Dyskurs Prawniczy i Administracyjny

2/2026

Dyskurs Prawniczy i Administracyjny
2/2026

REDAKCJA NAUKOWA TOMU
Martyna Łaszewska-Hellriegel i Ewa Szewczyk

Zielona Góra 2026

UNIWERSYTET ZIELONOGÓRSKI

INSTYTUT NAUK PRAWNYCH
UNIWERSYTET ZIELONOGÓRSKI

RECENZJE NUMERU

http://www.dyskurs.inp.uz.zgora.pl

REDAKCJA

Katarzyna Hadasik
Aldona Reich

REDAKCJA TECHNICZNA

Arkadiusz Sroka

PROJEKT OKŁADKI

Jacek Papla

ISSN 2657-926X

Oficyna Wydawnicza Uniwersytetu Zielonogórskiego
65-246 Zielona Góra, ul. Podgórna 50, tel. (68) 328 78 64
www.ow.uz.zgora.pl, sekretariat@ow.uz.zgora.pl

Spis treści

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Katarzyna Chałubińska-Jentkiewicz
Akademia Leona Koźmińskiego
ORCID: 0000-0003-0188-5704
kasiachalubinska@gmail.com

The Importance of Artificial Intelligence Tools in Combating Illegal and Harmful Content

Keywords: cybersecurity, illegal content, artificial intelligence, cyber threats, digital services

Summary. The convergence of various areas of our lives has led to a particular type of conflict regarding the scope and level of new regulations, especially in relation to digital content, where most issues concern new media and new technologies, including the liability of digital service providers. New models of content management (including online) are emerging, supported by new principles of the virtual organisation of the digital market. Changes in communication technology have fundamentally altered the way individuals and entire communities function. Online multimedia platforms providing electronic services are emerging, which require the use of modern technological solutions and in which the private sector most often invests. On the one hand, an open and free cyberspace allows for the exchange of cultures and experiences between nations, communities and citizens, facilitating interaction and the exchange of content, and consequently of knowledge, experiences and technology. On the other hand, this openness triggers activities that encroach upon the realm of rights and freedoms. This article addresses these elements of a new phase in the development of breakthrough technologies such as artificial intelligence, which require an analysis of the impact of new technological tools on the digital environment and the legal situation of the individual. This is particularly evident in procedures related to combating illegal and harmful content.

Znaczenie narzędzi sztucznej inteligencji w zwalczaniu nielegalnych i szkodliwych treści

Słowa kluczowe: cyberbezpieczeństo, treści nielegalne, sztuczna inteligencja, cyberzagrożenia, usługi cyfrowe

Streszczenie. Procesy konwergencji różnych obszarów naszego funkcjonowania przyczyniły się do powstania szczególnego rodzaju konfliktu w obszarze ustaleń co do zakresu i poziomu nowych regulacji, zwłaszcza w odniesieniu do treści cyfrowych, gdzie najwięcej zagadnień dotyka kwestii nowych mediów i nowych technologii, również w sferze odpowiedzialności dostawców usług cyfrowych. Pojawiają się nowe modele zarządzania treściami (także online), wspomagane przez nowe zasady wirtualnej organizacji rynku cyfrowego. Zmiana technologii komunikowania zasadniczo zmieniła zasady funkcjonowania jednostek i całych społeczności. Powstają internetowe platformy multimedialne świadczące usługi drogą elektroniczną, które wymagają zastosowania nowoczesnych rozwiązań technologicznych, a w które najczęściej inwestuje sektor prywatny. Z jednej strony otwarta i wolna cyberprzestrzeń pozwala na wymianę kultur i doświadczeń między państwami, społecznościami i obywatelami, umożliwiając interakcję oraz wymianę treści, a w konsekwencji wiedzy, doświadczeń i technologii. Z drugiej strony otwartość ta wyzwala aktywność wkraczającą w obszar praw i wolności.

Artykuł odnosi się do tych elementów nowego etapu rozwoju technologii przełomowych takich jak Sztuczna Inteligencja, które wymagają analizy wpływu zastosowania instrumentów nowych technologii na środowisko cyfrowe i prawną sytuację jednostki. Jest to szczególnie widoczne w procedurach związanych ze zwalczaniem treści nielegalnych oraz treści szkodliwych.

Introduction

One of the key objectives in law enforcement and combating cyber threats has become the regulation of liability for online activities. New ways of providing online services have given rise to entirely new types of service providers: network providers, hosting providers, online platform operators, including providers of online content sharing services and video-sharing platform operators, and digital content providers, who, under the new regulations, become liable for online activities, including in the context of combating illegal or harmful content. In the new regulatory ecosystem for the provision of e-commerce services, the issue of establishing operational measures and appropriate procedures has also arisen. A fundamental challenge in enforcing these measures is the verification of internet users' data, which enables the digital identification of infringers. Apparent anonymity online remains a fundamental challenge for law enforcement in the digital environment. These limitations are effectively overcome by technological solutions, such as those available to NASK PIB as a research organisation, but also as a technological and operational institution. The unique Polish model of cooperation between CERTU and the DNS registry enables a rapid response and a fundamental shift in the digital world when it comes to combating cybercrime.

The new technological order serves as both the premise and the subject of the changes under discussion, which have a fundamental impact on the crucial area of cyberspace and the entities operating within it, whose very existence is underpinned by freedom of speech and the right to privacy. It must be emphasised here, however, that the issue of restrictions on freedom of speech, in a democratic context, involves decisions to interfere with the content produced by religious organisations, owners of 'free' media, political groups and various other social groups that adopt a strictly defined way of thinking, which digital media facilitate, trapping them in an ideological bubble, not merely an information bubble, because this problem concerns not only the reception of specific information by internet users but also the sharing of digital content. This problem concerns not only the delivery of specific content to internet users but also the sharing of digital content by users. However, we should draw here on the words of the philosopher Onora O'Neill, who states

that "our media should be not only accessible, but also assessable"[1]. Although, as T. Garton Ash emphasises: "Truth be told, the most effective constraint on the media should be ourselves – readers, viewers and users"[2]. New technologies, including artificial intelligence (AI), enable the automatic, mathematical analysis of digital information such as text, sound, images or various other data. At the same time, automated data processing methods are becoming crucial for insights into issues related to in-depth analysis not only of the digital environment itself but also of the individual. Microsoft is actively researching how to automatically detect 'listeners' affective reactions during online presentations". Microsoft Teams analyses the facial expressions and gestures of participants in online events. Emotion recognition using ML systems is widely used by the police, nuclear power plant operators, airport security and psychiatrists in Russia, China, Japan and South Korea. There is a known police initiative in India to use systems to automatically 'detect anxiety on women's faces in order to alert officers', which is intended to help combat sexual harassment and violence in public spaces. And here, too, issues arise regarding decisions made on the basis of AI[3].

---

[1] O.O'Neill, as cited in: T. Garton Ash, *Free Speech: Ten Principles for a Connected World*, Kraków 2018, p. 301.

[2] T. Garton Ash, *op. cit.*, p. 379.

[3] See J. Spivak, C. Garvie, *A Taxonomy of Legislative Approaches to Face Recognition in the United States*, [in:] *Regulating Biometrics: Global Approaches and Urgent Questions*, ed. A. Kak, AI Now Institute, 1 September 2020, pp. 86-95, https://ainowinstitute.org/regulatingbiometrics.html [accessed on: 10.04.2026]; J. Cobbe, M.S. Ah Lee, J. Singh, *Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems. ACM Conference on Fairness, Accountability, and Transparency (FAccT' 21), 1-10 March 2021*, Virtual Event, Canada, ACM, New York, NY, USA, p. 4. See also D. Bergamini, *Need for democratic governance of artificial intelligence. Report for the Parliamentary Assembly of the Council of Europe*. Doc. 15150. 24 September 2020; I. Berle, *Face Recognition Technology: Compulsory Visibility and its Impact on Privacy and the Confidentiality of Personal Identifiable Images*, Springer, 2020; J. Buolamwini, T. Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, "Proceedings of Machine Learning Research" 2018, 81, pp. 1-15; C. Burt, *Concerns about biometric online proctoring expressed by students in Australia, the US and Canada*, Biometric Update, 3 July 2020, https://www.biometricupdate.com/202007/concerns-about-biometric-online-proctoring-expressed-by-students-in-australia-u-s-and-canada [accessed on: 10.04.2026]; L.A. Bygrave, L. Tosoni, *Article 4(14). Biometric Data*, [in:] *The EU General Data Protection Regulation (GDPR): A Commentary*, eds. C. Kuner, L.A. Bygrave, C. Docksey, Oxford University Press, Oxford 2020, pp. 207-216; CAHAI. Feasibility study. CAHAI(2020)23. Council of Europe, Strasbourg. 17 December 2020; D. Clifford, M. Richardson, N. Witzleb, *Artificial Intelligence and Sensitive Inferences: New Challenges for Data Protection Laws*, [in:] *Regulatory Insights on Artificial Intelligence: Research for Policy*, eds. M Findlay et al., Edward Elgar, Cheltenham, UK, 2021, pp. 19-45; C. Kind, *Containing the canary in the AI coalmine – the EU's efforts to regulate biometrics*, Ada Lovelace Institute Blog, 30 April 2021, https://www.adalovelaceinstitute.org/blog/canary-ai-coalmine-eu-regulate-biometrics/ [accessed on: 10.04.2026]; E. Kindt, *A First Attempt at Regulating Biometric Data in the European Union*, [in:] *Regulating Biometrics…*, pp. 62-69.

New Obligations for Service Providers and the Use of rtificial ntelligence

Examples of the use of so-called artificial intelligence (AI), deep learning (DL) and machine learning (ML) programmes are numerous and, in many respects, are becoming increasingly worrying. An example of such activity is the processing of personal data, *i.e.,* information that can be used to distinguish or track the identity of a natural person, such as name and surname, social security numbers, tax identification numbers and biometric data, linked to other personal or identifying data, such as date and place of birth and mother's maiden name. Personal data collected from consumers and end-users of digital services has become a valuable resource for technology providers focusing on artificial intelligence. The processing of various types of data reveals external behaviours and trends, as well as the internal preferences and biases of an individual, a community or a specific demographic group throughout all decision-making processes. Without proper data management, personal data can be exploited by malicious actors for unlawful impersonation, micro-targeted advertising and the potential manipulation of individual preferences. Sensitive personal data collected from citizens of one country may be shared or sold to businesses in another country. This grants foreign entities access to valuable political, economic, social and geographical data, opening the door for foreign entities to identify, exploit and expose another country's vulnerabilities. As the Internet of Things (IoT) market develops, large-scale machine-to-machine communication provides deep operational insights into the status of a machine, a production line, and sometimes the entire workflow. This IoT communication data can be used for monitoring, preventative maintenance and the overall improvement of industrial and manufacturing process efficiency. However, if mismanaged, sensitive or confidential operational data may be exposed, potentially resulting in the loss of valuable patents, trade secrets, designs, and even contracts. The current binary concept of personal data is already being challenged by mainstream technologies; however, artificial intelligence is blurring the distinction to such an extent that defining what is and what is not 'personal data' is becoming significantly more difficult. The emergence of AI is likely to lead to an environment in which all information generated by or related to an individual can be identified. Furthermore, if this data falls into the hands of malicious actors, it can be manipulated to produce false results or representations, negatively disrupting general business and industrial/manufacturing processes and exposing the entire company to serious risk. However, today we are focusing more on issues related to the use of AI for the benefit of society.

The new 'notice and take action' system proposed in the DSA[4] requires platforms to take a proactive approach to identifying illegal and harmful content, as well as to automatically detect such content. Blocking orders under the DSA help to reduce the scale of illegal online activity by capitalising on the fact that digital service providers – intermediaries – have the practical means to prevent further infringements of the law through the use of automated data processing systems. In accordance with Recital 50 of the DSA (50) Hosting service providers play a particularly important role in combating illegal content on the internet, as they store information provided by service users and at their request, and typically make that information available to other users, sometimes on a large scale. It is important that all hosting service providers, regardless of size, put in place easily accessible and user-friendly reporting and action mechanisms that facilitate the reporting to the hosting service provider of specific information deemed by the reporter to be illegal content ('a report'), on the basis of which the provider can decide whether it agrees with that assessment and whether it wishes to remove the illegal content or disable access to it ("action").

Such mechanisms should be easily identifiable, located in close proximity to the information in question, and at least as easy to find and use as mechanisms for reporting content that infringes the hosting service provider's terms of service. Subject to compliance with the requirements regarding notifications, individuals or entities should be able to report multiple items of potentially illegal content in a single notification in order to ensure the effective functioning of the notification and action mechanisms. Importantly, the notification mechanism should allow for the identification of the person or entity making the notification, but should not require it. For certain types of reported information, the identity of the person or entity making the report may be necessary to determine whether the information in question constitutes potentially illegal content. The obligation to put in place reporting and action mechanisms should apply, for example, to file storage and sharing services, website hosting services, ad servers and pastebin-type services, provided they qualify as hosting services. Such a system should provide for appropriate time limits that enable victims to exercise their right to a fair trial and other fundamental rights, although it must be noted that it is particularly difficult to establish time limits that are appropriate for the interactive online environment.

For example, the French legislator has entrusted the assessment of compliance with obligations and the factors determining them, such as the availability of blocking measures, to an independent public body. The implementation of these measures does not constitute a general monitoring obligation. However, moder-

[4] Regulation 2022/2065 on the Digital Services Act and amending Directive 2000/31/EC (Digital Services Act) OJ L 2022, 277.1.

ation is possible. The DSA, in Recital 96, defines the scope of supervision over content moderation. According to Recital (96), in order to adequately monitor and assess compliance by very large online platforms and very large search engines with the obligations under the DSA, the Digital Services Coordinator competent for the place of establishment or the Commission may request access to specific data, including data relating to algorithms. Such a requirement may include, for example, data necessary to assess the risks and potential harm caused by the systems of a very large online platform or a very large search engine, data concerning the accuracy, functioning and testing of algorithmic systems used for content moderation, recommendation systems or advertising systems, including, where applicable, training data and algorithms, or data concerning the processes and results of content moderation or internal complaint-handling systems. However, the imposed moderation of uploaded content is based on automatic recognition tools. Furthermore, the established content filtering mechanism explicitly states that the fair use rule shall be observed in cases where it is applied in the context of combating infringements of intellectual property rights. It must be borne in mind that, in this area too, the protection of fundamental rights and freedoms sets the limits of any new regulation. At the same time, it should be noted that restrictions on these rights and freedoms may be dictated by considerations of the protection of the public interest, which is defined, inter alia, by the rights and freedoms of others. Therefore, such access to data should not include requests for specific information on individual service users for the purpose of determining whether they comply with other applicable provisions of Union or national law. Article 8 of the DSA excludes the possibility of imposing a general obligation on digital service providers to monitor data within their network. This also applies to the active identification of facts or circumstances indicating illegal activity. According to the CJEU, this prohibition applies in particular to national provisions that would impose on a provider an obligation to actively monitor all data of every service user in order to prevent future infringements of intellectual property rights[5] . At the time, the view was taken that this could discourage entities from providing services electronically and, as a result, contribute to hampering the development of telecommunications networks, which form the basis for the functioning of the information society. Furthermore, a contentious issue was whether the adoption of the possibility of imposing the proposed obligation (*e.g.,* in the form of a protective order) would

[5] Judgment of the Court (Third Chamber) of 24 November 2011: Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM). Reference for a preliminary ruling: Cour d'appel de Bruxelles – Belgium .Information society – Copyright – Internet – Peer-to-peer software – Internet service providers – Introduction of a system for filtering electronic communications to prevent file-sharing that infringes copyright – No general obligation to monitor the information transmitted. Case C-70/10. European Court Reports 2011 I-11959.

amount to permitting preventive censorship on the internet, restricting freedom of speech, the freedom to express one's views, and freedom of expression, as well as infringing the right to privacy. Meanwhile, an order to establish a filtering system was issued, which, however, amounts to surveillance, and such surveillance is, moreover, unlimited in time, covering all future infringements. Interestingly, the CJEU held that the order to establish the disputed filtering system should be regarded as failing to meet the requirement to ensure an appropriate balance between, on the one hand, the protection of intellectual property rights enjoyed by copyright holders and, on the other hand, the protection of the freedom to conduct business enjoyed by economic operators – digital service providers.

Automated Processing Tools and the Fight Against Illegal and Harmful cCntent

As already highlighted above, the order in question has been deemed questionable in the context of freedom of expression, since, according to opponents of such a procedure, the system might not sufficiently distinguish between unlawful and lawful content, with the result that its implementation could lead to the blocking of lawful content. It should be noted, however, that freedom of expression is subject to restrictions arising from public interest objectives. There is no doubt that the question of illegality also depends on the application of statutory exceptions, which vary from one Member State to another. Furthermore, it was emphasised that in some Member States, for example, certain works may be in the public domain or published free of charge on a website by their authors. According to the Court of Appeal in Wrocław of 15 January 2010 I ACa 1202/09, one cannot agree with the assessment that the standard of conduct, professionalism, requires the administrator to continuously filter and remove statements that infringe the law or may infringe the law in an objective assessment, without prior notification of this fact. The analyses conducted here by researchers into the evolution and significance of systemic risk in the digital environment are of particular importance for reducing information asymmetry and creating a resilient risk mitigation system, as they provide insights into the risks to internet platform providers, internet search engine providers, digital services coordinators, other competent authorities, the Commission and society as a whole.

It is clear that technical measures and new technologies are intended to support the analysis of the online environment. In accordance with Article 17 of the Directive on Copyright and Related Rights in the Digital Single Market, if no authorisation has been granted for the use of a work, providers of online content-sharing services shall be liable for unauthorised acts of making available to the public,

including communication to the public, of copyright-protected works and other protected subject-matter, unless they demonstrate, inter alia, they acted promptly upon receiving a duly substantiated notice from rightholders to block access to the works or other protected subject-matter covered by the notice, or to remove them from their websites, and made every effort to prevent the future uploading of such content. Consequently, Article 22(7) of the Polish Act on Copyright and Related Rights also introduced the absence of an obligation for providers of online content-sharing services to monitor the lawfulness of the public making available of a work posted by a service user[6].

It should be emphasised that the legal ecosystem governing the operation of intermediaries is currently changing, with new obligations for intermediary service providers including, amongst other things, so-called content moderation and even filtering. The new tasks set out in the DSA fall within the scope of so-called 'content moderation', which means actions, whether automated or not, taken by intermediary service providers, the purpose of which is, in particular, to detect, identify and combat illegal content or information that is inconsistent with the terms of use of their services, transmitted by service users, including measures implemented that affect the availability, visibility and accessibility of such illegal content or information, such as the ranking of such content or information, demonetisation, blocking access to it or its removal, or which affect the ability of service users to transmit such information, such as the closure or suspension of a user's account (Article 3(t) of the DSA). In accordance with Recital (30):

> Intermediary service providers should not be subject to a monitoring obligation – either by law or in practice – in respect of general obligations. This does not apply to monitoring obligations in specific cases and, in particular, does not affect orders issued by national authorities in accordance with national legislation compatible with Union law, as interpreted by the Court of Justice of the European Union, and in accordance with the conditions laid down in the DSA. Nothing in this Regulation shall be interpreted as imposing a general monitoring obligation or a general obligation to actively seek facts, nor as imposing on providers a general obligation to take active measures in relation to illegal content

This issue determined the application of solutions concerning automated data processing for public interest purposes. At the same time, recital 46 explains that intermediaries should clearly state and update, in their terms of service, the reasons for which they may restrict the provision of their services. These terms and conditions should, in particular, include information on any policies, procedures, measures and tools used for content moderation purposes, including algorithmic decision-making and human verification, as well as on the internal rules of the complaints-handling system. They should also include easily accessible information

[6] Act of 4 February 1994 on copyright and related rights (*i.e.,* "Journal of Laws" of 2025, item 24).

on the right to terminate use of the service. This shift in approach is crucial for the implementation of EU measures aimed at creating new conditions for combating illegal content.

Meanwhile, in the case of the regulation of audiovisual media services set out in Directive 2010/13/EU, the use of filtering is permitted for the purpose of protecting the physical, mental and moral development of minors and the protection of human dignity, subject to the compatibility of such systems with the fundamental right to freedom of expression enshrined in the Charter of Fundamental Rights of the European Union. In accordance with Recital 60 of the Directive:

> The aim of such measures, such as the use of personal identification numbers (PINs), filtering or labelling systems, should therefore be to ensure an adequate level of protection of the physical, mental and moral development of minors and the protection of human dignity, particularly in the case of on-demand audiovisual media services. The importance of content filtering and labelling systems was already highlighted in the Recommendation on the protection of minors, human dignity and the right of reply, which sets out a number of possible measures for the protection of minors, such as the systematic provision to users who subscribe to a service from a provider of an effective, updatable and user-friendly filtering system, or the equipping of services intended for children with automatic filtering systems.

Member States shall ensure that all video-sharing platform providers within their jurisdiction apply such measures. These measures must be feasible and proportionate, taking into account the scale of the video-sharing platform service and the nature of the service provided. These measures shall not lead to ex ante control or to the filtering of content as it is uploaded to the platform. For the purposes of the protection of minors referred to in point (a) of paragraph 1 of this Article, the most harmful content shall be subject to the strictest access control measures.

These measures shall include, as appropriate:

a) providing users uploading videos to the platform with a function allowing them to declare whether the video contains audiovisual commercial communications, insofar as they have such knowledge or can reasonably be expected to have such knowledge;
b) the establishment and operation of transparent and user-friendly mechanisms on the video-sharing platform, enabling users to report or flag content to the provider of such a platform;
c) the establishment and operation of systems through which video-sharing platform providers inform platform users of the action taken in response to a report or notification;
d) the creation and operation of systems for verifying the age of users of video-sharing platforms in relation to content that may be harmful to the physical, mental or moral development of minors;

e) the establishment and operation of user-friendly systems enabling users of video-sharing platforms to rate content;
f) providing parental control systems, subject to end-user control, in relation to content that may be harmful to the physical, mental or moral development of minors.

Personal data of minors collected or otherwise generated by video-sharing platform providers pursuant to points (f) and (h) of the third subparagraph shall not be processed for commercial purposes, such as direct marketing, profiling or behavioural advertising. Member States shall establish the necessary mechanisms to assess the adequacy of the measures implemented by video-sharing platform providers. Member States shall entrust the assessment of these measures to national authorities or regulatory bodies. The Polish legal system has introduced obligations to apply technical measures to safeguard content harmful to children. The scope and mechanisms for restricting access to content hosted on all types of websites can take several forms. The most important of these are mechanisms based on content filtering at the provider's end (ISP), known as server-based, and mechanisms based on blocking the display of content by client software, known as client-based. In the case of online content-sharing services, if rightholders request that access to their individual works or other protected subject-matter be blocked or that such content be removed, they must duly justify their request. Complaints submitted under this mechanism shall be dealt with without undue delay, and decisions to block access to or remove posted content shall be subject to human review, which forms the basis for the exercise of the right to human intervention in situations where the service provider uses such automated data processing measures.

In the case of harmful content, and in particular deepfakes, the provisions of Article 18(e) of Regulation (EU) 2024/1083 of the European Parliament and of the Council of 11 April 2024 on the establishment of a common framework for media services in the internal market and amending Directive 2010/13/EU (European Media Freedom Act) Text with EEA relevance (OJ EU L 2024, item 1083), pursuant to which providers of very large online platforms shall introduce a function enabling users of their services to declare that they do not provide content generated by artificial intelligence systems without subjecting such content to human verification or editorial control, which means the possibility of creating content – albeit with human input. Similarly, the requirements regarding the transparency of AI systems also constitute an attempt at regulatory restrictions on the use of AI. In accordance with Article 50(2) of the AI Act[7], providers of AI systems, including

[7] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on the establishment of harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU)

general-purpose AI systems, that generate content in the form of synthetic sounds, images, video or text, shall ensure that the outputs of the AI system are labelled in a machine-readable format and are detectable as artificially generated or manipulated. Providers shall ensure the effectiveness, interoperability, robustness and reliability of their technical solutions to the extent that this is technically feasible, taking into account the specific characteristics and limitations of different types of content, implementation costs and the generally recognised state of the art, which may be reflected in relevant technical standards. This obligation does not apply to the extent that AI systems perform an auxiliary function in standard editing or do not substantially alter the input data or its semantics provided by the user, nor to the extent that this is permitted by law for the purposes of detecting, preventing them, conducting investigations into them or prosecuting their perpetrators. In addition, entities using an AI system that generates text published to inform the public about matters of public interest, or manipulates such text, shall disclose that the text has been artificially generated or manipulated. This obligation shall not apply where such use is permitted by law for the purposes of detecting, preventing, investigating or prosecuting criminal offences, or where the AI-generated content has been subject to human verification or editorial control and where a natural or legal person bears editorial responsibility for the publication of the content. This approach follows transparency rules, whilst allowing for the application of solutions without the specified restrictions in situations dictated by security considerations.

But are these measures, which can only support activities in the public interest aimed at detecting cybercrime, safe? Current regulations certainly indicate that AI tools are to be used primarily for the common good. However, models intended to serve the public good are perfectly capable of actions that go beyond this primary objective. This is the case with the recently described Claude Mythos Preview model, a new general-purpose language model[8]. This model operates in every field, but is also exceptionally effective in tasks related to computer security concerning operating systems. In this spirit, Project Glasswing was launched, an initiative aimed at using Mythos Preview to secure the world's most critical software and prepare the industry for the practices we will all need to implement to stay ahead of cybercriminals. Mythos Preview is capable of identifying and then exploiting

2019/2144, and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) Text with EEA relevance (OJ EU L 2024, p. 1689).

[8] Assessment of cybersecurity capabilities Claude Mythos Preview Nicholas Carlini, Newton Cheng, Keane Lucas, Michael Moore, Milad Nasr, Vinay Prabhushankar, Winnie Xiao, Hakeem Angulu, Evyatar Ben Asher, Jackie Bow, Keir Bradwell, Ben Buchanan, David Forsythe, Daniel Freeman, Alex Gaynor, Xinyang Ge, Logan Graham, Kyla Guru, Hasnain Lakhani, Matt McNiece, Mojtaba Mehrara, Renee Nichol, Adnan Pirzada, Sophia Porter, Andreas Terzis, Kevin Troy; 7 April 2026 https://red.anthropic.com/2026/mythos-preview/ [accessed on: 10.04.2026].

zero-day vulnerabilities in every major operating system and every major web browser, whenever the user requests it. The vulnerabilities detected are often subtle or difficult to detect. Many of them are ten or twenty years old, with the oldest being a 27-year-old flaw in OpenBSD – an operating system known primarily within the cybersecurity community. This model has been able to identify a number of vulnerabilities in the world's most popular cryptographic libraries, algorithms and protocols. Developers are already highlighting the need to develop cybersecurity (and other) safeguards that detect and block the model's most dangerous outputs. The future will reveal the scope of application for such models. Like any other, they will likely have dual applications.

Summary

Technological changes in the functioning of digital services require a fresh perspective on issues related to safeguarding public interest objectives. In the course of assessing what new directions have emerged in the pursuit of the public interest in the digital environment, it may turn out that entirely new regulation will be required here concerning the risks of using AI tools – including those subject to restrictions – as in the case of the transparency rule described above. This is because entirely new tasks will arise – including for the legislator – which will, however, require consideration of factors such as public morality, public safety, public order, and the rights and freedoms of others.

The current race to achieve 'Artificial General Intelligence' (AGI) is, in reality, a race for power. The winner of this race could gain absolute power over everyone (people, governments, organisations) who base their management of processes, finances or security on its foundation model. This dependence will be far deeper than dependence on Google's search engine. Search engines or internet platforms may manipulate us to a greater or lesser extent by suggesting certain results rather than others. But if we base any AI system on a model controlled by one of the mega corporations, we must accept that, to a small – though sometimes crucial – extent, we lose control over it. This also applies to state institutions. It is partly true to say that, after all, no one has full control over such a model because that is the nature of artificial neural networks. We will not even know what data the model was trained on (as most of this data is obtained in breach of the law), nor what security measures were applied during the training process without clear rules defining risks and procedures. In the same way that we add cognitive rules to the model, various hidden functions can be added, the existence of which will be known only to someone directly controlling the model in question. And this need not be a function supervised by its owner, the state or its institutions.

References

Bergamini D., *Need for democratic governance of artificial intelligence. Report for the Parliamentary Assembly of the Council of Europe*, doc. 15150, 24 September 2020.

Berle I., *Face Recognition Technology: Compulsory Visibility and its Impact on Privacy and the Confidentiality of Personal Identifiable Images*, Springer, 2020.

Buolamwini J., Gebru T., *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, "Proceedings of Machine Learning Research" 2018, 81, pp. 1-15.

Burt C., *Concerns about biometric online proctoring expressed by students in Australia, the US and Canada*, Biometric Update. 3 July 2020, https://www.biometricupdate.com/202007/concerns-about-biometric-online-proctoring-expressed-by-students-in-australia-u-s-and-canada.

Bygrave L.A., Tosoni L., *Article 4(14). Biometric Data*, [in:] *The EU General Data Protection Regulation (GDPR): A Commentary*, eds. C. Kuner, L.A. Bygrave, C. Docksey, Oxford University Press, Oxford 2020, pp. 207-216.

CAHAI. Feasibility study. CAHAI(2020)23. Council of Europe, Strasbourg. 17 December 2020.

Clifford D., Richardson M., Witzleb N., *Artificial Intelligence and Sensitive Inferences: New Challenges for Data Protection Laws*, [in:] *Regulatory Insights on Artificial Intelligence: Research for Policy*, eds. M Findlay et al., Edward Elgar, Cheltenham, UK, 2021, pp. 19-45.

Cobbe J., Ah Lee M.S., Singh *J., Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems. ACM Conference on Fairness, Accountability, and Transparency (FAccT' 21), 1-10 March 2021*, Virtual Event, Canada, ACM, New York, NY, USA.

Garton Ash T., *Free Speech: Ten Principles for a Connected World*, Kraków 2018.

Kind C., *Containing the canary in the AI coalmine – the EU's efforts to regulate biometrics* Ada Lovelace Institute Blog. 30 April 2021, https://www.adalovelaceinstitute.org/blog/canary-ai-coalmine-eu-regulate-biometrics/.

Kindt E., *A First Attempt at Regulating Biometric Data in the European Union* [in:] *Regulatory Insights on Artificial Intelligence: Research for Policy*, eds. M Findlay et al., Edward Elgar, Cheltenham, UK, 2021, pp. 62-69.

Spivak C., Garvie A., *Taxonomy of Legislative Approaches to Face Recognition in the United States*, [in:] *Regulating Biometrics: Global Approaches and Urgent Questions*, ed. A. Kak, AI Now Institute, 1 September 2020, pp. 86-95, https://ainowinstitute.org/regulatingbiometrics.html.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Izabela Gawłowicz
University of Zielona Gora
ORCID: 0000-0003-3591-3953
i.gawlowicz@wpa.uz.zgora.pl

From Outer Space to Algorithms: What Can Administrative Law Learn from International State Responsibility Regimes?

Keywords: administrative law, space law, international responsibility, algorithms

Summary. The ongoing process of digital transformation across all sectors of public life is fundamentally transformative, parallelling the revolutionary nature of the technologies that drive it. The development and deployment of artificial intelligence in public administration pose challenges to administrative law that are structurally analogous to those encountered by international law in relation to state activities in outer space: the absence of complete control, high levels of risk, the involvement of private actors, and the necessity of attributing ultimate responsibility to the state as the "ultimate guarantor".

The analogy advanced here is necessarily broad, given that the domains in question appear, *prima facie*, to be entirely distinct. Nevertheless, the effort to address questions such as who bears responsibility (and concomitant risk) for decisions made by AI, the legal status of AI and the possibility of conferring legal personality upon it, and the nature of the relationship between the state and private actors with regard to AI, reflects longstanding inquiries within the broader space sector. These inquiries encompass issues of accountability, the interaction between the state and private entities engaged in space activities, and the adequacy of the existing treaty framework governing such activities, including considerations for its potential modification where deficiencies are identified.

Od przestrzeni kosmicznej do algorytmów: jakie inspiracje prawo administracyjne może zaczerpnąć z międzynarodowych reżimów odpowiedzialności państwa?

Słowa kluczowe: odpowiedzialność państwa, odpowiedzialność algorytmu, odpowiedzialność w prawie kosmicznym, due dilligence w prawie kosmicznym i administracyjnym

Streszczenie. Proces transformacji cyfrowej, który obecnie zachodzi we wszelkich dziedzinach życia publicznego ma charakter przełomowy, tak, jak przełomowa jest technologia, która tę transformację uruchomiła. Rozwój i wykorzystanie sztucznej inteligencji w administracji publicznej stawia prawo administracyjne przed problemami strukturalnie podobnymi do tych, z którymi prawo międzynarodowe mierzyło się (i nadal się z nimi konfrontuje) w odniesieniu do działalności państw w przestrzeni kosmicznej: brakiem pełnej kontroli, wysokim stopniem ryzyk o różnym charakterze, aktywnością podmiotów prywatnych oraz koniecznością przypisania odpowiedzialności państwu jako „ostatecznemu gwarantowi". Zasygnalizowana tu, z konieczności luźna, analogia, tylko na pozór dotyczy sfer kompletne odległych i nie mających wspólnej części – w gruncie rzeczy poszukiwanie odpowiedzi

na takie pytania, jak: kto odpowiada (i tym samym, kto ponosi ryzyko) za decyzje podjęte przez AI, jaki jest status AI i czy można przypisać AI osobowość prawną, a także, jaka w perspektywie funkcjonowania AI jest relacja między państwem a podmiotami prywatnymi, odzwierciedla wieloletnie rozważania prowadzone w szeroko rozumianym sektorze kosmicznym. Rozważania te obejmują kwestie odpowiedzialności, interakcji między państwem a prywatnymi podmiotami prowadzącymi działalność w przestrzeni kosmicznej oraz adekwatności istniejących ram traktatowych regulujących tę działalność, w tym możliwości ich modyfikacji w przypadku stwierdzenia nieprawidłowości, do aktualnego stanu wiedzy, rozwoju technologicznego i cywilizacyjnego.

Introduction

The process of digital transformation currently unfolding across all spheres of public life is of a groundbreaking nature, much like the technology that has set it in motion. It affects all economic, social, and welfare systems and constitutes an integral component of the global strategy for sustainable development, as reflected in two key reports prepared in 2018 and 2025 by the research initiative The World in 2050 (TWI2050)[1]. In both reports, this process is described as a *digital revolution* and is attributed the role of a leading driving force of societal transformation. According to the authors of these reports, digital technologies are capable – faster than ever before – of supporting decarbonisation and the development of a circular economy, as well as the so-called sharing economy, dematerialisation, efficiency, and resource and energy sufficiency. They may also facilitate the monitoring and protection of ecological and other Earth systems, the safeguarding of global commons, and the promotion of sustainable behaviour. At the same time, these technologies, which also possess destabilising potential, have not yet been fully harnessed for the purposes of societal transformation towards sustainable development and a sustainable future. It appears self-evident that public administration bears the responsibility of taming these technologies[2] and embedding them in the service of humanity in the era of the Anthropocene[3]. It may seem trite to observe that the rapid development

[1] *The Digital Revolution and Sustainable Development: Opportunities and Challenges* chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.idos-research.de/uploads/media/TWI2050-for-web.pdf: https://pure.iiasa.ac.at/id/eprint/15913/ [accessed on: 15.03.2026].

[2] According to E. Szewczyk: "Contemporary public administration, confronted with an ongoing digital transformation, will constitute one of the driving forces behind the next major development in human history and a key engine of societal progress" (E. Szewczyk, *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne*, Warszawa 2025, *passim*).

[3] The Anthropocene is referred to as the geological epoch in which human activity has become the principal force shaping the Earth's environment, as well as an ethical-legal concept relating to the management of terrestrial, water, and outer space resources. This term—ambiguous and contested—is present in legal discourse, and its use reflects the need for a new approach to human agency in the world, as well as to responsibility for human actions (including in outer space and for climate change), a return to the conception of the Earth as a common good, and the framing of human rights as the legal basis for the protection of the natural environment. See more: D. Mnich; J. Karp, *Prawa człowieka w epoce i filozofii antropocenu*, "Ius Novum" 2025, Vol. 19, Nr 1, pp. 129-145; A special issue

of artificial intelligence has intensified the search for adequate normative solutions, which – according to a well-founded view among some experts –are currently taking shape under conditions of significant regulatory uncertainty, both at the national and international levels[4]. For instance, the European Artificial Intelligence Act (AI Act[5]) has been controversial already at the stage of negotiating its content – from the relatively broad definition of artificial intelligence it adopts, through the complex and analytically demanding classification of AI systems, in particular as high-risk systems entailing additional obligations for entities placing them on the market. As a consequence, the Act has been criticised for imposing excessive regulatory constraints on a dynamically developing technological sector, which may weaken Europe's competitiveness and lead to its relative slowdown in technological development processes[6]. Moreover, the concepts currently advanced remain fragmentary and reactive in nature and appear to constitute merely a preliminary stage in the development of a comprehensive regulatory framework[7]. A thorough and wide-ranging discussion on future regulations concerning the integration of humans and technology, whose normative articulation may become one of the principal challenges of the second half of the twenty-first century, still lies ahead. The increasingly widespread use of artificial intelligence in public administration gives rise to fundamental questions regarding the liability of the state for actions undertaken with the use of algorithmic systems. This is particularly pertinent in situations where administrative decisions are supported or made using AI-based tools, often designed and operated by private entities, while at the same time the effective control exercised by public authorities remains limited. The aim of the

of the journal "International Relations" 2024, 61/4 edited by M. Basu, the Author of the following considerations: *Reimagining International Relations in the Anthropocene* published in this very issue, was dedicated to the Anthropocene.

[4] J. Sobczak, *Czy świat algorytmów i sztucznej inteligencji będzie sprawiedliwy? Wykorzystywanie nowoczesnych technologii, algorytmów i modeli predykcyjnych w wymiarze sprawiedliwości*, [in:] *Human Rights Revolution in the Digital Era*, eds. M. Sitek, J.B. Navarro, Józefów 2021, p. 330.

[5] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) and the other Directives: 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 the texts available here: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng [accessed on: 15.03.2026].

[6] M. Ebers, V.R.S. Hoch, F. Rosenkranz, H. Ruschemeier, B. Steinrötter, *The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS),*available at: https://www.mdpi.com/2571-8800/4/4/43 [accessed on: 15.03.2026].

[7] P. Książek, *Prawo cyborgów (1). Wprowadzenie w problematykę*, "Głos Prawa/Przegląd Prawniczy Allerhanda" 2021, t. 4, nr 2 (8), poz. 29. As this author rightfully observes, artificial intelligence like the Internet or social media before it is proving to transform reality extremely quickly and so profoundly that, in effect, entirely new fields of study are emerging. A separate question is to what extent it is possible to ensure a symbiosis between these different spheres amid growing information chaos and the rapid, exponential development of technology.

present study is to propose an analytical comparative perspective that combines administrative law with the acquis of public international law[8], in particular with the regimes governing State responsibility for activities conducted in outer space. In both contexts, States face similar challenges: the constant fluctuation of the operational environment, a high level of risk and the unpredictability of certain consequences, technological complexity, the involvement of non-State actors, and difficulties in attributing specific conduct to public authorities. This study seeks to address the question of whether, and to what extent, solutions developed within public international law – such as the principle of attribution (jurisdiction), the duty of supervision, and the concept of due diligence – may serve as a source of inspiration for shaping administrative-law models of State responsibility for the use of artificial intelligence. The analysis encompasses selected systemic elements of the international legal order relating to outer space that exhibit analogies to the use of modern digital technologies, including artificial intelligence, in the functioning of public administration. For this reason, it is based on the application of the legal-dogmatic and comparative methods. The present considerations do not constitute a classical comparative analysis in the strict sense; rather, they aim to capture analogies of a conceptual nature which – consistent with the well-established methodological view that analogy constitutes an important cognitive tool – may serve as a point of departure for further reflection on contemporary legal challenges. The adoption of such a perspective is justified both by the distinctiveness of the normative orders under examination and by the need to maintain selectivity of argumentation within the constraints of length inherent in a scholarly article.

1. Artificial Intelligence Between Technology and Administrative Regulation

Artificial intelligence (AI) is not merely a technological challenge; from the perspective of the functioning of the State, it constitutes an administrative-law challenge. Designing normative solutions adequate to a reality undergoing rapid digital transformation requires the development of comprehensive conceptual frameworks—structures within which the law can respond systematically and consistently to emerging problems, while at the same time ensuring the protection of individual rights. Meanwhile, as noted by P. Książek, the body of law on artificial intelligence developed thus far is reactive and fragmented, lacking the necessary structure, coherence, and foundational underpinnings[9].

[8] Further as PIL.
[9] *Ibidem*.

State responsibility in the era of the algorithmisation of public administration is thereby reinforced as a key instrument for the protection of individual rights and for legitimising the exercise of public authority under conditions of digital transformation. From this perspective, the legal consequences of the use of artificial intelligence delineate a new "regulatory space," much as the exploration and use of outer space once led to the emergence of new normative structures. In order to substantiate this analogy as an analytical tool, it is necessary to point to the features common to both domains: the absence of full control, unpredictability and a high level of risk, the activity of diverse private actors and their complex relationships with the State, and the responsibility of the State as a "gatekeeper." Domains such as space activities and the use of artificial intelligence in public life also constitute regulatory areas characterised by a unique profile of ethical challenges, which must be taken into account in the law-making process. The term "regulatory space" is not a term of legal language and is only rarely used in the doctrine of law, typically in an intuitive manner without being assigned a precise meaning[10]. In the context of digital transformation, however, it appears to possess considerable analytical potential, as it does not refer to a territorially bounded domain but rather reflects the necessity of framing, in legal terms, the conduct of actors within a complex and rapidly evolving sphere of public life and environment. This sphere encompasses, inter alia, domestic and international legal regulations, the monitoring and governance of commercial, scientific, and security-related activities, and the attribution of responsibility for such activities. The concept of a "regulatory space" supports the assumption that the resources relevant to the exercise of regulatory authority and competences are dispersed or fragmented. They are not limited to formal legislative power or contractual relations, but also include information, capital, and organisational capacities. Control over these resources is distributed both among State authorities and between public and private actors. Under such conditions, certain entities such as corporations may acquire an informal hegemonic position by concentrating key resources, including organisational capabilities, technological capacity, and accumulated information. This informal position may, in turn, enable them to influence law-making or the application of law as a regulatory instrument.

[10] Relatively considerable attention has been devoted to this term and its meaning in the social sciences by L. Hancher, M. Moran, *Organizing Regulatory Space*, [in:] *Capitalism, Culture and Economic Regulation*, eds. L. Hancher, M. Moran, Clarendon Press 1989, pp. 272-277; and similar considerations can be found in M. Leach, *From Recklessness to Prudence. A Study of Regulatory Space Change in Indonesian Banking from 1997 to 2008*, Trinity 2015, pp. 11-23. In turn, M. Zachara argued already a decade ago that the regulatory space lying beyond the reach and control capacities of governments is expanding, creating a vacuum of authority that is readily filled by private actors (a point that is particularly relevant both in the context of outer space and the use of artificial intelligence in public life), see this Author's *Global governance. Ład międzynarodowy po zakończeniu stulecia Ameryki*, Kraków 2012, pp. 85 and 265.

In reality, therefore, the hierarchical subordination of regulated entities to the regulator is not always maintained within a regulatory space. This space is populated not only by regulated entities, but also by other stakeholders—both public and private—who, to varying degrees, possess relevant resources. The relationships between them are complex, dynamic, horizontal, and based on negotiated interdependence. The aforementioned dispersion of resources among heterogeneous actors within the same regulatory space deprives the regulator of a monopoly over both formal and informal power. In such a context, the negotiation of mutual relationships among active participants becomes a strategy not only for survival but also for effective action, while law-making and law enforcement constitute integral elements of this process[11]. It appears justified to apply the notion of a (new) "regulatory space" both to outer space and to those aspects of public administration that are based on algorithmic operations, in particular automated decision-making.

Ensuring sustainable development, a central objective and challenge of our time, requires the promotion of innovation across all spheres of life. The use of modern technologies in the functioning of public administration is therefore not only expected by contemporary societies but constitutes a necessity for the modern realisation of the rule of law. As aptly observed by E. Szewczyk, contemporary societies are transitioning from the industrial era (based on the rule of law) to the digital era (based on algorithms)[12], which entails confronting one of the most significant contemporary challenges in this field: the implementation widely discussed in legal doctrine of solutions involving the use of algorithms in the application of law, in particular for the automated issuance of administrative decisions. Referring to the previously outlined concept of a new regulatory space emerging or to be shaped in the context of integrating public administration with algorithmic tools, attention must be paid to the associated legal concerns, social implications, and axiological tensions inherent in these regulatory domains[13]. At the core lies the need to ensure appropriate legal safeguards both for decision-making processes involving algorithms and for State responsibility for their effects, over which full control cannot be exercised. These decision-making domains include social benefits, taxation, migration, and broadly understood security, encompassing personal,

[11] C. Scott, *Analysing Regulatory Space: Fragmented Resources and Institutional Design*, "Public Law" 2001, pp. 283-305.

[12] E. Szewczyk, *op. cit.*, p. LII.

[13] According to E. Szewczyk these concerns primarily revolve around the digital facilitation of totalitarian ambitions, elite domination, the deepening of social inequalities, complete surveillance, and the loss of freedom and social cohesion; the evolution of the artificial human and the blurring of boundaries between humans and machines; and whether, at the next stage of the digital revolution, sentient artificial beings with autonomous decision-making and reproductive capabilities could emerge (*ibidem*, p. 8).

digital, social, health, and financial dimensions. The use of algorithms necessitates compliance with standards applicable to the functioning of public authorities, including the principles of legality, transparency, and proportionality.

This is not a challenge faced solely by the Polish legislator – States are confronting it in various ways. On the one hand, the development of a normative model at the international level appears desirable; on the other hand, as illustrated by the analogy to international space law, there is a growing reluctance among States to assume treaty obligations, particularly in areas undergoing especially rapid and profound transformation.

2. State Responsibility and International Analogies

The question of who bears responsibility when a system fails is equally pertinent both: in the outer space where, due to factors such as environmental conditions, human activity cannot be fully controlled and may give rise to consequences that are unforeseeable even with the exercise of due diligence and in public administration operating with the use of algorithms with special regard of decision – making. *Prima facie*, in both cases, responsibility is attributed to the State.

2.1. State Responsibility in Complex Systems: Doctrinal Evolution from ICJ Jurisprudence to Algorithmic and Space Governance

In public international law, including international space law, it has not been possible to codify, by treaty, rules governing the responsibility of subjects of international law starting with states. There is, however, a functional consensus within the international community as to the constituent elements of such responsibility: states (which remain the most important group of subjects of international law) know which types of conduct give rise to responsibility, understand its consequences, are familiar with its forms of implementation, and accept these rules in practice. As aptly put by R. Sonnenfeld, responsibility constitutes the entirety of legal relations that arise in international law in connection with a breach of international obligations or with damage caused in the course of lawful activities[14]. Even though there has so far been no political will to fully codify the issue of responsibility and liability of states and other subjects in public international law, the doctrine of state responsibility is reflected in the jurisprudence of the International Court

[14] R. Sonnenfeld, *Podstawowe zasady odpowiedzialności międzynarodowej państwa*, [in:] *Odpowiedzialność państwa w prawie międzynarodowym*, ed. R. Sonnenfeld, Warszawa 1980, p. 22.

of Justice[15] and finds its expression in the Articles on Responsibility of States for Internationally Wrongful Acts[16].

The concept of state's responsibility in public international law has an allegedly simple, dual structure: a state incurs responsibility where conduct attributable to the state constitutes a breach of an international obligation. This dual requirement forms the core of international responsibility, as evidenced by the structure and normative content of ARSIWA as well as by the reasoning of the International Court of Justice. The ICJ has consistently affirmed that any internationally wrongful act entails responsibility, and that this finding has the status of general principle. Consequently, the Court's role is to determine whether conduct is attributable to the State and whether a breach of an international obligation can be established. According to the ICJ, attribution may be established on two distinct grounds. Attribution is straightforward and essentially automatic in respect of State organs, including armed forces operating extraterritorially, regardless of whether they act *ultra vires*, as reaffirmed by the ICJ in Armed Activities on the Territory of the Congo[17]. ICJ jurisprudence clearly establishes that a state exercising effective control over foreign territory incurs responsibility not only for its direct conduct, but also for failures to prevent violations within that sphere of control, thereby grounding responsibility in governance rather than in discrete acts. Doing so, ICJ already has moved beyond the classical paradigm of state responsibility (anchored in attributable conduct and breach) toward a broader conception based on omissions, foreseeability, and systemic failures of control.

The above-described approach ensures that the state evade international responsibility by invoking its internal law or the irregular conduct of its organs, since international law attributes such conduct to the State regardless of its conformity with domestic law or its *ultra vires* character. However, once conduct involves non-state actors, the Court adopts a markedly more restrictive stance. In Military and Paramilitary Activities in and against Nicaragua[18] the ICJ developed the "effective control" test. Under this standard, conduct of private actors is attributable to a state only if the state exercises control over the specific operations in which the wrongful acts occur while general support, financing, or even coordination is insufficient.

---

[15] Further as ICJ.

[16] Further as ARSIWA: United Nations Legislative Series, Materials on the Responsibility of States for Internationally Wrongful Acts (ST/LEG/SER B/25, U.N. Sales. No. E.12.V.12). The text has been reproduced in General Assembly resolution No 56/83 of 12 December 2001 and corrected by document A/56/49(vol. I)/Corr.4. See also the commentary and the other materials on ARSIWA in "Yearbook of the International Law Commission" 2001, vol. II (Part Two).

[17] *Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda*, Judgment – Merits) ICJ Reports 2005, paras 173-180.

[18] *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America*, Judgement – Merits) ICJ Rep 1986, paras 109-110 and 115.

Attribution is narrowly confined to situations of demonstrable operational control by the state over the conduct of non-state actors[19].

This restrictive attribution-based approach, however, generates an inherent structural tension. Where attribution is difficult to establish in multi-actor and causally opaque environments, there is a significant risk that harmful conduct escapes the scope of international responsibility. Such environments are characterized by fragmented agency, indirect chains of causation, and the diffusion of decision-making across institutional and technological layers, including state organs, delegated entities, and algorithmically mediated systems. The ICJ has responded to this problem in a gradual but doctrinally significant manner through the development of due diligence obligations. Already in the Corfu Channel Case[20], the Court articulated the principle that a state must not knowingly allow its territory to be used for acts contrary to the rights of other states. This obligation does not depend on strict attribution of wrongful conduct; rather, it imposes responsibility for failures of prevention in situations where harm is foreseeable and within the state's capacity to address. The same normative logic is further developed in the Bosnian Genocide Case[21], where the ICJ distinguished between responsibility for committing genocide and responsibility for failing to prevent it. The latter is governed by a due diligence standard, requiring states to employ all reasonably available means to prevent foreseeable harm within their sphere of influence or control.

It's worth to note that this insight becomes even more salient when one turns to contemporary forms of governance relying on algorithmic systems[22]. The comparison is instructive not because the classical framework is obsolete, but because its internal tensions become more visible under technological pressure. Algorithmic decision-making in public administration introduces features that strain the traditional framework. Decisions are no longer the product of singular, identifiable acts (or actors) but emerge from data-driven processes that are probabilistic, adaptive, and often opaque.

---

[19] What the ICJ confirmed in Bosnian Genocide Case, where the Court declined to attribute acts of genocide to Serbia despite substantial evidence of political and logistical support: *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro,* Judgment) ICJ Reports 2007, paras 398 and 437-376.

[20] *Corfu Channel (United Kingdom v Albania*, Judgment – Merits) ICJ Reports 1949.

[21] *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro*, Judgment) ICJ Reports 2007, paras 398 and 430.

[22] K. Creutz noted a few years ago that responsibility has become fragmented, and that the traditional bilateral approach to state responsibility also fits poorly with addressing certain modern legal challenges, including climate change or criminal responsibility regimes. According to this author, there is a tendency to broaden the scope of responsible actors beyond the state. See: K. Creutz, *The Tenacity of the Articles on State Responsibility as a General and Residual Framework: An Appraisal*, https://www.ejiltalk.org/the-tenacity-of-the-articles-on-state-responsibility-as-a-general-and-residual-framework-an-appraisal/ [accessed on: 28.04.2026].

From the perspective of ARSIWA, attribution appears straightforward at first: the system is deployed by a public authority and thus falls within the scope of state organs under Article 4. However, the situation becomes more complex when the system is developed by private entities, trained on external datasets, and operates with a degree of autonomy. Applying the effective control test from the Nicaragua case, it becomes difficult to argue that the state exercises control over various stages and specific outputs generated by the system. The result is a paradox the state clearly governs the system, yet may lack the level of control required for attribution under classical doctrine. Here the ICJ's due diligence jurisprudence becomes relevant. If one transposes the logic of Corfu Channel to the digital domain, the state's obligation is not limited to refraining from wrongful acts but extends to ensuring that systems under its jurisdiction do not produce harmful effects. The relevant question shifts from whether the state controlled a particular output to whether it took all reasonable steps to prevent foreseeable harm. In the context of algorithmic governance, such steps would include testing for bias, ensuring data quality, implementing human oversight, and establishing mechanisms for redress. Responsibility thus arises not from the act itself, but from the failure to adequately govern the conditions under which the act emerges. It's illustrated and developed in the following legal acts: the EU AI Act[23] and the OECD AI Principles[24]. These instruments adopt a fundamentally different approach to responsibility – instead of focusing on attribution of conduct and violation of obligation, they focus on risk management, continuous monitoring, and *ex ante* safeguards. Responsibility is distributed across multiple subjects: developers, deployers, operators, and is embedded in the lifecycle of the system. The objective is not merely to assign liability after harm occurs, but to prevent harm through structured oversight. The PIL's concept of responsibility shaped by ARSIWA and the ICJ's normative logic is retrospective and event-based, while algorithmic governance has to be prospective and process-oriented. PIL seeks to establish a causal link between conduct and harm, AI regulation has to accept uncertainty and operates on the basis of probabilistic risk. ARSIWA presupposes control as the basis of responsibility, algorithmic governance replaces control with accountability for managing systems that generate risk. In this case, the due diligence obligations already presented in the ICJ's jurisprudence may evolve into the primary mechanism for engaging responsibility. The states would be held responsible not because they caused harm in a direct sense, but because they failed to adequately regulate the systems through which harm materialized.

[23] Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L, 12.7.2024.

[24] https://www.oecd.org/en/topics/ai-principles.html [accessed on: 27.04.2026].

The question of responsibility is no longer simply about who acted, but about who designed, deployed, and failed to control the systems through which action occurs.

The other important elements of the classical framework of state responsibility remain resistant to adaptation. The law of reparation, anchored in the Factory at Chorzów[25] as the principle of full reparation, presupposes that harm can be identified, quantified, and even reversed. In the context of algorithmic systems, harm is often diffuse, collective, and ongoing. Compensation, as illustrated in DRC v Uganda Reparations[26], remains the dominant remedy in PIL, but it may be ill-suited to address systemic discrimination or loss of trust in public institutions.

From the provisions of international space law, including the two treaties forming its normative core: the Outer Space Treaty adopted in Moscow, London and Washington on the 27th of January 1967 (in particular Articles VI and VII)[27] and the Convention on International Liability for Damage Caused by Space Objects adopted in Moscow, London and Washington the 29th of November 1971 (with particular emphasis on Articles II, IV, V and VI)[28] it follows that a state bears responsibility for the acts of its own organs as well as private entities, even where it does not exercise full control over their activities. Under these provisions, in the event of damage caused beyond the surface of the Earth to a space object of one state by a space object of another launching state, the responsibility of the latter arises if the damage is due to its fault or the fault of persons for whom it is responsible. This reflects the essence of attributing responsibility to the state under international law for activities carried out in outer space. Such responsibility stems not from direct authorship, but from the role the state has assumed in this sphere. On the one hand, in outer space and more broadly beyond the limits of national jurisdiction, there is no equivalent of what domestic law would describe as a sovereign: there is no overarching legislative, executive, or enforcement authority empowered to regulate the conduct of, *inter alia*, private actors. Consequently, it is not feasible to attribute responsibility directly to a company at the international level or to enforce it within the decentralized system of the international legal order. On the other hand, at the time when the foundations of international space law were being established, the private space sector was virtually non-existent, and the only entity whose responsibility could be defined and enforced was (and,

[25] *Factory at Chorzów (Germany v Poland*, Judgment – Merits) PCIJ Series A No 17, 1928 (Sept. 13).

[26] *Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda* – Reparations) ICJ Reports 2022, para 108.

[27] UNTC 610/1967, No 8843, the text available also at: https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html [accessed on: 25.02.2026].

[28] UNTS 961/1983, No. 13810 the text available also at: https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/liability-convention.html [accessed on: 25.02.2026].

in this sphere, continues to be) the state. The international community has thus developed, in a sense, a substitute mechanism of responsibility based on risk rather than solely on fault, placing the state conceived as the guardian of access to outer space, at its centre. The burden of identifying the actual perpetrator of the damage and of ensuring effective reparation (in particular through the payment of compensation) is shifted to the level of internal relations between the state and the entity conducting the relevant activities. From the perspective of international law, such internal settlements, including any possible recourse mechanisms, are irrelevant, since responsibility vis-à-vis external actors is consistently attributed to the state. A state may delegate technologically complex activities, but it remains under a duty of due diligence[29], retains and must exercise supervision, bears responsibility, and must ensure the safety of those activities.

Finally, this structure of responsibility also stimulates the state: it not only prevents states from evading responsibility by shifting it onto private companies and from creating gaps in accountability, but also compels them to develop domestic legislation[30] and control mechanisms (such as registries, licensing systems, and supervisory frameworks). Space law is therefore, *inter alia*, a system for governing high-risk technologies, not merely a regime of classical liability for wrongful acts, and outer space is not only a domain of promising opportunities but also a common good of all humankind where any damage may have global consequences. Adopting a broad scope of responsibility is thus a more adequate solution than leaving it indeterminate, as it prevents situations in which no responsible subject can be identified.

[29] K. Szpyt, A. Bilski, *Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych*, "Prawo i Więź" 2025, nr 1(54), *passim*.

[30] It is, however, difficult not to note that this pressure does not always translate into practice—the Polish Act on Space Activity, for instance, only reached its final form in February of this year and was forwarded in March for the signature of the President of the Republic of Poland: https://www.sejm.gov.pl/Sejm10.nsf/PrzebiegProc.xsp?nr=2078 [accessed on: 13.03.2026] however, legislative working on it spanned over a dozen years, and its reception among space law experts has not been enthusiastic, as evidenced, for example, by the ambiguous opinion of the Legislative Council dated 13 June 2025, available at: https://www.gov.pl/web/radalegislacyjna/opinia-z-13-czerwca-2025-r-o-rzadowym-projekcie-ustawy-o-dzialalnosci-kosmicznej?utm_source=chatgpt.com [accessed on: 13.03.2026]. For instance, see some critical reflactions: K. Malinowska, *O polskim prawie kosmicznym i jego pożądanym kształcie*, "Ad Astra" 2022, nr 4; also Sz. Siuda, *Reglamentacja działalności kosmicznej w Polsce: zakres podmiotowy obowiązku uzyskania zezwolenia w świetle projektu ustawy z 12 maja 2025 r.*, "Studia Prawa Publicznego" 2025, nr 4(52).

2.2. State Responsibility for Its Own Conduct (and That of Other Entities) in the Use of Digital Technologies and AI

The use of artificial intelligence systems in public administration is inherently associated with a fragmentation of causation[31], whereby the decision-making process is dispersed across multiple actors (the model developer, the technology provider, the implementing entity, and the administrative authority itself) and across successive technological stages. Public institutions operate under increasing pressure, both financial and stemming from citizens' expectations regarding the quality and accessibility of digital services, as well as demands to implement solutions that simplify and accelerate administrative procedures. Under such conditions, identifying a single author of an administrative decision (a civil servant, the administrative authority, or an autonomous system itself) becomes difficult and, at times, even illusory. This gives rise to the risk of an accountability gap, consisting in the inability to attribute responsibility to a specific entity. The legal system, however, does not tolerate such gaps and does not accept situations in which the exercise of public authority falls outside a regime of responsibility. For this reason, it is necessary to designate an entity serving as the ultimate bearer of responsibility (analogous to the substitute liability mechanism developed in space law and referred to above, under which responsibility is borne exclusively by the State), while at the same time ensuring transparency and the protection of individual rights in decision-making processes involving AI. Responsibility thus naturally falls upon the public administrative authority as the entity vested, in the first instance, with democratic legitimacy and decision-making competences. This responsibility does not stem from direct authorship of the act, but rather from the role assumed by the administration in this sphere as organiser and user of AI systems, and therefore has a functional character.

2.3. Due Diligence and Standards of Integrity in Public Administration

In the analysed context, particular importance attaches to the concept of due diligence, which must be understood as an obligation on the part of public administrative authorities to undertake all reasonably required measures to prevent

[31] On the inadequacy of classical models of liability and the diffusion of responsibility in the use of AI patrz *Prawo sztucznej inteligencji*, eds. L. Luigi, M. Świerczyński, Warszawa 2020. Also see M. Świerczyński, *Sztuczna inteligencja a odpowiedzialność cywilna – uwagi na tle prawa UE*, M. Świerczyński, *Odpowiedzialność cywilna za szkodę wyrządzoną przez sztuczną inteligencję*, [in:] *Prawo sztucznej inteligencji*, s. 57-80 as well as M. Świerczyński, Z. Więckowski, *Sztuczna inteligencja w prawie międzynarodowym. Rekomendacje wybranych rozwiązań*, Warszawa 2021 and those Author's: *Liability for damages caused by artificial intelligence systems – main challenges to be addressed by the European Union conflict-of-laws regulations*, "Prawo w Działaniu" 2023, nr 54, pp. 200-2016.

violations of law and to minimise risks associated with the use of artificial intelligence systems. The notions of "due diligence" (*due diligence*) and "integrity" (*diligentia*), while not identical, are closely related. Due diligence focuses on the process of action undertaken by a public authority or entity and signifies that all reasonably required steps are taken to prevent harm or legal violations. It has an objective character, as it involves assessing whether the authority acted as could be expected of a competent entity in comparable circumstances. In administrative practice, this includes, *inter alia*, the preparation of decisions, procedural oversight, supervision of systems (including algorithmic systems), and risk assessment. Integrity, by contrast, relates more to the attitude of the actor, encompassing diligence, honesty, and compliance with the principles of professional ethics. It has a qualitative and subjective dimension, referring to the motivations of the authority and its commitment to the proper conduct of procedures. In this sense, due diligence may be regarded as a formalised dimension of integrity, combining the obligation of proper and rational conduct with the ethical and professional posture of public administration. In conditions of increasing autonomy of algorithmic systems, due diligence is not limited to the decision-making act itself, but extends to the entire lifecycle of the system from its selection and deployment to ongoing supervision and control of its operation. The concept of due diligence enables a functional linkage between the responsibility of public administration and the use of artificial intelligence systems, serving as a mechanism that compensates for the fragmentation of agency inherent in algorithmic processes. The analogy to international space law is not incidental once again, a similar legal mechanism operates across two distinct normative systems.

In administrative law doctrine, it is emphasised that public authorities are obliged to act with due diligence and integrity, which includes proper supervision and control over decision-making processes irrespective of the tools employed. In the context of the use of artificial intelligence in public administration, due diligence functions as a formalised expression of administrative integrity, encompassing both the obligation to undertake all reasonably required actions to ensure the correctness of decision-making processes and the systematic supervision, auditing, and control of algorithmic tools, while maintaining transparency and fairness in accordance with standards of administrative professionalism. Due diligence thus serves as an instrument of technological risk management, enabling the attribution of responsibility even in situations where direct agency remains dispersed. The adoption of such an approach contributes to ensuring effective protection of individual rights and transparency in administrative action. At the same time, it fulfils a preventive function, incentivising authorities to implement mechanisms of supervision, control, and auditing of algorithms. Consequently, the responsibility

of public administration for the use of AI constitutes – and must constitute – an instrument stabilising the legal system under conditions of increasing technological complexity and volatility[32].

Conclusions

The current digital transformation in public administration may be likened to a proton storm: just as charged particles penetrate the atmosphere and produce unpredictable effects, so to do digital technologies and AI systems, as they become embedded in the structures of administration and public life, accelerate its dynamics, introducing impulses and changes that necessitate new regulations, supervisory mechanisms, and the readaptation of existing law.

The experience of public international law concerning State responsibility for space activities may serve as a source of inspiration for administrative law, particularly in the context of the use of artificial intelligence systems. By analogy to the mandatory supervision mechanisms and the exclusive responsibility of the State in outer space, public administration should adopt an approach grounded in due diligence[33], encompassing both the organisation and control of decision-making processes and the systematic oversight of algorithmic tools. This concept makes it possible to attribute functional responsibility even in situations where technical agency is dispersed, while at the same time ensuring the protection of individual rights and the preservation of standards of administrative integrity.

As a result, the use of AI in public administration need not lead to an erosion of responsibility; rather, it calls for its adaptation to technological complexity, indicating directions for the further development of domestic legal frameworks and supervisory procedures. In space law, State responsibility constitutes a systemic necessity: broadly conceived, it compensates for the absence of a sovereign. In the sphere of digital technologies and AI in public administration, it performs a sta-

[32] See more: *Administracja publiczna i prawo administracyjne wobec współczesnych ryzyk*, eds. K. Kurczewska, K. Kurzępa-Dedo, Rzeszów–Szczecin 2018.

[33] See E. Chmura-Szczecińska, *Automatyzacja decyzji w prawie – pola i granice stosowania*, "Studia Prawnicze" 2021, nr 2(224), pp. 135-159; J. Sługocki, *Rzetelność jako konstytucyjny wymóg jakości działania administracji publicznej*, "Przegląd Prawa Konstytucyjnego" 2023, nr 3(73), pp. 11-23. As M. Kupiec aptly points out, the authority's implementation of the principle of conducting proceedings in a manner that builds trust in public administration does not preclude […] the use of standardized, publicly available templates of justifications in the case of routine, mass-issued decisions based on results generated by artificial intelligence systems. However, it is a condition that officials must be able to individualize such templates with regard to facts and legal grounds that deviate from the standard justification appropriate for a given type of administrative case. […], this Author's: *Uzasadnienie decyzji administracyjnej wydanej za pomocą systemu sztucznej inteligencji w świetle realizacji zasady budowania zaufania do władzy publicznej*, "Dyskurs Prawniczy i Administracyjny" 2025, nr 1, pp. 137-147.

bilising function for legal certainty under conditions of fragmented agency and represents a regulatory choice; in this sense, it also serves as an effective instrument of risk management, incentivising administrative authorities to ensure an adequate level of oversight over the technologies they employ.

At the same time, both in space law and in public administration utilising AI systems, specific legal challenges arise in attributing responsibility for decisions taken within multi-stage processes involving both public and private actors, and which are not fully subject to human control. In space law, this problem manifests itself in the need to ensure that the State effectively supervises the activities of space operators and is able to assume liability for potential damage vis-à-vis third parties, even where it does not directly control every action. In AI-enabled administration, by contrast, the challenge lies in defining standards of due diligence in the design, deployment, and ongoing monitoring of algorithmic systems, so that decisions remain lawful, protect individual rights, and are transparent.

In both cases, it becomes essential to develop mechanisms that combine the functional responsibility of public authorities with effective oversight of technological and procedural tools, rather than relying solely on general frameworks of liability.

References

Books and articles

Administracja publiczna i prawo administracyjne wobec współczesnych ryzyk, eds. K. Kurczewska, K. Kurzępa-Dedo, Rzeszów–Szczecin 2018.

Basu M., *Reimagining International Relations in the Anthropocene*, "International Relations" 2024, 61/4, pp. 355-363.

Chmura-Szczecińska E., *Automatyzacja decyzji w prawie – pola i granice stosowania*, "Studia Prawnicze" 2021, nr 2(224), pp. 135-159.

Ebers M., Hoch V.R.S, Rosenkranz F., Ruschemeier H., Steinrötter B., *The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)*, https://www.mdpi.com/2571-8800/4/4/43.

Hancher L., Moran M., *Organizing Regulatory Space*, [in:] *Capitalism, Culture and Economic Regulation*, eds. L. Hancher; M. Moran, Clarendon Press 1989 , pp. 148-172.

Książek P., *Prawo cyborgów (1). Wprowadzenie w problematykę*, "Głos Prawa/Przegląd Prawniczy Allerhanda" 2021, t. 4, nr 2(8), pp. 551-563.

Kupiec M., *Uzasadnienie decyzji administracyjnej wydanej za pomocą systemu sztucznej inteligencji w świetle realizacji zasady budowania zaufania do władzy publicznej*, "Dyskurs Prawniczy i Administracyjny" 2025, nr 1, pp. 142-147.

Leach M., *From Recklessness to Prudence. A Study of Regulatory Space Change in Indonesian Banking from 1997 to 2008*, Trinity 2015.

Malinowska K., *O polskim prawie kosmicznym i jego pożądanym kształcie*, "Ad Astra" 2022, nr 4.

Mnich D.; Karp J., *Prawa człowieka w epoce i filozofii antropocenu*, "Ius Novum" 2025, Vol. 19, Nr 1, pp. 129-145.

Prawo sztucznej inteligencji, eds. L. Luigi, M. Świerczyński, Warszawa 2020.

Scott C., *Analysing Regulatory Space: Fragmented Resources and Institutional Design*, "Public Law" 2001, pp. 329-353.

Sobczak J., *Czy świat algorytmów i sztucznej inteligencji będzie sprawiedliwy? Wykorzystywanie nowoczesnych technologii, algorytmów i modeli predykcyjnych w wymiarze sprawiedliwości*, [in:] *Human Rights Revolution in the Digital Era*, eds. M. Sitek, J.B. Navarro, Józefów 2021, pp. 305-334.
Szpyt K., Bilski A., *Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych*, "Prawo i Więź" 2025, nr 1(54), pp. 555-596.
Siuda S., *Reglamentacja działalności kosmicznej w Polsce: zakres podmiotowy obowiązku uzyskania zezwolenia w świetle projektu ustawy z 12 maja 2025 r.*, "Studia Prawa Publicznego" 2025, nr 4(52), pp. 127-137.
Sługocki J., *Rzetelność jako konstytucyjny wymóg jakości działania administracji publicznej*, "Przegląd Prawa Konstytucyjnego" 2023, nr 3(73), pp. 13-23.
Szewczyk E., *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne*, Warszawa 2025.
Świerczyński M., Więckowski Z., *Liability for damages caused by artificial intelligence systems – main challenges to be addressed by the European Union conflict-of-laws regulations*, "Prawo w Działaniu" 2023, nr 54, pp. 200-216.
Świerczyński M., Więckowski Z., *Sztuczna inteligencja w prawie międzynarodowym. Rekomendacje wybranych rozwiązań*, Warszawa 2021.
Zachara M., *Global governance. Ład międzynarodowy po zakończeniu stulecia Ameryki*, Kraków 2012.

Sources

Treaty on principles governing the activities of States in the exploration and use of outer space, including the moon and other celestial bodies 27.01.1967, UNTC 610/1967, No 8843.
Convention on International Liability for Damage Caused by Space Objects 29.03.1972, UNTS 961/1983, No. 13810.
Responsibility of States for Internationally Wrongful Acts 2001, ILC 53-rd Session, General Assembly Resolution No 56/83 of 12 December 2001 and corrected by document A/56/49(Vol. I)/Corr.4.
Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L, 12.7.2024. Opinion of the Legislative Council 13 June 2025.

Jurisprudence

Factory at Chorzów (Germany v Poland, Judgment – Merits) PCIJ Series A No 17, 1928 (Sept. 13).
Corfu Channel (United Kingdom v Albania, Judgment – Merits) ICJ Reports 1949.
Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America, Judgement – Merits) ICJ Reports 1986.
Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda, Judgment – Merits) ICJ Reports 2005.
Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro, Judgment) ICJ Reports 2007
Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v Uganda – Reparations) ICJ Reports 2022.

Others

United Nations Legislative Series, Materials on the Responsibility of States for Internationally Wrongful Acts (ST/LEG/SER B/25, U.N. Sales. No. E.12.V.12).
The Digital Revolution and Sustainable Development: Opportunities and Challenges chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.idos-research.de/uploads/media/TWI2050-for-web.pdf [accessed on: 10.03.2026].
AI Principles OECD 2019/2024.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Winfried Huck
Ostfalia Hochschule für angewandte Wissenschaften
ORCID: 0000-0001-5644-9067
w.huck@ostfalia.de

Artificial Intelligence in Administrative Law: Structural Deficits, Control Gaps and Platform Dependency. A Jurisprudential and Administrative Law Analysis

Keywords: artificial intelligence, VwVfG, methodological control, platform dependency, digital sovereignty, human-in-the-loop, Kelsen, Hart, AI Regulation, CLOUD Act

Summary. This paper analyses the deployment of artificial intelligence (AI) in public administration across four structurally distinct yet normatively interconnected levels. It connects jurisprudential foundations — Kelsen's Grundnorm and Hart's Rule of Recognition — with an administrative law review of the structural mismatches between the historically sedimented Federal Administrative Procedures Act (*Verwaltungsverfahrensgesetz*, VwVfG) and the logic of probabilistic language models; the European law framework established by the AI Regulation (EU) 2024/1689 and the AI Market Surveillance and Innovation Act (*Künstliches-Intelligenz-Marktüberwachungs* und *Innovationsförderungsgesetz*, KI-MIG); the institutional modernisation policy of the Federal Ministry for Digital Affairs and State Modernisation (*Bundesministerium für Digitales und Staatsmodernisierung*, BMDS); and the geopolitical dimension of digital dependency on externally controlled platform architectures. The central jurisprudential question is: whose normative claim to validity sustains the administrative decision when the underlying AI system is governed by an external, state-uncontrollable platform architecture?

Sztuczna inteligencja w prawie administracyjnym: deficyty strukturalne, luki kontrolne i zależność platformowa. Analiza z perspektywy teorii prawa i prawa administracyjnego

Słowa kluczowe: sztuczna inteligencja, niemiecka ustawa o postępowaniu administracyjnym (VwVfG), kontrola metodologiczna, zależność platformowa, suwerenność cyfrowa, człowiek w pętli decyzyjnej (human-in-the-loop), Kelsen, Hart, rozporządzenie w sprawie sztucznej inteligencji, CLOUD Act

Streszczenie. Niniejszy artykuł analizuje wykorzystanie sztucznej inteligencji (AI) w administracji publicznej na czterech strukturalnie odrębnych, lecz normatywnie powiązanych płaszczyznach. Łączy on podstawy teoretycznoprawne — Grundnorm Kelsena oraz Rule of Recognition Harta — z analizą administracyjnoprawną strukturalnych niezgodności między historycznie ukształtowaną federalną ustawą o postępowaniu administracyjnym (Verwaltungsverfahrensgesetz, VwVfG) a logiką probabilistycznych modeli językowych; z ramami prawa europejskiego wyznaczonymi przez rozporządzenie (UE) 2024/1689 w sprawie sztucznej inteligencji oraz ustawę o nadzorze rynku sztucznej inteligencji

i wspieraniu innowacji (Künstliches-Intelligenz-Marktüberwachungs- und Innovationsförderungsgesetz, KI-MIG); z instytucjonalną polityką modernizacyjną Federalnego Ministerstwa ds. Cyfryzacji i Modernizacji Państwa (Bundesministerium für Digitales und Staatsmodernisierung, BMDS); oraz z geopolitycznym wymiarem cyfrowej zależności od zewnętrznie kontrolowanych architektur platformowych. Centralne pytanie teoretycznoprawne brzmi: czyje normatywne roszczenie obowiązywania podtrzymuje decyzję administracyjną, gdy system AI stanowiący jej podstawę podlega zewnętrznej, niekontrolowalnej przez państwo architekturze platformowej?

I. Introduction: The Fourfold Structural Problem

German administrative law faces a fundamental challenge that unfolds across at least four interconnected dimensions. First, the structural foundation of German administrative decisions — the Federal Administrative Procedures Act (*Verwaltungsverfahrensgesetz*, VwVfG) — is not designed for probabilistic AI systems: it presupposes linear decision-making processes, personal attribution of responsibility and (mostly) text-based traceability[1]. Second, EU law, in the form of the AI Regulation, overlays national law with stringent requirements for whose integration into administrative procedural law there is still no normative framework, even though a draft implementing statute was submitted in Germany in February 2026 (the Government Draft KI-MIG). Third, the Federal Government has been operating autonomous administrative agents in 17 municipalities since March 2026 — through the Agentic AI Hub — without any coherent AI-specific reform agenda for German administrative law at federal and state level being discernible, and without Section 35a or Section 39 VwVfG being designed for these systems. Fourth — and this is the dimension that reaches deepest in jurisprudential terms — German public administration operates AI systems predominantly on US platform infrastructure subject to the Clarifying Lawful Overseas Use of Data Act (CLOUD Act), 18 U.S.C. § 2713: an extraterritorially operating access authorisation that structurally jeopardises the sovereign audit trail and thus the reconstructibility of the grounds for decisions. This appears to apply — as far as can be established — to the AI Opportunities Marketplace (MaKI) as well, which currently lists 265 AI systems[2] across German federal, state and municipal administration.

These four problem dimensions can be condensed into a single jurisprudential core question: *whose normative claim to validity sustains the administrative decision when the underlying AI system is governed by an external, state-uncontrollable platform architecture and national procedural law provides no normative guidance for the integration of such systems?* This paper addresses that question across four levels,

[1] A. Müller, Digitalisation of Paper in Public Administration, *RDi – Recht Digital* 2024, p. 557; A. Guckelberger, Germany's (Lag in) Administrative Digitalisation, *LTZ – LegalTech: Zeitschrift für die digitale Rechtsanwendung* 2023, p. 167.

[2] Stand 25.03.2026.

connecting the jurisprudential foundation with the administrative law analysis, the European regulatory dimension, and the geopolitical reality of digital dependency.

II. Jurisprudential Foundations: Methodological Control and the Basis of Validity

1. Normative Legal Facts: How Large Language Models Work — and Where Their Limits Lie

The legal-policy debate on AI deployment in public administration often suffers from an insufficient understanding of its object. Those who do not understand how a large language model functions technically cannot formulate the normative requirements for its deployment with precision, nor determine its inherent limits. This section establishes the technical foundations as normative legal facts[3].

a) Artificial Intelligence as a System Category (Article 3(1) AI Regulation)

Article 3(1) of the AI Regulation[4] defines an 'AI system' as a machine-based system designed to operate with varying degrees of autonomy, which may adapt after deployment, and which, from the inputs it receives, infers how to generate outputs — such as predictions, content, recommendations or decisions — that can influence physical or virtual environments. For the category of interest here — large language models deployed for text generation, legal research and decision preparation — the decisive characteristic is that these systems do not operate on a rule basis; they learn no explicit if-then structures[5] but instead extract statistical patterns from enormous volumes of text and reproduce those patterns in response to new inputs. This structural feature — statistical *pattern recognition* rather than rule-based decision logic — is the starting point for the normative requirements applicable to their deployment in legally bound contexts.

b) Large Language Models: Structure and Performance Limits

A large language model (LLM) is a transformer-architecture-based neural network[6] that encodes the statistical structure of human language in parametric

---

[3] A. Nussbaum, *Legal Facts Research: Its Significance for Scholarship and Teaching*, Tübingen (J.C.B. Mohr) 1914; M. Rehbinder, *The Founding of the Sociology of Law by Eugen Ehrlich*, Berlin (Duncker & Humblot) 1967.

[4] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (AI Act); consolidated version: https://eur-lex.europa.eu/eli/reg/2024/1689 [accessed on: 22.04.2026].

[5] *BeckOK AI Law*/J. Kirschke-Biller/A. Füllsack, 5th ed., as of 01.11.2025, AI Act Art. 3 para. 57–60.

[6] A. Vaswani et al., Attention Is All You Need, *NeurIPS* 2017; Y. Zhang et al., Large Language Models: A Survey, arXiv:2402.06196v2 (2024), https://arxiv.org/abs/2402.06196 [accessed on: 22.04.2026]; A. Hebing, E. Kastlunger, Artificial Intelligence in Sales and Marketing, *ZVertriebsR* 2026.

weight matrices through training on extensive text corpora. The model consists of hundreds of billions of numerical parameters. Once training is complete, these parameters are immutable; the model learns nothing new during subsequent use — it reconstructs what it extracted from texts during training.

What the model has thereby learned is not propositional (true) knowledge. It has learned the statistical relations between linguistic units, not the normative validity order and the inherent methodology interwoven in the various linguistic strata of the LLM. Legal texts, with their respective historical, cultural, methodological and systematic reference points at different levels of abstraction, constitute a specialised language. This specialised language can only be decoded through the legal methodology of the relevant legal system. From this basic structure two performance limits follow: the model 'hallucinates' — it occasionally produces formally coherent but professionally inaccurate statements. It also sometimes lacks access to the present; developments after the training cut-off are unknown, unless the gap between training completion and the present can be closed by means of the LLM itself conducting an internet search, which is increasingly the case.

c) Tokens and Vectorisation: How Language Models Process Text

An LLM does not process text as a sequence of words but as a sequence of *tokens* — the smallest units into which text is broken down by a prior tokenisation process[7]. The tokens are transformed into high-dimensional numerical vectors — so-called *embeddings*[8]. Semantically similar terms — such as *Ermessen* (administrative discretion) and *Beurteilungsspielraum* (margin of assessment), or *Verwaltungsakt* (administrative act) under Section 35 VwVfG and *acte administratif* in French administrative law — are positioned spatially close together in this high-dimensional space because they have appeared in similar linguistic contexts. This statistical proximity in the vector space does not, however, say anything about the respective doctrinal or historical content.

The attention mechanism of the transformer architecture determines which tokens are weighted most strongly when generating a new token. Multi-head attention[9] enables multiple relations to be captured simultaneously: normative rank, teleological purpose, and systematic context can be activated in parallel — without the model establishing an explicit hierarchy between them. The model can thus reproduce the outward form of legal argumentation without having grasped its inner basis of validity.

---

[7] L. Lemnitzer, K.-M. Würzner, The Word in Language Technology, [in:] *Handbook of Word and Vocabulary*, vol. 3, eds. U. Haaß-Zumkehr, W. Kallmeyer, G. Zifonun, Berlin–Boston (de Gruyter) 2015, p. 297.

[8] T.W. Dornis, Generative *AI Training and the TDM Fallacy*, GRUR 2024, p. 1681.

[9] P. Jin et al., *MoH: Multi-Head Attention as Mixture-of-Head Attention*, arXiv:2410.11842 (2024), https://arxiv.org/abs/2410.11842 [accessed on: 22.04.2026].

d) Probabilism, Training Data Bias and Their Legal Consequences

The outputs of an LLM are the result of probabilistic processes[10]. The model calculates a probability distribution for every possible continuation and selects the next token from that distribution. The output is statistically the most plausible continuation based on numerical sequences, not the legally compelling result. Plausibility is a property of the statistical distribution in the training corpus; professional correctness is a property of the relationship between a statement and the normative validity order. Both properties are categorially distinct.

This probabilistic basic structure has a consequence that is particularly significant for legal applications[11]:if English-language legal texts from common law jurisdictions dominate a training corpus numerically — which is empirically plausible for the major commercial models — then the LLM tends to activate common law concepts and methodology as statistically dominant patterns, even when the question posed concerns a civil law jurisdiction. Not because the model would confuse the systems, but because the quantitative dominance of one discursive style in the training corpus increases its statistical probability in the output. This phenomenon — *training data bias* — produces systematic, non-random distortions towards dominant legal discourses[12].

e) Methodology with Constraints: The Decisive Systemic Difference

An LLM that is fed legal texts still operates as a statistical language model on the basis of document structures through pattern recognition. It reproduces the methodology interwoven in those documents without clearly grasping it or drawing its implications. It cannot generate independent methodological insights because methodology is not a property of documents but an — usually unwritten — property of the legal order from which those documents derive.

The decisive systemic difference lies in the question of whether the system architecture imposes external *methodological constraints* on the model. Constraints are control parameters that instruct the model which interpretive framework applies at which normative level[13], how norm conflicts are to be resolved, which conflict rules — *lex superior, lex specialis, lex posterior* — are to be activated in

---

[10] F. Brauner, M. Murawski, M. Bick, *The development of a competence framework for artificial intelligence professionals using probabilistic topic modelling*, „Journal of Enterprise Information Management" 2025, S. 197.

[11] Are Large Language Models like ChatGPT offering legal service (Rechtsdienstleistungen) according to § 2 Abs. 1 Rechtsdienstleistungsgesetz (RDG), see F. Baumer, *Large Language Models auf dem Prüfstand des RDG*, „RDi" 2026, S. 8.

[12] D. Lacmanovic, M. Skare, *Artificial intelligence bias auditing*, „Review of Accounting and Finance" 2025, Bd. 24, Nr. 3, S. 375; M. van Bekkum, *Using sensitive data to de-bias AI systems*, „Computer Law & Security Review" 2025, Bd. 56.

[13] Z. Yang et al., *Plug in the safety chip: Enforcing constraints for LLM-driven robot agents*, [in:] *Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

which sequence, and where methodological uncertainties are to be identified as such. Without external methodological constraints, an LLM operates as a document-centred retrieval and generation system. With external constraints, it can be deployed as an instrument of methodologically controlled reconstruction of legal argumentation. This difference is categorial in nature: it determines whether the output can constitute a legally sound basis for a decision.

2. On the Normative Capacity of AI Systems in Law

The question of whether and to what extent AI can be deployed in legal — and thus administrative — argumentation[14] must be distinguished from the question of whether AI outputs can constitute legally binding grounds for decisions[15]. An LLM can conduct grammatical and systematic interpretation, form analogies and apply teleological reductions, because the evaluative logics on which these rest are implicitly present across the entire doctrinal discourse of a specific legal order. Legal methodology is formalisable in a precise sense. The Savigny canon — grammatical, systematic, historical and teleological interpretation[16] — consists of abstract rule structures that are describable independently of their concrete exercise and reconstructable by a sufficiently trained model[17].

The central thesis is thus: an LLM can generate methodologically correct, logically consistent and systemically tenable legal argumentation options under German law. This capacity is not trivial — and it has direct consequences for administrative practice: where AI prepares structured discretion, subsumes facts and formulates decision proposals, it can deliver legally sound inputs. However, in Germany the legal foundations for this are lacking. A comprehensive administrative reform placing AI at the c enter of federal and state government is currently not in sight[18]. Rather, it is ideas of acceleration that have recently been driving the agenda, including for administrative (procedural) law[19]. The Federal Government's guide-

[14] The German Government (BMDS) intends to implement legal administrative AI by the end of 2029, see AI Marketplace (MaKI): https://www.kimarktplatz.bund.de [accessed on: 22.04.2026]; I. Bantekas, E. Bratsiakou, *Automated Decision-Making Systems and Black Box Challenges Under European Union Administrative Law*, „Fordham International Law Journal" 2026, Bd. 49, Nr. 1, S. 1.

[15] Y. Chandra, S. Feng, *Algorithms for a new season? Mapping a decade of research on AI-driven digital transformation of public administration*, „Public Management Review" 2026, S. 620.

[16] F.C. von Savigny, *System of the Modern Roman Law*, vol. 1, Berlin (Veit) 1840, § 33, pp. 213 f.

[17] C. Möllers, The Significance of Language Models Such as ChatGPT Using the Example of Savigny's Canon of Interpretation, [in:] *Legal Methodology and Digitalisation*, Potsdam (Universitätsverlag Potsdam), eds. Susanne Hähnchen, 2024, pp. 55–93C.

[18] Cf. Resolution of the Conference of Minister-Presidents with the Federal Chancellor of 4 December 2025 https://bmds.bund.de/fileadmin/BMDS/Dokumente/2025-12-04-mpk-staatsmodernisierung-data_CDR.pdf [accessed on: 22.04.2026].

[19] Seventh Act Amending the Code of Administrative Court Procedure (7. VwGOÄndG) of 2 February 2026, https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/DE/2026_VwGOAendG.html [Zugriff: 22.04.2026].

lines for AI deployment in the federal administration (2025)[20] do not go beyond elementary prompting advice and provide no further analytical value.

3. The Normative Limit: Kelsen's Grundnorm and Hart's Rule of Recognition

Against the epistemological capacity of AI systems stands a normative limit that cannot be overcome by technical improvements, because it is not of a technical but of a jurisprudential nature. Hans Kelsen's *Pure Theory of Law*[21] develops the Grundnorm as a purely hypothetical presupposition — initially as a transcendental-logical prerequisite, later as a fiction — that must be assumed in order to ascribe coherence and validity to a legal order. H. L. A. Hart[22] develops the Rule of Recognition — traceable back to Kelsen — as the ultimate recognition rule: that rule which, as a matter of social practice, determines what standards are to be acknowledged as law of the community in question[23]. Both concepts share a fundamental characteristic: the basis of validity is not itself an object of legal application — it is its immediate presupposition.

An LLM can interpret norms, fill gaps, check coherence requirements and link arguments consistently. What remains categorially closed to it is of a different order: it cannot assume the Grundnorm, because this is a normative act presupposing the acknowledgement of an obligation. It cannot assent to the Rule of Recognition, because this is a form of constitutive legal-social practice in which only agents with capacity to act can participate. And it cannot assume responsibility for the binding force of a decision, since responsibility presupposes a normative status of the subject that an AI system does not possess — neither originally nor by attribution. An AI system can produce results equivalent to a human legal application, even identical in wording and argumentation. In normative terms, however, the result differs fundamentally despite external similarity: the difference lies not in the quality of the argument but in the inherent anchoring of the responsibility structure on which the argument rests — here, a human being; there, the probabilism of mathematical vectors.

---

[20] Federal Ministry of the Interior and Community, *Guidelines for the Use of Artificial Intelligence in the Federal Administration*, March 2025, pp. 27 f.

[21] H. Kelsen, *Pure Theory of Law*, 1st ed., Leipzig–Vienna (Deuticke) 1934; study edition, ed. M. Jestaedt, Tübingen (Mohr Siebeck) 2008, https://library.oapen.org/bitstream/id/40f462a5-662c-442f-a3fc-10fa272bb78e/external_content.pdf [accessed on: 22.04.2026].

[22] H.L.A.: Herbert Lionel Adolphus.

[23] H.L.A. Hart, *The Concept of Law*, 2nd ed., Oxford (Clarendon Press) 1994, ch. VI ("The Foundations of a Legal System" pp. 100 f.

4. Human in the Loop as a Jurisprudential Necessity

The requirement of a Human in the Loop (HitL)[24] is not a regulatory precondition and certainly not a safety net against 'hallucinations', but a jurisprudentially necessary structural condition of every normatively reliable AI-supported administrative decision. This finding is independent of Article 22 GDPR, which regulates the prohibition on fully automated decisions from a data protection perspective, and independent of Article 14 of the AI Regulation, which prescribes effective human oversight as a mandatory system requirement for high-risk AI systems[25]. Both provisions operationally regulate — on a regulatory level — a theoretically deeper-grounded requirement: normative binding force is structurally tied to human assumption of responsibility; or, put differently, the human assumption of professionally grounded responsibility is a *condition sine qua non* for the normative binding force of an administrative decision.

What is decisive, however, is that HitL is only effective where methodological competence and actual decision-making power coincide. A case officer who signs a KI-generated draft decision without substantive review formally satisfies Section 39(1) sentence 2 VwVfG but not materially: the statement of reasons does not contain the grounds that moved the human being, but the mathematical pattern of high probability activated by the LLM. The deceptive HitL facade — formally human final decision, materially algorithmic pre-determination — remains the central implementation risk of AI-assisted administrative action.

5. Methodological Control in the Europeanised Multi-Level System

The demands on methodological command intensify in the context of Europeanised administrative law[26]. National law of Member States is subordinate to Union law[27] and must accordingly be interpreted and applied in the light of the wording and objectives of Union law[28].Administrative law in EU Member States is no longer a purely national legal domain. From the obligation to implement directives under Article 288(3) TFEU and the principle of sincere cooperation (Article 4(3)

---

[24] *BeckOK AI Law*/B. Buchner, 5th ed., as of 01.02.2026, AI Act Art. 14 para ; S. Natarajan et al., *Human-in-the-loop or AI-in-the-loop? Automate or Collaborate?*, [in:] *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, Bd. 39, Nr. 27, S. 28594, https://doi.org/10.1609/aaai.v39i27.35083 [accessed on: 22.04.2026].

[25] *BeckOK Data Protection Law*/K. von Lewinski, 55th ed., as of 01.02.2026, GDPR Art. 22 para. 1.1.

[26] Stelkens/Bonk/Sachs/U. Stelkens, *Administrative Procedure Act* (VwVfG), 10th ed., Munich (C.H. Beck) 2022, European Administrative Law para. 36; I. Bantekas, E. Bratsiakou, *ibidem*.

[27] CJEU, Judgment of 15 July 1964, Case 6/64, *Costa v ENEL*, ECLI:EU:C:1964:66.

[28] Köhler/Feddersen/H. Köhler, *Act Against Unfair Competition* (UWG), 44th ed., Munich (C.H. Beck) 2026, Introduction para. 3.13.

TEU) follows the duty to interpret provisions of national law within the scope of a directive in conformity with that directive — in the light of its wording and purpose[29]. In energy and climate law, the international level additionally enters the picture, bringing its own methodology, inter alia under the Vienna Convention on the Law of Treaties (VCLT)[30] and a distinct doctrine of sources of law[31].

For AI systems in administration, this yields a system-architectural requirement: structurally, the LLM must be able to differentiate between diverse normative levels — at minimum between the constitutionally structured national law (for Germany: federal, state, local, mediate state administration) and the primary, secondary and tertiary law of the EU[32]. A model that cannot identify which normative framework applies at which level within a multi-level system cannot serve as a fit instrument for decision preparation but harbours the latent risk of level confusion. Section 39(1) sentence 2 VwVfG requires the statement of the essential legal grounds — in a multi-level system this includes the identification of the applicable level and its respective methodological framework.

The system-parameters argument is of particular doctrinal sharpness in this context: where an authority deploys an AI system on the basis of specific prompt structures, weighting rules and decision parameters that shape the selection of legally relevant facts and the structure of decision-making, these parameters are functionally a constitutive element of the decision and thus simultaneously a component of the grounds for decision. The principles of accuracy, clarity and completeness of the file (Section 24 VwVfG) presuppose this transparency[33]. What the administration cannot state as reasons it may not allow to pre-determine or even pre-shape algorithmically. System parameters that cannot be translated into a legally traceable argument are procedurally impermissible.

Only knowledge of the operative reasoning of an administrative act enables the person concerned to assess properly whether an appeal is advisable and has reasonable prospects of success[34]. If the person cannot identify the grounds, they cannot contest the decision. The obligation to give reasons is to that extent a concretisation of the constitutionally paramount guarantee of effective legal protection under

---

[29] CJEU, Judgment of 4 July 2006, Case C-212/04, *Adeneler*, ECR 2006, I-6057, para. 108; BGH, *NJW* 2009, p. 427.

[30] Vienna Convention on the Law of Treaties (VCLT) of 23 May 1969.

[31] Art. 38 ICJ Statute; M. Herdegen, *Public International Law*, 23rd ed., Munich (C.H. Beck) 2024, § 14 para.

[32] R. Streinz/P. Huber, *Treaty on European Union* (EUV), 3rd ed., Munich (C.H. Beck) 2018, Art. 12 para. 31.

[33] W. Huck, [in:] *Administrative Procedure Act* (VwVfG), eds. W. Huck, M. Müller, 4th ed., Baden-Baden (Nomos) 2025, § 24 paras. 54, 54b.

[34] BVerfGE 6, 32; BVerwG, *NVwZ* 1993, p. 677.

Article 19(4) Basic Law (*Grundgesetz,* GG), which assigns to the courts comprehensive review of the executive within the framework of the separation of powers[35].

III. Structural Deficits of the Federal Administrative Procedures Act

1. Structural Mismatch Problems of the VwVfG

The VwVfG was conceived as a statute for the uniform regulation of decision-making processes — the sum of provisions under which federal and state administrative authorities exercise their activities under public law[36]. It is the product of historical sedimentation of norms reaching back into the German and French administrative legal tradition[37]. Unlike, for example, the Baltic legal orders, which early and resolutely embraced digital administrative structures[38], the German legislature has consistently accommodated technological developments in a trailing fashion — most recently incorporating lessons from the pandemic era[39] — generating a path dependency that has structural, not merely technical, causes.

The only explicit provision on automation is Section 35a VwVfG. This permits the fully automated issuance of an administrative act where authorised by statute and where neither discretion nor a margin of assessment exists. The provision was conceived for deterministic if-then algorithms[40]. It does not distinguish between deterministic automation and AI-supported decision preparation, although these present categorially different risk profiles. Section 35a VwVfG was thus technologically outdated even upon its introduction and cannot answer the genuinely relevant question of how probabilistic AI can be legally reliably integrated into administrative decisions.

---

[35] BeckOK *Administrative Procedure Law*/P. Tiedemann, 70th ed., as of 01.01.2026, VwVfG § 39 para. 5.

[36] Federal Government Bill, Draft Administrative Procedure Act (VwVfG), BT-Drs. 7/910 of 18 July 1973, p. 28.

[37] F. Schoch/J.-P. Schneider/M. Knauff, *Administrative Procedure Act* (VwVfG), 7th supplementary instalment 2025, Munich (C.H. Beck), § 35 paras. 1–4.

[38] Deutsche Welle, *Estonia is Europe's Champion of Digitalisation*, 20 July 2025, https://www.dw.com/de/estland-digitalisierung-deutschland-analog-voll-digital-eurostack/a-73335359 [accessed on: 22.04.2026.

[39] Fifth Act Amending Provisions of Administrative Procedure Law (5. VwVfÄndG) of 4 December 2023, BGBl. I No. 344; S. Prell, N. Altmiks, *NVwZ* 2024, p. 105.

[40] C. Besio et al., Responsible Machines Without Accountability?, *Soziale Systeme* 2022, pp. 129–159.

2. Six Normative Gaps in Detail

a) Absence of a Conceptual Category

The VwVfG recognises the administrative act (Section 35 VwVfG)[41], the automated administrative procedure (Section 35a VwVfG)[42] and electronic communication (Section 3a VwVfG)[43]. It does not, however, know any category of algorithmically supported decision preparation through probabilistic systems. Learning systems, probabilistic forecasts and dynamic decision models find no statutory counterpart in the VwVfG.

b) Absence of a Normative Anchoring of HitL

HitL is at best derivable indirectly from the constitutional rule-of-law principle and Article 19(4) Basic Law, but is not normatively anchored in the VwVfG. There is no statutory duty of personal final decision-making, no clarification of the attribution of responsibility and no demarcation between delegable and non-delegable procedural steps. The Federal Constitutional Court has made clear that final decision-making powers that are not reviewable by courts require a statutory basis[44]. An AI system that factually becomes the final decision-making authority establishes a decision-making level that is not legitimated within the normative hierarchy and is thus in conflict with the judicial authority. The violation of the separation of powers immanent in the rule-of-law principle[45] and the violation of Article 19(4) Basic Law are thus apparent.

c) The Obligation to Give Reasons Is Structurally Unfit for AI

The obligation to give reasons under Section 39(1) sentence 2 VwVfG requires communication of the grounds that moved the authority to its decision. The verb *moved* (*bewogen*) denotes a personal act of cognition — not the reproduction of a system output[46]. System parameters of an AI that underlie an administrative decision must be incorporated into the statement of reasons in a normatively

---

[41] M. Müller, [in:] *Administrative Procedure Act* (VwVfG), eds. W. Huck, M. Müller, 4th ed., Baden-Baden (Nomos) 2025, § 35 para. 1.

[42] M. Müller, [in:] *Administrative Procedure Act* (VwVfG), eds. W. Huck, M. Müller, 4th ed., Baden-Baden (Nomos) 2025, § 35a para. 1.

[43] M. Müller, [in:] *Administrative Procedure Act* (VwVfG), eds. W. Huck, M. Müller, 4th ed., Baden-Baden (Nomos) 2025, § 3a para. 1.

[44] BVerfG, *NVwZ* 2011, p. 1062: "If the legislature wishes to restrict the judicial review of subjective rights it recognises, it must bear in mind that the final and binding interpretation of norms, as well as the review of their concrete application in individual cases, is in principle a matter for the courts. The legislature may not abandon the effectiveness of the judicial protection guaranteed by Art. 19(4) of the Basic Law by conferring on authorities final and binding decision-making powers over entire areas of law. The exemption of the application of law from judicial review always requires a sufficiently weighty substantive reason oriented towards the principle of effective legal protection".

[45] *BeckOK Basic Law*/Rux, 64th ed., as of 15.11.2025, Basic Law Art. 20 para. 174.

[46] M. Müller, [in:] *VwVfG*, Hrsg. W. Huck, M. Müller, 4. Aufl., Baden-Baden (Nomos) 2025, § 39 Rn. 4.

mediated form: not technically, but as a legal classification of algorithmic contributions, together with an account of how the AI's recommendation influenced the decision. Mere self-disclosure by the AI is insufficient: a prompt such as 'explain the criteria by which you decide' produces at best a descriptive system representation, not a normative justification.

d) Absence of Documentation Duties for AI-Relevant Parameters

Record-keeping and documentation are oriented towards human decision-making processes and text-based (electronic) materials. Which AI-relevant information — training data, model versions, system parameters, decision parameters, prompts, *etc.* — must be documented and how judicial review is structurally prepared is not regulated by statute. This threatens a de facto reduction in the intensity of judicial scrutiny, in conflict with Article 19(4) Basic Law[47].

e) Need to Reform Section 35a VwVfG: Structured Discretion

The blanket prohibition on automating discretionary decisions in Section 35a VwVfG is incompatible with the risk-based approach of the AI Regulation: whereas the AI Regulation differentiates by the risk potential of deployment, Section 35a VwVfG prohibits automation for discretion as a blanket rule — irrespective of the risk in any individual case. A sound reform would need to distinguish between structured discretion (standardised discretion in mass administrative proceedings with high case volumes) and genuine balancing discretion. For the former, AI-supported decision preparation could be permitted under strict conditions; the latter remains a personal core domain.

f) No Legislative Reform Process in Sight

A reform of the VwVfG is neither resolved nor announced. The Federal Government has been piloting autonomous administrative agents in 17 municipalities since March 2026 — in naturalisation proceedings under Sections 10 et seq. of the Nationality Act (*Staatsangehörigkeitsgesetz*, StAG), in housing entitlement proceedings, and in social administration[48]. The BMDS emphasises that its goal is to deploy AI in a lawful, efficient and scalable manner. For this purpose, it has established the AI Opportunities Marketplace (MaKI) as a transparency register and produced guidance on AI competence-building under the AI Regulation. These instruments are welcome; but they do not substitute for normative containment through procedural law. Scaling means: the same probabilistic models would be deployed in thousands of municipalities for equivalent administrative decisions — without normative standards in the VwVfG, without documentation duties, without secured reasoning structures. It appears as though administrative practice is about to outpace the legal framework.

---

47 BVerfG, *ibidem*.

48 BMDS, *Press Release: Agentic AI Hub – Piloting in Municipalities Commences*, 09 March 2026.

3. The Language-Competence/Methodology Paradox

a) The Apparent Strength: Enormous Language and Legal Material

Large language models are trained on text corpora of extraordinary scope, which contain a considerable proportion of legal texts: court judgments, statutes, commentaries, legal opinions and scholarly articles in numerous languages. This generates the widespread impression that an LLM possesses comprehensive legal knowledge. This assessment is not merely incomplete — its core component is structurally false. It rests on a confusion between *language competence* and *legal system competence*.

b) The Structural Paradox: Language Competence Is Not Legal System Competence

The language of a text — its lexical material, its syntactic structures, its argumentative patterns — is not a reliable indicator of the state embedding of the legal order from which that text derives. Language and legal order are categorically distinct attribution variables.

English is the official language of the United Nations[49] and of more than 58 states[50]. In those states fundamentally different legal orders apply, differing in system affiliation (common law, civil law, hybrid systems), historical genesis, doctrinal tradition, doctrine of precedent and interpretive canon. A judgment of the Supreme Court of the United States follows different methods from a judgment of the Supreme Court of Kenya, the Singapore Court of Appeal or the High Court of Malta — although all four are written in English. An LLM lacks any structural signal that could make these legal-systemic differences identifiable — because they do not manifest themselves in language as such, but in the state-anchored normative order, its institutional history and its doctrinal tradition formation.

The same structure applies to Spanish, which is the official language of more than twenty states. Spanish, Argentine, Chilean, Cuban and Mexican law — all on the same language, all with a Romance basic structure, but with fundamentally different constitutional histories, different roles of case law and divergent doctrinal leading decisions. The Chilean Civil Code by Andrés Bello (1855) has a different genesis, methodological tradition and subsequent interpretive practice from the Spanish Código Civil of 1889 — although both are written in the same language. The confusion between linguistic proximity and legal-systemic identity is struc-

[49] United Nations, *Official languages*, https://www.un.org/en/our-work/official-languages [accessed on: 22.04.2026].

[50] Wikipedia, *List of countries and territories where English is an official language*, https://en.wikipedia.org/wiki/List_of_countries_and_territories_where_English_is_an_official_language [accessed on: 22.04.2026].

turally particularly hazardous in the Spanish-speaking legal space, because the similarity of terminology conceals the depth of doctrinal divergence.

c) Doctrinal and Historical Sedimentation as Non-Linguistic Information

Every national legal order is the product of a historical sedimentation process in which influences of diverse provenance have been condensed into a doctrinally independent system. Polish administrative law is a hybrid of pre-socialist administrative legal traditions, socialist state law doctrine and the European integration process that began after 1989. The interpretation of Article 107 § 3 of the *Kodeks postępowania administracyjnego (*KPA — Polish Code of Administrative Procedure) cannot be handled methodologically correctly without knowledge of these strata. The language model has absorbed the linguistic material of this tradition — not the tradition itself, because tradition is an institutional and historical fact standing outside the medium of word and language. In the Europeanised administrative law context this applies with particular force to the interplay of climate law, sanctions law and national administrative law: the Paris Agreement, the European Climate Law and the Federal Climate Protection Act (*Bundes-Klimaschutzgesetz*, KSG)[51] follow their own validity logics and interpretive rules — international treaty law under Articles 31 et seq. VCLT, Union law under ECJ methodology with *effet utile*, national law under the Savigny canon overlain by the duty of conforming interpretation. An LLM trained on these texts has absorbed their linguistic patterns — not the normative hierarchy that governs among them.

d) The Consequence: Methodological Context Must Be Established Externally

LLMs possess enormous quantities of linguistic material from legal discourses, but this material has lost its legal-systemic context in the training process. The model has learned the language of law — not the methodology of the respective legal order.

The methodological context must be *established externally*. The system architecture must tell the model to which legal order a fact pattern belongs, which interpretive rules apply at that level, which normative hierarchy is relevant and which doctrinal leading decisions are to be treated as the framework. Without this external contextualisation — which cannot reside in the linguistic material of the training corpus, because it is a normative and institutional fact, not a linguistic one — the model is not capable of performing methodologically correct legal work.

For the deployment of AI in public administration the following applies: the administration cannot assume that a model trained on legal texts commands the methodology of German, Polish or Union administrative law. It must actively supply that methodology through the architecture of the system's control. The

[51] Federal Climate Protection Act (KSG) of 12 December 2019, BGBl. I p. 2513.

sequence is: first the methodological context, then the use of the model — not vice versa. This grounds the requirement of level separation that is operationally translated into administrative law terms in the following section.

4. Methodological Command in the Europeanised Multi-Level System

The demands on methodological command intensify considerably when one takes into account that administrative law in EU Member States is no longer a purely national domain. Numerous areas are overlaid by Union law: climate law (UN Agenda 2030,[52] Paris Agreement[53], European Climate Law[54],Emissions Trading Directive[55], Federal Climate Protection Act[56]), sanctions law and procurement law. These levels each follow their own validity logics and interpretive rules.

Union law has an interpretive methodology that is categorially different from the German one. The *effet utile* developed by the ECJ[57] and the duty of conforming interpretation overlay the Savigny canon, supplemented by Article 51(1) of the EU Charter of Fundamental Rights[58]. At the international level Article 31 et seq. VCLT applies[59].Each normative level carries its own methodological framework.

A methodologically uncontrolled system systematically confuses levels: it interprets international treaty law by ECJ methodology, treats IPCC reports[60] as legal norms or transfers the duty of conforming interpretation to primary law provisions. These confusions are formally difficult to detect in their outcome — and thereby particularly dangerous. The system architecture must structurally differentiate between normative levels, activate the applicable framework and indicate this level separation in the output.

---

[52] UN General Assembly, Resolution A/RES/70/1 – Transforming Our World: The 2030 Agenda for Sustainable Development, adopted on 25 September 2015; W. Huck, *Sustainable Development Goals: Article-by-Article Commentary*, Baden-Baden (Nomos) 2022.

[53] UNFCCC/CP/2015/L.9/Rev.1, Paris Agreement, 12 December 2015, https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf [accessed on: 22.04.2026].

[54] Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 establishing the framework for achieving climate neutrality (European Climate Law).

[55] Directive 2003/87/EC of the European Parliament and of the Council of 13 October 2003 establishing a system for greenhouse gas emission allowance trading (Emissions Trading Directive).

[56] Federal Climate Protection Act (KSG) of 12 December 2019, BGBl. I p. 2513.

[57] M. Potacs, Effet Utile as a Principle of Interpretation, *EuR – Europarecht* 2009, p. 465.

[58] C. Calliess/M. Ruffert/T. Kingreen, *EU Charter of Fundamental Rights*, 6th ed., Munich (C.H. Beck) 2022, Art. 51 paras. 4–6.

[59] T. Berner, Authentic Interpretation in Public International Law, *ZaöRV – Zeitschrift für ausländisches öffentliches Recht und Völkerrecht* 2016, p. 845.

[60] Intergovernmental Panel on Climate Change (IPCC), About, https://www.ipcc.ch/about/ [accessed on: 22.04.2026].

5. Procedural Differentiation: Where AI Is Applicable — and Where It Is Not

In bound mass proceedings with high norm density — charges, fees, straightforward benefit grants — AI-supported decision preparation is comparatively unproblematic. Structurally problematic, however, is AI deployment wherever discretionary decisions with fundamental rights relevance are required, where atypical constellations demand context-sensitive subsumption, or where the statement of reasons requires an evaluative overall assessment.[61] Section 40 VwVfG binds the exercise of discretion to the purpose of the empowering provision — in Europeanised administrative law regularly a purpose overlain by Union law, which a methodologically uncontrolled system cannot correctly grasp. The distinction is doctrinal in nature, not technical.

IV. European Law Overlay: The AI Regulation and the KI-MIG

1. The AI Regulation as Directly Applicable Union Law

Since 1 August 2024 the AI Regulation has applied as directly applicable Union law (cf. Article 288 TFEU). General applicability enters on 2 August 2026; the governance rules and obligations for general-purpose AI (GPAI) models have applied since 2 August 2025. For public administration, Article 6(2) in conjunction with Annex III of the AI Regulation is particularly relevant: law enforcement, migration, asylum, border control, justice and democracy, the grant of essential public services, and biometrics are classified as high-risk AI systems subject to the most stringent requirements. For high-risk AI systems Article 14 of the AI Regulation prescribes human oversight as a mandatory system requirement — the Union-law operationalisation of the jurisprudentially grounded HitL requirement, albeit without its concretisation in national procedural law.

On 19 November 2025 the European Commission proposed — as part of the Digital Omnibus — a shift in the deadlines for high-risk AI systems under Annex III of the AI Regulation by up to eighteen months to 2 December 2027[62]. The legislative procedure had not been completed as of March 2026; the currently applicable deadlines remain binding until an agreement is reached.

---

[61] N. Carstens, J. Thiele, Use of Artificial Intelligence in the Context of Public Participation in Plan Approval Procedures, *LKV – Landes- und Kommunalverwaltung* 2025, p. 385, first part, and LKV 2026, 13, second part.

[62] Proposal for a Regulation amending Regulations (EU) 2024/1689 and (EU) 2018/1139 (Digital Omnibus for AI), COM(2025) 836 final, https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52025PC0836 [accessed on: 22.04.2026].

2. The KI-MIG: Institutional Architecture of Market Supervision

The AI Market Surveillance and Innovation Act (*Künstliches-Intelligenz-Marktüberwachungs- und Innovationsförderungsgesetz*, KI-MIG), adopted by the Federal Government and submitted to the parliamentary procedure, is designed to establish the national institutional infrastructure for implementing the AI Regulation[63]. The Government Draft was adopted by the Cabinet on 11 February 2026[64]. The authority architecture follows the principle of institutional path-fidelity: existing sector-specific supervisory structures are loaded with competences under the AI Regulation rather than disrupted.

Under Section 2(1) KI-MIG, the Federal Network Agency (*Bundesnetzagentur,* BNetzA) is designated as the central market surveillance authority. For high-risk AI systems in particularly sensitive areas, an independent AI Market Surveillance Chamber is established under Section 2(5) in conjunction with Section 4 KI-MIG, which under Section 4(3) KI-MIG acts independently of instructions and reports annually to the Bundestag. In the financial sector, the Federal Financial Supervisory Authority (*Bundesanstalt für Finanzdienstleistungsaufsicht,* BaFin) becomes the competent market surveillance authority under Section 2(3) KI-MIG; the Federal Institute for Occupational Safety and Health (*Bundesanstalt für Arbeitsschutz und Arbeitsmedizin,* BAuA) assumes the function of the national focal point for notification to the European Commission under Section 7 KI-MIG. The Federal Office for Information Security (*Bundesamt für Sicherheit in der Informationstechnik*, BSI) communicates information on non-conformity with cybersecurity requirements under Section 10 KI-MIG. The Coordination and Competence Centre (Section 5 KI-MIG) is to secure coherent application of law across sectoral boundaries — without powers of instruction vis-à-vis the sectoral authorities. The fragmentation of competences across diverse subordinate authorities and their respective ministerial portfolios makes a degree of nuance in future application not unlikely — likely to produce divergences or coincidental convergences particularly in structurally analogous cases. The Federal Chancellery could then be the higher-ranking hierarchical level that smooths out different application practices through 'instructions'.

[63] Draft Act implementing Regulation (EU) 2024/1689 (Act on AI Market Supervision and Innovation Promotion – KI-MIG), Government Bill, https://bmds.bund.de/themen/kuenstliche-intelligenz/ki-regulierung [accessed on: 22.04.2026].

[64] T. Roth-Isigkeit, AI Supervision on the Home Straight, *KIR – Künstliche Intelligenz und Recht* 2025, p. 356.

3. The Regulatory Incongruence

The KI-MIG is an implementing statute for the AI Regulation — not a statute to modernise administrative procedural law. This gap is legally systematic in its significance. From 2 August 2026, the stringent requirements of the AI Regulation apply to virtually all high-risk deployment areas of public administration. The VwVfG remains unchanged. Public authorities as operators of AI systems within the meaning of Article 3(4) of the AI Regulation will be obliged under Union law to implement HitL measures — without national procedural law regulating their configuration. The norm collision between the Union-law operator duty and a VwVfG that remains legally unconcerned is real from August 2026 onwards and no longer merely hypothetical.

V. Digital Sovereignty as a Condition of Lawfulness

1. The ICC Precedent and Its Normative Structure

In February 2025, the US Government imposed sanctions on the International Criminal Court (ICC) in response to an arrest warrant against the Israeli Prime Minister. In immediate consequence, the Microsoft account of the Prosecutor Karim Khan was suspended; Microsoft informed the ICC that the service relationship would have to be terminated on account of the US sanctions. Whatever the disputed details of responsibility — Microsoft disputed having made an independent suspension decision — the rule-of-law finding stands: the connection between the normative mandate of an international institution and its operational infrastructure was interruptible by a private-law provider that itself is subject to a foreign statutory access regime[65].

This precedent is so far singular but reveals a structurally deficient state of affairs. In Germany, approximately 70% of public data are estimated to reside in US cloud environments. The Federal Government's Microsoft expenditure was 198 million euros in 2023 and 481.4 million euros in 2025[66], a rising trend[67] — notwithstand-

[65] K. Ambos, The Sanctioning of Law, *VerfBlog* 18 December 2025, https://verfassungsblog.de/sanctions-us-icc-united-states/ [accessed on: 22.04.2026]; J. Galbraith, U.S. Sanctions on the ICC, *VerfBlog* 07 June 2025; Opinio Juris, Potential Implications of US Sanctions on the ICC's Use of AI and Digital Evidence, 25 February 2025, https://opiniojuris.org/2025/02/25/artificial-sanctions-potential-implications-of-us-sanctions-on-the-iccs-use-of-ai-and-digital-evidence/ [accessed on: 22.04.2026].

[66] R. Lenhard, My Written Parliamentary Question Reveals: €481.4 Million for Microsoft Licences, https://rebecca-lenhard.de/news/meine-schriftliche-legt-offen-481-4-mio-euro-fur-microsoft-lizenzen/ [accessed on: 22.04.2026].

[67] Reply of the Federal Government to a Minor Interpellation, BT-Drs. 20/12864 of 6 September 2024.

ing the existence of the state-supported Centre for Digital Sovereignty of Public Administration (*Zentrum für Digitale Souveränität der Öffentlichen Verwaltung*, ZenDiS)[68] and openDesk[69] as alternatives.

2. The CLOUD Act as a Normative Problem of Lawfulness

The Clarifying Lawful Overseas Use of Data Act (CLOUD Act) of the United States of 23 March 2018 obliges US providers of electronic communication services under 18 U.S.C. § 2713 to disclose data in their possession, custody or control on governmental demand — irrespective of where the data are physically stored[70]. Microsoft confirmed before a French court in June 2025 that it could not guarantee digital sovereignty should US authorities demand access to data stored in Europe under the CLOUD Act[71].

The connection to the jurisprudential foundation in Section II is immediate: Kelsen's Grundnorm and Hart's Rule of Recognition identify the normative basis of validity of a legal order as that which is structurally non-delegable. If, however, the infrastructure on which AI-supported decision preparation takes place is subject to a foreign extraterritorial statutory access regime that can be activated at any time without judicial oversight by the affected state, then the operative foundation of the decision — the availability of the system, the integrity of the data, the reconstructibility of the audit trail — is not normatively secured. The decision may be formally lawful and its foundation nonetheless structurally unsovereign, remaining tainted by a shadow.

3. Consequences for Article 19(4) Basic Law

Effective legal protection under Article 19(4) sentence 1 Basic Law presupposes that the grounds for decision are reconstructible and fully accessible to state scrutiny. Gaps in the factual basis, in the decision system, represent a 'black box' in the normative chain of the obligation to give reasons resting on the facts and the application of law. An audit trail that resides on a platform accessible extraterritorially to US authorities under 18 U.S.C. § 2713 at any time is not a sovereign audit trail in the rule-of-law sense. Digital sovereignty remains a condition of lawfulness for

[68] ZenDiS – Centre for Digital Sovereignty of Public Administration, https://www.zendis.de/ [accessed on: 22.04.2026].

[69] openDesk – flexible office and collaboration suite for public administration, provided by ZenDiS on behalf of the Federal Ministry of the Interior and Community.

[70] J. Daskal, Congress Enacts the CLOUD Act, *American Journal of International Law* 2018, vol. 112, no. 3, p. 487, https://doi.org/10.1017/ajil.2018.61 [accessed on: 22.04.2026].

[71] Dr. Datenschutz, Microsoft Cannot Prevent US Access to EU Cloud, 24 July 2025, https://www.dr-datenschutz.de/microsoft-kann-us-zugriff-auf-eu-cloud-nicht-verhindern/ [accessed on: 22.04.2026].

AI-supported administrative decisions — one that follows from the foundational principles of the rule of law.

Methodological command without sovereign infrastructure falls short of the rule-of-law requirement for sovereign availability of the grounds for decisions as part of the exercise of public authority. Methodological command without sovereign infrastructure is constitutionally insufficient; sovereign infrastructure without methodological command is likewise insufficient. Both together are necessary — though not sufficient — conditions for the constitutional quality of AI-supported administrative decisions. This distinguishes the system sovereignty question categorially from the regulatory debate: it does not suffice to enact the right AI regulatory rules; it must also be ensured that the infrastructure on which those rules are implemented is normatively sovereign. Neutralisations or access restrictions of varying temporal and substantive scope generate gaps in a closed reasoning chain subject to judicial review. The administrative decision thereby becomes a truncated account of the facts and thus inadequate for the underlying discretionary or bound administrative decision.

4. The Germany Stack as an Incomplete Answer

The Germany Stack (*Deutschland-Stack*, D-Stack) of the BMDS is conceived as a national technology platform providing interoperable, reusable basic services. On 13 March 2026, the Federal Government and the Länder agreed on its joint implementation. The priority use of open-source offerings and European providers is stated as a goal; the use of ZenDiS and openDesk as alternatives to US infrastructure is institutionally envisaged. The D-Stack is thus the infrastructural response to the sovereignty problem — albeit one that is not binding. As long as the D-Stack is not mandatory and the administration continues to operate predominantly on US infrastructure, the lawfulness problem of the CLOUD Act persists.

VI. Eleven Theses

- **Thesis 1:** Epistemics and normativity must be categorially separated. An LLM can generate methodologically correct, logically consistent legal argumentation options. What remains structurally closed to it is the normative claim to validity: the capacity to assume the Grundnorm, to assent to the Rule of Recognition and to assume responsibility for the binding force of the decision;
- **Thesis 2:** Human-in-the-Loop (HitL) establishes the necessary connection between AI's epistemic performance and the normative claim to validity. It is both regulatorily mandated (Article 14 AI Regulation) and jurisprudentially necessary. Its effectiveness is given when the deciding official possesses

legal-methodological competence and the system additionally operates on sovereign infrastructure. The deceptive HitL facade — formally human final decision with materially algorithmic pre-determination — is the central implementation risk;

- **Thesis 3:** Language competence is not legal system competence. LLMs possess enormous quantities of linguistic material from legal discourses, but this material has lost its legal-systemic context in the training process. English is an official language in more than 58 states with different legal orders and their inherent methodologies; statistical proximity in the vector space between legal texts from different legal orders does not necessarily correlate with the doctrinal-methodological content of the respective legal concept. The methodological context must be established externally — as a normative system architecture with constraints functioning as guardrails;
- **Thesis 4:** The system parameters of an AI are simultaneously part of the grounds for decision and must be disclosed and justifiable in a normatively traceable manner. The administration cannot have its statement of reasons explained to it by the machine;
- **Thesis 5:** Administrative law is structurally Europeanised — and sometimes internationally influenced — and AI systems must map the specificities of the multi-level system. National administrative decisions are embedded in a multi-level system encompassing primary, secondary and tertiary Union law. Those who deploy AI in administration without mapping this level structure in system architecture deploy a methodologically unfit instrument. AI merely perpetuates the frozen methodology of its training data. The methodology architecture is an external steering performance by means of constraints not immanent to an LLM system;
- **Thesis 6:** Methodological control is a precondition of lawfulness. The obligation to give reasons under Section 39(1) sentence 2 VwVfG, the binding of discretion under Section 40 VwVfG and the guarantee of effective legal protection under Article 19(4) sentence 1 Basic Law require methodologically reconstructible decisions. AI without traceable methodological control is not a modernisation of administrative law — it promotes its hollowing out;
- **Thesis 7:** The system-parameters argument sets a categorial limit. What the administration cannot state as reasons normatively it may not allow to pre-determine algorithmically. System parameters that cannot be translated into a legally traceable argument are procedurally impermissible — irrespective of whether they produce technically correct results;
- **Thesis 8:** Digital sovereignty is a condition of lawfulness, not merely technology policy. An administrative authority that operates its AI system on US

infrastructure possesses no sovereign audit trail. Without a sovereign audit trail, effective legal protection under Article 19(4) sentence 1 Basic Law is structurally jeopardized;

- **Thesis 9:** The legislature is called upon: VwVfG reform as a matter of urgency. The reform must at minimum first establish a conceptual category for AI-supported decision preparation; second, anchor HitL as a statutory standard; third, extend Section 39 VwVfG to include disclosure duties for system parameters; fourth, introduce documentation duties for judicial reviewability; and fifth, replace Section 35a VwVfG with a differentiation between structured discretion and genuine balancing discretion;
- **Thesis 10:** The scientific task is set: jurisprudence, administrative law and system architecture must work together. What normative status do AI-generated arguments acquire in the administrative decision? Who bears responsibility when methodologically flawed AI outputs enter into legally binding administrative acts? What statutory foundations does an auditable, methodologically controlled AI architecture on sovereign infrastructure require?;
- **Thesis 11 (supplementary):** The regulatory incongruence between the AI Regulation (applicable from August 2026), the KI-MIG and the unreformed VwVfG is no longer hypothetical from August 2026. Public authorities as operators of high-risk AI systems are subject to Article 14 of the AI Regulation without national procedural law providing guidance on how to configure the required human oversight in the administrative procedure. This gap is a legislative deficiency to be remedied urgently.

VII. Summary

The current situation of German administrative law in its engagement with artificial intelligence is characterised by four structural problem areas. First, at the technological level, there is a marked disparity between the dynamism of actual AI deployment (including in the area of agentic AI systems) and the speed of legislative adjustment; in particular, a reform of the VwVfG that would be necessary for this purpose is currently not discernible. Second, at the normative level, a tension arises from the fact that the AI Regulation obliges public authorities from August 2026 to implement Human-in-the-Loop measures under Article 14, while the VwVfG leaves the specific applicability of AI in administrative procedure unregulated, producing a norm collision between the Union-law operator duty and national procedural law. Third, the Europeanised structure of administrative law itself generates heightened requirements on methodological control that an LLM can only meet through a system architecture with external methodological con-

straints — a structural demand that neither the KI-MIG nor the current VwVfG addresses. Fourth and finally, at the geopolitical level, there persists a sustained tension between the normative validity claim of German administrative law, the claim to judicial review under Article 19(4) Basic Law, and the factual embedding of the underlying AI infrastructure in foreign legal spheres — in particular through the applicability of the extraterritorially operative US CLOUD Act — raising fundamental questions of lawfulness that existing regulation does not resolve.

References

Literature

Ambos K., The Sanctioning of Law, *VerfBlog* 18 December 2025.

Bantekas I., Bratsiakou E., Automated Decision-Making Systems and Black Box Challenges Under European Union Administrative Law, *Fordham International Law Journal* 2026, vol. 49, no. 1, pp. 1 ff.

Baumer F., Large Language Models Under the Scrutiny of the RDG, *RDi – Recht Digital* 2026, pp. 8 ff.

Berner T., Authentic Interpretation in Public International Law, *ZaöRV – Zeitschrift für ausländisches öffentliches Recht und Völkerrecht* 2016, pp. 845 ff.

Besio C. et al., Responsible Machines Without Accountability?, *Soziale Systeme* 2022, pp. 129–159.

Brauner F., Murawski M., Bick M., The Development of a Competence Framework for Artificial Intelligence Professionals Using Probabilistic Topic Modelling, *Journal of Enterprise Information Management* 2025, pp. 197 ff.

Calliess C., Ruffert M., Kingreen T., *EU Charter of Fundamental Rights*, 6th ed., Munich (C.H. Beck) 2022.

Carstens N., Thiele J., Use of Artificial Intelligence in the Context of Public Participation in Plan Approval Procedures, *LKV – Landes- und Kommunalverwaltung* 2025, pp. 385 ff, first part, and LKV 2026, pp. 13, second part.

Chandra Y., Feng S., Algorithms for a New Season? Mapping a Decade of Research on AI-Driven Digital Transformation of Public Administration, *Public Management Review* 2026, pp. 620 ff.

Daskal J., Congress Enacts the CLOUD Act, *American Journal of International Law* 2018, vol. 112, no. 3, pp. 487 ff.

Dornis T.W., Generative AI Training and the TDM Fallacy, *GRUR – Gewerblicher Rechtsschutz und Urheberrecht* 2024, pp. 1681 ff.

Galbraith J., U.S. Sanctions on the ICC, *VerfBlog* 7 June 2025.

Guckelberger A., Germany's (Lag in) Administrative Digitalisation, *LTZ – LegalTech: Zeitschrift für die digitale Rechtsanwendung* 2023, pp. 167 ff.

Hart H.L.A., *The Concept of Law*, 2nd ed., Oxford (Clarendon Press) 1994.

Hebing A., Kastlunger E., Artificial Intelligence in Sales and Marketing, *ZVertriebsR – Zeitschrift für Vertriebsrecht* 2026, pp. 957 ff.

Herdegen M., *Public International Law*, 23rd ed., Munich (C.H. Beck) 2024.

Huck W., *Sustainable Development Goals: Article-by-Article Commentary*, Baden-Baden (Nomos) 2022.

Huck W., Müller M. (eds.), *Administrative Procedure Act (VwVfG) – Commentary*, 4th ed., Baden-Baden (Nomos) 2025.

Jin P. et al., MoH: Multi-Head Attention as Mixture-of-Head Attention, arXiv:2410.11842 (2024).

Kelsen H., *Pure Theory of Law*, 1st ed., Leipzig–Vienna (Deuticke) 1934; study edition, ed. M. Jestaedt, Tübingen (Mohr Siebeck) 2008.

Kirschke-Biller J., Füllsack A., [in:] *BeckOK AI Law*, 5th ed., Munich (C.H. Beck), as of 01.11.2025.

Köhler H., Feddersen J., Köhler M., *Act Against Unfair Competition* (UWG), 44th ed., Munich (C.H. Beck) 2026.
Lacmanovic D., Skare M., Artificial Intelligence Bias Auditing, *Review of Accounting and Finance* 2025, vol. 24, no. 3, pp. 375 ff.
Lemnitzer L., Würzner K.-M., The Word in Language Technology, [in:] *Handbook of Word and Vocabulary*, vol. 3, eds. U. Haaß-Zumkehr, W. Kallmeyer, G. Zifonun, Berlin–Boston (de Gruyter) 2015, pp. 297 ff.
Möllers C., The Significance of Language Models Such as ChatGPT Using the Example of Savigny's Canon of Interpretation, [in:] *Legal Methodology and Digitalisation*, Potsdam (Universitätsverlag Potsdam) 2024, pp. 55–93.
Müller A., Digitalisation of Paper in Public Administration, *RDi – Recht Digital* 2024, pp. 557 ff.
Natarajan S. et al., Human-in-the-loop or AI-in-the-loop? Automate or Collaborate?, [in:] *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, vol. 39, no. 27, pp. 28594 ff.
Nussbaum A., *Legal Facts Research: Its Significance for Scholarship and Teaching*, Tübingen (J.C.B. Mohr) 1914.
Potacs M., Effet Utile as a Principle of Interpretation, *EuR – Europarecht* 2009, pp. 465 ff.
Prell S., Altmiks N., The Fifth Act Amending Provisions of Administrative Procedure Law, *NVwZ – Neue Zeitschrift für Verwaltungsrecht* 2024, pp. 105 ff.
Rehbinder M., *The Founding of the Sociology of Law by Eugen Ehrlich*, Berlin (Duncker & Humblot) 1967.
Roth-Isigkeit T., AI Supervision on the Home Straight, *KIR – Künstliche Intelligenz und Recht* 2025, pp. 356 ff.
Rux J., [in:] *BeckOK Basic Law*, 64th ed., Munich (C.H. Beck), as of 15.11.2025.
Savigny F.C. von, *System of the Modern Roman Law*, vol. 1, Berlin (Veit) 1840.
Schoch F., Schneider J.-P., Knauff M., *Administrative Procedure Act* (VwVfG), Munich (C.H. Beck), 7th supplementary instalment 2025.
Stelkens P., Bonk H.J., Sachs M., Stelkens U., *Administrative Procedure Act* (VwVfG), 10th ed., Munich (C.H. Beck) 2022.
Streinz R., Huber P.M., *Treaty on European Union / Treaty on the Functioning of the European Union* (EUV/AEUV), 3rd ed., Munich (C.H. Beck) 2018.
Tiedemann P., [in:] *BeckOK Administrative Procedure Law*, 70th ed., Munich (C.H. Beck), as of 01.01.2026.
van Bekkum M., Using Sensitive Data to De-bias AI Systems, *Computer Law & Security Review* 2025, vol. 56.
Vaswani A. et al., Attention Is All You Need, [in:] *Proceedings of the 31st Conference on Neural Information Processing Systems* (NeurIPS), 2017.
von Lewinski K., [in:] *BeckOK Data Protection Law*, 55th ed., Munich (C.H. Beck), as of 01.02.2026.
Yang Z. et al., Plug in the Safety Chip: Enforcing Constraints for LLM-Driven Robot Agents, [in:] *Proceedings of the 2024 IEEE International Conference on Robotics and Automation* (ICRA), 2024.
Zhang Y. et al., Large Language Models: A Survey, arXiv:2402.06196v2 (2024).

Internet Sources

Federal Ministry of the Interior and Community (BMI), *Guidelines for the Use of Artificial Intelligence in the Federal Administration*, March 2025. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/moderne-verwaltung/ki/BMI25020-leitlinien-ki-bundesverwaltung.pdf?__blob=publicationFile&v=4 [22.04.2026]
Federal Ministry for Digital Affairs and State Modernisation (BMDS), Press Release: Agentic AI Hub – Piloting in Municipalities Commences, 9 March 2026. https://bmds.bund.de/themen/kuenstliche-intelligenz/ki-in-der-verwaltung/agentic-ai-hub [accessed on: 22.04.2026]

Federal Ministry for Digital Affairs and State Modernisation (BMDS), AI Regulation: Government Bill KI-MIG, https://bmds.bund.de/themen/kuenstliche-intelligenz/ki-regulierung [accessed on: 22.04.2026].

Deutsche Welle, Estonia is Europe's Champion of Digitalisation, 20 July 2025, https://www.dw.com/de/estland-digitalisierung-deutschland-analog-voll-digital-eurostack/a-73335359 [accessed on: 22.04.2026].

Dr. Datenschutz, Microsoft Cannot Prevent US Access to EU Cloud, 24 July 2025, https://www.dr-datenschutz.de/microsoft-kann-us-zugriff-auf-eu-cloud-nicht-verhindern/ [accessed on: 22.04.2026].

European Commission, Proposal for a Regulation amending Regulations (EU) 2024/1689 and (EU) 2018/1139 (Digital Omnibus for AI), COM(2025) 836 final, https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52025PC0836 [accessed on: 22.04.2026].

Intergovernmental Panel on Climate Change (IPCC), About, https://www.ipcc.ch/about/ [accessed on: 22.04.2026].

AI Marketplace (MaKI), https://www.kimarktplatz.bund.de [accessed on: 22.04.2026].

Lenhard R., My Written Parliamentary Question Reveals: €481.4 Million for Microsoft Licences, https://rebecca-lenhard.de/news/meine-schriftliche-legt-offen-481-4-mio-euro-fur-microsoft-lizenzen/ [accessed on: 22.04.2026].

Conference of Minister-Presidents with the Federal Chancellor, Resolution of 4 December 2025, https://bmds.bund.de/fileadmin/BMDS/Dokumente/2025-12-04-mpk-staatsmodernisierung-data_CDR.pdf [accessed on: 22.04.2026].

Opinio Juris, Potential Implications of US Sanctions on the ICC's Use of AI and Digital Evidence, 25 February 2025, https://opiniojuris.org/2025/02/25/artificial-sanctions-potential-implications-of-us-sanctions-on-the-iccs-use-of-ai-and-digital-evidence/ [accessed on: 22.04.2026].

United Nations, Official Languages, https://www.un.org/en/our-work/official-languages [accessed on: 22.04.2026].

Wikipedia, List of Countries and Territories Where English is an Official Language, https://en.wikipedia.org/wiki/List_of_countries_and_territories_where_English_is_an_official_language [accessed on: 22.04.2026].

ZenDiS – Centre for Digital Sovereignty of Public Administration, https://www.zendis.de/ [accessed on: 22.04.2026].

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Agnieszka Jankowska
Fundacja Digital Poland
Ministerstwo Cyfryzacji
ORCID: 0009-0001-8423-5426
agnieszka.jankowska1@op.pl

Cybersecurity Standards in Public Health. Regulatory and Systemic Challenges in the Era of Data-driven Healthcare

Keywords: cybersecurity, public health, health data, digital regulation

Summary. The article examines cybersecurity in public health in the context of the dynamic development of digital technologies and the growing role of health data as a strategic resource. The starting point is the assumption that despite numerous legal regulations, technical standards and guidelines at the national and EU levels, the healthcare system operates in conditions of significant regulatory and organizational fragmentation. The varying levels of digital maturity among healthcare entities, limited financial and human resources and insufficient implementation coordination translate into a fragmented and inconsistent approach to cyber risk management. The article discusses the importance of health data for the innovation and competitiveness of the European Union and identifies key cyber threats in the healthcare sector. The conclusion presents directions for systemic change, pointing to the need for a more integrated regulatory model combining legal, technical and organizational standards with investments in competencies, infrastructure and building institutional resilience.

Standardy cyberbezpieczeństwa w zdrowiu publicznym. Wyzwania regulacyjne i systemowe w erze opieki zdrowotnej opartej na danych

Słowa kluczowe: cyberbezpieczeństwo, zdrowie publiczne, dane zdrowotne, regulacje cyfrowe

Streszczenie. Artykuł analizuje problematykę cyberbezpieczeństwa w zdrowiu publicznym w kontekście dynamicznego rozwoju technologii cyfrowych oraz rosnącej roli danych zdrowotnych jako strategicznego zasobu. Punktem wyjścia jest założenie, że mimo licznych regulacji prawnych, norm technicznych i wytycznych na poziomie krajowym i unijnym, system ochrony zdrowia funkcjonuje w warunkach znacznego rozproszenia regulacyjnego i organizacyjnego. Zróżnicowany poziom dojrzałości cyfrowej podmiotów leczniczych, ograniczone zasoby finansowe i kadrowe oraz niedostateczna koordynacja wdrożeń przekładają się na fragmentaryczne i niespójne podejście do zarządzania ryzykiem cybernetycznym. Artykuł omawia znaczenie danych zdrowotnych dla innowacji i konkurencyjności Unii Europejskiej, a także identyfikuje kluczowe zagrożenia cybernetyczne w sektorze ochrony zdrowia. W konkluzji przedstawiono pożądane kierunki zmian systemowych, wskazując na potrzebę bardziej zintegrowanego modelu regulacyjnego, łączącego standardy prawne, techniczne i organizacyjne z inwestycjami w kompetencje, infrastrukturę i budowę odporności instytucjonalnej.

1. Introduction: The Data Era

Technological progress, accompanied by the pervasive digitalization of social, economic and administrative processes, has resulted in the exponential growth of data generated, processed and stored in the digital form. The contemporary data ecosystem is characterized not only by quantitative expansion, but by qualitative transformation as well: data has become a strategic resource underpinning innovation, automation and decision-making across both public and private sectors. The global volume of data is increasing at an unprecedented pace, driven primarily by the rapid deployment of artificial intelligence, the expansion of the Internet of Things (IoT), widespread adoption of cloud computing and the ongoing digital transformation of industries and public services.

According to the latest estimates from Statista.com reports, the global volume of data reached 149 zettabytes in 2024 and is projected to grow to 181 zettabytes by the end of 2025[1]. On a daily basis hundreds of millions of terabytes of data are generated worldwide (throughout 2024 approximately 402.74 million terabytes of data were generated every day) and recent analyses suggest that the volume of stored data doubles roughly every four years[2]. This dynamic reflects the deepening digitization of global activities, encompassing consumer services, industrial production, healthcare, finance, public administration and critical infrastructure. At the same time, it exposes structural limitations in existing data storage architectures, cybersecurity mechanisms and governance frameworks.

Artificial intelligence plays a central role in this process. Machine learning models, large language models and generative AI systems generate and process vast datasets during both training and deployment phases. Applications such as automated content generation, recommendation systems, predictive analytics and decision-support tools rely on continuous data inflows and feedback loops to optimize performance. This constant circulation of data not only amplifies data volumes but also increases legal and technical risks related to data integrity, confidentiality and accountability, particularly where AI systems operate on sensitive or high-risk datasets. A significant proportion of global data production originates from digital platforms, especially social media ecosystems. Platforms facilitating user-generated content, such as video-sharing, image-based and short-form content services, handle billions of uploads and interactions daily. Each user action contributes to expanding pools of structured and unstructured data, while emerging functionalities

[1] https://www.statista.com/statistics/871513/worldwide-data-created/ [accessed on: 01.02.2026].

[2] International Data Corporation, https://www.businesswire.com/news/home/20200513005075/en/IDCs-Global-StorageSphere-Forecast-Shows-Continued-Strong-Growth-in-the-Worlds-Installed-Base-of-Storage-Capacity [accessed on: 02.02.2026].

such as live streaming, immersive media and augmented reality further intensify data generation. Equally important is the role of the Internet of Things, which has transformed everyday objects and industrial equipment into continuous data producers. Smart home systems, wearable health devices, industrial sensors and autonomous or semi-autonomous machines generate real-time data streams that are often processed in distributed and interconnected environments. Additionally, e-commerce and digital service platforms further contribute to the data surge by collecting granular transactional and behavioral information. Searches, clicks, purchases and user reviews generate datasets that feed algorithmic systems for personalization, dynamic pricing and targeted advertising.

The healthcare sector has become one of the most data-intensive domains of the digital economy, encompassing both private and public healthcare systems. The rapid expansion of electronic health records, medical imaging, wearable devices, and real-time patient monitoring has transformed health data into a cornerstone of contemporary healthcare governance and innovation. Market analyses indicate that the global big data in healthcare sector, valued at over USD 27 billion in 2024, is expected to grow dynamically over the coming decade, reflecting the accelerating integration of data-driven technologies into clinical practice, public health management and biomedical research. The Global Big Data in healthcare market is projected to grow by over 18% annually, reaching even USD 75 billion by 2030[3]. Today, the healthcare industry generates nearly one-third of the world's data volumes. According to the World Health Organization, global health data generation reached 2.314 exabytes in 2023, with projections indicating a threefold increase by 2030, which urgently requires solid infrastructure to manage and leverage this expanding digital footprint[4].

The societal potential of health data is considerable. Large-scale health datasets enable earlier and more accurate diagnoses, evidence-based clinical decision-making, improved disease surveillance and more efficient allocation of public healthcare resources. At the same time, the scale and sensitivity of healthcare data bring significant limitations and risks. Health data is inherently personal and often highly sensitive, making it a prime target for cyberattacks, including ransomware incidents, data breaches and integrity manipulation. The increasing reliance on interconnected digital infrastructures, cloud services and AI-driven analytics expands the attack surface and amplifies the potential impact of cybersecurity failures. In public healthcare systems, such incidents may disrupt essential services, undermine

[3] https://www.techsciresearch.com/report/big-data-in-healthcare-market/4009.html [accessed on: 02.02.2026].

[4] J. Thomason, "Data, digital worlds, and the avatarization of health care", *Global Health Journal*, vol. 8, no. 1 (March 2024), p. 2.

public trust and directly endanger patient safety. From a legal and regulatory perspective, the rapid growth of healthcare data exposes structural tensions between innovation, data protection and cybersecurity obligations. Existing legal frameworks often struggle to keep pace with data-intensive technologies, particularly in areas such as secondary use of health data, cross-border data flows, accountability for AI-driven decision-making and cybersecurity standards for critical health infrastructure. Consequently, while the expansion of healthcare data represents a significant opportunity for societal benefit, it simultaneously necessitates robust, sector-specific data governance and cybersecurity regimes capable of addressing the growing risks inherent in data-driven healthcare ecosystems. In this context, cybersecurity becomes no longer merely a technical issue, but of fundamental normative and systemic significance.

The purpose of this article is to analyze cybersecurity standards in public health amid dynamic digitalization and the growing significance of health data. Despite the existence of numerous legal regulations, technical standards and guidelines at both national and European Union levels, the healthcare system operates within a context of substantial regulatory and organizational fragmentation. This fragmentation stems from the multiplicity of healthcare entities, their varying levels of digital maturity, limited financial and human resources, and the absence of coherent coordination in cyber risk management. Concurrently, the rapid expansion of health data processing, driven by advancements in artificial intelligence, personalized medicine, and European data ecosystems, presents substantial opportunities to enhance the quality of healthcare and the European Union's competitiveness. However, this development intensifies the tension between innovative potential and the system's actual capacity to ensure security. The article argues that the current regulatory model, characterized by its inherent complexity, fails to provide an adequate response to the scale and dynamism of cyber threats, which is a challenge that is exacerbated for healthcare entities. This justifies the need to develop a comprehensive approach that integrates legal frameworks, institutional mechanisms, investments in infrastructure and human capital and cohesive implementation standards, thereby enabling the secure utilization of health data as a strategic resource.

2. Health Data as a Strategic Resource: The European Health Data Space

Data constitute the foundation of the modern internet economy, but companies operating in more traditional industries and sectors are also increasingly relying on data to develop their business activities. Today, there is virtually no area of the economy that has not been affected by digital transformation. The scale of data

generation on a daily basis, particularly in the era of rapidly developing artificial intelligence, creates new potential for economies, societies and the global system as a whole. Data-driven innovation can deliver substantial benefits across multiple sectors, including healthcare through improved personalized medicine, transportation through new mobility solutions and environmental protection through contributions to energy transformation. On the one hand, the growing economic and societal potential of data, and on the other, strong competitive pressure from global players based in the United States and China, have motivated the European Commission to develop a comprehensive EU data strategy. According to the European Commission, improved access to and reuse of data could generate social and economic benefits amounting to between 1% and 2.5% of EU GDP. New common rules and data-sharing mechanisms are expected to have a broad impact on the EU economy and society as a whole. Annual savings in the EU healthcare sector alone are estimated at EUR 120 billion, while EU-27 GDP gains by 2028 are projected to reach EUR 270 billion. Furthermore, whereas in 2018 the value of the EU data economy was estimated at EUR 301 billion (2.4% of EU GDP), by 2025 it was assessed to grow to nearly EUR 830 billion, representing approximately 5.8% of EU GDP[5].

Europe seeks to capture the benefits of improved data use through a comprehensive approach to the data economy, aimed at increasing both the use of data and the demand for data-enabled products, services and processes throughout the Single Market[6]. The European Data Strategy is designed to create conditions for enhanced data utilization across the economy and society, while maintaining appropriate control and security mechanisms, particularly with respect to sensitive data and high standards of cybersecurity, privacy and data protection. Two legislative acts form the core pillars of this strategy: the Data Governance Act, which aims to increase trust in voluntary data-sharing mechanisms, and the Data Act, which establishes rules governing who can access and use which data and for what purposes, across the EU. While the former focuses on facilitating data availability and reuse by regulating data intermediaries and governance structures, the latter seeks to ensure fairness in the distribution of value generated from data. Key element of the EU data strategy is the creation of Common European Data Spaces across a number of strategic economic sectors and domains of public interest. These data spaces are intended to make more data available for access and reuse within a trustworthy and secure environment, benefiting European businesses and citizens

[5] https://digital-strategy.ec.europa.eu/en/factpages/data [accessed on: 03.02.2026].

[6] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European strategy for data, COM/2020/66, Brussels, 19.02.2020.

across fourteen sectors. The first and most significant of these is the healthcare sector. Health data offer enormous potential not only for business development, but above all for advancing scientific research and developing innovative methods of disease treatment. Consequently, at the EU level, steps have been taken to establish a framework for a European Health Data Space[7], intended to serve as a critical bridge between health data ecosystems and the development of advanced technologies, including artificial intelligence. This framework is designed to enable data laboratories and AI factories to use anonymized and synthetic datasets within trusted processing environments[8].

The European Health Data Space (EHDS) constitutes the first common EU data space dedicated to a specific sector under the European Data Strategy. In the coming years, the EHDS is expected to significantly affect, inter alia, patients' access to and control over their medical records, the work of hospitals and healthcare professionals, the development of research and innovation based on advanced technologies and the overall functioning of healthcare systems in the Member States, including Poland[9]. It aims to establish a unified framework for the use and exchange of electronic health data across the EU. At the same time, it enables certain categories of data to be reused for purposes of public interest, policymaking and scientific research. This Regulation also introduces a harmonized legal and technical framework for electronic health record systems, fostering interoperability, innovation and the smooth functioning of the internal market. The EHDS is built on three core pillars: (1) empowering individuals to take control of their health data, ensuring immediate and easy electronic access to their personal health information; (2) supporting the use of health data to improve healthcare delivery across the EU, by strengthening data exchange between Member States, avoiding duplication and enabling healthcare providers to access complete patient medical histories, thereby improving treatment outcomes while saving time and resources; and (3) enabling the EU to fully harness the potential of secure data exchange, as well as the primary and secondary use of health data.

As part of these efforts, the EU expects to achieve savings of EUR 5.5 billion over the next decade through improved access to and exchange of health data within the healthcare sector. An additional EUR 5.4 billion in savings may result from

[7] Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, https://eur-lex.europa.eu/eli/reg/2025/327/oj/eng [accessed on: 07.02.2026].

[8] European Commission, *Communication from the Commission to the European Parliament and the Council: European Data Union Strategy – Unlocking Data for Artificial Intelligence*, Brussels, 19 November 2025, COM(2025) 835 final, pp. 7–8.

[9] A. Ryś„Europejska Przestrzeń Danych Dotyczących Zdrowia – nowe wyzwania i możliwości". *Med. Prakt.*, 2025; no. 3, pp. 110–119.

better use of health data in research, innovation and policymaking, contributing to an estimated 20–30% increase in market growth[10]. The EHDS establishes a robust legal framework for the use of health data for scientific research, innovation, public health purposes, and regulatory policymaking. Under clearly defined conditions, researchers, innovators, public institutions and industry actors will gain access to large volumes of high-quality health data. This access is crucial for the development of life-saving treatments, vaccines, and medical devices, as well as for improving access to healthcare and strengthening the resilience of health systems. Responsible and sustainable governance of health data will support innovation and progress in healthcare, although ensuring data security and compliance with personal data protection principles still remains a major challenge. Achieving the objectives of the EHDS will therefore require cooperation, commitment, and adequate infrastructure. Successful implementation of the EHDS depends above all on consistent and uniform application of this Regulation across all Member States, as well as on ensuring interoperability between information systems to enable scale in scientific and commercial activities. A major challenge lies in the significant disparities between Member States, not only in terms of digital infrastructure development, but also in their readiness to adopt innovation. Countries such as Finland, Denmark, and France already offer examples of mature, system-level digital health solutions[11]. Nevertheless, a number of issues still need to be addressed to ensure effective implementation of the EHDS, including the establishment of clear rules fostering a data-sharing culture, allocation of sufficient resources for upskilling healthcare personnel, investment in technical infrastructure, development of procedures and standards ensuring high data quality and interoperability across healthcare providers. Health systems' capacity to contribute high-quality data to the EHDS, to close the loop between primary and secondary data use and to reintegrate research outputs and innovation back into healthcare practice will ultimately determine whether all EU citizens can benefit equally from the system. Cultural differences, varying perceptions of health data sharing and heterogeneous levels of health and digital literacy across and within Member States must not be underestimated, as they necessitate tailored and intensified efforts in education, communication and stakeholder engagement, strengthening public trust in the use of medical data in both clinical practice and research [12].

[10] The impact of the European Health Data Space Regulation and the Open Data Directive on citizens, European Commission, data.europa.eu, Publications Office of the European Union, July 2025.

[11] A. Ryś, *op. cit.*, p. 117.

[12] Implementing the European Health Data Space across Europe, eitHealth, Think Tank, April 2024, https://eithealth.eu/wp-content/uploads/2024/04/EIT_Health_ThinkTank_Implementing_the_ EHDS_across_Europe_23.04.24.pdf [accessed on: 07.02.2026].

3. Regulatory and Standardization Framework for Cybersecurity in Public Health

Cybersecurity in public health in Poland is currently shaped by a multilayered and fragmented regulatory framework, encompassing provisions of medical law, personal data protection law, public administration digitalization, as well as technical and organizational norms and standards. The obligations imposed on healthcare providers and public institutions stem from multiple normative sources, often pursuing different objectives and applying to different scopes. Below, the most significant regulations and standards relevant to cybersecurity in public health are outlined.

3.1. Medical Law as the Foundation for the Protection of Health Information

The primary regulatory domain determining the manner in which information is processed and protected in the healthcare sector is medical law. The Act of 6 November 2008 on Patients' Rights and the Patient Ombudsman establishes the fundamental principle of confidentiality of patient-related information[13]. Article 13 guarantees patients the right to have all information concerning them, including information about their health status, diagnosis and prognosis, kept confidential by persons performing medical professions. Although the Act does not explicitly refer to cybersecurity, it imposes an obligation on healthcare providers to ensure the confidentiality, integrity and availability of patient health information, which, in the context of digital healthcare systems, directly translates into requirements concerning the security of ICT systems.

This regulation is supplemented by the Regulation of the Minister of Health of 6 April 2020 on medical documentation, which specifies detailed rules for the creation, storage, and processing of medical records, including records maintained in electronic form[14]. The Regulation introduces requirements aimed at protecting medical documentation against damage, loss or unauthorized access; however, similarly to the Patients' Rights Act, it leaves considerable discretion as to the choice of specific technical and organizational measures. In particular paragraphs 86–89 of the Regulation impose an obligation to ensure the integrity, reliability and continuous availability of medical documentation, including through access control mechanisms and safeguards against unauthorized modification of data. At the same time, the Regulation does not define minimum cybersecurity standards

[13] Act of 6 November 2008 on Patients' Rights and the Patients' Rights Ombudsman, Journal of Laws 2024, item 581.

[14] Regulation of the Minister of Health of 6 April 2020 on the types, scope and models of medical documentation and the manner of its processing, Journal of Laws 2024, item 798.

nor does it explicitly refer to recognized technical norms, which in practice leads to significant disparities in the level of protection of electronic medical records across healthcare entities.

Of key importance for the digital infrastructure of public healthcare is also the Act of 28 April 2011 on the Healthcare Information System[15]. This Act establishes the framework for the operation of central ICT systems, such as the P1 platform, and regulates data exchange between participants in the healthcare system. While it emphasizes the need to ensure the security of data processed within healthcare information systems, it does not, like other sectoral acts, introduce a uniform and detailed catalogue of minimum cybersecurity requirements applicable to all system participants.

3.2. Personal Data Protection as a Pillar of Health Information Security

A second fundamental regulatory pillar in the area of public health cybersecurity is personal data protection law. Under the General Data Protection Regulation (GDPR)[16], data concerning health are classified as a special category of personal data, implying a heightened standard of protection. Article 32 of GDPR explicitly obliges controllers and processors to implement "appropriate technical and organizational measures" to ensure the security of processing, including system resilience, the ability to restore data availability promptly and regular testing of safeguards. At the national level, the GDPR is complemented by the Act of 10 May 2018 on the protection of personal data[17], which regulates, inter alia, the competences of the President of the Personal Data Protection Office (UODO) as well as domestic supervisory and sanctioning mechanisms. A significant practical development in the application of the GDPR in the healthcare sector is the "Code of Conduct for the Healthcare Sector", approved by the President of the UODO on 11 December 2023[18]. As the second code issued under Article 40 GDPR for this sector and the first applicable to both public and private entities, it constitutes an attempt to specify the general GDPR obligations in the context of healthcare activities.

---

[15] Act of 28 April 2011 on the Information System in Healthcare, Journal of Laws 2023, item 2465, as amended.

[16] 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), OJ L 119, 4 May 2016, p. 1, as amended.

[17] Act of 10 May 2018 on the Protection of Personal Data, Journal of Laws 2019, item 1781, as amended.

[18] Code of Conduct for the Health Care Sector adopted pursuant to Article 40 of Regulation (EU) 2016/679, applicable to healthcare providers and data processors, 11 December 2023, https://uodo.gov.pl/pl/file/4525 [accessed on: 06.02.2026].

Although this Code does not constitute universally binding law, its normative and interpretative relevance for health data security standards is substantial, particularly in the context of data controller responsibility. Application of the Code may serve as evidence of compliance with GDPR obligations by controllers and processors (implementation of the accountability principle). The Code provides a set of guidelines aimed at enhancing the level of personal data protection, consistent with GDPR and national legislation, including principles of data processing under Article 5 of GDPR, data pseudonymization, controller obligations, data protection by design and by default, as well as measures ensuring the security of processing.

3.3. State Digitalization and Interoperability of Healthcare Systems

Another regulatory area affecting cybersecurity in public health concerns the digitalization of public administration. The Act of 17 February 2005 on the computerization of entities performing public tasks[19] establishes a general framework for ICT systems used by public entities, including healthcare institutions performing public tasks. The Act introduces obligations to ensure data security and system compliance with specified standards. Particularly important in this context is the Regulation of the Council of Ministers of 21 May 2024 on the National Interoperability Framework (KRI)[20], which defines minimum requirements for ICT systems and electronic information exchange. Although KRI is not dedicated exclusively to the healthcare sector, in practice it establishes a baseline level of security and interoperability for systems used in public healthcare. At the same time, the Regulation remains general in nature and does not fully reflect the specific risks associated with the processing of medical data.

3.4. Cybersecurity as a Horizontal Regulatory Domain

The regulation most directly addressing cybersecurity issues is the Act of 5 July 2018 on the National Cybersecurity System (KSC)[21], which implements the NIS Directive into Polish law. This Act applies, inter alia, to operators of essential services, a category that in certain cases includes healthcare entities. KSC introduces obligations related to risk management, incident reporting and cooperation

[19] Act of 17 February 2005 on Computerisation of Activities of Entities Performing Public Tasks, Journal of Laws 2024, item 307.

[20] Regulation of the Council of Ministers of 21 May 2024 on the National Interoperability Framework, minimum requirements for public registers and electronic information exchange, and minimum requirements for ICT systems, Journal of Laws 2024, item 773.

[21] Act of 5 July 2018 on the National Cybersecurity System, Journal of Laws 2023, item 913, as amended. At the end of January 2026, the Sejm of the Republic of Poland adopted an amendment to this Act and submitted it for signature by the President of the Republic of Poland; however, at the time of writing, the Act had not yet been signed.

with competent cybersecurity authorities. The importance of this framework will increase with the implementation of the NIS 2 Directive[22], which should have been transposed into Polish law by 18 October 2024. NIS 2 expands the scope of regulated entities and significantly strengthens cybersecurity management requirements, explicitly covering many healthcare entities as essential or important entities. From a public health perspective, this entails a substantial reinforcement of legal obligations related to cyber resilience, while simultaneously raising questions regarding regulatory coherence and the practical feasibility of implementation in the healthcare sector.

3.5. Non-binding Standards and Guidelines in the Healthcare Sector

Legal regulations are complemented by technical and organizational standards and guidelines. Of particular relevance are ISO standards, especially ISO 27799, which constitutes a sector-specific extension of ISO/IEC 27001 and addresses information security management in healthcare. This standard reflects the specificity of medical data and clinical processes, offering practical guidance on protecting the confidentiality, integrity, and availability of health information.

At the national level, an important role is played by guidelines issued by the Centre for eHealth (CeZ), particularly those relating to electronic medical record systems (EDM). Although formally non-binding, these guidelines effectively shape technical standards applied by healthcare providers using the national eHealth infrastructure. At the same time, the absence of a uniform, mandatory catalogue of minimum cybersecurity requirements for the entire healthcare sector leads to significant disparities in security levels across entities. At the EU level, an important reference point is provided by recommendations and analyses issued by the European Union Agency for Cybersecurity (ENISA), including documents on cyber hygiene in the health sector[23]. ENISA identifies healthcare as one of the sectors most exposed to cyber threats and highlights good practices related to risk management, staff training and incident response.

An analysis of the applicable regulations and standards demonstrates that cybersecurity in public health is shaped in a fragmented and poorly coordinated manner. Although numerous legal acts, guidelines and technical standards apply

[22] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148 (NIS 2 Directive).

[23] ENISA – European Union Agency for Cybersecurity, Cyber Hygiene in the Health Sector, September 2025, https://www.enisa.europa.eu/sites/default/files/2025-09/ENISA%20Cyber%20Hygiene% 20in%20the%20Health%20Sector.pdf [accessed on: 06.02.2026].

to the sector, they do not form a coherent and unequivocal system of minimum cybersecurity requirements. The absence of centrally defined, mandatory standards and imposed technical solutions leaves healthcare entities with significant discretion, resulting in uneven levels of protection and difficulties in assessing regulatory compliance. This situation is further exacerbated by limited coordination between sector-specific regulations, horizontal legal instruments and soft law documents, as well as by insufficient financial resources, competency gaps and inadequate cybersecurity training. Consequently, the healthcare system remains particularly vulnerable to cyber incidents, the effects of which may extend beyond personal data breaches and directly threaten the continuity of healthcare services and patient safety.

4. Cybersecurity Threats in the Healthcare Sector and the Healthcare Data Trilemma: Between Value, Utility and Security

For several years now, Poland has been experiencing a steadily increasing number of cyberattacks, placing it among the most frequently targeted EU Member States. At the turn of 2025 and 2026, Polish government institutions were repelling nearly 3,188 attacks per week, significantly more than countries such as the Czech Republic, Austria or Germany[24]. In 2025, the number of reports registered by the CERT Poland exceeded 658,000, representing an increase of nearly 10% compared to 2024 (601,000) and over 70% compared to 2023 (371,000). The number of cybersecurity incidents recorded by CERT Poland in 2025 reached a record level of more than 260,800 incidents, which constitutes an increase of nearly 152% compared to 2024, when 103,400 incidents were recorded, and over 300% compared to 2023, when nearly 80,000 incidents were registered. In 2024, the highest number of incidents was recorded in public administration (1,911 cases), followed by the education sector (579 cases) and the healthcare sector (440 cases). Among the 57 serious incidents recorded in 2024, 44 occurred in the banking and financial market infrastructure sector, 11 concerned the healthcare sector and 2 were related to the transport sector. The number of serious incidents in 2024 increased by 43% compared to 2023[25].

In recent years, the healthcare sector has increasingly become a victim of cyberattacks and at a significantly higher rate than any other critical sector in the European Union. Medical data currently constitute one of the most attractive targets for cybercriminals. Their particular value stems from the fact that they combine

[24] Checkpoint Research, Cybersecurity Report 2026, https://www.checkpoint.com/security-report/ [accessed on: 07.02.2026].

[25] CSIRT NASK (CERT Polska), *Monthly Summary Report No. 4/2025*, December 2025, pp. 5–6; CERT Polska – NASK, *Annual Report on the Activities of CERT Polska 2023*.

information about an individual's health status with identification, financial and insurance data, making them exceptionally useful for various forms of cybercrime. Unlike payment card data, which can be blocked or replaced, medical data are permanent in nature, therefore once stolen, they may be exploited for many years for identity theft, credit fraud, insurance fraud, blackmail or sophisticated phishing attacks. Their potential commercial value should also not be underestimated, particularly in the context of creating behavioral profiles, conducting statistical analyses or generating personalized offers, which further increases interest in these resources among unauthorized actors. At the same time, the ongoing process of digitalization has revolutionized healthcare in recent years, making an increasing number of patient services available online. Examples include the e-prescription and e-referral systems, the implementation of the Internet Patient Account (IKP) and the rapid development of telemedicine. These initiatives demonstrate that modern technologies in medicine create new opportunities and have a tangible impact on doctors' work and patient care, contributing inter alia to improved quality of treatment and time savings for medical staff. Nevertheless, the healthcare system remains at an early stage of building a comprehensive e-health ecosystem. Significant challenges lie ahead, related to the further digitalization of medical services and the intensified development and use of systems based on cutting-edge technologies, including artificial intelligence. In this context, the proper use of such technologies in the provision of healthcare services will be crucial, particularly with regard to ensuring the security of privacy and health data, as the effective functioning of these systems requires the collection and processing of large datasets analyzed, among others, by artificial intelligence algorithms[26].

The attractiveness of the healthcare sector to cybercriminals is further reinforced by the relatively low level of technical and organizational security in many medical institutions. These entities often operate under conditions of chronic underfunding, limited human resources and a shortage of specialized IT personnel, which in practice leads to delays in implementing updates, a lack of systematic threat monitoring and insufficient staff training. As a result, cyberattacks on healthcare facilities are not only frequent but also relatively effective, causing disruptions and delays in the provision of health- and life-saving services[27]. Cybersecurity has thus ceased to be merely an IT issue and has become a systemic and organizational challenge.

[26] P. Kaźmierczyk, D. Lukosek, M. Czarnuch, A. Horyń, *Medical Data in Physicians' Practice: Current State and Desired Changes*, Report of the Network of Physician Innovators, Supreme Medical Chamber (Naczelna Izba Lekarska), 2023, pp. 23–27.

[27] https://commission.europa.eu/topics/digital-economy-and-society/cybersecurity-healthcare_pl [accessed on: 07.02.2026].

At the beginning of 2025, the Centre for eHealth (CeZ) published the results of the eighth edition of the survey on the level of digitalization of entities providing healthcare services. The aim of the survey was to obtain a comprehensive picture of the level of digitalization of healthcare entities, their readiness to create and share electronic medical documentation, their use of modern technologies (including telemedicine and artificial intelligence), as well as to assess needs and barriers in the area of IT, including cybersecurity[28]. The results indicate a decidedly insufficient level of cybersecurity in medical entities. Dedicated cybersecurity teams operate in only 27.5% of healthcare facilities, with a significantly higher prevalence in hospitals (41.1%) than in outpatient care providers (25.8%). Most of these teams consist of no more than five individuals, suggesting limited human resources and organizational capacity for comprehensive management of ICT security. Moreover, many facilities have not implemented even basic security measures: 60% do not monitor vulnerabilities, nearly one in ten lacks antivirus software and fewer than 60% regularly test backup copies and data recovery procedures.

The low prevalence of dedicated cybersecurity teams, combined with their small size, means that many healthcare facilities lack human resources capable of effectively managing cyber risk. With limited internal teams and a small number of specialists, facilities often rely on external service providers, which may raise challenges related to oversight and information security management. Furthermore, as many as 82.1% of surveyed facilities identified insufficient financial resources for IT investments as the main barrier to further digitalization, including cybersecurity, representing a significant increase compared to the previous survey cycle. Cybersecurity activities are most often financed from the entities' own funds (84.8%), less frequently with support from the National Health Fund (NFZ; 35.6%) or EU funds (10.6%). The lack of adequate budgets results in delays in implementing monitoring tools, risk analysis, intrusion detection systems and advanced security solutions, which are crucial from the perspective of cyber resilience. At the same time, the reported needs in the area of cybersecurity confirm that awareness of cyber risk and recognition of competency gaps are becoming increasingly widespread, particularly among hospitals (91.2% reporting needs in this area). The frequently expressed demand for intensive and widespread training of medical and administrative staff points to a significant deficit in competencies related to incident analysis and response, further exacerbating the sector's vulnerability to cyber threats. It is therefore unsurprising that more than half of healthcare facilities (53%) assess their level of digital maturity as average, while nearly 25% consider it low or very low.

[28] Centre for e-Health (Centrum e-Zdrowia), *VIII Survey on the Level of Digitalisation of Entities Performing Medical Activities*, March 2025, https://cez.gov.pl/pl/page/publikacje [accessed on: 08.02.2026], pp. 5–6; The study covered 11,185 healthcare entities.

In light of the results of the digitalization survey conducted by the Centre for eHealth, the examples of successful cyberattacks on medical facilities in Poland in recent years come as little surprise[29]. Within just eight months of 2025, the healthcare sector recorded the same number of cybersecurity incidents as in the entire year of 2024, namely 946 incidents. The most common attack methods included computer fraud (311 cases), vulnerable services (238 cases, of which as many as 150 were directly related to the exploitation of software vulnerabilities), and ransomware malware (90 cases)[30]. The number of ransomware attacks targeting healthcare facilities has been increasing worldwide at an alarming rate. Over a five-year period, such attacks increased more than sevenfold, from 69 cases reported in 2020 to 506 cases in 2024[31]. This highly unfavorable trend indicates that cybercriminals are increasingly targeting healthcare entities, exploiting both their weaker security posture and their critical societal role. In cases where hospital systems are infected with malware, attackers often remain undetected within the hospital's IT infrastructure for many months, analyzing its architecture, escalating privileges and exfiltrating data. Data indicate that the average time to identify an incident exceeds 270 days, giving attackers over half a year of unrestricted activity[32]. After this period, they reveal their presence by encrypting key systems (such as HIS, LIS, RIS, PACS or patient registration systems), effectively paralyzing the facility's operations[33]. The consequences are severe, particularly for patients, as they lead to the cancellation of surgeries and other planned procedures, necessitate the transfer of patients to other hospitals and, in extreme cases, may pose a real threat to life. From the perspective of medical institutions, such incidents also result in

[29] For example, in March 2025, hackers attacked the Ministry of Interior and Administration hospital in Krakow, paralyzing the hospital's main computer system handling medical records. The attackers encrypted files and demanded a ransom for their decryption, which forced a temporary suspension of admissions and a return to paper records. In April 2025, the entity administering the IKP reported a vulnerability: manual manipulation of the URL in the browser allowed unauthorized access to other patients' medical records. The user who reported the issue managed to access the medical records of four other individuals this way. In November 2023, hackers attacked data stored on ALAB Laboratoria servers using ransomware, which blocked access to the computer system, and then demanded a ransom to restore it to its original state. Cybercriminals seized approximately 200,000 patient records from 2017-2023.

[30] Number of data protection incidents in the healthcare sector: 2021 – 150, 2022 – 251, 2023 – 405, 2024 – 1028, I-VIII 2025 – 946. Centrum eZdrowia, www.cez.gov.pl [accessed on: 30.01.2026].

[31] https://www.cez.gov.pl/pl/page/o-nas/aktualnosci/csirt-cez-ostrzega-wzrasta-liczba-incydentow-cyberbezpieczenstwa-w-ochronie-zdrowia [accessed on: 28.01.2026].

[32] https://www.totalassure.com/blog/average-time-to-detect-cyber-attack-2025 [accessed on: 07.02.2026].

[33] Healthcare continues to experience the longest detection times across all industries requiring nearly *4*0 additional days beyond the global average. 2025 Data Breach Investigations Report, Healthcare Snapshot,https://www.verizon.com/business/resources/infographics/2025-dbir-healthcare-snapshot.pdf [accessed on: 07.02.2026].

data loss and significant financial damage[34]. A further critical gap remains the lack of business continuity plans and post-incident procedures. Meanwhile, regulations such as the NIS2 Directive explicitly require not only technical safeguards but also prepared scenarios for incident response and system recovery.

Medical data are becoming an increasingly attractive target for cybercriminals also due to their intrinsic value and monetization potential. They are highly valuable on the black market because they contain critical information that can be readily exploited for social engineering, phishing attacks and fraud. Such data may be used for insurance fraud, identity theft, as well as the illicit acquisition of medications through fraudulent prescriptions or abuses related to falsified sick leave certificates, vaccination certificates or other medical documentation.

5. Towards a Systemic Model of Cybersecurity in Public Health: Key Recommendations for Systemic Improvement

In recent years, we have witnessed an unprecedented increase in the volume of data generated and processed in digital form. Data have become one of the key resources of the modern economy, forming the foundation of numerous business models, research and development projects and innovation processes. This trend is particularly evident in the healthcare sector, where medical data are used not only for the ongoing treatment of patients but also for scientific research, the improvement of therapies, the planning of health policies and the deployment of solutions based on artificial intelligence and personalized medicine. The dynamic growth of digital data therefore represents a significant social opportunity, while at the same time generating new and increasingly serious risks. Parallel to technological development and the growing dependence of healthcare systems on ICT infrastructure, the scale of cyberattacks has been increasing, particularly those targeting sensitive data such as health-related information. Medical data belong to the most valuable categories of information on the black market, as they may be used for criminal purposes, blackmail or manipulation. At the same time, medical institutions remain among the least protected entities within the digital ecosystem. This situation results from the accumulation of several structural factors, including chronic underfunding of the healthcare system, limited budgets for IT and cybersecurity investments, shortages of specialized personnel and insufficient level of knowledge and awareness of threats. A particularly important element of this landscape is the human factor. Experience shows that even the most advanced technologies do not provide real security if they are not properly implemented,

[34] The average cost of a data breach in the global healthcare sector is estimated at $7.42 million in 2025,www.cez.gov.pl [accessed on: 08.02.2026].

configured and used. It is the human being who ultimately determines the effectiveness of technical solutions, both at the system design stage and during everyday operation. The lack of clear procedures, insufficient training and the routine neglect of cyber hygiene principles by medical and administrative staff significantly increase the vulnerability of healthcare institutions to incidents. Cybercriminals skillfully exploit these weaknesses, relying not only on technical attack methods but also on social engineering, phishing and manipulation of user behavior.

Awareness of these challenges led in 2025 to an initiative undertaken by the Council for Digitalization operating under the Minister of Digital Affairs[35], aimed at formulating recommendations to improve the level of cybersecurity in the healthcare sector[36]. These recommendations respond to a broader need to enhance coordination of activities, reduce regulatory fragmentation and bridge the gap between growing technological requirements and the actual organizational capabilities of healthcare entities. They demonstrate that improving cybersecurity cannot rely solely on technology but requires coherent regulatory, organizational and educational measures that take into account the specific characteristics of the healthcare sector and its key role in public security. The recommendations of the Council for Digitalization issued in 2025 should be viewed as an attempt to provide a systemic response to the shortcomings of the current cybersecurity regulatory model in the healthcare sector. Their significance lies not so much in introducing entirely new normative solutions, but rather in indicating the necessary directions of change within the framework of existing legal instruments.

The Council for Digitalization clearly points to the need to improve the coordination of cybersecurity activities in the healthcare sector at the central level. The absence of a single authority responsible for defining directions, priorities and standards leads to a dispersion of competences and inconsistency in actions taken. The recommendations move toward a clearer delineation of the roles of individual public institutions and the strengthening of cooperation mechanisms between government administration, healthcare entities and institutions responsible for cybersecurity. To this end, the Council recommended the establishment of a working group tasked with identifying security gaps and vulnerabilities within the healthcare sector and preparing legal, organizational and technical solutions aimed at enhancing the security of patient data.

---

[35] More information about the Council for Digital Affairs (its composition, tasks and recommendations) can be found at: https://www.gov.pl/web/cyfryzacja/rada-do-spraw-cyfryzacji [accessed on: 08.02.2026].

[36] Council for Digitalisation (Rada do Spraw Cyfryzacji) at the Ministry of Digital Affairs, *Statement of the Council for Digitalisation on Cybersecurity in the Healthcare Sector*, April 2025, https://www.gov.pl/attachment/cd8ef083-f370-49b7-b7c7-b17c97ebd404 [accessed on: 07.02.2026].

Another key postulate concerns the development and implementation of a uniform, mandatory catalogue of minimum cybersecurity requirements applicable to all entities within the healthcare system. The recommendations emphasize that the current freedom in selecting technical and organizational measures leads to significant disparities in security levels. The applied standards should take into account the specific nature of medical data, the role of critical systems and the growing cyber threats, while at the same time remaining feasible for implementation by smaller entities. Consequently, the most urgent actions identified include the immediate definition of minimum requirements for commercial medical practice software applications. Although the Act on the Healthcare Information System imposes on service providers and entities maintaining medical registers an obligation to ensure compliance of their ICT systems with minimum technical and functional requirements[37], service providers tend to focus primarily on functionality while marginalizing technical security aspects. For this reason, the Council recommended imposing a legal obligation to ensure specific minimum cybersecurity requirements for the systems of service providers and entities maintaining medical registers. Such requirements should include, inter alia, mandatory two-factor authentication, enforced use of strong passwords, account lockout after several unsuccessful login attempts, automatic session timeouts, role-based access control, encryption of network traffic, access control mechanisms, logging and monitoring of audit trails and automatic alerts in the event of detected suspicious activity.

An important area of the recommendations concerns the enhancement of digital competencies and awareness of cyber threats among medical staff, administrative personnel and management. In its position, the Council emphasizes that the human factor remains one of the weakest links in the system and that the lack of regular and practical training significantly increases vulnerability to social engineering attacks. The recommendations advocate for systemic, cyclical educational programs tailored to different roles within healthcare organizations. Cybersecurity should be treated as an integral element of risk management in healthcare entities, therefore, beyond a permanent educational component, medical institutions should be required to implement procedures related to business continuity planning, including regular testing of contingency plans. The recommendations stress the need to move away from a reactive approach toward a model based on risk analysis, resilience testing and continuous improvement of procedures. This includes not only prevention but also preparedness for incidents and response procedures in the event of a cybersecurity incident. Furthermore, healthcare sector entities should apply

[37] The Public Information Bulletin of the Minister of Health specifies the technical and functional requirements for service providers and entities maintaining medical records.

basic security measures to the devices and information systems they use, such as regular updating of operating systems.

In conclusion, the recommendations of the Council for Digitalization reveal the need to transition from a fragmented and reactive model of cybersecurity regulation in healthcare to a systemic, integrated and proactive approach. Their implementation requires not so much the multiplication of new obligations as better coordination of existing regulations, clarification of standards and genuine organizational and financial support for healthcare entities. In this sense, they constitute an important starting point for further *de lege ferenda* discussion on the future shape of cybersecurity in public health.

The scale, velocity, and heterogeneity of contemporary data generation underscore an increasingly visible gap between technological capabilities and existing legal and regulatory frameworks. In the absence of coherent, sector-specific and adaptive data governance and cybersecurity standards, the exponential growth of data, particularly sensitive health data, amplifies systemic vulnerabilities within digital public services. This reality calls for a deliberate transition from fragmented and reactive regulatory approaches towards a coordinated, risk-based and forward-looking model of cybersecurity governance in public health. Only such approach can ensure that the benefits of big data and advanced digital technologies are realized without undermining fundamental rights, public trust and the operational resilience of healthcare systems and their underlying digital infrastructures.

References

Sources

Act of 17 February 2005 on Computerisation of Activities of Entities Performing Public Tasks, Journal of Laws 2024, item 307.

Act of 6 November 2008 on Patients' Rights and the Patients' Rights Ombudsman, Journal of Laws 2024, item 581.

Act of 28 April 2011 on the Information System in Healthcare, Journal of Laws 2023, item 2465, as amended.

Act of 5 July 2018 on the National Cybersecurity System, Journal of Laws 2023, item 913, as amended.

Act of 10 May 2018 on the Protection of Personal Data, Journal of Laws 2019, item 1781, as amended.

Centre for e-Health (Centrum e-Zdrowia), *VIII Survey on the Level of Digitalisation of Entities Performing Medical Activities*, March 2025, https://cez.gov.pl/pl/page/publikacje [accessed on: 08.02.2026]. Code of Conduct for the Health Care Sector adopted pursuant to Article 40 of Regulation (EU) 2016/679, applicable to healthcare providers and data processors, 11 December 2023, https://uodo.gov.pl/pl/file/4525 [accessed on: 06.02.2026].

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, European strategy for data, COM/2020/66, Brussels, 19.02.2020.

Council for Digitalisation (Rada do Spraw Cyfryzacji) at the Ministry of Digital Affairs, *Statement of the Council for Digitalisation on Cybersecurity in the Healthcare Sector*, April 2025, https://www.gov.pl/ attachment/cd8ef083-f370-49b7-b7c7-b17c97ebd404 [accessed on: 07.02.2026].

Checkpoint Research, Cybersecurity Report 2026, https://www.checkpoint.com/security-report/ [accessed on: 07.02.2026].
CERT Polska – NASK, *Annual Report on the Activities of CERT Polska,* April 2023.
CSIRT NASK (CERT Polska), *Monthly Summary Report No. 4/2025*, December 2025, pp. 5–6.
Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148 (NIS 2 Directive), OJ L 333, 27.12.2022, pp. 80–152. ENISA – European Union Agency for Cybersecurity, "Cyber Hygiene in the Health Sector", September 2025, https://www.enisa.europa.eu/sites/default/files/2025-09/ENISA%20Cyber %20Hygiene% 20in%20the%20Health%20Sector.pdf [accessed on: 06.02.2026].
European Commission, *Communication from the Commission to the European Parliament and the Council: European Data Union Strategy – Unlocking Data for Artificial Intelligence*, Brussels, 19 November 2025, COM(2025) 835 final, pp. 7–8.
European Commission, Data, https://digital-strategy.ec.europa.eu/en/factpages/data [accessed on: 03.02.2026].
International Data Corporation, https://www.businesswire.com/news/home/2020051300 5075/en/IDCs-Global-StorageSphere-Forecast-Shows-Continued-Strong-Growth-in-the-Worlds-Installed -Base-of-Storage-Capacity [accessed on: 02.02.2026].
Implementing the European Health Data Space across Europe, eitHealth Think Tank, https://eithealth.eu/wp-content/uploads/2024/04/EIT_Health_ThinkTank_Implementing_the _EHDS_across_Europe_23.04.24.pdf., April 2024 [accessed on: 07.02.2026].
Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, https://eur-lex.europa.eu/eli/reg/2025/327/oj/eng [accessed on: 07.02.2026].
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), OJ L 119, 4 May 2016, p. 1, as amended. Regulation of the Council of Ministers of 21 May 2024 on the National Interoperability Framework, minimum requirements for public registers and electronic information exchange, and minimum requirements for ICT systems, Journal of Laws 2024, item 773.
Regulation of the Minister of Health of 6 April 2020 on the types, scope and models of medical documentation and the manner of its processing, Journal of Laws 2024, item 798.
Statista Global Data Platform, https://www.statista.com/statistics/871513/worldwide-data-created/ [accessed on: 01.02.2026].
TechSci Research LLC, https://www.techsciresearch.com/report/big-data-in-healthcare-market/4009.html [accessed on: 02.02.2026].
J. Thomason, "Data, Digital Worlds, and the Avatarization of Health Care", *Global Health Journal*, Vol. 8, Issue 1, March 2024, pp. 1–3. https://www.cez.gov.pl/pl/page/o-nas/aktualnosci/csirt-cez-ostrzega-wzrasta-liczba-incydentow-cyberbezpieczenstwa-w-ochronie-zdrowia [accessed on: 28.01.2026].
https://commission.europa.eu/topics/digital-economy-and-society/cybersecurity-healthcare_pl [accessed on: 07.02.2026].
https://www.totalassure.com/blog/average-time-to-detect-cyber-attack-2025 [accessed on: 07.02.2026].
2025 Data Breach Investigations Report, Healthcare Snapshot, https://www.verizon.com/ business/resources/infographics/2025-dbir-healthcare-snapshot.pdf [accessed on: 07.02.2026].

Literature

European Commission, *The Impact of the European Health Data Space Regulation and the Open Data Directive on Citizens*, data.europa.eu, Publications Office of the European Union, July 2025.

Kaźmierczyk P., Lukosek D., Czarnuch M., Horyń A., *Medical Data in Physicians' Practice: Current State and Desired Changes*, Report of the Network of Physician Innovators, Supreme Medical Chamber, 2023.

Ryś A., "European Health Data Space – New Challenges and Opportunities", *Med. Prakt.*, 2025, no. 3, pp. 110–119.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Martyna Kaczmarczyk
Uniwersytet Warmińsko-Mazurski w Olsztynie
ORCID: 0000-0001-6169-9466
martyna.kaczmarczyk@uwm.edu.pl

Między innowacją a bezpieczeństwem: o zasadach regulujących europejski system sztucznej inteligencji

Słowa kluczowe: sztuczna inteligencja, prawa podstawowe, cyberbezpieczeństwo, prawo Unii Europejskiej

Streszczenie. Dynamiczny rozwój technologii sztucznej inteligencji stwarza nowe możliwości innowacji, jednocześnie rodząc istotne wyzwania prawne, etyczne i społeczne. W odpowiedzi na te wyzwania Unia Europejska wprowadziła kompleksowe ramy regulacyjne mające na celu zapewnienie, że systemy sztucznej inteligencji są projektowane i wykorzystywane w sposób bezpieczny, przejrzysty i godny zaufania. Niniejszy artykuł analizuje zasady przewodnie regulujące funkcjonowanie sztucznej inteligencji w europejskim środowisku prawnym, ze szczególnym uwzględnieniem przepisów zawartych w AI Act. W opracowaniu omówiono podejście oparte na analizie ryzyka, które klasyfikuje systemy AI w zależności od ich potencjalnego wpływu na bezpieczeństwo, prawa podstawowe oraz interes społeczny. Szczególną uwagę poświęcono kluczowym zasadom regulacyjnym, takim jak przejrzystość działania systemów, nadzór człowieka, odpowiedzialność podmiotów tworzących i wykorzystujących AI, jakość danych oraz zarządzanie ryzykiem. W artykule przedstawiono również obowiązki nakładane na dostawców i twórców systemów wysokiego ryzyka, w tym wymagania dotyczące dokumentacji technicznej, procedur oceny zgodności oraz monitorowania systemów po ich wdrożeniu. Przeprowadzona analiza wskazuje, że europejskie podejście regulacyjne dąży do znalezienia równowagi między wspieraniem innowacji technologicznych a ochroną praw podstawowych i budowaniem zaufania społecznego do nowych technologii. Ustanowienie jasnych standardów prawnych dla projektowania i stosowania systemów sztucznej inteligencji ma sprzyjać odpowiedzialnemu rozwojowi AI, jednocześnie ograniczając potencjalne zagrożenia wynikające z wykorzystania zaawansowanych technologii algorytmicznych.

Between innovation and security: about principles governing the European artificial intelligence system

Keywords: artificial intelligence, fundamental rights, cybersecurity, European Union law

Summary. Rapid development of artificial intelligence technologies has created new opportunities for innovation while simultaneously raising significant legal, ethical and societal concerns. In response to these challenges European Union has introduced a comprehensive regulatory framework aimed at ensuring that artificial intelligence systems are developed in a safe, transparent and trustworthy manner. Article analyzes guiding principles governing artificial intelligence within the European regulatory environment, with particular reference to the provisions of the AI Act. Study examines the risk-based approach adopted by the regulation, which classifies AI systems according to their potential impact on fundamental rights, safety and societal interests. Special attention is devoted to

key regulatory principles such as transparency, human oversight, accountability, data quality and risk management. Article also discusses the obligations imposed on developers and providers of high-risk AI systems, including requirements related to documentation, compliance assessment and post-market monitoring. Analysis demonstrates that the European approach seeks to balance technological innovation with the protection of fundamental rights and public trust. By establishing clear legal standards for the design and deployment of AI systems, the regulation aims to foster responsible innovation while minimizing potential harms associated with the use of advanced algorithmic technologies. Article concludes that the European regulatory model may serve as an important reference point for global discussions on the governance of artificial intelligence.

Wprowadzenie

Rozwój technologii cyfrowych w ostatnich latach doprowadził do znaczącego wzrostu roli systemów sztucznej inteligencji w wielu obszarach życia społecznego, gospodarczego oraz administracyjnego. Rozwiązania oparte na sztucznej inteligencji znajdują coraz szersze zastosowanie między innymi w medycynie, finansach, edukacji, transporcie, przemyśle czy administracji publicznej. Systemy te wspierają procesy decyzyjne, analizę dużych zbiorów danych oraz automatyzację wielu zadań, które dotychczas wymagały bezpośredniego udziału człowieka. W konsekwencji sztuczna inteligencja staje się jednym z kluczowych czynników transformacji cyfrowej współczesnych społeczeństw oraz gospodarek.

Jednocześnie dynamiczny rozwój technologii sztucznej inteligencji rodzi liczne wyzwania o charakterze prawnym, etycznym i społecznym. Coraz częściej podnoszone są kwestie związane z przejrzystością działania algorytmów, odpowiedzialnością za decyzje podejmowane przez systemy autonomiczne, ochroną danych osobowych czy ryzykiem występowania uprzedzeń algorytmicznych. W kontekście rosnącej autonomii systemów AI pojawiają się również pytania dotyczące możliwości skutecznego nadzoru nad ich funkcjonowaniem oraz zapewnienia zgodności ich działania z podstawowymi wartościami demokratycznego państwa prawa, w tym z poszanowaniem praw podstawowych jednostki.

W odpowiedzi na te wyzwania instytucje europejskie podjęły działania zmierzające do stworzenia spójnych i kompleksowych ram prawnych regulujących rozwój oraz wykorzystanie technologii sztucznej inteligencji. Jednym z najważniejszych elementów tych działań jest przyjęcie aktu prawnego znanego jako AI Act[1]. Regulacja ta stanowi pierwszą na świecie próbę całościowego uregulowania funkcjonowania systemów sztucznej inteligencji na poziomie ponadnarodowym. Jej głównym celem jest stworzenie ram prawnych umożliwiających bezpieczny

[1] Parlament Europejski i Rada Unii Europejskiej. Rozporządzenie (UE) 2024/1689 z dnia 13 czerwca 2024 r. ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (AI Act), Dz.Urz. UE L 1689, 12.07.2024.

i odpowiedzialny rozwój technologii AI, przy jednoczesnym wspieraniu innowacyjności i konkurencyjności gospodarki europejskiej.

Podstawą przyjętego modelu regulacyjnego jest podejście oparte na analizie ryzyka, w ramach którego systemy sztucznej inteligencji klasyfikowane są w zależności od poziomu potencjalnego zagrożenia, jakie mogą stwarzać dla bezpieczeństwa użytkowników, praw podstawowych jednostki oraz interesu publicznego. W zależności od stopnia ryzyka regulacja przewiduje zróżnicowany zakres obowiązków nakładanych na podmioty projektujące, rozwijające i wdrażające systemy AI. Szczególnie rygorystyczne wymagania dotyczą systemów wysokiego ryzyka, które mogą mieć istotny wpływ na sytuację prawną lub społeczną jednostki, np. w obszarze zatrudnienia, edukacji, ochrony zdrowia czy usług finansowych[2].

Istotnym elementem europejskiego podejścia do regulacji sztucznej inteligencji jest także promowanie tzw. godnej zaufania sztucznej inteligencji (*trustworthy AI*), której funkcjonowanie powinno opierać się na zasadach przejrzystości, bezpieczeństwa, odpowiedzialności oraz nadzoru człowieka[3]. W tym kontekście regulacja ma na celu nie tylko ograniczenie potencjalnych zagrożeń związanych z wykorzystaniem technologii AI, lecz również budowanie społecznego zaufania do nowych technologii oraz wspieranie ich odpowiedzialnego wdrażania w różnych sektorach gospodarki.

Celem niniejszego artykułu jest przedstawienie i analiza zasad przewodnich regulujących funkcjonowanie systemów sztucznej inteligencji w europejskim porządku prawnym. Szczególna uwaga zostanie poświęcona kluczowym zasadom regulacyjnym wynikającym z przepisów wspomnianego aktu prawnego, takim jak przejrzystość działania systemów AI, nadzór człowieka nad procesami decyzyjnymi, odpowiedzialność podmiotów rozwijających i wdrażających technologie sztucznej inteligencji, a także wymogi dotyczące jakości danych oraz zarządzania ryzykiem. Analiza ta ma na celu ukazanie, w jaki sposób europejski model regulacyjny stara się równoważyć potrzebę wspierania innowacji technologicznych z koniecznością zapewnienia bezpieczeństwa, ochrony praw podstawowych oraz budowania zaufania społecznego do systemów sztucznej inteligencji.

Podejście oparte na ryzyku

Jednym z kluczowych elementów europejskiego modelu regulacyjnego jest podejście oparte na analizie ryzyka, stanowiące fundament regulacji zawartej w AI Act[4].

[2] AI Act, art. 8-15.

[3] Komisja Europejska, *Communication on Fostering a European Approach to Trustworthy AI*, COM (2019) 168 final, 2019, https://digital-strategy.ec.europa.eu/en/library/communication-fostering-european-approach-trustworthy-ai [dostęp: 21.01.2026].

[4] AI Act, art. 6.

Podejście to zakłada, że zakres obowiązków prawnych oraz poziom kontroli nad systemami sztucznej inteligencji powinny być uzależnione od stopnia zagrożenia, jakie dany system może stwarzać dla bezpieczeństwa, praw podstawowych jednostki oraz interesu publicznego.

Istotą podejścia opartego na ryzyku jest klasyfikacja systemów sztucznej inteligencji według poziomu potencjalnego ryzyka ich zastosowania. W ramach przyjętego modelu wyróżnia się cztery podstawowe kategorie systemów AI: systemy stwarzające niedopuszczalne ryzyko, systemy wysokiego ryzyka, systemy o ograniczonym ryzyku oraz systemy o minimalnym ryzyku[5]. Najbardziej restrykcyjne regulacje dotyczą systemów uznanych za stwarzające niedopuszczalne ryzyko, których stosowanie zostało w zasadzie zakazane ze względu na możliwość poważnego naruszenia praw podstawowych, np. poprzez manipulowanie zachowaniem użytkowników lub stosowanie niektórych form biometrycznej identyfikacji w przestrzeni publicznej.

Szczególną kategorię stanowią systemy wysokiego ryzyka, które mogą mieć istotny wpływ na sytuację prawną lub społeczną jednostki. Do tej grupy zaliczane są m.in. systemy wykorzystywane w procesach rekrutacyjnych, systemy oceny zdolności kredytowej, rozwiązania stosowane w sektorze ochrony zdrowia czy w zarządzaniu infrastrukturą krytyczną. W odniesieniu do tych systemów regulacja przewiduje szereg obowiązków dla podmiotów rozwijających i wdrażających te systemy. Obejmują one między innymi konieczność przeprowadzania analizy ryzyka, zapewnienia odpowiedniej jakości danych wykorzystywanych do trenowania modeli, prowadzenia szczegółowej dokumentacji technicznej oraz zapewnienia nadzoru człowieka nad działaniem systemu[6].

Systemy o ograniczonym ryzyku podlegają przede wszystkim wymogom transparentności. W praktyce oznacza to obowiązek informowania użytkowników o tym, że mają do czynienia z systemem sztucznej inteligencji, co ma szczególne znaczenie w przypadku *chatbotów* czy systemów generujących treści[7]. Natomiast systemy o minimalnym ryzyku, takie jak filtry antyspamowe czy algorytmy rekomendacyjne, mogą być stosowane bez dodatkowych wymogów regulacyjnych, ponieważ uznaje się, że nie stwarzają one istotnego zagrożenia dla użytkowników lub społeczeństwa[8].

---

[5] *Ibidem*.

[6] A. Pajdo, *Zakres obowiązków podmiotów odpowiedzialnych za systemy sztucznej inteligencji wysokiego ryzyka w świetle aktu o SI – Rozporządzenia 2024/1689*, „Roczniki Nauk Prawnych" 2025, nr 2(35), s. 47.

[7] AI Act, art. 50; S. Westerstrand, *Fairness in AI systems development: EU AI Act compliance and beyond*, „Information and Software Technology" 2025, vol. 187, s. 7.

[8] AI Act, art. 53; M. Ebers, *Truly Risk-based Regulation of Artificial Intelligence: How to Implement the EU's AI Act*, „European Journal of Risk Regulation" 2024, vol. 16, is. 2, s. 6-7.

Podejście oparte na analizie ryzyka stanowi próbę znalezienia równowagi pomiędzy potrzebą zapewnienia bezpieczeństwa a koniecznością wspierania rozwoju innowacji technologicznych. Zamiast wprowadzać jednolite, restrykcyjne regulacje dla wszystkich systemów sztucznej inteligencji, model ten umożliwia elastyczne dostosowanie poziomu regulacji do rzeczywistego stopnia zagrożenia. Dzięki temu możliwe jest jednoczesne promowanie rozwoju technologii AI oraz ograniczanie potencjalnych negatywnych skutków ich wykorzystania. W konsekwencji podejście oparte na ryzyku stanowi jeden z najważniejszych elementów współczesnej regulacji sztucznej inteligencji w Europie. Jego skuteczność będzie jednak zależeć od praktycznego wdrożenia przepisów oraz od dalszego rozwoju badań naukowych analizujących wpływ tych regulacji na rozwój technologii oraz ochronę praw jednostki.

Wymóg transparentności

W miarę jak systemy sztucznej inteligencji stają się coraz bardziej powszechne w codziennym życiu i w kluczowych sektorach gospodarki, rośnie znaczenie zasady transparentności w regulacji ich funkcjonowania. Transparentność stanowi jeden z fundamentalnych filarów europejskiego modelu regulacyjnego, określonego w AI Act, i ma na celu zapewnienie, że użytkownicy i interesariusze systemów AI są świadomi sposobu ich działania, a także potencjalnych konsekwencji podejmowanych decyzji[9].

Wymóg transparentności obejmuje kilka kluczowych aspektów. Przede wszystkim systemy sztucznej inteligencji powinny informować użytkowników, gdy mają do czynienia z AI, co jest szczególnie istotne w przypadku interaktywnych narzędzi, takich jak *chatboty*, systemy rekomendacyjne czy generatory treści. Brak takiej informacji może prowadzić do wprowadzania użytkowników w błąd oraz utrudniać świadome korzystanie z technologii. Transparentność wymaga także udostępnienia informacji dotyczących źródeł danych, metod analitycznych i kryteriów podejmowania decyzji przez algorytmy, co umożliwia audyt i ocenę zgodności systemu z zasadami etycznymi i prawnymi[10].

W kontekście systemów wysokiego ryzyka transparentność staje się nie tylko narzędziem informacyjnym, ale także mechanizmem odpowiedzialności i nadzoru. Umożliwia to podmiotom regulacyjnym, użytkownikom i ekspertom technicznym ocenę, czy systemy AI działają zgodnie z obowiązującymi standardami bezpieczeń-

[9] K. Vladimirova, *The Principle of Transparency Under The Artificial Intelligence Act*, „Law Journal of New Bulgarian University" 2024, nr 2 (vol. 20), s. 44.

[10] M. Wyszomirska, *Implementation of the AI Act (EU) into the Polish legal system, its impact on this system and on selected strategic sectors*, „Safety & Fire Technology" 2025, nr 2(66), s. 120.

stwa, jakości danych i ochrony praw podstawowych. Transparentność przyczynia się również do budowania zaufania społecznego do technologii, co jest niezbędne dla szerszego wdrażania innowacyjnych rozwiązań w sektorach takich jak zdrowie, edukacja czy finanse.

W literaturze akademickiej wymóg transparentności systemów AI dopiero zaczyna być analizowany w kontekście praktycznego wdrażania regulacji i oceny ich skuteczności. Badacze zwracają uwagę na wyzwania związane z równoważeniem pełnej przejrzystości z ochroną własności intelektualnej, złożonością algorytmów oraz ryzykiem nadużyć informacyjnych[11]. Pomimo wyzwań transparentność pozostaje jednym z kluczowych elementów odpowiedzialnego rozwoju sztucznej inteligencji, zapewniającym zarówno ochronę użytkowników, jak i możliwość prowadzenia badań nad wpływem AI na społeczeństwo.

Podsumowując, wymóg transparentności w regulacji systemów AI nie jest jedynie formalnym obowiązkiem prawnym, lecz fundamentem budowania zaufania, nadzoru i odpowiedzialności w ekosystemie sztucznej inteligencji. Prawidłowe wdrożenie jest niezbędne dla zapewnienia, że rozwój technologii będzie przebiegał w sposób etyczny, bezpieczny i zgodny z wartościami demokratycznymi.

Nadzór człowieka

Kolejnym kluczowym elementem europejskiego podejścia regulacyjnego do sztucznej inteligencji, określonego w AI Act, jest zasada nadzoru człowieka (*human oversight*)[12]. Wymóg ten wynika z potrzeby zapewnienia, że decyzje podejmowane przez systemy AI, zwłaszcza te o wysokim ryzyku, są kontrolowane i weryfikowane przez człowieka. Celem jest ochrona praw podstawowych, ograniczenie ryzyka błędnych lub dyskryminujących decyzji oraz zwiększenie odpowiedzialności podmiotów wykorzystujących technologię[13].

Nadzór człowieka może przybierać różne formy w zależności od charakteru systemu i jego zastosowania. W przypadku systemów wysokiego ryzyka, np. w medycynie, finansach czy rekrutacji, nadzór obejmuje możliwość interwencji człowieka w procesie decyzyjnym, weryfikację wyników generowanych przez AI oraz, w razie potrzeby, wyłączenie systemu. Mechanizmy te mają na celu nie tylko zapobieganie

[11] M. Namysłowska *et al.*, *O etycznych, prawnych i społecznych konsekwencjach stosowania systemów sztucznej inteligencji w państwach członkowskich Unii Europejskiej. Uwagi na tle projektu rozporządzenia w sprawie sztucznej inteligencji*, „Przegląd Sejmowy" 2023, nr 6(179), s. 66-67.

[12] AI Act, art. 14.

[13] J. Kulesza, *Europe's Regulatory Sovereignty in the Age of Artificial Intelligence. A Human-Centric Response to the US-China Technological Rivalry*, „Sprawy Międzynarodowe" 2025, t. 78, nr 4, s. 177 i nast.

szkodom, ale również umożliwienie wykrywania błędów algorytmicznych i monitorowania wpływu systemów AI na prawa jednostki[14].

Wymóg nadzoru człowieka wiąże się ściśle z innymi zasadami regulacyjnymi, takimi jak transparentność i odpowiedzialność. Aby człowiek mógł skutecznie kontrolować system, musi mieć dostęp do informacji o sposobie działania algorytmu, kryteriach podejmowania decyzji oraz jakości danych wykorzystanych do trenowania modelu. W literaturze naukowej temat nadzoru człowieka w AI jest nadal w początkowej fazie badań, zwłaszcza w kontekście jego praktycznej realizacji w różnorodnych sektorach gospodarki i administracji publicznej. Wyzwania obejmują m.in. określenie optymalnego stopnia interwencji człowieka, uniknięcie nadmiernego obciążenia operacyjnego oraz minimalizowanie ryzyka automatyzacji decyzji bez odpowiedniego nadzoru[15].

Nadzór człowieka jest fundamentem odpowiedzialnego stosowania systemów sztucznej inteligencji. Zapewnia równowagę między automatyzacją procesów a ochroną praw jednostki, wzmacnia zaufanie społeczne do technologii oraz umożliwia skuteczne monitorowanie i kontrolę działania AI. Jego prawidłowe wdrożenie stanowi kluczowy warunek bezpieczeństwa, etycznego stosowania i legalności systemów sztucznej inteligencji w środowisku europejskim.

Jakość danych

Innym ważnym elementem skutecznej i odpowiedzialnej regulacji systemów sztucznej inteligencji, określonej w AI Act, jest wymóg zapewnienia wysokiej jakości danych wykorzystywanych do trenowania, testowania i działania modeli AI[16]. Dane stanowią fundament rzetelnego funkcjonowania systemów sztucznej inteligencji, ponieważ decyzje podejmowane przez algorytmy są bezpośrednio zależne od jakości informacji, na których zostały oparte. Błędne, niekompletne lub stronnicze dane mogą prowadzić do decyzji niezgodnych z prawem, dyskryminujących lub potencjalnie szkodliwych dla jednostek i społeczeństwa.

Regulacja europejska kładzie nacisk na to, aby dane używane w systemach AI były dokładne, reprezentatywne, kompletne i wolne od uprzedzeń. Oznacza to konieczność stosowania zaawansowanych procedur weryfikacji danych, oczyszczania zbiorów danych z błędów i anomalii, a także monitorowania jakości danych w czasie rzeczywistym. Uwzględnienie różnorodności demograficznej, społecznej i geograficznej w zbiorach danych jest istotne nie tylko dla poprawności działania

[14] M. Ebers, *op. cit.*, s. 16-17.

[15] B. Kaczmarek-Templin, *Sztuczna inteligencja (AI) i perspektywy jej wykorzystania w postępowaniu przed sądem cywilnym*, „Studia Prawnicze. Rozprawy i Materiały" 2022, nr 2(31), s. 74-75.

[16] AI Act, art. 10.

algorytmów, lecz także dla ograniczenia ryzyka powstawania uprzedzeń algorytmicznych, które mogłyby prowadzić do nierównego traktowania użytkowników lub grup społecznych[17].

Wysoka jakość danych ma także fundamentalne znaczenie dla przejrzystości i kontroli systemów AI. Tylko systemy o danych wysokiej jakości pozwalają na skuteczną kontrolę nad algorytmami oraz identyfikację potencjalnych błędów i ryzyk. Transparentność źródeł danych oraz sposobu ich przetwarzania umożliwia organom nadzoru, ekspertom technicznym i użytkownikom weryfikację zgodności działania systemu z obowiązującymi standardami prawnymi i etycznymi. W kontekście systemów wysokiego ryzyka, takich jak narzędzia wykorzystywane w medycynie, edukacji czy finansach, zapewnienie wysokiej jakości danych jest warunkiem umożliwiającym nadzór człowieka nad procesem decyzyjnym i minimalizowanie ryzyka szkód dla jednostek.

Literatura naukowa w obszarze jakości danych w AI jest wciąż w fazie rozwoju. Badania wskazują, że głównym wyzwaniem nie jest jedynie zapewnienie poprawności i kompletności danych, lecz także wykrywanie ukrytych uprzedzeń i asymetrii wynikających z historycznych nierówności społecznych lub specyfiki procesów zbierania danych[18]. Problemem pozostaje także brak jednolitych standardów oceny jakości danych oraz metodologii audytu stosowanych w różnych sektorach gospodarki i administracji publicznej. W praktyce twórcy systemów AI muszą balansować między pełną przejrzystością a ochroną własności intelektualnej, złożonością algorytmów oraz ochroną danych osobowych[19].

Zasada jakości danych jest również ściśle powiązana z innymi kluczowymi wymaganiami regulacyjnymi, takimi jak transparentność, nadzór człowieka i zarządzanie ryzykiem. System oparty na niskiej jakości danych ogranicza skuteczność nadzoru, utrudnia identyfikację uprzedzeń oraz zwiększa ryzyko podejmowania decyzji niezgodnych z prawem lub wartościami społecznymi. Dlatego zapewnienie wysokiej jakości danych jest niezbędne nie tylko dla prawidłowego działania systemów AI, lecz także dla budowania społecznego zaufania do technologii oraz odpowiedzialnego rozwoju innowacji w Unii Europejskiej.

Jakość danych stanowi fundament odpowiedzialnego i bezpiecznego stosowania systemów sztucznej inteligencji. Zapewnienie odpowiedniego standardu danych pozwala na rzetelne podejmowanie decyzji przez algorytmy, ogranicza ryzyko dyskryminacji, wspiera transparentność i umożliwia skuteczny nadzór człowieka.

---

[17] K. Sienkiewicz-Małyjurek, *Możliwości i problemy zastosowania sztucznej inteligencji w zarządzaniu kryzysowym,* „Bezpieczeństwo. Teoria i Praktyka" 2024, nr 1, s. 44-47.

[18] R. Rejmaniak, *Bias in Artificial Intelligence Systems*, „Białostockie Studia Prawnicze" 2021, nr 3 (vol. 26), s. 26-29.

[19] M. Namysłowska *et al.*, *op. cit.*, s. 50.

Niewiele jest jeszcze badań akademickich w tym obszarze, co wskazuje na potrzebę dalszych analiz i opracowania standardów jakości danych, które mogłyby być stosowane w praktyce przez twórców systemów AI oraz organy regulacyjne w różnych sektorach życia społecznego i gospodarczego.

Bezpieczeństwo i zarządzanie ryzykiem

Bezpieczeństwo i zarządzanie ryzykiem stanowią fundament odpowiedzialnego rozwoju i stosowania systemów sztucznej inteligencji. W regulacji europejskiej zagadnienia te zajmują centralne miejsce, szczególnie w kontekście systemów wysokiego ryzyka, których działanie może mieć istotny wpływ na prawa jednostki, bezpieczeństwo publiczne oraz interes społeczny[20]. Skuteczna implementacja regulacji z zakresu AI wymaga nie tylko innowacyjnych rozwiązań technologicznych, lecz także systematycznego zarządzania potencjalnymi zagrożeniami i zapewnienia bezpieczeństwa użytkowników.

Podejście oparte na bezpieczeństwie w AI obejmuje wszystkie etapy cyklu życia systemu – począwszy od projektowania, rozwoju, aż po wdrożenie i eksploatację. Na etapie projektowania konieczne jest przewidywanie możliwych scenariuszy zagrożeń oraz opracowywanie mechanizmów minimalizujących ryzyko awarii, błędów algorytmicznych czy niezamierzonych konsekwencji decyzji systemu. W fazie rozwoju i testowania systemu kluczowe znaczenie mają procedury walidacji, testowania odporności systemu na błędy oraz symulacje scenariuszy krytycznych. Po wdrożeniu systemu istotne jest natomiast monitorowanie jego działania w czasie rzeczywistym, identyfikacja potencjalnych nieprawidłowości oraz szybka reakcja w przypadku wystąpienia incydentów[21].

Zarządzanie ryzykiem w regulacji AI polega na systematycznej identyfikacji, ocenie i minimalizacji potencjalnych zagrożeń związanych z funkcjonowaniem systemu. AI Act wprowadza podejście oparte na klasyfikacji ryzyka, które pozwala dostosować poziom wymogów regulacyjnych do potencjalnego wpływu systemu na bezpieczeństwo i prawa jednostki. Systemy wysokiego ryzyka są objęte szczegółowymi obowiązkami dotyczącymi dokumentacji technicznej, audytu algorytmów, jakości danych oraz mechanizmów nadzoru człowieka, co pozwala na ograniczenie prawdopodobieństwa powstania szkód lub naruszeń praw podstawowych[22].

---

[20] AI Act, art. 9.

[21] D. Skoczylas, *Sztuczna inteligencja a dobrostan człowieka i ochrona praw podstawowych – rozważania na gruncie aktu w sprawie sztucznej inteligencji*, „Studia Prawnoustrojowe" 2025, nr 1(67), s. 357-358.

[22] AI Act, art. 11-12.

Bezpieczeństwo systemów AI wiąże się ściśle z innymi zasadami regulacyjnymi, takimi jak jakość danych, transparentność oraz nadzór człowieka. Niedokładne lub stronnicze dane mogą prowadzić do błędnych decyzji algorytmów. Brak transparentności ogranicza możliwość monitorowania działania systemu, a brak nadzoru człowieka zwiększa ryzyko niekontrolowanych skutków decyzji automatycznych. Dlatego skuteczne zarządzanie ryzykiem wymaga szerokiego podejścia, integrującego wszystkie powyższe aspekty oraz regularnej oceny ryzyka w trakcie całego cyklu życia systemu sztucznej inteligencji.

W literaturze przedmiotu dopiero zaczyna być szerzej analizowane zagadnienie bezpieczeństwa i zarządzania ryzykiem w kontekście praktycznej implementacji regulacji AI. Wskazuje się m.in. na potrzebę opracowania standardów audytu bezpieczeństwa, wytycznych w zakresie monitorowania systemów oraz metod oceny skuteczności działań minimalizujących ryzyko[23]. Szczególne wyzwania dotyczą dynamicznie rozwijających się modeli AI, takich jak generatywne systemy uczenia maszynowego, w przypadku których tradycyjne mechanizmy kontroli mogą być niewystarczające[24].

Podsumowując, bezpieczeństwo i zarządzanie ryzykiem stanowią nieodłączny element odpowiedzialnego wykorzystania sztucznej inteligencji. Wdrażanie skutecznych mechanizmów w tym zakresie jest niezbędne dla ochrony użytkowników i społeczeństwa, budowania zaufania do technologii oraz zapewnienia zgodności z europejskimi wartościami prawnymi i etycznymi. Przyszłe badania w tym obszarze będą kluczowe dla opracowania standardów i wytycznych, które umożliwią harmonijne połączenie innowacji technologicznych z bezpieczeństwem i ochroną praw jednostki.

Odpowiedzialność i dokumentacja

Odpowiedzialność za działania systemów sztucznej inteligencji oraz dokumentacja stosowania AI stanowią istotny wymóg prawny stawiany przez europejską regulację[25]. Zasady te mają na celu zapewnienie, że twórcy, dostawcy i użytkownicy systemów AI ponoszą odpowiedzialność za działanie technologii, a jednocześnie umożliwiają audyt, kontrolę i weryfikację ich funkcjonowania. W dobie rosnącej autonomii algorytmów odpowiedzialność prawna i dokumentacja stają się narzędziami ochrony użytkowników oraz budowania zaufania społecznego do nowych

[23] A. Pajdo, *op. cit.*, s. 49.
[24] C. Gutierrez, A. Aguirre, R. Uuk, C. Boine, M. Franklin, *A Proposal for a Definition of General Purpose Artificial Intelligence Systems*, „Digital Society" 2023, vol. 2, art. 36, s. 2-4.
[25] AI Act, art. 13-14.

technologii, a także mechanizmem ograniczania ryzyka prawnego i finansowego dla podmiotów rozwijających AI.

Podstawowym wymogiem w zakresie odpowiedzialności jest możliwość jednoznacznej identyfikacji podmiotów odpowiedzialnych za projektowanie, rozwój i wdrażanie systemów sztucznej inteligencji. Regulacja AI Act szczególnie podkreśla obowiązki dotyczące systemów wysokiego ryzyka, które mogą znacząco wpływać na prawa jednostki, bezpieczeństwo publiczne lub interes społeczny. Twórcy i dostawcy systemów zobowiązani są do wdrożenia procedur weryfikacji, testowania i monitorowania algorytmów oraz do prowadzenia szczegółowej dokumentacji technicznej[26]. Dokumentacja ta umożliwia audyt systemu, analizę ryzyka oraz przypisanie odpowiedzialności w przypadku błędów lub incydentów.

Dokumentacja techniczna obejmuje szeroki zakres informacji, w tym: architekturę systemu, wykorzystane metody uczenia maszynowego, szczegóły dotyczące zbiorów danych (źródła, jakość, reprezentatywność), wyniki testów i walidacji systemu, procedury monitorowania oraz mechanizmy zarządzania ryzykiem. W przypadku systemów wysokiego ryzyka zebrany materiał powinien również dokumentować interakcję człowieka z systemem, strategie nadzoru oraz procedury reagowania na sytuacje krytyczne[27]. Dzięki temu możliwe jest nie tylko identyfikowanie przyczyn ewentualnych błędów, ale także weryfikacja, czy system spełnia wymagania prawne i etyczne.

W literaturze zagadnienia odpowiedzialności i dokumentacji w AI dopiero zaczynają być szerzej analizowane[28]. Złożoność współczesnych systemów AI, w tym modeli generatywnych i autonomicznych, stwarza nowe wyzwania w zakresie przypisywania odpowiedzialności. W praktyce często trudne jest jednoznaczne określenie, który podmiot – twórca algorytmu, dostawca danych, integrator systemu czy użytkownik końcowy – odpowiada za skutki decyzji podejmowanych przez system. Ponadto dynamiczny charakter uczenia się algorytmów i zmienność danych utrudniają prowadzenie stałej dokumentacji oraz ocenę ryzyka w czasie rzeczywistym.

Odpowiedzialność i dokumentacja są ściśle powiązane z innymi zasadami regulacyjnymi, takimi jak transparentność, nadzór człowieka, jakość danych i bezpieczeństwo. Bez rzetelnej dokumentacji trudno zapewnić skuteczny nadzór człowieka,

---

[26] J. Laux, H. Ruschemeier, *Automation Bias in the AI Act: On the Legal Implications of Attempting to De-Bias Human Oversight of AI*, „European Journal of Risk Regulation" 2025, nr 16, s. 1522-1523.

[27] F. Doshi-Velez, M. Kortz, *Accountability of AI Under the Law: The Role of Explanation*, Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper 2017, s. 11-12.

[28] Zob. M. Świerczyński, Z. Więckowski, *Liability for damages caused by artificial intelligence systems – main challenges to be addressed by the European Union conflict-of-laws regulations*, „Prawo w Działaniu" 2023, t. 54, s. 200-216.

identyfikację uprzedzeń w danych oraz ocenę bezpieczeństwa systemu. Właściwa dokumentacja umożliwia również organom regulacyjnym kontrolę zgodności systemów z przepisami, a także służy jako podstawa do wprowadzenia mechanizmów odpowiedzialności cywilnej lub administracyjnej. W tym kontekście dokumentacja i odpowiedzialność pełnią zarówno funkcję prewencyjną, jak i kontrolną, zmniejszając ryzyko nieetycznych lub niebezpiecznych działań AI.

Odpowiedzialność i dokumentacja są niezbędnymi elementami bezpiecznego stosowania systemów sztucznej inteligencji. Zapewniają mechanizmy weryfikacji, audytu i kontroli, umożliwiają przypisanie odpowiedzialności podmiotom tworzącym i wdrażającym systemy AI, a także wspierają budowanie zaufania społecznego do technologii. W kontekście unijnej regulacji zasady te pozwalają łączyć innowacyjność z bezpieczeństwem, ochroną praw jednostki i zgodnością z wartościami demokratycznymi. Dalsze badania naukowe w tym obszarze są niezbędne dla opracowania standardów dokumentacji, audytu i mechanizmów odpowiedzialności, które będą mogły być stosowane w praktyce przez twórców systemów AI, organy regulacyjne oraz użytkowników końcowych, tworząc spójny i bezpieczny ekosystem sztucznej inteligencji.

Specjalne zasady dla modeli ogólnego przeznaczenia

Modele ogólnego przeznaczenia (ang. *General Purpose AI*, GPAI) stanowią jeden z najbardziej innowacyjnych i dynamicznie rozwijających się obszarów współczesnej sztucznej inteligencji[29]. W odróżnieniu od systemów wyspecjalizowanych, które zostały zaprojektowane do realizacji określonych zadań, modele GPAI charakteryzują się dużą elastycznością i możliwością adaptacji do różnych zastosowań w wielu sektorach gospodarki, administracji publicznej i życia codziennego. Ta wszechstronność niesie ze sobą zarówno ogromny potencjał innowacyjny, jak i zwiększone ryzyko wystąpienia nieprzewidzianych skutków, co wymaga wprowadzenia specjalnych zasad regulacyjnych.

Europejska regulacja sztucznej inteligencji przewiduje dla modeli ogólnego przeznaczenia wiele wymogów, które mają na celu zapewnienie ich bezpiecznego i odpowiedzialnego stosowania[30]. Ze względu na możliwość wykorzystywania modeli GPAI w wielu różnych kontekstach, kluczowe jest przewidywanie potencjalnych scenariuszy ich użycia oraz wprowadzenie mechanizmów ograniczających ryzyko. W praktyce oznacza to, że twórcy i dostawcy modeli muszą prowadzić

[29] I. Triguero, D. Molina, J. Poyatos, J. Del Ser, F. Herrera, *General Purpose Artificial Intelligence Systems (GPAIS): Properties, Definition, Taxonomy, Open Challenges and Implications*, „Information Fusion" 2024, vol. 103, s. 2-3.

[30] AI Act, rozdz. II-III.

szczegółową analizę ryzyka obejmującą różnorodne konteksty zastosowań, a także przewidywać możliwe interakcje systemu z innymi technologiami i procesami decyzyjnymi.

Wśród najważniejszych wymogów dla modeli ogólnego przeznaczenia wyróżnia się: rozszerzony nadzór człowieka, kompleksową dokumentację techniczną, szczegółową ocenę ryzyka oraz zapewnienie jakości danych[31]. Rozszerzony nadzór człowieka oznacza, że decyzje podejmowane przez modele GPAI muszą podlegać kontroli człowieka w każdym istotnym obszarze zastosowania, a mechanizmy interwencji muszą umożliwiać szybkie reagowanie na błędy lub nieprzewidziane skutki działania systemu. Dokumentacja techniczna obejmuje informacje o architekturze modelu, metodach uczenia, zbiorach danych, możliwościach adaptacyjnych, ograniczeniach systemu oraz procedurach monitorowania i reagowania na incydenty[32].

Bezpieczeństwo i przejrzystość stanowią szczególnie istotny aspekt stosowania modeli GPAI. Ich uniwersalność zwiększa ryzyko błędów lub nadużyć, dlatego dokumentacja i transparentność działania modelu są niezbędne dla zapewnienia odpowiedzialności podmiotów wdrażających technologię oraz dla ochrony użytkowników. Transparentność umożliwia audyt działania systemu, identyfikację potencjalnych błędów oraz ocenę wpływu decyzji algorytmów na prawa jednostki. W literaturze naukowej zwraca się uwagę, że brak właściwych mechanizmów nadzoru i monitorowania może prowadzić do poważnych konsekwencji prawnych, społecznych i etycznych, w tym do dyskryminacji, naruszenia prywatności lub niekontrolowanego wpływu na procesy decyzyjne[33].

Specjalne zasady dla modeli GPAI są również ściśle powiązane z innymi kluczowymi wymaganiami regulacyjnymi, takimi jak jakość danych, nadzór człowieka, zarządzanie ryzykiem i odpowiedzialność. Modele ogólnego przeznaczenia muszą być projektowane w sposób umożliwiający identyfikację i minimalizowanie ryzyka wynikającego z błędnych lub stronniczych danych, a także umożliwiać audyt i monitorowanie systemu w różnych zastosowaniach. Integracja tych zasad pozwala tworzyć systemy AI, które są nie tylko innowacyjne, ale także bezpieczne, etyczne i zgodne z wartościami europejskimi.

Zagadnienie regulacji prawnej modeli ogólnego przeznaczenia dopiero zaczyna być szerzej analizowane w literaturze przedmiotu. Badacze podkreślają, że elastyczność i uniwersalność GPAI sprawiają, że tradycyjne mechanizmy nadzoru i audytu mogą być niewystarczające. Istnieje potrzeba opracowania nowych metod oceny ryzyka, standardów dokumentacji oraz wytycznych dotyczących nadzoru człowieka,

[31] AI Act, art. 14-15.
[32] J. Laux, H. Ruschemeier, *op. cit*., s. 1525.
[33] M. Świerczyński, Z. Więckowski, *op. cit*., s. 209-211.

które uwzględniają wieloaspektowe zastosowania tych modeli[34]. Dalsze badania naukowe w tym obszarze będą kluczowe dla zrozumienia, w jaki sposób zapewnić odpowiedzialne stosowanie GPAI w dynamicznie zmieniającym się środowisku technologicznym.

Modele ogólnego przeznaczenia reprezentują obszar sztucznej inteligencji o ogromnym potencjale innowacyjnym, jak również podwyższonym poziomie ryzyka. Specjalne zasady przewidziane w AI Act – obejmujące rozszerzony nadzór człowieka, szczegółową dokumentację, ocenę ryzyka i jakość danych – mają na celu zapewnienie bezpieczeństwa, odpowiedzialności i zgodności z prawami jednostki. Skuteczne wdrożenie tych zasad wymaga ścisłej współpracy między twórcami systemów, organami regulacyjnymi i użytkownikami końcowymi, a także bieżącego monitorowania i oceny skutków stosowania GPAI. W kontekście dynamicznego rozwoju technologii dalsze badania naukowe są niezbędne do opracowania skutecznych metod nadzoru, audytu i regulacji, które pozwolą w pełni wykorzystać potencjał modeli ogólnego przeznaczenia w sposób bezpieczny, etyczny i zgodny z wartościami europejskimi.

Podsumowanie

Rozwój technologii sztucznej inteligencji stanowi jedno z najważniejszych wyzwań współczesnych systemów prawnych i społecznych. Rosnąca skala zastosowań systemów AI w różnych sektorach gospodarki oraz życia publicznego powoduje konieczność tworzenia odpowiednich mechanizmów regulacyjnych, które pozwolą na bezpieczne i odpowiedzialne wykorzystywanie technologii. W tym kontekście szczególne znaczenie ma inicjatywa legislacyjna Unii Europejskiej w postaci AI Act, której celem jest ustanowienie kompleksowych ram prawnych dla projektowania, wdrażania oraz wykorzystywania systemów sztucznej inteligencji.

Przeprowadzona analiza wskazuje, że europejski model regulacyjny opiera się przede wszystkim na podejściu opartym na analizie ryzyka, które umożliwia dostosowanie zakresu obowiązków regulacyjnych do potencjalnego wpływu danego systemu AI na prawa podstawowe jednostki, bezpieczeństwo oraz interes publiczny. Kluczowe znaczenie mają w tym zakresie zasady przejrzystości działania systemów sztucznej inteligencji, nadzoru człowieka nad procesami decyzyjnymi, odpowiedzialności podmiotów rozwijających i wdrażających technologie AI, a także wymogi dotyczące jakości danych i zarządzania ryzykiem. Zasady te mają na celu stworzenie ram umożliwiających rozwój technologii przy jednoczesnym ograniczeniu potencjalnych zagrożeń związanych z ich wykorzystaniem.

[34] P. Chyc, *Europejskie regulacje prawne dotyczące funkcjonowania sztucznej inteligencji*, „Prawo i Więź" 2025, nr 6(59), s. 207-209.

Należy jednak podkreślić, że problematyka prawnej regulacji sztucznej inteligencji oraz zasad funkcjonowania systemów AI wciąż znajduje się na stosunkowo wczesnym etapie rozwoju w literaturze przedmiotu. Ze względu na dynamiczny charakter postępu technologicznego oraz relatywnie niedawne przyjęcie europejskich regulacji prawnych, zagadnienia te dopiero zaczynają być szerzej analizowane w badaniach akademickich. W konsekwencji wiele kwestii związanych z praktycznym stosowaniem nowych regulacji, ich skutecznością oraz wpływem na rozwój innowacji pozostaje nadal otwartych i wymaga dalszych pogłębionych badań.

Można zatem stwierdzić, że europejskie podejście do regulacji sztucznej inteligencji stanowi istotny krok w kierunku budowy odpowiedzialnego i bezpiecznego systemu prawnego z zakresu nowych technologii. Jednocześnie dalszy rozwój refleksji naukowej nad zasadami regulującymi funkcjonowanie systemów AI będzie odgrywał kluczową rolę w ocenie skuteczności przyjętych rozwiązań prawnych oraz w kształtowaniu przyszłych kierunków rozwoju regulacji w tym obszarze.

Literatura

Chyc P., *Europejskie regulacje prawne dotyczące funkcjonowania sztucznej inteligencji*, „Prawo i Więź" 2025, nr 6 (59).

Doshi-Velez F., Kortz M., *Accountability of AI Under the Law: The Role of Explanation*, Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper 2017.

Ebers M., *Truly Risk-based Regulation of Artificial Intelligence: How to Implement the EU's AI Act*, „European Journal of Risk Regulation" 2024, vol. 16, is. 2.

Gutierrez C., Aguirre A., Uuk R., Boine C., Franklin M., *A Proposal for a Definition of General Purpose Artificial Intelligence Systems*, „Digital Society" 2023, vol. 2, article 36.

Kaczmarek-Templin B., *Sztuczna inteligencja (AI) i perspektywy jej wykorzystania w postępowaniu przed sądem cywilnym*, „Studia Prawnicze. Rozprawy i Materiały" 2022, nr 2(31).

Kulesza J., *Europe's Regulatory Sovereignty in the Age of Artificial Intelligence. A Human-Centric Response to the US-China Technological Rivalry*, „Sprawy Międzynarodowe" 2025, t. 78, nr 4.

Laux J., Ruschemeier H., *Automation Bias in the AI Act: On the Legal Implications of Attempting to De-Bias Human Oversight of AI*, „European Journal of Risk Regulation" 2025, nr 16.

Namysłowska M. *et al.*, *O etycznych, prawnych i społecznych konsekwencjach stosowania systemów sztucznej inteligencji w państwach członkowskich Unii Europejskiej. Uwagi na tle projektu rozporządzenia w sprawie sztucznej inteligencji*, „Przegląd Sejmowy" 2023, nr 6(179).

Pajdo A., *Zakres obowiązków podmiotów odpowiedzialnych za systemy sztucznej inteligencji wysokiego ryzyka w świetle aktu o SI – Rozporządzenia 2024/1689*, „Roczniki Nauk Prawnych" 2025, nr 2(35).

Rejmaniak R., *Bias in Artificial Intelligence Systems*, „Białostockie Studia Prawnicze" 2021, nr 3 (vol. 26).

Skoczylas D., *Sztuczna inteligencja a dobrostan człowieka i ochrona praw podstawowych – rozważania na gruncie aktu w sprawie sztucznej inteligencji*, „Studia Prawnoustrojowe" 2025, nr 1(67).

Sienkiewicz-Małyjurek K., *Możliwości i problemy zastosowania sztucznej inteligencji w zarządzaniu kryzysowym*, „Bezpieczeństwo. Teoria i Praktyka" 2024, nr 1.

Świerczyński M., Więckowski Z., *Liability for damages caused by artificial intelligence systems – main challenges to be addressed by the European Union conflict-of-laws regulations*, „Prawo w Działaniu", 2023, t. 54.

Triguero I., Molina D., Poyatos J., Del Ser J., Herrera F., *General Purpose Artificial Intelligence Systems (GPAIS): Properties, Definition, Taxonomy, Open Challenges and Implications*, „Information Fusion" 2024, vol. 103.

Vladimirova K., *The Principle of Transparency Under The Artificial Intelligence Act*, „Law Journal of New Bulgarian University" 2024, nr 2 (vol. 20).

Westerstrand S., *Fairness in AI systems development: EU AI Act compliance and beyond*, „Information and Software Technology" 2025, vol. 187.

Wyszomirska M., *Implementation of the AI Act (EU) into the Polish legal system, its impact on this system and on selected strategic sectors*, „Safety & Fire Technology" 2025, nr 2(66).

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Martyna Łaszewska-Hellrigel
Uniwersytet Zielonogórski
ORCID: 0000-0002-2212-371X
m.laszewska-hellriegel@wpa.uz.zgora.pl

Between the Algorithm's Recommendation and the Individual's Will: Two Practical Case Studies on AI in Healthcare and Elder Care

Keywords: artificial intelligence, autonomy, informed consent, refusal of treatment, elder care, dignity of risk, public administration

Summary. This article reworks the discussion on artificial intelligence in healthcare and elder care from a predominantly conceptual debate into a case-oriented legal and bioethical analysis. Its central claim is that the most important practical risks do not arise where an AI system formally replaces the human decision-maker, but where it quietly restructures the conditions under which patients, professionals, relatives and care institutions act. The article therefore analyses two model case studies grounded in real and recurring patterns of practice: first, the refusal of life-sustaining treatment in the presence of an algorithmic recommendation for escalation of therapy; second, the gradual intensification of AI-supported home monitoring of an older person in the name of safety. In both cases, the decisive issue is not whether the technology is technically accurate, but whether legal and administrative arrangements preserve meaningful autonomy, contestability, proportionality and responsibility. The argument proceeds by connecting practical case analysis with medical-law principles concerning competent refusal, informed consent, dignity, privacy and the continuing relevance of the dignity of risk. The article concludes that a rights-sensitive use of AI requires procedural safeguards embedded in institutional design: functional transparency, documented dialogue, renewable consent, protected deviation from automated outputs, and accessible routes of contestation.

Paternalizm algorytmiczny czy upodmiotowiona autonomia? Teoretycznoprawne refleksje na temat sztucznej inteligencji w opiece nad osobami starszymi i problematyce zgody

Słowa kluczowe: sztuczna inteligencja, autonomia, świadoma zgoda, odmowa leczenia, opieka nad osobami starszymi, godność ryzyka, administracja publiczna

Streszczenie. W miarę jak systemy sztucznej inteligencji zyskują na znaczeniu w ochronie zdrowia i usługach społecznych, ich zastosowanie w opiece nad osobami starszymi skłania do ponownego podjęcia teoretycznoprawnej debaty na temat autonomii, świadomej zgody oraz ewoluującej roli prawa w regulowaniu wspomaganego maszynowo podejmowania decyzji. Artykuł analizuje normatywne implikacje narzędzi algorytmicznych stosowanych w opiece nad osobami w podeszłym wieku, ze szczególnym uwzględnieniem napięcia między ochroną autonomii jednostki a promowaniem jej dobrostanu za pośrednictwem systemów zautomatyzowanych.

Koncentrując się na środkowoeuropejskich porządkach prawnych – w tym Polski i Słowacji – w szerszym kontekście prawa unijnego, artykuł analizuje, w jaki sposób oparte na AI systemy monitorowania, triaż i rekomendacji mogą wpływać na zdolność prawną, zastępcze podejmowanie decyzji oraz wykonywanie woli przez osoby starsze. Szczególną uwagę poświęcono modelom zgody w zakresie dawstwa narządów (*opt-in* i *opt-out*) oraz dyrektywom opiekuńczym, badając, czy algorytmicznie kształtowane rozwiązania domyślne osłabiają, czy też wspierają rzeczywisty wybór.

Artykuł odwołuje się do prawa porównawczego, bioetyki oraz teorii autonomii relacyjnej, aby ocenić, kiedy AI stanowi formę paternalizmu, a kiedy może służyć jako narzędzie prawnego upodmiotowienia. Kończy się rekomendacjami dotyczącymi włączenia zasad przejrzystości, proporcjonalności i kontroli użytkownika do standardów prawnych regulujących algorytmiczne podejmowanie decyzji w opiece nad osobami starszymi, postulując wyważoną równowagę między efektywnością technologiczną a poszanowaniem godności człowieka.

1. Introduction

Artificial intelligence now appears in healthcare and care administration not as a distant prospect, but as part of the ordinary infrastructure of decision-making. Risk prediction tools, triage systems, diagnostic support interfaces, remote monitoring solutions, care-navigation applications and behaviour-detection systems are increasingly presented as instruments of assistance. Their promise is familiar: they are said to improve accuracy, reduce human error, allocate scarce resources more efficiently and make care more continuous, anticipatory and personalised. Yet, the legal and bioethical challenge begins precisely at the point where the language of assistance sounds most reassuring. An instrument can assist without being neutral. It can support a decision while also shaping the field in which that decision is made. In contemporary data-rich environments, that shaping power becomes especially strong because the system does not merely display information; it sorts, highlights, ranks, alerts, prompts and normalises[1].

For that reason, the central question is not simply whether AI 'decides'. In many practical settings, it does not decide in a formal sense. A physician still signs the order, a patient still gives or refuses consent, a family member still approves the installation of a monitoring device, and a care institution still adopts the policy. But between formal decision-making power and real normative influence, there is a wide and legally important space. The architecture of options can be curated in advance; what counts as prudent, safe, explainable or professionally defensible can be pre-structured; and the burdens of justification can be redistributed so that deviating from the machine's recommendation becomes institutionally costly. This is why current debates on autonomy cannot stop at the classical image of a face-

[1] K. Yeung, *'Hypernudg": Big Data as a mode of regulation by design, Information*, "Communication & Society" 2017, vol. 20(1), pp. 118-136.

to-face encounter between two human agents. Autonomy must also be analysed at the level of design, workflow and governance[2].

This article proceeds from the view that case-based analysis is particularly useful for this task. General discussions about 'autonomy', 'fairness' or 'human oversight' are important, but they often remain too abstract to show how pressure is actually generated in practice. The crucial transformations occur in concrete situations: when a competent patient refuses a beneficial intervention; when the record is written in a way that casts refusal as irrational; when a nurse feels unable to contradict a risk score; when an older person who initially agreed to a simple alert system gradually becomes the object of continuous surveillance; or when relatives invoke safety to legitimise a level of control that alters the meaning of home itself. In those moments, law and bioethics must move from slogans to criteria[3].

The article therefore centres on two practical case studies. They are modelled cases rather than reports of one single lawsuit, but each is firmly rooted in patterns that are already visible in medical practice, elder care, scholarship on automation bias and the growing literature on digital surveillance in care settings. The first case concerns a competent patient's refusal of invasive treatment in circumstances where an AI-supported clinical system strongly recommends escalation. The second concerns an older person living at home whose initially limited monitoring arrangement is repeatedly expanded in the name of safety. Both cases are analytically productive because they involve a tension between will and welfare, but they do so in different institutional forms. In the first, the conflict appears acutely and dramatically at the bedside. In the second, it unfolds gradually through a series of apparently small administrative and technical adjustments.

The argument developed below is straightforward but demanding in its implications. AI does not threaten autonomy, only when it openly replaces human judgement. It also threatens autonomy when it silently alters the practical conditions of choice, when it privileges one understanding of rationality over others, and when institutions organise accountability in ways that reward compliance with the system and penalise justified resistance. From that perspective, the most important legal task is not merely to ask whether AI may be used, but under what procedural conditions its use remains compatible with the rights, dignity and agency of the person whose life is being governed by the system.

---

[2] R.H. Thaler, C.R. Sunstein, *Libertarian Paternalism*, "American Economic Review" 2003, vol. 93(2), pp. 175-179; K. Yeung, *'Hypernudge': Big Data as a mode of regulation*…, pp. 118-136.

[3] O'Neill's critique remains relevant here because formal consent may coexist with structurally weak autonomy: O.O'Neill, *Autonomy and Trust in Bioethics*, Cambridge 2002, p. 49.

2. Why a Case-oriented Approach is Needed

A case-oriented approach is valuable because AI governance in healthcare and elder care is rarely exhausted by questions of technical performance. A prediction can be statistically impressive and still produce a legally distorted encounter. A monitoring system can detect risk accurately and still damage privacy, agency and trust. Conversely, a system may be imperfect yet remain acceptable if its institutional use preserves contestability, limits intrusiveness and leaves room for humanly intelligible reasoning. The law therefore needs categories that are sensitive not only to outcomes, but also to the way in which outcomes are reached.

The first conceptual point concerns autonomy. In medical law, respect for autonomy has never meant that a competent person must choose what others consider objectively best. It means, rather, that a person who can understand relevant information, deliberate and communicate a decision retains authority over certain deeply personal choices, including choices that others regard as mistaken or dangerous. The problem introduced by AI is that systems built on statistical optimisation can subtly recode this relationship. Once a system defines the medically recommended path with unusual confidence, refusal may cease to appear as one legitimate option, among others, and instead come to be treated as a deviation from the rational baseline[4].

The second point concerns rights theory. A will-based understanding of rights highlights the importance of normative control: the point of a right is that the right-holder can decide whether to authorise, refuse, waive or insist. An interest-based understanding emphasises the protection of welfare, well-being and objective interests. Both perspectives illuminate AI-supported care. If one looks only through the lens of welfare, algorithmic steering may seem justified whenever it improves measurable outcomes. If one looks only through the lens of will, one may fail to appreciate genuine vulnerability, dependency and cognitive burden. The practical challenge is therefore not to choose one theory once and for all, but to prevent welfare-oriented design from quietly erasing the person's normative authority[5].

The third point concerns rationality. Healthcare law traditionally protects competent decisions even when they are not medically prudent, because competence is not identical with substantive wisdom. Yet, AI systems are built precisely to predict, compare and optimise. They derive authority from the capacity to identify the option most strongly associated with survival, adherence, reduced cost or re-

[4] P.S. Appelbaum, *Assessment of Patients' Competence to Consent to Treatment*, "New England Journal of Medicine" 2007, vol. 357(18), pp. 1834-1840.

[5] J. Raz, *Legal Rights*, "Oxford Journal of Legal Studies" 1984, vol. 4(1), pp. 1-21.

duced risk. That makes them powerful tools, but also potentially paternalistic ones. Once optimisation criteria are embedded in workflow, rationality can migrate from a procedural notion-did the person understand and decide? To a substantive one-did the person choose what the model identified as the most beneficial option? [6]

Case studies expose this migration in a way that abstract principle cannot. They show where the recommendation appears, who sees it, how it is documented, how dissent is framed, who bears the burden of explanation and what institutional incentives are at work. They thus make visible an often neglected truth: many rights violations in technologically mediated care are not produced by a dramatic single act, but by the cumulative interaction of interface design, professional psychology, family anxiety and organisational routines.

3. Case Study One: Refusal of Treatment Against an Algorithmic Recommendation

3.1. Factual Structure of the Case

A patient with severe respiratory failure is admitted to hospital. He is alert, oriented and able to communicate consistently. He is informed that his condition may deteriorate and that invasive ventilation would likely prolong life. The hospital uses an AI-supported clinical deterioration system trained on prior patient data. On the basis of physiological variables and comparable trajectories, the system assigns the patient a very high-risk score and generates a prominent recommendation that escalation of treatment should not be delayed. The treating team, already under pressure because intensive-care beds are limited and delayed decisions are closely reviewed, receives repeated alerts. The patient nevertheless states, calmly and repeatedly, that he does not want intubation. He accepts oxygen, symptom relief and further conversation, but refuses invasive ventilation.

At first glance, the legal structure of the case appears familiar. It resembles the classical problem of a competent refusal of life-sustaining treatment. Yet, the presence of the algorithm changes the normative atmosphere. The physician is no longer merely dealing with a clinical disagreement or a tragic conflict between prognosis and patient will. The refusal now stands against a quantified, repeatable and institutionally visible recommendation produced by a system marketed as reducing error and delay. The machine does not have legal personality and does

[6] On the risk that digitally mediated environments turn influence into design-based steering, see K. Yeung, *'Hypernudge': Big Data as a mode of regulation by design, Information,* "Communication & Society" 2017, vol. 20(1), pp. 118-136.

not consent on anyone's behalf. Nevertheless, it exerts pressure by changing what counts as responsible action inside the organisation.

The most significant feature of the case is therefore not the alert itself but the redistribution of justificatory burden. Before the introduction of the system, the patient's refusal had to be taken seriously because the clinician needed positive reasons to override it, and those reasons would normally be unavailable where capacity is intact. After the introduction of the system, however, the clinician who accepts the refusal may feel required to justify not only her own judgement but also her departure from the algorithmic output. The patient's will, while formally unchanged, becomes practically harder to respect.

3.2. Legal Analysis

Medical law has long treated competent refusal as a core expression of bodily autonomy. The underlying principle is not that refusal is always wise, but that invasive treatment without valid consent constitutes a profound interference with bodily integrity. Capacity is thus the threshold question, not agreement with medical advice. The standard view in contemporary medical law is that an adult patient with decision-making capacity may refuse treatment even where the likely consequence is death. That proposition would lose much of its meaning if technological systems could indirectly downgrade the refusal by branding it as statistically irrational or institutionally non-compliant[7].

The case law on refusal and consent confirms the same orientation. In Re B, the English court recognised that a competent adult may refuse life-sustaining treatment even when others consider the decision tragic or misguided. Later jurisprudence on informed consent, including Montgomery, reinforced the broader move away from professional paternalism by insisting that the patient is not merely the object of medical expertise but a legal agent entitled to information material to her own decision. These authorities are not about AI, but they become highly relevant once algorithmic recommendations begin to alter the informational and institutional environment in which consent and refusal are recorded[8].

From this perspective, the legal problem is not whether the model may be consulted. It usually may. The problem is whether consultation quietly becomes normative displacement. Several pathways of displacement deserve particular attention. First, documentation may subtly change its vocabulary: a refusal may be

[7] P.S. Appelbaum, *Assessment of Patients' Competence to Consent to Treatment*, "New England Journal of Medicine" 2007, vol. 357(18), pp. 1834-1840.

[8] Montgomery v Lanarkshire Health Board [2015] UKSC 11; Re B (Adult: Refusal of Medical Treatment) [2002] 2 All ER 449; see also M. Stauch, *Comment on Re B (Adult: Refusal of Medical Treatment) [2002] 2 All England Reports 449*, "Journal of Medical Ethics" 2002, vol. 28(4), pp. 232-233.

described less as an autonomous choice and more as non-compliance with the recommended pathway. Secondly, professional discussions may begin from the assumption that the system's output expresses the most responsible option, so that respect for refusal appears as an exception. Thirdly, retrospective review mechanisms may reward adherence to the alert and expose deviation to internal scrutiny. Under those conditions, the right formally survives while its exercise becomes structurally disfavoured.

A further issue is the quality of information disclosed to the patient. If the clinician tells the patient simply that 'the computer says your risk is very high', the patient may be exposed to a particularly authoritative form of persuasion without receiving an intelligible account of what the system actually does, what kind of data it relies on, how confident the output is, or whether comparable clinicians sometimes disagree with it. The informational asymmetry becomes deeper because the aura of objectivity attached to the model may be greater than the patient's ability to interrogate it. Respect for autonomy in such conditions requires more than passing on the alert. It requires translating the role of the system into terms the patient can meaningfully use.

Legally, then, the crucial point is that capacity, consent and refusal remain person-centred categories. They do not collapse into algorithmic optimisation. A hospital that uses AI in a manner that effectively penalises clinicians for honouring a competent refusal risks reintroducing paternalism in technologically updated form. It also risks blurring responsibility: if the intervention proceeds against the person's will, the relevant legal wrong cannot be neutralised by the fact that the system recommended it.

3.3. Bioethical Analysis

The bioethical difficulty of the case lies in the tension between welfare and authority. The system may indeed predict that ventilation would prolong life; the clinicians may sincerely believe that treatment is in the patient's best interests; relatives may be distressed by the prospect of a refusal. None of that is ethically irrelevant. But beneficence does not erase agency. Respect for persons includes accepting that a competent individual may evaluate burdens, dependence, fear, invasiveness, future quality of life and personal meaning differently from the clinical team. To insist that the medically beneficial option must prevail is to replace moral dialogue with a narrower idea of optimisation[9].

[9] For a concise account of the tension between autonomy, individuality and consent, see O.O'Neill, *Autonomy and Trust in Bioethic*s, Cambridge 2002, p. 40.

This is where the notion of trust becomes decisive. Trust in healthcare is not strengthened when the professional appears merely to transmit an opaque recommendation backed by institutional pressure. It is strengthened when the patient can see that technology is being used as an aid to conversation rather than as a device that settles the matter in advance. O'Neill's critique of thin and over-individualised conceptions of autonomy remains useful here: autonomy cannot be secured by ritualised consent talk if the surrounding structure renders disagreement practically futile[10].

The case also illuminates a familiar cognitive danger. Human operators often over-rely on automated systems, especially when those systems are presented as safety-enhancing and when institutional cultures treat them as markers of diligence. Research on human-automation interaction has repeatedly shown that reliance on automation can generate both omission and commission errors: people may fail to act when the system does not prompt them, and they may follow the prompt, even when independent judgement should lead elsewhere. In the clinical setting, that dynamic does not merely threaten technical accuracy. It can also transform the ethical stance of the clinician by making resistance appear reckless[11].

Accordingly, the ethical failure in this case would not consist only in physically imposing treatment. It could already occur earlier, at the point where the conversation becomes subtly coercive, where refusal is repeatedly reclassified as confusion, where documentation narrows the patient's reasons into a problem of non-adherence, or where clinicians are made to feel that honouring refusal is professionally indefensible. AI magnifies these risks because it gives optimisation a visible interface and an administrative trace.

3.4. Administrative Implications

For public administration and institutional governance, the first case demonstrates that human oversight cannot be reduced to the mere presence of a human final signatory. Oversight has to be organised. A clinician must be able to depart from the model's recommendation without entering a zone of institutional suspicion. Documentation systems should therefore record not only that an alert was generated, but also that competent refusal remains legally decisive and that deviation from the algorithmic pathway may be justified by patient choice. Hospitals that

[10] O.O'Neill, *Autonomy and…*, pp. 44.

[11] R. Parasuraman, V. Riley, *Humans and Automation: Use, Misuse, Disuse, Abuse*, "Human Factors" 1997, vol. 39(2), pp. 230-253; K. Goddard, A. Roudsari, J.C. Wyatt, *Automation bias - a hidden issue for clinical decision support system use*, "Studies in Health Technology and Informatics" 2011, vol. 164, pp. 17-22; L.J. Skitka, K.L. Mosier, M. Burdick, *Does automation bias decision-making?*, "International Journal of Human-Computer Studies" 1999, vol. 51(5), pp. 991-1006.

fail to build such space into workflow risk producing a de facto hierarchy in which the model outranks the patient.

The case also shows why the language of explanation matters. Institutions should not confine themselves to broad statements that AI is 'used to support decisions'. In disputes about consent and refusal, the patient and the record should reflect the actual function of the system: whether it predicts deterioration, whether it recommends a threshold for escalation, whether it influences triage order, and whether clinicians are expected to document disagreement. Functional transparency is a precondition of meaningful contestation.

4. Case Study Two: AI-supported Home Monitoring and the Gradual Expansion of Control

4.1. Factual Structure of the Case

An older woman lives alone in her own home. She has early cognitive impairment but remains able to manage most daily activities and clearly states that remaining at home matters deeply to her. Her daughter, worried after two minor incidents of disorientation, arranges a home-monitoring package through a care provider. At the outset, the system is modest: a door sensor, medication reminders and an alert if no movement is detected for an unusual period. The older woman agrees, partly because she wants to reassure her family and avoid institutionalisation.

Over the next months, the arrangement changes. The provider introduces predictive analytics that classify patterns of activity as normal or abnormal. Nighttime wandering alerts are activated. Audio prompts are added. Then the daughter requests camera-based monitoring in the hallway, arguing that she needs to know whether her mother has fallen. The provider explains that the upgraded package is safer and easier to manage. The older woman expresses discomfort, saying that she feels watched, but is told that the changes are necessary because her risk profile has worsened. Nothing dramatic happens in a single step. Rather, a series of individually defensible adjustments gradually transforms the home into a monitored environment governed by risk detection.

This second case differs from the first in tempo and tone. There is no single bedside refusal and no immediate life-or-death confrontation. Yet, the normative stakes are no less serious. What is at issue is whether safety can become a solvent that dissolves privacy, spontaneity and the practical authority of the older person over the terms of her own everyday life. The case is especially important because the justifications for intervention are plausible, emotionally compelling and often motivated by genuine concern.

4.2. Legal Analysis

The legal significance of the case lies first in the continuing relevance of consent, and secondly in the limits of treating an initial consent as a blanket authorisation for future escalation. Agreement to a basic assistive arrangement does not automatically legitimise every later increase in intrusiveness. When the technology changes function-moving from reminders to behavioural inference, from passive sensing to camera-based observation, or from emergency response to continuous pattern analysis-the normative object of consent also changes. The person's earlier agreement cannot simply be stretched to cover the new system.

Data protection law helps to clarify this point, even though the issue is broader than data governance alone. Principles such as purpose limitation, data minimisation and proportionality point against the easy expansion of surveillance whenever more data become technically available. In a home-care context, moreover, personal data are inseparable from lived experience. The hallway camera does not merely 'process data'; it modifies the conditions of domestic life, alters the sense of being alone, and changes the relation between dependence and intimacy. Legal analysis must therefore avoid the reductionist error of treating surveillance only as an information-management problem[12].

The case also raises an important issue of capacity and vulnerability. Early cognitive impairment may justify supportive communication, repeated discussion and tailored safeguards. It does not justify the assumption that the person's present preferences have lost normative significance. Too often, systems of care slide from 'supporting decision-making' to 'pre-empting objection'. The mere invocation of safety is insufficient to justify that slide. If the person understands the relevant change well enough to express resistance to more intrusive monitoring, that resistance must remain a legally meaningful fact.

Finally, the institutional role of family members and providers requires scrutiny. Relatives frequently act out of fear, guilt and responsibility. Care providers often operate in environments where adverse events trigger blame and where surveillance appears as a defensible risk-management tool. The law should recognise those pressures without letting them displace the subject of care herself. Otherwise, home care becomes an administrative arrangement negotiated around the older person rather than with her.

[12] European Parliament Regulation (EU) 2016/679 Council of 27 April 2016 (General Data Protection Regulation).

4.3. Bioethical Analysis

The bioethical core of the second case is captured by the idea of the dignity of risk. The phrase is old, but its relevance has grown rather than diminished. Perske's classic insight was that overprotection can degrade persons by depriving them of the chance to make choices that carry some measure of risk. Contemporary debates in elder care have restated the same thought in rights-based terms: if all significant risk is designed out of life, what remains may be physically safer but existentially diminished[13][14].

AI surveillance technologies intensify this dilemma because they permit an unprecedented continuity of observation. The question is no longer only whether a care worker checks on someone too often. It is whether the background environment itself becomes interpretive and interventionist-constantly sorting behaviour into risk categories, generating alerts, prompting relatives to intervene and nudging the person toward pre-approved patterns of conduct. Such systems can be defended as benign because they often operate quietly. That quietness is precisely what makes them ethically powerful.

The language of safety should therefore be handled with care. Safety is not a value that automatically trumps all others. In the life of an older person who wishes to remain at home, values such as privacy, routine, intimacy, solitude, unpredictability, self-presentation and the ordinary right not to be constantly assessed may be central to well-being. A morally serious analysis cannot define welfare in purely actuarial terms. Joseph's recent argument that elder-care technologies should begin from a commitment to dignity rather than mere efficiency is especially apt in this respect[15].

The case also reveals the ethical inadequacy of one-off consent. In dynamic monitoring environments, the person does not confront a single intervention but an evolving regime. Ethical legitimacy must therefore be renewable. Every increase in intrusiveness changes the balance between protection and self-determination. To insist on renewed discussion is not bureaucratic formalism; it is the practical expression of respect for agency under conditions of technological drift.

---

[13] M. Foundas, *Dignity of risk in residential aged care: a call to reframe understandings of risk*, "Medical Journal of Australia" 2025, vol. 223(4), pp. 187-188.

[14] R. Perske, *The Dignity of Risk and the Mentally Retarded*, "Mental Retardation" 1972, vol. 10(1), pp. 24-27.

[15] J. Joseph, *Designing for dignity: ethics of AI surveillance in older adult care, Frontiers in Digital Health* 2025, vol. 7, article 1643238.

4.4. Administrative Implications

Administratively, the second case shows that care providers need escalation rules tied to proportionality rather than to mere technical capability. The fact that the system can infer more, store more or show more does not answer whether it should. Providers should therefore separate emergency functions from lifestyle surveillance, distinguish fall detection from behaviour scoring, and document why a less intrusive option is insufficient before migrating to a more invasive one. Without such discipline, AI-supported home care becomes vulnerable to what may be called incremental normalisation: each addition looks minor, but together they produce a different moral world.

Equally important is the availability of contestation. Older persons and their relatives should know who has authority to change the settings, how those changes are reviewed, and how disagreement is handled. This is where public-administration concerns and private-care practices meet. Even where care is delivered through mixed or private arrangements, the governance logic of rights, proportionality and review remains indispensable. Otherwise, the most consequential decisions are hidden inside procurement choices, service packages and dashboard settings rather than treated as matters of rights-bearing significance.

5. Cross-case Lessons for Law and Public Administration

The two cases are institutionally distinct, yet they converge on a single lesson: the central legal question is not whether AI is useful, but whether the surrounding procedure preserves meaningful human agency. In the first case, the threat arises from the authority of optimisation in a high-stakes clinical encounter. In the second, it arises from the slow sedimentation of surveillance into everyday care. But in both, the decisive danger is the same. The person remains formally present as decision-maker while the substantive terms of the decision are increasingly organised elsewhere.

Several safeguards follow from this insight. The first is functional transparency. Individuals should be told, in intelligible terms, what role the system plays in the relevant decision. Not every person needs a technical explanation of model architecture; what matters is whether the person can understand the practical function of the system in the pathway that affects her. In the treatment-refusal case, that means explaining that the system predicts deterioration and prompts escalation. In the home-monitoring case, it means explaining whether the technology merely alerts after a fall or also classifies patterns of behaviour. Functional opacity is incompatible with meaningful self-determination.

The second safeguard is documented dialogue rather than documentation of compliance. Records should show that the person's reasons were heard, not simply that the recommended pathway was communicated. In clinical settings, this means recording the patient's values, burdens and expressed priorities, not only the risk score. In care settings, it means documenting objections to monitoring changes and the reasons why less intrusive alternatives were considered sufficient or insufficient. Documentation that only mirrors the system's categories distorts the normative landscape.

The third safeguard is renewable consent tied to changes in intrusiveness. This is particularly important in elder care, but it also matters in healthcare whenever tools are repurposed or extended. A person may agree to one form of assistance and reasonably object to another. Institutions should therefore treat material functional expansion as a new normative event, rather than as a mere update of the same service. This point aligns both with rights-sensitive care ethics and with legal principles that resist overbroad consent[16].

The fourth safeguard is protected deviation from automated outputs. Human oversight is real only if a professional can disagree without triggering a presumption of negligence. Where every deviation from the alert requires exceptional justification, the system effectively governs. Healthcare institutions and care providers should instead normalise reasoned departure. The requirement should be explanation, not obedience. This approach is also more consistent with current European regulation, which increasingly treats human oversight, accountability and risk management as design obligations rather than rhetorical aspirations[17].

The fifth safeguard is accessible contestation. Persons affected by AI-supported decisions need a route by which they can object, seek review and ask for reconfiguration. In public administration, procedural justice depends not only on the initial decision but also on the availability of challenge. In technologically mediated care, this means that rights should not evaporate into software settings. If a patient objects to the algorithmic framing of refusal, or an older person objects to camera-based monitoring, there must be a humanly accountable place to take that objection.

These safeguards do not amount to hostility toward innovation. On the contrary, they identify the conditions under which innovation can be integrated without hollowing out the legal subject. The choice is not between technology and autonomy. It is between technologies embedded in procedures that respect contestable

[16] Regulation (EU) 2016/679 of the European Parliament; Council General Data Protection Regulation of 27 April 2016.

[17] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (AI Act).

agency and technologies embedded in procedures that quietly substitute optimisation for judgement.

6. Conclusion

The current regulatory and ethical discussion on AI often asks whether humans remain 'in the loop'. The two case studies analysed in this article show why that formula is insufficient. A human may remain formally in the loop while the practical meaning of choice is nonetheless thinned out by design, workflow and institutional pressure. In the treatment-refusal case, the danger lies in converting statistical recommendation into a new baseline of clinical reasonableness that marginalises competent refusal. In the home-monitoring case, the danger lies in letting safety rationales gradually redefine home as a site of continuous behavioural supervision.

The deeper lesson is that autonomy in technologically mediated care is a procedural achievement. It depends on how information is framed, how reasons are recorded, how dissent is treated, whether consent is revisited when systems expand, and whether professionals are genuinely permitted to depart from the machine's pathway. Legal and administrative analysis should therefore resist the temptation to measure legitimacy by accuracy alone. A highly accurate system can still produce a legally and morally defective practice if it reorganises agency in the wrong way.

For a journal concerned with law and administration, that conclusion has a practical consequence. The task is not merely to classify AI as permissible or impermissible in the abstract. The task is to govern concrete decision environments. When practical case analysis is taken seriously, the relevant question becomes clearer: not whether AI can assist care, but whether institutions are willing to design assistance in a way that leaves the person whose life is at stake recognisable as an author, not merely an input.

References

Appelbaum P.S., *Assessment of Patients' Competence to Consent to Treatment*, "New England Journal of Medicine" 2007, vol. 357(18), pp. 1834-1840.

Foundas M., *Dignity of risk in residential aged care: a call to reframe understandings of risk*, "Medical Journal of Australia" 2025, vol. 223(4), pp. 187-188.

Goddard K., Roudsari A., Wyatt J.C., *Automation bias - a hidden issue for clinical decision support system use*, "Studies in Health Technology and Informatics" 2011, vol. 164, pp. 17-22.

Joseph J., *Designing for dignity: ethics of AI surveillance in older adult care*, "Frontiers in Digital Health" 2025, vol. 7, article 1643238.

O'Neill O., *Autonomy and Trust in Bioethics*, Cambridge 2002.

Parasuraman R., Riley V., *Humans and Automation: Use, Misuse, Disuse, Abuse*, "Human Factors" 1997, vol. 39(2), pp. 230-253.

Perske R., *The Dignity of Risk and the Mentally Retarded*, "Mental Retardation" 1972, vol. 10(1), pp. 24-27.
Raz J., *Legal Rights*, Oxford Journal of Legal Studies 1984, vol. 4(1), pp. 1-21.
Skitka L.J., Mosier K.L., Burdick M., *Does automation bias decision-making?* "International Journal of Human-Computer Studies" 1999, vol. 51(5), pp. 991-1006.
Stauch M., *Comment on Re B (Adult: Refusal of Medical Treatment) [2002] 2 All England Reports 449,* "Journal of Medical Ethics" 2002, vol. 28(4), pp. 232-233.
Thaler R.H., Sunstein C.R., *Libertarian Paternalism*, "American Economic Review" 2003, vol. 93(2), pp. 175-179.
Yeung K., *'Hypernudg": Big Data as a mode of regulation by design,* "Information, Communication & Society" 2017, vol. 20(1), pp. 118-136.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Petr Mrkývka
Uniwersytet Masaryka Brno
ORCID: 0000-0001-7954-3603
Petr.Mrkyvka@law.muni.cz

Dana Šramková
Uniwersytet Masaryka Brno
ORCID: 0000-0002-5717-7507
Dana.Sramkova@law.muni.cz

Michal Janovec
Uniwersytet Masaryka Brno
ORCID: 0000-0002-6565-4043
Michal.Janovec@law.muni.cz

Ewa Szewczyk
Uniwersytet Zielonogórski
ORCID: 0000-0002-2980-6564
e.szewczyk@wpa.uz.zgora.pl

AI in the Czech Public Administration, Selected Topics

Keywords: artificial intelligence, public administration

Summary. The implementation of artificial intelligence (AI) is undoubtedly a growing trend in the public sector. This is because the public sector has access to a vast amount of data. For this reason alone, the potential for using AI is significantly higher here than in the private sector. It is fair to say that artificial intelligence 'feeds' on data, and its task is to process it at a speed beyond human capability. It is therefore clear that its development in this area is inevitable. Issues related to cost-effectiveness, efficiency and effectiveness mean that technological innovations driven by artificial intelligence systems are appearing with increasing frequency in public administration. Incidentally, it is precisely this kind of administration – that is, a modern administration – that contemporary society expects. There is no doubt that artificial intelligence algorithms can be successfully used to automate routine administrative processes. This article aims to provide the reader with an overview of the state of advancement of the Czech public administration in the use of groundbreaking technologies.

The introduction of generative artificial intelligence into the Czech public administration is intended to help streamline decision-making, optimise internal processes, improve accessibility and simplify bureaucratic procedures. In this regard, the Czech Republic is focusing on the development of e-government services, such as the eDoklady system – digital documents on a mobile phone. Four public sectors in particular appear to be ideal candidates for experimenting with artificial intelligence: the judiciary, healthcare, education, and infrastructure and municipal services[1].

[1] F. Křepelka, *Artificial intelligence in the Czech government: political and economic hype, bureaucratic paralysis, and demarcation with the European Union,* Italian Journal of Public Law, vol. 17, issue 2/2025, p. 603.

Development in this area must take security issues into account, particularly data security. For this reason, in July 2025, following warnings from the National Cyber and Information Security Agency (NUKIB), the Czech government banned the use of services provided by the Chinese company DeepSeek AI in public administration, citing data security risks as well as concerns over potential surveillance by Beijing. Czech experts have highlighted the risk of unauthorised access to user data, as DeepSeek – which offers a large AI-based language model – is obliged, under Chinese law, to cooperate with the authorities there[2].

AI w czeskiej administracji publicznej, zagadnienia wybrane

Słowa kluczowe: sztuczna inteligencja, administracja publiczna

Streszczenie. Wdrażanie sztucznej inteligencji (Artificial Intelligence, AI) jest niewątpliwie rosnącym trendem w sferze publicznej. Sektor publiczny dysponuje bowiem olbrzymią ilością danych. Już chociażby z tego powodu potencjał wykorzystania AI jest tu znacznie wyższy niż w sferze prywatnej. Zasadnie można stwierdzić, że sztuczna inteligencja „karmi się" danymi, a jej zadaniem jest ich przetwarzanie, w tempie niedostępnym dla człowieka. Oczywistym więc jest, że jej rozwój w tym obszarze jest nieuchronny. Kwestie związane chociażby z oszczędnością oraz wydajnością czy skutecznością powodują, że innowacje technologiczne napędzane przez systemy sztucznej inteligencji pojawiają się coraz częściej w administracji publicznej. Nawiasem mówiąc takiej właśnie administracji – czyli administracji nowoczesnej – oczekuje współczesne społeczeństwo. Nie ulega wątpliwości, że algorytmy sztucznej inteligencji mogą być z powodzeniem wykorzystywane do automatyzacji rutynowych procesów administracyjnych. Artykuł ma na celu przybliżenie czytelnikowi stanu zaawansowania czeskiej administracji publicznej w korzystaniu z przełomowych technologii.

Wprowadzenie generatywnej sztucznej inteligencji do czeskiej administracji publicznej ma pomóc usprawnić proces decyzyjny, zoptymalizować procesy wewnętrzne, poprawić dostępność i uprościć procedury biurokratyczne. W tym zakresie Czechy stawiają na rozwój usług e-government, takich jak system eDoklady czyli cyfrowe dokumenty w telefonie komórkowym. Zwłaszcza cztery sektory publiczne wydają się być szczególnie idealnymi kandydatami do prowadzenia eksperymentów z wykorzystaniem sztucznej inteligencji: sądownictwo, służba zdrowia, edukacja oraz infrastruktura i usługi komunalne[3].

Rozwój w tym obszarze musi uwzględniać kwestie bezpieczeństwa, zwłaszcza bezpieczeństwa danych. Z tego względu w lipcu 2025 r. czeski rząd, po ostrzeżeniach Krajowej Agencji Bezpieczeństwa Cybernetycznego i Informatycznego (NUKIB), zakazał korzystania z usług chińskiej firmy DeepSeek AI w administracji publicznej, powołując się na zagrożenia dla bezpieczeństwa danych, jak również w obawie o potencjalny nadzór ze strony Pekinu. Czescy eksperci wskazali na ryzyko nieautoryzowanego dostępu do danych użytkowników, ponieważ firma DeepSeek - oferująca duży model językowy oparty na sztucznej inteligencji - zgodnie z chińskimi przepisamiprawnymi, jest zobowiązana do współpracy z tamtejszymi organami państwowymi[4].

[2] https://pl.euronews.com/next/2025/07/10/czechy-blokuja-sztuczna-inteligencje-w-administracji-publicznej, dostęp do chińskiego chatbota zablokowały także inne kraje – między innymi Włochy oraz Australia, [accessed on: 01.04.2026].

[3] F. Křepelka, *Artificial intelligence in the Czech government: political and economic hype, bureaucratic paralysis, and demarcation with the European Union,* Italian Journal of Public Law, vol. 17, issue 2/2025, p. 603.

[4] https://pl.euronews.com/next/2025/07/10/czechy-blokuja-sztuczna-inteligencje-w-administracji-publicznej, dostęp do chińskiego chatbota zablokowały także inne kraje – między innymi Włochy oraz Australia [accesseed on: 04.01.2026].

The Etymology of the Word 'Robot' – The Czech Contribution to GlobalTechnological Vocabulary

Even before the term 'artificial intelligence' had taken hold, the word 'robot' – which had appeared several decades earlier and rapidly gained popularity as a Europeanism – was probably the first term associated with the field of automated yet intelligent operation[5]. Czech artists made a particular contribution in this field. The whole world owes them the introduction of the word "robot" into common usage on a transnational scale. The word originates from the Czech language and means hard labour, including forced labour. It came into wider circulation following the translation of Karel Čapek's play R.U.R., staged in Prague in 1920[6]. Josef Čapek derived the word from the noun 'robota', meaning hard labour, even drudgery. This classic neologism, typical of the creators of that era, would probably have been forgotten had it not been for the poet's brother, Karel, who used it in his utopian comedy, the science-fiction novel 'R.U.R.' (full title: 'Rossumovi Univerzální Roboti'), to name the humanoid machines featured in it. In "R.U.R.", robots are beings made of laboratory tissue, highly intelligent, resembling humans in appearance, yet devoid of emotions and needs. Over time, the word 'robot' came to refer to mechanical devices. The English translation of this novel (Rossum's Universal Robots) made it famous worldwide, particularly through theatrical productions[7].

Reasons for Using Artificial Intelligence in Public Administration

Regardless of the country in which artificial intelligence is used in public administration, the underlying motivations are similar. These include: increasing efficiency (*e.g.,* because artificial intelligence can process information much faster than humans), improving decision-making and policymaking (*e.g.,* a higher level of accuracy/consistency may be achievable with automated rather than human decision-making), giving officials more time for work requiring creativity – rather than performing routine tasks (*e.g.,* simple, repetitive tasks can be carried out by

[5] The term "artificial intelligence" was coined during a workshop held in 1956 in Dartmouth, in which J. McCarthy participated. This date is regarded as a symbolic moment marking the birth of a new digital discipline, , ebook.pwn.pl/wp-content/uploads/2022/04/ Sztuczna-inteligencja_ebook_Final.pdf?utm_source=google&utm_medium=cpc&utm_campaign=0048_ebookownia&gclid=EAIaIQobChMIkZmI7dvdhwMVTFyRBR37HR-0mEAAYAiAAEgI6N_D_BwE, p. 27 [accessed on: 04.06.2026].

[6] https://blog.archive.org/2021/03/24/major-scifi-discovery-hiding-in-plain-sight-at-the-internet-archive/ [accesed on: 06.04.2026].

[7] https://elportal.pl/blog/ciekawostki/904-etymologia-slowa-robot [accesed on: 04.07.2026].

artificial intelligence, allowing officials to undertake tasks requiring creativity or social intelligence) and improving the speed and quality of public services[8].

The Czech Innovation Strategy

The Czech Republic regards artificial intelligence (AI) as a key technology and is committed to contributing to its development. According to the AI Readiness Index, published annually by Oxford Insights, which assesses the effectiveness of governments in creating conditions for AI development in terms of strategy, regulation and investment, the Czech Republic consistently ranks around 30th in the world[9].

The Czech government is implementing the National Artificial Intelligence Strategy (NAIS) as part of the Innovation Strategy for 2019–2030 (approved by Government Resolution No. 520 of 24 July 2024)[10].

The aim of these measures is to make the country a leader in innovation, including through the implementation of AI in the public sector. Details relating to the public sector are set out in Chapter III of the Strategy, focusing on the introduction of modern e-government services, including in the transport and health sectors. The issues of key importance to this study are set out in the section entitled 'Artificial Intelligence in Public Administration and Public Services'. The vision for this key area is: "A modern and efficient public administration utilising artificial intelligence, whose employees possess the necessary skills and adhere to ethical and security standards, thereby ensuring the provision of high-quality and efficient services in the Czech Republic". The aim of this area is to use artificial intelligence to support decision-making, reduce errors and ensure that services provided to citizens are as user-friendly as possible. Furthermore, this area aims to simplify decision-making processes, improve the efficiency of public administration, and reduce the staffing and administrative burden associated with carrying out certain tasks. In accordance with the Strategy, the sub-objectives are intended to help achieve this vision and goal. In order for these to be realised, the public administration must have access to both high-quality data and a robust digital

[8] L. Lane, L .Haitsma, T. Nowak, Tommaso De Mari Casareto dal Verme, E. Marchesini, R. Carbone, S. Ranise, O. Davidsson, S. Molinari, *AI and Public Administration: The (legal) limits of algorithmic governance, PUBLICATION IS FUNDED BY THE EUROPEAN UNION'S JUSTICE PROGRAMME (2021-2027), 2025, p. 11.*

[9] https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/czech-republic/ [accessed on: 03.04.2026].

[10] https://mpo.gov.cz/en/business/digital-economy/artificial-intelligence/ [accessed on: 02.04.2026].

infrastructure. In this regard, the training of public administration staff is no less important so that they understand the possibilities and limitations of using artificial intelligence. High security standards are emphasised, linking this area to the key area entitled: 'Security aspects of artificial intelligence'. With regard to the implementation of artificial intelligence in public administration, it is crucial to focus on the ethical and transparent use of artificial intelligence.

By 2026, the Strategy envisages the implementation of groundbreaking AI projects in public administration to make life easier for citizens and businesses, streamline operations and increase the added value of public administration. By 2035, the Strategy envisages a significant strengthening of the Czech Republic's position on the international stage in terms of competitiveness and technological and industrial capabilities across all key sectors, underpinned by close cooperation between businesses and the public administration. According to the assumptions set out in the Strategy, the introduction of artificial intelligence into public administration aims to simplify the lives of citizens and businesses and increase its efficiency, in particular by boosting productivity, improving service quality, enforcing tax collection and detecting fraud. This will be achieved through the maximum use of open-source technologies, where their nature permits, whilst maintaining the protection of intellectual property rights and conditions conducive to investment and cooperation with private entities.

The Strategy is complemented by other initiatives and sectoral strategies. These include: the National Space Plan (to be implemented between 2020 and 2025), which relates to the National AI Strategy in the area of AI use in data processing, and the National Health Framework (2030), which covers the development of AI applications for use in healthcare[11].

Chatbots and Smart Cities

Chatbots are widely used in the public sector. They have become intuitive guides and helpers for navigating the often highly complex websites of public institutions[12]. Before the chatbot was introduced, citizens had to look for information in written instructions or obtain it by calling the telephone number provided. Since its introduction, the chatbot has handled the majority of enquiries. It is a relatively inexpensive and easy-to-implement technology capable of performing a variety of

[11] https://stip.oecd.org/stip/interactive-dashboards/policy-initiatives/2025%2Fdata%2FpolicyInitiatives%2F24171 [accessed on: 03.04.2026].

[12] Wawrosz, P. & Jurásek, M. (2022) Wykorzystanie chatbotów w Republice Czeskiej ze szczególnym uwzględnieniem czeskiej administracji publicznej. *International Journal of Public Administration, Management and Economic Development*, 7 (2), pp. 62–78. Source: https://www.ijpamed.eu/index.php/journal/article/view/76.

tasks, such as customer service, booking appointments and answering questions. The use of multilingual chatbots or chatbots combining (languages) with English, which may be useful for people who do not have a good knowledge of Czech. Their usefulness largely depends on whether the chatbot's communication with users is as close as possible to human interaction.

Artificial intelligence is becoming an integral part of modern cities and public administration. One of the AI-based solutions in the Czech transport sector, introduced in Prague, involves improving road condition forecasting through the use of satellite imagery and sensor data. Appropriate models enable the analysis and forecasting of traffic and congestion to prevent accidents and other traffic complications. AI-powered technology helps Prague's road authorities plan targeted repairs and optimise long-term investments. Furthermore, it reduces traffic disruption and improves safety and travel efficiency. This ensures a much more thoughtful approach to urban mobility planning and infrastructure lifecycle management[13].

Another initiative in this area involves the Prague 6 district office, which has been using an advanced chatbot based on Citymind's technology since June 2025. This tool assists residents with matters such as document replacement and navigating the office's structure. Over time, the chatbot has also been rolled out in smaller towns. The automation of this type of service enables communication in multiple languages (including, since the war in Ukraine, in Ukrainian) and 24/7 availability. Chatbots in the Czech public administration are also used to analyse data from the Czech Statistical Office and cadastral registers. Their operations assist officials with spatial planning and crisis management. This solution benefits officials, who, thanks to the chatbot, can focus on more specialised matters rather than standard enquiries. In addition, the chatbot generates reports on the most frequently asked questions by residents, providing officials with valuable data for further service improvement[14].

One of the latest achievements is the 'Public Administration Contact Centre' project (reg. no. CZ.31.5.0/0.0/0.0/24_112/0011231, funded under the National Recovery Plan, component 1.7 – Digital transformation of public administration). The project focuses on the development of the so-called Public Administration Contact Centre (hereinafter referred to as "PACC") with services supported by artificial intelligence. These primarily involve the development of a chatbot, voice bot and mail bot to serve PACC customers. The project aims to address the current

[13] https://digital-strategy.ec.europa.eu/pl/events/genai-meets-public-administrations-info-days-funding-opportunities-and-future-adoption-strategies [accessed on: 01.04.2026].

[14] https://www.citymind.tech/en/case-studies/prague6-oict#:~:text=On%20Wednesday%2012%20June%202025%2C%20a%20new,cooperated%20as%20one%20of%20the%20pilot%20cities [accessed on: 01.04.2026].

situation where there is no single central point providing comprehensive assistance and information on public administration services and solutions related to the use of digital services. Citizens often struggle to find the necessary information and frequently encounter problems when using government information systems. The PACC project offers a solution in the form of a universal, multi-channel Public Administration Contact Centre, emphasising the use of artificial intelligence in citizens' interactions with public administration. This solution will enable users to obtain the necessary information and support quickly and effectively. The project aims to automate routine enquiries using AI-based voice and chatbots, which will speed up and streamline service delivery[15].

eDoklady

At the start of 2024, the Czech Republic launched the eDoklady app, which allows electronic identity documents to be stored digitally on a mobile phone. The rollout of digital documents for Czech citizens was divided into three phases. Initially, it covered central government bodies (all ministries and a number of other state institutions). After a few months, institutions such as the police, courts, tax offices, employment offices and registry offices were included in the programme. Following the completion of the final stage, the use of e-Documents is now possible in other institutions where the law requires identity verification. The eDoklady app represents a significant step towards the introduction in the Czech Republic of the European digital identity wallet, which will be valid in all European Union member states and contain a range of documents and certificates[16].

The gateway to digital services is the Citizen Portal (*Portál občana*), which facilitates communication between citizens and public institutions and can be linked to the eDoklady app. On the Citizen Portal, Czechs can manage their documents and other data from state registers and databases. They can also create a data box, which they can use to communicate electronically with public administration offices, the eRecepta portal, the Vaccination Portal, the Czech Social Security Administration or the Labour Office.

[15] J. Nešpor, A. Jareš, E. Klimentová, *AI and public administration: transforming governance with caution*, ACTA UNIVERSITATIS CAROLINAE – IURIDICA 1 2026, p. 107.

[16] https://ies.lublin.pl/komentarze/proces-cyfryzacji-w-republice-czeskiej-edoklady-i-dipsy/ [accessed on: 01.04.2026].

Digitalisation in the Public Health Sector

Another area that has undergone digitalisation is the healthcare sector. Initiatives in this area aim to improve access to medical services, manage patient data and integrate IT systems across different healthcare institutions. Czech citizens can also use prescriptions issued online.

AI in Customs Law

Artificial intelligence is currently being used in the economic sector for automated risk assessment, document verification and the classification of goods. This technology enables goods to be assigned the correct tariff codes, thereby improving efficiency, speed and accuracy in cross-border trade.

Baltic AI Gigafactory

AI GigaFactories (AIGF) are to be based on the world-class European supercomputing network, EuroHPC. These will be large-scale AI computing infrastructure facilities, designed to create, train and deploy very large AI models and applications on an unprecedented scale (*e.g.,* AI models with hundreds of trillions of parameters). They are intended to be state-of-the-art, large-scale artificial intelligence and data storage centres, specifically designed for the development, training and deployment of next-generation AI models and applications, such as models with hundreds of trillions of parameters. A total of 13 such facilities are planned across Europe. Poland, Lithuania, Latvia and Estonia have formed the Baltic AIGF initiative, which involves the construction of one or two such facilities in our country[17]. In autumn 2025, the countries participating in the Baltic AI Gigafactory began collaborating with the Czech Republic on the integration of projects that had been submitted to Brussels several months earlier. The Czech Republic's inclusion in this initiative will lend the project additional political and economic weight. According to the Deputy Minister of Industry and Trade of the Czech Republic, the country will become a technological leader in this field. This will enable the creation of the most advanced AI models, which will be used in healthcare, transport, and security, as well as for research, to support universities, and to create an entire technological ecosystem. This will lead to the development of data

[17] https://pwr.edu.pl/uczelnia/aktualnosci/gigafabryka-ai-moze-powstac-we-wroclaw-iu-silne-wsparcie-dla-projektu-pwr-13985.html [accessed on: 07.04.2026].

centres, the creation of facilities for training models, and will ensure cooperation with research centres[18].

The Automatic Issuance of Administrative Decisions in Czech Administrative Proceedings

The use of artificial intelligence algorithms to issue administrative decisions is quite common in many countries. This is the case in several Scandinavian countries, Canada and Germany. In these countries, algorithms have already replaced humans in handling certain administrative matters and issuing the final decisions on those matters.

At the end of 2024, a bill was tabled in the Czech Parliament to amend the Czech Code of Administrative Procedure by introducing a new Article 15a, which would pave the way for the automatic performance of certain acts in administrative proceedings[19].

This initiative is undoubtedly timely and forms part of a broader global trend. At present, however, the Czech Code of Administrative Procedure does not provide for the possibility of issuing administrative decisions automatically, without the involvement of an official. For the time being, automated tools may only be used in practice to support the activities of the public administration.

The proposed Article 15a would read as follows:

> Unless the nature of the case under consideration, the protection of the rights of the persons concerned or the protection of the public interest requires that a procedural act be performed by an official, that act may be performed automatically, without the involvement of an official. In particular, such an act may not be performed in this manner if it requires the exercise of discretionary powers or concerns a decision on an appeal.

The relevant literature emphasises that a reform of this scale should not be carried out by means of an ad hoc parliamentary amendment, but through a comprehensive government initiative. The legislator should conduct a full review of existing administrative procedures to determine which are suitable for automation and which are not. These procedures should be clearly and exhaustively listed in the legislation, or at the very least, criteria should be established to enable authorities and citizens to unequivocally determine when automation is permitted. Further-

[18] https://cyberdefence24.pl/technologie/polska-i-czechy-nawiazaly-wspolprace-ws-baltic-ai-gigafactory [accessed on: 07.04.2026].

[19] Act No. 500/2004 Coll., the Administrative Procedure Code.

more, human oversight would be desirable and should be incorporated into the system, at the very least[20].

Conclusions

Public administration, which is bringing its services closer to the public through digitalisation, is perceived by citizens as innovative and modern. The latest analyses clearly indicate that, whilst interest in artificial intelligence is growing in the Czech Republic, practical implementations have not yet reached a large scale and remain uneven. In this respect, the path to the digitalisation of public administration in Poland and the Czech Republic looks quite similar[21]. There are many examples of groundbreaking technologies being implemented in the public sector, but these have not yet resulted in systemic solutions. The potential available to public administration in this area has not yet been 'tapped' by AI. Numerous matters falling within the remit of public administration, involving straightforward facts, which are currently handled on a massive scale, will in future be dealt with without the involvement, or with only minimal involvement, of an official or human being. At present, one might rather speak of a sort of 'early spring' for artificial intelligence in public administration, which so far relies only to a limited extent on automation.

References

Lane L., Haitsma L., Nowak T., De Mari Casareto dal Verme T., Marchesini E., Carbone R., Ranise S., Davidsson O., Molinari S.*AI and Public Administration: The (legal) limits of algorithmic governance,* PUBLICATION IS FUNDED BY THE EUROPEAN UNION'S JUSTICE PROGRAMME (2021-2027), 2025.

Sharp V., Nešpor J., Klimentova E., *Automation of administrative proceedings in the Czech Republic: critical reflections on the draft "adm amendment" to the administrative procedure code,* STUDIA IURIDICA Cassoviensia vol. 13, 2025.

Szewczyk E., *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne,* Warszawa 2025.

Wawrosz, P. & Jurásek, M. (2022) Wykorzystanie chatbotów w Republice Czeskiej ze szczególnym uwzględnieniem czeskiej administracji publicznej. *International Journal of Public Administration, Management and Economic Development*, *7* (2).

[20] V. Sharp, J. Nešpor, E. Klimentova, *Automation of administrative proceedings in the Czech Republic: critical reflections on the draft "adm amendment" to the administrative procedure code,* STUDIA IURIDICA Cassoviensia vol. 13, 2025, pp. 177, 184.

[21] Szerzej na temat zastosowania AI w administracji publicznej w Polsce E. Szewczyk, *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne,* Warszawa 2025, p. 8.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Michał Tadeusz Najman
Akademia Mazowiecka w Płocku
ORCID: 0000-0001-8207-8458
michalnajman0@gmail.com

Przewlekłość recenzowania rozpraw doktorskich i fikcyjność odpowiedzialności recenzentów za nią

Słowa kluczowe: przewlekłość, recenzja, szkolnictwo wyższe, stopień naukowy doktora

Streszczenie. Postępowanie o nadanie stopnia naukowego doktora jest specyficznym postępowaniem administracyjnym prowadzonym przez uprawnione do tego podmioty. Jednym z jego etapów jest recenzja osiągnięcia naukowego stanowiącego podstawę do nadania stopnia naukowego. Znaczenie recenzji w postępowaniu o nadanie stopnia naukowego doktora jest kluczowe dla prawidłowości, obiektywności i rzetelności awansu naukowego. Zapewnia ona gwarancję nobilitacji naukowej z uwagi na przygotowywane przez wiele lat osiągnięcie naukowe, a nie inne okoliczności. Stanowi zabezpieczenie przed deprecjacją statusu społecznego nadawanego stopnia naukowego. Od tego, czy recenzja jest pozytywna, zależy powodzenie nadania stopnia naukowego. Jest ona zatem etapem kluczowym dla szkolnictwa wyższego. Determinuje to jednocześnie istotne znaczenie recenzentów w procesie recenzowania. Wymaga się od nich stosowania jednolitych i obiektywnych kryteriów oceny. Nieco mniej uwagi poświęca się natomiast sprawności etapu recenzyjnego. Na kwestię tę zwrócił uwagę ustawodawca w postępowaniach o nadanie stopnia naukowego doktora habilitowanego i profesora. Jednakże w postępowaniu o nadanie stopnia naukowego doktora kwestia ta jest nieuregulowana w tym sensie, że nie przewiduje się sankcji prawnej za naruszenie terminu na sporządzenie recenzji. Celem artykułu jest wskazanie roli i znaczenia recenzentów w postępowaniu o nadanie stopnia doktora ze szczególnym uwzględnieniem skutków prawnych zarówno dla recenzentów, jak i dla ubiegających się o nadanie stopnia w przypadku przewlekłości postępowania recenzyjnego.

The prolixity of the review process for doctoral thesis and the fictitious nature of the reviewers' responsibility for it

Keywords: prolixity, review, tertiary education, PHD

Summary. The procedure for obtaining a doctoral degree is a specific administrative procedure conducted by authorized entities. One of its stages is the review of the scientific achievement that forms the basis for awarding the degree. The importance of the review in giving procedure is crucial to the accuracy, objectivity, and reliability of academic advancement. It guarantees academic ennoblement based on the scientific achievement developed over many years, and not on other circumstances. It protects against the depreciation of the social status of the awarded academic degree. The success of the academic degree award depends on the review's positive outcome. Therefore, it is a key stage in higher education. This simultaneously determines the significant role of reviewers in the review process. They are required to apply uniform and objective evaluation criteria. Less attention, however, is paid to the efficiency of the review phase. This issue was highlighted by the legislator in the procedures for awarding the academic degrees of doktor habilitowany and profesor. However, in

the doctoral degree award process, this issue is unregulated in that there are no legal penalties for failing to meet the review deadline. The aim of this article is to highlight the role and importance of reviewers in the doctoral degree award process, with particular emphasis on the legal consequences for both reviewers and applicants if the review process is excessively lengthy.

Wprowadzenie

Od 2022 r. Rzecznik Praw Obywatelskich wskazuje, że „w postępowaniach o nadanie stopni i tytułów naukowych występuje problem opóźnień w sporządzaniu recenzji. Dzieje się tak mimo przepisów przewidujących, że recenzentem nie może zostać osoba, która w ostatnich 5 latach dwukrotnie nie dochowała terminu sporządzenia recenzji. Brak jednoznacznych rozwiązań co do sposobu gromadzenia i udostępniania informacji o terminowości recenzji stawia pod znakiem zapytania praktyczne znaczenie tych regulacji”[1]. Należy jedynie nadmienić, że wskazane przepisy przewidujące, iż recenzentem nie może zostać osoba, która w ostatnich 5 latach dwukrotnie nie dochowała terminu sporządzenia recenzji, odnoszą się do postępowań o nadanie stopnia naukowego doktora habilitowanego i profesora. Nie znajdują zastosowania do recenzji rozpraw doktorskich. Wskazane konstatacje RPO uwidaczniają problem polskiego szkolnictwa wyższego w postaci rażącej nieterminowości sporządzania recenzji w postępowaniach awansowych.

Celem artykułu jest zwrócenie uwagi na szczególne znaczenie procesu recenzyjnego w postępowaniach awansowych oraz wskazanie regulacji stanowiących podstawę i zakres odpowiedzialności prawnej recenzentów w przypadku naruszenia terminów sporządzenia recenzji. Przedstawiono również skutki prawne dla osoby ubiegającej się o nadanie stopnia naukowego doktora w przypadku przewlekłości procesu recenzyjnego.

1. Postępowanie w sprawie nadania stopnia naukowego doktora

Postępowanie w sprawie nadania stopnia naukowego doktora regulowane jest ustawie z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (dalej jako ustawa/ustawa o szkolnictwie wyższym)[2]. Stosownie do art. 178 ust. 3 ustawy w postępowaniach w sprawie nadania stopnia doktora oraz w sprawie nadania stopnia doktora habilitowanego, w zakresie nieuregulowanym w ustawie, stosuje się odpowiednio przepisy k.p.a.[3] Omawiane postępowanie uznaje się za szczegól-

[1] https://bip.brpo.gov.pl/pl/contetnt/rpo-opoznienia-recenzje-postepowania-awansowe-mein-kolejne-odpowiedz [dostęp: 25.11.2025].

[2] T.j. Dz.U. z 2024 r. poz. 1571, 1871, 1897, z 2025 r. poz. 619, 620, 621, 622, 1162.

[3] Ustawa z dnia 14 czerwca 1960 r. Kodeks postępowania administracyjnego, t.j. Dz.U. z 2024 r. poz. 572, z 2025 r. poz. 769.

ny rodzaj postępowania administracyjnego[4]. Prowadzone jest ono przez podmiot doktoryzujący w rozumieniu art. 185 ust. 1 ustawy, tj. uczelnię, instytut PAN, instytut badawczy, instytut międzynarodowy albo CMKP, które w danej dyscyplinie posiadają kategorię naukową A+, A lub B+ albo uprawnienie nadane w trybie określonym w art. 226a ust. 1 ustawy.

Postępowanie w sprawie nadania stopnia doktora wszczyna się na wniosek osoby spełniającej wymagania do nadania stopnia doktora[5]. Do wniosku dołącza się rozprawę doktorską wraz z pozytywną opinią promotora lub promotorów (art. 189 ustawy). Dalsze czynności w postępowaniu regulowane są m.in. na mocy przepisów stosowanych w toku postępowania (szczegółowe przepisy ustawy, względnie k.p.a.), lecz przede wszystkim na mocy uchwał senatu lub rady naukowej podmiotu doktoryzującego, określających m.in. sposób wyznaczania i zmiany promotora, promotorów lub promotora pomocniczego w przypadku osób ubiegających się o nadanie stopnia doktora w trybie eksternistycznym; zasady ustalania wysokości opłaty za postępowanie w sprawie nadania stopnia doktora w trybie eksternistycznym oraz zwalniania z tej opłaty; tryb złożenia rozprawy doktorskiej sposób wyznaczania recenzentów (art. 192 ust. 2 ustawy). Po złożeniu wniosku organ przeprowadza tzw. postępowanie wstępne, w którym bada właściwość, wymogi formalne oraz czy nie występują przesłanki odmowy wszczęcia postępowania[6].

---

[4] M. Bogusz, *Charakter prawny uchwał rady wydziału w przewodzie doktorskim*, „Gdańskie Studia Prawnicze" 2015, nr 33, s. 73-81; J. Tarno, *Rola odpowiedniego stosowania przepisów k.p.a. w postępowaniach w sprawach stopni naukowych (wybrane zagadnienia)*, „Zeszyty Naukowe Sądownictwa Administracyjnego" 2011, nr 6, s. 18-35; S. Łazuk, *Wybrane problemy stosowania Kodeksu postępowania administracyjnego w procedurze awansu naukowego*, [w:] *Specyfika postępowań administracyjnych w sprawach z zakresu szkolnictwa wyższego i nauki*, red. J. Tarno, A. Szrot, P. Pokorny, Lublin 2016, s. 87-94.

[5] Art. 186 ust 1 ustawy – Stopień doktora nadaje się osobie, która: 1) posiada tytuł zawodowy magistra, magistra inżyniera albo równorzędny lub posiada dyplom, o którym mowa w art. 326 ust. 2 pkt 2 lub art. 327 ust. 2, dający prawo do ubiegania się o nadanie stopnia doktora w państwie, w którego systemie szkolnictwa wyższego działa uczelnia, która go wydała; 2) uzyskała efekty uczenia się dla kwalifikacji na poziomie 8 PRK, przy czym efekty uczenia się w zakresie znajomości nowożytnego języka obcego są potwierdzone certyfikatem lub dyplomem ukończenia studiów, poświadczającymi znajomość tego języka na poziomie biegłości językowej co najmniej B2; 3) posiada w dorobku co najmniej: a) 1 artykuł naukowy opublikowany w czasopiśmie naukowym lub w recenzowanych materiałach z konferencji międzynarodowej, które w roku opublikowania artykułu w ostatecznej formie były ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 ust. 2 pkt 2 lit. b, lub b) 1 monografię naukową wydaną przez wydawnictwo, które w roku opublikowania monografii w ostatecznej formie było ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 ust. 2 pkt 2 lit. a, albo rozdział w takiej monografii, lub c) dzieło artystyczne o istotnym znaczeniu; 4) przedstawiła i obroniła rozprawę doktorską; 5) spełniła inne wymagania określone przez podmiot doktoryzujący.

[6] M. Sieniuć, *Jednostka w postępowaniach w sprawach nadania stopnia i tytułu naukowego. Studium z prawa administracyjnego procesowego*, Łódź 2019, s. 81-83; P. Waszkiewicz, *Nadawanie stopnia doktora*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. A. Jakubowski, Warszawa 2023, s. 590-591.

W przypadku spełnienia wymogów formalnych, po złożeniu wniosku wraz z rozprawą doktorską powoływani są recenzenci. Stosownie do art. 190 ust. 2 ustawy w postępowaniu w sprawie nadania stopnia doktora wyznacza się trzech recenzentów spośród osób niebędących pracownikami podmiotu doktoryzującego oraz uczelni, instytutu PAN, instytutu badawczego, instytutu międzynarodowego, Centrum Łukasiewicz, instytutu Sieci Łukasiewicz albo CMKP, których pracownikiem jest osoba ubiegająca się o stopień doktora. Recenzentem może być osoba posiadająca co najmniej stopień doktora habilitowanego (art. 190 ust. 4 ustawy)[7].

Kolejnym etapem jest sporządzenie recenzji przez recenzentów. Zgodnie z art. 190 ust. 3 ustawy recenzenci sporządzają recenzje rozprawy doktorskiej w terminie dwóch miesięcy od dnia jej doręczenia.

Po sporządzeniu recenzji wydawane jest postanowienie o dopuszczeniu lub odmowie dopuszczeni do obrony, na które przysługuje zażalenie. Jeśli w podmiocie doktoryzującym określono dodatkowe wymagania na podstawie art. 192 ust. 3 ustawy, to ich weryfikacja może odbywa się również po sporządzeniu recenzji. Kolejno organ prowadzący postępowanie sprawdza rozprawę w Jednolitym Systemie Antyplagiatowym (*Uniform Anti-plagiarism System*) (art. 188 ust. 4 ustawy).

Następnym etapem postępowania o nadanie stopnia doktora jest obrona rozprawy doktorskiej. Może być do niej dopuszczona osoba, która uzyskała pozytywne recenzje od co najmniej dwóch recenzentów oraz spełniła wymagania do nadania stopnia doktora w podmiocie doktoryzującym (art.191 ust. 1 ustawy o szkolnictwie wyższym).

Po przeprowadzeniu obrony podmiot doktoryzujący podejmuje decyzję w sprawie nadania stopnia doktora lub odmowy nadania stopnia doktora (art. 193 ust. 1 ustawy o szkolnictwie wyższym).

Na marginesie należy wskazać, że postępowanie o nadanie stopnia doktora habilitowanego (art. 281-226 ustawy o szkolnictwie wyższym) oraz profesora (art. 227-231 ustawy o szkolnictwie wyższym) są zbliżone proceduralnie, jednakże kluczowa różnica z punktu widzenia dalszego wywodu polega na wprowadzeniu dodatkowych wymogów dla recenzentów wniosku o nadanie stopnia doktora habilitowanego lub wniosku o nadanie tytułu profesora, jakim jest to, że recenzentem nie może zostać osoba, która w okresie ostatnich 5 lat dwukrotnie nie dochowała terminu sporządzenia recenzji (dla wniosków o nadanie stopnia doktora habilitowanego – 8 tygodni (art. 221 ust. 7 ustawy o szkolnictwie wyższym)

[7] W art. 190 ust. 5 ustawy wskazano wyjątek od tej reguły. Stosownie do niego promotorem i recenzentem może być osoba niespełniająca warunków określonych w ust. 4, która jest pracownikiem zagranicznej uczelni lub instytucji naukowej, jeżeli organ, o którym mowa w art. 178 ust. 1, uzna, że osoba ta posiada znaczące osiągnięcia w zakresie zagadnień naukowych, których dotyczy rozprawa doktorska.

oraz dla wniosku o nadanie tytułu profesora – 3 miesięcy (art. 229 ust. 2 ustawy o szkolnictwie wyższym).

W przypadku odmowy nadania stopnia doktora, ubiegającemu się o nadanie stopnia naukowego przysługuje odwołanie do Rady Doskonałości Naukowej, która rozpoznaje je w terminie nie dłuższym niż 6 miesięcy (art. 193 ust. 1-4 ustawy). W toku postępowania Rada Doskonałości Naukowej obowiązana jest powołać co najmniej dwóch recenzentów w postępowaniu odwoławczym. W Poradniku Rady Doskonałości Naukowej z 2022 r. „Recenzje w Postępowaniach o awans naukowy" (dalej jako: Poradnik)[8] wskazuje się, że są oni zobowiązani do sporządzenia „opinii zakończonej jednoznaczną konkluzją o uznaniu albo braku uznania zasadności wniesionego odwołania" (s. 18-19). Zarówno ustawa, jak i Poradnik milczą co do wymogów dla recenzentów w postępowaniu odwoławczym[9]. Należy jednak przyjąć, że będą one tożsame z wymogami dla recenzentów w poszczególnych postępowaniach (art. 190 ust. 4 i 5 ustawy lub art. 221 ust. 5 i 6 ustawy). W ustawie nie wskazano również terminu na sporządzenie recenzji w postępowaniu odwoławczym, choć można wywodzić go pośrednio z art. 193 ust. 4 ustawy. Termin ten jest ruchomy i każdorazowo zależy od decyzji Rady Doskonałości Naukowej. Ustawodawca pozostawia pole decyzyjne Radzie Doskonałości Naukowej, przenosząc na nią obowiązek wyznaczenia recenzentowi terminu na sporządzenie opinii w postępowaniu odwoławczym w taki sposób, aby była ona w stanie dotrzymać terminu na załatwienie sprawy. Brak wyznaczenia takiego terminu recenzentowi przez Radę Doskonałości Naukowej rodzić może wobec niej skutki prawne, które opisane zostaną w dalszej części.

2. Rola recenzentów w postępowaniach awansowych

W przytoczonym Poradniku wskazano, że „recenzenci pełnią jedną z najistotniejszych ról w postępowaniach o awans naukowy. Przyjęty w polskim systemie prawnym model awansów naukowych wiąże się z dokonywaniem eksperckiej oceny na każdym jego szczeblu – od postępowania w sprawie nadania stopnia doktora aż do postępowania w sprawie nadania tytułu profesora. To właśnie merytoryczna ocena, która dokonywana jest przez recenzentów, stanowi podstawowy materiał, który uwzględniany jest przez podmioty doktoryzujące oraz podmioty habilitujące, a także przez Radę Doskonałości Naukowej w ramach podejmowanych

[8] https://www.rdn.gov.pl/dobre-praktyki.poradnik-recenzje-w-postepowaniach-o-awans-naukowy.html [dostęp: 25.11.2025].

[9] B. Sitek, A. Woźniak, *Komentarz do art. 238*, [w:] *Prawo zatrudnienia pracowników szkół wyższych. Komentarz do wybranych przepisów Prawa o szkolnictwie wyższym i nauce*, red. K. Baran, Warszawa 2026, s. 420-421.

rozstrzygnięć" (s. 2). W dalszej części podkreślono, iż „Przepisy ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce precyzyjnie określają rolę, jaką przypisuje się recenzentom w postępowaniach o awans naukowy. Należy podkreślić, że – w zależności od danego postępowania – rola ta będzie kształtowała się nieco odmiennie. Niewątpliwie można natomiast wyróżnić pewne cechy wspólne, niezależne od rodzaju postępowania i trybu, które będą odnosiły się do roli recenzentów. Należą do nich: 1) obiektywność; 2) rzetelność; 3) dokonywanie oceny zgodnie z najlepszą posiadaną wiedzą; 4) poufność" (s. 5).

W literaturze poświęconej instytucji recenzji niewiele uwagi poświęca się jej znaczeniu społecznemu[10]. W zasadzie nie zwraca się uwagi na istotę weryfikacji osiągnięć, w szczególności naukowych, przez podmioty charakteryzujące się szczególnymi właściwościami, takimi jak wiedza, doświadczenie życiowe oraz dorobek zawodowy i naukowy[11].

Recenzja służy utrzymaniu znaczenia danej instytucji społecznej[12]. Gwarantuje ona innym jednostkom, że osoba otrzymująca stopnień naukowy dostąpiła go ze względu na osiągnięcie naukowe. Jednocześnie przemawia za tym publiczny charakter recenzji (art. 188 ust. 1 pkt. 1 lit. b ustawy). Recenzja ma stanowić ocenę osiągnięcia naukowego przez pryzmat obiektywnych kryteriów podlegających weryfikacji przez czynnik społeczny. Społeczeństwo, polegając na wiedzy i doświadczeniu recenzentów, winno ustalić stopień rozwoju naukowego kandydata do stopnia naukowego. Powyższe chronić ma społeczne postrzeganie osób posiadających stopień naukowy, co z kolei ma sprzyjać rozwojowi nauki i zachęcać jednostki do podejmowania badań naukowych – abstrahując od pobudek ambicjonalnych[13]. Innymi słowy to, że coś podlega recenzji, wymusza na twórcy dążenie do zaspokojenia oczekiwań społecznych[14]. Świadczy jednocześnie o istnieniu standardów społecznych dla recenzowanego dzieła i związanych z nim dalszych korzyści (materialnych i niematerialnych) w przypadku recenzji pozytywnych[15]. Stopień naukowy

[10] M. Krauz, *Krytyk i/a recenzent – znaczenie pojęć w słowniku i w tekście*, „SŁOWO. Studia językoznawcze" 2021, nr 12, s. 86-98.

[11] N. Black, S. Van Rooyen, F. Godlee, R. Smith, S. Evans, *What makes a good reviewer and a good review for a general medical journal*, „Jama" 1998, nr 280(3), s. 231-233; A. Scott, *Peer review and the relevance of scienc*e, „Futures" 2007, nr 39(7), s. 830-832.

[12] T. Pfeiffer, L. Tran, C. Krumme, D. Rand, *The value of reputation*, „Journal of the Royal Society Interface" 2012, nr 9(76), s. 2792-2794; J. Wang, *The Importance of Reviewers*, „Global Spine Journal" 2015, nr 5(2), s. 83.

[13] S. Cannegieter, T. Lisman, What is in it for the reviewer?, „Journal of Thrombosis and Hemostasis" 2024, nr 22(3), s. 579-580; P. Herzog, *The Importance of Peer Review: Recommendations for Reviewers and Authors*, „Review of Religious Research" 2023, nr 65(1), s. 3-6.

[14] A. Bahishti, The importance of review articles & its prospects in scholarly literature [w:] „Extensive Reviews" 2021, nr 1(1), s. 1-6.

[15] F. Bloom, *The importance of reviewers*, „Science"1999, nr 283(5403), s. 789; P. Nightingale, A. Scott, *Peer review and the relevance gap: ten suggestions for policy-makers*, „Science and Public Policy" 2007, nr 34(8), s. 548.

stanowi zatem gratyfikację społeczną za wkład jednostki w rozwój nauki, a recenzja ma dawać potwierdzenie, że gratyfikacja ta jest słuszna[16].

3. Odpowiedzialność recenzentów za naruszenie terminu sporządzenia recenzji

Odpowiedzialność recenzentów za naruszenie terminu sporządzenia recenzji dysertacji doktorskiej rozpatrywać można na gruncie ustawy o szkolnictwie wyższym i nauce, na gruncie przepisów k.p.a. o przewlekłości postępowania, na gruncie prawa pracy oraz na gruncie prawa cywilnego.

Jak wskazuje Paweł Dańczak termin na sporządzenia recenzji rozprawy doktorskiej ma charakter instrukcyjny, co oznacza, że jego przekroczenie nie wpływa negatywnie na ważność tak sporządzonej recenzji[17]. Tożsame stanowisko zajmuje Marcin Dokowicz, który wskazuje, że ustawodawca nie przewidział sankcji związanych z niedotrzymaniem tego terminu. Jak podkreśla ten autor, w poprzednim stanie prawnym umowa sporządzana pomiędzy podmiotem doktoryzującym a recenzentem musiała przewidywać kary umowne związane z niedotrzymaniem tego terminu. Obecne przepisy nie przewidują możliwości stosowania takich kar[18]. Podobnie Przemysław Mroczkowski stwierdza, że termin ten ma wyłącznie instrukcyjny charakter, wszak w ślad za jego ewentualnym uchybieniem nie stoją żadne negatywne konsekwencje, np. w postaci konieczności zmiany recenzenta, który spóźniałby się ze sporządzeniem recenzji, zmniejszeniem jego wynagrodzenia, czy w końcu sankcji, jaką ustawodawca wprowadził w przypadku postępowania habilitacyjnego oraz postępowania o nadanie tytułu profesora, gdzie dwukrotne niedochowanie terminu przez recenzenta w okresie 5 ostatnich lat uniemożliwia powołanie danej osoby do pełnienia funkcji recenzenta w tych postępowaniach. Trudno powiedzieć, czy brak podobnej regulacji na gruncie spraw dotyczących nadania stopnia doktora to zamierzony zabieg ustawodawcy, czy też efekt jego niedopatrzenia. Niewątpliwie jednak brak jakichkolwiek środków dyscyplinujących recenzentów do terminowego sporządzenia recenzji nie zasługuje na aprobatę, gdyż opieszałość w tym zakresie nie tylko może prowadzić do przedłużenia postępowania, lecz może stać się również instrumentem służącym jego świadomemu

[16] R. Walker, B. Barros, R. Conejo, K. Neumann, M. Telefont, *Personal attributes of authors and reviewers, social bias and the outcomes of peer review: a case study*, „F1000Research" 2015, nr 4(21), s. 9-11.

[17] P. Dańczak, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. A. Jakubowski, wyd. 1, Warszawa 2023, art. 190.

[18] M. Dokowicz, *Komentarz do art. 190*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Woźnicki, Warszawa 2019, s. 502.

przewlekaniu[19]. Również Hubert Izdebski[20] oraz Wojciech Kiełbasiński[21] wskazują na brak sankcji prawnych za przewlekłość recenzowania rozpraw doktorskich. Tym samym ustawa o szkolnictwie wyższym nie przewiduje szczególnych sankcji wobec recenzentów za naruszeni dwumiesięcznego terminu sporządzenia recenzji rozprawy doktorskiej.

Na gruncie postępowania administracyjnego środkiem służącym przeciwdziałaniu bezczynności organu i przewlekłości postępowania jest ponaglenie uregulowanie w art. 37 k.p.a. Stosownie do wskazanego przepisu stronie służy prawo do wniesienia ponaglenia, jeżeli: 1) nie załatwiono sprawy w terminie określonym w art. 35 lub przepisach szczególnych ani w terminie wskazanym zgodnie z art. 36 § 1 (bezczynność); 2) postępowanie jest prowadzone dłużej, niż jest to niezbędne do załatwienia sprawy (przewlekłość). Organ właściwy do rozpoznania ponaglenia rozpatruje je w terminie siedmiu dni od dnia jego otrzymania. Organ rozpatrujący ponaglenie wydaje postanowienie, w którym: 1) wskazuje, czy organ rozpatrujący sprawę dopuścił się bezczynności lub przewlekłego prowadzenia postępowania, stwierdzając jednocześnie, czy miało ono miejsce z rażącym naruszeniem prawa; 2) w przypadku stwierdzenia bezczynności lub przewlekłości: a) zobowiązuje organ rozpatrujący sprawę do załatwienia sprawy, wyznaczając termin do jej załatwienia, jeżeli postępowanie jest niezakończone, b) zarządza wyjaśnienie przyczyn i ustalenie osób winnych bezczynności lub przewlekłości, a w razie potrzeby także podjęcie środków zapobiegających bezczynności lub przewlekłości w przyszłości. Organem właściwym do w toku postępowania doktoryzującego do rozpoznania ponaglenia jest Rada Doskonałości Naukowej (art. 232 ust. 2a ustawy). Aby móc stosować ponaglenie jako środek mobilizujący recenzentów postępowanie recenzyjne w postępowaniu o nadanie stopnia doktora traktować należy jako samodzielne postępowanie. Nie należy wówczas opowiadać się za niemożliwością stosowania art. 35 k.p.a. z uwagi na brak ustawowego terminu na nadanie stopnia doktora przez organ prowadzący postępowanie, jak również brak przesłanki prowadzenia postępowania dłużej niż to niezbędne z uwagi na konieczność uzyskania recenzji. Powyższe odnieść należy również to postępowań odwoławczych, tj. dopuszczalności wniesienia ponaglenia wobec Rady Doskonałości Naukowej do Ministra właściwego do spraw szkolnictwa wyższego i nauki (art. 232 ust. 3 ustawy). W przypadku

[19] P. Mroczkowski, [w:] *Prawo o szkolnictwie wyższym i nauce dla studenta i doktoranta – komentarz praktyczny*, red. W. Kiełbasiński, B. Pietrzyk-Tobiasz, M. W. Kuliński, wyd. 1, Warszawa 2024, art. 190.

[20] H. Izdebski, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Zieliński H. Izdebski, wyd. 3, LEX/el. 2023, art. 190, https://sip.lex.pl/#/commentary/587785502/731978/izdebski-hubert-zielinski-jan-michal-prawo-o-szkolnictwie-wyzszym-i-nauce-komentarz-wyd-iii [dostęp: 16.01.2026].

[21] W. Kiełbasiński, [w:] *Prawo o szkolnictwie wyższym…*, red. W. Kiełbasiński, M. Kuliński, B. Pietrzyk-Tobiasz, s. 454-456.

postępowań odwoławczych możliwość wniesienia ponaglenia jest od strony formalnej mniej problematyczna. Art. 193 ust. 4 ustawy wskazuje, że Rada Doskonałości Naukowej rozpatruje odwołanie w terminie 6 miesięcy. Przewlekłość recenzentów uniemożliwiająca dochowanie wskazanego terminu skutkować może stwierdzeniem bezczynności lub przewlekłości Rady Doskonałości Naukowej.

W doktrynie dość jednoznacznie podkreśla się, że recenzent w postępowaniu o nadanie stopnia naukowego nie jest biegłym[22]. Jako uzasadnienie takiego stanowiska podnosi się związanie recenzją podmiotu doktoryzującego (art. 191 ust. 1 i art. 221 ust. 9c ustawy), a także brak dopuszczalności merytorycznej kontroli recenzji oraz jej obligatoryjność, jak również fakt, że recenzenci nie posiadają informacji specjalnych, gdyż pozostali członkowie komisji również posiadają tożsame informacje danej dziedzinie naukowej, przynajmniej pod względem formalnym. Również Rada Doskonałości Naukowej jasno wskazuje, że recenzenci nie są biegłymi (s. 2-3 Poradnika), a co za tym idzie, przepisy dotyczące biegłych z k.p.a. nie znajdują zastosowania. Jednak w orzecznictwie odmiennie: opowiadając się za wykładnią funkcjonalną i celowościową, podnosi się możliwość mobilizowanie opieszałych recenzentów właśnie poprzez stosowanie przepisów o biegłych, m.in. poprzez badanie, czy recenzja spełnia warunki formalne i czy w ogóle może być podstawą prowadzenia postępowania[23]. Zaznaczano nawet, że choć wyłączone jest do recenzenta stosowanie (wprost) przepisów k.p.a. o biegłych, to mimo wszystko pełni on jednak rolę eksperta. A zatem błędny jest wniosek, że nie może być on pod żadnym względem traktowany jako biegły w rozumieniu art. 84 § 1 k.p.a.[24] Niewątpliwe jest bowiem, że recenzja jest opinią specjalisty z danej dziedziny, którą zajmuje się kandydat do tytułu naukowego. Wobec tego jest ona przecież dowodem, i to – z uwagi na przedmiot wspomnianego postępowania – szczególnego rodzaju. Dlatego musi podlegać wszechstronnej ocenie przy podejmowaniu uchwały przez wspomniany organ. Do tego postępowania mają przecież odpowiednie zastosowanie przepisy art. 77 § 1, 80 i 107 § 3 k.p.a., zgodnie z którymi organ administracji

---

[22] Z. Kmieciak, J. Wegner, *Odpowiednie stosowanie przepisów kodeksu postępowania administracyjnego w sprawach nadawania stopni naukowych i tytułu profesora*, „Państwo i Prawo" 2021, nr 3, s. 5; A. Bochentyn, *Dowód z opinii biegłego w jurysdykcyjnym postępowaniu administracyjnym*, Warszawa 2020, s. 298-299; P. Dańczak, *op. cit.*; M. Bogusz, M. Jędrzejczak, *Zagadnienie wyłączenia członka organu doktoryzującego i członka komisji doktorskiej z powodu pełnienia funkcji recenzenta w postępowaniu doktorskim*, [w:] *Administracja publiczna wobec procesów zmian w XXI wieku. Księga jubileuszowa Profesora Jerzego Korczaka*, red. P. Lisowski, Wrocław 2024, s. 902; M. Bogusz, *Zagadnienie dopuszczalności zbiegu ról procesowych członka komisji habilitacyjnej i członka organu habilitującego*, „Prawo i Więź" 2025, nr 5(58), s. 258; J. Tarno, *op. cit.*, s. 27. Odmiennie: M. Sieniuć, *op. cit.*, s. 242.

[23] Wyrok WSA w Warszawie z 29.06.2021 r., II SA/Wa 686/21, LEX nr 3293247; wyrok WSA w Warszawie z 3.12.2021 r., II SA/Wa 1081/21, LEX nr 3308515.

[24] M. Trzebiatowski, *Kryteria prawne nadawania tytułu profesora w zakresie sztuki – analiza normatywno-empiryczna*, „Prawo i Więź" 2022, nr 4(42), s. 384-386.

publicznej ma obowiązek w sposób wyczerpujący zebrać i rozpatrzyć cały materiał dowodowy i na podstawie jego całokształtu ocenić wykazanie określonych okoliczności (podejmując przy tym wszelkie niezbędne czynności dla dokładnego wyjaśnienia – art. 7 k.p.a.), a swoje rozstrzygnięcie należycie uzasadnić przez wskazanie podstawy faktycznej i prawnej[25]. Uwidacznia to dysonans pomiędzy orzecznictwem i doktryną. Słusznie zaznacza się, że w obecnym stanie prawnym brak jest możliwości stosowania sankcji przewidzianych dla biegłych (art. 88 k.p.a.) wobec recenzentów w omawianym postępowaniu, lecz prawdziwość tego stwierdzenia uznać można przy wyłącznym stosowaniu jedynie wykładni językowej. Szczegółowe odrębności recenzji od opinii i jej szczególna, wiążąca moc nie pozbawiają jej charakteru opinii. Fakt, że recenzent nie jest biegłym, a sporządzona przez niego recenzja nie podlega kontroli merytorycznej tak jak opinia biegłego w postępowaniu administracyjnym, nie oznacza, że nie można stosować wobec niego przepisów odnoszących się do biegłych. Tym samym należy postulować, aby status recenzenta został uregulowany w tym zakresie, przynajmniej co do jednoznacznej dopuszczalności stosowania do recenzenta przepisów przewidzianych dla biegłych, względnie winna zostać wprowadzona odrębna regulacja w tym zakresie, co zostanie omówione w dalszej części.

Innym środkiem prawnym mającym mobilizować organy administracji do przestrzegania terminów załatwienia sprawy jest skarga określona w art. 227 k.p.a. Organ rozpoznaje skargę bez zbędnej zwłoki, nie później jednak niż w ciągu miesiąca.

Organem właściwym do rozpoznawania skarg oraz ponagleń w toku postępowania doktoryzującego, tj. organem wyższego stopnia wobec podmiotu doktoryzującego, jest Rada Doskonałości Naukowej, stosownie do art. 232 ust. 2a ustawy o szkolnictwie wyższym, w przypadku zaś ponaglenia Rady Doskonałości Naukowej organem właściwym będzie Minister właściwy do spraw szkolnictwa wyższego i nauki (art. 232 ust. 3 ustawy).

Przewlekłość postępowania recenzyjnego w prawie administracyjnym rodzić może również skutki finansowe dla podmiotu doktoryzującego, a tym samym pośrednio wobec recenzentów. Stosownie do art. 149 ustawy z dnia 30 sierpnia 2002 r. o postępowaniu przed sądami administracyjnymi (dalej jako: „p.p.s.a.")[26] wojewódzki sąd administracyjny, uwzględniając skargę na bezczynność lub przewlekłe prowadzenie postępowania przez organy w sprawach określonych w art. 3 § 2 pkt. 1-4 albo na przewlekłe prowadzenie postępowania w sprawach określonych w art. 3 § 2 pkt. 4a3 wskazanej ustawy, w tym na czynności z zakresu administracji publicznej dotyczące uprawnień lub obowiązków wynikających z przepisów prawa, z wyłączeniem aktów lub czynności podjętych w ramach postępowania administra-

[25] Wyrok NSA z 20.11.2018 r., I OSK 92/17, LEX nr 2603692.
[26] T.j. Dz.U. z 2024 r. poz. 935, 1685, z 2025 r. poz. 769, 1427.

cyjnego określonego w ustawie z dnia 14 czerwca 1960 r. – Kodeks postępowania administracyjnego (Dz.U. z 2024 r. poz. 572), postępowań określonych w działach IV, V i VI ustawy z dnia 29 sierpnia 1997 r. – Ordynacja podatkowa (Dz.U. z 2023 r. poz. 2383 i 2760), postępowań, o których mowa w dziale V w rozdziale 1 ustawy z dnia 16 listopada 2016 r. o Krajowej Administracji Skarbowej (Dz.U. z 2023 r. poz. 615 z późn. zm.), oraz postępowań, do których mają zastosowanie przepisy powołanych ustaw: 1) zobowiązuje organ do wydania w określonym terminie aktu, interpretacji albo do dokonania czynności; 2) zobowiązuje organ do stwierdzenia albo uznania uprawnienia lub obowiązku wynikających z przepisów prawa; 3) stwierdza, że organ dopuścił się bezczynności lub przewlekłego prowadzenia postępowania. Jednocześnie sąd stwierdza, czy bezczynność organu lub przewlekłe prowadzenie postępowania przez organ miały miejsce z rażącym naruszeniem prawa. Sąd może ponadto orzec o istnieniu lub nieistnieniu uprawnienia lub obowiązku, jeżeli pozwala na to charakter sprawy oraz niebudzące uzasadnionych wątpliwości okoliczności jej stanu faktycznego i prawnego. Sąd jest również uprawniony orzec z urzędu albo na wniosek strony o wymierzeniu organowi grzywny w lub przyznać od organu na rzecz skarżącego sumę pieniężną do wysokości do wysokości pięciokrotnego przeciętnego wynagrodzenia miesięcznego w gospodarce narodowej w roku poprzednim, ogłaszanego przez Prezesa Głównego Urzędu Statystycznego na podstawie odrębnych przepisów. Wymaga podkreślenia, że wskazany przepis jedynie pośrednio dotyczyć może opieszałych recenzentów. Sąd administracyjny nałożyć może karę jedynie na organ, który kolejno mógłby dochodzić jej w postępowaniu regresowym od recenzenta, przy czym postępowanie takie obarczone byłoby istotnym ryzkiem i m.in. usiłowaniem wykazywania przez recenzentów braku winy w niedochowaniu terminu. Determinowałoby to konieczność opowiedzenia się orzecznictwa, czy przy zachowaniu terminu na sporządzenie recenzji opowiadać należy się za jego wykładnią obiektywną, subiektywną lub obiektywno-subiektywną, bądź też podnoszeniem przez recenzentów winy w wyborze, tj. ustaleniu czy recenzent odpowiada za naruszenie terminu na sporządzenie opinii niezależnie od swojej winy, czy też wina w zawinieniu jest elementem kluczowym w uznaniu odpowiedzialności. Kwestii takiej nie rozstrzyga obowiązujący art. 221 ust. 7 ustawy. Innymi słowy, nie wiadomo, czy naruszenie terminu na sporządzenie recenzji o nadanie stopnia naukowego doktora habilitowanego odnosi się do naruszenia obiektywnego, czy też brane są pod uwagę czynniki subiektywne recenzenta, tj. przede wszystkim brak winy w niedochowaniu terminu[27]. H. Izdebski wskazuje, że sankcja ta nie może dotyczyć przypadków, za które recenzent nie może ponosić

[27] Por. T. Jędrzejewski, *Komentarz do art. 221*, [w:] *Prawo o szkolnictwie wyższym*…, red. J. Woźnicki, s. 581.

odpowiedzialności[28]. Podważyć należy zatem prawidłowość istniejącej regulacji z uwagi na jej niezupełność i brak uwzględnienia ochrony interesu recenzentów.

Rozpatrując skutki prawne niewykonania recenzji rozprawy doktorskiej wobec recenzenta, należy wskazać, że zgodnie z art. 183 ustawy o szkolnictwie wyższym nauczyciel akademicki oraz pracownik naukowy nie może bez uzasadnionej przyczyny uchylić się od pełnienia m.in. w postępowaniu w sprawie nadania stopnia doktora funkcji promotora, promotora pomocniczego czy recenzenta. Przepis ten dotyczy wyłącznie nauczycieli akademickich i pracowników naukowych czynnych zawodowo. Ustawa nie wprowadza jednak zakazu, by promotorem, promotorem pomocniczym lub recenzentem nie mógł zostać emerytowany nauczyciel akademicki lub pracownik naukowy, z tym że osoby takie mogą bezwarunkowo odmówić przyjęcia takiej funkcji. Osoby czynne zawodowe powinny natomiast swoją odmowę zawsze uzasadnić, wskazując na istnienie uzasadnionej przyczyny odmowy. Brak wykonania recenzji w określonym terminie może zostać uznany za naruszenie obowiązków pracowniczych i rodzić skutki prawne w relacji pomiędzy pomiotem zatrudniającym recenzenta i recenzentem. Odpowiedzialność dyscyplinarna w przypadku nauczycieli akademickich stanowi rodzaj odpowiedzialności pracowniczej[29], na co wskazuje pośrednio art. 302 ust. 4 ustawy. Zgodnie z art. 275 ust. 1 ustawy o szkolnictwie wyższym nauczyciel akademicki podlega odpowiedzialności dyscyplinarnej za przewinienie stanowiące czyn uchybiający obowiązkom nauczyciela akademickiego lub godności zawodu nauczyciela akademickiego[30]. Sporządzenie recenzji z przekroczeniem ustawowego terminu stanowi w każdym przypadku naruszenie obowiązków nauczyciela akademickiego. Pośrednio podkreśla to także Kodeks Etyki Pracownika Naukowego PAN z 5 grudnia 2025 roku[31], który w § 49 pkt. 4 wskazuje, iż inicjowanie, realizacja lub tolerowanie działań

[28] H. Izdebski, *op. cit.*

[29] Z. Sypniewski, *Odpowiedzialność dyscyplinarna jako odpowiedzialność prawa pracy*, [w:] *Odpowiedzialność pracownicza. Materiały XI Zjazdu Szkoły Prawa Pracy*, red. W. Sanetra, Wrocław 1984, s. 108. Por. K. Wiliński, A. Szatkowska, *Odpowiedzialność dyscyplinarna nauczycieli akademickich*, „Studia Iuridica" 2018, nr 78, s. 497-516; M. Rogacka-Rzewnicka, *Charakter odpowiedzialności dyscyplinarnej studentów i pracowników wyższych uczelni. Refleksje aksjologiczno-prawne*, „Studia Iuridica" 2020, nr 84, s. 7-22.

[30] Podobnie w art. 94 ust. 1 ustawy z dnia 30 kwietnia 2010 r. o Polskiej Akademii Nauk (t.j. Dz.U. z 2020 r. poz. 1796 z późn. zm.) wskazuje się, że do podstawowych obowiązków pracowników naukowych należy wykonywanie zadań statutowych jednostek naukowych, w szczególności prowadzenie badań naukowych i prac rozwojowych, ogłaszanie i upowszechnianie ich wyników oraz udział w realizacji podjętych przez jednostkę zadań w zakresie kształcenia; a także art. 51 ustawy z dnia 30 kwietnia 2010 r. o instytutach badawczych (t.j. Dz.U. z 2024 r. poz. 534 z późn. zm.), zgodnie z którym pracownicy naukowi i badawczo-techniczni ponoszą odpowiedzialność dyscyplinarną za rażące naruszenie obowiązków lub uchybienie godności pracownika nauki.

[31] https://publikacje.pan.pl/book/156751/kodeks-etyki-pracownika-naukowego-code-of-ethics-for-researchers?_gl=1*lcpmmy*_ga*MTE1ODg2NTM5Ny4xNzY5MjQ1MzA5*_ga_TKV678S29R*czE3NjkyNDUzMDkkbzEkZzAkdDE3NjkyNDUzMDkkajYwJGwwJGgw [dostęp: 15.01.2026].

mających na celu wywieranie negatywnego wpływu na relacje między pracownikami naukowymi, a także na ich kariery lub awanse naukowe stanowi naruszenie zasad etyki pracownika naukowego. Do takich działań należą w szczególności sporządzenie nierzetelnych recenzji prac dyplomowych lub doktorskich bądź osiągnięć w postępowaniach habilitacyjnych lub profesorskich oraz w procedurach zatrudnieniowych w instytucjach prowadzących badania. Podkreślenia wymaga też, że komisja dyscyplinarna jest związana opinią komisji do spraw etyki w nauce PAN (art. 278 ust. 7 ustawy). W praktyce jednak odpowiedzialność taka nie jest stosowana. Odpowiedzialność ta nie mogłaby zostać zastosowana wobec np. recenzentów zagranicznych[32]. Zaznaczenia wymaga też, że wskazana odpowiedzialność recenzenta ma charakter pośredni, gdyż sporządzenie recenzji nie wynika ze stosunku zatrudnienia, lecz z uchwały organu doktoryzującego. Tym samym wszczęcie postępowania dyscyplinarnego wobec nieterminowego recenzenta wymaga aktywności organu doktoryzującego.

W ramach sankcji dyscyplinarnej wobec nieterminowych recenzentów P. Mroczkowski wskazuje, że – dla przykładu – możliwym rozwiązaniem jest nałożenie na podmioty doktoryzujące obowiązku zgłaszania każdego przypadku opóźnienia w sporządzeniu rozprawy doktorskiej do Rady Doskonałości Naukowej i sporządzenie swoistej „czarnej listy" recenzentów. W obecnym stanie prawnym takie wykazy nierzetelnych recenzentów mogłyby być tworzone przez same podmioty doktoryzujące na swój własny użytek[33].

W cytowanym już Poradniku zawarty został wzór umowy dotyczących sporządzenia recenzji rozpraw doktorskiej (poradnik s. 27-30). W § 2 ust. 8 wskazanego wzoru zaznaczono, iż w sytuacji złożenia recenzji wykonanej w sposób nieprawidłowy lub niepełny Zamawiający może wskazać zastrzeżenia, zażądać od Recenzenta poprawienia lub uzupełnienia recenzji oraz wyznaczyć termin na ich dokonanie, wstrzymując przy tym wypłatę wynagrodzenia do czasu poprawienia lub uzupełnienia recenzji, pod rygorem – w przypadku ich niewykonania – uznania umowy za niewykonaną. Podstawą uznania umowy za niewykonaną może być także rażące naruszenie terminu wykonania niniejszej umowy (s. 28 Poradnika). W doktrynie[34] podkreśla się, że obecny stan prawny nie stwarza obowiązku prawnego zawierania umów cywilnoprawnych w celu uregulowania stosunku zobowiązaniowego pomiędzy recenzentem a podmiotem doktoryzującym. Wskazuje się, że wskazany wzór umowy w poradniku stanowi relikt poprzedniego stanu prawnego, tj. art. 30 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz

[32] P. Mroczkowski, *op. cit.*

[33] *Ibidem*.

[34] H. Izdebski, *op.cit.*; *Prawo o szkolnictwie wyższym…*, red. W. Kiełbasiński, M. Kuliński, B. Pietrzyk-Tobiasz, s. 446-467; M. Dokowicz, *op. cit.*, s. 480-481.

o stopniach i tytule w zakresie sztuki[35], zgodnie z którym jednostka organizacyjna przeprowadzająca przewód doktorski, postępowanie habilitacyjne lub o nadanie tytułu zawiera z recenzentem umowę, która określa w szczególności termin sporządzenia recenzji [...], wysokość wynagrodzenia oraz kary umowne za niedotrzymanie jej warunków. Obecnie wynagrodzenie recenzenta w postępowaniu w sprawie nadania stopnia doktora wynosi 27% wynagrodzenia profesora. P. Mroczkowski wskazuje, iż prawo do wynagrodzenia z tytułu sporządzonej recenzji – określone w jednoznacznej wysokości poprzez odwołanie do minimalnego wynagrodzenia zasadniczego profesora – ma charakter publicznego prawa podmiotowego, a co za tym idzie, nie jest dopuszczalne jego uszczuplenie w przypadku opóźnienia w sporządzeniu recenzji[36]. Z kolei Marcin Dokowicz[37] i Łukasz Kierznkowski[38] zdają się postulować wprowadzenie sankcji finansowych dla nieterminowych recenzentów. Widoczne są zatem dwa odmienne stanowiska – RDN oraz doktryny – w przedmiocie możliwości obniżenia wynagrodzenia recenzenta w przypadku nienależytego (tu: nieterminowego) sporządzenia recenzji. Istniejący stan i stanowisko RDN wynikać może z dwóch czynników. Albo Rada Doskonałości Naukowej w sposób niezupełny i nieuwzględniający zmian prawnych stworzyła cytowany Poradnik, wskazując w nim wzór umowy cywilnoprawnej, albo też jest to podjęta przez nią próba uzupełnienia niepełnej regulacji w omawianym zakresie. Niepełność regulacji objawia się brakiem możliwości pociągnięcia recenzentów do odpowiedzialności na gruncie ustawy w przypadku uchybienia terminowi sporządzenia recenzji określonemu w art. 190 ust. 3 ustawy. Zawarcie umowy z recenzentem w obecnym stanie prawnym i obniżenie mu wynagrodzenia określonego ustawowo na mocy przepisów umownych za nieterminowe sporządzenie recenzji naruszałoby art. 184 ust. 1 i 3 ustawy. Znamienne jest jednak, że jedynie ust. 2 i 4 art. 184 ustawy nakazują wypłatę wynagrodzenia w określonym terminie. Ustawa nie nakazuje natomiast wypłaty wynagrodzenia recenzentowi. Tym samym należy przyjąć, że ustawodawca milcząco nakazuje indywidualne uregulowanie kwestii wypłaty wynagrodzenia recenzenta, w tym w szczególności terminu wypłaty. Ustawa mówi wszak jedynie o tym, że wynagrodzenie recenzentowi „przysługuje" i w jakiej wysokości, ale nie nakazuje jego wypłaty. Idąc dalej, podnieść można, iż ustawodawca uznaje wypłatę wynagrodzenia jedynie w przypadku terminowego sporządzenia recenzji, co jednak nie znajduje uzasadnienia w istniejącej regulacji i nie jest stosowane w praktyce.

[35] T.j. Dz.U. z 2017 r. poz. 1789 z późn. zm.
[36] P. Mroczkowski, *op. cit.*.
[37] M. Dokowicz, *op. cit.*, s. 480.
[38] Ł. Kierznowski, [w:] *Stopnie naukowe i stopnie w zakresie sztuki. Komentarz*, Warszawa 2021, art. 184, https://sip.lex.pl/#/commentary/587931746/729186/kierznowski-lukasz-stopnie-naukowe-i-stopnie-w-zakresie-sztuki-komentarz [dostęp: 15.01. 2026].

Na koniec należy podkreślić również możliwość odpowiedzialności podmiotu doktoryzującego oraz Rady Doskonałości Naukowej w przypadku postępowania odwoławczego na podstawie art. 417 k.c. i ewentualnych roszczeń regresowych w stosunku do recenzenta. Pociągnięcie do odpowiedzialności podmiotu doktoryzującego wymaga wydania aktu stwierdzającego, że doszło do przewlekłości prowadzenia postępowania, jak m.in. postanowienie wydane w wyniku rozpatrzenia ponaglenia, względnie zaświadczenie[39].

4. Skutki prawne dla ubiegających się o stopień doktora

Jednostką najbardziej poszkodowaną w przypadku naruszenia przez recenzentów terminu na sporządzenie recenzji jest doktorant. Niezależnie od wyniku pozostaje on w zawieszeniu co do rozstrzygnięcia postępowania lub ewentualnego rozpoczęcia przygotowań do wejścia na inną ścieżkę naukową. Zatem należy poddać analizie, czy posiada on uprawnienia do egzekwowania stosowania sankcji wobec recenzenta opóźniającego się ze sporządzeniem recenzji dysertacji doktorskiej.

Na gruncie odpowiedzialności dyscyplinarnej może on domagać się wszczęcia w podmiocie zatrudniającym recenzenta wystosowania wobec niego stosownych środków mobilizujących lub wszczęcia postępowania dyscyplinarnego.

Na gruncie administracyjnoprawnym doktorant może wnieść ponaglenie do Rady Doskonałości Naukowej, a także skargę do wojewódzkiego sądu administracyjnego.

Na gruncie cywilnoprawnym może domagać się odszkodowania na zasadach ogólnych, w tym, w sytuacji gdy ponosi koszty postępowania o nadanie stopnia doktora, o obniżenie wynagrodzenia recenzenta.

Choć dostępne środki prawne mogą rodzić dotkliwe konsekwencje dla opieszałego recenzenta, to przedstawione uprawnienia doktoranta pozostają jednak fikcyjne z uwagi na rozpatrywaną materię. Próby ponaglenia recenzenta obarczone są ryzykiem otrzymania recenzji negatywnej. Pokrzywdzony doktorant nie będzie skłonny do przymuszenia recenzenta do terminowego wykonania swoich obowiązków z obawy przed możliwością niedopuszczenia go do dalszych etapów postępowania awansowego.

Należy również nadmienić, że tożsame skutki dla doktoranta wywołać może reakcja podmiotu doktoryzującego, co dodatkowo utrudnia egzekwowanie terminu wynikającego z art. 190 ust. 3 ustawy o szkolnictwie wyższym. Tym samym dostępne środki prawne egzekucji terminowości sporządzenia recenzji są niemożliwe do zastosowania z uwagi na wysokie prawdopodobieństwo pokrzywdzenia ubiegającego

[39] S. Gajewski, A. Jakubowski, *Nieakt w prawie administracyjnym*, „Zeszyty Naukowe Sądownictwa Administracyjnego" 2013, nr 6, s. 73-88.

się o nadanie stopnia doktora. Stanowi to naruszenie zasady sprawiedliwości proceduralnej poprzez naruszenie gwarancji procesowych ubiegającego się o stopnień naukowy doktora poprzez brak możliwości zapewnienia realizacji zasady szybkości postępowania, tak od strony formalnoprawnej, jak i faktycznej.

5. Wnioski

Rola recenzentów w postępowaniu awansowym jest istotna. Determinuje ona możliwość dopuszczenia kandydata do stopnia naukowego do dalszych etapów postępowania. Odpowiedzialność recenzenta za nieterminowe sporządzenie recenzji może mieć charakter zarówno dyscyplinarny, cywilnoprawny, administracyjnoprawny, jak i na gruncie prawa pracy. Opieszały recenzent może zostać usankcjonowany za nienależyte wykonywanie zadań wynikających ze stosunku zatrudnienia, jak i za nienależyte wykonanie zobowiązania wynikającego z umowy o sporządzenie recenzji. Odpowiadać może również majątkowo wobec podmiotu doktoryzującego w przypadku stwierdzenia przewlekłości lub bezczynności postępowania awansowego. W odniesieniu do doktoranta recenzent może odpowiadać również za szkody na zasadach ogólnych. Niemniej jednak ani doktorant, ani podmiot doktoryzujący nie będą skłonni do pociągania nieterminowego recenzenta do wskazanych sankcji z uwagi na istotne ryzyko pokrzywdzenia doktoranta. Chociaż zasady etyki zawodowej zobowiązują recenzenta do zachowania obiektywizmu przy sporządzaniu recenzji, to praktyka i realia społeczne dają podstawy sądzić, że w przypadku próby wymuszenia na recenzencie terminowego wykonania swoich zobowiązań może on podjąć zachowania nacechowane odwetowością za naruszenie jego autorytetu jako specjalisty.

Należy zatem postulować, aby wprowadzone zostały środki respektowania przestrzegania przeprowadzana były z urzędu przez organ wyższego stopnia, tj. Radę Doskonałości Naukowej. Wymagane, konieczne zmiany regulacji wymagają gruntownego przemyślenia. Najprostszym rozwiązaniem wydaje się recypowanie tożsamej sankcji, co w przypadku postępowań o nadanie stopnia naukowego doktora habilitowanego. Jednakże, jak zaznaczono, rozwiązanie to obarczone jest wadą braku uwzględnienia realizacji zasady sprawiedliwości proceduralnej wobec recenzentów naruszających ustawowy termin sporządzenia recenzji z przyczyn od nich niezależnych. Konieczność ochrony wszystkich uczestników postępowania o nadanie stopnia doktora przemawia jednak za wprowadzeniem regulacji bardziej kompleksowej, która nie rodziłaby potencjalnego zagrożenia dla doktoranta z uwagi na stosowane wobec recenzenta środki mobilizujące. Jedną z możliwości wydaje się być ustawowe obniżenie wynagrodzenia recenzenta w uzależnieniu od okresu opóźnienia. Może to jednak stanowić okoliczność na godzenie się recenzentów na

otrzymanie niższego wynagrodzenia w zamian za wydłużenie terminu sporządzenia recenzji. Tym samym optymalnym rozwiązaniem wydaje się być zaangażowanie RDN w proces kontroli terminowości sporządzania recenzji w postępowaniach o nadanie stopnia doktora. To z kolei wymaga stworzenia jednolitego elektronicznego systemu rejestracji postępowań w sprawie o nadanie stopnia naukowego doktora, w którym organy uprawnione do przeprowadzenia takich postępowań obowiązane byłyby do wprowadzania danych powołanych recenzentów wraz z datami rozpoczęcia biegu terminu na sporządzenie recenzji (wynikającymi z przepisów o doręczeniu zgodnie z k.p.a. i zgodnie z art. 190 ust. 2 ustawy o szkolnictwie wyższym). System powinien automatycznie informować organ prowadzący postępowanie awansowe o wystąpieniu opóźnienia po stronie recenzenta. W dalszej kolejności konieczne są zmiany uprawnień mobilizujących organów prowadzących postępowania awansowe oraz Rady Doskonałości Naukowej. Rozważyć można obligatoryjne wymierzanie z urzędu sankcji wobec nieterminowych recenzentów przez wskazane organy na podstawie danych dostępnych w jednolitym rejestrze, względnie automatycznego przypisywania opóźnienia recenzentowi w systemie. Sankcje te powinny jednak podlegać kontroli instancyjnej i możliwości uchylenia się od nich przez recenzentów, gdy opóźnienie w sporządzeniu recenzji nie wynikało z ich winy. Tożsame uprawnienia winny zostać przyznane Radzie Doskonałości Naukowej wobec recenzentów w postępowaniu odwoławczym, z tym że w tym przypadku powinna istnieć możliwość w systemie modyfikacji terminu sporządzenia recenzji (opinii) według zaleceń Rady Doskonałości Naukowej ustalanych indywidualnie. Kolejną kwestią wymagającą zmian w prawie jest ustalenie terminu wypłaty wynagrodzenia recenzentowi.

Poczynione refleksje potwierdzają niedostateczny stan gwarancji proceduralnych w postępowaniu o nadania stopnia naukowego doktora. Istniejący stan prawny zezwala na recenzowanie rozpraw doktorskich z naruszeniem art. 190 ust. 3 ustawy o szkolnictwie wyższym bez realnych możliwości wymierzenia jakichkolwiek sankcji wobec opieszałych recenzentów. Niezbędne jest zatem podjęcie stosownych działań legislacyjnych w tym zakresie.

Literatura

Bahishti A., *The importance of review articles & its prospects in scholarly literature*, „Extensive Reviews" 2021, nr 1(1).

Black N., Van Rooyen S., Godlee F., Smith R., Evans S., *What makes a good reviewer and a good review for a general medical journal*, „Jama" 1998, nr 280(3).

Bloom F., *The importance of reviewers*, „Science" 1999, nr 283(5403).

Bochentyn A., *Dowód z opinii biegłego w jurysdykcyjnym postępowaniu administracyjnym*, Warszawa 2020.

Bogusz M., *Charakter prawny uchwał rady wydziału w przewodzie doktorskim*, „Gdańskie Studia Prawnicze" 2015, nr 33.
Bogusz M., *Zagadnienie dopuszczalności zbiegu ról procesowych członka komisji habilitacyjnej i członka organu habilitującego*, „Prawo i Więź" 2025, nr 5(58).
Bogusz M., Jędrzejczak M., *Zagadnienie wyłączenia członka organu doktoryzującego i członka komisji doktorskiej z powodu pełnienia funkcji recenzenta w postępowaniu doktorskim*, [w:] *Administracja publiczna wobec procesów zmian w XXI wieku. Księga jubileuszowa Profesora Jerzego Korczaka*, red. P. Lisowski, Wrocław 2024, s. 897-907.
Cannegieter S., Lisman T., *What is in it for the reviewer?*, „Journal of Thrombosis and Hemostasis" 2024, nr 22(3).
Dańczak P., [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. A. Jakubowski, wyd. 1, 2023, art. 190, Warszawa 2023.
Dokowicz M., *Komentarz do art. 184*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Woźnicki, Warszawa 2019.
Dokowicz M., *Komentarz do art. 190*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Woźnicki, Warszawa 2019.
Gajewski S., Jakubowski A., *Nieakt w prawie administracyjnym*, „Zeszyty Naukowe Sądownictwa Administracyjnego" 2013, nr 6/2013.
Herzog P., *The Importance of Peer Review: Recommendations for Reviewers and Authors*, „Review of Religious Research" 2023, nr 65(1).
Izdebski H., [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Zieliński, H. Izdebski, wyd. 3, LEX/el. 2023, art. 190, https://sip.lex.pl/#/commentary/587785502/731978/izdebski-hubert-zielinski-jan-michal-prawo-o-szkolnictwie-wyzszym-i-nauce-komentarz-wyd-iii.
Jędrzejewski T., *Komentarz do art. 221*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. J. Woźnicki, Warszawa 2019.
Kiełbasiński W., [w:] *Prawo o szkolnictwie wyższym i nauce dla studenta i doktoranta – komentarz praktyczny*, red. W. Kiełbasiński, M. Kuliński, B. Pietrzyk-Tobiasz, Warszawa 2024.
Kierznowski Ł., [w:] *Stopnie naukowe i stopnie w zakresie sztuki. Komentarz*, Warszawa 2021, art. 184, https://sip.lex.pl/#/commentary/587931746/729186/kierznowski-lukasz-stopnie-naukowe-i-stopnie-w-zakresie-sztuki-komentarz.
Kmieciak Z., Wegner J., *Odpowiednie stosowanie przepisów kodeksu postępowania administracyjnego w sprawach nadawania stopni naukowych i tytułu profesora*, „Państwo i Prawo" 2021, nr 3.
Kodeks Etyki Pracownika Naukowego PAN z 5 grudnia 2025 roku, https://publikacje.pan.pl/book/156751/kodeks-etyki-pracownika-naukowego-code-of-ethics-for-researchers?_gl=1*lcpm-my*_ga*MTE1ODg2NTM5Ny4xNzY5MjQ1MzA5*_ga_TKV678S29R*czE3NjkyNDUzMDkkbzEkZzAkdDE3NjkyNDUzMDkkajYwJGwwJGgw.
Krauz M., *Krytyk i/a recenzent – znaczenie pojęć w słowniku i w tekście*, „SŁOWO. Studia Językoznawcze" 2021, nr 12.
Łazuk S., *Wybrane problemy stosowania Kodeksu postępowania administracyjnego w procedurze awansu naukowego*, [w:] *Specyfika postępowań administracyjnych w sprawach z zakresu szkolnictwa wyższego i nauki*, red. J. Tarno, A. Szrot, P. Pokorny, Lublin 2016.
Mroczkowski P., [w:] *Prawo o szkolnictwie wyższym i nauce dla studenta i doktoranta – komentarz praktyczny*, red. W. Kiełbasiński, B. Pietrzyk-Tobiasz, M. W. Kuliński, wyd. 1, Warszawa 2024, art. 190.
Nightingale P., Scott A., *Peer review and the relevance gap: ten suggestions for policy-makers*, „Science and Public Policy" 2007, nr 34(8).
Pfeiffer T., Tran L., Krumme C., Rand D., *The value of reputation*, „Journal of the Royal Society Interface" 2012, nr 9(76).
Poradnik Rady Doskonałości Naukowej z 2022 roku „Recenzje w Postępowaniach o awans naukowy", https://www.rdn.gov.pl/dobre-praktyki.poradnik-recenzje-w-postepowaniach-o-awans-naukowy.html.
Rogacka-Rzewnicka M., *Charakter odpowiedzialności dyscyplinarnej studentów i pracowników wyższych uczelni. Refleksje aksjologiczno-prawne*, „Studia Iuridica" 2020, nr 84.
Scott A., *Peer review and the relevance of science*, „Futures" 2007, nr 39(7).

Sieniuć M., *Jednostka w postępowaniach w sprawach nadania stopnia i tytułu naukowego. Studium z prawa administracyjnego procesowego. Studium z prawa administracyjnego procesowego*, Łódź 2019.
Sitek B., Woźniak A., *Komentarz do art. 238*, [w:] *Prawo zatrudnienia pracowników szkół wyższych. Komentarz do wybranych przepisów Prawa o szkolnictwie wyższym i nauce*, red. K. Baran, Warszawa 2026.
Sypniewski Z., *Odpowiedzialność dyscyplinarna jako odpowiedzialność prawa pracy*, [w:] *Odpowiedzialność pracownicza. Materiały XI Zjazdu Szkoły Prawa Pracy*, red. W. Sanetra, Wrocław 1984.
Tarno J., *Rola odpowiedniego stosowania przepisów k.p.a. w postępowaniach w sprawach stopni naukowych (wybrane zagadnienia)*, „Zeszyty Naukowe Sądownictwa Administracyjnego" 2011, nr 6.
Trzebiatowski M., *Kryteria prawne nadawania tytułu profesora w zakresie sztuki – analiza normatywno-empiryczna*, „Prawo i Więź" 2022, nr 4(42).
Walker R., Barros B., Conejo R., Neumann K., Telefont M., *Personal attributes of authors and reviewers, social bias and the outcomes of peer review: a case study*, „F1000Research" 2015, nr 4(21).
Wang J., *The Importance of Reviewers*, „Global Spine Journal" 2015, nr 5(2).
Waszkiewicz P., *Nadawanie stopnia doktora*, [w:] *Prawo o szkolnictwie wyższym i nauce. Komentarz*, red. A. Jakubowski, Warszawa 2023.
Wiliński K., Szatkowska A., *Odpowiedzialność dyscyplinarna nauczycieli akademickich*, „Studia Iuridica" 2018, nr 78.

Akty prawne

Ustawa z dnia 14 czerwca 1960 r. Kodeks postępowania administracyjnego, t.j. Dz.U. z 2024 r. poz. 572, z 2025 r. poz. 769.
Ustawa z dnia 30 sierpnia 2002 r. o postępowaniu przed sądami administracyjnymi, t.j. Dz.U. z 2024 r. poz. 935, 1685, z 2025 r. poz. 769, 1427.
Ustawa z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki, t.j. Dz.U. z 2017 r. poz. 1789 z późn. zm. (uchylony).
Ustawa z dnia 30 kwietnia 2010 r. o Polskiej Akademii Nauk, t.j. Dz.U. z 2020 r. poz. 1796 z późn. zm.
Ustawa z dnia 30 kwietnia 2010 r. o instytutach badawczych, t.j. Dz.U. z 2024 r. poz. 534 z późn. zm.
Ustawa z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce, t.j. Dz.U. z 2024 r. poz. 1571, 1871, 1897, z 2025 r. poz. 619, 620, 621, 622, 1162.

Orzecznictwo

Wyrok NSA z 20.11.2018 r., I OSK 92/17, LEX nr 2603692.
Wyrok WSA w Warszawie z 29.06.2021 r., II SA/Wa 686/21, LEX nr 3293247.
Wyrok WSA w Warszawie z 3.12.2021 r., II SA/Wa 1081/21, LEX nr 3308515.

Strony internetowe

https://bip.brpo.gov.pl/pl/contetnt/rpo-opoznienia-recenzje-postepowania-awansowe-mein-kolejne-odpowiedz.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Bartłomiej Nałęcz
Uniwersytet Pomorski w Słupsku
ORCID: 0000-0002-6974-1435
bartlomiej.nalecz@upsl.edu.pl

Uznawanie treści erotycznych udostępnianych online za usługę kulturalną. Glosa krytyczna do wyroku NSA z 16.10.2025 r., sygn. akt I FSK 1342/22

Słowa kluczowe: usługi kulturalne, erotyka, pornografia, treści erotyczne online, VAT

Streszczenie. Glosa krytyczna analizuje wyrok NSA z 16 października 2025 r. (I FSK 1342/22), w którym uznano, że transmitowane online spektakle o tematyce erotycznej mogą stanowić usługę kulturalną zwolnioną z VAT. W tekście wskazano problemy, z jakimi wiążą się przyjęte przez sąd kryteria odróżniania erotyki od pornografii. Tekst podkreśla płynność granicy zachodzącej między tymi zjawiskami oraz trudności interpretacyjne wynikające z cyfrowej konsumpcji treści seksualnych. W konkluzji wysunięto postulat doprecyzowania przepisów prawa podatkowego w zakresie objętym rozważaniami z glosowanego orzeczenia.

Recognition of erotic content made available online as a cultural service. Critical commentary on the judgment of the Supreme Administrative Court of 16 October 2025, ref. no. I FSK 1342/22

Key words: cultural services, erotica, pornography, online erotic content, VAT

Summary. The critical commentary analyses the judgment of the Supreme Administrative Court of 16 October 2025 (I FSK 1342/22), in which it was held that erotic performances transmitted online may constitute a cultural service exempt from VAT. The paper points out the problems associated with the criteria adopted by the court to distinguish between erotica and pornography. It emphasises the fluid boundary between these phenomena and the interpretative difficulties arising from the digital consumption of sexual content. In conclusion, the author postulates the need to clarify the tax law provisions concerning the issues addressed in the commented judgment.

Teza

Nie powinno budzić wątpliwości, że tematyka erotyczna jest elementem życia codziennego, stosunków międzyludzkich, jak również immanentną częścią kultury – obecna jest w literaturze, teatrze, filmie, sztuce, np. japońskie obrazy shunga, przedstawiające japońską sztukę erotyczną. Obrazy erotyczne (w tym nagość) były i są wykorzystywane przez artystów w wielu rodzajach sztuki (np. malarstwie, te-

atrze) i nie stanowią elementu pornografii, nie mają na celu wzbudzenia w widzu pożądania seksualnego. Jeżeli zatem spektakl o tematyce erotycznej nie prezentuje penalizowanych zachowań, nie polega wyłącznie na przedstawieniu aktów seksualnych w sposób naruszający społeczne normy obyczajowe, a intencją twórcy nie jest wywołanie podniecenia seksualnego u odbiorcy, to nie można wykluczyć jego kulturalnego charakteru, jak również i tego, że odbywa się w interesie publicznym[1].

W glosowanym orzeczeniu[2] Naczelny Sąd Administracyjny dokonał analizy charakteru prawnego transmitowanych online erotycznych spektakli/pokazów artystycznych. W efekcie podjętych rozważań uznano je za formę usługi kulturalnej zwolnionej z podatku VAT na podstawie art. 43 ust. 1 pkt 33 lit. b ustawy o VAT[3]. Chronologicznie sprawę zainicjowała interpretacja indywidualna wydana przez Dyrektora Krajowej Informacji Skarbowej[4]. Stwierdzono w niej, że wprawdzie wnioskodawca jest twórcą i wykonawcą w rozumieniu prawa autorskiego, to jednak sama działalność (erotyczne spektakle online) nie stanowi jeszcze usługi kulturalnej i wobec tego nie korzysta ze zwolnienia podatkowego. Jak uargumentowano, zasadniczy ich cel sprowadza się do wywołania pobudzenia seksualnego, a nie realizacji interesu publicznego. Interpretację tę uchylił Wojewódzki Sąd Administracyjny[5]. W następstwie złożenia skargi kasacyjnej Naczelny Sąd Administracyjny zbadał prawidłowość interpretacji uchylającej, uznając, że organ błędnie zinterpretował pojęcie usług kulturalnych. Tym samym skargę oddalono, wskazując m.in., że usługi kulturalne w rozumieniu przywołanej regulacji podatkowej należy wykładać szeroko, sama zaś tematyka erotyczna nie wyklucza kulturalnego charakteru usługi, o ile nie ma ona charakteru pornograficznego ani nie narusza norm penalizowanych.

Rozważany pogląd w swoim zasadniczym przekazie podtrzymuje wcześniejsze stanowiska wyrażane w kontekście prawnopodatkowej oceny erotycznych spektakli/pokazów erotycznych[6]. Ponowne wyartykułowanie dotychczas prezentowanych tez stanowi potwierdzenie i utrwalenie kształtowanej w tym zakresie linii orzeczniczej. Jej analiza ujawnia jednak obszary problemowe, koncentrujące się wokół dwóch

[1] Teza wyodrębniona przez autora glosy jako element argumentacji zaakceptowanej przez sąd kasacyjny.

[2] Wyrok NSA z 16.10.2025 r., I FSK 1342/22, LEX nr 3986992.

[3] Ustawa z dnia 11 marca 2004 r. o podatku od towarów i usług, t.j. Dz.U. z 2025 r. poz. 775 z późn. zm.

[4] Pismo z dnia 16.09.2021 r. wydane przez: Dyrektor Krajowej Informacji Skarbowej, 0114-KDIP4-2.4012.419.2021.3.AS, https://sip.mf.gov.pl.

[5] Wyrok WSA w Gliwicach z 19.04.2022 r., I SA/Gl 1527/21, LEX nr 3346729.

[6] Twierdzenia o tym, że erotyka jest obecna w kulturze i sztuce, jak również o braku podstaw do wykluczania dzieł o tematyce erotycznej z kręgu wytworów z zakresu kultury, sformułowano również m.in. w: wyroku WSA w Krakowie z 12.05.2022 r., I SA/Kr 1574/21, LEX nr 3443552; wyroku WSA w Krakowie z 29.07.2022 r., I SA/Kr 1553/21, LEX nr 3395127.

zasadniczych zagadnień. Po pierwsze, miejsca, w którym przebiega granica między erotyką jako formą ekspresji artystycznej mieszczącej się w polu znaczeniowym pojęcia usług kulturalnych, a pornografią, która nie korzysta ze zwolnienia podatkowego. Po drugie, kryteriów normatywnych i funkcjonalnych służących dokonywaniu takiej kwalifikacji, aby uniknąć arbitralności ocen, zapewniając spójność wykładni przepisów prawa podatkowego. Wedle propozycji z glosowanego orzeczenia rozwikłaniu tych kwestii ma służyć test treści, która nie może prezentować penalizowanych zachowań oraz polegać wyłącznie na przedstawieniu aktów seksualnych z naruszeniem społecznych norm obyczajowych, a intencja twórcy nie sprowadza się do wywołania podniecenia seksualnego u odbiorcy.

Weryfikując przyjętą przez sąd metodę rozgraniczania erotyki i pornografii, trzeba mieć na uwadze, że operują one podobną treścią. Niewątpliwie motywem obu zjawisk jest seksualność. Wszystkie te trzy pojęcia, tj. erotykę, pornografię i seksualność, trudno jednak jednoznacznie zaklasyfikować czy nawet wyraźnie od siebie oddzielić[7]. Należy zatem zakładać pewną ich zbieżność. W glosowanym orzeczeniu jednak takiego rozdziału dokonano. Jednocześnie stanowczo spozycjonowano erotykę jako wycinek obszaru usług kulturalnych, odmawiając takiego zaszeregowania pornografii. Tym samym za kryterium rozróżniające przyjęto charakter rzeczonej usługi, która niejako poza swoim obrębem lokuje pornografię.

W kolejnym kroku należałoby ustalić znaczenie pojęcia usługi kulturalnej, którego wyostrzenie rodzi autonomiczne trudności. Niełatwo bowiem o konkretniejsze wskazania, gdy łączy się ją z upowszechnianiem oraz ochroną kultury rozumianej jako materialne i niematerialne dziedzictwo ludzkości, zwłaszcza w obszarze historii, nauki i sztuki[8]. W tak szeroką definicję usług możliwe jest wpisanie każdego wytworu ludzkiej działalności[9]. W pewnym stopniu precyzujące, lecz kłopotliwe, jest też wiązanie ich z jednoznacznie pozytywnymi wartościami jak prawda, sprawiedliwość, wolność czy równość[10]. Należałoby wówczas przyznać, że erotyka jest

[7] Za świadectwo tego posłużyć może definicja pornografii zaprezentowana przez Janusza Wojciechowskiego, zgodnie z którą są nią: „pisma, druki, przedstawienia, filmy, zdjęcia, obrazy, rysunki i inne przedmioty treści nieprzyzwoitej, obliczone na wywołanie podniecenia erotycznego u odbiorców" (J. Wojciechowski, *Kodeks karny. Komentarz. Orzecznictwo*, Warszawa 1997, s. 357). W tej propozycji podniecenie erotyczne utożsamia się z podnieceniem seksualnym. Pornografia ma zatem doprowadzić do czegoś, co ma mieć wymiar erotyczny. W takim ujęciu to, co pornograficzne, jest też erotyczne, zaś to, co erotyczne, jest też seksualne. To z kolei uprawnia twierdzenie, że to, co pornograficzne, jest seksualne.

[8] A. Bartosiewicz, *VAT. Komentarz*, Warszawa 2025, s. 730.

[9] Również prezentowanie aktów seksualnych online.

[10] Taką charakterystykę, formułowaną na podstawie źródeł encyklopedycznych, można spotkać w interpretacjach wprawdzie indywidualnych, lecz przedstawianych przez organy odpowiedzialne za zapewnienie jednolitej informacji podatkowej i w Polsce (tak w Piśmie z dnia 11.07.2025 r., wydanym przez Dyrektora Krajowej Informacji Skarbowej, 0114-KDIP4-2.4012.403.2025.1.KS, Zwolnienie usług kulturalnych od podatku VAT, https://sip.mf.gov.pl).

czymś w istocie pozytywnym, a z punktu widzenia społecznego pożądanym. Taka konstatacja nie przekonuje, gdy uwzględni się ścisły związek erotyki z seksualnością oraz zjawiska osiągania satysfakcji w tej sferze. Nie można bowiem pomijać, że sama swoboda jednostki w sferze seksualnej nie ogranicza się wyłącznie do celów prokreacyjnych czy relacyjnych. Nierzadko łączy się również z chęcią dostarczenia sobie rozrywki lub przyjemności[11]. To wprowadza do sfery wolności seksualnej rozumianej jako prawo do osiągania satysfakcji seksualnej. W tym obszarze należy natomiast przewidywać niestandardowe rozwiązania, które niekoniecznie spotykać się muszą z akceptacją większości społeczeństwa[12]. Utrwalenie takich aktów seksualnych może jednak prowadzić do ich zakwalifikowania jako pornografii, która nie podlega penalizacji[13]. Polskie prawo karne nie zakazuje bowiem obcowania z pornografią jako taką, lecz tylko z określonymi jej formami[14]. Jeśli zatem treść przedstawiająca wskazane niestandardowe rozwiązanie nie poddaje się takiej kwalifikacji, należy na poziomie ogólnym uznać ją za legalną pornografię. To z kolei zbliża ją do erotyki, która z perspektywy legalności pozostaje kategorią tożsamą.

W wątpliwość podać należy również prawidłowość kryterium ustalającego, czy dana treść narusza społeczne normy obyczajowe, co ma odróżniać pornografię od treści erotycznej traktowanej jako usługa kulturalna. Wydaje się bowiem, że istotowo pokrywa się ona z rozstrzyganiem, czy analizowany materiał jest penalizowany, czy nie. Dąży bowiem do udzielenia tej samej odpowiedzi, którą uzyskać można w efekcie zastosowania wyżej omówionego kryterium. Typizacje dotyczące pornografii są bowiem przez ustawodawcę konstruowane jako przestępstwa skierowane przeciwko obyczajności[15]. Innymi słowy, uzasadnienie kryminalizacji danej treści pornograficznej tkwi w ochronie wartości społecznych w postaci – obok wolności

[11] L. Boul, R. Hallam-Jones, K.R. Wylie, *Sexual Pleasure and Motivation*, „Journal of Sex & Marital Therapy" 2008, nr 35, s. 25-39.

[12] Przykładem zachowań mieszczących się w tej sferze są różnego rodzaju nietypowe praktyki seksualne, np. sadomasochizm czy asfiksjofilia. Podejmowane są one za zgodą partnerów, choć mogą wiązać się z ryzykiem utraty życia lub spowodowania ciężkiego uszczerbku na zdrowiu – za M. Grudecki, *Kontratypy pozaustawowe w polskim prawie karnym*, Warszawa 2021, s. 164.

[13] W tym świetle można również owe niestandardowe rozwiązania zrównać z charakterystyką znamienia treści pornograficznej sformułowaną przez Sąd Najwyższy, zgodnie z którą stanowią ją: „czynności seksualne człowieka sprzeczne z przyjętymi w społeczeństwie wzorcami zachowań seksualnych" – wyrok SN z 23.11.2010 r., IV KK 173/10, LEX nr 667510.

[14] Choćby wskazanych art. 202 § 3 ustawy z dnia 6 czerwca 1997 r. Kodeks karny (t.j. Dz.U. z 2025 r. poz. 383 z późn. zm.), tj. k.k., w tym m.in. prezentujących przemoc. W tym miejscu wypada odnotować, że w doktrynie i orzecznictwie często przyjmuje się, iż nawet przemoc pozorowana w pornografii podpada pod kryminalizację, ale pogląd ten wymaga krytycznej weryfikacji – szerzej zob. B. Nałęcz, *Inscenizacja przemocy w pornografii a zakres kryminalizacji art. 202 § 3 k.k.*, „Prawo i Więź" 2025, nr 6, s. 817-840.

[15] Wskazuje na to już tytuł Rozdziału XXV k.k.

seksualnej – właśnie norm obyczajowych[16]. Za tym idą poważne dla analizowanej tezy konsekwencje. Kryterium naruszenia norm obyczajowych nie może pełnić roli instrumentu odróżniającego pornografię od erotyki, albowiem pełni już funkcję aksjologicznej racji kryminalizacji scharakteryzowanych w przepisach typizujących dotyczących określonych w nich treści pornograficznych[17].

Niezależnie od sygnalizowanych nieścisłości weryfikacja treści przeprowadzona na podstawie tych kryteriów potwierdza znany w literaturze pogląd, zgodnie z którym pornografię i erotykę ma odróżniać objęcie rozpowszechniania tej pierwszej ograniczeniami ustawowymi, podczas gdy rozpowszechnianie treści erotycznych takimi ograniczeniami nie jest obciążone[18]. W jego świetle omawiane zagadnienie sprowadza się do problemu odróżniania niekryminalizowanej pornografii i erotyki. Granica między nimi wciąż jawi się jednak jako nader płynna. Jednym ze sposobów rozróżnienia może być przyjęte w orzecznictwie kryterium rodzaju satysfakcji, jaką odbiorca czerpie z takich treści. W tej koncepcji erotykę ma cechować wyższy poziom estetyki względem tego właściwego dla pornografii[19]. Ma ona tę zaletę, że obciąża pornografię charakteryzującą ją negatywną konotacją.

Inną rzeczą jest racjonalny powód takiego obciążania. Nie sposób bowiem ignorować trwającego procesu przeobrażania kanonów estetycznych w treściach seksualnych. Wyrazem tego jest fakt, że przenikanie treści o cieniu pornograficznym do kultury masowej nie jest zjawiskiem czekającym na odkrycie[20]. Kryterium

---

[16] Por. P. Kozłowska, M. Mozgawa, *Prawnokarne aspekty rozpowszechniania pornografii (analiza dogmatyczna i praktyka ścigania)*, „Prokuratura i Prawo" 2002, nr 3, s. 15.

[17] Konsekwencją tego jest również twierdzenie, że niestandardowe praktyki seksualne, które w efekcie utrwalenia przybierają postać treści pornograficznej, różnią się od aktywności naruszających normy obyczajowe. Stąd bardziej właściwe jest twierdzenie, że te pierwsze są wyborem wynikającym z preferencji rozrywki lub przyjemności danego człowieka, który w istocie nie może godzić w normy obyczajowe.

[18] Tak P. Ślęzak, *Zasady rozpowszechniania fotografii i wizerunku*, Warszawa 2025, s. 70. Potwierdzać zdaje się to również wyróżnianie tzw. erotyki dziecięcej, czyli treści prezentujących małoletnich w seksualnym kontekście, lecz niespełniających kryteriów pornografii dziecięcej. Zaliczać się mają do nich materiały prezentujące dzieci, które służą celom seksualnym danej osoby, przykładowo fotografia przedstawiające małoletnią, pozującą nago lub będącą częściowo ubrane, przyjmującą pozę nieadekwatną do wieku, tj. na osoby starsze (zob. A. Thannhäuser, J. Pawlik-Czyniewska, *Z praktyki powoływania biegłych i opiniowania w sprawach dotyczących czynów zabronionych z art. 202 k.k.*, [w:] *Hominum causa omne ius constitutum sit. Księga jubileuszowa Profesora Piotra Hofmańskiego*, red. P. Czarnecki, S. Głogowska, A. Górski, A. Sakowicz, A. Światłowski, Warszawa 2024, s. 269).

[19] Tak w innym orzeczeniu NSA, wydanym w dacie glosowanego orzeczenia, w którym stwierdzono, że: „usługi, które wykonywał Wnioskodawca, były usługami kulturalnymi polegającymi na wystawianiu artystycznych pokazów/spektakli o charakterze erotycznym, które nie stanowiły pornografii, ale sztukę erotyki różnych kultur, np. azjatyckiej. Erotyka przedstawiana w pokazach/spektaklach była ukazywana ze smakiem i kunsztem, odbiorca miał czerpać z nich wiedzę i satysfakcję bez odczucia oglądania filmu pornograficznego" – wyrok NSA z 16.10.2025 r., I FSK 1559/22, LEX nr 4011526.

[20] Z pewnością niepokojące, lecz dające się zaobserwować jest zjawisko sukcesywnego oddalania się pornografii od klasyfikowania jej jako twórczości z marginesu. Zaczyna nawet wpływać na estetykę m.in. mediów, reklamy i mody. Zacierania granicy między pornografią a kulturą masową dowodzą

estetyki jako odróżniające erotykę i pornografię jest też problematyczne ze względu na możliwy brak sposobności jednoznacznego zakwalifikowania danej treści jako mniej estetycznej (pornograficznej). Z pewnością istnieją bowiem takie akty twórczości, które odznaczają się wyraźnym potencjałem pobudzającym seksualnie, być może nawet zasadniczo ukierunkowanych w tym aspekcie, lecz ich kompozycja lub narracja zawiera walory artystyczne, co z kolei uprawnia do potraktowania ich zarówno jako treści erotycznej, jak i pornograficznej[21]. To zaś stwarza podstawy dla polemiki, nawet falsyfikowania twierdzenia, że to cel, jakim ma być pobudzenie lub zaspokojenie seksualne odbiorcy danej treści, stanowi zasadniczy determinant klasyfikowania danej treści jako pornograficznej[22]. Wyższy poziom estetyczny, niezależnie od tego, co przedstawia, ma uzasadniać zaliczenie jej do kategorii artystycznych[23].

Niezależnie od powyższego, nie wydaje się uzasadnione ograniczanie aspektu podmiotowego interakcji online wyłącznie do jednej strony. Erotyka ma bowiem pobudzać lub intrygować seksualnie także osoby odbierające ją wspólnie, co odróżnia ją od pornografii postrzeganej jako treść konsumowana, co do zasady, indywidualnie[24]. Stąd takie poszerzenie jawi się jako nie tylko uprawnione, ale i konieczne. Tym niemniej włączenie do analizy zagadnienia woli obu stron interakcji online niewątpliwie prowadzi do dalszego jej komplikowania[25].

---

zresztą badania wskazujące na istnienie zjawiska określanego jako *pornification* lub *mainstreaming of pornography*. Polega ono właśnie na przenikaniu pornografii do mediów szerszego odbioru – szerzej na ten temat, wciąż jednak postulując rozgraniczanie pojęcia pornografizacji i seksualizacji, zob. M.Tyler, K. Quek, *Conceptualizing Pornographication: A Lack of Clarity and Problems for Feminist Analysis*, Sexualization, „Media, & Society" kwiecień-czerwiec 2016, s. 1-14; także J. Sobczak, *Prostytucja w Internecie*, [w:] *Prostytucja*, red. M. Mozgawa, Warszawa 2014, s. 214.

[21] Pod. J. Levinson, *Erotic Art and Pornographic Pictures*, „Philosophy and Literature Johns Hopkins University Press" 2005, t. 29, nr 1, s. 228-240.

[22] Co pozostaje właściwe dla subiektywno-moralistycznej koncepcji definiowania pornografii – zob. M. Filar, *Przestępstwa seksualne w polskim prawie karnym*, Toruń 1985, s. 113-118.

[23] Por. M. Rodzynkiewicz, *Rozdział XXV Przestępstwa przeciwko wolności seksualnej i obyczajności*, [w:] *Kodeks karny. Część szczególna. Komentarz do art. 117-277 Kodeksu karnego*, red. A. Zoll, Kraków 1999, s. 572.

[24] Tak N. Pope, K. Voges, K.-A. Kuhn, E. Bloxsome, *Pornography and Erotica: Definitions and Prevalence*, International nonprofit and social marketing conference social entrepreneurship, social change and sustainability 2007, s. 168.

[25] W literaturze podnosi się, że wpływ pornografii internetowej należy analizować poprzez mechanizmy psychologiczne, które kierują zachowaniami seksualnymi. Materiały erotyczne działają jako bodźce seksualne, które mogą wywoływać pobudzenie fizjologiczne oraz reakcje emocjonalne i poznawcze. Te z kolei prowadzić do określonych zachowań seksualnych, wzmacniać istniejące preferencje lub wpływać na kształtowanie nowych zainteresowań seksualnych, niejako konsumowanych w drodze obcowania z pornografią. W świetle tego trzeba stwierdzić, że skutki kontaktu z pornografią nie są takie same dla wszystkich. Reakcje użytkowników zależą od wielu czynników. Dla niektórych osób pornografia może pełnić funkcję edukacyjną lub rozrywkową, a u innych wzmacniać poznane wzorce zachowań seksualnych lub orientować postrzeganie relacji seksualnych – szerzej zob. W.A. Fisher,

W analizach przeprowadzonych w glosowanym wyroku zabrakło również rozważań na temat traktowania konsumowania treści w sposób cyfrowy. Wydaje się bowiem, że oglądanie produkcji o tematyce seksualnej, w tym również pornograficznych, przeglądanie zdjęć, korzystanie z kamer erotycznych, *sexting*, erotyczne czaty i *role-play*, subskrypcje platform typu OnlyFans pozostaje jakościowo tożsame. Stąd uprawnione jest ich zaklasyfikowanie jako jeden rodzaj konsumpcji[26]. Takie podejście dostarcza argumentów za traktowaniem usług online wyłącznie jako formy przekazu treści mogących mieć charakter erotyczny lub pornograficzny.

Podsumowanie

Konkludując, w glosowanym orzeczeniu wątpliwości budzi nie tylko samo odwołanie się do bardzo szerokiej i nieostrej znaczeniowo kategorii usług kulturalnych, albowiem nie sposób tego wykluczyć. Podważyć należy zaproponowane kryteria, które wydają się być instrumentami niewydolnymi dla zachowania odpowiedniej precyzji interpretacyjnej. Kryterium naruszenia norm obyczajowych w istocie powiela aksjologiczne uzasadnienie kryminalizacji ustawowo określonych form pornografii. Z tego powodu nie może pełnić z odpowiednią transparentnością funkcji rozgraniczającej tytułowe pojęcia. Trudności pogłębia także brak ogólnego zakazu obcowania z pornografią. To niejako przenosi problem na grunt odróżnienia erotyki od pornografii niekryminalizowanej. Proponowane w literaturze kryterium estetyki nie jest tu jednak przydatne, albowiem nie prowadzi do jednoznacznych rozstrzygnięć. Szczególnie w dobie zmieniających się kanonów estetycznych i przenikania treści seksualnych do kultury masowej. W efekcie możliwe jest identyfikowanie treści, które jednocześnie cechuje artystyczność i wyraźny potencjał pobudzający seksualnie, co utrudnia ich kategoryczną kwalifikację. Kolejne komplikacje pojawiają się w odniesieniu do treści funkcjonujących w przestrzeni cyfrowej, gdzie tradycyjny podział na indywidualną konsumpcję pornografii i wspólne przeżywanie erotyki traci na znaczeniu. W konsekwencji granica między erotyką online a legalną pornografią staje się jeszcze bardziej nieostra. To z kolei nie tylko ogranicza przydatność wskazanych kryteriów jako narzędzia kwalifikacji prawnej, lecz sprawia, że stanowią one instrumenty pogłębiające trudności interpretacyjne.

Szukając bardziej przekonującego rozwiązania, właściwe wydaje się odwrócenie tendencji z glosowanego wyroku. To wymaga przyjęcia, że dla treści określanych

A. Barak, *Internet pornography: A social psychological perspective on Internet sexuality*, „Journal of Sex Research" 2001, nr 38, s. 312-323.

[26] W literaturze wymienione aktywności traktuje się jako zbiór jednorodnych aktywności określany *mianem online sexual activities* – tak A. Wéry, J. Billieux, *Online sexual activities: An exploratory study of problematic and non-problematic usage patterns in a sample of men*, „Computers in Human Behavior" 2016, nr 56, s. 257-266.

jako erotyczne, o ile nie mieszczą się w zakresie penalizowanych form pornografii, należałoby w celach prawnopodatkowych rozważyć co najmniej dwa rozwiązania. Pierwsze wymaga prawnopodatkowego ich kwalifikowania jako pornografii niepenalizowanej. Pozwalałoby uniknąć odwoływania się do nieostrego pojęcia usług kulturalnych oraz zapewnić większą spójność w kwalifikowaniu treści niejednoznacznych. Taki wariant jest jednak trudny do zaakceptowania jako potencjalnie naruszający zasadę *in dubio pro tributario*[27] wywodzoną z zakazu wykładni rozszerzającej w obszarze prawa podatkowego[28]. Postulować należy zatem drugie rozwiązanie *de lege ferenda*, polegające na pilnym rozstrzygnięciu wątpliwości w warstwie tekstowej. Trudno jednak o konkretniejsze propozycje. Przyjęty kierunek wymagałby z pewnością uwzględnienia sygnalizowanych w tekście przeobrażeń w kanonach kultury. Ponadto niezbędna byłaby szersza refleksja, obejmująca również podejście prawodawcy do kwestii podatkowych[29].

Literatura

Bartosiewicz A., *VAT. Komentarz*, Warszawa 2025.
Boul L., Hallam-Jones R., Wylie K.R., *Sexual Pleasure and Motivation*, „Journal of Sex & Marital Therapy" 2008, nr 35.
Filar M., *Przestępstwa seksualne w polskim prawie karnym*, Toruń 1985.
Fisher W.A., Barak A., *Internet pornography: A social psychological perspective on Internet sexuality*, „Journal of Sex Research" 2001, nr 38.
Grudecki M., *Kontratypy pozaustawowe w polskim prawie karnym*, Warszawa 2021.
Gutowski M., Kardas P., *Wykładnia i stosowanie prawa w procesie opartym na Konstytucji*, Warszawa 2017.
Kozłowska P., Mozgawa M., *Prawnokarne aspekty rozpowszechniania pornografii (analiza dogmatyczna i praktyka ścigania)*, „Prokuratura i Prawo" 2002, nr 3.
Levinson J., *Erotic Art and Pornographic Pictures*, „Philosophy and Literature Johns Hopkins University Press" 2005, t. 29, nr 1.
Morawski L., *Zasady wykładni prawa*, Toruń 2006.
Nałęcz B., *Inscenizacja przemocy w pornografii a zakres kryminalizacji art. 202 § 3 k.k.*, „Prawo i Więź" 2025, nr 6.
Pope N., Voges K., Kuhn K.-A., Bloxsome E., *Pornography and Erotica: Definitions and Prevalence*, International nonprofit and social marketing conference social entrepreneurship, social change and sustainability 2007.

[27] Zasadę tę postrzega się jako prawnopodatkowy odpowiednik zasady *in dubio pro reo* funkcjonującej w prawie karnym. Stosownie do niej niedające się usunąć wątpliwości interpretacyjne nie powinny być rozstrzygane na niekorzyść podatnika. W doktrynie bywa ona również ujmowana od strony negatywnej, jako zakaz interpretowania wątpliwości na korzyść fiskusa (*in dubio contra fiscum*) – zob. M. Gutowski, P. Kardas, *Wykładnia i stosowanie prawa w procesie opartym na Konstytucji*, Warszawa 2017, s. 293.

[28] L. Morawski, *Zasady wykładni prawa*, Toruń 2006, s. 178.

[29] To wymaga rozstrzygnięcia podejścia nie tylko do kreowania polityki podatkowej państwa, ale również do koncepcji kształtowania prawa. Istotnie, należałoby przesądzić, czy prawo podatkowe kształtować powinno rzeczywistość społeczną, co łączy się z instrumentalnym użyciem podatków, czy raczej dostosowywać się do istniejących w społeczeństwie zjawisk.

Rodzynkiewicz M., *Rozdział XXV Przestępstwa przeciwko wolności seksualnej i obyczajności*, [w:] *Kodeks karny. Część szczególna. Komentarz do art. 117-277 Kodeksu karnego*, red. A. Zoll, Kraków 1999.
Sobczak J., *Prostytucja w Internecie*, [w:] *Prostytucja*, red. M. Mozgawa, Warszawa 2014.
Ślęzak P., *Zasady rozpowszechniania fotografii i wizerunku*, Warszawa 2025.
Thannhäuser A., Pawlik-Czyniewska J., *Z praktyki powoływania biegłych i opiniowania w sprawach dotyczących czynów zabronionych z art. 202 k.k.*, [w:] *Hominum causa omne ius constitutum sit. Księga jubileuszowa Profesora Piotra Hofmańskiego*, red. P. Czarnecki, S. Głogowska, A. Górski, A. Sakowicz, A. Światłowski, Warszawa 2024.
Tyler M., Quek K., *Conceptualizing Pornographication: A Lack of Clarity and Problems for Feminist Analysis*, „Sexualization, Media, & Society" kwiecień-czerwiec 2016.
Wéry A., Billieux J., *Online sexual activities: An exploratory study of problematic and non-problematic usage patterns in a sample of men*, „Computers in Human Behavior" 2016, nr 56.
Wojciechowski J., *Kodeks karny. Komentarz. Orzecznictwo*, Warszawa 1997.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Agnieszka Narożniak
Uniwersytet WSB Merito w Poznaniu
ORCID: 0000-0002-3091-9849
agnieszka.narozniak@poznan.merito.pl

Justyna Matusiak
Uniwersytet WSB Merito w Poznaniu
ORCID: 0000-0002-2008-3865
justyna.matusiak@poznan.merito.pl

Elektroniczny dokument pobytowy. Charakter prawny, zastosowanie i perspektywy rozwoju

Słowa kluczowe: dokument pobytowy, elektroniczny dokument pobytowy, cudzoziemcy, ochrona czasowa

Streszczenie. Artykuł podejmuje problematykę elektronicznych dokumentów pobytowych dla cudzoziemców w kontekście procesów cyfryzacji administracji publicznej i wyzwań związanych z masowymi migracjami spowodowanymi sytuacjami kryzysami. Celem artykułu jest zatem zarysowanie koncepcji elektronicznego dokumentu pobytowego oraz określenie jego charakteru prawnego i mechanizmu działania na styku dwóch płaszczyzn: praktycznej – związanej z wdrożeniem i funkcjonowaniem nowatorskiego rozwiązania jakim dokument Diia.pl wydawany obywatelom Ukrainy korzystającym z ochrony czasowej w Polsce, oraz teoretycznej – dotyczącej miejsca tego instrumentu w systemie prawa publicznego. W pierwszej części artykułu przedstawione zostało pojęcie i cechy dokumentu pobytowego, z uwzględnieniem przepisów polskiego prawa o cudzoziemcach oraz prawa unijnego. Dalsza części artykułu koncentruje się na polityczno-prawnym kontekście wprowadzenia Diia.pl. Szczególne miejsce w analizie zajmuje charakter prawny elektronicznego dokumentu Diia.pl. Autorki konstatują, że stanowi on dokument pobytowy *sui generis*, gdyż – oprócz swej elektronicznej postaci – różni się także pod innymi względami od tradycyjnych dokumentów pobytowych wydawanych cudzoziemcom. W artykule przedstawione zostają również problemy praktyczne ujawnione w trakcie wdrażania dokumentu. Autorki podkreślają jednak, że mimo tych trudności Diia.pl stanowi rozwiązanie zasługujące na pozytywną ocenę, mogące stanowić wzór dla przyszłych rozwiązań w UE. W końcowej części artykułu autorki formułują wnioski i rekomendacje dotyczące potencjalnego rozwoju instytucji elektronicznego dokumentu pobytowego. Zwracają uwagę, że w świetle obecnych regulacji prawnych instrument ten należy traktować jako rozwiązanie szczególne, stosowane przede wszystkim w sytuacjach kryzysowych. Jednocześnie wskazują na potrzebę stworzenia standardów prawnych i technicznych dla elektronicznych dokumentów pobytowych, które zapewnią bezpieczeństwo i uznawalność takich dokumentów w obrocie prawnym, a równocześnie będą stanowiły gwarancję ochrony praw jednostki.

Electronic Residence Document.
Legal Nature, Application and Development Prospects

Keywords: residence document, electronic residence document, foreigners, temporary protection

Summary. This article addresses the issue of electronic residence documents for foreigners within the context of the digitisation of public administration, as well as the challenges associated with mass migration caused by crisis situations. The article aims to outline the concept of an electronic residence document, defining its legal nature and how it operates at the intersection of two levels: the practical, relating to the implementation and functioning of innovative solutions such as the Diia.pl document issued to Ukrainian citizens benefiting from temporary protection in Poland; and the theoretical, concerning the place of this instrument in the public law system. The first part of the article outlines the concept and characteristics of a residence document, considering the relevant provisions of Polish and EU law relating to foreigners. The subsequent sections focus on the political and legal context surrounding the introduction of Diia.pl. A special place in the analysis is given to the legal nature of the electronic Diia.pl document. The authors conclude that, apart from its electronic form, it differs in other respects from traditional residence documents issued to foreigners, and is therefore a *sui generis* residence document. The article also presents practical problems that were revealed during the implementation of the document. Nevertheless, the authors emphasise that, despite these challenges, Diia.pl is a solution that merits positive evaluation and could serve as a model for future initiatives within the EU. In the final part of the article, the authors draw conclusions and make recommendations regarding the potential development of electronic residence documents. They point out that, given current legal regulations, this instrument should be considered a special solution primarily intended for use in crisis situations. At the same time, they highlight the need to establish legal and technical standards for electronic residence documents to ensure the security and recognition of such documents in legal transactions while guaranteeing the protection of individual rights.

Wprowadzenie

Nieodłączną cechą globalnego społeczeństwa jest rozwój transgranicznych przemieszczeń ludności owocujących długoterminowymi pobytami cudzoziemców w poszczególnych państwach. Jeśli cudzoziemiec uzyskuje prawo pobytu w danym państwie przez dłuższy czas, zwykle skutkuje to wydaniem odpowiedniego dokumentu, który potwierdza uzyskane uprawnienia, a który w związku z tym można określić jako „dokument pobytowy". Zwykle nie wiąże się to ze szczególnymi wyzwaniami organizacyjnymi lub złożonymi problemami prawnymi. Pewne komplikacje mogą jednak się pojawiać w sytuacjach intensywnego napływu dużej liczby osób na obszar danego terytorium – w szczególności osób poszukujących ochrony międzynarodowej. Z jednej strony powstaje wówczas pilna potrzeba wydania odpowiednich dokumentów potwierdzających status danej osoby, z drugiej – administracja mierzy się z ogromnymi nieraz wyzwaniami organizacyjnymi, logistycznymi i prawnymi związanymi z kryzysową sytuacją.

Spektakularnym przykładem takiej sytuacji był tzw. europejski kryzys migracyjny z roku 2015, gdy odnotowano wzmożony napływ migrantów, głównie z obszaru Środkowego Wschodu. Jeszcze większą liczbę osób poszukujących ochrony w pań-

stwach europejskich przyniósł rosyjski atak na Ukrainę w lutym 2022 r. i wojna, która była jego skutkiem. Według stanu na listopad 2023 r. w Unii Europejskiej (UE) z ochrony korzystało ok. 4,2 mln osób z Ukrainy: najwięcej w Niemczech – ponad 1,2 mln, Polsce – ponad 950 tys. i Czechach – blisko 370 tys.[1] Jeśli chodzi o Polskę, to tak masowy napływ osób poszukujących ochrony był absolutnie bezprecedensowy. Kryzys roku 2015 dotknął Polskę bowiem tylko w nieznacznym stopniu (złożono wówczas jedynie nieco ponad 10 tys. wniosków[2]), natomiast w latach 2022-2024 Polska zapewniała ochronę dla około 20-25% ogółu osób z Ukrainy korzystających z ochrony na terytorium Unii Europejskiej. Charakterystyka migracji w obu przypadkach była odmienna również z innych względów, np. w roku 2015 około 70% migrantów stanowili dorośli mężczyźni, podczas gdy w roku 2022 były to głównie kobiety z dziećmi[3]. Jednakże wiele problemów powstających w tych sytuacjach miało podobny charakter. Do uniwersalnych wyzwań stających przed władzami państw przyjmujących było (i jest) ewidencjonowanie ruchów migrantów na terenie UE oraz rejestrowanie pobytu cudzoziemców na wybranym terytorium. Jako jedno z rozwiązań w tym zakresie zaproponowano wprowadzanie elektronicznych dokumentów dla osób poszukujących schronienia. Asumpt do sięgnięcia po nie dał już kryzys z roku 2015, gdy w Niemczech wprowadzono możliwość wydawania migrantom elektronicznych kart zdrowia (*elektronische Gesundheitskarte*)[4]. W Polsce w trend ten wpisały się nowatorskie rozwiązania związane z elektronicznym dokumentem pobytowym Diia.pl dla obywateli Ukrainy korzystających z ochrony czasowej.

Celem niniejszego artykułu jest zaprezentowanie koncepcji elektronicznego dokumentu pobytowego, jego charakteru prawnego i sposobu wdrożenia na przykładzie Diia.pl w Polsce, z uwzględnieniem aspektów praktycznych i teoretycznych omawianego zjawiska. Pierwsza część artykułu jest poświęcona polskiej regulacji prawnej dotyczącej dokumentów pobytowych wydawanych cudzoziemcom. Już na wstępie można zaznaczyć, że nie zawiera ona własnej definicji legalnej dokumentu pobytowego, lecz odsyła w tym względzie do prawa UE. Konsekwentnie, w ko-

---

[1] *Uchodźcy z Ukrainy w UE*, https://www.consilium.europa.eu/pl/infographics/ukraine-refugees-eu/ [dostęp: 17.10.2025].

[2] *Asylum in the EU Member States. Record number of over 1.2 million first time asylum seekers registered in 2015, Syrians, Afghans and Iraqis: top citizenships*, Eurostat news release 44/2016 – 4 March 2016, https://ec.europa.eu/eurostat/documents/2995521/7203832/3-04032016-APEN.pdf/790eba01-381c-4163-bcd2-a54959b99ed6 [dostęp: 17.10.2025].

[3] A. Szachoń-Pszenny, *Szczyt kryzysu migracyjnego w 2015 r. a szczyt kryzysu uchodźczego w 2022 r. – próba analizy porównawczej wpływu na „obszar bez granic" UE*, „Studia Politologiczne" 2023, nr 68, s. 68-69.

[4] A. Koch, N. Biehler, N. Knapp, D. Kipp, *Integrating refugees: Lessons from Germany since 2015-16. Background paper to the World Development Report 2023: Migrants, Refugees and Societies. April 2023*, s. 34, https://www.worldbank.org/en/publication/wdr2023/backgroundpapers [dostęp: 18.10.2025].

lejnej części opracowania zaprezentowana zostanie charakterystyka dokumentu pobytowego z perspektywy prawa UE, a następnie nakreślony polityczny i prawny kontekst wprowadzenia elektronicznego dokumentu pobytowego Diia.pl. Na tym tle omówione zostaną regulacje stanowiące podstawę prawną jego wprowadzenia, mechanizm działania, zasady wydawania oraz jego charakter prawny, jak również wybrane problemy, które zaobserwowano przy jego wdrażaniu. Rozważania zwieńczą wnioski i rekomendacje płynące z przeprowadzonej analizy w perspektywie potencjalnego rozwoju wykorzystania elektronicznych dokumentów pobytowych w przyszłości. Jak bowiem wskazuje się w literaturze, doświadczenia niemieckie z lat 2015-2016 pozwoliły w dużo lepszym stopniu przygotować się na migracje z roku 2022 związane z wojną w Ukrainie[5]. Analogicznie, biorąc pod uwagę – z jednej strony – wyzwania związane z powszechną cyfryzacją w obszarze administrowania państwem, z drugiej – prognozy dotyczące wzrostu przymusowych migracji w przyszłości, analiza doświadczeń z wykorzystaniem Diia.pl i wysnucie właściwych wniosków może przyczynić się w przyszłości do udoskonalenia procedur, w tym do sprawnego reagowania na nieuniknione kryzysy.

1. Rodzaje dokumentów wydawanych cudzoziemcom

Rozważania o dokumencie pobytowym dla porządku wypada rozpocząć od stwierdzenia, że organy administracji publicznej wydają różnego rodzaju dokumenty zarówno swym obywatelom, jak i cudzoziemcom. Z perspektywy dalszych rozważań istotne jest zwrócenie uwagi, że wśród dokumentów wydawanych cudzoziemcom przez organy państw pobytu można wskazać różne ich typy.

Pierwszy rodzaj to dokumenty, których funkcją jest wyłącznie potwierdzenie tożsamości danej osoby. W Polsce do tej kategorii należy polski dokument tożsamości cudzoziemca. Dokument ten wydaje się w szczególnych przypadkach wskazanych w przepisach ustawy z dnia 13 grudnia 2013 r. o cudzoziemcach[6]. W okresie swojej ważności potwierdza on tożsamość cudzoziemca podczas jego pobytu na terytorium Rzeczypospolitej Polskiej, lecz nie potwierdza jego obywatelstwa, nie uprawnia do przekroczenia granicy ani nie zwalnia z obowiązku uzyskania tytułu pobytowego.

Drugi typ to dokumenty podróży, a więc – najogólniej mówiąc – dokumenty, których podstawową funkcją jest potwierdzenie uprawnienia do przekraczania granicy[7]. Należy tu wymienić tzw. Genewski Dokument Podróży wydawany oso-

[5] *Ibidem*, s. 36.

[6] Art. 260 Ustawy z dnia 13 grudnia 2013 r. o cudzoziemcach; dalej: „uoc", „ustawa o cudzoziemacach".

[7] Zob. art. 3 pkt 3 uoc.

bom posiadającym status uchodźcy[8], jak i różnego rodzaju dokumenty wydawane na podstawie wewnętrznych uregulowań (w Polsce są to polski dokument podróży dla cudzoziemca[9] oraz tymczasowy polski dokument podróży dla cudzoziemca[10]). Polski dokument podróży dla cudzoziemca w okresie swojej ważności (tj. jednego roku) uprawnia cudzoziemca do wielokrotnego przekraczania granicy, natomiast tymczasowy polski dokument podróży dla cudzoziemca uprawnia do (odpowiednio) jednego wjazdu lub wyjazdu w okresie ważności wynoszącym maksymalnie 7 dni.

Kolejnym rodzajem dokumentów wydawanych cudzoziemcom są dokumenty będące przedmiotem niniejszego opracowania, a mianowicie te, których zasadnicza – choć z reguły nie jedyna – funkcja polega na potwierdzeniu przyznanego w danym państwie prawa pobytu (dokumenty pobytowe).

W polskim systemie prawnym do kategorii dokumentów pobytowych zalicza się przede wszystkim kartę pobytu wydawaną cudzoziemcom, którym udzielono zezwolenia na pobyt czasowy, zezwolenia na pobyt stały lub zezwolenia na pobyt rezydenta długoterminowego Unii Europejskiej albo zgody na pobyt ze względów humanitarnych[11]. Oprócz funkcji potwierdzenia prawa pobytu wynikającego z wydanego zezwolenia lub zgody karta pobytu w myśl przepisów ustawy potwierdza tożsamość oraz uprawnia cudzoziemca, wraz z dokumentem podróży, do wielokrotnego przekraczania granicy bez konieczności uzyskania wizy[12]. Kartę pobytu wydaje się także na innych podstawach prawych cudzoziemcom, którym nadano status uchodźcy lub objęto ich ochroną uzupełniającą[13]. Z kolei cudzoziemcy będący obywatelami Unii Europejskiej otrzymują dokument nazwany „zaświadczeniem o zarejestrowaniu pobytu obywatela UE", a następnie, po uzyskaniu prawa pobytu stałego – „dokument potwierdzający prawo stałego pobytu", natomiast członkowie ich rodzin – odpowiednio – kartę pobytową/kartę stałego pobytu[14]. Karta pobytowa i karta stałego pobytu potwierdzają – odpowiednio – prawo pobytu/prawo stałego pobytu i uprawniają wraz z ważnym dokumentem podróży do wielokrotnego przekraczania granicy bez wizy[15].

---

[8] Dokument przewidziany w Konwencji dotyczącej statusu uchodźców, sporządzonej w Genewie dnia 28 lipca 1951 r., Dz.U. z 1991 r. poz. 515 i 516; zob. też art. 89i i nast. Ustawy z dnia 13 czerwca 2003 r. o udzielaniu cudzoziemcom ochrony na terytorium Rzeczypospolitej Polskiej, dalej: „uouco", „ustawa o udzielaniu cudzoziemcom ochrony".

[9] Art. 252 i nast. uoc.

[10] Art. 267 i nast. uoc.

[11] Art. 240 uoc.

[12] Art. 242 uoc.

[13] Art. 89i ust. 1 i 2 uouco.

[14] Art. 20, 27, 48 Ustawy z dnia 14 lipca 2006 r. o wjeździe na terytorium Rzeczypospolitej Polskiej, pobycie oraz wyjeździe z tego terytorium obywateli państw członkowskich Unii Europejskiej i członków ich rodzin, t.j. Dz.U. z 2025 r. poz. 224; dalej: „uoobUE".

[15] Art. 30 ust. 1, art. 54 uoobUE.

Oprócz wyżej wymienionych istnieje także grupa dokumentów o szczególnym charakterze, niepoddających się prostej kategoryzacji. Wśród dokumentów wydawanych w Polsce wymienić tu trzeba dokument „zgoda na pobyt tolerowany", który potwierdza podstawę prawną pobytu cudzoziemca, a także jego tożsamość, jednak nie potwierdza obywatelstwa ani nie uprawnia do przekraczania granicy[16]. Innym dokumentem o szczególnym charakterze jest tymczasowe zaświadczenie tożsamości cudzoziemca wydawane cudzoziemcom ubiegającym się o objęcie ochroną międzynarodową, które w okresie swojej ważności, wbrew sugestii płynącej z nazwy, nie tylko potwierdza tożsamość cudzoziemca, lecz także uprawnia do pobytu na terytorium Rzeczypospolitej Polskiej do zakończenia postępowania w sprawie udzielenia ochrony międzynarodowej decyzją ostateczną[17]. Specyficzny dokument, o którym jeszcze będzie mowa dalej, stanowi zaświadczenie potwierdzające korzystanie z ochrony czasowej, stanowiące wyłączny dowód korzystania z ochrony czasowej w Rzeczypospolitej Polskiej, które w okresie swej ważności poświadcza prawo jego posiadacza do pobytu w Polsce oraz uprawnia go, wraz z dokumentem podróży, do wielokrotnego przekraczania granicy bez konieczności uzyskania wizy[18].

Wreszcie, wśród omawianych dokumentów należy wymienić przywołany w tytule niniejszego opracowania i we wstępie elektroniczny dokument pobytowy – Diia.pl. Zostaną mu poświęcone szersze rozważania obejmujące zarówno genezę, jak i prawne ramy oraz praktyczne problemy dotyczące jego wprowadzenia. Przedtem jednak zasadne wydaje się przybliżenie regulacji dotyczących dokumentu pobytowego jako wyodrębnionej instytucji prawnej.

2. Pojęcie i cechy dokumentu pobytowego

Dla zaprezentowania charakterystyki dokumentu pobytowego jako instytucji prawnej w kontekście polskim i europejskim miarodajne są w szczególności akty normatywne należące do systemu prawa Unii Europejskiej.

W pierwszej kolejności, jako istotną z omawianej perspektywy, należy wskazać definicję tego pojęcia zawartą w Rozporządzeniu Rady (WE) nr 1030/2002 z dnia 13 czerwca 2002 r. ustanawiającym jednolity wzór dokumentów pobytowych dla obywateli państw trzecich[19]. W myśl art. 1 ust. 2 lit. a tego aktu, „dokument pobytowy" w tym kontekście oznacza upoważnienie wydane przez władze danego

[16] Art. 273-274 uoc.
[17] Art. 55 i nast. uouco.
[18] Art. 110 uouco.
[19] Dz.U. UE L 157 z 15.6.2002, p. 1-7 ze zm.(wersja skonsolidowana aktualna na dzień 21.10.2025: https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:02002R1030-20171121); dalej: „rozporządzenie 1030/2002".

państwa członkowskiego zezwalające obywatelowi państwa trzeciego na legalny pobyt na danym terytorium, przy czym terminem tym nie obejmuje się niektórych krótkoterminowych tytułów pobytowych, m.in. wiz[20]. Przez „obywatela państwa trzeciego" rozumie się, zgodnie z ugruntowaną terminologią prawa UE, osobę, która nie jest obywatelem Unii. Co istotne, zgodnie z postanowieniami rozporządzenia 1030/2002 dokumenty pobytowe wydawane przez państwa członkowskie obywatelom państw trzecich sporządza się według jednolitego wzoru określonego w załączniku do tego aktu. W polskim prawie o cudzoziemcach nie ma odrębnej definicji legalnej dokumentu pobytowego, natomiast można znaleźć bardzo liczne odesłania do „dokumentu pobytowego w rozumieniu art. 1 ust. 2 lit a" tego rozporządzenia[21].

Jeśli chodzi o dokumenty pobytowe wydawane obywatelom Unii i członkom ich rodzin, to stosuje się do nich postanowienia Dyrektywy 2004/38/WE z dnia 29 kwietnia 2004 r. w sprawie prawa obywateli Unii i członków ich rodzin do swobodnego przemieszczania się i pobytu na terytorium Państw Członkowskich[22] oraz Rozporządzenie Rady (UE) 2025/1208 z dnia 12 czerwca 2025 r. w sprawie poprawy zabezpieczeń dowodów osobistych obywateli Unii i dokumentów pobytowych wydawanych obywatelom Unii i członkom ich rodzin korzystającym z prawa do swobodnego przemieszczania się[23]. W przypadku dokumentów pobytowych wydawanych obywatelom Unii rozporządzenie 2025/1208 wskazuje minimalny zakres informacji, które muszą być w nim zawarte[24]. Natomiast w przypadku dokumentów pobytowych wydawanych członkom rodziny obywateli Unii, którzy sami są obywatelami państw trzecich, rozporządzenie 2025/1208 odsyła do wspomnianego wyżej Rozporządzenia 1030/2002, nakazując stosować ten sam jednolity format (z tym zastrzeżeniem, że na dokumencie umieszcza się informację o statusie członka rodziny obywatela UE)[25].

---

[20] Rozporządzenie stanowi, że oprócz wiz dokumentami pobytowymi w jego rozumieniu nie są także: zezwolenia wydawane do czasu rozpatrzenia wniosku o azyl, wniosku o wydanie dokumentu pobytowego lub wniosku o przedłużenie ważności dokumentu pobytowego; zezwolenia wydawane w wyjątkowych przypadkach w celu przedłużenia dozwolonego pobytu o okres nieprzekraczający jednego miesiąca; upoważnienia na okres pobytu nieprzekraczający sześciu miesięcy, wydawane przez Państwa Członkowskie niestosujące przepisów art. 21 Konwencji wykonawczej do Układu z Schengen z 14 czerwca 1985 r. między Rządami Państw Unii Gospodarczej Beneluksu, Republiką Federalną Niemiec i Republiką Francuską o stopniowym znoszeniu kontroli na wspólnych granicach, Dz.U. UE L 239 z 22.9.2000, s. 19.

[21] W uoc tych odesłań jest 82.

[22] Dz.U. UE L 158 z 30.4.2004, p. 77-123 ze zm. (wersja skonsolidowana aktualna na dzień 21.10.2025); dalej „dyrektywa 2004/38/WE".

[23] Dz.U. UE L, 2025/1208; dalej: „rozporządzenie 2025/1208".

[24] Art. 6 rozporządzenia 2025/1208.

[25] Art. 7 rozporządzenia 2025/1208.

Kolejnym aktem ważnym dla wyjaśnienia pojęcia dokumentu pobytowego jest Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/399 z dnia 9 marca 2016 r. w sprawie unijnego kodeksu zasad regulujących przepływ osób przez granice (kodeks graniczny Schengen)[26]. Zgodnie z art. 2 pkt 16 dokumenty pobytowe w rozumieniu kgS to nie tylko dokumenty pobytowe wydane przez państwa członkowskie zgodnie z jednolitym wzorem określonym w rozporządzeniu 1030/2002 oraz karty pobytowe wydane zgodnie z dyrektywą 2004/38/WE, lecz także inne dokumenty wydane przez państwo członkowskie obywatelom państw trzecich, uprawniające do pobytu na jego terytorium, które zostały odpowiednio zgłoszone, a następnie opublikowane zgodnie z procedurą przewidzianą w art. 39 kgS. Przepis ten wymaga, aby państwa członkowskie zgłaszały Komisji Europejskiej wykaz dokumentów pobytowych wraz z wzorami dokumentów nieodpowiadających jednolitemu formatowi, Komisja zaś ma obowiązek opublikować i rozpowszechnić te informacje. Pojęcie dokumentów pobytowych w rozumieniu kgS nie obejmuje natomiast dokumentów tymczasowych ważnych do momentu rozpatrzenia pierwszego wniosku o wydanie właściwego dokumentu pobytowego lub wniosku o nadanie statusu uchodźcy oraz wiz wydanych przez państwo członkowskie zgodnie z jednolitym formularzem określonym w odpowiednim akcie prawa UE.

Z zarysowanych wyżej uregulowań wynika, że – co do zasady – immanentną cechą dokumentu pobytowego w prawie UE jest posiadanie określonej przepisami, ujednoliconej w skali UE formy. Zasadniczą formą określoną w rozporządzeniu 1030/2002 jest karta wykonana w całości z poliwęglanu lub równorzędnego polimeru syntetycznego (o trwałości wynoszącej co najmniej dziesięć lat). Karta wykonana jest zgodnie ze specyfikacją techniczną określoną w załączniku do rozporządzenia, w kolorach przewidzianych w tym akcie i zawiera wymagane przezeń dane oraz zabezpieczenia. Od tej zasady można jednak odstąpić, pod warunkiem dopełnienia wymogu formalnego powiadomienia i upublicznienia tej informacji. Omawiane przepisy nie przewidują w żadnym miejscu *expressis verbis* opcji wprowadzenia dokumentu pobytowego w formie wyłącznie elektronicznej, choć – nie ulega wątpliwości – również nie zabraniają takiego rozwiązania przy spełnieniu wskazanego warunku.

W praktyce, w roku 2024 Polska, Litwa i Włochy były jedynymi państwami Europejskiej Sieci Migracyjnej (EMN – European Migration Network), które wydawały w określonym zakresie obywatelom państw trzecich zezwolenia na pobyt i/lub dokumenty mające wyłącznie postać elektroniczną[27]. Nie zgłaszano planów

[26] Dz.U. UE L 77 z 23.3.2016, s. 1-52 ze zm. (wersja skonsolidowana aktualna na dzień 21.10.2025); dalej: „kgS".

[27] *European Migration Network Inform, Digitalisation of identity documents and residence permits issued to third-country nationals, March 2024*, s. 4-6, https://home-affairs.ec.europa.eu/document/

wprowadzania tego typu rozwiązań na szerszą skalę[28]. Powody takiego stanu rzeczy są różne. Przykładowo Bułgaria nie jest w stanie zapewnić prawidłowej i sprawnej weryfikacji elektronicznych dokumentów na granicy ze względu na braki techniczne i niemożność wprowadzania danych z tego formatu do krajowego systemu informatycznego. Niektóre państwa UE uważają, że wprowadzenie jedynie wersji elektronicznej dokumentów wymagałoby zmian legislacyjnych, aby zapewnić, że nie zostaną naruszone prawa podstawowe (takie stanowisko prezentują np. Estonia i Niemcy)[29]. Doświadczenia związane z wykorzystaniem na szeroką skalę elektronicznego dokumentu pobytowego przez Polskę doskonale wpisują się w tę dyskusję i pozwalają zweryfikować niektóre twierdzenia i obawy.

3. Polityczny i prawny kontekst wprowadzenia elektronicznego dokumentu pobytowego w Polsce

Elektroniczny dokument pobytowy – Diia.pl – został wprowadzony w Polsce w 2022 r. w związku z bezprecedensową liczbą osób poszukujących schronienia na jej terytorium po rosyjskim ataku na Ukrainę. Jego nazwa nawiązuje do ukraińskiego portalu Дія (https://diia.gov.ua) i aplikacji mobilnej, za pośrednictwem których realizowane są usługi elektroniczne świadczone przez państwo[30].

W dniu wybuchu wojny, jak i w chwili redagowania niniejszego tekstu, kluczowym aktem normatywnym regulującym kwestie udzielania ochrony w Polsce była (i jest) przywoływana już ustawa z dnia 13 czerwca 2003 r. o udzielaniu cudzoziemcom ochrony na terytorium Rzeczypospolitej Polskiej. Ustawa ta przewiduje cztery formy ochrony: status uchodźcy, ochronę uzupełniającą, azyl (polityczny) i ochronę czasową. Z perspektywy genezy wykorzystania w Polsce elektronicznego dokumentu pobytowego, istotne znaczenie ma ta ostatnia instytucja.

Przepisy dotyczące ochrony czasowej stanowią formę wdrożenia do polskiego porządku prawnego postanowień Dyrektywy Rady 2001/55/WE z dnia 20 lipca 2001 r. w sprawie minimalnych standardów przyznawania tymczasowej ochrony na wypadek masowego napływu wysiedleńców oraz środków wspierających równowagę wysiłków między Państwami Członkowskimi, związanych z przyjęciem

download/bf63b9cd-1c72-4303-97f5-36a4ebd52a4c_en?filename=2024_EMN_inform_document_digitalisation.pdf [dostęp: 21.10.2025].

[28] *Ibidem*, s. 7.

[29] *Ibidem*, s. 8-9.

[30] Jak wskazuje T. Gajowniczek, to ukraińskie słowo oznacza „akt", „czyn", „działanie", nazwa portalu zaś stanowi akronim od Держава і я - „państwo i ja"; *E-gov w Ukrainie – zarys problemu*, „Środkowoeuropejskie Studia Polityczne" 2022, nr 4, s. 186; zob. też I. Pankevytch, I. Bulat, *Prawne aspekty e-administracji w Ukrainie: stan i perspektywy rozwoju*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3, s. 200.

takich osób wraz z jego następstwami[31]. Tymczasowa ochrona stanowi w jej świetle szczególnego rodzaju procedurę stosowaną w przypadku „zagrożenia wystąpienia bądź rzeczywistego wystąpienia masowego napływu wysiedleńców z państw trzecich, którzy nie mogą powrócić do państwa pochodzenia", a przesłanką jej uruchomienia jest „uzasadnione przypuszczenie, że system udzielania azylu nie będzie w stanie udźwignąć ciężaru takiego napływu bez negatywnych skutków dla działania tego systemu, ze szkodą dla zainteresowanych i innych osób poszukujących ochrony"[32]. Stwierdzenie, że przesłanka ta w danym przypadku zaistniała, następuje w drodze decyzji Rady Unii Europejskiej. Dyrektywa określa ponadto obowiązki państw członkowskich względem osób objętych ochroną, w tym w zakresie dokumentowania ich prawa pobytu. W ślad za postanowieniami dyrektywy 2001/55/WE polska ustawa o udzielaniu cudzoziemcom ochrony przewiduje, że ochrony czasowej można udzielić „cudzoziemcom masowo przybywającym do Rzeczypospolitej Polskiej, którzy opuścili swój kraj pochodzenia lub określony obszar geograficzny, z powodu obcej inwazji, wojny, wojny domowej, konfliktów etnicznych lub rażących naruszeń praw człowieka […] bez względu na to, czy ich przybycie miało charakter spontaniczny, czy też było wynikiem pomocy udzielonej im przez Rzeczpospolitą Polską lub społeczność międzynarodową"[33]. Udziela się jej się na podstawie i w granicach określonych w decyzji Rady Unii Europejskiej, przez okres określony każdorazowo w tej decyzji[34].

Do czasu wybuchu wojny w Ukrainie w 2022 r. tak skonstruowany mechanizm ochrony czasowej nigdy nie był uruchamiany w państwach członkowskich UE. Rosyjska inwazja na Ukrainę spowodowała ucieczkę milionów jej mieszkańców na terytorium UE, a w rezultacie Rada uznała przesłankę wprowadzenia ochrony tymczasowej za spełnioną. Wydana 4 marca 2022 r. Decyzja wykonawcza Rady (UE) 2022/382 stwierdza istnienie masowego napływu wysiedleńców z Ukrainy w rozumieniu art. 5 dyrektywy 2001/55/WE i skutkuje wprowadzeniem tymczasowej ochrony[35]. W sytuacji, w jakiej znalazła się Polska w związku z gwałtownie rosnącą liczbą osób uciekających z Ukrainy w poszukiwaniu schronienia przed skutkami inwazji, rozwiązania przewidziane w dyrektywie 2001/55/WE i w ustawie o udzielaniu cudzoziemcom ochrony na terytorium Rzeczypospolitej Polskiej okazały się jednak niewystarczające[36]. Aktem, który miał zapewnić odpowiednie,

[31] Dz.U. UE L 212 z 7.8.2001, p. 12-23, dalej „dyrektywa 2001/55/WE".
[32] Art. 2 lit. a) dyrektywy 2001/55/WE.
[33] Art. 106 ust. 1 uouco.
[34] Art. 107 ust.1 uouco.
[35] Dz.U. UE L 71 z 4.3.2022, p. 1-6.
[36] Zob. Uzasadnienie do Rządowego projektu ustawy o pomocy obywatelom Ukrainy w związku z konfliktem zbrojnym na terytorium tego państwa, Druk sejmowy nr 2069, Sejm IX kadencji, https://www.sejm.gov.pl/sejm9.nsf/druk.xsp?nr=2069 [dostęp: 21.10.2025].

szczegółowe ramy prawne, stała się uchwalona 12 marca 2022 r. specjalna ustawa o pomocy obywatelom Ukrainy w związku z konfliktem zbrojnym na terytorium tego państwa[37]. Stała się ona podstawą do wyodrębnienia specyficznego statusu prawnego cudzoziemców, do których znajduje zastosowanie. Status ten jest różny od statusu osoby korzystającej z ochrony tymczasowej w kształcie przewidzianym w ustawie o udzielaniu cudzoziemcom ochrony na terytorium Rzeczypospolitej Polskiej (do której zresztą dodano przepis, który wyraźnie stanowi, że przepisy rozdziału o ochronie czasowej mają zastosowanie w zakresie nieuregulowanym odrębnymi przepisami[38]). Takie rozwiązanie jest dopuszczalne w prawie unijnym, gdyż standard przewidziany w dyrektywie 2001/55/WE stanowi standard minimalny, co oznacza, że poszczególne państwa członkowskie mogą ustanowić dalej idące uprawnienia osób objętych ochroną tymczasową[39].

W rezultacie na terytorium Polski powstały dwa reżimy ochrony czasowej: jeden w kształcie nadanym ustawą o udzielaniu cudzoziemcom ochrony na terytorium Rzeczypospolitej Polskiej, drugi – na podstawie specustawy z 2022 r.[40] Reżim specustawy objął przede wszystkim obywateli Ukrainy, którzy przybyli do Polski z terytorium Ukrainy w związku z działaniami wojennymi prowadzonymi na terytorium tego państwa oraz obywateli Ukrainy posiadających Kartę Polaka[41], którzy wraz z najbliższą rodziną przybyli do Polski z powodu tych działań wojennych (choć niekoniecznie z terytorium Ukrainy). Ponadto ustawa znalazła zastosowanie także do nieposiadających obywatelstwa ukraińskiego członków rodziny obywatela Ukrainy. Do tego grona zaliczono małżonka obywatela Ukrainy, małoletnie dziecko obywatela Ukrainy oraz małoletnie dziecko małżonka obywatela Ukrainy – o ile przybyli oni na terytorium Rzeczypospolitej Polskiej z terytorium Ukrainy w związku z działaniami wojennymi prowadzonymi na terytorium tego państwa

---

[37] T.j. Dz.U. z 2025 r. poz. 337 z późn. zm.; dalej: „specustawa z 2022 r.".

[38] Art. 106 ust. 4 uouco dodany na mocy art. 80 pkt 2 specustawy z 2022 r. Zgodnie z art. 2 ust. 6 specustawy z 2022 r. obywatela Ukrainy, o którym mowa w ust. 1, uznaje się za osobę korzystającą w Rzeczypospolitej Polskiej z ochrony czasowej w rozumieniu art. 106 ust. 1 uouco. Jednak, w myśl regulacji zawartej w ust. 7, obywatelowi temu, jako osobie korzystającej w Rzeczypospolitej Polskiej z ochrony czasowej, przysługują uprawnienia określone w specustawie. Do ochrony czasowej, z której korzysta obywatel Ukrainy, o którym mowa w ust. 1 specustawy, nie stosuje się natomiast przepisów rozdziału 3 działu III uouco (art. 2 ust. 8 specustawy).

[39] Zob. art. 3 ust. 5 dyrektywy 2001/55/WE.

[40] Zob. *Ustawa o pomocy obywatelom Ukrainy w związku z konfliktem zbrojnym na terytorium tego państwa. Komentarz*, red. K. Szmid, P. Sawicki, Warszawa 2022, Legalis, Komentarz do art. 1; o relacjach tych regulacji, w szczególności w zakresie wykonywania pracy, zob. M. Górski, *Wpływ specustawy ukraińskiej na sytuację prawną migrantów przymusowych z Ukrainy na polskim rynku pracy / The impact of the Ukrainian special law on the legal situation of forced migrants from Ukraine on the Polish labour market*, Centre of Migration Research, CMR Working Papers 135/193, September 2023, p. 32-37, https://www.migracje.uw.edu.pl/publikacje/wplyw-specustawy-ukrainskiej-na-sytuacje-prawna-migrantow-przymusowych-z-ukrainy-na-polskim-rynku-pracy/ [dostęp: 21.10.2025].

[41] Zob. ustawa z dnia 7 września 2007 r. o Karcie Polaka, t.j. Dz.U. z 2023 r. poz. 192 z późn. zm.

i nie są obywatelami polskimi ani obywatelami innych niż Rzeczpospolita Polska państw członkowskich Unii Europejskiej[42]. Wypada zwrócić uwagę, że przepisy specustawy nie stały się podstawą prawną pobytu dla obywateli Ukrainy i członków ich rodzin już posiadających zezwolenie na pobyt w Polsce albo ubiegających się lub objętych inną formą ochrony (międzynarodowej lub krajowej) w Polsce, bądź też korzystających z ochrony czasowej na terenie innego państwa członkowskiego UE przyznanej z powodu działań wojennych prowadzonych na terytorium Ukrainy[43]. Innymi słowy, specustawa nie uchyliła istniejących podstaw prawnych pobytu, a jedynie uzupełniła w istotny sposób ich katalog.

Z perspektywy rozważań zawartych w niniejszym artykule podkreślenia wymaga także różnica pomiędzy beneficjentami ochrony czasowej uregulowanej w ustawie o udzielaniu cudzoziemcom ochrony a beneficjentami ochrony przewidzianej w specustawie w zakresie dokumentowania pobytu. W tym pierwszym przypadku przewidziano wydawanie przywołanych już wyżej zaświadczeń potwierdzających korzystanie z ochrony czasowej, a w drugim – elektroniczny dokument pobytowy Diia.pl. W praktyce to właśnie osoby przebywające w Polsce na podstawie statusu nadanego specustawą stanowiły od początku jej obowiązywania przemożną większość korzystających z ochrony. Nie uległo to zmianie w kolejnych latach; przykładowo w samym roku 2025 od stycznia do 1 października wydano 22 zaświadczenia, a równocześnie, w tym samym okresie, na ochronę czasową zarejestrowano w Rejestrze PESEL (na podstawie przepisów specustawy z 2022 r.) 113 900 osób[44]. Dla porównania, w roku 2024 Szef Urzędu do Spraw Cudzoziemców wydał 49 zaświadczeń o korzystaniu z ochrony czasowej, a w Rejestrze PESEL na ochronę czasową zarejestrowano 165 724 osoby[45].

4. Regulacja prawna elektronicznego dokumentu pobytowego w Polsce

Jak wspomniano wcześniej, na podstawie art. 39 kgS państwa członkowskie zgłaszają Komisji Europejskiej wykaz dokumentów pobytowych wraz z wzorami dokumentów nieodpowiadających jednolitemu formatowi, Komisja zaś ma obowiązek opublikować i rozpowszechnić te informacje. W związku z tym 12 lipca 2022 r. Polska notyfikowała Komisji Europejskiej nowy dokument pobytowy wydawany w związku ze szczególną sytuacją na Ukrainie, czyli dokument elektroniczny Diia.

[42] Art. 1 specustawy z 2022 r.

[43] Art. 2 ust. 1-3 specustawy z 2022 r.

[44] Urząd do Spraw Cudzoziemców, Raport na temat obywateli Ukrainy (stan na 1 października 2025 r.), https://www.gov.pl/web/udsc/sytuacja-dotyczaca-ukrainy [dostęp: 21.10.2025].

[45] Urząd do Spraw Cudzoziemców, Raport na temat obywateli Ukrainy (stan na 31 grudnia 2024 r.), https://www.gov.pl/attachment/63a33038-f3c7-4430-8776-159d3fbfb637 [dostęp: 21.10.2025].

pl w aplikacji mobilnej mObywatel, potwierdzający m.in. status beneficjenta art. 2 ust. 1 specustawy z 2022 r.[46]

Podstawą prawną wprowadzenia elektronicznego dokumentu pobytowego Diia.pl w Polsce stał się art. 10 specustawy z 2022 r. Istotą przyjętego rozwiązania było umożliwienie ministrowi właściwemu do spraw informatyzacji wykorzystania istniejącej aplikacji mObywatel do uruchomienia usługi elektronicznej dla obywateli Ukrainy. Nie tworzono zatem od podstaw nowej struktury ani w ujęciu normatywnym, ani technicznym, choć oczywiście konieczne było odpowiednie dostosowanie przepisów prawa krajowego oraz spełnienie wymogów prawa UE, zagwarantowanie zdolności systemu do potwierdzania autentyczności dokumentu elektronicznego oraz wprowadzenie silnych mechanizmów szyfrowania[47]. Dla porządku wypada wyjaśnić, że aplikacja mObywatel to oprogramowanie przeznaczone dla urządzeń mobilnych, w którym są udostępniane usługi świadczone przez podmioty publiczne oraz niepubliczne[48]. Dzięki aplikacji mObywatel obywatel Ukrainy, którego pobyt na terytorium RP uznaje się za legalny na podstawie przepisów specustawy z 2022 r., po uwierzytelnieniu może pobrać, przechowywać i prezentować dokument elektroniczny zawierający dane pobrane z rejestru obywateli Ukrainy, którym nadano numer PESEL oraz weryfikować integralność i pochodzenie dokumentu elektronicznego. Założeniem tego rozwiązania była tymczasowość związana z ograniczonym z definicji okresem obejmowania cudzoziemców ochroną czasową. W związku z przedłużającą się wojną w Ukrainie ochrona czasowa jest systematycznie przedłużana; w chwili redagowania niniejszego tekstu – po ponad trzech i pół roku od uchwalenia specustawy – omawiane rozwiązania nadal funkcjonują[49].

Jak wskazano wyżej, użytkownikami Diia.pl mogą być wyłącznie osoby, dla których podstawą prawną pobytu na terytorium Polski są przepisy specustawy z 2022 r. Do uzyskania dokumentu konieczne jest, by danej osobie został nadany numer PESEL. Numer PESEL nadaje się obywatelom Ukrainy przebywającym w Polsce na podstawie specustawy z 2022 r. na ich wniosek złożony w wybranym urzędzie gminy, przy czym status obywatela Ukrainy korzystającego z tej formy

---

[46] Aktualizacja wykazu dokumentów pobytowych, o których mowa w art. 2 pkt 16 rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/399 w sprawie unijnego kodeksu zasad regulujących przepływ osób przez granice (kodeks graniczny Schengen), 2022/C 304/05.

[47] Zob. Serwis Rzeczypospolitej Polskiej, Diia.pl, https://www.gov.pl/web/cyfryzacja-badania-i-projektowanie-mobywatel20/diiapl [dostęp: 21.10.2025].

[48] Art. 2 pkt 1 Ustawy z dnia 26 maja 2023 r. o aplikacji mObywatel, t.j. Dz.U. z 2024 r. poz. 1275 z późn. zm., dalej „ustawa o aplikacji mObywatel".

[49] Ochrona tymczasowa w UE wg stanu na dzień 31.10.2025 r. obowiązuje do dnia 4 marca 2027 r., jednak UE podejmuje działania mające na celu przygotowanie jej stopniowego wycofywania (zob. Zalecenie Rady z dnia 16 września 2025 r. w sprawie skoordynowanego podejścia do stopniowego wycofywania tymczasowej ochrony dla wysiedleńców z Ukrainy, Dz.U. UE C, C/2025/5129, 23.9.2025).

ochrony oznacza się w rejestrze PESEL symbolem UKR[50]. Jeśli obywatel Ukrainy posiadał już wcześniej numer PESEL (w związku ze swoim wcześniejszym pobytem w Polsce), a przebywa aktualnie na podstawie specustawy z 2022 r., składa wniosek o nadanie statusu UKR. Odpowiedni wniosek powinien zostać złożony niezwłocznie po przybyciu do Polski[51]. Ustalenie, że dany obywatel Ukrainy przybył na terytorium Rzeczypospolitej Polskiej z terytorium Ukrainy w związku z działaniami wojennymi prowadzonymi na terytorium tego państwa opiera się zatem na stwierdzeniu, że obywatelowi temu nadano numer PESEL z adnotacją (statusem) UKR[52]. Warunkiem uzyskania dokumentu Diia.pl jest ponadto posiadanie profilu zaufanego. Profil zaufany to środek identyfikacji elektronicznej zawierający zestaw danych identyfikujących i opisujących osobę fizyczną, która posiada pełną lub ograniczoną zdolność do czynności prawnych, wydany wnioskodawcy po potwierdzeniu danych z wniosku ze stanem faktycznym[53]. Zakładanie profilu zaufanego zostało do pewnego stopnia sprzężone z procedurą uzyskiwania numeru PESEL, poprzez regulację przewidującą możliwość podania we wniosku o nadanie numeru PESEL danych poczty elektronicznej oraz numer telefonu komórkowego i wyrażenia zgody na potwierdzenie profilu zaufanego[54]. Minister właściwy do spraw informatyzacji automatycznie potwierdza profil zaufany dla wnioskodawcy. Natomiast w sytuacji, w której obywatel Ukrainy we wniosku o nadanie numeru PESEL nie wyraził zgody na wprowadzenie danych do rejestru danych kontaktowych lub na potwierdzenie profilu zaufanego albo nie było to możliwe z przyczyn technicznych, może złożyć w dowolnym organie wykonawczym gminy odrębny wniosek obejmujący taką zgodę, który zawiera żądanie o wprowadzenie numeru PESEL, imienia i nazwiska, adresu poczty elektronicznej oraz numeru telefonu komórkowego do rejestru danych kontaktowych oraz potwierdzenie profilu zaufanego[55]. Dalsze kroki do uzyskania dokumentu to zainstalowanie aplikacji mobilnej mObywatel na telefonie komórkowym spełniającym określone wymogi techniczne i dodanie dokumentu Diia.pl.

Jeśli chodzi o funkcje dokumentu Diia.pl, to w pierwszym rzędzie należy zwrócić uwagę na cechę decydującą o przypisaniu mu charakteru dokumentu pobytowego, a mianowicie możliwość potwierdzenia i weryfikacji posiadania tytułu prawnego do pobytu w Polsce. Trzeba w tym miejscu podkreślić, że Diia.pl – wbrew

---

[50] Art. 8 pkt 24a lit. d) ustawy z dnia 24 września 2010 r. o ewidencji ludności, t.j. Dz.U. z 2025 r. poz. 274 z późn. zm.

[51] Art. 4 ust. 2 specustawy z 2022 r.

[52] Wyrok NSA z 22.11.2023 r., II OSK 1062/23, OSP 2025, nr 3, poz. 31.

[53] Zob. art. 3 pkt 14 ustawy z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne, t.j. Dz.U. z 2024 r. poz. 1557 z późn. zm.

[54] Art. 4 ust. 5 specustawy z 2022 r.

[55] Art. 9 specustawy z 2022 r.

pojawiającym się niekiedy w różnego rodzaju publikacjach sformułowaniom – nie stanowi zezwolenia na pobyt w ścisłym znaczeniu[56]. Można w tym miejscu powtórzyć w ślad za Naczelnym Sądem Administracyjnym, że „objęcie określonej grupy obywateli Ukrainy ochroną czasową […] nie zostało uzależnione od wydania indywidulanego aktu administracyjnego […]. Przesłanki pozytywne nabycia ochrony czasowej na podstawie ustawy o pomocy to legalne przybycie przez obywatela Ukrainy na terytorium Rzeczypospolitej Polskiej z terytorium Ukrainy w związku z działaniami wojennymi prowadzonymi na terytorium tego państwa. Przesłanki negatywne nabycia ochrony czasowej obejmują natomiast przede wszystkim posiadanie tytułów pobytowych przez obywatela Ukrainy w chwili przybycia na terytorium RP"[57]. Potwierdzeniem uzyskania w ten sposób ochrony jest nadanie numeru PESEL z dopiskiem UKR, co z kolei stanowi czynność materialno-techniczną. Dokument elektroniczny Diia.pl pozwala na uzewnętrznienie faktu korzystania z ochrony czasowej w postaci umożliwiającej weryfikację w różnych sytuacjach, natomiast sam w sobie nie jest tytułem pobytowym.

Kolejna ważna funkcja dokumentu Diia.pl wywodzi się z treści art. 10 ust. 7 specustawy z 2022 r. Stanowi on, że elektroniczny dokument pobytowy wraz z dokumentem podróży uprawnia obywatela Ukrainy do wielokrotnego przekraczania granicy bez konieczności uzyskania wizy. W ramach Strefy Schengen uprawnia on (przy spełnieniu pozostałych warunków wjazdu i pobytu) do pobytu przez 90 dni w ciągu każdego 180-dniowego okresu[58]. Można zatem stwierdzić, że w tym zakresie Diia.pl pełni takie funkcje jak karta pobytu wydawana cudzoziemcom posiadającym prawo pobytu w Polsce na podstawie ustawy o cudzoziemcach. Równocześnie warto zauważyć, że w odniesieniu do posiadaczy Diia.pl przewidziano szerszy zakres uprawnień niż te, z których mogą korzystać obywatele polscy w ramach aplikacji mObywatel. Ustawa o aplikacji mObywatel, będąca fundamentem omawianych rozwiązań, przewiduje bowiem możliwość wydawania obywatelom polskim mobilnego dokumentu mObywatel, który stwierdza tożsamość i obywatelstwo polskie użytkownika na terytorium Rzeczypospolitej Polskiej w relacjach

---

[56] Tak np. Kancelaria Prezesa rady Ministrów, Rada do spraw Współpracy z Ukrainą, Polska Pomoc Ukrainie 2022-2023, s. 7, https://www.gov.pl/web/premier/rada-do-spraw-wspolpracy-z-ukraina2 [dostęo: 22.10.2025]; Serwis Rzeczypospolitej Polskiej, mObywatel 2.0, Diia.pl, https://www.gov.pl/web/cyfryzacja-badania-i-projektowanie-mobywatel20/diiapl; Warmińsko-Mazurski Urząd Wojewódzki w Olsztynie, *Diia.pl – pierwsze w UE w pełni cyfrowe pozwolenie na pobyt!,* 11.08.2022 r., https://www.gov.pl/web/uw-warminsko-mazurski/diiapl---pierwsze-w-ue-w-pelni-cyfrowe-pozwolenie-na-pobyt [dostęp: 22.10.2025].

[57] Wyrok NSA z 22.11.2023 r., II OSK 1062/23, OSP 2025, nr 3, poz. 3; zob. też R. Rogala, *Dopuszczalność zrzeczenia się ochrony czasowej. Glosa do wyroku Naczelnego Sądu Administracyjnego z dnia 22 listopada 2023 r., II OSK 1062/23, OSP 2025, nr 3, s. 31*, LEX.

[58] Zob. art. 6 ust. 1 kgS.

wzajemnej fizycznej obecności stron, ale nie uprawnia do przekraczania granicy państwowej[59].

Jeśli chodzi o funkcję dokumentu Diia.pl dotyczącą potwierdzania tożsamości, przepisy przewidują specyficzne rozwiązanie, różne od tego, jakie zastosowano w regulacjach dotyczących takich dokumentów jak dowód osobisty czy karta pobytu. W tych przypadkach przepisy ustaw – odpowiednio ustawy o dowodach osobistych[60] i ustawy o cudzoziemcach – stanowią *expressis verbis*, że wymienione dokumenty potwierdzają tożsamość obywatela lub cudzoziemca. Natomiast specustawa z 2022 r. takiego przepisu nie zawiera, lecz ustanawia delegację dla Rady Ministrów do określenia, w drodze rozporządzenia, procedur, w których dokument Diia.pl może być wykorzystywany do stwierdzania tożsamości obywatela Ukrainy. Wskazuje się przy tym, że Rada Ministrów, wydając rozporządzenie, winna mieć na uwadze ułatwienie sprawnego i bezpiecznego załatwiania spraw urzędowych obywatelom Ukrainy. Aktualnie obowiązujące rozporządzenie z dnia 5 października 2023 r. zawiera obszerny i przedmiotowo zróżnicowany katalog takich procedur[61]. Dokument może być zatem wykorzystywany m.in. zarówno do uzyskiwania pomocy lub korzystania z licznych uprawnień przewidzianych w specustawie dla jej beneficjentów, legitymowania się wobec Policji, Straży Granicznej i innych służb, okazywania przy kontroli granicznej, jak i udzielenia pełnomocnictwa pocztowego czy odbioru przesyłek rejestrowanych. Równocześnie jednak może powstać pytanie, czy w tym stanie prawnym *a contario* należałoby przyjąć, że w każdej niewymienionej w rozporządzeniu sytuacji formalne stwierdzenie tożsamości obywatela Ukrainy wymaga innego dokumentu, w szczególności dokumentu podróży. Wydaje się, że przy enumeratywnym wyliczeniu katalogu procedur, w których dokument potwierdzający status pobytowy może być wykorzystany do potwierdzenia tożsamości, a tym samym wykluczenia z niego innych procedur, może powstać w praktyce wiele wątpliwości w poszczególnych przypadkach. Bardziej operatywne wydaje się brzmienie Regulaminu aplikacji mObywatel, w którym zastosowano odwrotną regułę: dokument Diia.pl może być używany na terytorium Polski w każdej sytuacji życiowej podlegającej prawu polskiemu, z wyjątkiem sytuacji, gdy z przepisów prawa (ustaw, rozporządzeń) wynika obowiązek przedstawienia innych dokumentów stwierdzających tożsamość użytkownika, stwierdzających prawo użytkownika do

[59] Art. 2 pkt 8 oraz art. 7 ust. 5 pkt 1 ustawy o aplikacji mObywatel.

[60] Art. 4 ust. 1 ustawy z dnia 6 sierpnia 2010 r. o dowodach osobistych, t.j. Dz.U. z 2022 r. poz. 671 z późn. zm.; dalej: „ustawa o dowodach osobistych".

[61] Rozporządzenie Rady Ministrów z dnia 5 października 2023 r. w sprawie określenia procedur oraz usług, w których dokument elektroniczny oraz certyfikat mogą być wykorzystywane do stwierdzania tożsamości obywatela Ukrainy, Dz.U. z 2023 r. poz. 2156.

pobytu na terytorium RP lub innych praw przysługujących użytkownikowi[62]. Należy przy tym zwrócić uwagę, że powyższe brzmienie Regulaminu – choć wydaje się operatywne – jest równocześnie w pewnym zakresie sprzeczne z brzmieniem rozporządzenia z 5 października 2023 r., zakreślając potencjalnie szerszy zakres wykorzystania dokumentu.

Nakreślając charakterystykę dokumentu Diia.pl, należy nadmienić, że w Regulaminie aplikacji mObywatel podkreśla się, iż korzystanie z Diia.pl nie jest obowiązkowe. Różni to ów dokument od karty pobytu, którą – co do zasady – wydaje się z urzędu po udzieleniu cudzoziemcowi zezwolenia na pobyt, a dopiero jej wymiana lub wydanie kolejnej następuje na wniosek[63]. Oczywiście możliwa jest sytuacja, w której cudzoziemiec nie stawia się po odbiór karty pobytu, jednak samo jej pierwsze wydanie nie wymaga jakiegokolwiek działania ze strony cudzoziemca. Jest to również sytuacja odmienna od regulacji ustawy o dowodach osobistych, która przewiduje obowiązek posiadania dowodu osobistego przez pełnoletniego obywatela polskiego zamieszkującego w Polsce[64]. W przypadku Diia.pl, to obywatel Ukrainy musi dokonać szeregu czynności związanych z pobraniem i uruchomieniem aplikacji, a następnie pobraniem dokumentu. Wymaga to posiadania odpowiedniego sprzętu i kompetencji cyfrowych, co może niekiedy stanowić problem (o czym będzie mowa dalej), jednak w braku alternatywy wobec Diia.pl, wydaje się – pomimo braku formalnego obowiązku – praktycznie nieodzowne dla zwykłego codziennego funkcjonowania w Polsce. W szczególności, o ile potwierdzenie tożsamości jest możliwe na innej podstawie (przede wszystkim na podstawie dokumentu podróży), to nie ma alternatywnego względem Diia.pl dokumentu potwierdzającego status ochronny wynikający ze specustawy z 2022 r. Istnieje jedynie możliwość uzyskania na ogólnych zasadach zaświadczenia zawierającego pełny lub częściowy zakres danych przetwarzanych w rejestrze PESEL[65].

Dokument Diia.pl podlega unieważnieniu przez ministra właściwego do spraw informatyzacji w dwóch przypadkach. Najbardziej oczywistą okolicznością uzasadniającą takie działanie jest zgon obywatela Ukrainy odnotowany w rejestrze PESEL. Ponadto, unieważnienie następuje w przypadku utraty statusu ochronnego przyznanego na podstawie specustawy z 2022 r., co wyraża się zmianą statusu UKR. W praktyce następuje to w szczególności w związku z opuszczeniem terytorium Polski na okres przekraczający określony w ustawie czas.

---

[62] Zgodnie z § 2 ust. 3 załącznika Nr 8 do Regulaminu korzystania z aplikacji mObywatel, regulamin jest dostępny na stronach Ministerstwa Cyfryzacji, https://mc.bip.gov.pl/aplikacja-mobywatel/informacje-o-aplikacji-mobywatel.html [dostęp: 2.10.2025].

[63] Art. 229-230, 240-241 uoc.

[64] Art. 5 ust. 2 ustawy o dowodach osobistych.

[65] Art. 45 ustawy o ewidencji ludności.

5. Praktyczne problemy wdrażania elektronicznego dokumentu pobytowego w Polsce

Wdrożenie dokumentu Diia.pl zasadniczo odbyło się sprawnie, zważywszy na szybkie tempo i okoliczności. Nie udało się jednak całkowicie uniknąć pewnych problemów w praktyce. Dotyczą one przede wszystkim zapewnienia szeroko rozumianej dostępności i powszechności oraz bezpieczeństwa korzystania z dokumentu[66].

Jedną z pierwszych zidentyfikowanych kwestii problematycznych okazała się sytuacja osób małoletnich. W pierwotnym brzmieniu specustawy z 2022 r. nie przewidziano bowiem odpowiednich rozwiązań względem małoletnich poniżej 13. roku życia. Dopiero zmiany wprowadzone ustawą z dnia 9 marca 2023 r. o zmianie ustawy o cudzoziemcach oraz niektórych innych ustaw[67] pozwoliły na udostępnienie w publicznej aplikacji mobilnej funkcjonalności umożliwiającej pobranie, przechowywanie i prezentację dokumentu elektronicznego zawierającego pobrane z rejestru obywateli Ukrainy aktualne dane dotyczące dziecka obywatela Ukrainy pozostającego pod jego władzą rodzicielską oraz weryfikację integralności i pochodzenia tego dokumentu elektronicznego.

O ile problem dokumentu dla małoletnich można było rozwiązać stosunkowo łatwo, to aktualna pozostaje systemowa kwestia wykluczenia cyfrowego. Osoby wykluczone cyfrowo to osoby nieposiadające intelektualnych bądź materialnych środków, aby skorzystać z rozwiązań elektronicznych, np. osoby słabo wykształcone, bezrobotne, nieaktywne zawodowo, zamieszkujące tereny wiejskie, osoby starsze, trudna młodzież, imigranci czy mniejszości narodowe[68]. Wdrażanie innowacyjnych rozwiązań technologicznych powinno być dokonywane w środowisku zapewniającym zapobieganie wykluczeniu społecznemu. To ostatnie zadanie we współczesnych warunkach rozwojowych jest utożsamiane z pojęciem włączenia społecznego, a dotyczy tych jednostek czy grup społecznych, które nie są dostatecznie przygotowane do funkcjonowania w społeczeństwie informacyjnym[69]. Tymczasem w praktyce osoby starsze, osoby z niepełnosprawnością i osoby bez kompetencji cyfrowych lub środków materialnych pozwalających na posiadanie odpowiedniej klasy smartfona z wymaganym oprogramowaniem stają się równocześnie pozbawione możliwości korzystania z aplikacji mObywatel, a w dalszej perspektywie z wielu uprawnień

---

[66] Zob. też: J. Matusiak, A. Narożniak, *Diia.pl: Navigating legal, technological and human rights dimensions of an electronic document for Ukrainian citizens in Poland*, "Smart Cities and Regional Development Journal" 2025, Vol. 9, No 2, s. 24-27. https://doi.org/10.25019/nq2yrm66.

[67] Dz.U. z 2023 r. poz. 547 z późn. zm.

[68] C. Codagnone, D. Osimo, *Beyond i2010. E-Government current challenges and future scenarios*, [w:] P.G. Nixon, V.N. Koutrakou, R. Rawal, *Understading E-Government in Europe. Issues and Challenges*, New York 2010, p. 40, 43.

[69] K. Chałubińska-Jentkiewicz, M. Nowikowska, *Bezpieczeństwo, tożsamość, prywatność – aspekty prawne*, Legalis 2020.

gwarantowanych przepisami prawa. Problem ten zauważył m.in. Rzecznik Praw Obywatelskich, zwracając się równocześnie do Ministra Spraw Wewnętrznych i Administracji o rozważenie możliwości wydawania zaświadczeń o ochronie czasowej w formie papierowej, tak aby rozwiązanie proponowane w ramach aplikacji miało jedynie charakter fakultatywny. W odpowiedzi z Ministerstwa wskazano, że „przyjęcie rozwiązania zakładającego, że obywatelom Ukrainy, których pobyt na terytorium RP jest uznawany za legalny na podstawie specustawy, nie wydaje się tradycyjnych dokumentów pobytowych (w formacie tożsamym lub zbliżonym do karty pobytu) stanowiło jeden z przejawów konieczności takiego ukształtowania zadań organów władzy publicznej w obszarze obsługi uchodźców wojennych z Ukrainy, aby były one możliwe do realizacji przez te organy”[70]. Nie stwierdzono wprawdzie wprost, że uzupełnienie systemu o wydawanie dokumentów w tradycyjnej formie byłoby niemożliwe, ale wyraźnie zasugerowano, że wiązałoby się to z nadmiernym wysiłkiem organizacyjnym. Występowanie problemu wykluczenia cyfrowego z kolei rodzi pytanie o realizację zasady równości wobec prawa, co było już analizowane w różnym kontekście w nauce prawa, w tym w odniesieniu do postulowanych działań administracji publicznej[71].

Kolejna grupa problemów dotyczy możliwości korzystania z wszystkich funkcjonalności dokumentu i związana jest z uwarunkowaniami o charakterze technicznym. Zainstalowanie aplikacji mObywatel oraz pobranie dokumentu Diia.pl z oczywistych względów wymaga dostępu do Internetu. Z dokumentów dodanych wcześniej do aplikacji można korzystać niezależnie od dostępu do Internetu, przeglądając i okazując dokumenty w trybie *off-line*. Nie jest natomiast możliwe ich zaktualizowanie i potwierdzenie za pomocą kodu QR w tym trybie. W wielu sytuacjach może to być problematyczne, gdyż – jak wskazuje Ministerstwo Cyfryzacji – weryfikacja kodem QR to obecnie „najbardziej niezawodna metoda potwierdzania autentyczności dokumentu”, a samo okazanie go na ekranie smartfona może być niewystarczające, z uwagi na spotykane w praktyce przypadki fałszowania tego cyfrowego obrazu[72]. Inna potencjalna trudność wiąże się z problemami technicz-

[70] Zob. też *Aplikacja potwierdzająca status uchodźcy z Ukrainy nie dla osób poniżej 13. roku życia. Odpowiedź MSWiA,* Biuletyn Informacji Publicznej Rzecznika Praw Obywatelskich, 31.07.2023 i 24.03.2023; https://bip.brpo.gov.pl/pl/content/rpo-mswia-diia-pl-maloletni-odpowiedz [dostęp: 22.10.2025].

[71] M. Błażewski, *Administracja publiczna wobec wykluczenia cyfrowego*, [w:] J. Blicharz, M. Błażewski, T. Kocowski, A. Kordik, *Zasada równości wobec prawa a wykluczenie społeczne. Wybrane problemy i zagadnienia*, Uniwersytet Wrocławski 2023, s. 57-59, https://doi.org/10.34616/146424; E.M. Kwiatkowska, *Zdalne rozprawy – zapewnienie równości wobec zjawiska wykluczenia cyfrowego*, „Krytyka Prawa. Niezależne Studia nad Prawem” 2023, nr 15/3, s. 205-221.

[72] Ministerstwo Cyfryzacji, Kod QR – skuteczna metoda weryfikacji mDowodu, 25.02.2025, https://www.gov.pl/web/cyfryzacja/kod-qr--skuteczna-metoda-weryfikacji-mdowodu [dostęp: 23.10.2025].

nymi działania samej aplikacji, co może mieć charakter przejściowy, ale pomimo to może skutecznie uniemożliwiać posiadaczowi skorzystanie z jego uprawnień. Ten rodzaj problemów, o charakterze technicznym, dotyczy również scalania danych przechowywanych w rejestrach elektronicznych. Przykładowo, obywatel Ukrainy powracający do swego państwa celem uzyskania nowego dokumentu podróży, przekracza granicę na podstawie dwóch różnych dokumentów („stary" paszport przy wyjeździe, „nowy" przy powrocie). Jeśli dane o przekroczeniu granicy na podstawie różnych dokumentów nie są ze sobą scalone, skutkiem może być utrata nadanego wcześniej statusu UKR (w związku z uwidocznionym w systemie wyjazdem na okres przekraczający 30 dni i brakiem uwidocznienia powrotu). Pokłosiem takich problemów może być utrata dokumentu elektronicznego Diia.pl. Wskazane w tym miejscu problemy umacniają podstawy do formułowania postulatów dotyczących zapewnienia alternatywnych możliwości potwierdzania statusu pobytowego i posiadanych uprawnień z nim związanych.

Inny potencjalnie problematyczny obszar obejmuje ochronę danych osobowych i ochronę przed tzw. kradzieżą tożsamości. Przechowywanie danych w postaci elektronicznego dokumentu na jednym urządzeniu mobilnym może powodować obawę o ich bezpieczeństwo w razie kradzieży, zgubienia lub zniszczenia smartfona z aplikacją. Należy więc zwrócić uwagę na jakość systemów zabezpieczeń i odpowiednie przygotowane skutecznych sposobów reagowania na wypadek tego rodzaju zdarzeń. Zgodnie z art. 13 ust. 1 ustawy o aplikacji mObywatel minister właściwy do spraw informatyzacji ma obowiązek unieważnić certyfikaty użytkowników aplikacji mObywatel, jeżeli istnieje uzasadnione podejrzenie naruszenia ich bezpieczeństwa. Jest to kluczowy krok w procesie zabezpieczania informacji użytkowników. Decyzja o unieważnieniu powinna być adekwatna do rodzaju podejrzewanego naruszenia, co oznacza, że minister zobowiązany jest indywidualnie zbadać charakter i zakres zagrożenia, a zawieszenie świadczenia usług związanych z tymi certyfikatami jest konieczne do momentu, gdy bezpieczeństwo certyfikatów zostanie przywrócone[73]. Minister ogłasza niezwłocznie w Dzienniku Urzędowym Rzeczypospolitej Polskiej „Monitor Polski" informacje odpowiednio o unieważnieniu certyfikatów użytkowników aplikacji mObywatel i zawieszeniu oraz przywróceniu świadczenia usług związanych z tymi certyfikatami. Taka publikacja zapewnia, że wszyscy zainteresowani, w tym użytkownicy aplikacji mObywatel, mają świadomość podjętych – z uwagi na bezpieczeństwo – działań. Podobny tryb informowania obowiązuje w przypadku przywrócenia świadczenia usług związanych z uprzednio zawieszonymi lub unieważnionymi certyfikatami[74]. Poczucie bezpieczeństwa mają zapewnić

[73] M. Adamczyk, [w:] *Aplikacja mObywatel. Komentarz*, red. B. Kwiatek, A. Skóra, Warszawa 2025, LEX, art. 13.

[74] *Ibidem*.

również rozwiązania zawarte w Regulaminie korzystania z aplikacji mObywatel. Można tutaj zaliczyć takie rozwiązania jak brak możliwości uruchamiania aplikacji na niektórych urządzeniach, np. zmodyfikowanych i umożliwiających ukrywanie przełamania zabezpieczeń producenta urządzenia lub producenta oprogramowania (np. *rooting*, *jailbreaking*) czy w określonym środowisku, np. powszechnie uznanym za niebezpieczne. Aplikacja chroniona jest hasłem, nie przewiduje się usługi jego przywracania i jest ono podawane każdorazowo po uruchomieniu aplikacji, a trzykrotne wprowadzenie błędnego hasła powoduje czasową blokadę dostępu do aplikacji. Z innych wprowadzonych rozwiązań należy wymienić specjalny numer telefonu, pod którym można zgłosić unieważnienie pobranych na urządzenie certyfikatów, np. w przypadku kradzieży telefonu czy zabezpieczenie w postaci możliwości weryfikacji autentyczności dokumentu za pomocą kodu QR. Znowelizowana ustawa o aplikacji mObywatel[75] przewiduje, że w przypadku przywrócenia świadczenia usług związanych z wykorzystaniem certyfikatu podstawowego, który został unieważniony, użytkownikowi aplikacji mObywatel wydaje się nowy dokument mObywatel na jego żądanie złożone w tej aplikacji.

Zakończenie

Podsumowując powyższe rozważania dotyczące pierwszego elektronicznego dokumentu pobytowego dla cudzoziemców w Polsce, jakim jest Diia.pl, wypada przede wszystkim odnieść się do jego charakteru prawnego. Nie ulega wątpliwości, że stanowi on dokument pobytowy w rozumieniu przepisów prawa UE. Notyfikacja faktu tymczasowego wykorzystania go jako dokumentu poświadczającego legalność pobytu w Polsce w trybie art. 39 kgS pozwoliła bowiem na to, by dokument zasadniczo odbiegający od standardowej formy dokumentu pobytowego mógł uzyskać taki status. Równocześnie jednak należy stwierdzić, że jest to dokument pobytowy *sui generis* w kontekście prawa polskiego, gdyż – oprócz swej elektronicznej postaci – różni się także pod innymi względami od tradycyjnych dokumentów pobytowych wydawanych cudzoziemcom, jak i od podstawowego dokumentu tożsamości wydawanego obywatelom polskim, jakim jest dowód osobisty.

Odnosząc się do samej idei stworzenia i tymczasowego wykorzystania elektronicznego dokumentu pobytowego w sytuacji kryzysowej, jaką był masowy napływ osób uciekających na terytorium Polski przed skutkami rosyjskiej inwazji na Ukrainę, należy ocenić ją pozytywnie. Regulacje dotyczące dokumentu i praktyka jego wykorzystania wskazują, że koncepcja obarczona była pewnymi niedoskonałościami, jednak trzeba dostrzec, że polskie władze dopełniły starań o ich systema-

[75] Ustawa o zmianie ustawy o aplikacji mObywatel oraz niektórych innych ustaw z dnia 9 lipca 2025 r., Dz.U. z 2025 r., poz. 1019; nowelizacja wchodzi w życie z końcem 2025 r.

tyczne usuwanie. W konsekwencji, można śmiało rekomendować wykorzystanie elektronicznego dokumentu pobytowego w przypadku wystąpienia podobnych sytuacji kryzysowych.

Uwzględniając dotychczasowe doświadczenia można także sformułować pewne rekomendacje służące jak najpełniejszemu zabezpieczeniu praw jednostki i sprawności działania administracji. W szczególności wypada zalecić, by elektroniczny dokument pobytowy nie stanowił jedynego dostępnego środka weryfikacji tożsamości i uprawnień osoby objętej ochroną, lecz by tworzono – w miarę możliwości – alternatywne rozwiązania dla osób wykluczonych cyfrowo. Nie powinno dojść do sytuacji, w których ograniczenia finansowe lub technologiczne stanowiłyby o pozbawieniu danej osoby dokumentu lub możliwości korzystania z niego. Powinno to dotyczyć także osób małoletnich i ubezwłasnowolnionych, których sytuację należy brać pod szczególną rozwagę. Oczywistym, lecz niezwykle istotnym zaleceniem jest zapewnienie, by infrastruktura, w ramach której wykorzystywany jest elektroniczny dokument pobytowy, działała sprawnie, gwarantując zarówno bezpieczeństwo danych, jak i ich odpowiedni do potrzeb przepływ między instytucjami krajowymi i międzynarodowymi. Wydaje się również konieczne zadbanie o skonstruowanie takich ram prawnych, w których jasno wskazane zostałyby funkcje dokumentu. Regulacje dotyczące uprawnień i obowiązków osoby legitymującej się dokumentem, jak i samego obowiązku (albo fakultatywności) jego posiadania powinny znaleźć się w ustawie, nie zaś w aktach wykonawczych, regulaminach czy instrukcjach użytkowania rozwiązań technicznych. Zasadne byłoby zatem stworzenie standardów prawnych i technicznych dla elektronicznych dokumentów pobytowych jako elementu budowania zaufania do nowych, elektronicznych form identyfikacji.

W obecnych uwarunkowaniach prawnych, rozwiązanie polegające na wprowadzeniu elektronicznego dokumentu pobytowego jest i powinno być traktowane jako rozwiązanie szczególne, przydatne w sytuacjach kryzysowych. Nie ma przekonujących argumentów, które przemawiałyby za odejściem od spójnego w skali UE systemu dokumentowania prawa pobytu cudzoziemców. Trzeba jednak brać pod uwagę to, że w niedalekiej przyszłości uwarunkowania te mogą ulec zmianom. Już teraz trwają prace nad rozwojem europejskiej tożsamości cyfrowej i koncepcją „EU digital travel". Różne formy cyfryzacji, łącznie z wykorzystaniem narzędzi sztucznej inteligencji, wkraczają też od lat w dziedzinę procedur legalizacji pobytu cudzoziemców w Europie i na świecie. Nie można wykluczyć, że zmiany te obejmą też stopniowo dokumenty pobytowe. Wydaje się zatem, że prace nad rozwojem koncepcji elektronicznych dokumentów pobytowych dla cudzoziemców powinny być kontynuowane w ramach UE z myślą zarówno o potencjalnych sytuacjach kryzysowych, jak i zmianach technologicznych zachodzących we współczesnym świecie.

Literatura

Adamczyk M., [w:] *Aplikacja mObywatel. Komentarz*, red. B. Kwiatek, A. Skóra, Warszawa 2025, LEX, art. 13.

Błażewski M., *Administracja publiczna wobec wykluczenia cyfrowego*, [w:] *Zasada równości wobec prawa a wykluczenie społeczne. Wybrane problemy i zagadnienia*, red. J. Blicharz, M. Błażewski, T. Kocowski, A. Kordik, Uniwersytet Wrocławski 2023.

Chałubińska-Jentkiewicz K., Nowikowska M., *Bezpieczeństwo, tożsamość, prywatność – aspekty prawne*, Legalis 2020.

Codagnone C., Osimo D., *Beyond i2010. E-Government current challenges and future scenarios*, [w:] *Understading E-Government in Europe. Issues and Challenges*, red. P.G. Nixon, V.N. Koutrakou, R. Rawal, New York 2010.

Gajowniczek T, *E-gov w Ukrainie – zarys problemu*, „Środkowoeuropejskie Studia Polityczne" 2022, nr 4.

Górski M., *Wpływ specustawy ukraińskiej na sytuację prawną migrantów przymusowych z Ukrainy na polskim rynku pracy / The impact of the Ukrainian special law on the legal situation of forced migrants from Ukraine on the Polish labour market*, Centre of Migration Research, Working Papers 2023, nr 135/193.

Koch A., Biehler N., Knapp N., Kipp D., *Integrating refugees: Lessons from Germany since 2015-16. Background paper to the World Development Report 2023: Migrants, Refugees and Societies. April 2023*, https://www.worldbank.org/en/publication/wdr2023/backgroundpapers.

Kwiatkowska E.M., *Zdalne rozprawy – zapewnienie równości wobec zjawiska wykluczenia cyfrowego*, „Krytyka Prawa. Niezależne Studia nad Prawem" 2023, nr 15/3.

Matusiak J., Narożniak A., *Diia.pl: Navigating legal, technological and human rights dimensions of an electronic document for Ukrainian citizens in Poland*, „Smart Cities and Regional Development Journal" 2025, Vol. 9, No 2.

Pankevytch I., Bulat I., *Prawne aspekty e-administracji w Ukrainie: stan i perspektywy rozwoju*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3.

Rogala R., *Dopuszczalność zrzeczenia się ochrony czasowej. Glosa do wyroku Naczelnego Sądu Administracyjnego z dnia 22 listopada 2023 r., II OSK 1062/23, OSP 2025, nr 3, s. 31*, LEX.

Szachoń-Pszenny A., *Szczyt kryzysu migracyjnego w 2015 r. a szczyt kryzysu uchodźczego w 2022 r. – próba analizy porównawczej wpływu na „obszar bez granic" UE*, „Studia Politologiczne" 2023, nr 68.

Ustawa o pomocy obywatelom Ukrainy w związku z konfliktem zbrojnym na terytorium tego państwa. Komentarz, red. K. Szmid, P. Sawicki, Warszawa 2022, Legalis, art. 1.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Przemysław Pasierb
Uniwersytet Zielonogórski
Ministerstwo Cyfryzacji
ORCID: 0000-0003-3001-2566
przemyslaw.pasierb@onet.eu

Legal Regulation of New Technologies in the Context of Public Authority and State Governance

Keywords: artificial intelligence, public administration, digital technology

Summary. This article examines the role of new technologies, including artificial intelligence, in the exercise of public authority. The dynamic development of digitalisation and the processes taking place within contemporary societies indicate the necessity of equipping public administration with modern technological tools that enable the effective performance of tasks related to state governance. Particular attention must be paid to the fact that the scope of data processed by public administration creates a strong need for adequate safeguards ensuring the security of such data. Consequently, technological tools implemented in public institutions should meet the highest standards of authorisation and cybersecurity, while public officials should possess a high level of competence enabling the proper use of such solutions. At the same time, the state bears responsibility for establishing legal frameworks and clear guidelines regulating the use of artificial intelligence in public administration and providing a basis for the adoption of further internal regulations within administrative structures.

Regulacje prawne w zakresie nowych technologii w perspektywie sprawowania władzy publicznej i zarządzania państwem

Słowa kluczowe: Sztuczna inteligencja, administracja publiczna, nowe technologie

Streszczenie. Przedmiotem niniejszego opracowania jest wykorzystanie nowych technologii w tym sztucznej inteligencji w sprawowaniu władzy publicznej. Dynamika rozwoju cyfryzacji oraz procesów zachodzących w społeczeństwie wskazuje na potrzebę wyposażenia administracji publicznej w nowe narzędzia oparte o najnowsze rozwiązania informatyczne umożliwiające efektywne wykonywanie zadań z zakresu zarządzania państwem. Istotny jest fakt, że zakres danych przetwarzanych przez administrację stanowi o potrzebie bezwzględnego zabezpieczenia danych. W związku z powyższym dostarczane narzędzia powinny cechować się najwyższymi poziomami autoryzacji, a kadra urzędnicza wysokim poziomem wiedzy o korzystaniu z nich. Zadaniem państwa jest również stworzenie ram prawnych i jasnych wytycznych określających zasady użytkowania sztucznej inteligencji w urzędach oraz dające podstawy do dalszych wewnętrznych regulacji.

Introduction

The exercise of public authority by state bodies constitutes a broad and general concept, as it encompasses the performance of tasks at the local, regional and national levels. It should therefore be emphasised that activities falling within the notion of state governance may relate both to the responsibilities of local self-government units and to those performed directly by the Council of Ministers.

For this reason, the identification of the objectives of public administration requires the determination of the target group or the sphere of activity to which administrative actions are directed. The scope of issues encountered by public administration on a daily basis is extensive and multifaceted. Some of these responsibilities are defined, inter alia, in Articles 6–8 of the ustawa o samorządzie gminnym[1], Article 4 and, where applicable, Article 4a of the ustawa o samorządzie powiatowym[2], Articles 11 and 14 of the ustawa o samorządzie województwa[3], as well as in the ustawa o działach administracji rządowej[4], not to mention numerous specific statutes regulating particular areas of administrative activity.

Given the multiplicity of tasks performed by public administration, ensuring effective planning mechanisms becomes essential. At the same time, it should be noted that the role of public administration itself is evolving. It is no longer limited to individual administrative bodies providing services to citizens; increasingly, it involves active participation in planning and shaping public policy programmes[5]. Attention should also be drawn to the autonomy of regions in designing their own development policies, which results from the decentralisation of public authority. As a consequence, regional centres have gained greater capacity to make decisions concerning their own socio-economic development, which ultimately contributes to the development of the state as a whole[6]. National strategies, which fall within the competence of members of the Council of Ministers, are also subject to consultation within the framework of cooperation between central government administration and local self-government authorities[7].

---

[1] Ustawa z dnia 8 marca 1990 r. o samorządzie gminnym (Dz. U. z 2025 r. poz. 1153 i 1436).

[2] Ustawa z dnia 5 czerwca 1998 r. o samorządzie powiatowym (Dz. U. z 2025 r. poz. 1684).

[3] Ustawa z dnia 5 czerwca 1998 r. o samorządzie województwa (Dz. U. z 2025 r. poz. 581 i 1535).

[4] Ustawa z dnia 4 września 1997 r. o działach administracji rządowej (Dz. U. z 2025 r. poz. 1275 i 1846).

[5] A. Z. Kamiński, *Administracja publiczna we współczesnym państwie*, "Przegląd Socjologiczny" 2004, 53(2), p. 135.

[6] A.W. Jabłoński, *Uwarunkowania i kierunki rozwoju administracji publicznej w Polsce po 1989 roku*, [in:] *Administracja publiczna i polityki publiczne. Tom poświęcony pamięci Profesora Ryszarda Herbuta*, Wałbrzych 2025 r., p. 45-56.

[7] D. Sześciło, M. Niziołek, *Administracja współpracująca. Zarządzanie publiczne jako kooperacja*, [in:] *Administracja i zarządzanie publiczne. Nauka o współczesnej administracji*, red. D. Sześcidło, Warszawa 2014, p. 284.

The extensive responsibilities of contemporary public administration, viewed not only from the perspective of providing services to citizens but also from the perspective of supporting development and governance processes, create a need for the state to possess modern technological tools and access to a broad range of information resources necessary for the effective exercise of the competences entrusted to public authorities.

I. Legal Foundations for use od new tchnologies in public administration

In accordance with the well-established principle derived from Article 7 of the Constitution of the Republic of Poland[8], the activities of public administration must be strictly grounded in law. Consequently, while the aforementioned provisions and specific statutes provide the legal basis for undertaking actions in particular areas, it is equally necessary to consider the legal foundations for the use of new digital tools designed to facilitate the performance of administrative tasks.

The primary legal framework governing the use of artificial intelligence is the AI Act[9], which establishes regulatory guidelines for both private and public entities, including local self-government units[10]. The provisions of the regulation are further complemented by national legislation[11]. Nevertheless, the provision of public services and the conduct of state governance with the use of artificial intelligence require not only legal regulation but also the development of an appropriate expert and research environment.

For this reason, the IDEAS Research Institute, dedicated to research involving artificial intelligence, has been established[12]. According to the regulation establishing the institute, it operates under the supervision of both the Minister of

[8] Konstytucja Rzeczpospolitej Polskiej z dnia 2 kwietnia 1997 r. (Dz. U. nr 78 poz. 483 z późn. zm.).

[9] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 24 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Dz. Urz. UE. z 2024 r. str. 1689).

[10] K. Szpyt, A. Bilski, *Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych*, "Prawo i Więź" 2025, no. 1 vol. 54, p. 583.

[11] In Poland, a draft Act on Artificial Intelligence Systems (UC71) (projekt ustawy o systemach sztucznej inteligencji) is currently under consideration. The current stage of the legislative process indicates that the draft is being examined within the framework of intra-governmental proceedings. At present, consultations are ongoing within the Stały Komitet Rady Ministrów (status as of 20 February 2026).

[12] Rozporządzenie Rady Ministrów z dnia 25 lutego 2025 r. w sprawie utworzenia Instytutu Badawczego IDEAS (Dz. U. poz. 238).

Digital Affairs and the Minister of National Defence. The division of supervisory responsibilities reflects the scope of tasks performed by the institute. Breakthrough technologies, including artificial intelligence, have a direct impact on issues related to national defence; therefore, the involvement of the Minister of National Defence contributes to strengthening the security of the state. The supervisory responsibilities of the minister include, inter alia, the development of autonomous systems, advanced data analysis technologies and the security of digital solutions, which are essential for the protection of critical infrastructure, military operations and responses to contemporary security threats[13].

At the same time, the supervision exercised by the Minister of Digital Affairs is equally important, as artificial intelligence constitutes one of the key pillars of Poland's digital transformation. AI technologies support the development of smart cities, digital public administration and innovative services for citizens and entrepreneurs. Through cooperation with the Ministry of Digital Affairs, the IDEAS Research Institute may contribute to the development of ethical and effective technological solutions aimed at improving both the accessibility and the quality of public services. Furthermore, the supervisory role of the minister enables the integration of artificial intelligence research with broader digital transformation initiatives, including the development of data infrastructure, the implementation of advanced computing platforms and cybersecurity initiatives within public administration[14].

It should also be emphasised that the scope of activities carried out by the institute is not limited solely to areas assigned to ministers under the ustawa o działach administracji rządowej. According to the regulation establishing the institute, research involving artificial intelligence may also be conducted in the fields of natural sciences, engineering and technology, as well as social sciences. The institute may also undertake research aimed at strengthening Poland's position in the international arena and supporting the development of public administration services[15].

Within the legal framework there are also provisions which, although not directly related to the use of artificial intelligence itself, enable advanced analytical processes that had not previously been employed in the development of public policies and administrative tasks. As a result, the provision of public services and

[13] Justification to the draft rozporządzenia Rady Ministrów w sprawie utworzenia Instytutu Badawczego IDEAS (RD171), https://legislacja.rcl.gov.pl/projekt/12394003/katalog/13108636#13108636 (13.02.2026 r.).

[14] *Ibidem*.

[15] § 2 ust. 2 pkt 1-5 rozporządzenia IDEAS.

the management of state governance may be supported by comprehensive data and information resources made available to public authorities.

II. Technological solutions

It is a natural consequence that the organisation of state governance, of which public administration constitutes an essential element, has been developed primarily on the basis of formal procedures. These procedures, in turn, maintain their continuity thanks to the civil service corps, which ensures compliance with the applicable guidelines. Nevertheless, in order to perform the assigned tasks, appropriate tools are necessary, including those based on information systems.

The implementation of technologically advanced solutions within public administration requires the central authorities not only to introduce amendments in the field of legislation, but also to modify the operational schemes of the decision-making structures functioning at particular levels of the administrative hierarchy. At the same time, the awareness of public officials should be gradually increased through centrally organised training programmes and workshops.

It should be emphasised that any procedural changes will not be effective without properly trained public officials, understood as individuals who consciously and competently use the new technologies provided by those managing public organisational units. Furthermore, attention must also be paid to issues related to security and the protection of information, including the protection of processed data. For this reason, a cautious and pragmatic approach to the use of new technologies is justified, particularly with regard to solutions provided by commercial entities.

In response to the emerging expectations within public administration, models based on domestic technological solutions are being developed, including those involving the use of artificial intelligence. In this context, reference should be made to PLLuM (Polish Large Language Model), which was created in order to adapt a language model to standards and realities based on the grammar of the Polish language. The model is primarily intended to understand the aforementioned rules of the Polish language, but also the socio-cultural context and the specific characteristics of administrative language. The PLLuM library has been developed on the basis of resources originating from scientific institutions, public repositories and entities that voluntarily supported its development[16].

[16] https://www.gov.pl/web/cyfryzacja/pllum--polski-model-jezykowy-jak-dziala-i-do-czego-moze-sie-przydac [accessed on: 12.02.2026].

Experts and research institutions participated in the project, including those from Politechnika Wrocławska, the Instytut Podstaw Informatyki Polskiej Akademii Nauk, and the Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy[17]. The objective of the project was not only to create a language model capable of learning and evolving within the framework of large language models, but also to develop a prototype of a virtual assistant intended to support the work of public administration[18].

Thanks to the involvement of the project team, it was possible to develop an intelligent assistant specialised in retrieving information on the basis of queries formulated in natural language and contextual analysis and generating final responses. On the basis of this solution, a service dedicated to Polish public administration has been developed[19].

III. Use of modern technologies by public officials

As has already been emphasised in this article, regardless of the introduction of soft organisational changes in the organisation of work or the implementation of solutions requiring amendments to legislation, it is necessary to ensure an adequate level of awareness and competence among public officials. Persons responsible for introducing new technologies into the activities of public administration must first conduct a review of the institutional environment and determine the current needs for potential interventions or support for the administrative staff.

For this reason, the Ministry of Digital Affairs, recognising the necessity of identifying the aforementioned factors, undertook the task of examining the needs of public officials in the context of the use of artificial intelligence tools. This task was entrusted to a specialised research institution: Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy (NASK-PIB), which committed itself to carrying out the project entitled "AI in Public E-Administration: The Administrative and Institutional Perspective – Diagnosis of the Needs and Attitudes of Public Officials and Implementation Possibilities with Consideration of Practical Conditions"[20].

According to the study prepared by NASK-PIB, the objective of the project was to collect opinions among public officials concerning the problems related to

[17] https://science.nask.pl/research-areas/projects/7573 [accessed on: 12.02.2026].
[18] *Ibidem*.
[19] https://pllum.org.pl/blog/posts/zakonczenie-projektu-pllum [accessed on: 13.02.2026].
[20] The project was implemented on the basis of targeted grant agreement No. 5/WB/DBI/2025 concluded between the Minister of Digital Affairs and Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy (NASK-PIB).

the implementation of artificial intelligence in public administration and, where possible, to develop new administrative processes[21].

The study was conducted via the Internet, and invitations to participate in the survey were sent using the ePUAP and e-Delivery systems. A total of 6,093 questionnaires were obtained, the content of which indicates, for example, that artificial intelligence may most frequently be used in the work of public offices for the translation and editing of content in foreign languages (77% of responses)[22]. The next most frequently indicated uses were searching for and analysing existing documentation (72%) and the creation of large datasets (69%)[23]. The lowest ranking concerned the drafting of legal acts, justifications and administrative decisions (42%)[24].

These findings correspond with the report prepared by the Urząd Ochrony Danych Osobowych, which also emphasised that artificial intelligence is used primarily in operations aimed at improving everyday tasks, such as writing texts, handling clients and supporting communication[25].

As demonstrated by the research conducted by NASK-PIB, attitudes towards new technologies, particularly artificial intelligence, among managerial staff are diversified. It is possible to observe a favourable attitude, for example within ministries, where the approach towards officially available tools may be described as positive. At the same time, there is also a critical approach manifested in reluctance to use unofficial solutions[26].

Reluctance to use solutions that have not been authorised by the organisation is, of course, justified, particularly in light of the above-mentioned report of the Urząd Ochrony Danych Osobowych and the issue of the security of processed data. Even where such information does not constitute personal data, it may still include internal organisational information. It should also be noted that the greatest vulnerabilities occur among entities processing large volumes of data. In this context, the UODO identifies universities, local self-government units, educational institutions and smaller entrepreneurs[27].

---

[21] Final report of the product developed under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).

[22] *Ibidem*.

[23] *Ibidem*.

[24] *Ibidem*.

[25] Raport Strategiczny – Badanie potrzeb organizacji w zakresie wykorzystania sztucznej inteligencji i ochrony danych osobowych, p. 4, https://uodo.gov.pl/pl/589/4058 [accessed on: 16.02.2026].

[26] Annex No. 1 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).

[27] Raport Strategiczny – Badanie potrzeb organizacji…, p. 14.

Consequently, it has been observed that awareness of the use of artificial intelligence in public administration is in a certain sense twofold. On the one hand, there is widespread knowledge of the advantages and opportunities associated with such tools, for example in terms of increased productivity, new possibilities or practical assistance. On the other hand, a certain weakness may be observed due to the numerous threats associated with their use, such as the risk of data leakage[28]. Nevertheless, it should be emphasised that employees of state institutions demonstrate a high level of awareness of the responsibility associated with access to and use of information of state significance. For this reason, the level of concern regarding data security when using artificial intelligence tools is higher among public officials than within society in general[29].

An important observation is that employees of central administration perceive greater potential for the use of artificial intelligence in almost every area of activity. The greatest difference concerns document archiving and the assignment of file reference numbers (15 percentage points). Interestingly, local government officials more frequently indicated that artificial intelligence could be used for the drafting of legal acts. However, as the authors of the study emphasise, this response was relatively rarely indicated by respondents overall[30].

In light of the above, the conducted study led to the following conclusion: digital transformation at the municipal level and the implementation of new technologies require two different types of solutions, as two distinct working ecosystems exist. Cities resemble corporate environments and therefore require technological "multitools" enabling mass data processing and the automation of hundreds of documents. In rural municipalities, by contrast, a public official performs multiple roles, and the supporting technology must therefore function as a virtual assistant that facilitates knowledge retrieval and provides support in nearly every administrative matter[31].

It should also be emphasised that the problem of technology being treated merely as an additional layer of work must be addressed. Situations still occur in which so-called "double work" is performed, meaning that tasks are carried out both within ICT systems and in paper form. Consequently, the implementation of another technological solution, even more advanced, would remain ineffective, as it would not in practice reduce the workload of public officials. In some cases,

[28] Annex No. 1…
[29] *Ibidem*.
[30] Final report…
[31] Annex No. 2 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).

it may even become an additional solution requiring time and effort rather than providing genuine support[32].

At the municipal level, a phenomenon of decision-making paralysis can be observed due to the lack of clear guidelines regarding the use of artificial intelligence tools within administrative units. Public officials expect clear legal grounds defining the possibility of using AI. As the authors of the study point out: "Local officials are paralysed by legal risk. No one wants to be the first to 'let AI in' and expose the municipality to problems. There is a widespread expectation of central certification – a kind of 'security seal' issued by the state or a trusted provider. The ministry should promote solutions that relieve mayors of the burden of deciding on compliance with the GDPR, for example by guaranteeing secure cloud infrastructure or conducting audits. Without such a 'security stamp', decision-making at the local level will remain frozen"[33].

At the county level, two areas requiring urgent intervention can be identified. The first concerns relieving employees of routine manual tasks, which currently constitute a waste of the intellectual potential of administrative staff who must repeatedly rewrite data between systems. This indicates that the problems identified at the municipal and county levels are largely similar and relate to the continued engagement of employees in paper-based tasks that could already be performed digitally[34].

The second clearly visible need for intervention concerns the architecture of IT systems in terms of interoperability. At present, a lack of sectoral cooperation and the closed nature of supplier environments constitute barriers to the implementation of new technological solutions[35]. At the same time, the current delays in implementing artificial intelligence within public administration, when compared with the private sector, should not be interpreted as administrative inefficiency. The objectives of these sectors differ significantly. While the private sector aims primarily at maximising profits, the fundamental mission of public administration is to serve society and protect the public interest[36].

Nevertheless, the study also indicates that at the county level there is a clear expectation that formal guidelines defining the legal framework for the use of

---

[32] *Ibidem*.

[33] *Ibidem*.

[34] Annex No. 3 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).

[35] Annex No. 3…

[36] J. Blicharz, L. Zacharko, *Wdrażanie technologii sztucznej inteligencji w administracji publicznej - kilka refleksji*, [in:] *Administracja publiczna wobec procesów zmian w XXI wieku. Księga jubileuszowa Profesora Jerzego Korczaka*, red. P. Lisowski, W. Jakimowicz, Wrocław 2024, p. 354-355.

artificial intelligence should be introduced, as well as the designation of an entity that could act as a leader in the standardisation of such solutions[37].

Conclusion

The overall subject of state governance constitutes an extremely extensive field, as in addition to performing activities related to the day-to-day functioning of the state, public authorities are also responsible for developing public policies. These activities are characterised by their broad scope, and their range continues to expand over time. Furthermore, society itself constitutes an active organism and a complex structure subject to numerous interdependencies and external stimuli. As a result, the authorities responsible for state governance or for the development of public policies are in practice unable to determine in advance what kinds of challenges they may face in either the distant or even the near future.

A clear example of such a phenomenon occurred in the spring of 2020, when a profound negative transformation of social life took place, accompanied by an unprecedented modification of the scope of activities carried out by state authorities. For this reason, the state must possess not only a specialised corps of public officials who, by virtue of their competencies, are capable of undertaking actions necessary to ensure the functioning of the state, but also modern tools supporting them in the performance of these tasks.

Nevertheless, the introduction of new tools must be conditioned by certain standards. Above all, security considerations must be preserved, which means that solutions used by public administration should meet appropriate authorisation standards. At the same time, as demonstrated by research conducted by NASK-PIB, there exists an expectation at the local level for clear guidelines concerning the use of new technologies. Consequently, several essential elements may be identified with regard to the introduction of artificial intelligence in public administration:

1) the establishment of general legal frameworks providing overall guidelines regarding the possibility of using modern technological tools within public administration;
2) following the introduction of such general frameworks, the adoption of sectoral or internal regulations specifying in detail the scope of the use of available technological solutions;
3) the development of domestic language models and ICT systems, or the establishment of cooperation with service providers guaranteeing the highest security standards, where possible;

[37] Annex No. 3…

4) the differentiation of technological solutions according to their intended users, taking into account the different needs of central administration, counties, municipalities, as well as institutions such as inspections or educational supervisory authorities;
5) the elimination of the phenomenon of so-called "double work", meaning the simultaneous handling of cases in both paper and digital form;
6) training and education of public officials in the use of the tools provided, so that such tools become genuine support in the performance of administrative duties.

Provided that the above conditions are fulfilled, the performance of tasks related to state governance may acquire greater dynamism and flexibility in responding to external variables that directly affect the functioning of the state. The execution of tasks by public authorities is possible only with the support of an efficient and well-qualified civil service corps, and when combined with modern technologies provided by the state, the resilience of public administration to constantly emerging challenges becomes relatively higher.

Legal provisions as well as internal organisational regulations must therefore be adapted to the realities currently prevailing, particularly with regard to the principles governing the performance of tasks by public administration. Positive developments worthy of emphasis include events that have taken place in Poland in recent years, such as the development of domestic language models intended for use in public administration, advanced legislative work concerning the regulation of artificial intelligence, and the establishment of a research institute dedicated to the study of artificial intelligence.

The above clearly demonstrates that the emerging technological challenges are recognised by state authorities, which are actively responding to the needs arising from contemporary socio-economic reality.

References

Konstytucja Rzeczpospolitej Polskiej z dnia 2 kwietnia 1997 r. (Dz. U. nr 78 poz. 483 z późn. zm.).
Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 24 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Dz. Urz. UE. z 2024 r. str. 1689).
Ustawa z dnia 8 marca 1990 r. o samorządzie gminnym (Dz. U. z 2025 r. poz. 1153 i 1436).
Ustawa z dnia 4 września 1997 r. o działach administracji rządowej (Dz. U. z 2025 r. poz. 1275 i 1846).
Ustawa z dnia 5 czerwca 1998 r. o samorządzie powiatowym (Dz. U. z 2025 r. poz. 1684).
Ustawa z dnia 5 czerwca 1998 r. o samorządzie województwa (Dz. U. z 2025 r. poz. 581 i 1535).
Rozporządzenie Rady Ministrów z dnia 25 lutego 2025 r. w sprawie utworzenia Instytutu Badawczego IDEAS (Dz. U. poz. 238).

Justification to the draft Act rozporządzenia Rady Ministrów w sprawie utworzenia Instytutu Badawczego IDEAS (RD171)https://legislacja.rcl.gov.pl/projekt/12394003/katalog/13108636#13108636 (13.02.2026 r.).
Draft Act on Artificial Intelligence Systems (UC71).
Blicharz J., Zacharko L., *Wdrażanie technologii sztucznej inteligencji w administracji publicznej - kilka refleksji*, [in:] red. P. Lisowski, W. Jakimowicz, Administracja publiczna wobec procesów zmian w XXI wieku. Księga jubileuszowa Profesora Jerzego Korczaka, Wrocław 2024, p. 349-356.
Jabłoński A. W., *Uwarunkowania i kierunki rozwoju administracji publicznej w Polsce po 1989 roku*, [in:] *Administracja publiczna i polityki publiczne. Tom poświęcony pamięci Profesora Ryszarda Herbuta*, red. B. Detyna, P. Szymaniec, Wałbrzych 2025, p. 39-56.
Kamiński A. Z., *Administracja publiczna we współczesnym państwie*, "Przegląd Socjologiczny" 2004, 53(2), p. 133–149.
Sześciło D., Niziołek M., *Administracja współpracująca. Zarządzanie publiczne jako kooperacja*, [in:] *Administracja i zarządzanie publiczne. Nauka o współczesnej administracji*, red. D. Sześcidło, Warszawa 2014, p. 284.
Szpyt K., Bilski A., *Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych*, "Prawo i Więź" 2025, no. 1, vol. 54, p. 565-596.
Final report of the product developed under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).
Annex No. 1 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).
Annex No. 2 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).
Annex No. 3 to the final report of the project implemented under targeted grant agreement No. 5/WB/DBI/2025 (as of the date of submission of this article, the report has not yet been published by the Ministry of Digital Affairs).
Raport Strategiczny – Badanie potrzeb organizacji w zakresie wykorzystania sztucznej inteligencji i ochrony danych osobowych, p. 4, https://uodo.gov.pl/pl/589/4058 (16.02.2026 r.).
https://www.gov.pl/web/cyfryzacja/pllum--polski-model-jezykowy-jak-dziala-i-do-czego-moze-sie-przydac [accessed on: 12.02.2026].
https://science.nask.pl/research-areas/projects/7573 [accessed on: 12.02.2026].
https://pllum.org.pl/blog/posts/zakonczenie-projektu-pllum [accessed on: 13.02.2026].

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Amelia Pic
Uniwersytet Zielonogórski
ORCID: 0009-0001-6463-8797
a.ameliapic@gmail.com

Zastępstwo przez emerytowanego notariusza – zagadnienia wybrane

Słowa kluczowe: notariat, emerytowany notariusz, zastępstwo przez emerytowanego notariusza, sprawowanie zastępstwa

Streszczenie. Niniejszy artykuł przedstawia wybrane zagadnienia dotyczące sprawowania zastępstwa przez emerytowanego notariusza. W pierwszej kolejności omówiona została ewolucja przepisów regulujących wyznaczenie emerytowanego notariusza do zastępstwa wraz z wykazaniem wątpliwości interpretacyjnych pojawiających się na gruncie wykładni językowej art. 21 i 22 Prawa o notariacie. Omówiono także kwestię odpowiednich kompetencji emerytowanego notariusza wykonującego obowiązki w ramach zastępstwa w odniesieniu do wymogów dotyczących notariusza przy dokonywaniu czynności notarialnych, w szczególności obowiązku stałego podnoszenia kwalifikacji zawodowych. Zwrócono także uwagę na brak ograniczeń w zakresie podejmowanych przez emerytowanego notariusza czynności, które mogłyby uchybiać powadze wykonywanego zawodu lub kolidować z wykonywanymi obowiązkami w ramach zastępstwa.

Substitution by a retired notary – selected issues

Keywords: notariat, retired notary, substitution by a retired notary, acting as a substitute

Summary. This article presents selected issues concerning the performance of substitution duties by a retired notary. First, it discusses the evolution of the legal provisions governing the appointment of a retired notary as a substitute, while highlighting interpretative doubts arising from the linguistic interpretation of Articles 21 and 22 of the Notarial Law. The article also examines the issue of the qualifications and competencies of a retired notary performing duties within the framework of such substitution, particularly in relation to the requirements imposed on notaries when carrying out notarial acts, including the obligation of continuous professional development. Attention is also drawn to the absence of restrictions on activities undertaken by a retired notary that could undermine the dignity of the profession or conflict with the duties performed in the course of substitution.

1. Powołanie do zastępstwa

1.1. Ewolucja regulacji dotyczących zastępstwa notariusza pod rządami ustawy z 1991 r.

Sprawowanie zastępstwa przez emerytowanego notariusza zostało unormowane w ścisłym związku z regulacjami dotyczącymi procedury wyznaczenia zastępstwa w razie nieobecności notariusza i chociaż odnoszą się one także do innych podmiotów wskazanych przez ustawodawcę, wymagają uprzedniego przedstawienia i wyjaśnienia jako niezbędnej podstawy do dalszych rozważań niniejszej pracy. W pierwszej kolejności koniecznym jest zatem omówienie treści art. 21 ustawy z dnia 14 lutego 1991 r. – Prawo o notariacie[1], który od początku regulował zasady czasowego zastępstwa nieobecnego notariusza. W pierwotnym brzmieniu przepis ten przewidywał trzy podstawowe sytuacje: krótkotrwałą nieobecność do 3 dni (§ 1), planowaną nieobecność dłuższą niż 3 dni, lecz nieprzekraczającą 3 miesięcy (§ 2), oraz nieobecność nagłą, wynikającą z nieprzewidzianych przyczyn, gdy notariusz nie zdążył wyznaczyć zastępcy (§ 3). W każdej z tych sytuacji prawo do wyznaczenia zastępcy przysługiwało w pierwszej kolejności notariuszowi, a w przypadku niedokonania stosownego wyznaczenia przez notariusza – prezesowi rady izby notarialnej. Ustawodawca dopuszczał, aby funkcję zastępcy pełnił asesor notarialny z kancelarii notariusza lub innej kancelarii w tej samej izbie, a także przewidziano możliwość wyznaczenia do zastępstwa emerytowanego notariusza (§ 4).

Istotną zmianę tej regulacji wprowadzono ustawą z dnia 13 czerwca 2013 r. o zmianie ustaw regulujących wykonywanie niektórych zawodów[2]. Nowelizacja miała charakter systemowy i wiązała się z likwidacją instytucji asesora notarialnego jako etapu poprzedzającego możliwość powołania na notariusza. W efekcie art. 21 został istotnie przekształcony – zarówno pod względem struktury, jak i treści. W nowym brzmieniu art. 21 § 1 przewiduje możliwość wyznaczenia zastępstwa spośród zastępców notarialnych – nowej kategorii zawodowej wprowadzonej do ustawy w miejsce asesorów – zatrudnionych w kancelarii notariusza, notariuszy współprowadzących kancelarię w trybie art. 4 § 3 pr.not., a także zastępców notarialnych wpisanych do wykazu izby notarialnej oraz notariuszy mających siedzibę na jej obszarze, przy czym wyznaczenie osoby spoza kancelarii wymaga porozumienia z prezesem rady izby.

Znowelizowany przepis uchyla całkowicie § 2 – co oznacza dalszą rezygnację z formalnego zróżnicowania zasad wyznaczania zastępstwa w zależności od

[1] Ustawa z dnia 14 lutego 1991 r. – Prawo o notariacie, Dz.U. Nr 22, poz. 91, dalej: pr.not.

[2] Ustawa z dnia 13 czerwca 2013 r. o zmianie ustaw regulujących wykonywanie niektórych zawodów, Dz.U. z 2013 r. poz. 829.

długości planowanej nieobecności[3]. Jednocześnie § 3 wprowadza ograniczenie zakresu zastępstwa w przypadku jego wyznaczenia przez prezesa izby w sytuacjach nagłych – do czynności enumeratywnie wymienionych w art. 79 pkt 6a i 7 pr.not. oraz do wydawania rzeczy przyjętych na przechowanie. Tak sformułowany przepis stanowi ograniczenie kompetencji zastępcy wyznaczonego z urzędu, które nie istniało w poprzednim stanie prawnym. Paragraf 4 zachowuje możliwość powierzenia zastępstwa notariuszowi emerytowanemu, co stanowi kontynuację wcześniejszego rozwiązania.

Nowelizacja art. 21 odzwierciedla dwie istotne tendencje legislacyjne: konsekwencję reformy struktury ścieżki zawodowej w notariacie (likwidacja asesury i wprowadzenie instytucji zastępcy notarialnego)[4] oraz dążenie do większej kontroli i formalizacji procedur zastępstwa, szczególnie w sytuacjach nagłych. Ustawodawca ograniczył uznaniowość i zakres swobody działania podmiotów mogących sprawować zastępstwo, wprowadzając mechanizmy nadzorcze ze strony organów samorządu oraz uszczegóławiając katalog dopuszczalnych czynności. Działanie ustawodawcy można interpretować jako próbę wzmocnienia gwarancji bezpieczeństwa obrotu prawnego i nadzoru nad czynnościami notarialnymi dokonywanymi przez osoby inne niż notariusz prowadzący kancelarię.

Uzupełnieniem regulacji dotyczącej krótkoterminowego zastępstwa jest art. 22 pr.not., który odnosi się do sytuacji o charakterze trwałym lub długookresowym – takich jak zaprzestanie prowadzenia kancelarii przez notariusza bądź jego zawieszenie w czynnościach służbowych. Przepis ten, w pierwotnym jego brzmieniu[5], przewidywał, że „w razie zaprzestania prowadzenia kancelarii przez notariusza lub zawieszenia go w czynnościach, rada właściwej izby notarialnej wyznacza czasowo do pełnienia obowiązków notariusza zastępcę spośród asesorów notarialnych danej izby".

Powyższa regulacja miała na celu zapewnienie ciągłości działalności kancelarii notarialnej w sytuacji, gdy notariusz z przyczyn formalnych lub faktycznych nie mógł jej prowadzić. Wyznaczenie zastępcy należało do kompetencji rady izby notarialnej, a jego grono było ograniczone do asesorów notarialnych wpisanych na listę tej izby. Nowelizacja wprowadzona ustawą zmieniającą z dnia 30 czerwca 2005 r.[6]

---

[3] Podtrzymana została regulacja w zakresie § 2 zastosowana wcześniejszą nowelizacją art. 21 ustawy – Prawo o notariacie wprowadzona art. 3 pkt. 6 ustawy z dnia 30 czerwca 2005 r. zmieniająca ustawę – Prawo o adwokaturze i niektóre inne ustawy (Dz.U. Nr 163, poz. 1361 z późn. zm.), która pozostaje w opracowaniu jedynie odnotowana, bez podejmowania jej szczegółowej analizy z uwagi na marginalne znaczenie dla omawianej problematyki niniejszego artykułu.

[4] Zob. A. Oleszko, R. Pastuszko, *Sprawczość kompetencyjna zastępcy notarialnego*, „Rejent" 2023, nr 4, s. 102-104.

[5] T.j. Dz.U. z 2002 r. Nr 42, poz. 369.

[6] Ustawa z dnia 30 czerwca 2005 r. zmieniająca ustawę – Prawo o adwokaturze i niektóre inne ustawy, Dz.U. Nr 163, poz. 1361, art. 3 pkt 7.

złagodziła obligatoryjność tego rozwiązania – rada izby „może" wyznaczyć (a nie „wyznacza") zastępcę – oraz ograniczyła jego stosowanie jedynie do przypadków zaprzestania prowadzenia kancelarii „z innych powodów niż śmierć notariusza lub zawieszenia go w czynnościach". Oznaczało to zawężenie zakresu stosowania przepisu i nadanie radzie większej swobody decyzyjnej.

Jednocześnie utrzymano, że funkcję zastępcy może pełnić wyłącznie asesor notarialny danej izby. Najbardziej zasadnicze zmiany wprowadzono ustawą z dnia 13 czerwca 2013 r.[7], której celem była deregulacja dostępu do niektórych zawodów oraz gruntowna przebudowa ścieżki zawodowej w notariacie. Nowe brzmienie art. 22 znosiło odniesienie do asesora notarialnego i wprowadzało szerszy katalog potencjalnych zastępców: zastępców notarialnych, notariuszy oraz notariuszy emerytowanych danej izby notarialnej. Co istotne, przepis zawężał zakres czynności, które mogą być wykonywane przez wyznaczoną osobę – wyłącznie do tych określonych w art. 79 pkt 6a i 7 pr.not. oraz czynności związanych z wydawaniem rzeczy przyjętych na przechowanie.

1.2. Wyznaczenie zastępstwa przez notariusza

W ustawie prawo o notariacie została przewidziana możliwość zastępstwa notariusza przy dokonywaniu czynności notarialnych w sytuacji gdy notariusz nie może sprawować swojego urzędu. Ustawodawca w regulacji instytucji zastępstwa wskazał, że zastępstwo może sprawować także notariusz emerytowany. Określono trzy możliwości ustanowienia emerytowanego notariusza jako zastępcy.

Przede wszystkim wyboru emerytowanego notariusza do wykonywania zastępstwa może dokonać sam notariusz poprzez świadome wskazanie konkretnej osoby[8]. Ustawodawca dokonuje wyodrębnienia i podziału podmiotów mogących sprawować zastępstwo, ustalając, że notariusz jest obowiązany do poinformowania prezesa rady właściwej izby notarialnej o wyznaczeniu do zastępstwa osoby spośród a) zastępców notarialnych zatrudnionych w kancelarii notariusza; b) notariuszy ze wspólnej kancelarii prowadzonej w trybie art. 4 § 3 pr.not. Kolejno, zgodnie z treścią przepisu, możliwe jest wyznaczenie do zastępstwa a) osoby spośród zastępców notarialnych z wykazu właściwej izby notarialnej; b) notariuszy mających siedzibę kancelarii notarialnej w tej izbie, lecz w porozumieniu z prezesem rady właściwej izby notarialnej. Podział ten zastosowano w celu pozostawienia samorządowi notarialnemu pewnego rodzaju nadzoru i możliwości ingerencji przy wyborze osoby z dwóch konkretnych grup podmiotów, określając, że do skutecznego wyznaczenia

[7] Ustawa z dnia 13 czerwca 2013 r. o zmianie ustaw regulujących wykonywanie niektórych zawodów, Dz.U. poz. 829, art. 6 pkt 7.

[8] Art. 21 § 1 w zw. z § 3 ustawy z dnia 14 lutego 1991 r. Prawo o notariacie, Dz.U. z 2024 r. poz. 1001.

zastępcy może dojść poprzez porozumienie. Wskazanie takiej formy jest jednoznaczne z wyrażeniem obopólnej zgody notariusza, jak i prezesa właściwej rady notarialnej.

Kwestię zastępstwa emerytowanego notariusza ustawodawca reguluje w § 4 art. 21 pr.not., wskazując jedynie, że notariusz może wyznaczyć do zastępstwa emerytowanego notariusza. Emerytowani notariusze zostali więc uwzględnieni jako kolejna grupa podmiotów mogących zastępować notariusza, jednak nie określono, w jakim trybie wyznaczenia mogą być umocowani do zastępstwa. Ustawodawca nie precyzuje, czy notariusz przy wyznaczeniu jednego z emerytowanych notariuszy do zastępstwa w trybie art. 21 § 1 pr.not powinien tylko zawiadomić prezesa właściwej rady notarialnej, czy też adekwatne byłoby dokonanie umocowania za wcześniejszym porozumieniem, jak w przypadku zastępstwa przez zastępcę notarialnego wybranego z właściwego wykazu czy też innego notariusza mającego siedzibę kancelarii w tej samej izbie. Forma, jaką ustawodawca reguluje zastępstwo emerytowanego notariusza, jest na tyle szeroko ujęta, że budzi wątpliwości co do skutecznego powołania do zastępstwa emerytowanego notariusza, jak i jego odpowiedniego umocowania. Skoro ustawodawca wszelkie pozostałe możliwości wyznaczenia zastępstwa poprzez samego notariusza reguluje, ustalając ich formę, to odpowiednio sytuacja emerytowanego notariusza wyznaczonego do zastępstwa powinna być określona chociażby przez wskazanie, czy skuteczne jego ustanowienie ma być dokonane tylko za uprzednim poinformowaniem, czy też porozumieniem z prezesem właściwej rady izby notarialnej. Jeśli przewidujemy możliwość takiego wyznaczenia, emerytowani notariusze jako grupa osób zakwalifikowanych do powołania na zastępcę powinna zostać określona w sposób bardziej dokładny, poprzez umiejscowienie tego typu podmiotu w podziale, jaki zastosował ustawodawca w art. 21 § 1 pr.not.

Jednym z rozwiązań, które zasługuje na rozważenie, jest wprowadzenie obowiązku uprzedniego zawarcia pisemnego porozumienia pomiędzy notariuszem prowadzącym kancelarię a konkretnym notariuszem emerytowanym, potencjalnie przewidywanym do zastępstwa w razie nieobecności. Treść takiego porozumienia powinna jasno określać wolę obu stron do podjęcia zastępstwa oraz gotowość emerytowanego notariusza do świadczenia czynności notarialnych w przewidzianym zakresie. Zawarcie takiego porozumienia miałoby charakter prewencyjny i organizacyjny, dając podstawę do szybkiego działania w sytuacjach nagłych, bez konieczności prowadzenia doraźnych ustaleń personalnych lub formalnych. Tego rodzaju praktyka mogłaby znacząco usprawnić proces umocowania notariusza emerytowanego, a ponadto sprzyjałaby przejrzystości działań podejmowanych w ramach zastępstwa. Pozwoliłaby także prezesowi rady właściwej izby notarialnej na wcześniejszą weryfikację osoby przewidzianej do zastępstwa pod kątem jej

predyspozycji zawodowych, aktualnych kompetencji oraz zachowania wysokiego standardu usług notarialnych. Co więcej, istnienie takiego dokumentu mogłoby stanowić dodatkowy punkt odniesienia przy ocenie prawidłowości wykonywania zastępstwa i ewentualnej odpowiedzialności za działania podjęte w jego ramach.

Zastępstwo notariusza, w sytuacji gdy on sam dokonuje wyboru osoby zastępującej, nie zostało zakreślone ramami co do czynności, których zastępstwo może obejmować. Ustawodawca nie wskazuje również ram czasowych wyznaczenia zastępstwa, a także warunków formalnych takiego umocowania[9]. Chociaż brak owych regulacji dotyczy każdej grupy podmiotów mogących sprawować zastępstwo, to w odniesieniu do emerytowanego notariusza jest on szczególnie istotny z uwagi na indywidualny status tej kategorii, wynikający z faktu nieprzynależności do samorządu notarialnego. Z uwagi na brak wskazania zakresu czynności, które mogłyby być wykonywane w ramach zastępstwa, należy przyjąć, że zastępstwo może wedle uznania[10] obejmować wszelkie czynności notarialne. Należy podkreślić, że ustalenie ram wykonywanego zastępstwa przez obie strony porozumienia wiąże się z odpowiednią, należycie sformułowaną treścią, która w dostateczny sposób wykaże zakres umocowania. Jest to istotne z uwagi na specyfikę czynności notarialnych, które są złożonym procesem obejmującym całokształt zachowań uczestników postępowania notarialnego, jak i samego notariusza lub w ramach zastępstwa – emerytowanego notariusza. Niewątpliwie okoliczności każdej indywidualnej sprawy, która będzie mogła zostać przeprowadzona w ramach zastępstwa, może wymagać podjęcia różnych kolejnych działań przez osobę sprawującą zastępstwo, co do których powinna ona mieć odpowiednie umocowanie[11].

1.3. Wyznaczenie zastępstwa przez samorząd notarialny

W sytuacji, gdy notariusz nie wyznaczy zastępstwa, a z powodu nieprzewidzianych przyczyn jest on nieobecny, zastępcę wyznacza prezes rady właściwej izby notarialnej. Również w takich okolicznościach ustawodawca przewiduje możliwość ustanowienia emerytowanego notariusza do sprawowania zastępstwa podczas nieobecności notariusza prowadzącego kancelarię. Tego typu rozwiązanie ma służyć zapewnieniu ciągłości funkcjonowania kancelarii notarialnej oraz ochronie

[9] Zob. R. Wrzecionek, *Pojęcie i charakter prawny wyznaczenia zastępstwa przez notariusza – ocena aktualnych regulacji i postulaty de lege ferenda*, „Rejent" 2016, nr 6, s. 138-14.

[10] W kontekście woli notariusza co do formy i treści wykonywanego zastępstwa za uznaniem osoby wykonywującej zastępstwo, tj. emerytowanego notariusza, ale także innego notariusza lub zastępcy notarialnego.

[11] Szerzej u A. Oleszko, R. Pastuszko, *Tryb dokonywania czynności notarialnych jako postępowanie o charakterze cywilnoprawnym*, „Rejent" 2010, nr 7-8, s. 85-88.

interesów prawnych osób korzystających z jej usług[12]. Ustawodawca postanowił umożliwić wyznaczenie zastępcy także spośród notariuszy emerytowanych, co w zasadzie poszerza grupę podmiotów mogących takie zastępstwo sprawować, jednak w rzeczywistości wyznaczenie zastępcy przez organ samorządu notarialnego w okolicznościach nagłej nieobecności notariusza powinno przede wszystkim być skutecznym i pewnym rozwiązaniem służącym zachowaniu wszelkich cech zawodu zaufania publicznego.

Zauważyć należy, że Krajowa Rada Notarialna prowadzi, zgodnie z informacjami przekazywanymi przez rady izb notarialnych, listę notariuszy oraz zastępców notarialnych, która zawiera imię i nazwisko, właściwą izbę notarialną, datę i numer decyzji o powołaniu notariusza lub o jego odwołaniu, datę umieszczenia w wykazie zastępców notarialnych lub skreślenia z tego wykazu, a także informacje o zawieszeniu w czynnościach zawodowych[13]. Wybór kandydata do pełnienia zastępstwa spośród grupy aktywnych zawodowo notariuszy, jak i zastępców notarialnych dzięki ich sklasyfikowaniu jest w rzeczywistości ułatwiające dla weryfikacji odpowiedniego kandydata. Wybór emerytowanego notariusza będzie pociągać za sobą konieczność skutecznego sprawdzenia, czy konkretna osoba będzie spełniała konieczne wymogi dla zapewnienia należytego sprawowania tej funkcji. Ustawodawca nie reguluje żadnego kryterium, przy którym taki wybór będzie relatywnie odpowiednim. Pozostawienie tak szerokiej regulacji, może mieć negatywne skutki w postaci wykreowania błędnej praktyki, która w rzeczywistości nie będzie służyć rozwiązaniu tego typu problemu.

Nieobecność notariusza, zwłaszcza nagła i nieprzewidziana, zaburza funkcjonowanie działalności notarialnej prowadzonej przez notariusza, w tym także naraża bezpieczeństwo obrotu prawnego czy też interes klientów kancelarii. Forma wyznaczenia zastępstwa w trybie art. 21 § 3 pr.not. z pewnością ma za zadanie zapobiec powyższym problemom. Nie można w tej kwestii pominąć także faktu, że wyznaczenie zastępstwa w tych okolicznościach powinno być bezzwłoczne. Emerytowany notariusz jest nieaktywny zawodowo, a więc zweryfikowanie jego możliwości i adekwatności w wyborze na zastępcę może być zbyt przewlekłe lub też niedostatecznie przeprowadzone z uwagi na utrudnienia związane z dodatkowymi czynnościami sprawdzającymi. W nagłych przypadkach wybór kandydata musi być gwarantem zachowania ciągłości funkcjonowania kancelarii zgodnie z wymaganiami określonymi w ustawie.

---

[12] Zob. bliżej W. Boć, M. Leśniak, *Notariusz jako przedsiębiorca szczególnego rodzaju*, „Acta Universitatis Wratislaviensis" 2011, No 3304; „Przegląd Prawa i Administracji" LXXXIV, s. 17 w odniesieniu do stanowiska M. Szydło, *Swoboda prowadzenia działalności gospodarczej i swoboda świadczenia usług w prawie Unii Europejskiej*, Toruń 2005, s. 55.

[13] Art. 41a § 1 pr.not.

Notariusze i zastępcy notarialni objęci są przez ustawę pr.not. odpowiednimi wymaganiami[14], które weryfikowane są przez samorząd notarialny. W ramach wyznaczenia do zastępstwa emerytowanego notariusza ustalenie kandydata powinno opierać się na weryfikacji bez zbędnej zwłoki i w uzasadniany co do wyboru konkretnej osoby sposób. Uwzględnienie emerytowanych notariuszy do wykonywania zastępstwa poprzez wyznaczenie konkretnej osoby przez jednostkę samorządu notarialnego w rzeczywistości poszerza grupę możliwych kandydatów mogących sprawować takie zastępstwo, jednak wybór ten, zwłaszcza w przypadku wyboru emerytowanego notariusza, powinien sprowadzać się do istotnej przyczyny, dla której został on wyznaczony.

2. Emerytowany notariusz wyznaczany do zastępowania

2.1. Możliwość sprawowania zastępstwa

Przejście na emeryturę oznacza formalne zaprzestanie praktyki w danej dziedzinie, co może wiązać się z dalszą niezdolnością do wykonywania określonych czynności. Ustalenie, że emerytowani notariusze należą do grupy podmiotów mogących sprawować zastępstwo lub być wyznaczeni do takiego zastępstwa, powinno zakreślać także szczególne normy regulujące tę instytucję. Czynni przedstawiciele samorządu notarialnego podlegają *de facto* przepisom ustawowym, ale także etyce zawodowej oraz wewnętrznej organizacji samorządu ustalanej na mocy prawa. W sytuacji, gdy emerytowany notariusz zostanie wyznaczony do pełnienia zastępstwa, jego odpowiedzialność za podjęcie tego zobowiązania jest wciąż nieuregulowana dostatecznie, by móc jednoznacznie określić jej treść. Notariusz będący na emeryturze może nie być w stanie należycie wykonywać wyznaczonych czynności lub też całkowicie być niezdolnym do podjęcia zastępstwa.

Z uwagi na zmianę statusu notariusza w momencie przejścia na emeryturę należy rozważyć sposób stosowania względem emerytowanych notariuszy sprawujących zastępstwo przepisów prawa o notariacie odnoszących się do notariuszy. Niewątpliwie założenie, że zakończenie aktywnego wykonywania zawodu przez notariuszy nie uchybia faktowi posiadania doświadczenia i odpowiednich kwalifikacji, jest argumentem przemawiającym za możliwością powierzenia im określonych czynności zastępczych[15]. Natomiast sama regulacja wykonywania zastępstwa przez emerytowanego notariusza budzi wiele wątpliwości z uwagi na brak adekwatnych ustaleń w zakresie ciągłego poszerzania wiedzy.

[14] Patrz art. 11 pr.not. oraz art. 76 pr.not.
[15] Zob. wyrok WSA w Warszawie z 7.09.2006 r., sygn. VI SA/Wa 1168/06, Lex nr 255855.

Istotnym aspektem w zawodzie notariusza jest stałe podnoszenie swoich kwalifikacji zawodowych. Wymóg ten został ujęty w art. 17 pr.not., a także w § 9 Kodeksu Etyki Zawodowej Notariusza[16]. Rozwój w zawodzie notariusza dla efektywnego wykonywania pracy dotyczyć może nie tylko wiedzy teoretycznej w zakresie obwiązujących norm prawnych, ale także wiedzy technicznej dotyczącej chociażby wszelkiego rodzaju systemów teleinformatycznych[17], których stosowanie w praktyce przyczynia się do usprawnienia wykonywanej pracy oraz podnoszenia jakości świadczonych usług. Biorąc pod uwagę dynamiczny rozwój technologii, jest to celowe ukształtowanie tego wymogu. Emerytowani notariusze niekorzystający z różnego rodzaju szkoleń w tym zakresie mogą więc nie być bieżąco w tematyce regulacji systemów niezbędnych do wykonywania czynności notarialnych z uwagi na brak możliwości praktycznego ich stosowania, co w konsekwencji przejawiać się może także brakiem znajomości w ich funkcjonowaniu. Z uwagi na wchodzące w życie liczne zmiany dotyczące czynności praktycznych w zawodzie notariusza, jak i przepisów prawa, wymagane jest od przedstawicieli tego zawodu stałe kształcenie i podnoszenie kwalifikacji zawodowych[18].

Przyjęcie założenia, że emerytowany notariusz w konsekwencji przejścia na emeryturę nie posiada odpowiednich kwalifikacji zawodowych jest zbyt daleko idące, jednakże ustawodawca wykazuje konieczność podnoszenia kwalifikacji, nawiązując do jakości świadczonych usług, jednocześnie przyznając prawo do wykonywania czynności notarialnych emerytowanym notariuszom, co do których realizacja niniejszego wymogu nie jest możliwa z uwagi na ich status. Ustalając konieczność ciągłego poszerzania kwalifikacji jako elementu niezbędnego w wykonywaniu zawodu notariusza, należałoby ustalić odpowiednie wymagania, które musi realizować emerytowany notariusz mogący w ramach zastępstwa dokonywać czynności notarialnych.

---

[16] Uchwała Nr 19 Krajowej Rady Notarialnej z 12.12.1997 r. – Kodeks Etyki Zawodowej Notariusza, Dz.U. z 1997 r., 12.12.

[17] Systemy, które zostały uwzględnione w przepisach ustawy Prawa o notariacie dotyczące prowadzenia list notariuszy i zastępców notarialnych (art. 41a), możliwości złożenia wniosku o wpis w księdze wieczystej (art. 92 § 4), Centralne Repozytorium Elektronicznych Wypisów Aktów Notarialnych (art. 92a), rejestracja aktu poświadczenia dziedziczenia (art. 95h), system teleinformatyczny do prowadzenia Rejestru Spadkowego (art. 95i), wpis europejskiego poświadczenia spadkowego do Rejestru Spadkowego (art. 95y).

[18] A. Oleszko, *Prawo o notariacie. Komentarz*, t. 1: *Ustrój notariatu*, Warszawa 2016, art. 17, s. 466-467.

2.2. Zakres odpowiedniego stosowania regulacji o czynnościach notarialnych dokonywanych przez emerytowanego notariusza

Sprawowanie zastępstwa polega na wykonaniu obowiązków za inną osobę, a w kontekście czynności notarialnych także w ramach kompetencji, jakie posiada notariusz. Chociaż oczywistym wydawać się może, że sprawowanie zastępstwa w istocie służy dokonaniu tych czynności, które sam notariusz wykonałby, gdyby jego obecność była możliwa, to przy ustaleniu, że zastępstwo sprawować będzie emerytowany notariusz, pojawia się problem w określeniu, czy przepisy dotyczące sposobu wykonywania czynności notarialnych powinny być analogicznie stosowane do podmiotu nieaktywnego zawodowo, a także spoza grupy podmiotów tworzących samorząd notarialny.

Ustawodawca *in principio* określa, że notariusz działa jako osoba zaufania publicznego i korzysta z ochrony przysługującej funkcjonariuszom publicznym[19]. W przeważającej mierze kształtuje to indywidualną i szczególną pozycję notariusza przy wykonywaniu czynności notarialnych, zważywszy, że ma to swoje odniesienie do wszelkiej odpowiedzialności[20]. Ustawa pr.not. reguluje zakres czynności notarialnych, które w ramach zastępstwa można dokonać tylko w określonych przypadkach[21], w pozostałym zakresie nie przedstawiono żadnych ograniczeń. Analizując tę kwestię w perspektywie dążenia do zachowania ciągłości pracy, bez uszczerbku dla klientów korzystających z usług kancelarii, wydawać się to może uzasadnione, jednakże biorąc pod uwagę emerytowanych notariuszy, jako podmioty o niedostatecznym sformalizowaniu w przepisach dotyczących notariatu, można stwierdzić, że wymagane powinno być ustalenie tych obowiązków i praw, które w istocie przysługiwałyby oraz obowiązywały emerytowanych notariuszy w sposób jednoznacznie wskazany.

Ustawodawca nie określa, jakie cechy powinien posiadać emerytowany notariusz. W rzeczywistości może stanowić o tym założenie, że emerytowany notariusz prowadził niegdyś działalność notarialną, więc musiał spełniać wszelkie cechy przypisane notariuszom. Problem natomiast może pojawić się w sytuacji, gdy forma ukształtowanych do tej pory cech lub wymogów ulegnie zmianie. Ustawodawca odnosi się do wielu kwestii, które dotyczą dokonywania czynności notarialnych, a co do których nałożony jest szczególny wymóg ich stosowania lub przestrzegania.

---

[19] Art. 2 § 1 pr.not.

[20] Zob. M. Nowocień, *Odmowa dokonania czynności notarialnej*, Warszawa 2018, s. 19 oraz P. Marquardt, *Notariusz jako funkcjonariusz publiczny*, „Krakowski Przegląd Notarialny", rok IV, nr 22, s. 95-101., a także A. Oleszko, *Ustrój polskiego notariatu w świetle orzecznictwa Trybunału Konstytucyjnego*, „Studia Iuridica Lublinensia" 2014, nr 22, s. 552-553.

[21] Wykonywanie zastępstwa na mocy art. 21 § 3 pr.not. oraz art. 22 pr.not.

Przykładowo zachowanie w tajemnicy okoliczności sprawy zostało przewidziane jedynie w odniesieniu do notariusza[22]. Analogicznie obowiązek czuwania nad należytym bezpieczeństwie praw i słusznych interesów stron oraz innych osób, dla których czynność ta może powodować skutki prawne[23], a także obowiązek udzielania wszelkich wyjaśnień dotyczących dokonywania czynności notarialnej[24]. Kwestię sporną stanowi także depozyt papierów wartościowych oraz pieniędzy, które notariusz w ramach swojej działalności ma prawo przyjąć na przechowanie[25]. Wszelkie określone powyżej wymagania w zakresie pracy notariusza są niezbędnym elementem całokształtu działalności notarialnej i nie powinny budzić wątpliwości, czy wykonywanie czynności notarialnych w ramach zastępstwa w odniesieniu do norm zawartych w przepisach pr.not. winno być także stosowane wobec emerytowanego notariusza sprawującego zastępstwo. Gdyby jednak niezachowanie wymogów w ustawie wskazanych nastąpiło, podkreślenia wymaga istotny brak regulacji odpowiedzialności emerytowanego notariusza w tym zakresie.

3. Podjęcie przez emerytowanego notariusza sprawującego zastępstwo dodatkowego zatrudnienia lub zajęcia

W ustawie pr.not. przewidziano ograniczenia stosowane wobec notariuszy w zakresie należytego zachowania staranności zawodowej, a także powagi zawodu. Jednym z ograniczeń, jakie wprowadza ww. ustawa, wynika z treści art. 19 pr.not., zgodnie z którym notariusz nie może podejmować zatrudnienia bez uprzedniej zgody rady właściwej izby notarialnej, z wyjątkiem zatrudnienia w charakterze pracownika naukowo-dydaktycznego, dydaktycznego lub naukowego, chyba że wykonywanie tego zatrudnienia przeszkadza w pełnieniu jego obowiązków. Ustawodawca w § 2 doprecyzowuje, że notariuszowi nie wolno także podejmować zajęcia, które mogłoby kolidować w pełnieniu obowiązków albo uchybiać powadze wykonywanego zawodu. Przepis wprowadza ponadto zakaz zajmowania się handlem, przemysłem, pośrednictwem i doradztwem w interesach, a nadto § 3 określa obowiązek zawiadomienia prezesa rady właściwej izby notarialnej o zamiarze podjęcia takiego zatrudnienia lub zajęcia.

Norma zawarta w art. 19 ma charakter gwarancyjny i etyczny. Jej celem jest zapewnienie, aby osoba wykonująca zawód zaufania publicznego, jakim jest notariusz, zachowała pełną niezależność oraz bezstronność przy dokonywaniu czynności

[22] Art. 18 § 1 pr.not.
[23] Art. 80 § 2 pr.not.
[24] Art. 80 § 3 pr.not.
[25] Rozważania w tym kontekście poczynione przez J. Sawarzyński, *Podstawy odmowy dokonania czynności notarialnej*, Poznań 2021, s. 280-281 w nawiązaniu do wypowiedzi A.J. Szereda, *Czynności notarialne. Komentarz do art. 79-112 Prawa o notariacie*, Warszawa 2018, s. 517.

notarialnych. Zakaz podejmowania dodatkowych zajęć ma chronić powagę urzędu i autorytet notariatu, a także zapobiegać konfliktom interesów mogącym zagrozić wiarygodności notarialnego poświadczenia w obrocie prawnym[26]. W judykaturze podkreśla się, że art. 19 pr.not. jest wyrazem zasady, iż notariusz – jako osoba pełniąca funkcje publiczne – powinien koncentrować się wyłącznie na czynnościach notarialnych, a każda inna aktywność zawodowa wymaga zgody organu samorządu, który ocenia, czy nie zostanie naruszona powaga i bezstronność zawodu[27].

Obostrzenie w zakresie zatrudnienia lub podejmowania się dodatkowego zajęcia ustawodawca przewiduje także dla aplikantów notarialnych i zastępców notarialnych[28]. Wątpliwości interpretacyjne pojawiają się jednak w odniesieniu do emerytowanych notariuszy, którzy nie są już członkami samorządu notarialnego i nie podlegają jego nadzorowi. Rozważenia wymaga, czy wobec takiej osoby powinny mieć zastosowanie ograniczenia z art. 19 pr.not., skoro formalnie utraciła ona status notariusza, a ustawa nie reguluje jej obowiązków etycznych i zawodowych w czasie wykonywania zastępstwa.

Emerytowany notariusz może więc podejmować dowolne zajęcia zawodowe, gospodarcze czy społeczne bez obowiązku uzyskania zgody rady izby. Nie podlega też obowiązkowi informowania o swojej dalszej działalności zawodowej, jeśli takowej się podejmie. W sytuacji, gdy zostaje on powołany do zastępstwa, powstaje ryzyko, że brak takich mechanizmów kontrolnych doprowadzi do konfliktu interesów lub podważenia bezstronności czynności notarialnych. Skoro notariusz w stanie spoczynku – powołany do zastępstwa – wykonuje czynności o identycznym charakterze prawnym jak notariusz czynny, powinien również podlegać podobnym standardom zachowania, jak i ograniczeniom.

Tymczasem brak takiego obowiązku prowadzi do sytuacji, w której emerytowany notariusz może jednocześnie wykonywać działalność gospodarczą, zasiadać w organach spółek lub uczestniczyć w przedsięwzięciach o charakterze komercyjnym, a jednocześnie sprawować funkcję zastępcy notariusza dokonującego czynności w imieniu państwa. Taka sytuacja może wręcz osłabiać gwarancję bezstronności i niezależności notariusza, które stanowią istotę urzędu notarialnego. Ponadto może dojść do sytuacji, w której dla obejścia ograniczeń ustawowych, jakie ustawodawca nakłada na notariuszy w zakresie dokonywanych czynności notarialnych[29], wyznaczony do ich realizacji zostanie emerytowany notariusz w ramach pełnienia

[26] P. Marquardt, [w:] *Prawo o notariacie. Komentarz*, red. W. Gonet, wyd. 2, Warszawa 2025, art. 19.

[27] Zob. wyrok SN z 27.09.2012 r., sygn. SDI 23/12, Lex nr 1360316; postanowienie SN z 29.09.2005 r., SDI 23/05, Lex nr 568824; wyrok WSA w Warszawie z 8.12.2022 r., sygn. VI SA/Wa 775/22, Lex nr 3783147; wyrok SN z 14.10.2014 r., sygn. SDI 27/14, Biul.PK 2014, nr 10, poz. 30-32.

[28] Art. 78 pr.not.

[29] Przykładowo ograniczenia wynikające z art. 19 pr.not, 84 pr.not., art. 86 pr.not.

czasowego zastępstwa. Jest to istotny aspekt na gruncie późniejszego ustalenia odpowiedzialności za działania naruszające fundamentalne cechy tego zawodu. W konsekwencji, jeśli działania emerytowanego notariusza nie są sprzeczne z obowiązującym prawem, mogą one prowadzić do obniżenia autorytetu zawodu oraz podważać zaufanie publiczne do instytucji notariatu.

Rozszerzenie ograniczenia wynikającego z art. 19 pr.not. na grupę emerytowanych notariuszy mogłoby natomiast budzić zastrzeżenia konstytucyjne, bowiem art. 65 ust. 1 Konstytucji RP[30] gwarantuje każdemu wolność wyboru i wykonywania zawodu, a ograniczenia tej wolności mogą wynikać wyłącznie z ustawy i tylko w zakresie niezbędnym dla ochrony ważnych wartości publicznych. Ustalenie, iż pełnienie funkcji notariusza, bez względu na aktywne wykonywanie zawodu, jak i przejście w stan spoczynku, musiałoby wiązać się z dożywotnimi ograniczeniami w sferze zawodowej oraz być trwale pozbawionym prawa do podejmowania zajęć sprzecznych z powagą tego zawodu. Niewątpliwie takie rozwiązanie byłoby przejawem nieposzanowania praw i wolności człowieka.

Dla adekwatnego zapobieżenia powyższym zagrożeniom zasadne wydaje się rozważenie stosowania ograniczenia, jednakże tylko w wąskim zakresie. Zasługuje na aprobatę stanowisko Adama Tkaczyńskiego, wyrażające pogląd, iż zasadne byłoby stosowanie instytucji art. 19 pr.not. tylko w sytuacji, gdy emerytowany notariusz podejmie się sprawowania zastępstwa. W przeciwnym razie nie będzie on podlegał pod zakres wynikający z powyższej normy[31]. Warunkowana w ten sposób legitymacja do podjęcia się czynności notarialnych przez emerytowanego notariusza powinna stanowić gwarancję należytego dokonania czynności, także w odniesieniu do bezstronności osoby zastępującej. W odmiennym przypadku brak jest podstaw do podważenia działań dokonanych z naruszeniem wartości, których ochronę stanowi ograniczenie w sferze zawodowej, jak i zajęć dodatkowych względem notariusza.

Rozważenia wymaga także ustalenie, w jaki sposób stosowanie art. 19 pr.not wobec emerytowanych notariuszy mogłoby przebiegać w aspekcie praktycznego sposobu weryfikowania niemożliwości sprawowania zastępstwa z powodu zajmowania stanowiska, pracy, zajęcia lub innych czynności sprzecznych z powagą funkcji notariusza, jak i zagrażających niezależności zastępcy podczas sprawowania zastępstwa. Proponowanym rozwiązaniem może być wprowadzenie obowiązku złożenia oświadczenia przez emerytowanego notariusza wyznaczonego do sprawowania zastępstwa o zajmowanych stanowiskach, przynależnościach i powiązaniach

[30] Konstytucja Rzeczypospolitej Polskiej z dnia 2 kwietnia 1997 r., Dz.U. Nr 78, poz. 483 z późn. zm.

[31] A. Tkaczyński, *Ograniczenia w podejmowaniu dodatkowych zajęć i zatrudnienia przez notariuszy*, „Acta Universitatis Wratislaviensis" 2013, No 3508, Prawo CCCXV/1, s. 293.

zawodowych. Takie oświadczenie mogłoby podlegać ocenie przez radę właściwej izby notarialnej, która weryfikowałaby, czy dana osoba daje rękojmię bezstronności przy dokonywaniu czynności notarialnych w ramach zastępstwa.

Podsumowanie

Przepisy regulujące funkcjonowanie notariatu zbyt ogólnikowo określają pozycję emerytowanego notariusza w kontekście możliwego sprawowania zastępstwa. Niejednoznaczne jest, czy emerytowany notariusz w ramach zastępstwa podlega takim samym regulacjom prawnym co notariusz. By gwarancja należytego wykonywania czynności notarialnych przez notariusza emerytowanego nie budziła wątpliwości, konieczne są zmiany w regulacji przepisów dotyczących funkcjonowania emerytowanych notariuszy w działalności notarialnej. Samo wykazanie przez ustawodawcę prawa do możliwego wyznaczenia na zastępstwo emerytowanego notariusza nie jest wystarczające z uwagi zawiłość tej funkcji i oddziaływanie na wiele aspektów wiążących się z wymaganym przez pr.not. standardem dokonywania wszelkich czynności notarialnych. Konieczne jest szczególne określenie zasad, na mocy których możliwości przewidziane w przepisach regulujących notariat będą zapewniały odpowiednie ich stosowanie dla zachowania należytego bezpieczeństwa obrotu prawnego, interesu stron postępowania notarialnego oraz samego emerytowanego notariusza sprawującego zastępstwo.

Literatura

Boć W., Leśniak M., *Notariusz jako przedsiębiorca szczególnego rodzaju*, „Acta Universitatis Wratislaviensis" 2011, No 3304, „Przegląd Prawa i Administracji" LXXXIV.
Marquardt P., [w:] *Prawo o notariacie. Komentarz*, red. W. Gonet, wyd. 2, Warszawa 2025.
Marquardt P., *Notariusz jako funkcjonariusz publiczny*, „Krakowski Przegląd Notarialny" rok IV, nr 22.
Nowocień M., *Odmowa dokonania czynności notarialnej*, Warszawa 2018.
Oleszko A., *Prawo o notariacie. Komentarz*, t 1: *Ustrój notariatu*, Warszawa 2016, art. 17.
Oleszko A., *Tryb dokonywania czynności notarialnych jako postępowanie o charakterze cywilnoprawnym*, „Rejent" 2010, nr 7-8.
Oleszko A., *Ustrój polskiego notariatu w świetle orzecznictwa Trybunału Konstytucyjnego*, „Studia Iuridica Lublinensia" 2014, nr 22.
Oleszko A., Pastuszko R., *Sprawczość kompetencyjna zastępcy notarialnego*, „Rejent" 2023, nr 4.
Sawarzyński J., *Podstawy odmowy dokonania czynności notarialnej*, Poznań 2021.
Szereda A.J., *Czynności notarialne. Komentarz do art. 79-112 Prawa o notariacie*, Warszawa 2018.
Szydło M., *Swoboda prowadzenia działalności gospodarczej i swoboda świadczenia usług w prawie Unii Europejskiej*, Toruń 2005.
Tkaczyński A., *Ograniczenia w podejmowaniu dodatkowych zajęć i zatrudnienia przez notariuszy*, „Acta Universitatis Wratislaviensis" 2013, No 3508, „Prawo" CCCXV/1.
Wrzecionek R., *Pojęcie i charakter prawny wyznaczenia zastępstwa przez notariusza – ocena aktualnych regulacji i postulaty de lege ferenda*, „Rejent" 2016, nr 6.

Orzecznictwo

Postanowienie SN z 29.09.2005 r., sygn. SDI 23/05, LEX nr 568824.
Wyrok SN z 27.09.2012 r., sygn. SDI 23/12, LEX nr 1360316.
Wyrok SN z 14.10.2014 r., sygn. SDI 27/14, Biul.PK 2014, nr 10, poz. 30-32.
Wyrok WSA w Warszawie z 7.09.2006 r., sygn. VI SA/Wa 1168/06, LEX nr 255855.
Wyrok WSA w Warszawie z 8.12.2022 r., sygn. VI SA/Wa 775/22, LEX nr 3783147.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Johan Schweigl
Faculty of Law, Masaryk University, Czech Republic
ORCID: 0009-0001-7866-5542
Johan.Schweigl@law.muni.cz

From Internal Standards to the AI Act: Reshaping Financial Processes for the Era of Algorithmic Accountability

Keywords: EU AI Act, Algorithmic Hygiene, Duty of Care, Financial Regulation, Agentic AI

Summary. This paper looks at how internal governance in financial institutions changes as the sector moves from voluntary standards to the mandatory EU AI Act. It traces the path from the traditional "duty of care" to what could be called "algorithmic accountability". The argument is that as Agentic AI takes over cognitive workflows such as AML monitoring and risk management, the traditional focus on human behaviour becomes less relevant. The paper proposes the concept of "Algorithmic Hygiene" as a way of describing the continuous oversight of how software is configured and how it performs. The conclusion is that institutional stability now depends ongoing beyond formal "tick-the-box" compliance toward a real technical audit of an institution's "algorithmic DNA".

Od standardów wewnętrznych do aktu o AI: przekształcanie procesów finansowych na potrzeby ery algorytmicznej odpowiedzialności

Słowa kluczowe: unijny AI Act, higiena algorytmiczna, należyta staranność, regulacja finansowa, Agentic AI.

Streszczenie. Niniejszy artykuł analizuje transformację wewnętrznego ładu instytucji finansowych w okresie przejścia od dobrowolnych standardów do obowiązkowego unijnego AI Act. Przedstawia ewolucję tradycyjnej zasady „należytej staranności" w sformalizowane ramy „odpowiedzialności algorytmicznej". W oparciu o metodę analityczną wzbogaconą o praktyczne obserwacje z sektora bankowego, autor wskazuje na zasadniczą zmianę w obszarze kontroli wewnętrznej: w miarę jak Agentic AI automatyzuje procesy poznawcze, takie jak monitorowanie AML czy zarządzanie ryzykiem, tradycyjny nadzór nad zachowaniem człowieka staje się drugorzędny. Autor wprowadza pojęcie „higieny algorytmicznej" jako proaktywnego procesu skupionego na technicznej integralności i konfiguracji oprogramowania. W konkluzji artykuł stwierdza, że stabilność instytucjonalna zależy obecnie od wyjścia poza formalne podejście „tick-the-box" w kierunku rzeczywistego technicznego audytu „algorytmicznego DNA" instytucji.

1. Introduction: The Role of AI in the Financial Services Ecosystem

Discussions about the integration of Artificial Intelligence (AI) into the financial sector usually focus on technological innovation or on its broad macroeconomic impact. The most important changes, however, take place inside financial institutions themselves, where new technology meets established legal principles and strict regulatory expectations. This paper looks beyond the general development of AI as a tool for efficiency, and focuses on how this shift reshapes internal processes, control mechanisms, and questions of organisational responsibility.

2. Background and Context

In theoretical terms, this paper combines two areas of law: financial law and an emerging field that is now often referred to as AI law. Although the focus is on banking and payment regulation, the underlying logic – how legal systems set rules for autonomous technologies – has a much wider reach. As AI is entering all areas of human activity, AI law will likely develop into an independent legal field, with cross-cutting principles applicable to many sectors. Financial regulation is simply among the first areas where these principles must be tested in practice.

AI has been reshaping the global economy over the last several years, and is sometimes compared to a new "industrial revolution". Recent studies show that AI may transform labour markets and economic production across all sectors. In advanced economies, around 60% of jobs are already exposed to AI in some way, which gives a sense of the scale of the shift[1]. Some studies suggest that, unlike previous waves of technology, the current adoption of Generative AI will hit highly skilled professions where cognitive tasks dominate. By 2030, it may be possible to automate more than 50% of the cognitive skills needed across many jobs. Even managers and experts in finance now face a much higher level of automation than was expected before the arrival of Generative AI[2].

Findings of a study by the McKinsey Global Institute[3] show a major shift in how automation works. Earlier waves of technology mostly replaced physical and

[1] M. Cazzaniga, F. Jaumotte, L. Li, G. Melina, J. A. Panton, C. Pizzinelli, E. Rockall, M. M. Tavares, *Gen-AI: Artificial Intelligence and the Future of Work*, IMF Staff Discussion Note SDN/2024/001, International Monetary Fund, January 2024, p. 2, available at: https://www.imf.org/-/media/files/publications/sdn/2024/english/sdnea2024001.pdf [accessed on: 01.05.2026].

[2] Technologické centrum Praha, *Policy Brief: Dopady generativní umělé inteligence na společnost v ČR*, Policy Brief M2, 2024, available at: https://stratin.tc.cas.cz/vystupy/2024/M2/Policy%20Brief_DopadyGenAI_2024_final.pdf [accessed on: 01.05.2026].

[3] McKinsey Global Institute, *The economic potential of generative AI: The next productivity frontier*, McKinsey & Company, June 2023, available at: https://www.mckinsey.com/capabilities/mck-

repetitive tasks. Generative AI now affects high-wage "knowledge workers" whose roles were previously seen as safe from automation. The financial sector relies heavily on professional expertise and is therefore among the first to face these changes. According to Bloomberg's analysis of the McKinsey data, banks alone could see a productivity boost of around $200 to $340 billion per year, which could mean a 9% to 15% jump in operating profits[4].

The EU AI Act[5] is a major piece of legislation that aims to address the main challenges posed by AI in different areas of human activity. For highly regulated entities, however, the AI Act is not a completely new start. Businesses, and not only financial institutions, are already required to operate under a strict "duty of care" (fiduciary duty). This principle, which is part of the legal orders of the EU Member States, requires financial institutions to manage risks and protect client assets even without specific AI legislation.

The financial sector is relatively unique compared to other industries for several reasons:

– Management of Third-Party Assets: financial firms manage money that belongs to their clients, which creates a higher level of responsibility;
– Systemic Importance: the failure of a major bank can have consequences for the stability of the entire financial system;
– Public Trust: financial services depend on the trust of the public.

Regulated financial institutions had, to some extent, already adapted their internal rules for the use of AI before the AI Act came into force. The new legal requirements therefore mostly formalise and unify what already existed in internal standards, rather than introducing entirely new rules.

3. Problem Statement and Research Goal

The AI Act takes the rules that previously sat under the general principle of "good faith" and turns them into a mandatory "legal minimum". The risk is that compliance becomes a "tick-the-box" exercise, where the principles set by law are reflected in internal processes only on paper, without proper controls and audits. For example,

insey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier [accessed on: 01.05.2026].

[4] B. Baschuk, *Biggest Losers of AI Boom Are Knowledge Workers, McKinsey Says*, Bloomberg News, 14 June 2023, available at: https://www.bloomberg.com/news/articles/2023-06-14/biggest-losers-of-ai-boom-are-knowledge-workers-mckinsey-says [accessed on: 01.05.2026].

[5] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 12 July 2024.

when an employee approves a decision made by an AI system, that person should actually understand how the result was reached, not just sign off.

Another challenge is that financial institutions need to balance the major laws applying to them. They have to follow the AI Act and at the same time respect the GDPR[6] (for privacy) and DORA[7] (for digital security). Balancing these three areas is not easy for any compliance department.

The goal of this paper is to look at how internal controls in financial institutions are changing in the era of AI. The paper first describes how firms managed technology risks in the past under the duty of care. It then sets out the basic principles of the EU AI Act and explains how they create a new mandatory standard for the industry. Finally, it discusses where internal governance is heading, with a focus on what the paper calls "Algorithmic Hygiene" – a state where compliance and audits look directly at the technical health of the algorithms.

4. Methodology and Paper Structure

The paper combines several methods. The analytical method is used to review the requirements of the EU AI Act, DORA and the GDPR, and to identify how these regulations overlap and what specific obligations they impose on financial institutions. The comparative method is used to compare the "pre-regulatory" era based on the duty of care with the new mandatory standards of the AI Act. The case-based approach also plays a part, since the paper draws on the author's experience and observations from the banking and payment industry.

Section 5 covers the "pre-regulatory" era and the duty of care. Section 6 analyses the core principles of the EU AI Act. Section 7 looks at where internal governance is heading, with a focus on "Algorithmic Hygiene" and the new role of internal controls. Section 8 summarises the findings and reflects on supervisory implications.

This paper aims to address three research questions:

RQ1: How does the AI Act change the way financial institutions manage internal risks, compared to the previous era,based on the duty of care?

RQ2: What new internal control mechanisms become necessary when AI agents take over tasks previously performed by humans?

---

[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), OJ L 119, 4 May 2016, p. 1.

[7] Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011, OJ L 333, 27 December 2022, p. 1.

RQ3: Can the traditional Three Lines of Defence model still work in the era of agentic AI, or does it need to be rebuilt?

The research has several limitations. First, the AI Act has only recently been adopted and there is almost no case law that would clarify how the rules will be applied in practice. Second, sectoral guidance such as the BaFin document is national-level soft law, and similar guidance from other Member States may differ. Third, the concept of "Algorithmic Hygiene" introduced here is a normative proposal of the author and has not yet been empirically tested across institutions. The paper is therefore meant to open a discussion rather than to present a finished framework.

5. Governance Before the AI Act: Internal Risk Management Practices

The financial services sector has long been among the leaders in digital transformation.[8] However, the adoption of new technology has not been uniform across the industry. Large banks have moved more slowly because their systems are usually more complex, with more legacy and a more rigid corporate culture. This opened space for fintech companies and neobanks. Many of these new players were originally built as technology firms and used automation as the core of their business model, not just as a side tool.

The shift toward automation was driven by the need for operational efficiency and competitive pressure. Because no specific AI legislation existed at the time, the sector relied on internal standards and guidelines. Internal governance was not about following an AI law but about keeping the business stable and managing risk. Financial institutions understood that internal processes which were left unregulated could cause serious financial and reputational damage, and so they started to integrate technology risks into their existing risk management frameworks.

5.1. Duty of Care and Risk Management

The fundamental legal principle of the "Duty of Care" and its corporate equivalent, "Prudent Management" (known in the Czech legal system as *péče řádného hospodáře*), is rather universal across the legal orders of the EU Member States. Despite some variations in terminology, it is one of the basic pillars on which corporate law stands. Corporate law theory and case law have developed this principle in detail, and the AI Act now translates parts of that established theory into concrete

[8] M. Fernández, N. Charnay, *AI and banking: Leaders will soon pull away from the pack*, S&P Global Ratings, 28 October 2024, available at: https://www.spglobal.com/ratings/en/research/articles/241028-ai-and-banking-leaders-will-soon-pull-away-from-the-pack-13354388 [accessed on: 01.05. 2026].

obligations for the use of AI systems. In the financial sector, the statutory bodies of financial firms have a fiduciary obligation to act in the best interests of their clients and to keep the institution stable[9]. Failing to do so properly, they would face consequences.

In the financial sector, a duty of care is understood as an obligation for providers to avoid conflicts of interest, act with loyalty, and protect the customer's interests[10] .As S&P Global notes, the banking sector has been a "pioneer" of machine learning and deep learning for over a decade[11]. Even during this "pre-regulatory" period, the move from rule-based systems to more autonomous models was guided by the general principle that firms must act honestly and professionally. Any new technology had to be assessed in terms of risk management to make sure it would not cause unreasonable harm or loss to the customer. The internal governance of technology, including AI, was therefore already part of professional conduct standards, long before the current unified regulation. The internal approaches taken by individual institutions, however, varied widely in quality and depth.

5.2. Early Limitations

When trying to gain efficiency through automation, the early adoption of AI in the financial sector faced several barriers. These were not only technical but also legal and ethical, and they made institutions cautious. The first major issue was data privacy and the risk of "data leakage". Under the GDPR, businesses including financial firms have a strict duty to protect their clients' personal data. In the early stages of the AI boom, many institutions were rightly worried that putting sensitive information into external AI models could mean a loss of control over banking secrecy. For a payment institution or a bank, any unauthorised exposure of client data is a red line, with consequences ranging from heavy fines to a loss of public trust.

Another important issue was the so-called "Black Box" effect. Many advanced AI models, in particular those used in banking, are designed to solve complex tasks. As a result, it is hard to understand exactly how they reach a specific result[12].In

[9] O.O. Cherednychenko, C. Busch (eds.), *A Bank's Duty of Care: Perspectives from European and Comparative Law*, European Review of Private Law, 2017, p. 397, available at: https://pure.eur.nl/en/publications/a-banks-duty-of-care-perspectives-from-european-and-comparative-l-2/[accessed on: 01.05.2026].

[10] Financial Services Consumer Panel, *Incorporating a Duty of Care into the Financial Services & Markets Act*, FSCP Position Paper, June 2015, available at: https://www.fca.org.uk/panels-assets/consumer-panel/publication/fscp_position_paper_on_duty_of_care_2015.pdf[accessed on: 01.05. 2026].

[11] M. Fernández, N. Charnay, *op. cit*.

[12] W. Knight, *The Dark Secret at the Heart of AI*, MIT Technology Review, 11 April 2017, available at: https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/ [accessed on: 01.05.2026].

a heavily regulated environment, this lack of transparency is a serious problem. For example, when a model rejects a loan application, the compliance department, internal audit and the regulator all need to know the specific reasons. Without explainability, such models had to be limited to back-office tasks and were rarely used for critical, client-facing decisions.

There were also technical and operational risks, sometimes described as human-machine misalignment. There was a real concern that AI agents could optimise their goals in ways that ignore human values[13]. In addition, older banks often struggled with legacy systems and poor data quality, which made it difficult to integrate modern AI tools. Shneiderman argues that, given these risks, the focus needs to shift from autonomous agents to Human-Centred AI. According to him, the answer lies in designing systems that allow a high level of automation while keeping a high level of human control. For the financial sector, this means treating AI not as a replacement for human judgement but as a tool that supports human expertise and is reliable and safe to use[14].

6. The AI Act as a Legal Minimum and Process Transformation

The entry into force of the AI Act marks a shift for the financial services sector, although it is not a paradigm shift. The earlier era based on the duty of care and on voluntary best practices has now been replaced by mandatory regulatory principles. For highly regulated entities such as banks and payment institutions, the AI Act does not replace existing responsibilities. It formalises them into a strict and unified legal minimum.

Banks now have to make sure that their automated processes follow the core AI Act principles and that they are transparent and traceable. Institutions are moving from a "best effort" regime to enforceable compliance, where every algorithmic decision is subject to oversight. The change covers the full lifecycle of AI systems – from the initial design and the choice of training data, through automated pre-decisions, to ongoing supervision. This creates room for a new standard of integrity, where technical reliability matters as much as financial stability.

6.1. Standardisation of Rules: Core Principles of the AI Act

The aim of the AI legislation framework, as set out in its introductory recitals, is to create a unified EU market for safe and trustworthy AI. The framework should

[13] M. Fernández, N. Charnay, *op. cit*.

[14] B. Shneiderman, *Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy*, arXiv:2002.04087 [cs.HC], 2020, available at: https://doi.org/10.48550/arXiv.2002.04087 [accessed on: 01.05.2026].

support innovation while keeping a high level of protection for fundamental rights, safety, and democratic values. It complements existing Union laws such as the GDPR and consumer protection law by adding specific obligations on transparency, technical documentation, and a human-centred approach. For financial institutions, the AI Act turns internal best practices into a standardised, mandatory legal floor. The following principles describe what firms must now incorporate into their internal control systems.

Risk Management System (Article 9). The Act requires a continuous process of risk management throughout the lifecycle of a high-risk AI system. AI-specific risks therefore need to be considered as part of the standard risk assessment. Risk management is essential for the stability of an institution, and AI risks have to be properly assessed and, where necessary, mitigated. This systemic approach is further developed in sectoral guidance, such as the recent BaFin Guidance on ICT Risks in the Use of AI (2026) in Germany. BaFin states that AI systems must be fully integrated into the existing ICT risk management framework under DORA.[15]

Data Governance and Quality (Article 10). To prevent algorithmic bias, in particular in credit scoring, the Act requires the use of high-quality training, validation and testing data. Article 10 says that these data sets must be relevant, representative, and as free of errors as possible.

Transparency and Provision of Information (Article 13). This principle is the legal answer to the "Black Box" problem described above. High-risk AI systems must be designed in a way that allows users to understand the system's output, or at least the process by which the result was reached. Firms must provide detailed technical documentation and "instructions for use" so that the algorithm is not a mystery to those who operate it[16]. The BaFin guideline mentioned above also addresses this point. It stresses that the management body has a strategic responsibility and must have enough AI literacy to understand how algorithmic decisions affect the firm's risk profile and capital stability[17].

Human Oversight (Article 14). One of the most important principles for the financial sector is the requirement of meaningful human oversight. The aim is to prevent automation bias, where human operators stop questioning the machine. According to Constantino, meaningful human oversight requires that the persons

[15] BaFin (*Bundesanstalt für Finanzdienstleistungsaufsicht), Guidance on ICT Risks in the Use of AI at Financial Entities*, version of 23 January 2026, pp. 9–11, available at: https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Meldung/2025/neu/meldung_2025_12_18_orientierungshilfe_ikt_risiken_en.html [accessed on: 01.05. 2026].

[16] W. Knight, *op. cit*.

[17] BaFin, *op. cit*.

in charge have the necessary competence and authority to intervene, override or even shut the system down when it deviates from its intended path[18].

Accuracy, Robustness, and Cybersecurity (Article 15). High-risk systems must be resilient against errors and against attempts to alter their performance, such as "data poisoning". In the financial context, this principle overlaps with DORA, since AI is increasingly treated as a key part of a firm's ICT systems.

6.2. Data Governance and the Shield of Banking Secrecy

The practical use of AI in financial institutions faces a basic dilemma: AI models need a lot of data to work well, but banks are bound by the strictest standards of banking secrecy and data protection (the GDPR). In the past, this often led to a zero-trust approach, where real client data was almost never used in AI systems and only general or anonymised data was used instead. With the AI Act, in particular Article 10, the focus moves toward a more practical approach. Any system that processes or has access to client data must go through a serious technical and legal due diligence. If the system passes, new options for the use of AI become available. The process should normally include:

– Data Minimisation and Localisation: ensuring that only necessary data is processed and, where possible, kept within the institution's secure infrastructure to prevent data leakage to public models or third-party cloud providers;
– Purpose Limitation: under the GDPR and banking laws, client data can only be used for specific, authorised purposes (*e.g.,* fraud prevention or credit scoring). Using this data to train a general-purpose AI without explicit consent or legal basis would constitute a major breach of requirements on data privacy;
– Vetting of AI Vendors: if an institution uses an external AI provider (*e.g.,* via API), the onboarding must confirm that the vendor does not use the institution's input data to train its own models, as it would constitute a direct violation of banking secrecy.

As pointed out in the BaFin guideline[19], data security in AI is not just about encryption – it is about the integrity of the entire data pipeline.

6.3. Regulatory Interactions

Another aspect connected with the AI Act is the interaction between the AI Act, the GDPR and DORA. The GDPR has been in force for almost a decade, and

[18] J. Constantino, *Exploring Article 14 of the EU AI proposal: accountability challenges of the human in the loop when supervising high-risk AI systems in public administration*, SSRN, 2022, pp. 31, 43, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4254940 [accessed on: 01.05.2026].

[19] BaFin, *op. cit.*

DORA for the last couple of years. These two regulations are two important pillars of financial regulation, although their reach goes well beyond the financial sector. The AI Act will now complete the trio. Although these are often seen as three separate regulatory areas, the current trend, driven by the European Commission's "Digital Omnibus" package[20], is moving toward radical simplification. The goal is to reduce the administrative burden on firms by up to 25% by 2029 by removing overlapping requirements.

These three areas are inextricably intertwined and the process of bringing them together will raise some difficult questions. The trend toward simplification, however, is clear.

7. The Future Perspective: Algorithmic Hygiene and Internal Controls

The rapid progress of AI automation is not just an upgrade of existing tools – it changes the way institutions operate at a fundamental level. As AI systems become more advanced, they are increasingly moving from supporting roles to taking over entire workflows previously handled by human staff. The clearest examples can be found in high-volume areas such as customer service, complaint handling, routine risk management tasks, the updating of contractual terms, back-office administration and parts of client onboarding and AML (Anti-Money Laundering) monitoring.

The wide integration of AI across the banking business leads to a complete change of internal processes. In this new setting, the traditional focus of internal control, namely the supervision of human behaviour and manual data entry, becomes less central in some areas. The key factor for the safety and stability of the institution is now how the AI is configured and how it functions internally. When a human is removed from a process, the "control" must be built directly into the way the algorithm is set up. A new approach to institutional health is therefore needed, which the paper calls Algorithmic Hygiene – a continuous effort to keep the institution's automated processes clean, transparent and compliant.

7.1. The Need for Algorithmic Hygiene

With the arrival of Agentic AI, the focus of internal controls has to shift from supervising human behaviour to direct oversight of the integrity and functioning

[20] European Commission, *Commission proposes Digital Omnibus to simplify EU digital rules*, IP/25/2718, Brussels, November 2025, available at: https://ec.europa.eu/commission/presscorner/ [accessed on: 01.05. 2026].

of the software. We are moving beyond simple automation of isolated tasks toward the deployment of autonomous assistants that can make decisions with a direct impact on clients and on the stability of the institution, although still within the limits set by the human-in-the-loop principle. Algorithmic Hygiene therefore becomes a central element of regulatory compliance.

Algorithmic hygiene is not a one-off audit but a continuous process of keeping "software health". In practice, it will probably focus on:

– System Integration and ICT Interaction: monitoring how the AI agent is embedded within the institution's IT systems and ensuring its interactions with other ICT systems do not create vulnerabilities;
– Output Measurement and Error Statistics: implementing rigorous, real-time evaluation of the agent's performance. Statistical deviation from expected results would serve as the primary red flag for internal control intervention;
– Quality Control of "Pre-fixed" Solutions: given the high costs of custom AI development, financial institutions will largely rely on purchasing pre-fixed, certified solutions tailored to specific use cases. Algorithmic hygiene here acts as a continuous quality assurance mechanism for these external modules.

A new element of internal controls will be the testing of agents through non-standard tasks. Instead of just monitoring routine operations, control departments will likely confront agentic assistants with specific, unusual or edge-case scenarios. The way the agent handles such challenges, and whether it stays within legal and ethical limits, will probably become part of modern controls and audits.

7.2. Redefining the Three Lines of Defence for the Agentic Era

The arrival of Agentic AI does not call for a radical restructuring of internal organisational frameworks, but rather for a targeted update of the Three Lines of Defence (3LoD) model. The focus of control shifts to the software itself: how it is configured, how it is integrated into the institution, and how it performs in real conditions. The evaluation has two key phases: tool onboarding (the pre-production phase) and ongoing control (the production phase), and both phases run across all three lines.

The First Line (Business & IT Operations). Responsibility starts at the selection and onboarding of the AI solution. Since institutions will increasingly rely on pre-fixed external solutions, the first line must check that the configuration of the tool fits the specific needs of the given process, such as AML monitoring or client onboarding. During the operational phase, the focus moves to continuous hygiene: monitoring how the software functions and analysing error statistics. The first line is responsible for making sure that the algorithm's outputs meet the agreed KPIs and do not show dangerous deviations.

The Second Line (Risk & Compliance). The second line carries out an independent assessment at the onboarding stage and verifies that the setup follows both the relevant legislation (the AI Act, the GDPR) and the internal risk appetite. During periodic controls, it uses methods such as the assignment of non-standard tasks. By looking at how the AI agent deals with unusual situations, the second line checks the underlying logic of the software. This line will also benefit from the simplifications brought by the Digital Omnibus, which should allow more efficient reporting and cross-regulatory monitoring.

The Third Line (Internal Audit). Internal audit completes the system by periodically reviewing how well the whole algorithmic hygiene framework works. The audit focus will likely shift from tracking individual human actions to auditing the software as a whole. Internal audit will check whether the onboarding and ongoing control processes in the first and second lines are working properly and whether the reported error statistics are credible. Auditing the integrity of the algorithm and its integration into the institution will likely become the way of confirming that the institution is acting with a duty of care.

8. Conclusion: New Trends in Financial Governance and Supervision

The move toward broad AI automation marks an important turning point in the internal governance of financial institutions. As this paper has shown, the implementation of the AI Act and DORA is not just additional bureaucracy. It is a logical continuation of the traditional "duty of care". In a setting where the human factor is gradually moving out of critical processes, accountability does not disappear. It changes shape and becomes "algorithmic accountability".

Recent legislative developments show a clear trend toward simpler and more harmonised rules. This gives institutions more space for effective risk management and is an important condition for the practical use of "algorithmic hygiene". Instead of being buried in formal administrative tasks, institutions can focus on the actual health and configuration of the software and on how it is set up within their internal processes.

Algorithmic accountability is gradually becoming the standard against which professional financial management will be measured. Institutions that combine technological progress with the integrity of the algorithms they use will be the ones best placed for the future.

8. Policy Implications

For supervisory practice, one risk in particular deserves attention. The human-in-the-loop principle, which is central to Article 14 of the AI Act, can easily turn into a "tick-the-box" exercise. A human formally approves a decision produced by an AI

system without understanding how the system reached it. Supervisory authorities should therefore look not only at whether human oversight exists on paper, but also at whether the persons in charge have the competence and time to challenge the algorithm's output. Without this, human oversight is only a formality.

8.2. Open Questions and Future Research

Several questions remain open and would deserve further work. First, the concept of Algorithmic Hygiene needs to be tested in practice across institutions of different size and type. Second, the relationship between the AI Act and sectoral regulation, such as the CRR and CRD for banks, is not yet fully clear and may lead to overlaps that future case law or supervisory practice will have to deal with. Third, the role of supervisory authorities (the EBA, national regulators) in the technical audit of algorithms is still an open question – it is not yet settled whether supervisors will themselves carry out algorithmic audits or rely on internal and external auditors. These questions go beyond the scope of this paper and should be addressed in follow-up research.

References

Baschuk B., *Biggest Losers of AI Boom Are Knowledge Workers, McKinsey Says*, Bloomberg News, 14 June 2023, https://www.bloomberg.com/news/articles/2023-06-14/biggest-losers-of-ai-boom-are-knowledge-workers-mckinsey-says [accessed on: 01.05.2026].

BaFin (*Bundesanstalt für Finanzdienstleistungsaufsicht*), *Guidance on ICT Risks in the Use of AI at Financial Entities*, version of 23 January 2026, https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Meldung/2025/neu/meldung_2025_12_18_orientierungshilfe_ikt_risiken_en.html [accessed on: 01.05.2026].

Cazzaniga M., Jaumotte F., Li L., Melina G., Panton J. A., Pizzinelli C., Rockall E., Tavares M.M., *Gen-AI: Artificial Intelligence and the Future of Work*, IMF Staff Discussion Note SDN/2024/001, International Monetary Fund, January 2024, https://www.imf.org/-/media/files/publications/sdn/2024/english/sdnea2024001.pdf [accessed on: 01.05.2026].

Cherednychenko O.O., Busch C. (eds.), *A Bank's Duty of Care: Perspectives from European and Comparative Law*, European Review of Private Law, 2017, https://pure.eur.nl/en/publications/a-banks-duty-of-care-perspectives-from-european-and-comparative-l-2/ [accessed on: 01.05.2026].

Constantino J., *Exploring Article 14 of the EU AI proposal: accountability challenges of the human in the loop when supervising high-risk AI systems in public administration*, SSRN, 2022, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4254940 [accessed on: 01.05.2026].

European Commission, *Commission proposes Digital Omnibus to simplify EU digital rules*, IP/25/2718, Brussels, November 2025.

Fernández M., Charnay N., *AI and banking: Leaders will soon pull away from the pack*, S&P Global Ratings, 28 October 2024, https://www.spglobal.com/ratings/en/research/articles/241028-ai-and-banking-leaders-will-soon-pull-away-from-the-pack-13354388 [accessed on: 01.05.2026].

Financial Services Consumer Panel, *Incorporating a Duty of Care into the Financial Services & Markets Act*, FSCP Position Paper, June 2015, https://www.fca.org.uk/panels-assets/consumer-panel/publication/fscp_position_paper_on_duty_of_care_2015.pdf [accessed on: 01.05.2026].

Knight W., *The Dark Secret at the Heart of AI*, MIT Technology Review, 11 April 2017, https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/ [accessed on: 01.05.2026].
McKinsey Global Institute, *The economic potential of generative AI: The next productivity frontier*, McKinsey & Company, June 2023, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier [accessed on: 01.05.2026].
Shneiderman B., *Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy*, arXiv:2002.04087 [cs.HC], 2020, https://doi.org/10.48550/arXiv.2002.04087.
Technologické centrum Praha, *Policy Brief: Dopady generativní umělé inteligence na společnost v ČR*, Policy Brief M2, 2024, https://stratin.tc.cas.cz/vystupy/2024/M2/Policy%20Brief_DopadyGenAI_2024_final.pdf [accessed on: 01.05.2026].

Legal sources

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), OJ L 119, 4 May 2016.
Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector (DORA), OJ L 333, 27 December 2022.
Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act), OJ L, 12 July 2024.

AI Use Disclosure

The author used an artificial intelligence language model to assist with linguistic editing, translation refinement, and the structural organisation of this paper. While the AI was employed to enhance the clarity and stylistic quality of the text, all conceptual ideas, legal analyses, original terminology (including the concept of "Algorithmic Hygiene"), and final conclusions remain the sole responsibility of the author.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Patricia Sharp
Uniwersytet Zielonogórski
ORCID 0009-0007-9547-563X
20000967@stud.uz.zgora.pl

AI Advancements and GDPR Obligations Regarding the Content of Personal Data Processing Notices: Challenges for Public Administration in Poland

Keywords: personal data processing notice, GDPR transparency, artificial intelligence, public administration, automated decision-making

Summary. The integration of artificial intelligence systems into public administration in Poland raises significant compliance challenges under the EU General Data Protection Regulation (GDPR). This article examines the content requirements for personal data processing notices, the primary transparency instrument under GDPR Articles 13 and 14, as they apply to AI systems deployed by Polish public bodies. The article argues that AI systems create genuine and previously underappreciated challenges for fulfilling transparency obligations, particularly the requirement to provide "meaningful information about the logic involved" under Article 13(2)(f) of the GDPR. The analysis identifies three temporally and functionally distinct disclosure obligations arising from the intersection of GDPR, the EU AI Act, and Poland's Administrative Procedure Code, and demonstrates that processing notices under GDPR Article 13 occupy an irreplaceable ex ante role that cannot be substituted by procedural transparency requirements.

Rozwój AI a obowiązki RODO w zakresie treści klauzul informacyjnych: wyzwania dla administracji publicznej w Polsce

Słowa kluczowe: klauzula informacyjna o przetwarzaniu danych osobowych, przejrzystość RODO, sztuczna inteligencja, administracja publiczna, zautomatyzowane podejmowanie decyzji

Streszczenie. Integracja systemów sztucznej inteligencji z administracją publiczną w Polsce generuje istotne wyzwania w zakresie zapewnienia zgodności z unijnym Ogólnym Rozporządzeniem o Ochronie Danych (RODO). Artykuł podejmuje analizę wymogów treściowych klauzul informacyjnych dotyczących przetwarzania danych osobowych, stanowiących podstawowy instrument realizacji zasady przejrzystości wynikającej z art. 13 i 14 RODO, w kontekście systemów AI wykorzystywanych przez polskie organy władzy publicznej. W artykule wykazano, że zastosowanie systemów sztucznej inteligencji prowadzi do powstania rzeczywistych oraz dotychczas niedostatecznie rozpoznanych trudności w zakresie realizacji obowiązków informacyjnych, w szczególności w odniesieniu do wymogu przekazania „istotnych informacji o zasadach podejmowania decyzji", o którym mowa w art. 13 ust. 2 lit. f RODO. Analiza wskazuje, że specyfika systemów AI, w tym ich złożoność techniczna oraz ograniczona przejrzystość modeli decyzyjnych, utrudnia spełnienie standardu informacyjnego przewidzianego w przepisach o ochronie danych osobowych. W artykule zidentyfikowano trzy odrębne obowiązki ujawnieniowe wynikające z interakcji między RODO, unijnym Aktem w sprawie

sztucznej inteligencji (AI Act) oraz przepisami polskiego Kodeksu postępowania administracyjnego. Wykazano, że klauzule informacyjne sporządzane na podstawie art. 13 RODO pełnią autonomiczną i niezastępowalną funkcję *ex ante,* której nie mogą zastąpić późniejsze mechanizmy transparentności o charakterze proceduralnym.

Introduction

A personal data processing notice is a document or notification provided to a person whose personal data are being processed by a controller. The provision of such notices is mandated by the EU's General Data Protection Regulation (GDPR). The GDPR imposes an information obligation to provide the data subject with specified information concerning the processing of their data. This obligation, implemented through the provision of a data processing notice containing information relevant to the collection of personal data and their subsequent processing, constitutes one of the principal safeguards of an individual's informational self-determination and serves to enable the effective exercise of the rights conferred on the data subject in relation to the processing of personal data. Personal data processing notices are expected to be clear to understand for the recipient.

The drive for the integration of artificial intelligence (AI) systems into public administration in Poland has been gaining momentum, with official support from Polish public bodies. As stated in the Policy for the Development of Artificial Intelligence in Poland since 2020[1], AI and the public sector is one of the six main areas of interest. Actions are aimed at supporting the public sector in carrying out AI-related procurements, coordination enhancement, and the further development of programs such as GovTech Poland.

The Policy explicitly states:

> AI-based solutions can improve the efficiency of central government and local government bodies. The constantly improving technical capabilities makes process automation a more and more attractive avenue pursued by public administration. Thanks to advances in artificial intelligence, processes that had to be done by numerous clerks just a couple of years ago can be partially automated these days.

An example of how Poland is advancing AI in public administration is the recent launching of models like PLLuM (Polish Large Language Model)[2] for automating tasks, citizen communication, and reducing bureaucracy. The first pilot

[1] Policy for the Development of Artificial Intelligence in Poland from 2020, Appendix to the Resolution No. 196 of the Council of Ministers of 28 December 2020 (item 23), available at: https://www.gov.pl/attachment/928200fa-b1a6-4c0c-b3a8-d1fbf1e1175a [accessed on: 09.02.2026].

[2] Ministerstwo Cyfryzacji, *PLLuM — polski model językowy, jak działa i do czego może się przydać*, available at: https://www.gov.pl/web/cyfryzacja/pllum-polski-model-jezykowy-jak-dziala-i-do-czego-moze-sie-przydac [accessed on: 09.02.2026].

project was carried out in the city of Częstochowa[3], in southern Poland. Developed by Polish research institutions, working on behalf of the Ministry of Digital Affairs, it is capable of understanding Polish grammar, socio-cultural context, and official administrative language.

HIVE AI is a direct continuation of the *Polish Large Language Model (PLLuM)* project[4]. The PLLuM was a government-backed AI project started in 2023 with an aim to produce a language model that could more effectively process and produce texts in the Polish language. The follow-on initiative, HIVE AI aims to facilitate the integration of Polish language models into the mObywatel app (Poland's official digital ID and documentation phone app), develop a clerical assistant for the Polish Ministry of Digital Affairs, and conduct a pilot implementation of LLMs in a municipal or provincial office[5].

Poland facilitates public access to government services through a project known as eGovernment, the HIVE AI project strongly aligns with its goals. An implement consulting group study commissioned by Google in partnership with SGH Warsaw School of Economics found that: "Generative AI has significant potential to enhance productivity in public administration, creating 7% more value in money, equivalent to an annual contribution of PLN 8 billion"[6].

However, the rapid integration of artificial intelligence systems into public administration in Poland (as well as across the European Union) also represents a challenge to established data protection principles and to public bodies who must still comply with the GDPR.

This challenge is further strenuous in Poland by the absence of detailed supervisory authority guidance comparable to that available in some other EU Member States. While France's CNIL (the French national data protection enforcement agency) has published comprehensive, AI-specific guidance on processing notice content, Poland's President of the Personal Data Protection Office (PUODO) has

[3] Ministerstwo Cyfryzacji, Polski model językowy PLLuM trafi do samorządów, available at: https://www.gov.pl/web/cyfryzacja/polski-model-jezykowy-pllum-trafi-do-samorzadu [accessed on: 09.02.2026].

[4] National Information Processing Institute — National Research Institute (OPI PIB), HIVE AI: Development and pilot implementation of large language models in Polish public administration, available at: https://opi.org.pl/en/project/hive-ai-rozwoj-i-pilotazowe-wdrozenie-duzych-modeli-jezykowych-w-polskiej-administracji-publicznej/ [accessed on: 09.02.2026].

[5] NASK, *About HIVE AI Project*, available at: https://www.nask.pl/en/institute-for-institutions/hive-ai/about-project [accessed on: 09.02.2026].

[6] Implement Consulting Group, *The AI opportunity for eGovernment in Poland* (2025), commissioned by Google in partnership with SGH Warsaw School of Economics, available at: https://cms.implementconsultinggroup.com/media/uploads/articles/2025/The-AI-opportunity-for-eGovernment-in-Poland/2025-The-AI-opportunity-for-eGovernment-in-Poland-English.pdf [accessed on: 09.02.2026].

not yet delivered equivalent resources, leaving public bodies deploying systems like PLLuM and HIVE AI without clear implementation standards.

It is obvious that the GDPR was designed for and adopted in an extensively different technological reality than today's world, in which there is an ever-increasing reliance on language models. It provides that the protection of natural persons in relation to the processing of personal data is a fundamental right, one that should not be diminished by the omnipresence of AI.

Any information referred to in Articles 13 and 14 of the GDPR and any communication under Articles 15 to 22 and 34 of the GDPR relating to personal data processing must be provided in a concise, transparent, intelligible and easily accessible form, using clear and plain language. Public bodies using AI systems are therefore challenged with the task of providing a notice that contains "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing" (GDPR Article 13(2)(f)) about systems that are not fully understood even by their creators[7].

Public bodies are also faced with dealing with a population with varying levels of digital literacy[8], and the need to comply not only with GDPR but also the AI Act[9] and Poland's Administrative Procedure Code (*Kodeks postępowania administracyjnego*), which imposes its own transparency and justification requirements.

In this context, the article provides an analysis that examines whether processing notices used by public administration are subject to change with a further emergence and use of AI.

The article further addresses the following interconnected questions:

– What specific content must processing notices contain when AI systems are used by Polish public administration?
– Why do modern AI systems create particular difficulties for meeting content requirements put in by the GDPR?
– What practical frameworks or solutions can enable Polish public bodies to achieve compliance?

This article is based on doctrinal legal analysis, technical assessment, and policy evaluation. The analysis was performed in the followings stages: textual interpreta-

[7] T. Constantino, "AI's Biggest Secret — Creators Don' Understand It'", *Forbes*, 8 May 2025, available at: https://www.forbes.com/sites/torconstantino/2025/05/08/ais-biggest-secret---creators-dont-understand-it-experts-split/ [accessed on: 09.02. 2026].

[8] 44% of citizens hold at least basic digital skills. European Commission, Digital Skills & Jobs Platform, *Poland digital skills map*, available at: https://digital-skills-jobs.europa.eu/en/european-interactive-map/poland [accessed on: 09.02. 2026].

[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act), OJ L 2024/1689, 12.7.2024, entered into force 1 August 2024.

tion of the GDPR and AI Act provisions, Court of Justice of the European Union (CJEU) case law, European Data Protection Board guidelines, and the Polish Personal Data Protection Office guidance documents.

The analysis concentrates on public administration in Poland. The paper does not address private sector applications of AI, except where they provide comparative insights. The legal analysis within focuses on data protection, AI regulation, and administrative procedure; it does not extensively address related areas such as substantive administrative law.

This paper specifically addresses the content requirements of processing notices under GDPR Articles 13 and 14 as they apply to AI systems, with particular emphasis on the enhanced obligations for automated decision-making and profiling under Article 13(2)(f).

The research also incorporates a comparative analysis of transparency approaches in the EU and examination of enforcement actions by data protection authorities to understand how content requirements are interpreted in practice.

1. Legal Framework for Processing Notice Content

The GDPR establishes transparency as a fundamental principle of data protection. Article 5(1)(a) requires that personal data be "processed lawfully, fairly and in a transparent manner in relation to the data subject". This principle is not merely aspirational; it creates enforceable obligations implemented through specific requirements in Articles 12-14.

The transparency principle serves multiple functions: it enables data subjects to understand and verify the lawfulness of processing. Article 12(1) GDPR establishes how transparency obligations must be fulfilled, requiring that information be provided in a manner that is "concise, transparent, intelligible and easily accessible". The information must use "clear and plain language", which rules out purely technical descriptions accessible only to specialists, and must be provided "in writing, or by other means, including, where appropriate, by electronic means".

Article 12(1) does not specify a maximum length for processing notices, recognizing that the amount of information needed depends on the complexity of the processing. However, the requirement for conciseness means controllers cannot simply provide exhaustive technical documentation and claim compliance. The Article 29 Working Party (now the European Data Protection Board) has endorsed a "layered approach" where brief, essential information is provided at the point of

data collection, with more detailed information available through links or separate documents[10].

Articles 13 and 14 specify what information must be provided to data subjects, with Article 13 applying when data is collected directly from the individual and Article 14 applying when data is obtained from other sources. Both articles require disclosure of: the identity and contact details of the controller; the contact details of the data protection officer (where applicable); the purposes of processing; the legal basis for processing; the categories of personal data concerned (Article 14 only); the recipients or categories of recipients; information about international transfers if applicable; the retention period or criteria for determining it; the existence of data subject rights; the right to lodge a complaint with the supervisory authority; and where data provision is a statutory or contractual requirement, whether the individual is obliged to provide it and the consequences of failing to do so.

2. Understanding the Phrase "At Least in Those Cases"

Article 13(2)(f) requires controllers to provide data subjects with information about:

> the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, ***at least in those cases***, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".

Article 14(2)(g) contains identical language for situations where data is not obtained directly from the data subject.

In the GDPR, the phrase **at least in those cases** appears primarily in the context of transparency requirements regarding automated decision-making and profiling (Articles 13, 14, and 15) and has been subject to debate.

A narrow reading would limit the enhanced disclosure obligation to situations covered by Article 22, that is solely automated decisions that produce legal effects or similarly significantly affect the data subject.

A broader reading would extend it to all automated decision-making and profiling, regardless of whether Article 22 applies.

The Article 29 Working Party adopted the broad interpretation in its Guidelines on Automated Individual Decision-Making and Profiling, stating that: "informa-

[10] "In the digital context, in light of the volume of information which is required to be provided to the data subject, a layered approach may be followed by data controllers where they opt to use a combination of methods to ensure transparency. WP29 recommends in particular that layered privacy statements/ notices should be used to link to the various categories of information which must be provided to the data subject, rather than displaying all such information in a single notice on the screen, in order to avoid information fatigue". Article 29 Working Party, *Guidelines on transparency under Regulation 2016/679*, WP260 rev.01, adopted 29 November 2017, last revised 11 April 2018, p. 19.

tion must be provided whenever there is automated decision-making (including profiling)[11]", not only when Article 22 applies.

The CJEU's February 2025 ruling in Case C-203/22[12] also provides support for the broad interpretation. While the case concerned Article 22 rights specifically, the Court emphasized that the GDPR's transparency requirements exist to enable data subjects to "exercise their rights effectively"[13].

For Polish public administration, this means processing notices must contain the enhanced information whenever AI systems are used in decision-making, even if human oversight means Article 22 does not technically apply.

3. Specific Challenges for AI Systems

Article 13(2)(f) provides for the disclosure of the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

The phrase "meaningful information" is a significant challenge for AI systems. The term in question is not used anywhere else in the GDPR. Recital 71 GDPR also provides only minimal guidelines that the right to information should not "adversely affect the rights and freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software". However, Recital 71 also emphasizes that this limitation should not "result in the data subject being refused all information".

The phrase "at least in those cases, meaningful information about the logic involved", referring to "automated decision-making, including profiling", may, on the one hand, support the view that the information obligation encompasses both the "meaningful principles" of automated decision-making based on profiling and the "meaningful principles" of profiling itself, even where it does not lead to automated decision-making.

---

[11] Article 12, recital 58, Articles 13 and 14, and Article 22(1) of the GDPR. In this regard, please also see the European Data Protection Board, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, WP251 rev.01, available at: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053 [accessed on: 09.02.2026], p. 17.

[12] Court of Justice of the European Union, judgment of 27 February 2025, Case C-203/22, EU:C:2025:117, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX-:62022CJ0203 [accessed: on 09.02.2026].

[13] "In particular, in the specific context of the adoption of a decision based solely on automated processing, the main purpose of the data subject's right to obtain the information provided for in Article 15(1)(h) of the GDPR is to enable him or her effectively to exercise the rights conferred on him or her by Article 22(3) of that regulation, namely the right to express his or her point of view on that decision and to contest it", *Ibid.*, 45–50 (right to explanation under Art. 15(1)(h) GDPR).

On the other hand, Recital 73 GDPR provides that Union or Member State law may introduce restrictions on certain rights and obligations, including those relating to decisions based on profiling, where this is necessary and proportionate in a democratic society to safeguard public security, including the protection of human life, especially in response to natural or manmade disasters, the prevention, investigation and prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security. In this respect, the EU legislator refers not to profiling as such, but to decisions based on profiling.

Accordingly, it should be concluded that the object of protection envisaged by the EU legislator in the GDPR is not profiling per se but profiling that results in a decision-making process. Otherwise, profiling constitutes merely one method of processing personal data. A similar conclusion follows from Recital 24 of the GDPR, which addresses profiling while again emphasizing that it concerns profiling "in order to take decisions concerning the data subject or for analyzing or predicting personal preferences, behaviors and attitudes." Importantly, the information provided on this basis should include the specific procedures and rules applied to the automated use of personal data relating to the individual concerned in order to produce a particular outcome (as held by the Court of Justice of the European Union in its judgment of 27 February 2025, C-203/22, EU:C:2025:117), rather than generic descriptions[14]. The Court in Case C-203/22 stated:

> Those requirements cannot be satisfied either by the mere communication of a complex mathematical formula, such as an algorithm, or by the detailed description of all the steps in automated decision-making, since none of those would constitute a sufficiently concise and intelligible explanation.

Currently, the notion of the "meaningful principles" of automated decision-making based on profiling and of profiling itself is interpreted in light of the purpose of the information obligation. That purpose is to ensure transparency with regard to all circumstances surrounding profiling (see X. Konarski)[15]. From this perspective, informing the data subject about the specific computer program used for profiling

[14] "Thus, the 'meaningful information about the logic involved' in automated decision-making, within the meaning of Article 15(1)(h) of the GDPR, must describe the procedure and principles actually applied in such a way that the data subject can understand which of his or her personal data have been used in what way in the automated decision-making at issue, with the complexity of the operations to be carried out in the context of automated decision-making not being capable of relieving the controller of the duty to provide an explanation" *Ibid*., para. 62 (specific procedures, not algorithms).

[15] P. Barta, M. Kawecki, P. Litwiński [in:] P. Litwiński (ed.), *Ustawa o ochronie danych osobowych. Komentarz* [in:] *Ogólne rozporządzenie o ochronie danych osobowych. Ustawa o ochronie danych osobowych. Wybrane przepisy sektorowe. Komentarz*, edn. 2, Warszawa 2025, art. 13, Legalis.

does not appear to be of material relevance. Rather, the information should be structured so as to explain how profiling operates and to indicate its consequences for the data subject, thereby enabling the individual to challenge an automated decision (see W. Chomiczewski)[16].

These traditional views made considerable sense for conventional profiling systems. If a controller uses a standard database query or statistical regression implemented in common software, the choice of software is largely irrelevant—the same logic could be implemented in different programs with identical results An insured person who knows that their retirement benefit is calculated according to a statutory formula does not gain meaningful additional information from knowing whether the calculation is performed in Excel, within the pension office's internal system, or in any other software.

However, AI systems challenge this distinction between program and logic. For AI systems, the specific program often is the logic in a way that was not true for traditional profiling. The difference in explainability is inherent to the model architecture it is not just about implementation details.

For traditional profiling using explicit rules ("if income < threshold AND age > 65, then..."), the logic is transparent regardless of the software implementing it. For AI systems that learn from data, the specific model embeds assumptions and patterns from its training data that may not be visible from general descriptions[17]. Differences are embedded in specific programs and fundamentally affect their behavior[18]. A data subject cannot meaningfully challenge a decision without understanding which model was used and how it was trained.

The EU AI Act creates a regulatory framework where the specific AI system's classification determines legal obligations:

– High-risk systems have extensive requirements (conformity assessment, registration, human oversight);
– Non-high-risk (minimal) systems have minimal requirements;
– General-purpose AI systems have different obligations.

A data subject cannot verify whether appropriate safeguards apply without knowing which specific AI system was used. If a public body claims to use a high-risk system subject to rigorous oversight, but actually uses a simpler non-regulated tool, this is material information that affects rights and remedies.

---

[16] *Ibid*.

[17] J. Burrell "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms" (September 15, 2015). Available at: SSRN: https://ssrn.com/abstract=2660674 or http://dx.doi.org/10.2139/ssrn.2660674, DOI: 10.1177/2053951715622512 [accessed on: 09.02.2026].

[18] M.H. Jarrahi, V. Glaser, "Not All AI Systems Are Created Equal: A Typology of AI Systems' Performance and Affordance", *Proceedings of HICSS 2025*, DOI: 10.24251/HICSS.2025.655.

The EU AI Act, which entered into force on August 1, 2024, creates a risk-based regulatory framework for AI systems. Annex III of the Act lists "high-risk AI systems", many of which are commonly deployed in public administration. These include AI systems used for: evaluating eligibility for public assistance benefits and services; assessing the creditworthiness of natural persons or establishing their credit score (when used by public authorities); dispatching or establishing priority in emergency services; evaluating students" performance in educational institutions; and making decisions in relation to migration, asylum, and border control management.

The classification of an AI system as "high-risk" triggers extensive regulatory obligations, including transparency requirements that affect processing notice content. Importantly, the AI Act's definition of "deployer" (Article 3(4)) encompasses public authorities that use AI systems for administrative purposes even when they did not develop the system themselves. This means Polish public bodies using commercial or centrally developed AI systems are "deployers" that are subject to AI Act obligations.

Article 13 of the AI Act requires providers of high-risk AI systems to provide deployers with instructions for use that include: the identity and contact details of the provider; the characteristics, capabilities, and limitations of the AI system's performance; the level of accuracy, robustness, and cybersecurity of the system; known or foreseeable circumstances in which the system may not perform as intended; specifications for input data quality and formatting; the expected lifetime of the system and necessary maintenance; and descriptions of the human oversight measures built into the system.

For processing notice content, this creates an indirect but important obligation. Public bodies – as deployers cannot provide meaningful information about the logic involved in automated processing if they lack detailed information about how the AI system functions.

Article 13 of the AI Act ensures that deployers receive this information from providers, which they can then include in appropriate disclosure for data subjects. However, the information providers supply to deployers is typically technical and detailed, precisely the kind of information that C-203/22 ruled cannot simply be passed on to data subjects. Deployers must therefore engage in a translation exercise, converting provider documentation into explanations accessible to ordinary citizens.

Poland's Administrative Procedure Code (KPA) further creates transparency and justification requirements that, while related to GDPR processing notice obligations, serve fundamentally different functions – a distinction that becomes more significant when AI systems are deployed.

The processing notice under GDPR Art. 13 operates independently of and prior to administrative proceedings. Its primary function is not procedural transparency (enabling participation in a specific case) but enabling data subjects to understand how their personal data will be processed and to exercise their data protection rights before processing begins or while it is ongoing[19]. This distinction was preserved when Poland adapted the KPA to GDPR. Art. 2a KPA, which addresses the procedural modalities of providing GDPR Art. 13 information within administrative contexts—when, how, and in what form the notice must be given—but does not alter the substantive content requirements of Art. 13, which derive directly from the GDPR itself[20].

Polish legal commentary recognizes that the information obligation under Art. 2a KPA is distinct from regulated procedural obligations because it protects a different interest: personal data rights rather than procedural participation rights[21]. The notice serves data subjects regardless of whether they become parties to administrative proceedings, and it must be provided at the point of data collection, not merely when a decision is rendered.

For AI systems, this temporal and functional separation has critical implications. A processing notice may be the primary—or only—mechanism through which data subjects learn that AI is being used and how it functions:

– Not all data subjects become parties to proceedings. A person whose data is processed for profiling or risk assessment may never receive an individual administrative decision triggering KPA Art. 107 justification requirements. Art. 13 GDPR information is supposed to enable [data subjects] to make informed decisions about the processing of their data, making the processing notice the primary transparency mechanism for those who never enter formal proceedings;
– The notice enables ex ante exercise of rights that cannot be exercised ex post. Data subjects who understand from the processing notice that AI profiling will occur can exercise rights such as objection (GDPR Art. 21) (in the case of a public authority, this may only concern processing based on Article 6(1)(e) GDPR)[22] before the AI processes their data. Once a decision is made, these preventive rights are moot; only reactive rights (appeal, complaint) remain;

---

[19] B. Adamiak [in:] B. Adamiak, J. Borkowski, *Kodeks postępowania administracyjnego. Komentarz*, edn. 19, Warszawa 2024, art. 2a, Legalis.

[20] *Ibid.*

[21] P. Litwiński [in:] P. Litwiński (ed.), *Kodeks postępowania administracyjnego. Komentarz* [in:] *Ogólne rozporządzenie o ochronie danych osobowych. Ustawa o ochronie danych osobowych. Wybrane przepisy sektorowe. Komentarz*, edn. 2, Warszawa 2025, art. 2a, Legalis.

[22] P. Gołaszewski, K. Wąsowski [in:] R. Hauser, M. Wierzbowski (eds.), *Kodeks postępowania administracyjnego. Komentarz*, edn. 8, Warszawa 2023, art. 2a, Legalis.

– The notice reaches broader audiences than individual justifications. A processing notice published on an agency website or provided at first contact informs all potential data subjects about AI use. Individual decision justifications under KPA Art. 107 inform only those who receive adverse decisions—a much narrower group. The notice thus serves a systemic transparency function distinct from case-specific justification;
– The notice must explain AI logic while it remains abstract. At the point of data collection (when Art. 13 applies), the public body must explain how the AI system functions in general, not merely how it processed a specific person's data. This is actually more demanding than ex post justification because it requires explaining the system's methodology without reference to concrete facts.

The KPA's procedural transparency requirements (Arts. 10, 107) complement but do not substitute for processing notices. Article 10's right to express views on evidence presupposes that the party already knows. Article 107's justification is to explain how the AI analyzed the specific party's data but cannot inform those never subject to individual decisions. The processing notice thus occupies a distinct and irreplaceable position in the transparency framework.

This functional independence becomes even more significant when considering the additional requirements resulting from the AI Act.

Article 13 of the AI Act requires providers to furnish deployers (which public bodies would be considered to be) with detailed documentation about high-risk AI systems. Article 86 of the AI Act grants data subjects a right to obtain explanations of individual decisions.

These requirements create three temporally and functionally distinct disclosure moments:

1. Ex ante general disclosure (GDPR Art. 13 processing notice): Before/at data collection, explaining the AI system in abstract;
2. During-proceeding disclosure (KPA Art. 10): During the administrative process, enabling party participation;
3. Ex post individual disclosure (KPA Art. 107 + AI Act Art. 86): After decision, explaining how AI processed specific data.

Each serves a distinct purpose; none can substitute for the others. The processing notice's role is therefore not diminished but rather elevated in the context of AI: it becomes the foundational disclosure upon which subsequent procedural transparency depends.

The intersection creates potentially reinforcing obligations: GDPR Art. 13(2)(f) requires ex ante disclosure of AI logic in processing notices, while KPA Art. 107 requires ex post justification of how AI contributed to the specific decision, and

Art. 10 may require interim disclosure enabling party participation. Together, these create a higher transparency standard than either regime alone would impose.

The challenge of explaining "the logic involved" is compounded when AI systems are trained on large, diverse datasets. Data protection authorities have begun providing practical guidance on how controllers can satisfy Article 13(2)(f) disclosure requirements without overwhelming data subjects with unusable detail.

The French data protection authority *Commission Nationale de l'Informatique et des Libertés* (CNIL) has issued specific guidance addressing this challenge in the context of AI systems. The CNIL recognizes that different levels of specificity are appropriate depending on the nature and scale of data sources[23]:

This guidance is directly relevant to Polish public bodies deploying AI systems like PLLuM or HIVE AI. These large language models are trained on an extensive sample of Polish-language that may include personal data from diverse online sources.

The CNIL has issued specific guidance addressing how to inform data subjects about AI training data sources[24]. The CNIL distinguishes between two scenarios:

Individual information (when controllers can identify and contact data subjects): Controllers must provide all Article 13/14 GDPR information, including identity and contact details, purpose and legal basis, recipients, retention period, and data subject rights. For indirect collection, controllers must additionally specify categories of personal data and "a precise indication of the source(s) of the data (including whether or not they are publicly available sources)"[25].

General information notice (when individual contact is impossible or disproportionate): Controllers must publish an information notice (*e.g.,* on a website) that includes as much information as would be provided individually, plus a statement that the controller cannot identify individuals to respond to right requests.[26]

Within these frameworks, different levels of specificity apply based on data sources:

– **Limited sources:** Precise information required, including dataset names and hyperlinks. Example: "This training data comes from a publicly accessible dataset (hyperlink) containing 70,000 images...published on [social network] between 2020 and 2021[27]";

---

[23] CNIL (*Commission nationale de l'informatique et des libertés*), *AI system development: CNIL's recommendations to comply with the GDPR*, 5 January 2026, available at: https://www.cnil.fr/en/ai-system-development-cnils-recommendations-to-comply-gdpr [accessed on: 09.02.2026].

[24] CNIL, *Informing data subjects*, available at: https://www.cnil.fr/en/informing-data-subjects [accessed on: 09.02.2026].

[25] *Ibid.*

[26] *Ibid.*

[27] *Ibid.*

- **Numerous sources:** "Overall information, such as source categories, or even the names of a few main or typical sources, is generally sufficient" Example: "The data was collected by scraping websites specialized in the field studied...[28]";
- **General-purpose AI models:** Article 53 AI Act requires "a sufficiently detailed summary" listing main datasets and explaining other data sources used[29].

The CNIL guidance thus operationalizes the principle established in C-203/22: explanations must enable data subjects to "understand the envisaged method and the main characteristics of the processing"[1] without requiring disclosure of every technical detail.

For AI training data, this means explaining *what kinds* of data sources were used, *what types* of personal data they contained, and *how* data subjects can learn more or exercise their rights—not exhaustively cataloguing every data point.

Conclusion

This paper has examined the content requirements for personal data processing notices when AI systems are deployed in Polish public administration, analyzing the intersection of GDPR transparency obligations, AI Act requirements, and Polish legal frameworks.

First, the existing legal framework creates extensive and specific content obligations. GDPR Articles 13 and 14 require baseline disclosures applicable to all processing, while Article 13(2)(f) mandates enhanced content when automated decision-making or profiling occurs: meaningful information about the logic involved, the significance of the processing, and the envisaged consequences. The EU AI Act adds further obligations, particularly Article 86's right to explanation of individual decisions and Article 13's requirement that providers furnish deployers with detailed system information. Polish law, including the Administrative Procedure Code, creates additional transparency expectations rooted in administrative law principles.

Second, modern AI systems create genuine technical challenges for meeting these content requirements. While simple, interpretable models can be explained straightforwardly, many AI systems that deliver superior accuracy rely on complex architectures that resist transparent explanation. The accuracy-interpretability trade-off means that choosing models that are easy to explain may sacrifice performance, potentially providing inferior service to citizens. This creates a fundamental tension: the law requires explanation of "the logic involved" but for some AI sys-

[28] *Ibid.*
[29] *Ibid.*

tems, the true logic cannot be explained in terms comprehensible to non-specialists without oversimplification that borders on misrepresentation.

Third, Polish public administration faces distinctive implementation challenges. These include resource constraints that limit investment in sophisticated AI systems with built-in interpretability; legacy IT infrastructure that may not integrate well with modern transparency tools; diverse populations with varying levels of digital literacy, requiring multiple explanation formats and languages; and coordination challenges among multiple supervisory authorities (PUODO for data protection, the proposed AI Commission, and sectoral regulators). The intersection of GDPR, AI Act, and KPA requirements creates particular complexity: processing notices must satisfy data protection law while also fulfilling administrative law expectations for justification and transparency.

PUODO should develop detailed guidance specifically addressing processing notice content for AI systems. This guidance should include: concrete examples of adequate and inadequate disclosures for common AI use cases in public administration; templates that public bodies can adapt to their specific circumstances; clear articulation of when simplified explanations are acceptable and when more detailed technical information must be provided; guidance on the layered transparency approach, including what must be in each layer; and clarification of how PUODO will assess compliance given the technical limitations of certain AI systems.

Public bodies need training in both data protection law and AI technology, enabling data protection officers and technical staff to collaborate effectively in drafting processing notices that are legally compliant and technically accurate.

The deployment of PLLuM and HIVE AI demonstrates commitment to AI-enabled public administration. Equally, detailed commitment to transparency is now required. Processing notices that adequately explain these systems—their training data sources, processing logic, and envisaged consequences—will determine whether Poland's AI transformation enhances or undermines citizens' trust in public institutions.

Empirical examination of processing notices issued by Polish public bodies deploying AI systems represents a promising avenue for future research. Such analysis could assess whether notices provide meaningful information about AI logic, how technical complexity is translated for ordinary citizens, and whether layered transparency approaches are effectively implemented. Comparative analysis across different public bodies and AI system types would reveal emerging best practices and common compliance challenges.

References

EU Legislation and Official Documents

Article 29 Working Party, *Guidelines on transparency under Regulation 2016/679*, WP260 rev.01, adopted 29 November 2017, last revised 11 April 2018

European Data Protection Board, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, WP251 rev.01.

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 (General Data Protection Regulation), OJ L 119, 4.5.2016, p. 1.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act), OJ L 2024/1689, 12.7.2024.

Polish Legislation and Official Documents

Act of 14 June 1960 – Code of Administrative Procedure (KPA), consolidated text, Journal of Laws 2024, item 572.

Policy for the Development of Artificial Intelligence in Poland from 2020, Appendix to the Resolution No. 196 of the Council of Ministers of 28 December 2020 (item 23), available at: https://www.gov.pl/attachment/928200fa-b1a6-4c0c-b3a8-d1fbf1e1175a [accessed on: 09.022026].

Polish and European Government and Institutional Sources

Ministerstwo Cyfryzacji, *PLLuM — polski model językowy, jak działa i do czego może się przydać*, available at: https://www.gov.pl/web/cyfryzacja/pllum-polski-model-jezykowy-jak-dziala-i-do-czego-moze-sie-przydac [accessed on09.02.2026].

Ministerstwo Cyfryzacji, *Polski model językowy PLLuM trafi do samorządów* (pilot w Częstochowie), available at: https://www.gov.pl/web/cyfryzacja/polski-model-jezykowy-pllum-trafi-do-samorzadu [accessed: 9 February 2026].

National Information Processing Institute - National Research Institute (OPI PIB), *HIVE AI: Development and pilot implementation of large language models in Polish public administration*, available at: https://opi.org.pl/en/project/hive-ai-rozwoj-i-pilotazowe-wdrozenie-duzych-modeli-jezykowych-w-polskiej-administracji-publicznej/ [accessed on: 09.02 2026].

NASK, *About HIVE AI Project*, available at: https://www.nask.pl/en/institute-for-institutions/hive-ai/about-project [accessed on: 09.02.2026].

Implement Consulting Group, *The AI opportunity for eGovernment in Poland* (2025), commissioned by Google in partnership with SGH Warsaw School of Economics, available at: https://cms.implementconsultinggroup.com/media/uploads/articles/2025/The-AI-opportunity-for-eGovernment-in-Poland/2025-The-AI-opportunity-for-eGovernment-in-Poland-English.pdf [accessed on: 09.022026].

European Commission, Digital Skills & Jobs Platform, *Poland digital skills map*, available at: https://digital-skills-jobs.europa.eu/en/european-interactive-map/poland [accessed on: 09.02.2026].

Case Law

Court of Justice of the European Union, judgment of 27 February 2025, Case C-203/22, EU:C:2025:117, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62022CJ0203 [accessed on: 09.02.2026].

Literature

Adamiak B., [in:] B. Adamiak, J. Borkowski, *Kodeks postępowania administracyjnego. Komentarz*, edn. 19, Warszawa 2024, art. 2a, Legalis.

Barta P., Kawecki M., Litwiński P., [in:] P. Litwiński (ed.), *Ustawa o ochronie danych osobowych. Komentarz* [in:] *Ogólne rozporządzenie o ochronie danych osobowych. Ustawa o ochronie danych osobowych. Wybrane przepisy sektorowe. Komentarz*, edn. 2, Warszawa 2025, art. 13, Legalis.

Burrell J., "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms" (September 15, 2015). Available at SSRN: https://ssrn.com/abstract=2660674 or http://dx.doi.org/10.2139/ssrn.2660674, DOI: 10.1177/2053951715622512, [accessed on: 09.02.2026].

Constantino T., "AI's Biggest Secret — Creators Don't Understand It", *Forbes*, 8 May 2025, available at: https://www.forbes.com/sites/torconstantino/2025/05/08/ais-biggest-secret-creators-dont-understand-it-experts-split/ [accessed on: 09.02.2026].

Gołaszewski P., Wąsowski K., [in:] R. Hauser, M. Wierzbowski (eds.), *Kodeks postępowania administracyjnego. Komentarz*, edn. 8, Warszawa 2023, art. 2a, Legalis.

Jarrahi, M.H., Glaser V., "Not All AI Systems Are Created Equal: A Typology of AI Systems' Performance and Affordance", *Proceedings of HICSS 2025*, DOI: 10.24251/HICSS.2025.655.

Litwiński, P., [in:] P. Litwiński (ed..), *Kodeks postępowania administracyjnego. Komentarz* [in:] *Ogólne rozporządzenie o ochronie danych osobowych. Ustawa o ochronie danych osobowych. Wybrane przepisy sektorowe. Komentarz*, edn. 2, Warszawa 2025, art. 2a, Legalis.

Regulatory Guidance

Commission nationale de l'informatique et des libertés, *AI system development: CNIL's recommendations to comply with the GDPR*, 5 January 2026, available at: https://www.cnil.fr/en/ai-system-development-cnils-recommendations-to-comply-gdpr [accessed on: 09.02.2026].

Commission nationale de l'informatique et des libertés, *Informing data subjects*, available at: https://www.cnil.fr/en/informing-data-subjects [accessed on: 9 09.02.2026].

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Jacek Sobczak
Uniwersytet Vizja
ORCID: 0000-0002-2231-8824
j.sobczak@vizja.pl

Ksenia Kakareko
Uniwersytet Warszawski
ORCID: 0000-0003-3707-4479
k.kakareko@uw.edu.pl

Sztuczna inteligencja w wykonywaniu władzy publicznej a demokratyczna rozliczalność organu. Między legalizmem, nadzorem człowieka i kontrolą sądową

Słowa kluczowe: sztuczna inteligencja, władza publiczna, demokratyczna rozliczalność, nadzór człowieka, kontrola sądowa

Streszczenie. Artykuł analizuje demokratyczną rozliczalność użycia sztucznej inteligencji w wykonywaniu władzy publicznej. Punktem wyjścia nie jest pełna automatyzacja decyzji administracyjnych, lecz mniej widoczne sytuacje, w których system AI wpływa na analizę materiału, klasyfikację spraw, przygotowanie projektu rozstrzygnięcia, uzasadnienia lub komunikatu organu. Główna teza zakłada, że demokratyczna rozliczalność jest kategorią szerszą niż zwykła rozliczalność prawna, gdyż obejmuje zachowanie związku między wykonywaniem kompetencji publicznej a demokratycznie legitymowanym organem, rzeczywisty nadzór człowieka, możliwość wyjaśnienia znaczenia systemu dla sprawy oraz skuteczną kontrolę instancyjną i sądową. W artykule odróżniono obowiązki dostawcy i podmiotu stosującego system na tle AI Act, a także wskazano, że ślad decyzyjny stanowi jedynie jeden z instrumentów dokumentacyjnych służących realizacji szerszego standardu. Metoda badawcza ma charakter dogmatycznoprawny i opiera się na analizie Konstytucji RP, Kodeksu postępowania administracyjnego, prawa o postępowaniu przed sądami administracyjnymi, RODO oraz AI Act.

Artificial intelligence in the exercise of public authority and the democratic accountability of the administrative body. Between legality, human oversight and judicial review

Keywords: artificial intelligence, public authority, democratic accountability, human oversight, judicial review

Summary. The article analyses the democratic accountability of the use of artificial intelligence in the exercise of public authority. Its point of departure is not the full automation of administrative decisions, but less visible situations in which an AI system affects the analysis of case materials, case classification, the drafting of decisions, statements of reasons, or public communications. The main argument is that democratic accountability is broader than ordinary legal accountability, because it includes preserving the link between the exercise of public powers and the democratically legitimised

authority, ensuring genuine human oversight, enabling an explanation of the system's role in the case, and securing effective administrative and judicial review. The article distinguishes the obligations of providers and deployers under the AI Act and argues that the decision trace is only one documentary instrument within a broader standard. The analysis is doctrinal and is based on the Polish Constitution, the Code of Administrative Procedure, the Law on Proceedings before Administrative Courts, the GDPR, and the AI Act.

Wprowadzenie

Dyskusja o sztucznej inteligencji w administracji publicznej zbyt często koncentruje się na pytaniu, czy możliwe jest automatyczne wydanie decyzji administracyjnej. Jest to zagadnienie ważne, ale nie wyczerpuje problemu. Prawnie doniosłe może być również posłużenie się systemem na etapach wcześniejszych: przy porządkowaniu materiału, klasyfikacji spraw, przygotowaniu projektu pisma, redakcji uzasadnienia, nadawaniu sprawom kolejności albo komunikacji z obywatelem. W takich sytuacjach końcowe działanie nadal przybiera klasyczną postać aktu, pisma albo czynności materialno-technicznej podpisanej przez człowieka, jednak część jego zaplecza mogła zostać ukształtowana z udziałem systemu[1].

W takim układzie pytanie o sztuczną inteligencję nie jest już pytaniem wyłącznie o technikę organizacji urzędu, lecz o warunki wykonywania władzy publicznej. Jeżeli system współkształtuje tok działania organu, to pojawia się problem możliwości przypisania działania, realności nadzoru człowieka, zakresu informacji należnej jednostce oraz możliwości późniejszej kontroli instancyjnej i sądowej. Nie wystarcza więc odpowiedź, że organ „korzysta z narzędzia". Trzeba jeszcze ustalić, czy użycie tego narzędzia nie rozluźnia związku między kompetencją publiczną a demokratycznie legitymowanym podmiotem władzy[2].

Celem artykułu jest wykazanie, że użycie sztucznej inteligencji w wykonywaniu władzy publicznej wymaga szczególnego standardu demokratycznej rozliczalności. Standard ten jest szerszy niż zwykła rozliczalność prawna, ponieważ obejmuje nie tylko możliwość przypisania odpowiedzialności za wadliwe działanie organu, lecz także zachowanie publicznej uchwytności sposobu wykonywania kompetencji, rzeczywisty nadzór człowieka (*human oversight*), możliwość wyjaśnienia znaczenia systemu dla sprawy (*explainability*) oraz skuteczną kontrolę instancyjną i sądową.

---

[1] G. Sibiga, W. Wiewiórowski, *Automatyzacja rozstrzygnięć i innych czynności w sprawach indywidualnych załatwianych przez organ administracji publicznej*, [w:] *Informatyzacja postępowania sądowego i administracji publicznej*, red. J. Gołaczyński, Warszawa 2010, s. 123-145; F. Geburczyk, *Automatyzacja załatwiania spraw w administracji samorządowej a gwarancje procesowe jednostek. Uwagi de lege ferenda w kontekście ogólnego rozporządzenia o ochronie danych (RODO)*, „Samorząd Terytorialny" 2021, nr 5, s. 21-32.

[2] Por. C. Coglianese, D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, „Georgetown Law Journal" 2017, vol. 105, s. 1147-1223.

Ślad decyzyjny nie stanowi w tej konstrukcji osi samodzielnej, lecz jeden z instrumentów dokumentacyjnych służących realizacji standardu[3].

Tekst ma charakter dogmatycznoprawny. Punktem odniesienia są Konstytucja RP, Kodeks postępowania administracyjnego, prawo o postępowaniu przed sądami administracyjnymi, RODO oraz AI Act. Poza zakresem pozostają empiryczna analiza wdrożeń, zamówienia publiczne i szczegółowe zagadnienia cyberbezpieczeństwa. Artykuł koncentruje się na pytaniu, jakie warunki muszą być spełnione, aby użycie AI przez organ administracji nie rozbijało związku między wykonywaniem kompetencji publicznej, odpowiedzialnym organem i możliwością późniejszej kontroli[4].

1. Demokratyczna rozliczalność a zwykła rozliczalność prawna

Pojęcie rozliczalności należy odróżnić od klasycznej odpowiedzialności prawnej. Odpowiedzialność prawna aktualizuje się zazwyczaj po naruszeniu normy i prowadzi do zastosowania określonych konsekwencji. Rozliczalność obejmuje etap wcześniejszy: obowiązek zdania sprawy z działania, wyjaśnienia jego podstaw i poddania go ocenie. W tym sensie *accountability* opisuje relację szerszą niż sama sankcja za naruszenie prawa[5].

Dla dalszego wywodu szczególne znaczenie ma rozróżnienie zaproponowane przez Marka Bovensa między *accountability as virtue* a *accountability as mechanism*. W pierwszym ujęciu akcent pada na cnoty dobrego zarządcy: przejrzystość, odpowiedzialność, rzetelność i gotowość do samokontroli. W drugim chodzi o uchwytny układ relacyjny, w którym określony aktor jest zobowiązany przed określonym forum do przedstawienia, wyjaśnienia i obrony swego działania, a forum może działanie to ocenić oraz wyciągnąć konsekwencje. Niniejszy artykuł przyjmuje właśnie drugie, mechanizmowe rozumienie rozliczalności, ponieważ tylko ono pozwala przełożyć problem użycia AI przez organ na język obowiązków proceduralnych, kontroli i odpowiedzialności instytucjonalnej[6].

---

[3] Na temat dokumentacyjnych warunków odtwarzalności użycia AI w administracji por. A. Szot, *Algorithm in the Structure of Administrative Decision-Making: A Model of Explainability and Burden of Proof for the Purposes of an Effective Remedy*, „Studia Iuridica Lublinensia" 2025, vol. 34, nr 5, s. 257-276.

[4] Podstawę normatywną analizy tworzą przede wszystkim: Konstytucja RP, k.p.a., ustawa z 30.08.2002 r. – Prawo o postępowaniu przed sądami administracyjnymi (dalej p.p.s.a.), RODO oraz rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 (AI Act).

[5] M. Bovens, *Analysing and Assessing Accountability: A Conceptual Framework*, „European Law Journal" 2007, vol. 13, nr 4, s. 447-468; R. Mulgan, *Holding Power to Account. Accountability in Modern Democracies*, Basingstoke 2003, s. 8-14.

[6] M. Bovens, *Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism*, „West European Politics" 2010, vol. 33, nr 5, s. 946-967, zwł. s. 948-951; T. Schillemans, *Moving Beyond the*

Rozliczalność może mieć wymiar prawny, administracyjny, polityczny albo organizacyjny. W niniejszym artykule chodzi jednak o jej odmianę demokratyczną. Nie wyczerpuje się ona w odpowiedzi na pytanie, kto poniesie konsekwencje błędu. Dotyczy również tego, czy można jeszcze ustalić, że to właśnie organ publiczny rzeczywiście wykonywał kompetencję, a nie tylko nadał urzędową formę wynikowi procesu, którego zasadniczy przebieg został ukształtowany przez system obliczeniowy. Demokratyczna rozliczalność wymaga zatem nie tylko późniejszej odpowiedzialności za naruszenie prawa, lecz także wcześniejszej publicznej uchwytności sposobu wykonywania kompetencji.

Tak rozumiana demokratyczna rozliczalność znajduje zakorzenienie w art. 1 i 2 Konstytucji RP. Państwo prawa nie może być wyłącznie skuteczne; musi być także normatywnie uchwytne i podatne na ocenę z perspektywy obywatela. Z kolei art.4 i 7 Konstytucji RP przypominają, że władza publiczna nie jest anonimowym procesem technicznym, lecz wykonywaniem kompetencji powierzonych organom działającym na podstawie i w granicach prawa[7].

Demokratyczna rozliczalność obejmuje więc co najmniej cztery elementy: możliwość przypisania działania organowi, uchwytność poznawczą procesu, możliwość wyjaśnienia roli systemu w sprawie oraz kontrolowalność działania przez jednostkę i sąd. Dopiero ta łączna konstrukcja pozwala ocenić, czy użycie AI nie osłabia samego standardu wykonywania władzy publicznej. W dalszych częściach artykułu elementy te zostaną rozwinięte osobno[8].

2. AI w wykonywaniu władzy publicznej jako problem ustrojowy

Administracja publiczna od dawna korzysta z ekspertyz, modeli statystycznych i systemów informatycznych. Sama obecność narzędzia technicznego nie jest więc niczym nowym. Nowość AI polega na tym, że system może współkształtować tok, treść albo wynik działania organu, a więc wejść w samą strukturę wykonywania kompetencji publicznej[9].

Z punktu widzenia art. 4 i 7 Konstytucji RP oznacza to powstanie pytania, czy formalny podmiot odpowiedzialności pozostaje zarazem rzeczywistym podmiotem decyzji. Jeżeli wpływ systemu na działanie organu staje się realny, lecz zarazem

Clash of Interests: On Stewardship Theory and the Relationships Between Central Government Departments and Public Agencies, „Public Management Review" 2013, vol. 15, nr 4, s. 541-562.

[7] *Konstytucja Rzeczypospolitej Polskiej. Komentarz,* red. L. Garlicki, M. Zubik, t. 1, Warszawa 2016, komentarz do art. 1 i 2; M. Florczak-Wątor, [w:] *Konstytucja RP*, t. 1: *Komentarz. Art. 1-86*, red. M. Safjan, L. Bosek, Warszawa 2016, komentarz do art. 7.

[8] A. Szot, *op. cit.*, s. 257-276.

[9] C. Coglianese, D. Lehr, *op. cit.*, s. 1156-1162.

słabo uchwytny, problem przestaje mieć wyłącznie charakter organizacyjny. Staje się problemem ustrojowym, ponieważ dotyczy związku między demokratyczną legitymacją organu a sposobem wykonywania przezeń kompetencji[10].

Nie chodzi tu wyłącznie o przypadki pełnej automatyzacji. W praktyce znaczenie AI ujawnia się częściej na etapach wcześniejszych: selekcji materiału, klasyfikacji spraw, redakcji projektu rozstrzygnięcia albo komunikatu. To właśnie te mniej widoczne sytuacje szczególnie łatwo wymykają się późniejszej kontroli, choć mogą wywierać istotny wpływ na sytuację jednostki. Stąd potrzeba standardu, który obejmuje nie tylko produkt finalny, ale również drogę dojścia do działania organu[11].

3. Nadzór człowieka i problem fasadowości

Jednym z najczęściej powtarzanych argumentów na rzecz dopuszczalności użycia AI w administracji jest twierdzenie, że człowiek pozostaje w procedurze i zachowuje kontrolę nad wynikiem. Argument ten jest trafny tylko wtedy gdy udział człowieka ma charakter rzeczywisty. Sama obecność człowieka w procesie nie przesądza jeszcze o zachowaniu kontroli nad działaniem organu[12].

Trzeba odróżnić obecność formalną, techniczną akceptację oraz rzeczywistą możliwość oceny, zmiany lub odrzucenia wyniku systemu. Jeżeli urzędnik nie zna funkcji systemu, nie rozumie charakteru przedstawionego mu wyniku albo nie ma realnej swobody jego zakwestionowania, jego udział zachowuje zewnętrzną formę odpowiedzialności, ale nie stanowi realnego nadzoru. Taka sytuacja prowadzi do fasadowości nadzoru człowieka i osłabia możliwość przypisania działania organowi[13].

Na tle zasady legalizmu nie można zaakceptować modelu, w którym organ sprowadza się do roli nośnika podpisu albo pieczęci. Władza publiczna musi pozostać rzeczywistym wykonawcą kompetencji, a nie tylko formalną osłoną dla wyniku obliczeniowego. Dlatego demokratyczna rozliczalność wymaga nie samej obecności człowieka, lecz materialnie uchwytnego nadzoru człowieka. W dalszej części artykułu będzie to miało znaczenie dla oceny obowiązków *deployera*, zakresu informacji należnej jednostce i sądowej kontrolowalności działania organu[14].

---

[10] Art. 4 i 7 Konstytucji RP; por. także *Konstytucja Rzeczypospolitej Polskiej. Komentarz*, red. P. Tuleja, Warszawa 2023, komentarz do art. 2 i 7.

[11] J. Blicharz, *Inteligentne miasta i sztuczna inteligencja. Wybrane aspekty teoretycznoprawne*, Wrocław 2023, s. 88-97.

[12] G. Sibiga, *Problem fikcji czynnika ludzkiego w rozstrzygnięciach podejmowanych automatycznie na podstawie prawa administracyjnego*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3, s. 111-126.

[13] *Ibidem*; por. też J. Sawicki, *The Role of Decision-Making Algorithms in Public Administration – Legal and Ethical Challenges*, „Zeszyt Prawniczy UAM" 2025, nr 15, s. 99-111.

[14] Art. 7 Konstytucji RP; por. B. Adamiak, J. Borkowski, *Kodeks postępowania administracyjnego. Komentarz*, Warszawa 2024, komentarz do art. 7.

4. AI Act i rozkład obowiązków między dostawcą a podmiotem stosującym system

AI Act nie może być czytany wyłącznie jako akt skierowany do projektantów i producentów technologii. Rozporządzenie odróżnia dostawcę (*provider*) od podmiotu stosującego system (*deployer*), a w praktyce administracyjnej to właśnie organ publiczny będzie najczęściej *deployerem*. Oznacza to, że nie może on ograniczyć się do twierdzenia, iż odpowiedzialność za wszystkie ryzyka i obowiązki regulacyjne spoczywa po stronie podmiotu zewnętrznego[15].

Provider odpowiada przede wszystkim za zgodność systemu jako produktu: jego projekt, dokumentację techniczną, ocenę zgodności i instrukcję użycia. *Deployer* odpowiada natomiast za sposób użycia systemu w konkretnym kontekście instytucjonalnym. Z punktu widzenia wykonywania władzy publicznej ma to znaczenie fundamentalne, bo problemem nie jest tylko zgodność systemu jako rozwiązania technicznego, lecz również to, czy organ używa go w sposób zachowujący realny nadzór człowieka i możliwość późniejszej kontroli[16].

Szczególne znaczenie ma art. 26 AI Act, który określa obowiązki *deployerów* systemów wysokiego ryzyka. Obejmują one m.in. stosowanie systemu zgodnie z instrukcją, zapewnienie odpowiednich środków organizacyjnych, wyznaczenie osób sprawujących *human oversight*, monitorowanie działania systemu oraz przechowywanie automatycznie generowanych logów, o ile pozostają one pod kontrolą *deployera*. Przepis ten pokazuje, że prawodawca unijny wiąże rozliczalność nie tylko z fazą projektową, lecz także z fazą użycia systemu przez podmiot publiczny[17].

Nie mniej ważny jest art. 27 AI Act, przewidujący ocenę wpływu na prawa podstawowe (*fundamental rights impact assessment*) dla określonych systemów wysokiego ryzyka, oraz art. 86 AI Act, przyznający osobie dotkniętej decyzją prawo do uzyskania od *deployera* jasnych i znaczących wyjaśnień co do roli systemu w procedurze decyzyjnej. Oba przepisy wzmacniają tezę, że użycie AI przez organ nie może pozostawać poza zakresem obowiązków dokumentacyjnych, organizacyjnych i informacyjnych. AI Act nie tworzy samodzielnie całego standardu demokratycznej rozliczalności, ale wyraźnie go wzmacnia[18].

---

[15] AI Act, art. 3 pkt 3 i 4.

[16] H. Graux *et al.*, *Interplay between the AI Act and the EU digital legislative framework*, European Parliament, PE 778.575, 2025, s. 21-24.

[17] AI Act, art. 26.

[18] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13.06.2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji, art. 27 i 86; motyw 93 i 96 rozporządzenia 2024/1689.

5. Możliwość wyjaśnienia i zakres informacji potrzebnej jednostce

Przejrzystość techniczna nie jest tym samym co możliwość wyjaśnienia znaczenia systemu dla sprawy jednostki. Obywatel nie musi znać pełnej architektury modelu ani kodu źródłowego, lecz może potrzebować informacji o tym, czy AI została użyta, w jakim celu i jaki wpływ wywarła na działanie organu. To właśnie taki poziom informacji ma znaczenie z punktu widzenia ochrony proceduralnej i prawa do sądu[19].

Możliwość wyjaśnienia należy więc rozumieć funkcjonalnie. Chodzi nie o pełen opis technologii, lecz o taki zakres wyjaśnienia, który pozwala jednostce zrozumieć, czy i w jaki sposób wynik systemu wpłynął na jej sprawę. Właśnie dlatego możliwość wyjaśnienia trzeba łączyć z art. 61 Konstytucji RP, art. 9 i 11 k.p.a. oraz odpowiednimi gwarancjami wynikającymi z RODO i AI Act. Nie jest to prawo do poznania całej technologii, ale prawo do informacji użytecznej z punktu widzenia obrony własnej sytuacji prawnej[20].

Tak rozumiana możliwość wyjaśnienia znajduje mocne potwierdzenie w orzecznictwie unijnym. W wyroku z 27 lutego 2025 r. w sprawie C-203/22 CK przeciwko Dun & Bradstreet Austria GmbH i Magistrat der Stadt Wien Trybunał Sprawiedliwości podkreślił, że osoba, wobec której zastosowano mechanizm oceny opartej na zautomatyzowanym przetwarzaniu, musi otrzymać informacje pozwalające zrozumieć logikę postępowania i zakwestionować wynik, przy czym tajemnica przedsiębiorstwa nie może prowadzić do wyłączenia istoty tego uprawnienia. Oznacza to, że obowiązek wyjaśnienia nie może zostać zastąpiony ani czysto technicznym opisem systemu, ani lakonicznym zapewnieniem, że użyto narzędzia wspierającego. Informacja musi być na tyle konkretna, by miała znaczenie dla pozycji prawnej jednostki.

W polskim porządku konstytucyjnym podobny kierunek wzmacnia art. 61 Konstytucji RP, odczytywany w świetle wyroku Trybunału Konstytucyjnego z 20 marca 2006 r., K 17/05, w którym podkreślono, że ograniczenia prawa do informacji muszą odpowiadać wymogom z art. 31 ust. 3 Konstytucji RP i nie mogą niweczyć samej istoty uprawnienia informacyjnego. W sprawach wspieranych przez AI oznacza

---

[19] J. Blicharz, *Zautomatyzowane podejmowanie decyzji oparte na technologii algorytmicznej: kilka refleksji na temat prawa do wyjaśnienia w kontekście RODO*, [w:] *Stosowanie prawa administracyjnego. Księga jubileuszowa Profesora Andrzeja Matana*, red. G. Łaszczyca, Warszawa 2024, s. 123-130; wyrok TS z 27.02.2025 r., C-203/22, CK przeciwko Dun & Bradstreet Austria GmbH i Magistrat der Stadt Wien, ECLI:EU:C:2025:117.

[20] Art. 61 Konstytucji RP; art. 9 i 11 k.p.a.; S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, „International Data Privacy Law" 2017, vol. 7, nr 2, s. 76-99; wyrok TK z 20.03.2006 r., K 17/05, OTK-A 2006, nr 3, poz. 30.

to, że organ nie może zasłaniać się technologicznym charakterem narzędzia w celu uniknięcia wyjaśnienia roli systemu w działaniu władzy publicznej. Zakres informacji może podlegać ograniczeniom, lecz samo prawo do informacji o znaczeniu systemu dla sprawy nie może stać się iluzoryczne[21].

Tak rozumiana możliwość wyjaśnienia pozostaje warunkiem zarówno realnego nadzoru człowieka, jak i późniejszej kontroli sądowej. Jeżeli organ nie potrafi wyjaśnić znaczenia systemu dla swojej sprawy inaczej niż przez ogólne stwierdzenie, że narzędzie wspierało proces, rozliczalność ma charakter wyłącznie formalny. Dlatego możliwość wyjaśnienia jest jednym z kluczowych elementów demokratycznej rozliczalności, ale nie może zostać utożsamiona z pełną przezroczystością technologiczną[22].

6. Sądowa kontrolowalność działania organu wspieranego przez AI

Jednym z najważniejszych sprawdzianów demokratycznej rozliczalności jest odpowiedź na pytanie, czy działanie organu wspieranego przez AI pozostaje realnie kontrolowalne przez sąd administracyjny. Sąd nie ma badać technologii jako takiej, lecz legalność działania organu. Musi jednak mieć dostateczny materiał, aby ocenić, czy organ zachował samodzielność działania, dochował obowiązków informacyjnych i oparł się na materiale nadającym się do kontroli *ex post*[23].

Nie oznacza to, że sąd ma każdorazowo rekonstruować całą architekturę modelu albo badać kod źródłowy. Równie błędne byłoby jednak założenie, że sam podpis człowieka pod końcowym aktem eliminuje problem kontroli. Sąd potrzebuje takiego materiału, który pozwala odpowiedzieć, czy AI była jedynie narzędziem pomocniczym, czy też współkształtowała elementy istotne z punktu widzenia legalności rozstrzygnięcia albo sytuacji jednostki.

W tym kontekście szczególnego znaczenia nabiera art. 133 § 1 p.p.s.a., zgodnie z którym sąd wydaje wyrok na podstawie akt sprawy, oraz art. 145 § 1 p.p.s.a. określający skutki stwierdzonych naruszeń prawa. Jeżeli wpływ AI na tok sprawy nie został w aktach w żaden sposób utrwalony, sąd może zostać postawiony wobec formalnie kompletnego rozstrzygnięcia, którego rzeczywistego zaplecza poznawczego nie sposób ocenić. W takim układzie problem dotyczy już nie samej techniki

---

[21] Art. 61 Konstytucji RP; art. 9 i 11 k.p.a.; S. Wachter, B. Mittelstadt, L. Floridi, *op. cit.*; wyrok TK z 20.03.2006 r., *op. cit.*

[22] L. Edwards, M. Veale, *Slave to the Algorithm? Why a Right to an Explanation Is Probably Not the Remedy You Are Looking For,* „Duke Law and Technology Review" 2017, vol. 16, s. 18-84.

[23] Konstytucja RP, art. 45, 77, 78 i 184; ponadto art. 1, 133 § 1 i art. 145 § 1 p.p.s.a.

działania organu, lecz granic materiału procesowego, na podstawie którego sąd ma wykonać konstytucyjną funkcję kontroli administracji[24].

Brak śladu dokumentującego wpływ AI nie zawsze będzie prowadził do wadliwości aktu, ale nie może być z góry traktowany jako nieistotny. Jeżeli system odegrał rolę czysto techniczną i nie oddziaływał na ustalenia faktyczne, kwalifikację prawną ani treść rozstrzygnięcia, brak szczegółowego odtworzenia jego działania może pozostać uchybieniem organizacyjnym. Inaczej należy ocenić sytuację, w której AI wpływała na dobór materiału, klasyfikację zgłoszeń, propozycję uzasadnienia, ocenę ryzyka albo kolejność rozpoznania spraw. Wówczas brak informacji o tym wpływie może oznaczać naruszenie przepisów postępowania mogące mieć istotny wpływ na wynik sprawy w rozumieniu art. 145 § 1 pkt 1 lit. c p.p.s.a.; w skrajnych sytuacjach może także prowadzić do takiego zubożenia uzasadnienia, że utrudni ocenę legalności aktu na gruncie art. 145 § 1 pkt 1 lit. a p.p.s.a. w związku z naruszeniem przepisów postępowania administracyjnego dotyczących wyjaśnienia sprawy i uzasadnienia rozstrzygnięcia.

Powyższy wniosek wzmacnia orzecznictwo unijne dotyczące decyzji opartych na zautomatyzowanym przetwarzaniu. W wyroku z 7 grudnia 2023 r. w sprawie C-634/21 OQ przeciwko Land Hessen Trybunał Sprawiedliwości uznał, że zautomatyzowana ocena zdolności kredytowej może stanowić decyzję opartą wyłącznie na automatycznym przetwarzaniu w rozumieniu art. 22 RODO, jeżeli w praktyce determinuje zachowanie podmiotu korzystającego z tej oceny. Chociaż wyrok zapadł na tle sektora prywatnego, jego *ratio* ma znaczenie także dla administracji: jeżeli wynik systemu w praktyce współokreśla rozstrzygnięcie albo istotny element działania organu, nie można sprowadzać go do niewidocznej technicznej okoliczności ubocznej. To właśnie w takich przypadkach brak materiału pozwalającego odtworzyć wpływ systemu osłabia realność sądowej kontroli[25].

W praktyce oznacza to, że sądowa kontrolowalność działania organu wspieranego przez AI zależy nie od pełnej technicznej przezroczystości systemu, lecz od istnienia dostatecznego materiału pozwalającego ocenić samodzielność działania organu. Tym samym brak odpowiedniej dokumentacji wpływu AI może mieć doniosłość procesową nie dlatego, że sąd musi badać sam system, ale dlatego, że bez tego materiału nie jest w stanie ocenić legalności działania organu w zwykłym,

---

[24] M. Jaśkowska, A. Kabat, [w:] *Prawo o postępowaniu przed sądami administracyjnymi. Komentarz*, red. R. Hauser, M. Wierzbowski, Warszawa 2024, komentarz do art. 133 i 145 p.p.s.a.; M. Kupiec, *Uzasadnienie decyzji administracyjnej wydanej za pomocą systemu sztucznej inteligencji w świetle realizacji zasady budowania zaufania do władzy publicznej*, „Dyskurs Prawniczy i Administracyjny" 2025, nr 1, s. 137-147; wyrok NSA z 19.04.2021 r., III OSK 836/21, OSP 2022, nr 1, poz. 9.

[25] F. Geburczyk, *op. cit.*, s. 26-30; A. Szot, *op. cit.*, s. 269-273; wyrok TS UE z 7.12.2023 r., C-634/21, OQ przeciwko Land Hessen, ECLI:EU:C:2023:957.

klasycznym sensie. Właśnie dlatego sądowa kontrolowalność pozostaje jednym z głównych kryteriów dopuszczalnego polegania na AI przez organ administracji.

7. Ślad decyzyjny jako instrument dokumentacyjny demokratycznej rozliczalności

Demokratyczna rozliczalność wymaga nie tylko norm proceduralnych i obowiązków organizacyjnych, lecz także instrumentów pozwalających odtworzyć rzeczywisty przebieg działania organu. W tym sensie ślad decyzyjny należy traktować jako instrument dokumentacyjny służący realizacji szerszego standardu, a nie jako samodzielną teorię użycia AI przez administrację. Jego rolą jest utrwalenie tego, co w przeciwnym razie mogłoby pozostać wyłącznie wewnętrznym i niewidocznym elementem działania urzędu[26].

Ślad decyzyjny można rozumieć jako utrwalony zespół informacji pozwalających ustalić, czy AI została użyta, na jakim etapie, w jakim celu, z jakim wynikiem i jaki był zakres późniejszej weryfikacji człowieka. Nie chodzi tu o pełną dokumentację techniczną ani o pełną jawność systemu, lecz o taki poziom utrwalenia, który pozwala połączyć wymaganie odpowiedzialności organu z wymogiem wyjaśnienia wpływu systemu i późniejszej kontroli. Z tego względu ślad nie wyczerpuje demokratycznej rozliczalności, ale może stanowić jej najważniejsze narzędzie dowodowe i organizacyjne[27].

Jego zakres nie musi być identyczny w każdej sprawie. Powinien zależeć od rzeczywistego znaczenia AI dla działania organu i od wagi sprawy dla sytuacji jednostki. To odróżnia ślad decyzyjny od nadmiarowego formalizmu i pozwala zachować go w granicach rozsądnego wymogu dokumentacyjnego. W nowej strukturze artykułu jego wartość polega nie na tym, że zastępuje całą analizę, lecz na tym, że pozwala ją praktycznie urzeczywistnić[28].

8. Granice demokratycznej rozliczalności: proporcjonalność, tajemnice chronione, dane i bezpieczeństwo

Demokratyczna rozliczalność nie oznacza pełnej jawności wszystkiego, co wiąże się z użyciem AI przez organ. Zakres wymaganej dokumentacji, jawności i wyjaśnienia musi pozostawać proporcjonalny do znaczenia systemu dla sprawy oraz do ciężaru

[26] Por. G. Sibiga, W. Wiewiórowski, *op. cit.*, s. 123-145; A. Szot, *op. cit.*, s. 257-260.

[27] Na temat dokumentacyjnych wymogów odtwarzalności użycia systemu por. A. Szot, *op. cit.*, s. 261-266; G. Sibiga, op. cit., s. 118-121.

[28] J. Sawicki, *op. cit.*, s. 105-108.

ingerencji w sytuację jednostki. Nie każda sprawa wymaga tego samego poziomu dokumentowania i ujawnienia.

Punktem odniesienia dla tego ważenia jest art. 31 ust. 3 Konstytucji RP. Skoro każda ingerencja ograniczająca prawa i wolności musi być konieczna w demokratycznym państwie, służyć prawnie legitymowanemu celowi oraz pozostawać proporcjonalna *sensu stricto*, to również zakres dopuszczalnego polegania na AI powinien być oceniany przez pryzmat tych trzech pytań: czy użycie systemu jest przydatne i potrzebne dla realizacji zadania publicznego, czy nie da się osiągnąć celu środkami mniej dotkliwymi dla jednostki, oraz czy korzyści z użycia AI pozostają we właściwej proporcji do ciężaru, jaki rozwiązanie to nakłada na prawa, sytuację procesową lub pozycję informacyjną jednostki[29].

Przy ocenie tej proporcjonalności należy brać pod uwagę co najmniej wagę sprawy, stopień wpływu AI na wynik, zakres ingerencji w dane lub prawa jednostki, możliwość korekty przez człowieka oraz dostępność skutecznego środka kontroli. Kryteria te nie tworzą autonomicznego testu oderwanego od Konstytucji, lecz konkretyzują zastosowanie art. 31 ust. 3 Konstytucji RP do sytuacji, w których organ posługuje się systemem sztucznej inteligencji. Nie mogą przy tym służyć jako wygodny argument za zatarciem roli systemu tam, gdzie wpływ AI był realny i istotny dla sprawy[30].

Istotnym elementem proporcjonalności pozostaje też rozróżnienie między obowiązkiem zachowania informacji a obowiązkiem ich udostępnienia. Organ może być zobowiązany zachować więcej niż ostatecznie ujawnia stronie lub opinii publicznej, ponieważ granice ujawnienia wyznaczają prywatność, dane osobowe, tajemnice ustawowo chronione i bezpieczeństwo systemu. Ograniczenia te nie znoszą jednak samej potrzeby dokumentowania. Przeciwnie, dopiero istnienie materiału pozwala ustalić, co może zostać ujawnione, co wymaga anonimizacji, a co powinno pozostać chronione. Właśnie taki kierunek wynika także z orzecznictwa Trybunału Konstytucyjnego dotyczącego prawa do informacji publicznej, które wymaga, by ograniczenia tego prawa pozostawały wyjątkowe i zgodne z konstytucyjnym testem proporcjonalności.

Granic demokratycznej rozliczalności nie wyznacza zatem sama techniczna złożoność AI, lecz relacja między ochroną jednostki, wymaganiami państwa prawa i uzasadnionymi ograniczeniami jawności. Im większy jest wpływ systemu na sytuację prawną jednostki, tym słabsze staje się uzasadnienie dla ograniczania dokumentowania, wyjaśnienia i kontroli. Proporcjonalność nie osłabia więc standardu

[29] Art. 31 ust. 3 Konstytucji RP jako wzorzec ważenia ograniczeń i wymagań proceduralnych; por. wyrok TK z 20.03.2006 r., *op. cit.*

[30] O potrzebie ważenia wpływu systemu na prawa jednostki zob. AI Act, art. 27; por. też H. Graux *et al.*, *op. cit.*, s. 23.

demokratycznej rozliczalności, lecz porządkuje jego zastosowanie w zależności od ciężaru sprawy[31].

Podsumowanie

Sztuczna inteligencja w wykonywaniu władzy publicznej nie może być oceniana wyłącznie przez pryzmat sprawności i modernizacji urzędu. Wpływa ona na warunki przypisywania działania organowi, realność nadzoru człowieka, zakres informacji należnej jednostce oraz możliwość kontroli instancyjnej i sądowej. Właśnie dlatego użycie AI wymaga szczególnego standardu demokratycznej rozliczalności.

Standard ten jest szerszy niż zwykła odpowiedzialność prawna. Obejmuje możliwość przypisania działania organowi, możliwość wyjaśnienia roli systemu dla sprawy, realność nadzoru człowieka, kontrolowalność sądową oraz odpowiednie instrumenty dokumentacyjne. AI Act wzmacnia tę konstrukcję, ale jej fundament pozostaje zakorzeniony także w Konstytucji RP, KPA i p.p.s.a.

Ślad decyzyjny pozostaje ważnym narzędziem dokumentacyjnym, lecz nie zastępuje całego standardu. Dopuszczalność użycia AI zależy więc nie od samego faktu wykorzystania technologii, lecz od tego, czy nie rozrywa ona związku między kompetencją publiczną, odpowiedzialnym organem i możliwością późniejszej kontroli. Ten związek należy uznać za rdzeń demokratycznej rozliczalności użycia sztucznej inteligencji w wykonywaniu władzy publicznej.

Literatura

Adamiak B., Borkowski J., *Kodeks postępowania administracyjnego. Komentarz*, Warszawa 2024.

Bidziński M., Chmaj M., Szustakiewicz P., *Ustawa o dostępie do informacji publicznej. Komentarz*, Warszawa 2023.

Blicharz J., *Inteligentne miasta i sztuczna inteligencja. Wybrane aspekty teoretycznoprawne*, Wrocław 2023.

Blicharz J., *Zautomatyzowane podejmowanie decyzji oparte na technologii algorytmicznej: kilka refleksji na temat prawa do wyjaśnienia w kontekście RODO*, [w:] *Stosowanie prawa administracyjnego. Księga jubileuszowa Profesora Andrzeja Matana*, red. G. Łaszczyca, Warszawa 2024.

Bovens M., *Analysing and Assessing Accountability: A Conceptual Framework*, „European Law Journal" 2007, vol. 13, nr 4.

Bovens M., *Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism*, „West European Politics" 2010, vol. 33, nr 5.

Coglianese C., Lehr D., *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, „Georgetown Law Journal" 2017, vol. 105.

Edwards L., Veale M., *Slave to the Algorithm? Why a Right to an Explanation Is Probably Not the Remedy You Are Looking For*, "Duke Law and Technology Review" 2017, vol. 16.

[31] E. Gudowska-Natanek, G. Kuca, *Ustawa o dostępie do informacji publicznej. Komentarz*, Warszawa 2024, komentarz do art. 5; M. Bidziński, M. Chmaj, P. Szustakiewicz, *Ustawa o dostępie do informacji publicznej. Komentarz*, Warszawa 2023, komentarz do art. 5; wyrok TK z 20.03.2006 r., *op. cit*.

Florczak-Wątor M., [w:] *Konstytucja RP*, t. 1: *Komentarz*, red. M. Safjan, L. Bosek, Art. 1-86, Warszawa 2016.
Konstytucja Rzeczypospolitej Polskiej. Komentarz, t. 1, red. L. Garlicki, M. Zubik, Warszawa 2016.
Geburczyk F., *Automatyzacja załatwiania spraw w administracji samorządowej a gwarancje procesowe jednostek. Uwagi de lege ferenda w kontekście ogólnego rozporządzenia o ochronie danych (RODO)*, „Samorząd Terytorialny" 2021, nr 5.
Graux H. *et al.*, *Interplay between the AI Act and the EU digital legislative framework*, European Parliament, PE 778.575, 2025.
Gudowska-Natanek E., Kuca G., *Ustawa o dostępie do informacji publicznej. Komentarz*, Warszawa 2024.
Jaśkowska M., Kabat A., [w:] *Prawo o postępowaniu przed sądami administracyjnymi. Komentarz*, red. R. Hauser, M. Wierzbowski, Warszawa 2024.
Kupiec M., *Uzasadnienie decyzji administracyjnej wydanej za pomocą systemu sztucznej inteligencji w świetle realizacji zasady budowania zaufania do władzy publicznej*, „Dyskurs Prawniczy i Administracyjny" 2025, nr 1.
Mulgan R., *Holding Power to Account. Accountability in Modern Democracies*, Basingstoke 2003.
Sawicki J., *The Role of Decision-Making Algorithms in Public Administration - Legal and Ethical Challenges*, „Zeszyt Prawniczy UAM" 2025, nr 15.
Schillemans T., *Moving Beyond the Clash of Interests: On Stewardship Theory and the Relationships Between Central Government Departments and Public Agencies*, „Public Management Review" 2013, vol. 15, nr 4.
Sibiga G., *Problem fikcji czynnika ludzkiego w rozstrzygnięciach podejmowanych automatycznie na podstawie prawa administracyjnego*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3.
Sibiga G., Wiewiórowski W., *Automatyzacja rozstrzygnięć i innych czynności w sprawach indywidualnych załatwianych przez organ administracji publicznej*, [w:] *Informatyzacja postępowania sądowego i administracji publicznej*, red. J. Gołaczyński Warszawa 2010.
Szot A., *Algorithm in the Structure of Administrative Decision-Making: A Model of Explainability and Burden of Proof for the Purposes of an Effective Remedy*, „Studia Iuridica Lublinensia" 2025, vol. 34, nr 5.
Konstytucja Rzeczypospolitej Polskiej. Komentarz, red. P. Tuleja, Warszawa 2023.
Wachter S., Mittelstadt B., Floridi L., *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, „International Data Privacy Law" 2017, vol. 7, nr 2.

Akty normatywne i orzecznictwo

Konstytucja Rzeczypospolitej Polskiej z 2.04.1997 r., Dz.U. Nr 78, poz. 483 ze zm.
Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27.04.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE, Dz.Urz. UE L 119.
Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13.06.2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji, Dz.Urz. UE L 2024/1689.
Ustawa z 14.06.1960 r. – Kodeks postępowania administracyjnego, t.j. Dz.U. z 2025 r. poz. 1691.
Ustawa z 30.08.2002 r. – Prawo o postępowaniu przed sądami administracyjnymi, t.j. Dz.U. z 2024 r. poz. 935 ze zm.
Ustawa z 6.09.2001 r. o dostępie do informacji publicznej, t.j. Dz.U. z 2022 r. poz. 902.
Wyrok NSA z 19.04.2021 r., III OSK 836/21, OSP 2022, nr 1, poz. 9.
Wyrok NSA z 26.05.2022 r., III OSK 1189/21, ONSAiWSA 2022, nr 6, poz. 92.
Wyrok TK z 20.03.2006 r., K 17/05, OTK-A 2006, nr 3, poz. 30.
Wyrok TS z 7.12.2023 r., C-634/21, OQ przeciwko Land Hessen, ECLI:EU:C:2023:957.
Wyrok TS z 27.02.2025 r., C-203/22, CK przeciwko Dun & Bradstreet Austria GmbH i Magistrat der Stadt Wien, ECLI:EU:C:2025:117.

UNIWERSYTET ZIELONOGÓRSKI

DYSKURS PRAWNICZY I ADMINISTRACYJNY 2/2026 ISSN 2657-926X

Mihaela Miruna Tudorascu
University "1 Decembrie 1918" of Alba Iulia, Romania
ORCID: 0000-0003-4641-4578
miruna.tudorascu@uab.ro

Raluca Oana Ivan
University "1 Decembrie 1918" of Alba Iulia, Romania
ORCID: 0000-0002-8019-3601
raluca.ivan@uab.ro

Personal Data Protection in Digitalized Public Administration: Challenges and Solutions in the Context of AI

Keywords*:* personal data protection, artificial intelligence, public administration, GDPR, EU AI Act

Summary. The rapid digitalization of public administration, accelerated by artificial intelligence (AI) technologies, promises enhanced efficiency and improved citizen services but simultaneously introduces profound risks to personal data protection. This paper examines these challenges within the context of Romania's ongoing governance reforms, proposing actionable solutions aligned with the General Data Protection Regulation (GDPR) and the EU Artificial Intelligence Act. Through a qualitative research methodology combining EU-wide regulatory analysis with Romania-focused case studies, this study addresses key tensions between AI-driven automation, such as predictive analytics, automated decision-making systems, and biometric identification, and fundamental privacy rights. With Romania advancing its Digital Governance Framework amid EU mandates, unresolved issues including algorithmic bias, surveillance concerns, and enforcement gaps threaten both public trust and regulatory compliance. This analysis offers policymakers and practitioners practical pathways to balance technological innovation with robust data protection through Privacy by Design principles, enhanced oversight mechanisms, and transparent algorithmic governance.

Ochrona danych osobowych w cyfryzowanej administracji publicznej: wyzwania i rozwiązania w kontekście AI

Słowa kluczowe: ochrona danych osobowych, sztuczna inteligencja, administracja publiczna, RODO/ /GDPR, unijny akt w sprawie sztucznej inteligencji

Streszczenie. Szybka cyfryzacja administracji publicznej, przyspieszona przez technologie sztucznej inteligencji (AI), obiecuje zwiększenie efektywności oraz poprawę jakości usług świadczonych obywatelom, lecz jednocześnie rodzi poważne ryzyka dla ochrony danych osobowych. Artykuł analizuje te wyzwania w kontekście trwających reform zarządzania publicznego w Rumunii, proponując praktyczne rozwiązania zgodne z ogólnym rozporządzeniem o ochronie danych osobowych (RODO/ /GDPR) oraz unijnym aktem w sprawie sztucznej inteligencji.

W oparciu o jakościową metodologię badawczą, łączącą analizę regulacji obowiązujących w Unii Europejskiej ze studiami przypadków dotyczącymi Rumunii, opracowanie podejmuje kluczowe napięcia między automatyzacją opartą na AI – taką jak analityka predykcyjna, zautomatyzowane systemy decyzyjne oraz identyfikacja biometryczna – a podstawowymi prawami do prywatności. W sytuacji, gdy Rumunia rozwija swoje ramy cyfrowego zarządzania zgodnie z wymogami UE, nierozwiązane problemy, w tym stronniczość algorytmiczna, obawy dotyczące nadzoru oraz luki w egzekwowaniu przepisów, zagrażają zarówno zaufaniu publicznemu, jak i zgodności regulacyjnej.

Analiza wskazuje decydentom politycznym i praktykom możliwe kierunki działania pozwalające pogodzić innowacje technologiczne z solidną ochroną danych osobowych poprzez stosowanie zasad privacy by design, wzmocnienie mechanizmów nadzoru oraz przejrzyste zarządzanie algorytmiczne.

Introduction

The rapid digitalization of public administration, accelerated by AI technologies, promises enhanced efficiency and citizen services but introduces profound risks to personal data protection. This paper examines these challenges in the context of Romania's ongoing governance reforms, proposing actionable solutions aligned with GDPR and the EU AI Act. It addresses key tensions between AI-driven automation, such as predictive analytics and biometric systems, and fundamental privacy rights, focusing on Romania's public sector where digital initiatives like ROeID[1] and government cloud platforms process vast citizen data. While literature examines GDPR implementation in Western European contexts, there is limited empirical research on how Central and Eastern European countries with less developed digital infrastructure and institutional capacity address the combined challenges of GDPR and AI Act implementation. Notably, systematic analysis is lacking on the interaction between legacy IT systems and emerging AI governance requirements, as well as on the strategies used by resource-constrained data protection authorities to enforce regulations in rapidly digitalizing environments. This study considers the gap by examining Romania's unique position as a nation pursuing ambitious digitalization amid notable capacity constraints.

With Romania advancing its Digital Governance Framework amid EU mandates, unresolved issues like algorithmic bias and enforcement gaps threaten trust and compliance, this analysis fills that void through targeted research questions and a case study. It offers policymakers practical pathways to balance innovation with protection[2].

The research questions of our research are:

[1] Romania's national Single Sign-On (SSO) platform for digital identity, enabling citizens to access e-government services with one set of secure credentials. Launched by the Authority for Digitalization of Romania, it supports online authentication for platforms like ghișeul.ro and is designed for EU-wide interoperability under eIDAS regulations [accessed on: 26.01.2026].

[2] Romania and the EU AI Act: Ready or Just Warming Up? https://eurocloud.org/news/article/romania-and-the-eu-ai-act-ready-or-just-warming-up/ [accessed on: 26.01.2026].

RQ1: What are the primary challenges to personal data protection in AI-driven digitalized public administration, such as bias, surveillance, and unauthorized processing?
RQ2: How does Romania's implementation of GDPR and the EU AI Act address these challenges in its public sector digitalization efforts?

AI in public administration amplifies risks like massive personal data processing without adequate safeguards, algorithmic bias affecting decisions, and surveillance via tools like facial recognition. In digitalized systems, challenges include data breaches from interconnected platforms and lack of transparency in automated decisions. Globally, regulations like GDPR and EU AI Act (effective 2024) impose strict requirements, but enforcement gaps persist[3].

Our research will follow a qualitative approach, combining EU-wide insights on AI privacy risks (*e.g.,* data breaches, bias) and Romania-focused analysis via legal documents, reports, and case studies. We realize a literature review on global/EU challenges, Romania case study, and synthesis of solutions with recommendations.

General Aspects

The analysis of personal data protection in AI-driven public administration requires a multi-layered theoretical framework that integrates legal, technological, and governance perspectives. This section establishes the conceptual foundation for our analysis by examining the intersection of data protection law, AI ethics, and public administration theory[4].

The Legal Framework: GDPR and the EU AI Act. The General Data Protection Regulation (GDPR), effective since May 2018, establishes the foundational legal framework for personal data protection within the European Union. Its core principles, lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity, and confidentiality, provide a comprehensive regime for protecting individual privacy rights. Particularly relevant to AI deployment are Articles 22 (automated decision-making), 25 (data protection by design and by default), and 35 (data protection impact assessments)[5].

---

[3] Hisham Hamed Al-Kasasbeh, Nasir Albalawee, Haneen A. Al-Khawaja, Ali Khaled Qtaishat, *Legal Challenges of Using AI and Big Data in Public Administration: Administrative Liability, Data Protection, and Public Services Efficiency* in International Journal of Sustainable Development and Planning, vol. 20, no. 6, June, 2025, pp. 2639-2650.

[4] R.M. Gellert (2015), Understanding data protection as risk regulation. *Journal of Internet Law*, *18*(11), pp. 3-15.

[5] European Data Protection Board, *Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679 (wp251rev.01)*, 2018.

The EU Artificial Intelligence Act, provisionally agreed in December 2023 and entering into force in stages from 2024-2027, complements GDPR by establishing a risk-based regulatory framework specifically for AI systems. The Act categorizes AI applications according to risk levels (unacceptable, high, limited, minimal) and imposes corresponding obligations on developers and deployers. For public administration, many AI systems fall into the high-risk category, triggering requirements for conformity assessments, risk management systems[6], data governance, transparency, human oversight, and robustness measures.

Privacy by Design and Default. Privacy by Design (PbD), conceptualized by Ann Cavoukian[7] and codified in GDPR Article 25, represents a proactive approach to privacy protection requiring that data protection measures be embedded into technology systems from their inception rather than added as afterthoughts. The seven foundational principles of PbD, proactive not reactive, privacy as default, privacy embedded into design, full functionality, end-to-end security, visibility and transparency, and respect for user privacy, provide a comprehensive framework for developing privacy-preserving AI systems in public administration.

This study employs a qualitative research methodology combining doctrinal legal analysis with empirical case study investigation. The methodological approach is designed to provide both theoretical depth and practical insights into the implementation challenges of data protection in AI-driven public administration. First, a comprehensive literature review examines academic publications, policy documents, and regulatory guidelines addressing AI deployment in public administration and personal data protection within the European Union. This review encompasses scholarly articles from legal, computer science, and public administration journals; EU legislative texts and preparatory materials; guidance documents from the European Data Protection Board (EDPB); reports from civil society organizations; and technical standards for AI governance.

Second, the study conducts doctrinal analysis of relevant legal instruments, including GDPR provisions, the EU AI Act, Romanian implementation legislation (Law 190/2018 on data protection measures), and decisions of the Romanian National Supervisory Authority for Personal Data Processing (ANSPDCP). This analysis identifies gaps, ambiguities, and potential conflicts in the regulatory framework.

Third, case study methodology examines specific AI implementations in Romanian public administration, including the ROeID digital identity system, au-

[6] European Commission, The Artificial Intelligence Act. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024.

[7] Ann Cavoukian, *Privacy by Design: The 7 Foundational Principles*, Information and Privacy Commissioner of Ontario, Canada, 2011.

tomated tax assessment systems, AI-powered social benefit allocation platforms, and biometric border control systems. Data sources include official government documentation, ANSPDCP investigation reports, media coverage, and publicly available technical specifications.

This research acknowledges several limitations. Access to detailed technical specifications of AI systems deployed in Romanian public administration is restricted due to security and commercial confidentiality concerns. The rapidly evolving nature of both AI technology and regulatory frameworks means that findings represent a snapshot of current conditions. Additionally, the study focuses primarily on legal and institutional dimensions, with limited examination of citizen perspectives and experiences.

Challenges to Personal Data Protection in AI Driven Public Administration

The deployment of AI systems in public administration creates multiple challenges to personal data protection, ranging from technical vulnerabilities to governance deficits. This section analyzes the primary challenges identified through our research.

Massive Personal Data Processing Without Adequate Safeguards. AI systems in public administration require vast quantities of personal data for training, validation, and operational deployment[8]. Unlike traditional administrative systems that process data for specific, clearly defined purposes, AI systems often aggregate data from multiple sources, creating comprehensive citizen profiles that extend far beyond the original collection purpose. In Romania, the integration of various governmental databases, like tax records, health information, social benefits data, educational records, into centralized platforms creates unprecedented concentrations of personal information.

The principle of data minimization, fundamental to GDPR, becomes particularly challenging in AI contexts. Machine learning algorithms typically perform better with larger datasets, creating institutional pressures to collect and retain maximum amounts of data rather than minimum necessary amounts. Furthermore, the lack of clear purpose limitation, AI systems may be repurposed for uses not anticipated at the time of data collection, undermines citizen control over their personal information.

[8] Tal Zarsky, The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making, *Science, Technology, & Human Values*, vol. 41, no. 1, 2016, pp. 118-132.

Algorithmic Bias and Discriminatory Decision-Making. Algorithmic bias represents one of the most pernicious threats to fundamental rights in AI-driven administration. Bias can enter AI systems through multiple pathways[9], like in historical data reflecting past discrimination, proxy variables that correlate with protected characteristics, skewed training datasets, or flawed model assumptions. In public administration contexts, where AI systems may determine access to benefits, assess tax liabilities, or make law enforcement decisions, algorithmic bias can perpetuate and amplify existing social inequalities.

Surveillance Capabilities and Biometric Systems. The deployment of AI-powered surveillance technologies, particularly facial recognition[10] and other biometric systems, raises profound privacy concerns. While the EU AI Act prohibits certain uses of real-time remote biometric identification in public spaces, exceptions exist for law enforcement purposes, and post-deployment (retrospective) facial recognition faces fewer restrictions. In Romania, biometric systems have been deployed for border control, national ID cards, and increasingly for access to government buildings and services.

Lack of Transparency and Explainability. Many AI systems, particularly those based on deep learning neural networks, function as 'black boxes'[11] where the reasoning process behind specific decisions remains opaque even to system developers. This opacity conflicts with GDPR's transparency requirements and citizens' rights to understand decisions affecting them. In administrative contexts, where decisions may have significant consequences, as denial of benefits, tax assessments, eligibility determinations, the inability to explain algorithmic reasoning undermines both accountability and the right to effective remedy. Romania's public administration faces particular challenges in achieving algorithmic transparency. Many AI systems are procured from private vendors under contracts that protect proprietary algorithms as trade secrets. Public officials often lack the technical expertise to evaluate or explain AI system operations. Citizens have limited means to challenge automated decisions, and judicial review mechanisms have not yet adapted to address algorithmic decision-making.

Data Breaches and Cybersecurity Vulnerabilities. The interconnected nature of digital government platforms creates significant cybersecurity risks. Centralized

---

[9] Andrew D. Selbst et al., Fairness and abstraction in sociotechnical systems, in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 59-68.

[10] Karen Yeung, Algorithmic Regulation: A Critical Interrogation (May 23, 2017). TLI Think! Paper 62/2017, Regulation & Governance, Forthcoming, King's College London Law School Research Paper no. 2017-27, Available at SSRN: https://ssrn.com/abstract=2972505.

[11] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, Why a right to explanation of automated decision-making does not exist in the general data protection regulation, *International Data Privacy Law*, vol. 7, no. 2, 2017, pp. 76-99.

databases containing comprehensive citizen information present attractive targets for cyberattacks. Romania[12] has experienced multiple data breaches affecting government systems, exposing personal information of millions of citizens. AI systems introduce additional vulnerabilities through their complexity, dependence on cloud infrastructure, and potential for adversarial attacks that manipulate model behavior. The concentration of personal data in AI training datasets means that a single breach can expose vast quantities of sensitive information.

Romania's Implementation of GDPR and the EU AI Act in Public Sector Digitalization

Romania has invested heavily in digitalizing public services[13] (health, education, administration) under its National Recovery and Resilience Plan (PNRR)[14].

As an EU member it fully applies the GDPR (Regulation (EU) 2016/679), implemented by national Law 190/2018 and is beginning to implement the EU AI Act. This report examines how GDPR compliance is carried out in Romanian healthcare, education, and public administration, outlines national measures (existing or proposed) for the EU AI Act, surveys examples of AI/digital tools in public services and their legal treatment and discusses challenges and impacts on citizens' privacy, trust, and administrative efficiency.

GDPR Compliance in Public Administration. Romanian public bodies, from central ministries to city halls, are explicitly designated in national law as "public authorities" subject to GDPR (Law 190/2018 defines processing for public-interest tasks)[15]. In practice, all government agencies must have DPOs, register data flows, and follow ANSPDCP directions. For example, public-sector GDPR compliance often involves frequent audits by ANSPDCP, if violations are found, public entities receive remediation plans with strict timelines (*e.g.,* 90 days to remedy)[16]. Breaches in the public sector (*e.g.,* leaked employment or tax records) must be reported to ANSPDCP within 72 hours, and indeed the Authority regularly publishes sanctions against state bodies that neglect privacy.

---

[12] National Supervisory Authority for Personal Data Processing, Annual Activity Report 2022, Bucharest, 2023.

[13] Maria Albu, Local Administration in The Era of Digitalization: From E-Government to Smart Governance. Case Study: Alba Iulia Smart City, in Annales Universitatis Apulensis – Series Jurisprudența, nr.28/2025, pp. 9-19.

[14] Romania Country Commercial Guide at https://www.trade.gov/country-commercial-guides/romania-digital-economy [accessed on: 05.02.2026].

[15] ROMANIA Data Protection Laws of the World https://www.dlapiperdataprotection.com/index.html?t=law&c=RO [accessed on: 05.02.2026].

[16] *Ibidem*.

Compliance in government IT projects is also emphasized. Large digitalization projects (like new e-Government portals and the national government cloud) include GDPR assessments. A U.S. trade report notes that as an EU member, "Romania adheres to GDPR, which imposes strict data privacy and protection requirements" and that aligning systems with these rules often requires significant investment in security and legal expertise. In recent years, key initiatives, such as a secure government private cloud (launched 2025), explicitly build in privacy and security audits[17]. Nonetheless, legacy issues (fragmented databases, varying technical capacity across ministries) mean ANSPDCP still finds gaps. As one summary notes, Romania's regulators maintain "high standards" and require extensive privacy measures for public institutions[18] (including employee monitoring and CCTV rules).

EU AI Act: National Strategy and Legislative Measures. With the EU AI Act now in force, Romania is drafting national measures to implement it. In July 2024, the government approved a National AI Strategy 2024–2027, which serves as a roadmap for AI adoption and governance. The Strategy proposes creating a dedicated authority under the Digitalization Agency (ADR) to oversee AI implementation, and calls for national guidelines on safe AI use[19]. Separately, Parliament is considering draft laws on AI. One proposal (PL-x 184/2025) would establish substantive rules for "responsible AI" in both private and public spheres. Key provisions under discussion include in public administration, any AI system used in decision-making must be transparent and subject to regular independent audits, fully automated decisions (without human oversight) would be prohibited, in judiciary and law enforcement, AI could not be the sole basis for judicial rulings or enforcement decisions, in healthcare, AI tools for diagnostics/treatment would need scientific validation and official approval *etc.* As sanctions we found abusive or illegal AI use (*e.g.,* violating these rules) would carry heavy penalties, proposed fines range up to €20 million and even criminal sanctions[20].

These proposals mirror EU AI Act obligations but go further in places (*e.g.,* criminal penalties). They have raised legal debates, as experts note some draft provisions may overlap or conflict with the EU Act's regime. Officially, Romania

[17] Romania Digital Transformation: Launch of Major Procurement for Government Private Cloud Migration https://www.trade.gov/market-intelligence/romania-digital-transformation-launch-major-procurement-government-private [accessed on: 05.02.2026].

[18] Data Privacy Management in Romania - Compliance & Law, https://responsum.eu/privacy-management-in-romania/ [accessed on: 05.02.2026].

[19] Romania and the EU AI Act: Ready or Just Warming Up? | International Network of Privacy Law Professionals https://inplp.com/latest-news/article/romania-and-the-eu-ai-act-ready-or-just-warming-up/ [accessed on: 05.02.2026].

[20] *Ibidem.*

has not yet designated its national authorities under the AI Act as required. To comply, Member States must name a "market surveillance" body, a "notification" authority, and a fundamental rights authority. Romania has notified the EU that agencies like the Ombudsman and Data Protection Authority will monitor high-risk AI systems[21], but as of late 2025, no single national AI supervisory body has been formally established, despite an EU deadline of Aug 2025. In short, Romania's approach is evolving, the government has set strategy and is working on laws, but institutional and regulatory steps remain to be finalized.

As a study case, several high-profile digital services illustrate AI and data use in government. In 2024-2025, Romania launched its Government Private Cloud project, migrating dozens of state IT systems into a secure centralized cloud[22]. This aims to improve efficiency, resilience and security of citizen services (tax filing, business registrations, *etc.*) while ensuring personal data are better protected. Another example is the ongoing rollout of a digital ID (eID) for citizens, which will replace paper documents with secure electronic identification. By law this system is built with GDPR safeguards (data encryption, user consent). On the AI front, McKinsey reports that over half of Romanian civil servants have already begun using AI tools (*e.g.,* automated text summarizers, translation, chatbot assistants) in their daily work[23]. Some municipalities and agencies are piloting chatbots for citizen inquiries, always with disclaimers that personal requests still go to human agents. In administrative law enforcement and benefits, no fully automated AI decisions exist, cases must be reviewed by officials. These steps reflect how GDPR and draft AI laws influence practice like any use of AI by a public agency is planned with a DPIA and human supervision, and citizens are informed when an AI tool (like a virtual assistant) is involved.

Key Challenges in Integrating Regulations. Romania's digital transformation faces several hurdles in applying GDPR and AI rules, like Technical and Infrastructure Gaps. Romania ranks low on the EU Digital Economy and Society Index, indicating shortfalls in broadband coverage and IT infrastructure[24]. Government reports and OECD reviews note that limited internet in rural areas and outdated legacy

---

[21] *Ibidem*.

[22] Romania Digital Transformation: Launch of Major Procurement for Government Private Cloud Migration https://www.trade.gov/market-intelligence/romania-digital-transformation-launch-major-procurement-government-private [accessed on: 05.02.2026]

[23] Martin H. Thelle et al., The opportunity of AI in public administration in Romania, Capturing the benefits of AI in public administration, *An implement Consulting Group by Google*, October 2024.

[24] Cristian Constantin FRANCU, Ștefan SĂVA, Global Perspectives on Digital and AI Legislation: A Comparative Study of Data Protection, AI Governance, and Healthcare Innovations with a Focus on Romania, DOI: 10.2478/picbe-2025-0191, *Proceedings of the 19th International Conference on Business Excellence 2025*, Sciendo, pp. 2469-2481.

systems hinder seamless data sharing and secure cloud adoption[25]. Public sector IT projects often suffer delays, as slower connectivity and old hardware complicate GDPR compliance (*e.g.,* harder to encrypt or update aging servers).

Widespread lack of digital literacy impedes regulation. For example, fewer than 25% of Romanians use any e-government service (EU average is 75%)[26]. In schools and local offices, many staff members still have only basic IT skills, so GDPR concepts (data subject rights, DPIAs) are not fully understood. Studies report that even after GDPR became law, many institutions did not implement all requirements[27]. Similarly, public servants may lack training on AI/ML risks. The OECD and EU note persistent deficits in Romania's "digital skills" and the need for user-centered design[28].

Fragmentation of rules creates uncertainty. Romania's own AI draft law and the EU AI Act cover similar ground, experts warn some national provisions (like criminal sanctions) may conflict with EU rules. At present, Romania's AI Act implementation relies on existing authorities (*e.g.,* ANSPDCP for data rights, Ombudsman for discrimination), but the formal designation of AI supervisory bodies was delayed[29]. This transitional gap makes compliance responsibilities unclear for AI developers and public agencies. Likewise, new local laws on topics like telework or e-commerce must be carefully aligned with GDPR, requiring constant legal review.

Public trust is a challenge. Surveys show Romanian citizens are very cautious about data-driven tools. About 70% are "very" or "moderately" concerned about how AI accesses their data[30]. Most insist on human oversight (*e.g.,* over 60% prefer to interact with a real person rather than a chatbot for government services). These attitudes pressure officials to proceed carefully. Mistrust also applies to public-sector

[25] AI in Romania: The potential for the public sector | McKinsey, https://www.mckinsey.com/industries/public-sector/our-insights/the-transformative-potential-of-ai-in-romanias-public-sector [accessed on: 05.02.2026].

[26] *Ibidem* [accessed on: 08.02.2026].

[27] Anisoara Belbe, Implementation and Application of GDPR In Romanian Educational Institutions, in *The Annals of the University of Oradea. Economic Sciences* TOM XXXI, 1st Issue, July 2022, pp. 160-171.

[28] AI in Romania: The potential for the public sector | McKinsey, https://www.mckinsey.com/industries/public-sector/our-insights/the-transformative-potential-of-ai-in-romanias-public-sector [accessed on: 08.02.2026].

[29] Romania and the EU AI Act: Ready or Just Warming Up? | International Network of Privacy Law Professionals https://inplp.com/latest-news/article/romania-and-the-eu-ai-act-ready-or-just-warming-up/ [accessed on: 08.02.2026].

[30] Study: Romanians do not trust AI without human intervention, https://www.romaniajournal.ro/business/study-romanians-do-not-trust-ai-without-human-intervention-seven-out-of-ten-prioritise-data-protection/ [accessed on: 08.02.2026].

data use: any news of a data leak (like a hospital breach) erodes confidence that GDPR is enforced.

Ensuring compliance can be costly. Businesses and agencies must invest in new systems, security audits, and legal staff. A U.S. report notes GDPR compliance "requires significant investment" in data systems and monitoring[31]. Romania's public budgets are constrained, so modernization projects (cloud migration, smart classrooms) sometimes run behind schedule. Limited funding can leave smaller agencies struggling to install recommended encryption or hire qualified DPOs.

In summary, Romania must continue bridging gaps in technology, skills, and legal alignment to fully integrate GDPR and the AI Act into its digital agenda.

Impacts on Citizens: Privacy, Trust, and Efficiency. The interplay of GDPR and AI rules shapes Romanian citizens' experiences in three key ways, like Data Privacy and Rights**,** citizens benefit from strong legal guarantees. Under GDPR, individuals can access and delete their data held by any public agency, and special rules protect sensitive health or student records. For instance, when using telemedicine or school portals, citizens must give explicit consent for data processing, and agencies publish transparency notices online. Trust in AI and Services**,** Romanian citizens are wary but pragmatic. Recent surveys[32] indicate a strong demand that any AI in healthcare, education, or administration be "strictly governed by law" with human oversight. Citizens recognize the benefits of digitalization on the other hand, they appreciate that online tax filing or e-health portals can save time. Importantly, trust hinges on transparency, officials communicate how personal data are used and allow complaint channels. Effective oversight (*e.g.,* ANSPDCP's high-profile actions) helps build confidence that violations have consequences. Administrative efficiency**,** digital tools promise and deliver faster services. Under GDPR and AI rules, projects must meet privacy standards, but when designed properly they greatly improve efficiency. In practice, citizens experience this as shorter wait times and one-stop services. For instance, the rollout of digital ID and e-signatures has enabled fully online applications for permits and benefits. Over time, harmonizing GDPR and AI Act compliance is expected to enhance service quality by ensuring systems are robust: a well-regulated, secure digital infrastructure both protects citizens and makes agencies more responsive.

In conclusion, Romania's public sector is rapidly evolving its digital landscape under the twin imperatives of privacy and innovation. GDPR has been embedded

[31] Romania – Digital Economy, https://www.trade.gov/country-commercial-guides/romania-digital-economy [accessed on: 08.02.2026].

[32] Study: Romanians do not trust AI without human intervention, https://www.romaniajournal.ro/business/study-romanians-do-not-trust-ai-without-human-intervention-seven-out-of-ten-prioritise-data-protection/ [accessed on: 08.02.2026].

via national laws, DPOs, and sector-specific rules in healthcare, education, and administration. The AI Act's requirements are being addressed through a national AI strategy, draft legislation, and ongoing institutional reforms. Early case studies, from e-health records to AI chatbots, illustrate both the promise and complexity of these changes. Key challenges remain in skills, infrastructure, and legal clarity. As Romania continues its digital transformation, the impact on citizens will depend on maintaining trust and rights while harnessing technology for better, faster public services.

General Conclusions

The rapid digitalization of Romania's public administration, driven by the National Recovery and Resilience Plan (PNRR), presents a fundamental paradox. While AI and digital tools (such as the Government Private Cloud and ROeID) offer significant improvements in administrative efficiency and citizen services, they introduce profound risks to fundamental privacy rights. The concentration of vast amounts of personal data in centralized systems creates unprecedented vulnerabilities regarding data breaches and unauthorized processing.

Romania occupies a complex transitional space regarding regulatory compliance, like the General Data Protection Regulation is well-established within the public sector. Public authorities are explicitly subject to the law, mandated to appoint Data Protection Officers (DPOs), and face active auditing and sanctions from the National Supervisory Authority (ANSPDCP) for non-compliance. Implementation of the EU AI Act is still evolving. While the government has approved a National AI Strategy (2024–2027) and is drafting legislation that may impose stricter penalties than the EU itself, formal institutional structures, such as designated supervisory bodies, were delayed beyond EU deadlines.

Successful governance is hindered by outdated legacy IT systems, limited rural connectivity, and a significant lack of digital literacy among public servants, which complicates the practical application of privacy laws. Sustainable digital governance requires moving beyond reactive compliance to a proactive *Privacy by Design* approach. To maintain public trust, AI systems must ensure transparency and retain human oversight for critical administrative decisions.

Based on the limitations and gaps identified in the text, here are concrete proposals for future research: investigate the direct correlation between digital literacy levels and citizens' trust in automated government decisions (*e.g.,* tax assessments), moving beyond legal analysis to capture the user experience, conduct a multi-year study tracking the practical enforcement of Romania's upcoming AI laws (2024-2027) to determine if strict proposed penalties (up to €20M) are effectively applied

or remain symbolic, compare Romania's digital governance with other EU nations facing similar rural connectivity gaps to identify best practices for maintaining data privacy despite aging legacy infrastructure.

References

AI in Romania: The potential for the public sector | McKinsey, https://www.mckinsey.com/industries/public-sector/our-insights/the-transformative-potential-of-ai-in-romanias-public-sector [accessed on: 05.02.2026].

Al-Kasasbeh Hisham Hamed, Albalawee Nasir, Al-Khawaja Haneen A., Qtaishat Ali Khaled, *Legal Challenges of Using AI and Big Data in Public Administration: Administrative Liability, Data Protection, and Public Services Efficiency* in International Journal of Sustainable Development and Planning, vol. 20, no. 6, June, 2025, pp. 2639-2650.

Belbe Anisoara, Implementation and Application of GDPR In Romanian Educational Institutions, in *The Annals of the University of Oradea. Economic Sciences* TOM XXXI, 1st Issue, July 2022, pp. 160-171.

Cavoukian Ann, *Privacy by Design: The 7 Foundational Principles*, Information and Privacy Commissioner of Ontario, Canada2011.

Data Privacy Management in Romania – Compliance & Law, https://responsum.eu/privacy-management-in-romania/ [accessed on: 05.02.2026].

European Commission, The Artificial Intelligence Act. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024.

European Data Protection Board, *Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679 (wp251rev.01)*, 2018.

Francu Cristian Constantin, Sava Ștefan, Global Perspectives on Digital and AI Legislation: A Comparative Study of Data Protection, AI Governance, and Healthcare Innovations with a Focus on Romania, DOI: 10.2478/picbe-2025-0191, *Proceedings of the 19th International Conference on Business Excellence 2025*, Sciendo, pp. 2469-2481.

Gellert, R. M. (2015). Understanding data protection as risk regulation. *Journal of Internet Law*, *18*(11), 3-15.

National Supervisory Authority for Personal Data Processing, Annual Activity Report 2022, Bucharest, 2023.

Romania and the EU AI Act: Ready or Just Warming Up? https://eurocloud.org/news/article/romania-and-the-eu-ai-act-ready-or-just-warming-up/ [accessed on: 26.01.2026].

Romania Data Protection Laws of the World https://www.dlapiperdataprotection.com/index.html?t=law&c=RO [accessed on: 05.02.2026].

Romania Digital Transformation: Launch of Major Procurement for Government Private Cloud Migration

https://www.trade.gov/market-intelligence/romania-digital-transformation-launch-major-procurement-government-private [accessed on: 05.02.2026].

Romania – Digital Economy, Romania Country Commercial Guide at https://www.trade.gov/country-commercial-guides/romania-digital-economy [accessed on: 05.02.2026].

Selbst Andrew D. et al., Fairness and abstraction in sociotechnical systems, în *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 59-68.

Study: Romanians do not trust AI without human intervention, https://www.romaniajournal.ro/business/study-romanians-do-not-trust-ai-without-human-intervention-seven-out-of-ten-prioritise-data-protection/ [accessed on: 08.02.2026].

Thelle Martin H. et al., The opportunity of AI in public administration in Romania, Capturing the benefits of AI in public administration, *An implement Consulting Group by Google*, October 2024.

Wachter Sandra, Brent Mittelstadt, Luciano Floridi, Why a right to explanation of automated decision-making does not exist in the general data protection regulation, *International Data Privacy Law*, vol. 7, no 2, 2017, pp. 76-99.

Yeung Karen, Algorithmic Regulation: A Critical Interrogation (May 23, 2017). TLI Think! Paper 62/2017, Regulation & Governance, Forthcoming, King's College London Law School Research Paper No. 2017-27, Available at: SSRN: https://ssrn.com/abstract=2972505 [accessed on: 25.01.2026].

Zarsky Tal, The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making, *Science, Technology, & Human Values*, vol. 41, no 1, 2016, pp. 118-132.