Exe: A Nuclear Option for the Digital Age. *Catholic University Journal of Law & Technology*, *25*(1), 94.

Manderieux, L. (2010). Secured Transactions as a Tool for Better Use of Intellectual Property Rights and of Intellectual Property Licensing (including Patent Licensing). *UNIDROIT Uniform Law Review*, *2010-1*, 447.

Martinez, R. (2019). Artificial Intelligence: Distinguishing Between Types & Definitions. *Nevada Law Journal*, *19*(3), 1015–1041.

McCarthy, J. (2008). The Well-Designed Child. *Artificial Intelligence*, *172*(18). https://doi.org/10.1016/j.artint.2008.10.001

Mizaras, V. (2012). Lithuania, In R. M. Hilty, & S, Ne´risson (Eds), *Balancing copyright – a survey of national approaches* (pp. 623–644). Springer, Berlin.

Omohundro, S. M. (2008). The Basic AI Drives. In Pei Wang et al. (Eds.), *Artificial General Intelligence 2008: Proceedings Of The First Agi Conference* (p. 483).

Pila, J., & Torremans, P. (2019). *European Intellectual Property Law*. Oxford University Press.

Rai, Arti Kaur. (1999). Regulating Scientific Research: Intellectual Property Rights and the Norms of Science. *Northwestern University Law Review*, *94*, 77. https://doi.org/10.2139/ssrn.172032

Ricketson, S., & Ginsburg, J. (2005). *International copyright and neighbouring rights: The Berne Convention and beyond* (2d ed.). New York: Oxford University Press.

Rosati, E. (2014). Copyright in the EU: in search of (in)flexibilities. *Journal of Intellectual Property Law & Practice*, *9*(7), 585–598. https://doi.org/10.1093/jiplp/jpu034

Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3d ed.). Pearson.

Salzberger, E. (2006). Economic Analysis of the Public Domain In L. Guibault, & P. B. Hugenholtz (Eds.), *The Future of the Public Domain* (Ch. III, pp. 27–59). Kluwer Law Intl.

Senftleben, M., & Buijtelaar, L. (2020). Robot creativity: an incentive-based neighboring rights approach. *European Intellectual Property Review*, *42*, 797–806. https://doi.org/10.2139/ssrn.3707741

Sganga, C., & Scalzini, S. (2017). From Abuse of Right to European Copyright Misuse: A New Doctrine for EU Copyright Law. *International Review of Intellectual Property and Competition Law (IIC)*, *48*(4), 405–435. https://doi.org/10.1007/s40319-017-0584-z

Sobel, B. L. W. (2017). Artificial Intelligence's Fair Use Crisis. *Columbia Journal of Law & The Arts*, *41*(1), 45–97. https://doi.org/10.7916/jla.v41i1.2036

Spector, L. (2006). Evolution of artificial intelligence. *Artificial Intelligence*, *170*(18), 1251–1253. https://doi.org/10.1016/j.artint.2006.10.009

Suchman, L., & Weber, J. (2016). Human-Machine Autonomies. In N. Bhuta, S. Beck, R. Geib, H. Yan Liu, & C. Kreb (Eds.), *Autonomous Weapon Systems: Law, Ethics, Policy* (pp. 39, 40). Cambridge: Cambridge University Press.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *49*(236), 433–460. https://doi.org/10.1093/mind/lix.236.433

Van Asselt, M. B. A., & Renn, O. (2011). Risk Governance. *Journal of Risk Research*, *14*(4), 431–449. https://doi.org/10.1080/13669877.2011.553730

van Eechoud, M. (2012). Along the road to uniformity: diverse readings of the Court of Justice Judgments on copyright works. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, *3*(1), 60–80.

van Gompel, S. (2014). Creativity, autonomy and personal touch. A critical appraisal of the CJEU's originality test for copyright. In M. van Eechoud (Ed.), *The work of authorship* (pp. 95–143). Amsterdam: University Press.

Walter, M., & von Lewinski, S. (2010). *European copyright law: a commentary*. New York: Oxford University Press.

Xalabarder, R. (2016). The Role of the CJEU in Harmonizing EU Copyright Law. *International Review of Intellectual Property and Competition Law – IIC*, *47*, 635–639. https://doi.org/10.1007/s40319-016-0509-2

Xu, M, David, J. M., & Kim, S. H. (2018). The Fourth Industrial Revolution: Opportunities and Challenges. *International Journal of Financial Research*, *9*(2), 90–95. https://doi.org/10.5430/ijfr.v9n2p90

Yanisky-Ravid, S., & Liu, X. (2018). When Artificial Intelligence Systems Produce Inventions: The 3A Era and an Alternative Model for Patent Law. *Cardozo Law Review*, *39*, 2215–2263.

Yong, Wan, & Hongxuyang, Lu. (2021). Copyright protection for AI-generated outputs: The experience from China. *Computer Law & Security Review*, *42*. https://doi.org/10.1016/j.clsr.2021.105581

Zurth, P. (2021). Artificial Creativity? A Case against Copyright Protection for AI-Generated Works. *UCLA Journal of Law & Technology*, *25*(2).

Сведения об авторе

Фемистоклис Цимас – PhD в области государственного международного права и политологии, ассистент преподавателя, Школа права, Университет Фракии имени Демокрита
Адрес: Греция, PS 66100, г. Комотини, Университетский кампус
E-mail: themis.tzimas@gmail.com
ORCID ID: https://orcid.org/0000-0002-0454-8220
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56615756300
Google Scholar ID: https://scholar.google.com/citations?user=XYVSDaIAAAAJ

Конфликт интересов

Автор сообщает об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.41.51 / Охрана авторских прав
Специальность ВАК: 5.1.3 / Частно-правовые (цивилистические) науки

История статьи

Дата поступления – 20 октября 2024 г.
Дата одобрения после рецензирования – 8 ноября 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:343.3/.7:004.056:004.9
EDN: https://elibrary.ru/rtolza
DOI: https://doi.org/10.21202/jdtl.2025.3

Crimes in the Age of Artificial Intelligence: a Hybrid Approach to Liability and Security in the Digital Era

Neelkanth Bhatt

Lukhdhirji Engineering College, Morbi, India

Keywords

artificial intelligence,
crime,
criminal legislation,
criminal liability,
criminality,
digital technologies,
law,
PESTEL technique,
product quality,
security

Abstract

Objective: to study the applicability of existing norms on product quality liability and negligence laws to crimes related to artificial intelligence. The author hypothesizes that the hybrid application of these legal mechanisms can become the basis for an effective regulatory system under the rapid technological development.

Methods: the research includes a comprehensive approach based on the PESTEL analysis (political, economic, social, technological, environmental and legal factors), the "five whys" root cause analysis, and cases from various countries. This multi-level approach allows not only identifying key problems, but also proposing adapted solutions that take into account the specifics of crimes related to artificial intelligence.

Results: the research shows that the existing norms on product quality and negligence are not effective enough to regulate crimes related to artificial intelligence. The main obstacles are technological complexity, lack of precedents, lack of consumer awareness, and jurisdictional issues. The author concludes that effective regulation requires a global system that includes clear principles of responsibility, strict safety standards, and constant adaptation to new challenges.

Scientific novelty: the paper represents a unique approach to the crimes related to artificial intelligence through the prism of hybrid application of existing legal mechanisms. It offers a new perspective on the problem, combining theoretical analysis with practical recommendations based on case study.

Practical significance: recommendations for legislators and regulators were developed. The author emphasizes the need to create specialized agencies, introduce educational programs for citizens and employees, and to provide funding for research in the field of explicable artificial intelligence and security standards. These measures are aimed at forming a stable regulatory system capable of effectively countering crimes related to the use of artificial intelligence. The work opens up new horizons for further research on the regulation of AI technologies and emphasizes the need for international cooperation and an interdisciplinary approach.

For citation

Bhatt, N. (2025). Crimes in the Age of Artificial Intelligence: a Hybrid Approach to Liability and Security in the Digital Era. *Journal of Digital Technologies and Law, 3*(1), 65–88. https://doi.org/10.21202/jdtl.2025.3

Contents

Introduction
1. Review of Literature
2. Methodology
3. Results and Discussion
   3.1. PESTEL Analysis
   3.2. Root Cause Analysis
   3.3. Case Studies: How do countries handle AI crimes?
      3.3.1. USA's Greyall Episode
      3.3.2. UK's British Airways Data Breach Incident
      3.3.3. India's Aadhaar Data Leak Case
      3.3.4. China's Tencent's Deepfakes Episode
      3.3.5. Russia's Sovereign Strategy for AI
   3.4. Key Observations and Insights
Conclusions
References

Introduction

Artificial Intelligence (AI) assisted crimes, or AI crimes (AICs) is a new but well-known term applied to refer to crimes that utilize AI for criminal activities. This comprises AI as a tool for criminals, generating forged deepfakes for fraud or social engineering for deception and manipulation[1]. AI technologies may also be used by criminals intending

[1] Center for AI Crime. (2023). About AI Crimes. https://clck.ru/3Gpr34

to bypass security systems or to manipulate their decision-making (King et al, 2021). AI's emergent erudition allows unprecedented and potentially more widespread crimes while making it intricate to establish necessary safeguards[2].

As AI is gaining potency, its potential for criminal misuse also gets bigger. This would lead to spanking new types of criminal activities that increase with leaps and bounces, and more victims. Are we equipped to deal with this? A rushed regulation now could result in an outdated regulation in the coming days.

The applicability of existing laws to novel AIC is a complex subject. Whilst some argue that the existing legal frameworks are versatile and can adjust to these novel offenses for instance, AI-generated deepfakes used for financial fraud can be dealt with fraud laws (criminal laws). However, others draw attention to its limitations in dealing with such offences. Sukhodolov et al. (2020) believes that the existing laws may not adequately address AIC's aspects like criminal intent, which is customarily applied to human beings. The obscure nature of AI algorithms seldom allows us to associate responsibility (Sukhodolov et al., 2020).

Lack of clarity on liability, responsibility gap, inadequate legal regime, difficulty in determining fault, and jurisdictional challenges may be identified as some of the important reasons why the existing criminal laws would stand deficient in handling AICs. Further, the application of criminal laws for crimes associated with AI technologies or products necessitates the establishment of elements of intent and attribution. It needs to be proved beyond doubt that the AIC was committed with 'mens rea'- a guilty mind and that the AIC ought to be associated with the subject committing it. But, what if they unintentionally harm humans? It is also quite difficult to attribute the offence to the programmer, manufacturer, or user of AI. This confusion would only get in the way of prosecution and effective deterrence. Efforts are therefore required to enact a more robust legal framework for this ever-evolving AI technology.

Contrary to criminal laws, civil laws primarily focus on duty and foreseeability which can be more readily applied to AICs. The onus here is on granting compensation to victims and not imprisonment or harsh punishment to criminals. Considering the nature of AICs, if this is allowed, the victims shall receive compensation but these won't necessarily deter future crimes. However, AI harms not always rise to the level of heinous crimes. In principle, existing traditional tort laws such as the 'Product Liability' Law and 'Negligence' Law are quite capable of handling the AICs as these laws

[2] Markoff, J. (2016, October 23). As Artificial Intelligence Evolves, So Does Its Criminal Potential. The New York Times. https://clck.ru/3Gpr5Z

aim to balance public safety, responsibility, and growth. Yet, to what degree is a matter of investigation?

In case of defective products, the customer's rights are protected by 'Product Liability' laws as manufacturers, distributors, and sellers are held responsible for such acts. Likewise, if an AI system malfunctions or causes damages or harm the developer or manufacturer may be held accountable for such defects (Scherer, 2015). On the other hand, 'Negligence' laws necessitate individuals to take due care in exercising actions to prevent harm. It is possible to apply these laws to AICs when individuals fail to prevent the misuse of AI systems for criminal activities (Zhao, 2024). Independent decision-making, learning capabilities, and the lack of human involvement in criminal acts are some of the unique challenges that are posed by AI systems. Therefore, a specialized legal framework to counter AICs establishing clear guidelines for the determination of liability, responsibility and jurisdiction is the need of the hour.

Ideally, for a win-win situation for all AI stakeholders, the AI regulating framework shall be based on 'PEEC' doctrine, i.e. on considerations of 'public interest' and 'principles of environmental sustainability', 'economic development' and 'criminal law' (Neelkanth Bhatt & Jaikishen Bhatt, 2023).

The present study hypothesizes that "a hybrid application of existing 'Product Liability' Law (PLL) and 'Negligence' Law (NL) offers a robust legal framework for artificial intelligence crimes". The study aims to validate this idea through systematic investigation and undertake a case-study-based protocol to analyze the effectiveness of PLL and NL approaches for AICs. Through this, the present study hopes to lead into the discussion on how AI can best be regulated for a flourishing society.

1. Review of Literature

The pertinent prerequisites of existing criminal law make it difficult to regulate AICs (Qatawneh et al., 2023; Abbot & Sarch, 2019; Shestak et al., 2019). Application of traditional principles of 'Mens Rea' & 'Actus Reus' is difficult in cases of AICs (Abbot & Sarch, 2019; Shestak et al., 2019). World over, there is a strong consensus for the creation of legislative reforms for AICs. This includes ideas to cover AICs through criminal laws (Qatawneh et al., 2023; Neelkanth Bhatt & Jaikishen Bhatt, 2023; Shestak et al., 2019; Khisamova & Begishev, 2019). Few suggest modest changes to existing laws, whereas others are of the view of changing them drastically (Abbot & Sarch, 2019; Khisamova & Begishev, 2019). To mitigate associated risks, there is an urgent requirement for standardization and certification in the design, development and deployment of AI

technologies (Khisamova & Begishev, 2019; Broadhurst et al., 2019). There is also significant concern regarding the potential for AI to infringe on fundamental rights and perpetuate biases; AI can play the dual role of crime enabler and preventer (Broadhurst et al., 2019; Ivan & Manea, 2022). Shestak et al. (2019) & Khan et al. (2021) discuss various models of AI liability under certain conditions, though the independent actions of AI forms often make the applicability of laws quite complex. Though AI systems have the potential to enable crimes, considering their future, a lot of uncertainty can be associated with them (King et al., 2021). The existing legal framework is scanty in the determination of culpability in crimes involving the use of AI as a tool (Dremliuga & Prisckina, 2020). AI technologies can perhaps meet the criteria for criminal liability, still, additional regulatory efforts are essential to address these challenges (Lagioia & Sartor, 2019).

A focused legal framework for addressing AICs must consider several vital basics, including the establishment of unambiguous guidelines for liability and responsibility in cases involving AI, the implementation of vigorous standards for the development and deployment of AI systems, and formation of regulatory bodies to oversee and enforce these principles. In addition, such a framework ought to incorporate instruments for continuous monitoring and its updations to match the rapid technological advancements, coping with international cooperation to address the global nature of AICs (Binns, 2018; Calo, 2019; Gless, 2019). The framework is supposed to prioritize protective actions, including mandatory safety audits and ethical impact assessments for the development of AI (Jobin et al., 2019). It has to adapt to the evolving nature of AI systems, with channels for ongoing review and revision.

Leveraging an amalgamated tool that blends product liability principles with negligence law has the potential to address the crucial nitty-gritty of a dedicated AI crime framework. Product liability laws, having focus on design defects and standards of safety (Solum, 2020), could assign responsibility to developers for inbuilt flaws in AI systems. Negligence law, underlining the duty of care (Kingston, 2016), could hold humans accountable for projected risks taking place due to improper deployment or use of AI systems. This hybrid approach could offer a comprehensive means for assigning culpability and promoting preventative actions in the development & use of AI systems.

At present, there is a paucity of established legal frameworks exclusively addressing AICs across both developed and developing nations. The European Union (EU) by way of General Data Protection Regulation (GDPR) and AI Act Proposal is taking significant

strides in addressing the challenges posed by AI technologies[3], [4]. The United States lacks overarching regulation but various agencies have guidelines[5]. Singapore has proposed a model AI governance framework[6]. In 2021, China imposed several sector-specific regulations in the form of Guiding Opinions on Regulating Scientific and Technological Activities in the Field of Artificial Intelligence (Li, 2023).

The Russian approach focuses on development and support and not on stricter regulations. Russia proposes a road map of the development of breakthrough technologies 'Neurotechnology and AI' through 'National Strategy for the Development of Artificial Intelligence for the period until 2030'[7]. India too, does not have overarching regulation instead there are reports by four Committees on various aspects of AI outlining recommendations for the ethical development of AI systems[8]. There are no comprehensive legal regulations on AI technology in Japan, though the Protection of Personal Information Act of 2020, deals with some aspects relating to AI systems[9]. South Korea too has a Framework for Ethical Development and Use of Artificial Intelligence (2020) which are rather non-binding guidelines[10].

The EU's proposal of the AI act incorporates adherence to strict safety standards and clear lines of liability in case of harm emphasizing risk management of AI systems. The USA has numerous guidelines to deal with AICs all focusing on fairness, accountability, and reduction of harm. The principles of current set of standards/guidelines in both EU & USA aligns with 'Product Liability Laws' and 'Negligence Laws'.

2. Methodology

The present study employs a rational, logical, comprehensive, and multi-layered approach to scientifically examine the hypothesis. To gain valuable insight and a proper understanding of the external factors influencing the hypothesis 'PESTEL' (Political,

3 European Commission. https://goo.su/y3Zuwv

4 Iapp.org (International Association of Privacy Professionals). Global AI Law and Policy Tracker. https://clck.ru/3Gpsti

5 National Institute of Standards and Technology (NIST). https://clck.ru/3GpszK

6 Singapore's AI Governance webpage. https://goo.su/uBEdf

7 Russia: Current status and development of AI regulations. (2024, May 24). Data Guidance. https://clck.ru/3GptAx

8 Government of India. (2018). Reports of various Committees on Artificial Intelligence. https://goo.su/H9dNS

9 Personal Information Protection Commission, Japan. https://clck.ru/3GptNq

10 Ministry of Science and ICT (MSIT), South Korea (2020). Framework for Ethical Development and Use of Artificial Intelligence. https://clck.ru/3GptTi

Economic, Social, Technological, Environmental, and Legal) analysis has been carried out for the identification of factors that are beyond immediate control but could significantly impact the hypothesis. Rather than addressing the symptoms, this study attempts to address the real issue by delving deeper into systematic identification of the root cause of the problem by performing the 'Five Whys' analysis.

Further, to evaluate perceptions and to broaden our perspective the study utilized various case studies for comparing existing 'Product Liability Laws' and 'Negligence Laws' across various countries which was followed by examining successful solutions that were effectively implemented in other fields too. This integrated analysis allowed to propose solutions tailored to address the identified root cause. It also facilitated the adaption of these solutions to their potential effects.

This multi-faceted robust methodology not only examined a wider context but also endorsed a thorough investigation of the proposed hypothesis that just went beyond superficial analysis and offered a versatile understanding of the issue and its potential solutions that are required for proposing a robust legal framework for AI. The methodology's strength lies in this holistic approach which allows theoretically sound and practically viable solutions.

3. Results and Discussion

3.1. PESTEL Analysis

Purely quantitative methods in strategic planning very seldom offer the distinct advantage required to test a hypothesis. Qualitative methods perform excellently well when it is desired to measure internal performance or capture market trends, but it has very limited potential to capture the broader environment. Conversely, 'PESTEL' allows systematic examination of Political, Economic, Social, Technological, Environmental, and Legal factors with a holistic view of external factors contributing to a company's success (Yüksel, 2012). This approach allows to counter the dynamism of AI systems and enables a comprehensive understanding of associated potential threats and opportunities of the hypothesis.

Detailed Comparison of Specific Sections/Articles and Penalties as covered by PLL and NL of various countries is presented in Fig. 1. This comparison lays the foundation to perform further analysis. A Comprehensive 'PESTEL' analysis of 'Product Liability Laws' & 'Negligence Laws' across various countries is presented at Table 1.

Fig. 1. Comparison of Existing Laws across Countries

Table 1. PESTEL Analysis on Existing Product Liability & Negligence Laws of Various Countries

Factors	USA	UK	India	China	Russia
Political	1. Pro-Consumer Political Change	1. Stable political base, strong support for consumer rights	1. Growing focus on consumer protection & rights	1. Centralized political control enables swift changes in regulations	1. Strong political will for consumer protection, at times inconsistent implementation
	2. Tug-of-war among Consumers, Legal System & Industries	2. Changes in Regulations due to Brexit	2. Bureaucratic hurdles to strong implementation	2. Strong government will for technological advancement with consumer protection	2. Government control over systems
Economic	1. High litigation costs	1. Regulatory compliance burden	1. Economic burden of compensation and fines	1. Economic penalties adversely impact business	1. Significant economic fines and sanctions for damages
	2. High economic incentives for compliance	2. Adverse economic impact on business due to recalls and compensations	2. Compared to the Western world lower cost of litigation	2. High cost of compliance with stringent product safety laws	2. Suspension of business activities due to non-compliance

End of Table1

Social	1. Higher consumer awareness & activism	1. Strong movements for consumer rights	1. Media and government efforts are pivotal in creating consumer awareness	1. Growing product demands with a high awareness level	1. Consumer awareness and activism are growing
	2. Class action lawsuits are a powerful tool for protecting consumers' rights	2. High public awareness of safety and product issues	2. Societal push for stringent regulations	2. Influence of social media on public opinion and regulatory norms	2. Growing public demand for harsher regulations and effective implementation
Technological	1. Advancement in technology influences product design and safety	1. High technological innovations impacting the safety of products	1. Technological advancements for product safety	1. Rapid growth in AI and consumer electronics	1. Technological advancements for product manufacturing and safety
	2. Increasing the use of AI for compliance monitoring and defect detection	2. Adoption of AI and IoT for regulatory compliance	2. Growing use of AI for regulation	2. Integration of technology for regulative measures	2. Novel technological adoptions for regulations
Environmental	1. Environmental considerations for product liability	1. Strong environmental regulations affecting product standards	1. Increasing regulations for environmental protection for various products	1. Stringent environmental laws to regulate products	1. Eco-compliance for product liability
	2. Emphasis on eco-friendly and sustainable products	2. Focus on environmental sustainability	2. Efforts to reduce harmful environmental effects of products	2. Government emphasis on green and sustainable products	2. Emphasis on compliance with environmental standards for products
Legal	1. Comprehensive framework to deal with legal aspects of products	1. Strict liability by way of Consumer Protection Act 1987	1. Consumer Protection Act, 2019 with ample provisions on product liability	1. Strict liability provisions in Product Quality Law and Tort Liability Law	1. Strict provision for liability and negligence in Civil Code and Consumer Protection Law
	2. Strict liability and well-defined compensatory and punitive norms	2. Strong compensatory damages and recall orders	2. Fines, compensation, and imprisonment for violations	2. Compensatory damages, administrative fines and recalls	2. Compensatory and moral damages and suspension of business as penalty

The PESTEL analysis demonstrates an intricate global setting for product liability, especially for AI. The USA, UK, China, and Russia claim wide-ranging legal frameworks covering products, and enforcement challenges transpire in India. Strong political support for consumer protection can be observed in the US and UK. Economically, US companies are burdened with high litigation costs, while India faces a compliance burden. Technological advancements in the UK, USA, and China assist compliance, nonetheless, enforcement gaps are quite evident in some regions. The USA and UK set a high bar

through stringent environmental regulations, yet enforcement varies worldwide. The emergent universal focus on consumer safety offers an opportunity for coherent legal standards, but differing versions and actions pose a risk.

The analysis also revealed unambiguous opportunities for global regulations of product safety especially for AI systems that are driven by rising consumer awareness and concerns with technological advancements. The PESTEL analysis was conducted for validation of the hypothesis. It can be inferred that even a hybrid application of PLL and NL would require fine-tuning to these prevalent regulations to address AICs. The existing laws have been designed for physical products that may not holistically cover artificially intelligent systems. Enforcement challenges and the rapid pace of development in the field of AI would hinder the effectiveness of ordinary regulations. In these contexts, the new holistic framework that lays AI-specific liability regimes with a focus on transparency and enhanced revelation of how the AI system works would facilitate harmonizing global standards to deal with AICs for ensuring its consistent enforcement across countries.

3.2. Root Cause Analysis

Root Cause Analysis (RCA) is a decisive tool for testing a particular hypothesis, mostly when dealing with multifaceted phenomena (Barsalou, 2014). This analysis allows us to systematically investigate the cause-and-effect relationship for any observation. It helps identification of flaws in the hypothesis and allows adjustments required to ensure the accuracy of the adopted research design (Barsalou, 2014). This convergent process reinforces the overall investigation and leads to more substantial inferences.

Fig. 2 shows the Ishikawa diagram (cause and effect diagram) for the ineffectiveness of existing PLL & NL in dealing with AICs.

The diagram noticeably demonstrates that the rapid evolution of AI technologies, its mismatch with existing regulations, lack of transparency and accountability, and lack of globally acceptable enforcement mechanism renders the existing framework of PLL and NL deficient in handling AICs.

The ‘Five Whys’ technique is a powerful tool requiring minimal resources or training for uncovering the root cause of problems across various disciplines (Barsalou & Starzynska, 2023). The technique involves asking «why» five times in succession which allows a structured and logical tool for identification of vital factors that contribute to the issue (Pugna et al., 2016). Repeated questioning peels off any superficial causes leading to a deeper root cause responsible for the problem on hand.

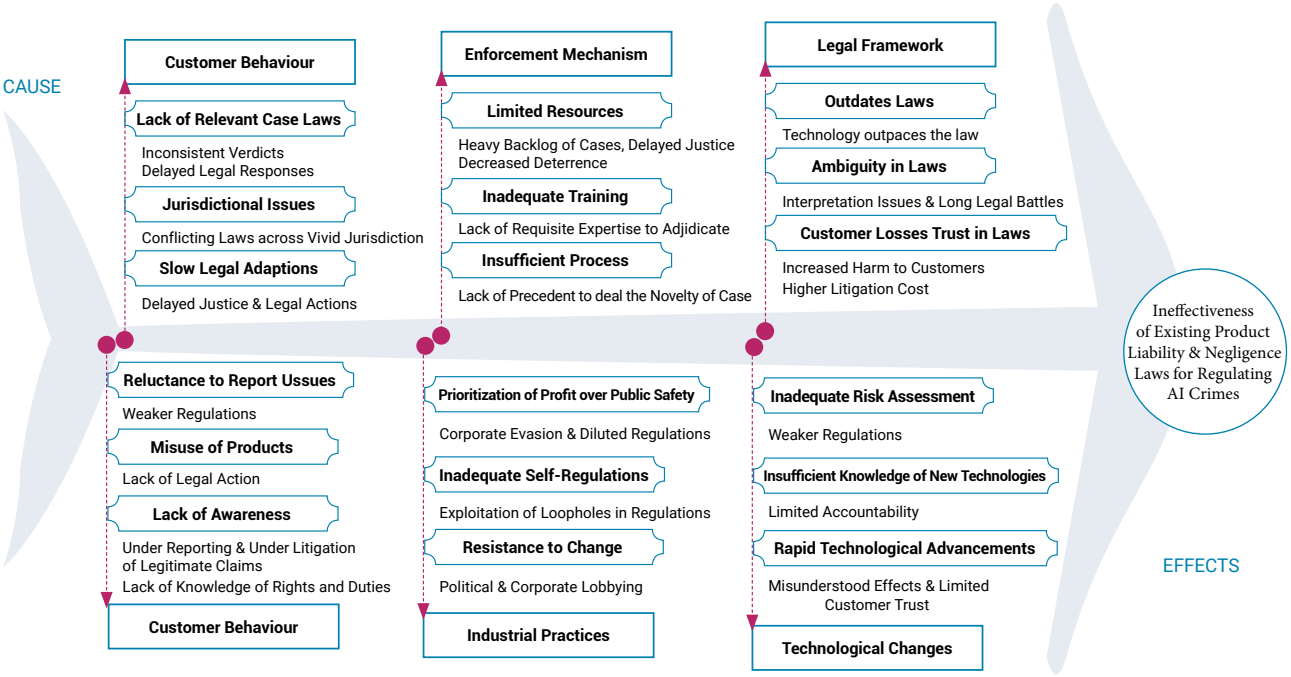


Fig. 2. Cause & Effect Relationship of Existing PLL & NL with AICs

Fig. 3 shows a systematic 'Five Why' analysis of the ineffectiveness of PLL and Fig. 4 shows a systematic 'Five Why' analysis of the ineffectiveness of NL in dealing with AICs.

The 'Five Why' analysis on PLL and NL finds these existing laws ineffective in dealing with unique and unexpected AICs and thus disapproves the hypothesis. The rapid speed of technological advancements and global market dynamics surpass the ability of existing legal frameworks to acclimate, leading to dearth of established standards, deficient regulation, unsatisfactory judicial expertise, limited consumer awareness, commercial exploitation of legal loopholes, jurisdictional challenges, huge judicial backlogs, evolving risk perceptions, underfunded judicial and regulatory bodies and economic pressures that prioritize business interests over consumer protection are the chief issues that needs prompt remedying and adjustments for leveraging the existing PLL & NL for AICs.

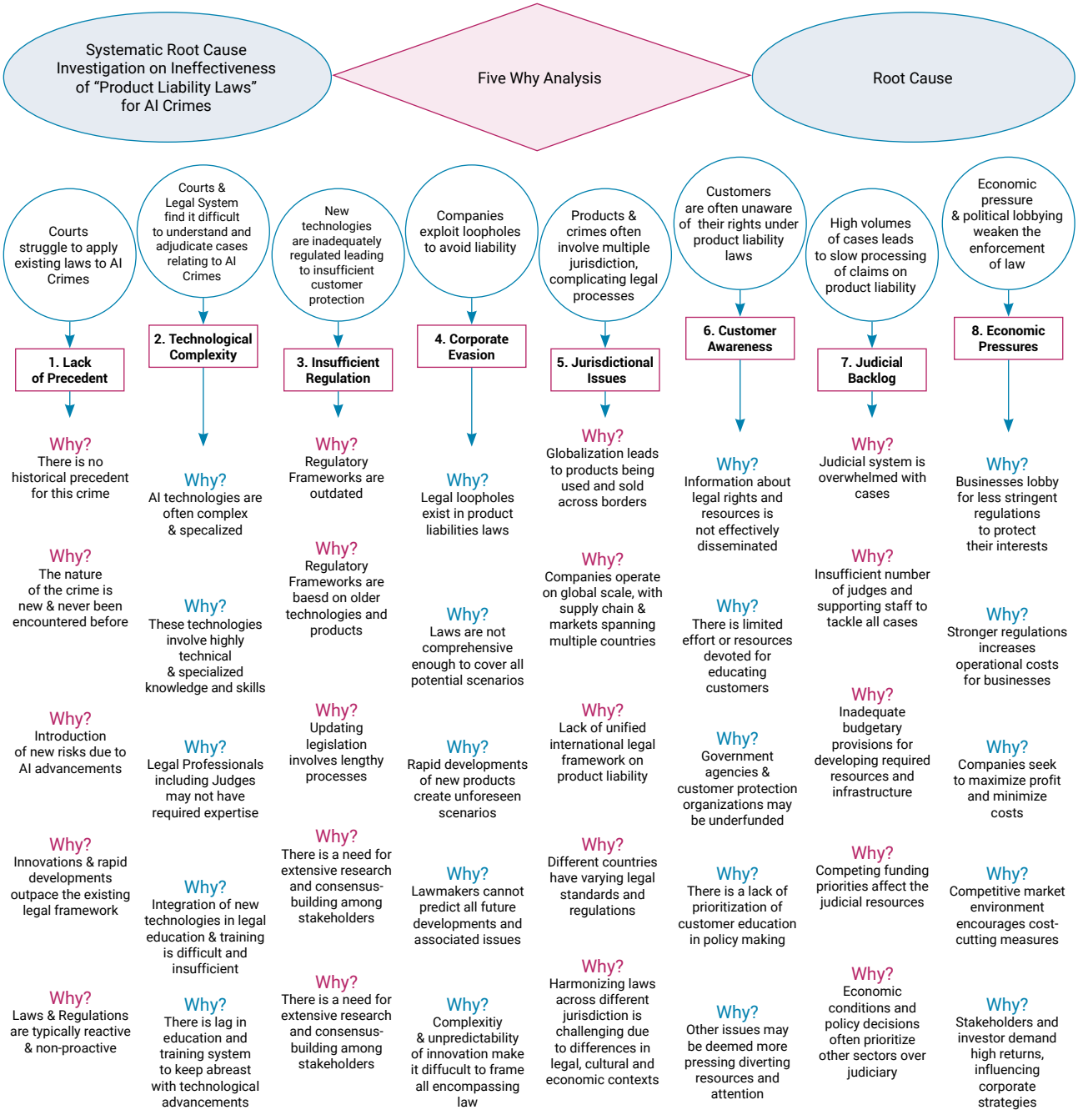


Fig. 3. Root Cause Analysis of Ineffectiveness of PLL in Dealing AICs

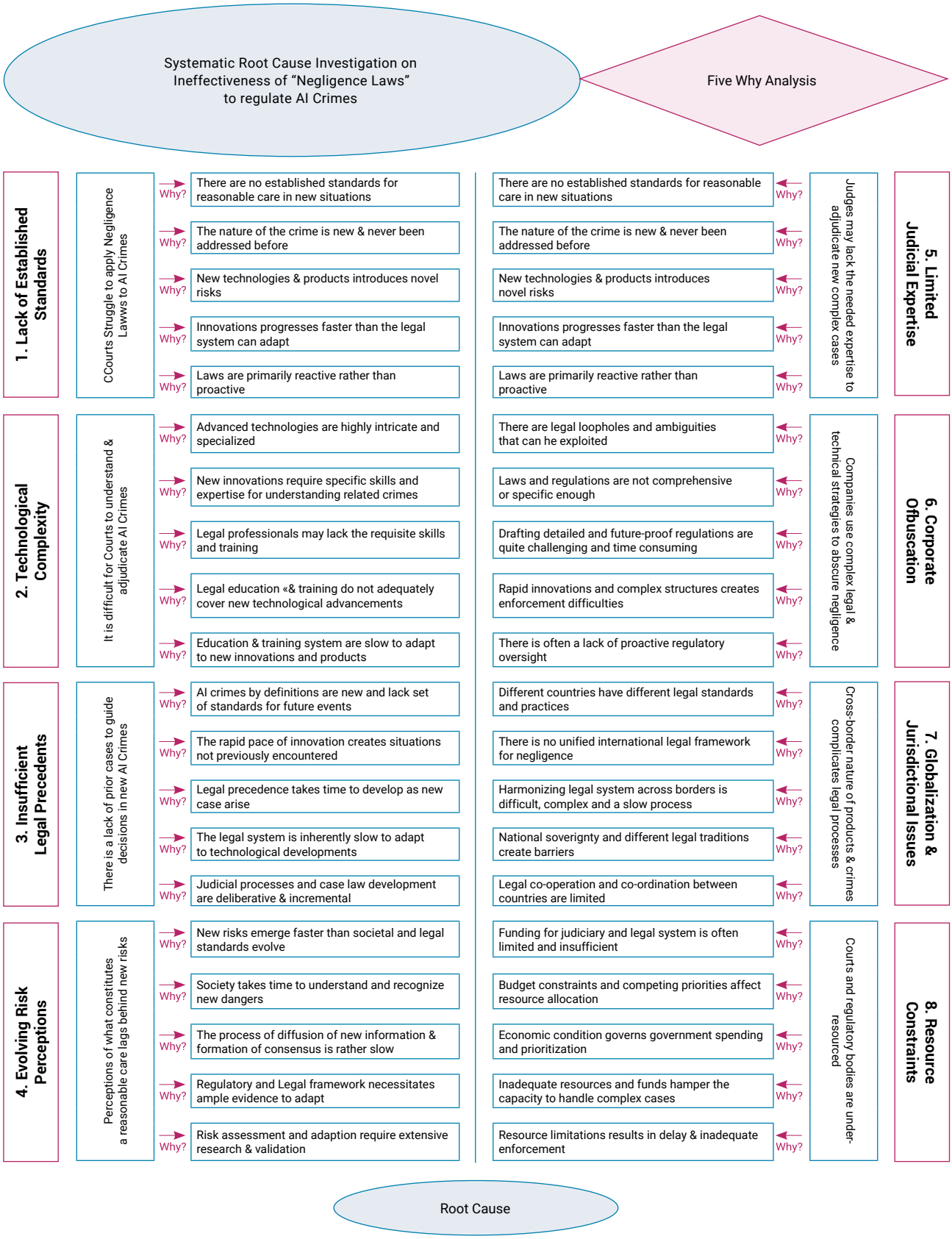


Fig. 4. Root Cause Analysis of NL in Dealing with AICs

3.3. Case Studies: How do countries handle AI crimes?

3.3.1. USA's Greyball Episode

This is a case where a company 'Uber' employed "Greyball" – an AI-driven tool to evade law enforcement in cities where their services were not allowed[11]. The tool identified, targeted, and served law enforcement officials with a fake version of the app to dodge detection. The US Department of Justice and numerous other local authorities initiated the investigation against 'Uber' for this deemed intentional misconduct. Uber agreed to cease using the tool and had to bear reputational damages and increased regulatory scrutiny. This is a classic case of regulatory authorities dealing with intentional misconduct of AI systems and imposition of penalties that would serve as deterrence to such future acts.

3.3.2. UK's British Airways Data Breach Incident

During 2018, the 'British Airways' website utilizing AI systems suffered a data breach where over 400,000 customers' personal and financial details were compromised[12]. The company was found negligent in protecting customers' data by the Information Commissioner's Office (ICO). The company's cooperation in the matter meant that they were penalized only £20 million against the original proposed fine of £183 million. This is a typical case where the authorities focused on laying reasonable and proportionate fines as a means to promote self-regulating measures without fear of harsh punishment.

3.3.3. India's Aadhaar Data Leak Case

Security lapses and negligence in the management of AI-driven databases resulted in the leakage of millions of citizens' personal information by the 'Aadhaar'- 'India's Biometric Identity System'[13]. Severe criticism and legal challenges were faced by the Unique Identification Authority of India (UIDAI) due to the said incident. Stricter compliance norms and enhanced security features were introduced post this incident. However, the incident failed to attract any financial penalties due to existing legal frameworks. This is a perfect demonstration accentuating the need to have a robust framework to effectively deal with unintentional harm caused by AI systems.

---

11 Greyball: how Uber used secret software to dodge the law. (2017, March 4). The Guardian. https://clck.ru/3Gq9ZC

12 BA fined record £20m for customer data breach. (2020, October 16). The Guardian. https://clck.ru/3Gq9d8

13 Aadhaar data leak exposes cyber security flaws. (2023, March 29). The Hindu Business Line. https://clck.ru/3Gq9hv

3.3.4. China's Tencent's Deepfakes Episode

During 2019, 'Tencent' had to deal with a huge controversy over its deepfakes generating AI tool. It was perceived that the tool had been misused for fraud and misinformation. This resulted in swift regulatory actions and China's new regulations on AI & Deepfake Technology were implemented. The new regulations require clear labeling and restricting the ill use of these technologies. The company complied by adjustment to their tool's functionality. This is a unique case of proactive legislation which is intended to keep content regulation and censorship efforts a step ahead of emerging new AI technologies and stricter enforcement for the prevention of intentional misconduct and misuse of AI technologies[14].

3.3.5. Russia's Sovereign Strategy for AI

Russia has announced plans to avoid Western dominance over AI technologies[15]. The dominance of certain countries in the development of AI would potentially reflect region-specific biases and this could render digital discrimination and negatively affect the sovereignty of a country. The AI development strategy adopted by Russia is unique as it aims to preserve national identity and cultural heritage in the development of AI technologies. In Russia, contrary to US & UK the development is not led by the government or the private sector but by the state-owned firms (Petrella et al., 2021). Russia through 'Digital Sandboxes' has introduced a novel experimental legal regime for AI development where companies are allowed to work on AI systems that are not currently regulated by existing legislations; facilitating opportunities for these companies to see how developed AI performs in real-life situations in Moscow and subsequently throughout Russia[16].

3.4. Key Observations and Insights

The 'PESTEL Analysis', 'Root Cause Analysis', and the 'Case Studies' have revealed that the advocated hypothesis does not hold good in cases of complex crimes related to AI technologies. Given the complexities of AICs, a hybrid approach leveraging the existing framework to deal with AICs is extremely challenging. To deal with AICs we ought to establish a robust international framework to accommodate several contentious

---

14 Kharpal, A. (2022, Dec 22). China is about to get tougher on deepfakes in an unprecedented way. Here's what the rules mean. CNBC. https://clck.ru/3GqAah

15 Putin to boost AI in Russia to Fight 'Unacceptable and Dangerous' Western Monopoly. (2023, November 24). VAO. https://clck.ru/3GqAsU

16 Mondaq. https://clck.ru/3GqAyM

issues. Firstly, the framework must clearly define all current and potential AICs. Secondly, it has to have a wide-ranging proposed set of actions, and procedures for prosecutorial authorities, and thirdly, harsh penalties for criminal conduct capable of accommodating rapid technological developmental pace and global market dynamics with incorporation of traditional elements of 'mens rea' and 'actus reus'. The framework must promote enforcement, and compliance while being fair to defendants, and customer awareness. The framework must encourage harmony between national and international bodies and enhance jurisdictional effectiveness for upholding transboundary stakeholders' justice, accountability, and rights.

A few more thought-provoking ideas drawn from vivid sectors need to be investigated to further understand the level of complexity posed by evolving AI systems. Let us first consider AI systems analogous to a 'gun' where only human actors are held solely responsible for its use. This idea cannot withstand the legal test due to unforeseen consequences of highly evolving AI. Assigning 'strict liability' to developers and 'personhood' to certain AI is yet another potential approach to regulating AI. However, certain AIs are built to evolve and make their own decisions which makes it extremely difficult to regulate AI even through this idea. Yet another idea is to consider the unforeseen and unintended act of AI analogous to an 'act of god', but this idea also lacks the test of intention which is not present in cases of 'act of god'. However, this idea rendered important lessons like taking proactive measures similar to safety checks and developing ethical guidelines for AI. In another consideration, regulating AI in a manner by which authorities deal with 'infectious diseases', it is possible to link similarities in risk management and public education. However, the concept lacks intentionality and the pace of change usually associated with AI. To close the considerations of ideas, we can regulate AI in a manner analogous to 'nuclear weapon' regulations, stressing international cooperation and having adequate safety norms while recognizing the distinctive challenges posed by the accessibility and rapid evolution of AI systems.

The foregoing discussion suggests that we ought to focus on explainable AI, robust safety standards, gradual advancement with oversight, and adapting legal framework. This would help to ensure that human actors are held responsible throughout the development and deployment stages of AI. The goal should be to create a responsible AI system where responsibilities are clear and adequate proactive measures are taken to minimize the risk of unforeseen harm and to ensure that AI remains a tool for good.

Establishing such a framework is a time-consuming task that becomes even more difficult as a strong transboundary consensus has to be built for its effective enforcement to cover extra-jurisdictional crimes committed through AI systems. Till then, all countries

allowing the use of AI technologies have to adapt their existing legal framework to address AICs. Such adaption measures shall essentially be in the form of:

1. Modernizing of definition of crime to include AI-induced crime, whether intentional or unintentional.
2. Setting up core principles for the development and deployment of AI.
3. Phased implementation of AI regulations, starting with clear guidelines and evolving alongside AI advancements.
4. Encouraging developers to create only transparent and explainable AI systems.
5. Mandatory public disclosure and collaboration to inculcate societal and ethical considerations into account for the development of the regulatory framework.
6. Mandatory requirements of raising public awareness for AI developers and users.
7. Establishing independent and dedicated bodies to monitor AI development and deployment.
8. Mandatory funding for research on explainable AI, development of safety standards, and studying the societal implications of AI.

Conclusions

This study aimed to explore the appropriateness of a hybrid application of existing 'Product Liability' Law and 'Negligence' Law for artificial intelligence crimes. Through a systematic investigation using 'PESTEL', 'Root Cause Analysis' and 'Case Studies' approach the study delved deeper into validating the hypothesis and gained valuable insights into the requirements of a legal framework for AI systems.

There are a lot of complexities of AI accountability. While the responsibilities of programmer remains crucial, the ever-evolving nature of AI systems necessitates a multi-layered framework. The unique features of AI demand a unique approach. Since AI technologies have been increasingly used across international boundaries, if AI is to benefit society, it has to have international cooperation, robust safety standards, and unending adaptation. Focusing on core regulating principles with a phased implementation and prioritizing transparency, accountability and proactive measures such as public education, having specialized dedicated regulating bodies and adequate funds for continued research for responsible AI would certainly ensure a future where AI serves humanity only for good.

The study demonstrated a highly scientific qualitative unprecedented approach to address the issue of the development of a regulatory framework for AI and to draw pertinent inferences. The study substantially contributes to the existing literature by proposing apt considerations and measures for a robust AI framework. Future research on explainable AI and the development of safety standards for AI would provide a more comprehensive understanding of the required AI regulations.

References

Abbott, R., & Sarch, A. (2019). Punishing Artificial Intelligence: Legal Fiction or Science Fiction. *UC Davis Law Review*, *53*, 323–384. https://doi.org/10.2139/SSRN.3327485

Barsalou, M. A. (2014). *Root cause analysis: A step-by-step guide to using the right tool at the right time*. New York: CRC Press. https://doi.org/10.1201/b17834

Barsalou, M., & Starzyńska, B. (2023). Inquiry into the Use of Five Whys in Industry. *Quality Innovation Prosperity*, *27*(1), 62–78. https://doi.org/10.12776/qip.v27i1.1771

Bhatt, N., & Bhatt, J. (2023). Towards a Novel Eclectic Framework for Administering Artificial Intelligence Technologies: A Proposed 'PEEC' Doctrine. *EPRA International Journal of Research and Development (IJRD)*, *8*(9), 27–36. https://doi.org/10.13140/RG.2.2.11434.18888

Binns, R. (2018). Algorithmic accountability and public reason. *Philosophy & Technology*, *31*(4), 543–556. https://10.1007/s13347-017-0263-5

Broadhurst, R., Maxim, D., Brown, P., Trivedi, H., & Wang, J. (2019). Artificial Intelligence and Crime. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3407779

Calo, R. (2019). Artificial intelligence policy: A primer and roadmap. *UC Davis Law Review*, *51*(2), 399–435. https://dx.doi.org/10.2139/ssrn.3015350

Dremliuga, R., & Prisekina, N. (2020). The Concept of Culpability in Criminal Law and AI Systems. *Journal of Programming Languages*, *13*(3), 256. https://doi.org/10.5539/jpl.v13n3p256

Gless, S. (2019). AI in the Courtroom: a comparative analysis of machine evidence in criminal trials. *Georgetown Journal of International Law*, *51*(2), 195–253.

Ivan, D., & Manea, T. (2022). AI Use in Criminal Matters as Permitted Under EU Law and as Needed to Safeguard the Essence of Fundamental Rights. *International Journal of Law in Changing World*, *1*(1), 17–32. https://doi.org/10.54934/ijlcw.v1i1.15

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Khan, K., Ali, A., Khan, Z., & Siddiqua, H. (2021). Artificial Intelligence and Criminal Culpability. In *2021 International Conference on Innovative Computing (ICIC), IEEE* (pp. 1–7). https://doi.org/10.1109/icic53490.2021.9692954

Khisamova, Z., & Begishev, I. (2019). Criminal Liability and Artificial Intelligence: Theoretical and Applied Aspects. *Russian Journal of Criminology*, *13*(4), 564–574. https://doi.org/10.17150/2500-4255.2019.13(4).564-574

King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2021). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. In J. Cowls, & J. Morley (Eds.), *The 2020 Yearbook of the Digital Ethics Lab. Digital Ethics Lab Yearbook*. Springer, Cham. https://doi.org/10.1007/978-3-030-80083-3_14

Kingston, J. K. (2016). Artificial Intelligence and Legal Liability. In M. Bramer, & M. Petridis (Eds.), *Research and Development in Intelligent Systems XXXIII. SGAI 2016*. Springer, Cham. https://doi.org/10.1007/978-3-319-47175-4_20

Lagioia, F., & Sartor, G. (2019). AI Systems under Criminal Law: A Legal Analysis and A Regulatory Perspective. *Philosophy & Technology*, *33*, 433–465. https://doi.org/10.1007/s13347-019-00362-x

Li, Yao (2023). Specifics of Regulatory and Legal Regulation of Generative Artificial Intelligence in the UK, USA, EU and China. Law. *Journal of the Higher School of Economics*, *16*(3), 245–267 (in Russ.). https://doi.org/10.17323/2072-8166.2023.3.245.267

Petrella, S., Miller, C., & Cooper, B. (2021). Russia's artificial intelligence strategy: the role of state-owned firms. *Orbis*, *65*(1), 75–100. https://doi.org/10.1016/j.orbis.2020.11.004

Pugna, A., Negrea, R., & Miclea, S. (2016). Using Six Sigma Methodology to Improve the Assembly Process in an Automotive Company. *Procedia – Social and Behavioral Sciences*, *221*, 308–316. https://doi.org/10.1016/J.SBSPRO.2016.05.120

Qatawneh, I., Moussa, A., Haswa, M., Jaffal, Z., & Barafi, J. (2023). Artificial Intelligence Crimes. *Academic Journal of Interdisciplinary Studies*, *12*(1), 143–150. https://doi.org/10.36941/ajis-2023-0012

Scherer, M. U. (2015). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, *29*(2), 353. https://doi.org/10.2139/ssrn.2609777

Shestak, V., Volevodz, A., & Alizade, V. (2019). On the Possibility of Doctrinal Perception of Artificial Intelligence as the Subject of Crime in the System of Common Law: Using the Example of the U.S. Criminal Legislation. *Russian Journal of Criminology*. *13*(4), 547–554. https://doi.org/10.17150/2500-4255.2019.13(4).547-554

Solum, L. B. (2020). Legal personhood for artificial intelligences. In *Machine ethics and robot ethics* (pp. 415–471). Routledge.

Sukhodolov, A., Bychkov, A., & Bychkova, A. (2020). Criminal Policy for Crimes Committed Using Artificial Intelligence Technologies: State, Problems, Prospects. *Journal of Siberian Federal University*, *13*(1), 116–122. https://doi.org/10.17516/1997-1370-0542

Yüksel, I. (2012). Developing a Multi-Criteria Decision Making Model for PESTEL Analysis. *International Journal of Biometrics*, *7*(24), 52. https://doi.org/10.5539/IJBM.V7N24P52

Zhao, S. (2024). *Principle of Criminal Imputation for Negligence Crime Involving Artificial Intelligence*. Springer Nature. https://10.1007/978-981-97-0722-5

Author information

Neelkanth Bhatt – PhD (Engineering), Assistant Professor, Department of Civil Engineering, Lukhdhirji Engineering College
Address: Sama Kanthe, Morbi, Gujarat 363642, India
E-mail: neelkanth78bhatt@gmail.com
ORCID ID: https://orcid.org/0000-0003-0315-2985
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58919442100
WoS Researcher ID: https://www.webofscience.com/wos/author/record/KRO-8652-2024
Google Scholar ID: https://scholar.google.com/citations?user=L7K-e3IAAAAJ

Conflict of interest

The author declare no conflict of interest.

Financial disclosure

The research had no sponsorship.

Acknowledgements

The author is grateful to Shri. Jaikishen Bhatt, Retired Social Security Officer, Employees' State Insurance Corporation, Ahmedabad (Gujarat, India) and Prof. Bipin Pandit, Retired Professor of Civil Engineering, Lukhdhirji Engineering College, Morbi (Gujarat, India) for their expert help with language, writing and meticulous proofreading which significantly improved the clarity and the quality of the work.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – September 2, 2024
Date of approval – September 20, 2024
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:343.3/.7:004.056:004.9
EDN: https://elibrary.ru/rtolza
DOI: https://doi.org/10.21202/jdtl.2025.3

Преступления в эпоху искусственного интеллекта: гибридный подход к ответственности и безопасности в цифровую эру

Нилкант Бхатт

Инженерный колледж Лухдхирджи, Морби, Индия

Ключевые слова

безопасность,
искусственный интеллект,
качество продукции,
метод PESTEL,
право,
преступление,
преступность,
уголовная ответственность,
уголовное законодательство,
цифровые технологии

Аннотация

Цель: Изучение применимости существующих норм об ответственности за качество продукции и законов о халатности к преступлениям, связанным с использованием искусственного интеллекта. Автор выдвигает гипотезу о том, что гибридное применение этих правовых механизмов может стать основой для создания эффективной системы регулирования в условиях стремительного развития технологий.

Методы: комплексный подход, основанный на анализе PESTEL (политические, экономические, социальные, технологические, экологические и правовые факторы), методе анализа первопричин «Пять почему» и изучении кейсов из различных стран. Такой многоуровневый подход позволяет не только выявить ключевые проблемы, но и предложить адаптированные решения, учитывающие специфику преступлений, связанных с искусственным интеллектом.

Результаты: исследование демонстрирует, что существующие нормы об ответственности за качество продукции и халатности недостаточно эффективны для регулирования преступлений, связанных с искусственным интеллектом. Основными препятствиями являются технологическая сложность, отсутствие прецедентов, недостаточная осведомленность потребителей и юрисдикционные проблемы. Автор приходит к выводу, что для эффективного регулирования необходима глобальная система, включающая четкие принципы ответственности, строгие стандарты безопасности и постоянную адаптацию к новым вызовам.

Научная новизна: заключается в уникальном подходе к изучению преступлений, связанных с искусственным интеллектом, через призму гибридного применения существующих правовых механизмов. Исследование предлагает новый взгляд на проблему, сочетая теоретический анализ с практическими рекомендациями, основанными на изучении реальных кейсов.

Практическая значимость: состоит в разработке конкретных рекомендаций для законодателей и регулирующих органов. Автор подчеркивает необходимость создания специализированных органов, внедрения образовательных программ для граждан и сотрудников, а также обеспечения финансирования исследований в области объяснимого искусственного интеллекта и стандартов безопасности. Эти меры направлены на формирование устойчивой системы регулирования, способной эффективно противостоять преступлениям, связанным с использованием искусственного интеллекта. Работа открывает новые горизонты для дальнейших исследований в области регулирования технологий искусственного интеллекта, подчеркивая необходимость международного сотрудничества и междисциплинарного подхода.

Для цитирования

Бхатт, Н. (2025). Преступления в эпоху искусственного интеллекта: гибридный подход к ответственности и безопасности в цифровую эру. *Journal of Digital Technologies and Law*, *3*(1), 65–88. https://doi.org/10.21202/jdtl.2025.3

Список литературы

Ли, Яо. (2023). Особенности нормативно-правового регулирования генеративного искусственного интеллекта в Великобритании, США, Евросоюзе и Китае. *Право. Журнал Высшей школы экономики*, *16*(3), 245–267. EDN: https://elibrary.ru/yitzoa DOI: https://doi.org/10.17323/2072-8166.2023.3.245.267

Шестак, В. А., Волеводз, А. Г., Ализаде, В. А. (2019). О возможности доктринального восприятия системой общего права искусственного интеллекта как субъекта преступления: на примере уголовного законодательства США. *Всероссийский криминологический журнал*, *13*(4), 547–554. https://doi.org/10.17150/2500-4255.2019.13(4).547-554

Abbott, R., & Sarch, A. (2019). Punishing Artificial Intelligence: Legal Fiction or Science Fiction. *UC Davis Law Review*, *53*, 323–384. https://doi.org/10.2139/SSRN.3327485

Barsalou, M. A. (2014). *Root cause analysis: A step-by-step guide to using the right tool at the right time*. New York: CRC Press. https://doi.org/10.1201/b17834

Barsalou, M., & Starzyńska, B. (2023). Inquiry into the Use of Five Whys in Industry. *Quality Innovation Prosperity*, *27*(1), 62–78. https://doi.org/10.12776/qip.v27i1.1771

Bhatt, N., & Bhatt, J. (2023). Towards a Novel Eclectic Framework for Administering Artificial Intelligence Technologies: A Proposed 'PEEC' Doctrine. *EPRA International Journal of Research and Development (IJRD)*, *8*(9), 27–36. https://doi.org/10.13140/RG.2.2.11434.18888

Binns, R. (2018). Algorithmic accountability and public reason. *Philosophy & Technology*, *31*(4), 543–556. https://10.1007/s13347-017-0263-5

Broadhurst, R., Maxim, D., Brown, P., Trivedi, H., & Wang, J. (2019). Artificial Intelligence and Crime. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3407779

Calo, R. (2019). Artificial intelligence policy: A primer and roadmap. *UC Davis Law Review*, *51*(2), 399–435. https://dx.doi.org/10.2139/ssrn.3015350

Dremliuga, R., & Prisekina, N. (2020). The Concept of Culpability in Criminal Law and AI Systems. *Journal of Programming Languages*, *13*(3), 256. https://doi.org/10.5539/jpl.v13n3p256

Gless, S. (2019). AI in the Courtroom: a comparative analysis of machine evidence in criminal trials. *Georgetown Journal of International Law*, *51*(2), 195–253.

Ivan, D., & Manea, T. (2022). AI Use in Criminal Matters as Permitted Under EU Law and as Needed to Safeguard the Essence of Fundamental Rights. *International Journal of Law in Changing World*, *1*(1), 17–32. https://doi.org/10.54934/ijlcw.v1i1.15

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Khan, K., Ali, A., Khan, Z., & Siddiqua, H. (2021). Artificial Intelligence and Criminal Culpability. In *2021 International Conference on Innovative Computing (ICIC), IEEE* (pp. 1–7). https://doi.org/10.1109/icic53490.2021.9692954

Khisamova, Z., & Begishev, I. (2019). Criminal Liability and Artificial Intelligence: Theoretical and Applied Aspects. *Russian Journal of Criminology*, *13*(4), 564–574. https://doi.org/10.17150/2500-4255.2019.13(4).564-574

King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2021). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. In J. Cowls, & J. Morley (Eds.), *The 2020 Yearbook of the Digital Ethics Lab. Digital Ethics Lab Yearbook*. Springer, Cham. https://doi.org/10.1007/978-3-030-80083-3_14

Kingston, J. K. (2016). Artificial Intelligence and Legal Liability. In M. Bramer, & M. Petridis (Eds.), *Research and Development in Intelligent Systems XXXIII. SGAI 2016*. Springer, Cham. https://doi.org/10.1007/978-3-319-47175-4_20

Lagioia, F., & Sartor, G. (2019). AI Systems under Criminal Law: A Legal Analysis and A Regulatory Perspective. *Philosophy & Technology*, *33*, 433–465. https://doi.org/10.1007/s13347-019-00362-x

Petrella, S., Miller, C., & Cooper, B. (2021). Russia's artificial intelligence strategy: the role of state-owned firms. *Orbis*, *65*(1), 75–100. https://doi.org/10.1016/j.orbis.2020.11.004

Pugna, A., Negrea, R., & Miclea, S. (2016). Using Six Sigma Methodology to Improve the Assembly Process in an Automotive Company. *Procedia – Social and Behavioral Sciences*, *221*, 308–316. https://doi.org/10.1016/J.SBSPRO.2016.05.120

Qatawneh, I., Moussa, A., Haswa, M., Jaffal, Z., & Barafı, J. (2023). Artificial Intelligence Crimes. *Academic Journal of Interdisciplinary Studies*, *12*(1), 143–150. https://doi.org/10.36941/ajis-2023-0012

Scherer, M. U. (2015). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, *29*(2), 353. https://doi.org/10.2139/ssrn.2609777

Solum, L. B. (2020). Legal personhood for artificial intelligences. In *Machine ethics and robot ethics* (pp. 415–471). Routledge.

Sukhodolov, A., Bychkov, A., & Bychkova, A. (2020). Criminal Policy for Crimes Committed Using Artificial Intelligence Technologies: State, Problems, Prospects. *Journal of Siberian Federal University*, *13*(1), 116–122. https://doi.org/10.17516/1997-1370-0542

Yüksel, I. (2012). Developing a Multi-Criteria Decision Making Model for PESTEL Analysis. *International Journal of Biometrics*, *7*(24), 52. https://doi.org/10.5539/IJBM.V7N24P52

Zhao, S. (2024). *Principle of Criminal Imputation for Negligence Crime Involving Artificial Intelligence*. Springer Nature. https://10.1007/978-981-97-0722-5

Сведения об авторе

Нилкант Бхатт – PhD в области инженерных наук, кафедра гражданского проектирования, Инженерный колледж Лухдхирджи, г. Морби, Индия
Адрес: Индия, 363642, г. Морби, Гуджарат, Сама Канте
E-mail: neelkanth78bhatt@gmail.com
ORCID ID: https://orcid.org/0000-0003-0315-2985
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58919442100
WoS Researcher ID: https://www.webofscience.com/wos/author/record/KRO-8652-2024
Google Scholar ID: https://scholar.google.com/citations?user=L7K-e3IAAAAJ

Конфликт интересов

Автор сообщает об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Благодарность

Автор выражает благодарность господину Джайкишен Бхатт, сотруднику Службы социального обеспечения в отставке (Государственная корпорация страхования работников, Ахмадабад, Гуджарат, Индия) и господину Бипин Пандит, профессору гражданского строительства в отставке (Инженерный колледж Лухдхирджи, Морби, Гуджарат, Индия) за их квалифицированную помощь в формулировании, написании и тщательной корректуре работы, что значительно повысило ее точность и качество.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.77 / Уголовное право
Специальность ВАК: 5.1.4 / Уголовно-правовые науки

История статьи

Дата поступления – 2 сентября 2024 г.
Дата одобрения после рецензирования – 20 сентября 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:177.5:316.647.82:004.8:004.652
EDN: https://elibrary.ru/mbwjxf
DOI: https://doi.org/10.21202/jdtl.2025.4

Constitutional-Legal Aspect of Creating Large Language Models: the Problem of Digital Inequality and Linguistic Discrimination

Ilya G. Ilin

Saint Petersburg State University, Saint Petersburg, Russia

Keywords

artificial intelligence,
constitutional rights,
digital inequality,
digital technologies,
generative artificial intelligence,
human rights,
large language models,
law,
linguistic discrimination,
natural language processing

Abstract

Objective: to study the impact of digital inequality on the implementation of constitutional human rights; to identify the risks of linguistic discrimination associated with the development and use of large language models.

Methods: formal-legal and comparative-legal methods, as well as the method of theoretical modeling. These approaches are complemented by general scientific methods of cognition, allowing for a comprehensive analysis of the legal, technological and social aspects of the issue.

Results: the research found that, in relation to large language models, digital inequality arises due to the uneven digitalization of languages and manifests itself in limited access to natural language processing technology. In turn, unequal access to this technology can negatively affect the implementation of constitutionally guaranteed rights and can be viewed from the viewpoint of equality and non-discrimination concepts. The author emphasizes that unequal access to natural language processing technologies can exacerbate existing social and economic inequalities and create new forms of discrimination.

Scientific novelty: hidden and indirect forms of discrimination are analyzed that manifest themselves in artificial intelligence systems, especially in generative models. While direct forms of discrimination can be detected in predictive algorithms, generative models create more subtle but no less significant cumulative effects. These effects contribute to the formation of social stereotypes and inequalities in areas such as professional activity, gender and ethnicity. The author also draws attention to the fact that with the increasing autonomy of artificial intelligence, traditional approaches

to discrimination detection are becoming less effective, which requires the development of new analysis and regulation methods.

Practical significance: the results provide a basis for identifying and assessing the legal risks associated with unequal access to digital products using natural language processing. This contributes to the improvement of legal regulation in the field of the development and use of artificial intelligence technologies. The article offers recommendations for lawmakers, regulators, and technology developers aimed at minimizing the risks of digital inequality and linguistic discrimination.

For citation

Ilin, I. G. (2025). Constitutional-Legal Aspect of Creating Large Language Models: the Problem of Digital Inequality and Linguistic Discrimination. *Journal of Digital Technologies and Law, 3*(1), 89–107. https://doi.org/10.21202/jdtl.2025.4

Contents

Introduction
1. Digitalization of languages as a source of digital inequality: a technical and legal analysis
2. Linguistic discrimination as a form of digital inequality
3. Qualification problem and criteria for assessing linguistic discrimination
Conclusions
References

Introduction

Large Language Models (LLM) are generative artificial intelligence models used in natural language processing technology (NLP). They allow a computer to efficiently process text data, demonstrating the ability to "understand" text at a deep level, create coherent and contextually relevant responses to queries, translate from one language to another, and generate texts that meet certain stylistic and content requirements (Glauner, 2024). Examples of large language models include BERT[1], GPT-3[2], and related digital products such as Google Assistant and ChatGPT.

[1] Bidirectional Encoder Representations from Transformers (BERT) is a large language model developed by Alphabet Inc. (USA), based on the Transformer architecture. It is trained on a bidirectional context – it can analyze and "understand" text both from left to right and from right to left. For more information about the BERT model, see (Devlin J. et al., 2018).

[2] Generative Pre-trained Transformer (GPT) is a series of large language models developed by Openway (USA), based on the Transformer architecture. It is taught without a "teacher", does not require adaptation and can be used and adapted for a wide range of tasks. For more information about the GPT model, see (Yenduri G. et al., 2023). For more information about the Transformer architecture, see (Vaswani, 2017).

Large language models are trained on vast arrays of linguistic data, including structured linguistic corpora: databases containing a variety of texts (books, text transcriptions, translations, etc.) and audio files (audiobooks, broadcast recordings, podcasts, and other audio content). The structure and representativeness of such data, their volume and format determine the learning process and the accuracy of understanding the context (Ilin, 2024), while defects[3] or insufficient data can lead to incorrect model functioning and generally hinder the technology development (Hacker, 2021). Thus, the possibility of creating a high-quality language model directly depends on the volume, representativeness and other qualitative characteristics of the training data for a particular language.

At the same time, the levels of digitalization of languages – the volume of existing linguistic corpora and the data for their creation – differ significantly. For some languages or dialects, data may be extremely limited or non-existent. This hinders the development of accurate and effective language models, slowing down their digital development and limiting their integration into modern technologies. For example, if a data set is not comprehensive enough and does not cover all variants of a particular language, the model may process incoming requests incorrectly or inaccurately, and in some cases may not function at all. Differences in pronunciation, vocabulary, and grammar can lead to errors in recognizing and analyzing text or speech, and reduce the quality of results.

The inability to create a full-fledged language model for particular languages or dialects makes unavailable many digital products for speakers of these languages or significantly worsens the quality of their functioning compared to how the same technologies function for speakers of languages with a high level of digitalization. As a result, digital inequality arises, when access to modern technologies is unevenly distributed among different linguistic communities, which, in turn, increases the risk of discrimination.

The article aims to analyze the constitutional and legal aspects of creating large language models in the context of digital inequality and linguistic discrimination. To achieve this goal, we will investigate how digital inequality affects constitutional human rights, as well as analyze the risks of linguistic discrimination associated with the creation of large language models.

The article contains the main results of the corresponding research, as well as directions for further study. The research paper is divided into three thematic sections, supplemented by an introduction and conclusion. The first section analyzes the problem of digital inequality given the different levels of digitalization of languages – the volume and representativeness of linguistic data. The second section examines linguistic discrimination as a potential form of digital inequality, with an emphasis on unequal

[3] In this context, a data defect includes both the data not meeting certain technical criteria and metrics, for example, the criteria of representativeness, volume, purity, etc. (quality defect) and a legal defect – the use of data in violation of the applicable legal regime. For example, a violation of the personal data regime during their processing as part of the language model. For more information about the impact of data quality on creating large language models, see (Ilin, 2024).

access to natural language processing (NLP). In the third section, the problem of linguistic discrimination is conceptualized in relation to other human rights and in the context of the digital technologies development.

1. Digitalization of languages as a source of digital inequality: a technical and legal analysis

Digital inequality is a form of social inequality characterized by unequal access to information technologies and by varying levels of skills in using them among individuals and social groups (Mushakov, 2022). This phenomenon encompasses a wide range of factors, including differences in technical equipment, access to Internet, digital literacy, and educational opportunities, which in turn leads to social and economic division (Rogers, 2016). The need to address gaps in access to digital technologies in order to achieve a more equal and inclusive society has been repeatedly highlighted both at the national[4] and international[5] levels.

In the context of creating large language models, the problem of digital inequality manifests itself in the limited ability of speakers of low-digitalization languages to use digital products in their language. This leads to unequal access of individuals or social groups to natural language processing (NLP) technologies. As a result, there may appear restrictions on access to information, education, and social services for native speakers of such languages. For example, the ability to understand text contextually and generate appropriate responses contributes to the active use of this technology in areas such as education and healthcare (Jiang et al., 2023; Sohail & Zhang, 2024). The lack of support for particular languages in these areas may negatively affect the exercising of the corresponding constitutionally guaranteed rights: the right to access education[6] and medical care[7]. This may limit the availability and quality of these services. In this regard, it seems logical to consider the problem of digital inequality from the viewpoint of constitutional and legal relations, i.e. the concepts of “equality” and prohibition of discrimination[8].

---

4 Decree of the Government of the Russian Federation No. 313 dated 15.04.2014 (2014). Here and further, all references to documents, regulations and judicial practice are given by SPS ConsultantPlyus refence system. https://clck.ru/3GP8do

5 Geneva Declaration of Principles (Building the Information Society: A Global Challenge in the New Millennium) (UN) of December 12, 2003 https://clck.ru/3GP8fD ; Tunisian Programme for the Information Society (UN) dated November 15, 2005. https://clck.ru/3GP8ge

6 The Constitution of the Russian Federation, adopted by popular vote on 12.12.1993 with amendments approved during the nationwide vote on 01.07.2020 (hereinafter referred to as the Constitution of the Russian Federation). Art. 43. https://clck.ru/3GP8hh

7 Constitution of the Russian Federation. Art. 41. https://clck.ru/3GP8jK

8 In both cases, the issue of equality of rights is considered, but the right to non-discrimination has a narrower content and in this sense stems from the general right to equality. For more information, see (Talapina, 2022).

We can also agree with some researchers that constitutional norms ensuring equality before the law and access to services should take into account and eliminate inequality in access to digital resources, as this directly affects the ability of citizens to exercise their rights and freedoms in the digital era (Mushakov, 2022).

To create an effective language model, a set of training data is needed that must meet criteria such as volume, representativeness[9], and other qualitative characteristics. These parameters directly depend on the level of digitalization of a particular language, since the higher the degree of digitalization, the more diverse and high-quality data can be used to train the model.

Digitalization of a language in a broad sense is the transformation of data into appropriate electronic linguistic corpora. For this purpose, text data (for example, files, transcriptions, abstracts), speech data (for example, audio recordings, phonetic and intonation abstracts) and multimodal data (i.e. data combining several types, for example, video and text, images and text, etc.) are used (Dash & Arulmozi, 2018). It should be noted that this process not only contributes to technological development and digital transformation of society, but also plays an important role in preserving national and cultural identity (Kelli et al., 2016). For example, digitalization of minority languages can significantly contribute to the preservation of the cultural heritage of small nations.

Despite the importance of digitalization for technological progress and the high social significance of this process, the level of digitalization of languages and their dialects remains uneven. There are economic, technical, and legal factors that limit or hinder the digitalization of languages.

The economic factors are related to the fact that languages have different economic potential (Alarcón, 2022; Monteith & Sung, 2023). Hence, digitalization requires significant resources, including time, finance, etc. In this regard, the development of linguistic corpora for some languages may be economically unfeasible. The technical factors are directly related to creating linguistic corpora. Such factors may include errors in data collection, flaws in the corpora design and limitations of existing datasets, errors in metadata, etc. (Solovyev & Akhtyamova, 2019; Doğruöz et al., 2023; Li et al., 2024). The legal factors are related to the presence of regulatory restrictions on access to training data and the need to comply with the relevant legal regime when using them for training.

In previous works, the author discussed in detail the issues of regulating access to training data (Ilin, 2024), as well as compliance with their legal regimes, such as the personal data regime (Ilin, 2020) and the intellectual property regime

[9] Given the multifaceted meaning of the term “representativeness” (for more details see (Chasalow & Levy, 2021)), it is important to note that in the context of this article, the volume of linguistic data means their quantity, while representativeness means their diversity, i.e. the degree to which various styles, dialects, time periods and contexts are covered.

(Ilin, 2022; Ilin & Kelli, 2019, 2024). The central topic of those studies was the conflict between equally protected human rights when using training data, such as the right to non-discrimination[10] and the right to privacy, to protection of personal and family secrets[11]. Overcoming this problem is necessary both at the conceptual level (removing regulatory barriers to data access, taking into account the balance of private and public interests) and in practical terms (creating conditions for the dissemination and exchange of linguistic data, for example, by developing an institution of reusing the data accumulated in government information systems (hereinafter referred to as GIS) or involving higher educational institutions to create linguistic corpora and digitalize the language).

According to the analytical report of the Accounts Chamber of the Russian Federation[12], by 2020, more than 800 federal state information systems were operating in Russia, providing data exchange between government agencies in various areas of public life. These systems cover a wide range of information, including statistics, as well as information on healthcare, education, and other key sectors. In this context, the use of GIS data to create linguistic corpora seems to be particularly promising. Despite the varying levels of development of these systems, one may expect that the data collected in them will have the necessary qualitative characteristics, and their diversity can provide the necessary representativeness and volume (Ilin, 2024). However, given the risks associated with legal restrictions on the use of data, their reuse should be carried out in accordance with uniform principles and regulations. These should include legislative standards and control mechanisms that take into account the specifics of each data type and its compliance with the purposes of its initial collection.

Another possible solution to the problem of access and lack of linguistic data is to involve higher education institutions to creating and subsequently disseminating linguistic corpora. The participation of universities in language digitalization can also be justified by the social significance of this activity. As an example of successful cooperation between commercial organizations and higher education institutions in the field of natural language processing, we can mention the joint academic program of the Center for Speech Technologies Group of Companies with the ITMO National Research University (Ilin & Dedova, 2019).

However, although this solves the problem of creating linguistic corpora, the issue of their further distribution remains open. For example, for various reasons, a university may not be interested in further dissemination of the linguistic corpus or may not have

[10] Constitution of the Russian Federation. Art. 19. https://clck.ru/3GPBg6

[11] Constitution of the Russian Federation. Art. 23. https://clck.ru/3GPBhb

[12] Center for Advanced Governance (2020). Assessment of the openness of government information systems in Russia: analytical report. https://clck.ru/3GPBjT

the necessary resources for this and, accordingly, may not distribute it. If a university operates as an entrepreneurial university and commercializes its results, for example, via a spin-off company, one may also question the possibility to employ the doctrine of the free use of works[13] when processing linguistic data. All these questions require further careful analysis, both from a legal and other points of view.

2. Linguistic discrimination as a form of digital inequality

Since, in relation to large language models development, digital inequality leads to unequal access of individuals or social groups to natural language processing (NLP) technologies (the inability to fully use this technology in their language), the problem of digital inequality should primarily be considered in the context of linguistic discrimination.

The problem of discrimination by artificial intelligence systems, although not new, remains relevant today. The development and active implementation of artificial intelligence systems in various spheres of life opens up new areas for discussion of this problem. Examples include discrimination by artificial intelligence systems in the field of labor relations (Morin, 2024), the impact of profiling[14] on human dignity (Orwat, 2024), the potential impact of artificial intelligence on discrimination based on ethnicity, religion and gender (Ozkul, 2024), etc.

In addition, with the increasing autonomy of artificial intelligence systems and the development of generative models, discrimination begins to take on an implicit character, which allows classifying its manifestations into direct and indirect ones. For example, unlike the obvious cases of discrimination observed in predictive crime analytics systems such as those based on PredPol[15] and COMPAS[16] algorithms, discrimination in generative artificial intelligence systems may be less apparent. For example, these systems may preferentially create images of white men in response

[13] Civil Code of the Russian Federation (part 4) of 18.12.2006 No. 230-FZ. Art. 1274. https://clck.ru/3GPBnL

[14] Profiling is a technique of intellectual data analysis that can be automated or semi-automated and aims to create classes or categories of characteristics from large datasets. In this process, data is collected, analyzed using various algorithms such as machine learning, and used to create profiles describing typical characteristics or behavioral patterns of groups or individuals. For more information, see (Bosco et al., 2015).

[15] PredPol (Predictive Policing) is a predictive analytics system used by police and designed to predict crimes. PredPol's main goal is to use historical data on crime to create maps of "hot spots" – areas where crimes are most likely to occur. For more information, see (Browning & Arrigo, 2021).

[16] COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a predictive analytics system designed to assess the risk of recidivism among convicts. Its main goal is to analyze data on offenses, behavior, and the social record of suspects in order to predict the likelihood of their reoffending. The system is used in judicial practice to help making decisions on sentencing and release conditions. For more information, see (Engel et al., 2024).

to repeated requests for examples of people employed in important professions, potentially leading to cumulative discriminatory effects (Hacker et al., 2024). In such cases, discrimination becomes difficult to detect, as it may not be explicit or obvious, but nevertheless has a significant impact on the representation and perception of various groups in society.

The National Strategy for the Development of Artificial Intelligence up to 2030[17] (hereinafter referred to as the Strategy) emphasizes that the protection of human rights and freedoms is one of the main principles of the development and use of artificial intelligence technology[18], while "non-discrimination" is highlighted as one of the main principles of the development of legal regulation of public relations in the sphere of development and use of artificial intelligence technology[19].

Article 2 of the Universal Declaration of Human Rights (1948)[20] prohibits discrimination, including on the basis of language. A similar provision is contained in Article 1 (3) of the UN Charter[21], and is also reflected in paragraph 2 of Article 19 of the Constitution of the Russian Federation, according to which the state guarantees equality of human and civil rights and freedoms, regardless of language.

In the field of language discrimination, several key aspects are identified, related to its recognition, legal protection and public perception. One of the main problems is the lack of recognition of linguistic discrimination at the international level. For example, discrimination based on voice often goes unnoticed (Baugh, 2023), which can be critical when interacting with speech and voice recognition technology and related digital products: interactive response systems and voice assistants.

The UN Human Rights Committee[22] has repeatedly addressed the issue of linguistic discrimination, but its judicial practice is underdeveloped and does not provide reliable protection for linguistic minorities (Möller, 2011). The regulatory framework at various

---

17 The National Strategy for the Development of Artificial Intelligence up to 2030 was approved by Decree of the Russian President "ated 10.10.2019 No. 490 «On the development of artificial intelligence in the Russian Federation" (hereinafter referred to as the National Strategy for the Development of Artificial Intelligence up to 2030).

18 National Strategy for the Development of Artificial Intelligence up to 2030. 19 (a). https://clck.ru/3Ghyfz

19 National Strategy for the Development of Artificial Intelligence up to 2030. cl. 51 (10) (d). https://clck.ru/3Ghyfz

20 The Universal Declaration of Human Rights (adopted by the UN General Assembly on 10.12.1948). https://clck.ru/3GPBqd

21 Charter of the United Nations Organization (adopted in San Francisco on 26.06.1945). https://clck.ru/3GPBsN

22 The UN Human Rights Committee was established on the basis of the International Covenant on Civil and Political Rights, which was adopted by the UN General Assembly in 1966 and entered into force in 1976. The Committee is the body that monitors the fulfillment of the obligations assumed under this Covenant by participating states. The Committee considers reports from the states on how they respect the rights enshrined in the Covenant, as well as individual complaints about violations of rights (if the state has recognized the Committee's jurisdiction in this matter). More about the Committee: https://clck.ru/3GPBvJ

levels also often does not take into account all the nuances of linguistic discrimination. Legislation at the international, regional and national levels, as a rule, does not provide sufficient protection for the rights of linguistic minorities, which leads to gaps in the legal protection of crime victims (Chilingaryan et al., 2020).

Discrimination based on language can be defined as any unjustified distinction or restriction that weakens or excludes the possibility of exercising rights enshrined in international or national regulations based on language affiliation. At the same time, it should be added that states also have positive obligations to protect and promote linguistic rights as part of their obligation to respect human rights[23]. Therefore, in the context of creating large language models, it seems necessary to expand the definition of linguistic discrimination to include actions aimed at hindering the preservation or development of minority languages. The essence of the first part of the definition is that linguistic discrimination occurs when a person experiences worse treatment than others in a similar situation due to insufficient or complete lack of proficiency in the official language established in a given state or region. The second part refers to a deeper aspect of this problem – the states' fulfillment of legal obligations, established by international conventions and national legislation, to protect and promote minority languages. This said, it should be noted that the expansion of the linguistic discrimination concept is more likely to reflect the perspective sought by judicial practice and scientific discussion, rather than the current perception of the problem by law enforcers and lawyers.

3. Qualification problem and criteria for assessing linguistic discrimination

Ambiguity in the definition of linguistic discrimination makes law enforcement difficult and raises questions about the criteria used in assessing these situations. As noted earlier, linguistic discrimination occurs when people are treated differently because of their language proficiency or accent, which often leads to limited access to opportunities and rights (Mironova, 2019). However, linguistic discrimination is a multifaceted problem that differs from other forms of discrimination, such as racial or religious, and depends on various factors. An analysis of existing practice allows us to identify a number of key factors for determining linguistic discrimination. The first factor is the number of native speakers: the level of discrimination is often determined by the prevalence of language in society. For example, in Cameroon, the English-speaking minority faces systemic discrimination due to its small size compared to the French-speaking majority (Donard, 2023).

[23] For example, the obligations arising from Federal Law No. 273-FZ of December 29, 2012 "On education in the Russian Federation", Federal Law No. 74-FZ of 17.06.1996 "On national cultural autonomy".

Another important factor is the ability of the state to support multilingualism. The more actively the state creates conditions for learning and using multiple languages, the lower the likelihood of linguistic discrimination. For example, research shows that support for multilingualism in educational institutions helps to reduce discrimination based on language (Page, 2023).

The use of minority languages in public life is also of great importance. When these languages do not receive institutional support, their speakers are often marginalized, which reinforces existing social inequalities.

In addition, it should be borne in mind that linguistic discrimination may overlap with other forms of discrimination, such as racial, religious, or ethnic. In such cases, people are subjected to complex forms of discrimination, which significantly exacerbates the problem (Drożdżowicz & Peled, 2024). In order to illustrate the complexity of the problem, let us briefly consider some of these intersections.

Failure to provide equal access to services in the mother tongue may violate the right to equality, creating barriers that hinder full participation in social life[24]. These barriers, for example, can affect the right to education[25] by limiting access to educational resources and materials in the mother tongue, which can reduce the quality of education and limit educational opportunities.

In addition, linguistic discrimination affects the right to freedom of expression[26]. People should be able to express their opinions freely in the language they prefer, and restrictions on this may be seen as a violation of this fundamental right. Linguistic discrimination also affects cultural rights, as language is a key element of cultural identity and expression. Restricting the use of a minority language in cultural and social contexts can undermine the cultural rights of these communities and their ability to preserve and develop their cultural identity.

Access to justice can also be hampered by language barriers, as the need to understand and participate in court proceedings in one's native language is critical to ensuring fair justice[27]. Language barriers may prevent the correct understanding of charges, court procedure, or legal decisions, which can lead to unfair outcomes.

---

24 D.H. and Others v. Czech Republic: Judgment of the Grand Chamber of the European Court of Human Rights of November 13, 2007 (complaint No. 57325/00).

25 Communication No. 760/1997. J.G.A. Diergaardt (late Captain of the Rehoboth Baster Community) et al. v. Namibia, Views of 25 July 2000, CCPR/C/69/D/760/1997.

26 Communication No. 221/1987. Yves Cadoret and Hervé Le Bihan v. France, Views of 11 April 1991, CCPR/C/41/D/221/1987; Communication No. 219/1986. Dominique Guesdon v. France, Views of 25 July 1990, CCPR/C/39/D/219/1986.

27 For example, the court's refusal to provide the accused with the text of the indictment translated into the Karachay language led to the cancellation of the verdict due to violations of the norms of criminal and criminal-procedural law by the preliminary investigation authorities. For more information, see "Review of the cassation practice of the Judicial Board for Criminal Cases of the Supreme Court of the Russian Federation of 2003" (2004). Bulletin of the Supreme Court of the Russian Federation, 9.

Thus, although it is possible to identify factors for assessing linguistic discrimination, the legal qualification of such cases in the context of digital technologies causes certain difficulties. For example, it is necessary to find out whether errors in the language model can be considered a manifestation of discrimination. Such errors are often difficult to detect, as discrimination may be hidden, which makes it less obvious for analysis. Discrimination in models can be the result of algorithmic or human bias. Algorithmic bias occurs due to limitations or distortions in the data on which the model is trained, whereas human bias can manifest itself in the developing and configuring algorithms (Kharitonova et al., 2021). Both forms of bias can not only affect the accuracy and fairness of decisions, but also maintain or exacerbate existing social inequalities, ultimately leading to discrimination. The distinction between errors and discrimination requires in-depth analysis, as errors may be accidental or may result from systemic biases. It is important to understand how bias, both algorithmic and human, affects the decision-making process and how it is integrated into algorithms and models. This understanding is necessary to develop more equitable and inclusive digital systems.

Conclusions

This article aims to analyze the constitutional-legal aspects of creating large language models in the context of digital inequality and linguistic discrimination. The study found that digital inequality in the context of large language models is due to the uneven digitalization of languages and manifests itself in limited access to natural language processing technologies. Such unequal access can negatively affect the implementation of constitutionally guaranteed rights and requires consideration through the prism of such concepts as "equality" and prohibition of discrimination. In turn, the identification and legal qualification of linguistic discrimination when creating large language models is a difficult task, since biases in models can be hidden and have a cumulative discriminatory effect. Discrimination may be caused by both algorithmic and human bias. Algorithmic bias occurs due to limitations or distortions in the data on which the model is trained, while human bias can manifest itself in developing and configuring algorithms. Distinguishing between these categories and assessing their impact on decision-making are becoming important areas for future research aimed at developing mechanisms to ensure equal access to digital technologies and the protection of language rights.

References

Alarcón, A. A. (2022). The economics of language. In Miquel Àngel Pradilla Cardona (Ed.), *Catalan Sociolinguistics: State of the art and future challenges* (pp. 173–182). https://doi.org/10.1075/ivitra.32.12ala

Baugh, J. (2023). Linguistic profiling across international geopolitical landscapes. *Daedalus*, *152*(3), 167–177. https://doi.org/10.1162/daed_a_02024

Bosco, F., Creemers, N., Ferraris, V., Guagnin, D., & Koops, B. J. (2015). Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. In S. Gutwirth, R. Leenes, P. de Hert (Eds.), *Reforming European data protection law* (Vol. 20, pp. 3–33). https://doi.org/10.1007/978-94-017-9385-8_1

Browning, M., & Arrigo, B. (2021). Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice*, *46*(2), 298–316. https://doi.org/10.1007/s12103-020-09557-x

Chasalow, K., & Levy, K. (2021). Representativeness in statistics, politics, and machine learning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 77–89). https://doi.org/10.1145/3442188.3445872

Chilingaryan, K., Meshkova, I., & Sheremetieva, O. (2020). International legal protection of linguistic minorities. *International Journal of Psychosocial Rehabilitation*, *24*(6), 9750–9758. https://doi.org/10.37200/IJPR/V24I6/PR26097

Dash, N. S., & Arulmozi, S. (2018). *History, features, and typology of language corpora*. Springer Singapore. https://doi.org/10.1007/978-981-10-7458-5

Devlin, J., Chang, Ming-Wei, Lee, Kenton, & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Doğruöz, A. S., Sitaram, S., & Yong, Z. X. (2023). *Representativeness as a forgotten lesson for multilingual and code-switched data collection and preparation*. arXiv preprint arXiv:2310.20470 (pp. 5751–5767).

Donard, K. (2023). Legal protection of linguistic minority under discrimination: the case of anglophone Cameroon. *International Journal of Business and Technology*, *11*(2), Article 1.

Drożdżowicz, A., & Peled, Y. (2024). The complexities of linguistic discrimination. *Philosophical Psychology*, *37*(6), 1459–1482. https://doi.org/10.1080/09515089.2024.2307993

Engel, C., Linhardt, L., & Schubert, M. (2024). Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism. In *Artificial Intelligence and Law*. https://doi.org/10.1007/s10506-024-09389-8

Glauner, P. (2024). Technical foundations of generative AI models. In *Legal Tech – Zeitschrift für die digitale Anwendung*, *1*, 24–34.

Hacker, P. A (2021). Legal framework for AI training data—from first principles to the Artificial Intelligence Act. *Law, Innovation and Technology*, *13*(2), 257–301. https://doi.org/10.1080/17579961.2021.1977219

Hacker, P., Mittelstadt, B., Zuiderveen Borgesius, F., Wachteret, S. (2024). *Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It*. arXiv preprint arXiv:2407.10329. https://doi.org/10.2139/ssrn.4877398

Ilin, I. (2020). The Voice and Speech Processing within Language Technology Applications: Perspective of the Russian Data Protection Law. *Legal Issues in the Digital Age*, *1*, 99–123. https://doi.org/10.17323/2713-2749.2020.1.99.123

Ilin, I. (2022). Legal Regime of the Language Resources in the Context of the European Language Technology Development. In Z. Vetulani, P. Paroubek, M. Kubis (Eds.), *Human Language Technology. Challenges for Computer Science and Linguistics. LTC 2019. Lecture Notes in Computer Science* (vol. 13212, pp. 367–376). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-05328-3_24

Ilin, I. (2024). Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data. *Legal Issues in the Digital Age*, *2*, 36–56. https://doi.org/10.17323/2713-2749.2024.2.36.56

Ilin, I., & Dedova, M. (2019). Academic Entrepreneurship in the Field of Language Resource Creation and Dissemination. In A. Riviezzo, M. Rosaria Napolitano, & A. Garofano (Eds.), *The ESU 2019 Conference and Doctoral Programme, Naples (Italy), 8–14 September 2019. Electronic Conference Proceedings* (pp. 193–200).

Ilin, I., & Kelli, A. (2019). The use of human voice and speech in language technologies: the EU and Russian intellectual property law perspectives. *Juridical International*, *28*, 17–27. https://doi.org/10.12697/ji.2019.28.03

Ilin, I., & Kelli, A. (2024). Natural Language, Legal Hurdles: Navigating the Complexities in Natural Language Processing Development and Application. *Journal of the University of Latvia*. *Law*, *17*, 44–67. https://doi.org/10.22364/jull.17.03

Jiang, X., Yan, L., Vavekanand, R., & Hu, M. (2023). Large Language Models in Healthcare Current Development and Future Directions. *Generative AI Research*, *2*, 12. https://doi.org/10.20944/preprints202407.0923.v1

Kelli, A., Vider, K., Pisuke, H., & Siil, T. (2016). Constitutional values as a basis for the limitation of copyright within the context of digitalisation of the Estonian language. In *Constitutional Values in Contemporary Legal Space* (Vol. II, pp. 126–139).

Kharitonova, Yu. S., Savina, V. S., & Pagnini, F. (2021). Artificial Intelligence's Algorithmic Bias: Ethical and Legal Issues. Perm University Herald. *Juridical Sciences*, *53*, 488–515. (In Russ.). https://doi.org/10.17072/1995-4190-2021-53-488-515

Li, X., Dou, Zh., Zhou, Yu., & Liu, F. (2024). CorpusLM: Towards a unified language model on corpus for knowledge-intensive tasks. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 26–37). https://doi.org/10.1145/3626772.3657778

Mironova, M. V. (2019). Formation of the term "Linguistic discrimination" in modern sociolinguistics. In *New Language. New World. New Thinking: collection of works of the 2nd Annual international scientific-practical conference* (pp. 555–558). Moscow: Diplomatic Academy of Ministry of Foreign Affairs of the Russian Federation. (In Russ.).

Möller, J. T. (2011). Case Law of the UN Human Rights Committee relevant to Members of Minorities and Peoples in the Arctic Region. *The Yearbook of Polar Law Online*, *3*(1), 27–56. https://doi.org/10.1163/22116427-91000054

Monteith, B., & Sung, M. (2023). Unleashing the Economic Potential of Large Language Models: The Case of Chinese Language Efficiency. *TechRxiv*. *June 07*. https://doi.org/10.36227/techrxiv.23291831.v1

Morin, S. L. (2024). AI Discrimination in Hiring. In D. Norman (Ed.), *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology* (pp. 64–74). IGI Global. https://doi.org/10.4018/979-8-3693-1906-2.ch004

Mushakov, V. (2022). Constitutional human rights in the context of bridging the digital divide. *Vestnik of the St. Petersburg University of the Ministry of Internal Affairs of Russia*, *2022*(1). (In Russ.). https://doi.org/10.35750/2071-8284-2022-1-69-73

Orwat, C. (2024). Algorithmic Discrimination From the Perspective of Human Dignity. *Social Inclusion*, *12*, 1–18. https://doi.org/10.17645/si.7160

Ozkul, D. (2024). Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination. *Social Inclusion*, *12*, 1–3. https://doi.org/10.17645/si.8942

Page, C. (2023). Academic language development and linguistic discrimination: Perspectives from internationally educated students. *Comparative and International Education*, *52*(2), 39–53. https://doi.org/10.5206/cie-eci.v52i2.15000

Rogers, S. E. (2016). Bridging the 21st century digital divide. *TechTrends*, *60*(3), 197–199. https://doi.org/10.1007/s11528-016-0057-0

Sohail, A., & Zhang, L. (2024). *Integrating large language models into the psychological sciences*. https://doi.org/10.1007/s12144-025-07438-2

Solovyev, V. D., & Akhtyamova, S. (2019). Linguistic Big Data: Problem of Purity and Representativeness. In *21st International Conference on Data analytics and management in data intensive domains, DAMDID/RCDL 2019* (pp. 193–204).

Talapina, E. (2022). Artificial Intelligence Processing and Risks of Discrimination. *Law Journal of the Higher School of Economics*, *1*, 4–27. (In Russ.). https://doi.org/10.17323/2072-8166.2022.1.4.27

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. https://doi.org/10.48550/arXiv.1706.03762

Yenduri, G., Ramalingam, M., Chemmalar Selvi, G., Supriya, Y., Srivastava, G., Maddikunta, P. K. R. et al. (2023). Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. In *IEEE Access* (Vol. 12, pp. 54608–54649). https://doi.org/10.1109/access.2024.3389497

Author information

Ilya G. Ilin – Master of Law (information technologies), postgraduate student, Faculty of Law, Saint Petersburg State University
Address: 22nd line of Vasilievsky Island, 7199106 Saint Petersburg, Russian Federation
E-mail: i.g.ilin@spbu.ru
ORCID ID: https://orcid.org/0000-0003-1076-2765
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57765898000
WoS Researcher ID: https://www.webofscience.com/wos/author/record/FDF-0979-2022
Google Scholar ID: https://scholar.google.com/citations?user=YruuMK0AAAAJ
RSCI Author ID: https://elibrary.ru/author_profile.asp?authorid=1253542

Conflict of interest

The author declares no conflict of interest.

Financial disclosure

The research had no sponsorship.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – November 15, 2024
Date of approval – November 25, 2024
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:177.5:316.647.82:004.8:004.652
EDN: https://elibrary.ru/mbwjxf
DOI: https://doi.org/10.21202/jdtl.2025.4

Конституционно-правовой аспект создания больших языковых моделей: проблема цифрового неравенства и языковой дискриминации

Илья Геннадьевич Ильин

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Ключевые слова

большие языковые модели,
генеративный искусственный интеллект,
искусственный интеллект,
конституционные права,
обработка естественного языка,
права человека,
право,
цифровое неравенство,
цифровые технологии,
языковая дискриминация

Аннотация

Цель: исследование влияния цифрового неравенства на реализацию конституционных прав человека, а также выявление рисков языковой дискриминации, связанных с разработкой и использованием больших языковых моделей.

Методы: формально-юридический и сравнительно-правовой методы, а также метод теоретического моделирования. Эти подходы дополняются общенаучными методами познания, что позволяет провести комплексный анализ правовых, технологических и социальных аспектов проблемы.

Результаты: было установлено, что применительно к большим языковым моделям цифровое неравенство возникает из-за неравномерного уровня цифровизации языков и проявляется в ограниченном доступе к технологии обработки естественного языка. В свою очередь, неравный доступ к указанной технологии может негативно влиять на реализацию конституционно гарантированных прав и может быть рассмотрен с точки зрения концепций «равенства» и запрета на дискриминацию. Автор подчеркивает, что неравный доступ к технологиям обработки естественного языка может усугублять существующие социальные и экономические неравенства, создавая новые формы дискриминации.

Научная новизна: заключается в анализе скрытых и косвенных форм дискриминации, которые проявляются в системах искусственного интеллекта, особенно в генеративных моделях. В отличие от прямых форм дискриминации, которые могут быть выявлены в предсказательных алгоритмах, генеративные модели создают более тонкие, но не менее значимые кумулятивные эффекты. Эти эффекты способствуют формированию социальных стереотипов и неравенства в таких областях, как профессиональная деятельность, гендерная и этническая принадлежность. Автор также обращает внимание на то, что с увеличением

автономности искусственного интеллекта традиционные подходы к выявлению дискриминации становятся менее эффективными, что требует разработки новых методов анализа и регулирования.

Практическая значимость: состоит в том в том, что его результаты предоставляют основу для выявления и оценки правовых рисков, связанных с неравным доступом к цифровым продуктам, использующим технологии обработки естественного языка. Это способствует совершенствованию правового регулирования в сфере разработки и использования технологий искусственного интеллекта. Статья предлагает рекомендации для законодателей, регулирующих органов и разработчиков технологий, направленные на минимизацию рисков цифрового неравенства и языковой дискриминации.

Для цитирования

Ильин, И. Г. (2025). Конституционно-правовой аспект создания больших языковых моделей: проблема цифрового неравенства и языковой дискриминации. *Journal of Digital Technologies and Law*, *3*(1), 89–107. https://doi.org/10.21202/jdtl.2025.4

Список литературы

Миронова, М. В. (2019). Становление термина «языковая дискриминация» в современной социолингвистике. В сб. *New Language. New World. New Thinking: сборник материалов II Ежегодной международной научно-практической конференции* (с. 555–558). Москва: Дипломатическая академия Министерства иностранных дел Российской Федерации. https://elibrary.ru/bjegvs

Мушаков, В. Е. (2022). Конституционные права человека в контексте проблемы преодоления цифрового разрыва. *Вестник Санкт-Петербургского университета МВД России*, *1*(93), 69–73. EDN: https://elibrary.ru/elrbud. DOI: https://doi.org/10.35750/2071-8284-2022-1-69-73

Талапина, Э. В. (2022). Обработка данных при помощи искусственного интеллекта и риски дискриминации. Право. *Журнал Высшей школы экономики*, *1*, 4–27. EDN: https://elibrary.ru/pwepsj. DOI: https://doi.org/10.17323/2072-8166.2022.1.4.27

Харитонова, Ю. С., Савина, В. С., Паньини, Ф. (2021). Предвзятость алгоритмов искусственного интеллекта: вопросы этики и права. *Вестник Пермского университета. Юридические науки*, *53*, 488–515. EDN: https://elibrary.ru/eukcpy. DOI: https://doi.org/10.17072/1995-4190-2021-53-488-515

Alarcón, A. A. (2022). The economics of language. In Miquel Àngel Pradilla Cardona (Ed.), *Catalan Sociolinguistics: State of the art and future challenges* (pp. 173–182). https://doi.org/10.1075/ivitra.32.12ala

Baugh, J. (2023). Linguistic profiling across international geopolitical landscapes. *Daedalus*, *152*(3), 167–177. https://doi.org/10.1162/daed_a_02024

Bosco, F., Creemers, N., Ferraris, V., Guagnin, D., & Koops, B. J. (2015). Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. In S. Gutwirth, R. Leenes, P. de Hert (Eds.), *Reforming European data protection law* (Vol. 20, pp. 3–33). https://doi.org/10.1007/978-94-017-9385-8_1

Browning, M., & Arrigo, B. (2021). Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice*, *46*(2), 298–316. https://doi.org/10.1007/s12103-020-09557-x

Chasalow, K., & Levy, K. (2021). Representativeness in statistics, politics, and machine learning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 77–89). https://doi.org/10.1145/3442188.3445872

Chilingaryan, K., Meshkova, I., & Sheremetieva, O. (2020). International legal protection of linguistic minorities. *International Journal of Psychosocial Rehabilitation*, *24*(6), 9750–9758. EDN: https://elibrary.ru/dgcwtx. DOI: https://doi.org/10.37200/IJPR/V24I6/PR26097

Dash, N. S., & Arulmozi, S. (2018). *History, features, and typology of language corpora*. Springer Singapore. https://doi.org/10.1007/978-981-10-7458-5

Devlin, J., Chang, Ming-Wei, Lee, Kenton, & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Doğruöz, A. S., Sitaram, S., & Yong, Z. X. (2023). *Representativeness as a forgotten lesson for multilingual and code-switched data collection and preparation*. arXiv preprint arXiv:2310.20470 (pp. 5751–5767).

Donard, K. (2023). Legal protection of linguistic minority under discrimination: the case of anglophone Cameroon. *International Journal of Business and Technology*, *11*(2), Article 1.

Drożdżowicz, A., & Peled, Y. (2024). The complexities of linguistic discrimination. *Philosophical Psychology*, *37*(6), 1459–1482. https://doi.org/10.1080/09515089.2024.2307993

Engel, C., Linhardt, L., & Schubert, M. (2024). Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism. In *Artificial Intelligence and Law*. https://doi.org/10.1007/s10506-024-09389-8

Glauner, P. (2024). Technical foundations of generative AI models. In *Legal Tech-Zeitschrift für die digitale Anwendung*, *1*, 24–34.

Hacker, P. A (2021). Legal framework for AI training data—from first principles to the Artificial Intelligence Act. *Law, Innovation and Technology*, *13*(2), 257–301. https://doi.org/10.1080/17579961.2021.1977219

Hacker, P., Mittelstadt, B., Zuiderveen Borgesius, F., Wachteret, S. (2024). *Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It*. arXiv preprint arXiv:2407.10329. https://doi.org/10.2139/ssrn.4877398

Ilin, I. (2022). Legal Regime of the Language Resources in the Context of the European Language Technology Development. In Z. Vetulani, P. Paroubek, M. Kubis (Eds.), *Human Language Technology. Challenges for Computer Science and Linguistics. LTC 2019. Lecture Notes in Computer Science* (vol. 13212, pp. 367–376). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-05328-3_24

Ilin, I., & Dedova, M. (2019). Academic Entrepreneurship in the Field of Language Resource Creation and Dissemination. In A. Riviezzo, M. Rosaria Napolitano, & A. Garofano (Eds.), *The ESU 2019 Conference and Doctoral Programme, Naples (Italy), 8–14 September 2019. Electronic Conference Proceedings* (pp. 193–200).

Ilin, I., & Kelli, A. (2024). Natural Language, Legal Hurdles: Navigating the Complexities in Natural Language Processing Development and Application. *Journal of the University of Latvia*. *Law*, *17*, 44–67. https://doi.org/10.22364/jull.17.03

Ilin, I., & Kelli, A. (2019). The use of human voice and speech in language technologies: the EU and Russian intellectual property law perspectives. *Juridical International*, *28*, 17–27. https://doi.org/10.12697/ji.2019.28.03

Ilin, I. (2020). The Voice and Speech Processing within Language Technology Applications: Perspective of the Russian Data Protection Law. *Legal Issues in the Digital Age*, *1*, 99–123. EDN: https://elibrary.ru/axbzzq. DOI: https://doi.org/10.17323/2713-2749.2020.1.99.123

Ilin, I. (2024). Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data. *Legal Issues in the Digital Age*, *2*, 36–56. EDN: https://elibrary.ru/azkzba. DOI: https://doi.org/10.17323/2713-2749.2024.2.36.56

Jiang, X., Yan, L., Vavekanand, R., & Hu, M. (2023). Large Language Models in Healthcare Current Development and Future Directions. *Generative AI Research*, *2*, 12. https://doi.org/10.20944/preprints202407.0923.v1

Kelli, A., Vider, K., Pisuke, H., & Siil, T. (2016). Constitutional values as a basis for the limitation of copyright within the context of digitalisation of the Estonian language. In *Constitutional Values in Contemporary Legal Space* (Vol. II, pp. 126–139).

Li, X., Dou, Zh., Zhou, Yu., & Liu, F. (2024). CorpusLM: Towards a unified language model on corpus for knowledge-intensive tasks. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 26–37). https://doi.org/10.1145/3626772.3657778

Möller, J. T. (2011). Case Law of the UN Human Rights Committee relevant to Members of Minorities and Peoples in the Arctic Region. *The Yearbook of Polar Law Online*, *3*(1), 27–56. https://doi.org/10.1163/22116427-91000054

Monteith, B., & Sung, M. (2023). Unleashing the Economic Potential of Large Language Models: The Case of Chinese Language Efficiency. *TechRxiv*. *June 07*. https://doi.org/10.36227/techrxiv.23291831.v1

Morin, S. L. (2024). AI Discrimination in Hiring. In D. Norman (Ed.), *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology* (pp. 64–74). IGI Global. https://doi.org/10.4018/979-8-3693-1906-2.ch004

Orwat, C. (2024). Algorithmic Discrimination From the Perspective of Human Dignity. *Social Inclusion*, *12*, 1–18. https://doi.org/10.17645/si.7160

Ozkul, D. (2024). Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination. *Social Inclusion*, *12*, 1–3. https://doi.org/10.17645/si.8942

Page, C. (2023). Academic language development and linguistic discrimination: Perspectives from internationally educated students. *Comparative and International Education*, *52*(2), 39–53. https://doi.org/10.5206/cie-eci.v52i2.15000

Rogers, S. E. (2016). Bridging the 21st century digital divide. *TechTrends*, *60*(3), 197–199. https://doi.org/10.1007/s11528-016-0057-0

Sohail, A., & Zhang, L. (2024). *Integrating large language models into the psychological sciences*. https://doi.org/10.1007/s12144-025-07438-2

Solovyev, V. D., & Akhtyamova, S. (2019). Linguistic Big Data: Problem of Purity and Representativeness. In *21st International Conference on Data analytics and management in data intensive domains, DAMDID/RCDL 2019* (pp. 193–204). EDN: https://elibrary.ru/tqmgbu

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. https://doi.org/10.48550/arXiv.1706.03762

Yenduri, G., Ramalingam, M., Chemmalar Selvi, G., Supriya, Y., Srivastava, G., Maddikunta, P. K. R. et al. (2023). Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. In *IEEE Access* (Vol. 12, pp. 54608–54649). https://doi.org/10.1109/access.2024.3389497

Сведения об авторе

Ильин Илья Геннадьевич – магистр права в области информационных технологий, аспирант юридического факультета, Санкт-Петербургский государственный университет
Адрес: 199106, Россия, г. Санкт-Петербург, 22-я линия В.О., 7
E-mail: i.g.ilin@spbu.ru
ORCID ID: https://orcid.org/0000-0003-1076-2765
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57765898000
WoS Researcher ID: https://www.webofscience.com/wos/author/record/FDF-0979-2022
Google Scholar ID: https://scholar.google.com/citations?user=YruuMK0AAAAJ
РИНЦ Author ID: https://elibrary.ru/author_profile.asp?authorid=1253542

Конфликт интересов

Автор сообщает об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.15 / Конституционное (государственное) право
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 15 ноября 2024 г.
Дата одобрения после рецензирования – 25 ноября 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:341:57.087.1:316.642.4
EDN: https://elibrary.ru/joupzt
DOI: https://doi.org/10.21202/jdtl.2025.5

Behavioral Biometrics in the European Union: Legal Challenges and Technological Prospects

Baurzhan Rakhmetov
M. Narikbayev KAZGUU University, Astana, Kazakhstan

Kazbek Khaizabekov
University of Padova, Padova, Italy

Keywords

artificial intelligence,
behavioral biometrics,
data protection,
digital technologies,
European Union,
facial recognition,
law,
legal regulation,
legislation,
privacy

Abstract

Objective: to study the historical development of the European Union legislation on behavioral biometrics; to identify the features of the European approach to the regulation of behavioral biometrics, to assess its advantages and disadvantages.

Methods: general scientific methods of analysis and comparison, with an emphasis on the study of legal texts such as directives, regulations and conventions. To ensure a comprehensive understanding of the issue, the authors also consider the technical aspects of behavioral biometrics, which allows for a comprehensive analysis of both legal norms and the technological processes underlying them.

Results: the research demonstrates that the European Union regulatory legal framework on biometrics does not clearly distinguish between behavioral and physical biometrics technologies. This leads to ambiguity in understanding the risks and opportunities associated with the use of behavioral biometrics. The authors emphasize that the insufficiently specific legislation creates significant difficulties for regulators, technology developers, and end users.

Scientific novelty: the article is the first comprehensive study of the historical development of European Union legislation on behavioral biometrics. The work reveals the key characteristics of the European

✉ Corresponding author

approach, its strengths and weaknesses, and compares it with the United States' regulatory practice. The study reveals the key aspects that require further regulation: from a clear definition of behavioral biometrics to the development of comprehensive mechanisms to ensure transparency and accountability in the use of these technologies. Given that behavioral biometrics is a relatively new and rapidly developing technology, the research is important for understanding current challenges and prospects for its regulation.

Practical significance: the research is multifaceted and relevant for experts in digital technologies: legal scholars, law enforcement officers, legislators, and developers of artificial intelligence and biometrics technologies.

For citation

Rakhmetov, B., & Khaizabekov, K. (2025). Behavioral Biometrics in the European Union: Legal Challenges and Technological Prospects. *Journal of Digital Technologies and Law, 3*(1), 108–124. https://doi.org/10.21202/jdtl.2025.5

Contents

Introduction
1. Evolution of Behavioral Biometrics Regulation in the European Union
2. Specific Features of Behavioral Biometrics Regulation in the European Union
   2.1. Characteristics of the European Union's Approach to Behavioral Biometrics Regulation
   2.2. Advantages and Disadvantages of the European Union's Approach to Behavioral Biometrics Regulation
3. Comparative Analysis of the European Union's Approach to Regulating Behavioral Biometrics and the US Experience
Conclusions
References

Introduction

The regulation of behavioral biometrics in the European Union (EU) highlights the importance of addressing data protection and privacy amid a rapidly evolving technological landscape. Biometrics, which includes physical, physiological, and behavioral characteristics used to identify individuals, has been receiving more attention as privacy concerns grow around the world. The EU has continuously updated its legislation to protect personal data since the adoption of the 1981 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108). Key documents such as the Directive 95/46/EC and the General Data Protection Regulation (GDPR) have helped the EU set a global benchmark for the protection of personal data, particularly in the sensitive area of biometric data. Recent developments in the form of the 2024

European Union Artificial Intelligence Act (EU AI Act) have further expanded the scope of these regulations, particularly regarding the use of behavioral biometrics and artificial intelligence in sensitive areas such as security and privacy.

Despite these advances, the regulation of behavioral biometrics presents various challenges, especially concerning the distinctions between physical and behavioral data. The EU's approach to regulating biometric data has significantly influenced privacy laws globally, yet it has faced criticism for its lack of clarity and specificity in certain areas. This article examines the evolution of EU regulations on behavioral biometrics, analyzing key legislation, its influence on data protection, and existing challenges. It also compares the EU regulatory approach with the United States, where the lack of a national law has resulted in less comprehensive regulation on biometrics.

1. Evolution of Behavioral Biometrics Regulation in the European Union

To get a more comprehensive picture of how biometrics are regulated in the EU, it is essential to understand the context and conditions that contributed to the emergence and implementation of its legal framework (Carrigan & Coglianese, 2011). With the rapid advances in electronic data processing in the 1960s and 1970s, there was a severe necessity to strengthen privacy protection measures, especially in the automatic collection of personal data. This was echoed in the EU and led to the adoption of the Convention 108 and the Directive 95/46/EC (De Hert, 2013). These documents were the first legally binding international normative acts regulating data protection, including biometric data.

Convention 108, signed on January 28, 1981, obliged EU countries to introduce a series of specific changes in their domestic legislation by principles such as fair and lawful collection and automatic processing of data and the presence of concrete, explicit, and legitimate purposes for storing such data; it is not allowed to use data that is inconsistent with these objectives or where the storage process takes more time than needed. These principles also encompass the adequacy, relevance, and not excessiveness of the data. Generally, it is the responsibility of controllers, under the provisions of Convention 108, to manage the processing of personal data[1].

There is now in force a modernized version of the document, known as Convention 108+, Article 6 of which stipulates that the processing of biometric data for personal identification is permitted as long as appropriate guarantees are in place to guard against risks that endanger the individual's interests, rights, and fundamental freedoms, including the risk of discrimination. At the same time, biometric data used for unambiguous identification belong to the category of sensitive data, therefore their processing must

[1] Convention 108 and Protocols: Background. (n.d.). Council of Europe Portal. https://clck.ru/3Ge8xN

be accompanied by specific guarantees. It requires separate or joint consent of the data subject, a law defining the objectives, methods, and specific conditions under which data processing may be utilized, confidentiality, measures based on risk analysis, and security precautions[2].

Directive 95/46/EC, adopted on October 24, 1995, was also an important document in the regulation of biometrics that cannot be omitted. The primary focus was on the safeguarding of the individual's fundamental rights and freedoms during the processing of data within the EU and the free movement of such data. The key responsibility for data protection rested with the supervisory authorities established by each state that adopted Directive 95/46/EC. These are independent bodies that have the power to advise on administrative measures and regulations as well as to initiate legal proceedings if breaches of data protection requirements are found. Although Directive 95/46/EC does not specifically mention the processing of biometric data, Article 29 established a Working Party to provide consultations and opinions on the operation and regulation of biometrics[3]. For example, in 2003, a "Working Document on Biometrics," which examines how the provisions of Directive 95/46/EC apply to the use of biometric technologies, was issued[4]. In 2012, the Working Party also published an opinion on revised guidelines on principles and recommendations for enhancing privacy and data protection in biometric applications[5].

Notwithstanding, today, Directive 95/46/EC is considered no longer in force due to its replacement by the General Data Protection Regulation (GDPR), which has ushered in a new stage of development in the regulation of personal data. GDPR was adopted in 2016 and entered into force in 2018. GDPR, as well as Directive 95/94/EC, applies to all countries in the EU but does not require them to change their domestic laws. All organizations both inside and outside the EU must comply with the GDPR. Meanwhile, the GDPR requires organizations based outside the EU, which provide goods or services that track the behavior and process and store data of EU citizens, to identify their representatives in the EU. In turn, controllers and processors also have certain obligations. Controllers should always remember to follow the steps necessary for effective data protection; they should only process data that is within the scope of their duties and not allow access to it to anyone other than those who are obliged to process it (Nguyen, 2018).

2 Convention 108 +. (2018). Council of Europe. https://clck.ru/3Ge94r

3 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of Such Data. (1995). Official Journal of the European Union. https://clck.ru/3Ge97y

4 Working Document on Biometrics. (2003). The Working Party. https://clck.ru/3Ge9AL

5 Opinion on Developments in Biometric Technologies. (2012). The Working Party. https://clck.ru/3Ge9D4

In the parlance of regulators, the term biometric data first appeared with the introduction of GDPR. Article 4 characterizes biometric data as "personal data resulting from specific technical processing relating to the physical, physiological or behavioral characteristics of a natural person"[6]. It is worth noting that the GDPR foresees a category of special data requiring a higher level of protection, which also comprises biometric data. In accordance with Article 9, the processing of biometric data – the objective of which is to determine identity, health, or sexual life and orientation – is strictly prohibited, except under certain conditions. For example, the explicit consent of the subject allows for circumventing the above prohibition[7] (Meden et al., 2021).

Finally, the EU AI Act, adopted in March 2024, is the most recent and highly relevant regulation to date that also applies to biometrics. Nowadays, AI has a tremendous impact on the advancement of biometrics technologies. In combination with biometrics, AI systems contribute to reducing human error and accelerating decision making (Rawat et al., 2023). Therefore, the EU AI Act incorporates several key considerations designed to regulate biometrics, including behavioral biometrics. Notably, this document covers the following aspects related to biometrics: biometric data, emotion recognition system, biometric categorization system, remote biometric identification system, real-time remote biometric identification system, and post-remote biometric identification system[8]. Among all of them, emotion recognition systems and real-time remote biometric identification systems refer to behavioral biometrics (Xefteris et al., 2016; Alsaadi, 2021; Revett, 2008). For example, an emotion recognition system aims to process characteristics such as gaze tracking, mood, facial movement and expression, gait, and heartbeat. In this regard, the EU AI Act imposes a ban on the use of emotion recognition technologies in the workplace and schools, on predictive policing if it is based on human profiling and personal characteristics assessment, and on AI that involves manipulation of people's behavior or vulnerabilities. As for real-time remote biometric identification systems, this technology can be used subject to strict safeguards and limitations[9].

---

6 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). (2016). Official Journal of the European Union. https://clck.ru/3Ge9Gn

7 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). (2016). Official Journal of the European Union. https://clck.ru/3Ge9Gn

8 Santalu, N. (2023). Biometrics Under the EU AI Act. The International Association of Privacy Professionals. https://clck.ru/3Ge9Jz

9 Holistic AI Team. (2024). Prohibited AI Practices Under the EU AI Act. https://clck.ru/3Ge9Lf

2. Specific Features of Behavioral Biometrics Regulation in the European Union

2.1. Characteristics of the European Union's Approach to Behavioral Biometrics Regulation

The approach that characterizes the regulation of biometrics in the EU can be considered as risk-oriented. Behavioral biometrics legislation is accompanied by strong restrictive measures intended to protect privacy and civil liberties and to combat bias and discriminatory technologies. Collecting, processing, and storing behavioral data is a riskier process, especially in comparison with other types of personal data (Rezaee, 2025). The point is that the accuracy of analyzing such data does not allow to fully determine a person's identity but only reveals specific patterns related to his or her character and habits. It is important to consider that factors such as high stress levels or physical condition do not make it possible for behavioral biometrics to accurately capture a person's behavior. In turn, more accurate profiling of characteristic behavior requires the collection of a significant amount of behavioral data. Furthermore, since behavioral data collection is always ongoing, it necessitates the storage of significant amounts of such information, which also poses additional risks to data privacy[10] (Sharma & Elmiligi, 2022).

Behavioral biometrics are a relatively emerging technology that is actively gaining momentum today but still accompanied by certain challenges and risks. To mitigate them, the GDPR makes it compulsory to obtain data subjects' consent to the handling and gathering of their biometric data, including behavioral ones. Previously, the GDPR had already classified biometric data as sensitive. However, the recently passed EU AI Act has expanded the categorization system by adding risk levels such as unacceptable risk, high risk, limited risk, low or minimal risk. To determine the level of risk, it is essential to ascertain the nature and the extent of AI application (Arcila, 2024). The category of unacceptable risk that prohibits use includes AI systems that imply social scoring based on behavior or personal traits and manipulation of people's behavior or vulnerabilities. The use of real-time remote biometric identification is also not allowed, unless this technology can contribute to locating missing individuals, preventing life-threatening situations, including a foreseeable terrorist attack, and identifying criminal suspects. The emotion recognition system falls into the limited risk category, but its application is not allowed in educational institutions or the workplace except for medical or safety reasons[11].

---

[10] Makhani, F. (2022). Beyond Fingerprints: Exploring Behavioral Biometrics For Secure Identity Verification. VikingCloud. https://clck.ru/3Ge9SS

[11] High-Level Summary of the AI Act. (2024). Future of Life Institute. https://clck.ru/3Ge9UK

2.2. Advantages and Disadvantages of the European Union's Approach to Behavioral Biometrics Regulation

One of the advantages of the EU's approach is that its regulatory experience has significantly influenced other countries that are creating and developing their own data protection and biometrics legislation. Greenleaf (2012) examined 39 countries outside Europe and found that there was a wide range of specific similarities with Convention 108 in the legislation of 33 countries. Some of the reasons for this phenomenon include the fact that countries are thereby demonstrating their commitment to become part of European privacy laws. Overall, Argentina, Cabo Verde, Mauritius, Mexico, Morocco, Senegal, Tunisia, and Uruguay decided to accede to Convention 108[12].

Convention 108+, which replaced Convention 108 in 2018, has also succeeded in becoming an influential benchmark in global data protection regulation practices. In addition to EU member states, the updated protocol was signed by the United Kingdom (then a member of the EU), Uruguay, Cabo Verde, Mauritius, Mexico, Senegal, and Tunisia. Furthermore, Argentina, Burkina Faso, and Morocco were also welcomed to the Convention 108+[13]. Nonetheless, it is the GDPR that has served as the primary benchmark for data protection regulation around the world. As of 2020, countries such as Brazil, Canada, and South Korea have enacted laws similar to the GDPR (Chen et al., 2022). Many African countries including Tanzania, Eswatini, Rwanda, Uganda, and Nigeria have established several new data protection regulations that contain common principles with the GDPR[14]. It is worth noting that not only countries but also organizations around the world have been impacted by GDPR requirements. Since the European regulation mandates strict data protection safeguards, organizations whose activities extend to European citizens have had to make significant changes to comply with GDPR (Li et al., 2019; Chen et al., 2022).

The EU's approach to regulating biometrics, including behavioral biometrics, is distinguished by a robust degree of data protection and privacy. According to Article 9 of the GDPR, biometric data is classified as sensitive data requiring special protection and privacy requirements. This implies that, in general, the processing of such data is only permitted under strict compliance with certain conditions. For instance, it is imperative to acquire the explicit consent of the data subject – the individual whose data is being utilized. Furthermore, the GDPR has given data subjects the right to examine information held by organizations, and to withdraw consent for data collected by organizations. The obligations of an organization collecting and processing data from European citizens include expressing an interest in collecting personal information, justifying the reason

12 Chart of Signatures and Ratifications of Treaty 108. (n.d.). Council of Europe. https://clck.ru/3Ge9a5

13 Baker, J. (2018). What Does the Newly Signed 'Convention 108+' Mean for UK Adequacy? The International Association of Privacy Professionals. https://clck.ru/3Ge9ch

14 Wu, J., & Hayward, M. (2023). International Impact of the GDPR Felt Five Years on. Pinsent Masons. https://clck.ru/3Ge9gi

to possess this information, and presenting their identity to data subjects. Overall, the GDPR requires organizations to limit data processing, as well as possession and transfer of data between platforms, providing appropriate means of protecting and disposing of data after a set period. It is becoming clear that the GDPR approach can be considered user-centered, which has a positive effect on individual responsibility, reducing security risks and increasing privacy measures (Aseri, 2020).

There are severe sanctions for violating the above biometric data policies. Generally, there are two levels of administrative fines for non-compliance with the GDPR: 1) up to 10 million euros or 2 % of annual global turnover, whichever is greater; 2) up to 20 million euros or 4 % of annual global turnover, whichever is greater. The amount of the fine is determined depending on the specific provisions of the GDPR; it will be less if data security is breached and more if people's privacy rights are violated. For example, under the application of the GDPR, Meta[15] was fined €1.2 billion by the Irish Data Protection Commission in 2023 for sending European users' personal information to the US without proper data protection mechanisms. Before that, companies such as Amazon, TikTok, WhatsApp, Google, and others were also sanctioned[16].

The most significant drawback of the European approach is the failure to categorize and be specific in some aspects. In particular, European regulators do not consider the point that different types of biometrics use different types of data; only behavioral biometrics collects data such as keystroke dynamics, mouse movements, touchscreen inputs, eye movements, gesture, and gait (Eberz et al., 2017; Cheung & Vhaduri, 2020). Instead, the European legal framework contains only generic interpretations and guidelines pertaining to physical, psychological, and behavioral features. This shortcoming can be traced both in the past, meaning Convention 108 and Directive 95/46/EC, and up to the present, referring to Convention 108+, GDPR, and EU AI Act. The fact is that the dynamic nature of behavioral data, which does not allow it to be forecasted, modeled, or fabricated as easily, makes it non-adaptable and unsuitable for current EU regulations that cover only physical biometrics. Companies and financial institutions located in the EU, which have already started a consistent implementation of behavioral biometrics, are still guided by regulations for collecting people's physical data[17] (Kindt, 2018).

The problem of vagueness is also evident in other important provisions of regulations concerning the use of biometrics. For example, the GDPR does not make a significant distinction between the primary comparative functions of biometric technologies, specifically between 'verification' and 'identification.' Verification involves the use

[15] The organization is recognized as extremist, its functioning is prohibited in the territory of the Russian Federation.

[16] 20 Biggest GDPR Fines So Far. (2024). Data Privacy Manager. https://clck.ru/3Ge9me

[17] Özal, M. (2020). 'Behavioral Biometrics': A Brief Introduction from the Perspective of Data Protection Law. CiTiP Blog. https://clck.ru/3Ge9pA

of biometric data on a one-to-one (1:1) basis, while 'identification' involves a one-to-many (1:n) basis. It is worth noting that, according to the Council of Europe and national data protection supervisory authorities, the verification function is more secure than the identification method because it does not involve a database. In contrast, the use of biometric identification requires extensive collection and storage of biometric information in databases. Also, it should be noted that biometric identification introduces further risks due to probability-based matching, which adversely affects the level of accuracy. Accordingly, European regulators should objectively consider the relative risks of both verification and identification while introducing appropriate rules for the application of the two main functions of biometric technologies[18].

3. Comparative Analysis of the European Union's Approach to Regulating Behavioral Biometrics and the US Experience

To better comprehend the regulatory landscape for behavioral biometrics in the EU, it may be useful to examine the experience of the United States and compare it with the EU's approach. Most notably, the United States differs from the EU in the way that, instead of unified national legislation as in the EU, biometrics are regulated at the state level[19].

Numerous states, including Illinois, Texas, and Arkansas, have various laws aimed at regulating biometrics and biometric data. To begin with, it is worth noting Illinois, which introduced the Biometric Information Privacy Act (BIPA) in 2008, becoming the very first state to start regulating the collection, use, and storage of biometric data. Under BIPA, companies are obliged to acquire written consent from data subjects prior to collecting their biometric information and to limit scanning methods such as retinal or iris scanning, fingerprinting, voiceprints, or facial and hand geometry scanning. In other words, other biological and behavioral data are not considered biometric identifiers under this law (Illman, 2017). Texas provides similar definitions of biometric identifiers related to biometrics in its 2009 law under Section 503.001[20].

The State of Arkansas has enacted the Biometric Data Act, which focuses solely on biological parameters, thereby excluding behavioral data. This regulation defines biometric data as information about a person's biological characteristics, such as fingerprints, facial or eye scans, DNA, and other unique biological features utilized for identification purposes[21].

---

18 Kindt, E. (2020). A First Attempt at Regulating Biometric Data in the European Union. AI Now Institute. https://clck.ru/3Ge9ti

19 Biometric Data Protection (Privacy – EU, UK and US). (2021). https://clck.ru/3Ge9w2

20 2023 Texas Statutes Business and Commerce Code. https://clck.ru/3GePrv

21 Arkansas Personal Information Protection Act. https://clck.ru/3GeRBo

Among the various interpretations of biometric data in different United States laws, the California Consumer Privacy Act (CCPA) of 2018 notably broadens the understanding of biometric data. Under the CCPA, biometric information comprises "physiological, biological, or behavioral characteristics of an individual, including DNA, that can be used alone or in combination with other data to establish individual identity."[22] This encompasses not only traditional biometric identifiers but also behavioral patterns such as keystrokes, gait rhythms, sleep habits, health, or exercise data that can identify a person[23]. Distinctively, under the CCPA, citizens of the State have greater visibility and control over their biometric data, including the rights to general disclosure, requests for information, deletion of information, and "equal service and prices" (Ghelardi, 2020).

Additionally, The American Privacy Rights Act of 2024, successor to the American Data Privacy and Protection Act of 2021, aims to establish clear national data rights and protections. The legislation was introduced by lawmakers in both the House and Senate in April 2024, and was approved by the Subcommittee on Data, Innovation and Commerce a month later. Now, it will have to pass through a full committee and both houses of Congress prior to potentially gaining enactment into law. This legislation defines biometric information as data derived from the technological processing of unique biological, physical, or physiological characteristics, including fingerprints, facial scans, and gait, among others. Importantly, the bill was modeled on the GDPR operating within the EU[24].

In comparison with the EU legislation, the American one is less elaborated. Laws are not adopted at the federal level, but at the state level, which indicates the need for a more comprehensive approach (Neace, 2020). It also becomes apparent that these regulations, like those in the EU, do not clearly distinguish between physical and behavioral biometrics. While some laws mention behavioral characteristics, there is still no explicit legislative definition or regulation of behavioral biometrics within these frameworks. Consequently, issues related to behavioral biometrics remain inadequately addressed, requiring further legislative attention and development. Moreover, the need to focus closely on behavioral biometrics, including characteristics such as hand movements and gaze direction, has been clearly articulated by the executive branch under President Joe Biden's Executive Order on Artificial Intelligence[25].

---

22 California Consumer Privacy Act. https://clck.ru/3GeRFp

23 What is the California Consumer Privacy Act (CCPA)? (2024). TermsFeed. https://clck.ru/3GeAJh

24 Wright, V. (2024). The American Privacy Rights Act (APRA): Everything You Need to Know. BigID. https://clck.ru/3GeALE ; Pınarbaşı, A. T. (2024). The American Privacy Rights Act (APRA): Everything You Need to Know. Didomi. https://clck.ru/3GeANc

25 Brunetti, F. (2024). Behavioral Characteristics as a Biometric: Something to Keep an Eye (Scan) on. The International Association of Privacy Professionals. https://clck.ru/3GeATg

Conclusions

To sum up, the regulation of behavioral biometrics within the EU has evolved significantly, shaped by key legislative frameworks such as Convention 108, Directive 95/46/EC, and most recently the 2018 General Data Protection Regulation and the 2024 EU AI Act. These regulations have established a solid foundation for protecting personal data, especially biometric data categorized as sensitive information. The introduction of the GDPR's definition of biometric data and its strict rules for processing has set a global standard, influencing not only European countries but also legal practices across the world. However, certain challenges remain, for example, in distinguishing between physical and behavioral biometrics and in addressing the complexities of biometric technologies like verification and identification.

Comparing the EU's approach to the United States' approach highlights the EU's more comprehensive and unified regulatory framework, in contrast to the fragmented state-level laws in the US. Although the US has made progress through state regulations such as Illinois' Biometric Information Privacy Act (BIPA) and the more recent American Privacy Rights Act, the lack of the unified national approach raises concerns. For instance, behavioral biometrics are still inadequately addressed in the US. As technologies such as AI continue to evolve and become increasingly interconnected with biometric technologies, it is important to highlight that the EU and the US should continue to strengthen their regulations to safeguard personal data and promote the ethical use of biometrics. Yet given the novelty of behavior biometrics, further research of legal regulation of personal data is required.

References

Alsaadi, E. (2021). *Study on Most Popular Behavioral Biometrics, Advantages, Disadvantages and Recent Applications: A Review*. https://doi.org/10.13140/RG.2.2.28802.09926

Arcila, B. B. (2024). AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight? *Computer Law & Security Review*, *54*, 106012. https://doi.org/10.1016/j.clsr.2024.106012

Aseri, A. (2020). The Implication of the European Union's General Data Protection Regulation (GDPR) on the Global Data Privacy. *Journal of Theoretical and Applied Information Technology*, *98*(4), 692–702.

Carrigan, C., & Coglianese, C. (2011). The Politics of Regulation: From New Institutionalism to New Governance. *Annual Review of Political Science*, *14*(1), 107–129. https://doi.org/10.1146/annurev.polisci.032408.171344

Chen, C., Frey, C. B., & Presidente, G. (2022). Privacy Regulation and Firm Performance: Estimating the GDPR Effect Globally. *Oxford Martin School*.

Cheung, W., & Vhaduri, S. (2020). Continuous Authentication of Wearable Device Users from Heart Rate, Gait, and Breathing Data. *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 587–592. https://doi.org/10.1109/BioRob49111.2020.9224356

De Hert, P. (2013). Biometrics and the Challenge to Human Rights in Europe. Need for Regulation and Regulatory Distinctions. In: P. Campisi (Eds.), *Security and Privacy in Biometrics* (pp. 369–413). Springer London. https://doi.org/10.1007/978-1-4471-5230-9_15

Eberz, S., Rasmussen, K. B., Lenders, V., & Martinovic, I. (2017). Evaluating Behavioral Biometrics for Continuous Authentication: Challenges and Metrics. *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 386–399. https://doi.org/10.1145/3052973.3053032

Ghelardi, E.-M. (2020). Closing the Data Gap: Protecting Biometric Information Under the Biometric Information Privacy Act and the California Consumer Protection Act. *St. John's Law Review*, *94*(3).

Greenleaf, G. (2012). The Influence of European Data Privacy Standards Outside Europe: Implications for Globalization of Convention 108. *International Data Privacy Law*, *2*(2), 68–92. https://doi.org/10.1093/idpl/ips006

Illman, E. J. (2017). Data Privacy Laws Targeting Biometric and Geolocation Technologies. *The Business Lawyer*, *73*(1), 191–198.

Kindt, E. (2018). Having Yes, Using No? About the New Legal Regime for Biometric Data. *Computer Law & Security Review*, *34*(3), 523–538. https://doi.org/10.1016/j.clsr.2017.11.004

Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of Global Information Technology Management*, *22*(1), 1–6. https://doi.org/10.1080/1097198X.2019.1569186

Meden, B., Rot, P., Terhorst, P., Damer, N., Kuijper, A., Scheirer, W. J., Ross, A., Peer, P., & Struc, V. (2021). Privacy–Enhancing Face Biometrics: A Comprehensive Survey. *IEEE Transactions on Information Forensics and Security*, *16*, 4147–4183. https://doi.org/10.1109/TIFS.2021.3096024

Neace, G. (2020). Biometric Privacy: Blending Employment Law with the Growth of Technology. *UIC Law Review*, *53*(1).

Nguyen, F. Q. (2018). The Standard for Biometric Data Protection. *Journal of Law & Cyber Warfare*, *7*(1), 61–84.

Rawat, Y., Gupta, Y., Khothari, G., Mittal, A., & Rautela, D. (2023). The Role of Artificial Intelligence in Biometrics. *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, 622–626. https://doi.org/10.1109/ICECAA58104.2023.10212224

Revett, K. (2008). *Behavioral Biometrics: A Remote Access Approach*. Wiley.

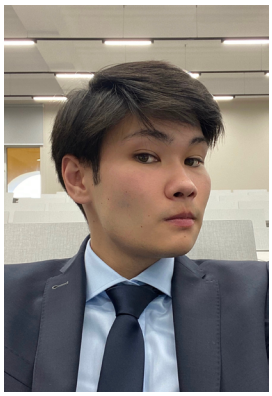
Rezaee, K. (2025). Machine Learning and Facial Recognition for Down Syndrome Detection: A Comprehensive review. *Computers in Human Behavior Reports*, *17*, 100600. https://doi.org/10.1016/j.chbr.2025.100600

Sharma, M., & Elmiligi, H. (2022). Behavioral Biometrics: Past, Present and Future. *Recent Advances in Biometrics*. IntechOpen. https://doi.org/10.5772/intechopen.102841

Xefteris, S., Doulamis, N., Andronikou, V., Varvarigou, T., & Cambourakis, G. (2016). Behavioral Biometrics in Assisted Living: A Methodology for Emotion Recognition. *Engineering, Technology & Applied Science Research*, *6*(4), 1035–1044. https://doi.org/10.48084/etasr.634

Authors information

Baurzhan Rakhmetov – PhD (Politics and International Relations), Assistant Professor, International School of Economics, M. Narikbayev KAZGUU University
Address: 8 Korgalzhyn street, 010000, Astana, Kazakhstan
E-mail: b_rakhmetov@kazguu.kz
ORCID ID: https://orcid.org/0000-0003-3948-9977
WoS Researcher ID: https://www.webofscience.com/wos/author/record/32537389
Google Scholar ID: https://scholar.google.com/citations?user=hanfRwoAAAAJ



Kazbek Khaizabekov – Master's Student, European and Global Studies, Department of Political Science, Law, and International Studies, University of Padova
Address: Via VIII Febbraio, 2, 35122 Padova PD, Italy
E-mail: khaizabekovk@gmail.com
ORCID ID: https://orcid.org/0009-0009-8241-8016
WoS Researcher ID: https://www.webofscience.com/wos/author/record/MIN-4357-2025

Authors' contributions

The authors have contributed equally into the concept and methodology elaboration, validation, formal analysis, research, selection of sources, text writing and editing, project guidance and management.

Conflict of interest

Baurzhan Rakhmetov is a member of the Editorial Board of the Journal; the article has been reviewed on general terms.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – September 19, 2024
Date of approval – October 23, 2024
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:341:57.087.1:316.642.4
EDN: https://elibrary.ru/joupzt
DOI: https://doi.org/10.21202/jdtl.2025.5

Поведенческая биометрия в Европейском Союзе: правовые вызовы и технологические перспективы

Бауржан Рахметов ✉

Университет КАЗГЮУ имени М. С. Нарикбаева, Астана, Казахстан

Казбек Хайзабеков

Падуанский университет, Падуя, Италия

Ключевые слова

Европейский Союз, законодательство, защита данных, искусственный интеллект, конфиденциальность, поведенческая биометрия, право, правовое регулирование, распознавание лиц, цифровые технологии

Аннотация

Цель: изучение исторического развития законодательства Европейского союза в области поведенческой биометрии, выявление особенностей европейского подхода к регулированию поведенческой биометрии, а также оценка его преимуществ и недостатков.

Методы: общенаучные методы анализа и сравнения с акцентом на изучение юридических текстов, таких как директивы, регламенты и конвенции. Для обеспечения всестороннего понимания проблемы авторы также рассматривают технические аспекты поведенческой биометрии, что позволяет провести комплексный анализ как правовых норм, так и технологических процессов, лежащих в их основе.

Результаты: нормативная правовая база Европейского союза в области биометрии недостаточно четко разграничивает технологии поведенческой и физической биометрии. Это приводит к неоднозначности в понимании рисков и возможностей, связанных с использованием поведенческой биометрии. Авторы подчеркивают, что отсутствие конкретики в законодательстве создает значительные трудности для регулирующих органов, разработчиков технологий и конечных пользователей.

Научная новизна: заключается в том, что она представляет собой первое комплексное исследование исторического развития законодательства Европейского союза в области поведенческой биометрии. В статье раскрываются ключевые характеристики европейского подхода, его сильные и слабые стороны, а также проводится сравнительный анализ с опытом регулирования в Соединенных Штатах. В исследовании детально раскрываются ключевые аспекты, требующие дальнейшего законодательного урегулирования: от четкой дефиниции

✉ Контактное лицо

поведенческой биометрии до разработки комплексных механизмов обеспечения прозрачности и подотчетности при использовании данных технологий. Учитывая, что поведенческая биометрия является относительно новой и быстро развивающейся технологией, такое исследование имеет важное значение для понимания современных вызовов и перспектив ее регулирования.

Практическая значимость: определяется его многогранным характером и актуальностью для широкого круга специалистов в сфере цифровых технологий: от ученых-правоведов, правоприменителей и законодателей до разработчиков технологий искусственного интеллекта и биометрии.

Для цитирования

Рахметов, Б., Хайзабеков, К. (2025). Поведенческая биометрия в Европейском союзе: правовые вызовы и технологические перспективы. *Journal of Digital Technologies and Law*, *3*(1), 108–124. https://doi.org/10.21202/jdtl.2025.5

Список литературы

Alsaadi, E. (2021). *Study on Most Popular Behavioral Biometrics, Advantages, Disadvantages and Recent Applications: A Review*. https://doi.org/10.13140/RG.2.2.28802.09926

Arcila, B. B. (2024). AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight? *Computer Law & Security Review*, *54*, 106012. https://doi.org/10.1016/j.clsr.2024.106012

Aseri, A. (2020). The Implication of the European Union's General Data Protection Regulation (GDPR) on the Global Data Privacy. *Journal of Theoretical and Applied Information Technology*, *98*(4), 692–702.

Carrigan, C., & Coglianese, C. (2011). The Politics of Regulation: From New Institutionalism to New Governance. *Annual Review of Political Science*, *14*(1), 107–129. https://doi.org/10.1146/annurev.polisci.032408.171344

Chen, C., Frey, C. B., & Presidente, G. (2022). Privacy Regulation and Firm Performance: Estimating the GDPR Effect Globally. *Oxford Martin School*.

Cheung, W., & Vhaduri, S. (2020). Continuous Authentication of Wearable Device Users from Heart Rate, Gait, and Breathing Data. *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 587–592. https://doi.org/10.1109/BioRob49111.2020.9224356

De Hert, P. (2013). Biometrics and the Challenge to Human Rights in Europe. Need for Regulation and Regulatory Distinctions. In: P. Campisi (Eds.), *Security and Privacy in Biometrics* (pp. 369–413). Springer London. https://doi.org/10.1007/978-1-4471-5230-9_15

Eberz, S., Rasmussen, K. B., Lenders, V., & Martinovic, I. (2017). Evaluating Behavioral Biometrics for Continuous Authentication: Challenges and Metrics. *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 386–399. https://doi.org/10.1145/3052973.3053032

Ghelardi, E.-M. (2020). Closing the Data Gap: Protecting Biometric Information Under the Biometric Information Privacy Act and the California Consumer Protection Act. *St. John's Law Review*, *94*(3).

Greenleaf, G. (2012). The Influence of European Data Privacy Standards Outside Europe: Implications for Globalization of Convention 108. *International Data Privacy Law*, *2*(2), 68–92. https://doi.org/10.1093/idpl/ips006

Illman, E. J. (2017). Data Privacy Laws Targeting Biometric and Geolocation Technologies. *The Business Lawyer*, *73*(1), 191–198.

Kindt, E. (2018). Having Yes, Using No? About the New Legal Regime for Biometric Data. *Computer Law & Security Review*, *34*(3), 523–538. https://doi.org/10.1016/j.clsr.2017.11.004

Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of Global Information Technology Management*, *22*(1), 1–6. https://doi.org/10.1080/1097198X.2019.1569186

Meden, B., Rot, P., Terhorst, P., Damer, N., Kuijper, A., Scheirer, W. J., Ross, A., Peer, P., & Struc, V. (2021). Privacy–Enhancing Face Biometrics: A Comprehensive Survey. *IEEE Transactions on Information Forensics and Security*, *16*, 4147–4183. https://doi.org/10.1109/TIFS.2021.3096024

Neace, G. (2020). Biometric Privacy: Blending Employment Law with the Growth of Technology. *UIC Law Review*, *53*(1).

Nguyen, F. Q. (2018). The Standard for Biometric Data Protection. *Journal of Law & Cyber Warfare*, *7*(1), 61–84.

Rawat, Y., Gupta, Y., Khothari, G., Mittal, A., & Rautela, D. (2023). The Role of Artificial Intelligence in Biometrics. *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, 622–626. https://doi.org/10.1109/ICECAA58104.2023.10212224

Revett, K. (2008). *Behavioral Biometrics: A Remote Access Approach*. Wiley.

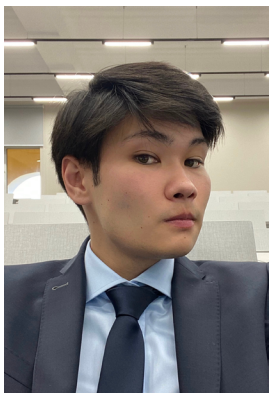
Rezaee, K. (2025). Machine Learning and Facial Recognition for Down Syndrome Detection: A Comprehensive review. *Computers in Human Behavior Reports*, *17*, 100600. https://doi.org/10.1016/j.chbr.2025.100600

Sharma, M., & Elmiligi, H. (2022). Behavioral Biometrics: Past, Present and Future. *Recent Advances in Biometrics*. IntechOpen. https://doi.org/10.5772/intechopen.102841

Xefteris, S., Doulamis, N., Andronikou, V., Varvarigou, T., & Cambourakis, G. (2016). Behavioral Biometrics in Assisted Living: A Methodology for Emotion Recognition. *Engineering, Technology & Applied Science Research*, *6*(4), 1035–1044. https://doi.org/10.48084/etasr.634

Сведения об авторах

Рахметов Бауржан – PhD в области политологии и международных отношений, ассистент-профессор, Международная школа экономики, Университет КАЗГЮУ имени М.С. Нарикбаева
Адрес: Казахстан, 010000, г. Астана, Коргалжинское шоссе, 8
E-mail: b_rakhmetov@kazguu.kz
ORCID ID: https://orcid.org/0000-0003-3948-9977
WoS Researcher ID: https://www.webofscience.com/wos/author/record/32537389
Google Scholar ID: https://scholar.google.com/citations?user=hanfRwoAAAAJ



Хайзабеков Казбек – магистрант в области европеистики и глобальных исследований, кафедра политологии, права и международных исследований, Падуанский университет
Адрес: Италия, 35122, г. Падуя, ул. 8 февраля, 2
E-mail: khaizabekovk@gmail.com
ORCID ID: https://orcid.org/0009-0009-8241-8016
WoS Researcher ID: https://www.webofscience.com/wos/author/record/MIN-4357-2025

Вклад авторов

Авторы внесли равный вклад в разработку концепции, методологии, валидацию, формальный анализ, проведение исследования, подбор источников, написание и редактирование текста, руководство и управление проектом.

Конфликт интересов

Бауржан Рахметов является членом редакционной коллегии данного журнала; статья прошла рецензирование на общих условиях.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.87 / Международное право
Специальность ВАК: 5.1.5 / Международно-правовые науки

История статьи

Дата поступления – 19 сентября 2024 г.
Дата одобрения после рецензирования – 23 октября 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:34.05:004.056.5
EDN: https://elibrary.ru/xujvtm
DOI: https://doi.org/10.21202/jdtl.2025.6

Universal Information Security Governance System: Organizational and Legal Principles

Musadaq Ahmed Hadi ✉
University of Technology, Baghdad, Iraq

Mohammed Najm Abdulredha
University of Baghdad, Baghdad, Iraq

Keywords

cybersecurity,
digital technologies,
information protection,
information security governance,
information security,
information technologies,
law,
legal regulation,
legislation,
organizational structure

Abstract

Objective: to develop universal organizational and legal principles for building an information security governance system that will allow each organization to create its own effective information security governance system, taking into account its unique business goals and tasks.

Methods: the research integrates the key elements of information security governance, such as vision, strategy, goals, policies, standards, processes, and matrices. Vision and goals set the direction of an organization's development; policies and standards provide a conceptual framework for information protection; processes allow for systematic achievement of objectives; and matrices provide tools for evaluating and monitoring the entire structure. The proposed principles are consistent with international standards, regulatory requirements, and best practices in the field of information security.

Results: the research showed that the developed information security governance system allows for a clear distribution of roles and responsibilities among the employees, ensuring effective implementation of the governance system. The authors also analyzed the existing principles of information security, integrating them into a security strategy that meets the corporate goals. The proposed universal system complies with regulatory legal requirements and can be adapted for organizations of any scale and profile.

✉ Corresponding author

Scientific novelty: the paper represents a practical approach to the implementation of an information security governance system based on the authors' experience, international standards, control systems and legal acts. Unlike existing approaches, the proposed system is flexible and can be adapted to any organization, which makes it a universal tool for information security governance.

Practical significance: the research provides a structured approach to creating a universal information security governance system that can be used by organizations lacking knowledge and resources to implement such initiatives. The authors propose a general structure that can be adapted depending on the organization's assets, the employees' training and awareness of information security issues. This makes the paper a valuable resource for professionals seeking to increase information security in their organizations.

For citation

Hadi, M. A., & Abdulredha, M. N. (2025). Universal Information Security Governance System: Organizational and Legal Principles. *Journal of Digital Technologies and Law, 3*(1), 125–142. https://doi.org/10.21202/jdtl.2025.6

Contents

Introduction
1. Why do organizations should have ISG framework? What is the aim of it?
2. Job titles, roles and responsibilities in organizations
   2.1. Standard method of governing organizations
   2.2. Standards and security frameworks of organizations
3. Top ISG/Cybersecurity frameworks
4. A proposed information security governance framework
Conclusions
References

Introduction

Historically information security (infosec) was started when ancient Egyptians, Greeks and Romans were practicing techniques to secure their messages such as Cryptography. One of the first and most famous people to secure message communications was Julius Caesar. He was invented and used the Caesar cipher to secure his private communications for military purposes. After that, there were many contributions were made to confront this challenge and it became more necessary by agencies which a major portion of their duty is to guarantee infosec. Afterwards, many techniques were invented in the middle ages such as steganography which hides data throughout date as a part of security through obscurity (Rao & Nayak, 2014; Hadi et al., 2023; Wu et al., 2021).

In 1889, British government enacted the Official Secrets Act by created a framework and codified classification schemes to secure and control sensitive data. Moreover, Cyber

schools and security Government Codes were established as mandatory need in 1919. These codes were then applied and put into implementation in World War I to secure sensitive data communications. By this time, a lot of securing methods and algorithms were invented such as classification algorithms, Cryptography algorithms code-breaking algorithms were (Ohki et al., 2009).

In World War II, one of the most important information security devices were designed and developed by German called Enigma Machine. It was electro-mechanical device that used for encryption and decryption messages coding warfare. After that, a mathematician and cryptanalyst Alan Turing who was working in British Government Code and Cipher School gained notoriety to solve the mystery of the German code and decipher it. In this era, many technological advancements in infosec, securing communication, encryption, and computer science were developed to made it easier to share an information and sensitive data (Rastogi & von Solms, 2005).

Between 1960 and 1990, infosec was significantly developed as digital electronic and information technology advanced. During this period, first mainframe computers was invented also, time-sharing systems became more important with more data protection mechanisms access development. Furthermore, ARPANET networked systems development with the design of Data Encryption Standard (DES) which was based on symmetric-key encryption standard. Afterwards, Local Area Networks (LANs) and Personal Computers (PCs) came to the scene and organizations started to implement Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS) and Firewalls to protect their information all over networks. By the end of the 1980s, infosec became an important aspect of computer and network operations which laid the groundwork to the era of Cybersecurity[1] (Bendovschi, 2015; Johnston & Hale, 2009).

From 1990 to 2024, infosec was rapidly advanced and due to the expansion of security concepts and security appliances. The main reasons behind that the invention of the internet in the 1990s that created real challenges in the world of infosec such as converting normal protocols into securing protocols (HTTP to HTTPS) by adding security protocols (SSL). Moreover, new technologies and other challenges were discovered such as IoT, blockchain, cloud-computing, quantum-computers and Artificial Intelligence these created many attack-vectors by involving cyberattacks which made infosec of an organization harder to be achieved. Eventually, the mandatory need to assign and design an ISG to every organization is essential to secure organization data (Corriss, 2010; Moulton & Coles, 2003; AlGhamdi et al., 2020).

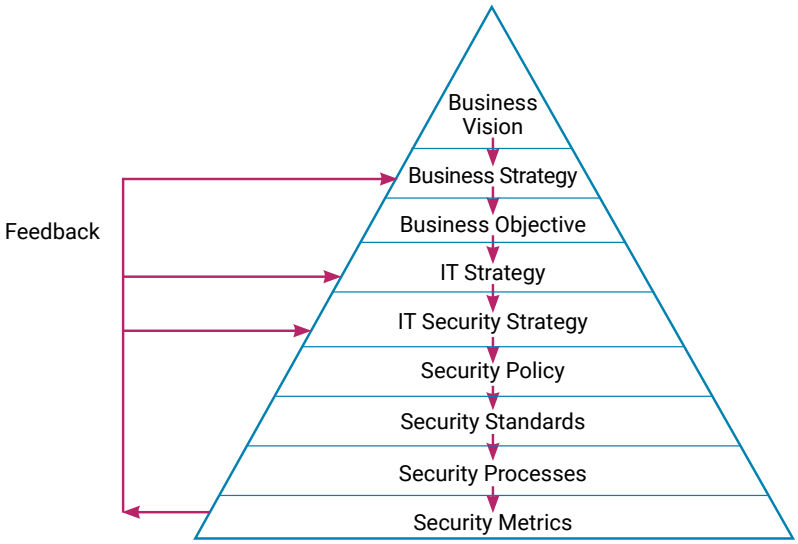
Nowadays, the activities that represent the main focus of the infosec is the management within all assets such as people, risk, incident, vulnerability and also business continuity plan. On the other hand, metrics and other instruments are are used to measure and evaluate in order to achieve an effective ISG program by monitoring

[1] Gregory, P. H. (2018). CISM®: Certified Information Security Manager Exam Guide. New York: McGraw Hill Education.

process and procedures od organization improvement. However, addressing infosec as a critical business challenge is necessary because the lack in securing data leads to serious issue in organizations which indicates the lack of commitment and knowledge of senior executives and boards of directors. Moreover, public and private organizations have challenges to handle infosec at board-level due to lack of knowledge or cybersecurity skills[2] (Carcary et al., 2016; Rocha Flores et al., 2014).

First of all, employees, seniors even board members should appreciate the importance of having ISG, otherwise it cannot be applied effectively. Second, a solid IT governance program must exist in order for ISG to be implemented successfully. A good ISG framework need to be supported by a good IT framework and they should be integrated together in order to achieve organization objectives and goals. Strategically, IT framework contributes to the overall performance of the organization by supporting operational efficiency. An organized method for overseeing IT processes and assets is offered by ISG architecture. Eventually, to put it simply, the cooperation between ISG and IT governance is necessary to guarantee that infosec of public and private organizations are not only put into its place but also align with their overall business aims (S. H. von Solms & R. von Solms, 2010).

In this paper, an efficient ISG (included Cybersecurity) is proposed for public and private organizations that helps them to secure and control their assets and IT. Since, infosec assigns the broader desire of protecting all forms of information, Cybersecurity is a branch of infosec that specified to secure digital information. Therefore, what is applied in the ISG framework includes necessarily Cybersecurity. Moreover, the proposed ISG framework should work under laws and regulations which makes it successively implemented (Fig.). Eventually, a cooperation between IT management and ISG should exist in order to make it easier and serve the organization needs.



An organization vision path from top to bottom

Source: (H., von S.S. & Solms, 2010).

[2] Gregory, P. H. (2022). CISM Certified Information Security manager all-in-one exam guide. New York: McGraw Hill.

The rest of this paper is structured as follows: In Section 1, a key question is asked about the aim and the reason beyond the ISG with details. In Section 2, roles and responsibilities are introduced with RACI chart and Organization elements. In Section 3, Some of top ISG frameworks are presented to consider them as example to design a customized ISG framework that inspired by them. In Section 4, the proposed ISG framework is presented with laws and regulations plus the actions that should be provided by the organization to ensure a good ISG framework is Implemented. the ISG framework is concluded at the end.

1. Why do organizations should have ISG framework? What is the aim of it?

In this section, the mandatory needs of organizations having its own ISG is presented. Generally, ISG becomes very necessary for Public and private organizations due to the rapid advancements in technologies and social media era as discussed earlier in the previous section. Consequently, many government organizations in different sectors are highly depend on their information and IT infrastructure. Therefore, this dependency could reach a level where organizations keep focusing on products/services that information-related to keep their business operations. Moreover, it is necessary for these organizations to secure their assets and data by recognizing their business priority and apply Confidentiality, Integrity, and Availability (CIA) concepts[3].

The aim of designing ISG framework is to work along with their business needs which should contain a strategy to secure/control infosec efficiently. Furthermore, formal security controls are established to explain and ensure activities and desired results. IT and security programs are structured and executed consistently in order to support business priorities. Meanwhile, formal controls and measurement of processes provides managing with clear insights into security governing of the organization. By aligning security management along with procedures that is used in corporate and IT governance ensures efficiency of ISG program. Security management should be integrated into IT and corporate governance processes. Eventually, through security management, strategic planning the proposed ISG program can ensure an overall governance in both public and private organization (Rebollo et al., 2015).

After all, for effective ISG framework implementation, the C-level executives of organization should take the responsibility of data protection in their organizations, here are some activities that should be included in the organization ISG framework (Rebollo et al., 2015):

a) risk management: organization risks should me managed to mitigate exist and future risks of the organization. However, in some cases management should compromise and accept a certain level of risk to sustain the organization functionality.

[3] Gregory, P. H. (2020). CISA Certified Information Systems Auditor all-in-one exam guide. New York: McGraw-Hill.

b) compliance: organization should have restricted to laws and regulations applied in their country also, it should have their own policies and standards to protect their data and assets.

c) incident Response Management: organization executives should establish strategy to handle incidents in order to control sudden event, minimize its impact and support the organization's capability to mitigate the after effect.

d) business Continuity Plan (BCP): it ensures that the organization should stay functional during and after any incident or disaster. Also, this is essentially including a Disaster Recovery Plan (DRP) to maintain the operations in the organization.

e) Disaster Recovery Plan (DRP): it is a part of BCP which is focused on restoring the organization data, infrastructure, IT systems, after incident or the disaster. DRP should have emergency team which have done backups (mirrors), multiple sites (hot, cold) and documentation.

f) security Awareness: it is important to keep all members of the organization at a certain level of awareness in infosec and what ISG framework brings to the table by subjecting multiple training programs thought a year especially IT and management staffs.

Through these activities, C-level executives are played important roles in managing and directing the organization's information systems, ensuring resilience against potential threats and fostering a strategic approach to security governance.

2. Job titles, roles and responsibilities in organizations

The department of infosec in any organization is imaged of being the «department of no» and viewed as an obstacle to business activities. This image emerged from infosec managers who were occasionally overly cautious about risks, oversighting organizations in terms of expand, introduce innovative products and services. As a result, this reputation creates a hesitation among IT members and other business units to interact with security professionals without fearing that cooperation may impede their job. Moreover, a good implementation of ISG happens when organization members grasps their duties and restrict to roles and responsibilities. Also, organizations should establish formal roles and responsibilities that assigns every employee to instructions on how to preserve organization data and assets. Consequently. these roles should be tied to job titles by indicating an employee's place within the organization. Job titles are valued by the organizations in order to ensure that everyone should be rolled based on their titles. Generally, job titles are attached with employee's position which reflects their authority level, here are some job titles that are listed in order of seniority (Nicho, 2018):

a. Chairman, Board of Directors.

b. Member, Board of Directors.

c. Chief Executive Officer.

d. President.

e. Executive Vice President.

f. Senior Vice President.

g. Vice President.

h. Executive Director.

i. Senior Director.

j. Director.

k. Senior Manager.

l. Manager.

m. Supervisor.

The above list covered some of seniority ranks but in larger organizations there are other titles such as first (e.g. first vice president), general (e.g. general manager) and assistant (e.g. assistant director). Further, the responsibilities are much like roles defining the tasks expected from someone. In infosec, organizations assign specific roles and responsibilities to employees and team members in order to guarantee the organization's ISG strategy and goals.

2.1. Standard method of governing organizations

Many organizations used non-standard methods for governing infosec such as doing some security experience here and there. However, there is a standard method which used widely to define roles and responsibilities in organizations known as the Responsible-Accountable-Consulted-Informed (RACI) chart. It is designed to assign roles to employees and teams to perform tasks and activities. Moreover, the chart basically describes who to do what in that organization. For instance, assigns a manager for a project plus that manager should work as security analyst. Also, it gives responsibilities to each employee at any seniority level as follow (Bettwy et al., 2016):

i. Responsible: Any employee who is responsible of a task.

ii. Accountable: Any employee who is responsible of result of a task.

iii. Consulted: Any employee who has experience and can be consulted in a topic.

iv. Informed: Any employee who gets prior notice during or before an action.

Table 1 contains an example of assign roles and responsibilities in an organization. First, employees in organizations must have their own roles and belong to a team. Moreover, every employee should get a specific training course that give him a set of skills in order to accomplish their tasks. In addition, RACI chart urging employees in the organization to have their own tasks and this is called Separation of Duties (SoD). SoD means that no single employee has the full control of a critical process of activity that may affect the organization's functionality. For example, the provisioning of employee

account, the provisioner, approver and requester must not be at the same department as a part of preventing the conflicts of interest (Von Solms et al., 2011).

Table 1. Assign roles and responsibilities in RACI chart

Activity	Responsible	Accountable	Consulted	Informed
Audit user account	IA	IAM	AO	IT SD, IT SM, EUM
Provision user account	IT SD	IT SM	AO	IT SD, EUM, ST
Approve user account	AO	COO	EUM, ST	EU, IA, IT SD
Request user account	EU	EUM	IT SD, EUM	AO, ST

* EU: End User, EUM: End User Manager, SD: Service Desk, EUM: End User Manager, AO: Asset Owner, ST: Security Team, IA: Internal Audit, SM: Service Manager, IAM: Internal Audit Manager.

Source: (Von Solms et al., 2011)

2.2. Standards and security frameworks of organizations

A successive organization is the one that operates under well designed ISG framework which based on standards that works along with its vision and strategy. However, design a strategy is not easy which includes policies, standards, process and matrices that supports the overall vision of the organization. In this section, many standards are presented that applied in world-wide organizations such as Google, Meta[4], Amazon, …, etc. Moreover, if security professional along with C-level executives are decided to design ISG framework to an organization and selected a control framework alone, it is often considered as a mistake. Arguably, this ISG framework should exemplified and take advantage of some world-wide standards and security controls to avoid mistakes/ issues in order to provide a good start in governing the organization. Many standards and control frameworks are listed in Table 2 that can be useful as a start point to design ISG framework for public and private organizations (Tan et al., 2010; Fazlida & Said, 2015; Ula et al., 2017).

Table 2. Some standards and control frameworks

No.	Standards and frameworks	Explanation
1	National Institute of Standards and Technology (NIST)	NIST is used by U.S. Department of Commerce to standardize economic security, innovation, industry and technology comprehensive
2	International Organization for Standardization/ International Electrotechnical Commission (ISO/IEC)	ISO/IEC two main international standards that are used for infosec, technology, industry and business practices

[4] The organization is recognized as extremist, its functioning is prohibited in the territory of the Russian Federation.

End of Table 2

No.	Standards and frameworks	Explanation
3	Control Objectives for Information and Related Technologies (COBIT)	COBIT is a framework that is used to governing IT and managing enterprises by set some guidelines and best practices for IT organizations
4	Payment Card Industry Data Security Standard (PCI-DSS)	PCI-DSS is a combination of infosec standards that established to protect sensitive information of payment cards
5	Health Insurance Portability and Accountability Act (HIPAA)	HIPAA is an act in U.S. law that enacted to protect information of people's health privacy plus to secure medical data
6	Information Security Forum (ISF)	ISF is a combination of security controls and best practices that used to manage risks in infosec
7	Information Technology Infrastructure Library (ITIL)	ITIL is an IT service framework is designed to for management purposes that included managing IT infrastructure, environment, services and processes

Source: (Ula et al., 2017).

3. Top ISG/Cybersecurity frameworks

In this section, some of the top countries in ISG and cybersecurity field are presented. Meanwhile, there are some Arabic countries that are listed with the highly ranked and powerful countries such as United Arab Emirates and Saudi Aribia. These countries are highly recommended examples to be follow in such sensitive and valuable field. The following are the countries including their ISG frameworks, Laws and regulations to control Information (Shingarev & Kazakova, 2021; Creemers, 2023; Priyadarshini & Cotton, 2022; Carr & Tanczer, 2018; Singh & Alshammari, 2020; Al Neaimi et al., 2015):

Russia's ISG framework: it implies laws such as Federal Service for Technical and Export Control (FSTEC) which was enacted to secure products, Federal Law No. 149-FZ to protect data and the role of Federal Security Service (FSB) in Cybersecurity. Consequently, local laws of data urged that Russian citizens' information has to be stored in the country (Shingarev & Kazakova, 2021).

China's ISG framework: it includes ministries like Ministry of Public Security (MPS) which oversights infosec in China. In addition, the Chinese Cybersecurity Law enacted in 2017 with the National Security Law (NSL). Eventually, the law of Cyberspace Administration of China (CAC) that controls internet regulations (Creemers, 2023).

US's ISG framework: it involves laws such as Federal Information Security Management Act (FISMA) and Cybersecurity Enhancement Act (CEA) of 2014. Also, National Institute of Standards and Technology (NIST) which established standards. Cybersecurity and Infrastructure Security Agency (CISA) that coordinated agencies such as the National Security Agency (NSA) that handles Cybersecurity activities (Priyadarshini & Cotton, 2022).

UK's ISG framework: it contains acts and strategy National Cyber Security Strategy (NCSC) and Data Protection Act (DPA) 2018. Moreover, the Computer Misuse Act (CMA) in 1990 assigns cyber offenses. The UK government involves international cybersecurity cooperation (Carr & Tanczer, 2018).

Saudi Arabia's ISG framework: it involves laws such as the Saudi Arabia Cybersecurity Law (SACL) in 2019 that controlled by the Communications, Space and Technology Commission (CITC) and National Cybersecurity Authority (NCA). Saudi Central Bank (SAMA) oversight the financial transactions in the financial sector (Singh & Alshammari, 2020).

UAE's ISG framework: it includes laws like UAE Cybersecurity Law of 2019 and oversight by the National Electronic Security Authority (NESA) and Telecommunications and Digital Government Regulatory Authority (TRDA). Dubai English Speaking College (DESC) that oversights the cybersecurity in Dubai (Al Neaimi et al., 2015).

4. A proposed information security governance framework

Before proposing any new framework, or development an existing framework, it is imperative to explain two fundamental concepts: Governance and Corporate Governance. Governance pertains to protect the interests of owners by guiding, managing and supervising on their behalf, with the Board of Directors acting as their representatives. Corporate Governance is defined as response to the separation between management and ownership within private and public organizations. Moreover, it aims to maintain this separation by providing incentives to both the management and board to pursue aims which are in line with the interests of the company and its shareholders.

The proposed framework for ISG should include some elements to ensure an effective protection and management of an organization's information assets, achieving both discipline between owners and management plus granting owners the authority to oversee the organization. Additionally, by establishing a secure environment for sharing and storing information, organizations can not only enhance productivity, consumer benefits and business efficiency but also support security measures. Conversely, any insecure work environment presents significant risks, potentially resulting in substantial harm to corporations and governments, with possible adverse effects on citizens and consumers. This is particularly critical for businesses operating in crucial organizations such as finance, electricity generation, banking, or healthcare, where the stakes are exceptionally high. Table 3 includes the key questions essential for establishing effective ISG.

Table 3. Some important questions/actions for effective ISG

Actors/Actions	Corporate Executives	Business Unit Head	Senior Manager	CIO/CISO
Governance/Business Drivers	What am I required to do? What am I afraid not to do?			
Roles and Responsibilities	How do I accomplish my objectives?			
Metrics/Audit	How effectively do I achieve my objectives? What adjustments do I need to make?			

The ISG framework serves as a tool to implement the strategy and vision of the C-level executives to achieve high performance of business operations and decision-making in organizations. It falls under their purview to manage as part of their oversighting the organization and protect its data and assets by guaranteeing the efficient integration of infosec throughout their organization.

To design an effective ISG framework that can be globally affirmed and accepted, there are some global laws and regulations that should be taken into consideration. Consequently, these laws and regulations can be a great advantage due to their structures and well-designed by use it as a law experiences of other countries to legislate and enact our Laws and Regulations. Table 4 shows some key laws and regulations that has been used over the globe in governing infosec.

Table 4. Global laws and regulations examples[5]

No.	Laws and Regulations	Explanation
1	General Data Protection Regulation (GDPR)	It mandates organizations to protect the personal data of individuals within the European Union (EU) and imposes strict requirements for data privacy and security.
2	California Consumer Privacy Act (CCPA)	It applies to businesses that collect personal information of California residents and requires them to implement measures to protect the privacy and security of such information.
3	Sarbanes-Oxley Act (SOX)	It requires companies to establish and maintain internal controls over financial reporting, which includes measures to protect the integrity and confidentiality of financial data.
4	Federal Information Security Management Act (FISMA)	It is a US federal law that establishes security requirements for federal information systems and provides a framework for managing cybersecurity risks in federal agencies.
5	Cybersecurity Maturity Model Certification (CMMC)	It developed by the U.S. Department of Defense (DoD) to assess and enhance the cybersecurity posture of defense contractors and subcontractors.
6	Data Protection Laws (DPL)	Various countries have enacted their own data protection laws, such as the Personal Information Protection and Electronic Documents Act (PIPEDA) in Canada and the Personal Data Protection Act (PDPA) in Singapore.

Conclusions

In this work, a new ISG (includes cybersecurity) framework is proposed to protect information and assets of public and private organizations by taking advantage of some laws and regulations. This framework can be compared to the existed frameworks that have been implemented in the world-wide organizations. In addition, it focuses on cooperation between continuous improvement and risk management which aligns with the business model of the organization that includes regulations and laws requirements.

[5] Manning, W. (2010). CISM Certified Information Security Manager certification exam preparation course in a book for passing the CISM: The how to pass on your first try certification study guide. Brisbane, Australia: Emereo Pty Ltd.

Meanwhile, any organization should have its own ISG framework and a committee (BoD) to implement it. For successive ISG in a country, the committees in all organizations should be connected to each other by a higher committee of ISG or Cybersecurity that can implement the overall governance. Furthermore, this ISG framework acts as a weapon to implement governance of infosec plus ensures that the overall process works along with the business goals and objectives effectively. Finally, this ISG framework offers a real security program which can be applied by the authors to any private and public organization.

References

Al Neaimi, A., Ranginya, T., & Lutaaya, P. (2015). A framework for effectiveness of cyber security defenses, a case of the United Arab Emirates (UAE). *International Journal of Cyber-Security and Digital Forensics*, *4*(1), 290–301. https://doi.org/10.17781/p001502

AlGhamdi, S., Win, K. T., & Vlahu-Gjorgievska, E. (2020). Information security governance challenges AND CRITICAL SUCCESS FACTORS: Systematic review. *Computers & Security*, *99*, 102030. https://doi.org/10.1016/j.cose.2020.102030

Bendovschi, A. (2015). Cyber-attacks – trends, patterns and security countermeasures. *Procedia Economics and Finance*, *28*, 24–31. https://doi.org/10.1016/s2212-5671(15)01077-1

Bettwy, G., Beevers, M., & Williams, M. (2016). *CISM Review Manual Certified Information Security manager*. Rolling Meadows, Ill: ISACA.

Carcary, M., Renaud, K., McLaughlin, S., & O'Brien, C. (2016). A framework for Information Security Governance and Management. *IT Professional*, *18*(2), 22–30. https://doi.org/10.1109/mitp.2016.27

Carr, M., & Tanczer, L. M. (2018). UK cybersecurity industrial policy: An analysis of drivers, market failures and interventions. *Journal of Cyber Policy*, *3*(3), 430–444. https://doi.org/10.1080/23738871.2018.1550523

Corriss, L. (2010). Information security governance. *Proceedings of the 2010 Workshop on Governance of Technology, Information and Policies* (pp. 35–41). https://doi.org/10.1145/1920320.1920326

Creemers, R. (2023). Cybersecurity law and regulation in China: Securing the smart state. *China Law and Society Review*, *6*(2), 111–145. https://doi.org/10.1163/25427466-06020001

Fazlida, M. R., & Said, J. (2015). Information security: Risk, governance and implementation setback. *Procedia Economics and Finance*, *28*, 243–248. https://doi.org/10.1016/s2212-5671(15)01106-5

Hadi, M. A., Abdulredha, M. N., & Hasan, E. (2023). Introduction to ChatGPT: A new revolution of artificial intelligence with machine learning algorithms and cybersecurity. *Science Archives*, *04*(04), 276–285. https://doi.org/10.47587/sa.2023.4406

Johnston, A. C., & Hale, R. (2009). Improved security through information security governance. *Communications of the ACM*, *52*(1), 126–129. https://doi.org/10.1145/1435417.1435446

Moulton, R., & Coles, R. S. (2003). Applying information security governance. *Computers & Security*, *22*(7), 580–584. https://doi.org/10.1016/s0167-4048(03)00705-3

Nicho, M. (2018). A process model for implementing Information Systems Security governance. *Information & Computer Security*, *26*(1), 10–38. https://doi.org/10.1108/ics-07-2016-0061

Ohki, E., Harada, Y., Kawaguchi, Sh., Shiozaki, T., & Kagaya, T. (2009). Information security governance framework. *Proceedings of the first ACM workshop on Information security governance* (pp. 1–6). https://doi.org/10.1145/1655168.1655170

Priyadarshini, I., & Cotton, C. (2022). Cyber laws in the United States. *Cybersecurity (*pp. 157–237). https://doi.org/10.1201/9781003187127-6

Rao, U. H., & Nayak, U. (2014). History of computer security. *The InfoSec Handbook* (pp. 13–25). https://doi.org/10.1007/978-1-4302-6383-8_2

Rastogi, R., & von Solms, R. (2005). Information security governance – a re-definition. *Security Management, Integrity, and Internal Control in Information Systems* (pp. 223–236). https://doi.org/10.1007/0-387-31167-x_14

Rebollo, O., Mellado, D., Fernández-Medina, E., & Mouratidis, H. (2015). Empirical evaluation of a cloud computing information security governance framework. *Information and Software Technology*, *58*, 44–57. https://doi.org/10.1016/j.infsof.2014.10.003

Rocha Flores, W., Antonsen, E., & Ekstedt, M. (2014). Information security knowledge sharing in organizations: Investigating the effect of behavioral information security governance and national culture. *Computers & Security*, *43*, 90–110. https://doi.org/10.1016/j.cose.2014.03.004

Shingarev, A., & Kazakova, A. (2021). The Russian Federation's approach to Cybersecurity. *The Oxford Handbook of Cyber Security* (pp. 672–684). https://doi.org/10.1093/oxfordhb/9780198800682.013.44

Singh, H. P., & Alshammari, T. S. (2020). An institutional theory perspective on developing a cyber security legal framework: A case of Saudi Arabia. *Beijing Law Review*, *11*(03), 637–650. https://doi.org/10.4236/blr.2020.113039

Tan, T. C., Ruighaver, A. B., & Ahmad, A. (2010). Information security governance: When compliance becomes more important than security. *Security and Privacy – Silver Linings in the Cloud* (pp. 55–67). https://doi.org/10.1007/978-3-642-15257-3_6

Ula, M., Ula, M., & Fuadi, W. (2017). A method for Evaluating Information Security Governance (ISG) components in banking environment. *Journal of Physics: Conference Series*, *812*, 012031. https://doi.org/10.1088/1742-6596/812/1/012031

Von Solms, S. H., & von Solms, R. (2010). *Information security governance*. New York, United States: Springer.

Von Solms, R., Thomson, K.-L., & Maninjwa, P. M. (2011). Information security governance control through comprehensive policy architectures. *Information Security for South Africa* (pp. 1–6). https://doi.org/10.1109/issa.2011.6027522

Wu, M., Aranovich, R., & Filkov, V. (2021). Evolution and differentiation of the cybersecurity communities in three social question and answer sites: A mixed-methods analysis. PLOS ONE, 16(12). https://doi.org/10.1371/journal.pone.0261954

Authors information

Musadaq Ahmed Hadi – MSc. (Control Engineer), Control and Systems Engineering Department, University of Technology
Address: Al-Wehda, Baghdad, Iraq
E-mail: musadaq.ahmed@alshaab.edu.iq
ORCID ID: https://orcid.org/0000-0002-3884-495X
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57117413800
WoS Researcher ID: https://www.webofscience.com/wos/author/record/JZT-4113-2024
Google Scholar ID: https://scholar.google.com/citations?user=wcrA7n8AAAAJ

Mohammed Najm Abdulredha – MSc. (Computer Science), Department of Computer Science, University of Baghdad
Address: Al-Jadriya, Baghdad, Iraq
E-mail: mohammed.najm.422@gmail.com
ORCID ID: https://orcid.org/0009-0007-8441-3505
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57226709471
WoS Researcher ID: https://www.webofscience.com/wos/author/record/KIJ-3538-2024
Google Scholar ID: https://scholar.google.ru/citations?user=-oXQXKEAAAAJ

Authors' contributions

The authors have contributed equally into the concept and methodology elaboration, validation, formal analysis, research, selection of sources, text writing and editing, project guidance and management.

Conflict of interest

The authors declare no conflict of interest.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – April 16, 2025
Date of approval – May 4, 2025
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:34.05:004.056.5
EDN: https://elibrary.ru/xujvtm
DOI: https://doi.org/10.21202/jdtl.2025.6

Универсальная система управления информационной безопасностью: организационно-правовые принципы

Мусадак Ахмед Хади ✉

Технологический университет, Багдад, Ирак

Мохаммед Наджм Абдулредха

Багдадский университет, Багдад, Ирак

Ключевые слова

законодательство,
защита информации,
информационная безопасность,
информационные технологии,
кибербезопасность,
организационная структура,
право,
правовое регулирование,
управление информационной безопасностью,
цифровые технологии

Аннотация

Цель: разработка универсальных организационно-правовых принципов построения системы управления информационной безопасностью, которые позволят каждой организации создать собственную эффективную систему управления информационной безопасностью с учетом ее уникальных бизнес-целей и задач.

Методы: основаны на интеграции ключевых элементов управления информационной безопасностью, таких как видение, стратегия, цели, политики, стандарты, процессы и матрицы. Видение и цели задают направление развития организации, политики и стандарты обеспечивают концептуальную основу для защиты информации, процессы позволяют систематически достигать поставленных задач, а матрицы предоставляют инструменты для оценки и контроля всей структуры. Предложенные принципы согласуются с международными стандартами, нормативными требованиями и лучшими практиками в области информационной безопасности.

Результаты: разработанная система управления информационной безопасностью позволяет четко распределить роли и обязанности среди сотрудников организации, обеспечивая эффективное внедрение системы управления. Авторы также анализируют существующие принципы безопасности информационных технологий, интегрируя их в стратегию безопасности, которая соответствует целям организации. Предложенная универсальная система соответствует нормативным правовым требованиям и может быть адаптирована для использования в организациях любого масштаба и профиля.

✉ Контактное лицо

Научная новизна: заключается в представлении практического подхода к внедрению системы управления информационной безопасностью, основанного на опыте авторов, а также на мировых стандартах, системах контроля и правовых актах. В отличие от существующих подходов предлагаемая система является гибкой и может быть адаптирована под специфику любой организации, что делает ее универсальным инструментом для управления информационной безопасностью.

Практическая значимость: состоит в предоставлении структурированного подхода к созданию универсальной системы управления информационной безопасностью, который может быть использован организациями, испытывающими недостаток знаний и ресурсов для реализации подобных инициатив. Авторы предлагают общую структуру, которая может быть адаптирована в зависимости от активов организации, уровня подготовки сотрудников и их осведомленности в вопросах информационной безопасности. Это делает настоящую работу ценным ресурсом для специалистов, стремящихся повысить уровень защиты информации в своих организациях.

Для цитирования

Хади, М. А., Абдулредха, М. Н. (2025). Универсальная система управления информационной безопасностью: организационно-правовые принципы. *Journal of Digital Technologies and Law*, *3*(1), 125–142. https://doi.org/10.21202/jdtl.2025.6

Список литературы

Al Neaimi, A., Ranginya, T., & Lutaaya, P. (2015). A framework for effectiveness of cyber security defenses, a case of the United Arab Emirates (UAE). *International Journal of Cyber-Security and Digital Forensics*, *4*(1), 290–301. https://doi.org/10.17781/p001502

AlGhamdi, S., Win, K. T., & Vlahu-Gjorgievska, E. (2020). Information security governance challenges AND CRITICAL SUCCESS FACTORS: Systematic review. *Computers & Security*, *99*, 102030. https://doi.org/10.1016/j.cose.2020.102030

Bendovschi, A. (2015). Cyber-attacks – trends, patterns and security countermeasures. *Procedia Economics and Finance*, *28*, 24–31. https://doi.org/10.1016/s2212-5671(15)01077-1

Bettwy, G., Beevers, M., & Williams, M. (2016). *CISM Review Manual Certified Information Security manager*. Rolling Meadows, Ill: ISACA.

Carcary, M., Renaud, K., McLaughlin, S., & O'Brien, C. (2016). A framework for Information Security Governance and Management. *IT Professional*, *18*(2), 22–30. https://doi.org/10.1109/mitp.2016.27

Carr, M., & Tanczer, L. M. (2018). UK cybersecurity industrial policy: An analysis of drivers, market failures and interventions. *Journal of Cyber Policy*, *3*(3), 430–444. https://doi.org/10.1080/23738871.2018.1550523

Corriss, L. (2010). Information security governance. *Proceedings of the 2010 Workshop on Governance of Technology, Information and Policies* (pp. 35–41). https://doi.org/10.1145/1920320.1920326

Creemers, R. (2023). Cybersecurity law and regulation in China: Securing the smart state. *China Law and Society Review*, *6*(2), 111–145. https://doi.org/10.1163/25427466-06020001

Fazlida, M. R., & Said, J. (2015). Information security: Risk, governance and implementation setback. *Procedia Economics and Finance*, *28*, 243–248. https://doi.org/10.1016/s2212-5671(15)01106-5

Hadi, M. A., Abdulredha, M. N., & Hasan, E. (2023). Introduction to ChatGPT: A new revolution of artificial intelligence with machine learning algorithms and cybersecurity. *Science Archives*, *04*(04), 276–285. https://doi.org/10.47587/sa.2023.4406

Johnston, A. C., & Hale, R. (2009). Improved security through information security governance. *Communications of the ACM*, *52*(1), 126–129. https://doi.org/10.1145/1435417.1435446

Moulton, R., & Coles, R. S. (2003). Applying information security governance. *Computers & Security*, *22*(7), 580–584. https://doi.org/10.1016/s0167-4048(03)00705-3

Nicho, M. (2018). A process model for implementing Information Systems Security governance. *Information & Computer Security*, *26*(1), 10–38. https://doi.org/10.1108/ics-07-2016-0061

Ohki, E., Harada, Y., Kawaguchi, Sh., Shiozaki, T., & Kagaya, T. (2009). Information security governance framework. *Proceedings of the first ACM workshop on Information security governance* (pp. 1–6). https://doi.org/10.1145/1655168.1655170

Priyadarshini, I., & Cotton, C. (2022). Cyber laws in the United States. *Cybersecurity (*pp. 157–237). https://doi.org/10.1201/9781003187127-6

Rao, U. H., & Nayak, U. (2014). History of computer security. *The InfoSec Handbook* (pp. 13–25). https://doi.org/10.1007/978-1-4302-6383-8_2

Rastogi, R., & von Solms, R. (2005). Information security governance – a re-definition. *Security Management, Integrity, and Internal Control in Information Systems* (pp. 223–236). https://doi.org/10.1007/0-387-31167-x_14

Rebollo, O., Mellado, D., Fernández-Medina, E., & Mouratidis, H. (2015). Empirical evaluation of a cloud computing information security governance framework. *Information and Software Technology*, *58*, 44–57. https://doi.org/10.1016/j.infsof.2014.10.003

Rocha Flores, W., Antonsen, E., & Ekstedt, M. (2014). Information security knowledge sharing in organizations: Investigating the effect of behavioral information security governance and national culture. *Computers & Security*, *43*, 90–110. https://doi.org/10.1016/j.cose.2014.03.004

Shingarev, A., & Kazakova, A. (2021). The Russian Federation's approach to Cybersecurity. *The Oxford Handbook of Cyber Security* (pp. 672–684). https://doi.org/10.1093/oxfordhb/9780198800682.013.44

Singh, H. P., & Alshammari, T. S. (2020). An institutional theory perspective on developing a cyber security legal framework: A case of Saudi Arabia. *Beijing Law Review*, *11*(03), 637–650. https://doi.org/10.4236/blr.2020.113039

Tan, T. C., Ruighaver, A. B., & Ahmad, A. (2010). Information security governance: When compliance becomes more important than security. *Security and Privacy – Silver Linings in the Cloud* (pp. 55–67). https://doi.org/10.1007/978-3-642-15257-3_6

Ula, M., Ula, M., & Fuadi, W. (2017). A method for Evaluating Information Security Governance (ISG) components in banking environment. *Journal of Physics: Conference Series*, *812*, 012031. https://doi.org/10.1088/1742-6596/812/1/012031

Von Solms, S. H., & von Solms, R. (2010). *Information security governance*. New York, United States: Springer.

Von Solms, R., Thomson, K.-L., & Maninjwa, P. M. (2011). Information security governance control through comprehensive policy architectures. *Information Security for South Africa* (pp. 1–6). https://doi.org/10.1109/issa.2011.6027522

Wu, M., Aranovich, R., & Filkov, V. (2021). Evolution and differentiation of the cybersecurity communities in three social question and answer sites: A mixed-methods analysis. PLOS ONE, 16(12). https://doi.org/10.1371/journal.pone.0261954

Сведения об авторах

Хади Мусадак Ахмед – магистр наук (инженер систем управления), кафедра управления и системной инженерии, Технологический университет
Адрес: Аль-Вейда, г. Багдад, Ирак
E-mail: musadaq.ahmed@alshaab.edu.iq
ORCID ID: https://orcid.org/0000-0002-3884-495X
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57117413800
WoS Researcher ID: https://www.webofscience.com/wos/author/record/JZT-4113-2024
Google Scholar ID: https://scholar.google.com/citations?user=wcrA7n8AAAAJ

Абдулредха Мохаммед Наджм – магистр компьютерных наук, кафедра компьютерных наук, Багдадский университет
Адрес: Аль-Джадрийя, г. Багдад, Ирак
E-mail: mohammed.najm.422@gmail.com
ORCID ID: https://orcid.org/0009-0007-8441-3505
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57226709471
WoS Researcher ID: https://www.webofscience.com/wos/author/record/KIJ-3538-2024
Google Scholar ID: https://scholar.google.ru/citations?user=-oXQXKEAAAAJ

Вклад авторов

Авторы внесли равный вклад в разработку концепции, методологии, валидацию, формальный анализ, проведение исследования, подбор источников, написание и редактирование текста, руководство и управление проектом.

Конфликт интересов

Авторы сообщают об отсутствии конфликта интересов.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.19.61 / Правовое регулирование информационной безопасности
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 16 апреля 2024 г.
Дата одобрения после рецензирования – 4 мая 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.

Research article
UDC 34:004:17:004.8:342.7
EDN: https://elibrary.ru/egkppn
DOI: https://doi.org/10.21202/jdtl.2025.7

Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection

Pedro Miguel Alves Ribeiro Correia ✉
University of Coimbra, Coimbra, Portugal

Ricardo Lopes Dinis Pedro
University of Lisbon, Lisbon, Portugal

Susana Videira
University of Lisbon, Lisboa, Portugal

Keywords

artificial intelligence,
data protection,
ethical regulation,
ethics,
fundamental rights,
healthcare,
human rights,
law,
legal regulation,
predictive analytics

Abstract

Objective: to identify key ethical, legal and social challenges related to the use of artificial intelligence in healthcare; to develop recommendations for creating adaptive legal mechanisms that can ensure a balance between innovation, ethical regulation and the protection of fundamental human rights.

Methods: a multidimensional methodological approach was implemented, integrating classical legal analysis methods with modern tools of comparative jurisprudence. The study covers both the fundamental legal regulation of digital technologies in the medical field and the in-depth analysis of the ethical, legal and social implications of using artificial intelligence in healthcare. Such an integrated approach provides a comprehensive understanding of the issues and well-grounded conclusions about the development prospects in this area.

Results: has revealed a number of serious problems related to the use of artificial intelligence in healthcare. These include data bias, non-transparent complex algorithms, and privacy violation risks. These problems can undermine public confidence in artificial intelligence technologies and exacerbate inequalities in access to health services. The authors conclude that the integration of artificial intelligence into healthcare should take into account fundamental rights, such as data protection and non-discrimination, and comply with ethical standards.

✉ Corresponding author

Scientific novelty: the work proposes effective mechanisms to reduce risks and maximize the potential of artificial intelligence under crises. Special attention is paid to regulatory measures, such as the impact assessment provided for by the Artificial Intelligence Act. These measures play a key role in identifying and minimizing the risks associated with high-risk artificial intelligence systems, ensuring compliance with ethical standards and protection of fundamental rights.

Practical significance: adaptive legal mechanisms were developed, that support democratic norms and respond promptly to emerging challenges in public healthcare. The proposed mechanisms allow achieving a balance between using artificial intelligence for crisis management and human rights. This helps to build confidence in artificial intelligence systems and their sustained positive impact on public healthcare.

For citation

Correia, P. M. A. R., Pedro, R. L. D., & Videira, S. (2025). Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection. *Journal of Digital Technologies and Law, 3*(1), 143–180. https://doi.org/10.21202/jdtl.2025.7

Contents

Introduction
1. The use of Artificial Intelligence in the Combat Against Pandemics and Why it Fails
2. Examples of Artificial Intelligence Failures: Lessons Learned and the Path Forward
3. The Problem with Models: Garbage In Garbage Out
4. The Sustainability Problem (almost) no One Talks About
5. Governance as Key: How State Measures and Data Availability Reinforce some Organizational Values and Contribute to the Sustainability of the National Health System
6. Governance as Key: How Governance Reinforces some Organizational Valuesand Contributes to the Sustainability of Crisis Management
7. Critical Reflections
8. AI and Fundamental Rights
9. Pandemics and Fundamental Rights
10. Possible Relationships between AI and Fundamental Rights in Combating Pandemics
Conclusions
References

Introduction

As practitioners, institutions and governments consider the looming shadow of disease and pandemics, can artificial intelligence (AI) become our panacea and our light? What limits can fundamental rights present when fighting pandemics with the use of AI?

Throughout history, humanity has faced the devastating wrath of epidemics and pandemics. From the bubonic plague to the Spanish flu, these outbreaks have reshaped societies and challenged our very existence. Today, as we navigate an increasingly connected world, the threat of new and rapidly spreading diseases looms large. Globalization is increasingly becoming a double-edged sword, fostering collaboration but also facilitating the swift travel of pathogens across borders (Jones et al., 2008; Morse et al., 2012).

Artificial intelligence is now viewed as a new, emerging weapon in this age-old fight. This revolutionary, powerful technology (or amalgamation of technologies), once relegated to the realm of science fiction, holds immense potential to revolutionize how humans combat epidemics and pandemics. Poised to become a powerful weapon in our arsenal for combatting disease. By means of artificial intelligence techniques one might analyze vast amounts of data to predict outbreaks and outcomes, accelerate drug discovery, and even personalize treatment strategies (Syrowatka et al., 2021). Or so it is believed.

For instance, artificial intelligence can be used for forecasting the unforeseen, as it can sift through voluminous information from social media, news reports, and even satellite imagery, potentially identifying early warning signs of an outbreak before it explodes into a full-blown pandemic. It is not farfetched to imagine an artificial intelligence system detecting a surge in searches for flu-like symptoms in a specific region, triggering an immediate investigation that could potentially nip an outbreak in the bud (Wong et al., 2023).

Another instance is the use of artificial intelligence to accelerate the quest for cures (if that, and not only chronic disease mitigating solutions, perpetual cash cows, are still pursued by pharmaceutical companies). Medical chemical compounds discovery (including vaccines) has traditionally been a slow and arduous process, often taking years to yield results, if at all. Artificial intelligence can be used to revolutionize this process by analyzing molecular structures to identify potential drug candidates or repurpose existing drugs for new significantly reducing the time it takes to get life-saving treatments into the hands of patients (Matsuzaka & Yashiro, 2022).

Yet another instance is the tailoring of treatments, diverging from the one-size-fits-all approach, as artificial intelligence could be developed to be capable of analyzing patients' individual genetic makeups and medical histories to predict how they might respond to different treatment options, paving the way for personalized medicine and allowing doctors to tailor treatment plans for maximum effectiveness (Topol, 2019).

An additional example for the use of artificial intelligence, in this context, would be predicting the trajectory of an outbreak or a pandemic, due to the capability of this models to analyze historical data and disease characteristics, allowing public health officials to strategically allocate resources and implement targeted containment measures (Ferguson et al., 2006).

On final example (of numerous others not listed here) is the use of artificial intelligence to enhance contact tracing, as it can analyze contact tracing data and travel information to pinpoint individuals at high risk of infection, therefore helping healthcare workers prioritize testing and quarantine measures, potentially containing the spread of the pathogen (Fetzer & Graeber, 2021).

The path forward, however, is not without its challenges. Biases in data, the «black box» or "gray box" nature of complex algorithms, and the ever-present risk of over-reliance on artificial intelligence all pose potential and significant pitfalls (DeCamp & Tilburt, 2019).

Despite this perfunctory listing of potentialities, this text will delve not so much into the stimulating possibilities that arose with the advent of artificial intelligence in the fight against disease, epidemics, or pandemics, but, much more, into the critical considerations and cautions users should take into account as the use of these approaches is evermore pondered. We will examine the various challenges artificial intelligence faces before it can become, if ever, a beacon of hope in a world ever shadowed by the threat of widespread, generalized, and acute disease outbreaks.

In addition, the impact of AI on fundamental rights and, in particular, the use of AI to combat a pandemic and its relationship with the preservation of fundamental rights are also analysed.

1. The use of Artificial Intelligence in the Combat Against Pandemics and Why it Fails

However, this potential for good comes intertwined with challenges that must be addressed. George Box famously stated that "all models are wrong, but some are useful" (Box, 1979). It should come as no surprise that, despite its strengths, artificial intelligence models' effectiveness is hampered by several limitations that can turn those solutions into a double-edged sword.

One of the major Achilles heels of artificial intelligence models is that those same models are only as good as the data they're trained on. Inaccurate, incomplete, or biased data can lead to unreliable and potentially harmful outputs (Gianfrancesco et al., 2018). Limited access to real-time health data in some regions or privacy concerns can further restrict artificial intelligence's capabilities.

Another one is commonly known as the "black box enigma" or, in a diluted formulation the "grey box" enigma. Understanding the inner workings of intricate artificial intelligence algorithms poses an extremely difficult task, impeding humans' ability to grasp their

decision-making processes and this opacity can, in turn, undermine confidence and complicate efforts to detect and rectify underlying biases and various other types of problems (Mittelstadt et al., 2016).

Yet another one is the temptation of over-reliance on artificial intelligence models' significant capabilities without recognizing their fundamental limitations and, thus, relying too heavily on artificial intelligence without paying due attention to fundamental public health strategies such as contact tracing, vaccination efforts, and community awareness campaigns, all of them established approaches that must continue to play a vital role in effectively managing disease outbreaks (Silva et al., 2022).

2. Examples of Artificial Intelligence Failures: Lessons Learned and the Path Forward

The COVID-19 pandemic has been a testing ground for artificial intelligence in pandemic response, with mixed results. Some examples follow below.

Early in the pandemic, some artificial intelligence models drastically overestimated the spread of the virus due to limited initial data and the rapidly evolving situation. This led to unnecessary panic and resource allocation. In other instances, artificial intelligence powered chatbots designed to answer public health questions were overwhelmed by the surge in demand and provided inaccurate information in some cases. This highlights the need for robust training data and clear limitations set for artificial intelligence applications (Bajwa et al., 2021; Gürsoy & Kaya, 2023).

It can, therefore, be argued that only by acknowledging the artificial intelligence limitations and focusing on responsible development, can humanity harness its power for a healthier future. Actions as prioritizing data quality and responsible data collection practices or addressing data bias and ensuring data privacy are crucial for building trustworthy artificial intelligence models. The development of robust methods to mitigate bias in artificial intelligence algorithms and techniques based on fairness testing and data augmentation can also help identify and address potential biases from the start. Investing in explainable artificial intelligence (also known as XAI) research can help stakeholders understand how artificial intelligence models arrive at their conclusions, fostering trust and enabling early detection of potential problems (Jobin et al., 2019).

The promotion of balanced approaches, where artificial intelligence is used alongside traditional public health interventions and complements, not replaces, established public health measures could be the path forward. By addressing these challenges and fostering responsible artificial intelligence development, agents can leverage the power of artificial intelligence to create a world better prepared for future pandemics (Benke & Benke, 2018). Artificial intelligence can be a powerful weapon in humanity's arsenal, but only if wielded wisely, as evidenced underneath.

3. The Problem with Models: Garbage In Garbage Out

The concept of «garbage in, garbage out» is a fundamental principle in artificial intelligence, particularly in machine learning (Breiman, 2001).

It helps to understand what are the main types of «garbage» data that exist. Firstly, data can be inaccurate. This includes errors in spelling, typos, factual mistakes, or outdated information (Halevy et al., 2009). One can easily imagine an artificial intelligence trained on news articles with many typos – it might struggle to understand language properly. Secondly, data can be incomplete. Missing values or data points can skew the model's understanding (Little & Rubin, 2019). For example, an artificial intelligence for predicting customer churn (why patients choose not to return to a hospital, for instance) might miss crucial data points if customer feedback isn't collected. Thirdly, data can be biased. Data that unfairly represents a certain group can lead to discriminatory outcomes (Berk, 1983). An artificial intelligence used for hiring decisions trained on resumes with mostly male applicants for nursing jobs might favor male candidates in the future. Fourthly and lastly, data can be irrelevant. Information not relevant to the task at hand can confuse the model (Greiner et al., 1997). An artificial intelligence for sentiment analysis (understanding emotions in text) in a psychiatric institution might be overwhelmed by irrelevant emojis in a dataset.

It also helps to understand what might be the consequences of «garbage in». First, there is the perpetuation of bias, as artificial intelligence can amplify existing societal biases if the training data reflects those biases (Bazarkina & Pashentsev, 2020). This can lead to unfair outcomes in areas like loan approvals, facial recognition, and criminal justice predictions. Next, there is the reduction of accuracy and reliability, as models trained on inaccurate data will produce unreliable outputs (Shin & Park, 2019). Imagine an artificial intelligence for pathology prediction trained on faulty temperature readings from patients – its diagnostics would be inaccurate. Furthermore, resources can be wasted, as time and money spent training models on bad data are significant resource drains (Hulten, 2018).

Most useful is to know the chief techniques for combating «garbage in». On one hand, it pays off to invest on data cleaning and curation. Techniques like data validation, error correction, and filtering are used to ensure data quality. This can be a labor-intensive process, but crucial for reliable artificial intelligence (Wang & Shi, 2011). On another hand, the use of data augmentation, that is, creating synthetic data to supplement existing datasets can help address issues like incomplete data (Mumuni & Mumuni, 2022). For example, generating realistic-looking images with diverse faces can help reduce bias in facial recognition. Another instance is the use of algorithmic bias detection methods to identify and mitigate bias in artificial intelligence algorithms themselves. This can involve analyzing the model's decision-making process to uncover hidden biases

(Kordzadeh & Ghasemaghaei, 2022). Explainable artificial intelligence constitutes another technique that focuses on making artificial intelligence models more transparent, allowing humans to understand how the model arrives at its conclusions, helping to identify potential biases or errors (Arrieta et al., 2020).

Addressing the «garbage in, garbage out» problem is critical for building trustworthy and ethical artificial intelligence in the future. As artificial intelligence becomes more integrated into everyone's lives, ensuring data quality and mitigating bias is essential (Jobin et al., 2019). In this regard, some efforts are well underway. Standardization and regulations won't solve it by themselves but can help. Developing guidelines and regulations for responsible artificial intelligence development and deployment can improve data quality and fairness[1]. Public education and awareness are also being pursued. Raising awareness about the potential pitfalls of artificial intelligence and the importance of responsible development can foster public trust (Kandlhofer et al., 2023). Additionally, collaboration between artificial intelligence developers and experts, on an interdisciplinary basis, including data scientists, ethicists, and policymakers is crucial for building robust and responsible artificial intelligence systems (Bisconti et al., 2023). This is another trend on the rise that can prevent or moderate succumbing to the pitfalls of «garbage in, garbage out».

4. The Sustainability Problem (almost) no One Talks About

Sustainability of artificial intelligence solutions will be a key factor.

Energy consumption is paramount. Training a large language model like GPT-3 can consume the same amount of energy as several cars in their lifetime. Some studies estimate the energy consumption of training a single large language model to be around 1.5 MWh[2]. Data centers housing artificial intelligence systems are estimated, conservatively, to consume 1% to 3% of global electricity[3].

Also, water consumption is a growing concern. Data centers rely heavily on water for cooling, with estimates suggesting they use up to 1.7 billion gallons of water per year in the United States of America alone[4]. The water footprint of artificial intelligence can be significant, even for individual users. A single query on a large language model can require enough water to fill a small bottle[5].

---

1 European Commission. (2021). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Legislative Instruments (COM(2021) 206 final). https://clck.ru/3DzaGK

2 Luccioni, S. (2023, April 12). The mounting human and environmental costs of generative AI. Ars Technica. https://clck.ru/3DzaKM

3 AI Now Institute. (2023). Algorithmic Accountability: Moving Beyond Audits. https://clck.ru/3DzaLM

4 Meredith, S. (2023, December 6). A 'thirsty' generative AI boom poses a growing problem for Big Tech. CNBC. https://clck.ru/3DzaLy ; Microsoft (2022). 2022 Environmental Sustainability Report. https://clck.ru/3DzaLy

5 Ibid.

However, data storage limitations will probably be the ultimate limiting factor (Susskind, 2020). The amount of data generated globally is growing exponentially, doubling roughly every two years. Current storage technologies like hard disk drives are reaching their physical limitations in terms of miniaturization and storage capacity[6].

It's important to note that this data is constantly evolving as technology advances. Researchers are actively developing more energy-efficient artificial intelligence models, water-saving cooling systems for data centers, and new data storage technologies with higher capacities (Chen, 2016).

Predicting exactly when the entire Earth will be needed to store data is difficult. Data growth is exponential, but storage technology is also constantly evolving. However, it is safe to say the entire planet's physical space for data storage will not be needed in the near future. The world is witnessing a race between exponential growth and Moore's Law (Theis & Wong, 2017). Data creation is indeed growing exponentially, doubling roughly every two years. Even if the growth slows down to 20% per year (it is actually closer to 70%), it would still be an unsustainable growth rate, in the long term. Existing infrastructure and energy limitations would make it increasingly difficult to maintain such a pace[7]. However, storage capacity is also increasing rapidly, following a trend similar to Moore's Law (doubling of transistor density on integrated circuits roughly every two years). Data storage isn't a one-to-one process. Compression techniques can significantly reduce the physical space needed to store information. While data growth might outpace storage capacity at some point, advancements in storage technology like solid-state drives and advancements in data compression techniques can help bridge the gap (Chen, 2016). Arguments are also made that not all data needs forever storage in these systems. A significant portion of data doesn't require permanent storage. Logs, temporary files, and certain types of entertainment content can be deleted after a set period. Focusing on efficient, sustainable data management and prioritizing what gets stored permanently can significantly reduce storage needs (Arass & Souissi, 2018). Nonetheless, in this case one starts to consider that the supposed superintelligences are being brought down to human level. Less than perfect memory implies imperfect solutions, more mistakes made and less reliability: in other words, human like performance. Entirely new but yet distant new storage technologies might delay this inevitability, as researchers are exploring alternative solutions with much higher capacities than traditional hard drives. These include technologies like DNA Storage to achieve vast amounts of data stored in a very compact space (as all life form do in their genome). While still in its early stages, DNA storage holds immense potential for long-term data archiving (Goldman et al., 2013). They include holographic storage, as well, using laser technology

6 Rydning, D., Reinsel, J., & Gantz, J. (2018). The digitization of the world from edge to core. International Data Corporation. https://clck.ru/3DzaNN

7 Ibid.

to store data in three dimensions, offering much higher density than traditional methods (Lin et al., 2020).

Despite all this wishful thinking, Vopson (2020), estimates, in a convincingly rigorous calculation, for several information annual rate growth scenarios, how many years it would take for the entire mass of the Earth to be dedicated to data storage. His approach, as it is based on the number of atoms available, becomes mostly independent of storage efficiency management techniques and eventual, hypothetical new storage technologies. The values attained: 4500 years at 1% growth, 918 years at 5% growth, 246 years at 20% growth, and circa 110 years at 50% growth. The author refers to this eventuality as the impending information catastrophe. Considering that the crust of the Earth comprises only 0.7% of the planet's total volume and that that, even if wildly farfetched, is the part of the volume humanity has a more realistic chance to utilize in its entirety, humans are less than a century away (more likely 30 to 50 year away, or even less) from Vopson's information catastrophe. Studies by the AI Now Institute[8] and the Stanford Institute for Human-Centered Artificial Intelligence[9] might not agree on the exact numbers or not share the same approach but highlight the same trajectory.

5. Governance as Key: How State Measures and Data Availability Reinforce some Organizational Values and Contribute to the Sustainability of the National Health System

It can be argued that the missing link for adequately bridging the gap between traditional health measures and practices, and an artificial intelligence approach to pandemic response is good governance. State measures aligning to the principles of good governance are in a privileged position to become the foundation for artificial intelligence integration with more conventional methods.

Correia et al. (2020a) explore established public health measures that form the bedrock of a robust national health system during pandemics. For one, these authors address lockdowns, contact tracing, and vaccination campaigns as measures that can be implemented to slow the spread of viruses, protect vulnerable populations, and achieve herd immunity, embodying the value of prioritizing public health and demonstrating thar government's responsibility should be towards its citizens. A second important point stressed by these authors is that data plays a crucial role in monitoring infection rates, tracking resource allocation, and understanding patient outcomes, informing decision-making, and that, in turn, reinforces the value of evidence-based practice, and ultimately contributes to the efficient use of resources within the health system.

---

8 AI Now Institute. (2023). Algorithmic Accountability: Moving Beyond Audits. https://clck.ru/3DzaQC

9 Stanford Institute for Human-Centered Artificial Intelligence. (2023). Sustainability and AI. https://clck.ru/3DzaRd

A path, then, emerges. The path of sustainability through efficiency, where artificial intelligence can amplify the impact of traditional measures. While conventional measures are still essential and probably will always be so, artificial intelligence offers the potential to significantly enhance their effectiveness and further contribute to the sustainability of national health systems. One immediate application of this idea can be materialized in que use of artificial intelligence models to analyze historical data, identify patterns, and predict the emergence or spread of future disease outbreaks, epidemics, and pandemics, allowing public health officials to take proactive measures like early warning systems, stockpile of vital supplies, and the strategical deployment of resources. In particular, this optimization of resource allocation by means of artificial intelligence algorithms can be readily put to use in the analysis of real-time data on infection rates, hospital capacity, and material resources availability, allowing for the dynamic allocation of medical staff, equipment, and critical supplies to the areas facing the biggest strain (Correia et al., 2021, 2022), and, therefore, ensuring efficient resource management. More advanced and, consequently, delicate applications encompass personalized treatment plans that require the analysis of individual patient data like medical history and genetic makeup, where artificial intelligence can potentially assist medical professionals in tailoring treatment plans for maximum effectiveness, contributing to faster recovery times, improved patient outcomes, and reduced strain on healthcare resources (Jiang et al., 2017).

It becomes obvious, then, that the effectiveness of governance in pandemic response (whether making use of artificial intelligence or not), hinges on the availability of high-quality, comprehensive data, data gathered through traditional measures and methods like contact tracing and patient records (Wu et al., 2022). However, if governance is good, it will address data privacy concerns regarding the collecting and use of patient data, including for artificial intelligence development, not neglecting the assurance of data anonymization and robust data security protocols to maintain public trust (Smidt & Jokonya, 2021).

Effective use of artificial intelligence in public health requires seamless data sharing between different healthcare institutions and interoperability (O'Reilly-Shah et al., 2020). The need for standardized data formats and secure communication channels to facilitate this data exchange is paramount (Sass et al., 2020).

In conclusion, a symbiotic relationship for sustainable health systems can be implemented. Traditional public health measures, data availability, and artificial intelligence are not separate entities but can be interconnected elements in the fight against pandemics. The existing data infrastructure and experience with traditional measures create a fertile ground for artificial intelligence integration (Baclic et al., 2020). By leveraging the power of artificial intelligence in conjunction with established practices, national health systems can achieve greater efficiency, personalize treatment approaches,

and ultimately ensure their long-term sustainability in the face of future occurrences (Gunasekeran et al., 2021). That is equivalent to affirm that robust management practices and sound organization values can pave the way for future artificial intelligence integration in this crucial domain.

6. Governance as Key: How Governance Reinforces some Organizational Values and Contributes to the Sustainability of Crisis Management

The statement above holds immense significance for the context of artificial intelligence applied to the fight against pandemics. That is because effective governance practices provide the framework and guiding principles for utilizing artificial intelligence responsibly and ethically in crisis management, with pandemics being a prime example.

Governance emphasizes open communication and holding decision-makers accountable for their actions. This is vital for building public trust in artificial intelligence powered solutions used during pandemics, like contact tracing apps. Clear explanations about how artificial intelligence is being used and how data privacy is protected are crucial to avoid public apprehension (Galetsi et al., 2022). Good governance also fosters collaboration, including data sharing, and coordination between diverse stakeholders, including resource allocation, vaccine development, and communication strategies. This collaboration and cooperation comprise government agencies, healthcare institutions, research bodies, and private sector entities (Bulled, 2023).

Good governance translates into specific organizational values that have the potential to shape artificial intelligence development and use in pandemics. Well applied, it promotes equity and fairness, adequate distribution of resources and strives to bridge the digital divide. This, in turn, can ensure that artificial intelligence tools do not exacerbate existing social inequalities (Margetts, 2022). For example, artificial intelligence powered contact tracing apps should be accessible to all demographics and should not unfairly target certain populations. Governance can also be determinant in establishing robust data privacy and security protocols. These types of actions can contribute to protects citizens' information while allowing for responsible data collection and use for artificial intelligence development in pandemic response. Striking a balance between data-driven insights and data security is crucial during a pandemic (Zhang et al., 2022). Moreover, good governance fosters evidence-based decision making, a culture of relying on data and scientific evidence to inform decisions. This aligns perfectly with the core principle of artificial intelligence, which uses data analysis to generate insights and recommendations for, amongst others, public health officials (Rubin et al., 2021).

In addition, effective and sustainable pandemic response requires a forward-looking approach and long-term planning. Good governance practices contribute to the sustainability of crisis management in several ways. Governance is tasked with ensuring long-term investment in infrastructure and maintaining the hardware, software, and expertise needed for artificial intelligence development and deployment in public health. This includes investment in research and development, training programs for artificial intelligence specialists within healthcare institutions, and establishing robust data management systems (Balog-Way & McComas, 2022). Governance also promotes the futureproofing of strategies, the development of flexible frameworks that can adapt to evolving threats and pandemics with novel characteristics. This ensures that artificial intelligence remains relevant and useful for future public health challenges. For example, artificial intelligence algorithms for pandemic prediction need to be adaptable to handle new virus strains and variations. Good governance has, as well, the ability to build and strengthen public trust in government institutions and their use of artificial intelligence during a pandemic (Romano et al., 2021). This trust fuels cooperation with artificial intelligence suggested initiatives, like contact tracing and symptom-tracking apps.

Correia et al. (2020b), explore the principles that create a strong foundation for solutions in pandemic response. By fostering collaboration, prioritizing ethical values, and ensuring long-term sustainability, governance practices pave the way to become a powerful weapon in the arsenal for combatting pandemics and building a more resilient future for public health. The authors propose a six-dimensions, eight-hypothesis model (already validated for very specific circumstances) represented on figure 1.

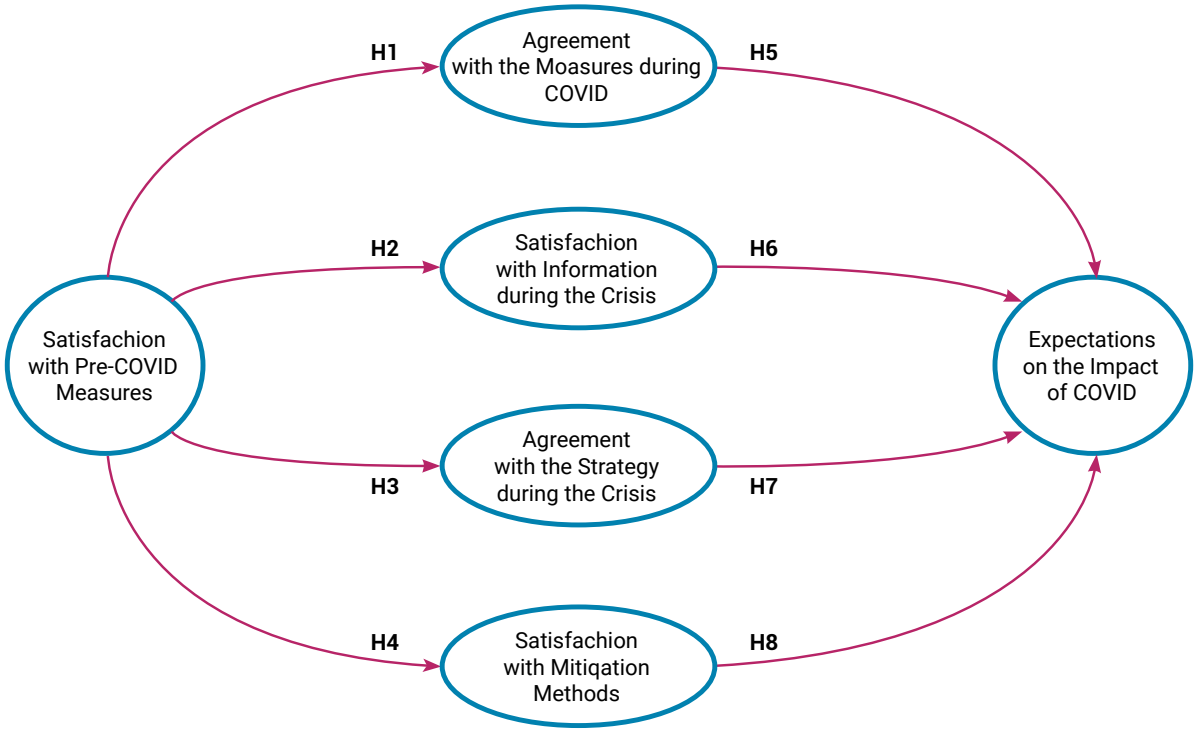


Figure 1. Crisis management: COVID-19 structural model (Correia et al., 2020b)

It can be reasoned that the incorporation on these inputs (models, dimensions and relations between dimensions), the human factors, if you will, is the link that has been missing in the creation of a supportive environment, through effective governance, for artificial intelligence to be leveraged responsibly and ethically, while also increasing their performance in predicting, preparing and warning for outbreaks, in predicting and preparing for individuals' responses to public health measures, in optimizing real time strategic resource allocation based on infection rates, in developing targeted, personalized interventions for high-risk individuals, and in tracking, containing, isolating and limiting spread of pathogens.

The linkage of good governance, crisis management models and artificial intelligent solutions, and the synergies thus generated hold immense potential to improve our preparedness and response to future pandemics, ultimately saving lives and ensuring a more sustainable future for global health.

7. Critical Reflections

Artificial intelligence represents one of the greatest technological innovations of the modern era, with the potential to fundamentally transform society in many aspects (Bostrom, 2014). However, this transformation also brings with it significant challenges and concerns about the fragility of modern human societies.

Digital inequality presents a pressing concern as the integration of artificial intelligence into various sectors may widen socioeconomic gaps, accentuating disparities between individuals equipped with access and proficiency in leveraging technological advancements and those marginalized without such resources (Eubanks, 2018). This phenomenon has the potential to intensify preexisting inequities while giving rise to novel manifestations of digital exclusion.

The advent of artificial intelligence driven automation also poses a significant risk of displacing numerous conventional occupations, particularly those characterized by routine and predictable duties. This transition holds the potential to precipitate widespread job loss, plunging affected individuals into a state of mass unemployment and instigating profound existential crises rooted in the displacement by technological innovations. The health sector appears to be no exception to this peril (Hazarika, 2020).

Moreover, artificial intelligence algorithms, especially in crucial domains such as justice, healthcare, and finance, are prone to manipulation and bias, posing a risk of unfair and harmful consequences, particularly for marginalized and vulnerable communities (Obermeyer, 2019).

Furthermore, concerns regarding data privacy and security are heightened by the extensive collection and analysis of data powering artificial intelligence algorithms, with lack of transparency and control over data usage posing as a predictable consequence and a significant threat to public trust in technology, public policy, and public institutions (Larsson & Heintz, 2020).

Additionally, societies around the world have to deal, increasingly, with polarization and misinformation, manifested by the so-called misuse of digital platforms driven by artificial intelligence, that can undermine social cohesion and erode trust in democratic institutions, leading to an increasingly fragmented and divisive society (Kavanagh & Rich, 2018).

One must not ignore, as well, the potential dangers of becoming overly reliant on technology. The more humanity depends on artificial intelligence for decision-making and task completion, the weaker human's ability to function independently becomes (Bostrom, 2014). This vulnerability to disruptions and systemic failures in artificial intelligence could have devastating consequences.

Also, artificial intelligence is creeping in, or, in a more technical terminology, the erosion of human autonomy and agency is ever more noticeable (Ettlinger, 2022). The increasing integration of artificial intelligence into people lives could lead to the erosion of human autonomy and agency, as we increasingly leave important decisions to automated systems. This raises questions about who controls technology and who can one trust to make decisions that affect one's life.

Finally, this first layer, superficial challenges analysis must include not only immediate concerns, but the often-mentioned existential risks, that comprise long-term fears about the development of artificial intelligence, including the common scenarios of artificial superintelligences that surpass human control and threaten the survival of humanity (Bostrom, 2014).

In essence, the convergence of artificial intelligence with the delicate nature of contemporary human societies prompts profound inquiries into ethics, governance, fairness, and human principles. A comprehensive and cooperative approach is vital to tackle these matters, ensuring that artificial intelligence advancement and implementation prioritize human welfare and enduring sustainability (Jobin et al., 2019).

However, it is possible to add several layers of depth in a conscious quest to use artificial intelligence in a responsible, productive, and secure manner.

Artificial intelligence appears to be quite fragile when stressed in just the right way. When stressed in just the right way, artificial intelligence systems can indeed exhibit fragility or vulnerability. This fragility can manifest in various ways, depending on the context and nature of the artificial intelligence systems. For example, adversarial attacks can happen, involving intentionally manipulating input data to artificial intelligence systems in a way that causes them to make mistakes or produce incorrect outputs. These attacks can exploit vulnerabilities in the artificial intelligence algorithms, such as deep learning neural networks, leading to unexpected and potentially harmful behavior (Ruan et al., 2021). Another example, similar to one previously addressed through a different lens, is the use of unrepresentative data to train artificial intelligence models, leading to the above-mentioned biased decisions or predictions. This bias can be exacerbated under certain conditions or when the artificial intelligence system encounters new, unseen data that

differs significantly from the training data sets. As a result, the artificial intelligence system may fail to generalize effectively, leading to fragility in its performance (Navigli et al., 2023). One more example consists in what can be called catastrophic forgetting, as some artificial intelligence systems, particularly those based on artificial neural networks, can exhibit catastrophic forgetting when exposed to new data. This phenomenon occurs when the artificial intelligence system forgets previously learned information as it learns new information, leading to a loss of performance or accuracy over time. This fragility can limit the system's ability to adapt to changing environments or tasks (Kirkpatrick et al., 2017). Yet another example is model fragility, given that artificial intelligence models can be fragile to small changes in input data or model parameters. For example, slight perturbations to input images can cause image recognition models to misclassify objects, leading to potentially dangerous consequences in applications such as autonomous vehicles or medical diagnosis (Chen et al., 2020). One final example of artificial intelligence fragility is the inherent system complexity. As artificial intelligence systems become more complex and interconnected, they can become increasingly fragile to disruptions or failures in individual components. A failure in one part of the system can cascade into other parts, leading to system-wide failures or breakdowns (Chen et al., 2020). Complexity in a double-edged sword as many systems derive their power from their complexity and ability to process vast amounts of data. However, this complexity can also be a vulnerability, as it increases the surface area for potential attacks and makes the system more difficult to understand and secure. This fragility underscores the importance of robustness and resilience in artificial intelligence system design and deployment. Addressing the fragility of artificial intelligence systems requires careful attention to design, testing, and validation processes, as well as ongoing monitoring and maintenance. It also highlights the need for transparency, accountability, and ethical considerations in the development and deployment of artificial intelligence technologies. By addressing these challenges, individuals, institutions, and societies can work towards creating artificial intelligence systems that are more robust, reliable, and trustworthy in a wide range of applications. Just imagine the pandemonium that would result of such types of fragilities occurring separately or in chain, in the health sector, during an epidemic or pandemic crisis.

An additional convolution surfaces when one comprehends that artificial intelligence pathways, even if they appear robust, can often be undermined by the presence of choke points, which are critical junctures where failure or disruption can have cascading effects on the entire system. This duality, where robustness and fragility coexist, is inherent to many complex artificial intelligence systems (Zhou et al., 2024). For instance, as seen before, artificial intelligence models often involve complex networks of interconnected components, such as layers in deep neural networks or nodes in graph-based models. While this interconnectedness can enhance robustness by allowing for redundancy and fault tolerance, it also introduces choke points where failure or disruption in a critical

component can propagate throughout the network (Villegas-Ch et al., 2024). Another instance is the emergence of critical dependencies. Certain components of artificial intelligence applications may serve as critical dependencies, upon which the functionality of the entire system relies. These choke points can include specific layers or nodes in neural networks that play pivotal roles in processing or decision-making. If these critical dependencies fail or malfunction, it can lead to a breakdown in the system's performance (Macrae, 2022). Still another instance is the tendency of artificial intelligence usages to be sensitive to input data, especially in uses such as image recognition or natural language processing. Small perturbations or adversarial inputs at choke points within the pathway can lead to significant changes in the system's output. This sensitivity underscores the fragility of the pathway to specific types of input manipulation (Dhingra & Gupta, 2017). One final instance that illustrates the dangers of choke points are trade-offs in Design. These trade-offs are between robustness and efficiency. Strategies aimed at enhancing robustness, such as adding redundancy or error correction mechanisms, may introduce additional choke points or computational overhead. Conversely, optimizations for efficiency may inadvertently increase the system's fragility by reducing redundancy or resilience. Addressing these specific problems requires a multi-faceted approach that involves identifying and mitigating choke points and enhancing robustness through redundancy and diversity (Goodfellow et al., 2016). By understanding these delicate balances, researchers, engineers and practitioners can work towards creating more resilient and trustworthy artificial intelligence technologies.

Artificial intelligence intricacy expands when the concept of losing control acquires a broader and more nuanced meaning. That is to say, when the concept reflects the complexities and challenges associated with the development and deployment of artificial intelligence technologies. One such manifestation relates to autonomy and decision-making (Wallach & Allen, 2008). As artificial intelligence systems become increasingly autonomous and capable of making decisions without direct human intervention, there is a concern about losing control over the outcomes of these decisions. This can be particularly relevant in high-stakes applications such as autonomous vehicles, or medical appliances where the actions of artificial intelligent systems can have real-world, life-threatening consequences. Another manifestation is the inclination of artificial intelligent systems, especially those based on deep learning and neural networks, to be highly opaque, making it difficult for humans to understand or interpret their internal workings. This lack of transparency can lead to a loss of control over how artificial intelligent systems arrive at their decisions, raising concerns about accountability and trust (Chiao, 2019). One other manifestation is patent in the fact that artificial intelligence systems can exhibit emergent behavior, where complex patterns or behaviors arise from the interactions of simple components. This emergent behavior can be difficult to predict or control, leading to uncertainty about the behavior of the systems in novel or unanticipated situations. That, in itself, can lead to unintended consequences, as the actions or decisions of artificial

intelligence functions produce outcomes that were not anticipated or intended by their creators. This can occur due to unexpected interactions with the environment, or a range of other factors (Bostrom, 2014). Last but not least, the ethical and societal implications must be considered. Losing control over artificial intelligent technologies can also bear broader ethical and societal implications, such as the impact of artificial intelligence on employment, privacy, security, and inequality, as addressed in the course of this text (Thomsen, 2019). These concerns highlight the need for responsible artificial intelligence development and governance to ensure that artificial intelligent technologies are deployed in ways that benefit society as a whole and not only oligarchies, big tech companies and the ruling elites. Addressing these issues requires a holistic approach that encompasses technical, ethical, and regulatory considerations. This includes promoting transparency and accountability in artificial intelligent systems and engaging in ongoing dialogue and collaboration between stakeholders to mitigate risks and maximize the benefits.

Another complication rises as one ponders on the premise of reliability, particularly during times of stress or uncertainty. This premise may warrant urgent reevaluation. One illustration of ill placed premises refers to interconnectedness, as artificial intelligence systems often comprise intricate networks of interconnected components, each contributing to overall functionality. When stress or unexpected conditions arise, such as adversarial attacks, data anomalies, or environmental changes, the complexity of these systems can amplify the likelihood of failure. This underscores the need for a more nuanced understanding of reliability beyond traditional measures (Macrae, 2022). Another illustration is unpredictability. The emergent behavior exhibited by artificial intelligent systems can lead to unpredictability in their responses to stressors. Even minor perturbations or variations in input data can trigger unexpected outcomes, highlighting the challenges of ensuring reliability under diverse conditions. This unpredictability underscores the importance of robustness testing and scenario planning to identify and mitigate potential failure points (Bostrom, 2014). These exact concerns were expressed, in different combinations, above. Yet another illustration of this class of phenomena deals with adaptive and evolving environments, given that artificial intelligence systems operate within dynamic and everchanging settings, where conditions may change rapidly and unpredictably. In such environments, the notion of reliability as a static attribute becomes inadequate. Instead, reliability must be viewed as a dynamic property that adapts to changing circumstances, requiring continuous monitoring, adaptation, and feedback mechanisms (Sundar, 2020). One final illustration focusses on the closely intertwined relation between these types of systems and the human-machine interaction. Human operators play a critical role in monitoring system performance, interpreting outputs, and intervening when necessary. However, under stress or high-pressure situations, human operators may also be prone to errors or cognitive biases, further complicating the reliability of the systems (Hoff & Bashir, 2015). In light of this set of challenges, rethinking the premise of reliability in artificial intelligence necessitates a shift towards more adaptive, resilient,

and context-aware approaches. This may involve incorporating principles of uncertainty quantification, robustness engineering, and human-centered design into the development and deployment of artificial intelligence models and applications. By embracing a broader understanding of reliability and proactively addressing the factors that contribute to failure, humanity can strive towards more trustworthy and dependable artificial intelligence technologies. Preparing and having contingency, artificial intelligence free, plans that allow society to continue functioning in case of technological crisis or collapse should be of the utmost priority. Would most, if not all, developed countries still have a justice system tomorrow, if the internet failed now that justice has been "dematerialized"? Do we really want to take such risks, ones that may bring societies to a halt or to altogether collapse?

A deeper yet layer, one more hurdle to be surpassed is the necessity to shift from a premise of reliability to one of risk regarding artificial intelligence. That is equivalent to a move from fixity to nimbleness in responding to changing circumstances. One good example in the understanding of risk. Reliability is often associated with the notion of deterministic outcomes and predictable behavior. However, in complex and dynamic environments, such as those encountered by artificial intelligence systems, complete reliability may be unattainable. Recognizing this, the focus shifts towards understanding and managing risk — the likelihood and impact of adverse events or uncertainties (Bigham et al., 2019). This embracing of risk implies acknowledging the inherent uncertainty in artificial intelligence systems and embracing adaptive strategies to cope with it. Rather than striving for absolute reliability, artificial intelligence systems should be designed to be resilient and responsive in the face of changing circumstances. This may involve incorporating mechanisms for real-time monitoring, dynamic adjustment, and learning from experience (Syed et al., 2023). This nimbleness in responding to changing circumstances requires agility and flexibility in these applications, incorporating the ability to quickly assess risks, identify opportunities, and adapt behavior or decision-making strategies accordingly. Agile artificial intelligence systems will be increasingly capable of dynamically allocate resources, prioritize tasks, and adjust to new information or objectives as they arise. Shifting from a mindset of reliability to a mindset of risk necessitates the development of robust risk management frameworks. These frameworks must provide a systematic approach to identifying, assessing, mitigating, and monitoring risks throughout the lifecycle of artificial intelligence usages. By proactively managing risks, organizations can enhance resilience and reduce the likelihood of adverse outcomes (Jobin et al., 2019). Naturally, a risk-aware artificial intelligence use needs to be centered in a continuous learning and improvement capacity, leveraging feedback loops, experimentation, and data-driven insights to iteratively enhance performance and adapt to evolving challenges. This iterative approach will increasingly allow artificial intelligence systems to refine their strategies over time and become more effective in managing risks. The final piece of this particular puzzle will be put in place by recognizing the limitations of artificial intelligent

systems in navigating complex and uncertain environments, and, as a consequence, increasing emphasis on human-in-the-loop approaches (Russell & Norvig, 2021). By integrating human judgment, expertise, and oversight, artificial intelligent systems can more adequately complement human decision-making, mitigate risks, and enhance overall system performance (Parasuraman & Riley, 1997). In general, the shift from a premise of reliability to one of risk reflects a broader recognition of the inherent uncertainties and complexities of real-world applications. By embracing risk and fostering nimbleness in responding to changing circumstances, artificial intelligent models can better navigate uncertain terrain, adapt to evolving challenges, and ultimately deliver greater value and impact in diverse domains, to which healthcare is no stranger.

The next in-depth layer is one in which artificial intelligence can appear to disguise weaknesses as strengths, especially when it comes to certain types of machine learning models or algorithms. The most blatant example is the imbalance between generalization and overfitting, occurring when a model learns to perform well on the training data but fails to generalize to new, unseen data (Igual & Seguí, 2024). This can give the illusion of strength because the model appears to perform exceptionally well on the data it was trained on. However, when exposed to new data, the weaknesses of the model become apparent as it fails to make accurate predictions. A visual recognition model can easily learn to identify ties, on pictures, and associate that to males, if trained with a biased Wall Street executive data set. This can happen because artificial intelligent models aim to generalize patterns from training data to make predictions on unseen data. While strong generalization is desirable, over-reliance on specific patterns in the training data can lead to overfitting, where the model fails to generalize effectively. Recognizing the balance between generalization and overfitting is crucial for ensuring robustness. Thoroughly testing artificial intelligence systems on diverse datasets, scrutinizing their decision-making processes, and mitigating biases and vulnerabilities, allows users to uncover, and address weaknesses disguised as strengths, leading to more reliable and trustworthy solutions.

As the analysis keeps going deeper, it arrives at the realization thar artificial intelligence cultivates instability. A collective that has come to expect that things have and will operate without interruption and that blockages are either serendipitous or negligible in impact is, undoubtedly, much less prepared when these principles are threatened. The dependency on artificial intelligence is becoming increasingly integrated into various aspects of society and there is a growing dependency on their functionality. Individual, organizations, governments, and supranational institutions (like the United Nations and the World Health Organization) rely evermore on artificial intelligence for decision-making, automation, and optimization of processes (Bostrom, 2014). However, this dependency can create instability if these functions experience failures or disruptions. This widespread adoption of artificial intelligent solutions may foster a collective expectation of continuity and seamless operation. When these solutions function as expected,

they reinforce the perception that contrarieties are minimal. However, this expectation can lead to complacency and vulnerability if systems encounter unexpected challenges or malfunctions (Parasuraman et al., 2000). It follows that, when artificial intelligence applications fail or encounter blockages, the impact can be significant, especially if they are relied upon for critical tasks or services. Disruptions in technological-driven processes can disrupt supply chains, financial markets, communication networks, and other essential functions and critical systems (like nuclear power stations), leading to economic losses, social unrest, or even safety risks. Cultivating resilience and adaptability in the face of this potential instability requires proactive measures to anticipate and mitigate risks, as mentioned prior. This may involve diversifying technological dependencies, building redundancy into critical systems, and developing human-centric approaches to decision-making and problem-solving (Bigham et al., 2019; Jobin et al., 2019). In summary, artificial intelligence technologies also pose challenges related to stability and resilience. By recognizing the potential for instability inherent in these systems and taking proactive measures to address risks, stakeholders can better navigate the complexities of an artificial intelligence driven world and build more robust and sustainable systems.

At last arrived at the deepest place, the center of and coldest place in, the universe, according to Aristotle, the center of Earth and Hell, according to Dante, one must contemplate artificial intelligence in the light of what can be called Luciferian semiotics, a voyage into the symbolic or metaphorical implications of artificial intelligence. In various mythologies and belief systems, Lucifer (Luciferian symbolism) is often associated with themes of rebellion, enlightenment, and the pursuit of knowledge. Lucifer is often depicted as the carrier of the light (Hanegraaff, 2013). The term Luciferian may thus connote the pursuit of knowledge or power that challenges established norms or authority structures. Semiotics refers to the study of signs and symbols and their interpretation. In the context of artificial intelligence, semiotics encompasses the symbolic meanings associated with artificial intelligence, including notions of intelligence, autonomy, and control (Binder, 2024). Artificial intelligence is often perceived as a symbol of strength and capability, given its ability to process vast amounts of data, make complex decisions, and automate tasks with efficiency. This perception of strength may be reinforced by the impressive feats accomplished by these solutions in various domains. However, as shown throughout the text, objects or entities that appear strong may, in fact, possess vulnerabilities or weaknesses that are not immediately apparent. This reversal of expectations can be seen as a manifestation of the Luciferian symbolism, where the pursuit of knowledge or power leads to a reevaluation of established truths or assumptions. Artificial intelligence models and systems, despite their perceived strength, can exhibit vulnerabilities or limitations in certain contexts. These weaknesses may become more visible over time as artificial intelligence technologies are subjected to scrutiny, experimentation, and real-world deployment. The exploration of the

Luciferian semiotics raises broader ethical and philosophical questions (Bostrom, 2014) about the nature of power, knowledge, and control in the age of artificial intelligence. It prompts reflections on the unintended consequences of technological advancement and the need for responsible stewardship of all technologies. In summary, the Luciferian semiotics applied to artificial intelligence invites us to consider the symbolic meanings and implications of artificial intelligence, including the ways in which perceptions of strength and power may be subverted or challenged by deeper exploration and understanding. It underscores the importance of critical inquiry and ethical reflection in navigating the complexities of artificial intelligence and its impact on societies, public policies, and political systems.

Given all the above, the use of artificial intelligence in the combat against pandemics, quid juris? Maybe, probably, humanities best course of action is to continue to rely on the human factor. Humans make more errors, there is no question about that. But most of those errors are small and inconsequent. They may originate individual tragedies but not global ones. Artificial intelligent models to deal with outbreaks, epidemics, and pandemics, as well as other artificial intelligence reliant medical applications might be almost error free. But that one error may doom all.

8. AI and Fundamental Rights

The impact of AI on fundamental rights is so relevant that it has not gone unnoticed by the AI Act, which in article 27 requires that high-risk systems must be subject to an impact assessment on fundamental rights. The main purpose of the impact assessment of AI systems is to identify and mitigate the potential risks that these systems may pose to people's fundamental rights. This is especially important when it comes to high-risk AI systems, which have the potential to significantly affect people's lives and well-being. To summarise, in view of the negative impacts of AI, we have not opted to abandon AI, but rather, in order to take advantage of its positive impacts, to classify AI systems according to risk and through their prior, concomitant and a posteriori control.

There are several relationships that can be established between AI and fundamental rights, in particular the impact that AI can have on the fulfilment of fundamental rights and, in a different sense, the impact that AI can have on the violation of fundamental rights. In any case, it is reasonable to believe that, as the FRA[10] points out, even in a restricted context, the lack of a large body of empirical data on the wide range of rights involved in AI makes it fragile to provide the necessary safeguards to ensure that the use of AI is effectively in line with fundamental rights.

[10] FRA – European Union Agency for Fundamental Rights, Getting the future right – Artificial intelligence and fundamental rights – Report, Publications Office of the European Union, 2020.

The main argument in favour of using AI is efficiency (Pedro, 2023). As far as the challenges are concerned, the main concerns are the violation of fundamental rights. Thus, among the main candidates for fundamental rights potentially harmed by AI are the right to the protection of personal data (Gómez Abeja, 2022) and the right to non-discrimination (Gómez Abeja, 2022), the right to effective judicial protection (Shaelou & Razmetaeva, 2023), the right to freedom of information, the right to suffrage and the right of access to public information (Gómez Abeja, 2022).

Returning to the work of the FRA[11], the use of AI can have an impact on fundamental rights, imposing the need to guarantee the non-discriminatory use of AI (right to non-discrimination); the requirement to process data lawfully (right to personal data protection); and the possibility of lodging complaints about AI-based decisions and lodging appeals (right to an effective remedy and to an impartial tribunal).

Finally, it should also be emphasised that the relationship between AI and fundamental rights can be richer, at least in the following dimensions: confronting implicit fundamental rights (Gómez Colomer, 2023), such as the principle of the rule of law and the principle of the natural judge, and the emergence of «new» or «renewed» fundamental rights (Shaelou & Razmetaeva, 2023), such as the right to be forgotten (Gómez Abeja, 2022), the «right not to be subject to automatic decisions and automatic treatment» in the broad sense (Shaelou & Razmetaeva, 2023); the «right to influence your digital footprint» (Shaelou & Razmetaeva, 2023), and new rights, such as the 'right not to be manipulated', the 'right to be informed neutrally online' and the 'right to meaningful human contact', the 'right not to be measured, analysed or trained' (Shaelou & Razmetaeva, 2023).

9. Pandemics and Fundamental Rights

The pandemic situation, as happened with COVID-19, called for the admission of public legal regimes of exceptionality (alongside normality regimes), which is nothing new – before all time (Gomes & Pedro, 2020) – just by looking at the Latin brocardo «Necessitas non habet legem, sed ipsa sibi facit legem». It was this brocardo that justified extraordinary powers in Roman law, exercisable in cases where it was necessary to deal with an unforeseeable situation that required an immediate decision, with no possibility of postponement.

The need for a legal system of exceptionality, and therefore its mobilisation, has become more evident in recent times. The configuration of the current risk society (Beck, 1986) and the fact that we live in a globalised world (economically and socially) in which, despite physical distance, everything seems to be close by, as the (still) current health crisis caused by the outbreak of COVID-19 - which in the space of a few months has spread from its source (China – Wuhan city) to the whole world (Pedro, 2022) – has greatly contributed to this.

[11] Ibid.

In the face of disasters of this kind, public law could not and cannot remain indifferent, in other words, given the damaging effects that public disasters have on the «salus populi», it is easily understandable that public bodies must use all the means at their disposal to restore normality (Alvarez Garcia, 1996). Therefore, in order to guarantee the rule of law, it is essential to provide for regimes that are flexible enough to respond to public interests that are threatened – regimes that make it possible to respond to states of public necessity, or, in other words, public law regimes of exceptionality.

Within a framework of real normality, public law is governed by the principle of the legality of public action – which therefore corresponds to a framework of legal normality. The problem arises whenever reality temporarily changes in a radical way, creating situations of imminent or real danger for the community and in which the public law of normality does not offer an adequate response, and the idea of maintaining the democratic rule of law imposes the need for exceptional legal regimes to come into play – «jus extremae necessitatis» – so that normality is restored in the short term and the legal regimes of normality return to force. What is at stake is an alternative legality, an exceptional legality of exceptionality (Correia, 1987) – a substitute and temporary legality.

Thus, as a rule, in exceptional situations, a state of siege or a state of emergency, respecting the principle of proportionality, some fundamental rights can be suspended. Despite this permission, it should be noted that not all fundamental rights can be suspended, as is the case with the rights to life, personal integrity, personal identity, civil capacity and citizenship, the non-retroactivity of criminal law, the right of defence of defendants and freedom of conscience and religion.

10. Possible Relationships between AI and Fundamental Rights in Combating Pandemics

In democratic states governed by the rule of law, the consideration of the use of AI to combat pandemics generally involves respect for fundamental rights. This requires, on the one hand, consideration of the impact that the use of AI has on certain fundamental rights, taking into account the risks that each specific AI system entails and, on the other hand, that the context of a pandemic, as happened with COVID-19, calls for a legality of exception, which may allow for the restriction of certain fundamental rights, with a view to safeguarding values such as public health, in order to restore a situation of normality.

Conclusions

Struggling diseases and pandemics today requires a comprehensive approach, uniting the efforts of doctors, medical organizations and states. Artificial intelligence is a promising tool, capable of radically changing the methods of counteracting epidemics and pandemics. Its potential lies in analyzing big data, forecasting disease outbursts, accelerating development of medication, personalization of treatment and optimization

of resources distribution. Examples of AI use, such as early detection of outburst through analysis of social networks data, facilitation of search for medication and improved contact tracing, demonstrate its significance in struggling global threats to health.

However, introduction of AI into healthcare is accompanied by a number of challenges. These include data bias problems, algorithm complexity, risks of excessive dependence on technologies, and ethical dilemmas related to fundamental rights. Using AI to struggle pandemics requires observing a balance between innovations, ethics and human rights protection, including the right to privacy, freedom and equal access to medical assistance.

Hence, AI, despite its revolutionary capabilities, is not a panacea. Its use should be accompanied by a critical analysis of potential risks and requires developing legal and ethical mechanisms to ensure safe and fair use of technologies. Only after addressing these aspects, AI may become an effective tool to struggle diseases without threatening fundamental rights and freedoms.

References

Arass, M., & Souissi, N. (2018). Data lifecycle: from big data to SmartData. In *2018 IEEE 5th International Congress on Information Science and Technology* (pp. 80–87). IEEE. https://doi.org/10.1109/CIST.2018.8596547

Alvarez Garcia, V. (1996). *El concepto de necesidad en derecho público* (1st ed.). Madrid: Civitas. (In Spanish).

Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, *46*(6), 161–168. https://doi.org/10.14745/ccdr.v46i06a02

Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, *8*(2), e188-e194. https://doi.org/10.7861/fhj.2021-0095

Beck, U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp Verlag.

Balog-Way, D., & McComas, K. (2022). COVID-19: Reflections on trust, tradeoffs, and preparedness. In *COVID-19* (pp. 6–16). Routledge.

Bazarkina, D. Y., & Pashentsev, E. N. (2020). Malicious use of artificial intelligence. *Russia in Global Affairs*, *18*(4), 154–177. https://doi.org/10.31278/1810-6374-2020-18-4-154-177

Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *International Journal of Environmental Research and Public Health*, *15*(12), 2796. https://doi.org/10.3390/ijerph15122796

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, *48*(3), 386–398. https://doi.org/10.2307/2095230

Bigham, G., Adamtey, S., Onsarigo, L., & Jha, N. (2019). Artificial Intelligence for Construction Safety: Mitigation of the Risk of Fall. In K. Arai, S. Kapoor, R. Bhatia (Eds.). *Intelligent Systems and Applications*. Springer. https://doi.org/10.1007/978-3-030-01057-7_76

Binder, W. (2024). Technology as (dis-)enchantment. AlphaGo and the meaning-making of artificial intelligence. *Cultural Sociology*, *18*(1), 24–47. https://doi.org/10.1177/17499755221138720

Bisconti, P., Orsitto, D., Fedorczyk, F., Brau, F., Capasso, M., De Marinis, L., ... & Schettini, C. (2023). Maximizing team synergy in AI-related interdisciplinary groups: an interdisciplinary-by-design iterative methodology. *AI & Society*, *38*(4), 1443–1452. https://doi.org/10.1007/s00146-022-01518-8

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Box, G. (1979). Robustness in the strategy of scientific model building. In R. Launer & G. Wilkinson (Eds.), *Robustness in Statistics* (pp. 201–236). Academic Press. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231. https://doi.org/10.1214/ss/1009213726

Bulled, N. (2023). "Solidarity:" A failed call to action during the COVID-19 pandemic. *Public Health in Practice*, *5*, 100379. https://doi.org/10.1016/j.puhip.2023.100379

Chen, A. (2016). A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, *125*, 25–38. https://doi.org/10.1016/j.sse.2016.07.006

Chen, J., Zhang, R., Han, W., Jiang, W., Hu, J., Lu, X., Liu, X., & Zhao, P. (2020). Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment. *Journal of Advanced Transportation*, 2020, 6649867. https://doi.org/10.1155/2020/6649867

Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, *15*(2), 126–139. https://doi.org/10.1017/S1744552319000077

Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020a). The combat against COVID-19 in Portugal: How state measures and data availability reinforce some organizational values and contribute to the sustainability of the National Health System. *Sustainability*, *12*(18), 7513. https://doi.org/10.3390/su12187513

Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020b). The combat against COVID-19 in Portugal, Part II: how governance reinforces some organizational values and contributes to the sustainability of crisis management. *Sustainability*, *12*(20), 8715. https://doi.org/10.3390/su12208715

Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2022). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. *European Journal of Applied Business Management*, *8*(1), 1–12.

Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2021). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. In *European Consortium for Political Research General Conference* (pp. 1–18). United Kingdom.

Correia, J. M. C. (1987). *Legalidade e autonomia contratual nos contratos administrativos* (pp. 283, 768). Lisboa: Almedina.

DeCamp, M., & Tilburt, J. (2019). Why we cannot trust artificial intelligence in medicine. *The Lancet Digital health*, *1*(8), e390. https://doi.org/10.1016/S2589-7500(19)30197-9

Dhingra, M., & Gupta, N. (2017). Comparative analysis of fault tolerance models and their challenges in cloud computing. *International Journal of Engineering & Technology*, *6*(2), 36–40. https://doi.org/10.14419/ijet.v6i2.7565

Ettlinger, N. (2022). *Algorithms and the Assault on Critical Thought: Digitalized Dilemmas of Automated Governance and Communitarian Practice* (1st ed.). Routledge. https://doi.org/10.4324/9781003109792

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: Picador, St Martin's Press.

Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. *Nature*, *442*(7101), 448–452. https://doi.org/10.1038/nature04795

Fetzer, T., & Graeber, T. (2021). Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(33), e2100814118. https://doi.org/10.1073/pnas.2100814118

Galetsi, P., Katsaliaki, K., & Kumar, S. (2022). The medical and societal impact of big data analytics and artificial intelligence applications in combating pandemics: A review focused on Covid-19. *Social Science & Medicine*, *301*, 114973. https://doi.org/10.1016/j.socscimed.2022.114973

Gianfrancesco, M., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, *178*(11), 1544–1547. https://doi.org/10.1001/jamainternmed.2018.3763

Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., & Birney, E. (2013). Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, *494*(7435), 77–80. https://doi.org/10.1038/nature11875

Gomes, C. A., & Pedro, R. (Coords.). (2020). *Direito administrativo de necessidade e de excepção*. Lisboa: AAFDL.

Gómez Abeja, L. (2022). Inteligencia artificial y derechos fundamentales. In F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (1.ª ed., pp. 91–114, 93). Murcia: Ediciones Laborum. (In Spanish).

Gómez Colomer, J.-L. (2023). *El juez-robot: La independencia judicial en peligro*. Valencia: Tirant lo Blanch. (In Spanish).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Greiner, R., Grove, A., & Kogan, A. (1997). Knowing what doesn't matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, *97*(1–2), 345–380. https://doi.org/10.1016/S0004-3702(97)00048-9

Gunasekeran, D., Tseng, R., Tham, Y., & Wong, T. (2021). Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digital Medicine*, *4*(1), 40. https://doi.org/10.1038/s41746-021-00412-9

Gürsoy, E., & Kaya, Y. (2023). An overview of deep learning techniques for COVID-19 detection: methods, challenges, and future works. *Multimedia Systems*, *29*(3), 1603–1627. https://doi.org/10.1007/s00530-023-01083-0

Hanegraaff, W. (2013). *Western Esotericism: A Guide for the Perplexed*. Bloomsbury Publishing.

Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2), 8–12. https://doi.org/10.1109/MIS.2009.36

Hazarika, I. (2020). Artificial intelligence: opportunities and implications for the health workforce. *International Health*, *12*(4), 241–245. https://doi.org/10.1093/inthealth/ihaa007

Hoff, K., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hulten, G. (2018). Building Intelligent Systems: A Guide to Machine Learning Engineering. Apress.

Igual, L., & Seguí, S. (2024). *Supervised learning. In Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications* (pp. 67–97). Springer International Publishing.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jones, K., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993. https://doi.org/10.1038/nature06536

Kandlhofer, M., Weixelbraun, P., Menzinger, M., Steinbauer-Wagner, G., & Kemenesi, Á. (2023). Education and Awareness for Artificial Intelligence. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 3–12). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44900-0_1

Kavanagh, J., & Rich, M. (2018). *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. RAND Corporation. https://doi.org/10.7249/RR2314

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, *114*(13), 3521–3526. https://doi.org/10.1073/pnas.1611835114

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. E*uropean Journal of Information Systems*, *31*(3), 388–409. https://doi.org/10.1080/0960085X.2021.1927212

Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1469

Lin, X., Liu, J., Hao, J., Wang, K., Zhang, Y., Li, H., ... & Tan, X. (2020). Collinear holographic data storage technologies. *Opto-Electronic Advances*, *3*(3), 190004. https://doi.org/10.29026/oea.2020.190004

Little, R. J., & Rubin, D. B. (2019). Statistical analysis with missing data. John Wiley & Sons.

Macrae, C. (2022). Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. *Risk Analysis*, *42*(9), 1999–2025. https://doi.org/10.1111/risa.13850

Margetts, H. (2022). Rethinking AI for good governance. *Daedalus*, *151*(2), 360–371. https://doi.org/10.1162/daed_a_01922

Matsuzaka, Y., & Yashiro, R. (2022). Applications of Deep Learning for Drug Discovery Systems with BigData. *BioMedInformatics*, *2*(4), 603–624. https://doi.org/10.3390/biomedinformatics2040039

Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Morse, S., Mazet, J., Woolhouse, M., Parrish, C., Carroll, D., Karesh, W., Zambrana-Torrelio, C., Lipkin, W., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet*, *380*(9857), 1956–1965. https://doi.org/10.1016/S0140-6736(12)61684-5

Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, *16*, 100258. https://doi.org/10.1016/j.array.2022.10025

Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, *15*(2), 10. https://doi.org/10.1145/3597307

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

O'Reilly-Shah, V., Gentry, K., van Cleve, W., Kendale, S., Jabaley, C., & Long, D. (2020). The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. *Anesthesia & Analgesia*, *131*(2), 340–344. https://doi.org/10.1213/ANE.0000000000004945

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Pedro, R. (2022). Traços gerais da indemnização civil extracontratual pública em contextos de excecionalidade. In *Impactos da pandemia da Covid-19 nas estruturas do direito público* (pp. 379–413). Coimbra: Almedina. (In Portuguese).

Pedro, R. (2023). Inteligência artificial e arbitragem de direito público: Primeiras reflexões. In R. Pedro, & P. Caliendo (Coords.), *Inteligência artificial no contexto do direito público: Portugal e Brasil* (1.ª ed., pp. 105–127). Coimbra: Almedina. (In Portuguese).

Romano, A., Spadaro, G., Balliet, D., Joireman, J., van Lissa, C., Jin, S., ... & Leander, N. P. (2021). Cooperation and trust across societies during the COVID-19 pandemic. *Journal of Cross-Cultural Psychology*, *52*(7), 622–642. https://doi.org/10.1177/00220221209889

Ruan, W., Yi, X., & Huang, X. (2021). Adversarial robustness of deep learning: Theory, algorithms, and applications. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4866–4869). https://doi.org/10.48550/arXiv.2108.10451

Rubin, O., Errett, N., Upshur, R., & Baekkeskov, E. (2021). The challenges facing evidence-based decision making in the initial response to COVID-19. *Scandinavian Journal of Public Health*, *49*(7), 790–796. https://doi.org/10.1177/140349482199722

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Sass, J., Bartschke, A., Lehne, M., Essenwanger, A., Rinaldi, E., Rudolph, S., ... & Thun, S. (2020). The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Medical Informatics and Decision Making*, *20*, 341. https://doi.org/10.1186/s12911-020-01374-w

Shin, D., & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. https://doi.org/10.1016/j.chb.2019.04.019

Shaelou, S. L., & Razmetaeva, Y. (2023). Challenges to fundamental human rights in the age of artificial intelligence systems: Shaping the digital legal order while upholding rule of law principles and European values. *ERA Forum*, *24*(3), 567–587. https://doi.org/10.1007/s12027-023-00777-2

Silva, M., Flood, C., Goldenberg, A., & Singh, D. (2022). Regulating the Safety of Health-Related Artificial Intelligence. *Healthcare Policy*, *17*(4), 63–77. https://doi.org/10.12927/hcpol.2022.26824

Smidt, H., & Jokonya, O. (2021). The challenge of privacy and security when using technology to track people in times of COVID-19 pandemic. *Procedia Computer Science*, *181*, 1018–1026. https://doi.org/10.1016/j.procs.2021.01.281

Sundar, S. (2020). Rise of machine agency: A framework for studying the psychology of human – AI interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Susskind, D. (2021). A world without work: Technology, automation and how we should respond. New Technology, *Work and Employment*, *36*(1), 114-117. https://doi.org/10.1111/ntwe.12186

Syed, R., Ulbricht, M., Piotrowski, K., & Krstic, M. (2023). A Survey on Fault-Tolerant Methodologies for Deep Neural Networks. *Pomiary Automatyka Robotyka*, *27*(2), 89–98. https://doi.org/10.14313/PAR_248/89

Syrowatka, A., Kuznetsova, M., Alsubai, A., Beckman, A., Bain, P., Craig, K., ... & Bates, D. (2021). Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digital Medicine*, *4*(1), 96. https://doi.org/10.1038/s41746-021-00459-8

Theis, T., & Wong, H. (2017). The end of Moore's law: A new beginning for information technology. *Computing in Science & Engineering*, *19*(2), 41–50. https://doi.org/10.1109/MCSE.2017.29

Thomsen, K. (2019). Ethics for artificial intelligence, ethics for all. *Paladyn, Journal of Behavioral Robotics*, *10*(1), 359–363. https://doi.org/10.1515/pjbr-2019-0029

Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Villegas-Ch, W., Jaramillo-Alcázar, A., & Luján-Mora, S. (2024). Evaluating the Robustness of Deep Learning Models against Adversarial Attacks: An Analysis with FGSM, PGD and CW. *Big Data and Cognitive Computing*, *8*(1), 8. https://doi.org/10.3390/bdcc8010008

Vopson, M. (2020). The information catastrophe. *AIP Advances*, *10*(8), 085014. https://doi.org/10.1063/5.0019941

Wallach W., & Allen, C. (2008). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press.

Wang, S., & Shi, W. (2011). Data Mining and Knowledge Discovery. In W. Kresse, D. Danko (Eds.), *Springer Handbook of Geographic Information*. Springer Handbooks. https://doi.org/10.1007/978-3-540-72680-7_5

Wong, F., de la Fuente-Nunez, C., & Collins, J. (2023). Leveraging artificial intelligence in the fight against infectious diseases. *Science*, *381*(6654), 164–170. https://doi.org/10.1126/science.adh1114

Wu, D., Xu, H., Yongyi, W., & Zhu, H. (2022). Quality of government health data in COVID-19: definition and testing of an open government health data quality evaluation framework. *Library Hi Tech*, *40*(2), 516–534. https://doi.org/10.1108/LHT-04-2021-0126

Zhang, Q., Gao, J., Wu, J., Cao, Z., & Dajun, D. (2022). Data science approaches to confronting the COVID-19 pandemic: a narrative review. *Philosophical Transactions of the Royal Society A*, *380*(2214), 20210127. https://doi.org/10.1098/rsta.2021.0127

Zhou, J., Zheng, W., Wang, D., & Coit, D. W. (2024). A resilient network recovery framework against cascading failures with deep graph learning. *Journal of Risk and Reliability*, *238*(1), 193–203. https://doi.org/10.1177/1748006X22112886

Authors information

Pedro Miguel Alves Ribeiro Correia – PhD in Social Sciences (Specialty in Public Administration), Invited Associate Professor, Faculty of Law, University of Coimbra; Visiting Full Professor, ICET/CUA/UFMT, Barra do Garças
Address: Pátio da Universidade, 3004-528 Coimbra, Portugal;
Avenida ValdonVarjão, n. 6390, Barra do Garças – MT, CEP: 78605-091, Brazil
E-mail: pcorreia@fd.uc.pt
ORCID ID: https://orcid.org/0000-0002-3111-9843
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58223408400
WoS Researcher ID: https://www.webofscience.com/wos/author/record/B-2753-2015
Google Scholar ID: https://scholar.google.pt/citations?user=KABKPuUAAAAJ

Ricardo Lopes Dinis Pedro – PhD (Law), Researcher, Lisbon Public Law Research Centre, Faculty of Law, University of Lisbon
Address: Alameda da Universidade, 1649-014 Lisbon, Portugal
E-mail: ricardopedro@fd.ulisboa.pt
ORCID ID: https://orcid.org/0000-0001-6339-5140
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57879177700
WoS Researcher ID: https://www.webofscience.com/wos/author/record/AEN-4511-2022
Google Scholar ID: https://scholar.google.com/citations?hl=en&user=oJ1ImgUAAAAJ

Susana Videira – PhD (Law), Associate Professor, Faculty of Law, University of Lisbon; Scientific and Pedagogical Coordinator, Law Degree and the Master's Degree in Judicial Law, European University
Address: Faculdade de Direito da Universidade de Lisboa, Alameda da Universidade, 1649-014 Lisbon, Portugal; Universidade Europeia, Estrada da Correia, n.º 53, 1500-210, Lisbon, Portugal
E-mail: susanavideira@fd.ulisboa.pt
ORCID ID: https://orcid.org/0000-0002-9246-2557

Authors' contributions

The authors have contributed equally into the concept and methodology elaboration, validation, formal analysis, research, selection of sources, text writing and editing, project guidance and management.

Conflict of interests

The authors declare no conflict of interests.

Financial disclosure

Regarding the participation of the Author Ricardo Pedro, it should be noted that, to the exact extent of his participation, the work is financed (or partially financed) by national funds through FCT–Foundation for Science and Technology, I.P., under the project UIDP/04310/2020. This work was also supported by Portuguese national funds through FCT–Foundation for Science and Technology, I.P., under project UIDB/04643/2020.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – June 15, 2025
Date of approval – June 27, 2025
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Научная статья
УДК 34:004:17:004.8:342.7
EDN: https://elibrary.ru/egkppn
DOI: https://doi.org/10.21202/jdtl.2025.7

Искусственный интеллект в здравоохранении: баланс инноваций, этики и защиты прав человека

Педро Мигель Алвес Рибейро Коррейя ✉
Коимбрский университет, Коимбра, Португалия

Рикардо Лопес Динис Педро
Лиссабонский университет, Лиссабон, Португалия

Сусана Видейра
Лиссабонский университет, Лиссабон, Португалия

Ключевые слова

защита данных,
здравоохранение,
искусственный интеллект,
права человека,
право,
правовое регулирование,
предиктивная аналитика,
фундаментальные права,
этика,
этическое регулирование

Аннотация

Цель: определить ключевые этические, правовые и социальные вызовы, связанные с использованием искусственного интеллекта в здравоохранении, а также разработать рекомендации для создания адаптивных правовых механизмов, способных обеспечить баланс между инновациями, этическим регулированием и защитой фундаментальных прав человека.

Методы: в ходе исследования был реализован многоаспектный методологический подход, интегрирующий классические правовые методы анализа с современными инструментами сравнительного правоведения. Данное исследование охватывает как фундаментальные основы правового регулирования цифровых технологий в медицинской сфере, так и глубокий анализ этических, правовых и социальных импликаций внедрения искусственного интеллекта в систему здравоохранения. Такой комплексный подход позволил обеспечить всестороннее понимание проблематики и сформировать обоснованные выводы относительно перспектив развития данной области.

Результаты: выявлен ряд серьезных проблем, связанных с использованием искусственного интеллекта в здравоохранении. К ним относятся необъективность данных, непрозрачность сложных алгоритмов и риски нарушения неприкосновенности частной жизни. Эти проблемы могут подорвать доверие общества к технологиям искусственного интеллекта и усугубить неравенство в доступе к медицинским услугам. Авторы приходят к выводу, что интеграция искусственного интеллекта в систему здравоохранения должна осуществляться с учетом фундаментальных прав, таких как защита данных и запрет дискриминации, а также соответствовать этическим нормам.

✉ Контактное лицо

Научная новизна: состоит в предложении эффективных механизмов управления для снижения рисков и максимизации потенциала искусственного интеллекта в кризисных ситуациях. Особое внимание уделяется регулятивным мерам, таким как оценка влияния, предусмотренная Законом об искусственном интеллекте. Эти меры играют ключевую роль в выявлении и минимизации рисков, связанных с высокорисковыми системами искусственного интеллекта, обеспечивая соблюдение этических норм и защиту основных прав.

Практическая значимость: заключается в разработке адаптивных правовых механизмов, которые поддерживают демократические нормы и оперативно реагируют на возникающие вызовы в области общественного здравоохранения. Предложенные механизмы позволяют достичь баланса между использованием искусственного интеллекта для управления кризисными ситуациями и сохранением прав человека. Это способствует укреплению доверия к системам искусственного интеллекта и их устойчивому положительному влиянию на общественное здравоохранение.

Для цитирования

Коррейя, П. М. А. Р., Педро, Р. Л. Д., & Видейра, С. (2025). Искусственный интеллект в здравоохранении: баланс инноваций, этики и защиты прав человека. *Journal of Digital Technologies and Law*, *3*(1), 143–180. https://doi.org/10.21202/jdtl.2025.7

Список литературы

Arass, M., & Souissi, N. (2018). Data lifecycle: from big data to SmartData. In *2018 IEEE 5th International Congress on Information Science and Technology* (pp. 80–87). IEEE. https://doi.org/10.1109/CIST.2018.8596547

Alvarez Garcia, V. (1996). *El concepto de necesidad en derecho público* (1st ed.). Madrid: Civitas. (In Spanish).

Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, *46*(6), 161–168. https://doi.org/10.14745/ccdr.v46i06a02

Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, *8*(2), e188-e194. https://doi.org/10.7861/fhj.2021-0095

Beck, U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp Verlag.

Balog-Way, D., & McComas, K. (2022). COVID-19: Reflections on trust, tradeoffs, and preparedness. In *COVID-19* (pp. 6–16). Routledge.

Bazarkina, D. Y., & Pashentsev, E. N. (2020). Malicious use of artificial intelligence. *Russia in Global Affairs*, *18*(4), 154–177. https://doi.org/10.31278/1810-6374-2020-18-4-154-177

Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *International Journal of Environmental Research and Public Health*, *15*(12), 2796. https://doi.org/10.3390/ijerph15122796

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, *48*(3), 386–398. https://doi.org/10.2307/2095230

Bigham, G., Adamtey, S., Onsarigo, L., & Jha, N. (2019). Artificial Intelligence for Construction Safety: Mitigation of the Risk of Fall. In K. Arai, S. Kapoor, R. Bhatia (Eds.). *Intelligent Systems and Applications*. Springer. https://doi.org/10.1007/978-3-030-01057-7_76

Binder, W. (2024). Technology as (dis-)enchantment. AlphaGo and the meaning-making of artificial intelligence. *Cultural Sociology*, *18*(1), 24–47. https://doi.org/10.1177/17499755221138720

Bisconti, P., Orsitto, D., Fedorczyk, F., Brau, F., Capasso, M., De Marinis, L., ... & Schettini, C. (2023). Maximizing team synergy in AI-related interdisciplinary groups: an interdisciplinary-by-design iterative methodology. *AI & Society*, *38*(4), 1443–1452. https://doi.org/10.1007/s00146-022-01518-8

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Box, G. (1979). Robustness in the strategy of scientific model building. In R. Launer & G. Wilkinson (Eds.), *Robustness in Statistics* (pp. 201–236). Academic Press. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231. https://doi.org/10.1214/ss/1009213726

Bulled, N. (2023). "Solidarity:" A failed call to action during the COVID-19 pandemic. *Public Health in Practice*, *5*, 100379. https://doi.org/10.1016/j.puhip.2023.100379

Chen, A. (2016). A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, *125*, 25–38. https://doi.org/10.1016/j.sse.2016.07.006

Chen, J., Zhang, R., Han, W., Jiang, W., Hu, J., Lu, X., Liu, X., & Zhao, P. (2020). Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment. *Journal of Advanced Transportation*, 2020, 6649867. https://doi.org/10.1155/2020/6649867

Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, *15*(2), 126–139. https://doi.org/10.1017/S1744552319000077

Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020a). The combat against COVID-19 in Portugal: How state measures and data availability reinforce some organizational values and contribute to the sustainability of the National Health System. *Sustainability*, *12*(18), 7513. https://doi.org/10.3390/su12187513

Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020b). The combat against COVID-19 in Portugal, Part II: how governance reinforces some organizational values and contributes to the sustainability of crisis management. *Sustainability*, *12*(20), 8715. https://doi.org/10.3390/su12208715

Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2022). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. *European Journal of Applied Business Management*, *8*(1), 1–12.

Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2021). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. In *European Consortium for Political Research General Conference* (pp. 1–18). United Kingdom.

Correia, J. M. C. (1987). *Legalidade e autonomia contratual nos contratos administrativos* (pp. 283, 768). Lisboa: Almedina.

DeCamp, M., & Tilburt, J. (2019). Why we cannot trust artificial intelligence in medicine. *The Lancet Digital health*, *1*(8), e390. https://doi.org/10.1016/S2589-7500(19)30197-9

Dhingra, M., & Gupta, N. (2017). Comparative analysis of fault tolerance models and their challenges in cloud computing. *International Journal of Engineering & Technology*, *6*(2), 36–40. https://doi.org/10.14419/ijet.v6i2.7565

Ettlinger, N. (2022). *Algorithms and the Assault on Critical Thought: Digitalized Dilemmas of Automated Governance and Communitarian Practice* (1st ed.). Routledge. https://doi.org/10.4324/9781003109792

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: Picador, St Martin's Press.

Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. *Nature*, *442*(7101), 448–452. https://doi.org/10.1038/nature04795

Fetzer, T., & Graeber, T. (2021). Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(33), e2100814118. https://doi.org/10.1073/pnas.2100814118

Galetsi, P., Katsaliaki, K., & Kumar, S. (2022). The medical and societal impact of big data analytics and artificial intelligence applications in combating pandemics: A review focused on Covid-19. *Social Science & Medicine*, *301*, 114973. https://doi.org/10.1016/j.socscimed.2022.114973

Gianfrancesco, M., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, *178*(11), 1544–1547. https://doi.org/10.1001/jamainternmed.2018.3763

Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., & Birney, E. (2013). Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, *494*(7435), 77–80. https://doi.org/10.1038/nature11875

Gomes, C. A., & Pedro, R. (Coords.). (2020). *Direito administrativo de necessidade e de excepção*. Lisboa: AAFDL.

Gómez Abeja, L. (2022). Inteligencia artificial y derechos fundamentales. In F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (1.ª ed., pp. 91–114, 93). Murcia: Ediciones Laborum. (In Spanish).

Gómez Colomer, J.-L. (2023). *El juez-robot: La independencia judicial en peligro*. Valencia: Tirant lo Blanch. (In Spanish).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Greiner, R., Grove, A., & Kogan, A. (1997). Knowing what doesn't matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, *97*(1–2), 345–380. https://doi.org/10.1016/S0004-3702(97)00048-9

Gunasekeran, D., Tseng, R., Tham, Y., & Wong, T. (2021). Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digital Medicine*, *4*(1), 40. https://doi.org/10.1038/s41746-021-00412-9

Gürsoy, E., & Kaya, Y. (2023). An overview of deep learning techniques for COVID-19 detection: methods, challenges, and future works. *Multimedia Systems*, *29*(3), 1603–1627. https://doi.org/10.1007/s00530-023-01083-0

Hanegraaff, W. (2013). *Western Esotericism: A Guide for the Perplexed*. Bloomsbury Publishing.

Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2), 8–12. https://doi.org/10.1109/MIS.2009.36

Hazarika, I. (2020). Artificial intelligence: opportunities and implications for the health workforce. *International Health*, *12*(4), 241–245. https://doi.org/10.1093/inthealth/ihaa007

Hoff, K., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hulten, G. (2018). Building Intelligent Systems: A Guide to Machine Learning Engineering. Apress.

Igual, L., & Seguí, S. (2024). *Supervised learning. In Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications* (pp. 67–97). Springer International Publishing.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jones, K., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993. https://doi.org/10.1038/nature06536

Kandlhofer, M., Weixelbraun, P., Menzinger, M., Steinbauer-Wagner, G., & Kemenesi, Á. (2023). Education and Awareness for Artificial Intelligence. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 3–12). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44900-0_1

Kavanagh, J., & Rich, M. (2018). *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. RAND Corporation. https://doi.org/10.7249/RR2314

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, *114*(13), 3521–3526. https://doi.org/10.1073/pnas.1611835114

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. E*uropean Journal of Information Systems*, *31*(3), 388–409. https://doi.org/10.1080/0960085X.2021.1927212

Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1469

Lin, X., Liu, J., Hao, J., Wang, K., Zhang, Y., Li, H., ... & Tan, X. (2020). Collinear holographic data storage technologies. *Opto-Electronic Advances*, *3*(3), 190004. https://doi.org/10.29026/oea.2020.190004

Little, R. J., & Rubin, D. B. (2019). Statistical analysis with missing data. John Wiley & Sons.

Macrae, C. (2022). Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. *Risk Analysis*, *42*(9), 1999–2025. https://doi.org/10.1111/risa.13850

Margetts, H. (2022). Rethinking AI for good governance. *Daedalus*, *151*(2), 360–371. https://doi.org/10.1162/daed_a_01922

Matsuzaka, Y., & Yashiro, R. (2022). Applications of Deep Learning for Drug Discovery Systems with BigData. *BioMedInformatics*, *2*(4), 603–624. https://doi.org/10.3390/biomedinformatics2040039

Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Morse, S., Mazet, J., Woolhouse, M., Parrish, C., Carroll, D., Karesh, W., Zambrana-Torrelio, C., Lipkin, W., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet*, *380*(9857), 1956–1965. https://doi.org/10.1016/S0140-6736(12)61684-5

Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, *16*, 100258. https://doi.org/10.1016/j.array.2022.10025

Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, *15*(2), 10. https://doi.org/10.1145/3597307

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

O'Reilly-Shah, V., Gentry, K., van Cleve, W., Kendale, S., Jabaley, C., & Long, D. (2020). The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. *Anesthesia & Analgesia*, *131*(2), 340–344. https://doi.org/10.1213/ANE.0000000000004945

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Pedro, R. (2022). Traços gerais da indemnização civil extracontratual pública em contextos de excecionalidade. In *Impactos da pandemia da Covid-19 nas estruturas do direito público* (pp. 379–413). Coimbra: Almedina. (In Portuguese).

Pedro, R. (2023). Inteligência artificial e arbitragem de direito público: Primeiras reflexões. In R. Pedro, & P. Caliendo (Coords.), *Inteligência artificial no contexto do direito público: Portugal e Brasil* (1.ª ed., pp. 105–127). Coimbra: Almedina. (In Portuguese).

Romano, A., Spadaro, G., Balliet, D., Joireman, J., van Lissa, C., Jin, S., ... & Leander, N. P. (2021). Cooperation and trust across societies during the COVID-19 pandemic. *Journal of Cross-Cultural Psychology*, *52*(7), 622–642. https://doi.org/10.1177/00220221209889

Ruan, W., Yi, X., & Huang, X. (2021). Adversarial robustness of deep learning: Theory, algorithms, and applications. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4866–4869). https://doi.org/10.48550/arXiv.2108.10451

Rubin, O., Errett, N., Upshur, R., & Baekkeskov, E. (2021). The challenges facing evidence-based decision making in the initial response to COVID-19. *Scandinavian Journal of Public Health*, *49*(7), 790–796. https://doi.org/10.1177/140349482199722

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Sass, J., Bartschke, A., Lehne, M., Essenwanger, A., Rinaldi, E., Rudolph, S., ... & Thun, S. (2020). The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Medical Informatics and Decision Making*, *20*, 341. https://doi.org/10.1186/s12911-020-01374-w

Shin, D., & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. https://doi.org/10.1016/j.chb.2019.04.019

Shaelou, S. L., & Razmetaeva, Y. (2023). Challenges to fundamental human rights in the age of artificial intelligence systems: Shaping the digital legal order while upholding rule of law principles and European values. *ERA Forum*, *24*(3), 567–587. https://doi.org/10.1007/s12027-023-00777-2

Silva, M., Flood, C., Goldenberg, A., & Singh, D. (2022). Regulating the Safety of Health-Related Artificial Intelligence. *Healthcare Policy*, *17*(4), 63–77. https://doi.org/10.12927/hcpol.2022.26824

Smidt, H., & Jokonya, O. (2021). The challenge of privacy and security when using technology to track people in times of COVID-19 pandemic. *Procedia Computer Science*, *181*, 1018–1026. https://doi.org/10.1016/j.procs.2021.01.281

Sundar, S. (2020). Rise of machine agency: A framework for studying the psychology of human – AI interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Susskind, D. (2021). A world without work: Technology, automation and how we should respond. New Technology, *Work and Employment*, *36*(1), 114-117. https://doi.org/10.1111/ntwe.12186

Syed, R., Ulbricht, M., Piotrowski, K., & Krstic, M. (2023). A Survey on Fault-Tolerant Methodologies for Deep Neural Networks. *Pomiary Automatyka Robotyka*, *27*(2), 89–98. https://doi.org/10.14313/PAR_248/89

Syrowatka, A., Kuznetsova, M., Alsubai, A., Beckman, A., Bain, P., Craig, K., ... & Bates, D. (2021). Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digital Medicine*, *4*(1), 96. https://doi.org/10.1038/s41746-021-00459-8

Theis, T., & Wong, H. (2017). The end of Moore's law: A new beginning for information technology. *Computing in Science & Engineering*, *19*(2), 41–50. https://doi.org/10.1109/MCSE.2017.29

Thomsen, K. (2019). Ethics for artificial intelligence, ethics for all. *Paladyn, Journal of Behavioral Robotics*, *10*(1), 359–363. https://doi.org/10.1515/pjbr-2019-0029

Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Villegas-Ch, W., Jaramillo-Alcázar, A., & Luján-Mora, S. (2024). Evaluating the Robustness of Deep Learning Models against Adversarial Attacks: An Analysis with FGSM, PGD and CW. *Big Data and Cognitive Computing*, *8*(1), 8. https://doi.org/10.3390/bdcc8010008

Vopson, M. (2020). The information catastrophe. *AIP Advances*, *10*(8), 085014. https://doi.org/10.1063/5.0019941

Wallach W., & Allen, C. (2008). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press.

Wang, S., & Shi, W. (2011). Data Mining and Knowledge Discovery. In W. Kresse, D. Danko (Eds.), *Springer Handbook of Geographic Information*. Springer Handbooks. https://doi.org/10.1007/978-3-540-72680-7_5

Wong, F., de la Fuente-Nunez, C., & Collins, J. (2023). Leveraging artificial intelligence in the fight against infectious diseases. *Science*, *381*(6654), 164–170. https://doi.org/10.1126/science.adh1114

Wu, D., Xu, H., Yongyi, W., & Zhu, H. (2022). Quality of government health data in COVID-19: definition and testing of an open government health data quality evaluation framework. *Library Hi Tech*, *40*(2), 516–534. https://doi.org/10.1108/LHT-04-2021-0126

Zhang, Q., Gao, J., Wu, J., Cao, Z., & Dajun, D. (2022). Data science approaches to confronting the COVID-19 pandemic: a narrative review. *Philosophical Transactions of the Royal Society A*, *380*(2214), 20210127. https://doi.org/10.1098/rsta.2021.0127

Zhou, J., Zheng, W., Wang, D., & Coit, D. W. (2024). A resilient network recovery framework against cascading failures with deep graph learning. *Journal of Risk and Reliability*, *238*(1), 193–203. https://doi.org/10.1177/1748006X22112886

Сведения об авторах

Коррейя Педро Мигель Алвес Рибейро – PhD в области общественных наук (государственное управление), приглашенный доцент, юридический факультет, Коимбрский университет; приглашенный профессор, ICET/CUA/UFMT, Барра до Гарсас
Адрес: Португалия, 3004-528, г. Коимбра, Патио да Универсидаде; Бразилия, 78605-091, Авенида Валдон Варжан, 6390, Барра до Гарсас – MT, CEP
E-mail: pcorreia@fd.uc.pt
ORCID ID: https://orcid.org/0000-0002-3111-9843
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58223408400
WoS Researcher ID: https://www.webofscience.com/wos/author/record/B-2753-2015
Google Scholar ID: https://scholar.google.pt/citations?user=KABKPuUAAAAJ

Рикардо Лопес Динис Педро – PhD в области права, научный сотрудник, Лиссабонский исследовательский центр в области публичного права, юридический факультет, Лиссабонский университет
Адрес: Португалия, 1649-014, г. Лиссабон, Аламеда де Универсидаде
E-mail: ricardopedro@fd.ulisboa.pt
ORCID ID: https://orcid.org/0000-0001-6339-5140
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57879177700
WoS Researcher ID: https://www.webofscience.com/wos/author/record/AEN-4511-2022
Google Scholar ID: https://scholar.google.com/citations?hl=en&user=oJ1ImgUAAAAJ

Сусана Видейра – PhD в области права, доцент, юридический факультет, Лиссабонский университет; координатор по науке и образованию, Европейский университет в Лиссабоне
Адрес: Португалия, 1649-014, г. Лиссабон, Аламеда де Универсидаде; Португалия, 1500-210, г. Лиссабон, Эстрада да Коррейя, 53
E-mail: susanavideira@fd.ulisboa.pt
ORCID ID: https://orcid.org/0000-0002-9246-2557

Вклад авторов

Авторы внесли равный вклад в разработку концепции, методологии, валидацию, формальный анализ, проведение исследования, подбор источников, написание и редактирование текста, руководство и управление проектом.

Конфликт интересов

Авторы сообщают об отсутствии конфликта интересов.

Финансирование

Участие автора Рикардо Лопес Динис Педро частично финансировалось Фондом науки и технологий Португалии (Foundation for Science and Technology, FCT) в рамках проекта UIDP/04310/2020. Исследование также было поддержано тем же Фондом в рамках проекта UIDB/04643/2020.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.15.59 / Права и свободы человека и гражданина
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 15 июня 2024 г.
Дата одобрения после рецензирования – 27 июня 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.