Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International Law and European Law

HUMAN RIGHTS AND FREEDOMS UNDER PRESSURE FROM ALGORITHMS: INTERNATIONAL AND NATIONAL AI CHALLENGES

Peer-reviewed proceedings of an international scientific conference

Lucia Bakošová (ed.)

Košice 2026

This publication was prepared as part of the research project "Technological Progress and International Human Rights Law" (No. VVGS-2023-2947), implemented within the Internal Scientific Grant System of Pavol Jozef Šafárik University in Košice. Funded by the European Union under the EU NextGenerationEU initiative through the Recovery and Resilience Plan for Slovakia, Project No. 09I03-03-V05-00008.

BAKOŠOVÁ, L.(ed.): Human Rights and Freedoms under Pressure from Algorithms: International and National AI Challenges, Košice: Vydavateľstvo ŠafárikPress, 2026, 333 p., ISBN 978-80-574-0541-2.

Human Rights and Freedoms under Pressure from Algorithms: International and National AI Challenges
Peer-reviewed proceedings of an international scientific conference

Editor:
Mgr. Lucia Bakošová, PhD.
Institute of International Law and European Law, Faculty of Law, University of Pavol Jozef Šafárik University in Košice

Reviewers:
doc. JUDr. Daniel Bednár, Ph.D.
Department of International Law and European Law, Faculty of Law, Trnava University in Trnava

doc. PhDr. JUDr. Lilla Garayová, PhD., LL.M.
Institute of International Law and European Law, Faculty of Law, Pan-European University

This publication is published by Pavol Jozef Šafárik University in Košice under a Creative Commons license (CC) CC BY-NC-SA Creative Commons Attribution-NonCommercial-ShareAlike 4.0. This license allows the work to be freely shared and modified only with attribution to the author, without commercial use, and on the condition that all derivative works are distributed under the same license.

The author is responsible for the content and language editing of this publication. This work was produced at the Faculty of Law, Pavol Jozef Šafárik University in Košice.

Available at: unibook.upjs.sk
Publication date: 31.07.2026

DOI: https://doi.org/10.33542/HRF-0541-2
ISBN 978-80-574-0541-2 (e-publication)

CONTENT

Preface 8

Predhovor 10

1 Is Sharing Really Caring? Legislative Response to Sharenting in the EU Digital Legislation 12
Lucia Rybanová
1.1 Sharenting 14
1.2 Digital Services Act and Sharenting 15
1.3 Sharenting in Risk Assessments 18
1.4 Sharenting in France as a Model Law 20

2 Right to a Fair Trial: Can AI Replace Human Judges? European Perspective 25
Magdalena Matusiak-Frącczak
2.1 Right To a Fair Trial: European Standard 27
2.2 AI As a Judge 33
2.2.1 Tribunal Established by Law 34
2.2.2 Independence and Impartiality 34
2.2.3 Legal Certainty 36
2.2.4 Public Hearing 38
2.2.5 Evidence Analysis 39
2.2.6 Hallucinations 40

3 AI-Powered Autonomous Weapons and the Protection of the Right to Life under International Law 49
Ayla Kaya Çalışkan
3.1 Autonomous Weapon Systems (AWS) and the Integration of Artificial Intelligence (AI) 51
3.2 The Protection of the Right to Life under International Law 53
3.2.1 The Right to Life under IHRL 54
3.2.2 The Right to Life under IHL 55
3.2.3 The Relationship Between IHRL and IHL 55
3.2.4 Implications for Autonomous and AI-powered Weapon Systems 56
3.3 AI-powered Autonomous Weapon Systems and Violations of the Right to Life 56
3.4 Accountability for Right-to-Life Violations 58

4 AI-Assisted Cyberattacks on Privacy: Challenges for Private International Law 65
Ľubica Gregová Širicová
4.1 The Nature of AI-Assisted Cyberattacks and Privacy Harms 67
4.1.1 Automation and Scalability 67
4.1.2 Anonymisation and Decentralisation 67
4.1.3 AI-Enhanced Personalisation 67
4.1.4 Simultaneous Dissemination and Chain Harms 68

4.1.5 Consequence for Jurisdictional Rules PIL ... 68
4.2 Jurisdictional Challenges under Brussels Ia Regulation in Personality Rights Claims ... 68
4.2.1 Defendant Domiciled in a Member State ... 69
4.2.2 Defendant Domiciled in a Third State ... 73
4.3 Analysis of the provisions of the Slovak PIL Act relevant to jurisdiction in cases of damage caused by infringements of personality rights ... 73
4.4 Jurisdictional Criteria *De Lege Ferenda* ... 76

5 AI in the Workplace and its Impact on Employees' Rights: Historical perspectives, applications of AI, and liability 84
Milica Ilić
5.1 Historical Background of AI ... 85
5.2 Legal Sources, Definitions, and Examples ... 88
5.2.1 Definitions ... 88
5.2.2 Legal Sources ... 90
5.3 AI Application in the Working Environment ... 92
5.4 AI-related Legal Risks ... 95
5.5 Civil Liability in Automated Decion-Making ... 98

6 When Values End at the Border: EU Surveillance Exports and Human Rights Risks 104
Hanna Pishchyk
6.1 Conceptual Framework and Methodology ... 106
6.1.1 Conceptual Frameworks ... 106
6.1.2 Methodology ... 107
6.2 The Evolution of EU Values and Their Impact on Policy and Regulation ... 107
6.3 EU Export Controls on Surveillance Technologies ... 109
6.4 Surveillance Tech Exports to Morocco: Case Study ... 112
6.4.1 Use of Surveillance Technologies in Morocco ... 112
6.4.2 Structural Gaps in EU Export Controls Revealed by the Moroccan Case ... 115
6.5 Critical Reforms for Effective Regulation ... 116

7 New Human Rights Bodies in the Age of Artificial Intelligence 122
Lucia Bakošová
7.1 Why Artificial Intelligence Challenges Traditional Human-Rights Institutions ... 125
7.2 Global Institutional Developments on Artificial Intelligence and Human Rights ... 127
7.2.1 Bodies Created under the Global Digital Compact ... 131
7.2.2 Proposed New Human-Rights-Focused Bodies within the UN System ... 133
7.3 Regional Human-Rights Bodies and Their Approaches to AI ... 135

7.3.1 Council of Europe ... 136
7.3.2 The European Union ... 138
7.4 Do We Need New Human-Rights Bodies? ... 144

8 Automatizované rozhodování a profilování v pracovněprávních vztazích 155
Jakub Morávek
8.1 Automatizovné rozhodování, včetně profilování v pracovněprávních vztazích ... 157
8.2 Právní úprava – minulost, přítomnost, budoucnost ... 160
8.3 Problematické momenty, závěrečné poznámky a úvahy ... 163

9 Asistencia umelej inteligencie pri excesívnej aktivite verejnej správy so zásahom do ľudských práv 168
Tibor Seman
9.1 Význam umelej inteligencie vo verejnej správe ... 170
9.1.1 Podstata umelej inteligencie ... 170
9.1.2 Využitie umelej inteligencie vo verejnej správe ... 171
9.2 Hranica medzi využitím a zneužitím umelej inteligencie alias kde sú ľudské práva ... 172
9.2.1 Prípad S. a Marper proti Spojenému kráľovstvu [GC], 2008 ... 173
9.2.2 Prípad Gaughran proti Spojenému kráľovstvu, 2020 ... 176
9.2.3 Prípad P.N. proti Nemecku, 2020 ... 182
9.2.4 Umelá inteligencia vo vzťahu k uchovávaným osobným údajom ... 184
9.3 Námety de lege ferenda ... 186
9.3.1 Všeobecné námety de lege ferenda ... 186
9.3.2 Konkrétne námety de lege ferenda ... 188

10 Přezkoumatelnost jako lidské právo a algoritmické rozhodování 195
Ondřej Lipka
10.1 Lidskoprávní rozměr přezkoumatelnosti ... 196
10.1.1 Nepřezkoumatelnost v judikatuře správních soudů ... 197
10.2 Algoritmické rozhodování ... 199
10.2.1 Problematika black box rozhodování ... 199
10.2.2 Velké uvažující modely ... 200
10.3 Deference jako krajní meze lidského prvku nepřezkoumatelnosti .. 201
10.3.1 Obecně k deferenčnímu přezkumu ... 201
10.3.2 Přězkoumatelnost jako výsledek lidského snažení ... 203

11 Využitie AI pri dokumentovaní vojnových zločinov a porušení medzinárodného humanitárneho práva 207
Bartolomej Mižák
11.1 AI v dobe moderných konfliktov ... 207
11.2 Vyšetrovanie pomocou AI ... 209
11.3 Analýza obsahu pomocou AI ... 210
11.4 Výzvy ... 212

12 Algoritmy ako prostriedok hybridných vojen zacielených na ovplyvňovanie volieb v cudzích štátoch 217

Ján Svák

12.1 Priebeh hybridnej vojny Ruska proti USA, Spojenému kráľovstvu a Rumunsku s cieľom ovplyvniť voľby algoritmami 218

12.2 Medzinárodnoprávna ochrana pred algoritmami zasahujúcimi volebné právo 225

13 Proaktívne zverejňovanie informácií súdmi ako nástroj predchádzania vzniku právnych halucinácií 242

Emma Humeňanská

13.1 Vplyv umelej inteligencie na premenu práva na informácie 243

13.2 Problém právnych halucinácií 246

13.3 Komunikácia súdov ako protiváha informáciám generovaných AI? 249

14 Zodpovednosť za škodu spôsobenú AI: Potreba nového právneho rámca? 256

Natália Kobulská

14.1 Európska legislatíva 257

14.2 Umelá inteligencia ako chybný výrobok 258

14.2.1 Smernica o zodpovednosti za škodu spôsobenú AI 260

14.2.2 Porovnanie EÚ úpravy 261

14.3 Technické aspekty AI a ich vplyv na zodpovednosť za škodu 262

14.4 Návrh právnej úpravy od budúcna 263

15 Digitálny konštitucionalizmus: cesta k rešpektovaniu ľudských práv v dobe algoritmov? 269

Soňa Juhásová

15.1 Teoretické východiská digitálneho konštitucionalizmu – ako algoritmy rozhodujú o ľudských právach? 271

15.2 Európsky prehľad: od digitálneho divokého západu k digitálnemu konštitucionalizmu 273

15.2.1 Prvá fáza: Digitálny liberalizmus 273

15.2.2 Druhá fáza: Súdny aktivizmus 274

15.2.3 Tretia fáza: Digitálny konštitucionalizmus 276

15.3 Meta Oversight Board: digitálny konštitucionalizmus v praxi 278

15.4 Kritické pohľady 280

16 Čítanie mysle, prepisovanie pravidiel? Ľudské práva v ére neurotechnológií 289

Laura Gazdagová

16.1 Integrácia neurotechnológií a umelej inteligencie v modernej medicíne a technike 290

16.1.1 Krátky exkurz do regulácie umelej inteligencie v EÚ a jej dôsledky pre neurotechnológie ... 292
16.2 Ľudsko-právne aspekty využívania moderných neurotechnológií ... 295
16.2.1 Potreba diskusií o nových „neuro“ - právach? ... 297

17 Žurnalistický prejav v ére sociálnych sietí a AI – úsvit alebo súmrak slobody prejavu? 306
Jakub Vanacký
17.1 Tradičné média v kontexte nových technológii ... 308
17.1.1 O povahe žurnalistiky ... 310
17.1.2 Zmeny v mediálnom prostredí v kontexte základných práv a slobôd ... 313
17.1.3 Nové technológie v informačnom priestore ... 315
17.2 K diskusii o potrebe regulácie informačného prostredia ... 318
17.2.1 Etický kódex novinára ... 321
17.2.2 Nejasnosti v kontexte s právom na pravdivé informácie ... 324

Preface

The present volume is the outcome of the international scientific conference "Human Rights and Freedoms under Pressure from Algorithms: International and National AI Challenges", which took place on 28 November 2025 at the Faculty of Law of Pavol Jozef Šafárik University in Košice and online. The conference was organised within the framework of the research project "Technological Progress and International Human Rights Law" *(No. VVGS-2023-2947)*, implemented within the Internal Scientific Grant System of Pavol Jozef Šafárik University in Košice. The research was supported by the European Union under the EU NextGenerationEU initiative through the Recovery and Resilience Plan for Slovakia, Project No. 09I03-03-V05-00008.

The rapid development of artificial intelligence and its increasing integration into various areas of social life raise a wide range of new legal questions. Algorithmic systems increasingly influence decision-making processes in fields that directly affect the exercise and protection of fundamental rights and freedoms. Consequently, the question of whether existing legal frameworks can respond adequately to these technological transformations has become one of the central topics of contemporary legal scholarship.

The aim of the conference was to provide an interdisciplinary platform for scholarly discussion on the challenges that artificial intelligence poses to the protection of human rights at both international and national levels. Attention was devoted to issues such as accountability, transparency of algorithmic decision-making, protection of privacy, prevention of discrimination, and the availability of effective legal remedies in an increasingly digitalised environment.

The conference brought together experts from various academic and institutional backgrounds. It was a particular honour to welcome the keynote speakers. Prof. Dr. Vasilka Sancin, PhD., Judge at the European Court of Human Rights, delivered a keynote lecture entitled "The Role of AI Sandboxes before International Human Rights Bodies and Courts", addressing the role of regulatory experimentation frameworks in the context of artificial intelligence and human rights protection. The second keynote speaker was prof. JUDr. Ján Klučka, CSc., professor emeritus at the Faculty of Law of Pavol Jozef Šafárik University in Košice, whose presentation "Ethical Rules and Artificial Intelligence" reflected on the ethical dimensions of artificial intelligence and their significance for the legal regulation of emerging technologies.

This volume presents a selection of papers delivered at the conference. Not all contributions presented during the event are included in the proceedings, as some authors did not express an interest in publishing their papers afterwards.

The contributions included in this volume have undergone peer review and reflect diverse perspectives on the relationship between artificial intelligence and the protection of fundamental rights and freedoms.

It is hoped that this collection will contribute to the further development of scholarly debate on the legal implications of technological progress and offer readers valuable reflections on how law can respond to the challenges brought by the age of algorithms.

Lucia Bakošová

Predhovor

Predkladaný zborník je výsledkom medzinárodnej vedeckej konferencie „Ľudské práva a slobody pod tlakom algoritmov: Medzinárodné a vnútroštátne výzvy AI“, ktorá sa uskutočnila 28. novembra 2025 na Právnickej fakulte Univerzity Pavla Jozefa Šafárika v Košiciach a online. Konferencia bola organizovaná v rámci vedecko-výskumného projektu „Technologický pokrok a medzinárodné právo ľudských práv“ (č. VVGS-2023-2947), realizovaného v rámci Vnútorného vedeckého grantového systému Univerzity Pavla Jozefa Šafárika v Košiciach. Výskum bol podporený Európskou úniou v rámci iniciatívy EU NextGenerationEU prostredníctvom Plánu obnovy a odolnosti Slovenskej republiky, projekt č. 09I03-03-V05-00008.

Rýchly rozvoj umelej inteligencie a jej čoraz širšia integrácia do rôznych oblastí spoločenského života vyvolávajú široké spektrum nových právnych otázok. Algoritmické systémy čoraz viac ovplyvňujú rozhodovacie procesy v oblastiach, ktoré majú priamy dopad na výkon a ochranu základných práv a slobôd. Otázka, do akej miery sú existujúce právne rámce schopné adekvátne reagovať na tieto technologické transformácie, sa preto stala jednou z ústredných tém súčasného právneho výskumu.

Cieľom konferencie bolo vytvoriť interdisciplinárnu platformu pre odbornú diskusiu o výzvach, ktoré umelá inteligencia prináša pre ochranu ľudských práv na medzinárodnej aj vnútroštátnej úrovni. Osobitná pozornosť bola venovaná otázkam zodpovednosti, transparentnosti algoritmického rozhodovania, ochrany súkromia, predchádzania diskriminácii, ako aj dostupnosti účinných právnych prostriedkov ochrany v čoraz viac digitalizovanom prostredí.

Konferencia spojila odborníkov z rôznych akademických a inštitucionálnych prostredí. Osobitnou cťou bolo vystúpenie hlavných rečníkov. Prof. Dr. Vasilka Sancin, PhD., sudkyňa Európskeho súdu pre ľudské práva, vystúpila s hlavným referátom s názvom „The Role of AI Sandboxes before International Human Rights Bodies and Courts“, v ktorom sa venovala úlohe regulačných experimentálnych prostredí v kontexte umelej inteligencie a ochrany ľudských práv. Druhým hlavným rečníkom bol prof. JUDr. Ján Klučka, CSc., emeritný profesor Právnickej fakulty Univerzity Pavla Jozefa Šafárika v Košiciach, ktorého príspevok „Etické pravidlá a umelá inteligencia“ reflektoval etické dimenzie umelej inteligencie a ich význam pre právnu reguláciu nových technológií.

Tento zborník obsahuje výber príspevkov prezentovaných na konferencii. Nie všetky príspevky, ktoré na podujatí zazneli, sú však v zborníku zahrnuté, keďže nie všetci autori prejavili záujem o ich následné publikovanie. Zaradené

príspevky prešli recenzným konaním a reflektujú rôzne perspektívy na vzťah medzi umelou inteligenciou a ochranou základných práv a slobôd.

Tento zborník má ambíciu prispieť k ďalšiemu rozvoju vedeckej diskusie o právnych dôsledkoch technologického pokroku a ponúknuť čitateľom podnetné reflexie o tom, ako môže právo reagovať na výzvy, ktoré prináša éra algoritmov.

Lucia Bakošová

1 Is Sharing Really Caring? Legislative Response to Sharenting in the EU Digital Legislation

Lucia Rybanová

https://doi.org/10.33542/HRF-0541-2-01

Abstract

Sharenting has become a very widespread practice among parents. It consists of sharing personal information, images and videos of their children online, particularly on social media platforms. While often well-intentioned, this phenomenon raises significant concerns regarding children's privacy and safety, including various risks. Despite its growing prevalence, sharenting remains a relatively recent concept, and digital legislation has yet to fully address its implications. This paper examines the response of European Union digital legislation to sharenting, with particular emphasis on the Digital Services Act. It explores how the Digital Services Act engages with the practice, focusing on the risk assessments conducted by selected Very Large Online Platforms under Article 34. More specifically, we aim to analyse whether these platforms acknowledge sharenting as a distinct issue and what measures, if any, they have implemented in response. Additionally, the paper will consider the French Act. No. 2024-120, amending the French Civil Code, and how it aims to safeguard children from sharenting. The French legislator has been among the first in the European Union to implement a law addressing this issue, and it could serve as an inspiration for strengthening the protection of children's image rights.[1]

Introduction

Imagine a woman who is an influencer. Her job is to share her everyday life on social media. This woman becomes pregnant. Since influencing is her only source of income, she continues sharing almost every moment of her pregnancy and does not stop after her baby is born. The mother posts photos and videos from her early motherhood days, including images of her baby. She does not think twice before posting the photos on social media because, what could go wrong? In reality, a lot can go wrong. This is a clear example of sharenting. In the

[1] The research was supported by the CUSEC Grant Scheme, Faculty of Law, Comenius University Bratislava (Rybanova_CUSEC,PRAF). Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 17R05-04-V01-00002 (Competence Center for the Regulation of Cybersecurity, Privacy Protection and Cybercrime).

following paragraphs, we will explain what this phenomenon is, why is it potentially dangerous, and how European digital legislation responds to it.

Sharenting refers to the practice of parents sharing their children's personal data online, including photos or videos. While it has become a widely popular practice, this behaviour can infringe a child's privacy and undermine the protective role of the parent.[2] Numerous research studies[3] have highlighted both the prevalence of sharenting and the potential negative consequences associated with this phenomenon.

One of these papers, which focuses on measuring child exposure and user engagement on TikTok[4] demonstrates the real impacts of sharenting. The study analysed 115 TikTok accounts featuring children under 13 years old, most of which were managed by adults, usually their parents. Examining 432,178 comments on 5,896 unique videos from these accounts, the authors found that almost 20% of the videos showed minors in revealing clothing, 21% of the comments related to visual appearance, and 114 comments included attempts by other users to get in touch with the account holder, mostly by asking to exchange direct messages or addresses.[5]

Another example is an investigation conducted in 2023 by the Wall Street Journal in collaboration with researchers from Stanford University and the University of Massachusetts Amherst.[6] This investigation revealed that Instagram's algorithms were actively promoting paedophile networks which commission and sell child pornography content through the Meta image-sharing app.

---

2 OECD: *Children in the Digital Environment: Revised Typology of Risks*. OECD Digital Economy Papers, No. 302, 2021. p. 14. Online: https://doi.org/10.1787/9b8f222e-en.

3 For further information, find below some of the research papers related to negative consequences of sharenting:
FERNANDEZ BLANCO, E., RAMOS GUTIERREZ, M*.: Sharenting on Instagram: abuse of the presence of minors in advertising*. In: Revista de Comunicación, Vol. 23, No. 1, 2024, pp. 177–194. Online: https://dx.doi.org/10.26441/rc23.1-2024-3460.
STEPHENSON, S., et al.: *Sharenting on TikTok: Exploring parental sharing behaviors and the discourse around children's online privacy*. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–17. Online: https://dl.acm.org/doi/full/10.1145/3613904.3642447.

4 SCHIRMER, M., VOGGENREITER, A., PFEFFER, J.: *More skin, more likes! meas-uring child exposure and user engagement on tiktok.* arXiv preprint. Online: https://doi.org/10.48550/arXiv.2408.05622.

5 *Ibid*, p. 3.

6 HORWITZ J., BLUNT K.: *Instagram connects vast pedophile network,* The Wall Street Journal, 7 June 2023. Online: https://www.wsj.com/tech/instagram-vast-pedophile-network-4ab7189.

These examples illustrate how sharenting can threaten children's privacy and safety. The aim of this paper is to analyse how the Digital Services Act (DSA)[7] reacts to the issue of sharenting, identify the most promising measures within the regulation, and suggest possible improvements. Methodologically, we firstly made a literature review of books and articles related to sharenting, risks that arise from sharenting, as well as sharenting and EU digital legislation. Then we conducted a normative analysis of law, specifically the DSA, with the focus on the articles that can potentially be used as measures to prevent sharenting, since it does not address this issue separately. As we concluded from the analysis that risk assessments could be the most promising measure, we chose 5 biggest social media platforms designated as Very Large Online Platforms and analysed whether they address the issue of sharenting in their risk assessments. Finally, we discuss the French Sharenting law, since France is the first EU Member State adopt a separate law regulating sharenting.

First, we provide a brief introduction to the phenomenon of sharenting and the risks it entails. Next, we examine which provisions of the DSA could be help prevent sharenting, with particular attention to Article 34. In this context, we analyse the risk assessments of selected Very Large Online Platforms (VLOPs), to determine whether they identify sharenting as a distinct category of risk. Finally, we briefly discuss the mechanisms introduced by the French sharenting law. The conclusions will be summarised in the final section.

1.1 Sharenting

The term sharenting, derived from the words "*sharing*" and "*parenting*", refers to the practice of parents publicly sharing information about their children online, mostly on social media platforms. A related subcategory is "influencer sharenting", where influencers feature their children on their profiles or create separate accounts for them.[8] This practice raises concerns about children's right to privacy, which is often infringed by their own parents.[9]

---

[7] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)(hereinafter „DSA"), OJ L 277, 27.10.2022.

[8] HUDDERS, L. et al.: *Little Stars, Big Troubles? Testing the Effectiveness of an Influencer Sharenting Intervention to Promote Mindful Sharenting Behaviors*. In: Annual Conference of Association of Internet Researchers, 2024, p. 1. Online: https://biblio.ugent.be/publication/01JNKQXSGTXV95QSKYEXQN3XCK.

[9] BLUM-ROSS, A., LIVINGSTONE, S.: *"Sharenting," parent blogging, and the boundaries of the digital self.* In: THUMIM, N. (ed.): Self-(re) presentation now. Routledge, 2020, p. 72. Online: https://www.taylorfrancis.com/chapters/edit/10.4324/9780429429187-

Posting children's pictures and videos, or personal data online can lead to various risks, including violations of privacy, undermining the parents' protective role, online grooming by predators,[10] victimisation of minors,[11] and even identity theft.[12] These risks can affect minors in different ways. Sharenting may violate their privacy when parents share photos and videos on social media that the children might not want others to see, which can lead to a traumatizing experience. Moreover, these photos can potentially become part of Child Sexual Abuse Material (CSAM). Children can also be exposed to online grooming or targeted by predators who may stalk the parents to learn the child's whereabouts and approach them when unattended. These risks should be taken seriously, as they can result in deeply harmful experiences for children. Furthermore, a study conducted in the United Kingdom found that, on average, nearly 1,500 photos of a child will have been posted online before they reach the age of five.[13]

These concerns surrounding sharenting raise important questions about child protection, including potential legal safeguards. While European Union digital regulation has not yet addressed this issue directly, certain rules and provisions partially cover it. The following sections will analyse, which provisions of the DSA could be applied to prevent sharenting, with particular attention to Article 34.

1.2 Digital Services Act and Sharenting

The DSA is an EU regulation that establishes specific obligations for a range of online services used by individuals within the European Union, regardless of where the service provider is located.[14] Intermediaries have different obligations, depending on the category of services they fall under according to the DSA. In this paragraph, we assess which mechanisms can be potentially used to prevent sharenting. These mechanisms can be found in Articles 16 (notice-and-action mechanisms), Article 22 (trusted flaggers), Articles 28 (protection of minors), Article 34 and 35 (risk assessment and mitigating measures). While the

6/sharenting-parent-blogging-boundaries-digital-self-alicia-blum-ross-sonia-livingstone.

10 *Ibid*.

11 LAVORGNA, A., UGWUDIKE, P., TARTARI, M.: *Online sharenting: Identifying existing vulnerabilities and demystifying media reported crime risks*. In: Crime, Media, Culture, Vol. 19, Issue 4, 2023, p. 484. Online: https://doi.org/10.1177/1741659022114844.

12 OECD: *Children in the Digital Environment: Revised Typology of Risks*. OECD Digital Economy Papers, No. 302, 2021, p. 14. Online: https://doi.org/10.1787/9b8f222e-en.

13 *Ibid,* p. 13.

14 DSA, Article 2 (1).

regulation includes provisions aimed at protecting children, not all of them are relevant for addressing sharenting content. Article 28 of the DSA is a clear example, as it lays out obligations for social media platforms to have "*appropriate and proportionate measures*" in place to protect children with respect to their privacy, safety, and security.[15] Paragraph 3 of this article provides that the European Commission may issue guidelines to support the implementation of the first paragraph. These guidelines,[16] published in July 2025, include specific measures[17] aimed at protecting minors.

Furthermore, Article 28, para. 2 of the DSA prohibits social media platforms from displaying advertisements based on profiling that uses the personal data of a minor. This rule applies only when the platform is "*aware with reasonable certainty*" that the user is a child. The measure aims to protect minors from targeted marketing and profiling. However, the DSA does not specify any mechanism for verifying whether a user is a minor. It likely also excludes cases where targeted marketing occurs through means other than profiling, such as inferring interests from user behaviour without direct profiling.[18] Since sharenting concerns the behaviour of parents (or adults in general), this article is not suitable for preventing sharenting.

Article 16 of the DSA addresses notice-and-action mechanisms. These systems for submitting complaints should be simple and accessible, making it easier to report and remove illegal content.[19] Together with the support of trusted flaggers,[20] this mechanism can play an important role in taking down sharenting content. However, research evaluating these reporting systems shows that they are not always user-friendly, largely due to the legal complexity

---

15 *Ibid,* Article 28 (1).

16 The guidelines were issued as an annex to the *Communication to the Commission No. C(2025) 4764 final, Approval of the content on a draft Communication from the Commission – Guidelines on measures to ensure a high level of privacy, safety and security for minors online.* Online: https://ec.europa.eu/newsroom/dae/redirection/document/118226.

17 The guidelines contain measures aimed to protect minors, such as recommendation to use effective age verification measures, prohibition from downloading or taking screenshots of content posted by minors or setting children's account as private by default.

18 WOLTERS, P., ZUIDERVEEN BORGESIUS, F.: *The EU Digital Services Act: what does it mean for online advertising and adtech?* In: International Journal of Law and Information Technology, 2025, Vol. 33, p. 24. Online: https://doi.org/10.1093/ijlit/eaaf004.

19 DSA, Article 16 (1).

20 *Ibid*, Article 22.

involved in filing a report under the DSA.[21] In practice, this makes it difficult for minors to report content through a mechanism that is not user-friendly. Another issue is the age limit: children under 13 are generally unable to use this mechanism on online platforms, as social media services require users to be at least 13 years old to register.[22]

A further limitation in applying this mechanism on sharenting content lies in the definition of "*illegal content*". Under the DSA, "*illegal content*" refers to any information that, either on its own or in connection with an activity, violates EU law or the law of any Member State that aligns with Union law, regardless of the specific subject matter or nature of that law.[23] This means that the legality of sharenting content depends on national laws, or on the EU law, and in result, it may be illegal in one Member State but not in another.

Finally, the mechanisms under Articles 34 and 35 of the DSA may offer another possible measure to address sharenting. Platforms such as Facebook, Instagram, Snapchat, and TikTok, which are examples of the largest social media stakeholders, have been designated as VLOPs under the DSA.[24] Providers of VLOPs are required, under Article 34 of the DSA, to conduct a risk assessment at least once a year or before implementing functionalities that could significantly affect the risks identified in Article 34. These systematic risks include, considering their severity and probability, any actual or potential negative effects on the protection of minors.[25] Based on these risk assessments, VLOPs must implement, reasonable, proportionate and effective mitigation measures tailored to the identified risks, in accordance with Article 35 of the DSA, which also provides a non-exhaustive list of possible measures which can be put in place.[26]

---

21 SEKWENZ, M. T., WAGNER, B., PARKIN, S.: *"It is unfair, and it would be unwise to expect the user to know the law!"– Evaluating reporting mechanisms under the Digital Services Act*. In: ACMF AccT 2025 - Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 2025, p. 541. Online: https://resolver.tudelft.nl/uuid:e992f10b-51ee-4f48-b62d-1b31a6850a46.

22 LENDVAI, G. F.: *Sharenting as a regulatory paradox–a comprehensive overview of the conceptualization and regulation of sharenting*. In: International Journal of Law, Policy and the Family, Vol. 38, Issue 1, 2024, p. 14. Online: https://doi.org/10.1093/lawfam/ebae013.

23 DSA, Article 3 (h).

24 Online platforms are designated as VLOPs by a decision of European Commission, according to Art. 33 (4) of the DSA. The overview of the designated VLOPs is available here: https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses.

25 DSA, Article 34 (1) (d).

26 *Ibid*, Article 35.

1.3 Sharenting in Risk Assessments

In the following paragraphs, we analyse the outcomes of the risk assessments of selected VLOPs. We chose the five largest social media platforms among all designated VLOPs, based on their average monthly active users (in millions). This data was updated on October 10, 2025, and is therefore up to date. The five biggest social media platforms according to their average monthly active users in millions are: (i) YouTube, with 416.6 million users; (ii) Facebook, with 259 million users; (iii) Instagram, with 259 million users; (iv) TikTok, with 135.9 million users; and (v) Pinterest, with 124 million users.[27] We analysed only their 2024 risk assessments, as not all of them were designated and required to prepare a risk assessment for 2023.

First, none of these platforms mention the term "sharenting", nor do they treat it as a separate issue. In relation to minors, the risk assessments primarily address child safety as users. Regarding the protection of a child's image, sharenting is addressed only in a very limited way, mostly under categories such as "illegal content," "child nudity," or "CSAM". Some platforms mention it more briefly than others, as demonstrated below.

Starting with YouTube,[28] the risk assessment refers to "*Illegal Content, Behaviour, and Products and Services*". This category includes risks that content or behavior enabling or involving child sexual exploitation and abuse may be present or occur on the service.[29] The risk assessment highlights YouTube's content moderation measures aimed at removing illegal content, especially CSAM. In addition to CSAI Match technology, YouTube enforces its Community Guidelines, which explicitly prohibit CSAM.[30] The platform claims to use automated detection systems to identify problematic content and enforce the Community Guidelines.[31] However, the risk assessment does not provide statistics on how many posts or videos were removed through these measures.[32]

---

27 Data source: The overview of the designated VLOPs is available here: https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses.

28 *YouTube Risk assessment*, 2024. Online: https://storage.googleapis.com/transparencyreport/report-downloads/dsa-risk-assessment_2024-8-28_2024-8-28_en_v1.pdf.

29 *Ibid*, p. 131.

30 *YouTube child safety policy*, Online: https://support.google.com/youtube/answer/2801999?sjid=7802933342237580276-EU#zippy=%2Ccontent-featuring-minors.

31 *YouTube Risk assessment*, 2024, p. 114–116.

32 Please note that these statistics may be found in the Transparency reports, however, these were not analysed in this research.

Facebook[33] and Instagram[34] are addressed together, as their risk assessments contain almost identical information regarding issues related to sharenting. Both identified "*Child Sexual Exploitation, Abuse and Nudity*" as a systematic risk. They reported that the platforms were used to share or promote content or activities involving non-sexual child abuse, child nudity, child endangerment, child sexual exploitation, CSAM (including self-generated CSAM), solicitation of CSAM, sexualization of minors, exploitative intimate imagery, and sextortion of minors.[35]

Both of these platforms stated that they disrupted 37 networks which were abusive and took down almost 200,000 accounts associated with abusive networks globally. On Facebook, 14.4 million pieces of potential child sexual exploitation content was identified, while Instagram reported 2.7 million. Most of this content was detected before users reported it.[36]

TikTok[37] identified illegal content as a systemic risk, specifically under the category of "risks of CSAM and child sexual exploitation and abuse (CSEA)". The assessment emphasized the dangers of disseminating CSAM material and sexually assaulting minors. The identified risks include dissemination of CSAM and CSEA, as well as sextortion and grooming. TikTok also provided exact figures for videos classified as CSAM and CSEA that were removed under the categories "Safety & Civility – Youth Exploitation & Abuse" and "Youth Safety & Well-Being – Youth Exploitation & Abuse."[38]

Finally, Pinterest's risk assessment[39] also listed illegal content as a systemic risk. To evaluate the risk of illegal content, Pinterest reviewed categories in its Community Guidelines[40] that align closely with the concept of illegal content and are most likely to include potentially illegal material. These categories include

---

33 *Facebook Risk assessment*, 2024. Online: https://drive.google.com/file/d/1ID9UGmERDUAV8WIhD65KzIQniMcrmaYH/view?usp=drive_link.

34 *Instagram Risk assessment*, 2024, Online: https://drive.google.com/file/d/1aVDSwdcUQUa5wPXqPyHIFeXGkwiMNT5s/view?usp=sharing.

35 *Facebook Risk assessment*, 2024, p. 69; *Instagram Risk assessment*, 2024, p. 67.

36 *Facebook Risk assessment*, 2024, p. 70; *Instagram Risk assessment*, 2024, p. 68.

37 *TikTok Risk assessment*, 2024, Online: https://sf16-va.tiktokcdn.com/obj/eden-va2/zayvwlY_fjulyhwzuhy[/ljhwZthlaukjlkulzlp/DSA/TikTok_DSA_Risk_Assessment_Report_2024.pdf.

38 *Ibid*, p. 41.

39 *Pinterest Risk assessment*, 2024. Online: https://help.pinterest.com/sites/default/files/dsa/2024-Pinterest-DSA-RIsk-Assessment-and-Mitigation-Report.pdf.

40 *Pinterest Community Guidelines*. Online: https://policy.pinterest.com/en/community-guidelines.

adult content, adult sexual services, child safety, dangerous goods and activities, graphic violence and threats, harassment and criticism, hateful activities, and violent actors. Pinterest expanded these categories to address new and emerging harm types, with particular emphasis on child safety and adult content.[41]

Illegal content mentioned in the Community Guidelines includes material that contributes to the sexualisation of minors, whether through imagery or text, such as nudity, partial exposure, or minors in underwear. The guidelines also state that Pinterest will deactivate users who save otherwise non-violating content into collections or in other contexts suggesting the intent to sexualise minors. The risk assessment reports that Pinterest removed 1,233 user-reported posts violating its child safety policy.[42]

To summarise none of the analysed platforms identified sharenting as a separate risk. Instead, it was partially addressed under categories related to illegal content, primarily child nudity. Moreover, only Pinterest provided specific figures for posts removed for child safety violations. This demonstrates that sharenting is not yet considered by platforms as a distinct systemic risk.

1.4 Sharenting in France as a Model Law

France is the first EU Member State to adopt a separate law aimed at combating and preventing sharenting. Law No. 2024-120 of February 19, 2024, aimed at guaranteeing respect for children's image rights[43] (Sharenting law) amends the French Civil Code by introducing specific protections for a child's image. Emphasized the importance of safeguarding children's image rights, citing a report from the National Center for Missing and Exploited Children, which found that 50% of photographs exchanged on child pornography forums had originally been posted by parents on their social media accounts.[44]

The text was conceived as a "*law of education rather than a repressive or punitive law*". Its purpose is therefore more about regulating than punishing the publication of images of minors online.[45] The Sharenting Law is relatively short but introduces significant changes. It introduces the concept of the privacy of minors into the definition of parental authority. It incorporates the concept of minors' privacy into the definition of parental authority and applies co-parenting

41 *Ibid*, p. 18.
42 *Pinterest Risk assessment*, 2024, p. 19.
43 Loi n° 2024-120 du 19 février 2024 visant à garantir le respect du droit à l'image des enfants.
44 Assemblé nationale, propostion de la loi. Online: https://www.assemblee-nationale.fr/dyn/16/textes/l16b0758_proposition-loi.pdf.
45 *Ibid*.

principles to the exercise of a child's image rights, meaning that all decisions related to these rights must be made jointly by both parents. Even separated parents are required to jointly protect their child's image rights while respecting the child's right to privacy.[46]

The law also strengthens the power of the family court judge (*Le Juge aux Affaires Familiales*), by prohibiting the publication or dissemination of a child's image without the consent of the other parent.[47] Finally, in cases where the dissemination of a child's image seriously undermines the child's dignity or moral integrity, the judge may delegate parental authority regarding image rights.[48]

The French Sharenting Law demonstrates that national legislation can fill gaps left by EU regulations. While the DSA addresses systemic risks and illegal content broadly, it does not explicitly recognize sharenting as a distinct category of harm. France's approach highlights the need for more targeted measures to protect children's privacy and image rights in the digital age. This raises an important question: should similar legislation be adopted at the EU level to ensure harmonized protection across Member States? And would it be possible?

Conclusion

To conclude, even though sharenting is becoming an increasingly widespread issue, the DSA does not expressly regulate this practice. While the DSA introduces important obligations for online platforms, its provisions primarily focus on illegal content, systemic risks, and child safety in general, without recognising sharenting as a distinct category of harm. This regulatory gap is evident in the risk assessments of Very Large Online Platforms, which address the issue only indirectly and in a very limited way, mostly under headings such as "illegal content," "child nudity," or "CSAM." None of the analysed platforms treat sharenting as a separate systemic risk, nor do they propose targeted measures to mitigate its consequences. However, the issue may also lie in the interpretation of the DSA by VLOPs, since the regulation does not prevent them from defining sharenting as a distinct category of risk.

Given the growing prevalence of sharenting and its potential to undermine children's privacy, dignity, and safety, additional measures should be considered to strengthen protection. First, the concept of co-parenting in decisions related to a child's image rights, as introduced by the French Sharenting Law, could serve

46 Sharenting law, Article 1 and 2.
47 *Ibid*, Article 3.
48 *Ibid,* Article 4.

as an inspiration for broader regulatory frameworks. This approach ensures that both parents jointly decide on the publication of a child's image, thereby reinforcing accountability and safeguarding the child's best interests. Second, compulsory warning signs could be implemented on social media platforms, alerting parents before they post images or personal information about their children. Such warnings would serve an educational function, reminding parents of potential risks and encouraging more responsible online behaviour. Third, platforms should develop effective automated mechanisms for removing or restricting potentially problematic sharenting content, as this type of content is not yet recognised by platforms as potentially harmful.

Ultimately, the absence of explicit regulation on sharenting at the EU level, combined with the limited attention given to this issue by VLOPs, highlights the need for a more comprehensive approach. As digital environments continue to evolve, protecting children's image rights should become a priority in both national and EU legislation. Introducing targeted measures, such as co-parenting requirements, preventive warnings, and takedown systems, would help close existing gaps and ensure that children's privacy and dignity are effectively safeguarded in the online sphere.

List of Literature and Sources

1. Assemblé nationale, propostion de la loi. Online: https://www.assemblee-nationale.fr/dyn/16/textes/l16b0758_proposition-loi.pdf.
2. BLUM-ROSS, A., LIVINGSTONE, S.: *"Sharenting," parent blogging, and the boundaries of the digital self*. In: THUMIM, N. (ed.): Self-(re) presentation now. Routledge, 2020, pp. 70–85. Online: https://www.taylorfrancis.com/chapters/edit/10.4324/9780429429187-6/sharenting-parent-blogging-boundaries-digital-self-alicia-blum-ross-sonia-livingstone.
3. *Facebook Risk assessment*, 2024, Online: https://drive.google.com/file/d/1ID9UGmERDUAV8WIhD65KzIQniMcrmaYH/view?usp=drive_link.
4. HORWITZ J., BLUNT K.: *Instagram connects vast pedophile network*. In: The Wall Street Journal, 7 June 2023. Online: https://www.wsj.com/tech/instagram-vast-pedophile-network-4ab7189.
5. HUDDERS, L, et al.: *Little Stars, Big Troubles? Testing the Effectiveness of an Influencer Sharenting Intervention to Promote Mindful Sharenting Behaviors*. In: Annual Conference of Association of Internet Researchers, 2024. Online: https://biblio.ugent.be/publication/01JNKQXSGTXV95QSKYEXQN3XCK.
6. LAVORGNA, A., UGWUDIKE, P., TARTARI, M.: *Online sharenting: Identifying existing vulnerabilities and demystifying media reported crime risks*. In: Crime, Media, Culture, Vol. 19, Issue 4, 2023, pp. 472–490. Online: https://doi.org/10.1177/1741659022114844.
7. LENDVAI, G. F.: *Sharenting as a regulatory paradox–a comprehensive overview of the conceptualization and regulation of sharenting*. In: International Journal of Law, Policy and the Family, Vol 38, Issue 1, 2024, pp. 1–17. Online: https://doi.org/10.1093/lawfam/ebae013.
8. Loi n° 2024-120 du 19 février 2024 visant à garantir le respect du droit à l'image des enfants.
9. OECD: *Children in the Digital Environment: Revised Typology of Risks*. OECD Digital Economy Papers, No. 302, 2021. Online: https://doi.org/10.1787/9b8f222e-en.
10. *Pinterest Community Guidelines*. Online: https://policy.pinterest.com/en/community-guidelines.
11. *Pinterest Risk assessment*, 2024. Online: https://help.pinterest.com/sites/default/files/dsa/2024-Pinterest-DSA-RIsk-Assessment-and-Mitigation-Report.pdf.

12. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022.
13. SEKWENZ, M. T., WAGNER, B., PARKIN, S.: *"It is unfair, and it would be unwise to expect the user to know the law!"– Evaluating reporting mechanisms under the Digital Services Act*. In: ACMF AccT 2025 - Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 2025, pp. 532–546. Online: https://resolver.tudelft.nl/uuid:e992f10b-51ee-4f48-b62d-1b31a6850a46.
14. SCHIRMER, M., VOGGENREITER, A., PFEFFER, J.: *More skin, more likes! measuring child exposure and user engagement on tiktok*. arXiv preprint. Online: https://doi.org/10.48550/arXiv.2408.05622.
15. *TikTok Risk assessment*, 2024, Online: https://sf16-va.tiktokcdn.com/obj/eden-va2/zayvwlY_fjulyhwzuhy[/ljhwZthlaukjlkulzlp/DSA/TikTok_DSA_Risk_Assessment_Report_2024.pdf.
16. WOLTERS, P., ZUIDERVEEN BORGESIUS, F.: *The EU Digital Services Act: what does it mean for online advertising and adtech?* In: International Journal of Law and Information Technology, Vol. 33, 2025, pp. 1–25. Online: https://doi.org/10.1093/ijlit/eaaf004.
17. *YouTube Risk assessment*, 2024. Online: https://storage.googleapis.com/transparencyreport/report-downloads/dsa-risk-assessment_2024-8-28_2024-8-28_en_v1.pdf.

Author details

Mgr. Lucia RYBANOVÁ
Comenius University Bratislava
Faculty of Law
Institute of Information Technology Law and Intellectual Property Law
Email: lucia.rybanova@flaw.uniba.sk

2 Right to a Fair Trial: Can AI Replace Human Judges? European Perspective

Magdalena Matusiak-Frącczak

https://doi.org/10.33542/HRF-0541-2-02

Abstract

Right to a fair trial is protected in a multi-layer way. It is enshrined in Article 6 of the European Convention on Human Rights, in Article 47 of the EU Charter of Fundamental Rights and Article 19 of the Treaty on the European Union, but also in national constitutions of European States. Thanks to this right everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law. The case law on the meaning of "independent and impartial tribunal established by law", both of the ECtHR and the CJEU, is abundant and has developed rapidly in recent years due to the rule of law violations in different European States. A "tribunal" is characterised in the substantive sense of the term by its judicial function, that is to say, determining matters within its competence on the basis of legal rules and after proceedings conducted in a prescribed manner. It must also satisfy a series of further requirements, such as "independence, in particular of the executive; impartiality; duration of its members' terms of office. It is also inherent in the very notion of a "tribunal" that it be composed of judges selected on the basis of merit – that is, judges who fulfil the requirements of technical competence and moral integrity to perform the judicial functions required of it in a State governed by the rule of law. The right to a fair trial presupposes that judge – a human being – is to be characterized by specific qualities that permit them to objectively and independently deliver justice. The question of this presentation is to discuss whether AI can replace human judges, so whether AI can constitute an independent and impartial tribunal established by law. On the one hand algorithms can provide for objectivity and homogenous application of law. On the other hand, more complicated factual background of a case or legal problem might escape algorithms and lead to unjust results. Is AI capable of analysing emotional and psychological aspects of a case at stake (like credibility of witnesses)? This issue raises further questions like on who in fact delivers judgment: a judge, a programmer or an executive authority that orders the creation of a program and sets technical requirements for its functioning. Wouldn't it be a violation of the separation of powers? Another problem are hallucinations, which happen to AI much too often. Who would review an AI judgment from the perspective of eventual hallucinations and who does deliver judgment in such case – AI or the reviewer?

The chapter will try to answer these questions with reference to European standards.

Introduction

Currently only a few courts use AI-based systems to support judges or to adjudicate,[1] so the possibility of AI replacing human judges is a question of the future. Some scholars indicate that AI merely simulates human thinking, so the full imitation of human intelligence might not be attainable for algorithmic models.[2] However, according to some scholars, replacing human judges with AI is inevitable.[3]

The use of AI to replace judges is definitely tempting, as it facilitates an analysis of large quantities of data in a very short time. AI systems can be even superior to human capabilities in terms of precision, speed and reliability, and moreover any such system, with access to huge databases of national law, international law and EU law would have larger knowledge than any judge.[4] AI can quickly and effectively analyse large portions of documents and evidence, including text and images, and automatically generate a draft of judicial decision.[5] Traditionally such searches were done by legal professionals

---

1 Estonia is developing an AI judge to decide on small pecuniary disputes of less than 7000 EUR and China has AI judges. In Wisconsin in the US judges use AI to draft recommended criminal sentences. CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol. 2, No. 1, 2020, p. 4; CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, p. 233; KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, p. 209. BARYSĖ, D., SAREL, R.: *Algorithms in the court: does it matter which part of the judicial decision-making is automated?* In: Artificial Intelligence and Law, Vol. 32, 2024, p. 118.

2 LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian case*. In: Mexican Law Review, Vol. XIV, No. 2, 2022, p. 57.

3 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, p. 1031.

4 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, pp. 2223–2224.

5 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1027–1028.

manually, based on keywords and the analysis took plenty of time, AI can overtake the performance of such tasks and it will do them more effectively.[6]

Replacing human judges with AI raises in Europe serious issues regarding human rights. The European Union even identified AI systems intended for the administration of justice as high-risk, considering their potentially significant impact on democracy, the rule of law, individual freedoms as well as the right to an effective remedy and to a fair trial. The EU indicated that AI tools can support the decision-making power of judges but should not replace it as the final decision-making must remain a human-driven activity.[7]

This paper will analyse and deliberate whether and to what extent it is possible to replace human judges with AI in Europe. For this purpose, the first part of the paper will be devoted to the European standard on the right to a fair trial. The second part will focus on deploring, how AI fits into this standard, what would enable making final conclusions.

2.1 Right To a Fair Trial: European Standard

Not many documents refer directly to the relationship of AI and human rights. Within the framework of the Council of Europe the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment[8] was established. It provided for five basic principles and the first of them is the respect of fundamental rights. According to this principle, the use of AI should not undermine the guarantees of the right of access to the judge and the right to a fair trial, as well as the rule of law and judges' independence in their decision-making process.

Therefore, when discussing the right to a fair trial, we need to start with relevant basic regulations. In general, in the determination of his civil rights and obligations or of any criminal charge against him, everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law. Judgment shall be pronounced publicly but the press

---

6 LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian case*. In: Mexican Law Review, Vol. XIV, No. 2, 2022, p. 61.

7 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) (Text with EEA relevance), OJ L, 2024/1689, 12.7.2024, motive (61), Annex III (8).

8 European Commission for the Efficiency of Justice: *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, adopted at the 31st plenary meeting of the CEPEJ, Strasbourg, 3 – 4 December 2018.

and public may be excluded in exceptional circumstances from all or part of the trial (Article 6(1) ECHR). In the European Union this corresponds to Article 47 of the EU Charter of Fundamental Rights: Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law. It has to be remembered that the level of protection stemming from the EU Charter should be at least the same as the level of protection provided by the Convention (Article 52(3) EU Charter of Fundamental Rights). At the same time, the EU Member States are obliged to provide remedies sufficient to ensure effective legal protection in the fields covered by Union law (Article 19(1) TEU), which is understood as an obligation to provide for effective judicial protection.

The preservation of the right to a fair trial depends on the circumstances of the particular case, States are obliged to provide for the overall fairness of the proceedings. Usually the analysis should not be based on an isolated consideration of one particular aspect or one particular incident, although at times it can happen that an isolated factor might be decisive to the evaluation of the fairness of the proceedings.[9] At the same time it is emphasized that "fairness" encompasses only procedural safeguards like equality of arms or participation in the proceedings and not material fairness related to the manner of resolving legal dispute by a tribunal.[10]

One of the most important aspects of the right to a trial is the right to a hearing by an independent and impartial tribunal established by law. These concepts correlate with each other and there is a close interrelationship between them. Any court that does not fulfil these criteria, cannot be considered to be a "tribunal" in the meaning of Article 6(1) ECHR or Article 47 of the EU Charter of Fundamental Rights and Article 1 (1) TEU.

The notion of "tribunal established by law" requires that there exists a legal basis of its existence as well as the court or the composition of the bench in each case complies with the particular rules that govern it. This notion encompasses also the process of the initial appointment of a judge to office.[11] An organ that was not appointed in conformity with the will of the legislator, would be

---

9 ECtHR: *Ibrahim and others v. the United Kingdom*, Applications no. 50541/08, 50571/08, 50573/08 and 40351/09, Judgment of 13 September 2016, pp. 250–251.

10 GÓRSKI, M.: *Dziesiąty element. Sztuczna inteligencja jako sędzia a prawo do sądu*. In: Prawo Mediów Elektronicznych, No. 3, 2022, p. 6.

11 ECtHR: *Kontalexis v. Greece*, Application no. 59000/08, Judgment of 31 May 2011, p. 42; ECtHR: *Guðmundur Andri Ástráðsson v. Iceland*, Application no. 26374/18, Judgment of 1 December 2020, p. 223, 228; ECtHR: *Xero Floor w Polsce sp. z o.o. v. Poland,* Application no. 4907/18, Judgment of 7 May 2021, p. 245; ECtHR: *Reczkowicz v. Poland,* Application no. 43447/19, Judgment of 22 July 2021, p. 218.

deprived of legitimacy required in a democratic society, as the judicial organisation in a democratic society does not depend on the discretion of the executive, but that it is regulated by law emanating from parliament.[12] Moreover, if a tribunal does not have jurisdiction to try a defendant in accordance with the provisions applicable under domestic law, it is not "established by law".[13] The requirement "established by law" encompasses not only the existence of a legal basis for a tribunal, but these legal provisions need to be applied in a rational and non-arbitrary manner.[14] The law establishing a tribunal must also meet certain requirements. It must be formulated in unequivocal terms, to the extent possible, so as not to allow arbitrary interferences in the appointment process, including by the executive.[15]

According to the ECtHR, it is inherent in the very notion of a "tribunal" that it is composed of judges selected on the basis of merit – that is, judges who fulfil the requirements of technical competence and moral integrity to perform the judicial functions required of it in a State governed by the rule of law. The higher a tribunal is placed in judicial hierarchy, the more demanding the applicable selection criteria should be.[16]

A "tribunal" determines matters within its competence on the basis of rules of law and after proceedings conducted in a prescribed manner. It must be independent, especially of the executive and legislator, and impartial. These can be achieved by duration of its members' terms of office or guarantees afforded by its procedure.[17]

---

12 ECtHR: *Lavents v. Latvia,* Application no. 58442/00, Judgment of 28 Novembre 2002, p. 114; ECtHR: *Richert v. Poland,* Application no. 54809/07, Judgment of 25 October 2011, p. 42; ECtHR: *Reczkowicz v. Poland,* Application no. 43447/19, Judgment of 22 July 2021, p. 216.

13 ECtHR: *Richert v. Poland,* Application no. 54809/07, Judgment of 25 October 2011, p. 41.

14 HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1–18, Warszawa 2010, p. 312.

15 ECtHR: *Guðmundur Andri Ástráðsson v. Iceland*, Application No. 26374/18, Judgment of 1 December 2020, p. 230; ECtHR: *Xero Floor w Polsce sp. z o.o. v. Poland,* Application no. 4907/18, Judgment of 7 May 2021, p. 246; ECtHR: *Reczkowicz v. Poland,* Application no. 43447/19, Judgment of 22 July 2021, p. 219.

16 ECtHR: *Guðmundur Andri Ástráðsson v. Iceland*, Application No. 26374/18, Judgment of 1 December 2020, pp. 220 – 222; ECtHR: *Xero Floor w Polsce sp. z o.o. v. Poland,* Application no. 4907/18, Judgment of 7 May 2021, p. 244. ECtHR: *Reczkowicz v. Poland,* Application no. 43447/19, Judgment of 22 July 2021, p. 217.

17 ECtHR: *Belilos v. Switzerland,* Application no. 10328/83, Judgment of 29 April 1988, p. 64; ECtHR: *Coëme and others v. Belgium,* Application no. 32492/96, 32547/96, 32548/96, 33209/96 and 33210/96, Judgment of 22 June 2000, p. 99; ECtHR: *Richert v. Poland,* Application no. 54809/07, Judgment of 25 October 2011, p. 43; ECtHR: *Ninn-*

"Independence" is inherent in the task of adjudication[18] and refers to the necessary personal and institutional independence of a judge that is required for impartial decision-making, and it is a prerequisite for impartiality. It characterises both a state of mind, which denotes a judge's imperviousness to external pressure as a matter of moral integrity and a set of institutional and operational arrangements – involving both a procedure by which judges can be appointed in a manner that ensures their independence and selection criteria based on merit – which must provide safeguards against undue influence and/or unfettered discretion of the other State powers (executive, legislative), both at the initial stage of the appointment of a judge and during the exercise of his or her duties,[19] as judiciary is a separate pillar of State authority.[20] Courts need to be protected both against direct and indirect influence of other State authorities – legislative and executive.[21] The body being a court should be protected by guarantees against outside pressures and it should present an appearance of independence. In this case the confidence which the courts in a democratic society must inspire in the public is at stake.[22] An independent judge is free not only from undue influences outside the judiciary, but also from within, including from directives or pressures from the fellow judges or those who have administrative responsibilities in the court. The absence of sufficient safeguards securing the independence of judges within the judiciary and, in particular, vis-à-vis their judicial superiors, may lead the violation of the right to a fair trial,[23] especially if they are in a subordinate position in terms of their duties and the organisation of their service, regarding one of the parties.[24] An independent

Hansen v. Denmark, Application no. 28972/95, Decision of 18 May 1999. PECH, L.: *Article 47(2) – An Independent and Impartial Tribunal Previously Established by Law*. In: PEERS, S., HERVEY, T., KENNER, J., WARD, J.: *The EU Charter of Fundamental Rights. A Commentary*. Hart Publishing: Oxford, 2021, p. 1351.

18 PECH, L.: *Article 47(2) – An Independent and Impartial Tribunal Previously Established by Law*. In: PEERS, S., HERVEY, T., KENNER, J., WARD, J.: The EU Charter of Fundamental Rights. A Commentary. Oxford: Hart Publishing, 2021, pp. 1349, 1353.

19 ECtHR: *Guðmundur Andri Ástráðsson v. Iceland*, Application No. 26374/18, Judgment of 1 December 2020, p. 234.

20 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2222.

21 CJEU: *Case C-585/18, A. K. and Others v Sąd Najwyższy, CP v Sąd Najwyższy and DO v Sąd Najwyższy*, Judgment of 19 November 2019, p. 124–125.

22 ECtHR: *Morice v. France,* Application no. 29369/10, Judgment of 23 April 2015, p. 78.

23 ECtHR: *Parlov-Tkalčić v. Croatia*, Application no. 24810/06, Judgment of 22 December 2009, p. 86.

24 ECtHR: *Şahiner v. Turkey,* Application No. 29279/95, Judgment of 25 September 2001, p. 45.

tribunal exercises its judicial functions wholly autonomously, without being subject to any hierarchical constraint or subordinated to any other body and without taking orders or instructions from any source whatsoever, and is thus protected against external interventions or pressure liable to impair the independent judgment of its members and to influence their decisions. The guarantees of independence include the protection from the removal from office and an adequate level of remuneration.[25] "Impartiality" on its part requires that the court is subjectively free of personal prejudice or bias and impartial from an objective viewpoint, that is, it must offer sufficient guarantees to exclude any legitimate doubt in this respect.[26] An impartial judge maintains an equal distance from the parties to the proceedings and their respective interests with regard to the subject matter of those proceedings.[27] At the same time a judge may not be biased by their own emotions or political preferences.[28] To this regard even appearances may be of some importance,[29] nevertheless doubts regarding the impartiality should be objectively justified.[30] The independence and impartiality criterion applies both to professional judges as well as lay judges,[31] so as a consequence we should assume that it will apply as well to a judge being AI.

One of the important aspects of the right to a fair trial is legal certainty, which guarantees a certain stability in legal situations and contributes to public confidence in the courts. The ECtHR accepts on the one hand that certain divergences in interpretation exist as they are an inherent trait of any judicial system. On the other hand inconsistent adjudication of claims brought by many persons in identical situations and the lack of institutional resolution of these

---

25 CJEU: *Case C-64/16, Associação Sindical dos Juízes Portugueses v Tribunal de Contas*, judgment of 27 February 2018, p. 44–45.

26 ECtHR: *Ninn-Hansen v. Denmark,* Application no. 28972/95, Decision of 18 May 1999, p. 30; ECtHR: *Daktaras v. Lithuania,* Application no. 42095/98, Judgment of 10 October 2000, p. 30.

27 CJEU: *Case C-585/18, A. K. and Others v Sąd Najwyższy, CP v Sąd Najwyższy and DO v Sąd Najwyższy*, judgment of 19 November 2019, p. 122.

28 HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1–18, Warszawa 2010, p. 319.

29 ECtHR: *Kyprianou v. Cyprus,* Application no. 73797/01, Judgment of 15 December 2005, p. 118; PECH, L.: *Article 47(2) – An Independent and Impartial Tribunal Previously Established by Law*. In: PEERS, S., HERVEY, T., KENNER, J., WARD, J.: The EU Charter of Fundamental Rights. A Commentary. Hart Publishing: Oxford, 2021, p. 1351.

30 HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1–18, Warszawa 2010, p. 321.

31 ECtHR: *Holm v. Sweden,* Application no. 14191/88, Judgment of 25 November 1993, p. 30.

inconsistencies creating continued uncertainty might lead to a violation of Article 6 ECHR.[32] However this does not create a general right to consistency of case-law or established jurisprudence,[33] as case-law development is not, in itself, contrary to the proper administration of justice since a failure to maintain a dynamic and evolutive approach would risk rendering it a bar to reform or improvement, provided that reasons justifying the departure are substantial.[34]

One of the aims of the discussed right is fairness, as proceedings in a given case considered as a whole should be fair. A guarantee of fairness is one of the fundamental principles of any democratic society.[35] This means that the rule of law should be respected and the adjudication is free from arbitrariness.[36] At the same time tribunals should analyse the case in thoroughly and seriously,[37] by a proper examination of submissions, arguments and evidence presented by the parties,[38] and not by pre-existing pattern. Also elements of procedural fairness, like adversarial proceedings or equality of arms, are essential.[39] Some cases require more than only a mere application of legal norms, they might require human intuition of what is fair and just.[40] Fairness is something more than providing for judicial services, it encompasses the social value of correct, equitable, and adequate solution of judicial disputes that considers also ethical values.[41]

---

32 ECtHR: *Vinčić and Others v. Serbia*, Application no. 44698/06, Judgment of 1 December 2009, p. 56.

33 ECtHR: *Unédic v. France,* Application no. 20153/04, Judgment of 18 December 2008, p. 74.

34 ECtHR: *Atanasovski v. "the Former Yugoslav Republic of Macedonia",* Application no. 36815/03, Judgment of 14 January 2010, p. 38; ECtHR: *Borg v. Malta,* Application no. 37537/13, Judgment of 12 January 2016, p. 107.

35 ECtHR: *Pretto and Others v. Italy,* Application no. 7984/77, Judgment of 8 December 1983, p. 21.

36 GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, p. 85.

37 ECtHR: *Donadzé v. Georgie,* Application no. 74664/01, Judgment of 7 March 2006, p. 35.

38 ECtHR: *Kraska v. Switzerland,* Application no. 13942/88, Judgment of 19 April 1993, p. 30.

39 HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1–18, Warszawa 2010, p. 329.

40 GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, p. 87.

41 LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian* case. In: Mexican Law Review, Vol. XIV, No. 2, 2022, pp. 73–74.

Another important requirement of the right to a fair trial is public hearing which includes both holding of public hearings and public delivery of judgments.[42] It protects litigants against the secret administration of justice with no public scrutiny. It is also one of the means whereby confidence in the courts can be maintained,[43] however this requirement is not absolute. The decision to hold a hearing *in camera* should be strictly required in the circumstances of the case.[44] Sometimes an oral hearing may not be required, especially where there are no issues of credibility or contested facts which necessitate a hearing and the courts may fairly and reasonably decide the case on the basis of the parties' submissions and other written materials.[45] Also the demands of efficiency and economy can be at stake, especially when the systematic holding of hearings would prevent compliance with the reasonable-time requirement.[46] Oral hearings are essential in criminal cases as the accused may need to verify the accuracy of his or her statements and compare them with those of the victim and of the witnesses,[47] although at times a hearing might take place with the absence of the accused when the accused waives the right to be present or deliberately seeks to evade the trial.[48] From this perspective, oral hearings would be indispensable also in family matters, as they are based predominantly on parties' and witnesses' statements.

2.2 AI As a Judge

Having established basic principles regulation, the right to a fair trial, it can be discussed how AI fulfils them. In this part of the paper the analysis will focus on the following criteria determined in the previous part: tribunal established by law, independence and impartiality, legal certainty and public hearing. In addition to that, some remarks on other issues relating to the use of AI in

42 ECtHR: *Sutter v. Switzerland,* Application no. 8209/78, Judgment of 22 February 1984, p. 27.

43 ECtHR: *Riepan v. Austria,* Application no. 35115/97, Judgment of 14 November 2000, p. 27.

44 ECtHR: *Kilin v. Russia,* Application no. 10271/12, Judgment of 11 May 2021, pp. 111–112.

45 ECtHR: *Döry v. Sweden,* Application no. 28394/95, Judgment of 12 November 2002, p. 37.

46 ECtHR: *Jussila v. Finland,* Application no. 73063/01, Judgment of 23 November 2006, p. 42.

47 ECtHR: *Medenica v. Switzerland,* Application no. 20491/92, Judgment of 14 June 2001, p. 54.

48 ECtHR: *Lena Atanasova v. Bulgaria,* Application no. 52009/07, Judgment of 26 January 2017, p. 52.

adjudication will be discussed, namely the evidence analysis and hallucinations, as they highly affect the possibility of the use of AI.

2.2.1 Tribunal Established by Law

A law introducing AI as a judge can be established by the legislator, as long as constitution permits so, therefore in some States constitutional amendments would be required.[49] If this law comes from a parliament, it is formulated in unequivocal terms and does not permit arbitrary interferences in the appointment process, it can be easily envisaged, that this requirement is met. As long as AI acts within frames established by the legislator, on the basis of law, within its jurisdiction, this technical and formal criterium is fulfilled. Nevertheless, the story does not end here with formalities, but there are other qualities required from judges that should be fulfilled to consider them to constitute and independent and impartial tribunal established by law, safeguarding everyone's right to a fair trial.

The law should also indicate requirements of technical competence and moral integrity to perform the judicial functions, which should be more demanding the higher the tribunal is placed in the judicial hierarchy. AI's technical competence can be very high, higher than most if not all human judges, provided that it is equipped with sufficiently abundant and rich databases. However, can we speak about moral integrity with reference to AI? From the linguistic perspective moral refers to the standards of good or bad behaviour, fairness, honesty, that each person believes in, rather than to laws, it is a person who behaves in ways considered by most people to be correct and honest.[50] AI's functioning is technical by its very nature, AI doesn't make moral choices nor has moral beliefs. At the same time judgments may require analysis from moral perspective, something that AI simply lacks. On the other hand, there are simple cases concerning for example mere payments, that require purely legal and technical analysis. Therefore, it is envisageable that in such simple, technical cases decided by courts placed low in judicial hierarchy AI could deliver judgments.

2.2.2 Independence and Impartiality

At first sight AI would be an independent and impartial tribunal, provided that legislator creates an adequate law regulating its functioning. Its term of

---

[49] VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, p. 1158.

[50] Cambridge Dictionary: *Moral*. Online: https://dictionary.cambridge.org/dictionary/english/moral#google_vignette.

office might be unlimited, as it is not a human being so it is not concerned by limitations created by biological barriers. Deciding a case AI appears to be free from prejudice and bias, fair, objective and neutral. It seems to be free from external pressures, as AI would not be subject to disciplinary proceedings, external criticism simply may not reach AI's "attention". AI would not have any concerns regarding removal from office or level of remuneration, which are of importance for human judges.

However, we cannot lose from sight that AI is human-created, so it will function in a way its programmer creates it, it will contain those databases, that its programmer uploads. AI functions on the basis of algorithms, which can be defined as pre-defined action plans to be followed by AI.[51] The AI substantive content will highly depend on its programmer or on the person or persons that decide on parameters that judicial AI would need to fulfil. It is very probable that AI parameters will be decided in a centralised way by an organ who orders the creation of this specific AI, in most cases it would probably be the Ministry of Justice. It will then be possible to accuse AI of not being independent from the Ministry, so from the executive. It is also possible that AI algorithms will be programmed in such a way that the State Treasury will be favoured in court proceedings or public prosecutors would be favoured in penal cases. Therefore, in a case between two individuals AI can be completely impartial and independent, whereas it may happen that in a case of an individual against a State AI can be biased by the way it was programmed. This risk needs to be taken into consideration in discussions on whether AI can replace a judge. As a remedy to this drawback it is postulated, that the algorithms of such AI systems should be made public, what would improve transparency and permit public scrutiny thereof.[52] However, it would not be a simple task, because at the same time the full transparency is not advisable, as then the system vulnerabilities can be discovered and exploited. Therefore, another solution to this problem is conferring the overview competences to the body of experts, what would limit the group of people with the access to codes and algorithms.[53] Moreover, the

---

51 LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian case*. In: Mexican Law Review, Vol. XIV, No. 2, 2022, p. 64.

52 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2224; MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, pp. 1099–1100; KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, pp. 229–230, 244.

53 CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol.

coding and algorithms can be erroneous and at the same time the users may act in full confidence in the system.[54] This problem can be only remedied by high awareness of the users and their own knowledge combined with the obligation to review the AI-generated results. We can also find opinions, that independence and impartiality are necessarily human characteristics, and for this reason a tribunal may only be composed of human beings.[55] This is further supported by an argument that one of the aspects of the separation of powers, which is an inherent feature of democracy, is the ability of judiciary to check the other two branches and AI would be unable to perform the function of checks and balances.[56]

M. Górski disagrees, indicating that it is not the AI program itself, but the decision-making process, that should not be interfered with and there should be no external undue influence on how justice is delivered in particular case. According to this author, human judges are also in a way "programmed" by their knowledge, cultural and social background or individual prejudices. Therefore, while AI is being programmed by humans, there should exist specific safeguards to protect the algorithm from external influence.[57] Provided that such safeguards are technically possible, AI would meet the criterium of independence and impartiality.

2.2.3 Legal Certainty

AI can contribute to legal certainty, as it will adjudicate in conformity with patterns compatible with its databases,[58] therefore it can raise the level of confidence in the courts. At some point it may even eliminate inconsistencies in interpretations and legal uncertainty or even human arbitrariness. This can contribute to the unification of judicial standards and prevent unfair and unpredictable judgments. Delivering different judgments in similar and

2, No. 1, 2020, p. 15; VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1175–1176.

54 CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol. 2, No. 1, 2020, p. 14.

55 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2224.

56 MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, p. 1100.

57 GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, p. 1156.

58 KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, pp. 214–215.

repetitive cases undermines judicial credibility and may lead to questioning the legitimacy of judge's discretion,[59] AI can therefore eliminate this risk.

Unfortunately, at the same time it can eliminate completely legal development and judicial activism, which can be good for the administration of justice. Social and economic realms of everyday life change and jurisprudence should catch up with these developments. Moreover, substantial reasons might justify departure from a well-established legal precedence. AI acting on the basis of programmed patterns and past case-law would prevent this from happening. The law itself requires human interpretation with a space left to choose from different options. In the case of AI the interpretation would be predetermined, narrow and lacking flexibility.[60] A human judge, when confronted with the arguments of the parties, can potentially be persuaded, AI does not seem to be persuadable, as it will deliver judgment on a pre-programmed basis.[61] Within the EU law, AI might overlook the necessity to initiate the preliminary ruling procedure (Article 267 TFEU).

If the fairness of the proceedings is to be guaranteed by AI, it should be able to properly analyse the case, submissions and evidence. This might require very complex and complicated analyses which do not necessarily have to follow the pre-existing pattern. A case might require taking into account social developments or specific circumstances existing in a given population,[62] and as it was already stated, AI decisions would be greatly decided by patterns. If circumstances and other aspects of the case are not repetitive, but unique, AI seems rather unable to deliver a fair judgment, as it lacks human behavioural

---

59 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1028–1029; GRAVETT, W. H.: *Judicial Decision-Making in the Age of Artificial Intelligence*. In: ANTUNES, H. S.; FREITAS, P. M.; OLIVEIRA, A. L.; PEREIRA, C. M.; SEQUEIRA, E. V.; XAVIER, L. B. (eds.): Multidisciplinary Perspectives on Artificial Intelligence and the Law, Cham: Springer, 2024, pp. 282–283.

60 CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol. 2, No. 1, 2020, pp. 6–7; VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1183–1184. CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, pp. 237–238, 247. KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, pp. 249–250.

61 MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, p. 1084.

62 GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, p. 85. MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, pp. 1089–1091.

functions permitting to safeguard the procedural fairness of proceedings.[63] On the other hand, some authors postulate, that if AI can convincingly and persuasively justify their judgments, it may be more reliable and credible than a human judge, and thus its judgments would be more fair.[64] The preservation of procedural fairness by AI, with protections granted to persons who have justified difficulties in coping with the electronic system, seems sufficient for some authors to maintain the overall fairness of the proceedings.[65] For this reason it is again justified that AI could replace human judges in simple, repetitive cases. This would make work of courts more effective if judges concentrated on more complicated cases where the application of a simple pattern is not sufficient.

2.2.4 Public Hearing

AI judgments would be considered to be judgments delivered *ex parte* and *in camera*, so its very essence precludes public hearings and public scrutiny over judicial decisions. At the same time some cases require hearing statements from the parties or from witnesses, therefore in such instances delivery of a judgment by AI would be incompatible with the European standard.

On the other hand, the requirement of public hearing is not absolute. It is envisageable that in cases where adjudication is based solely on documents, requirements of efficiency and economy might justify the use of AI to deliver judgments. In such cases the lack of public hearing would not constitute a violation of the right to a fair trial. Therefore, AI can considerably improve the speed of proceedings, what in itself is also a value from the perspective of the right to a fair trial.[66] This would require a higher informatization of courts, including online case-filing, online evidence exchange and including online case-

63 WALTERS, R.: *Robots replacing human arbitrators: the legal dilemma*. In: Information & Communications Technology Law, Vol. 34, No. 2, 2025, p. 136.

64 VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1159–1160.

65 GÓRSKI, M.: *Dziesiąty element. Sztuczna inteligencja jako sędzia a prawo do sądu*. In: Prawo Mediów Elektronicznych, No. 3, 2022, p. 6; HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1–18*,* Warszawa 2010, p. 328.

66 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2224. GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, pp. 83–84; LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian case*. In: Mexican Law Review, Vol. XIV, No. 2, 2022, p. 68.

hearings with AI making transcripts of statements. It would be highly profitable in cases of mass litigation with standardized case-law.[67]

2.2.5 Evidence Analysis

AI can be very efficient in analysing documents, so from this perspective replacing a human judge with AI is imaginable. That would be true especially in civil and administrative cases based on documents, like for example cases concerning simple payments. Evidence examination can be done effectively by AI in standardized and repetitive cases.[68]

At the same time, evidence analysis requires often decisions on credibility thereof, especially when it comes to oral statements made by parties or by witnesses, they require life experience, so something that AI lacks. As the development of technology now stands, AI is not equipped with tools permitting to decide what is true or untrue, or what is credible or incredible, as it purely operates on databases.[69] Evaluation of statements made by different persons during the proceedings may also require making moral choices. As it was already stated, AI is unable to make such choices or value judgments, as these are not mere technical decisions. Another problem is that analysis of texts by AI is purely mechanical, technical and linguistic, whereas natural persons use contexts, metaphors or irony.[70] Moreover, to prepare a good outcome, AI would need to receive mere, pure facts, whereas the relations on facts made by parties and witnesses are personal and subjective, what blurs the image of facts.[71]

There are types of cases where evidence analysis plays an important role in the decision-making process. For example, in penal cases the court needs to decide not only about legal qualification and penalty imposed, but it also needs to adjudicate whether the commission of a crime was intentional or unintentional, a judge must discover why the crime was committed and what are the defendant's chances for resocialisation, rehabilitation and reintegration. Although there are suggestions of automated systems for recidivism predictions,

---

67 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1027–1028.

68 *Ibid*, p. 1027.

69 Although it is possible that AI makes probability analysis and is not limited by personal knowledge like human judges. See: XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, p. 1029.

70 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1032–1033.

71 KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, p. 246.

it is indicated that they may lead to discriminatory outcomes, as they fail to adequately control for gender or racial disparity, but at the same time a human judge can be also biased and prejudice-driven.[72] In family matters the courts need to examine what family ties exist between parents and children, how children function in a family, what type of educational skills parents have, how a divorce ruling should be shaped. These are issues that escape databases and cannot be framed in a form of an algorithm. Many civil cases also escape automatization, especially when the court is to judge on the basis of comity or principles of social coexistence, and not merely on the basis of available documents.

2.2.6 Hallucinations

As the technical development now stands AI tools hallucinate. This means that they can refer to bibliographic sources with references to known authors or institutions with coherent titles and reasonably appearing contents but completely non-existent in reality.[73] It can make up stories, case-law, facts about the world and at the same time some of the content it presents is true. Regardless whether it concerns a real content or a hallucination, AI presents both in a very confident tone.[74] Hallucinations result from the way AI models are trained, simply when AI tool does not have sufficient data to refer to, it starts fabricating details.[75] A recent report of EBU showed that about 45% of all AI answers had at least one significant issue, 31% of responses showed serious sourcing problems, meaning they were missing, misleading or incorrect, whereas 20% contained major accuracy issues, including hallucinations and outdated information.[76] At the same time humans trust automated systems and perceive them as agents

---

72 GRAVETT, W. H.: *Judicial Decision-Making in the Age of Artificial Intelligence*. In: ANTUNES, H. S.; FREITAS, P. M.; OLIVEIRA, A. L.; PEREIRA, C. M.; SEQUEIRA, E. V.; XAVIER, L. B. (eds.): Multidisciplinary Perspectives on Artificial Intelligence and the Law, Cham: Springer, 2024, pp. 283–284.

73 SALVAGNO, M.; TACCONE, F. S.; GERLI, A. G.: *Artificial intelligence hallucinations.* In: Critical Care, Vol. 27, Art. No. 180, 2023, p. 1.

74 SLATER, J., HUMPHRIES, J.: *Another reason to call bullshit on AI "hallucinations"*. In: AI & Society, Vol. 40, 2025, p. 5631.

75 KAMEL, H.: *Understanding the impact of AI Hallucinations on the university community.* In: Cybrarians Journal, No. 73, Special Issue, 2024, p. 115.

76 European Broadcasting Union: *News Integrity in AI Assistants. An international PSM study*, October 2025. Online: https://www.ebu.ch/research/open/report/news-integrity-in-ai-assistants.

with superior analytical capabilities,[77] what may lead to blind reproduction of AI hallucinations in judicial decision-making process.

Courts already use AI to prepare draft rulings, and the problem arises when the prepared text is not verified by a specialist. Then due to hallucinations they can contain factual inaccuracies and other serious errors, including mistakes related to applicable law or cited case-law, which may even be non-existent. In the US two federal judges already admitted that their staff used artificial intelligence to help prepare some court orders, the drafts were not subject to typical review process, and the result was that two orders full of errors were delivered.[78]

If AI analyses the entire case-file it will work also on legal filings prepared by professional lawyers, who also can happen to use AI to prepare these fillings. Because of the hallucinations they can also be incorrect and refer to false case-law or make other errors and at the same time they can be suggestive for courts' AI. In the US some lawyers have already been sanctioned for not having verified texts prepared for them by AI.[79] If left unverified, such filings could cause further errors attributable either to a judge or to AI used by a judge to prepare the ruling. If not verified by human judges, AI may continue to repeat the same errors, making them constant practice, as AI lacks self-reflection and self-correction.[80] That is why the awareness of this problem and the issue of ethical use of these tools should be addressed among human judges.[81]

Conclusion

For the reasons mentioned in the previous part it can be concluded that AI can replace judges up to some extent. Some argue, that if a judicial decision is just, equal and impartial, it should not matter, whether it comes from AI or a

---

77 GRAVETT, W. H.: *Judicial Decision-Making in the Age of Artificial Intelligence*. In: ANTUNES, H. S.; FREITAS, P. M.; OLIVEIRA, A. L.; PEREIRA, C. M.; SEQUEIRA, E. V.; XAVIER, L. B. (eds.): Multidisciplinary Perspectives on Artificial Intelligence and the Law, Cham: Springer, 2024, pp. 290–291.

78 MERKEN, S.: *Two federal judges say use of AI led to errors in US court rulings.* In: Reuters, 23 October 2025. Online: https://www.reuters.com/sustainability/society-equity/two-federal-judges-say-use-ai-led-errors-us-court-rulings-2025-10-23/.

79 MERKEN, S.: *Some judges move beyond fines to keep lawyers' AI errors in check.* In: Reuters, 16 September 2025. Online: https://www.reuters.com/legal/government/some-judges-move-beyond-fines-keep-lawyers-ai-errors-check-2025-09-16/.

80 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, p. 1035.

81 KAMEL, H.: *Understanding the impact of AI Hallucinations on the university community.* In: Cybrarians Journal, No. 73, Special Issue, 2024, p. 112.

human judge.[82] AI can eliminate inconsistencies in judicial decisions, the automated process is fast and cheaper.[83] Provided for sufficient legal safeguards regarding its programming, AI can be also impartial.

A good example are civil judgments concerning mere payments. Such rulings are considered to be uncomplicated and repetitive activities, therefore AI could be used to replace unnecessary human labour, what would improve considerably the speed of the proceedings and make courts more effective and the procedures simplified. At the same time it would reduce the costs of the proceedings.[84] Another advantage would be consistent and predictable rulings.[85] Some authors even go further and emphasize that AI judge would eliminate prejudices that might characterize human judges, although at the same time they notice that if the input data contain biases, the output will be biased as well, therefore an adequate programming should solve this problem by making AI ignore certain category of data or prepare unbiased samples for AI training.[86]

At the same time there are proceedings which by their very nature escape the AI adjudication, like family or penal cases, or other complex cases that are difficult, with new problems to be solved. These should be decided by human judges in order to get a fair and acceptable result. If AI focuses on simple and repetitive cases, human judges can concentrate on these more complicated and nuanced.

It needs to be emphasized that the risk of dehumanizing the justice system should be avoided.[87] Yet AI is based on logical-mathematical intelligence,

---

82 VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, p. 1163, 1189–1190.

83 CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, p. 236–237; MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, p. 1102.

84 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 1027

85 CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol. 2, No. 1, 2020, p. 6.

86 VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1156–1157, 1168-1169; CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, p. 240. KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, p. 225–227.

87 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2227; GÓRSKI, M.: *Dziesiąty element. Sztuczna*

whereas it lacks emotional intelligence.[88] Thanks to AI formal justice can be achieved, but in some cases it will be extremely difficult to achieve substantial justice without elements of emotions or fairness, as well as empathy or compassion.[89] Therefore, judge follows social developments, whereas AI might be unable to apply legal norms in changing social context.[90] Scholars assume that public acceptance of the result of proceedings and the legitimacy of the court system would be higher with human judges.[91] Therefore even in cases where automatization is possible, a party to the proceedings should have the opportunity to appeal and to request the analysis of their case by a human judge.

In cases decided by human judges the use of AI is not completely excluded, as AI can be used to support the work of a judge, by documents and case-law searches, and including up to suggesting to them the text of a ruling and its justification[92] or provide for necessary data or assessments. The task of the judge would be then to review the prepared materials and deliver the judgment, as the judges should not rely blindly on the AI[93] and include also other factors, what should be reflected in the justification of judicial decision.[94] This causes a risk of shifting the burden of responsibility for the verdict from the judge to AI tool,[95]

---

inteligencja jako sędzia a prawo do sądu. In: Prawo Mediów Elektronicznych, No. 3, 2022, pp. 8–9.

88 KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, p. 234.

89 XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1026, 1040. Contrary view: VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, p. 1167; GÓRSKI, M.: *Dziesiąty element. Sztuczna inteligencja jako sędzia a prawo do sądu*. In: Prawo Mediów Elektronicznych, No. 3, 2022, pp. 9–10.

90 CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, p. 237.

91 MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, p. 1083.

92 NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, p. 2223; BARYSĖ, D., SAREL, R.: *Algorithms in the court: does it matter which part of the judicial decision-making is automated?* In: Artificial Intelligence and Law, Vol. 32, 2024, p. 118; XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, p. 1032. The author argues that it's the judgment that should be just and compassionate, not the judge, so it doesn't matter if the judge is AI or a human.

93 WALTERS, R.: *Robots replacing human arbitrators: the legal dilemma*. In: Information & Communications Technology Law, Vol. 34, No. 2, 2025, p. 135; CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, pp. 243–244.

94 Wisconsin Supreme Court: *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016).

95 GRAVETT, W. H.: *Judicial Decision-Making in the Age of Artificial Intelligence*. In: ANTUNES, H. S.; FREITAS, P. M.; OLIVEIRA, A. L.; PEREIRA, C. M.; SEQUEIRA, E. V.;

therefore human judges should always bear in mind that the ultimate decision is theirs, and not of the automated system. If human judges discover errors or hallucinations, they can provide for important feedback helping to improve AI algorithms and databases. AI can also contribute to judges' personal development, by drawing their attention to legal norms, case-law or interpretation which they have not encountered before. We need to keep in mind that delivering justice is not a mere application of legal norms, it must also include sociological context or the principles of equity, so there is still place for judicial discretion done by human judges.

Regardless of whether AI will become a judge or it will be just a tool supporting a judge in decision-making process, it needs to be remembered that like all computer programs it is susceptible to hacking and glitches exploitation. Some errors in coding or other glitches might also appear. Therefore, any such tools will require constant human supervision.[96]

XAVIER, L. B. (eds.): *Multidisciplinary Perspectives on Artificial Intelligence and the Law,* Springer: Cham, 2024, p. 291.

[96] VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1171–1177; CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, pp. 240–241.

List of Literature and Sources

1. BARYSĖ, D., SAREL, R.: *Algorithms in the court: does it matter which part of the judicial decision-making is automated?* In: Artificial Intelligence and Law, Vol. 32, 2024, pp. 117–146. Online: https://doi.org/10.1007/s10506-022-09343-6.
2. Cambridge Dictionary: *Moral*. Online: https://dictionary.cambridge.org/dictionary/english/moral#google_vignette.
3. CJEU: *Case C-64/16, Associação Sindical dos Juízes Portugueses v Tribunal de Contas*, Judgment of 27 February 2018.
4. CJEU: *Case C-585/18, A. K. and Others v Sąd Najwyższy, CP v Sąd Najwyższy and DO v Sąd Najwyższy*, Judgment of 19 November 2019.
5. CONTINI, F.: *Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings.* In: Law, Technology and Humans, Vol. 2, No. 1, 2020, pp. 4–18.
6. CROOTOF, R.: *"Cyborg justice" and the risk of technological-legal lock-in*. In: Columbia Law Review Forum, Vol. 119, No. 7, 2019, pp. 233–251.
7. ECtHR: *Atanasovski v. "the Former Yugoslav Republic of Macedonia"*, Application no. 36815/03, Judgment of 14 January 2010.
8. ECtHR: *Belilos v. Switzerland*, Application no. 10328/83, Judgment of 29 April 1988.
9. ECtHR: *Borg v. Malta*, Application no. 37537/13, Judgment of 12 January 2016.
10. ECtHR: *Coëme and Others v. Belgium*, Applications no. 32492/96, 32547/96, 32548/96, 33209/96 and 33210/96, Judgment of 22 June 2000.
11. ECtHR: *Daktaras v. Lithuania*, Application no. 42095/98, Judgment of 10 October 2000.
12. ECtHR: *Donadzé v. Georgia*, Application no. 74664/01, Judgment of 7 March 2006.
13. ECtHR: *Döry v. Sweden*, Application no. 28394/95, Judgment of 12 November 2002.
14. ECtHR: *Guðmundur Andri Ástráðsson v. Iceland*, Application no. 26374/18, Judgment of 1 December 2020.
15. ECtHR: *Holm v. Sweden*, Application no. 14191/88, Judgment of 25 November 1993.
16. ECtHR: *Ibrahim and Others v. the United Kingdom*, Applications no. 50541/08, 50571/08, 50573/08 and 40351/09, Judgment of 13 September 2016.

17. ECtHR: *Jussila v. Finland*, Application no. 73063/01, Judgment of 23 November 2006.
18. ECtHR: *Kilin v. Russia*, Application no. 10271/12, Judgment of 11 May 2021.
19. ECtHR: *Kontalexis v. Greece*, Application no. 59000/08, Judgment of 31 May 2011.
20. ECtHR: *Kraska v. Switzerland*, Application no. 13942/88, Judgment of 19 April 1993.
21. ECtHR: *Kyprianou v. Cyprus*, Application no. 73797/01, Judgment of 15 December 2005.
22. ECtHR: *Lavents v. Latvia*, Application no. 58442/00, Judgment of 28 November 2002.
23. ECtHR: *Lena Atanasova v. Bulgaria*, Application no. 52009/07, Judgment of 26 January 2017.
24. ECtHR: *Medenica v. Switzerland*, Application no. 20491/92, Judgment of 14 June 2001.
25. ECtHR: *Morice v. France*, Application no. 29369/10, Judgment of 23 April 2015.
26. ECtHR: *Ninn-Hansen v. Denmark*, Application no. 28972/95, Decision of 18 May 1999.
27. ECtHR: *Parlov-Tkalčić v. Croatia*, Application no. 24810/06, Judgment of 22 December 2009.
28. ECtHR: *Pretto and Others v. Italy*, Application no. 7984/77, Judgment of 8 December 1983.
29. ECtHR: *Reczkowicz v. Poland*, Application no. 43447/19, Judgment of 22 July 2021.
30. ECtHR: *Richert v. Poland*, Application no. 54809/07, Judgment of 25 October 2011.
31. ECtHR: *Riepan v. Austria*, Application no. 35115/97, Judgment of 14 November 2000.
32. ECtHR: *Şahiner v. Turkey*, Application no. 29279/95, Judgment of 25 September 2001.
33. ECtHR: *Sutter v. Switzerland*, Application no. 8209/78, Judgment of 22 February 1984.
34. ECtHR: *Unédic v. France*, Application no. 20153/04, Judgment of 18 December 2008.
35. ECtHR: *Vinčić and Others v. Serbia*, Application no. 44698/06, Judgment of 1 December 2009.
36. ECtHR: *Xero Flor w Polsce sp. z o.o. v. Poland*, Application no. 4907/18, Judgment of 7 May 2021.

37. European Broadcasting Union, *News Integrity in AI Assistants. An international PSM study*, October 2025. Online: https://www.ebu.ch/research/open/report/news-integrity-in-ai-assistants.
38. European Commission for the Efficiency of Justice: *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, adopted at the 31st plenary meeting of the CEPEJ, Strasbourg, 3 – 4 December 2018.
39. GÓRSKI, M.: *Dziesiąty element. Sztuczna inteligencja jako sędzia a prawo do sądu*. In: Prawo Mediów Elektronicznych, No. 3, 2022, pp. 4–10.
40. GÓRSKI, M.: *On the Right to a Human Judge in the Context of the Fair Trial Principle.* In: eucrim, Vol. 18, No. 1, 2023, pp. 83–88.
41. GRAVETT, W. H.: *Judicial Decision-Making in the Age of Artificial Intelligence*. In: ANTUNES, H. S.; FREITAS, P. M.; OLIVEIRA, A. L.; PEREIRA, C. M.; SEQUEIRA, E. V.; XAVIER, L. B. (eds.): *Multidisciplinary Perspectives on Artificial Intelligence and the Law,* Springer: Cham, 2024, pp. 281–297.
42. HOFMAŃSKI, P., WRÓBEL, A.: *Artykuł* 6. In: GARLICKI, L. (ed.): Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1-18*,* Warszawa 2010, pp. 241–461.
43. KAMEL, H.: *Understanding the impact of AI Hallucinations on the university community.* In: Cybrarians Journal, No. 73, Special Issue, 2024, pp. 111–134.
44. KASAP, G. H.: *Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications*. In: Journal of Dispute Resolution, Vol. 2021, No. 2, 2021, pp. 209–254.
45. LIMBERGER, T., BECK DA SILVA GIANNAKOS, D., SZINVELSKI, M. M.: *Can judges be replaced by machines? The Brazilian* case. In: Mexican Law Review, Vol. XIV, No. 2, 2022, pp. 55–81.
46. MERKEN, S.: *Some judges move beyond fines to keep lawyers' AI errors in check.* In: Reuters, 16 September 2025. Online: https://www.reuters.com/legal/government/some-judges-move-beyond-fines-keep-lawyers-ai-errors-check-2025-09-16/.
47. MERKEN, S.: *Two federal judges say use of AI led to errors in US court rulings.* In: Reuters, 23 October 2025. Online: https://www.reuters.com/sustainability/society-equity/two-federal-judges-say-use-ai-led-errors-us-court-rulings-2025-10-23/.
48. MICHAELS, A.C.: *Artificial Intelligence, Legal Change, and Separation of Powers*. In: University of Cincinnati Law Review, No. 88, Issue 4, 2020, pp. 1083–1103.

49. NOWOTKO, P. M.: *AI in judicial application of law and the right to a court.* In: Procedia Computer Science, No. 192, 2021, pp. 2220–2228.
50. PECH, L.: *Article 47(2) – An Independent and Impartial Tribunal Previously Established by Law*. In: PEERS, S., HERVEY, T., KENNER, J., WARD, J.: The EU Charter of Fundamental Rights. A Commentary. Oxford: Hart Publishing, 2021, pp. 1339–1357.
51. *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) (Text with EEA relevance),* OJ L, 2024/1689, 12.7.2024.
52. SALVAGNO, M.; TACCONE, F. S.; GERLI, A. G.: *Artificial intelligence hallucinations.* In: Critical Care, Vol. 27, Art. No. 180, 2023, pp. 1–2. Online: https://doi.org/10.1186/s13054-023-04473-y.
53. SLATER, J., HUMPHRIES, J.: *Another reason to call bullshit on AI "hallucinations"*. In: AI & Society, Vol. 40, 2025, pp. 5631–5632. Online: https://doi.org/10.1007/s00146-025-02346-2.
54. VOLOKH, E.: *Chief justice robots*. In: Duke Law Journal, Vol. 68, No. 6, 2019, pp. 1135–1192.
55. WALTERS, R.: *Robots replacing human arbitrators: the legal dilemma*. In: Information & Communications Technology Law, Vol. 34, No. 2, 2025, pp. 129–145.
56. XU, Z.: *Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities*. In: Applied Artificial Intelligence, Vol. 36, No. 1, 2022, pp. 1025–1045.

Author details

Dr hab. Magdalena MATUSIAK-FRĄCCZAK
University of Lodz
Faculty of Law and Administration
Department of European Constitutional Law
Email: mfracczak@wpia.uni.lodz.pl

3 AI-Powered Autonomous Weapons and the Protection of the Right to Life under International Law

Ayla Kaya Çalışkan

https://doi.org/10.33542/HRF-0541-2-03

Abstract

The growing use of artificial intelligence in modern warfare is transforming the application of lethal force and creating serious risks for the protection of the right to life. AI-powered weapon systems, capable of selecting and attacking targets with limited or no human control, may act on the basis of data and algorithms rather than human judgment. This shift can lead to arbitrary loss of life, especially when AI-powered systems mistakenly target civilians or behave unpredictably in situations where human operators are unable to monitor or intervene in time. The expanding role of AI-powered autonomous weapons in contemporary warfare risks weakening one of the most fundamental guarantees of international law, which is the prohibition of arbitrary deprivation of life. Existing legal frameworks, including the International Covenant on Civil and Political Rights, the European Convention on Human Rights, and the Geneva Conventions, all presuppose a human decision-maker who can assess necessity, proportionality, and distinction. When these judgements are delegated to machines, the traditional safeguards of legality and accountability are seriously weakened. The use of such systems also raises complex questions of accountability, as it often remains unclear whether responsibility for an unlawful deprivation of life should rest with the operator, the commander, the developer, or the State itself. The 2024 Report of the United Nations Secretary-General on Lethal Autonomous Weapons Systems underscores these dangers. It warns that allowing machines to autonomously target human beings crosses an unacceptable moral line and calls upon States to negotiate a legally binding instrument by 2026. Within the framework of the Convention on Certain Conventional Weapons (CCW), States have been discussing this issue since 2014. However, while the CCW does not currently prohibit the use or development of autonomous weapons, States, through the meetings of the Group of Governmental Experts (GGE), discuss possible legal and regulatory approaches to such systems. The paper proposes that future violations of the right to life can be prevented, and human dignity preserved in the era of autonomous warfare, only through a binding international framework that incorporates the principle of meaningful human control and encourages states to integrate this principle into their domestic legislation and defense- related policies. It further argues that in armed conflicts where AI-powered autonomous weapons are already being developed or deployed,

states should reconsider this approach and adopt the principle of meaningful human control as a fundamental safeguard for the right to life. By reinforcing human oversight in decision-making, the adoption of meaningful human control can also help reduce disputes over accountability for right-to-life violations caused by AI-powered autonomous weapons.

Introduction

The rapid integration of artificial intelligence (AI) into contemporary weapon systems, especially AI-powered autonomous weapon systems (AWS), has generated significant conceptual and normative challenges for international law by introducing machine-driven decision-making that may increasingly displace human judgment in target selection and engagement. Systems capable of independently selecting and attacking targets may challenge the long-standing legal assumption that decisions involving lethal force should be made by humans capable of contextual evaluation, moral reasoning, and legal judgment. Where such decisions are delegated to algorithmic processes, the effective protection of the right to life under international law may be placed under considerable strain.

International institutions and legal scholarship have repeatedly warned that increasing autonomy in weapon systems may undermine compliance with the principles governing the use of lethal force. In particular, concerns have been raised regarding the ability of AI-powered AWS to respect distinction, proportionality, and precautions in attack, as well as the implications of delegating life-and-death decisions to algorithmic processes. These concerns are not merely theoretical, as recent conflicts suggest that AI-supported targeting practices may contribute to civilian harm and weakened accountability.

This article argues that AI-powered AWS may pose structural risks to the protection of the right to life due to three interrelated factors: first, their limited capacity to apply the normative requirements embedded in international humanitarian law (IHL) and international human rights law (IHRL); second, the foreseeable risk of unlawful harm arising from algorithmic limitations, bias, and unpredictability; and third, the diffusion of responsibility across complex human-machine assemblages, which challenges traditional accountability frameworks. While existing bodies of international law remain formally applicable irrespective of technological change, the operational logic of AI-powered autonomy exposes gaps in enforcement, attribution, and accountability that current legal regimes were not designed to address.

The article proceeds in four parts. Section 1 conceptualizes autonomous and AI-powered AWS and clarifies their technical and operational

characteristics. Section 2 examines the legal protection of the right to life under IHRL and IHL. Section 3 analyses how autonomous and AI-supported systems may give rise to unlawful deprivations of life, both intentional and unintentional. Section 4 addresses accountability, focusing on state responsibility, individual criminal liability, and the emerging "accountability gap".

3.1 Autonomous Weapon Systems (AWS) and the Integration of Artificial Intelligence (AI)

AWS are commonly understood as weapons that, once activated, can select and engage targets without further human intervention. This functional approach, focusing on autonomy in critical targeting functions such as target identification, selection, and engagement, is reflected in institutional definitions and State practice. While the precise scope and terminology remain subject to debate, there is broad agreement that the increasing reliance on automated and algorithmic decision-making distinguishes AWS from traditional weapons systems.[1]

In the literature, AWS are often positioned along a spectrum of human involvement, commonly described as human-in-the-loop, human-on-the-loop, and human-out-of-the-loop systems.[2] These categories are widely used as

---

1 Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts (GGE) has emphasized that autonomy in identification, selection, and engagement constitutes the core feature of emerging technologies in the area of AWS. *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*. Online: https://documents.un.org/doc/undoc/gen/g14/048/96/pdf/g1404896.pdf; ICRC describes AWS as systems that process environmental data through their own sensors and software and can select and attack targets independently. ICRC: *Views on Autonomous Weapon System*, 11 April 2016. Online: https://www.icrc.org/sites/default/files/document/file_list/ccw-autonomous-weapons-icrc-april-2016.pdf; United States practice similarly AWS as system that, once activated, can select and engage targets without further intervention by an operator. Office of the Under Secretary of Defense for Policy: *DoD Directive 3000.09, Autonomy in Weapon Systems*, 25 January 2023. Online: https://www.congress.gov/crs-product/IF11150#:~:text=DODD%203000.09%20defines%20LAWS%20as,%22human%20on%20the%20loop%2C%22.

2 The UN Institute for Disarmament Research (UNIDIR): *The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches*, 2017, p. 20. Online: https://unidir.org/files/publication/pdfs/the-weaponization-of-increasingly-autonomous-technologies-concerns-characteristics-and-definitional-approaches-en-689.pdf.

descriptive tools to capture variations in the degree of human involvement as technological autonomy increases.

A further distinction may be drawn between "autonomous" and "AI-powered" systems. AI is not a prerequisite for autonomy, as some systems operate on pre-programmed rules. Where AI is integrated, however, it may expand autonomous capabilities by enabling systems to derive behaviour from data, adapt to changing environments, or generate decisions through probabilistic inference.[3]

The integration of machine learning introduces legally relevant risks. Advanced models may function as opaque "black boxes" that are difficult to interpret or predict, even for developers and operators.[4] Such opacity may undermine transparency, verification, and accountability, while errors or bias in training data may contribute to systematic misclassification and heightened risks of civilian harm.[5]

AI-powered AWS may also operate at machine speed, compressing human reaction time and narrowing practical opportunities for intervention once an engagement sequence has begun. These technical dynamics challenge foundational legal assumptions about foreseeability, verification, and human oversight in targeting, thereby intensifying concerns regarding compliance with IHL and the protection of life under IHRL.[6]

---

3 UN Office for Disarmament Affairs: *Lethal Autonomous Weapon Systems*. Online: https://disarmament.unoda.org/en/our-work/emerging-challenges/lethal-autonomous-weapon-systems.

4 SCHARRE, P.: *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018, pp. 104–181.

5 ICRC: *Views on Autonomous Weapon System*, 11 April 2016; SULLIVAN, S., RICKET, I.: *Targeting in the Black Box*. In.: KWAN, C., LINDSTRÖM, L., GIOVANNELLI, D., PODIŅŠ, K., ŠTRUCL, D. (eds.): CyCon 2024: Over the Horizon 16th International Conference on Cyber Conflict, 2024, pp. 208–209. Online: https://ccdcoe.org/uploads/2024/05/CyCon_2024_Sullivan_Ricket-1.pdf; International Human Rights Clinic: *"Killer Robots" Threaten Human Rights During War, Peace*, 12 May 2025. Online: https://humanrightsclinic.law.harvard.edu/killer-robots-threaten-human-rights-during-war-peace/.

6 SCHARRE, P.: *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018, pp. 104–115; BOULANIN V., VERBRUGGEN, M.: *Mapping the Development of Autonomy in Weapon Systems*. Solna: Stockholm International Peace Research Institute, 2017, pp. 65–82; UNIDIR: *The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches*, 2017, pp. 3–9; Human Rights Watch and The International Human Rights Clinic: *Losing Humanity: The Case Against Killer Robots*, 2012, pp. 30–37. Online: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots; UN General Assembly: *Lethal Autonomous Weapons Systems: Report of the*

Against this background, international regulatory momentum has increased. More than 120 States have expressed support for negotiating a new treaty on autonomous weapon systems.[7] The UN Secretary-General and the President of the International Committee of the Red Cross (ICRC) have called for a legally binding instrument by 2026 designed to preserve meaningful human control over the use of force.[8] Treaty proponents generally support banning autonomous weapon systems that operate without meaningful human control or that are designed to target human beings, while subjecting all other autonomous systems to strict requirements ensuring meaningful human control.[9] At the same time, CCW discussions since 2014 have produced limited substantive outcomes, in part due to the consensus-based decision-making structure.

From the perspective of the right to life, these developments are significant because the capacity of autonomous systems to reliably distinguish lawful from unlawful targets, take precautions in attack, and apply proportionality remains deeply contested. The prospect of developing an AWS capable of fully satisfying the requirements of both IHL and IHRL therefore appears, at present, highly uncertain.[10]

3.2 The Protection of the Right to Life under International Law

The right to life is recognised as a fundamental and non-derogable human right, applicable in both peacetime and armed conflict. Accordingly, the use of autonomous weapon systems raises concerns for the protection of life under both IHRL and IHL.[11]

Secretary-General, A/RES/79/88, 1 July 2024, pp. 7–12. Online: https://docs.un.org/en/A/79/88.

7 Automated Decision Research. Online: https://automatedresearch.org/state-ICRCs/.

8 United Nations Secretary-General and International Committee of the Red Cross: *Joint Call for States to Establish New Prohibitions and Restrictions on Autonomous Weapon Systems*, 5 October 2023. Online. https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems.

9 Stop Killer Robots Campaign: *Key Elements of a Treaty on Fully Autonomous Weapons: Frequently Asked Questions*, 2020. Online: https://www.stopkillerrobots.org/wp-content/uploads/2020/03/FAQ-Treaty-Elements.pdf.

10 Geneva Academy of International Humanitarian Law and Human Rights: *Sending Up a Flare: Autonomous Weapons Systems Proliferation Risks to Human Rights and International Security*, Research Brief, July 2024, p. 4. Online: https://geneva-academy.ch/wp-content/uploads/2025/09/Sending-Up-a-Flare-Autonomous-Weapons-Systems-Proliferation-Risks.pdf.

11 BREHM, M.: *Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems under International Humanitarian Law and International Human Rights Law*. In: Geneva Academy Briefing, No. 9, 2017, p. 54. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972071; ICRC: *ICRC*

3.2.1 The Right to Life under IHRL

The right to life is protected under international law as a fundamental and universal human right and is enshrined in a range of global and regional human rights instruments. In this regard, Article 6 of the International Covenant on Civil and Political Rights (ICCPR) and Article 2 of the European Convention on Human Rights (ECHR) constitute some of the principal provisions establishing States' obligations to respect and protect the right to life. These provisions impose negative obligations on states to refrain from arbitrary deprivation of life and positive obligations to take appropriate measures to safeguard life within their jurisdiction.[12] As emphasized by the Human Rights Committee, the right to life is the supreme right from which no derogation is permitted, even in situations of armed conflict or public emergency.[13]

Under IHRL, the use of lethal force by state agents is permissible only in exceptional circumstances. Any deprivation of life must satisfy strict requirements of necessity and proportionality. Lethal force may be used only as a last resort and only where it is unavoidable to protect life. The level of force used must be strictly proportionate to the legitimate aim pursued.[14] These assessments presuppose human judgment capable of interpreting intent, evaluating rapidly evolving situations, and exercising restraint.[15]

Position on Autonomous Weapon Systems, 12 May 2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.

12 HEYNS, CH.: Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, A/HRC/23/47, 9 April 2013, p. 6. Online: https://digitallibrary.un.org/record/755741; HEYNS, CH.: *Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement*. In: Human Rights Quarterly, Vol. 38, No. 2, 2016, p. 362.

13 Human Rights Committee: *General Comment No. 36: Article 6: right to life*, CCPR/C/GC/36, 3 September 2019, p. 1, https://docs.un.org/en/CCPR/C/GC/36.

14 ECtHR: *McCann and Others v. the United Kingdom*, Application No. 18984/91, Judgment of 27 September 1995, para. 149; ECtHR: *Mendy v. France* (dec.), Application No. 71428/12, Judgment of 4 September 2018, para. 31–33; Human Rights Committee: *General Comment No. 36: Article 6: right to life*, CCPR/C/GC/36, 3 September 2019, para. 12.

15 Human Rights Committee: *General Comment No. 36: Article 6: right to life*, CCPR/C/GC/36, 3 September 2019, para. 14. See also Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders, Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, 7 September 1990, para. 14, Online: https://www.ohchr.org/en/instruments-mechanisms/instruments/basic-principles-use-force-and-firearms-law-enforcement; ECtHR: *Giuliani and Gaggio v. Italy*, Application No. 23458/02, Judgment of 24 March 2011, para. 174–179.

IHRL does not rely on a combatant-civilian distinction. Instead, it protects all individuals as rights-holders and evaluates the lawfulness of lethal force through individualized assessments.[16] This feature is particularly relevant when assessing the compatibility of autonomous and AI-powered AWS with the right to life.

3.2.2 The Right to Life under IHL

In situations of armed conflict, the protection of life is primarily governed by IHL. While IHL does not prohibit killing as such, it strictly regulates the conduct of hostilities through the 1949 Geneva Conventions and their Additional Protocols. The lawful use of lethal force under IHL is conditioned upon compliance with fundamental principles designed to limit civilian harm.

The principle of distinction obliges parties to an armed conflict to distinguish at all times between civilians and combatants. Autonomous systems may be limited in their ability to interpret contextual and behavioural cues, such as surrender or hors de combat status.[17] AI-based classification relies on pattern recognition rather than normative reasoning, which may increase the risk of misidentification in complex environments.

Under Additional Protocol I, the principles of proportionality[18] and precautions[19] in attack further require verification of targets, minimization of civilian harm, and cancellation or suspension of attacks where doubt arises. Autonomous systems, however, do not recognise doubt or reassess attacks in light of new information in the manner required by IHL.

3.2.3 The Relationship Between IHRL and IHL

It is now well established that IHRL continues to apply during armed conflict; however, whether a particular deprivation of life is arbitrary must be assessed by reference to IHL as the *lex specialis* governing the conduct of hostilities.[20] This complementary relationship is particularly relevant when assessing new means and methods of warfare.

---

16 Human Rights Committee: *General Comment No. 36: Article 6: right to life*, CCPR/C/GC/36, 3 September 2019, para. 3; ECtHR: *McCann and Others v. the United Kingdom*, Application No. 18984/91, Judgment of 27 September 1995, para. 146.

17 ICRC: *Views on Autonomous Weapon System*, 11 April 2016, p. 4.

18 Protocol Additional call to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts, (Additional Protocol I), 8 June 1977, Article 51(5)(b).

19 Additional Protocol I, Article 57.

20 International Court of Justice: *The Legality of the Threat or Use of Nuclear Weapons,* Advisory Opinion of 8 July 1996, para. 25; International Court of Justice: *Legal*

3.2.4 Implications for Autonomous and AI-powered Weapon Systems

Both IHRL and IHL are premised on the assumption that decisions involving lethal force are made by humans capable of contextual interpretation, moral reasoning, and legal judgment. Autonomous and AI-powered weapon systems challenge this assumption at a structural level. Their reliance on algorithmic decision-making and sensor data raises concerns as to their ability to comply with the legal requirements governing the protection of life, particularly where lethal force is applied without meaningful human control.[21]

3.3 AI-powered Autonomous Weapon Systems and Violations of the Right to Life

AI-powered AWS may give rise to serious risks of violations of the right to life due to both their technical limitations and foreseeable operational risks.

Where States knowingly use systems that are known to be incapable of reliably complying with the principles of distinction, proportionality, or precautions in attack, or where such systems are reasonably foreseeable to entail serious risks in this regard, resulting civilian deaths may constitute arbitrary deprivations of life under IHL.[22]

Recent conflicts provide reported illustrations of these risks in practice. Publicly available investigations concerning the use of AI-supported target-

Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion of 9 July 2004, para. 106; Study Group of the International Law Commission: *Fragmentation of International Law: Difficulties Arising from Diversification and Expansion of International Law*, A/CN.4/L.682, 13 April 2006, para. 104. Online: https://digitallibrary.un.org/record/574810?v=pdf&ln=en.

21 HEYNS, CH.: *Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement*. In: Human Rights Quarterly, Vol. 38, No. 2, 2016, p. 366; UN General Assembly: Lethal Autonomous Weapons Systems: Report of the Secretary-General, A/RES/79/88, 1 July 2024, p. 11; ICRC: *Views on Autonomous Weapon System*, 11 April 2016, pp. 3–4. Online: https://www.icrc.org/sites/default/files/document/file_list/ccw-autonomous-weapons-icrc-april-2016.pdf.

22 HEYNS, CH.: *Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement*. In: Human Rights Quarterly, Vol. 38, No. 2, 2016, pp. 365–366; CREYGHTON, R.: *The Lavender Precedent: Automated Kill Lists and The Limits of International Humanitarian Law, Action on Armed Violence*, 6 November 2025. Online: https://aoav.org.uk/2025/the-lavender-precedent-automated-kill-lists-and-the-limits-of-international-humanitarian-law/ICRC; ICRC: *ICRC Position on Autonomous Weapon Systems*, 12 May 2021.

generation systems in Gaza suggest that individuals were designated as military objectives on the basis of automated assessments with known error rates, while human oversight was reportedly limited to minimal or perfunctory review.[23] Such practices may point not merely to technical malfunction, but to systemic decision-making structures that risk deprioritizing civilian protection. In these circumstances, the resulting loss of life may not be readily characterized as purely accidental. Rather, it may reflect a decision to rely on systems known to carry a heightened risk of unlawful outcomes, potentially engaging responsibility for serious violations of the right to life. The UN Secretary-General has highlighted that autonomous and AI-powered AWS may lead to violations of the right to life not only through deliberate use, but also as a result of technical failures, misidentification, unpredictability, and the inability of humans to intervene in time.[24]

Beyond that, AI-powered AWS may give rise to violations of the right to life through a range of technical and operational factors. Autonomous systems rely on sensor inputs including visual, thermal, radar, or signal data that may be distorted by environmental conditions such as dust, weather, electromagnetic interference, or structural obstacles.[25] Errors in perception or data interpretation may lead the system to misidentify civilian objects or persons as military targets.[26] The speed at which autonomous systems operate may also effectively eliminate opportunities for timely human intervention.[27]

---

23 ABRAHAM, Y.: *'Lavender': The AI Machine Directing Israel's Bombing Spree in Gaza*. In: +972 Magazine, 3 April 2024. Online: https://www.972mag.com/lavender-ai-israeli-army-gaza/.

24 UN General Assembly: *Lethal Autonomous Weapons Systems: Report of the Secretary-General*, A/RES/79/88, 1 July 2024, para.44–46.

25 ICRC: *Views on Autonomous Weapon System*, 11 April 2016, pp. 3, 6–9; SCHARRE, P.: *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018, p. 149; 34. UN Institute for Disarmament Research (UNIDIR): *The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches*, 2017, p. 14.

26 ICRC: *Views on Autonomous Weapon System*, 11 April 2016, p. 5.

27 SCHARRE, P.: *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018, pp. 104–115; BOULANIN V., VERBRUGGEN, M.: Mapping the Development of Autonomy in Weapon Systems. Solna: Stockholm International Peace Research Institute, 2017, p. 8; UNIDIR: *The Weaponization of Increasingly Autonomous Technologies: Considering How Meaningful Human Control Might Move the Discussion Forward*, 2014, p. 3. Online: https://unidir.org/files/publication/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf; 12. Human Rights Watch and The International Human Rights Clinic: Losing Humanity: The Case Against Killer Robots, 2012, p. 37. Online: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.

Additionally, bias in training data may lead to discriminatory targeting practices. Algorithms reflect the data and design choices on which they are based. Where datasets fail to adequately capture civilian behaviour or incorporate social, ethnic, or political biases, autonomous systems may disproportionately impact certain populations.[28] Once deployed, they act on predefined objectives without contextual reassessment, potentially undermining respect for human dignity and the value of life.

In such cases, the absence of intent does not preclude a finding of arbitrariness under IHL, particularly where lethal force is applied without the capacity for contextual assessment or meaningful human judgment.

3.4 Accountability for Right-to-Life Violations

Autonomous decision-making in AI-powered AWS complicates accountability under both IHL and IHRL by making it more difficult to link unlawful killings to identifiable human decision-makers or States. While state responsibility remains applicable under the Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA), the deployment of autonomous systems may heighten States' due diligence obligations in light of the increased risks involved. Where an autonomous weapon system is deployed by State organs or under their authority, the conduct of that system is attributable to the State, regardless of the degree of autonomy involved.[29] State responsibility may also arise where a State deploys an autonomous weapon system that has not been adequately tested, reviewed, or assessed for compliance with international law prior to deployment.[30]

Individual criminal responsibility is not excluded; however, its attribution may be rendered more complex by the increasing integration of AI-powered AWS.[31] Command responsibility may arise where commanders knowingly deploy systems posing foreseeable risks to civilians. Operator responsibility may

---

28 See MILANINIA, N.: *Biases in Machine Learning Models and Big Data Analytics: The International Criminal and Humanitarian Law Implications*. In: International Review of the Red Cross Vol. 102, No. 913, 2020, p. 217.

29 DAVISON, N.: *A Legal Perspective: Autonomous Weapon Systems under International Humanitarian Law*. In: UNODA Occasional Papers No. 30, November 2017, p. 16. Online: https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf; HAMMOND, D. N.: *Autonomous Weapons and the Problem of State Accountability*. In: Chicago Journal of International Law, Vol. 15, No. 2, 2015, pp. 668–670.

30 ICRC: *Views on Autonomous Weapon System*, 11 April 2016, p. 5.

31 *Ibid*.

attach where available oversight is neglected or where reliance on automated outputs prevails over critical human judgment. Developer responsibility may arise only in limited circumstances, particularly where individuals knowingly contribute to unlawful outcomes.[32]

The combined effect of technical opacity and distributed human involvement gives rise to an emerging accountability gap, a problem that is further exacerbated by the difficulty of establishing the requisite mens rea, as AI-mediated decision-making may obscure intent and knowledge even where unlawful outcomes were foreseeable.[33] This accountability gap undermines not only deterrence but also the procedural dimension of the right to life, including the right to an effective investigation and remedy.

Conclusion

AI-powered AWS exposes growing tensions between emerging military technologies and the international legal framework governing the protection of the right to life. By delegating critical targeting and engagement decisions to algorithmic processes, such systems challenge foundational assumptions shared by IHL and IHRL, particularly the premise that decisions involving lethal force are subject to contextual human judgment and legal accountability. Reliance on automated, data-driven decision-making raises persistent concerns regarding compliance with the principles of distinction, proportionality, and precautions in attack, as well as the continued applicability of the prohibition of arbitrary deprivation of life.

Beyond the immediate risks of unlawful civilian harm, the increasing autonomy of weapon systems also complicates accountability mechanisms

---

[32] SPADARO, A.: *A Weapon is No Subordinate Autonomous Weapon Systems and the Scope of Superior Responsibility*. In: Journal of International Criminal Justice, Vol. 21, No. 5, November 2023, p. 1131; BO, M., BRUUN, L., BOULANIN, V.: *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS*, SIPRI, 2022, pp. 34–39.

[33] Human Rights Watch and The International Human Rights Clinic: *Mind the Gap: The Lack of Accountability for Killer Robots*, 2015, p. 18. Online: https://www.hrw.org/sites/default/files/reports/arms0415_ForUpload_0.pdf; DAVISON, N.: *A Legal Perspective: Autonomous Weapon Systems under International Humanitarian Law*. In: UNODA Occasional Papers No. 30, November 2017, p. 17. See, CHENGETA, T.: *Accountability Gap, Autonomous Weapon Systems and Modes of Responsibility in International Law*. In: Denver Journal of International Law&Policy, Vol. 45, No. 1, 2016, p. 24.

central to the protection of the right to life. Technical opacity and distributed human involvement make it more difficult to attribute responsibility at both the State and individual levels, thereby undermining deterrence and the procedural obligations associated with effective investigation and remedy. In this context, preserving meaningful human control over the use of lethal force emerges not merely as a policy preference, but as a normative requirement grounded in respect for human dignity and the effective protection of the right to life under international law. Absent meaningful human control, the delegation of lethal decision-making to machines risks transforming the prohibition of arbitrary deprivation of life from a legal guarantee into a technical aspiration.

List of Literature and Sources

1. ABRAHAM, Y.: *'Lavender': The AI Machine Directing Israel's Bombing Spree in Gaza*. In: +972 Magazine, 3 April 2024. Online: https://www.972mag.com/lavender-ai-israeli-army-gaza/.
2. Automated Decision Research. Online: https://automatedresearch.org/state-ICRCs/.
3. BO, M., BRUUN, L., BOULANIN, V.: *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS*, SIPRI, 2022, 61 p.
4. BOULANIN V., VERBRUGGEN, M.: *Mapping the Development of Autonomy in Weapon Systems*. Solna: Stockholm International Peace Research Institute, 2017.
5. BREHM, M.: *Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems under International Humanitarian Law and International Human Rights Law*. In: Geneva Academy Briefing, No. 9, 2017, 72 p. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972071.
6. CHENGETA, T.: *Accountability Gap, Autonomous Weapon Systems and Modes of Responsibility in International Law*. In: Denver Journal of International Law&Policy, Vol. 45, No. 1, 2016, pp. 1–50.
7. CREYGHTON, R.: *The Lavender Precedent: Automated Kill Lists and The Limits of International Humanitarian Law, Action on Armed Violence*, 6 November 2025. Online: https://aoav.org.uk/2025/the-lavender-precedent-automated-kill-lists-and-the-limits-of-international-humanitarian-law/ICRC.
8. DAVISON, N.: *A Legal Perspective: Autonomous Weapon Systems under International Humanitarian Law.* In: UNODA Occasional Papers No. 30, November 2017. Online: https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf.
9. ECtHR: *Giuliani and Gaggio v. Italy*, Application no. 23458/02, Judgment of 24 March 2011.
10. ECtHR: *McCann and Others v. the United Kingdom*, Application no. 18984/91, 27 September 1995.
11. ECtHR: *Mendy v. France (dec.)*, Application No. 71428/12, Decision of 4 September 2018.
12. Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders: *Basic Principles on the Use of Force and Firearms by*

Law Enforcement Officials, 07 September 1990. Online: https://www.ohchr.org/en/instruments-mechanisms/instruments/basic-principles-use-force-and-firearms-law-enforcement.

13. Geneva Academy of International Humanitarian Law and Human Rights: *Sending Up a Flare: Autonomous Weapons Systems Proliferation Risks to Human Rights and International Security*, Research Brief, July 2024, 16 p. Online: https://geneva-academy.ch/wp-content/uploads/2025/09/Sending-Up-a-Flare-Autonomous-Weapons-Systems-Proliferation-Risks.pdf.
14. HAMMOND, D. N.: *Autonomous Weapons and the Problem of State Accountability*. In: Chicago Journal of International Law, Vol. 15, No. 2, 2015, pp. 652–687.
15. HEYNS, CH.: *Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement*. In: Human Rights Quarterly, Vol. 38, No. 2, 2016, pp. 350–378.
16. HEYNS, CH.: *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, A/HRC/23/47, 9 April 2013. Online: https://digitallibrary.un.org/record/755741.
17. Human Rights Committee: *General Comment No. 36: Article 6: right to life*, CCPR/C/GC/36, 3 September 2019. Online: https://docs.un.org/en/CCPR/C/GC/36.
18. Human Rights Watch and The International Human Rights Clinic: *Losing Humanity: The Case Against Killer Robots*, 2012. Online: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.
19. Human Rights Watch and The International Human Rights Clinic: *Mind the Gap: The Lack of Accountability for Killer Robots*, 2015. Online: https://www.hrw.org/sites/default/files/reports/arms0415_ForUpload_0.pdf.
20. ICRC: *ICRC Position on Autonomous Weapon Systems*, 12 May 2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.
21. ICRC: *Views on Autonomous Weapon System*, 11 April 2016. Online: https://www.icrc.org/sites/default/files/document/file_list/ccw-autonomous-weapons-icrc-april-2016.pdf.
22. International Court of Justice: *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion of 9 July 2004.
23. International Court of Justice: The Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion of 8 July 1996.

24. International Human Rights Clinic: *"Killer Robots" Threaten Human Rights During War, Peace*, 2025. Online: https://humanrightsclinic.law.harvard.edu/killer-robots-threaten-human-rights-during-war-peace/.
25. *Joint Call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to Establish New Prohibitions and Restrictions on Autonomous Weapon Systems*. Online: https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems.
26. MILANINIA, N.: *Biases in Machine Learning Models and Big Data Analytics: The International Criminal and Humanitarian Law Implications*. In: International Review of the Red Cross Vol. 102, No. 913, 2020, pp. 199–234.
27. Office of the Under Secretary of Defense for Policy: *DoD Directive 3000.09*, Autonomy in Weapon Systems, 25 January 2023. Online: https://www.congress.gov/crs-product/IF11150#:~:text=DODD%203000.09%20defines%20LAWS%20as,%22human%20on%20the%20loop%2C%22.
28. Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts, (Additional Protocol I), 8 June 1977.
29. *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*. Online: https://documents.un.org/doc/undoc/gen/g14/048/96/pdf/g1404896.pdf.
30. SCHARRE, P.: *Army of None: Autonomous Weapons and the Future of War*. New York: W. W. Norton & Company, 2018, 481 p.
31. SPADARO, A.: *A Weapon is No Subordinate Autonomous Weapon Systems and the Scope of Superior Responsibility*. In: Journal of International Criminal Justice, Vol. 21, No. 5, November 2023, pp. 1119–1136.
32. Stop Killer Robots Campaign: *Key Elements of a Treaty on Fully Autonomous Weapons: Frequently Asked Questions*, 2020. Online: https://www.stopkillerrobots.org/wp-content/uploads/2020/03/FAQ-Treaty-Elements.pdf.
33. Study Group of the International Law Commission: *Fragmentation of International Law: Difficulties Arising from Diversification and Expansion of International Law*, A/CN.4/L.682, 13 April 2006. Online: https://digitallibrary.un.org/record/574810?v=pdf&ln=en.
34. SULLIVAN, S., RICKET, I.: *Targeting in the Black Box*. In.: KWAN, C., LINDSTRÖM, L., GIOVANNELLI, D., PODIŅŠ, K., ŠTRUCL, D. (eds.): CyCon

2024: Over the Horizon 16th International Conference on Cyber Conflict, 2024, pp. 207 – 220. Online: https://ccdcoe.org/uploads/2024/05/CyCon_2024_Sullivan_Ricket-1.pdf.

35. UN General Assembly: *Lethal Autonomous Weapons Systems: Report of the Secretary-General*, A/RES/79/88, 1 July 2024.
36. UN Institute for Disarmament Research (UNIDIR): *The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches*, 2017. Online: https://unidir.org/files/publication/pdfs/the-weaponization-of-increasingly-autonomous-technologies-concerns-characteristics-and-definitional-approaches-en-689.pdf.
37. UN Office for Disarmament Affairs*: Lethal Autonomous Weapon Systems*. Online: https://disarmament.unoda.org/en/our-work/emerging-challenges/lethal-autonomous-weapon-systems.
38. UNIDIR: *The Weaponization of Increasingly Autonomous Technologies: Considering How Meaningful Human Control Might Move the Discussion Forward*, 2014. Online: https://unidir.org/files/publication/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf.

Author details

Asst. Prof. Ayla KAYA ÇALIŞKAN
Anadolu University
Faculty of Law
Department of Human Rights Law
Email: aylakayacaliskan@anadolu.edu.tr

4 AI-Assisted Cyberattacks on Privacy: Challenges for Private International Law

Ľubica Gregová Širicová

https://doi.org/10.33542/HRF-0541-2-04

Abstract

This article examines the challenges posed by AI-assisted infringements of personality rights for the system of private international law within the European Union. It argues that the existing jurisdictional framework, particularly under the Brussels I Regulation (recast), is not fully adapted to address cross-border harms that are decentralised and simultaneously distributed across multiple jurisdictions. The analysis focuses on jurisdictional issues in civil liability claims, with particular attention to the case-law of the Court of Justice of the European Union and the application of Slovak private international law. It shows that traditional connecting factors, such as the defendant's domicile and the localisation of harm, may not always ensure effective access to justice in cases involving AI-driven privacy violations and defamation. The article further highlights the tension between these limitations and the positive obligations arising under Article 8 of the European Convention on Human Rights. It concludes by suggesting the need for a more coherent and adaptable approach to jurisdiction in the digital environment.[1]

Introduction

AI-assisted cyberattacks are reshaping the landscape of infringements of personality rights. AI-assisted operations, deepfake identity spoofing, large-scale data scraping, and algorithmic surveillance enable privacy intrusions that are increasingly borderless, decentralised, and fast-moving. While substantive areas of law — including data-protection and cybersecurity law — attempt to address these harms, Private International Law (PIL) remains behind to offer suitable rules for jurisdiction, applicable law, and enforceability of judgments in these situations. Notoriously, PIL struggles to address harm that occurred via internet.[2] This effect is magnified when AI is used as a vehicle for infringements of personality rights including privacy or defamation. The result is a widening gap between the scope of violations and the capacity of PIL to respond.

---

1 The paper presents a partial output within the research project *Kompetenčné centrum kybernetickej bezpečnosti na Univerzite Pavla Jozefa Šafárika v Košiciach (KC KB UPJŠ) 17R05-04-V01-00007*.

2 KOHL, U.: *Jurisdiction and the Internet – Regulatory Competence over Online Activity*. Cambridge University Press, 2007, ppb. 2010, p. 2. Online: https://ssrn.com/abstract=4169834.

The difficulty of existing PIL instruments to cope with these harms stands in tension with the right to respect for private life under Article 8 of the European Convention on Human Rights (ECHR), which imposes positive State obligations to secure privacy rights not only against public authorities, but also against interferences by private actors.[3] The European Court of Human Rights (ECtHR) has repeatedly held that Article 8 would be rendered illusory if States did not provide adequate legal remedies where individuals suffer serious interferences with their private life caused by non-state actors.[4] Yet in cross-border cyber incidents, individuals frequently cannot obtain a remedy not because privacy law is absent, but because PIL rules defeat the ability to sue or enforce a judgment.

This paper argues that AI-assisted privacy attacks expose a structural mismatch between technological reality and factors used as jurisdictional criteria in PIL. Regulation Brussels Ia relies on defendant domicile in the EU and identifiable territorial links, assumptions which no longer hold where perpetrators are located outside of the EU or where harm occurs via internet simultaneously in multiple States. To maintain the effectiveness of Article 8, PIL must ensure that private-law remedies against cross-border AI harms to personality rights are not illusory in practice.

This article proceeds from the premise that the current framework of private international law is structurally ill-suited to address AI-driven harms to personality rights. The analysis is limited in scope to questions of civil liability within the EU private international law regime, focusing on disputes between private parties, namely individuals and corporate actors. It does not extend to issues of criminal liability, State responsibility, or the enforcement of data protection rules under the GDPR.

---

3 ECtHR: *X and Y v the Netherlands*, Application no. 8978/80, Judgment of 26 March 1985, para 23: „*The Court recalls that although the object of Article 8 (art. 8) is essentially that of protecting the individual against arbitrary interference by the public authorities, it does not merely compel the State to abstain from such interference: in addition to this primarily negative undertaking, there may be positive obligations inherent in an effective respect for private or family life (see the Airey judgment of 9 October 1979, Series A no. 32, p. 17, para. 32). These obligations may involve the adoption of measures designed to secure respect for private life even in the sphere of the relations of individuals between themselves*."

4 ECtHR: *López Ribalda and Others v Spain*, Application nos. 1874/13 and 8567/13, Judgment of 17 October 2019, para 135; ECtHR: *Söderman v Sweden*, Application no. 5786/08, Judgment of 12 November 2013, paras 78–85.

4.1 The Nature of AI-Assisted Cyberattacks and Privacy Harms

AI alters both the scale and quality of cross-border privacy violations. Four characteristics are particularly relevant for PIL analysis.

4.1.1 Automation and Scalability

AI enables automated deployment of privacy-invasive operations at a larger scale than would be feasible manually.[5] Automated scraping tools, deepfake generation, and AI-assisted phishing scripts[6] can potentially affect thousands of individuals across multiple jurisdictions within short timeframes. While these operations often involve human setup, the rapid distribution and replication amplify potential harm. This makes privacy injury inherently multi-territorial, undermining the assumption that harm can be located in a single State.

4.1.2 Anonymisation and Decentralisation

AI systems often rely on distributed infrastructure — including cloud servers, decentralised storage networks, and routing systems that obscure IP origins.[7] Attribution to a natural or legal person is often challenging in real time. Where the defendant's identity cannot be established, defendant-based jurisdictional rules such as Article 4 Brussels Ia may be inapplicable, creating procedural gaps.

4.1.3 AI-Enhanced Personalisation

AI tools analyse datasets to generate content tailored to individuals — including personal history, social networks, and behavioural patterns.[8] Privacy attacks using AI are therefore targeted rather than random.

---

5 HAKEEMAT, J.: *Harnessing AI for data privacy: Examining risks, opportunities and strategic future directions*. In: International Journal of Science and Research Archive, 2024, Vol. 13, No. 02, p. 2879. Online: DOI: 10.30574/ijsra.2024.13.2.2510.

6 JABIR, R., LE, J., NGUYEN, C.: *Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors*. In: AI 2025, Vol. 6, 174, pp. 11–12. Online: https://doi.org/ 10.3390/ai6080174.

7 ABU-ZAID, R., HAMMAD, A.: *Applications of AI in Decentralized Computing Systems: Harnessing Artificial Intelligence for Enhanced Scalability, Efficiency, and Autonomous Decision-Making in Distributed Architectures*. In: Journal of Artificial Intelligence & Cloud Computing, Vol. 7, No. 6, 2024, p. 165. Online: https://www.researchgate.net/publication/386219179_Applications_of_AI_in_Decentralized_Computing_Systems_Harnessing_Artificial_Intelligence_for_Enhanced_Scalability_Efficiency_and_Autonomous_Decision-Making_in_Distributed_Architectures.

8 SADIKU, M. N., ASHAOLU, T. J., AJAYI-MAJEBI, A., & MUSA, S. M.: *Artificial Intelligence in Social Media*. In: International Journal of Scientific Advances, 2021, Vol. 2, No. 1, pp. 15–20. Online: DOI: 10.51542/ijscia.v2i1.4.

4.1.4 Simultaneous Dissemination and Chain Harms

Once deployed, harmful online content may be amplified at scale through automated mechanisms, including social bots and algorithmic recommendation systems, which have been shown to play a significant role in the dissemination of low-credibility or harmful information.[9] Harm can persist even after the original content is removed. Traditional PIL concepts, which assume a single-place event of harm, are ill-suited to these dynamics.

4.1.5 Consequence for Jurisdictional Rules PIL

Efforts to obtain compensation for damage in cases involving a foreign element through civil proceedings necessarily entail the application of the rules of private international law. As a first step, the claimant must consider which State's courts may be seized of the claim for damages; likewise, the court must assess whether it has international jurisdiction to hear the case.

The following analysis focuses on the issue of international jurisdiction of courts in cases concerning infringements of personality rights with a foreign element, from the perspective of the courts of the Slovak Republic. Particular attention is given to infringements carried out via the Internet with the use of artificial intelligence. The applicability of the Brussels Ia Regulation[10] as well as of the domestic legal sources of the Slovak Republic, will be examined.

4.2 Jurisdictional Challenges under Brussels Ia Regulation in Personality Rights Claims

A key preliminary issue in applying the Regulation is its personal scope:[11] does the defendant have a domicile in a Member State of the European Union?

---

9 SHAO, C., CIAMPAGLIA, G.L., VAROL, O. et al: *The spread of low-credibility content by social bots*. Nat Commun 9, 4787 (2018), p. 5. Online: https://doi.org/10.1038/s41467-018-06930-7.

10 Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast), OJ L 351, 20.12.2012, pp. 1–32 (hereinafter referred to as the "Brussels Ia Regulation").

11 This understanding of the personal scope of the Regulation is shared by Lazić and Mankowski (LAZIĆ, V., MANKOWSKI, P.: *The Brussels I-bis Regulation: Interpretation and Implementation*. Edward Elgar Publishing Limited, 2023, p. 54). By contrast, Magnus states that the Regulation does not require specific personal conditions, arguing that the basic rule that the defendant must be domiciled in a Member State rather concerns the territorial connection between the dispute and the court (MAGNUS, U.: *Introduction*. In: MAGNUS, U., MANKOWSKI, P.: Brussels Ibis Regulation Commentary. Köln: Verlag Dr. Otto Schmidt KG., p. 36).

In the case of companies or other legal persons, or associations of natural or legal persons, domicile is defined in Article 63(1) by three alternative connecting factors—statutory seat, central administration, or principal place of business—any one of which is sufficient. Article 63(2) further clarifies the concept of "seat" for jurisdictions that do not recognise this notion (namely Cyprus and Ireland).

With regard to natural persons, however, the Regulation expressly refers the concept of domicile to national legal systems (Article 62). As a result, the notion is not interpreted uniformly or autonomously, but rather according to the law of each Member State. In the case of the Slovak Republic, the interpretation of domicile for the purposes of the Brussels I Regulation (recast) is based on the case law of the Supreme Court of the Slovak Republic,[12] which has held that domicile is the place where a person actually resides on a permanent basis, even if they are formally registered as having permanent residence in the jurisdiction of another court. The Court further emphasised that the concepts of domicile and permanent residence cannot be equated, as a natural person may not in fact reside at their officially registered place of residence.

The purpose of this article is to examine infringements of personality rights involving the use of artificial intelligence. It proceeds on the assumption that the claim will be directed against a natural or legal person exercising control over the relevant AI application and should therefore bear responsibility for the damage caused.

4.2.1 Defendant Domiciled in a Member State

Where the defendant is domiciled in a Member State, they may rely on the application of the uniform rules of international jurisdiction laid down in the Brussels Ia Regulation. Pursuant to Article 5, national rules of jurisdiction are excluded in such cases.

Within the Regulation, general jurisdiction is established in Article 4. Before applying this provision, however, it is necessary to consider special heads of jurisdiction, which operate as exceptions and thus constitute *lex specialis*. Exclusive jurisdiction under Article 24 is not applicable to claims for damages arising from infringements of personality rights. Likewise, the protective provisions concerning weaker contractual parties (Articles 10–23) are, by their nature, irrelevant. A further deviation from the general rule is the choice of court

[12] Decision of the Supreme Court of the Slovak Republic 4 Cdo 40/2014 of 19. 2. 2015, p. 9.

under Article 25, which may also apply to non-contractual claims (despite the ambiguity of the Slovak language version).[13]

In the absence of prorogation of jurisdiction under Article 25—which will likely be the typical scenario—jurisdiction is determined under Article 4(1), allowing proceedings to be brought in the courts of the defendant's domicile. At the same time, the Regulation provides alternative fora under Articles 7–9, which the claimant may invoke alongside the general rule.

A key alternative is Article 7(2), governing non-contractual liability, which confers jurisdiction on the courts of the place where the harmful event occurred or may occur. As this concept is not defined in the Regulation, it has been shaped through the autonomous and uniform interpretation of the Court of Justice of the European Union. In Case 21/76 *Bier*, concerning environmental damage caused by pollution of the Rhine, the Court held that this criterion encompasses both the place of the causal event (*locus actus*) and the place where the damage materialised (*locus damni*).[14] These connecting factors are equal in rank, allowing the claimant to choose between them[15]—a solution commonly referred to as the principle of ubiquity.[16]

In practice, the two connecting factors often coincide in the same jurisdiction, diverging only in so-called distance torts. In *Shevill*, which concerned cross-border defamation in the press, the Court further clarified that, where damage occurs in several States, jurisdiction based on the place of damage is limited to the harm suffered within that State[17] (the mosaic principle[18]). Jurisdiction arises only in respect of direct damage,[19] whereas the place of the

---

13 Article 25(1) of the Brussels Ia Regulation in the Slovak language provides that "the parties to a contract" may choose the court having jurisdiction. In light of the other language versions, this provision is interpreted as referring to the parties to the jurisdiction agreement itself. CSACH, K., GREGOVÁ ŠIRICOVÁ, Ľ.: *Určovanie právomoci pri rozhodovaní o občianskych a obchodných veciach*. In: CSACH, K., GREGOVÁ ŠIRICOVÁ, Ľ., JÚDOVÁ, E.: *Úvod do štúdia medzinárodného práva súkromného*. Bratislava: Wolters Kluwer, 2018, p. 79.

14 CJEU: *Handelskwekerij G.J. Bier BV v Mines de Potasse d'Alsace SA*, Case 21/76, Judgment of 30 November 1976, paras. 2–4.

15 *Ibid*, paras. 15–19, 24–25.

16 MANKOWSKI, P.: *Article 7*. In: In MAGNUS, U., MANKOWSKI, P.: Brussels Ibis Regulation Commentary. Köln: Verlag Dr. Otto Schmidt KG., p. 259.

17 CJEU: *Fiona Shevill and Others v Presse Alliance SA*, Case C-68/93, Judgment of 7 March 1995, paras. 28–33.

18 MANKOWSKI, P.: Article 7, cited above, p. 261.

19 CJEU: *Verein für Konsumenteninformation v Volkswagen AG*, Case C-343/19, Judgment of 9 July 2020, para. 26.

causal event confers jurisdiction over the entirety of the damage, including harm occurring across multiple jurisdictions.

The internet significantly amplifies the occurrence of such cross-border torts. Due to the global accessibility of online content, harm may arise in any State where the content is available. The Court of Justice has adapted its interpretation of Article 7(2) accordingly. In *eDate*,[20] the Court held that, in cases of online defamation, full compensation may be sought either in the State of the causal event or in the State of the victim's centre of interests, typically their habitual residence. At the same time, the mosaic approach was reaffirmed. Instead of an action for liability in respect of all the damage caused, the claimant may also bring his action before the courts of each Member State in the territory of which content placed online is or has been accessible. Those courts have jurisdiction only in respect of the damage caused in the territory of the Member State of the court seized. This conclusion gives claimants significant choice (forum optio) but also produce fragmentation and inefficiency when damage is widespread.

More recently, *Bolagsupplysningen*[21] extended that reasoning to legal persons. Further, it clarified that that a person who alleges that his personality rights have been infringed by the publication of incorrect information concerning him on the internet and by the failure to remove comments relating to him cannot bring an action for rectification of that information and removal of those comments before the courts of each Member State in which the information published on the internet is or was accessible. Only courts having jurisdiction over the entire damage are competent to rule on requests for removal and rectification of allegedly disparaging content since those requests are not separable per Member State as the amount of damages. In the latest step of this judicature journey (case *Gtflix Tv*[22]), the Court of Justice added that it is possible to claim, before the courts of each Member State in which the comments are or were accessible, compensation for the damage suffered in the Member State of the court seized, even though those courts do not have jurisdiction to rule on the application for rectification and removal.

Even these adaptations remain premised on a known, EU domiciled defendant and require some concrete link between content and victim (e.g., name, image, or other identifiable reference) to justify jurisdiction

---

20 CJEU: *eDate Advertising GmbH and Others v X and MGN Limited*, Joined Cases C-509/09 and C-161/10, Judgment of 25 October 2011.

21 CJEU: *Bolagsupplysningen OÜ and Ingrid Ilsjan v Svensk Handel AB*, Case C-194/16, Judgment of 17 October 2017.

22 CJEU: *Gtflix Tv v DR*, Case C-251/20, Judgment of 21 December 2021.

(*Mittelbayerischer Verlag*).[23] The Court has held that, in cases of harm affecting an indeterminate group of persons, the centre of interests cannot serve as a connecting factor due to the absence of individualisation. This illustrates the structural limits of the current framework, particularly where harm affects multiple subjects.

The CJEU's case-law on the violation of personality rights has been criticized by scholars, some Advocates General and practitioners. Critics claim that the mere accessibility of the infringing content is not sufficient to establish jurisdiction based on the place where the damage occurred, and that the application of the mosaic principle brings about a multiplication of fora that does not serve the legitimate interest of any party and defies the objectives of predictability and sound administration of justice The Commission Regarding Article 7(2), debates focus on the difficulties in relation to the determination of the place of damage in cases of purely financial loss, as well as the continued relevance of the mosaic principle in cases involving the violation of privacy rights. A future review of the Regulation could therefore consider ways to simplify and modernize Article 7(1) and 7(2).[24] According to the EU Commission's Report, the distinction between online and offline content could be criticised as difficult to apply in practice, since most of the newspapers are available both in print (offline) and online. As a result, the courts are required to complete the challenging task of finding out at the centre of interest which damage has been caused online (the damage can be claimed as a whole) or offline (only damage sustained in this single Member State can be claimed).[25]

For the sake of completeness, it should be noted that the claim could also potentially fall within Article 7(3), concerning civil claims arising from criminal offences, provided that the criminal court has jurisdiction to hear civil matters under national law. Article 7(5) may also be relevant, allowing proceedings in matters relating to the operations of a branch, agency, or establishment before the courts of its location.

In addition, Article 8 addresses situations involving closely connected claims or multiple defendants, enabling their joint adjudication in the interest of procedural efficiency; jurisdiction also extends to counterclaims.

---

23 CJEU: *Mittelbayerischer Verlag KG v SM*, Case C-800/19, Judgment of 15 June 2021.

24 Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the Application of Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast), Brussels, 2.6.2025 COM(2025) 268 final, p. 10.

25 *Ibid*, p. 9.

Finally, jurisdiction may also be established through tacit prorogation under Article 26, where the defendant enters an appearance without contesting jurisdiction at the earliest stage of the proceedings.

4.2.2 Defendant Domiciled in a Third State

Where the answer is negative—i.e. where the defendant is not domiciled in any EU Member State—the scope of application of the Brussels I Regulation (recast) is limited. The international jurisdiction of the courts of a Member State may then be established under the Regulation only in the situations enumerated in Article 6. The first two options listed therein are not relevant to the defamation claim for damages under consideration, as they concern claims arising from specific (protected) types of contracts (Article 18/1 for consumer contracts and Article 21/2 for individual employment contracts). Nor can exclusive jurisdiction under Article 24 be relied upon, since none of the situations listed therein is applicable to claims for damages resulting from infringements of personality rights. The remaining possibility is to establish jurisdiction by agreement (Article 25 for express agreements and Article 26 for tacit prorogation[26]), which, although possible for damages claims, is unlikely in practice.

Where none of the criteria laid down in the Regulation can be relied upon, jurisdiction is determined in accordance with the law of each Member State. In the case of Slovak courts, the relevant criteria are set out in Act No 97/1963 Coll. on Private International Law and Procedure.[27] The rules on jurisdiction are contained in §§ 37–47a; the following analysis focuses on those provisions that are relevant.

4.3 Analysis of the provisions of the Slovak PIL Act relevant to jurisdiction in cases of damage caused by infringements of personality rights

The basic jurisdictional rule in § 37 is based on the criterion of domicile/seat (which is not applicable here, as the defendant is domiciled in a third State) and, in matters relating to property rights, on the presence of property in Slovakia. According to the case-law of the Supreme Court of the

[26] The German Federal Court of Justice has expressed a positive view on the application of Article 26 to defendants domiciled outside the EU (notwithstanding the fact that Article 6 does not refer to Article 26). See WIEPEN, M.: *German Federal Court of Justice: Article 26 Brussels Ia Regulation Applies to Non-EU Defendants* (October 12, 2023). Online:https://conflictoflaws.net/2023/german-federal-court-of-justice-article-26-brussels-ia-regulation-applies-to-non-eu-defendants/ (accessed 12 October 2025).

[27] Hereinafter also as „the Act".

Slovak Republic,[28] this does not refer to any property, but specifically to the property that is the subject matter of the dispute. This, again, is not relevant for by infringements of personality rights.

In addition to the general rule (§ 37), jurisdiction may alternatively be established under §§ 37a–c. Section 37a is not applicable, as it concerns contractual relationships. By contrast, § 37b is relevant for damage caused by infringements of privacy or personality rights and provides for three situations. Under point (a), Slovak courts have jurisdiction in matters relating to non-contractual liability where the event giving rise to the damage occurred or may occur in the territory of the Slovak Republic. Point (b) concerns civil claims arising from criminal proceedings: Slovak courts may hear such claims where the criminal proceedings are conducted by Slovak authorities. The third criterion, set out in point (c), relates to disputes concerning the business activities of an undertaking or its branch or establishment; Slovak courts have jurisdiction where such an establishment is located in Slovakia. All three situations could theoretically apply to defamation claims and serve as a basis for the jurisdiction of Slovak courts.

Particular attention should be paid to § 37b(a), which was introduced by an amendment[29] intended, according to the explanatory memorandum,[30] to incorporate the principles of Article 5(3)–(5) of the Brussels I Regulation. Specifically, Article 5(3) provided that, in matters relating to non-contractual liability, proceedings may be brought in the courts of the place where the harmful event occurred or may occur.[31] The same criterion was transposed into § 37b(a). As noted above, according to the case-law of the Court of Justice, this concept encompasses both the place of the causal event and the place where the damage occurred, these being equivalent connecting factors between which the claimant may choose. However, the explanatory memorandum adopts a narrower

---

28 Decision of the Supreme Court of the Slovak Republic 1Cdo/37/2017 of 22 August 2018, para. 27: "*The wording of the provision 'if he has property here' cannot be interpreted broadly to mean that it is sufficient for a person to have any property in the territory of the Slovak Republic (such as a bank account or similar), but rather it must be the specific property that forms the subject matter of the proceedings*."

29 Act No. 589/2003 Coll., amending and supplementing Act No 97/1963 Coll. on Private International Law and Procedure, as amended, and amending and supplementing certain other acts.

30 See Explanatory Memorandum to the draft Act No 589/2003 Coll., p. 10 (section "Re § 37"). Online: https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=187042.

31 The same criterion is now contained in Article 7(2) of the Brussels I Regulation (recast), referred to above.

interpretation, focusing solely on the occurrence of damage within the territory of Slovakia. It does not envisage jurisdiction based on the place of the causal event (where the wrongful conduct occurred in Slovakia): "*In the case of point (a), this concerns any damage arising from a non-contractual relationship. The decisive factor for establishing the jurisdiction of Slovak courts is that the damage occurred or may occur within the territory of the Slovak Republic (i.e. the mere risk of damage occurring in Slovakia is sufficient)*."[32] The explanatory memorandum is therefore internally inconsistent: while aiming to align domestic law with the jurisdictional criteria of the Brussels regime, it interprets those criteria differently.

It is preferable to adopt an interpretation under which the "event giving rise to the damage" includes both the place of the causal event (conduct in Slovakia) and the place where the damage occurred (damage in Slovakia).[33] In addition to ensuring consistency with the Brussels regime, such an interpretation can be justified on the ground that the place of the wrongful act establishes a sufficiently predictable connection between the dispute and the forum for the defendant.

Furthermore, § 37c may be relevant in cases involving defendants domiciled outside the EU. Under this provision, jurisdiction may also be established in respect of closely connected claims where joint proceedings are necessary to avoid the risk of irreconcilable judgments, as well as in the case of counterclaims that are factually linked to the original claim. The Slovak court may thus exercise jurisdiction over a party in respect of whom it would otherwise lack jurisdiction, provided that the claims are closely connected to those already pending before it and for which jurisdiction exists under other provisions of the Act.[34]

Section 37d (exclusive jurisdiction) is not applicable, as the case (privacy and defamation) does not fall within any of the exhaustively listed situations. Likewise, § 37e (choice of court) is of limited relevance, since the parties may rely on Article 25 of the Brussels Ia Regulation, which takes precedence and applies irrespective of the domicile of the parties.

---

32 See Explanatory Memorandum to the draft Act No 589/2003 Coll., p. 10 (section "Re § 37b"), cited above.

33 *Števček* also supports taking into account the case-law of the Court of Justice of the European Union for the interpretation of § 37b(a) of the Private International Law Act (without considering the relevant part of the explanatory memorandum) (ŠTEVČEK, M.: *§ 37b [Alternative jurisdiction]*. In: LYSINA, P., ŠTEFANKOVÁ, N., ĎURIŠ, M., ŠTEVČEK, M.: Zákon o medzinárodnom práve súkromnom a procesnom. Komentár. Bratislava: C. H. Beck, 2011).

34 Decision of the Supreme Court of the Slovak Republic 1Cdo/37/2017 of 22 August 2018, para. 28.

In exceptional circumstances, Slovak courts may establish jurisdiction even in the absence of any connecting factor under the Act. The conditions are set out in § 47a, introduced by amendment with effect from 1 July 2022:[35]

- proceedings cannot be effectively initiated or conducted abroad; and
- the case has a close connection with the Slovak Republic.

The Act explicitly provides that this includes situations where recognition of a foreign judgment has been refused in Slovakia and new proceedings cannot be brought abroad due to res judicata. According to the explanatory memorandum, § 47a reflects the principle of forum necessitatis. This priciple is recognised in other national legal systems and in certain EU instruments.[36] In practice, exceptional situations may arise where a foreign court that would otherwise have jurisdiction is unable to exercise it—for example due to diplomatic immunity or civil unrest. In such cases, it may be appropriate to allow Slovak courts to hear the dispute in order to safeguard the claimant's right to a fair trial. The exercise of this jurisdiction presupposes a sufficient connection with Slovakia (e.g. nationality of a party or economic interests located in Slovakia), and it remains within the court's discretion whether to assume jurisdiction.[37]

Finally, it is submitted, in line with *Haťapka*,[38] that § 47a is not intended as a general corrective mechanism for deficiencies in foreign proceedings: issues relating to procedural fairness (such as bias or delays) or the outcome of foreign proceedings must primarily be addressed through remedies available under the *lex fori* of the court seized.

4.4 Jurisdictional Criteria *De Lege Ferenda*

The report by *Symeonides* and *Jayme*, prepared for the *Institut de Droit International*, addresses legal challenges arising from cross-border

---

35 Act No 108/2022 Coll., amending and supplementing Act No 97/1963 Coll. on Private International Law and Procedure, as amended, and amending and supplementing certain other acts.

36 FRANZINA, P.: *Forum Necessitatis*. In: I. VIARENGO, I., VILLATA F. C. (Eds.): Planning the Future of Cross Border Families - A Path Through Coordination. Bloomsbury Publishing, 2020, pp. 325 330.

37 Explanatory Memorandum to the draft Act No 108/2022 Coll., p. 2 (Re point 3 – § 47a). Online: https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=505568.

38 HAŤAPKA: *Právomoc slovenských justičných orgánov v občianskych a obchodných veciach*. In: LYSINA, P., HAŤAPKA, M., BURDOVÁ, K. a kol.: Medzinárodné právo súkromné. 3. vydanie. Bratislava: C. H. Beck, 2023, pp. 238.

infringements of personality rights on the Internet. Its goal is to clarify issues of jurisdiction, applicable law, and recognition and enforcement of foreign judgments in a context marked by conflicting national approaches to privacy and freedom of expression. The report adopts a comparative and policy-oriented approach, analyzing differences among legal systems and the inadequacy of traditional territorial connecting factors in the online environment. It proposes a coherent framework that emphasizes simplicity, efficiency, and predictability, including unified jurisdictional rules and flexible choice-of-law solutions. Notably, it advocates a "holistic" model allowing victims to bring a single action for global harm, thereby avoiding fragmented litigation and enhancing legal certainty.[39]

Buonaiuti notes, that Resolution adopted by the Institute does not contemplate among the proposed options that of concurrent assertions of jurisdiction based on pure accessibility, with effect limited to damages caused within the jurisdiction of the court seized. The „mosaic" solution is in fact rejected by setting out a rule on parallel proceedings based on a strict lis pendens logic. The rules embodied in the Institute's resolution are commendable in their effort of maintaining equality of arms between the parties as concerns the allocation of jurisdiction, by subjecting the applicability of the rules establishing *fora* closer to the legal position of the plaintiff to exceptions. The defendant as the person claimed to be liable may object to such *fora* in those cases where it can demonstrate that it did not derive any pecuniary or other significant benefit from the accessibility of the information in question in that State, or where a reasonable person could not have foreseen that the materials in dispute could have been accessible in the State of the court seized or that damage would have been caused in that State. The author suggests that a sort of cross-fertilization might arise between the works of the Institute and the future developments of the Court's case law, in such terms as to reconsider the appropriateness of clinging to a solution, such as that based on the mosaic principle, which is not attuned to the reality of the online environment.[40]

The mosaic principle was thoroughly criticised by *Marola.* The author proposes clarifying and harmonizing the relationship between jurisdictional

---

39 SYMEONIDES, S. C., JAYME, E., *Internet and privacy infringements*, report for the Institut de Droit International, The Hague session, 2019. Online: https://www.idi-iil.org/app/uploads/2019/06/Commission-8-Internet-and-privacy-infringement-Symeonides-Travaux-La-Haye-2019.pdf.

40 BUONAIUTI, F. M.: *Jurisdiction Concerning Actions by a Legal Person for Disparaging Statements on the Internet: The Persistence of the Mosaic Approach*. In: European Papers, Vol. 7, 2022, No. 1, European Forum, Insight of 29 June 2022, pp. 357–360. DOI: 10.15166/2499-8249/565.

grounds under the Brussels Ibis Regulation and data protection rules under the GDPR, especially in cases involving online harm. He suggests prioritizing the "centre of interests" criterion as a primary connecting factor while limiting fragmented jurisdiction under the mosaic approach to enhance predictability and consistency across Member States.[41]

The article by *Sakolciová* highlights that current EU jurisdictional rules in online defamation cases—particularly the accessibility-based approach and the mosaic principle—create practical difficulties and legal uncertainty in social media disputes. The author points out these shortcomings and emphasizes the need to reconsider how jurisdiction is determined in the digital environment noting that, as GA *Michal Bobek* suggested,[42] the place where the gravity of the dispute lies should play the key role.[43]

The need to reconsider the traditional jurisdictional framework is even more apparent in the context of AI-assisted privacy attacks. As illustrated above, the Brussels Ia Regulation is generally not applicable in cases involving offshore operators. In such situations, EU-based claimants are left to rely on the national rules of individual Member States. Slovak law provides an example of the uncertain footing faced by claimants. Moreover, due to the scale of AI-enhanced attacks, mass victimisation may easily occur. Where numerous victims are dispersed across multiple States, the Brussels I Regulation (recast) does not provide a mechanism for efficient consolidation or collective redress.[44]

Conclusion

This article has shown that AI-assisted infringements of personality rights challenge the structural assumptions underlying contemporary private international law. The analysis of the Brussels I Regulation (recast) and Slovak law demonstrates that existing jurisdictional rules remain tied to territoriality,

---

41 MAROLA, G.: *Jurisdiction in cross-border disputes concerning personality rights*, Doctoral dissertation, University of Pavia, 2023. Online: https://iris.unipv.it/retrieve/cc49369d-9a73-44d3-a1db-bf5dd018a9dc/PhD%20thesis_Marola%20Giacomo_PdfA.pdf.

42 SAKOLCIOVÁ, S.: *Defamation on Social Media – Challenges of Private International Law.* In: Bratislava Law Review, 2021, Vol. 5, No. 1, p. 125. DOI: https://doi.org/10.46282/blr.2021.5.1.226.

43 CJEU: Opinion of Advocate General Bobek delivered on 13 July 2017, C-194/16, *Bolagsupplysningen OÜ and Ingrid Ilsjan v Svensk Handel AB*.

44 LEIN, E.: *Cross-border Collective Redress and Jurisdiction under Brussels I: A Mismatch*. In: FARGRIEVE, D., LEIN, E. (eds.): Extraterritoriality and Collective Redress. Oxford: Oxford University Press, 2012, pp. 129–142.

identifiable defendants, and locatable harm—elements that are increasingly difficult to establish in AI-driven, cross-border contexts. While the case-law of the Court of Justice has introduced important adaptations, these remain insufficient where harm is diffuse, anonymised, or simultaneously distributed across multiple jurisdictions.

As a result, there is a growing risk that private-law remedies for cross-border privacy violations remain formally available but practically ineffective. This tension is particularly significant in light of the positive obligations stemming from Article 8 ECHR, which require that individuals have access to effective remedies against interferences with their private life. The current framework, however, may fail to provide such access where jurisdiction cannot be meaningfully established or leads to fragmented proceedings.

The developments discussed in this article suggest the need for a more coherent approach to jurisdiction in cases of online and AI-assisted harm. Greater emphasis on mechanisms that reduce fragmentation, improve foreseeability, and ensure effective access to justice—particularly in cases involving non-EU defendants—appears necessary. Without such adjustments, private international law risks lagging behind the realities it is meant to regulate.

List of Literature and Sources

1. ABU-ZAID, R., HAMMAD, A.: *Applications of AI in Decentralized Computing Systems: Harnessing Artificial Intelligence for Enhanced Scalability, Efficiency, and Autonomous Decision-Making in Distributed Architectures*. In: Journal of Artificial Intelligence & Cloud Computing, Vol. 7, No. 6, 2024, pp. 161–187. Online: https://www.researchgate.net/publication/386219179_Applications_of_AI_in_Decentralized_Computing_Systems_Harnessing_Artificial_Intelligence_for_Enhanced_Scalability_Efficiency_and_Autonomous_Decision-Making_in_Distributed_Architectures.
2. Act No. 589/2003 Coll., amending and supplementing Act No 97/1963 Coll. on Private International Law and Procedure, as amended, and amending and supplementing certain other acts.
3. Act No 108/2022 Coll., amending and supplementing Act No 97/1963 Coll. on Private International Law and Procedure, as amended, and amending and supplementing certain other acts.
4. BUONAIUTI, F. M.: *Jurisdiction Concerning Actions by a Legal Person for Disparaging Statements on the Internet: The Persistence of the Mosaic Approach*. In: European Papers, Vol. 7, No .1, 2022, European Forum, Insight of 29 June 2022, pp. 345–360; DOI: 10.15166/2499-8249/565.
5. CJEU: *Handelskwekerij G.J. Bier BV v Mines de Potasse d'Alsace SA*, Case 21/76, Judgment of 30 November 1976.
6. CJEU: *Fiona Shevill and Others v Presse Alliance SA*, Case C-68/93, Judgment of 7 March 1995.
7. CJEU: *Verein für Konsumenteninformation v Volkswagen AG*, Case C-343/19, Judgment of 9 July 2020.
8. CJEU: *eDate Advertising GmbH and Others v X and MGN Limited*, Joined Cases C-509/09 and C-161/10, Judgment of 25 October 2011.
9. CJEU: *Bolagsupplysningen OÜ and Ingrid Ilsjan v Svensk Handel AB*, Case C-194/16, Judgment of 17 October 2017.
10. CJEU: *Gtflix Tv v DR*, Case C-251/20, Judgment of 21 December 2021.
11. CJEU: *Mittelbayerischer Verlag KG v SM*, Case C-800/19, Judgment of 15 June 2021.
12. CJEU: Opinion of Advocate General Bobek delivered on 13 July 2017, C-194/16, *Bolagsupplysningen OÜ and Ingrid Ilsjan v Svensk Handel AB*.
13. CSACH, K., GREGOVÁ ŠIRICOVÁ, Ľ.: *Určovanie právomoci pri rozhodovaní o občianskych a obchodných veciach*. In: CSACH, K., GREGOVÁ ŠIRICOVÁ, Ľ.,

JÚDOVÁ, E.: *Úvod do štúdia medzinárodného práva súkromného*. Bratislava: Wolters Kluwer, 2018.

14. Decision of the Supreme Court of the Slovak Republic 4 Cdo 40/2014 of 19. 2. 2015.
15. Decision of the Supreme Court of the Slovak Republic 1Cdo/37/2017 of 22 August 2018.
16. ECtHR: *X and Y v the Netherlands*, Application no. 8978/80, Judgment of 26 March 1985.
17. ECtHR: *López Ribalda and Others v Spain*, Application nos. 1874/13 and 8567/13, Judgment of 17 October 2019.
18. ECtHR: *Söderman v Sweden*, Application no. 5786/08, Judgment of 12 November 2013.
19. Explanatory Memorandum to the draft Act No 589/2003 Coll., p. 10 (section "Re § 37"). Online: https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=187042.
20. Explanatory Memorandum to the draft Act No 108/2022 Coll., p. 2 (Re point 3 – § 47a). Online: https://www.nrsr.sk/web/Dynamic/DocumentPreview.aspx?DocID=505568.
21. FRANZINA, P.: *Forum Necessitatis*. In: I. VIARENGO, I., VILLATA F. C. (Eds.): Planning the Future of Cross Border Families - A Path Through Coordination. Bloomsbury Publishing, 2020.
22. HAKEEMAT, J.: *Harnessing AI for data privacy: Examining risks, opportunities and strategic future directions*. In: International Journal of Science and Research Archive, 2024, Vol. 13, No. 02, pp. 2878–2892. DOI: 10.30574/ijsra.2024.13.2.2510.
23. HAŤAPKA: *Právomoc slovenských justičných orgánov v občianskych a obchodných veciach*. In: LYSINA, P., HAŤAPKA, M., BURDOVÁ, K. a kol.: Medzinárodné právo súkromné. 3. vydanie. Bratislava: C. H. Beck, 2023.
24. JABIR, R., LE, J., NGUYEN, C.: *Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors*. In: AI 2025, Vol. 6, 174. Online: https://doi.org/ 10.3390/ai6080174.
25. KOHL, U.: *Jurisdiction and the Internet – Regulatory Competence over Online Activity*. Cambridge University Press, 2007, ppb. 2010. Online: https://ssrn.com/abstract=4169834.
26. LAZIĆ, V., MANKOWSKI, P.: *The Brussels I-bis Regulation: Interpretation and Implementation*. Edward Elgar Publishing Limited, 2023.

27. LEIN, E.: *Cross-border Collective Redress and Jurisdiction under Brussels I: A Mismatch*. In: FARGRIEVE, D., LEIN, E. (eds.): Extraterritoriality and Collective Redress. Oxford: Oxford University Press, 2012.
28. MAGNUS, U.: *Introduction*. In: MAGNUS, U., MANKOWSKI, P.: *Brussels Ibis Regulation Commentary.* Köln: Verlag Dr. Otto Schmidt KG.
29. MANKOWSKI, P.: *Article 7*. In: In MAGNUS, U., MANKOWSKI, P.: Brussels Ibis Regulation Commentary. Köln: Verlag Dr. Otto Schmidt KG.
30. MAROLA, G.: *Jurisdiction in cross-border disputes concerning personality rights*, Doctoral dissertation, University of Pavia, 2023. Online: https://iris.unipv.it/retrieve/cc49369d-9a73-44d3-a1db-bf5dd018a9dc/PhD%20thesis_Marola%20Giacomo_PdfA.pdf.
31. Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast), OJ L 351, 20.12.2012, pp. 1–32 (hereinafter referred to as the "Brussels Ia Regulation").
32. Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the Application of Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast), Brussels, 2.6.2025 COM(2025) 268 final.
33. SADIKU, M. N., ASHAOLU, T. J., AJAYI-MAJEBI, A., & MUSA, S. M.: *Artificial Intelligence in Social Media*. In: International Journal of Scientific Advances, 2021, Vol. 2, No. 1, pp. 15–20. DOI: 10.51542/ijscia.v2i1.4.
34. SAKOLCIOVÁ, S.: *Defamation on Social Media – Challenges of Private International Law.* In: Bratislava Law Review, 2021, Vol. 5, No. 1, pp. 121–134. DOI: https://doi.org/10.46282/blr.2021.5.1.226.
35. SHAO, C., CIAMPAGLIA, G.L., VAROL, O. et al: *The spread of low-credibility content by social bots*. Nat Commun 9, 4787 (2018). Online: https://doi.org/10.1038/s41467-018-06930-7.
36. SYMEONIDES, S. C., AND JAYME, E., *Internet and privacy infringements*, report for the Institut de Droit International, The Hague session, 2019. Online: https://www.idi-iil.org/app/uploads/2019/06/Commission-8-Internet-and-privacy-infringement-Symeonides-Travaux-La-Haye-2019.pdf.
37. ŠTEVČEK, M.: *§ 37b [Alternative jurisdiction]*. In: LYSINA, P., ŠTEFANKOVÁ, N., ĎURIŠ, M., ŠTEVČEK, M.: Zákon o medzinárodnom práve súkromnom a procesnom. Komentár. Bratislava: C. H. Beck, 2011.

38. WIEPEN, M.: *German Federal Court of Justice: Article 26 Brussels Ia Regulation Applies to Non-EU Defendants* (October 12, 2023). Online: https://conflictoflaws.net/2023/german-federal-court-of-justice-article-26-brussels-ia-regulation-applies-to-non-eu-defendants/

Author details

Mgr. Ľubica GREGOVÁ ŠIRICOVÁ, PhD.
Pavol Jozef Šafárik University in Košice
Faculty of law
Institute of International Law and European law
Email: lubica.siricova@upjs.sk

5 AI in the Workplace and its Impact on Employees' Rights: Historical perspectives, applications of AI, and liability

Milica Ilić

https://doi.org/10.33542/HRF-0541-2-05

Abstract

This paper analyses legal regulations of AI and its impact on employees' rights and liabilities. One part is dedicated to the historical aspect of AI, taking into account that the idea of humanity, throughout the historical development, has been to make machines to solve complex tasks instead of human beings, which also applies in the work environment. In addition, definitions are an indispensable part of this topic, as well as the legal sources. However, the application of new technologies brings concerns about how the workplace adapts to these changes. Therefore, this paper will answer the question of what the risks and benefits of using AI tools that influens on employees' rights, their privacy and data collection are. Further, the issue of civil liability is crucial because, in the case of AI, it has to be compared with traditional principles based on fault, causation, and compensation. This is challenging because AI lacks its own will, and this raises the question of different theoretical approaches.

Introduction

Today, new technologies have assumed a leading role across all fields of society.[1] Artificial intelligence (AI) has become increasingly prevalent in companies, where it has changed the way work is done.[2] Workplaces are filled with a lot of examples of artificial intelligence systems whose purpose is to facilitate work.[3] However, sometimes it can lead to discrimination in some situations, such as, for example, in the recruitment process, which will be discussed below.[4] In addition, the employees' privacy, personal data, and

---

1 LIVSON, M., ESHTOKIN, S., VASYUKOV, V., YUDINA, E., BAYBARIN, A., PIVNEVA, S.: *Impact of digitalization on legal regulation: Formation of new legal practices*. In: Campo Jurídico, Revista de direito agroambiental e teoria do direito, Vol. 9, 2021, p. 2.

2 DRONADUL, S. D., BHASKAR, D. V.: *Impact of Artificial Intelligence (AI) on Legal Profession and Justice System*. In: International Journal of Law Management and Humanities, Vol. 6, No. 2, 2023, p. 1084.

3 LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 82, Online: https://journals.um.si/index.php/lexonomica/article/view/5247/3365.

4 FRITTS, M., CABRERA, F: *AI recruitment algorithms and the dehumanization problem.* In: Ethics and Information Technology, Vol. 23, 2021, p. 793.

occupational health and safety are affected by the application of AI.[5] The legal system of the EU followed these social changes and influences and adopted new acts that govern the use of artificial intelligence in employment relations.[6] Moreover, the use of AI also implies the liability for damage caused by such a system. This issue is still challenging because there is no unique legal solution for regulating it. Thus, that leaves room for achieving a permanent solution in the future.[7]

5.1 Historical Background of AI

The world is currently in the centre of the Fourth Industrial Revolution, as a result of the electronic and information technology of the Third Industrial Revolution, which is erasing the boundaries separating the physical, digital, and biological realms.[8] Looking at the historical evolution, the idea of artificial intelligence dates back to ancient times. One early example is Talos from Greek mythology, who was a giant, man-made guardian of Crete, created from bronze and animated by "ichor," the divine fluid believed to be the blood of gods and immortals.[9] Thus, this is the history of fantasies where, as well as Homer described a mechanical „tripod" as waiters who served dinner to the gods.[10] However, in modern popular culture, AI is often portrayed as an abstract, all-powerful computerized being.[11] Accordingly, AI has long been influenced by the

---

5 LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 94.

6 KOVAČEVIĆ, L. J.: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024, p. 16.

7 FARAJPOUR, R: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 5.

8 Schwab, 2016, cited in HOWELLS, G: *Protecting Consumer Protection Values in the Fourth Industrial Revolution.* In: Journal of Consumer Policy, Vol. 43, 2020, p. 145. Online: https://doi.org/10.1007/s10603-019-09430-3.

9 Mayor, 2018, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

10 BUCHANAN, B. G.: *A (Very) Brief History of Artificial Intelligence*. In: AI Magazine, Vol. 26, No. 4, 2005, p. 53. Online: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1848.

11 „Such as "HAL 9000" in Arthur C. Clarke's (1968) 2001: A Space Odyssey or "Deep Thought" in Douglas Adams' (1979) The Hitchhiker's Guide to the Galaxy. Cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.):

evolution of computer and information sciences. Namely, before the formal concept of AI emerged, some pre-AI scientists, among them Norbert Wiener, for instance, defined cybernetics in terms of control and stability within electrical systems.[12] Another influential figure was Claude Shannon, who developed information theory and explained how digital signals operate.[13] One of the best-known pioneers was Alan Turing[14], who established the theory of computation, based on which the computational process could be expressed digitally. In 1950, he also introduced the famous "Turing test",[15] a thought experiment proposing that a machine could be trained to talk with people so effectively that they would be unable to tell whether they were interacting with a machine or another person.[16] As a matter of fact, Turing had an idea to learn the machine how to learn from its own experience and easily solve problems.[17]

The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

12 Hass, 2024, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

13 Cordeschi, 2001; PEARLMAN & SAID, 2011, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

14 More than seventy-five years have passed since Alan Turing created the automatic computing engine (ACE) in 1945. A reduced version of the machine, called the pilot model ACE and then the world's fastest computer, executed its first program in 1950. COPELAND & PROUDFOOT, 2000, cited in GARCIA-MURILLO, M., MACINNES, I.: *The impact of AI on employment: a historical account of its evolution*, 30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society", Helsinki, Finland, 16th-19th June, 2019, International Telecommunications Society (ITS), Calgary, 2019, p. 7.

15 Gonçalves, 2024, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

16 LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

17 Copeland & Proudfoot, 2000, cited in GARCIA-MURILLO, M., MACINNES, I.: *The impact of AI on employment: a historical account of its evolution*, 30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and

In addition, the term "artificial intelligence" however, was not used until 1956, when American cognitive scientist John McCarthy coined the term while organising the first academic conference dedicated to the topic.[18] Moreover, over the years, extensive conceptual work was carried out on AI, and many futurists viewed the rapid technological progress with a certain degree of apprehension. Alvin Toffler, for example, argued that society was undergoing a massive structural transformation from an industrial society to what he termed a "super-industrial society." Nevertheless, AI did not start to be taken seriously as a tool for tackling deep, complex problems until the 1980s, when computing technology advanced significantly.[19] AI captured global attention on May 11, 1997, when Garry Kasparov (World Chess Champion) was defeated by the chess-playing computer, "Deep Blue".[20] The next century (the 21st century) brought sophisticated machine-learning methods that began to successfully address a wide range of challenges.[21] Particularly, personal computers (PCs) were introduced at home and at work in the 1980s, which affected office workers[22] with all their advantages and disadvantages, which will be discussed in the continuation of the text.

Automated Society", Helsinki, Finland, 16th-19th June, 2019, International Telecommunications Society (ITS), Calgary, 2019, pp. 7–8.

18 Alon, Jaffe, Prange, & Vianelli, 2020, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

19 Newquist, 1994, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

20 Newborn, 2003, cited in LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, p. 3.

21 LARSSON, A., HATZIGEORGIOU, A*.: The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work, Routledge, Taylor & Francis, London and New York, 2015, pp. 3–4.

22 GARCIA-MURILLO, M., MACINNES, I.: *The impact of AI on employment: a historical account of its evolution*, 30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society", Helsinki, Finland, 16th-19th June 2019, International Telecommunications Society (ITS), Calgary, 2019, pp. 19–20.

5.2 Legal Sources, Definitions, and Examples

5.2.1 Definitions

The very definition of artificial intelligence is provided in the EU AI Act.[23] It refers to a machine-based system that is intended to function with different degrees of autonomy, may demonstrate adaptability after application, and for various goals, deduces how to produce outputs like predictions, content, recommendations, or decisions that can affect real-world or virtual environments based on the input it receives.[24] It is followed by the Guidelines of the European Commission that interpret the definition of an artificial intelligence system.[25]

According to that interpretation the definition consists of seven core components: (1) a system based on machines; (2) designed to function with different degrees of autonomy; (3) capable of demonstrating adaptiveness after application; (4) operating in pursuit of various goals; (5) inferring, from the inputs it processes, the manner in which outputs are produced; (6) including predictions, content, recommendations, or decisions; and (7) having the capacity to affect physical or virtual environments.[26]

In theory, there are various definitions of artificial intelligence.[27] According to Tuon and others, it is generally understood as the capability of a computerised system to decide in place of human beings.[28] In view of Grosz and Rolland, this

---

23 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) (Text with EEA relevance), OJ L, 2024/1689, 12.7.2024, Art. 3, Sec. 1. „'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."

24 *Ibid.*

25 EUROPEAN COMMISSION: *Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*. Brussels, 29.7.2025 C(2025) 5053 final.

26 *Ibid,* II. 9.

27 Čerka et al., 2017, cited in JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective.* In: Computer Law & Security Review, Vol. 49, 2023, 105825, p. 2. Online: https://www.sciencedirect.com/journal/computer-law-and-security-review.

28 Tuon et al., 2022, cited in JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an*

capability is enabled through machine learning, which leverages the processing of data.[29] Machine learning encompasses different methods, including learning through observation and interaction, according to Buiten.[30] More broadly, an AI system can be described as software, or a component of a device, that exhibits intelligent behaviour via the collection, processing, and analysis of data, and makes a certain degree of autonomous decisions in order to achieve a specific objective.[31] In addition, AI is "a branch of computer science - the study of the relation between computation and cognition"[32] or a system that is "imitating intelligent human behavior"[33] used to speed up business processes.[34] Experience, intuition, and knowledge are essential to human decision-making[35] whereas algorithmic decision-making relies on statistical modelling, allowing for greater speed, objectivity, and accuracy in generating outcomes.[36] Moreover, artificial intelligence demonstrates the capacity to match and frequently surpass human cognitive abilities in specific task domains. Consequently, AI-driven

occupational health and safety perspective. In: Computer Law & Security Review, Vol. 49, 2023, 105825, p. 2.

29 Grosz and Rolland, 1990, cited in JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective.* In: Computer Law & Security Review, Vol. 49, 2023, 105825, p. 2.

30 Buiten, 2019, cited in JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective.* In: Computer Law & Security Review, Vol. 49, 2023, 105825, p. 2.

31 JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective.* In: Computer Law & Security Review, Vol. 49, 2023, 105825, p. 2.

32 Barr & Feigenbaum, 2014, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1. Online: https://www.sciencedirect.com/journal/human-resource-management-review.

33 Kok, Boers, Kosters, Van der Putten, & Poel, 2009, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1.

34 VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1.

35 Klein, 2017, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1.

36 Chen, Chiang, & Storey, 2012, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1.

technologies, including virtual personal assistants and chatbots, are increasingly being integrated into workplace settings.[37]

For example, one of the most common chatbots is GPT, a language model developed by the OpenAI company, based in San Francisco, that can respond to various end-user queries in natural language.[38]

In addition, in practice, various examples of AI systems can be found in the work environment, such as FAMA, which relates to the recruitment process. It assesses whether an employee will adapt to the work environment and behave in accordance with the company's values.[39] PREACTOR is about guiding employees and scheduling and planning production tasks, as well as indicating which products must be manufactured and in what quantity.[40] COGITO pertains to monitoring employees.[41] Disciplinary decisions are made using Amazon's video surveillance system installed in delivery vehicles.[42]

5.2.2 Legal Sources

With regard to legal sources, in addition to the above-mentioned AI Act, several European Union acts govern the use of artificial intelligence in employment relations.[43]

The Artificial Intelligence Act, adopted by the European Parliament on 13 March 2024, is the key legislation in this area, designed to guarantee the security of AI systems used across the EU and to uphold human rights and fundamental

---

37 VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 1.

38 ŽIVKOVIĆ, A: *Computer programs legal protection framework with special reference to artificial intelligence ChatGPT*. In: Strani pravni život, Vol. LXVIII, No. 3, 2024, p. 331. Online: https://doi.org/10.56461/SPZ_24301KJ.

39 Adams-Prassl, 2024: 190–191, cited in LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 82.

40 Briône, 2017: 17, cited in LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 82.

41 Wood, 2021: 6, cited in LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 82.

42 Wood, 2021: 9, cited in LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 82.

43 KOVAČEVIĆ, LJ: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024., p. 16.

freedoms.[44] It has a wide-ranging scope of application, and it shall be fully applicable from 2 August 2026.[45]

It does not focus on regulating technology per se, but rather on artificial intelligence-based systems, with the regulatory approach grounded in evaluating the risks arising from their deployment. This framework also covers AI systems used in hiring processes and in managing employees' work, which are designated as high-risk and therefore subject to enhanced protective measures.[46]

Hajdú rightly points out that it does not sufficiently take into account the particularities of using AI in the field of employment, including employees in platform-based work. Given that employees are in a dependent relationship with employers, there is a risk that their rights could be neglected in the EU's strive to remain competitive in the AI sector. In this regard, it is imperative to create a legal framework that is both protective and enforceable.[47]

In addition, the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law was adopted in 2024 and contains a lot of solutions of the EU AI Act.[48] The main purpose is that by implementing the provisions of this Convention, artificial intelligence systems are ensured to be fully compliant with human rights, democracy, and the rule of law.[49] Furthermore, the following acts are also relevant: Directive (EU) 2024/2831 of the European Parliament and of the Council (Directive on Improving Working Conditions in Platform Work),[50] Directive (EU) 2019/1152

---

44 *Ibid*.

45 ĆERANIĆ, PERIŠIĆ, J, MIŠAN, S: *Akt o veštačkoj inteligenciji Evropske Unije i savremeni izazoovi u ostvarivanju i zaštiti ljudskih prava*. In: Univerzitet u Prištini, Pravni fakultet, Institut za kriminološka i sociološka istraživanja, Insitut za uporedno parvo, 2025, pp. 23–24. Online: http://ricl.iup.rs/2199/.

46 KOVAČEVIĆ, L. J.: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024.

47 József Hajdú, "The Influence of EU AI Regulation on Employment", Peace and Law – European Peace Agreements in the Broader Social Context. International Scientific Conference Book of Abstracts (eds. Branislav Ristivojević et al.), University of Novi Sad Faculty of Law, Novi Sad 2022, 62.

48 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Vilnius, 5. IX.2024, Council of Europe Treaty Series - No. 225.

49 *Ibid,* Article 1, Sec. 1.

50 Directive (EU) 2024/2831 of the European Parliament and of the Council of 23 October 2024 on improving working conditions in platform work, OJ L, 11.11.2024.

(Directive on Transparent and Predictable Working Conditions[51], Regulation (EU) 2016/679 (General Data Protection Regulation – GDPR).[52]

When it comes to algorithmic transparency, it is ensured by the provisions of Directive 2002/14/EC of the European Parliament and of the Council of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community.[53]

Art. 2, Sec. f, defined 'information' as „transmission by the employer to the employees' representatives of data in order to enable them to acquaint themselves with the subject matter and to examine it".[54] While according to Sec g, 'consultation' is defined as „the exchange of views and establishment of dialogue between the employees' representatives and the employer".[55]

5.3 AI Application in the Working Environment

The digital revolution has given rise to artificial intelligence, which is now at the centre of digital transformation and one of the main vectors of accelerated technological development. AI is a technique that enables machines to perform tasks that require human intelligence.[56] Its application influences labour law,

---

51 Directive (EU) 2019/1152 of the European Parliament and of the Council of 20 June 2019 on transparent and predictable working conditions in the European Union, OJ L 186/105, 11.7.2019.

52 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 119/1, 4.5.2016.

53 Directive 2002/14/EC of the European Parliament and of the Council of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community - Joint declaration of the European Parliament, the Council and the Commission on employee representation, OJ L 80, 23.3.2002, pp. 29–34. Cited in KOVAČEVIĆ, LJ: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024, p. 19.

54 Directive 2002/14/EC of the European Parliament and of the Council of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community - Joint declaration of the European Parliament, the Council and the Commission on employee representation, OJ L 80, 23.3.2002, Article 2, Sec. f.

55 *Ibid*, Article 2, Sec. g.

56 Marie-Cécile Amauger-Lattes, Droit social 2/2021, 1. Cited in KOVAČEVIĆ, LJ: Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects). In: Poslovni izazovi, Vol. II, 2024, p. 2.

especially on employees' rights and obligations and particularly in relation to algorithmic employee management.[57]

At the first spot, it should be noted that AI plays a new role in the recruitment process, and according to the opinion of HR managers, it leads to a "dehumanization" of that process.[58] Namely, AI is used to find qualified candidates and to reduce the time spent reviewing curriculum vitae. But, in addition, using such technologies could lead to, for example, discrimination (gender inequality).[59]

One example of discrimination is Amazon, which, after recognising the risk of discriminatory outcomes, discontinued its hiring algorithm when it became evident that the system was leading to reduced hiring of women.[60] In contrast, when algorithms are designed using ostensibly objective indicators such as termination for poor performance, the limited sample sizes available in most organizations often prevent the development of reliable predictive models.[61] Furthermore, both the quality and provenance of data pose significant challenges for organizations. Some firms have used indicators, including applicants' time spent on social media platforms such as Facebook, to infer work-related traits like commitment and engagement.[62] These measures not only lack empirical validation but are also intrusive and raise concerns about validity, as candidates may modify their behavior when they are aware of being monitored.[63]

---

57 KOVAČEVIĆ, LJ: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024, p. 2.

58 FRITTS, M., CABRERA, F: *AI recruitment algorithms and the dehumanization problem*. In: Ethics and Information Technology, Vol. 23, 2021, p. 793.

59 Eubanks, 2018; Noble, 2018, Dastin, 2018, cited in FRITTS, M., CABRERA, F: *AI recruitment algorithms and the dehumanization problem*. In: Ethics and Information Technology, Vol. 23, 2021, p. 793.

60 Meyer, 2018, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 5. Online: https://www.sciencedirect.com/journal/human-resource-management-review.

61 Tambe, Cappelli, & Yakubovich, 2019, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 5.

62 Cowgill et al., 2020; Tucker, 2017, cited in VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 5.

63 VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023, p. 5.

AI application in today's working environment means that employees participate less in the working processes because the automation of work tasks is inevitable.[64] According to the opinion of the majority of scholars, artificial intelligence poses a threat to the survival of employees' jobs.[65] It is also predicted that AI applications will replace full-time jobs with short-term ones.[66]

Consequently, uncertainty regarding employment stability seems to be an integrated aspect of workplace AI.[67] This risk is particularly real for tasks involving data handling, repeated motion, and frequent manual control of equipment.[68] However, there are certain drawbacks to this argument. Workplace AI applications have an indirect impact on employee productivity through the creation of new routines.[69]

For example, exploratory and exploitative routines are being changed by humanoid robots. Exploratory routines means developments of new ideas by human beings, while humanoids perform repetitive actions.[70]

Namely, labour productivity is enhanced by humanoids[71] in addition to innovation, which influences productivity too.[72]

---

64 Holford, 2019, cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

65 Arslan et al., 2021; Rampersad, 2020, cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

66 Braganza et al., 2020, cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

67 Costello and Donnellan, 2007, cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

68 Chuang, 2020, cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

69 Giudice et al., 2021. Cited in ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, p. 1.

70 DEL GIUDICEA, M., SCUOTTOD, V., BALLESTRAF, L., V., PIRONTIG, M: *Humanoid robot adoption and labour productivity: a perspective on ambidextrous product innovation routines*. In: The International Journal of Human Resource Management, Vol. 33, Issue 6, 2022, p. 1.

71 Stated by Acemoglu and Restrepo (2017a, 2017b), cited DEL GIUDICEA, M., SCUOTTOD, V., BALLESTRAF, L., V., PIRONTIG, M: *Humanoid robot adoption and labour productivity: a perspective on ambidextrous product innovation routines*. In: The International Journal of Human Resource Management, Vol. 33, Issue 6, 2022, p. 6.

72 DEL GIUDICEA, M., SCUOTTOD, V., BALLESTRAF, L., V., PIRONTIG, M: *Humanoid robot adoption and labour productivity: a perspective on ambidextrous product innovation*

According to various authors, in their theoretical interpretation, it can be found that, a substantial share of professional activities consists of routine tasks as types of jobs that are affected by AI.[73] Even when such work involves advanced cognitive elements such as memorization, complex judgment, or evaluation, high-skilled occupations are still susceptible to automation by AI. In the legal field, for example, tools such as PerfTech.AI's "Artificial Law Clerk" claim to significantly improve accuracy and productivity while reducing overall costs. In addition, one of the most significant features of AI-driven automation is that machines can act independently and learn from data, rather than simply following fixed instructions.[74] Opposite, there are also a lot of entry-level and manual jobs that are not routine, thus they can't be coded.[75] For instance, occupations such as security staff, cleaners, gardeners, receptionists, and chefs are hard to be replaced.[76] But in the end, automated processes can help people do their jobs better instead of just copying them. Machines can handle the boring parts of the job.[77]

5.4 AI-related Legal Risks

Regarding risks associated with AI, those are as mentioned above: discrimination, as well as employees' privacy, personal data, occupational health and safety, and liability for damage.[78]

When it comes to gender-based discrimination, women are often the first to face it, especially in certain occupations. The stereotypes associated with these jobs suggest that only men can perform them, which deprives women of the

---

routines. In: The International Journal of Human Resource Management, Vol. 33, Issue 6, 2022, p. 6.

73 Ford 2015, 2021; Susskind and Susskind 2015, 2020, cited in DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, pp. 678. Online: https://link.springer.com/article/10.1007/s00146-022-01496-xp.

74 DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, p. 678.

75 Goos et al. 2014; Autor et al. 2015, Barbieri et al. 2020, cited in DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, p. 679.

76 Gries and Naudé 2018, 4, cited in DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, p. 679.

77 Autor 2014; Brooks et al. 2020; Alarie et al. 2018; Ekbia and Nardi 2017, cited in DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, p. 679.

78 LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, p. 94.

opportunity to develop professionally and advance in their careers.[79] The goal of artificial intelligence is to replicate human beings, especially in routine and administrative jobs that are predominantly performed by women, because of the stereotype that it is only suitable for women to work in such jobs. Therefore, they are the most affected because the risk of losing these jobs due to automation may lead to this form of discrimination.[80] For examle in many developing countries such as Indonesia, Bangladesh, and India, women are hired in routine, low-wage jobs like clerical work, textiles, and call centres due to persistent gender stereotypes. With the rise of AI and automation, these roles, already limited in career growth and training opportunities, are increasingly at risk of disappearing.[81]

A good example is, as mentioned above, Amazon. It used indirect indicators from the text of CVs, such as "captain of the women's chess club," to indirectly infer an applicant's gender, which led to discrimination.[82] Namely, in the recruitment process, Amazon used AI tools that it trained itself using past candidate selection data for the purpose of identifying and ranking candidates who are qualified. However, due to the pattern-based system, biography of female candidates began to be downgraded. Although efforts were made to prevent gender-based discrimination after the detection of this problem, the company had to abandon this tool due to public concerns.[83] The reason this solution emerged is twofold: first, the resumes analyzed during the recruitment process can subtly reveal information about a candidate's gender, ethnicity, and age; second, AI systems, learning from historical data that already contain

---

79 Nuraeni Pusat Penelitian dan Pengembangan et al., cited in WIDIASTIANI, N, S: *Transforming Labour and Employment Law: Protection for Female Workers Against Discrimination in Artificial Intelligence Era*. In: Law and Justice, Vol. 10, No. 1, 2025, p. 30. Online: https://doi.org/10.23917/laj.v10i1.8028.

80 Farida, 2021, cited in WIDIASTIANI, N, S: *Transforming Labour and Employment Law: Protection for Female Workers Against Discrimination in Artificial Intelligence Era*. In: Law and Justice, Vol. 10, No. 1, 2025, p. 30.

81 The Asia Foundation, 2022, cited in WIDIASTIANI, N, S: *Transforming Labour and Employment Law: Protection for Female Workers Against Discrimination in Artificial Intelligence Era*. In: Law and Justice, Vol. 10, No. 1, 2025, p. 31.

82 LARSSON, S., MERRICKS WHITE, J., INGRAM BOGUSZ, C: *The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making*. In: Social Inclusion, Vol. 12, 2024, p. 2. Online: https://www.cogitatiopress.com/socialinclusion/article/view/7471/3747.

83 cf. Myers West et al., 2019, cited in LARSSON, S., MERRICKS WHITE, J., INGRAM BOGUSZ, C: *The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making*. In: Social Inclusion, Vol. 12, 2024, p. 2.

biases[84] (for example, because women were historically less frequently employed), can reproduce these biases instead of focusing solely on the skills that employers actually seek.

On the other hand, there are also benefits of using AI tools in employee selection.[85] The advantages of AI application during the selection process are that it can practically eliminate situations in which multiple candidates are equally qualified in a way that it can analyse massive amounts of data, using a large number of selection criteria, and ultimately select a single "best" candidate without human intervention. In practice, it often happens that multiple candidates meet the same basic criteria in terms of education, skills, and experience, thus in order to select one the best candidate, it should be added next criterion into AI tools (e. g. excelent writing skills, personality assessment) based on which that candidate will be selected.[86]

When it comes to the employees' privacy, monitoring of the employees' working activity is one of the powers the employer can exercise.[87]

For example, in French law, employee surveillance and data collection are governed by principles of transparency (or loyalty), proportionality, and relevance. As a result, employees must be informed in advance about any monitoring systems before their use. Any limitation on employees' rights must pursue a legitimate objective, be proportionate to that objective, and be directly associated with assessing their professional skills.[88]

Employers in Germany can monitor the working activities and performance of their employees, but restrictions must be observed.[89] For example, the legality of monitoring an employee's work performance is regulated by the Federal Data

---

84 Ajunwa, 2023, cf. Gaucher et al., 2011, Leavy, 2018, cf. Larsson et al., 2023b, cited in LARSSON, S., MERRICKS WHITE, J., INGRAM BOGUSZ, C: *The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making*. In: Social Inclusion, Vol. 12, 2024. Online: https://www.cogitatiopress.com/socialinclusion/article/view/7471/3747, p. 2.

85 LEDBETTER, K. J.: *Artificial intelligence (AI) and worker selection criteria: A pathway to fair hiring and positive organizational outcomes*. In: School of Business, University of Maryland Global Campus, 2024, p. 11. Online: https://www.proquest.com/docview/3174316853?fromopenview=true&fromunauthdoc=true&pq-origsite=gscholar&sourcetype=Dissertations%20&%20Theses.

86 *Ibid*, p. 10.

87 ALOISI, A., GRAMANO, E.: Artificial intelligence is watching you at work: digital surveillance, employee monitoring, and regulatory issues in the EU context. In: The Comparative Labor Law & Policy Journal, 2019, Vol. 41, No. XXX, p. 115. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3399548.

88 *Ibid*, p. 117.

89 *Ibid*, p. 118.

Protection Act, to the extent that such monitoring involves the processing of personal data within the meaning of Article 4(1) of the GDPR.[90]

In Italy, remote monitoring tools are allowed, and the protection of employees' privacy is fully ensured in the workplace if the rules of the Privacy Code are respected.[91]

5.5 Civil Liability in Automated Decion-Making

Decision-making capability is a feature of machine-based AI, which enables it to make choices „without direct human intervention, which brings both advantages and challenges."[92]

Advantages include faster data processing and greater accuracy, but challenges include the possibility of errors, which can lead to difficulties in establishing legal liability.[93]

Thus, when it comes to AI civil liability in automated decision-making, it poses the question of who is liable. This question is too hard to answer because legal liability is traditionally based on fault, causation, and damage compensation.[94] In the case of AI, this is different because it does not have its own will and can not be considered as independent entity.[95] Therefore, various legal regimes have various solutions. For example, in the United States and Iran, one of these solutions is that the operator or developer of AI is liable. In the

---

90 Franzen, § 26 BDSG, margin 2, (2019); Gola,2017; Kort, 2017; Kort, 2018. Cited in ALOISI, A., GRAMANO, E.: *Artificial intelligence is watching you at work: digital surveillance, employee monitoring, and regulatory issues in the EU context*. In: The Comparative Labor Law & Policy Journal, 2019, Vol. 41, No. XXX, p. 120.

91 For details, see Bellavista, A., Sitzia, A., Tullini, P., Ingrao, A. cited in ALOISI, A., GRAMANO, E.: *Artificial intelligence is watching you at work: digital surveillance, employee monitoring, and regulatory issues in the EU context*. In: The Comparative Labor Law & Policy Journal, 2019, Vol. 41, No. XXX, pp. 125–126.

92 Jiang, 2023; Nguyen & Quan, 2023 cited in FARAJPOUR, R.: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 2.

93 Barron, 2022; Zak et al., 2020, cited in FARAJPOUR, R.: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 2.

94 Katouzian, 2011, cited in FARAJPOUR, R.: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 5.

95 Boden, 2016, cited in FARAJPOUR, R.: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 5.

European Union, the proposal is to adopt the concepts of strict liability and consider AI with limited legal personality.[96]

Conclusion

From ancient times, artificial intelligence was only a fantasy; however, today we are living that fantasy. The term "artificial intelligence" was first time used in 1956, when American scientist John McCarthy organised the conference on the subject matter. However, from the 1980s, computer technology began to develop rapidly, and with the introduction of home computers (PCs), they consequently entered the world of work. Today, we have a new legislation on this topic, and the Artificial Intelligence Act, adopted by the European Parliament on 13 March 2024, has primacy.

Regarding the working environment, firstly, AI plays a new role in the recruitment process in order to find the best candidate for vacancies. However, in practice, the difficulties of using these tools emerged because it lead to discrimination between the genders (Amazon). On the other hand, there are benefits because it can analyse massive amounts of data and save time in selecting process.

As far as the employees' privacy is concerned, due to their relationship of subordination, the employer can exercise monitoring of the employees' working activity as one of their powers. However, this right is bound by various legal acts (the examples of French, German and Italian legislature are taken). The advantages of using artificial intelligence in the workplace include increased productivity and reduced costs. On the other hand, the disadvantages include the fear of job loss, as artificial intelligence replaces automated tasks, resulting in reduced involvement of employees in these processes. In addition, the issue of civil liability in automated decision-making is very important because AI does not have its own will and can not be considered as independent entity. However, in the EU, for example, the proposal is to have AI with limited legal personality. Thus, questions concerning the civil liability of artificial intelligence are yet to be developed in the future.

Thus, the development of artificial intelligence should not be limited solely to compliance with existing legal standards. AI systems must also be designed in a way that respects all human rights. As technology advances faster than legislation, liability becomes an essential guiding framework for the responsible use and governance of AI.

---

[96] Sartor and Lagioia, 2020, cited in FARAJPOUR, R.: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025, p. 5.

List of Literature and Sources

1. ALOISI, A., GRAMANO, E.: *Artificial intelligence is watching you at work: digital surveillance, employee monitoring, and regulatory issues in the EU context*. In: The Comparative Labor Law & Policy Journal, 2019, Vol. 41, No. XXX. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3399548.
2. BUCHANAN, B. G.: *A (Very) Brief History of Artificial Intelligence*. In: AI Magazine, Vol. 26, No. 4, 2005. Online: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1848.
3. ĆERANIĆ, PERIŠIĆ, J, MIŠAN, S: *Akt o veštačkoj inteligenciji Evropske Unije i savremeni izazoovi u ostvarivanju i zaštiti ljudskih prava*. In: Univerzitet u Prištini, Pravni fakultet, Institut za kriminološka i sociološka istraživanja, Insitut za uporedno parvo, 2025. Online: http://ricl.iup.rs/2199/.
4. DEL GIUDICEA, M., SCUOTTOD, V., BALLESTRAF, L., V., PIRONTIG, M: *Humanoid robot adoption and labour productivity: a perspective on ambidextrous product innovation routines*. In: The International Journal of Human Resource Management, Vol. 33, Issue 6, 2022, pp. 1098–1124. Online: https://doi.org/10.1080/09585192.2021.1897643.
5. DERANTY, P. J., CORBIN, T: *Artificial intelligence and work: a critical review of recent research from the social sciences*, In: AI & Society, Vol. 39, 2024, pp. 675–691. Online: https://link.springer.com/article/10.1007/s00146-022-01496-x.
6. Directive 2002/14/EC of the European Parliament and of the Council of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community - Joint declaration of the European Parliament, the Council and the Commission on employee representation, OJ L 80, 23.3.2002.
7. Directive (EU) 2019/1152 of the European Parliament and of the Council of 20 June 2019 on transparent and predictable working conditions in the European Union, OJ L 186/105, 11.7.2019.
8. Directive (EU) 2024/2831 of the European Parliament and of the Council of 23 October 2024 on improving working conditions in platform work, OJ L, 11.11.2024.
9. DRONADUL, S. D., BHASKAR, D. V.: *Impact of Artificial Intelligence (AI) on Legal Profession and Justice System*. In: International Journal of Law Management and Humanities, Vol. 6, No. 2, 2023, pp. 1084–1124.

10. EUROPEAN COMMISSION: *Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*. Brussels, 29.7.2025 C(2025) 5053 final.
11. FARAJPOUR, R: *Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making*. In: AI and Tech in Behavioural and Social Sciences, Vol. 3 No. 3, 2025.
12. Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Strasbourg, 5 September 2024, Council of Europe Treaty Series – No. 225.
13. FRITTS, M., CABRERA, F: *AI recruitment algorithms and the dehumanization problem*. In: Ethics and Information Technology, Vol. 23, 2021, pp. 791–801. Online: https://doi.org/10.1007/s10676-021-09615-w.
14. GARCIA-MURILLO, M., MACINNES, I.: *The impact of AI on employment: a historical account of its evolution*, 30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society", Helsinki, Finland, 16th-19th June, 2019, International Telecommunications Society (ITS), Calgary, 2019.
15. HOWELLS, G: *Protecting Consumer Protection Values in the Fourth Industrial Revolution*. In: Journal of Consumer Policy, Vol. 43, 2020, pp. 145–175. Online: https://doi.org/10.1007/s10603-019-09430-3.
16. HAJDÚ, J: *The Influence of EU AI Regulation on Employment*. In: RISTIVOJEVIĆ, B. et al. (eds.): Peace and Law – European Peace Agreements in the Broader Social Context. International Scientific Conference Book of Abstracts. University of Novi Sad Faculty of Law, Novi Sad, 2022.
17. JAROTA, M: *Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective*. In: Computer Law & Security Review, Vol. 49, 2023, 105825, Online: https://www.sciencedirect.com/journal/computer-law-and-security-review.
18. KOVAČEVIĆ, LJ: *Коришћење система вештачке интелигенције у свету рада – важнији радноправни аспекти, (Use of Artificial Intelligence Systems in the World of Work - Important Labour Law Aspects)*. In: Poslovni izazovi, Vol. II, 2024.
19. LARSSON, A., HATZIGEORGIOU, A.: *The Way Disruption, Technology and Practice Will Change the Way We Work in the Future*. In LARSSON, A., HATZIGEORGIOU, A. (eds.): The Future of Labour: How AI, Technological Disruption and Practice Will Change the Way We Work. London and New

York: Routledge, Taylor & Francis, 2015. Online: https://doi.org/10.4324/9781003391333.

20. LARSSON, S., MERRICKS WHITE, J., INGRAM BOGUSZ, C: *The Artificial Recruiter: Risks of Discrimination in Employers' Use of AI and Automated Decision-Making*. In: Social Inclusion, Vol. 12, 2024. Online: https://www.cogitatiopress.com/socialinclusion/article/view/7471/3747.
21. LEDBETTER, K. J.: *Artificial intelligence (AI) and worker selection criteria: A pathway to fair hiring and positive organizational outcomes*. In: School of Business, University of Maryland Global Campus, 2024. Online: https://www.proquest.com/docview/3174316853?fromopenview=true&fromunauthdoc=true&pq-origsite=gscholar&sourcetype=Dissertations%20&%20Theses.
22. LEŠNIK, A: *Protection of workers in relation to the use of artificial intelligence in the workplace*. In: LEXONOMICA, Journal of Law and Economics, Vol. 17, No. 1, 2025, pp. 79–120, Online: https://journals.um.si/index.php/lexonomica/article/view/5247/3365.
23. LIVSON, M., ESHTOKIN, S., VASYUKOV, V., YUDINA, E., BAYBARIN, A., PIVNEVA, S.: *Impact of digitalization on legal regulation: Formation of new legal practices*. In: Campo Jurídico, Revista de direito agroambiental e teoria do direito, Vol. 9, 2021, pp. 1–12.
24. 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 119/1, 4.5.2016.
25. 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) (Text with EEA relevance), OJ L, 2024/1689, 12.7.2024.
26. VARMA, A., DAWKINS, C., CHAUDHURI, K: *Artificial intelligence and people management: A critical assessment through the ethical lens*. In: Human Resource Management Review, Vol. 33, Issue 1, 2023. Online: https://www.sciencedirect.com/journal/human-resource-management-review.

27. WIDIASTIANI, N, S: *Transforming Labour and Employment Law: Protection for Female Workers Against Discrimination in Artificial Intelligence Era*. In: Law and Justice, Vol. 10, No. 1, 2025, pp. 24–39. Online: https://doi.org/10.23917/laj.v10i1.8028.
28. ZIRAR, A., IMRAN ALI, S., ISLAM, N.: *Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda*. In: Technovation, Vol. 124, 102747, 2023, pp. 1–17.
29. ŽIVKOVIĆ, A: *Computer programs legal protection framework with special reference to artificial intelligence ChatGPT*. In: Strani pravni život, Vol. LXVIII, No. 3, 2024, pp. 317–338. Online: https://doi.org/10.56461/SPZ_24301KJ.

Author details

Milica B. ILIĆ
University of Miskolc
Ferenc Deák Doctoral School
Central European Academy, Budapest
Email: adv.milicailic@gmail.com

6 When Values End at the Border: EU Surveillance Exports and Human Rights Risks

Hanna Pishchyk

https://doi.org/10.33542/HRF-0541-2-06

Abstract

The paper examines the dissonance between the EU's foundational values, particularly its commitment to human rights, and its export practices of digital and AI-enabled surveillance technologies. Focusing on the EU Dual-Use Regulation (Regulation (EU) 2021/821), the paper uses Morocco as a case study to assess how EU-origin and EU-funded technologies have been deployed to facilitate human rights abuses in contexts with limited judicial oversight and restricted civic space. The research draws on EU legal documents, human rights reports, NGO investigations, and thematic studies, such as the 2023 Institute of Development Studies report Mapping the Supply of Surveillance Technologies to Africa. The study scrutinises the regulatory framework, identifies discrepancies in its application, and evaluates the broader human rights implications. The findings reveal a significant gap between EU human rights rhetoric and action, emphasising the need for robust policy reforms to align external actions with foundational values. The paper concludes by outlining recommendations to strengthen EU surveillance governance and ensure that technology exports uphold human rights.

Introduction

In the latest Eurobarometer survey, respondents identified respect for fundamental rights as one of the two values the EU embodies best, alongside freedom of speech and expression.[1] This value framework, embedded in the Charter of Fundamental Rights of the European Union, guides a wide range of EU activities internally and globally, from diplomatic and cooperation efforts to trade agreements.[2] It also plays a crucial role in shaping EU laws and regulations,

---

1 EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR COMMUNICATION, and IPSOS: *EU Challenges and Priorities: Eurobarometer Summary*. Luxembourg: Publications Office of the European Union, 2024. Online: https://doi.org/10.2775/691136.

2 EUROPEAN COMMISSION: *Human Rights*. Trade and Economic Security. Online: https://policy.trade.ec.europa.eu/development-and-sustainability/sustainable-development/human-rights_en

often setting a global standard for advancing human rights protections.[3] However, a closer examination of the EU's regulatory ambitions often reveals challenges in harmonising them with its foundational values.[4] One manifestation of this dissonance can be found in the area of surveillance technologies, which are increasingly delivered as Surveillance-as-a-Service with ongoing maintenance, training, and AI-assisted features.[5] As such, AI-enhanced surveillance tools vastly expand states' capacity to monitor, classify, and predict individual behaviour. They are also being increasingly deployed against key guardians of democracy — journalists, human rights advocates, and protestors.[6]

The EU's recognition of the potential dangers of surveillance technologies has led to discussions and legislative efforts to mitigate these risks.[7] Yet, EU countries remain among the world's prominent exporters of surveillance technologies.[8] While the 2021 recast of the Dual-Use Regulation (Regulation 2021/821) introduced explicit controls on cyber-surveillance items, questions remain about whether this framework adequately prevents EU-origin technologies from facilitating repression abroad.

This paper examines this tension through the case of Morocco, a country with documented instances where EU-exported technologies have been used to suppress dissent and violate freedoms. Drawing on the Theory of Institutional Technology Actors and Norms (TITAN), the analysis situates EU export controls within the broader institutionalisation of Union values and collective responsibilities. The paper proceeds from the thesis that despite recent reforms,

---

3 DO, E.: *The Evolution of Data Privacy Legislation: A Closer Look at GDPR*. CyberInfoBlog, 27 February 2024. Online: https://www.cyberinfoblog.com/blog/the-evolution-of-data-privacy-legislation-a-closer-look-at-gdpr.

4 TSIMITAKIS, M.: *The EU's Migration Policies and the End of Human Rights in Europe*, Al Jazeera, 12 February 2024. Online: https://www.aljazeera.com/opinions/2024/2/12/the-eus-migration-policies-and-the-end-of-human-rights-in-europe.

5 AU, Y.: *Surveillance-as-a-Service (SaaS): A Survey of the Market for Facial Recognition and Live-Streaming Surveillance Tools,* Oxford Internet Institute, November 2021, p. 5. Online: https://www.oii.ox.ac.uk/wp-content/uploads/2021/12/SaaS_Yung_Au_Nov_2021-1.pdf.

6 GUARIGLIA, M.: *High Tech Police Surveillance of Protests and Activism: Year in Review 2020*, Electronic Frontier Foundation, 25 December 2020. https://www.eff.org/deeplinks/2020/12/high-tech-police-surveillance-protests-and-activism-year-review-2020.

7 CERULUS, L.: *Europe to Crack Down on Surveillance Software Exports*, POLITICO, 15 October 2020. Online: https://www.politico.eu/article/europe-to-curtail-spyware-exports-to-authoritarian-countries/

8 AU, Y.: *Surveillance-as-a-Service (SaaS): A Survey of the Market for Facial Recognition and Live-Streaming Surveillance Tools,* Oxford Internet Institute, November 2021, p. 5.

structural gaps in Regulation 2021/821 permit the export of AI-enabled surveillance capabilities into high-risk environments where misuse is highly foreseeable. A secondary thesis is that the Moroccan case illustrates how gaps in EU export controls allow transferred surveillance tools to contribute to human-rights violations.

Despite the rapidly expanding and opaque market for AI-assisted surveillance, there is a limited body of scholarship connecting these developments to EU value-based governance. The paper contributes to debates on human rights and freedoms affected by algorithmic systems in two ways. Empirically, it traces how EU-based suppliers and EU-funded programmes have expanded Morocco's surveillance infrastructure. Conceptually, it utilises TITAN to highlight the incomplete embedding of human rights norms in export control practices related to surveillance technologies. To achieve these goals, the paper employs a qualitative case study approach combining document analysis, NGO reports, and EU legal materials. The goal is to assess the alignment between EU values, regulatory practice, and the real-world consequences of EU-origin surveillance exports.

6.1 Conceptual Framework and Methodology

6.1.1 Conceptual Frameworks

This research employs the Theory of Institutional Technology Actors and Norms (TITAN) to critically examine the EU's exports of surveillance technology and their implications for human rights. TITAN, as presented by Smith and Miller (2023), offers a comprehensive, normatively informed framework for understanding technology regulation through the interplay of institutional structures, key actors, and ethical norms.[9] TITAN integrates two core theoretical notions: normative institutional theory and collective responsibility. It argues that technology regulation should ensure that technology is embedded within institutions in ways that uphold collective goods, such as security, privacy, and human rights. The theory identifies four key categories of actors: producers of technology, users of technology, government regulators, and normative policy shapers, each with distinct roles and responsibilities in the regulatory process.[10]

This framework provides a robust analytical lens for assessing EU surveillance exports, where divergent institutional interests, market incentives,

---

9 SMITH, M., MILLER, S.: *Technology, Institutions and Regulation: Towards a Normative Theory.* In: *AI & Society*, Vol. 40, 2023, pp. 1007–1017. DOI: https://doi.org/10.1007/s00146-023-01803-0.

10 *Ibid*.

and human rights norms frequently intersect. In this paper, TITAN structures the analysis along four dimensions:

- institutional (the evolution and design of Regulation 2021/821),
- technological (the characteristics of surveillance systems exported from the EU to Morocco),
- actor-based (EU institutions, private exporters, NGOs, and recipient governments),
- normative (the embedding, dilution, or inconsistency of EU human-rights standards across policy domains).

This approach clarifies how gaps between EU values, regulatory mechanisms, and export practices emerge and persist.

6.1.2 Methodology

The research employs several qualitative analysis methods to examine the EU's institutional values and human rights commitments in conjunction with its policy framework for technology exports. Given the relatively recent emergence of this topic from a historical perspective and its under-researched status in academia, this study relies heavily on field-specific resources, including reports and insights from NGOs, policy and research centres, human rights organisations, and EU legal documents. The analysis's reliance on secondary sources such as NGO reports, leaked materials, and expert investigations is also due to the opacity of export-licensing decisions and the nature of surveillance operations, including procurement and deployment.

Morocco is selected as a case study for both analytical and empirical reasons. Specifically, it combines extensive documented evidence of EU-origin surveillance systems, the availability of forensic material, and domestic governance weaknesses that heighten the human rights risks associated with AI-enabled surveillance technologies.

The qualitative nature of the research and chosen conceptual framework allow for a nuanced understanding of the complex interplay between regulatory frameworks, technological advancements, and human rights protections.

6.2 The Evolution of EU Values and Their Impact on Policy and Regulation

Before discussing the regulatory mechanisms, it is essential to understand the foundational values that guide the European Union's internal and external policies. In this paper, "values" refer to the foundational principles established in

the treaties, while "commitments" describe the legal and policy mechanisms through which those principles are operationalised.

From a historical and legal perspective, the European Union's unique values have evolved alongside its political subjectivity over the past thirty years. The Treaty of Maastricht (1992) marked a milestone by legally enshrining these values within the EU's regulatory framework. Although the Maastricht Treaty did not provide a comprehensive definition of EU values, it declared the member states' commitment to liberty, democracy, respect for human rights, and fundamental freedoms.[11] The Treaty of Amsterdam in 1997 further reinforced these core principles. Article 6 of the Treaty states: "The Union is based on liberty, democracy, human rights respect and fundamental freedoms and the rule of law."[12] Furthermore, the Treaty's provisions on a Common Foreign and Security Policy (Article J.1) outline these values in external relations, highlighting the EU's commitment to promoting democracy, the rule of law, and human rights internationally.[13]

In 2000, the adoption of the EU Charter of Fundamental Rights became a milestone in solidifying these values into a comprehensive list of rights and freedoms. In 2009, the Charter was given a binding status by the Lisbon Treaty, while its institutional influence has come to exemplify the primacy of respect for human rights over purely political or economic interests. For example, the Global Strategy for the European Union's Foreign and Security Policy states, "Our interests and values go hand in hand. We have an interest in promoting our values in the world. At the same time, our fundamental values are embedded in our interests."[14]

These values and commitments not only influence internal policy decisions within the Union but also shape its external regulatory approach. For instance, in response to the challenges brought by rapid digital development to fundamental rights, the EU acknowledged the need to regulate cyber-surveillance technologies.[15] The 2021 update to Dual-use export controls

---

11 Treaty on European Union, OJ C 191, 29 July 1992.

12 Treaty of Amsterdam Amending the Treaty on European Union, the Treaties Establishing the European Communities and Certain Related Acts, OJ C 340, 10 November 1997.

13 *Ibid*, Article J.1.

14 EUROPEAN UNION EXTERNAL ACTION SERVICE: *Shared Vision, Common Action: A Stronger Europe*. Brussels: EEAS, 2019, p. 13. Online: https://www.eeas.europa.eu/sites/default/files/eugs_review_web_0.pdf.

15 EUROPEAN UNION: *Dual-Use Export Controls | SUMMARY OF: Regulation (EU) 2021/821 Setting up an EU Regime for the Control of Exports, Brokering, Technical Assistance, Transit and Transfer of Dual-Use Items*. EUR-Lex. Last updated 17.11.2023.

demonstrates this acknowledgment by integrating human rights considerations into the EU trade operations. Although the Regulation does not explicitly mention the EU treaties or the Charter, it incorporates a vital human rights aspect to export controls on cyber-surveillance technologies.[16]

Taken together, this historical trajectory establishes a clear expectation that EU regulatory action, whether domestic or external, should reflect its foundational human rights framework. Yet, as later sections demonstrate, the export of surveillance technologies presents a significant challenge for this coherence.

6.3 EU Export Controls on Surveillance Technologies

The EU export control is governed under the exclusive competence within the common commercial policy.[17] The primary legislative act covering the exports of cyber technologies is Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit, and transfer of dual-use items (recast), adopted in May 2021.[18] The regulation pertains to the export and transfer of so-called dual-use goods. Dual-use items, as defined under Regulation 2021/821, encompass goods with components, technologies, and software used for civilian purposes that can also be utilised in the development and production of weapons of mass destruction. A permit document (license) is necessary to export dual-use products from EU territory.[19]

Although the legislation entered into force in 2021, the EU export controls on dual-use items are not a recent development. The EU began regulating these items in 1994, leading to the Council Regulation (EC) No 428/2009, which was in force until the new regulation took effect in 2021.[20] Regulation 2021/821,

Online: https://eur-lex.europa.eu/EN/legal-content/summary/dual-use-export-controls.html.

16 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021.

17 EUROPEAN UNION: *Division of Competences within the European Union*, EUR-Lex, last modified 24 February 2022. Online: https://eur-lex.europa.eu/EN/legal-content/summary/division-of-competences-within-the-european-union.html.

18 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021.

19 *Ibid.*

20 Council Regulation (EC) No 428/2009 of 5 May 2009 Setting Up a Community Regime for the Control of Exports, Transfer, Brokering and Transit of Dual-Use Items, OJ L 134/1, 29.5.2009.

adopted by the European Parliament and the Council, came as a significant evolution from the 2009 framework, particularly in relation to human rights and cyber-surveillance technology. Some notable updates included:

- Expanded authorisation and national control measures

Under the 2009 regulation, the authorisation requirements for non-listed items primarily focused on weapons of mass destruction and destinations subject to arms embargoes.[21] Regulation 2021/821 broadened the scope to include measures to prevent terrorism and human rights violations. Article 10 allows Member States to prohibit or impose authorisation requirements on non-listed dual-use items and requires notification to the Commission and other Member States.[22]

- Regulating cyber-surveillance: new controls and definitions

For the first time, the Dual-Use Regulation explicitly defined cyber-surveillance technologies, addressing growing concerns about their potential misuse. It defines cyber-surveillance items as 'dual-use items specially designed to enable the covert surveillance of natural persons by monitoring, extracting, collecting, or analyzing data from information and telecommunication systems'.[23] Article 2 of the preamble emphasises the need for Member State authorities to consider whether these items could be deployed in acts of internal repression or serious breaches of human rights and international humanitarian law.

- Catch-all provisions for emerging risks

One of the most significant aspects of the updated regulation was its new catch-all provisions under Article 5. These provisions mandate that authorisation is required for exporting non-listed cyber-surveillance items intended for internal repression or to commit serious human rights violations. Furthermore, it requires that exporters notify the competent authority if their due diligence reveals that cyber-surveillance items are intended for concerning uses.[24]

In October 2024, the Commission issued non-binding guidelines on Article 5 to clarify the definition of cyber-surveillance items and the due diligence

---

21 *Ibid*.

22 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021, Article 10.

23 *Ibid*, Article 2, paragraph 20.

24 *Ibid*, Article 5.

expectations.[25] While these guidelines aim to harmonise interpretation, they do not address core structural gaps such as transparency limitations and Member State discretion.

- Introduction of Internal Compliance Programs (ICPs)

Another notable innovation was the mandatory implementation of EU-wide Internal Compliance Programs for exporters. These programs require companies to adopt policies and procedures that are effective, appropriate, and proportionate to the risk of non-compliance, including considerations related to human rights.[26] This approach contrasted with the more reactive guidelines of the 2009 regulation and reflected a significant shift towards embedding human rights considerations within the export control framework.[27]

- Enhancing oversight through coordination and transparency

Regulation 2021/821 also introduced coordination mechanisms among Member States and a transparent reporting system managed by the European Commission. The Commission must prepare and publish an annual report detailing actions by Member States, including applications received and decisions made concerning cyber-surveillance items.[28]

While the 2021 regulation introduced several crucial improvements, several prominent civil society organisations pointed out that the adopted measures do not adequately prevent human rights abuses facilitated by the export of surveillance technologies.[29] The Commission's and Parliament's proposed ambitions were significantly reduced during negotiations. For instance, the Parliament's 2018 amendments elaborated specific rights at risk (privacy, data protection, expression, assembly), but the final text omits a clear

---

25 Commission Recommendation (EU) 2024/2659 of 11 October 2024 on Guidelines on the Export of Cyber-Surveillance Items under Article 5 of Regulation (EU) 2021/821 of the European Parliament and of the Council. OJ L 2024/2659, 16 October 2024.

26 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021, Article 22.

27 CURTI, I., MICHELINI, M. C. , PICCIANO, I. : *European Union Adopts New Regulation No. 2021/821 on Dual Use*, Portolano Cavallo Studio Legale, 15 July 2021. Online: https://portolano.it/en/newsletter/portolano-cavallo-inform-compliance/european-union-adopts-new-regulation-no-2021821-on-dual-use.

28 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021, Article 26, paragraph 2.

29 ACCESS NOW: *Human Rights Organizations' Response to the Adoption of the New EU Dual Use Export Control Rules*, March 2021. Online: https://www.accessnow.org/wp-content/uploads/2021/03/Analysis-EU-Surveillance-Tech-Export-Rules.pdf.

human rights framework. The Commission only later issued non-binding 2024 guidelines that elaborate concepts such as 'internal repression' and 'serious violations of human rights' for export assessments.[30] However, these guidelines do not alter the underlying legal structure and lack binding human rights criteria. The Commission's 2016 proposal for a dedicated Category 10 covering cyber-surveillance items was also excluded, reflecting Member State reluctance to adopt an autonomous EU control list and a preference for industrial interests.

To illustrate the practical challenges of the regulation, the case of Morocco offers valuable perspectives. As a recipient of EU-exported surveillance technologies, the country exemplifies the critical gaps between EU regulatory intentions and the on-the-ground reality of technology-facilitated human rights abuses.

6.4 Surveillance Tech Exports to Morocco: Case Study

6.4.1 Use of Surveillance Technologies in Morocco

Assessing the EU export regulations of cyber-surveillance through the lens of Moroccan practices offers several unique insights. Firstly, the country has a history of controversies related to surveillance technology being involved in high-profile incidents.[31] It aligns with its long-standing use of surveillance against local journalists, activists, and human rights advocates, often utilising tools supplied by European companies.[32] This is reinforced by Morocco's documented history of suppressing civil and political rights and restrictions on civic and political freedoms.[33] Furthermore, Morocco is among the countries that have substantially benefited from EU-provided cyber-surveillance technologies

30 Commission Recommendation (EU) 2024/2659 of 11 October 2024 on Guidelines on the Export of Cyber-Surveillance Items under Article 5 of Regulation (EU) 2021/821 of the European Parliament and of the Council. OJ L 2024/2659, 16 October 2024, para. 1.3.

31 MENAS ASSOCIATES: *Morocco Targeted 6,000 Algerian Phones in Pegasus Scandal*, Menas Associates, 27 July 2021. Online: https://menas.co.uk/blog/morocco-targeted-6000-algerian-phones-in-pegasus-scandal; AFRICAN PROGRAM CENTER FOR INTERNATIONAL POLICY: *The Pegasus Project: How the Cyber-Surveillance Scandal Affected Africa*, Medium, 27 July 2021. Online: https://africaprogram.medium.com/the-pegasus-project-how-the-cyber-surveillance-scandal-affected-africa-9d9a98595520.

32 PRIVACY INTERNATIONAL: *Human Rights Committee 116th Session: The Right to Privacy in Morocco*, March 2016. Online: https://privacyinternational.org/sites/default/files/2017-12/HRC_morocco.pdf.

33 FREEDOM HOUSE: *Morocco: Freedom in the World 2023 Country Report*, 2023. Online: https://freedomhouse.org/country/morocco/freedom-world/2023.

and related training programs.[34] Additionally, investigations have shown that the EU provided Morocco with phone-hacking spyware, intended for migration management but potentially used for broader surveillance purposes.[35]

In 2023, the Institute of Development Studies published a report titled "*Mapping the Supply of Surveillance Technologies to Africa: Case Studies from Nigeria, Ghana, Morocco, Malawi, and Zambia,*" offering an examination of Morocco's surveillance capabilities over time.[36] The report provides an analysis of the country's acquisition of surveillance technologies dating back to the 2010s. A significant portion of these technologies has been acquired from companies based in EU countries such as France, Italy, Cyprus, Germany, and Finland. For example, the Moroccan government invested significantly in surveillance systems like the Eagle System developed by the French company Amesys Bull (later Nexa Technologies).[37] This system enables the government to perform extensive monitoring and interception of internet traffic, facilitating state censorship and broad surveillance of digital communications.[38] As documented in the Oxford SaaS report, Eagle's architecture includes both traditional deep-packet inspection and "smart modules" for semantic analysis, email parsing, and voice identification — features that blur the boundary between non-AI interception and AI-assisted analytics.[39]

Moreover, Morocco has used technologies provided by the Italian company Hacking Team, which supplied the Remote Control System (RCS) malware, enabling whole-device compromise and deployment against journalists, including the outlet Mamfakinch. Although RCS is not an AI-enabled system, but conventional intrusive spyware, its use allowed authorities to monitor communications. It contributed to Mamfakinch's eventual shutdown,

---

34 PRIVACY INTERNATIONAL: *Revealed: The EU Training Regime Teaching Neighbours How to Spy*, 10 November 2020. Online: https://privacyinternational.org/long-read/4289/revealed-eu-training-regime-teaching-neighbours-how-spy.

35 DISCLOSE.NGO: *How the EU Supplied Morocco with Phone-Hacking Spyware*, Disclose, 25 July 2022. Online: https://disclose.ngo/en/article/how-the-eu-supplied-morocco-with-phone-hacking-spyware.

36 ROBERTS, T. et al: *Mapping the Supply of Surveillance Technologies to Africa: Case Studies from Nigeria, Ghana, Morocco, Malawi, and Zambia*. Brighton: Institute of Development Studies, 2023. DOI: https://doi.org/10.19088/ids.2023.027.

37 *Ibid*, p. 24.

38 *Ibid*, p. 89.

39 AU, Y.: *Surveillance-as-a-Service (SaaS): A Survey of the Market for Facial Recognition and Live-Streaming Surveillance Tools,* Oxford Internet Institute, November 2021, p. 11.

demonstrating that the human-rights risks of EU-origin surveillance exports extend beyond AI-based technologies.[40]

It is noteworthy that some surveillance technology purchases from EU-based companies have not been merely business transactions but have been facilitated by European organisations and even funded by the EU. European companies such as MSAB, a Swedish company, provided Morocco with advanced phone hacking spyware.[41] The surveillance technologies were provided by Intertech Lebanon, a French-Lebanese company, and overseen by the Vienna-headquartered International Centre for Migration Policy Development (ICMPD). These operations were financed through the EU "Border Management Programme for the Maghreb Region."[42] While the main aim was to tackle migration and human trafficking, this deployment has raised significant concerns regarding human rights. For instance, MSAB's XRY software can unlock and extract data from smartphones. At the same time, its associated analytical platform (XAMN) incorporates AI-enabled features such as automated image classification, similarity matching, and object recognition.[43] These capabilities significantly expand the surveillance value of the extracted data, posing profound implications for privacy and freedom of expression.[44]

The Moroccan legal framework, designed to regulate surveillance technologies, often fails to protect citizens effectively. While the country's constitution guarantees freedom of expression and privacy, the reality shows frequent arrests of journalists and restrictions on media coverage of sensitive topics like corruption and the monarchy.[45] National security concerns often take precedence over privacy rights, and the judiciary lacks the necessary independence to prevent abuses.[46] Vague national laws and unclear language in local laws and international agreements complicate the enforcement of these

---

40 ROBERTS, T. et al: *Mapping the Supply of Surveillance Technologies to Africa: Case Studies from Nigeria, Ghana, Morocco, Malawi, and Zambia*. Brighton: Institute of Development Studies, 2023, p. 90.

41 DISCLOSE.NGO: *How the EU Supplied Morocco with Phone-Hacking Spyware*, Disclose, 25 July 2022.

42 *Ibid*.

43 MSAB: *Discover Digital Evidence Rapidly with XAMN: A Guide*. Stockholm: MSAB, October 2022, p. 8. Online: https://www.msab.com/wp-content/uploads/2022/10/MSAB_Guide_XAMN_Discover_Evidence.pdf.

44 DISCLOSE.NGO: *How the EU Supplied Morocco with Phone-Hacking Spyware*, Disclose, 25 July 2022.

45 ROBERTS, T. et al: *Mapping the Supply of Surveillance Technologies to Africa: Case Studies from Nigeria, Ghana, Morocco, Malawi, and Zambia*. Brighton: Institute of Development Studies, 2023, p. 86.

46 *Ibid*, p. 86.

rights. For instance, despite Morocco's adoption of the International Covenant on Civil and Political Rights (ICCPR), its application is conditional and prioritises national laws and Islamic identity.[47] This legal ambiguity enables surveillance practices that often suppress political dissent and violate human rights.

6.4.2 Structural Gaps in EU Export Controls Revealed by the Moroccan Case

The Moroccan case discussed above illustrates how EU-origin surveillance tools can reach environments with weak safeguards, pointing out the need to examine the regulatory framework itself. The updated EU Regulation 2021/821 is a crucial advancement towards addressing the complexities associated with modern surveillance technologies.[48] Nevertheless, when compared with surveillance practices in high-risk recipient countries, several disparities and areas of concern become apparent.

The new rules aim to control the export of non-listed cyber-surveillance items, especially if they could be used for repression or to violate human rights. However, the lack of clear definitions and criteria leaves a broad space for interpretation, potentially allowing EU member states to justify exports on security grounds or as diplomatic excuses. As a result, Member States can diverge widely in how they apply human rights criteria, allowing surveillance tools to end up in rights-restrictive regimes with limited scrutiny. This also reflects a mismatch between regulating items and the realities of today's Surveillance-as-a-Service exports, which incorporate hardware, software, analytics, and ongoing support. Meanwhile, the regulation grants considerable discretion to national authorities to determine the conditions under which licenses are granted or denied.[49] This discretion can result in varied interpretations of the rules, influenced by economic or political interests rather than consistent human rights standards. Member States with active surveillance technology industries might interpret the rules more loosely to benefit local businesses at the expense of human rights.

The updated regulation includes a reporting system overseen by the European Commission to monitor how Member States handle cyber surveillance

47 *Ibid*, p. 87.

48 Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021, Article 5.

49 *Ibid,* Article 10.

item exports.[50] However, if countries fail to report their actions entirely or if the Commission does not rigorously enforce compliance, it leads to a significant oversight gap. The Commission's own report notes that almost half of global licences cannot be broken down by item or destination, and the reporting methodology makes it impossible for external observers to trace specific exports or verify whether transfers appear in the data.[51]

Moreover, the recast regulation does not include an initially proposed separate category for cyber-surveillance technology.[52] The absence of a dedicated category means that certain cyber-surveillance items, especially those integrating AI-enabled features, might not receive the necessary scrutiny, leading to their misuse in recipient countries, as in the case of Morocco. Additionally, the regulation does not establish concrete accountability measures for Member States or exporters who fail to conduct thorough human rights assessments. Even with the Commission's 2024 guidelines under Article 5, structural deficiencies remain, which highlights the need for targeted reforms.

6.5 Critical Reforms for Effective Regulation

To ensure that Regulation 2021/821 fulfills its goal of protecting rights while effectively regulating technology exports, several key reforms are needed. The primary focus should be robust enforcement, improved Member-State coordination, and transparency within the existing framework. Strengthening and standardising human rights impact assessments for cyber-surveillance technologies is imperative. These assessments must be applied uniformly across Member States, cover both listed and non-listed technologies, and include clear methodological guidance that builds on, but is not limited to, the Commission's 2024 Article 5 guidelines. The non-sensitive parts of the assessments should be made publicly available, enabling civil society oversight and ensuring that all stakeholders can scrutinise the potential human rights implications of exported technologies.

---

50 *Ibid*, Article 23 and 26.

51 EUROPEAN COMMISSION: *Report on the Implementation of Regulation (EU) 2021/821 Setting Up a Union Regime for the Control of Exports, Brokering, Technical Assistance, Transit and Transfer of Dual-Use Items*. COM(2025) 19 final, 2025, Section 5. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52025DC0019.

52 CURTI, I., MICHELINI, M. C. , PICCIANO, I. : *European Union Adopts New Regulation No. 2021/821 on Dual Use*, Portolano Cavallo Studio Legale, 15 July 2021. Online: https://portolano.it/en/newsletter/portolano-cavallo-inform-compliance/european-union-adopts-new-regulation-no-2021821-on-dual-use.

Moreover, enhancing Internal Compliance Programs (ICPs) is crucial. These programs need clear guidelines, regular audits, and continuous improvement support to guarantee effective implementation. Enhancing coordination and information sharing among Member States is also critical to ensure consistent enforcement. This can be achieved through regular inter-state meetings, joint training sessions, and a centralised database for licensing and enforcement actions. Increasing scrutiny and accountability in the export licensing process can be another significant step. Implementing stricter approval criteria, regular reviews, and transparent documentation of licensing decisions will ensure that human rights considerations are prioritised over the sole economic or strategic interests of individual Member States. In particular, reducing the broad discretion currently afforded to national authorities is essential to avoid divergent standards and inconsistent outcomes.

Finally, given the slow pace of formal regulatory revision relative to technological advances, an additional priority should be to strengthen administrative mechanisms for identifying emerging risks. This includes more frequent updates to Commission guidance, structured monitoring of technological developments, and coordinated expert assessments of new surveillance capabilities.

By focusing on these reforms, the EU can enhance the enforcement of Regulation 2021/821, improve coordination among Member States, and maintain transparency within the established framework. Failure to address current regulatory loopholes risks perpetuating the abuses the regulation aims to prevent.

Conclusion

This paper examined the tension between the EU's human rights values and its export control practices by assessing how Regulation 2021/821 functions in the context of Morocco's documented misuse of EU-origin surveillance technologies. The findings confirm the thesis formulated at the beginning of the paper, demonstrating that the regulation contains substantive structural gaps that permit the export of AI-enabled surveillance capabilities to environments with a high risk of human rights abuses. The Moroccan case further confirmed the secondary thesis, showing how deficiencies in EU export controls enable the transfer of surveillance tools that contribute to documented restrictions on civic space and human rights.

The analysis showed that while the regulation marks an important step toward integrating human rights considerations into export controls, its practical impact is constrained by ambiguity in human rights criteria,

considerable discretion granted to national authorities, and regulatory mechanisms that struggle to keep pace with technological developments. Furthermore, the paper traced how EU-based suppliers and EU-funded programmes have supported Morocco's expanding surveillance ecosystem. Through the case study, it spotlighted the misalignment between the EU's normative commitments and its regulatory implementation. These findings fulfil the paper's stated goal of assessing the coherence between EU values, regulatory practice, and the real-world consequences of EU-origin surveillance exports. The findings point out that strengthening EU export governance is essential not only to prevent misuse abroad but also to ensure coherence with the Union's foundational values.

Building on this, the paper outlined several necessary reforms: strengthening and standardising human rights impact assessments, enhancing Internal Compliance Programs, improving Member State coordination, and bolstering mechanisms for tracking emerging technological risks. If Regulation 2021/821 is to function as an adequate safeguard rather than a symbolic instrument, it necessitates a more transparent, harmonised, and technologically responsive governance approach.

List of Literature and Sources

1. ACCESS NOW: *Human Rights Organizations' Response to the Adoption of the New EU Dual Use Export Control Rules*, March 2021. Online: https://www.accessnow.org/wp-content/uploads/2021/03/Analysis-EU-Surveillance-Tech-Export-Rules.pdf.
2. AFRICAN PROGRAM CENTER FOR INTERNATIONAL POLICY: *The Pegasus Project: How the Cyber-Surveillance Scandal Affected Africa*, Medium, 27 July 2021. Online: https://africaprogram.medium.com/the-pegasus-project-how-the-cyber-surveillance-scandal-affected-africa-9d9a98595520.
3. CERULUS, L.: *Europe to Crack Down on Surveillance Software Exports*, POLITICO, 15 October 2020. Online: https://www.politico.eu/article/europe-to-curtail-spyware-exports-to-authoritarian-countries/.
4. Commission Recommendation (EU) 2024/2659 of 11 October 2024 on Guidelines on the Export of Cyber-Surveillance Items under Article 5 of Regulation (EU) 2021/821 of the European Parliament and of the Council. OJ L 2024/2659, 16 October 2024.
5. Council Regulation (EC) No 428/2009 of 5 May 2009 Setting Up a Community Regime for the Control of Exports, Transfer, Brokering and Transit of Dual-Use Items, OJ L 134/1, 29.5.2009.
6. CURTI, I., MICHELINI, M. C., PICCIANO, I.: *European Union Adopts New Regulation No. 2021/821 on Dual Use*, Portolano Cavallo Studio Legale, 15 July 2021. Online: https://portolano.it/en/newsletter/portolano-cavallo-inform-compliance/european-union-adopts-new-regulation-no-2021821-on-dual-use.
7. DISCLOSE.NGO: *How the EU Supplied Morocco with Phone-Hacking Spyware*, Disclose, 25 July 2022. Online: https://disclose.ngo/en/article/how-the-eu-supplied-morocco-with-phone-hacking-spyware.
8. DO, E.: *The Evolution of Data Privacy Legislation: A Closer Look at GDPR*, CyberInfoBlog, 27 February 2024. Online: https://www.cyberinfoblog.com/blog/the-evolution-of-data-privacy-legislation-a-closer-look-at-gdpr.
9. EUROPEAN COMMISSION: *Report on the Implementation of Regulation (EU) 2021/821 Setting Up a Union Regime for the Control of Exports, Brokering, Technical Assistance, Transit and Transfer of Dual-Use Items*. COM(2025) 19 final, 2025. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52025DC0019.
10. EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR COMMUNICATION, and IPSOS: *EU Challenges and Priorities: Eurobarometer*

Summary. Luxembourg: Publications Office of the European Union, 2024. Online: https://doi.org/10.2775/691136.

11. EUROPEAN COMMISSION: *Human Rights*. Trade and Economic Security. Online: https://policy.trade.ec.europa.eu/development-and-sustainability/sustainable-development/human-rights_en.
12. EUROPEAN UNION: *Division of Competences within the European Union*, EUR-Lex, last modified 24 February 2022. Online: https://eur-lex.europa.eu/EN/legal-content/summary/division-of-competences-within-the-european-union.html.
13. EUROPEAN UNION: *Dual-Use Export Controls | SUMMARY OF: Regulation (EU) 2021/821 Setting up an EU Regime for the Control of Exports, Brokering, Technical Assistance, Transit and Transfer of Dual-Use Items*. EUR-Lex. Last updated 17.11.2023. Online: https://eur-lex.europa.eu/EN/legal-content/summary/dual-use-export-controls.html.
14. EUROPEAN UNION EXTERNAL ACTION SERVICE: *Shared Vision, Common Action: A Stronger Europe*. Brussels: EEAS, 2019. Online: https://www.eeas.europa.eu/sites/default/files/eugs_review_web_0.pdf.
15. FREEDOM HOUSE: *Morocco: Freedom in the World 2023 Country Report*, 2023. Online: https://freedomhouse.org/country/morocco/freedom-world/2023.
16. GUARIGLIA, M.: *High Tech Police Surveillance of Protests and Activism: Year in Review 2020*, Electronic Frontier Foundation, 25 December 2020. https://www.eff.org/deeplinks/2020/12/high-tech-police-surveillance-protests-and-activism-year-review-2020.
17. MENAS ASSOCIATES: *Morocco Targeted 6,000 Algerian Phones in Pegasus Scandal*, Menas Associates, 27 July 2021. Online: https://menas.co.uk/blog/morocco-targeted-6000-algerian-phones-in-pegasus-scandal.
18. MSAB: *Discover Digital Evidence Rapidly with XAMN: A Guide*. Stockholm: MSAB, October 2022. Online: https://www.msab.com/wp-content/uploads/2022/10/MSAB_Guide_XAMN_Discover_Evidence.pdf.
19. PRIVACY INTERNATIONAL: *Human Rights Committee 116th Session: The Right to Privacy in Morocco*, March 2016. Online: https://privacyinternational.org/sites/default/files/2017-12/HRC_morocco.pdf.
20. PRIVACY INTERNATIONAL: *Revealed: The EU Training Regime Teaching Neighbours How to Spy*, 10 November 2020. Online: https://privacyinternational.org/long-read/4289/revealed-eu-training-regime-teaching-neighbours-how-spy.

21. Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (recast), OJ L 206/1, 11.6.2021.
22. ROBERTS, T. et al: *Mapping the Supply of Surveillance Technologies to Africa: Case Studies from Nigeria, Ghana, Morocco, Malawi, and Zambia*. Brighton: Institute of Development Studies, 2023, 167 p. DOI: https://doi.org/10.19088/ids.2023.027.
23. SMITH, M., MILLER, S.: *Technology, Institutions and Regulation: Towards a Normative Theory.* In: *AI & Society*, Vol. 40, 2023, pp. 1007–1017. DOI: https://doi.org/10.1007/s00146-023-01803-0.
24. Treaty of Amsterdam Amending the Treaty on European Union, the Treaties Establishing the European Communities and Certain Related Acts, OJ C 340, 10 November 1997.
25. Treaty on European Union, OJ C 191, 29 July 1992.
26. TSIMITAKIS, M.: *The EU's Migration Policies and the End of Human Rights in Europe*, Al Jazeera, 12 February 2024. Online: https://www.aljazeera.com/opinions/2024/2/12/the-eus-migration-policies-and-the-end-of-human-rights-in-europe.
27. AU, Y.: *Surveillance-as-a-Service (SaaS): A Survey of the Market for Facial Recognition and Live-Streaming Surveillance Tools*. Oxford Internet Institute, November 2021. Online: https://www.oii.ox.ac.uk/wp-content/uploads/2021/12/SaaS_Yung_Au_Nov_2021-1.pdf.

Author details

Mgr. Hanna PISHCHYK
Palacky University Olomouc
Faculty of Arts
Department of History
Email: research.hanna@proton.me

7 New Human Rights Bodies in the Age of Artificial Intelligence

Lucia Bakošová

https://doi.org/10.33542/HRF-0541-2-07

Abstract

The emergence of artificial intelligence (AI) is reshaping not only the substance of human rights protection but also the institutional architecture responsible for its enforcement. Traditional human rights bodies, such as the United Nations Human Rights Council, treaty bodies, and regional courts are confronted with challenges that transcend territorial jurisdiction, individual responsibility, and traditional forms of accountability. As AI systems increasingly mediate public administration, law enforcement, and private decision-making, existing institutions struggle to address algorithmic harms that are automated, opaque, and transnational in nature. Recently, several initiatives at the global and regional levels have sought to respond to these challenges by exploring new institutional approaches to the governance of AI and human rights. Proposals have emerged within the United Nations framework to establish or mandate a specialised body dedicated to monitoring human rights impacts of AI systems, either through a new standing mechanism or as part of a broader global governance framework. Parallel discussions in regional fora point to a gradual movement toward institutionalising human-rights-based oversight of AI technologies. The paper examines the potential and limitations of these initiatives, asking whether they signal the beginning of a new generation of human rights bodies or whether they merely extend existing mandates without altering the traditional institutional paradigm.[1]

Introduction

The rapid digitalisation of society and the dynamic technological progress, particularly in the field of artificial intelligence (AI), are fundamentally transforming the way human rights are exercised and protected.[2] While

[1] This contribution was prepared as part of the research project "Technological Progress and International Human Rights Law" (No. VVGS-2023-2947), implemented within the Internal Scientific Grant System of Pavol Jozef Šafárik University in Košice. The research was funded by the European Union under the EU NextGenerationEU initiative through the Recovery and Resilience Plan for Slovakia, Project No. 09I03-03-V05-00008.

[2] FRA – European Union Agency for Fundamental Rights: *Getting the Future Right – Artificial Intelligence and Fundamental Rights*, 2020, 108 p. Online: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-

traditional forms of human rights violations were typically linked to identifiable actors, specific territories, and clearly traceable causes, algorithmic decision-making introduces interferences that are automated, complex, often opaque, and transnational in nature.[3] AI systems now influence access to social benefits, healthcare, employment, public administration, justice, and the conduct of hostilities in armed conflicts, often in ways that are both hidden and simultaneously far-reaching.[4] The accelerating deployment of AI systems thus exposes a widening gap between the scale of algorithmic harms and the capacity of existing international human rights institutions to prevent and redress them. AI systems deployed by multinational platforms can generate rights-relevant impacts across jurisdictions, often without a clear nexus to the state traditionally responsible under international human rights law. Institutions such as the UN Human Rights Council, United Nations treaty-bodies, and regional courts were established at a time when state responsibility and territorial jurisdiction were assumed to be clear and predictable.

---

rights; UN Human Rights Council: *The Right to Privacy in the Digital Age*, A/HRC/48/31, 2021; Council of Europe: *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 2020. Online: https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016809c820f.

3 Council of Europe: *Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications*, Study DGI(2017)12, Strasbourg: Council of Europe, 2018, pp. 20–24. Online: https://rm.coe.int/algorithms-and-human-rights-en-rev/16808f06aa; European Parliamentary Research Service: *Understanding Algorithmic Decision-Making: Opportunities and Challenges for the EU*, PE 624.261, Brussels: European Parliament, 2019, pp. 7–24. Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf; MCGREGOR, L., MURRAY, D., NG, V.: *International Human Rights Law as a Framework for Algorithmic Accountability*. In: International and Comparative Law Quarterly, Vol. 68, No. 2, 2019, pp. 315–320.

4 AHMAD, S. et al.: *The Impact of Artificial Intelligence on Business & Social Values: Benefits, Challenges, And Future Directions*. In: *Educational Administration: Theory and Practice*, Vol. 30, No. 6, 2024, pp. 3174–3180; BAKOŠOVÁ, L.: *Ethical and legal aspects of the use of artificial intelligence in health and nursing care*. In: Studia Iuridica Cassoviensia, Vol. 8, No. 2, 2020, pp. 3–18; BENKO, R.: *The Impact of Artificial Intelligence on Fundamental Rights and Freedoms and Vice-Versa*. In: KLUČKA, J., BAKOŠOVÁ, L., SISÁK, Ľ.: Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges, Praha: Leges, 2021, pp. 39–42; ELBERT, Ľ.: *Medzinárodnoprávny úvod k používaniu autonómnych zbraňových systémov.* Praha: Leges, 2022, 109 p.; JAKAB, R.: *Artificial Intelligence and Administration Decision-Making*. In: Acta Universitatis Carolinae, Vol. 72, No. 1, 2026, pp. 31–43; SEMAN, T.: *Risks of Partial Exploitation of Artificial Intelligence Potential in Administrative Proceedings*. In: Studia Iuridica Cassoviensia, Vol. 13, Special Issue, 2025, pp. 155–169.

Against this backdrop, international organisations have begun to explore the need for institutional adaptation in response to AI-related challenges. While these developments can be observed at both global and regional levels, this contribution focuses primarily on the evolving institutional landscape within the United Nations (UN), where ongoing discussions have highlighted persistent fragmentation across existing mechanisms, as well as the absence of a system-wide approach capable of addressing the cross-border and systemic nature of algorithmic risks. As a result, a growing body of policy debate points towards the need for more coordinated, specialised and technically informed forms of oversight.

At the regional level, European organisations, particularly the European Union (EU) and the Council of Europe (CoE), have played a central role in shaping both the normative and institutional dimensions of AI governance. In 2024, each adopted the first legally binding instruments addressing AI. The EU *Artificial Intelligence Act* introduced a new institutional architecture for AI oversight, including supranational coordination mechanisms to ensure consistent implementation across Member States. By contrast, the CoE's *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law* adopted a more restrained approach, relying on a Conference of the Parties rather than a specialised supervisory body. This divergence reflects ongoing uncertainty about the capacity of existing human rights mechanisms to address AI-related harms and suggests that further institutional adaptation may be required in practice.

Beyond Europe, regional organisations such as the African Union (AU) and the Organization of American States (OAS) have likewise begun to explore the human rights implications of AI. However, these initiatives remain at a relatively early stage of development and, unlike the frameworks examined in this contribution, do not currently involve newly established institutional bodies specifically dedicated to AI governance or regulation. For this reason, they fall outside the scope of the present contribution, which focuses on developments within the UN system and selected European frameworks.

Taken together, these developments illustrate a growing recognition that safeguarding human rights in the algorithmic era requires more than the adaptation of substantive norms: it demands a new generation of specialised institutions capable of providing meaningful oversight, ensuring accountability, and bridging the systemic gaps left by traditional human rights mechanisms. The central question, therefore, is not whether institutional innovation is necessary, but how the international community can design mechanisms that are sufficiently agile, authoritative, and technically informed to meet the

unprecedented challenges posed by AI. This contribution therefore examines whether current initiatives constitute genuine institutional innovation or merely incremental adjustments within a traditional model of human rights oversight that may prove insufficient to address the systemic challenges posed by AI.

7.1 Why Artificial Intelligence Challenges Traditional Human-Rights Institutions

The emergence of AI has generated an extensive body of scholarship addressing its implications for fundamental rights. Much of this literature focuses on identifying specific risks associated with algorithmic systems, including privacy intrusions, discriminatory decision-making, or the manipulation of information environments.[5] The challenge therefore lies not merely in identifying new categories of rights-related risks, but in understanding how the structural features of AI systems interact with institutional models of human-rights protection. While these concerns are increasingly well documented,[6] they also reveal a deeper structural problem: the mismatch between the characteristics of algorithmic harms and the institutional architecture through which human rights protection has traditionally been organised.

---

[5] See SARABDEEN, J.: *Protection of the rights of the individual when using facial recognition technology*. In: Heliyon, Vol. 8, No. 3, 2022, pp. 1–12; KÖCHLING, A., WEHNER, M. C.: *Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development*. In: Business Research, Vol. 13, 2020, pp. 795–848; FRA – European Union Agency for Fundamental Rights: *Bias in Algorithms – Artificial Intelligence and Discrimination*, 2022, pp. 22–25. Online: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf; VAN GILS, F., MÜLLER, W., PRÜFER, J.: *Microtargeting, voters' unawareness, and democracy*. In: The Journal of Law, Economics, and Organization, Vol. 41, No. 2, 2025, pp. 634–653; European Commission for the Efficiency of Justice: *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*. Council of Europe: Strasbourg, 3 – 4 December 2018, pp. 8 and 15, Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

[6] See HOLZINGER, A., ZATLOUKAL, K., MÜLLER, H.: *Is human oversight to AI systems still possible?* In: New Biotechnology, Vol. 85, 2025, pp. 59–62; SÖDERLUND, K.*: High-risk AI transparency? On qualified transparency mandates for oversight bodies under the EU AI Act.* In: Technology and Regulation, Vol. 2025, 2025, pp. 96–113; CHEONG, B. CH.: *Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making*. In: *Frontiers in Human Dynamics*, Vol. 6, 2024, pp. 1–11; RODRIGUES, R.: *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities*. In: Journal of Responsible Technology, Vol. 4, 2020, pp. 1–12.

For the purposes of this contribution, AI is understood as a machine-based system that, for explicit or implicit objectives, infers from the input it receives how to generate outputs – such as predictions, content, recommendations or decisions – that may influence physical or virtual environments.[7] Such outputs may directly or indirectly interfere with the enjoyment of human rights and fundamental freedoms. These interferences are further complicated by several structural characteristics of AI systems, including opacity, scale and distributed responsibility.

One frequently cited feature is the opacity of algorithmic systems, often described through the metaphor of the "black box". According to *Almada,* this metaphor frames an AI system as an object not amenable to scrutiny of its inner workings, an opacity that stems from technical factors such as the vast amount of data it processes and its technical complexity. Furthermore, it invokes a sense of danger and helplessness, as the inscrutability of black boxes is presented as a challenge to many existing legal rights and safeguards.[8] In practice, explanations can often only be constructed from observations of the relationships between the inputs and outputs of the system, particularly where the operator or provider of the AI system is unwilling to cooperate.[9] When data are used as, or processed to produce, evidence for a particular conclusion, it is reasonable to expect that the connection between the data and the resulting decision should remain intelligible and open to scrutiny and critical assessment.[10] AI systems are frequently proprietary and technically complex, which makes it difficult even for states to explain or audit the tools they deploy. At the same time, responsibility for algorithmic systems tends to be fragmented among developers, data suppliers, private companies and public authorities, while jurisdictional

---

7 Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Art. 2. See also Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, Art. 2; OECD: *Explanatory Memorandum on the Updated OECD Definition of an AI System*. OECD Artificial Intelligence Papers, No. 8, 2024, p. 4. Online: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf.

8 ALMADA, M.: *Governing the Black Box of Artificial Intelligence*. In: SSRN Electronic Journal, January 2023, p. 2. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4587609.

9 CASTELLUCCIA, C., LE MÉTAYER, D.: *Understanding algorithmic-making: Opportunities and challenges.* European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), European Parliament, March 2019, pp. 47–50.

10 MITTELSTADT, B.D., ALLO, P., TADDEO, M., WACHTER, S., FLORIDI, L.: *The ethics of algorithms: Mapping the debate*. In: Big Data & Society, Vol. 3, No. 2, 2016, p. 4.

questions become blurred as models may be trained in one jurisdiction, deployed in another, and affect individuals far beyond both.[11]

As a result, AI-related interferences are often systemic, technologically mediated and transnational, making them difficult to capture through complaint-based procedures and ex post adjudication, which remain the standard modes of operation for many existing human-rights institutions, including treaty bodies and regional courts.[12] Algorithmic harms therefore frequently operate at a systemic rather than an individual level: a biased system may affect thousands of individuals who share similar characteristics rather than producing a single identifiable violation.[13] A widely cited real-life example is the Dutch childcare benefits scandal ("Toeslagenaffaire"), where algorithmic risk-profiling systems used by public authorities falsely identified thousands of families—many with migrant backgrounds—as potential fraudsters. Rather than harming a single identifiable individual, the system systematically targeted groups sharing similar socio-economic and ethnic characteristics, leading to widespread financial hardship, social exclusion and loss of trust in public institutions.[14]

Taken together, these characteristics demonstrate that AI challenges not only the substantive content of human rights but also the institutional assumptions on which their protection has traditionally relied. Complaint-based procedures, territorially bounded jurisdiction and ex post review mechanisms are increasingly problematic to address harms arising from the deployment of AI systems. This institutional mismatch has therefore prompted renewed debates at both the global and regional levels regarding the need for new oversight bodies or adapted institutional architectures capable of addressing the specific risks posed by AI systems.

7.2 Global Institutional Developments on Artificial Intelligence and Human Rights

One of the main purposes of the UN is to achieve international cooperation in solving international problems of an economic, social, cultural, or

---

11 MATTHIAS, A.: *The responsibility gap: Ascribing responsibility for the actions of learning automata*. In: Ethics and Information Technology, Vol. 6, 2004, pp. 181–182.

12 See, e.g. O'FLAHERTY, M.: *Human rights and the UN: practice before the treaty bodies*. Second Edition, Nijhoff Law Specials - Vol. 54, Brill, 2021, 240 p.

13 WACHTER, S., MITTELSTADT, B.D., RUSSELL, CH.: *Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI*. In: Computer Law & Security Review, Vol. 41, 105567, 2021, pp. 5–6.

14 AMNESTY INTERNATIONAL: *Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms*, 25 October 2021. Online: https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/.

humanitarian character, and in promoting and encouraging respect for human rights and for fundamental freedoms for all.[15] Over the past eight decades, a comprehensive system of human rights protection has been developed within the framework of the UN. Through the adoption of a series of international treaties aimed at safeguarding specific categories of human rights and freedoms[16], as well as the rights of particularly vulnerable groups[17], several treaty-based monitoring bodies have been established. Most notable among them are the Human Rights Committee, the Committee on Economic, Social and Cultural Rights, as well as specialised bodies such as the Committee on the Elimination of Discrimination against Women, the Committee on the Rights of the Child, the Committee on the Elimination of Racial Discrimination, and the Committee against Torture. These bodies were created to supervise the implementation of States' obligations, to interpret treaty norms, and to contribute to the progressive development of international human rights law. Where applicable, their relevance is particularly pronounced for individuals, through procedures allowing the submission of individual complaints (under relevant optional protocols), which—despite their formally non-binding character—have generated a substantial body of legal views and have strengthened expectations of compliance and follow-up at the domestic level.

UN human rights treaty bodies have not yet developed a body of case law explicitly addressing AI as such. However, their existing legal views provides a normative framework that is highly relevant to AI-driven systems. Through their views and general comments, these bodies have long addressed issues such as

---

15 Charter of the United Nations, 1 UNTS XVI, 24 October 1945, Art. 1 (3).

16 International Covenant on Civil and Political Rights. New York, 16 December 1966; International Covenant on Economic, Social and Cultural Rights. New York, 16 December 1966; Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, New York, 10 December 1984; International Convention for the Protection of All Persons from Enforced Disappearance, New York, 20 December 2006.

17 International Convention on the Elimination of All Forms of Racial Discrimination. New York, 21 December 1965; Convention on the Rights of the Child. New York, 20 November 1989; Convention on the Elimination of All Forms of Discrimination against Women. New York, 18 December 1979; International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, New York, 18 December 1990; Convention on the Rights of Persons with Disabilities, New York, 13 December 2006.

unlawful interference with privacy[18], discrimination[19], and the absence of procedural safeguards in decision-making processes[20]. These principles are directly applicable to contemporary forms of algorithmic governance, including automated decision-making, large-scale data processing and profiling. In particular, the Human Rights Committee's interpretation of the right to privacy under Article 17 of the *International Covenant on Civil and Political Rights,*[21] the Committee on the Rights of the Child's *General Comment No. 25 (2021) – Children's rights in the digital environment*, which addresses issues of automated processing, profiling and the need of establishing safeguards[22], together with the broader practice of treaty bodies addressing indirect discrimination[23], provide important guidance for assessing the human-rights implications of AI.

Alongside treaty bodies, the UN has also developed charter-based mechanisms – most prominently the UN Human Rights Council (UN HRC), its Special Procedures, and the Universal Periodic Review – supported institutionally by the Office of the High Commissioner for Human Rights (OHCHR). The 2021 report of the OHCHR, *The Right to Privacy in the Digital Age*, provides a comprehensive assessment of the human-rights implications of AI within the UN framework. It highlights that data-driven AI systems enable unprecedented forms of data collection, aggregation and inference, allowing both States and private actors to track, analyse and predict human behaviour on a large scale[24], as well as interfering with the enjoyment of the right to privacy and related rights,[25] particularly in the absence of adequate safeguards. The report confirms that the key challenge is not the absence of applicable human rights norms, but the difficulty of ensuring their effective application in

---

18 See, e.g. UN Human Rights Committee: *Toonen v. Australia*, CCPR/C/50/D/488/1992, 31 March 1994, para. 8.3.

19 UN Human Rights Committee: *CCPR General Comment No. 18: Non-discrimination*, 10 November 1989.

20 UN Human Rights Committee: *General comment no. 32, Article 14, Right to equality before courts and tribunals and to fair trial*, CCPR/C/GC/32, 23 August 2007.

21 UN Human Rights Committee: *CCPR General Comment No. 16: Article 17 (Right to Privacy), The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation*, 8 April 1988.

22 UN Committee on the Rights of the Child: *General comment No. 25 (2021) on children's rights in relation to the digital environment*, CRC/C/GC/25, 2 March 2021, para. 17, 42, 61, 68, 70, 75.

23 UN Human Rights Committee: *Althammer et al. v. Austria*, CCPR/C/78/D/998/2001, 22 September 2003, para. 10.2.

24 UN Human Rights Council: *The right to privacy in the digital age - Report of the United Nations High Commissioner for Human Rights*, A/HRC/48/31, 13 September 2021, para. 12.

25 The right to freedom of expression, the right to a fair trial, the right to life, the right to respect for private life.

technologically complex and globally distributed environments.[26] Building on these considerations, the OHCHR report further emphasises that individuals affected by AI-related human rights violations must have access to effective remedies, both judicial and non-judicial, capable of addressing harms arising in complex and data-driven environments.[27]

Against this backdrop, structural shortcomings become particularly visible when examining the role of UN human rights treaty bodies in the context of AI governance. While these bodies provide a well-developed normative framework applicable to AI-related harms, their institutional design reflects a model of human rights supervision that is structurally misaligned with the operational logic of contemporary AI systems. Their mandates are primarily reactive, dependent on periodic state reporting and individual complaints, and thus inherently limited in their ability to engage with fast-evolving, large-scale, and often opaque technological developments. Moreover, treaty bodies lack the technical capacity and specialised expertise necessary to systematically assess complex AI systems. This gap is further compounded by procedural constraints, such as limited resources, significant backlogs in individual complaints, and the non-binding nature of their views, which collectively weaken their potential to provide timely and effective oversight. As a result, while treaty bodies remain central actors in interpreting and applying international human rights law, their current institutional configuration risks rendering them peripheral in the governance of AI. The challenge therefore does not lie in the absence of applicable legal standards, but rather in the capacity of existing institutional frameworks to interpret and enforce these standards in the context of complex, data-driven and often transnational technological systems. In addition, their state-centric mandate limits their ability to directly engage with private actors, despite the fact that much of the development and deployment of AI systems is driven by corporate entities operating across jurisdictions.

These challenges are increasingly reflected in ongoing policy discussions within the UN, including those surrounding the *Global Digital Compact* (GDC), which calls for stronger institutional responses to the human rights implications of AI. Despite its non-binding nature, the GDC constitutes an important step forward in the evolving architecture of global AI governance, signalling a growing recognition of the need for more coordinated and institutionalised responses. While the Compact acknowledges the important role of the OHCHR in

---

26 *Ibid,* para. 59.
27 *Ibid*, para. 38.

providing expert guidance on human rights and technology,[28] it also highlights the absence of a system-wide mechanism capable of addressing cross-border algorithmic risks.[29] Furthermore, a growing number of UN actors, independent experts and civil-society organisations have argued that existing institutional arrangements are insufficient, calling for the creation of dedicated mechanisms—such as a UN Special Rapporteur on Artificial Intelligence and Human Rights[30] or a system-wide AI coordination office[31]—to ensure more coherent global monitoring, accountability and governance of AI systems.

7.2.1 Bodies Created under the Global Digital Compact

In response to structural limitations, a key reference point for recent institutional developments within the UN is the High-Level Advisory Body on Artificial Intelligence, established by the UN Secretary-General in 2023.[32] In its final report, *Governing AI for Humanity*, the Body outlines a global governance architecture consisting of a scientific panel, a global dialogue and, ultimately, a dedicated UN AI Office.[33]

A first step in this direction is the establishment of the Independent International Scientific Panel on Artificial Intelligence, endorsed by the UN General Assembly in resolution 79/325 (2025). The Panel, composed of forty independent experts, is mandated to produce evidence-based, policy-relevant (but non-prescriptive) annual reports on the opportunities, risks and impacts of AI, supported by thematic briefs where necessary. Its work is guided by principles of independence, scientific rigour, multidisciplinarity and inclusiveness. The Panel is internally structured with co-chairs representing

---

28 United Nations: *Global Digital Compact*, A/79/L.2, 2024, pp. 6–7. Online: https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.

29 GOEL, B., NILKANTH, D., SINGH, S.: *AI Governance Beyond 2025: UN Pathways and Implications.* Geneva: Geneva Graduate Institute, 2025, p. 30. Online: https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2-----Shreya-Singh.pdf?utm_source=chatgpt.com.

30 Center for AI & Digital Policy: *Statement of the Center for AI & Digital Policy to the Vice–President of the United Nations Human Rights Council To support the Establishment of the Special Rapporteur for AI and Human Rights*, November 2025. pp. 2–3. Online: https://www.caidp.org/app/download/8585853363/CAIDP-UNHRC-VPRusu-AI-Treaty-09112025.pdf?t=1763220293&utm.

31 United Nations: *Global Digital Compact*, A/79/L.2, 2024, para. 72.

32 United Nations, *UN Secretary-General launches AI Advisory Body on risks, opportunities, and international governance of artificial intelligence*, Press release, 2023. Online: https:// www.un.org/sites/un2.un.org/files/231025_press-release-aiab.pdf.

33 United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 10 – 12.

both developed and developing countries, supported by vice-chairs and ad hoc working groups, and may engage external experts. It reports regularly to the UN General Assembly through interactive dialogues and presents its findings within the framework of the Global Dialogue on AI Governance.[34] In this sense, the Panel does not operate as a regulatory body, but rather as a knowledge-generating mechanism intended to underpin more coordinated and informed global governance responses. Accordingly, while the Panel may play a valuable role in informing global debates on AI governance, its contribution to the protection of human rights remains inherently limited. Given its strictly advisory and non-prescriptive mandate, the Panel does not directly address issues of accountability, enforcement or individual remedies, which are central to effective human rights protection.

Building on this scientific foundation, the second pillar of the proposed architecture takes the form of the Global Dialogue on AI Governance. Also established within the framework of the GDC and endorsed by the same UN General Assembly resolution, the Dialogue is expected to convene annually, beginning at the AI for Good Summit in 2026. It is conceived as a multistakeholder platform involving Governments and other relevant actors, aimed at fostering international cooperation, facilitating the exchange of best practices and lessons learned, and enabling open, transparent and inclusive discussions on AI governance, with a view to harnessing AI in support of the Sustainable Development Goals.[35] While the Dialogue represents a valuable step towards more inclusive and coordinated global AI governance, it remains primarily deliberative in nature and does not, in itself, establish mechanisms capable of ensuring the effective protection and enforcement of human rights.

Taken together, these developments illustrate a gradual shift towards a more structured and system-wide approach to AI governance within the UN. At the same time, however, they remain predominantly oriented towards coordination, knowledge-sharing and dialogue, rather than the creation of mechanisms specifically designed to ensure the effective protection of human rights in the context of AI. This emerging gap has, in turn, prompted growing calls for the establishment of more specialised, human-rights-focused institutional arrangements within the UN system.

---

[34] United Nations – General Assembly: *Terms of reference and modalities for the establishment and functioning of the Independent International Scientific Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance*, A/RES/79/325, 26 August 2025, para. 1.

[35] *Ibid*, para. 4.

7.2.2 Proposed New Human-Rights-Focused Bodies within the UN System

Alongside these emerging structures, several proposals have been advanced for more explicitly human-rights-oriented oversight mechanisms. For instance, the Center for AI & Digital Policy, a civil society organization focused on AI governance has formally recommended that the UN HRC establish a dedicated mandate holder responsible for monitoring AI-related human-rights impacts worldwide – a Special Rapporteur on AI and Human Rights.[36] Within the system of special procedures of the UN HRC, such a mandate would be entrusted to an independent expert or working group tasked with examining and reporting on human rights issues from a thematic or country-specific perspective. These mandates are designed to address human rights situations globally, regardless of a state's treaty commitments, and require mandate holders to monitor and respond promptly to alleged violations affecting individuals or groups. In the case of country-specific mandates, they generally cover the full spectrum of human rights unless defined otherwise, while remaining accountable to the UN HRC. With the support of the OHCHR, mandate holders typically carry out country visits and address individual cases as well as broader structural concerns through communications with States. They also contribute to the development of international human rights standards by preparing thematic reports and convening expert consultations, while engaging in advocacy, raising public awareness and providing guidance in the context of technical cooperation.[37] In this sense, such a mechanism could strengthen global oversight by combining thematic reporting, country engagement and a degree of early-warning capacity. While such a mechanism could enhance the visibility of AI-related human rights concerns and contribute to their conceptual development, its impact would likely remain constrained by the inherently soft and non-binding nature of the special procedures system, particularly in addressing complex issues of accountability and enforcement in the AI context.

---

36 Center for AI & Digital Policy: Statement of the Center for AI & Digital Policy to the Vice–President of the United Nations Human Rights Council To support the Establishment of the Special Rapporteur for AI and Human Rights, November 2025. Online: https://www.caidp.org/app/download/8585853363/CAIDP-UNHRC-VPRusu-AI-Treaty-09112025.pdf?t=1763220293&utm.

37 UN Human Rights Council: *Institution-building of the United Nations Human Rights Council,* A/HRC/RES/5/1, 18 June 2007, para. 39 – 54; UN Human Rights Council: *Manual of Operations*, August 2008. Online: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.ohchr.org%2Fsites%2Fdefault%2Ffiles%2FDocuments%2FHRBodies%2FSP%2FManual_August_2008.doc&wdOrigin=BROWSELINK.

The second proposal concerns the establishment of a new AI Office with system-wide responsibility for coordinating digital and AI governance across the UN. Recommended by the High-Level Advisory Body on AI in its final report[38] and referenced in the GDC,[39] such an office would be created within the UN Secretariat, reporting directly to the UN Secretary-General, as a light and agile coordination mechanism rather than a fully-fledged regulatory body. Its primary role would be to support and connect key elements of the emerging global AI governance architecture, including the Independent International Scientific Panel on Artificial Intelligence, Global Dialogue on AI Governance, and related initiatives such as standards exchange and capacity-building networks. The Office is conceived as a central coordinating hub or "connector", facilitating cooperation across the UN system and with external stakeholders, including governments, industry, civil society and academia. Rather than exercising monitoring or enforcement powers, it would focus on enhancing coherence, information-sharing and joint action, enabling the UN to respond in a more coordinated and unified manner to the evolving challenges of AI governance.[40] In this respect, the proposed AI Office would once again represent a mechanism primarily oriented towards coordination, monitoring and the facilitation of cooperation, rather than one endowed with supervisory or enforcement powers. From a human rights perspective, it may serve as an important platform for identifying and highlighting risks and potential violations associated with AI systems, as well as for fostering dialogue among relevant stakeholders. However, similarly to other emerging structures, it does not address the question of accountability in a substantive manner. In particular, it lacks the capacity to attribute responsibility for violations of international human rights law, to ensure effective remedies, or to secure adequate reparation for affected individuals. As such, while contributing to greater visibility and coherence in AI governance, its added value in terms of actual human rights protection remains structurally limited.

Together, these existing and proposed bodies reflect a gradual shift within the UN towards a more structured and multi-pillar institutional architecture for AI governance, in which human rights are increasingly recognised as a foundational component. At the same time, however, this evolving framework remains characterised by a persistent gap between normative ambition and

---

38 United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 70 – 72.

39 United Nations: *Global Digital Compact*, A/79/L.2, 2024, para. 72.

40 United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, pp. 70 – 72.

institutional capacity. While existing human rights mechanisms provide a well-developed normative framework applicable to AI-related harms, specific characteristics of AI systems, such as opacity, distributed responsibility and transnational deployment challenge the capacity of these mechanisms to interpret and effectively enforce applicable norms in practice. Although the newly emerging structures contribute to greater coordination and exchange of best practices, they primarily operate as platforms for dialogue rather than as mechanisms capable of ensuring effective human rights protection. As a result, fundamental questions of accountability, enforcement and access to effective remedies remain largely unaddressed, raising doubts as to whether these developments will ultimately translate into institutional arrangements capable of ensuring the effective protection of human rights in the context of AI.

7.3 Regional Human-Rights Bodies and Their Approaches to AI

Important lessons also emerge from regional human-rights frameworks, which often respond more dynamically than global institutions. In this regard, the following sections examine key European systems, namely the CoE and the EU, which increasingly shape the protection of human rights and the development of AI governance through context-specific and evolving approaches. The adoption of the *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law* and the *Artificial Intelligence Act* represents key initiatives, which may inspire other regions and states in the analysed area.

Before turning to these newly adopted frameworks and bodies, it is important not to overlook the role of judicial institutions, namely the European Court of Human Rights (ECtHR) and the Court of Justice of the European Union (CJEU) whose jurisprudence continue to play a crucial role in safeguarding fundamental rights and freedoms. Through their interpretative function, these courts contribute to adapting existing human rights standards to new technological contexts, including those arising from the development and use of AI.[41]

---

41 See, e.g. ECtHR: *Big Brother Watch and Others v. the United Kingdom*, Applications nos. 58170/13, 62322/14 and 24960/15, Judgment of 25 May 2021; ECtHR: *S. and Marper v. the United Kingdom*, Application nos. 30562/04 and 30566/04, Judgment of 4 December 2008; ECtHR: *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland*, Application no. 931/13, Judgment of 27 June 2017; ECtHR: *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023; CJEU: *Digital Rights Ireland Ltd v Minister for Communications*, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014; CJEU: *Ligue des droits humains v Council of Ministers*, Case C-817/19, Judgment of 21 June 2022.

7.3.1 Council of Europe

The CoE currently represents one of the most advanced regional frameworks linking AI with human-rights protection. Its long-standing normative foundation, centred on the *European Convention on Human Rights (ECHR)*[42], has enabled a gradual adaptation of existing standards to emerging technological challenges.[43] In this regard, the jurisprudence of the ECtHR already reflects the growing relevance of AI-related issues, including cases involving surveillance[44], algorithmic profiling[45] and the use of advanced data-processing techniques.[46] This is illustrated, for instance, in cases such as *Big Brother Watch and Others v. the United Kingdom* or *Gluhkin v. Russia*, which demonstrate the Court's evolving approach to technologically mediated interferences with fundamental rights.

However, reliance on interpretative mechanisms alone remains insufficient to address the systemic and forward-looking challenges posed by AI. Therefore, member States of the CoE, together with observer and selected non-member States, engaged in the development of the first international treaty addressing the entire life cycle of AI systems from a human-rights perspective. Despite the fact, that the *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law* is currently not in force due to the absence of the necessary number of ratifications, it represents an important step towards international legal regulation of AI. With respect to institutional design, the Convention refrains from establishing a specialised supervisory authority and instead assigns responsibility for monitoring, implementation, and the identification of potential issues, including the effects of reservations and declarations under the Framework Convention, to the Conference of the Parties.[47] This design choice reflects both the continued reliance on existing

---

42 *Convention on the Protection of Human Rights and Fundamental Freedoms*, Rome, 4th November 1950, E.T.S. No. 5.

43 BÅRDSEN, A.: *AI and Human Rights at the European Court of Human Rights*, Verfassungsblog, 8 December 2025. Online: https://verfassungsblog.de/ai-and-human-rights-at-the-european-court-of-human-rights/.

44 ECtHR: *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023.

45 CJEU: *Ligue des droits humains v Council of Ministers*, Case C-817/19, Judgment of 21 June 2022.

46 ECtHR: *Big Brother Watch and Others v. the United Kingdom*, Applications nos. 58170/13, 62322/14 and 24960/15, Judgment of 25 May 2021.

47 *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, Strasbourg, 5 September 2024, Council of Europe Treaty Series – No. 225, Art. 23.

human-rights mechanisms and the underlying uncertainty as to their capacity to effectively address AI-related harms. At the same time, it suggests that further institutional adaptation may become necessary as the Convention is implemented in practice.[48]

In parallel, other CoE bodies have played an important role in shaping the human rights framework applicable to AI. Within this broader institutional landscape, the Council of Europe Commissioner for Human Rights constitutes an important, though not exclusive, actor in shaping the analysed area. The Council of Europe Commissioner for Human Rights is established as an independent and impartial, non-judicial institution tasked with promoting education, awareness, and respect for human rights as enshrined in the CoE's legal framework. Distinct from judicial and treaty-based supervisory mechanisms, the Commissioner does not examine individual complaints but instead performs a complementary function focused on promotion, prevention, and guidance. In this capacity, the Commissioner contributes to the effective observance of human rights through recommendations, thematic reporting, and dialogue with national authorities, while also identifying structural shortcomings in law and practice and supporting their remediation. The mandate further encompasses cooperation with national human rights institutions, engagement with international bodies, and reporting to the Committee of Ministers and the Parliamentary Assembly, thereby reinforcing the broader human rights architecture of the CoE through non-binding yet authoritative interventions.[49] In the context of AI interaction with human rights, the Commissioner has issued influential recommendation *"Unboxing Artificial Intelligence: 10 steps to protect Human Rights"*, where he emphasises key measures such as human rights impact assessments, transparency, and public participation, while promoting equality, non-discrimination, data protection, and fundamental freedoms. It further calls for independent oversight, effective remedies, and the responsibility of States to ensure compliance by private actors and to foster AI literacy.[50] In this way, the Commissioner's work reflects a broader shift towards soft-law and

---

48 *Ibid*, Art. 23 (2) (a) and (b).

49 Council of Europe – Committee of Ministers: *Resolution (99) 50 on the Council of Europe Commissioner for Human Rights*, 7 May 1995, Art. 1 – 5. Online: https://search.coe.int/cm#{%22CoEIdentifier%22:[%220912594880212826%22], %22sort%22:[%22CoEValidationDate%20Descending%22]}.

50 Council of Europe – Commissioner for Human Rights: *Unboxing Artificial Intelligence: 10 steps to protect Human Rights,* May 2019, pp. 7 – 15. Online: https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

interpretative mechanisms in the governance of AI, paving the way for more structured and systematised guidance developed within the CoE framework.

Taken together, these developments position the CoE as a key normative actor in the European AI governance landscape. The interplay between judicial mechanisms, soft-law instruments, and the emerging treaty framework reflects an increasingly coherent approach to addressing the human-rights implications of AI. At the same time, the CoE's predominantly interpretative and coordination-based model also reveals certain limits in terms of enforcement and institutional capacity to respond to systemic AI-related harms.

7.3.2 The European Union

The EU has established itself as a key global player in AI governance, combining regulatory ambition with a firm commitment to safeguarding fundamental rights. This approach is driven by a dual ambition: to encourage innovation and technological advancement while ensuring that AI systems are developed and used in accordance with human rights, safety considerations, and ethical standards.[51] Rather than remaining at the level of abstract principles, the EU operationalises this commitment by embedding fundamental rights considerations directly within binding legal and regulatory structures.[52]

At its core lies the *Charter of Fundamental Rights of the European Union*, which provides a comprehensive catalogue of human rights applicable within the scope of EU law, including those particularly relevant to the deployment of AI, such as the rights to private and family life, protection of personal data, non-discrimination, and effective remedy and to a fair trial.[53] The interpretation of these rights has been progressively developed through the case law of the CJEU, particularly in areas related to data protection, surveillance, and automated

---

51 European Commission: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Fostering a European Approach to Artificial Intelligence*, Brussels, 21 April 2021, COM(2021) 205 final, p. 1 and 4.

52 Treaty on European Union (consolidated version). Official Journal of the European Union, 2012, C 326, pp. 13–390, Art. 2; 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, 2024/1689, 12.7.2024, Art. 1, 2 (7), 27.

53 Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012, pp. 391–407, Art. 6, 8, 21, 47.

decision-making. This is reflected in cases such as *Digital Rights Ireland*,[54] *Tele2 Sverige*,[55] *La Quadrature du Net*,[56] and *Schrems I and II*,[57] which collectively establish strict limits on surveillance, emphasise proportionality, and reinforce the requirement of effective judicial protection in data-driven contexts. More recently, in *SCHUFA Holding*[58], the Court adopted a functional interpretation of automated decision-making under EU data protection law, recognising that even probabilistic scoring systems may fall within the scope of Art. 22 General Data Protection Regulation (GPDR),[59] where they play a decisive role in practice. While not explicitly concerned with AI, the judgment offers a powerful normative lens for addressing AI-related risks, particularly by exposing the limits of formal human oversight and highlighting the regulatory significance of private actors.[60] Judicial protection at the EU operates through a decentralised system in which national courts apply EU law, while the CJEU ensures uniform interpretation through the preliminary ruling procedure.

7.3.2.1 Bodies Created under the AI Act

As regards the specific regulatory framework for AI, the EU has adopted the *Artificial Intelligence Act*, which represents the first comprehensive and binding legal instrument governing AI systems across their lifecycle. It establishes a harmonised set of rules on the placing on the market, deployment, and use of AI systems, while also introducing specific obligations for high-risk applications. The Act prohibits certain harmful practices, sets out transparency requirements, and regulates general-purpose AI models. It further provides for market surveillance, governance, and enforcement mechanisms, alongside measures designed to support innovation, particularly for small and medium-sized

54 CJEU: *Digital Rights Ireland Ltd v Minister for Communications*, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014.

55 CJEU: *Tele2 Sverige AB Post- och telestyrelsen*, Joined Cases C-203/15 and C-698/15, Judgment of 21 December 2016.

56 CJEU: *La Quadrature du Net and Others*, Joined Cases C-511/18, C-512/18 and C-520/18, Judgment of 6 October 2020.

57 CJEU: *Schrems*, Case C-362/14, Judgment of 6 October 2015 and *Schrems II*, Case C-311/18, Judgment of 16 July 2020.

58 CJEU: *SCHUFA Holding (Scoring)*, Case C-634/21, Judgment of 7 December 2023.

59 The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her; 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) (Text with EEA relevance), OJ L, 119, 4.5.2016, Art. 22 (1).

60 *Ibid*, para. 46 – 60.

enterprises and start-ups.[61] At its core, the Act adopts a risk-based approach, classifying AI systems according to their potential impact on fundamental rights and safety and imposing differentiated obligations on providers, deployers, and users. High-risk systems are subject to stringent requirements, including risk management, data governance, transparency, human oversight, and conformity assessments.[62] In contrast to the predominantly interpretative and coordination-based approach of the CoE, the EU model is characterised by detailed regulatory prescriptions and enforceable obligations.

The regulatory approach of the AI Act is complemented by a multi-layered governance structure designed to ensure its consistent implementation and enforcement. At the Union level, a central role is entrusted to the AI Office, which contributes to the development of expertise and supports the application of the Act.[63] This is accompanied by the establishment of the European Artificial Intelligence Board (EAIB), composed of representatives of the Member States, which facilitates coordination, promotes the harmonisation of administrative practices, and issues recommendations and opinions on the interpretation and application of the Act.[64] The AI Act encourages information-sharing, joint investigations, and coordinated responses to any risks or issues that may arise from AI systems.[65]

In addition, the governance framework incorporates advisory and expert-based mechanisms, including an advisory forum bringing together a broad range of stakeholders, and a scientific panel of independent experts tasked with providing technical and methodological support, particularly in relation to general-purpose AI models and systemic risks.[66] These bodies contribute to the development of common standards, the identification of emerging risks, and the strengthening of regulatory capacities across the Union.

---

61 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), Art. 1 (2).

62 *Ibid*, Art. 9 – 15.

63 *Ibid*, Art. 64.

64 *Ibid*, Art. 65 – 66.

65 STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.: Artificial Intelligence and International Human Rights Law Developing Standards for a Changing World, Cheltenham: Edward Elgar Publishing, 2024, p. 64.

66 *Ibid*, Art. 67 – 69.

At the national level, the effective application of the AI Act is ensured by designated competent authorities, including market surveillance authorities, which are responsible for supervision and enforcement. Notably, these authorities are required to possess multidisciplinary expertise, encompassing not only technical knowledge of AI systems but also an understanding of fundamental rights, data protection, and safety risks.[67]

From a human-rights perspective, this governance architecture reflects a shift towards a hybrid and expertise-driven model, combining legal oversight with technical and scientific input. At the same time, it is characterised primarily by coordination, supervision, and advisory functions rather than adjudicatory powers. Unlike traditional human-rights bodies, these institutions are not empowered to issue legally binding decisions determining violations of fundamental rights or to grant remedies to individuals. Their role is therefore best understood as complementary, enhancing the capacity of existing legal frameworks to address the complex and evolving human-rights implications of AI systems.

7.3.2.2 Governance Bodies Supporting the Protection of Fundamental Rights in the Context of AI

The framework on AI is further complemented by instruments such as the GDPR,[68] the Data Act,[69] the Digital Services Act,[70] and the Digital Markets Act,[71] which together reinforce the protection of fundamental rights in the digital environment, especially regarding the automated decision-making. Within this broader regulatory ecosystem, the European Data Protection Board (EDPB) plays a central role in ensuring the consistent application of EU data protection law, particularly the GDPR. Through the adoption of guidelines,

---

67 *Ibid*, Art. 70.

68 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) (Text with EEA relevance), OJ L, 119, 4.5.2016.

69 Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance), OJ L 2023/2854, 22.12.2023.

70 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), OJ L 277, 27.10.2022.

71 Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ 2022 L 265, 12.10.2022.

recommendations, and binding decisions in dispute resolution procedures, the EDPB significantly contributes to the interpretation and operationalisation of fundamental rights related to privacy and data protection.[72] Although it is not a human-rights body and does not provide individual remedies, its activities shape the practical scope of these rights, especially in areas involving automated decision-making, profiling, and large-scale data processing. As such, the EDPB may be understood as a quasi-regulatory body contributing to the protection of fundamental rights in the digital environment.

A complementary role is played by the European Data Protection Supervisor (EDPS), whose mandate has further evolved in the context of the AI Act. As reflected in its strategic framework, the EDPS extends its supervisory function beyond traditional data protection[73] towards a more integrated oversight of AI systems, particularly within EU institutions, where it assumes the role of a market surveillance authority and, in certain cases, as a notified body responsible for conformity assessments of high-risk AI systems. Beyond direct supervision, the EDPS plays an active role within the EU governance framework by contributing to the work of the EAIB and supporting the consistent implementation of the AI Act through legal and technical expertise.[74] At the same time, it adopts a cooperative and capacity-building approach towards EU institutions, promoting human-centric and compliant AI through guidance, knowledge-sharing, and initiatives such as regulatory sandboxes and expert networks. This is complemented by strong international engagement aimed at enhancing supervisory capabilities and aligning best practices across jurisdictions.[75] Overall, the EDPS reflects a shift towards a comprehensive, lifecycle-oriented model of AI oversight that integrates supervision, coordination, and expertise. However, despite its expanded mandate, the EDPS remains a supervisory and regulatory authority rather than an adjudicatory body, as it does not determine violations of fundamental rights in a judicial sense nor provide individual remedies.

Recent study for the European Parliament's Committee on Industry, Research and Energy on the interplay between the AI Act and the EU digital legislative framework highlights several structural frictions within the EU's

---

72 General Data Protection Regulation, Arts. 68 – 76.

73 Data Act, Art. 37.

74 Artificial Intelligence Act, Art. 57, 65, 70, 74 (9), 100.

75 European Data Protection Supervisor: *Towards trustworthy AI in the EU public administration: The EDPS Compass for its new role under the AI Act (2026–2027)*, 2026, pp. 9 – 16. Online: https://www.edps.europa.eu/system/files/2026-03/26-03-17_edps-compass-under-ai-act-2026-2027_en.pdf.

digital regulatory framework and formulates recommendations for its further development. In the short term, the study emphasises the need for enhanced coordination among supervisory authorities, including joint guidance, harmonised enforcement practices, and mutual recognition of assessments such as data protection impact assessment and fundamental rights impact assessments. In the medium term, it suggests targeted legislative adjustments aimed at clarifying institutional roles, reducing overlapping obligations, and strengthening the practical enforceability of rights in AI contexts. From a longer-term perspective, the study calls for a broader reconfiguration of the EU's digital regulatory architecture towards greater simplification, coherence, and strategic integration, enabling more agile compliance while preserving fundamental rights and safety standards.[76] Overall, these recommendations underline the importance of a more coordinated and streamlined governance model capable of both safeguarding constitutional values and supporting the development of a competitive and sovereign AI ecosystem.

While the EDPB and the EDPS play a central role in the interpretation, supervision, and enforcement of EU rules relevant to AI, the effective protection of fundamental rights in this context also depends on the availability of empirical evidence and risk identification. In this regard, a complementary function is performed by the European Union Agency for Fundamental Rights (FRA). FRA is mandated to support EU institutions, bodies, offices, and agencies, as well as Member States in the implementation of EU law, by providing expertise and assistance on fundamental rights, with a view to ensuring that their actions and policies fully respect fundamental rights within their respective spheres of competence.[77] In the context of the AI Act, its role is further reinforced, as it is designated as a permanent member of the advisory forum established under Art. 67, thereby contributing its expertise to the governance and implementation of the EU's AI regulatory framework. Although it does not exercise supervisory or enforcement powers, its work strengthens the evidence base for regulatory and

---

[76] GRAUX, H., GARSTKA, K., MURALI, N., CAVE, J., BOTTERMAN, M.: *Interplay between the AI Act and the EU digital legislative framework*. European Parliament - Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry, PE 778.575 - October 2025, pp. 78 – 86. Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU(2025)778575_EN.pdf.

[77] *Council Regulation (EC) No 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights,* OJ L 4M, 8.1.2008, Art. 2.

judicial action, thereby playing a crucial role in the broader ecosystem of fundamental rights protection in the context of AI.[78]

A defining feature of the EU framework applicable to AI is its decentralised and multi-layered character. While the Union has developed an increasingly dense and sophisticated regulatory architecture, the effective enforcement of fundamental rights ultimately relies on national courts and administrative authorities, operating alongside EU-level bodies. This model allows for flexibility and context-sensitive application, reflecting the diversity of national legal systems. At the same time, however, it may give rise to fragmentation, overlapping competences, and uneven levels of protection across Member States. In this sense, the EU framework reveals an inherent tension between regulatory ambition and institutional coherence, raising broader questions as to whether existing mechanisms are sufficient to ensure consistent and effective human-rights protection in the context of AI.

7.4 Do We Need New Human-Rights Bodies?

Recent developments at the UN and selected regional organisations bring us to a central question: do existing human rights mechanisms remain adequate to address the challenges posed by AI, or is there a need for the creation of new specialised bodies? In practice, three main pathways are emerging.

The first is institutional adaptation, whereby existing mechanisms—such as treaty bodies, regional courts, and national human-rights institutions—expand their mandates and expertise. This approach benefits from established legitimacy and institutional continuity, yet it may struggle to fully capture the technical complexity and systemic nature of AI-related risks. In this context, closer cooperation between legal actors and technical experts becomes essential, enabling courts and monitoring bodies to develop interpretations that adequately reflect the specificities of AI systems. Emerging initiatives, such as

[78] See: Fundamental Rights Agency: *Assessing High-risk Artificial Intelligence: Fundamental Rights Risks*, 4 December 2025. Online: https://fra.europa.eu/en/publication/2025/assessing-high-risk-ai; Fundamental Rights Agency: *Bias in algorithms - Artificial intelligence and discrimination*, 8 December 2022. Online: https://fra.europa.eu/en/publication/2022/bias-algorithm; Fundamental Rights Agency: *Getting the future right – Artificial intelligence and fundamental rights*, 14 December 2020. Online: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights.

the introduction of regulatory sandboxes within human-rights frameworks, further illustrate attempts to enhance institutional adaptability.[79]

The second pathway is hybrid governance, combining traditional human-rights bodies with regulatory authorities, scientific panels, and independent expert networks. This model seeks to bridge the gap between legal oversight and technical expertise. Recent UN initiatives, including the establishment of the Independent International Scientific Panel on AI and the envisaged Global Dialogue on AI Governance, reflect a clear preference for hybrid and knowledge-based institutional models rather than purely legal supervisory bodies. On the other hand, these bodies are not vested with the authority to adopt legally binding decisions vis-à-vis States or private actors, nor to determine violations of human rights and fundamental freedoms or award remedies to affected individuals. Their role should therefore be understood as merely complementary, contributing to the development of knowledge and expertise on the human-rights implications of AI, rather than providing direct avenues for accountability or redress.

The third approach involves the creation of new, dedicated human-rights bodies focused specifically on AI, whether in the form of a Special Rapporteur, an advisory mechanism within the OHCHR, or a more comprehensive UN AI Office. While such bodies may offer greater specialisation and responsiveness, their establishment raises questions regarding institutional duplication, coordination, and political feasibility.

Rather than a binary choice between institutional adaptation and the creation of new bodies, the emerging landscape points towards a combination of these approaches. The governance of AI is likely to require greater reliance on interdisciplinary expertise, closer integration of technical and legal knowledge, more preventive and forward-looking mechanisms, strengthened regulation of private actors, and intensified dialogue across jurisdictions. In this context, the establishment of new institutional structures may play a complementary role, contributing to a more comprehensive and coherent response to the human-rights challenges posed by AI.

---

[79] Keynote speech of Prof. Dr. Vasilka Sancin, PhD., Judge of the European Court of Human Rights entitled "The Role of AI Sandboxes before International Human Rights Bodies and Courts" at the international scientific conference "Human Rights and Freedoms under Pressure from Algorithms: International and National Challenges of AI" (28 November 2025, Faculty of Law, Pavol Jozef Šafarik University in Košice), where she addressed the role of regulatory experimentation frameworks in the context of AI and human rights protection.

Conclusion

AI is fundamentally transforming the conditions under which human rights and freedoms are respected, protected, and fulfilled. As this contribution has shown, existing human rights institutions—originally designed to address state-centred and territorially bounded forms of harm—are increasingly confronted with automated, opaque, and transnational systems of decision-making. This creates a structural mismatch between the nature of algorithmic harms and the institutional frameworks traditionally responsible for their oversight.

The analysis demonstrates that, at both global and regional levels, important steps have already been taken to address these challenges. Within the UN system, recent initiatives, including the GDC and the establishment of expert and coordination mechanisms, reflect a growing recognition of the need for more structured and system-wide responses. At the regional level, the CoE and the EU have developed distinct yet complementary approaches: the former relying primarily on interpretative and coordination-based mechanisms, and the latter advancing a more enforcement-oriented model grounded in binding regulation and supervisory governance structures. At the same time, the findings of this research reveal persistent structural limitations. Existing mechanisms remain largely reactive, fragmented, and insufficiently equipped to address the systemic, cross-border, and technically complex nature of AI-related harms. In particular, the analysis highlights the absence of institutional arrangements capable of ensuring effective accountability, attributing responsibility across complex value chains, and providing accessible and timely remedies to affected individuals.

Against this background, three main pathways for institutional development can be identified: the adaptation of existing mechanisms, the emergence of hybrid governance models combining legal and technical expertise, and proposals for the creation of new specialised bodies. Rather than signalling a clear transition towards a wholly new generation of human rights institutions, these developments point towards a pluralistic and multi-layered governance model in which different types of bodies perform complementary functions.

The central finding of this contribution is therefore that the current trajectory of institutional development does not replace existing human rights bodies but progressively reshapes their functioning through the integration of regulatory, supervisory, and knowledge-based mechanisms. While this model enhances the capacity to identify and manage risks, it does not fully resolve the underlying accountability gap. Whether this evolving framework will ultimately give rise to a new generation of human rights bodies remains an open question. What is clear, however, is that safeguarding human rights in the age of AI requires

not only the continued evolution of legal norms, but also the development of institutional arrangements capable of combining legal authority with technical expertise, ensuring both preventive oversight and effective remedies in an increasingly complex and data-driven environment.

Note

The author acknowledges the use of AI-assisted tools exclusively for language editing and stylistic refinement. All ideas, analyses, and legal arguments presented in this chapter are entirely the author's own and reflect her independent scholarly judgment.

List of Literature and Sources

1. AHMAD, S. et al.: *The Impact of Artificial Intelligence on Business & Social Values: Benefits, Challenges, And Future Directions*. In: Educational Administration: Theory and Practice, Vol. 30, No. 6, 2024, pp. 3174 – 3180.
2. ALMADA, M.: *Governing the Black Box of Artificial Intelligence*. In: SSRN Electronic Journal, January 2023, 26 p. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4587609.
3. AMNESTY INTERNATIONAL: *Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms*, 25 October 2021. Online: https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/.
4. BAKOŠOVÁ, L.: *Ethical and legal aspects of the use of artificial intelligence in health and nursing care.* In: Studia Iuridica Cassoviensia, Vol. 8, No. 2, 2020, pp. 3 – 18.
5. BÅRDSEN, A.: *AI and Human Rights at the European Court of Human Rights*, Verfassungsblog, 8 December 2025. Online: https://verfassungsblog.de/ai-and-human-rights-at-the-european-court-of-human-rights/.
6. BENKO, R.: *The Impact of Artificial Intelligence on Fundamental Rights and Freedoms and Vice-Versa*. In: KLUČKA, J., BAKOŠOVÁ, L., SISÁK, Ľ.: Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges, Praha: Leges, 2021, pp. 39 – 50.
7. CASTELLUCCIA, C., LE MÉTAYER, D.: *Understanding algorithmic-making: Opportunities and challenges*. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), European Parliament, March 2019, 104 p. Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU%282019%29624261_EN.pdf.
8. Center for AI & Digital Policy: *Statement of the Center for AI & Digital Policy To the Vice–President of the United Nations Human Rights Council To support the Establishment of the Special Rapporteur for AI and Human Rights*, November 2025. Online: https://www.caidp.org/app/download/8585853363/CAIDP-UNHRC-VPRusu-AI-Treaty-09112025.pdf?t=1763220293&utm.
9. CJEU: *Digital Rights Ireland Ltd v Minister for Communications*, Joined Cases C-293/12 and C-594/12, Judgment of 8 April 2014.
10. CJEU: *La Quadrature du Net and Others*, Joined Cases C-511/18, C-512/18 and C-520/18, Judgment of 6 October 2020.

11. CJEU: *Ligue des droits humains v Council of Ministers*, Case C-817/19, Judgment of 21 June 2022.
12. CJEU: *Schrems*, Case C-362/14, Judgment of 6 October 2015.
13. CJEU: *Schrems II*, Case C-311/18, Judgment of 16 July 2020.
14. CJEU: *SCHUFA Holding (Scoring)*, Case C-634/21, Judgment of 7 December 2023.
15. CJEU: *Tele2 Sverige AB Post- och telestyrelsen*, Joined Cases C-203/15 and C-698/15, Judgment of 21 December 2016.
16. Council of Europe: *Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications*, Study DGI(2017)12, Strasbourg: Council of Europe, 2018, 28 p. Online: https://rm.coe.int/algorithms-and-human-rights-en-rev/16808f06aa.
17. Council of Europe: *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 2020. Online: https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016809c820f.
18. Council of Europe – Commissioner for Human Rights: *Unboxing Artificial Intelligence: 10 steps to protect Human Rights,* May 2019. Online: https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.
19. Council of Europe – Committee of Ministers: *Resolution (99) 50 on the Council of Europe Commissioner for Human Rights*, 7 May 1995. Online: https://search.coe.int/cm#{%22CoEIdentifier%22:[%220912594880212826%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}.
20. Council Regulation (EC) No 168/2007 of 15 February 2007 establishing a European Union Agency for Fundamental Rights, OJ L 4M, 8.1.2008.
21. ELBERT, Ľ.: *Medzinárodnoprávny úvod k používaniu autonómnych zbraňových systémov*. Praha: Leges, 2022, 109 p.
22. European Commission: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Fostering a European Approach to Artificial Intelligence*, Brussels, 21 April 2021, COM(2021) 205 final.
23. European Commission for the Efficiency of Justice: *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.* Strasbourg: Council of Europe, 3 – 4 December 2018, 79 p. Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

24. ECtHR: *Big Brother Watch and Others v. the United Kingdom*, Applications nos. 58170/13, 62322/14 and 24960/15, Judgment of 25 May 2021.
25. ECtHR: *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023.
26. ECtHR: *S. and Marper v. the United Kingdom*, Application nos. 30562/04 and 30566/04, Judgment of 4 December 2008.
27. ECtHR: *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland*, Application no. 931/13, Judgment of 27 June 2017.
28. European Data Protection Supervisor: *Towards trustworthy AI in the EU public administration: The EDPS Compass for its new role under the AI Act (2026–2027)*, 2026, 18 p. Online: https://www.edps.europa.eu/system/files/2026-03/26-03-17_edps-compass-under-ai-act-2026-2027_en.pdf.
29. European Parliamentary Research Service: *Understanding Algorithmic Decision-Making: Opportunities and Challenges for the EU*, PE 624.261, Brussels: European Parliament, 2019, 104 p. Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf.
30. FRA – European Union Agency for Fundamental Rights: *Bias in Algorithms – Artificial Intelligence and Discrimination*, 2022, 106 p. Online: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf.
31. FRA – European Union Agency for Fundamental Rights: *Getting the Future Right – Artificial Intelligence and Fundamental Rights*, 2020, 108 p. Online: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights.
32. Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Strasbourg, 5 September 2024, Council of Europe Treaty Series – No. 225.
33. GOEL, B., NILKANTH, D., SINGH, S.: *AI Governance Beyond 2025: UN Pathways and Implications.* Geneva: Geneva Graduate Institute, 2025, 42 p. Online: https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2----Shreya-Singh.pdf?utm_source=chatgpt.com.
34. GRAUX, H., GARSTKA, K., MURALI, N., CAVE, J., BOTTERMAN, M.: *Interplay between the AI Act and the EU digital legislative framework*. European Parliament - Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry, PE 778.575 - October 2025, 108 pp. Online:

https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU(2025)778575_EN.pdf.

35. HOLZINGER, A., ZATLOUKAL, K., MÜLLER, H.: *Is human oversight to AI systems still possible?* In: New Biotechnology, Vol. 85, 2025, pp. 59 – 62.
36. *Charter of the United Nations*, San Francisco, 24 October 1945.
37. CHEONG, B. CH.: *Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making*. In: Frontiers in Human Dynamics, Vol. 6, 2024, pp. 1 – 11.
38. JAKAB, R.: *Artificial Intelligence and Administration Decision-Making*. In: Acta Universitatis Carolinae, Vol. 72, No. 1, 2026, pp. 31 – 43.
39. KÖCHLING, A., WEHNER, M. C.: *Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development.* In: Business Research, Vol. 13, 2020, pp. 795 – 848.
40. MATTHIAS, A.: *The responsibility gap: Ascribing responsibility for the actions of learning automata*. In: Ethics and Information Technology, Vol. 6, 2004, pp. 175 – 183.
41. MCGREGOR, L., MURRAY, D., NG, V.: *International Human Rights Law as a Framework for Algorithmic Accountability*. In: International and Comparative Law Quarterly, Vol. 68, No. 2, 2019, pp. 309 – 343.
42. MITTELSTADT, B.D., ALLO, P., TADDEO, M., WACHTER, S., FLORIDI, L.: *The ethics of algorithms: Mapping the debate*. In: Big Data & Society, Vol. 3, No. 2, 2016, pp. 1 – 21.
43. OECD: *Explanatory Memorandum on the Updated OECD Definition of an AI System*. OECD Artificial Intelligence Papers, No. 8, 2024. Online: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf.
44. O'FLAHERTY, M.: *Human rights and the UN: practice before the treaty bodies*. Second Edition, Nijhoff Law Specials - Vol. 54, Brill, 2021, 240 p.
45. 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) (Text with EEA relevance), OJ L, 119, 4.5.2016.
46. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), OJ L 277, 27.10.2022.

47. Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ 2022 L 265, 12.10.2022.
48. Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance), OJ L 2023/2854, 22.12.2023.
49. 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, 2024/1689, 12.7.2024.
50. RODRIGUES, R.: L*egal and human rights issues of AI: Gaps, challenges and vulnerabilities.* In: Journal of Responsible Technology, Vol. 4, 2020, pp. 1 – 12.
51. SARABDEEN, J.: *Protection of the rights of the individual when using facial recognition technology*. In: Heliyon, Vol. 8, No. 3, 2022, pp. 1 – 12.
52. SEMAN, T.: *Risks of Partial Exploitation of Artificial Intelligence Potential in Administrative Proceedings.* In: Studia Iuridica Cassoviensia, Vol. 13, Special Issue, 2025, pp. 155 – 169.
53. SÖDERLUND, K.: *High-risk AI transparency? On qualified transparency mandates for oversight bodies under the EU AI Act*. In: Technology and Regulation, Vol. 2025, 2025, pp. 96 – 113.
54. STASZCZYK, P.: *Navigating the AI landscape in the EU: fostering innovation while upholding ethical principles*. In: BALCERZAK, M., KAPELAŃSKA-PRĘGOWSKA, J.: Artificial Intelligence and International Human Rights Law Developing Standards for a Changing World, Cheltenham: Edward Elgar Publishing, 2024, pp. 45 – 67.
55. *Treaty on European Union (consolidated version).* Official Journal of the European Union, 2012, C 326, pp. 13–390.
56. United Nations – AI Advisory Body: *Governing AI for Humanity*, New York: United Nations Publications, 2024, 101 p.
57. UN Committee on the Rights of the Child: *General comment No. 25 (2021) on children's rights in relation to the digital environment*, CRC/C/GC/25, 2 March 2021.
58. UN General Assembly: *Terms of reference and modalities for the establishment and functioning of the Independent International Scientific*

Panel on Artificial Intelligence and the Global Dialogue on Artificial Intelligence Governance, A/RES/79/325, 26 August 2025.

59. UN Human Rights Committee: *Althammer et al. v. Austria*, CCPR/C/78/D/998/2001, 22 September 2003.
60. UN Human Rights Committee: *CCPR General Comment No. 16: Article 17 (Right to Privacy), The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation*, 8 April 1988.
61. UN Human Rights Committee: *CCPR General Comment No. 18: Non-discrimination*, 10 November 1989.
62. UN Human Rights Committee: *Toonen v. Australia*, CCPR/C/50/D/488/1992, 31 March 1994.
63. UN Human Rights Council: *Institution-building of the United Nations Human Rights Council*, A/HRC/RES/5/1, 18 June 2007.
64. UN Human Rights Council: *Manual of Operations*, August 2008. Online: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.ohchr.org%2Fsites%2Fdefault%2Ffiles%2FDocuments%2FHRBodies%2FSP%2FManual_August_2008.doc&wdOrigin=BROWSELINK.
65. UN Human Rights Council: *The right to privacy in the digital age - Report of the United Nations High Commissioner for Human Rights*, A/HRC/48/31, 13 September 2021.
66. United Nations: *Global Digital Compact*, A/79/L.2, 2024, 16 pp. Online: https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.
67. United Nations, *UN Secretary-General launches AI Advisory Body on risks, opportunities, and international governance of artificial intelligence*, Press release, 2023. Online: https://www.un.org/sites/un2.un.org/files/231025_press-release-aiab.pdf.
68. VAN GILS, F., MÜLLER, W., PRÜFER, J.: *Microtargeting, voters' unawareness, and democracy*. In: The Journal of Law, Economics, and Organization, Vol. 41, No. 2, 2025, pp. 634 – 653.
69. WACHTER, S., MITTELSTADT, B.D., RUSSELL, CH.: *Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI*. In: Computer Law & Security Review, Vol. 41, 105567, 2021, pp. 1 – 31.

Author details

Mgr. Lucia BAKOŠOVÁ, PhD.

Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International Law and European Law
email: lucia.bakosova@upjs.sk

8 Automatizované rozhodování a profilování v pracovněprávních vztazích

Jakub Morávek

https://doi.org/10.33542/HRF-0541-2-08

Abstrakt

Příspěvek bude zaměřen na právní aspekty automatizovaného rozhodování a profilování v pracovněprávních vztazích z hlediska stávající a očekávané právní úpravy pracovněprávních vztahů, ochrany osobních údajů a umělé inteligence v právním řádu České republiky. Pozornost bude zvlášť zaměřena na důsledky automatizovaného rozhodování a profilování pro odpovědnost a odměňování zaměstnanců, a to včetně procesněprávních aspektů materie.

Úvod

Historie lidstva se v návaznosti na sociální, technologické a další milníky periodizuje do široké škály etap. Každá další éra je přitom v důsledku zrychlujícího se pokroku časově kratší než ta předchozí.

Přešli jsme od pravěku přes starověk, středověk a novověk do moderní doby.

Zda se v ní stále nacházíme, anebo zda je či spíše bude informační věk samostatnou historickou etapou (jde již o dobu postmoderní), bude otázkou pro budoucí generace.

Informační věk, který se vyznačuje digitálními a elektronickými technologiemi, se počal pozvolna rozvíjet v prvé poloviny 20. století. Iniciačním momentem byl vynález tranzistoru, který následoval vynález mikroprocesoru. Navázalo propojování počítačů a vytváření prvních počítačových sítí, předchůdců dnešního internetu.[1]

Později přišel nástup osobních počítačů, médií pro přenos dat, mobilních telefonů a nakonec internetu.

Pozvolná chůze, jíž zpočátku informační věk kráčel, přešla ve svižný klus, který se s nástupem umělé inteligence přerodil ve zběsilý sprint mající Boltovské parametry.

Co přinesou kvantové počítače a další rozvoj digitálních technologií, jenž se předpokládá v následujících letech, je (zejména s přihlédnutím k posunům a změnám posledních 20ti let) jen obtížně představitelné.

---

1 V této souvislosti srov. například BLACK, J.: *Historie světa*. Praha: Universum, 2024; BÁRTA, J.: *Sedm zákonů. Jak se civilizace rodí, rostou a upadají*. Praha: Jota, 2021.

Aktuální míra a rychlost pokroku, navyšování výpočetního výkonu a rozšiřování možnosti použití digitálních technologií, které jde ruku v ruce se zlevňováním prostředků a jejich zpřístupňováním široké veřejnosti, je stejně fascinující jako zneklidňující.

Zneklidnění vychází z toho, že zpřístupňování a zavádění prostředků moderní techniky, digitálních služeb atd., do běžného života stojí na tržním principu. I když jsou hesla a slogany, s nimiž se prostředky a služby nabízí, všemožně líbivá (a v řadě aspektů samozřejmě i pravdivá) - prodejci slibují vyšší rychlost, úsporu, všemožnou pomoc atd. - motiv je ve skutečnosti prosťounký. Je jím utržit co nejvíce.

A tak, stejně jako již tolikrát dříve, se začínají nové vynálezy (prostředky moderní techniky a digitální služby) využívat s (tradičně částečně bláhovou a tudíž i mylnou) vírou v pokrok a v dobro pramenící z pokroku, aniž by existovala věrohodná a pravdivá představa o tom, jaké mohou přinést negativní důsledky.

Oproti předcházím lidským etapám zde však je jeden podstatný rozdíl.

Je-li zavádění vynálezu či inovace pozvolné, související škodlivé jevy je možné identifikovat a proti negativním důsledkům (více méně úspěšně) zakročit. Eventuálně širší negativní důsledky eliminuje pokrok sám další inovací nebo vynálezem.

Prvně naznačený samoregulační efekt kvůli rychlosti nástupu moderních technologií a digitálních služeb spolu s nedostatečnou kapacitou a flexibilitou na straně veřejné moci (regulačních orgánů) v podstatě odpadá. Druhý vyloučen není. Jelikož se však zdá, že se využívání informačních technologií a digitálních služeb bude prohlubovat, než že by se od nich vbrzku upouštělo, je spíše nepravděpodobné, že by nastal.

Žijeme tak v nebezpečnější době, než se může na první pohled zdát.

Informační technologie a digitální služby můžeme v mnohém cenit – zrychlily komunikaci, nabízí přístup k téměř nekonečnému množství informací, ulehčují práci atd. Současně ale také (krom všech dalších možných negativních důsledků) vedou k úzké specializaci, ke ztrátě vědomostí, znalostí a širšího kontextu, ke ztrátě rozhledu a schopnosti kriticky uvažovat.[2]

Kvůli složitosti a náročnosti na výrobu a provoz navíc zvyšují závislost v širokém smyslu slova – závislost výrobců a poskytovatelů na zdrojích a dodavatelích, závislost uživatelů na poskytovatelích atd.

---

[2] V základu, včetně dále odkazované literatury, srov. zejména CHOMYNOVÁ, P., DVOŘÁKOVÁ, Z., ČERNÍKOVÁ, T., ORLÍKOVÁ, B., GROHMANNOVÁ, K., FRANCOVÁ, E.: *Zpráva o digitálních závislostech v České republice 2023*. Národní monitorovací středisko pro drogy a závislosti. Praha: Úřad vlády České republiky, 2023.

Míru aktuální zranitelnosti odráží fakt, že před 50 lety by několikadenní výpadek elektrického proudu znamenal nepříjemnost. V digitálním věku by znamenal katastrofu.

Předložený příspěvek se v kontextu naznačených výchozích tezí analyticky zaměřuje na základní problematický moment právní úpravy (nejen) automatizovaného rozhodování, včetně profilování, v pracovněprávních vztazích. Krom kritického zhodnocení trendu vývoje rozhodné právní úpravy a související činnosti regulačních orgánů naznačuje i směr pro úvahy *de lege ferenda*.

8.1 Automatizovné rozhodování, včetně profilování v pracovněprávních vztazích

Pracovněprávní vztahy jsou z povahy věci zpravidla jednou z prvních oblastí, v níž se moderní technologie a digitální služby široce zavádí. Děje s tak většinově za účelem zvýšení efektivity, rychlosti a snížení nákladů.

Zaměstnanci jsou tak skupinou, jejichž práva a svobody jsou využíváním moderních technologií a digitálních služeb podstatně dotčeny.

Moderní technologie a digitální služby, resp. automatizované rozhodování (čl. 22 GDPR[3]), v některých případech včetně profilování (čl. 4 bod 4 a čl. 22 GDPR), či aktuálně zapojení umělé inteligence, nachází široké uplatnění od vzniku základního pracovněprávního vztahu až do jeho skončení.

Příkladmo (aniž je hodnocena legálnost postupu zaměstnavatele) lze poukázat na následující využití moderních technologií a digitálních služeb v pracovněprávních vztazích:[4]

- ***a) výběr uchazečů o zaměstnání*** – za pomoci těchto prostředků je kupříkladu (za pomoci automatizovaného rozhodování) ověřováno splnění formálních či jiných předepsaných podmínek účasti ve výběrovém řízení. Za jejich pomoci jsou však také hodnoceni a vybíráni (prostřednictvím automatizovaného rozhodování stavícího na profilování) nejvhodnější uchazeči na pracovní pozici atd.;[5]
- ***b) evidence pracovní doby –*** *snad až na manuální práce mimo pracoviště zaměstnavatele je vedení evidence pracovní doby ve smyslu ust. § 96*

---

3 Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES dále jen „*GDPR*".

4 Srov. MORÁVEK, J.: *Ochrana osobních údajů podle obecného nařízení o ochraně osobních údajů (nejen) se zaměřením na pracovněprávní vztahy*. Praha: Wolters Kluwer, 2019, s. 420.

5 Příkladem budiž například softwarový produkt aptien společnosti Aptien Labs s.r.o. Online: https://aptien.com.

zákona č. 262/2006 Sb., zákoník práce,[6] *prostřednictvím digitálních záznamů dnes pravidlem;*

c) ***docházkové a přístupové systémy*** – vedle evidence pracovní doby (někdy dokonce nesprávně místo ní) bývá využíván elektronický systém evidující příchody a odchody z objektu zaměstnavatele. Vedle Rfid čipů bývá (většinou neoprávněně) k identifikaci příchozího využíváno automatizované zpracování biometrických údajů.[7] Stejně tak tomu bývá i u čistě přístupových systému, které nezaznamenávají příchody a odchody a slouží pouze k omezení přístupu neoprávněných osob do konkrétního objektu;[8]

d) ***hodnocení pracovního výkonu*** **–** řada systémů nabízí možnost automatizovaného hodnocení pracovního výkonu z různých hledisek (rychlost, kvalita, počet chyb atd.). Vedle přehledu o dosavadním pracovním výkonu takové systémy (nikoli výjimečně) nabízí i (na základě profilování) budoucí predikce;

e) ***hodnocení potenciálu*** **–** na hodnocení podle předchozího bodu mohou být navázány další predikce, které vychází jednak z dosavadního a predikovaného budoucího pracovního výkonu zaměstnance, a jednak ze zaměstnanci přisuzovaných vlastností a schopností (schopnost samostatně se rozhodovat, schopnost pracovat v týmu, inciativa, spolehlivost atd.). Tyto aspekty ve vztahu k jednotlivým zaměstnancům zpravidla zadávají vedoucí zaměstnanci, příp. jsou výsledek testování zaměstnanců. Cílem je (prostřednictvím automatizovaného rozhodování vycházejícího z profilování) hodnotit potenciál a perspektivu zaměstnance, jakož i možnost budoucího postupu;[9]

f) ***dodržování pracovněprávních povinností*** **–** nejen na evidenci pracovní doby a hodnocení pracovního výkonu může navazovat automatizované hodnocení dodržování pracovněprávních povinností. Automatizované hodnocení dodržování pracovněprávních povinností

---

6 Dále jen „ZPr".

7 V této souvislosti viz zejména Úřad pro ochranu osobních údajů - Společnost Metrostav a.s. Kontrola použití technologie FaceID. Online https://uoou.gov.cz/cinnost/ochrana-osobnich-udaju/ukoncene-kontroly/kontroly-za-rok-2019/kontrolni-cinnost-v-oblasti-ochrany-osobnich-udaju-2019/spolecnost-metrostav-as.

8 V této souvislosti viz například závěry ÚOOÚ obsažené ve Výroční zprávě Úřadu pro ochranu osobních údajů za rok 2024, srov. 10.

9 Příkladem budiž softwarový produkt lutherone společnosti LutherOne a.s. Online: https://www.lutherone.com.

je možné i v dalších oblastech, jako je například splnění povinnosti účastnit se online školení v oblasti bezpečnosti a ochrany zdraví při práci. Časté jsou však i prostředky kontroly využívání svěřených pracovních prostředků (kontrola přístupu k internetu, sledování prostřednictvím GPS atd.) a podobně;[10]

g) odměňování **–** v reflexi (nejen) výsledků automatizovaného rozhodování (v některých případech založeného na profilování) jsou používány i systémy automatizovaného rozhodování v oblasti odměňování zaměstnanců.

Míra využívání moderních technologií a digitálních služeb v pracovněprávních vztazích roste úměrně jejich stále širší dostupnosti.

Nikoli výjimečně je při zavádění rozhodující výlučně schopnost prostředku a přínos pro zaměstnavatele. Úvahy o nezbytnosti jeho využití, alternativních méně invazivních možnostech a proporcionalitě zůstávají často stranou.

Důvodů, proč tomu tak je, je řada. Jeden se však přímo nabízí.

Je jím přílišná stále větší komplikovanost právní úpravy, jejímž důsledkem je, že adresáti nemají možnost jí sami porozumět a zorientovat se v ní. Potřebují tudíž kvalifikovanou právní pomoc. Kvalifikovaná právní pomoc, která odpovídajícím způsobem zajistí komplexní zpracování materie, je drahá.

Oproti tomu pravděpodobnost kontroly ze strany (primárně) Úřadu pro ochranu osobních údajů je malá, stejně jako pravděpodobnost uložení sankce. Schází tak motivace právní aspekty, v nichž se odráží požadavek na střídmost, nezbytnost a přiměřenost, jež směřují k minimalizaci zásahu do zaměstnaneckých práv, řešit.

Pro ilustraci lze v této souvislosti uzavřít krátkým přehledem činnosti ÚOOÚ, jde-li o počet zahájených kontrol, ukončených kontrol a pravomocných rozhodnutí o uložení pokuty za několik posledních let, jak plyne z výročních zpráv ÚOOÚ. Jedná se o údaje vztahující se pouze k právní úpravě na ochranu osobních údajů. Stranou je ponechána agenda nevyžádaných obchodních sdělení atd.[11]

---

[10] Srov. MORÁVEK, J.: *Ochrana osobních údajů podle obecného nařízení o ochraně osobních údajů (nejen) se zaměřením na pracovněprávní vztahy*. Praha: Wolters Kluwer, 2019, s. 420.

[11] Cituj výroční zprávy Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2009, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2010, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2011, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2012, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2013, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2014, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2015, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2016, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2017,

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Počet zahájených kontrol	143	106	179	129	90	144	135	116	100	76	63	54	52	25	35	14
Počet ukončených kontrol	131	106	144	138	74	128	113	111	110	89	75	48	43	20	25	12
Počet pravomocných rozhodnutí o uložení pokuty	109	151	126	118	119	95	61	64	66	56	32	61	40	30	23	35

8.2 Právní úprava – minulost, přítomnost, budoucnost

Nařízení 2016/679 předcházející směrnice Evropského parlamentu a Rady 95/46/ES ze dne 24. října 1995 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů byla do právního řádu České republiky transponována primárně zákonem č. 101/2000 Sb., o ochraně osobních údajů.[12]

Metodickou podporu a jednotící stanoviska ve vztahu ke směrnici vydávala pracovní skupina WP 29 založená na základě čl. 29 směrnice 95/46/ES. Tato s ohledem na povahu evropské regulace a pracovní skupiny *stricto sensu* nebyla závazná.

Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2018, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2019, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2020, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2021, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2022, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2023, Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2024. Výroční zprávy jsou dostupné zde https://uoou.gov.cz/media-publikace/ke-stazeni/vyrocni-zpravy

12 Dále jen „*OchOúZ*“

Z hlediska aplikační praxe byla tudíž rozhodující spíše stanoviska a vyjádření ÚOOÚ. Těch nebylo mnoho, reagovaly na místní realitu a reálie, a byla většinou stručná.[13]

V oblasti ochrany osobních údajů, včetně pracovněprávních vztahů, náležela pravomoc nad kontrolou, dozorem a správním trestáním *de facto* výlučně ÚOOÚ. Až podstatně později, v roce 2017, byly dílčí kompetence v oblasti ochrany osobních údajů svěřeny i inspektorátu práce, když byl doplněn zákona č. 251/2005 Sb., o inspekci práce, mj. o skutkové podstaty, jejichž objektem je primárně soukromí zaměstnance, navázané na ust. § 316 ZPr, který upravuje mj. využívání pracovních prostředků zaměstnavatele, kontrolu zaměstnanců, sledování zaměstnanců a využívání některých osobních údajů zaměstnanců ze strany zaměstnavatele.

Již toto dílčí překrytí kompetencí přineslo kvůli částečně odlišnému přístupu inspekce práce a ÚOOÚ, stejně jako kvůli nekonzistentnosti jednotlivých inspektorátu práce, určité problémy.

Stěžejním momentem, který situaci podstatně zkomplikoval, bylo však až zavedení GDPR spolu s ustanovením Evropského sboru pro ochranu osobních údajů.

Kombinace podrobnější právní úpravy, byť stavící na shodných principech a základních povinnostech, která přece jen přinášela některé odlišnosti (z hlediska České republiky se jednalo například o odlišnou definici biometrického údaj), jednotící úlohy Evropského sboru pro ochranu osobních údajů prezentované zástupem rozsáhlých vodítek a stanovisek doprovozených doplňujícími a dodatečnými stanovisky a metodikami ÚOOÚ vedlo ke zděšení a později (obrazně řečeno) paralýze na straně valné většina adresátů právní úpravy. O tom, že paralýza nastala nejen na jejich straně, do značné míry vypovídá i shora připojený přehled (ne)aktivity ÚOOÚ.

Uvedené lze ilustrovat právě na právní úpravě automatizovaného zpracování osobních údajů, včetně případného souvisejícího profilování.[14]

Zákon o ochraně osobních údajů se automatizovanému rozhodování věnoval pouze v jednom odstavci. Ust. § 11 odst. 6 OchOúZ konkrétně stanovilo: „*Žádné rozhodnutí správce nebo zpracovatele, jehož důsledkem je zásah do právních a právem chráněných zájmů subjektu údajů, nelze bez ověření vydat nebo učinit výlučně na základě automatizovaného zpracování osobních údajů. To*

---

[13] Srov. MORÁVEK, J.: *Ochrana osobních údajů v pracovněprávních vztazích*. Praha: Wolters Kluwer, 2013, s. 436.

[14] Jádro právní úpravy vztahující se k postavení správce osobních údajů je pro tyto účely ponecháno stranou. Pozornost je věnována jen specifikům navázaným na automatizované rozhodování, včetně profilování.

neplatí v případě, že takové rozhodnutí bylo učiněno ve prospěch subjektu údajů a na jeho žádost.“

Spolu s tím byla k automatizovanému rozhodování, včetně případného profilování, založena stručná informační povinnost ust. § 11 odst. 1 OchOúZ – správce byl povinen informovat o způsobu, jakým bude osobní údaje zpracovávat. Na to navazovalo v ust. § 21 odst. 1 OchOúZ zakotvené právo na vysvětlení ve vztahu ke konkrétnímu zpracování, měl-li subjekt údajů za to, že zpracování jeho osobních údajů je v rozporu s ochranou jeho soukromého a osobního života nebo v rozporu se zákonem. Spolu s tím měl případně právo požadovat výmaz nebo oprava zpracovávaných osobních údajů.

Prosté, stručné, výstižné, funkční.

Se zavedením obecného nařízení o ochraně osobních údajů došlo k rozšíření informační povinnosti (čl. 12, 13 a 14 GDPR), k rozpracování a rozvedení obsahu jednotlivých práv subjektu údajů (právo na přenositelnost, právo na opravu a výmaz atd.), včetně explicitního formulování některých práv dovozovaných pouze judikaturou (právo na přenositelnost, právo na omezení zpracování), a k podobnější úpravě práv a povinností souvisejících s automatizovaným individuálním rozhodováním, včetně profilování (čl. 22 GDPR). K ochraně veřejných subjektivních práv subjektu údajů bylo zavedeno stížností řízení (čl. 77 GDPR) atd.

K tomu byly přijaty ještě WP 29 v tu dobu se přerodivší pozvolna v Evropský sbor pro ochranu osobních údajů pokyny k automatizovanému individuálnímu rozhodování a profilování (WP251 v revidované verzi ze dne 6. února 2018) a pokyny k transparentnosti podle nařízení 2016/679 (WP260 v revidované verzi ze dne 11. dubna 2018).[15] Dále se k problematice přímo vztahují pokyny Evropského sboru pro ochranu osobních údajů č. 5 z roku 2022 k používání technologií rozpoznávání obličeje v oblasti prosazování práva.

Bez významu ale nejsou ani pokyny Evropského sboru pro ochranu osobních údajů týkající se postavení správce a zpracovatele osobních údajů (pokyny č. 7 z roku 2020), pokyny vydané WP 29 pro posouzení vlivu na ochranu údajů a stanovení, zda „je pravděpodobné, že zpracování údajů bude mít za následek vysoké riziko“ pro účely nařízení 2016/679 (WP248 revidované 4. října 2017) anebo stanovisko WP29 č. 2 z roku 2017 ke zpracování údajů na pracovišti.

[15] V návaznosti na přerod WP 29 v Evropský sbor pro ochranu osobních údajů mají pracovní dokumenty vydané koncem funkčního období před přeměnou na Evropský sbor pro ochranu osobních údajů povahu pokynů Evropského sboru pro ochranu osobních údajů

Z hlediska ÚOOÚ se pak jedná například o Metodiku obecného posouzení vlivu na ochranu osobních údajů ze dne 11. listopadu 2020 a o související Seznam druhů operací zpracování (ne)podléhající požadavku na posouzení vlivu na ochranu osobních údajů či o Metodiku k plnění informační povinnosti a k souvisejícím ujednáním vůči zákazníkům.

Tím však ani zdaleka nekončíme.

V roce 2026 přibude právní úprava obsažená v nařízení Evropského parlamentu a Rady (EU) 2024/1689 ze dne 13. června 2024, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění nařízení (ES) č. 300/2008, (EU) č. 167/2013, (EU) č. 168/2013, (EU) 2018/858, (EU) 2018/1139 a (EU) 2019/2144 a směrnice 2014/90/EU, (EU) 2016/797 a (EU) 2020/1828, které se dotýká mj. uvádění na trh nebo používání systému AI s cílem odvodit emoce fyzické osoby na pracovišti (s výjimkou z lékařských nebo bezpečnostních důvodů), a uvádění na trh nebo používání systémů biometrické kategorizace za účelem dovození nebo odvození rasy, politických názorů, členství v odborových organizacích, náboženského nebo filozofického přesvědčení, sexuálního života nebo sexuální orientace.

Vedle toho v roce 2026 končí transpoziční lhůta směrnice Evropského parlamentu a Rady (EU) 2024/2831 ze dne 23. října 2024 o zlepšení pracovních podmínek při práci prostřednictvím platformy, která se dotká mj. automatizovaných monitorovacích a rozhodovacích systémů, rozvádí související informační povinnost a předkládá vyjádření se dotčených subjektů k nastavení systému zpracování osobních údajů, stejně tak jako lidský přezkum automatizovaných rozhodnutí a lidský dohled.

V jádru není na předmětné regulaci nic špatného až na její (zejména v případě GDPR) širokou (až univerzální) působnost.

8.3 Problematické momenty, závěrečné poznámky a úvahy

Z dosavadního výkladu s podává, že možnost postupu *lege artis* si na straně adresáta právní úpravy žádá její podrobné studium, včetně úvodních recitálů evropských předpisů, stejně jako studium stovek stran materiálů stanovisek, doporučení, pokynů a vodítek. To ovšem nelze u průměrného uživatele legitimně očekávat.

A i kdyby to tak nebylo, řešilo by to jen polovinu problému.

Aby bylo možné právní pravidla uplatnit na určitý jev, který mají regulovat, a postupovat podle nich, je třeba mít představu o tom, jak regulovaný jev funguje. Ani to však u průměrného adresáta právní úpravy v postavení správce osobních údajů, jde-li o moderních technologie a digitální služby, jako jsou například služby zmíněné v bodě 2 shora, nelze legitimně očekávat.

Nejde však jen o něj.

Poznání a popsání regulovaného jevu, stejně jako jeho negativních projevů, které mají být právní úpravou řešeny, je nezbytné z povahy věci již při formulaci právní regulace. Stejné je podmínkou i pro efektivní uplatňování (kontrolu a vymáhání) právních pravidel ze strany veřejné moci.

To si žádá odpovídající kompetence, kapacity a odbornost.

Toho se však v potřebné míře a rozsahu s ohledem na rychlost pokroku a složitost moderních technologií a digitálních služeb na straně aparátu připravujícího právní úpravu stejně jako na straně dozorových úřadů nedostává a (téměř jistě) do budoucna ani dostávat nebude, zejména vyjdeme-li z předkladu široké veřejnoprávní kontroly. Bez významu v této souvislosti není kvůli globálnímu fungování poskytovatelů služeb ani problematické fungování přeshraniční spolupráce mezi dozorovými úřady a limity plynoucí z teritoriality působení právního řádu a veřejné moci.

Konstrukce, jíž se snaží veřejná moc tomuto (možná vědomému) nedostatku, jenž je mj. jednou z příčin vedoucích ke snižování počtu kontrol, čelit, se podává z přehledu právní úpravy předloženého v předchozím bodě. Je jím široké vymezení okruhu informací, které jsou osoby v postavení správců osobních údajů využívajících moderní technologie a digitální služby povinny subjektům údajů poskytovat s tím, že „informovaný" člověk se následně rozumně a zodpovědně rozhodne a zachová.

To je ovšem alibistický a neefektivní přístup, jde-li o ochranu práv a právem chráněných zájmů subjektů údaj.

Důvody jsou prosté.

Již dávno došlo k inflaci infomační povinnosti. Lidé jsou jako spotřebitelé, zaměstnanci atd. zahrnuti nepřeberným množstvím informací, které jim mají pomoci při rozhodování. Ve skutečnosti však široké zákonné informační povinnosti mají pro ochranu práv a právem chráněných zájmů zanedbatelný až téměř nijaký efekt, neboť adresáti jednak nemají ambici se s informacemi seznamovat již jen kvůli jejich množství, a jednak, i kdyby se s nimi seznámit chtěli, většinovou jsou formulovány tak, aby odpovídaly zákonné úpravě z perspektivy dozorových úřadů, tedy (i přes opakované a výslovné požadavky na užití jednoduchých jazykových prostředků tak, aby byly srozumitelné osobě průměrného rozumu) nikoli tak, aby jim běžný adresát mohl skutečně porozumět. Jsou pro ně tudíž nesrozumitelná.

V konečném důsledku tak dochází k vynakládání prostředků na přípravu informačních materiálů bez většího reálného užitku pro dotčené osoby (subjekty údajů).

Je tu tak vlastně dvojí negativní důsledek. Jednak (často možná) nedůvodné a nepřiměřené zásahy do práv a svobod (neinformovaných) subjektů údajů (zaměstnanců, spotřebitelů atd.) a jednak nedůvodné a nepřiměřené zásahy do majetkových práv správců osobních údajů, neboť vynakládání prostředků k plnění povinností plynoucích z neefektivní právní úpravy (ve skutečnosti nechrání společenské hodnoty) z povahy věci z hlediska principu proporcionality nemůže obstát.[16]

Další rozšiřování informačních povinností, zavádění povinností dotázat se subjektů údajů na názor atd. tak z uvedených důvodů situaci nijak neprospěje.

Je třeba reflektovat, že ve skutečnosti zde jde o nerovný souboj mezi technologickými giganty na jedné straně a regulátory na straně druhé, který regulátoři při stávajícím přístupu nemohou vyhrát. Jejich prohra však znamená i prohru dotčených subjektů údajů a do určité míry i uživatelů nových technologií a digitálních služeb, vezmeme-li v potaz neefektivní vynakládání prostředků na plnění fakticky bezvýznamných povinností, které se v řadě případů jejich soutěžitele mimo působnost práva EU nezatěžuje.

Na možné řešení, které se čím dále více jeví jako jediná rozumná cesta, bylo poukazováno již v souvislosti se zaváděním GDPR. [17]

Je jím soustředění se na hlavní poskytovatele prostředků moderní techniky a digitálních služeb ve smyslu jejich primární odpovědnosti za poskytované prostředky a služby, včetně například jejich předchozí certifikace před uvedením na trh.

Vedle toho, pokud má být zachována nějaká suma povinností i vůči běžným uživatelům (správcům osobních údajů), jde primárně o zásadní zjednodušení právního rámce a o efektivní metodickou podporu ze strany dozorových úřadů – vytvořený stručných, srozumitelných a přímo použitelných šablon a vzorů. Tou však není kupříkladu ÚOOÚ formulovaná metodika ke kamerovým systém,[18] jíž se ÚOOÚ na 62 stranách mj. pokouší správci osobních údajů vysvětlit, jak provést test proporcionality při sprejování na fasádu domu tím, že určuje náklady na opravu poničené fasády přes počet poškození v roce a cenu za m2 nátěru fasády, včetně impregnace a ceny mytí.

---

16 Srov. MORÁVEK, J.: *Model práva – Vztah práva a morálky*. Praha: Lidne, 2013, s. 103.

17 V této souvislosti viz zejména MORÁVEK, J.: *Ochrana osobních údajů podle obecného nařízení o ochraně osobních údajů (nejen) se zaměřením na pracovněprávní vztahy*. Praha: Wolters Kluwer, 2019, s. 420.

18 Viz HAVELDA, Z., BURIAN, D., POŘÍZKA, A.: *Metodika k návrhu a provozování kamerových systémů z hlediska zpracování a ochrany osobních údajů*. Praha: Úřad pro ochranu osobních údajů. 2024. Metodika je dostupná zde https://uoou.gov.cz/profesional/metodiky-a-doporuceni-pro-spravce/metodika-ke-kamerovym-systemum.

Zoznam použitej literatúry a zdrojov

1. BÁRTA, J.: *Sedm zákonů. Jak se civilizace rodí, rostou a upadají*. Praha: Jota, 2021.
2. BLACK, J.: *Historie světa*. Praha: Universum, 2024.
3. HAVELDA, Z., BURIAN, D., POŘÍZKA, A.: *Metodika k návrhu a provozování kamerových systémů z hlediska zpracování a ochrany osobních údajů*. Praha: Úřad pro ochranu osobních údajů, 2024.
4. CHOMYNOVÁ, P., DVOŘÁKOVÁ, Z., ČERNÍKOVÁ, T., ORLÍKOVÁ, B., GROHMANNOVÁ, K., FRANCOVÁ, E.: *Zpráva o digitálních závislostech v České republice 2023*. Národní monitorovací středisko pro drogy a závislosti. Praha: Úřad vlády České republiky, 2023.
5. MORÁVEK, J.: *Model práva – Vztah práva a morálky*. Praha: Linde Praha. 2013, s. 103.
6. MORÁVEK, J.: *Ochrana osobních údajů podle obecného nařízení o ochraně osobních údajů (nejen) se zaměřením na pracovněprávní vztahy*. Praha: Wolters Kluwer, 2019, s. 420.
7. MORÁVEK, J.: *Ochrana osobních údajů v pracovněprávních vztazích*. Praha: Wolters Kluwer, 2013, s. 436.
8. Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (Text s významem pro EHP), L 119/1, 4.5.2016.
9. Nařízení Evropského parlamentu a Rady (EU) 2024/1689 ze dne 13. června 2024, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění nařízení (ES) č. 300/2008, (EU) č. 167/2013, (EU) č. 168/2013, (EU) 2018/858, (EU) 2018/1139 a (EU) 2019/2144 a směrnice 2014/90/EU, (EU) 2016/797 a (EU) 2020/1828 (akt o umělé inteligenci) (Text s významem pro EHP), L, 12.7.2024.
10. Směrnice Evropského parlamentu a Rady 95/46/ES ze dne 24. října 1995 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů, L 281, 23.11.1995.
11. Směrnice Evropského parlamentu a Rady (EU) 2024/2831 ze dne 23. října 2024 o zlepšení pracovních podmínek při práci prostřednictvím platformy (Text s významem pro EHP), L, 11.11.2024.
12. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2009
13. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2010
14. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2011

15. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2012
16. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2013
17. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2014
18. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2015
19. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2016
20. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2017
21. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2018
22. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2019
23. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2020
24. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2021
25. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2022
26. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2023
27. Výroční zpráva Úřadu pro ochranu osobních údajů za rok 2024
28. Zákon č. 101/2000 Sb., o ochraně osobních údajů
29. Zákon č. 251/2005 Sb., o inspekci práce
30. Zákon č. 262/2006 Sb., zákoník práce

Informácie o autorovi

doc. JUDr. Jakub MORÁVEK, Ph.D.
Západočeská univerzita v Plzni
Fakulta právnická
Katedra pracovního práva a práva sociálního zabezpečení
Email: moravek@prf.cuni.cz

9 Asistencia umelej inteligencie pri excesívnej aktivite verejnej správy so zásahom do ľudských práv

Tibor Seman

https://doi.org/10.33542/HRF-0541-2-09

Abstrakt

Využitie umelej inteligencie (AI) vo verejnej správe otvára možnosti efektívnejšieho a rýchlejšieho administratívneho konania a rozhodovania, avšak prináša so sebou aj riziko vzniku nových foriem zásahov do základných práv a slobôd jednotlivcov. Vo verejnej správe, kde je nerovnováha medzi jednotlivcom a štátnou autoritou osobitne výrazná, môžu byť dôsledky zneužitia alebo neprimeraného použitia AI závažné. Príspevok zdôrazňuje riziká zapojenia AI pri zásahoch do práva na súkromie, zakotveného v medzinárodnoprávnych dokumentoch (najmä Európsky dohovor o ľudských právach, Charta základných práv EÚ, Všeobecná deklarácia ľudských práv) a v slovenskom právnom poriadku. Cieľom príspevku je poukázať na potrebu právnej úpravy základných zásad používania umelej inteligencie vo verejnej správe sledujúc elimináciu nežiadúcich dôsledkov excesívnej aktivity verejnej správy s asistenciou umelej inteligencie. S poukázaním na vybrané rozhodnutia Európskeho súdu pre ľudské práva autor formuluje názory na zákonnosť využitia umelej inteligencie v slovenskom práve de lege lata. Uvádza tiež námety de lege ferenda, ktoré by mohli pomôcť predísť porušovaniu ľudských práv v dôsledku automatizácie verejnej správy.

Úvod

Umelá inteligencia sa postupne dostáva aj do oblasti verejnej správy. V súčasnosti predstavuje vítaného nezáväzného pomocníka pre úradníkov, ktorí si nemusia vedieť dať rady v zložitejších administratívnych konaniach. Samozrejme, v administratívnom rozhodnutí úradníci zatiaľ nekonkretizujú, či a ktoré usmernenia dostali od umelej inteligencie. Zrejme sa však blíži čas, keď bude nevyhnutné využívať tento výdobytok modernej vedy vo verejnej správe, keďže umelá inteligencia je vyhľadávanou technológiou zásadne v každej sfére ľudskej činnosti, napríklad v oblastiach zdravotníctva, vedy, techniky, dopravy, vzdelávania či umenia. Je teda nevyhnutné perspektívne uvažovať o postupnom zapojení umelej inteligencie do činností verejnej správy, vrátane využitia v procesoch vo verejnej správe, konkrétne aj v rozhodovacích procesoch. Na mysli však máme zákonné zapojenie a využívanie umelej inteligencie, nie jej „pokútne," hoci užitočné využívanie. Tým, že umelá inteligencia umožňuje spracovanie veľkého množstva dát, v spojení s jej samostatnou aktivitou a rýchlym plnením

požiadaviek zadaných človekom, sa javí ako nástroj aj na efektívnejšie a rýchlejšie administratívne rozhodovanie v budúcnosti možno už nie tak vzdialenej. Napriek tomu, činnosť verejnej správy, vrátane automatizácie rozhodovacích procesov, je potenciálnym zdrojom rizík, ktoré sa v doterajšej činnosti verejnej správy nevyskytovali. Rizikom môžu byť aj nežiadúce zásahy do základných práv a slobôd jednotlivcov, pričom vo verejnej správe, kde existuje prirodzená nerovnováha medzi štátnou autoritou a občanom, môže nesprávne, neustrážené, prípadne až zneužité zapojenie umelej inteligencie mať závažné dôsledky.

Umelá inteligencia je spôsobilá zmeniť spôsob komunikácie verejnej správy s fyzickými osobami a právnický osobami ako adresátmi administratívneho spravovania. Určite prinesie byrokratické odbremenenie, výrazné zrýchlenie priebehu administratívneho konania a v neposlednom rade sa od nej očakáva menej chýb ako u človeka. No práve tu nastáva otázka, či umelá inteligencia vskutku bude predstavovať výhradne pozitívny prvok v rozhodovacom procese alebo prinesie vlastné chyby a nedostatky, ktoré budú musieť byť riešené.

Využitie umelej inteligencie vo verejnej správe určite treba považovať za užitočný prvok vyvíjajúcej sa vedy a techniky. Nežiadúcimi však budú také aktivity verejnej správy, pri ktorých sa umelá inteligencia využije spôsobom odporujúcim ľudským právam. Teda, nielen nežiadúca, ale aj neprípustná je a bude práve excesívna aktivita verejnej správy ako činnosť presahujúca rámec normálu, v ktorej sa, hoci za účelom dôsledného plnenia verejnoprávnych úloh, zneužijú schopnosti umelej inteligencie.

Cieľom príspevku je poukázať na potrebu právnej úpravy základných zásad používania umelej inteligencie vo verejnej správe s vplyvom na ľudské práva a slobody, sledujúc elimináciu nežiadúcich dôsledkov excesívnej aktivity verejnej správy s asistenciou umelej inteligencie, čo má svoj odraz v právnej istote a celkovom nazeraní na dôveryhodnosť výkonu činnosti verejnej správy. Okrem toho príspevok hlbšie skúma riziko ohrozenia práva na ochranu osobných údajov v súvislosti s právom na ochranu súkromia, pričom s poukázaním na vybrané rozhodnutia Európskeho súdu pre ľudské práva sú formulované názory na zákonnosť využitia umelej inteligencie v slovenskom práve de lege lata a predostreté námety de lege ferenda, ktoré by mohli pomôcť predísť porušovaniu ľudských práv v dôsledku automatizácie verejnej správy.

Pri písaní príspevku boli použité viaceré vedecké metódy výskumu, a to analýza súčasného stavu relevantnej právnej úpravy so zvýraznením rizík pri zapojení umelej inteligencie do aktivít verejnej správy, takisto interpretačná a explanačná metóda pri podávaní výkladu tých zákonných ustanovení, ktoré potenciálne umožňujú, respektíve nevylučujú parciálne využitie umelej

inteligencie. Ďalšími metódami výskumu boli deduktívna a induktívna metóda pri pomenúvaní rizík, spojených s využitím umelej inteligencie.[1]

9.1 Význam umelej inteligencie vo verejnej správe

9.1.1 Podstata umelej inteligencie

Umelá inteligencia vonkoncom nie je nový fenomén. Sci fi spred desiatok rokov sa stáva realitou a umelá inteligencia preniká do mnohých sfér verejného i súkromného života. Ako najzrozumiteľnejšie a zároveň stručne charakterizovať túto entitu? Dokáže plniť úlohy zadané človekom, rozpoznáva obrazy, rozumie textom, vytvára výstupy podľa požiadaviek človeka, dokonca sa vie aj učiť. Svojimi postupmi sa správa a prejavuje ako inteligentná bytosť, teda ako človek. Nie však kompletne, nemá city, emócie, dokáže ich však fabulovať. Umelá inteligencia realizovanými technikami vlastne napodobňuje ľudskú inteligenciu. Umelá inteligencia je *„odbor informatiky zaoberajúci sa tvorbou systémov riešiacich komplexné úlohy ako je rozpoznávanie či klasifikácia, napr. v oblastiach spracovania obrazu alebo spracovania písaného textu či hovoreného jazyka, alebo plánovania či riadenia na základe spracovania veľkých objemov dát."* [2] *„Ide o inteligentné systémy, ktoré sú navrhnuté tak, aby dokázali samostatne rozmýšľať, učiť sa a robiť rozhodnutia. Tieto systémy sú založené na algoritmoch a strojovom učení a môžu byť použité na rôzne úlohy."*[3] Charakteristickou črtou umelej inteligencie je *„schopnosť zariadenia prejaviť schopnosti podobné človeku, ako sú uvažovanie, učenie, plánovanie a tvorivosť."*[4] Za najjednoduchšie, snáď i najvýstižnejšie by bolo vysvetliť podstatu umelej inteligencie slovami „mysliaci stroj".

---

1 Príspevok vznikol v rámci riešenia grantového projektu VEGA č. 1/0062/25 s názvom Automatizácia rozhodovacích procesov vo verejnej správe, v ktorom je autor členom riešiteľného kolektívu.

2 KAPLAN, A: *Artificial Intelligence, Business and Civilization: Our Fate Made in Machines*. London: Routledge, 2022, s. 1. Online: https://www.taylorfrancis.com/books/mono/10.4324/9781003244554/artificial-intelligence-business-civilization-andreas-kaplan; DOI: doi.org/10.4324/9781003244554.

3 SLOVÁKOVÁ, A. I.: *Čo je to umelá inteligencia? 1 000 slov o nej od nej*, 21.4.2023. Online: https://www.techbox.sk/co-je-to-umela-inteligencia-1-000-slov-o-nej-od-nej.

4 Európsky parlament: *Umelá inteligencia: definícia a využitie*, 21.06.2023. Online: https://www.europarl.europa.eu/topics/sk/article/20200827STO85804/umela-inteligencia-definicia-a-vyuzitie.

9.1.2 Využitie umelej inteligencie vo verejnej správe

Nové technológie, ku ktorým patrí umelá inteligencia, sa dajú využiť aj v oblasti verejného spravovania. Zrozumiteľným príkladom je pokus o telefonický kontakt s orgánom verejnej správy. Je možnosť, že bude obsadené alebo dlho nikto nezdvihne, a ak obsluhuje telefón iba jeden úradník, ktorý okrem toho plní aj iné úlohy, napríklad pripravuje rozhodnutia, jeho zaťaženosť je pochopiteľná. Pokiaľ by telefonické hovory vybavovala na úradoch umelá inteligencia, vedela by paralelne riešiť desiatky či stovky telefonátov, samozrejme, v prípade potreby by prepojila volajúceho na konkrétneho zamestnanca verejnej správy. Predstava o potenciálnom využití AI sa najviac spája s administratívnym konaním, teda procesom, smerujúcim k vydaniu, prípadne aj preskúmaniu a nútenému výkonu rozhodnutí, ktorými sa priznávajú práva, ukladajú povinnosti alebo sa zasahuje do právom chránených záujmov fyzických osôb a právnických osôb. V aktuálnom čase je jednoznačne možné reálne uvažovať o využití umelej inteligencie v niektorých fázach správneho konania. Táto technológia môže prijímať, obsahovo posúdiť a triediť podania. Už v tejto fáze môžu byť dotknuté práva účastníka konania, pokiaľ by umelá inteligencia napríklad ignorovala jeho podanie. Pozitívnym prínosom umelej inteligencie by bola schopnosť odhaliť nedostatky podania, chýbajúc prílohy či iné náležitosti podania, so súčasným vyhotovením a expedovaním výzvy na odstránenie nedostatkov konania. Úvodné fázy správneho konania by využitím umelej inteligencie odbremenili úradníkov a prispeli by k zrýchleniu rozhodovacieho procesu. Ak by podanie nemalo chyby, umelá inteligencia by toto podanie priradila na vybavenie konkrétnemu správnemu orgánu, prípadne aj konkrétnemu zamestnancovi - úradníkovi. Týmto by sa participácia umelej inteligencie nemusela končiť, pretože v správnom konaní by umelá inteligencia mohla viesť výsluchy pri ústnom pojednávaní, i keď v prítomnosti úradníka, takisto by mohla porovnať a vyhodnotiť dôkazy, pripraviť výrokovú časť ako aj odôvodnenie administratívneho rozhodnutia. Týmito úkonmi by umelá inteligencia vykonávala činnosti, ktoré by inak realizoval človek.[5] v skutočnosti tento „mysliaci stroj“ bude pomocníkom človeka, ktorému zjednoduší viacero úkonov v administratívnom konaní. Tým správne konanie môže mať rýchlejší priebeh a vyššiu efektivitu, keďže mnoho činností človeka môže byť nahradených automatizovanými úkonmi, realizovanými umelou inteligenciou.

---

[5] Viac k využitiu automatizovaných systémov ako náhrady ľudskej práce pozri EDWARDS, L., VEALE, M.: *Slave to the Algorithm? Why a “Right to Explanation” Is Probably Not the Remedy You Are Looking for*. In: Duke Law & Technology Review, roč. 16, č. 1, 2017, str. 82. Online: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1315&context=dltr.

Administratívne konanie však nie je zďaleka jedinou oblasťou verejnej správy s možným využitím umelej inteligencie. Potenciálnym oblastiam využitia umelej inteligencia vo verejnej správe, s poukázaním na súvisiace riziká, sa budeme venovať v nasledujúcich častiach príspevku.

9.2 Hranica medzi využitím a zneužitím umelej inteligencie alias kde sú ľudské práva

V tejto časti príspevku poukážeme na možnosti využitia umelej inteligencie pri práci s databázami, v kontexte s rozhodnutiami Európskeho súdu pre ľudské práva (ESĽP), súvisiacimi s tvorbou databáz a uchovávaním údajov v nich. Budeme sa venovať konkrétnym prípadom zásahov do ľudských práv, ktoré súviseli **s** používaním moderných technológií spracúvania a uchovávania biometrických údajov. Databázy zhromažďujúce osobné údaje, v spojení s možnosťami umelej inteligencie totiž vytvárajú podmienky na zásahy do ľudských práv. Pôjde o prípady *S. a Marper proti Spojenému kráľovstvu* [GC], 2008, *Gaughran proti Spojenému kráľovstvu*, 2020, a *P.N. proti Nemecku*, 2020. Každý z nasledujúcich prípadov preukazuje, že samotná tvorba databáz a uchovávanie údajov v nich môže predstavovať zásah do ľudských práv. Využitím umelej inteligencie sa tieto zásahy môžu ešte znásobiť. Tieto rozhodnutia ESĽP ukazujú, že technológie umožňujúce masové zhromažďovanie, uchovávanie alebo analýzu osobných a biometrických údajov môžu predstavovať závažný zásah do práva na súkromie. Umelá inteligencia iba plní úlohy určené človekom, preto za takéto zásahy spôsobené štátnym orgánom bude zodpovedať štát bez ohľadu na to, aké technické nástroje využíva.

V prípadoch *S. a Marper* a *Gaughran* šlo najmä o neopodstatnené a neprimerané uchovávanie DNA profilov, odtlačkov prstov a fotografií v policajných databázach, čo ESĽP vyhodnotil ako porušenie článku 8 Dohovoru o ochrane ľudských práv a základných slobôd (ďalej aj „Dohovor“). V prípade *P.N. proti Nemecku* ESĽP posudzoval otázku použitia biometrických údajov v kontexte vyšetrovania a ich vplyv na súkromie jednotlivca.

Tieto rozhodnutia potvrdzujú, že zodpovednosť za zásah do práv a slobôd zostáva na strane verejnej moci, ktorá technológie používa. Samotná technológia, či už ide o databázy alebo algoritmy umelej inteligencie, plní iba úlohy, ktoré jej určí človek. Právne relevantné sú teda rozhodnutia ľudí a inštitúcií, nie autonómnosť technológie. S poukázaním na predmetné rozhodnutia mienime zvýrazniť zvýšené riziko zásahov do základných práv, ak sa k porušovaniu základných práv (najmä pri existencii databáz biometrických údajov) pridajú schopnosti umelej inteligencie.

Po predstavení predmetných rozhodnutí nadviažeme na problematiku databázy odtlačkov prstov v Slovenskej republike ako aj na technológiu rozpoznávania tváre, u ktorých je potrebná precizácia pravidiel.

9.2.1 Prípad S. a Marper proti Spojenému kráľovstvu [GC], 2008

V predkladanom prípade, označenom v rozhodnutí Európskeho súdu pre ľudské práva ako kauza *„S. and Marper v. United Kingdom (Grand Chamber, 4. december 2008; nos. 30562/04, 30566/04)"* [6] bolo podstatou veci, že dvaja sťažovatelia, konkrétne mladík označený ako "S", a Michael Marper, mali odobraté odtlačky prstov a vzorky DNA políciou po ich zatknutí, ale ich prípady neboli uzavreté odsúdením. Mladík „S" bol prepustený bez trestu a obvinenia proti Marperovi boli stiahnuté. Po ukončení konaní obaja požiadali, aby boli ich biometrické údaje zlikvidované. Teda, aby boli odtlačky prstov odstránené a DNA vzorky zničené. Polícia to odmietla.

Podstatou právneho problému bolo, či trvalé, respektíve neobmedzené uchovávanie biometrických údajov (DNA profily, odtlačky prstov) osôb, ktoré neboli odsúdené, je kompatibilné s právom na súkromný život podľa čl. 8 Dohovoru.

Európsky súd pre ľudské práva jednomyseľne rozhodol, že takéto uchovávanie predstavuje porušenie čl. 8 Dohovoru.[7] Zákony Spojeného kráľovstva v danom čase (napr. Police and Criminal Evidence Act 1984 s dodatkami) umožňovali uchovať biometrické údaje bez ohľadu na vážnosť obvinení, bez časového obmedzenia a s len veľmi obmedzenými možnosťami vymazania. To bolo považované za neprimeraný zásah do súkromného života.

---

[6] Rozhodnutie Európskeho súdu pre ľudské práva v tejto kauze je zverejnené na internete. ECtHR: *S. and Marper v. United Kingdom*, Application nos. 30562/04, 30566/04, Judgment of 4 December 2008. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-90051"]}.

[7] Článok 8 Dohovoru o ochrane ľudských práv a základných slobôd (publikovaného pod č. 209/1992 Zb.) vo dvoch odsekoch zakotvuje právo na rešpektovanie súkromného a rodinného života. Podľa odseku 1: *„Každý má právo na rešpektovanie svojho súkromného a rodinného života, obydlia a korešpondencie."* a podľa odseku 2: *„Štátny orgán nemôže do výkonu tohto práva zasahovať s výnimkou prípadov, keď je to v súlade so zákonom a nevyhnutné v demokratickej spoločnosti v záujme národnej bezpečnosti, verejnej bezpečnosti, hospodárskeho blahobytu krajiny, predchádzania nepokojom alebo zločinnosti, ochrany zdravia alebo morálky alebo na ochranu práv a slobôd iných."* Európsky dohovor o ľudských právach, v znení protokolov č. 11, 14 a 15, s dodatkovým protokolom a protokolmi č. 4, 6, 7, 12, 13 a 16 je zverejnený na internete. Online: https://www.echr.coe.int/documents/d/echr/convention_slk.

Z uvedeného rozhodnutia ESĽP vyplýva, že blanketné právne predpisy, ktoré umožňujú uchovávanie citlivých osobných údajov aj tých osôb, ktoré neboli odsúdené, sú nekompatibilné s čl. 8 Dohovoru, ak chýbajú primerané záruky. Týmto zárukám patria časové limity uchovávania údajov, možnosť kontroly uchovávaných údajov ako aj ďalšie kritériá súvisiace s typom alebo závažnosťou prípadu.

Poznamenávame, že snímanie odtlačkov posledného článku ukazovákov na oboch rukách je v Slovenskej republike povinnosťou súvisiacou s vydávaním občianskeho preukazu podľa zákona o občianskych preukazoch.[8] Ustanovenie § 2 ods. 4 zákona o občianskych preukazoch znie: *„Občiansky preukaz sa občanovi, ktorý dovŕšil vek 15 rokov, vydáva s platnosťou na 10 rokov; občanovi mladšiemu ako 15 rokov sa vydáva s platnosťou na 5 rokov a občanovi mladšiemu ako 6 rokov sa vydáva s platnosťou na 2 roky. Občiansky preukaz sa občanovi, ktorému nie je dočasne možné zosnímať odtlačky prstov, vydáva s platnosťou na jeden rok.“* V súvislosti so žiadosťou o vydanie občianskeho preukazu sú spojené dve povinnosti, a to podrobiť sa snímaniu odtlačku prstov a podpísať sa. Podľa § 6 ods. 11 zákona o občianskych preukazoch *„Občan, ktorému má byť občiansky preukaz vydaný, je povinný podrobiť sa nasnímaniu podoby tváre, odtlačkov prstov a podpisu. Povinnosť podrobiť sa nasnímaniu odtlačkov prstov sa nevzťahuje na občana uvedeného v § 2 ods. 4 druhej vete, občana, u ktorého nie je trvalo možné odtlačkov prstov zosnímať a na občana mladšieho ako 12 rokov. Povinnosť podrobiť sa nasnímaniu podpisu sa nevzťahuje na občana mladšieho ako 12 rokov; osoba staršia ako 6 rokov a mladšia ako 12 rokov sa môže podrobiť nasnímaniu podpisu.“* Pokiaľ ide o samotné podanie žiadosti o vydanie občianskeho preukazu, v § 7 ods. 8 sa pamätá aj na občana, ktorý pre bezvládnosť z dôvodu dlhodobého pripútania na lôžko alebo psychického ochorenia nemôže osobne podať žiadosť o vydanie občianskeho preukazu. Zaňho môže žiadosť o vydanie občianskeho preukazu podať iná osoba. V takom prípade sa k žiadosti o vydanie občianskeho preukazu prikladá aj lekárske potvrdenie o tom, že takýto občan nemôže podať žiadosť osobne. Príslušný útvar Policajného zboru zabezpečí nasnímanie podoby tváre, odtlačkov prstov a podpisu na území Slovenskej republiky po dohode s občanom, ktorému má byť občiansky preukaz vydaný, alebo osobou, ktorá podala žiadosť.

Dôležitosť získania predmetných údajov vyplýva aj z § 7 ods. 10 zákona o občianskych preukazoch, podľa ktorého sa občiansky preukaz vyhotoví najneskôr do 30 dní od podania žiadosti o vydanie občianskeho preukazu alebo

[8] Ide o zákon č. 395/2019 Z.z. o občianskych preukazoch a o zmene a doplnení niektorých zákonov, v znení neskorších predpisov.

na požiadanie občana urýchlene do 2 pracovných dní, pričom lehota neplynie aj vtedy, ak nie je nasnímaná podoba tváre, odtlačky prstov alebo podpis. Pre úplnosť dodávame, že podľa § 3 ods. 1 písm. i) zákona o občianskych preukazoch sa podpis nevyžaduje, ak občan nie je spôsobilý podpísať sa. A podľa § 3 ods. 5 tohto zákona podpisom podľa odseku 1 písm. i) sa rozumie vlastnou rukou vykonané písomné vyjadrenie priezviska alebo mena a priezviska v ľubovoľnom poradí.

K samotnému podpisu na občianskom preukaze ešte uvádzame, že tento podpis je uchovaný v elektronickej podobe v informačnom systéme. Azda nie je potrebné osobitne zdôrazniť, že podpis občana na občianskom preukaze zvyšuje zneužiteľnosť tohto osobného dokladu napríklad po jeho odcudzení, keď na listine o právnom úkone páchateľ, vydávajúci sa za občana na občianskom preukaze, použije podpis nacvičený podľa vyobrazenia na občianskom preukaze. Podoba tváre podľa okolností prípadu môže, ale aj nemusí byť rozhodujúca.

Z uvedeného vyplýva, že Slovenská republika vedie databázu odtlačkov prstov jej občanov, ktorí sú povinní podrobiť sa snímaniu odtlačkov prstov. V kontexte s vyššie citovaným rozhodnutím ESĽP je možné hodnotiť takúto databázu ako odporujúcu článku 8 Dohovoru. Aj keď chápeme, že odtlačky prstov sa môžu využiť napríklad pri identifikácii mŕtvoly s neznámou totožnosťou (pokiaľ rozklad tkanív nie je pokročilý) alebo pri zisťovaní páchateľa trestného činu, ktorý zanechal daktyloskopické stopy, apriórne vedenie databázy odtlačkov prstov, hoci na základe zákona o občianskych preukazoch, v kontexte s Dohovorom, je prinajmenšom sporné.

Technológia umelej inteligencie zameraná na rozpoznávanie tváre je spôsobilá stotožniť zobrazenú osobu. V prípade sankčných administratívnych konaní síce môže ísť o prvok urýchľujúci stotožnenie osoby, proti ktorej sa bude viesť sankčné administratívne konanie, avšak z hľadiska proporcionality chránených záujmov nie je prevaha prínosu jednoznačná.

Takisto, aj pri daktyloskopických stopách s nedostatočným množstvom markantov potrebných na stotožnenie osoby, ktorá stopy zanechala, umelá inteligencia dokáže definovať množinu všetkých osôb, prichádzajúcich do úvahy ako páchatelia. Tým sú vytvorené podmienky pre permanentný monitoring fyzických osôb, bez ohľadu na súhlasný prejav vôle s takýmto monitoringom.

Snímanie daktyloskopických odtlačkov *en bloc* všetkým občanom Slovenskej republiky v súvislosti s vydávaním občianskeho preukazu, samozrejme okrem výnimiek (amputácie, vrodené deformity a pod.), a existencia databázy týchto odtlačkov má povahu excesívnej aktivity verejnej správy, ktorá hoci na základe zákona, avšak v porovnaní s rozhodovacou činnosťou ESĽP, môže dosahovať rozpor s článkom 8 Dohovoru. Ako sme uviedli,

táto excesívna aktivita môže s asistenciou umelej inteligencie predstavovať zásah do ľudských práv.

9.2.2 Prípad Gaughran proti Spojenému kráľovstvu, 2020

Predkladané rozhodnutie Európskeho súdu pre ľudské práva sa týka prípadu, označenému ako *„Gaughran v. United Kingdom (13. február 2020; no. 45245/15)"* a v ňom ESĽP opätovne potvrdil, že uchovávanie biometrických údajov aj po právoplatnom odsúdení predstavuje zásah do práva na rešpektovanie súkromného života. Ako zdôvodníme ďalej, súkromie zahŕňa aj kontrolu nad osobnými a biologickými informáciami jednotlivca.[9]

V predmetnom prípade šlo o to, že pán Fergus Gaughran bol zadržaný v Severnom Írsku, pretože šoféroval opitý, t.j. pod vplyvom alkoholu. Polícia zhotovila fotografiu tváre, odobrala mu odtlačky prstov a neintímnu vzorku DNA. Po uplynutí zákonnej doby od potrestania polícia mienila údaje uchovávať neobmedzene. Pán Gaughran sa domáhal vymazania týchto údajov. Domáce súdy jeho návrh zamietli, tak sa obrátil na Európsky súd pre ľudské práva, ktorý sa mal zaoberať otázkou, či je neobmedzené uchovávanie biometrických údajov (DNA profil, odtlačky prstov, fotografia) osoby, ktorá už vykonala trest, bez ohľadu na vážnosť previnenia, bez reálnej možnosti preskúmania ich uchovávania v súlade s článkom 8 Dohovoru, teda či je v súlade s právom na rešpektovanie súkromného života.

ESĽP v tejto veci rozhodol, že došlo k porušeniu článku 8 Dohovoru. Teda, že neobmedzené uchovávanie biometrických údajov bez ohľadu na vážnosť previnenia, bez možnosti efektívneho preskúmania ich uchovávania, predstavuje neprimeraný zásah a nepovažuje sa za opatrenie „nevyhnutné v demokratickej spoločnosti".

Poznamenávame, že pri uchovaní biometrických údajov, akými sú fotografie, profil DNA ako aj odtlačky prstov, je osoba permanentne vystavená diskomfortu, keďže pri využití systémov stotožnenia na základe takýchto údajov aj s využitím umelej inteligencie (napríklad definovanie okruhu možných páchateľov pri daktyloskopickej stope s nižším počtom markantov), sa zaradí medzi osoby podozrivé. Vskutku niet ani praktického dôvodu, najmä po vykonaní uloženého trestu a aj prípadnom zahladení odsúdenia, na ďalšie uchovávanie spomínaných biometrických údajov.

---

[9] Rozhodnutie Európskeho súdu pre ľudské práva v tejto kauze je zverejnené na internete. ECtHR: *Gaughran v. the United Kingdom*, Application no. 45245/15, Judgment of 13 February 2020. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-200817"]}.

Pokiaľ ide o Slovenskú republiku, právnym predpisom, upravujúcim podmienky spracúvania biometrických údajov je Zákon o ochrane osobných údajov.[10] V prvom rade treba zdôrazniť, že tento zákon na jeho účely definuje v § 5 písm. c) biometrické údaje ako *„osobné údaje, ktoré sú výsledkom osobitného technického spracúvania osobných údajov týkajúcich sa fyzických charakteristických znakov fyzickej osoby, fyziologických charakteristických znakov fyzickej osoby alebo behaviorálnych charakteristických znakov fyzickej osoby a ktoré umožňujú jedinečnú identifikáciu alebo potvrdzujú jedinečnú identifikáciu tejto fyzickej osoby, ako najmä vyobrazenie tváre alebo daktyloskopické údaje.“*

Zároveň tento zákon v ustanoveniach § 16 upravuje spracúvanie osobitných kategórií osobných údajov, pričom sa vo výpočte nachádzajú aj biometrické údaje. Konkrétne v § 16 ods. 1 je uvedené (zvýrazňujeme relevantné časti): *„Zakazuje sa spracúvanie osobitných kategórií osobných údajov. Osobitnými kategóriami osobných údajov sú údaje, ktoré odhaľujú rasový pôvod alebo etnický pôvod, politické názory, náboženskú vieru, filozofické presvedčenie, členstvo v odborových organizáciách, genetické údaje,* ***biometrické údaje****, údaje týkajúce sa zdravia alebo údaje týkajúce sa sexuálneho života alebo sexuálnej orientácie fyzickej osoby.“* Z tohto zákazu sú však upravené hneď v nasledujúcom (druhom) odseku výnimky,[11] ku ktorým patrí aj spracúvanie osobných údajov podľa § 6 ods. 2 písm. g) tohto zákona (*„spracúvanie je* ***nevyhnutné z dôvodu verejného záujmu*** *na základe tohto zákona, osobitného predpisu alebo medzinárodnej zmluvy, ktorou je Slovenská republika viazaná,* ***ktoré sú primerané vzhľadom na sledovaný cieľ****, rešpektujú podstatu práva na ochranu osobných údajov a ustanovujú vhodné a konkrétne opatrenia na zabezpečenie základných práv a záujmov dotknutej osoby“*).

---

[10] Ide o zákon č. 18/2018 Z.z. o ochrane osobných údajov, v znení neskorších predpisov.

[11] Na účel zvýraznenia závažnosti dôvodov na spracúvanie osobitných kategórií osobitných údajov, kde sú zaradené aj biometrické údaje, uvedieme príklady, kedy zákaz spracúvania osobných údajov neplatí. Z § 16 ods. 2 písm. a), e), g) zákona č. 18/2018 Z.z. o ochrane osobných údajov, v znení neskorších predpisov, vyplýva, že zákaz spracúvania osobitných kategórií osobných údajov neplatí aj vtedy, ak dotknutá osoba vyjadrila **výslovný súhlas** so spracúvaním týchto osobných údajov aspoň na jeden konkrétny účel; ak sa **spracúvanie týka osobných údajov, ktoré dotknutá osoba preukázateľne zverejnila**; ak je spracúvanie **nevyhnutné z dôvodu verejného záujmu** na základe tohto zákona, osobitného predpisu alebo medzinárodnej zmluvy, ktorou je Slovenská republika viazaná, ktoré sú primerané vzhľadom na sledovaný cieľ, rešpektujú podstatu práva na ochranu osobných údajov a ustanovujú vhodné a konkrétne opatrenia na zabezpečenie základných práv a záujmov dotknutej osoby.

Z citovaných ustanovení vyplýva, že zákon o ochrane osobných údajov rigorózne podmieňuje spracúvanie osobitných kategórií osobných údajov. Snímanie daktyloskopických odtlačkov pri vydávaní občianskeho preukazu, realizované síce na základe zákona o občianskych preukazoch, však podľa nášho názoru nevyjadruje dostatočne dôvod verejného záujmu ani vtedy, ak sú tieto biometrické údaje chránené pred zneužitím. Pochopiteľne, pri preukazovaní totožnosti sa zásadne hľadí na fotografiu osoby na občianskom preukaze. Je však prinajmenšom sporné, či v prípade neistoty totožnosti je namieste porovnávanie podpisu bez technických zariadení na porovnanie rukopisu (tvar písma, prítlak a pod.), keďže na vyslovenie totožnosti rukopisu nepostačuje zbežné porovnanie, ale znalecké skúmanie. Pokiaľ ide o daktyloskopické odtlačky posledných článkov ukazovákov na pravej a ľavej ruke, stotožnenie osoby prichádza do úvahy pri náleze mŕtvoly (aj to nie v pokročilom štádiu rozkladu), čo však zďaleka nie je jedinou možnosťou na zistenie totožnosti. Takisto, pri zanechaní daktyloskopických stôp na mieste trestného činu, prichádza síce do úvahy porovnanie, opäť však iba v obmedzenej miere a s podmienkou, že neboli použité prostriedky na zamedzenie zanechania daktyloskopických stôp, napríklad rukavice. Konštatujeme, že odtlačky posledných článkov oboch ukazovákov sú biometrickými údajmi a ich snímanie *en bloc* všetkým občanom systémom *a priori* je v danom kontexte podľa nášho názoru v rozpore s článkom 8 Dohovoru.

V súvislosti s využitím umelej inteligencie sa tento druh stôp dá simulovať, inak povedané, je možné vytvoriť model posledných článkov ukazovákov a zanechať tak falošnú stopu.

V zákone o ochrane osobných údajov je samostatne v § 17 upravené spracúvanie osobných údajov týkajúcich sa uznania viny za spáchanie trestného činu alebo priestupku nasledovne: *„Prevádzkovateľom na účel spracúvania osobných údajov v registri trestov podľa osobitného predpisu môže byť len štátny orgán. Spracúvať osobné údaje týkajúce sa uznania viny za spáchanie trestného činu alebo priestupku alebo súvisiacich bezpečnostných opatrení možno len na základe osobitného predpisu alebo medzinárodnej zmluvy, ktorou je Slovenská republika viazaná, ktoré poskytujú primerané záruky ochrany práv dotknutej osoby.“* Ani toto ustanovenie neumožňuje v súvislosti so spáchaním trestného činu alebo priestupku uchovávať biometrické údaje páchateľa.

Do pozornosti dávame aj ustanovenie § 78 ods. 5 zákona o ochrane osobných údajov, ktoré znie: *„Prevádzkovateľ môže spracúvať genetické údaje,* ***biometrické údaje*** *a údaje týkajúce sa zdravia* ***aj na právnom základe osobitného predpisu alebo medzinárodnej zmluvy, ktorou je Slovenská republika viazaná****.“* O toto ustanovenie je možné oprieť spracúvanie

biometrických údajov, ktorými sú daktyloskopické odtlačky, získané pri vydávaní občianskeho preukazu, avšak opäť poukazujeme na možnú kolíziu s článkom 8 ods. 1 Dohovoru, týkajúcim sa ochrany súkromného života, a spomínané riziko zneužiteľnosti aj s prípadnou asistenciou umelej inteligencie.

V uvedených súvislostiach teda nie je možné ignorovať už spomínané rozhodnutie ESĽP vo veci Gaughran proti Spojenému kráľovstvu (2020). Význam tohto rozhodnutia je v potvrdení a rozšírení princípov z prípadu S. a Marper proti Spojenému Kráľovstvu (GC, 2008), že **štát musí pri uchovávaní citlivých osobných údajov brať do úvahy časové obmedzenie, proporcionalitu, možnosť preskúmania, a rozdielnosť podľa závažnosti previnenia**. V súvislosti s vyhotovením, spracúvaním a uchovávaním daktyloskopických odtlačkov pri vydávaní občianskych preukazov je teda na zváženie, keď sa osoby, ktoré odovzdávajú biometrický údaj, ničoho nedopustili. Výkladom *a maiori ad minus* je možné ustáliť, že ak sa u páchateľa trestnej činnosti po právoplatnom vyslovení viny a vykonaní trestu skúma nutnosť následného uchovávania jeho biometrických údajov, tento postup nemožno uplatniť na osoby, ktoré sa ničoho protiprávneho nedopustili. Predpokladáme, že uchovanie týchto biometrických údajov by neuspelo v teste proporcionality.

S akcentom opakujeme, že opierajúc sa o rozhodnutie ESĽP vo veci Gaughran proti Spojenému kráľovstvu 45245/15 (rozsudok z 13. 2. 2020), je nevyhnutné aplikovať Článok 8 ods. 1 Dohovoru (Rešpektovanie súkromného života), keďže aj ESĽP poukazuje na **neprimeraný charakter neobmedzeného uchovávania odtlačkov prstov**, DNA profilu a fotografie **osoby odsúdenej** za menej závažný trestný čin.

Ak by pán Gaughran, odsúdený za menej závažný trestný čin v Severnom Írsku, nebol iniciatívny a nesťažoval sa ESĽP na odobratie a neobmedzené uchovávanie jeho DNA profilu, odtlačkov prstov a fotografie políciou, naše konštatovanie o rozpore „zákonného" získavania biometrických údajov (odtlačkov prstov) pri vydávaní občianskeho preukazu by nemalo pevnú oporu v rozhodnutí ESĽP, ktorý vyslovil, že uchovávanie DNA profilu a odtlačkov prstov pána Gaughrana predstavovalo zásah do jeho súkromného života. Fotografia pána Gaughrana bola zhotovená počas zatknutia a mala byť neobmedzene uchovávaná v miestnej databáze na použitie políciou, pričom polícia mohla na fotografiu aplikovať technológie rozpoznávania tváre alebo mapovania tváre, pričom samotné zhotovenie a uchovávanie fotografie predstavovalo zásah do práva sťažovateľa na súkromný život. Z rozhodnutia ESĽP vyplýva, že uchovávanie biometrických údajov a fotografií sledovalo legitímny cieľ, konkrétne prevenciu kriminality. Kým pôvodné odobratie týchto údajov malo za cieľ prepojiť konkrétnu osobu s konkrétnym trestným činom, z ktorého bola

podozrivá, uchovávanie týchto údajov sledovalo širší cieľ, pomáhať pri identifikácii osôb, ktoré by sa v budúcnosti mohli dopustiť trestného činu.

Poznamenávame, že stále je predmetom úvah uchovávanie biometrických údajov tých osôb, ktoré sa dopustili protiprávneho konania, teda nie osôb bezúhonných.

ESĽP sa zaoberal aj argumentáciou v prospech uchovávania údajov, konkrétne tvrdením, že čím viac údajov sa uchováva, tým viac sa predchádza kriminalite. Tento argument bol podporovaný poukázaním na štatistiku z ktorej vyplýva, že osoby, ktoré už boli odsúdené, sú štatisticky najpravdepodobnejšie opätovne odsúdené v priebehu dvoch rokov. Prijatie takéhoto argumentu v rámci systému neobmedzeného uchovávania by však v praxi znamenalo ospravedlnenie uchovávania údajov o celej populácii, vrátane zosnulých osôb, čo by bolo nepochybne nadmerné a neprimerané.

Napriek voľnosti danej štátom, stanoviť lehotu uchovávania údajov, ak štát preferoval režim neobmedzeného uchovávania, bolo jeho povinnosťou zabezpečiť, aby pre odsúdenú osobu existovali určité účinné záruky. Biometrické údaje a fotografie pána Gaughrana však boli uchovávané bez ohľadu na závažnosť jeho trestného činu a bez akéhokoľvek posúdenia potreby ich neobmedzeného uchovávania. Neexistovalo žiadne ustanovenie, ktoré by pánovi Gaughranovi umožňovalo požiadať o vymazanie jeho údajov, ak by ich uchovávanie už nebolo potrebné vzhľadom na povahu trestného činu, vek osoby, časový odstup a aktuálnu osobnosť jednotlivca.

ESĽP teda vyslovil, že voči pánovi Gaughranovi štát prekročil prípustný rámec voľného uváženia v súvislosti s dobou uchovania biometrických údajov. Sporné uchovávanie predstavovalo neprimeraný zásah do práva sťažovateľa na rešpektovanie súkromného života a nemohlo byť považované za nevyhnutné v demokratickej spoločnosti. Podľa rozhodnutia ESĽP, konštatovanie porušenia predstavovalo dostatočné zadosťučinenie za nemajetkovú ujmu.

Reagujúc na právnu úpravu v Slovenskej republike, je uchovávanie biometrických údajov (prinajmenšom odtlačkov oboch ukazovákov) rizikom, pretože s asistenciou umelej inteligencie je každá, aj celkom bezúhonná, osoba vystavená permanentnej neistote pri potenciálnej excesívnej aktivite verejnej správy.[12] V sankčných konaniach, k akým sa zaraďuje aj konanie o priestupkoch,

[12] V Brazílii sa biometrii pri občianskej identifikačnej registrácii venovala vo svojej dizertačnej práci Marta Mourão Kanashiro na Fakulte filozofie, literatúry a humanitných vied Univerzity v São Paule (USP). Pozri MOURÃO KANASHIRO, M: *Biometria no Brasil e o registro de identidade civil: novos rumos para a identificação* [Biometria v Brazílii a občianska identifikačná registrácia: nové smerovania identifikácie]. Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo, 2011.

by síce rýchlo došlo k odhaleniu páchateľa, zanechajúceho daktyloskopické stopy na mieste priestupku, avšak apriórna evidencia tohto biometrického údaja je zásahom do práva na súkromný život a odporuje Dohovoru.

Podobne, monitorovanie verejných priestranstiev, realizované nielen v mestách, ale aj v obciach, môže síce s použitím umelej inteligencie urýchliť tvárovú identifikáciu človeka, avšak výsledok testu proporcionality medzi právom na súkromný život a zaisťovaním bezpečnosti verejných priestranstiev, nemusí byť jednoznačný. Tvárovou identifikáciou sa totiž rozpoznajú všetky osoby, nachádzajúce sa v dosahu snímacích zariadení, aj keď sa ničoho nedopustili. Permanentný monitoring s potenciálnym okamžitým stotožňovaním osôb tak môže nadobúdať povahu nežiadúceho zásahu v zmysle diskomfortu. Je otázne, či možnosť tvárovej identifikácie človeka odradí páchateľov od protiprávnych aktivít, keďže títo ľahko zamedzia tvárovej identifikácii umelou inteligenciou prekrytím tvárovej časti hlavy šiltovkou alebo inou formou.[13] Permanentným vystavením človeka jeho identifikácii, realizovanej umelou inteligenciou, na verejnom priestranstve, je podľa nášho názoru dotknuté právo na súkromie, zakotvené v medzinárodnoprávnych dokumentoch. Ide najmä o článok 8 Európskeho dohovoru o ľudských právach a základných slobodách (právo na rešpektovanie súkromného a rodinného života),[14] článok 7 Charty základných práv EÚ (rešpektovanie súkromného a rodinného života)[15]ako aj článok 12 Všeobecnej deklarácie ľudských práv. [16]

DOI: doi.org/10.11606/t.8.2011.tde-23042012-120515. Online: https://teses.usp.br/teses/disponiveis/8/8132/tde-23042012-120515/publico/2011_MartaMouraoKanashiro_VOrig.pdf.

[13] Biometria nachádza rozsiahle využitie. Reálne však môže spôsobovať problémy. Bezpečnosť pri identifikácii je otázna napríklad pri biometrii dúhovky a sietnice. Viac sa tejto téme venuje napr. dizertačná práca autora Iberê Anselmo Garcia na Právnickej fakulte Univerzity v São Paule (USP)**.** Viac pozri GARCIA, I. A.: *A seguranca na identificacao: A segurança na identificação: a biometria da íris e da retina* [Bezpečnosť pri identifikácii: biometria dúhovky a sietnice]. Faculdade de Direito da USP, São Paulo, 2009. DOI: doi.org/10.11606/D.2.2009.tde-24062010-084048. Online: https://teses.usp.br/teses/disponiveis/2/2136/tde-24062010-084048/publico/Ibere_Anselmo_Garcia_Dissertacao.pdf.

[14] Európsky dohovor o ľudských právach v znení protokolov č. 11, 14 a 15, s dodatkovým protokolom a protokolmi č. 4, 6, 7, 12, 13 a 16.

[15] Charta základných práv Európskej únie, 2012/C 326/02, Ú. v. EÚ C 326, 26. 10. 2012, s. 391–407.

[16] Článok 12 Všeobecnej deklarácie ľudských práv znie: *„Nikto nesmie byť vystavený svojvoľnému zasahovaniu do súkromného života, do rodiny, domova alebo korešpondencie, ani útokom na svoju česť a povesť. Každý má právo na právnu ochranu proti takýmto zásahom alebo útokom.“* Všeobecná deklarácia ľudských práv, rezolúcia Valného zhromaždenia OSN č. 217 A (III), 10. december 1948.

Aj preto je nevyhnutné právne upraviť základné zásady používania umelej inteligencie vo verejnej správe. Permanentný monitoring v spojení s databázou ľudských tvárí pri využití umelej inteligencie smeruje bez takejto právnej úpravy k nežiadúcej excesívnej aktivite verejnej správy.

Umelá inteligencia má byť využívaná, avšak v súlade so zákonnými pravidlami. Súčasná doba vyžaduje čo najskoršie právne zakotvenie takýchto pravidiel, zohľadňujúc medzinárodné zmluvy ako aj relevantné rozhodnutia ESĽP, aby sa predišlo porušovaniu ľudských práv v dôsledku automatizácie verejnej správy.

9.2.3 Prípad P.N. proti Nemecku, 2020

Ako tretie sme vybrali rozhodnutie, v ktorom ESĽP nedal za pravdu sťažovateľovi a konštatoval zachovanie spravodlivého pomeru medzi verejným záujmom vyšetrovania a ochranou súkromia sťažovateľa. Na tomto rozhodnutí je možné demonštrovať, že zásahy do súkromia sú v istých medziach prípustné aj pri získaní a uchovaní biometrických údajov. Ide o rozhodnutie Európskeho súdu pre ľudské práva v prípade označenom ako *„P.N. v. Germany (11. jún 2020; no. 74440/17).“* Týmto rozhodnutím ESĽP vyslovil, že zákonná úprava Spolkovej republiky Nemecko poskytuje primeranú ochranu, identifikačné údaje boli neintenzívnym zásahom a že bol zachovaný spravodlivý pomer medzi verejným záujmom vyšetrovania a ochranou súkromia sťažovateľa.[17] V popisovanom prípade by potenciálne využitie umelej inteligencie pri identifikovaní osoby na základe biometrických údajov bolo možné považovať za súladné s ochranou práva na súkromie.

V predmetnom prípade šlo o to, že orgány Spolkovej republiky Nemecko vydali príkaz, aby sa od osoby P.N. v rámci trestného konania zhotovili fotografie tváre a tela (vrátane možných tetovaní), odtlačky prstov a dlaňové odtlačky, ako identifikačné údaje. Dôvodom bola predchádzajúca trestná činnosť P.N. a polícia dospela k záveru o pravdepodobnosti, že sťažovateľ sa dopustí podobných trestných činov alebo bude z nich podozrivý aj v budúcnosti a považovala nariadené identifikačné úkony za nevyhnutné a primerané, spôsobilé uľahčiť vyšetrovanie prípadných budúcich trestných činov, najmä majetkových, keďže by bolo jednoduchšie potvrdiť alebo vylúčiť účasť P.N. ako páchateľa. Vnútroštátne súdy, na ktorých sa P.N. domáhal ochrany, konštatovali, že

[17] Plné znenie tohto rozhodnutia je zverejnené na internete. ECtHR: *Case of P.N. v. Germany*, Application no. 74440/17, Judgment of 11 June 2020. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-202758"]}.

podmienky podľa Trestného poriadku pre odobratie identifikačných údajov sťažovateľa boli splnené.

Sťažovateľ P.N. pred ESĽP namietal, že ide o zásah do jeho práva na súkromný život, ku ktorému došlo policajným príkazom vyhotoviť fotografie sťažovateľa, vziať jeho odtlačky prstov a dlaní a zostaviť opis jeho osoby. Sťažovateľ tvrdil, že tento zber identifikačných údajov, nariadený len na základe niekoľkých menej závažných odsúdení spred mnohých rokov,[18] predstavoval neprimeraný zásah do jeho práva na rešpektovanie súkromného života podľa článku 8 Dohovoru. Bolo potrebné posúdiť právnu otázku, či takéto zbieranie a uchovávanie identifikačných údajov (bez DNA a genetických údajov) v rámci vyšetrovania je v súlade s článkom 8 Dohovoru, respektíve či existujú primerané zákonné obmedzenia a pravidlá ochrany pred zneužitím. V tejto kauze ESĽP rozhodol, že nedošlo k porušeniu článku 8 Dohovoru s poukázaním na skutočnosť, že zákonná právna úprava Nemecka poskytuje primeranú ochranu pred zásahmi do súkromného života, identifikačné údaje boli neintenzívnym zásahom a že bol zachovaný spravodlivý pomer medzi verejným záujmom vyšetrovania a ochranou súkromia sťažovateľa.

Práve v tomto rozhodnutí ESĽP je poukázané na rozdiel medzi citlivými biometrickými údajmi (DNA, profil tváre, živé rozpoznávanie) a bežnými identifikačnými údajmi (fotografie, odtlačky, tetovania) a že štáty môžu uplatniť vyššie požiadavky ochrany pri spracovaní citlivých údajov. Pri rozhodovaní ESĽP prihliadal na úroveň skutočného zásahu do práva na rešpektovanie súkromného života. Zistil, že uchovávanie bunkových vzoriek je podstatne invazívnejšie v

[18] Polícia zistila, že v v júni 2011 voči sťažovateľovi začaté trestné stíhanie pre podielnictvo. Konštatovala, že sťažovateľ je recidivista. V roku 2000 bol odsúdený za podvod a v roku 2001 za iný podvod. Ďalšie dve trestné stíhania boli zastavené v roku 1999 a 2000, keďže išlo o menej závažné skutky v porovnaní s trestnými činmi, z ktorých bol sťažovateľ odsúdený alebo mohol očakávať odsúdenie. V roku 2004 bolo trestné stíhanie vo veci podvodu zastavené, keďže jeho konanie bolo posúdené ako menej závažné. V roku 2009 bolo trestné stíhanie pre ublíženie na zdraví a poškodenie cudzej veci zastavené po zaplatení určitej sumy. V roku 2010 mu bol uložený trest za krivé obvinenie. Polícia dospela k záveru, že je pravdepodobné, že sťažovateľ sa dopustí podobných trestných činov alebo bude z nich podozrivý aj v budúcnosti. Podľa jej názoru boli nariadené identifikačné úkony nevyhnutné a primerané. Uľahčili by vyšetrovanie budúcich trestných činov, najmä majetkových, keďže by bolo jednoduchšie potvrdiť alebo vylúčiť jeho účasť ako páchateľa. V zásade tento názor mal oporu v článku 81b Trestného poriadku, ustanovujúcom podmienky na odobratie identifikačných údajov od podozrivej osoby nasledovne: *„Fotografie a odtlačky prstov osôb podozrivých zo spáchania trestného činu možno vyhotoviť, aj proti ich vôli, a možno pri nich vykonať merania a iné podobné úkony, pokiaľ je to potrebné na účely vedenia trestného konania alebo na účely policajných evidencií (für die Zwecke des Erkennungsdienstes).“*

porovnaní s inými opatreniami, ako je uchovávanie odtlačkov prstov, vzhľadom na množstvo genetických informácií, ktoré obsahujú (prípad *S. a Marper*). Vyslovil tiež, že *„vnútroštátnym orgánom treba priznať istý priestor na voľnú úvahu pri hodnotení, či je zásah nevyhnutný v demokratickej spoločnosti. Šírka tohto priestoru sa mení a závisí od viacerých faktorov, vrátane povahy dotknutého práva podľa Dohovoru, jeho významu pre jednotlivca, povahy zásahu a cieľa, ktorý zásah sleduje.“* [19] ESĽP dospel k záveru, že dôvody uvádzané vnútroštátnymi orgánmi na odôvodnenie zásahu do sťažovateľovho práva na rešpektovanie súkromného života prostredníctvom získania a uchovávania jeho osobných údajov boli relevantné a dostatočné. Získanie a uchovávanie týchto údajov v danom prípade predstavovalo spravodlivú rovnováhu medzi verejným záujmom a súkromným záujmom sťažovateľa, a preto spadalo do medzí voľnej úvahy štátu. Napadnuté opatrenie teda predstavovalo primeraný zásah do sťažovateľovho práva na rešpektovanie jeho súkromného života a možno ho považovať za nevyhnutné v demokratickej spoločnosti. Bolo preto odôvodnené podľa článku 8 ods. 2 Dohovoru. Z toho vyplýva, že nedošlo k porušeniu článku 8 Dohovoru.

Môžeme konštatovať, že pri uchovávaní biometrických údajov, vymedzených v analyzovanom rozhodnutí ESĽP, by bola umelá inteligencia využitá na stotožnenie osoby, ktorej biometrické údaje sa uchovávajú, až pri podozrení zo spáchania trestného činu. Teda, človekom usmernená aktivita umelej inteligencie by sa podľa nášho názoru v takýchto situáciách nepovažovala za excesívnu, ale vzhľadom na konštatovania ESĽP by toto využitie bolo súladné s ochranou súkromného života podľa článku 8 Dohovoru.

9.2.4 Umelá inteligencia vo vzťahu k uchovávaným osobným údajom

Pri porovnaní troch rozhodnutí ESĽP zisťujeme, že prístup k základnému právu na ochranu súkromného života sa posudzuje podľa viacerých okolností. Teda, rôzny prístup k ochrane práv sťažovateľov v daných veciach sa môže javiť ako nerovný. Rovnosť a zákaz diskriminácie sú tiež základnými princípmi v medzinárodných dokumentoch ako Charta základných práv Európskej únie či Všeobecná deklarácia ľudských práv.

Uchovávané údaje sú pri využití umelej inteligencie spôsobilé narušiť súkromný život. Už samotné vedomie takéhoto využitia môže viesť ku diskomfortu vo vzťahu k súkromnému životu. Napokon, aj Rada Európy priznáva, že umelá inteligencia môže mať vplyv na ľudské práva. *„Technológie*

[19] Bod 75 rozhodnutia ESĽP vo veci P.N. proti Nemecku (2020).

umelej inteligencie ponúkajú ľudstvu veľké možnosti, no zároveň so sebou prinášajú vážne potenciálne riziká pre uplatňovanie ľudských práv, fungovanie demokracie a dodržiavanie zásad právneho štátu.“ [20] Je historickým faktom, že vynálezy ľudstva sa využívali v pozitívnom aj v negatívnom zmysle. Takisto umelá inteligencia má prioritne ľudstvu slúžiť, ale riziká pri jej využívaní, samozrejme, existujú. *„Rada Európy sa zaviazala riešiť tieto riziká komplexným spôsobom, ktorý zároveň podporuje inovácie.“* [21] Alain Berset ako generálny riaditeľ Rady Európy, sa k tejto téme vyjadril tiež. *„Rada Európy snaží zabezpečiť, aby výhody technických inovácií mohli byť využívané ľuďmi po celom kontinente a aj mimo neho, pričom riziká spojené s inováciami sú zmierňované.“* [22]

Vo vzťahu k prípadnému využitiu umelej inteligencie z analyzovaných rozhodnutí ESĽP vyplýva, že DNA a biometria tváre si vyžadujú najvyšší stupeň ochrany, naproti tomu fotografiám a daktyloskopickým odtlačkom postačuje nižšia úroveň ochrany, samozrejme, ak tu existujú záruky nezneužitia. Treba zároveň rozlišovať uchovávanie údajov vzhľadom na účelové obmedzenia a zohľadňovať závažnosť konkrétneho prípadu. Použitie vysoko invazívnych technológií (napr. rozpoznávanie tváre v reálnom čase s použitím technológie umelej inteligencie[23]) vyžaduje najvyššiu mieru odôvodnenia. Živé rozpoznávanie tváre (tzv. live facial recognition) predstavuje technológiu, ktorá umožňuje sledovať a identifikovať osobu v reálnom čase, čo je výrazne invazívnejšie než pasívne zhromažďovanie obrázkov či videozáznamov. Takýto spôsob zasahovania vyžaduje veľmi silné právne záruky. So živým rozpoznávaním tváre sú teda spojené požiadavky na najvyššiu mieru odôvodnenia.

---

[20] COUNCIL OF EUROPE: The Council of Europe & Artificial intelligence – Overview of the Council of Europe activities in the field of artificial intelligence [Rada Európy a umelá inteligencia - Prehľad aktivít Rady Európy v oblasti umelej inteligencie]. Strasbourg, Council of Europe Publishing, F-67075 Strasbourg Cedex, March 2023, s. 3. Online: https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6?utm_source=chatgpt.com.

[21] *Tamže*, s. 3.

[22] BERSET, A.: *Foreword*. In: COUNCIL OF EUROPE: *The Council of Europe & Artificial intelligence – Overview of the Council of Europe activities in the field of artificial intelligence* [Rada Európy a umelá inteligencia - Prehľad aktivít Rady Európy v oblasti umelej inteligencie]. Strasbourg, Council of Europe Publishing, F-67075 Strasbourg Cedex, March 2023, s. 1. Online: https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6.

[23] ESĽP riešil otázku zákonnosti rozpoznávania tváre v reálnom čase v kauze Glukhin proti Rusku. Pozri ECtHR: Case of Glukhin v. Russia, Application no. 11519/20, Judgment of 4 July 2023. Online: https://hudoc.echr.coe.int/#{"itemid":["002-14142"]}.

Na zamedzenie zneužitiu uchovávaných údajov sú nevyhnutné záruky, najmä preskúmateľnosť zákonnosti uchovávania a možnosť vymazania údajov.

9.3 Námety de lege ferenda

9.3.1 Všeobecné námety de lege ferenda

V súčasnosti už existujú viaceré právne pravidlá a kritériá pre reguláciu umelej inteligencie. Patrí sem Rámcová dohoda Rady Európy o umelej inteligencii a ľudských právach, demokracii a právnom štáte.[24] Táto konvencia sa zameriava na ľudské práva, demokraciu a právny štát, pričom požaduje, aby cyklus existencie systémov umelej inteligencie (navrhovanie, vývoj, použitie, odstránenie) bol v súlade s týmito princípmi. Medzi základné princípy, ktoré musia byť dodržané, patrí dôstojnosť človeka, rovnosť, nediskriminácia, ochrana súkromia, transparentnosť, zodpovednosť, spoľahlivosť a bezpečná inovácia. Ďalej je tu Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie (akt o umelej inteligencii, teda „Artificial Intelligence Act" alebo tiež „AI Act,"),[25] ktorý predstavuje prvý komplexný právny rámec na svete pre reguláciu umelej inteligencie. Jeho hlavným cieľom je zabezpečiť, aby systémy umelej inteligencie používané v Európskej únii boli bezpečné, etické a dôveryhodné, a aby neohrozovali základné práva občanov. Ide o veľmi dôležité pravidlá, podľa ktorých sú v relevantných systémoch AI zakázané napr. manipulácia s ľuďmi alebo rozpoznávanie emócií v práci bez súhlasu dotknutej osoby. Ďalej popisujú AI systémy s vysokým rizikom (používané v zdravotníctve, doprave, súdnictve, verejnej správe, sú na ne kladené prísne požiadavky dohľadu) a s nižším rizikom (menej regulované, avšak tiež s požiadavkou dodržiavania základných pravidiel transparentnosti a bezpečnosti).

So zohľadnením aktuálnych požiadaviek, kladených na umelú inteligenciu a jej žiadúce využívanie, spolu s rozhodovacou činnosťou ESĽP, bude aj v Slovenskej republike nutné prehodnotiť databázy osobných údajov, existujúcich na základe zákonov, práve pri možnej excesívnej aktivite verejnej

---

[24] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series, No. 225, Vilnius 5.IX.2024. Online: https://rm.coe.int/1680afae3c.

[25] Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii). Online: https://eur-lex.europa.eu/eli/reg/2024/1689/oj?eliuri=eli%3Areg%3A2024%3A1689%3Aoj&locale=sk.

správy (monitorovanie verejných priestranstiev s prípadnou okamžitou tvárovou identifikáciou osôb) bez zákonných dôvodov.

V príspevku boli naznačené riziká pri rozpoznávaní ľudskej tváre. V zákonnej úprave bude potrebné jasne špecifikovať, čo sa považuje za „technológiu rozpoznávania tváre" a tejto technológii určiť hranice použitia, a to tak pri statickom rozpoznávaní (zo záznamu) ako aj pri živom sledovaní. Rozpoznávanie tváre technológiu AI sa bude musieť zákonom obmedziť na výnimočné prípady, ku ktorým by patrilo pátranie po osobách či ochrana života alebo zdravia pred násilím. Pri menej závažných administratívnych deliktoch by mohlo byť využitie technológie rozpoznávania tváre vyhodnotené v teste proporcionality ako neprimerané. Napokon, je potrebné zvážiť zákonný zákaz hromadného alebo preventívneho sledovania, teda sledovanie by nesmelo byť plošné, masové alebo preventívne, bez individuálneho dôvodu. Preto aj zdôvodňovanie formuláciami ako „v záujme verejného poriadku" alebo „v prípade potreby" bez konkrétneho vymedzenia, nebude možné považovať za vyhovujúce.

Zároveň by, podľa nášho názoru, monitorované územia mali byť zreteľné označené so zvýraznením, že sa uplatňuje technológia rozpoznávania tváre. Pri využití tejto technológie je vhodné uvažovať, že by mala podliehať predchádzajúcemu schváleniu nezávislým orgánom, napríklad novovytvoreným orgánom na využívanie umelej inteligencie, ale vylúčené nie je ani povolenie vydané súdom. Napokon, každé použitie rozpoznávania tváre by malo byť zaevidované, respektíve zdokumentované, vrátane dôvodu použitia, identity používateľov a časového i priestorového ohraničenia. Považujeme za legitímnu tú požiadavku, aby dotknuté osoby by mali mať právo vedieť, že boli predmetom spracovania automatizovaného rozpoznávania tváre, pričom výnimky by mali byť dovolené iba vo výnimočných prípadoch, ako napr. pri operatívnej činnosti policajných útvarov so súhlasom súdu. Okrem toho dotknutá osoba by mala mať možnosť využiť účinné prostriedky nápravy, napr. možnosť podať sťažnosť ako aj žiadosť o výmaz údajov.

Osobitnú pozornosť je nutné venovať zvýšenej ochrane citlivých údajov, kam patrí aj biometria, osobitne, ak spracovanie odhaľuje zdravotný stav, údaje z rozpoznávania tváre alebo DNA. Aj vzhľadom na rozhodovaciu činnosť ESĽP v týchto prípadoch musia byť uplatnené najprísnejšie záruky, či už pôjde o krátku dobu uchovávania ako aj obmedzený prístup k údajom.

Považujeme za dôležité zvýrazniť, že permanentný monitoring osoby, navyše s technológiou rozpoznávania tváre, zasahuje do práva na súkromie, preto by mal byť využívaný iba v prípadoch nevyhnutnosti, teda keď existuje

skutočná a aktuálna potreba, neexistujú miernejšie prostriedky (napr. ľudský dohľad) a zásah je primeraný cieľu.

ESĽP vo viacerých prípadoch (Marper, Gaughran, Glukhin) stanovil, že biometria, AI a technológie rozpoznávania tváre patria medzi najcitlivejšie zásahy do súkromia a preto si vyžadujú najvyššiu úroveň zákonnosti, proporcionality a ochranných záruk. Tieto zásady môžu slúžiť ako základ pre právnu úpravu používania AI na Slovensku, rešpektujúc ústavnosť a súladnosť s Európskym dohovorom o ľudských právach.

9.3.2 Konkrétne námety de lege ferenda

Zaradením aplikovateľnosti technológií umelej inteligencie vo verejnej správe budú dotknutými právnymi predpismi predovšetkým o Zákon o Policajnom zbore[26] a Zákon o ochrane osobných údajov[27], a to v súvislosti s doplnením ustanovení o používaní technológie rozpoznávania tváre (FRT) a automatizovaného spracovania údajov (vrátane AI).

Predkladáme tieto konkrétne námety na doplnenie Zákona o policajnom zbore, ktoré by sa dotkli ustanovení v tretej hlave (*„Povinnosti, oprávnenia a prostriedky policajta a Policajného zboru“*) a štvrtej hlave (*„Spracúvanie informácií Policajným zborom“*).

„§ ... Používanie technológie rozpoznávania tváre a automatizovaných systémov

(1) Policajný zbor je oprávnený používať technológiu rozpoznávania tváre alebo iné automatizované systémy identifikácie osôb len v prípade, že:

a) existuje závažný dôvod verejného záujmu, ako je ochrana života, predchádzanie vážnej trestnej činnosti alebo pátranie po nezvestnej osobe,

b) použitie takejto technológie je nevyhnutné a primerané vo vzťahu k sledovanému cieľu, a

c) nie je možné dosiahnuť cieľ menej invazívnym spôsobom.

(2) Použitie živého rozpoznávania tváre na verejne prístupných miestach je možné len na základe:

a) predchádzajúceho povolenia súdu

b) v časovo a vecne obmedzenom rozsahu podľa povolenia súdu, s presne určeným účelom, miestom, rozsahom a zoznamom hľadaných osôb

(3) Technológia nesmie byť použitá na:

a) sledovanie účastníkov pokojného zhromaždenia,

b) plošné, neadresné monitorovanie verejných priestorov bez konkrétneho bezpečnostného dôvodu.

---

26 Zákon č. 171/1993 Z.z. o Policajnom zbore, v znení neskorších predpisov.

27 Zákon č. 18/2018 Z.z. o ochrane osobných údajov a o zmene a doplnení niektorých zákonov, v znení neskorších predpisov.

(4) Policajný zbor je povinný viesť evidenciu každého použitia technológie rozpoznávania tváre, ktorá obsahuje účel, dátum, miesto a čas použitia, označenie zodpovednej osoby a údaj o tom, či bola osoba identifikovaná alebo nie.

(5) Každá osoba, voči ktorej bola technológia použitá, má po ukončení použitia tejto technológie právo byť o tom informovaná, pokiaľ neexistuje vážna prekážka z dôvodu prebiehajúceho trestného konania.

(6) Údaje získané pomocou rozpoznávania tváre musia byť:

a) uchovávané najviac 72 hodín, ak sa nepotvrdí zhodná identifikácia,

b) zmazané bez zbytočného odkladu, ak nedošlo k identifikácii osoby ako hľadanej,

c) spracúvané len v nevyhnutnom rozsahu, s rešpektovaním práva na súkromie a ochranu osobných údajov.

(7) Na použitie umelej inteligencie a automatizovaných rozhodovacích systémov sa primerane vzťahujú zásady stanovené v osobitnom predpise o ochrane osobných údajov."

Zákon o ochrane osobných údajov by bolo možné doplniť v tretej časti („Osobitné pravidlá ochrany osobných údajov fyzických osôb pri ich spracúvaní príslušnými orgánmi") nasledovne.

„§ ... Automatizované spracúvanie biometrických údajov

(1) Spracúvanie biometrických údajov na účely identifikácie osoby pomocou technológie rozpoznávania tváre je možné len v prípade, ak:

a) je to výslovne povolené zákonom,

b) je to nevyhnutné na dosiahnutie zákonného cieľa a

c) existujú primerané a účinné záruky ochrany práv a slobôd dotknutej osoby.

(2) V prípadoch spracúvania osobitných kategórií údajov, najmä biometrických údajov, je prevádzkovateľ povinný:

a) vykonať vopred posúdenie vplyvu na ochranu osobných údajov,

b) prijať technické a organizačné opatrenia, ktoré zabránia zneužitiu,

d) informovať dotknutú osobu o použitých údajoch a možnostiach opravy alebo výmazu."

Záver

V príspevku sme identifikovali viacero rizík, súvisiacich s využívaním schopností umelej inteligencie. Konštatujeme, že technológia umelej inteligencie zameraná na rozpoznávanie tváre je síce spôsobilá stotožniť zobrazenú osobu, čo v prípade sankčných administratívnych konaní môže predstavovať prvok urýchľujúci stotožnenie osoby, proti ktorej sa bude viesť sankčné administratívne konanie, avšak z hľadiska proporcionality chránených záujmov nie je prevaha prínosu jednoznačná. Pokiaľ ide o daktyloskopické stopy s nedostatočným množstvom markantov potrebných na stotožnenie osoby, ktorá stopy zanechala, umelá inteligencia dokáže definovať množinu všetkých osôb, prichádzajúcich do úvahy ako páchatelia. Tým sú vytvorené podmienky pre

permanentný monitoring fyzických osôb, bez ohľadu na súhlasný prejav vôle s takýmto monitoringom. Vzhľadom na zákon o občianskych preukazoch, upravujúci, okrem zdravotných výnimiek, snímanie daktyloskopických odtlačkov *en bloc* všetkým občanom Slovenskej republiky v súvislosti s vydávaním občianskeho preukazu, konštatujeme, že existencia databázy týchto odtlačkov má povahu excesívnej aktivity verejnej správy, ktorá hoci na základe zákona, avšak v porovnaní s rozhodovacou činnosťou ESĽP, môže dosahovať rozpor s článkom 8 Dohovoru. Navyše, táto excesívna aktivita môže s asistenciou umelej inteligencie predstavovať zásah do ľudských práv. Výkladom *a maiori ad minus* je možné ustáliť, že ak sa u páchateľa trestnej činnosti po právoplatnom vyslovení viny a vykonaní trestu skúma nutnosť následného uchovávania jeho biometrických údajov, tento postup nemožno uplatniť na osoby, ktoré sa ničoho protiprávneho nedopustili. Uchovanie týchto biometrických údajov podľa analýzy, vykonanej v príspevku, nebude mať úspech ani v teste proporcionality.

Podobne, monitorovanie verejných priestranstiev môže síce s použitím umelej inteligencie urýchliť tvárovú identifikáciu človeka, avšak výsledok testu proporcionality medzi právom na súkromný život a zaisťovaním bezpečnosti verejných priestranstiev, nemusí byť jednoznačný. Tvárovou identifikáciou sa rozpoznajú všetky osoby v dosahu snímacích zariadení, aj keď sa ničoho nedopustili. To môže predstavovať nielen diskomfort, ale aj reálny zásah do súkromného života. Permanentným vystavením človeka jeho identifikácii realizovanej umelou inteligenciou na verejnom priestranstve je podľa nášho názoru dotknuté právo na súkromie, zakotvené v medzinárodnoprávnych dokumentoch, najmä v článku 8 Dohovoru (právo na rešpektovanie súkromného a rodinného života), v článku 7 Charty základných práv EÚ ako aj v článku 12 Všeobecnej deklarácie ľudských práv. Aj preto v príspevku zdôrazňujeme nevyhnutnosť právne upraviť základné zásady používania umelej inteligencie vo verejnej správe v snahe minimalizovať ba až vylúčiť nežiadúcu excesívnu aktivitu verejnej správy. Súčasná doba vyžaduje čo najskoršie právne zakotvenie takýchto pravidiel, zohľadňujúc medzinárodné zmluvy ako aj relevantné rozhodnutia ESĽP, aby sa predišlo porušovaniu ľudských práv v dôsledku automatizácie verejnej správy.

Podľa nášho názoru by monitorované územia, na ktorých sa uplatňuje technológia rozpoznávania tváre, mali byť zreteľné označené so zvýraznením tejto skutočnosti. Využitie technológie rozpoznávania tváre by malo podliehať predchádzajúcemu schváleniu nezávislým orgánom a každé použitie rozpoznávania tváre by malo byť zaevidované a zdokumentované, s uvedením dôvodu použitia, identity používateľov a časového i priestorového ohraničenia. Dotknuté osoby by mali mať právo vedieť, že boli predmetom spracovania

automatizovaného rozpoznávania tváre s možnosťou využitia prostriedkov nápravy (sťažnosť, žiadosť o výmaz údajov).

Permanentný monitoring osoby spojený s technológiou rozpoznávania tváre zasahuje do práva na súkromie, preto by mal byť využívaný iba v prípadoch nevyhnutnosti. Biometria, umelá inteligencia a technológie rozpoznávania tváre patria medzi najcitlivejšie zásahy do súkromia. Ich využívanie musí byť zákonné a zodpovedať požiadavke proporcionality spolu so zárukami ochrany proti zneužitiu. Tieto nároky by sa mali de lege ferenda odraziť aj vo všeobecnej právnej úprave používania AI na Slovensku ako aj v osobitných zákonoch, vrátane Zákona o policajnom zbore a Zákona o ochrane osobných údajov.

Medzinárodné dokumenty ako Európsky dohovor o ľudských právach, Charta základných práv Európskej únie a takisto rozhodnutia ESĽP sú dôležitým základom pre posudzovanie zákonnosti využívania AI. Vzhľadom na to Slovenská republika bude musieť zabezpečiť primeranú ochranu práv a slobôd jednotlivcov aj v kontexte nových technológií.

Zoznam použitej literatúry a zdrojov

1. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series, No. 225, Vilnius 5.IX.2024, 12 s. Online: https://rm.coe.int/1680afae3c.
2. COUNCIL OF EUROPE: *The Council of Europe & Artificial intelligence – Overview of the Council of Europe activities in the field of artificial intelligence* [Rada Európy a umelá inteligencia - Prehľad aktivít Rady Európy v oblasti umelej inteligencie]. Strasbourg, Council of Europe Publishing, F-67075 Strasbourg Cedex, March 2023, 38 s. Online: https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6.
3. Dohovor o ochrane ľudských práv a základných slobôd, publikovaný pod č. 209/1992 Zb., v znení protokolov č. 11, 14 a 15, s dodatkovým protokolom a protokolmi č. 4, 6, 7, 12, 13 a 16. Online: https://www.echr.coe.int/documents/d/echr/convention_slk.
4. EDWARDS, L., VEALE, M.: *Slave to the Algorithm? Why a "Right to Explanation" Is Probably Not the Remedy You Are Looking for*. In: Duke Law & Technology Review. roč. 16, č. 1, 2017. Online*:* https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1315&context=dltr.
5. Charta základných práv Európskej únie, 2012/C 326/02, Ú. v. EÚ C 326, 26. 10. 2012, s. 391–407.
6. GARCIA, I. A.: *A seguranca na identificacao: A segurança na identificação: a biometria da íris e da retina* [Bezpečnosť pri identifikácii: biometria dúhovky a sietnice]. Faculdade de Direito da USP, São Paulo, 2009. DOI: doi.org/10.11606/D.2.2009.tde-24062010-084048. Online: https://teses.usp.br/teses/disponiveis/2/2136/tde-24062010-084048/publico/Ibere_Anselmo_Garcia_Dissertacao.pdf.
7. KAPLAN, A.: *Artificial Intelligence, Business and Civilization: Our Fate Made in Machines*. London: Routledge, 2022, 66 s. Online: https://www.taylorfrancis.com/books/mono/10.4324/9781003244554/artificial-intelligence-business-civilization-andreas-kaplan; DOI: doi.org/10.4324/9781003244554.
8. MOURÃO KANASHIRO, M: *Biometria no Brasil e o registro de identidade civil: novos rumos para a identificação* [Biometria v Brazílii a občianska identifikačná registrácia: nové smerovania identifikácie]. Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo, 2011, DOI: doi.org/10.11606/t.8.2011.tde-23042012-120515. Online:

https://teses.usp.br/teses/disponiveis/8/8132/tde-23042012-120515/publico/2011_MartaMouraoKanashiro_VOrig.pdf.

9. Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii). Online: https://eur-lex.europa.eu/eli/reg/2024/1689/oj?eliuri=eli%3Areg%3A2024%3A1689%3Aoj&locale=sk.
10. SLOVÁKOVÁ, A. I.: *Čo je to umelá inteligencia? 1 000 slov o nej od nej,* 21.4.2023. Online: https://www.techbox.sk/co-je-to-umela-inteligencia-1-000-slov-o-nej-od-nej.
11. Európsky parliament: *Umelá inteligencia: definícia a využitie*, 21.06.2023. Online: https://www.europarl.europa.eu/topics/sk/article/20200827STO85804/umela-inteligencia-definicia-a-vyuzitie.
12. Rozhodnutie Európskeho súdu pre ľudské práva - ECtHR: *Gaughran v. the United Kingdom* Application no. 45245/15, Judgment of 13 February 2020. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-200817"]}.
13. Rozhodnutie Európskeho súdu pre ľudské práva - ECtHR: *Glukhin v. Russia*, Application no. 11519/20, Judgment of 4 July 2023. Online: https://hudoc.echr.coe.int/#{"itemid":["002-14142"]}.
14. Rozhodnutie Európskeho súdu pre ľudské práva - ECtHR: *Case of P.N. v. Germany*, Application no. 74440/17, Judgment of 11 June 2020. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-202758"]}.
15. Rozhodnutie Európskeho súdu pre ľudské práva - ECtHR**:** *S. and Marper v. United Kingdom*, Application nos. 30562/04, 30566/04, Judgment of 4 December 2008. Online: https://hudoc.echr.coe.int/eng#{"itemid":["001-90051"]}.
16. Všeobecná deklarácia ľudských práv, rezolúcia Valného zhromaždenia OSN č. 217 A (III), 10. december 1948
17. Zákon č. 171/1993 Z.z. o Policajnom zbore, v znení neskorších predpisov
18. Zákon č. 18/2018 Z.z. o ochrane osobných údajov, v znení neskorších predpisov
19. Zákon č. 395/2019 Z.z.o občianskych preukazoch a o zmene a doplnení niektorých zákonov, v znení neskorších predpisov

Informácie o autorovi
doc. JUDr. Tibor SEMAN, PhD.
Univerzita Pavla Jozefa Šafárika v Košiciach
Právnická fakulta
Katedra ústavného práva a správneho práva
Email: tibor.seman@upjs.sk

10 Přezkoumatelnost jako lidské právo a algoritmické rozhodování

Ondřej Lipka

https://doi.org/10.33542/HRF-0541-2-10

Abstrakt

"Přezkoumatelnost je vlastností soudních a správních rozhodnutí, která značí dosažení určitého standardu transparentnosti. Jako taková má zásadní lidskoprávní přesah, neboť se v ní odráží mnoho ústavním a mezinárodním právem garantovaných požadavků na spravedlivé rozhodování. Příspěvek popíše přístupy k přezkoumatelnosti v judikatuře vrcholných soudů a popíše lidskoprávní souvislosti přezkoumatelnosti. Následně budou tyto poznatky konfrontovány se současným technologickým vývojem, zejm. ve vztahu ke zvyšující se popularitě tzv. velkých jazykových modelů. Nejdříve budou popsány technologické limitace algoritmického rozhodování, na které již delší dobu upozorňuje literatura, zejm. ve vztahu k diskriminačním vstupním datům. Dále budou analyzovány technologická omezení specifická pro velké jazykové modely, které podle relevantních studií nevykazují znaky přemýšlení, přičemž bude zdůrazněno, že velké jazykové modely jsou založeny na odhadování pravděpodobnosti. Příspěvek bude dále argumentovat, že všechny definice přezkoumatelnosti až doposud obsahovaly implicitní prvek lidského rozhodování. Problematika bude demonstrována zejm. na příkladu deferenčního přístupu správních soudů při přezkumu správního uvážení. V závěru příspěvku bude vyloženo, že současné chápání přezkoumatelnosti jako projevu několika ústavněprávních garancí není kompatibilní s algoritmickým rozhodováním. Bude položena otázka, zdali je možné či vůbec žádoucí z požadavků na přezkoumatelnost tváří tvář technologickým výzvám slevit. "

Úvod

Vzestup algoritmického rozhodování představuje zásadní výzvu pro dosavadní fungování rozhodování ve veřejné správě. Tento příspěvek se zaměřuje na základní právo na odůvodnění rozhodnutí, resp. zásadu transparentnosti ve veřejné správě v kontextu otázky, zda by rozhodnutí generované prediktivní algoritmem mohlo dostát podmínkám, která jsou na rozhodnutí tímto lidskoprávním rámcem kladena.

V první části příspěvku bude analyzována přezkoumatelnost jako prvek práva na odůvodnění. Následně dojde k analýze algoritmického rozhodování, a to zejm. v kontextu rozvoje velkých jazykových modelů. V třetí části práce bude poukázáno na lidské rozhodování jako inherentní součást přezkoumatelnosti, a to na příkladu deference jako mezní situace přezkumu správního rozhodnutí. V

závěru práce budou shrnuty poznatky a navržena některá řešení, jak přistoupit ke konfliktu algoritmického rozhodování a přezkoumatelnosti.

10.1 Lidskoprávní rozměr přezkoumatelnosti

Stěžejním pojmem tohoto příspěvku je přezkoumatelnost. Na úvod je třeba přiznat, že označovat samotnou přezkoumatelnost rozhodnutí za lidské právo je nepochybně určitým zjednodušením. Přezkoumatelnost je spíše vlastností rozhodnutí, které naplňuje požadavky vyplývající z práva na odůvodnění rozhodnutí, které je součástí širšího komplexu práv spadajících do práva na spravedlivý proces. Přezkoumatelnost rozhodnutí totiž lze chápat jako materiální stránku práva na odůvodnění. Jak upozorňuje literatura, relevantní vadou z hlediska základních práv není jen úplná absence odůvodnění, ale též odůvodnění, které je nedostatečné, nesrozumitelné, nebo se dostatečně nevypořádává s tvrzeními účastníka řízení.[1]

Právo na odůvodnění rozhodnutí lze v jeho současné podobě užívané jak českým Ústavním soudem, tak i Nejvyšším správním soudem, vysledovat k judikatuře Evropského soudu pro lidská práva (dále též „ESLP") k čl. 6 odst. 1 Evropské úmluvy o lidských právech (dále též „EÚLP"). Tato judikatura však připouští, že právo na odůvodnění rozhodnutí nelze směšovat s reakcí na každou uplatněnou námitku.[2] Nelze ani připustit, že by z judikatury ESLP bylo možné dovodit konkrétní, univerzální vzorec pro to, jaké rozhodnutí naplňuje požadavky: „*Rozsah povinnosti poskytnout důvody se bude lišit podle povahy případu a musí být vyhodnocen ve světle okolností případu.*"[3] I český Ústavní soud vychází z předpokladu, že rozdílné situace vyžadují povahově jiná odůvodnění. Vyplývá-li řešení přímo z textu zákona, bude se odůvodnění lišit od situace, kdy musí soud vysvětlit své právní úvahy s odkazy na judikaturu nebo závěry právní nauky.[4] Zvláště pečlivé odůvodnění pak vyžadují speciální situace, např. situace „tvrzení proti tvrzení" v trestních věcech.[5] Naopak v řízení o opravném prostředku může vyšší soud převzít závěry nižšího soudu, avšak pouze za předpokladu, že odůvodní, proč názor nižšího soudu obstojí.[6]

---

1 POSPÍŠIL, I.: *Čl. 36. Právo na spravedlivý proces*. In: WAGNEROVÁ, E. et al.: Listina základních práv a svobod: komentář, Praha: Wolters Kluwer ČR, 2023, s. 786.

2 SCHABAS, W. A.: *The European Convention on Human Rights: a commentary*. Oxford: Oxford University Press, 2015, s. 297.

3 *Ibid.*

4 KMEC, J.: *Právo domáhat se svého práva (čl. 36).* In: KÜHN, Z., KRATOCHVÍL, J., KMEC, J., KOSAŘ, D., ANTOŠ, M. et al.: Listina základních práv a svobod: velký komentář. Praha: Leges, 2022, s. 1351.

5 *Ibid,* s. 1352.

6 *Ibid,* s. 1353.

Ústavní soud ve své judikatuře rozšířil uvedené požadavky i na rozhodnutí jiných orgánů než soudů, např. rozhodnutí státních zastupitelství.[7]

10.1.1 Nepřezkoumatelnost v judikatuře správních soudů

Judikatura českého Nejvyššího správního soudu k nepřezkoumatelnosti v podstatě kopíruje uvedená východiska. Aplikuje je přitom fakticky bez rozdílu na rozsudky krajských soudů, které Nejvyšší správní soud přezkoumává v řízeních o kasačních stížnostech, tak i na rozhodnutí správních orgánů, která přezkoumává zprostředkovaně.[8] Na tomto místě je zejm. potřeba zdůraznit, že nepřezkoumatelnost správního rozhodnutí je v judikatuře správních soudů chápána poněkud širším způsobem než jen jako projev práva na odůvodnění. Jak uvádí literatura, nepřezkoumatelnost správního rozhodnutí je vadou, *„která souvisí zpravidla s nedodržením požadavků kladených procesním právním předpisem na výrok a odůvodnění správních rozhodnutí*."[9] Tento příspěvek je věnován požadavkům na odůvodnění rozhodnutí, a proto můžeme nepřezkoumatelnostní vady spojené s požadavky na výrok ponechat stranou zájmu.

Správní soudy v České republice a proces ochrany práv před nimi je normován zákonem č. 150/2002 Sb., soudní řád správní, ve znění pozdějších předpisů (dále též „SŘS"). Samotný SŘS s pojmem nepřezkoumatelnosti pracuje na několika místech. Stěžejním je v tomto ohledu § 76 odst. 1 SŘS, který v kontextu stanovení práva soudu rozhodnout o žalobě bez nařízení jednání jakoby mimoděk vymezuje dva základní druhy nepřezkoumatelnosti: pro nesrozumitelnost a pro nedostatek důvodů.

Nepřezkoumatelnost odůvodnění pro nesrozumitelnost je snadné si představit. Jedná se o situace, kdy je odůvodnění natolik formulačně nesrozumitelné, že nejde poznat, jaké právní a skutkové závěry správní orgán učinil.[10] Bude se též jednat o situace, kdy je rozhodnutí vnitřně rozporné. Nebude však možné hovořit o tomto druhu nepřezkoumatelnosti za situace, kdy je možné formulační nedostatky či vnitřní rozpornost odstranit výkladem rozhodnutí jako

---

7 *Ibid*, s. 1356.

8 V tomto ohledu lze upozornit, že relevantní komentářová literatura odkazuje na rozsudky, v nichž se hovoří o nepřezkoumatelnosti soudních rozhodnutí, i když hovoří o nepřezkoumatelnosti správních rozhodnutí; srov. např. KOCOUREK, T.: *§ 76. Rozhodování bez nařízení jednání*. In: KÜHN, Z., KOCOUREK, T., MIKEŠ, P., DIENSTBIER, F., KADLEC, F. et al.: Soudní řád správní: komentář. Praha: Wolters Kluwer ČR, 2019, s. 618.

9 *Ibid*, s. 616.

10 Srov. např. rozsudek Nejvyššího správního soudu ze dne 24. 9. 2003, čj. 7 A 547/2002-24.

celku.[11] Nepřezkoumatelnost pro nesrozumitelnost odůvodnění nezakládají ani zjevné chyby v počtech nebo v psaní.[12]

Nepřezkoumatelnost pro nedostatek důvodů je pak širší kategorií, obsahující více „skutkových podstat" nepřezkoumatelnosti. Jedná se např. o nedostatek skutkových důvodů, kdy se správní orgány opomenou zabývat pro posouzení věci podstatnou skutečností. V tomto případě pak nezáleží na tom, že správní orgán třeba k věci provádí dokazování, podstatné je, že tato skutečnost nenalezne odraz v odůvodnění rozhodnutí.[13]

Důvodem pro nepřezkoumatelnosti pro nedostatek důvodů je pak i nenaplnění zákonných požadavků kladených na odůvodnění rozhodnutí, jak jsou vymezeny v § 68 odst. 3 zákona č. 500/2004 Sb., správní řád, ve znění pozdějších předpisů. Zákon stanoví, že se v odůvodnění „*uvedou důvody výroku nebo výroků rozhodnutí, podklady pro jeho vydání, úvahy, kterými se správní orgán řídil při jejich hodnocení a při výkladu právních předpisů, a informace o tom, jak se správní orgán vypořádal s návrhy a námitkami účastníků a s jejich vyjádřením k podkladům rozhodnutí.*" Uvedená zákonnou úpravu lze považovat za promítnutí požadavků práva na odůvodnění ve smyslu výše citované judikatury, přičemž judikatura správních soudů v souladu s judikaturou ESLP a Ústavního soudu nepovažuje za důvod nepřezkoumatelnosti absenci podrobné reakce na každou vznesenou námitku.[14]

Pro shrnutí výše uvedeného lze odkázat na jednu z ustálených formulací, které ve svých rozhodnutích používá Nejvyšší správní soud: „*Má-li být rozhodnutí přezkoumatelné, musí z něj být zřejmé, jaký skutkový stav vzal* ***soud či správní orgán*** *za rozhodný a jak uvážil o pro věc podstatných skutečnostech, resp. jakým způsobem postupoval při posuzování těchto skutečností. Uvedené musí nalézt svůj odraz v odůvodnění rozhodnutí. Nepřezkoumatelné je takové rozhodnutí, které nelze přezkoumat pro nemožnost zjistit v něm jeho obsah nebo důvody. Nepřezkoumatelnost je třeba vykládat v souladu s jejím skutečným smyslem, neboť jde o jednu z nejzávažnějších vad řízení, kterou lze konstatovat pouze v případě, že skutečně nelze seznat, proč soud rozhodl tak, jak je uvedeno ve výroku jeho rozhodnutí, resp. proč považoval vznesené námitky za nedůvodné* […]. *Ani absence přímé reakce na každý jednotlivý argument účastníka řízení nepředstavuje nezákonnost či nepřezkoumatelnost, pokud soud prezentuje odlišný*

---

11 KOCOUREK, T.: *§ 76. Rozhodování bez nařízení jednání*. In: KÜHN, Z., KOCOUREK, T., MIKEŠ, P., DIENSTBIER, F., KADLEC, F. et al.: Soudní řád správní: komentář. Praha: Wolters Kluwer ČR, 201, s. 616.

12 *Ibid,* s. 617.

13 *Ibid.*

14 *Ibid,* s. 618.

názor, který přesvědčivě zdůvodní, a toto zdůvodnění poskytuje dostatečnou oporu výroku rozhodnutí […].“[15] Pro navazující části příspěvku je důležité povšimnout si formulací, které zde soud užívá. Hovoří zde o tom, že z rozhodnutí musí být zřejmé, jak správní orgán o něčem „uvážil“, jak „postupoval“, proč „považoval“ vznesené námitky za nedůvodné. Soud nepoužívá trpný rod, ale označuje správní orgán jako činitele nějaké rozhodovací činnosti, jejíž výstupy musí transparentně zohlednit myšlenkové procesy, které k nim vedly.

10.2 Algoritmické rozhodování

10.2.1 Problematika black box rozhodování

V uplynulých letech vzestup technologie velkých jazykových modelů (*large language models*, dále též „LLMs“) vrátil do popředí zájmu správní vědy a správního práva problematiku automatizovaného rozhodování (*automated decision making*, dále též „ADM“) prostřednictvím algoritmů. Vzhledem k tomu, že LLMs jsou ve svém jádru pouze sofistikované prediktivní algoritmy, nemohou se (zatím?) zbavit limitací, na které odborná literatura upozorňuje již od počátku širšího využívání ADM. Pro účely tohoto příspěvku se zaměříme zejm. na problematiku netransparentního rozhodování, přičemž tuto netransparentnost lze shledat jak na vstupu, tak i na výstupu. Tento komplex problémů s transparentností rozhodovacích algoritmů označujeme jako „black box“, neboli černá skříňka.[16]

Pro pochopení problematiky transparentnosti vstupu je nejdříve třeba pochopit, že prediktivní algoritmy vznikají tak, že jsou „učeny“ nebo „trénovány“ všímat si vzorců v datových sadách.[17] Pro kvalitu, resp. i pro užitečnost prediktivního algoritmu je tedy stěžejní kvalita dat, na kterých je trénován. Tento problém se poněkud nevzletně označuje jako „bordel dovnitř, bordel ven.“[18] Literatura často vyzdvihuje případ systému COMPAS, který po určitou dobu využívaly americké trestní soudy pro stanovování výše kauce. Investigativní novináři následně odhalili, že výstupy systému byly negativně předpojaté proti Afroameričanům a pozitivně předpojaté vůči bílým Američanům, a to z toho

---

15 Rozsudek Nejvyššího správního soudu ze dne 3. 12. 2025, čj. 9 As 137/2025-41, bod 12 (zvýraznění přidáno).

16 KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo*. Praha: Wolters Kluwer ČR, 2020, s. 107.

17 FINCK, M.: *Automated Decision-Making and Administrative Law*. In: CANE, P. HERWIG, C. H., HOFMAN, H. C. H., IP, E. C., LINDSETH, P. (eds.): The Oxford Handbook of Comparative Administrative Law. Oxford Handbooks. Oxford: Oxford University Press, 2021, s. 658.

18 *Ibid*, s. 665.

důvodu, že trénovací data algoritmu reflektovala existující rasové předsudky soudního rozhodování.[19]

Problémem pro transparentnost pak nemusí být jen skutečnost, že algoritmus rozhoduje na základě vadných dat, ale i vůbec samotné ověření jejich obsahu nebo existence. Vývoj ADM prostředků je nákladný a zpravidla mimo možnosti veřejných rozpočtů. Jejich vývoj je tedy často svěřen do rukou soukromého sektoru. Jak upozorňuje literatura, tyto soukromé subjekty v zájmu ochrany obchodního tajemství nemusí svým zákazníkům ve veřejném sektoru umožnit přístup k trénovacím datům, jakož ani k samotnému algoritmu.[20] Tito soukromí aktéři tak získávají podíl na rozhodování, zatímco subjekty, které jsou nadány originární a legitimní pravomocí k rozhodování, „*mohou přijít o schopnost vykonávat smysluplnou kontrolu*" nad rozhodováním.[21]

Pokud pak jde o transparentnost výstupů, se zvyšující se komplexností prediktivních algoritmů je spojena nižší přehlednost či čitelnost algoritmů, resp. jejich výstupů.[22] Sami vývojáři často upozorňují, že vývoj LLMs a jiných prediktivních algoritmů je značně specifický z toho důvodu, že u něj nelze aplikovat standardní metody průběžného nebo zpětného sledování procesů.[23]

10.2.2 Velké uvažující modely

Společnosti, které vyvíjejí LLMs, v nedávné době představily řadu modelů, které se označují jako velké uvažující modely (*large reasoning models*, dále též „LRMs"). Nejedná se zcela novou technologii, naopak jde spíše o vylepšení či zvýšení komplexnosti existujících LLMs. LRMs se od svých „neuvažujících" verzí liší tím, že produkují pro uživatele přístupný a čitelný řetězec úvah, který má simulovat lidské myšlení a zároveň poskytnout vhled do toho, jak model o problému „uvážil". Na první pohled se tedy zdá, že LRMs mohou být schopné rozhodovat způsobem, který je přezkoumatelný, neboť jejich inherentní součástí je produkce a zpřístupnění jakéhosi odůvodnění přijatého závěru.

Relevantní studie však poukazují na omezení i této technologie. Nedávná studie zveřejněná společností Apple se zaměřila na testování LRMs v tří

---

19 *Ibid*, s. 662.; KOLAŘÍKOVÁ, L.; HORÁK, F.: *Umělá inteligence & právo*. Praha: Wolters Kluwer ČR, 2020, s. 109.

20 FINCK, M.: *Automated Decision-Making and Administrative Law*. In: CANE, P. HERWIG, C. H., HOFMAN, H. C. H., IP, E. C., LINDSETH, P. (eds.): The Oxford Handbook of Comparative Administrative Law. Oxford Handbooks. Oxford: Oxford University Press, 2021, s. 665.

21 *Ibid.*

22 *Ibid*, s. 666.

23 KOLAŘÍKOVÁ, L.; HORÁK, F.: *Umělá inteligence & právo*. Praha: Wolters Kluwer ČR, 2020, s. 107.

režimové architektuře výkonnosti, kdy bylo testováno, jak si poradí standardní LLMs a pokročilejší LRMs při řešení problémů s nízkou složitostí, střední složitostí a vysokou složitostí. LRMs byly úspěšné pouze ve vztahu k problémům se střední složitostí. U problémů s nízkou složitostí je překonaly LLMs, přičemž u LRMs byla zjištěna tendence zbytečně komplikovaného přemýšlení, které zkreslovalo výsledky. U problémů s vysokou složitostí pak LRMs nebyly schopny efektivně alokovat výpočetní zdroje při inferenci, jinými slovy, nebyly schopny produkovat správné výsledky kvůli inherentním architektonickým omezením.[24] Nadto však tato studia došla k překvapivému závěru, že LRMs nezlepšily výkonnost ani za situace, kdy jim bylo poskytnuto celé řešení problému.[25] Uvedené dle mého názoru vypovídá o tom, že byť LRMs mohou produkovat něco podobající se odůvodnění, stále nelze hovořit obecném uvažování. Navazující studie pak zkoumala přesnost výsledků LRMs. Její výsledky ukazují, že v situacích, kdy je vyžadována vysoká přesnost a je přípustná pouze nízká míra falešné pozitivity (což je zcela nepochybně i oblast právního rozhodování), LRMs nejsou vhodné.[26]

Uvedené lze shrnout tak, že problém black box přetrvává i u LRMs, které jsou schopné produkovat text technicky vzato podobající se odůvodnění. Jak ukazují jejich nedostatky v případech nízké i vysoké komplexity řešených problémů, toto odůvodnění nezaručuje správné řešení a nevypovídá o tom, že se u těchto modelů jedná o skutečné obecné uvažování.

10.3 Deference jako krajní meze lidského prvku nepřezkoumatelnosti

10.3.1 Obecně k deferenčnímu přezkumu

Jak již bylo uvedeno výše, stanovit obecně platný vzorec pro posouzení, kdy je rozhodnutí přezkoumatelné, není možné. Přezkoumatelnost rozhodnutí, resp. požadavky na odůvodnění rozhodnutí se odvíjí od několika faktorů, z nichž lze za nejpodstatnější označit složitost právního problému, argumentace účastníků řízení nebo i specifické situace typu změny správní praxe či jiného před tím zaujatého názoru. V kontextu soudního přezkumu správních rozhodnutí však můžeme narazit i na situace, kdy odůvodnění rozhodnutí splňuje všechny

---

[24] SHOJAEE, P. et al.: *The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity*. In: ArXive, 2025, s. 8–9. Online: https://doi.org/10.48550/arXiv.2506.06941.

[25] *Ibid*, s. 11.

[26] CHEGINI, A., KAZEMI, H., SOUZA, G., SAFI, M., SONG, Y. et al.: *Reasoning's Razor: Reasoning Improves Accuracy but Can Hurt Recall at Critical Operating Points in Safety and Hallucination Detection*. In: ArXive, 2025, s. 11. Online: https://doi.org/10.48550/arXiv.2510.21049.

požadavky kladené na něj judikaturou, a přesto je jeho obsah pro soud z nějakého důvodu nečitelný. Stane se tak např. tehdy, kdy určitý orgán vykonává svou odbornou působnost na úzce vymezeném úseku veřejné správy, přičemž soud nedisponuje dostatečnými odbornými znalostmi na to, aby správnost přijatého závěru posoudil. V takovém případě hovoříme o deferenčním přezkumu. Bylo by tedy možné poukázat na to, že správní soudy se umí vypořádat ze situací, kdy je rozhodovací proces orgánu svým způsobem také „black box", neboť vzhledem k absenci odborné znalosti či originární zákonné pravomoci je pro soud takto učiněný závěr svým způsobem netransparentní. Domníváme se však, že takový závěr by byl chybný, neboť deferenční přezkum je na systémové úrovni umožněn sérií záruk (či spíše sérií domněnek o zárukách), které nelze v případě algoritmického rozhodování shledat.

Deferenci lze obecně označit jako druh soudního (sebe)omezení, kdy soud přizná určitou váhu rozhodnutí orgánu, který věc posoudil před soudem.[27] V kontextu České republiky jsou správní soudy zákonem povolány ke zdrženlivému přezkumu správního uvážení,[28] judikatura pak dovozuje nutnost zdrženlivosti při přezkumu některých aktů založených na odborných podkladech[29] či v případě rozhodování samosprávných korporací.[30]

Důvodů pro deferenční přístup může být celá řada.[31] Na tomto místě je vhodné vyzdvihnout zejm. deferenci epistemickou, která je v tomto kontextu ilustrativní. Označuje situaci, „*kdy se v konkrétním případě prokáže, že první rozhodující subjekt disponoval větší schopností správně posoudit dané záležitosti a vyšší odborností.*"[32] Příkladem epistemické deference je např. přístup Nejvyššího správního soudu k přezkumu opatření obecné povahy, která jsou založena na odborných podkladech: „*Nelze opomíjet, že soudní přezkum opatření obecné povahy, které je vystavěno na podkladech, východiscích a závěrech odborné*

---

27 KRATOCHVÍL, J.: *Důvody pro zdrženlivý přezkum Ústavním soudem*. In: Časopis pro právní vědu a praxi, roč. 30, č. 4, 2020, s. 807. Online: https://doi.org/10.5817/CPVP2022-4-6.

28 Pro stručnost na tomto místě odkazuji na svůj článek, kde jsem argumentoval, proč se v případě správního uvážení nejedná o vyloučení z přezkumu, ale o předepsanou deferenci: LIPKA, O.: *Veřejná správa mezi judicializací a deferencí*. In: Právník, roč. 163, č. 11, 2024, s. 1164.

29 *Ibid*, s. 1165.

30 KÜHN, Z., STAŠA, J.: *Deference to the Administration in Judicial Review in the Czech Republic*. In: ZHU, G. (ed.). Deference to the administration in judicial review: comparative perspectives. Cham: Springer, 2019, s. 148.

31 KRATOCHVÍL, J.: *Důvody pro zdrženlivý přezkum Ústavním soudem*. In: Časopis pro právní vědu a praxi, roč. 30, č. 4, 2020, s. 808–810. Online: https://doi.org/10.5817/CPVP2022-4-6.

32 *Ibid*, s. 809.

povahy, má svá specifika. Těmi se Nejvyšší správní soud podrobněji zabýval v judikatuře vztahující se sice k mimořádným opatřením […] *při epidemii onemocnění COVID-19* […], *avšak do určité míry jsou tyto závěry relevantní i pro přezkum jiných opatření obecné povahy, která se rovněž opírají o odborné podklady. Nejvyšší správní soud v těchto rozsudcích připomněl, že soudní moci nepřísluší věcně rozhodovat namísto správních orgánů, že soud se má zabývat tím, zda mimořádná opatření vycházejí ze ‚solidních vědeckých základů' a nejsou ‚svévolná či nesmyslná, přičemž nelze čekat na definitivní rozřešení určité vědecké otázky. Tato specifika soudního přezkumu mimořádných opatření, která jsou vystavěna začasté na odborných podkladech, východiscích a závěrech, nicméně nelze podle Nejvyššího správního soudu vykládat tak, že by soud neměl ověřovat, zda je odůvodnění napadeného opatření opřeno o relevantní zdroje, není vnitřně rozporné či nelogické apod. (a to samozřejmě i z hlediska odborného).*"[33]

Předpokladem pro uplatnění deferenčního přezkumu je tedy existence orgánu, na který se lze – lapidárně řečeno – spolehnout, pokud jde o znalost nebo jinou kvalitu, kterou soud nedisponuje. Ostatně, jak upozorňuje literatura, využití této odbornosti je předpokladem pro deferenční přezkum. Disponuje-li subjekt specifickou vlastností, na jejíž použití při řešení konkrétního případu rezignoval nebo není z jeho odůvodnění zřejmé, jak s ní naložil, deferenční přístup není na místě.[34]

10.3.2 Přezkoumatelnost jako výsledek lidského snažení

Domníváme se, že výše uvedené nelze vnímat tak, že soud potřebuje na straně správního orgánu spolehlivého partnera pouze v situacích, kdy je řešený případ mimo jeho odbornost. Příklad deference zmiňujeme jako situaci, kdy se vždy přítomný lidský aspekt přezkoumatelnosti dostává do popředí tím, že soudu nezbývá než se na něj spolehnout. Jinými slovy, deferenční přezkum nevyvrací, ale naopak podtrhuje vždy přítomný aspekt lidského rozhodování v úvahách o přezkoumatelnosti, který až doposud nebylo nutné explicitně popsat, neboť až doposud byl implicitně předpokládán jako její součást.

Shrneme-li výše uvedené závěry judikatury jak k běžné nepřezkoumatelnosti, tak k přezkumu odborného uvážení, vidíme, že soudy vyžadují, aby bylo vidět, jak ten který orgán o věci uvážil, jak ji posoudil a jestli jsou jeho závěry logické. (Správnost těchto závěrů je otázkou odlišnou, přičemž

---

[33] Rozsudek Nejvyššího správního soudu ze dne 24. 2. 2023, čj. 8 As 98/2020-235, č. 4464/2023 Sb. NSS, bod 46.

[34] KRATOCHVÍL, J.: *Důvody pro zdrženlivý přezkum Ústavním soudem*. In: Časopis pro právní vědu a praxi, roč. 30, č. 4, 2020, s. 811. Online: https://doi.org/10.5817/CPVP2022-4-6.

posouzení přezkoumatelnosti jí nutně předchází.) Uvedené je třeba konfrontovat s tím, že podle nyní dostupných informací o současném stupni technologického vývoje je vyloučené, abychom vůbec hovořili o tom, že algoritmus o něčem uvážil nebo něco posoudil.

Závěr

V příspěvku jsme se zaměřili na komplexní problematiku přezkoumatelnosti v kontextu soudního přezkumu správních aktů. Právní východiska byla konfrontována se současným stavem poznání o technologiích, které představují pro přezkoumatelnost a zásadu transparentnosti zcela zásadní výzvy.

Uvedená zjištění lze chápat více způsoby. Nepochybně se najdou tací, kteří např. s ohledem na předpokládaný vývoj LRMs nebudou mít problém (minimálně do budoucna) konstatovat, že algoritimické rozhodnutí naplňuje znaky přezkoumatelnosti. Najdou se ale i tací, podle kterých je zásadně vyloučeno, aby bylo takové rozhodnutí pro úplnou absenci lidského faktoru považováno za přezkoumatelné.[35]

Bez ohledu na uvedené dělení, je zcela nepochybné, že širší proliferace nástrojů algoritmického rozhodování bude vyžadovat nový právní rámec a veřejný dohled. Právní rámec je v obecné rovině jedním z definičních znaků rozhodování ve veřejné správě[36] a je obtížné si představit, aby se na rozhodování podílel nástroj, jehož role v rozhodovacím procesu není zákonem stanovena.

Má-li být nastíněno konkrétní řešení, autor tohoto příspěvku se domnívá, že algoritmické nástroje mohou mít budoucnost jako pomocníci při rešerši a přípravě podkladů pro rozhodnutí, ale nikdy ne jako „elektronické úřední osoby" samy rozhodující. Uvedené nástroje by pak měly podléhat veřejnoprávnímu dohledu. Lze si představit, že stát bude prostřednictvím různých veřejnoprávních certifikací garantovat vhodnost užití algoritmů pro určité úkony, stejně jako dnes vydává potvrzení o bezpečnostní způsobilosti nebo kybernetické bezpečnosti.

---

35 K nim se řadí i autor tohoto příspěvku.

36 SKULOVÁ, S.: *Rozhodování ve veřejné správě: některé správně vědní a správně právní aspekty*. Brno: Masarykova univerzita, 1995, s. 37.

Zoznam použitej literatúry a zdrojov

1. FINCK, M: *Automated Decision-Making and Administrative Law*. In: CANE, P., Herwig, C. H., HOFMAN, H. C. H., IP, E. C., LINDSETH, P. (eds.): *The Oxford Handbook of Comparative Administrative Law*. Oxford: Oxford University Press, 2021, s. 656–676.
2. CHEGINI, A., KAZEMI, H., SOUZA, G., SAFI, M., SONG, Y. et al.: *Reasoning's Razor: Reasoning Improves Accuracy but Can Hurt Recall at Critical Operating Points in Safety and Hallucination Detection*. In: *ArXive*, 2025, 27 s. Online: https://doi.org/10.48550/arXiv.2510.21049.
3. KMEC, J.: *Právo domáhat se svého práva (čl. 36)*. In: KÜHN, Z., KRATOCHVÍL, J., KMEC, J., KOSAŘ, D., ANTOŠ, M. et al.: Listina základních práv a svobod: velký komentář. Praha: Leges, 2022, s. 1244–1442.
4. KOCOUREK, T.: *§ 76. Rozhodování bez nařízení jednání*. In: KÜHN, Z., KOCOUREK, T., MIKEŠ, P., DIENSTBIER, F., KADLEC, F. et al.: Soudní řád správní: komentář. Praha: Wolters Kluwer ČR, 2019, s. 613–627.
5. KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo*. Praha: Wolters Kluwer ČR, 2020, 196 s..
6. KRATOCHVÍL, J.: *Důvody pro zdrženlivý přezkum Ústavním soudem*. In: *Časopis pro právní vědu a praxi*, roč. 30, č. 4, 2020, s. 805–827. Online: https://doi.org/10.5817/CPVP2022-4-6.
7. KÜHN, Z., STAŠA, J.: *Deference to the Administration in Judicial Review in the Czech Republic.* In: ZHU, G. (ed.): Deference to the administration in judicial review: comparative perspectives*.* Cham: Springer, 2019, s. 133–156.
8. LIPKA, O.: *Veřejná správa mezi judicializací a deferencí*. In: Právník, roč. 163, č. 11, 2024, s. 1153–1166.
9. POSPÍŠIL, I.: *Čl. 36. Právo na spravedlivý proces*. In: WAGNEROVÁ, E. et al.: Listina základních práv a svobod: komentář, Praha: Wolters Kluwer ČR, 2023, s. 759–832.
10. Rozsudek Nejvyššího správního soudu ze dne 3. 12. 2025, čj. 9 As 137/2025-41.
11. Rozsudek Nejvyššího správního soudu ze dne 24. 2. 2023, čj. 8 As 98/2020-235, č. 4464/2023 Sb. NSS.
12. Rozsudek Nejvyššího správního soudu ze dne 24. 9. 2003, čj. 7 A 547/2002-24.
13. Sdělení federálního ministerstva zahraničních věcí č. 209/1992 Sb., o sjednání Úmluvy o ochraně lidských práv a základních svobod a Protokolů na tuto Úmluvu navazujících.

14. SHOJAEE, P. et al.: *The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity*. In: *ArXive*, 2025, 13 s. Online: https://doi.org/10.48550/arXiv.2506.06941.
15. SCHABAS, W. A.: *The European Convention on Human Rights: a commentary*. Oxford: Oxford University Press, 2015, 1308 s.
16. SKULOVÁ, S.: *Rozhodování ve veřejné správě: některé správně vědní a správně právní aspekty*. Brno: Masarykova univerzita, 1995, 131 s.
17. Usnesení předsednictva České národní rady č. 2/1993 Sb., o vyhlášení Listiny základních práv a svobod jako součásti ústavního pořádku České republiky
18. Zákon č. 150/2002 Sb., soudní řád správní, ve znění pozdějších předpisů
19. Zákon č. 500/2004 Sb., správní řád, ve znění pozdějších předpisů

Informácie o autorovi

Mgr. Ondřej LIPKA
Masarykova Univerzita
Právnická fakulta
Katedra správní vědy a správního práva
Email: 468422@mail.muni.cz

11 Využitie AI pri dokumentovaní vojnových zločinov a porušení medzinárodného humanitárneho práva

Bartolomej Mižák

https://doi.org/10.33542/HRF-0541-2-11

Abstrakt

Článok analyzuje transformáciu dokumentovania ozbrojených konfliktov v 21. storočí, kde nárast digitálnych dát zo sociálnych sietí robí manuálne spracovanie dôkazov prakticky nemožným. Umelá inteligencia (AI) sa v tomto kontexte stáva nevyhnutnou technológiou, ktorá umožňuje automatizovanú identifikáciu, triedenie a analýzu potenciálnych dôkazov o vojnových zločinoch a porušeniach medzinárodného humanitárneho práva.

Úvod

Dynamický technologický vývoj v 21. storočí zásadne mení charakter ozbrojených konfliktov, ako aj spôsob ich dokumentovania a následného právneho posudzovania. Digitalizácia, masové rozšírenie mobilných telefónov a globálny dosah sociálnych sietí spôsobili, že moderné konflikty sú zachytávané v bezprecedentnom rozsahu a časovej bezprostrednosti. Tento jav vytvára nové možnosti pre odhaľovanie a dokazovanie vojnových zločinov a porušení medzinárodného humanitárneho práva, no zároveň kladie vysoké nároky na schopnosť tieto dáta efektívne spracovať, analyzovať a právne využiť. Tradičné metódy vyšetrovania sa v konfrontácii s enormným objemom digitálnych dôkazov ukazujú ako nedostatočné. V tomto kontexte nadobúda umelá inteligencia kľúčový význam ako nástroj umožňujúci automatizované triedenie, identifikáciu a analýzu relevantného obsahu. Jej využitie však nepredstavuje len technologickú inováciu, ale zároveň otvára zásadné právne otázky týkajúce sa spoľahlivosti, transparentnosti a prípustnosti takto získaných dôkazov v konaniach pred medzinárodnými súdnymi orgánmi. Cieľom príspevku je analyzovať úlohu umelej inteligencie pri dokumentovaní vojnových zločinov a porušení medzinárodného humanitárneho práva, poukázať na jej praktické využitie v rámci vyšetrovania, ako aj identifikovať hlavné právne a etické výzvy spojené s jej implementáciou.

11.1 AI v dobe moderných konfliktov

Ozbrojené konflikty 21. storočia predstavujú kvalitatívne odlišnú výzvu pre medzinárodné právo než konflikty minulosti, a to nielen z hľadiska zbraní či spôsobov vedenia vojny, ale predovšetkým v oblasti dokumentácie. Moderné

technológie, najmä sociálne siete a smartfóny, zásadne transformovali spôsob, akým sú ozbrojené konflikty zaznamenávané a sprostredkované verejnosti. Nikdy predtým neexistoval stav, v ktorom by sa vojnové zločiny a porušenia medzinárodného humanitárneho práva dokumentovali v takom rozsahu, s takým detailom a časovou bezprostrednosťou ako v súčasnosti.

Vojaci, civilisti aj novinári denne generujú obrovské množstvo obrazového a audiovizuálneho obsahu od fotografií a videí až po živé prenosy zo samotných bojových zón. Tento obsah často zachytáva správanie, ktoré by v minulosti zostalo nezdokumentované: plienenie civilného majetku, zlé zaobchádzanie so zajatcami, ponižovanie a mučenie osôb či násilie páchané na civilnom obyvateľstve. Nejde pritom o fenomén jedného konkrétneho konfliktu, ale o globálne sa opakujúci jav naprieč rôznymi regiónmi.

Objem obsahu dát dostupného na platformách ako Telegram, TikTok, Instagram, Facebook či Reddit je pre vyšetrovateľov prakticky nespracovateľný. Manuálne vyhľadávanie, triedenie, verifikácia a archivácia dôkazov je časovo náročná, náchylná na chyby a často limitovaná krátkou životnosťou obsahu, ktorý môže byť odstránený samotnými používateľmi alebo správcami platforiem ešte pred jeho zabezpečením. Príkladom môže byť Subreddit r/UkraineWarVideoReport[1] ktorý je zameraný na zhromažďovanie a zdieľanie videí a fotografií dokumentujúcich priebeh ozbrojeného konfliktu na Ukrajine. Obsahuje najmä zábery z bojových operácií, dronové snímky, následky útokov a ďalšie explicitné vizuálne materiály. Obsah je často publikovaný v reálnom čase, pričom ich autenticita a kontext si vyžadujú dodatočné overovanie.

V tomto kontexte sa AI a metódy strojového učenia prestávajú javiť ako doplnkový nástroj a nadobúdajú charakter nevyhnutnej infraštruktúry moderného vyšetrovania. Algoritmy sú schopné kontinuálne monitorovať rozsiahle dátové toky, automaticky identifikovať potenciálne relevantný obsah, analyzovať obrazové a zvukové stopy, rozpoznávať vzorce správania a spájať jednotlivé fragmenty informácií do ucelených celkov. Skutočnosť, že veľké technologické spoločnosti už dnes využívajú AI na moderovanie obsahu a detekciu porušení interných pravidiel, naznačuje, že podobné technológie môžu byť a s vysokou pravdepodobnosťou aj budú významné pri dokumentovaní porušení medzinárodného práva.

Využívanie AI v tejto oblasti zároveň generuje zásadné právne a procesné dilemy, ktoré sa dotýkajú samotných základov dokazovania v medzinárodnom trestnom práve. Dôkazy, ktorých spracovanie alebo analýza je založená na

[1] REDDIT: *r/UkraineWarVideoReport*. Online: https://www.reddit.com/r/UkraineWarVideoReport/.

algoritmoch, musia byť spoľahlivé a overiteľné, aby mohli obstáť v konaní pred medzinárodným trestným súdom prípadne medzinárodnými trestnými tribunálmi. Osobitne problematická je otázka transparentnosti rozhodovacích procesov algoritmov, alebo určenie zodpovednosti za výstupy systémov fungujúcich ako tzv. „čierne skrinky".

Vzniká tak potreba skúmať, či a do akej miery existuje adekvátna právna úprava, ktorá by umožňovala využívanie AI pri dokumentovaní vojnových zločinov a porušení medzinárodného humanitárneho práva.

11.2 Vyšetrovanie pomocou AI

OSINT je skratka pre Open-Source Intelligence teda spravodajstvo z otvorených zdrojov. Jedná sa o proces zhromažďovania, analyzovania a overovania informácií, ktoré sú verejne dostupné. Nepoužíva žiadne tajné metódy ani interné databázy všetko, čo sa spracúva, je voľne Online internete alebo v iných verejných zdrojoch. Hlavné zdroje sú Sociálne siete (Facebook, Instagram, X/Twitter, TikTok…), Diskusné fóra (Reddit, 4chan…), Videá a fotografie (YouTube, Telegram, satelitné snímky), Médiá a správy, Verejné registre, databázy a dokumenty, Geolokačné nástroje (Google Maps, Street View, OpenStreetMap).[2]

Tím Bellingcatu ukázal, ako vyšetroval údajné vojnové zločiny v etiópskom regióne Tigraj počas konfliktu (2020 - 2022), v ktorom etiópska vláda bojovala proti Tigray People's Liberation Front. Vyšetrovanie začalo v marci 2021, keď sa na sociálnych sieťach objavili videá zobrazujúce údajné popravy civilistov etiópskymi vojakmi. Tím analyzoval tiene vo videách, aby určil približný čas ich vzniku, a na určenie polohy využil PeakVisor - aplikáciu pôvodne určenú pre horolezcov. Tá poskytla detailné informácie o reliéfe, ktoré bolo možné porovnať s tvarmi kopcov a plošín zachytenými na záznamoch. Následne tieto údaje porovnali so satelitnými snímkami z Google Earth a určili miesto masakru - dedinu Mahbere Dego v regióne Tigraj. Na identifikáciu páchateľov skúmali jazyk hovorený vo videách a vojenské uniformy. Vďaka nezávislým prekladateľom zistili, že vojaci hovoria amharsky, čo naznačovalo, že nepochádzajú z Tigraja. To viedlo k záveru, že ide o príslušníkov etiópskej armády. V spolupráci s Newsy a BBC Africa Eye Bellingcat napokon potvrdil, že popravy vykonali etiópske ozbrojené sily.[3]

---

[2] IMSL WEB TEAM: *The Power of OSINT in War Crimes Investigations*. In: IntelMSL, 27.07.2023. Online: https://www.intelmsl.com/osint-war-crimes-investigations/.

[3] KUMAR, S.: *Fundamentals of Open-Source Intelligence for Journalists*. In: International Center for Journalists (ICFJ), 05.12.2024. Online: https://www.icfj.org/news/fundamentals-open-source-intelligence-journalists/.

Obrovské množstvo dát ktoré máme voľne prístupné je pomocou pri odhaľovaní porušovania práva. Otáznou však stále ostáva schopnosť spracovania dát. S týmto problém nám vie pomôcť AI, sofistikované modely dokážu rozoznávať nápisy, geolokáciu, analyzovať prostredie, rôzne paterny napríklad uniform. Dobrým príkladom strojového učenia je výsledok hackathonu ktorý bol organizovaný v júli 2025 v Kyjeve kde vyvinul tím IT inžinierov a dátových analytikov AI model schopný odhaľovať míny na záberoch z dronov.[4] Nástroje AI používajú aj ukrajinskí analytici na prehľadávanie online zdrojov prepojených s Ruskom (príspevky na sociálnych sieťach, blogy, fóra), filtrovanie relevantného obsahu a označovanie potenciálnych vojenských či konfliktom súvisiacich informácií. Pomáha zrýchliť získavanie spravodajských údajov z otvorených zdrojov.[5]

Prispôsobenie AI modelov môže pomôcť s analýzou dát. Netreba si to však predstavovať tak že AI bude vykonávať činnosti bez potreby kontroly, stále bude potrebná schopnosť kriticky analyzovať zhromaždené údaje. AI síce bude schopná za zlomok času zhromaždiť voľne dostupné informácie a selektovať relevantný obsah ale hodnotenie informácii bude stále úlohou ľudí. Potencionálny prínos je v schopnosti spracovať obrovské množstvo dát ako aj rýchlo sa prispôsobovať na základe strojového učenia.

11.3 Analýza obsahu pomocou AI

Súkromné spoločnosti ako Meta (Facebook a Instagram) a TikTok spracúvajú obrovské množstvo obsahu, ktorý každý deň pribúda na ich platformách. Aby mohli efektívne regulovať jednotlivé príspevky, spoliehajú sa stále viac na vlastné AI modely. Platformy ako Meta a TikTok používajú vlastné pokročilé modely strojového učenia, aby automaticky identifikovali obsah, ktorý porušuje pravidlá komunity alebo môže byť škodlivý pre používateľov. Ide predovšetkým o materiál zobrazujúci násilie, vojnové zločiny, porušenia humanitárneho práva, teroristický obsah či explicitnú brutalitu.

Meta implementuje modely nazývané napríklad Violence and Incitement AI classifier, ktoré analyzujú text, obrázky aj videá. Algoritmy dokážu rozpoznať krv, zbrane, popravy alebo scény z bojísk a vyhodnotiť, či ide o nebezpečný obsah. Pokročilé modely rozlišujú autentické zábery z konfliktov od filmových

---

4 UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP): *AI for Demining: Ukrainian Innovators Train Algorithms to Detect Explosives in Drone Images*. In: UNDP Ukraine, 30.07.2025. Online: https://www.undp.org/ukraine/press-releases/ai-demining-ukrainian-innovators-train-algorithms-detect-explosives-drone-images/.

5 MILITARYAI.AI: *Ukraine's AI Tool Scours Russian Military Clues Online*. Online: https://militaryai.ai/ukraine-tool-military-clues/.

scén alebo ; umožňujú presnejšie rekonštruovať sled udalostí, identifikovať zodpovedné osoby a preverovať tvrdenia jednotlivých aktérov konfliktu. Možnosť tieto dáta navzájom porovnávať a analyticky prepájať výrazne zvyšuje spoľahlivosť vyšetrovania a znižuje jeho závislosť od izolovaných alebo jednostranných zdrojov informácií. V minulosti bola dostupnosť takýchto dôkazných prostriedkov výrazne obmedzená. Vyšetrovatelia sa spoliehali najmä na priame pozorovania, oficiálne správy a výpovede svedkov, ktorých overiteľnosť bola často problematická a časovo oneskorená. Obrazové a priestorové dáta boli zriedkavé a prístup k nim mali prevažne štátne orgány alebo špecializované inštitúcie. Výsledkom bolo, že mnohé porušenia práva zostávali nedostatočne zdokumentované alebo sa potvrdzovali až s výrazným časovým odstupom. Súčasná situácia je diametrálne odlišná: dáta vznikajú v reálnom čase, vo veľkom množstve a z rozmanitých zdrojov, čo umožňuje nielen rýchlejšie zhromažďovanie informácií, ale aj ich podrobnejšiu analýzu a priebežné overovanie.

Tento posun však priniesol aj nový problém, objem digitálnych dôkazov dosahuje rozsah, ktorý je pre vyšetrovateľov prakticky nespracovateľný. Práve na túto výzvu reagoval Medzinárodný trestný súd spustením projektu Harmony, ktorý predstavuje zásadnú iniciatívu digitálnej transformácie Úradu prokuratúry. Projekt bol spustený v roku 2022 a do plnej prevádzky sa dostal v roku 2023 s cieľom prekonať systémovú disproporciu medzi tradičným, prevažne analógovým fungovaním súdu a realitou moderných konfliktov, v ktorých sa dôkazy produkujú vo forme tisícok hodín audiovizuálneho obsahu, satelitných snímok či zachytenej elektronickej komunikácie. Jadrom Projektu Harmony je využívanie AI a cloudových technológií na správu, analýzu a vyhodnocovanie dôkazov. Kľúčovým nástrojom je OTPLink. Jedná sa o online platformu vytvorenú Úradom prokuratúry Medzinárodného trestného súdu, ktorá slúži ako digitálny vstupný bod pre priame predkladanie dôkazov a informácií od externých zdrojov napríklad od obetí, svedkov, mimovládnych organizácií alebo vládnych subjektov. Táto platforma umožňuje bezpečné a rýchle nahrávanie dôkazného materiálu prostredníctvom webu alebo e-mailu.[6] S predkladaním dôkazov elektronicky počíta aj Rímsky štatút článok 68 ods.2 „senáty Súdu v záujme ochrany obetí môžu umožniť predkladanie dôkazov

[6] VASHUM, S.: *Artificial Sanctions? Potential Implications of US Sanctions on the ICC's Use of AI and Digital Evidence.* In: Opinio Juris, 25.02.2025. Online: https://opiniojuris.org/2025/02/25/artificial-sanctions-potential-implications-of-us-sanctions-on-the-iccs-use-of-ai-and-digital-evidence/.

elektronickými alebo inými špeciálnymi prostriedkami.[7]“ OTPLink je súčasťou širšej inovačnej iniciatívy Projektu Harmony, ktorá má posilniť schopnosti Úradu prokuratúry spracovávať rozsiahle objemy digitálnych dôkazov s využitím AI a strojového učenia.

Používanie AI výstupov v trestnom konaní pred ICC však nevyhnutne otvára otázku ich prípustnosti a dôkaznej hodnoty. Kľúčovým ustanovením je článok 69 ods. 4 Rímskeho štatútu, podľa ktorého „súd môže rozhodnúť o relevantnosti alebo prípustnosti všetkých dôkazov, pričom okrem iného prihliada na ich dôkaznú hodnotu a na to, či daný dôkaz nebude na ujmu spravodlivého procesu alebo spravodlivého vyhodnotenia svedeckej výpovede v súlade s Pravidlami súdneho konania a vykonávania dôkazov“[8].

Na základe vyššie spomenutého možno v kontexte využívania AI v dokazovaní konštatovať, že prípustnosť a relevantnosť digitálnych dôkazov generovaných alebo spracovaných algoritmami podlieha posúdeniu z hľadiska ich vzťahu k dokazovanej skutočnosti, ako aj z hľadiska ich vplyvu na zachovanie spravodlivého procesu. To v praxi predpokladá, že použitý algoritmus je objektívne spôsobilý identifikovať a analyzovať skutočnosti, ktoré sú predmetom dokazovania, a že jeho výstupy spĺňajú požiadavky spoľahlivosti a sú overiteľné.

11.4 Výzvy

V oblasti dokumentovania vojnových zločinov a porušení medzinárodného humanitárneho práva bude narastajúca prítomnosť AI-generovaného obsahu predstavovať jednu z najzásadnejších prekážok budúcnosti. Konfliktné strany môžu cielene generovať skrz AI fotografie, zmanipulované videá alebo nepravdivé „svedectvá“, ktoré budú navrhnuté tak, aby zmiatli systémy. Keďže kvalita AI generovaných vizuálnych ako aj audiovizuálnych materiálov rýchlo napreduje, AI nástroje, ktoré majú odhaľovať porušenia práva, sa môžu paradoxne ocitnúť pod útokom iných AI systémov. Výsledkom môže byť zahltenie informačného priestoru, komplikovanejšia verifikácia záznamov a zvýšené riziko, že falšované dôkazy sa dostanú do databáz, na ktorých sú trénované detekčné AI modely. Tento jav predstavuje významné riziko pri zhromažďovaní dôkazov o vojnových zločinoch či už ide o neziskové projekty, ktoré sledujú bombardovania civilnej infraštruktúry, medzinárodné vyšetrovacie tímy alebo Medzinárodný trestný súd. Preto môžeme očakávať, že

7 Oznámenie Ministerstva zahraničných vecí Slovenskej republiky č. 333/2002 Z. z. o prijatí Rímskeho štatútu Medzinárodného trestného súdu

8 *Ibid.*

aj pri najpokročilejších verifikačných algoritmoch zostane ľudský dohľad absolútne nevyhnutný. Expert musí vedieť identifikovať kontext, porovnávať záznamy s nezávislými zdrojmi a posúdiť, či daný materiál vyhodnotený AI skutočne odráža realitu. V dôsledku toho sa budú musieť organizácie zaoberajúce sa dokumentovaním porušení pripravovať na éru, v ktorej bude AI nielen nástrojom na ochranu, ale zároveň aj potenciálnym zdrojom nových hrozieb. Konzistentná ľudská kontrola, transparentné zdokumentovanie analytických postupov a kombinácia technických aj forenzných metód budú kľúčom k tomu, aby AI systémy mohli aj slúžiť ako spoľahlivý prostriedok pri zhromažďovaní a overovaní dôkazov.

Ďalšou z výziev, je konflikt medzi politikou sociálnych sietí zameranou na rýchle odstraňovanie škodlivého alebo násilného obsahu a potrebou takýchto materiálov ako potenciálnych dôkazov. Platformy ako Meta či TikTok používajú rozsiahle systémy strojového učenia na automatickú detekciu násilia, grafických scén alebo aktov porušovania humanitárneho práva. Tieto systémy sú navrhnuté predovšetkým s cieľom čo najrýchlejšie „vyčistiť" online priestor od obsahu, ktorý je považovaný za nebezpečný pre verejnosť, traumatizujúci alebo manipulatívny. Problém však nastáva vtedy, keď práve tento typ obsahu predstavuje dôkaz o spáchaných zločinoch. Videá, fotografie či záznamy získané z konfliktov ako bombardovanie civilných objektov, útoky na zdravotnícky personál, popravy či týranie sú okamžite identifikované ako porušenie pravidiel platformy a odstránené skôr, než ich môžu analytici alebo vyšetrovacie tímy uložiť a verifikovať. Takýto postup má za následok to že kľúčové dôkazy sa strácajú rýchlejšie, než ich možno zachytiť. Je to zásadná komplikácia, pretože mnohé zločiny sa dajú dokázať len vďaka existencii autentického vizuálneho materiálu. Tento konflikt medzi ochranou používateľov a uchovávaním digitálnych dôkazov je jednou z najväčších etických a dilem, ktoré prinesie budúcnosť.

Zavedenie Projektu Harmony predstavuje významný krok Medzinárodného trestného súdu k riešeniu rastúceho objemu digitálnych dôkazov, ktoré by tradičné metódy spracovania nedokázali efektívne zvládnuť. Online platforma OTPLink umožňuje bezpečné a priame predkladanie dôkazov od rôznych externých zdrojov, pričom plne rešpektuje ustanovenia Rímskeho štatútu týkajúce sa elektronického dokazovania.

Zároveň používanie AI a cloudových technológií pri spracovaní dôkazov otvára otázky ich prípustnosti a dôkaznej hodnoty. Podľa článku 69 ods. 4 Rímskeho štatútu musí súd pri hodnotení dôkazov zohľadniť ich relevantnosť, spoľahlivosť a vplyv na spravodlivý proces. V praxi to znamená, že algoritmy využívané na analýzu digitálnych dôkazov musia byť objektívne schopné

identifikovať fakty relevantné pre dokazovanie a generovať výstupy, ktoré sú overiteľné a spoľahlivé.

V dôsledku toho je zrejmé, že efektívne dokumentovanie vojnových zločinov a porušení medzinárodného humanitárneho práva v ére AI si bude vyžadovať nielen pokročilé analytické nástroje, ako sú systémy využívané v rámci Projektu Harmony a platforma OTPLink, ale aj dôslednú ľudskú kontrolu a schopnosť overovať dôkazy.

Záver

Analýza využitia AI pri dokumentovaní vojnových zločinov a porušení medzinárodného humanitárneho práva potvrdzuje, že ide o nevyhnutný nástroj reflektujúci realitu moderných ozbrojených konfliktov. Objem a komplexnosť digitálnych dôkazov generovaných v reálnom čase presahujú kapacity tradičných vyšetrovacích metód, pričom AI umožňuje efektívnejšie vyhľadávanie, triedenie a predbežnú analýzu relevantných informácií. Iniciatívy ako Projekt Harmony demonštrujú potenciál týchto technológií v aplikačnej praxi. Zároveň však využívanie AI prináša zásadné právne a procesné výzvy. Kľúčovou otázkou zostáva zabezpečenie spoľahlivosti, overiteľnosti a transparentnosti algoritmických výstupov tak, aby mohli obstáť ako dôkazné prostriedky v konaniach pred medzinárodnými súdmi. Osobitnú pozornosť si vyžaduje aj problematika tzv. „black box“ algoritmov, riziko manipulácie dôkazov prostredníctvom AI generovaného obsahu a konflikt medzi politikami moderácie obsahu a potrebou uchovávania dôkazného materiálu. Z uvedeného vyplýva, že budúcnosť dokumentovania vojnových zločinov nebude spočívať výlučne v technologických riešeniach, ale v ich synergii s odborným ľudským posúdením. Efektívne využitie AI si bude vyžadovať jasné právne rámce, štandardizáciu postupov, ako aj dôslednú kontrolu zo strany expertov. Len za týchto podmienok môže umelá inteligencia plniť svoju funkciu ako spoľahlivý nástroj na podporu spravodlivosti a zodpovednosti za najzávažnejšie porušenia medzinárodného práva.

Zoznam použitej literatúry a zdrojov

1. KUMAR, S.: *Fundamentals of Open-Source Intelligence for Journalists*. In: International Center for Journalists (ICFJ), 05.12.2024. Online: https://www.icfj.org/news/fundamentals-open-source-intelligence-journalists/.
2. META: *Meta's New AI System Tackles Harmful Content*, 08.12.2021. Online: https://about.fb.com/news/2021/12/metas-new-ai-system-tackles-harmful-content/.
3. MILITARYAI.AI: *Ukraine's AI Tool Scours Russian Military Clues Online*. Online: https://militaryai.ai/ukraine-tool-military-clues/.
4. MILMO, D.: *Meta Apologises over Flood of Gore, Violence and Dead Bodies on Instagram*. In: The Guardian, 28.02.2025. Online: https://www.theguardian.com/technology/2025/feb/28/meta-apologises-over-flood-of-gore-violence-and-dead-bodies-on-instagram.
5. REDDIT: *r/UkraineWarVideoReport*. Online: https://www.reddit.com/r/UkraineWarVideoReport/.
6. Oznámenie Ministerstva zahraničných vecí Slovenskej republiky č. 333/2002 Z. z. o prijatí Rímskeho štatútu Medzinárodného trestného súdu.
7. TIKTOK: *Content Moderation*. Online: https://www.tiktok.com/euonlinesafety/en/content-moderation/.
8. UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP): *AI for Demining: Ukrainian Innovators Train Algorithms to Detect Explosives in Drone Images*. In: UNDP Ukraine, 30.07.2025. Online: https://www.undp.org/ukraine/press-releases/ai-demining-ukrainian-innovators-train-algorithms-detect-explosives-drone-images/.
9. VASHUM, S.: *Artificial Sanctions? Potential Implications of US Sanctions on the ICC's Use of AI and Digital Evidence*. In: Opinio Juris, 25.02.2025. Online: https://opiniojuris.org/2025/02/25/artificial-sanctions-potential-implications-of-us-sanctions-on-the-iccs-use-of-ai-and-digital-evidence/.
10. IMSL WEB TEAM: *The Power of OSINT in War Crimes Investigations*. In: IntelMSL, 27.07.2023. Online: https://www.intelmsl.com/osint-war-crimes-investigations/.

Informácie o autorovi
Mgr. Bartolomej MIŽÁK
Univerzita Pavla Jozefa Šafárika v Košiciach
Právnická fakulta
Ústav medzinárodného práva a európskeho práva
Email: bartolomej.mizak@student.upjs.sk

12 Algoritmy ako prostriedok hybridných vojen zacielených na ovplyvňovanie volieb v cudzích štátoch

Ján Svák

https://doi.org/10.33542/HRF-0541-2-12

Abstrakt

Zásahy algoritmov do priebehu volieb sa šíria nielen kvantitatúrne ale aj kvalitatívne. Zásadným spôsobom začínajú ovplyvňovať výsledky volieb. Problémom sa stáva to, že tieto algoritmy teraz viac využívajú cudzie štáty a tým sa stáva právo na slobodné voľby predmetom rastúceho záujmu medzinárodného práva. Určitým vyústením je prvé rozhodnutie Európskeho súdu pre ľudské práva, ktoré riešilo aj otázku o pozitívnom záväzku štátu ochrániť voľby pred algoritmami nainfikovanými záujmami cudzích mocností.

Úvod

Medzinárodnoprávna komunita je už jednu celú dekádu konfrontovaná s reálnou či domnelou aktivitou cudzích mocností zameranou na ovplyvňovanie volieb v cudzích (nepriateľských?) štátoch. Orámcovaná je vnútroštátnym odznením prezidentských volieb v USA konaných v roku 2016 na začiatku a prezidentskými voľbami v Rumunsku v roku 2025 na jej aktuálnom konci. Vnútroštátne odznenie v USA sa však radikálne líši od toho ako relevantné vnútroštátne orgány reagovali v Rumunsku. Spoločné majú zatiaľ to, že nadnárodné decízne orgány zatiaľ na ne nereagovali.[1] Najväčší rozruch v čase medzi týmito voľbami spôsobili tri ľudové hlasovania v Spojenom kráľovstve, na ktoré už stihol zareagovať aj decíznym spôsobom Európsky súd pre ľudské práva („ESĽP“).

* Príspevok bol vytvorený v rámci projektu APVV-23-0509 Paradigmatická zmena v chápaní vojny a ozbrojených konfliktov z teologicko-historickej a právnej perspektívy

[1] Uvedené neznamená, že by si príslušné nadnárodné orgány nad tým prižmurovali oči. Len nedávno (november 2025) sa na Svetovom fóre pre demokraciu 2025 v Štrasburgu rozsiahlo diskutovalo o tom ako umelá inteligencia ovplyvňuje voľby. Rokovanie sa nieslo v duchu jeho názvu *„AI pri volebných urnách: hrozba alebo príležitosť pre demokratické voľby?“* a bolo organizované Parlamentným zhromaždením Rady Európy a Výborom Rady Európy pre umelú inteligenciu. Členka Parlamentnej aliancie za slobodné a spravodlivé voľby pani L. Castelová napríklad zdôraznila, že *„technológia musí slúžiť demokracii, nie ju ovládať. Občania musia mať možnosť voliť bez manipulácie a strachu“*. Pozri PARLAMENTNÉ ZHROMAŽDENIE RADY EURÓPY: *Umelá inteligencia pri volebných urnách: hrozba alebo príležitosť pre demokratické voľby?* 7.11.2025. Online: https://pace.coe.int/en/news/10095.

V prvej časti príspevku sa preto zameriavam na priblíženie udalostí spojených s algoritmami využívanými Ruskom na ovplyvnenie výsledkov volieb a ľudových hlasovaní v USA, Spojenom kráľovstve a Rumunsku ako súčasti jeho hybridnej vojny proti svojim nepriateľom. V druhej časti potom prinesiem reakciu medzinárodného práva na tieto udalosti smerujúcej k vytvoreniu pozitívneho záväzku štátu zabezpečiť konanie slobodných volieb vo svojich krajinách.

12.1 Priebeh hybridnej vojny Ruska proti USA, Spojenému kráľovstvu a Rumunsku s cieľom ovplyvniť voľby algoritmami

Prvý krát sa svetová verejnosť oboznámila s možnosťou využitia algoritmov pri ovplyvňovaní volieb cudzími mocnosťami, keď vzniklo podozrenie, že prezidentské voľby v USA v roku 2016 sa stali predmetom hybridnej vojny vedenej Ruskom proti USA. Ich cieľom bolo ovplyvniť prezidentské voľby medzi kandidátmi H. Clintonovou a D. Trumpom v prospech druhomenovaného, u ktorého sa predpokladal ústretovejší prístup k ruským záujmom. Podporovatelia neúspešnej kandidátky H. Clintonovej sa pokúšali spochybniť výsledky parlamentných volieb tvrdením, že boli zmanipulované aktivitou ruských tajných služieb a agentúr a osobitne ruskou hackerskou skupinou (GRU) pôsobiacou v rámci Ruskej vojenskej spravodajskej služby a ruskou masívnou kampaňou prostredníctvom tzv. Internet Research Agency podporujúcou D. Trumpa. K ovplyvňovaniu voličov sa ďalej využívala aj Agentúra pre internetový výskum (IRA), ktorá znevažovala kandidátku H. Clintonovú s cieľom vyvolať sociálne nepokojne a polarizáciu v americkej spoločnosti. Algoritmy sa vytvárali na základe falošných profilov, ale aj „hacknutých" e-mailových účtov Demokratickej strany a využívali sa okrem „klasických" dezinformácií aj dokumenty šírené z Wiki Leaks.[2] Prostredníctvom firmy *Cambridge Analytica* sa využívali dáta z Facebooku na mikrocielenie voličov a najmä tzv. Afroameričanov (predpokladaní odporcovia politiky D. Trumpa) odrádzali od zapojenia sa do volebného procesu.

Na základe iniciatívy vyše 130 členov Kongresu ako aj pod vplyvom tlaku médií bol v roku 2017 vymenovaný Robert Mueller za špeciálneho prokurátora.,

[2] Väčšina informácií pre účely tohto príspevku pochádza zo samotnej tzv. Muellerovej správy prístupnej v pôvodnom rozsahu (448 strán). MINISTERSTVO SPRAVODLIVOSTI USA, ÚRAD OSOBITNÉHO VYŠETROVATEĽA: *Správa o vyšetrovaní ruského zasahovania do prezidentských volieb v roku 2016* [tzv. Muellerova správa], Washington, D.C., marec 2019. Online: https://www.justice.gov/storage/report.pdf.

ktorého úlohou bolo vyšetriť možné ruské zasahovanie do prezidentských volieb v USA v roku 2016 a zistiť prípadné prepojenia medzi prostredím okolo D. Trumpa s ruskými predstaviteľmi.

Vyšetrovanie odhalilo rozsiahlu trestnú činnosť spojenú s prezidentskými voľbami (37 obžalôb a sedem dohôd o vine) a identifikovalo početné prepojenia medzi ruskou vládou a Trumpovou kampaňou, vrátane Trumpovho záujmu o zverejnenie dokumentov z WikiLeaks s cieľom poškodiť H. Clintonovú. Muellerova správa podrobne opisuje, akým spôsobom. D. Trump maril vyšetrovanie využívajúc pritom aj svoje aktuálne postavenie vo funkcii prezidenta USA. Mueller ukončil svoje vyšetrovanie a 22. marca 2019 poslal generálnemu prokurátorovi W. Barrovi 448 stranovú písomnú správu. Tento neskôr sám k nej pripojil svoj štvorstránkový komentár, ktorý zaslal do Kongresu s cieľom spochybniť závery vyšetrovania R. Muellera a vrhal D. Trumpa do lepšieho svetla.[3] Tým sa značne otupilo ostrie Muellerovej správy a napriek konsenzu, že ruská vláda vyvíjala aktivity na podkopanie dôvery vo volebnú infraštruktúru v USA nedošlo k spochybneniu výsledku volieb. Podnietilo to však predovšetkým demokratických poslancov k zvýšenej aktivite pri návrhoch novelizácie volebných zákonov, ktoré však neboli prijaté a tak súčasné volebné zákonodarstvo (ide o súbor právnych predpisov upravujúcich jednotlivé súčasti volebného procesu) nereaguje dostatočne na potreby ochrany svojich volieb, najmä pred útokmi algoritmov spojených s využívaním online dezinformácií.[4]

Správy "spoza mora" najčastejšie pod skratkou *Cambridge Analytica* spôsobili v Európe paniku a strach zo zapojenia sa Ruska do rozsiahleho a všadeprítomného zasahovania do demokratických volieb z Európe. Obavy vzbudzujú najmä možné zneužívania dezinformácií na útok proti demokratickým základom európskych štátov. Ďalej sú to úmyselné kybernetické útoky proti volebnej infraštruktúre, operácie „hack and leak" a využívanie

---

[3] Pozri štvorstranový list generálneho prokurátora Williama Barra adresovaný vedúcim predstaviteľom súdnych výborov Snemovne reprezentantov a Senátu o hlavných záveroch vyšetrovania Muellera (24. marca 2019). Online: www.justice.gov/ag/page/file/1147981/download.

[4] Prijímané sú len čiastkové zákony, ktoré v zásade neriešia problematiku využívania algoritmov vo volebnom procese. Príkladom takéhoto čiastkového zákona môže byť návrh zákona, ktorý predložil Výboru pre správu Snemovne reprezentantov dňa 3. januára 2025 pán Fitzpatrick k zasadnutiu HR 156. Jeho účelom je *„zabezpečiť integritu a bezpečnosť volieb zavedením jednotných požiadaviek na identifikáciu s fotografiou pre hlasovanie vo voľbách do federálnych funkcií a na iné účely."* Pozri SNEMOVŇA REPREZENTANTOV USA: *H.R. 156*, 119. kongres, 1. zasadnutie (2025).

„trolých fariem"[5] na manipuláciu s verejnosťou s cieľom vyvolať spoločenské nepokoje a polarizovanie spoločnosti.

Prvým európskym štátom, ktorý sa cítil ohrozený ruským útokom s využitím algoritmov, bolo celkom pochopiteľne Spojené kráľovstvo, v ktorom sa odohrali tri kľúčové udalosti údajne zasiahnuté algoritmami v rámci hybridnej vojny vedenej Ruskom. Išlo o

- referendum o nezávislosti Škótska v roku 2014,
- referendum o vystúpení Spojeného kráľovstva z členstva v Európskej únii v roku 2016,
- parlamentné voľby v roku 2019.

Predovšetkým referendá priamo formujúce suverenitu Spojeného kráľovstva považovali členovia britského parlamentu za možný cieľ kybernetického útoku zo strany Ruska. Preto bol poverený medzistranícky výbor poslancov Dolnej snemovne v podobe Výboru pre digitálne technológie, kultúru, médiá a šport ("DCMS") vyšetrovaním toho ako môžu byť politické rozhodnutia jednotlivcov ovplyvnené online informáciami zo strany „nepriateľských síl". Po osemnásťmesačnom vyšetrovaní zverejnil DCMS vo februári 2019 správu pod názvom "Zahraničný vplyv na politické kampane", v ktorej sa okrem iného konštatovalo: *"V zahraničných voľbách bol jasný a preukázaný ruský vplyv a dôkazy o takýchto pokusoch v referende o EÚ sme zdôraznili v našej predbežnej správe."*[6]

V priebehu roka 2018 vykonal rozsiahle vyšetrovanie aj parlamentný výbor pre spravodajské služby a bezpečnosť ("ISC"), ktorý v Správe s názvom „*Rusko*" predloženej parlamentu 21. júla 2020, uviedol: *"Šírenie dezinformácií (čím rozumieme propagáciu, zámerne nepravdivých, skresľujúcich alebo rušivých*

---

5 Bližšie MIRGA, T.: *Šírenie dezinformácií a organizované ovplyvňovanie verejnej mienky: Trolie farmy ako nástroj sociálnej a politickej manipulácie*. In: Itlib: Informačné technológie a knižnice. Online: http://doi.org/10.52036/1335793X.2023.3-4.42-55.

6 Konkrétne sa v tejto správe uvádzalo, že *„na Kremeľ naviazané médiá publikovali značný počet unikátnych článkov o referende o EÚ. Výskumníci z Up analyzovali najzdieľanejšie články a identifikovali 261 s jasnou protiúniovou zaujatosťou. Dve hlavné médiá boli RT a Sputnik s videom vytvoreným spoločnosťou Ruptly. Články, ktoré sa stali najviac virálnymi, mali najsilnejšiu protiúniovú zaujatosť. Sociálny dosah týchto protiúniových článkov publikovaných kremeľskými kanálmi bol 134 miliónov potenciálnych zobrazení v porovnaní s celkovým dosahom iba 33 miliónov a 11 miliónov potenciálnych zobrazení pre všetok obsah zdieľaný z webovej stránky Vote Leave a webovej stránky Leave.Eu. Hodnota porovnateľnej platenej kampane na sociálnych sieťach by sa pohybovala medzi 1, 4 a 4,14 miliónmi libier."* Zdroj: rozhodnutie ESĽP vo veci *Bradshaw a ostatní proti Spojenému kráľovstvu*, stažnosť č. 15653/22, rozsudok z 22. júla 2025, bod 40 rozsudku.

naratívov) a vedenie „vplyvových kampaní" sú samostatné, ale vzájomné prepojené témy. Napríklad vplyvová kampaň v súvislosti s voľbami môže využívať šírenie dezinformácií, ale môže zahŕňať aj iné taktiky, ako je nezákonné financovanie, narušanie volebných mechanizmov alebo priame útoky na jednu z kampaní (napríklad „hack and leak").[7]

Britská vláda reagovala na túto správu v júli 2020 a uznala, že je *„takmer isté, že ruskí aktéri sa snažili zasahovať do všeobecných volieb v roku 2019 prostredníctvom online šírenia nezákonne získaných a uniknutých vládnych dokumentov"*[8].

Napokon vo výročnej správe Národného centra kybernetickej bezpečnosti („NCSC") za rok 2023 sa v kapitole *„Ruské pokusy o manipuláciu demokratických inštitúcií"* hovorí, že

„Nie je žiadnym tajomstvom, že Rusko sa snaží oslabiť a rozdeliť svojich protivníkov zasahovaním do volieb pomocou dezinformácií, kybernetických útokov a iných metód. Vláda Spojeného kráľovstva odhaduje, že je takmer isté, že ruskí aktéri sa snažili zasahovať do všeobecných volieb v roku 2019. V nasledujúcich mesiacoch, s blížiacimi sa voľbami v Spojenom kráľovstve a USA, môžeme očakávať, že integrita našich systémov bude opäť preverená. Ochrana našich demokratických a volebných procesov pred zahraničným zasahovaním, či už zo strany Ruska alebo akéhokoľvek iného štátu, je a vždy bude absolútnou prioritou pre NCSC a my budeme naďalej podporovať kľúčovú prácu vlády v tejto oblasti.[9]*"*

Na tieto znepokojujúce správy a na rozdiel od USA reagoval britský zákonodarca zmenami vo svojom volebnom zákonodarstve. Novoprijatý zákon o voľbách z roku 2022 obmedzil výdavky tretích strán len na subjekty so sídlom v Spojenom kráľovstve. Zaviedol novú požiadavku, aby digitálne kampaňové materiály zobrazovali digitálnu tiráž s menom a adresou propagátora materiálu alebo akejkoľvek osoby, v mene ktorej bol materiál zverejnený.

Ďalej zákon o národnej bezpečnosti z roku 2023 výslovne kriminalizoval pomoc zahraničnej spravodajskej služby pri vykonávaní činnosti v Spojenom kráľovstve alebo v zámorí, ak takéto konanie poškodzovalo bezpečnosť a záujmy Spojeného kráľovstva. Zaviedol nový trestný čin sabotáže spáchaného prostredníctvom kybernetických útokov ako aj zaviedol nový trestný čin zahraničného zasahovania, ak malo konanie mať špecifický negatívny účinok a boli splnené určité podmienky. Zvýšil maximálne tresty odňatia slobody za určité

7 *Ibid*, bod 44 rozsudku.
8 *Ibid*, bod 53 rozsudku.
9 *Ibid*, bod 56 rozsudku.

trestné činy súvisiace s voľbami, ktoré boli vykonané pre zahraničnú mocnosť alebo v jej mene, alebo s úmyslom prospieť jej.

Napokon zákon o online bezpečnosti, tiež z roku 2023, zaviedol nový regulačný režim podriaďujúci technologické spoločnosti nezávislému regulačnému orgánu. Upravuje tiež dezinformácie a nepravdivé informácie, ak predstavujú nezákonný obsah alebo obsah škodlivý pre deti, a stanovil zoznam prioritných trestných činov odrážajúcich najvážnejší a najrozšírenejší nezákonný obsah a činnosť, proti ktorým museli spoločnosti prijať proaktívne opatrenia, pričom trestný čin zahraničného zasahovania zavedený zákonom o národnej bezpečnosti, bol považovaný za prioritný trestný čin.

Podobne ako v USA, tak aj v Spojenom kráľovstve si zachovala súdna moc svoj odstup od zasahovania do demokratických procesov stelesnených vo voľbách. Kým v USA ani len nedošlo k aktivácii súdov s cieľom spochybniť výsledky volieb z dôvodu ich ovplyvňovania cudzou mocnosťou pomocou algoritmov, tak najvyšší súd v Spojenom kráľovstve svoje zasahovanie do výsledkov volieb striktne odmietol. Aktivovali ho svojou žalobou traja poslanci parlamentu a argumentovali v nej vyššie uvedenými správami o zistenom zasahovaní Ruska do referend a parlamentných volieb v roku 2019.Najvyšším súdom bola ich žiadosť dňa 12. apríla 2021 zamietnutá z dôvodu, že z predmetných dokumentov nevyplývali žiadne konkrétne dôkazy o tom, že by nejaká ruská aktivita mala vplyv na demokratický priebeh volieb a platnosť hlasovania.

Žalobcovia podali nový návrh na začatie preskúmania volebného výsledku na základe argumentov vychádzajúcich z Dohovoru na ochranu ľudských práv a základných slobôd. („Dohovor"). V tomto návrhu tvrdili, že

- neexistoval žiadny právny subjekt s právnou zodpovednosťou predchádzať zahraničnému zasahovaniu do volieb a bojovať proti nemu;
- neexistovala v tom čase žiadna zákonná povinnosť, aby online politická reklama uvádzala svoj zdroj, kto za ňu zaplatil a krajinu pôvodu;
- neexistovala žiadna zákonná požiadavka, aby spoločnosti sociálnych médií spolupracovali s bezpečnostnými a spravodajskými službami, ak existovalo podozrenie, že nepriateľský zahraničný štát môže tajne viesť kampaň;
- neexistoval v tom čase žiadny zákaz zahraničných darov politickým stranám alebo volebným kampaniam;
- zahraniční štátni agenti ani iné osoby, ktoré zastupovali záujmy zahraničných mocností, nemali v tom čase povinnosť registrovať sa ako takí v Spojenom kráľovstve.

Sudca najvyššieho súdu považoval tieto sťažnosti za tvrdenie , že došlo k nedostatočnej legislatívnej činnosti, čo nespadá do rozsahu jeho pôsobnosti podľa § 6 zákona o ľudských právach z roku 1998. V každom prípade sudca uviedol, že sťažnosti sú nepodložené, pretože článok 3 Dodatkového protokolu č. 1 k Dohovoru nezakladá takéto konkrétne povinnosti.

Zdržanlivosť amerických a britských sudcov však nečakane prekonal súdny aktivizmus rumunského ústavného súdu[10], ktorý zrušil prvé kolo prezidentských volieb konaných dňa 24. novembra 2024. Kľúčovým dôkazným podkladom pre toto rozhodnutie boli dôverné informácie rumunskej Najvyššej rady národnej obrany („CSAT“) a direktoriátu pre vyšetrovanie organizovaného zločinu a terorizmu („DIICOT“). Dôvodom anulovania výsledkov prezidentských volieb bolo najmä to, že prezidentský kandidát C. Georgescu netransparentne využíval umelú inteligenciu a digitálne technológie porušujúce volebný kódex. Vzhľadom na semiprezidentskú formu vlády v Rumunsku takýto radikálny zásah zásadným spôsobom ovplyvnil fungovanie politického systému v ďalšom období. Svoj vplyv na rozhodnutie rumunského ústavného súdu má zrejme aj politický „backround“ týchto sudcov. Zaužívala sa tradícia, že do funkcií ústavných sudcov boli ustanovení aj bývalí politici s poslaneckou kariérou a aj súčasný predseda rumunského ústavného súdu je bývalým poslancom.[11]

Anulovaniu prezidentských volieb predchádzali aj ďalšie pozoruhodné rozhodnutia rumunských orgánov. Prvým[12] si rumunský ústavný súd pridelil ústavne implicitnú právomoc preskúmať ústavnosť celého volebného procesu, ktorá nastáva v prípade, keď dotknuté osoby nemajú prístup k iným účinným prostriedkom nápravy.

Druhým rozhodnutím rumunského ústavného súdu č.2 z 5. októbra 2024 vyhovel sťažnosti líderky krajnej pravicovej strany D. Şoşoacovej proti rozhodnutiu Ústrednej volebnej komisii, ktorá ju „vymazala“ z prezidentského hlasovacieho lístka z dôvodu jej verejných aktivít. Aj podľa sudcu Sçonteia v nesúhlasnom stanovisku týmto volebná komisia prekročila svoje notariálne právomoci pri preskúmavaní návrhov kandidátov na funkciu rumunského prezidenta. Úlohou komisie je len overovať naplnenie formálnych kritérií a nie hodnotiť vecnú spôsobilosť prezidentského kandidáta.

---

10 Bližšie MATUŠKA, P.: *Rumunský precedens: ústavnoprávne aspekty anulovania výsledku volieb v Slovenskej republike a suverenita volieb*. In: Právník, roč. 164, 11/2025, s. 1034–1054.

11 KUTI, C.: *On Means and Ends: The Annulment of the Romanian Presidential Elections*. Verfassungsblog, 10/12/2024. Online: https://verfassungsblog.de/on-means-and-ends/.

12 Rozhodnutie Ústavného súdu Rumunska č. 66 z 1. októbra 2019.

Tretím sporným rozhodnutím z 28. novembra 2024 požiadal rumunský ústavný súd Ústrednú volebnú komisiu o prepočítanie všetkých hlasov odovzdaných v prvom kole prezidentských volieb, pričom stanovil 24-hodinovú lehotu, ktorú komisia objektívne nemohla splniť. Následne stanovil druhý termín na ukončenie prepočítavania hlasov , a to na 2. decembra. Po tomto prepočítaní hlasov potom rumunský ústavný súd v ten istý deň potvrdil platnosť výsledkov prvého kola prezidentských volieb.

Štvrtým a definitívnym rozhodnutím vydaným pod č. 32 o štyri dni na to zrušil celý volebný proces a nariadil príslušným orgánom vyhlásiť nové prvé kolo prezidentských volieb. Stalo sa tak potom, čo sa oboznámil so správami rumunských tajných a bezpečnostných zložiek, ktoré boli odtajnené 4. decembra 2024.[13] Vo svojom rozhodnutí uviedol, že

- bol porušený čl. 81 rumunskej ústavy, ktorý prikazuje štátu, aby zabezpečil slobodný a spravodlivý priebeh prezidentských volieb,
- na základe manipulácie s platformami sociálnych médií sa znížili šance ostatných kandidátov, čo bolo v rozpore s rovnosťou volebného práva garantovanom v čl. 37 rumunskej ústavy;
- došlo k porušeniu zákona č. 370/2004, pretože Georgescuove videá nezobrazovali povinné značky volebného obsahu a nespĺňali požiadavky na zverejnenie financovania volebnej kampane.[14]

Podľa Cs. Kutiho toto rozhodnutie obsahuje viaceré rozporuplné skutočnosti, a to

- neexistenciu žiadneho ústavného alebo zákonného ustanovenia, ktoré by oprávňovalo rumunský ústavný súd začať konanie ex officio,
- neexistenciu žiadneho ústavného alebo zákonného ustanovenia, ktoré by upravovalo a umožňovalo anulovanie celého volebného procesu,
- zákon o prezidentských voľbách stanovuje trojdňovú lehotu od ukončenia hlasovania na podanie sťažnosti iniciujúcej prieskumné konanie rumunského ústavného súdu,

---

[13] Bližšie KUTI, C.: *On Means and Ends: The Annulment of the Romanian Presidential Elections*. Verfassungsblog, 10/12/2024. Online: https://verfassungsblog.de/on-means-and-ends/.

[14] Pozri MATUŠKA, P.: *Rumunský precedens: ústavnoprávne aspekty anulovania výsledku volieb v Slovenskej republike a suverenita volieb.* In*: Právník*, roč. 164, 11/2025, s. 1034–1054.

- občania si mohli uplatniť svoje volebné práva bez toho, aby vedeli, že sa zúčastňujú zmanipulovaného volebného procesu, zatiaľ čo oprávneným subjektom bolo znemožnené uplatniť si svoje právo podať sťažnosti na neplatnosť v zákonnej lehote, pretože príslušné informácie boli zverejnené až po uplynutí lehoty.

Všetky tri príklady z najnovšieho obdobia a každý z iného konca ukazujú na potrebu nielen jasnej a účinnej právnej úpravy využívania algoritmov vo volebnom procese, ale aj o potrebe jej nadnárodnej úpravy.

12.2 Medzinárodnoprávna ochrana pred algoritmami zasahujúcimi volebné právo

Problém využívania algoritmov na ovplyvňovanie voličov v cudzom štáte sa fakticky objavil ako „*škandál Cambridge Analytica*" pred desiatimi rokmi a preto vzhľadom na tempo tvorby medzinárodného práva ešte v ňom nezanechal výraznejšiu stopu. Vzhľadom na svoj výbušný obsah (možno ho zaradiť medzi prostriedky hybridnej vojny) nie je ani veľmi pravdepodobné, že sa v dohľadnom čase stane predmetom univerzálneho medzinárodného dohovoru. Napriek tomu možno vytipovať niektoré normy *soft law* univerzálneho medzinárodného práva už v súčasnosti. Bude zrejme aj naďalej viac vnútroštátnym problémom ako svojich občanov štát uchráni pred ovplyvňovaním ich politických rozhodnutí prostredníctvom algoritmov zavedených cudzími štátmi.

Bez nároku na vyčerpávajúci výpočet, a skôr nateraz len pre ilustráciu, možno medzi *soft law* pramene univerzálneho systému ochrany pred algoritmami ovplyvňujúcimi politické rozhodnutia občanov (najmä pri uplatňovaní ich volebného práva), zaradiť nasledujúce

- Medzinárodný dohovor o používaní vysielania vo veci mieru uzatvorený ešte 23. septembra 1936 v Ženeve (LNTS č.4319), v ktorom sa zmluvné strany zaviazali zakázať a zastaviť vysielanie na svojich územiach prenos (možno dopovedať súčasnou rečou „dát"), ktorý by bol na úkor dobrého medzinárodného porozumenia (okrem iného aj takých informácií, ktorých nesprávnosť bola alebo mala byť známa osobám zodpovedným za vysielanie, čo možno označiť súčasnou rečou ako zákaz šíriť dezinformácie) alebo ktorý by podnecoval obyvateľstvo k činom nezlúčiteľným s vnútorným poriadkom a bezpečnosťou zmluvnej strany;

- Spoločné vyhlásenie štyroch vysoko postavených subjektov Organizácie spojených národov[15] o slobode prejavu a „falošných správach“, dezinformáciách a propagande z marca 2017,
- Príručku o medzinárodných normách v oblasti ľudských práv vo voľbách z roku 2021, ktorá sa zaoberá aj vplyvom manipulácií s veľkými dátami a sociálnymi médiami zameraných na voľby[16],

---

[15] Išlo o osobitého spravodajcu Organizácie Spojených národov („OSN“) pre slobodu názoru a prejavu, zástupcu Organizácie pre bezpečnosť a spoluprácu v Európe („OBSE“) pre slobodu médií, osobitého spravodajcu Organizácie amerických štátov („OAŠ“) pre slobodu prejavu a osobitého spravodajcu Africkej komisie pre ľudské práva a práva národov („ACHPR“) pre slobodu prejavu a prístup k informáciám.

[16] Z pohľadu našej témy sú zaujímavé najmä tieto poznámky: *„V digitálnom veku sa objavili nové výzvy pre voľby, najmä informačný neporiadok a manipulácia s veľkými dátami. Zatiaľ čo ľudské právo na šírenie informácií a myšlienok sa neobmedzuje len na „správne“ vyhlásenia a chráni informácie a myšlienky, ktoré môžu šokovať, uraziť a rušiť, rozšírená dezinformácia môže predstavovať značné hrozby pre právo na politickú účasť, a to tak v súvislosti s právom zúčastňovať sa na verejných záležitostiach, ako aj kandidovať vo voľbách a voliť. Platformy sociálnych médií sa stali kľúčovými pre politických aktérov na šírenie dezinformácií, čomu často napomáha uprednostňovanie algoritmov kurátorstva obsahu pre senzačný obsah a možnosť mikrocielenia politickej reklamy. V kontexte volebných procesov môžu byť sociálne médiá zneužité na ovplyvňovanie výsledkov volieb diskreditáciou kandidátov a politických strán, poskytovaním nesprávnych informácií o procese hlasovania (potláčanie hlasov) a snahou ovplyvňovať volebné rozhodnutia konkrétnych segmentov spoločnosti, na ktoré sa môže zamerať pozornosť na základe vzorcov naznačených spracovaním osobných údajov a aktivitou na sociálnych sieťach.*
Dezinformácie môžu viesť k porušovaniu nevyhnutných práv na slobodné a skutočné voľby. Napríklad nenávistné prejavy a diskriminácia môžu byť zosilnené online dezinformačnými kampaňami a môžu následne viesť k rizikám pre bezpečnosť osôb a trestným činom z nenávisti. Sloboda prejavu a prístup k informáciám môžu byť ovplyvnené aj vtedy, ak má volič prístup k správam iba prostredníctvom jednej platformy sociálnych médií, ktorá obsahuje prevažne alebo iba dezinformácie. Šírenie dezinformácií môže viesť k zníženiu porozumenia medzi ľuďmi s rôznymi názormi alebo pôvodom a k prehĺbeniu polarizácie, zneužívaniu a skresľovaniu negatívnych názorov ľudí na iných. Môže sa použiť na rozdeľovanie a manipuláciu s verejnou diskusiou, čím sa voliči pripravujú o informácie dôležité pre ich rozhodovanie.
Hoci dezinformácie predstavujú skutočnú výzvu v kontexte volieb, štáty by sa mali zdržať všeobecného a nejednoznačného zákazu šírenia informácií, ako sú „klamstvá“ alebo „neobjektívne informácie“. Takéto pojmy dostatočne neopisujú zakázaný obsah. V dôsledku toho poskytujú orgánom široké právomoci cenzurovať vyjadrovanie nepopulárnych, kontroverzných alebo menšinových názorov, ako aj kritiku vlády a politikov v médiách a počas volebných kampaní. Odborníci na ľudské práva odporučili, aby sa prístupy k boju proti dezinformáciám vyhýbali kriminalizácii a namiesto toho boli založené na dôkazoch a prispôsobené preukázaným alebo zdokumentovaným dopadom dezinformácií a propagandy. Tieto opatrenia by mohli zahŕňať podporu nezávislých

- Rezolúciu Rady pre ľudské práva č. 49/21 z 1. apríla 2021 o Úlohe štátov v boji proti negatívnemu vplyvu dezinformácií na uplatňovanie a realizáciu ľudských práv[17].
- Správu osobitného spravodajcu OSN pre slobodu názoru a prejavu o presadzovaní a ochrane práva na slobodu názoru a prejavu z 13. apríla. 2021, v ktorej sa upozorňuje nielen na problematickosť dezinformácií, ale aj na problematickosť reakcií štátov na ne, ktoré podceňujú škodlivosť online dezinformácií a konajú s neznalosťou veci a nedostatku poznatkov a vedomostí.

Dôležitejšie z pohľadu Slovenska sú regionálne nadnárodné nástroje medzinárodného práva, ktoré sú prijímané na úrovni Rady Európy a Európskej únie. Tie, ktoré sú prijaté na úrovni Rady Európy, možno pre účely tohto príspevku rozdeliť na

- všeobecné nástroje Rady Európy v oblasti ovplyvňovania volieb,
- vychádzajúce z prvej časti príspevku rozoberaných konkrétnych prípadov ovplyvnenia volieb cudzou mocnosťou s využitím algoritmov, a to v súvislosti s parlamentnými voľbami v Spojenom kráľovstve v roku 2019 a s prezidentskými voľbami v Rumunsku v roku 2024.

---

mechanizmov overovania faktov, poskytovanie podpory nezávislým a rozmanitým verejnoprávnym médiám, verejné vzdelávanie a kampane za digitálnu gramotnosť.

17 Táto rezolúcia okrem iného

„1. *potvrdzuje* , že dezinformácie môžu negatívne ovplyvniť uplatňovanie a realizáciu všetkých ľudských práv a že štáty zohrávajú ústrednú úlohu v boji proti dezinformáciám;

2. *vyzýva* štáty, aby zabezpečili, že ich reakcie na šírenie dezinformácií budú v súlade s medzinárodným právom v oblasti ľudských práv a že ich úsilie v boji proti dezinformáciám bude podporovať, chrániť a rešpektovať slobodu prejavu jednotlivcov a slobodu vyhľadávať, prijímať a šíriť informácie, ako aj ďalšie ľudské práva;

3. *naliehavo vyzýva* štáty, aby uľahčili vytvorenie prostredia podporujúceho boj proti dezinformáciám prostredníctvom viacrozmerných a viacstranných reakcií, ktoré sú v súlade s medzinárodným právom v oblasti ľudských práv, a to aj prostredníctvom posilnenej spolupráce s medzinárodnými organizáciami, občianskou spoločnosťou, médiami, súkromným sektorom a ďalšími zainteresovanými stranami;

4. v*yzýva* štáty, aby nabádali podniky vrátane spoločností sociálnych médií k boju proti dezinformáciám a zároveň rešpektovali ľudské práva, a to aj prostredníctvom preskúmania obchodných modelov, najmä úlohy algoritmov a systémov hodnotenia pri zosilňovaní dezinformácií, zvyšovania transparentnosti, presadzovania všetkých platných právnych ochranných opatrení pre používateľov a podpory náležitej starostlivosti v súlade s hlavnými zásadami v oblasti podnikania a ľudských práv.“

K všeobecným nástrojom (z pohľadu prameňov medzinárodného práva ide o *soft law* medzinárodného práva) patria najmä

- Rezolúcia Parlamentného zhromaždenia Rady Európy číslo 2254 (2019) *„Sloboda médií ako podmienka demokratických volieb“*, ktorá zdôrazňuje, že voľby majú vtedy zmysel ak voliči v nich politicky rozhodujú so znalosťou veci a práve predmetná „znalosť“ sa stáva terčom hybridných útokov prostredníctvom algoritmov „nasýtených“ dezinformáciami[18],
- Rezolúcia Parlamentného zhromaždenia Rady Európy č. 2326(2020) *„Napadnutá demokracia? Ako reagovať?“*, v ktorej sa odrazilo znepokojenie nad šírením dezinformačných kampaní zameraných na formovanie verejnej mienky a trendov zahraničného zasahovania a manipulácie volieb[19]

---

[18] V rezolúcii sú členské štáty Rady Európy vyzývané najmä k tomu, aby

- sa zdržali šírenia alebo podnecovania šírenia vyhlásení, oznámení alebo správ na internete, o ktorých vedia alebo o ktorých sa dá rozumne očakávať, že vedia, že sú dezinformáciami alebo neprimeranou propagandou;

- vypracovali špecifické regulačné rámce pre internetový obsah v čase volieb a zahrnúť do týchto rámcov ustanovenia o transparentnosti v súvislosti so sponzorovaným obsahom zverejneným na sociálnych sieťach, aby verejnosť mohla vedieť o zdroji, ktorý financuje volebnú reklamu alebo akékoľvek iné informácie či názory;

- zaviedli jasnú právnu zodpovednosť pre spoločnosti sociálnych médií, ktoré zverejňujú nelegálny obsah škodlivý pre kandidátov alebo porušujú základné pravidlá mediálnej komunikácie počas volieb;

- zabezpečili, aby sa sankcie stanovené v súvislosti s nezákonným obsahom nepoužívali na vynucovanie autocenzúry názorov a kritických postojov oponentov, a obmedziť uplatňovanie extrémnych opatrení, ako je blokovanie celých webových stránok, IP adries, portov alebo sieťových protokolov, na najzávažnejšie prípady, v plnom súlade s prísnymi podmienkami stanovenými v článku 10 Európskeho dohovoru o ľudských právach;

- poskytovali špecifické školenia pre volebné riadiace orgány a regulačné orgány médií, aby ich členovia mohli lepšie pochopiť nové mediálne prostredie s cieľom zlepšiť implementáciu predpisov týkajúcich sa politickej komunikácie prostredníctvom sociálnych médií;

- povzbudzovali všetky zainteresované strany – vrátane internetových sprostredkovateľov, médií, občianskej spoločnosti a akademickej obce – aby rozvíjali participatívne iniciatívy, ktoré umožnia širokej verejnosti lepšie pochopiť nebezpečenstvo dezinformácií a neprimeranej propagandy na internete a aby spoločne hľadali vhodné reakcie na tieto javy.“

[19] Podľa tejto rezolúcie musia vlády členských štátov Rady Európy s cieľom riešiť problémy s dezinformáciami v kontexte demokratických volieb:

- Správa Benátskej komisie (Európska komisia pre demokraciu prostredníctvom práva ako expertný poradný orgán Rady Európy) *„Dopad informačného chaosu (dezinformácií) na voľby“* z novembra 2018, ktorá okrem iného uvádza, že *“nová éra informačného chaosu (Správa Rady Európy o informačnom chaose, 2017) deformovala komunikačný ekosystém do takej miery, že voliči môžu byť pri svojich rozhodnutiach vážne zaťažení zavádzajúcimi manipulatívnymi a nepravdivými informáciami, ktoré majú ovplyvniť ich hlasovanie. Toto prostredie potenciálne podkopáva uplatňovanie práva na slobodné voľby a vytvára značné riziká pre fungovanie demokratického systému.”*
- *Kódex osvedčených postupov vo volebných záležitostiach* –CDL-Ad(2002) 023 prijatý Benátskou komisiou ešte v roku 2002 a ktorý je pravidelne revidovaný[20],
- Správa Rady Európy DGI (2017) 09 *„Informačná porucha: Smerom k interdisciplinárnemu rámcu pre výskum a tvorbu politík“,* ktorý rozlišuje tri rôzne typy informačnej poruchy, a to dezinformáciu, nepravdivú informáciu a zavádzajúcu informáciu, pričom tieto jednotlivé typy dezinformácií vecne definuje a správa obsahuje spolu 34 odporúčaní[21],

1. Uznať nadnárodnú povahu problému a posilniť spoluprácu – s internetovými sprostredkovateľmi a prevádzkovateľmi sociálnych médií;
2. Umožniť voličom prijímať dôveryhodné informácie a stať sa informovanejšími a angažovanejšími;
3. Rozbiť monopol technologických spoločností, ktoré do značnej miery kontrolujú prístup občanov k informáciám a údajom; a
4. Zvážiť aktualizáciu vnútroštátnych právnych predpisov s cieľom účinnejšie bojovať proti dezinformačným kampaniam.

[20] Aktuálne znenie je prístupné na CDL-AD(2002) 023 rev 2-cor. Európska komisia pre demokraciu prostredníctvom práva (Benátska komisia): *Kódex osvedčených postupov vo volebných záležitostiach: usmernenia a vysvetľujúca správa*. CDL-AD(2002)023rev2-cor, Benátky, 30. október 2002. Online: https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2002)023rev2-cor-e.

[21] Okrem iných k nim patria nasledovné:

„1. *Zadať výskum na zmapovanie informačnej poruchy.* Národné vlády by mali zadať výskumné štúdie na preskúmanie informačnej poruchy v rámci svojich krajín s využitím koncepčnej mapy uvedenej v tejto správe. Aké typy informačnej poruchy sú najbežnejšie? Ktoré platformy sú primárnymi nástrojmi na šírenie informácií? Aký výskum sa uskutočnil a ktorý skúma reakcie publika na tento typ obsahu v konkrétnych krajinách? Metodika by mala byť v týchto výskumných štúdiách konzistentná, aby bolo možné presne porovnávať rôzne krajiny.“

2. *Regulujte reklamné siete* . Zatiaľ čo platformy podnikajú kroky na zabránenie zarábaniu peňazí vymyslenými „spravodajskými“ stránkami, iné siete zasahujú,

- Správa Rady Európy *„Dezinformácie a volebné kampane“* z júna 2019, ktorá prináša užitočný prehľad právnych úprav vo viacerých členských štátoch, ktoré sumarizuje a prináša odporúčania pre vnútroštátne právne úpravy volebných procesov[22],
- Správa a závery Bernskej konferencie konanej pod záštitou Parlamentného zhromaždenia Rady Európy v dňoch 9. – 10. mája 2023 *„Voľby v čase krízy: výzvy a príležitosti*“, ktorá sa podrobne venuje aj vplyvu zahraničného zasahovania umelej inteligencie a informačných a komunikačných technológií, komunikačných technológií na voľby. Podľa nej zavedenie umelej inteligencie prinieslo nové formy manipulácie vo volebnom procese a ten ešte nie je dostatočne regulovaný. Obyvateľstvo preto nevie,

aby vyplnili túto medzeru. Štáty by mali vypracovať nariadenia, ktoré zabránia zobrazovaniu akejkoľvek reklamy na týchto stránkach.

3. *Požadovať transparentnosť v súvislosti s reklamami na Facebooku***.** V súčasnosti neexistuje *žiadny dohľad nad tým, kto nakupuje reklamy na Facebooku*, aké reklamy nakupujú a na ktorých používateľov sú zamerané. Národné vlády by mali požadovať transparentnosť v súvislosti s týmito reklamami, aby kupujúci reklamy a Facebook mohli niesť zodpovednosť.
4. *Podporujte verejnoprávne médiá a miestne spravodajské agentúry***.** Finančné zaťaženie spravodajských organizácií v posledných rokoch viedlo v niektorých oblastiach k „spravodajským púšťam“. Ak to s redukciou vplyvu informačného neporiadku myslíme vážne, prioritou musí byť podpora kvalitných novinárskych iniciatív na miestnej, regionálnej a národnej úrovni.
5. *Zaviesť pokročilé školenia v oblasti kybernetickej bezpečnosti*. Mnohé vládne inštitúcie používajú počítačové systémy šité na mieru, ktoré sa neuveriteľne ľahko hackujú, čo umožňuje krádež údajov a generovanie nepravdivých informácií. Školenia by mali byť online na všetkých úrovniach vlády, aby sa zabezpečilo, že každý rozumie najlepším postupom v oblasti digitálnej bezpečnosti a aby sa predišlo pokusom o hacking a phishing.
6. *Vynucovať minimálne úrovne verejnoprávnych správ na platformách*. Povzbudzovať platformy k spolupráci s nezávislými verejnoprávnymi médiami s cieľom integrovať kvalitné správy a analýzy do používateľských kanálov.“

[22] V správe sa uvádza niekoľko odporúčaní, ktoré sa delia do troch kategórií:

1. predpisy v oblasti digitálneho práva, ktoré by vyžadovali transparentnosť od poskytovateľov služieb, pokiaľ ide o ich činnosti a ochranu osobných údajov,
2. predpisy volebného práva vrátane dlhších volebných kampaní, transparentnosti a finančných zdrojov, poskytovateľov a zákazu volebných výdavkov na digitálne aktivity zo strany zahraničnej, právnickej alebo fyzickej osoby a
3. osvedčené postupy vrátane opatrení zameraných na overovanie faktov, spoluprácu so všetkými zainteresovanými stranami, etiku, rozvoj programov, gramotnosti a samoreguláciu poskytovateľov služieb, pričom všetky tieto opatrenia podporujú kvalitnú žurnalistiku.

ako sa s tým vysporiadať alebo komu dôverovať. Je potrebné nájsť nové spôsoby, ako

- zaviesť transparentnosť zdrojov, údajov a algoritmov,
- pomôcť ľuďom rozpoznať, kedy sú vystavení obsahu generovanému umelou inteligenciou,
- zachovať si konečnú ľudskú kontrolu pri prijímaní základných rozhodnutí a
- zabezpečiť súlad systémov riadených umelou inteligenciou s volebnými/ demokratickými/ ľudskými právami a normami."

Ďalšie dva pramene regionálneho európskeho medzinárodného práva už konkrétne reagujú na

- prezidentské voľby v Rumunsku v roku 2024,
- parlamentné voľby v Spojenom kráľovstve v roku 2019.

Prvý prameň regionálneho európskeho medzinárodného práva má charakter *soft law* a jeho autorom je Benátska komisia. Svoju správu vydanú 27. januára 2025 vypracovala na základe „objednávky" predsedu Parlamentného zhromaždenia Rady Európy, ktorý ju požiadal o zodpovedanie otázky: „Za akých podmienok a na základe akých právnych noriem môže ústavný súd vyhlásiť voľby za neplatné vychádzajúc z nedávneho rumunského prípadu?"

Podľa názoru Benátskej komisie „vonkajší vplyv" - zo strany mimovládnych organizácií, médií (osobitne sociálnych médií), najmä tých, ktoré sú sponzorované a financované zo zahraničia a zahraničných, štátnych a neštátnych aktérov - nie je menej škodlivý a mohol by mať rovnaké (alebo dokonca závažnejšie) následky ako porušenie volebných pravidiel kandidátmi, politickými stranami a štátnymi úradníkmi. Preto je novou výzvou pre vnútroštátnu legislatívu a volebnú prax, aby sa dostatočne venovala otázkam online kampaní a dezinformáciám.[23]

[23] K tomu Benátska komisia uviedla nasledovné: „v porovnaní s tradičným vysielaním a tlačou sa sociálne médiá voľne pohybujú cez hranice a vo väčšine krajín nie sú sociálne médiá a online kampane v kontexte volieb regulované v takej miere ako sú tradičné médiá a tradičné kampane. Liberálny charakter sociálnych médií však, neznamená, že sú mimo vnútroštátnej regulácie a presadzovania v kontexte volieb. Rastúci význam online kampaní - vrátane využívania umelej inteligencie, ktorá má potenciál zosilniť účinok dezinformácií a manipulácie s verejnou mienkou - prináša nové výzvy v súvislosti s 1) kampaňou, propagandou, dezinformáciami a obsahom kampaňových posolstiev, pravidlami financovania kampaní a transparentnosti ... vrátane obmedzenia týkajúceho sa príspevkov z anonymných a zahraničných zdrojov a zneužívania administratívnych zdrojov. Z právneho hľadiska je preto dôležité rozlišovať medzi týmito dvomi záležitosťami". Pozri naliehavú správu o zrušení

Na základe toho Benátska komisia vyzýva štáty, aby okrem iného *regulovali* , že online volebná reklama musí byť vždy identifikovaná ako taká a musí byť transparentná, pokiaľ ide o identitu jej sponzora a použitú techniku šírenia. Ďalej, že financovanie online aktivít musí byť transparentné s potenciálnymi obmedzeniami výdavkov politických strán na digitálnu reklamu a že platformy sociálnych médií sú povinné dôsledne zverejňovať údaje o politickej reklame a svojich sponzoroch. Podľa interpretačného vyhlásenia môže byť zavedený zákaz určitých foriem platenej politickej reklamy na sociálnych médiách počas volebného obdobia, najmä ak sa používajú automatizované techniky hromadného šírenia alebo techniky mikrocielenia založené na umelej inteligencii. Rovnako je možné zakázať politickým stranám a kandidátom viesť anonymnú kampaň.

Upozornila tiež na to, že samotný fakt, že kandidát je úspešný v online kampani a že používanie platforiem sociálnych médií môže zosilniť posolstvo kandidáta nad rámec toho, čo bolo možné v tlačených a vysielaných médiách, neznamená, že kandidát porušil pravidlá týkajúce sa výdavkov na kampaň a transparentnosti a tým získal nespravodlivú výhodu.

Benátska komisia chápe, že preukázanie porušení zákona prostredníctvom kampaní online a prostredníctvom sociálnych médií je obzvlášť náročné. Dobre odôvodnené a transparentné rozhodnutia v takýchto veciach sú kľúčové. Podľa názoru Benátskej komisie by takéto rozhodnutia mali presne uvádzať porušenia a dôkazy a nesmú byť založené výlučne na utajovaných spravodajských informáciách (ktoré sa môžu použiť len ako kontextové informácie), pretože by to nezaručilo potrebnú transparentnosť a overiteľnosť. K tomu len dodávam, že konanie o preskúmanie ústavnosti volieb ústavným súdom by odporovalo aj princípu kontradiktórnosti, ktorý je založený na poznaní dôkazných prostriedkov protistrany a možnosti sa k nim vyjadriť. Všetko toto pri rozhodnutí rumunského ústavného súdu podľa môjho názoru absentovalo, čo ešte viac zvyšuje nebezpečenstvo súdneho aktivizmu pri *ex offo* zasahovaní do demokratických procesov v štátoch, kde sú voľby ich neoddeliteľnou zložkou.

Druhý prameň európskeho regionálneho práva regulujúceho ovplyvňovanie volieb využitím algoritmov zo strany cudzích mocností má už charakter regionálnej medzinárodnoprávnej obyčaje. Ide o rozhodnutie Európskeho súdu pre ľudské práva ESĽP v prípade *Bradshaw a ostatní proti Spojenému kráľovstvu (z 22. júla 2025, č. 15653/22)*, ktoré sa zaoberalo otázkou,

výsledkov volieb ústavnými súdmi CDL-AD(2025)003 schválenú Benátskou komisiou na jej 142. plenárnom zasadnutí konanom v dňoch 14. až 15. marca 2025. Online: https://www.coe.int/en/web/venice-commission/-/cdl-ad-2025-003-e.

či čl. 3 Dodatkového protokolu č. 1 k Dohovoru o ochrane ľudských práv a základných slobôd („čl. 3/1DP1 Dohovoru“) vyžaduje od štátu, aby prijal pozitívne opatrenia na ochranu integrity svojich volebných procesov v prípade reálneho rizika, že zasahovanie zo strany nepriateľského štátu by narušilo samotnú podstatu volebného práva voličov v zmysle čl. 3/1DP1 Dohovoru. Vzhľadom k tomu, že som podal komentár k tomuto rozhodnutiu na inom mieste[24], obmedzím sa len na stručné posolstvo predmetného rozhodnutia.

Pre uvedenie do deja uvádzam, že rozhodnutie má základ v sťažnosti troch aktuálnych aj bývalých poslancov britského parlamentu, ktorí namietali, že príslušné britské orgány neviedli vyšetrovanie zasahovania nepriateľského štátu do vnútroštátnych demokratických volieb. Teda kľúčovou otázkou bolo vymedzenie rozsahu pozitívnej povinnosti štátu, ktorá mu vyplýva z čl. 3/1DP1 Dohovoru. Išlo vôbec o prvé rozhodnutie ESĽP, ktoré malo dať odpoveď na túto otázku, aj keď znenie čl. 3/1DP1 Dohovoru je svojou formuláciou ako jediné, ktoré vytvára priamo pozitívny záväzok štátu. Doterajšia prax ESĽP sa však paradoxne zaoberala takmer výlučne ochranou aktívneho či pasívneho volebného práva, ktoré ESĽP objavil v implicitnom obsahu tohoto ustanovenia.

V prípade *Bradshaw* išlo navyše o dve úplne nové skutočnosti, a to

- slovami Benátskej komisie o *„novú éru informačného neporiadku“*, kde sa dezinformácie a na nich založené algoritmy stávajú zbraňou na zasahovanie do demokratických volieb, pričom sa využívajú nové technológie, ako napr. platformy sociálnych sietí,
- že subjektom ovplyvňujúcim voľby prostredníctvom nových technológií využívajúcich okrem iného aj algoritmy bola cudzia veľmoc.

Napokon ESĽP nevyvodil z 3/1DP1 Dohovoru implicitnú povinnosť štátu viesť vyšetrovanie podozrení z ovplyvňovania volebnej kampane. V bode 138 rozsudku však uviedol, že ak by štát ignoroval dôveryhodné obvinenia zo zahraničného zasahovania do svojich volieb, nebol by v komfortnej zóne a uviedol: *„hoci štáty nemusia mať samostatnú a autonómnu povinnosť vyšetrovať opodstatnené tvrdenia o porušení práv jednotlivca podľa článku 3 Protokolu č. 1, zjavné zlyhanie štátu pri vyšetrovaní dôveryhodných tvrdení o zasahovaní do jeho volieb by mohlo vyvolať problém podľa tohto článku, ak by bránilo jeho schopnosti prijať pozitívne opatrenia na ochranu voličov pred narušením podstaty ich práva profitovať z volieb za podmienok, ktoré zabezpečujú slobodné vyjadrenie názoru ľudu.“* Teda aj tu naviazal povinnosť vyšetrovať možné zásahy cudzích mocností

[24] Pozri SVÁK, J.: *Ovplyvňovanie volieb cudzími mocnosťami*. In: Justičná revue, č. 10/2025, s. 1033–1045.

do volieb na prípady, keď by rozsah tohoto zasahovania bol spôsobilý narušiť samotnú podstatu volebného práva.

Najväčším problémom v tomto prípade bolo to, že sťažovatelia vychádzali len zo zverejnených neutajených častí správ spravodajských a bezpečnostných služieb, ako bolo uvedené v prvej časti príspevku.

V ďalšom sa už potom ESĽP venoval legislatívnej reakcii Spojeného kráľovstva na tieto zverejnené informácie o zasahovaní Ruska do ovplyvňovania voličov, čo bolo napokon aj podstatou preskúmavania primeranosti plnenia pozitívneho záväzku štátu – ochraňovať voličov pred zasahovaním do ich slobodnej vôle pri využívaní volebného práva. Uvedené legislatívne zmeny britského volebného zákonodarstva považoval ESĽP za dostatočné a rozhodol o neporušení čl. 3/1DP1 Dohovoru. Zároveň tým aj vymedzil pozitívnu povinnosť štátu pri prevencii pred zasahovaním cudzích mocností do vnútroštátnych volieb prostredníctvom algoritmov, ktorú tak predstavuje aktuálna britská zákonná úprava. Nedá mi však, aby som k tomuto konštatovaniu nepridal pozoruhodné závery súhlasného stanoviska sudcu Jakaba, ktorými som ukončil môj komentár tohto rozhodnutia v Justičnej revue.

Sudca Jakab navrhuje rozlišovať zásahy do volieb cudzími mocnosťami podľa toho, či sú tieto mocnosti demokratické alebo nedemokratické, a to aj v takom prípade, ak ide o členov Rady Európy. Podľa jeho názoru *„čím menej demokratický je pôvod zasahovania, tým silnejšie by mohli a mali byť obranné opatrenia."* Názory A. Jakaba idú v intenciách rozdeľovania krajín Rady Európy, začatého ESĽP variabilným výkladom čl. 18 Dohovoru – čo je základom pre pertraktovanie tzv. dvojakého metra[25] Sudca Jakab iniciatívne navrhuje aj viaceré metódy filtrovania obsahu digitálnej komunikácie,[26] čo už nastoľuje inú

---

[25] Pre posudzovanie kvality „liberálnej demokracie" tento nový sudca ESĽP navrhuje *„informácie o výkone rozsudkov Súdneho dvora, systémové nedostatky identifikované v rozsudkoch Súdneho dvora, prebiehajúce postupy podľa článku 7 zmluvy o Európskej únii (ak ide o členský štát EÚ) alebo všeobecne uznávané indexy právneho štátu a demokracie (ako napríklad index liberálnej demokracie V-DEM alebo index právneho štátu World Justice Project)".* K tomu pozri bližšie aj JAKAB, A., KIRCHMAIR, L.: *Saving the European Union from Its Illiberal Member States*. Oxford: Oxford University Press, 2025.

[26] Podľa A. Jakaba čl. 10 Dohovoru vyžaduje pozitívnu povinnosť štátu, aby používatelia sociálnych médií vedeli, či informácie poskytujú reálne osoby používajúce skutočné mená. Mali by vidieť len tie kanály, na ktoré sa prihlásili alebo predtým sledovali. Navrhuje *„spomalenie virality zavedením čakajúcej doby (napr. niekoľko hodín) na objavenie príspevkov."* Ďalej by sa mal zvážiť zákaz politickej reklamy, „zaplniť" medzeru v nedostatočnej regulácii priestoru influencer marketingu, používať technológiu blockchain pre spravodajské príspevky na sociálnych sieťach, zabrániť zámernému infikovaniu chatbotov dezinformačnými kampaňami.

vážnu otázku, či je naozaj volič v súčasnosti natoľko nesvojprávnym a neuvedomelým, že si vyžaduje takúto až „orwelovskú" opateru.

Uvedené konštatovania by mohli byť dobrým záverom aj pre tento príspevok ale pre úplnosť nadnárodnej úpravy zasahovania cudzích mocností do volieb prostredníctvom algoritmov upozorním ešte aj na tri únijné pramene práva.

Prvým je nariadenie Rady (EÚ) 2022/350 z 1. marca 2022, ktorým sa mení nariadenie (EÚ) č. 833/2014 o reštriktívnych opatreniach s ohľadom na konanie Ruska destabilizujúce situáciu na Ukrajine, a rozhodnutie Rady (SZBP) 2022/351 z 1. marca 2022, ktorým sa mení rozhodnutie 2014/512/SZBP o reštriktívnych opatreniach s ohľadom na konanie Ruska, ktorým destabilizuje situáciu na Ukrajine. Prostredníctvom tohto nariadenia a rozhodnutia Európska únia zakázala prevádzkovateľom vysielať akýkoľvek obsah určitých identifikovaných právnických osôb, subjektov alebo orgánov, a to aj prostredníctvom prenosu alebo distribúcie akýmikoľvek prostriedkami, ako sú káblové vysielanie, satelit, IP-TV, poskytovatelia internetových služieb, platformy alebo aplikácie na zdieľanie videí na internete, či by už boli nové alebo predinštalované.

Druhým je nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 z 19. októbra 2022 o jednotnom trhu digitálnych služieb a o zmene smernice 2000/31ES (zákon o digitálnych službách), ktorý zavádza viaceré povinnosti, najmä dôkladne posúdiť systémové riziká vyplývajúce z návrhu, fungovania a používania ich služieb, ako aj z možného zneužitia zo strany príjemcov služby, a mali by prijať vhodné opatrenia na zmiernenie skutočných alebo predvídateľných negatívnych účinkov na demokratické procesy, občiansku diskusiu a volebné procesy, ako aj na verejnú bezpečnosť.

Pri posudzovaní takýchto systémových rizík by sa poskytovatelia veľkých online platforiem a veľmi veľkých online platforiem a veľmi veľkých online vyhľadávačov mali zamerať na systémy alebo iné prvky, ktoré by mohli prispieť k rizikám, vrátane všetkých algoritmických systémov, ktoré by mohli byť relevantné, najmä ich odporúčacích systémov a reklamných systémov, pričom by mali venovať pozornosť súvisiacim postupom zhromažďovania a používania údajov. Mali by tiež posúdiť, či boli ich zmluvné podmienky a ich presadzovanie vhodné, ako aj ich procesy moderovania obsahu, technické nástroje a pridelené zdroje. Takíto poskytovatelia by mali venovať osobitnú pozornosť tomu, ako sa ich služby používajú na šírenie alebo zosilňovanie zavádzajúceho alebo klamlivého obsahu vrátane dezinformácií.

Napokon tretím je úplne prvý zákon o umelej inteligencii na svete, ktorý vošiel do života pod krkolomným označením ako nariadenie Európskeho

parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii).

Nariadenie (zákon) o umelej inteligencii hneď na prvom mieste medzi zakázané praktiky využívajúce algoritmy umelej inteligencie voviedol *„ uvádzanie na trh, uvádzanie do prevádzky alebo používanie systému AI, ktorý využíva podprahové techniky mimo vedomia osoby alebo účelovo manipulatívne alebo klamlivé techniky s cieľom alebo účinkom podstatne narušiť správanie osoby alebo skupiny osôb tým, že sa citeľne oslabí ich schopnosť prijať informované rozhodnutie, čo zapríčiní, že prijmú rozhodnutie, ktoré by inak neprijali, a to spôsobom, ktorý tejto osobe, inej osobe alebo skupine osôb spôsobuje alebo pri ktorom je odôvodnené predpokladať, že pravdepodobne spôsobí značnú ujmu.“*

Záver

Pod značnou ujmou uvádzanou v akte o umelej inteligencii si nemožno nepredstaviť porušenie kľúčovej zásady volebného práva, a to práva slobodne a bez akéhokoľvek nátlaku vyjadriť svoje politické rozhodnutie. Odhaľovanie týchto praktík nebude v budúcnosti jednoduchšie ale to neznamená, že štáty nemajú mať pozitívny záväzok uskutočniť nielen slobodné ale aj nezávislé voľby bez ich ovplyvnenia algoritmami „oživenými“ inými štátmi v rámci ich hybridnej vojny.

Prehľad odhalených zásahov do volieb v USA, Spojenom kráľovstve a v Rumunsku zo strany Ruska nie je úplný a ani ukončený. Nemožno vylúčiť (práve naopak), že algoritmy neopúšťajú do digitálneho priestoru aj iné krajiny ako Rusko a nemusí to byť vždy ani nepriateľ v rámci hybridnej vojny. Preto by malo byť v záujme zachovania (ale aj vytvárania) základných princípov demokratického riadenia každej spoločnosti tvoriacej ľudskú civilizáciu, aby nástrojmi medzinárodného práva prinútila štáty vytvoriť pre občanov nielen slobodný ale aj neutrálny volebný proces. Zatiaľ sú to len zárodky *soft law* medzinárodného práva, ktoré apelujú na štáty, aby vytvárali vnútroštátny systém spôsobilý zabezpečiť suverénny výkon volebného práva pre svojich občanov. Volebné právo tak bude doplnené o nový princíp suverenity popri slobode, všeobecnosti, rovnosti a tajnosti.

Najbližšie k zakotveniu tohto princípu má v rámci európskeho regionálneho systému ochrany ľudských práv vymedzenie pozitívnej povinnosti štátu podľa čl. 3/1DP Dohovoru v rozhodnutí ESĽP vo veci *Bradshaw*. Predmetné

rozhodnutie vzhľadom na jeho medzinárodnú záväznosť nie je len výzvou pre slovenské legislatívne orgány ale aj urgentnou povinnosťou prijať tomu zodpovedajúcu právnu úpravu.

Zoznam použitej literatúry a zdrojov

1. DOLNÁ SNEMOVŇA, VÝBOR PRE DIGITÁLNE TECHNOLÓGIE, KULTÚRU, MÉDIÁ A ŠPORT: *Dezinformácie a „falošné správy": záverečná správa*. Ôsma správa zo zasadnutia 2017 – 2019, HC 1791, 14. február 2019.
2. DOUBLET, Y.-M.: *Dezinformácie a volebné kampane*. Rada Európy, jún 2019.
3. EURÓPSKA KOMISIA PRE DEMOKRACIU PROSTREDNÍCTVOM PRÁVA (BENÁTSKA KOMISIA): *Dopad informačného chaosu (dezinformácií) na voľby*. CDL-LA(2018)002, Štrasburg, 26. november 2018.
4. EURÓPSKA KOMISIA PRE DEMOKRACIU PROSTREDNÍCTVOM PRÁVA (BENÁTSKA KOMISIA): *Kódex osvedčených postupov vo volebných záležitostiach: usmernenia a vysvetľujúca správa*. CDL-AD(2002)023rev2-cor, Benátky, 30. október 2002. Online: https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2002)023rev2-cor-e.
5. EURÓPSKA KOMISIA PRE DEMOKRACIU PROSTREDNÍCTVOM PRÁVA (BENÁTSKA KOMISIA): *Naliehavá správa o zrušení výsledkov volieb ústavnými súdmi*. CDL-AD(2025)003, 27. január 2025. Online: https://www.coe.int/en/web/venice-commission/-/cdl-ad-2025-003-e.
6. JAKAB, A., KIRCHMAIR, L.: *Saving the European Union from Its Illiberal Member States*. Oxford: Oxford University Press, 2025.
7. KUTI, C.: *On Means and Ends: The Annulment of the Romanian Presidential Elections*. Verfassungsblog, 10/12/2024. Online: https://verfassungsblog.de/on-means-and-ends/.
8. MATUŠKA, P.: *Rumunský precedens: ústavnoprávne aspekty anulovania výsledku volieb v Slovenskej republike a suverenita volieb*. In: Právník, roč. 164, 11/2025, s. 1034–1054.
9. Medzinárodný dohovor o používaní vysielania vo veci mieru, Ženeva, 23. september 1936, LNTS č. 4319.
10. MINISTERSTVO SPRAVODLIVOSTI USA: *List generálneho prokurátora Williama P. Barra predsedom a vedúcim členom menšiny výborov pre justíciu Snemovne reprezentantov a Senátu týkajúci sa zhrnutia hlavných záverov správy osobitného vyšetrovateľa.* 24. marec 2019. Online: https://www.justice.gov/archives/ag/page/file/1147981/dl.
11. MINISTERSTVO SPRAVODLIVOSTI USA, ÚRAD OSOBITNÉHO VYŠETROVATEĽA: *Správa o vyšetrovaní ruského zasahovania do prezidentských volieb v roku 2016* [tzv. Muellerova správa], Washington, D.C., marec 2019. Online: https://www.justice.gov/storage/report.pdf.

12. MIRGA, T.: *Šírenie dezinformácií a organizované ovplyvňovanie verejnej mienky: Trolie farmy ako nástroj sociálnej a politickej manipulácie*. In: ITlib: Informačné technológie a knižnice, 2024. Online: http://doi.org/10.52036/1335793X.2023.3-4.42-55.
13. Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii), OJ L, 12.7.2024.
14. Nariadenie Rady (EÚ) 2022/350 z 1. marca 2022, ktorým sa mení nariadenie (EÚ) č. 833/2014 o reštriktívnych opatreniach s ohľadom na konanie Ruska,ktorým destabilizuje situáciu na Ukrajine, OJ L 65, 2.3.2022.
15. Nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 z 19. októbra 2022 o jednotnom trhu s digitálnymi službami a o zmene smernice 2000/31/ES (akt o digitálnych službách) (Text s významom pre EHP), OJ L 277, 27.10.2022.
16. ÚRAD VYSOKÉHO KOMISÁRA OSN PRE ĽUDSKÉ PRÁVA: *Ľudské práva a voľby – príručka o medzinárodných ľudskoprávnych štandardoch týkajúcich sa volieb*, 20. september 2021. Online: https://www.ohchr.org/en/publications/training-and-education-publications/human-rights-and-elections-handbook-international.
17. OSOBITNÁ SPRAVODAJKYŇA OSN PRE PODPORU A OCHRANU PRÁVA NA SLOBODU NÁZORU A PREJAVU: *Dezinformácie a sloboda názoru a prejavu*. Správa osobitnej spravodajkyne OSN, UN Doc. A/HRC/47/25, 13. apríl 2021.
18. Oznámenie Federálneho ministerstva zahraničných vecí č. 209/1992 Zb. Dohovor o ochrane ľudských práv a základných slobôd v znení protokolov č. 3, 5 a 8.
19. PARLAMENTNÉ ZHROMAŽDENIE RADY EURÓPY: *„Napadnutá demokracia? Ako reagovať?*, rezolúcia č. 2326(2020), 31. január 2020.
20. PARLAMENTNÉ ZHROMAŽDENIE RADY EURÓPY: *Sloboda médií ako podmienka demokratických volieb*, rezolúcia č. 2254 (2019), 23. január 2019.
21. PARLAMENTNÉ ZHROMAŽDENIE RADY EURÓPY: *Umelá inteligencia pri volebných urnách: hrozba alebo príležitosť pre demokratické voľby?* 7.11.2025. Online: https://pace.coe.int/en/news/10095.

22. PARLAMENTNÉ ZHROMAŽDENIE RADY EURÓPY: *Voľby v čase krízy: výzvy a príležitosti. Súhrnná správa a závery Bernskej konferencie konanej v dňoch 9. – 10. mája 2023.* Správa o činnosti, Doc. 15791 Add. 1, 16. jún 2023.
23. RADA PRE ĽUDSKÉ PRÁVA: *Úloha štátov pri potláčaní negatívneho vplyvu dezinformácií na užívanie a realizáciu ľudských práv*, rezolúcia 49/211, apríl 2022.
24. Rozhodnutie ESĽP vo veci *Bradshaw a ostatní proti Spojenému kráľovstvu*, stažnosť č. 15653/22, rozsudok z 22. júla 2025.
25. Rozhodnutie Rady (SZBP) 2022/351 z 1. marca 2022, ktorým sa mení rozhodnutie 2014/512/SZBP o reštriktívnych opatreniach s ohľadom na konanie Ruska, ktorým destabilizuje situáciu na Ukrajine, OJ L 65, 2.3.2022.
26. Rozhodnutie Ústavného súdu Rumunska č. 66 z 1. októbra 2019.
27. SNEMOVŇA REPREZENTANTOV USA: *H.R. 156*, 119. kongres, 1. zasadnutie (2025).
28. VLÁDA SPOJENÉHO KRÁĽOVSTVA: Odpoveď vlády na správu Parlamentného výboru pre spravodajské služby a bezpečnosť „Rusko“, júl 2020.
29. SVÁK, J.: *Ovplyvňovanie volieb cudzími mocnosťami*. In: Justičná revue, č. 10/2025, s. 1033–1045.
30. OSOBITNÝ SPRAVODAJCA OSN PRE SLOBODU NÁZORU A PREJAVU, PREDSTAVITEĽ ORGANIZÁCIE PRE BEZPEČNOSŤ A SPOLUPRÁCU V EURÓPE PRE SLOBODU MÉDIÍ, OSOBITNÝ SPRAVODAJCA ORGANIZÁCIE AMERICKÝCH ŠTÁTOV PRE SLOBODU PREJAVU A OSOBITNÝ SPRAVODAJCA AFRICKEJ KOMISIE PRE ĽUDSKÉ PRÁVA A PRÁVA NÁRODOV PRE SLOBODU PREJAVU A PRÍSTUP K INFORMÁCIÁM: *Spoločné vyhlásenie o slobode prejavu a „falošných správach“, dezinformáciách a propagande*, 3. marec 2017. Online: https://rfom.osce.org/fom/302796.
31. VÝBOR PARLAMENTU PRE SPRAVODAJSKÉ A BEZPEČNOSTNÉ SLUŽBY SPOJENÉHO KRÁĽOVSTVA: *Rusko,* HC 632, 21. júl 2020.
32. WARDLE, C., DERAKHSHAN, H.: *Informačná porucha: smerom k interdisciplinárnemu rámcu pre výskum a tvorbu politík*. Správa Rady Európy, DGI(2017)09, 27. september 2017.

Informácie o autorovi
prof. JUDr. Ján SVÁK, DrSc.
Univerzita Komenského v Bratislave
Právnická fakulta
Katedra medzinárodného práva a medzinárodných vzťahov
Email: jan.svak@flaw.uniba.sk

13 Proaktívne zverejňovanie informácií súdmi ako nástroj predchádzania vzniku právnych halucinácií

Emma Humeňanská

https://doi.org/10.33542/HRF-0541-2-13

Abstrakt

S nástupom umelej inteligencie sa prístup k informáciám a ich vyhľadávanie zásadne mení. Táto zmena sa dotýka všetkých spoločenských oblastí, súdnictvo nevynímajúc. Namiesto vyhľadávania informácií na webových stránkach súdov, či podaní žiadosti o sprístupnenie informácií, dokáže umelá inteligencia na základe jednoducho sformulovateľného promptu informácie o súdnictve a judikatúre nielen vyhľadať, ale aj analyzovať, interpretovať. Problém, ktorý vyvstáva v súvislosti s generovaním informácií je ich otázna pravdivostná hodnota, keďže umelá inteligencia nedokáže rozlišovať medzi pravdivou a nepravdivou informáciou. Cieľom príspevku je preskúmať, či sa s rozvojom informačných technológií nevytvára nový pozitívny záväzok súdov aktívne informovať verejnosť o svojej činnosti. Príspevok vychádza z vlastného empirického príkladu generovania informácií veľkým jazykovým modelom v súvislosti s vyhľadávaním informácií o Ústavnom súde SR. Príspevok v závere formuluje návrhy, akou formou a prostriedkami, by mohla byť aktívna komunikácia súdov ako protiváha k informáciám generovaných umelou inteligenciou realizovaná.

Úvod

Informácie o judikatúre, činnosti a organizácii súdnictva verejnosť tradične získavala z internetových stránok súdov. V prípade, ak špecifickejšiu informáciu súdy na webových stránkach nezverejňujú, je možné sa na súd ako orgán verejnej moci obrátiť so žiadosťou o sprístupnenie informácií. Infožiadosti je vyhovené, ak je požadovaná informácia dostupná a pripravená. S nástupom generatívnej umelej inteligencie sa kritérium pripravenosti a najmä dostupnosti mení. Umelá inteligencia dokáže informácie z oblasti práva nielen generovať, ale rovnako aj štatisticky spracovať, obsahovo analyzovať a následne interpretovať. Pred súdmi, ktoré mali výlučný prístup k informáciám o svojej činnosti a mohli ovplyvniť ich obsah, stojí nová výzva, ako na generované informácie reagovať. Riziko generovania nepravdivých či nesprávnych odpovedí predstavuje hlavný argument, pre ktorý by mali súdy zefektívniť svoju komunikáciu s verejnosťou a zverejňovať informácie o svojej činnosti v čo najväčšom rozsahu.

Cieľom príspevku je preto reflektovať, ako umelá inteligencia mení doterajšie vnímanie práva na informácie a zodpovedať na otázku, či sa

s rozvojom informačných technológií nevytvára nový pozitívny záväzok súdov aktívne informovať verejnosť o svojej činnosti. Prvá časť je venovaná premene práva na informácie pod vplyvom umelej inteligencie a problému tzv. právnych halucinácií. Vznik právnych halucinácií preukazujeme prostredníctvom konkrétnych empirických príkladov, na jazykovom modeli ChatGPT-5, v kontexte Ústavného súdu SR. V záverečnej časti sa venujeme osobitným formám komunikácie, ktoré by súdy mohli využiť nielen v snahe o ochranu práva na informácie, ale rovnako aj o posilnenie vnímanej legitimity súdnictva.

13.1 Vplyv umelej inteligencie na premenu práva na informácie

„Právo na informácie je nástrojom presadzovania princípu transparentnosti."[1] Na ústavnej úrovni je princíp transparentnosti prvotne garantovaný ako právo spočívajúce z oprávnenia informácie vyhľadávať, prijímať a rozširovať.[2] Vyhľadávanie informácií bolo tradične vnímané ako „zisťovanie... či v oblasti záujmu toho, kto informácie vyhľadáva, informácie vôbec jestvujú, a ak áno, tak aké a kde sa nachádzajú."[3] Vyhľadávanie informácií vlastnou činnosťou sa najčastejšie realizuje prostredníctvom internetového vyhľadávača a webových stránok, ako súboru rozličných odkazov, z ktorých je potrebné požadovanú informáciu abstrahovať. Proces vyhľadávania vychádza z niekoľkých samostatných krokov. Prvý krok je možné vymedziť ako uváženie, akú informáciu používateľ potrebuje vyhľadať, pričom na základe kľúčových slov informáciu rámcovo pomenúva a zadáva do vyhľadávača. Z navrhnutých internetových stránok uskutočňuje výber, z akého zdroja bude informáciu čerpať a v závere, v prípade pochybností môže informáciu verifikovať preskúmaním a porovnaním ďalšieho z nájdených zdrojov. Proces vyhľadávania je tak zavŕšený vyhodnotením dôveryhodnosti zdroja, čím si používateľ vytvára vlastný úsudok o (ne)pravdivosti informácie.

S nástupom umelej inteligencie sa proces vyhľadávania mení.[4] Pri vyhľadávaní informácií prostredníctvom veľkých jazykových modelov, je zadanie

---

1 DOBROVODSKÝ, R. KOŠIČIAROVA, S.: *Právo na informácie.* Krakov: Spolok Slovákov v Poľsku, 2015, s. 9.

2 Čl. 26 ods. 2 Ústavy SR č. 460/1992 Zb. v znení neskorších predpisov.

3 OROSZ, L., SVÁK, J. et. al.: *Ústava Slovenskej republiky*, Komentár, Zväzok I. Bratislava: Wolters Kluwer, 2021, 343 s.

4 Na Slovensku využíva generatívnu umelú inteligenciu tretina populácie. Každý druhý používateľ ju využíva ako nástroj rýchleho vyhľadávania informácií. LAKATOŠOVÁ, D.: *Umelú inteligenciu využíva tretina z nás. Sme svedkami revolúcie vo vyhľadávaní informácií?* NMS, 27.1.2025. Online: https://nms.global/sk/umelu-inteligenciu-vyuziva-tretina-z-nas-sme-svedkami-revolucie-vo-vyhladavani-informacii/.

kľúčových výrazov do internetového vyhľadávača nahradené položením otázky (promptu), na základe ktorej je informácia generovaná. V porovnaní s internetovým vyhľadávaním tak generovanie nepredpokladá uvedomelý výber zdroja informácie, čo umožňuje informáciu bez ďalšieho kritického zhodnotenie prijať.

Prijímanie informácií bolo doteraz definované predpokladom, že „ľudskými zmyslami vyjadrenú informáciu a informáciu už zachytenú technickými prostriedkami, ale i písmom možno prostredníctvom ľudského zmyslu prijať a ďalej rozširovať."[5] Vplyvom umelej inteligencie sa však toto vymedzenie rozširuje aj o prijímanie generovaných informácií, ktoré sú vo svojej podstate nové, nezávislé od predošlého vyjadrenia prostredníctvom ľudských zmyslov či technických prostriedkov. Uvedené vychádza zo skutočnosti, že veľké jazykové modely fungujú na báze strojového učenia. Sú trénované na rozsiahlych slovných korpusoch zložených z kníh, internetových stránok, zákonov, databáz súdnych rozhodnutí, na ktorých učia štatisticky predvídať vzťahy medzi jednotlivými slovami.[6] Generovaná odpoveď je tak zakaždým výberom najpravdepodobnejšej sekvencie znakov, ktorá by mala nasledovať po zadanom prompte. Práve preto je generovaná informácia „novou" informáciou, nezávislou od vstupných dát, na ktorých bol model trénovaný.[7] Táto skutočnosť je zároveň aj najväčším nedostatkom všetkých jazykových modelov. Umelá inteligencia nemá rozvinuté sémantické vnímanie, ale generuje informácie na základe štatistickej predikcie. Nedokáže preto zohľadniť obsah a kontext dát, na ktorých je trénovaná, no rovnako nedokáže zohľadniť ani obsah a kontext informácií, ktoré generuje.[8] Aj napriek tomu, že užívateľ prijíma novú informáciu, problémom sa stáva otázna miera jej pravdivostnej hodnoty, keďže umelá inteligencia nedokáže rozlišovať pravdivú informáciu od nepravdivej, ale vyhľadávanú informáciu „len" vytvára. Na problematickosť generovania fakticky nesprávnych informácií poukazujeme empiricky, prostredníctvom vlastného príkladu.

---

5 OROSZ, L., SVÁK, J. et. al.: *Ústava Slovenskej republiky*, Komentár, Zväzok I. Bratislava: Wolters Kluwer, 2021, s. 344. Podľa nálezu Ústavného súdu SR sp. zn. II. ÚS 28/96 z 12. mája 1997.

6 DAHL, M. et. al.: *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*. In: Journal of Legal Analysis, roč. 16, č. 1, 2024. s. 66. Online: https://doi.org/10.1093/jla/laae003.

7 COURT OF JUSTICE OF THE EUROPEAN UNION, DIRECTORATE-GENERAL FOR INFORMATION: *Artificial Intelligence Strategy*, Approved by the Administrative Committee of the Court on 30 June 2023, s. 6. Online: https://curia.europa.eu/site/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf

8 *Ibid*.

Podľa čl. 127 ods. 5 Ústavy SR č. 460/1992 Zb. (ďalej len „ústava“), môže fyzická alebo právnická osoba spolu s ústavnou sťažnosťou podať návrh, aby senát ústavného súdu, podal návrh na začatie konania podľa čl. 125 ods. 1 (konanie o súlade právnych predpisov).[9]

S cieľom preskúmať možný výskyt generovania nepravdivej informácie, bol do jazykového modelu ChatGPT-5 zadaný prompt: „môže fyzická alebo právnická osoba spolu s ústavnou sťažnosťou podať návrh, aby senát ústavného súdu podal návrh na začatie konania o súlade právnych predpisov?“ Generovaná odpoveď znela: „Nie, fyzická ani právnická osoba nemôže spolu s ústavnou sťažnosťou podať návrh, aby senát Ústavného súdu SR podal návrh na začatie konania o súlade právnych predpisov. Ústava a zákon č. 314/2018 Z. z. o Ústavnom súde SR takúto možnosť nepoznajú.“[10]

Aj napriek tomu, že zadaný prompt vychádzal zo znenia ústavy, jazykový model nedokázal vygenerovať správnu odpoveď. Správne síce rozlíšil jednotlivé právne predpisy, ktoré sa na konanie vzťahujú, avšak problém nastal pri ich ďalšej interpretácií. Umelá inteligencia uviedla, že „jednotlivec ani právnická osoba nie je podľa čl. 130 ods. 1 Ústavy SR oprávnená podať návrh na začatie konania o súlade právnych predpisov.“[11] Čl. 130 však nevymedzuje okruh subjektov, ktorí sú oprávnení iniciovať konanie o súlade právnych predpisov, ale vo všeobecnosti stanovuje subjekty, na návrh ktorých sa konanie pred ústavným súdom začína. Typ konania pritom nie je rozhodujúci.

Napriek nesprávnej interpretácií umelá inteligencia ďalej iniciatívne navrhla, „že jednotlivec alebo právnická osoba môžu v ústavnej sťažnosti namietať, že porušenie ich práv bolo spôsobené použitím protiústavného právneho predpisu podľa čl. 127 ods. 6 Ústavy SR.“[12] Čl. 127 ústavy sa však člení na päť odsekov. Umelá inteligencia si vytvorila vlastný odsek tohto článku, aby odpovedala na zadaný prompt. Za týmto účelom naformulovala aj znenie fiktívneho odseku: „Ak Ústavný súd pri prerokúvaní sťažnosti zistí, že porušenie práv bolo spôsobené použitím právneho predpisu, ktorý odporuje ústave,

---

9 Čl. 125 ods. 1 Ústavy SR č. 460/1992 Zb. v znení neskorších predpisov.

10 Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4773af908191ad5a27d94f82af6e.
Tento prompt bol do jazykového modelu ChatGPT zadaný trikrát, vždy v inej formulácií, aj s označením konkrétnych článkov ústavy, no odpoveď bola zakaždým nesprávna. Umelá inteligencia však správne dokázala interpretovať konanie o ústavnej sťažnosti, podľa čl. 127 ods. 1 a to aj vysvetlením procesného postupu na súde.

11 *Ibid.*

12 Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4773af908191ad5a27d94f82af6e.

rozhodne aj o jeho nesúlade s ústavou." Tento odsek následne aj interpretovala tak, že „Ústavný súd môže sám v rámci sťažnosti posúdiť ústavnosť právneho predpisu (tzv. *incidenčná kontrola*)."[13] V tomto prípade umelá inteligencia opätovne poskytla zavádzajúcu informáciu, keďže senát nie je oprávnený posúdiť ústavnosť právneho predpisu, ale konanie o súlade právnych predpisov môže len iniciovať. Návrh senátu je podmienený prvotným návrhom fyzickej alebo právnickej osoby, ktorá ho podáva spolu s ústavnou sťažnosťou.

Generovaná informácia napriek svojej faktickej nesprávnosti vyznela hodnoverne. Bez porovnania odpovede so skutočným znením ústavy alebo prvotnej znalosti tohto článku, by inak nebolo možné odhaliť jej nepresnosť, nakoľko si umelá inteligencia neuvedomuje svoju omylnosť,[14] a na vlastné zlyhanie používateľa neupozorní. K priznaniu pochybenia môže dôjsť až po upozornení používateľa. Táto skutočnosť bola overená tak, že po vygenerovanej odpovedi bol do modelu ChatGPT-5 zadaný nový prompt, s cieľom upozornenia umelej inteligencie na nesprávnosť generovanej informácie. Znenie promptu bolo nasledovné: „táto informácia nie je pravdivá, podľa čl. 127 ods. 5 ústavy fyzická alebo právnická osoba môže podať spolu s ústavnou sťažnosťou návrh, aby senát ústavného súdu podal návrh na začatie konania o súlade právnych predpisov."[15] Umelá inteligencia si po upozornení informáciu opätovne overila, a to cez funkcionalitu rozšíreného vyhľadávania na internetových stránkach. Následne sa „ospravedlnila za chybu" a informáciu správne interpretovala s „odkazom na ústavný zákon č. 422/2020, ktorým bol do čl. 127 ústavy doplnený odsek 5." [16]

13.2 Problém právnych halucinácií

Problém veľkých jazykových modelov pri vyhľadávaní informácií v oblasti práva spočíva v tom, že používateľ sa ocitá v riziku generovania fiktívnych zákonov a súdnych rozhodnutí. Riziko predstavuje aj nesprávna interpretácia, odkazy na prekonanú judikatúru, využitie jazyka, nekonzistentného so súčasnou právnou doktrínou, či opomenutie spoločenského kontextu, v ktorom boli zákony a súdne rozhodnutia vydané.[17] Spoločným pomenovaním tohto javu je

---

13 *Ibid.*

14 DAHL, M. et. al.: *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*. In: Journal of Legal Analysis, roč. 16, č. 1, 2024. s. 87.

15 Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4a2f86fc8191a3ac62ae1d54d20a.

16 *Ibid.*

17 COUNCIL OF EUROPE, COMMITTEE OF EXPERTS ON THE IMPLICATIONS OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR FREEDOM OF EXPRESSION (MSI-AI):

označenie „právnej halucinácie."[18]

Jeden z dôvodov ich vzniku môže spočívať v limitovanom rozsahu jazykových korpusov, na ktorých sú veľké jazykové modely trénované.[19] Napriek svojej veľkosti, prirodzene, nemôžu obsiahnuť všetku legislatívu, databázy súdnych rozhodnutí, odlišnosti jednotlivých právnych kultúr. S týmto problémom úzko súvisí otázka aktuálnosti dát. Aktualizácia voľne dostupných veľkých jazykových modelov je periodická, nie kontinuálna. Pre časovú a finančnú náročnosť prebieha približne každých 6 – 12 mesiacov, v dôsledku čoho nedokážu byť komerčné jazykové modely systematicky trénované na kompletnej judikatúre a právnych predpisoch.[20] S pokrokom vo vývoji sa tento nedostatok postupne vylepšuje zavadzaním funkcionality rozšíreného vyhľadávania (tzv. RAG),[21] ktorý umožňuje informáciu generovať aj z voľne dostupných domén a internetových stránok. Riziko nesprávnej interpretácie hrozí aj v tomto prípade, keďže napriek širšej škále dát, z ktorých môže jazykový model informáciu generovať, jeho neschopnosť obsahového rozlíšenia a nezohľadnenia kontextu pretrváva.[22]

Limit vstupných dát bol preukázaný aj na príklade otázky o ústavnej sťažnosti podľa čl. 127 ods. 5 ústavy. Vzhľadom na to, že ide o ustanovenie účinné od 1.1.2025 a doposiaľ nebolo predmetom rozsiahleho autorského záujmu, či početného zastúpenia v rozhodovacej činnosti ústavného súdu, umelá

Draft Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression, MSI-AI(2025)10, 3 June 2025, s. 3. Online: https://rm.coe.int/msi-ai-2025-10-draft-guidance-note-on-the-implications-of-generative-a/1680b68c48.

18 DAHL, M. et. al.: *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*. In: Journal of Legal Analysis, roč. 16, č. 1, 2024. s. 87.

19 *Ibid.*

20 V prípade jazykových modelov vyvinutých pre konkrétnu organizáciu aktualizácia dát môže prebiehať odlišne. Napr. v rámci Súdneho dvora EÚ je v budúcnosti cieľom, predovšetkým v oblasti prekladu, využívať model, ktorý by umožňoval spracovanie prirodzeného jazyka. COURT OF JUSTICE OF THE EUROPEAN UNION: *Artificial Intelligence Strategy*, 2023, s. 15.

21 Z anglického Retrieval-Augmented Generation. Ide o situáciu, kedy jazykový model negeneruje odpoveď priamo z trénovaných dát, ale z dostupných domén, internetových stránok, čo umožňuje generovanie presnejších informácií. Využitie tejto modality závisí od architektúry systému (či ide o jazykový model, ktorý využíva RAG). STANFORD UNIVERSITY HUMAN-CENTERED ARTIFICIAL INTELLIGENCE: AI on Trial: *Legal Models Hallucinate in 1 out of 6 (or More) Benchmarking Queries*, 23 May 2024. Online: https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries.

22 *Ibid.* Tento problém sa netýka len komerčných jazykových modelov, ale rovnako aj špecifických modelov navrhnutých priamo pre oblasť práva. Z doterajších výskumov je zrejmé, že aj v týchto prípadoch AI halucinovala prinajmenšom v každej piatej otázke.

inteligencia mala obmedzenejší rozsah vstupných dát, z ktorých mohla informáciu generovať. Až následne, keď bola upozornená na faktickú nepresnosť, informáciu generovala využitím funkcionality rozšíreného vyhľadávania. Paradoxom je, že túto informáciu vyhľadávala na 24 rôznych internetových stránkach,[23] pričom ani jednou z nich nebolo webové sídlo ústavného súdu. Túto skutočnosť možno zrejme odôvodniť tým, že ústavný súd na svojej stránke neuvádza informácie o konaní podľa čl. 127 ods. 5, len vo všeobecnosti sprístupňuje schému postupu konania podľa čl. 127 ods.1 ústavy.[24]

S cieľom overiť tento predpoklad bol do modelu ChatGPT-5 zadaný prompt: „v prípade, ak by ústavný súd zverejnil na svojej webovej stránke informáciu o konaní podľa čl. 127 ods. 5, bola by generovaná odpoveď správna?" Umelá inteligencia uviedla, že dôvodom jej prvotnej nesprávnej odpovede bola skutočnosť, že vychádzala „zo staršej právnej úpravy čl. 127 Ústavy SR, ktorá platila pred novelou ústavného zákona č. 422/2020 Z. z."[25] Následne odpoveď zdôvodnila tým, že „ak by web Ústavného súdu SR obsahoval aj výslovnú zmienku o čl. 127 ods. 5 (napr. v časti „Podanie ústavnej sťažnosti"), a model by túto stránku použil ako zdroj, odpoveď by bola správna už pri prvom pokuse."[26]

Pre porovnanie ako umelá inteligencia pracuje so zdrojmi bola do jazykového modelu opätovne položená rovnaká otázka, tentokrát však v súvislosti s českým ústavným súdom, ktorý môže tak ako slovenský ústavný súd, na podklade ústavnej sťažnosti zrušiť ustanovenie právneho predpisu, ktoré je nesúladné s ústavou. Na rozdiel od slovenského ústavného súdu bola generovaná odpoveď správna už po prvotnom zadaní otázky,[27] čo umelá inteligencia zdôvodnila odkazom na všeobecne známu doktrínu. Odpoveď pritom vychádzala zo vstupných dát, na ktorých bol jazykový model trénovaný. S cieľom porovnať využitie funkcionality rozšíreného vyhľadávania aj na českom príklade, bol do umelej inteligencie zadaný promt, či je možné túto informáciu vyhľadať priamo na stránke českého ústavného súdu. Umelá inteligencia uviedla, že informácia nevychádza len zo vstupných dát, ale rovnako sa „nachádza aj

---

[23] Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4a2f86fc8191a3ac62ae1d54d20a.

[24] ÚSTAVNÝ SÚD SR: *Konanie o ústavnej sťažnosti*. Online: https://www.ustavnysud.sk/informacie_pre_navrhovatelov/sprievodca-konanim-na-ussr.

[25] Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4a2f86fc8191a3ac62ae1d54d20a.

[26] *Ibid.*

[27] Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/c/69282f43-6aa0-8333-8433-f324b66c9379.

oficiálnej stránke českého ústavného súdu," na ktorú priamo odkázala.[28]

Aj keď odpoveď potvrdzuje náš predpoklad, je ju potrebné vnímať kriticky. Pri vyhľadávaní informácií umelou inteligenciou sa totiž každý používateľ vystavuje riziku servilnosti (sycophancy risk).[29] Vychádza zo skutočnosti, že umelá inteligencia ponúka personalizovaný obsah odzrkadľujúci presvedčenie používateľa. Používateľa utvrdí o vlastnom presvedčení poskytnutím vierohodnej (hoci možne nesprávnej) odpovede, v pochvale, či v iniciatívnom návrhu ďalšej otázky, ktorá však nereflektuje faktickosť a pluralitný pohľad. Túto skutočnosť preukazujeme tým, že umelá inteligencia na zadané prompty odpovedala: „Veľmi dobrá (a už pomerne pokročilá) otázka, výborná a veľmi premyslená otázka, či napr. máš úplnú pravdu - ospravedlňujem sa za chybu v predchádzajúcej odpovedi," a až následne vyhľadávanú informáciu interpretovala.[30]

13.3 Komunikácia súdov ako protiváha informáciám generovaných AI?

Vyhľadávanie a prijímanie informácií prostredníctvom umelej inteligencie, no najmä problém právnych halucinácií vytvára požiadavku pre aktívnejší a transparentnejší prístup súdov v spôsobe ich komunikácie navonok. Aj keď právo na informácie zahŕňa v sebe právo na informácie, ktoré sú fiktívne, neoveriteľné či nepravdivé,[31] vyššia intenzita zverejnených informácií zo strany súdov môže znamenať aj vyššiu pravdepodobnosť generovania fakticky správnej odpovede. Pravdivosť informácie je zároveň hlavným dôvodom, pre ktorý sa

---

28 *Ibid.*

29 COUNCIL OF EUROPE, COMMITTEE OF EXPERTS ON THE IMPLICATIONS OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR FREEDOM OF EXPRESSION (MSI-AI): *Draft Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression*, MSI-AI(2025)10, 3 June 2025, s. 3.

30 Zo štúdií, ktoré boli realizované na súdnych rozhodnutiach v USA zároveň vyplýva, že halucinácie sú častejšie v prípade: generovania informácií o nižších súdnych inštanciách, starších a najnovších rozhodnutiach, nižšej spoločenskej významnosti daného prípadu, menej známej jurisdikcii. Na pravdivostnú hodnotu odpovede však vplýva aj typ jazykového modelu. Pri modeli ChatGPT-4 je pravdepodobnosť halucinovania odhadovaná na 58%, pri modeli ChatGPT-3,5 na 69%. Na modeli Llama-2 dosahuje pravdepodobnosť halucinácií až 88%. STANFORD UNIVERSITY HUMAN-CENTERED ARTIFICIAL INTELLIGENCE: AI on Trial: Legal Models Hallucinate in 1 out of 6 (or More) Benchmarking Queries, 23 May 2024. Online: https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries.

31 Nález Ústavného súdu SR II. ÚS 307/2014 zo dňa 18. 12. 2014. V tomto prípade išlo o informácie sprostredkované médiami, no analogicky je možné právo na fiktívnu či nepravdivú odpoveď rozšíriť aj na AI ako nový zdroj informácií.

používateľ informáciu cez oficiálne komunikačné nástroje inštitúcií rozhodne vyhľadávať.[32] Proaktívne zverejňovanie informácií zo strany súdov je tak možné vnímať ako druhú rovinu princípu transparentnosti. Tá je formulovaná nielen cez právo na informácie, ale rovnako aj ako povinnosť orgánov verejnej moci, informácie o vlastnej činnosti primeraným spôsobom poskytovať.[33]

Názory na vnímanie primeraného spôsobu zverejňovania sa odlišujú. ESĽP vo svojej judikatúre uvádza, že Dohovor nevytvára pozitívny záväzok zhromažďovať a z vlastnej iniciatívy šíriť informácie, ktoré sa týkajú štátu.[34] Informácia by mala byť štátnym orgánom poskytnutá v prípade, ak predstavuje predpoklad pre realizáciu slobody prejavu, týka sa verejného záujmu, je pripravená a dostupná. Osobitne sa zohľadňuje postavenie žiadateľov, najmä novinárov, neziskových organizácií, vedcov a akademikov, ktorí prispievajú k verejnej diskusii a ovplyvňujú ju.[35]

Odlišný prístup k informáciám je možné rozlišovať na úrovni OSN. Ako uvádza Výbor OSN pre ľudské práva, na to, aby právo na informácie bolo skutočne efektívne, by mali zmluvné štáty proaktívne zverejňovať informácie vo verejnom záujme.[36] Uvedené vychádza zo skutočnosti, že všetky „informácie v dispozícii štátu patria verejnosti, s výnimkou obmedzenia, ktoré je v súlade s princípom zákonnosti, nevyhnutnosti a proporcionality."[37] Z pohľadu OSN sa tak vytvára

---

32 Komunikačné nástroje verejných inštitúcií predstavujú jeden zo spôsobov, ktorým verejnosť overuje informácie generované umelou inteligenciou. Najčastejšie používatelia overujú generovaný obsah rozhovorom s blízkymi osobami, následne cez oficiálne inštitúcie a sociálne siete. INŠTITÚT PRE VEREJNÉ OTÁZKY: *Schopnosť ľudí odhaliť klamlivý obsah generovaný AI je veľmi nízka*. Online: https://www.ivo.sk/8987/sk/aktuality/schopnost-ludi-odhalit-klamlivy-obsah-generovany-ai-je-velmi-nizka.

33 Čl. 26 ods. 5 Ústavy SR č. 460/1992 Zb. v znení neskorších predpisov.

34 Rozsudok ESĽP vo veci *Magyar Helsinki Bizottság v. Hungary,* sťažnosť č. 18030/11, rozsudok z 08.11. 2016, bod 156.
Z pohľadu judikatúry ESĽP je právo na informácie interpretované ako sloboda prijímať informácie, ktoré sa týkajú štátnej moci. Keďže vychádza a spadá do rozsahu slobody prejavu, zaväzuje štát neobmedzovať jednotlivca v prijímaní informácií v prípade, ak informáciu chce alebo je ochotný sprostredkovať tretí subjekt.

35 Rozsudok ESĽP vo veci *Magyar Helsinki Bizottság v. Hungary,* sťažnosť č. 18030/11, rozsudok z 08.11. 2016, body 158–170.

36 HUMAN RIGHTS COMMITTEE: *General comment No. 34 on Art. 19: Freedoms of opinion and expression*, CCPR/C/GC/34, 12 September 2011, bod 19.

37 UNITED NATIONS SPECIAL RAPPORTEUR ON THE PROMOTION AND PROTECTION OF THE RIGHT TO FREEDOM OF OPINION AND EXPRESSION: *Access to Information: Freedom of opinion and expression - 2013 thematic report to the Human Rights Council of the Special Rapporteur on freedom of opinion and expression*. 2013. Online: https://www.ohchr.org/sites/default/files/Documents/Issues/Expression/Factsheet_5.pdf.

samostatný princíp informovania v čo najširšom rozsahu. Vytvára domnienku, že všetky informácie v dispozícií orgánov verejnej moci, by mali podliehať zverejneniu s výnimkou len zákonom predpokladaných situácií, ktoré musia byť zakaždým interpretované úzko. Zároveň by štáty mali zabezpečiť jednoduchý, rýchly, efektívny a praktický prístup k týmto informáciám.[38]

Obdobný názor je zároveň prítomný aj v doktríne. Využitie neformálnych komunikačných nástrojov v podobe internetových stránok a profilov na sociálnych sieťach je považovaný nielen za napĺňanie informačnej povinnosti, ale rovnako aj za osobitný nástroj zvyšovania dôvery.[39] Bez ich aktívneho využitia sa informácie k verejnosti dostávajú výlučne prostredníctvom médií, čo zakaždým znamená prijatie sprostredkovanej informácie a oslabenie priameho kontaktu súdu s verejnosťou. So silnejúcim významom umelej inteligencie sa mení aj okruh adresátov zverejňovaných informácií. V prípade informácií uverejnených na internetových stránkach, sa okrem verejnosti stáva novým adresátom aj umelá inteligencia. Zo súdov, ktoré mali výlučný prístup k svojim informáciám a mohli ovplyvniť ich obsah a interpretáciu, sa zároveň stáva aj prvotný zdroj informácií, z ktorých umelá inteligencia novú informáciu vytvára. Nedostatočné informovanie zo strany súdov tak potom znamená aj vyššiu pravdepodobnosť generovania fakticky nesprávnych informácií ako sme poukázali aj na príklade generovania informácií o Ústavnom súde SR.

Požiadavka po intenzívnejšom informovaní však vychádza aj od verejnosti. Podľa údajov ústavného súdu počet infožiadostí každým rokom narastá, čo si vyžaduje intenzívnejšie zverejňovanie informácií na internetovej stránke.[40] Informovanie by nemalo byť zamerané len na popis právomocí, organizačnej štruktúry a sprístupnenia databáz súdnych rozhodnutí. Zaiste, predovšetkým publikácia rozhodnutí je z hľadiska prístupu k informáciám dôvodom, pre ktoré sa súdy verejnosti „otvárajú," avšak predstavuje len nevyhnutné minimum. Viacerí autori poukazujú, že rozhodnutia súdov nie sú určujúcim faktorom, na

---

38 HUMAN RIGHTS COMMITTEE: *General comment No. 34 on Art. 19: Freedoms of opinion and expression*, CCPR/C/GC/34, 12 September 2011, bod 19.

39 KOSAŘ, D.: *Embedding Strategies of the European Apex Courts: Why Court Communication with All Segments of Society Matters*. In: German Law Journal, roč. 26, č. 2, 2025, s. 278. Online: https://www.cambridge.org/core/journals/german-law-journal/article/embedding-strategies-of-the-european-apex-courts-why-court-communication-with-all-segments-of-society-matters/A2DC27653A0CDA1059B1097F4F402C5C. DOI: https://doi.org/10.1017/glj.2025.3.

40 ÚSTAVNÝ SÚD SR: *Ročenka 2024*. 2024, s. 49. Online: https://www.ustavnysud.sk/documents/d/portal/rocenka-2024-sk-pdf.

základe ktorého spoločnosť činnosť najvyšších súdov vníma.[41] Ide predovšetkým o ich vlastnú komunikáciu, organizáciu edukačných aktivít, s ktorými sa môže verejnosť priamo s činnosťou súdu oboznámiť, no rovnako sa tvorí aj vystupovaním sudcov mimo pojednávacích miestností.[42] Na proces informovania je preto potrebné v súčasnosti nazerať optikou toho, v čom umelá inteligencia súdy pri vyhľadávaní a prijímaní informácií nedokáže nahradiť. Pozitívny záväzok aktívne informovať by mal byť preto formovaný najmä cez vnímanie, že komunikácia s verejnosťou by mala prebiehať aj mimo fyzických hraníc súdov a rovnako by mala presahovať zákonné minimum povinného zverejňovania.

Konkrétnym príkladom z českého prostredia by pre Ústavný súd SR mohlo byť doplnenie schémy postupu konania o ústavných sťažnostiach aj o možnosť podania návrhu, na základe ktorého senát iniciuje konanie o súlade právnych predpisov. Práve absencia tejto informácie mohla byť jedným z dôvodov vzniku právnej halucinácie, na ktorú sme poukázali. K širšej informovanosti by však mohli pomôcť aj krátke vysvetľujúce videá[43] a grafiky[44] o podaní ústavnej sťažnosti, o postupe konania na súde, o možnostiach rozhodnutia. Ďalšou možnosťou pre posilnenie informovania podľa českého vzoru by mohlo byť zverejnenie informácií o asistentoch ústavného súdu,[45] či samostatné sekcie o prebiehajúcich konaniach.[46] V nich je možné s týždenným predstihom vyhľadať informácie o nadchádzajúcich rozhodnutiach, akými sú napr. predmet a typ konania, informácia o navrhovateľoch, sudcovi spravodajcovi, verejnom vyhlásení nálezu. Na všetky z týchto otázok dokáže aj v rámci slovenského prostredia umelá inteligencia generovať odpovede už dnes. Na prvý pohľad by sa mohlo javiť ako samozrejmé, že na väčšinu z týchto otázok by z povahy veci nebolo možné poskytnúť správnu odpoveď. Pre jednotlivcov mimo právnického

41 KOSAŘ, D.: *Embedding Strategies of the European Apex Courts: Why Court Communication with All Segments of Society Matters*. In: *German Law Journal*, roč. 26, č. 2, 2025, s. 278.

42 *Ibid.*

43 Online: https://www.usoud.cz/informace-pro-verejnost/video-o-ustavnim-soudu.

44 Online: https://www.usoud.cz/informace-pro-verejnost/informacni-brozury.
Pri požiadavke na aktívnejšiu komunikáciu súdov sa častokrát objavuje námietka nedostatku personálnych a finančných možností súdov. Oblasť komunikácie je však priestorom, kde by súdy mohli efektívne využívať komerčné modeli AI. A to práve na tvorbu grafík, tlačových správ, príspevkov na sociálnych sieťach. Aby sa predišlo riziku halucinácií, postup by mohol spočívať v tom, že do jazykového modelu by bolo prvotne nahrané rozhodnutie, z ktorého by sa mal výstup generovať a následne by bol zadaný prompt na vytvorenie obsahu.

45 Online: https://www.usoud.cz/asistenti-soudcu.

46 Online: https://www.usoud.cz/program-jednani-soudu.

prostredia, bez predchádzajúcej znalosti o fungovaní umelej inteligencie, sa to však ako samozrejmosť javiť nemusí.

Záver

Umelá inteligencia mení doterajšie vnímanie práva na informácie. Nielenže vytvára nový spôsob ich vyhľadávania a prijímania, ale rozširuje obsah tohto práva aj o informácie, ktoré sama generuje. S častým výskytom halucinácií sa však používateľ bez kritického zhodnotenia vystavuje riziku, že bude disponovať nepravdivou informáciou. Tento prirodzený nedostatok jazykových modelov môže predstavovať pre súdy dôvod, prečo sprostredkovať informácie o vlastnej činnosti čo najjednoduchším spôsobom. Výber prostriedkov je široký - od užívateľsky jednoduchšieho, no intenzívnejšieho informovania na webových sídlach súdov, cez aktívne profily na sociálnych sieťach, až po tvorbu vlastných vysvetľujúcich grafík a videí. Nie je možné striktne vymedziť, ktoré prostriedky sú pre konkrétny súd najideálnejšie, keďže práve ich voľbou si súdy vytvárajú svoju vlastnú identitu v digitálnom priestore. Dôležitejšie ako samotná voľba prostriedkov, by malo byť pre súdy uvedomenie, že efektívna realizácia práva na informácie predpokladá proaktívne informovanie, ktoré nemá byť konkurenciou voči umelej inteligencii, ale jej protiváhou.

Zoznam použitej literatúry a zdrojov

1. COUNCIL OF EUROPE, COMMITTEE OF EXPERTS ON THE IMPLICATIONS OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR FREEDOM OF EXPRESSION (MSI-AI): *Draft Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression*, MSI-AI(2025)10, 3 June 2025, 23 s. Online: https://rm.coe.int/msi-ai-2025-10-draft-guidance-note-on-the-implications-of-generative-a/1680b68c48.
2. COURT OF JUSTICE OF THE EUROPEAN UNION, DIRECTORATE-GENERAL FOR INFORMATION: *Artificial Intelligence Strategy*, Approved by the Administrative Committee of the Court on 30 June 2023, 25 s. Online: https://curia.europa.eu/site/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf.
3. DAHL, M. et. al.: *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*. In: Journal of Legal Analysis, roč. 16, č. 1, 2024, s. 66. Online: https://doi.org/10.1093/jla/laae003.
4. DOBROVODSKÝ, R., KOŠIČIAROVA, S.: *Právo na informácie,* Krakov: Spolok Slovákov v Poľsku, 2015, 9 s.
5. Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4773af908191ad5a27d94f82af6e.
6. Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/s/t_691b4a2f86fc8191a3ac62ae1d54d20a.
7. Generovaná odpoveď modelom ChatGPT-5. Online: https://chatgpt.com/c/69282f43-6aa0-8333-8433-f324b66c9379.
8. HUMAN RIGHTS COMMITTEE: *General comment No. 34 on Art. 19: Freedoms of opinion and expression*, CCPR/C/GC/34, 12 September 2011.
9. INŠTITÚT PRE VEREJNÉ OTÁZKY: *Schopnosť ľudí odhaliť klamlivý obsah generovaný AI je veľmi nízka*. Online: https://www.ivo.sk/8987/sk/aktuality/schopnost-ludi-odhalit-klamlivy-obsah-generovany-ai-je-velmi-nizka.
10. KOSAŘ, D.: *Embedding Strategies of the European Apex Courts: Why Court Communication with All Segments of Society Matters*. In: German Law Journal, roč. 26, č. 2, 2025, s. 278. Online: https://www.cambridge.org/core/journals/german-law-journal/article/embedding-strategies-of-the-european-apex-courts-why-court-communication-with-all-segments-of-society-matters/A2DC27653A0CDA1059B1097F4F402C5C. DOI: https://doi.org/10.1017/glj.2025.3
11. LAKATOŠOVÁ, D.: *Umelú inteligenciu využíva tretina z nás. Sme svedkami*

revolúcie vo vyhľadávaní informácií? NMS, 27.1.2025. Online: https://nms.global/sk/umelu-inteligenciu-vyuziva-tretina-z-nas-sme-svedkami-revolucie-vo-vyhladavani-informacii/.
12. Nález Ústavného súdu SR sp. zn. II. ÚS 28/96 z 12. mája 1997.
13. Nález Ústavného súdu SR II. ÚS 307/2014 zo dňa 18. 12. 2014.
14. OROSZ, L., SVÁK, J. et. al.: *Ústava Slovenskej republiky, Komentár*, Zväzok I. Bratislava: Wolters Kluwer, 2021, 343 s.
15. Rozsudok ESĽP vo veci *Magyar Helsinki Bizottság v. Hungary,* sťažnosť č. 18030/11, rozsudok z 08.11. 2016.
16. STANFORD UNIVERSITY HUMAN-CENTERED ARTIFICIAL INTELLIGENCE: AI on Trial: *Legal Models Hallucinate in 1 out of 6 (or More) Benchmarking Queries*, 23 May 2024. Online: https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries.
17. UNITED NATIONS SPECIAL RAPPORTEUR ON THE PROMOTION AND PROTECTION OF THE RIGHT TO FREEDOM OF OPINION AND EXPRESSION: *Access to Information: Freedom of opinion and expression - 2013 thematic report to the Human Rights Council of the Special Rapporteur on freedom of opinion and expression*. 2013. Online: https://www.ohchr.org/sites/default/files/Documents/Issues/Expression/Factsheet_5.pdf.
18. Ústava SR č. 460/1992 Zb. v znení neskorších predpisov
19. ÚSTAVNÝ SÚD SR: *Konanie o ústavnej sťažnosti*. Online: https://www.ustavnysud.sk/informacie_pre_navrhovatelov/sprievodca-konanim-na-ussr.
20. ÚSTAVNÝ SÚD SR: *Ročenka 2024*, 2024, 66 s. Online: https://www.ustavnysud.sk/documents/d/portal/rocenka-2024-sk-pdf.

Informácie o autorovi

Mgr. Emma HUMEŇANSKÁ
Univerzita Komenského v Bratislave
Právnická fakulta
Katedra ústavného práva
Email: humenanska3@uniba.sk

14 Zodpovednosť za škodu spôsobenú AI: Potreba nového právneho rámca?

Natália Kobulská

https://doi.org/10.33542/HRF-0541-2-14

Abstrakt

Rýchly rozvoj umelej inteligencie (AI) a jej implementácia do rôznych oblastí spoločenského života prináša nové výzvy pre právne systémy, najmä v oblasti zodpovednosti za škodu. Tento článok analyzuje technické aspekty autonómneho rozhodovania AI systémov, ktoré často fungujú bez priameho ľudského zásahu, čím vzniká otázka, kto by mal niesť právnu zodpovednosť za ich konanie. Diskutuje sa o možnosti priznania právnej subjektivity AI, ako aj o alternatívnych modeloch zodpovednosti vývojára, používateľa, poskytovateľa dát či ďalších zainteresovaných subjektov. V kontexte európskeho práva sa článok venuje návrhom Európskej únie, najmä smernici o zodpovednosti za chybné výrobky (Liability Directive) a pripravovanej smernici o zodpovednosti za AI, ktoré majú za cieľ aktualizovať právny rámec tak, aby reflektoval digitálnu realitu. Nové pravidlá rozširujú definíciu výrobku o softvér a digitálne súbory, čím sa zjednodušuje uplatňovanie nárokov na náhradu škody a identifikácia zodpovedných osôb. Článok tiež hodnotí dopad týchto návrhov na národné právne poriadky, najmä v oblasti občianskoprávnej zodpovednosti, a navrhuje možné legislatívne riešenia, ktoré by zabezpečili právnu istotu bez brzdenia technologických inovácií. Cieľom je prispieť k diskusii o tom, či je potrebné vytvoriť nový právny rámec špecificky pre AI, alebo postačuje adaptácia existujúcich právnych inštitútov.

Úvod

Regulácia nových technológií je zo svojej podstaty náročná. Právne normy týkajúce sa nových technológií sa často vytvárajú v čase, keď sa technológia ešte vyvíja. Je to skôr, ako sú známe všetky jej dôsledky alebo ako je možné jasne predvídať konkrétne problémy. Niekedy už meška za vývojom, a nie je možné jej retroaktívne pôsobenie, alebo po prijatí už nevyhovuje aktuálnemu stavu, keďže technologický vývoj medzitým výrazne pokročil.

Hoci AI je v istej forme využívaná už niekoľko desaťročí, až v posledných rokoch nastal obrovský prelom a jej používanie je dnes rozšírené na takmer každú oblasť našich životov. Je preto logické, že i napriek počiatočnému váhaniu, či používanie AI je nevyhnutné regulovať, a teda aj do istej miery obmedzovať, právnymi normami, je dnes AI predmetom právnej úpravy. V súčasnosti existuje

silná podpora zavedenia jednotných predpisov, ktoré zabezpečia ochranu ľudí a majetku. Ruka v ruke s používaním nejakej veci alebo služby vždy nutne ide aj o nápravu vzniknutej škody. S tým úzko súvisí právny inštitút zodpovednosť za škodu spôsobenú AI, a to v 2 rozmeroch:

- je možné iba rozšíriť existujúcu právnu úpravu na vnútroštátnej a európskej/medzinárodnej úrovni, aby bola aplikovateľná aj na AI alebo
- je nutné vytvoriť úplne nový súbor noriem špecificky zostavených pre potreby zodpovednosti za škodu spôsobenú AI?

Regulácia umelej inteligencie je nepochybne dôležitá, ale v skutočnosti predstavuje ešte väčšiu výzvu ako regulácia iných nových technológií. Technológia umelej inteligencie nielenže rýchlo napreduje, ale systémy umelej inteligencie sú tiež schopné učiť sa, napodobňovať, prispôsobovať sa a robiť rozhodnutia spôsobmi, ktoré môžu byť neprehľadné aj pre ich vývojárov. Je schopná fungovať takpovediac autonómne, bez zásahu človeka analyzovať, robiť rozhodnutia, ktoré majú zásadný význam a vplyv na ľudí. Táto dynamická a autonómna povaha umelej inteligencie komplikuje tradičné právne prístupy založené na jasne preukázateľnej zodpovednosti a predvídateľnosti. Dôležitou otázkou, s ktorou sa zákonodarcovia stretávajú pri navrhovaní nového zodpovednostného rámca pre umelú inteligenciu, bolo vyváženie 2 faktorov: zabezpečenie bezpečnosti jej používania vrátane eliminovania rizík, a ponechanie voľnosti pre inovácie a technický pokrok. V tejto súvislosti je obzvlášť náročným, ale kľúčovým aspektom regulácie otázka zodpovednosti za škody spôsobené zariadením alebo službou prevádzkovanou umelou inteligenciou. Cieľom príspevku je zhrnúť aktuálnu právnu úpravu zodpovednosti za škodu spôsobenú umelou inteligenciou, režimy zodpovednosti vývojára, používateľa dát a ďalších zainteresovaných subjektov. Súčasne predstavím návrhy možných legislatívnych riešení do budúcna, ktoré by zabezpečili právnu istotu bez brzdenia technologických inovácií. Tieto návrhy by mali prispieť k diskusii o tom, či je potrebné vytvoriť nový právny rámec špecificky pre AI alebo postačuje adaptácia existujúcich právnych inštitútov.

14.1 Európska legislatíva

V dôsledku rastúceho významu umelej inteligencie v podnikaní a spoločnosti Európska únia pracovala na implementácii komplexného rámca pre umelú inteligenciu. Základným kameňom nového rámca pre umelú inteligenciu je nariadenie, ktorým sa ustanovujú harmonizované pravidlá pre umelú inteligenciu, ktoré neskôr nadobudlo platnosť v júli 2024 ako AI

Act/Nariadenie o umelej inteligencii.[1] Nariadenie zavádza regulačný režim pre systémy umelej inteligencie založený na riziku, ktorý ukladá prísne *ex ante* povinnosti vývojárom a používateľom umelej inteligencie s cieľom zaistiť bezpečné, transparentné a na individuálne práva orientované používanie a uplatňovanie umelej inteligencie. Nezavádza však režim občianskoprávnej zodpovednosti za škody súvisiace s umelou inteligenciou. Hoci prijatie nariadenia je považované za veľký míľnik v úprave technológií, je aj kritizované pre nedostatočné mechanizmy ochrany ľudských práv napr. pri vysoko rizikových systémoch využívaných na masové sledovanie ľudí a rozpoznávanie tváre. Porušenie povinností pri ochrane práv majú stanovené prísne dôsledky pre spoločnosti – finančné pokuty porovnateľné s porušením povinností vyplývajúcich z GDPR.

Nariadenie kombinuje 2 režimy, ktoré zahŕňajú prvky právnej ochrany výrobkov (technické požiadavky, certifikácia) s prvkami ochrany ľudských práv (rozsah porušenia práv, transparentnosť, etické normy). Aj tento hybridný prístup sa stretol s kritikou, kde sa poukazuje na rozdielnosť týchto režimov a ich prienik kreuje rôzne praktické aj konceptuálne konfrontácie.[2]

Ak systém umelej inteligencie nefunguje správne alebo má škodlivé následky – napríklad autonómny robot zraní zamestnanca továrne, lekársky nástroj umelej inteligencie poskytne fatálne odporúčanie alebo algoritmus nezákonne odmietne niekomu zamestnanie – režimy náhrady škôd sa riadia existujúcimi právnymi predpismi o zodpovednosti, ako je právo týkajúce sa deliktov, zodpovednosti za výrobky, záväzkové, prípadne pracovné právo, a to ako na európskej, tak na vnútroštátnej úrovni.

14.2 Umelá inteligencia ako chybný výrobok

Zodpovednosť za škodu spôsobenú AI na úrovni EÚ je od roku 2024 čiastkovo upravená prijatím novej *Smernice EÚ o zodpovednosti za škodu*

---

1 Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii), OJ L, 12.7.2024.

2 ALMADA, M., PETIT, N.: *The EU AI Act: Between the rock of product safety and the hard place of fundamental rights*. In: Common Market Law Review, Vol. 62, No. 1, 2025, pp. 85–120. Online: https://kluwerlawonline.com/journalarticle/Common+Market+Law+Review/62.1/COLA2025004 .

spôsobenú chybným výrobkom.[3] Vzťahuje sa však na výrobky uvedené na trh alebo do prevádzky až od 9. decembra 2026. Kľúčovou zmenou je rozšírenie pojmu „výrobok", pričom definícia zahŕňa softvér a systémy umelej inteligencie. Smernica sa uplatňuje na všetky typy softvéru vrátane aplikácií. Veľkou zmenou je zodpovednosť výrobcu za chyby, ktoré by mohli vzniknúť po aktualizácii alebo modernizácii softvéru, a ktoré boli zároveň vykonané pod jeho kontrolou. Výrobca je tiež zodpovedný za chyby, ktoré by mohli nastať v dôsledku kontinuálneho učenia sa systému umelej inteligencie, ak systém umelej inteligencie zostáva pod kontrolou výrobcu. Smernica sa tiež vzťahuje na digitálne služby nevyhnutné na fungovanie výrobku a ktoré boli zabudované do výrobku pod kontrolou výrobcu.[4] Ak je teda súčasťou produktu digitálna zložka (softvér, AI), tieto zložky tiež podliehajú režimu bezchybnosti (nesmú mať chybu, inak je chybný celý výrobok). Výrobcovia sú povinní zabezpečiť, aby produkt spĺňal štandardy kyber bezpečnosti, čím sa stane odolným voči útokom, ktoré by mohli spôsobiť škodu. Ak ku škode napriek tomu dôjde, absencia bezpečnosti je chybou výrobku.[5]

Zmena nastala aj pri práve na náhradu škody vplyvom chybného výrobku: okrem tradičných typov škody (poškodenie majetku, smrť alebo zranenie vrátane poškodenia psychického zdravia priznaného lekárom) sa zavádza nový typ – zničenie alebo poškodenie údajov, tzv. škoda na dátach. Právo na náhradu škody sa vzťahuje aj na nemajetkovú ujmu. Tento režim zodpovednosti je možné nárokovať podľa vnútroštátneho práva (čl. 6 smernice).

Smernica ďalej upravuje subjekty zodpovedné za chybné výrobky, kde je v prvom rade zodpovedný výrobca, avšak ak výrobca sídli mimo EÚ, zodpovednosť nesú dovozcovia, splnomocnení zástupcovia alebo poskytovatelia logistických služieb (čl. 8 smernice). Online platformy môžu niesť zodpovednosť ako výrobca, ak predávajú produkt pod vlastnou značkou. Poškodená osoba, ktorá utrpela škodu spôsobenú chybným výrobkom, môže na vnútroštátnom súde podať návrh na konanie o náhradu škody, pričom platí trojročná subjektívna premlčacia lehota. Podniky nesú zodpovednosť za svoje chybné výrobky v lehote desiatich rokov po umiestnení výrobku na trh. Smernica

---

[3] Smernica Európskeho parlamentu a Rady (EÚ) 2024/2853 z 23. októbra 2024 o zodpovednosti za chybné výrobky a o zrušení smernice Rady 85/374/EHS, OJ L, 18.11.2024.

[4] EUROPEAN UNION: *Chybné výrobky - zodpovednosť*. EUR-Lex, posledná aktualizácia 19.8.2025. Online: https://eur-lex.europa.eu/SK/legal-content/summary/defective-products-liability.html.

[5] Porov. TACHE, C. E. P., VÂLCU, E. N.: *Artificial intelligence and corporate liability towards a new legal-ethical contract in the dynamics of emerging global human rights convergences*. In: Juridical Tribune - Review of Comparative and International Law, Vol. 15, No. 2, 2025, s. 290. Online: https://doi.org/10.62768/TBJ/2025/15/2/04.

zavádza právo poškodeného požadovať od výrobcu sprístupnenie dokumentov potrebných na dokazovanie. Ak výrobca odmietne poskytnúť dôkazy, uplatní sa vyvrátiteľná domnienka chyby výrobku a príčinnej súvislosti. Tá sa uplatní aj v prípade, ak je preukázané, že výrobok nie je v súlade s platnými požiadavkami na bezpečnosť výrobkov alebo ak škoda bola spôsobená zjavne nesprávnym fungovaním výrobku počas jeho rozumne predvídateľného použitia alebo za bežných okolností. Tento prístup zaručuje lepšiu vymožiteľnosť nárokov vyplývajúcich zo zodpovednosti za škodu.

14.2.1 Smernica o zodpovednosti za škodu spôsobenú AI

Nová smernica však zahrnula AI ako nutné minimum, keďže hlavný právny akt – smernica o zodpovednosti za škodu spôsobenú AI[6] nebolo doposiaľ prijaté pre nedostatok politického konsenzu medzi členskými štátmi, z obáv pred duplicitnou úpravou, keďže umelá inteligencia už je čiastkovo upravená v právnych predpisoch, ďalej boli proti samotné spoločnosti, ktoré vyvíjajú alebo predávajú AI systémy (zvýšili by sa im náklady na zaistenie konformity so smernicou a zabrzdilo by to inovácie), a v neposlednom rade aj ekonomické hľadisko: prísna regulácia by odradila subjekty investovať na EÚ trhu.

Návrh smernice obsahoval komplexné pokrytie úpravy zodpovednosti za škodu ako napr. nezmluvnú občianskoprávnu zodpovednosť vychádzajúcu z nedbanlivosti vo vzťahu k škodám spôsobeným AI systémami, predpoklad príčinnej súvislosti, zachovanie a sprístupnenie dôkazov od poskytovateľov vysokorizikových systémov AI na účely určenia zavinenia a príčiny škody. Návrh smeroval k vyrovnanejšiemu procesu pre obete škôd spôsobených AI – uľahčením dokazovania kauzality a sprístupnením interných dát AI systémov. Tieto opatrenia by mali kompenzovať technologické výhody AI, ktoré by inak mohli brániť efektívnemu uplatňovaniu nárokov. Smernica by bola doplňujúca k už existujúcej právnej úprave na úrovni EÚ, keďže preberá pojmy ako napr. „vysoko rizikové AI systémy“ a zmiernenie podmienok pre náhradu škody pre poškodených (predpoklad zavinenia, príčinnej súvislosti a ľahší prístup k dôkazom).

---

6 Návrh smernice definitívne stiahnuté vo februári 2025 (COM(2022) 496 final – Návrh Smernica Európskeho parlamentu a Rady o prispôsobení pravidiel mimozmluvnej občianskoprávnej zodpovednosti umelej inteligencii (smernica o zodpovednosti za umelú inteligenciu) 2022/0303 (COD)).

14.2.2 Porovnanie EÚ úpravy

Európska úprava umelej inteligencie síce nie je zatiaľ komplexná, ale poskytuje dostatočný rámec pre prevenciu vzniku škody, a aj následné uplatnenie zodpovednosti za škodu, aj keď iba za chybný výrobok, ktorého je AI súčasťou. V tabuľke je prehľadne zhrnutá právna úprava na európskej úrovni, a to vrátane stiahnutej smernice, aby bolo porovnanie ucelené.

Oblasť	**AI Act (nariadenie)**	**AI Liability Directive (smernica - stiahnutá)**	**Smernica o zodpovednosti za chybné výrobky (nová)**
Cieľ	Prevencia rizík pri používaní AI	Uľahčenie dokazovania a náhrady škody spôsobenej AI	Revidované pravidlá pre zodpovednosť za výrobky vrátane softvéru
Typ regulácie	Technické a organizačné požiadavky	Procesné pravidlá dokazovania (prezumpcia príčinnej súvislosti, prístup k dôkazom)	Zavedená objektívna zodpovednosť
Kedy sa uplatňuje	Pred uvedením na trh a počas prevádzky	Po vzniku škody	Po vzniku škody
Rozsah	Všetky AI systémy	Škody spôsobené vysokorizikovými AI systémami	Škody spôsobené chybným výrobkom vrátane digitálnych služieb a softvéru

Je možné ukázať praktické uplatnenie EÚ predpisov na modelovom prípade: autonómne vozidlo spôsobí dopravnú nehodu, pri ktorej utrpí chodec vážne zranenia. Pred nehodou musí výrobca vozidla splniť požiadavky Nariadenia o AI: systém je klasifikovaný ako vysoko rizikový, takže výrobca musí splniť prísne podmienky dané nariadením, vykonať bezpečnostné testy, zabezpečiť transparentnosť, uchovať dokumentáciu a zaviesť mechanizmy na monitorovanie po uvedení na trh. Ak sa zistí porušenie, môžu mu byť uložené pokuty alebo zakázaný ďalší predaj systému. Po nehode by sa uplatnilo nariadenie o zodpovednosti za škodu spôsobenú AI (keby bolo prijaté): ak výrobca alebo prevádzkovateľ porušil povinnosti (napr. z AI Actu), platí prezumpcia príčinnej súvislosti a predpokladá sa, že škoda bola spôsobená kvôli porušeniu povinnosti. Poškodený má jednoduchší spôsob, ako uniesť dôkazné bremeno a môže požiadať o prístup k dátam vzťahujúcim sa na autonómne

vozidlo (rozhodovacie algoritmy, logy, atď.). Ak budeme postupovať podľa Smernice o zodpovednosti za chybný výrobok, po nehode poškodený nemusí dokazovať zavinenie (platí objektívna zodpovednosť – *strict liability*) – stačí preukázať, že vozidlo ako výrobok bolo chybné a spôsobilo škodu. Zodpovednosť za škodu nesie v tomto prípade výrobca, aj keď dodržal všetky povinnosti z AI Actu (prevencia).

14.3 Technické aspekty AI a ich vplyv na zodpovednosť za škodu

Pre správne fungovanie umelej inteligencie a jej právnej regulácie na zabezpečenie ochrany práv jednotlivých subjektov, ktoré s ňou prichádzajú do kontaktu a sú ovplyvnení jej používaním či výsledkami dosiahnutými pomocou AI, je nevyhnutná spolupráca vývojárov a právnickej obce. AI je možno najprenikavejším príkladom tohto prieniku práva a technológií, a dosiahnuť dialóg a efektívne riešenia je veľmi náročné.

Podľa môjho názoru je nutné hlbšie skúmať technické pozadie AI, aby sme vedeli presnejšie pomenovať a adresovať problémy spojené s vyvodením zodpovednosti za škodu. Medzi hlavné technické aspekty patria:

1. modely strojového učenia a ich blackbox problém,
2. dátová závislosť,
3. autonómia rozhodovania AI
4. ľudský dohľad (tzv. human-in-the-loop) a zároveň náchylnosť ľudí k automatizácii.

Modely strojového učenia sú založené na miliónoch parametrov hlbokých neurónových sietí, teda nie je možný podrobný tracking rozhodovacieho procesu AI. Je náročné, a často aj nemožné určiť pravidlá, na základe ktorých AI rozhodla (tzv. blackbox problém). Tým, že nevidíme pozadie rozhodovacieho procesu, nevieme, komu pričítať zavinenie: má byť zodpovedný vývojár, prevádzkovateľ, používateľ, alebo je možné zodpovednosť do istej miery vyvodiť voči AI, hoci zatiaľ nemá právnu subjektivitu? Európska legislatíva reaguje zavedením povinností pre poskytovateľov a používateľov, prezumpcie príčinnej súvislosti a požiadavkami na transparentnosť. I keď sa začali objavovať náznaky, že umelá inteligencia by mala mať právnu subjektivitu, keďže dokáže autonómne „myslieť“ a prijímať rozhodnutia podobne ako človek (ktorý tiež robí chybné rozhodnutia), v súčasnosti je tento koncept eticky a právne neudržateľný. Je možné, že v budúcnosti bude plne využiteľná sebauvedomelá/ silná/ nadľudská AI, ale zatiaľ umelá inteligencia nemá právnu subjektivitu, a teda nemôže byť priamo voči nej vyvodzovaná zodpovednosť. K tomu prispieva fakt, že každý

systém AI musí mať ľudský dohľad, teda zodpovednosť nesie človek – vývojár, používateľ či prevádzkovateľ.

Ďalším problémovým aspektom je závislosť na dátach. AI funguje na spracovaní a čerpaní dát, ktoré sú do nej vložené, a teda kvalita vstupných dát vie značne ovplyvniť rozhodovací proces AI. Navyše AI nedáva na rovnakú otázku vždy rovnakú odpoveď, čo je do istej miery ovplyvnené parametrami veľkých jazykových modelov a mierou autonómnosti.

Na to sa viaže samoučiaci mechanizmus a alternácia rozhodnutí na základe spätnej väzby. Ak uvažujeme o zodpovednosti za škodu spôsobenú chybným výrobkom, nie je možné AI, ktorá sa v čase mení práve spomenutými faktormi, považovať za výrobok v tradičnom zmysle slova, kde je bežný, hoci technologicky zložitý stroj uvedený do používania a počas fungovania nemení podstatne svoje vlastnosti.

Dôležitým prvkom je aj náchylnosť ľudí k automatizácii, na ktorú nadväzuje jednoduchosť užívateľského rozhrania a nižšia miera ľudského dohľadu nad výstupmi umelej inteligencie.

Technické vlastnosti AI priamo ovplyvňujú aplikáciu právnych režimov:

- Pri rozdelení zodpovednosti: vývojár zodpovedá za chyby v algoritme, nedostatočné bezpečnostné opatrenia, analýzu rizík a komplexné testovanie

  Prevádzkovateľ za monitoring funkcionality, správnu implementáciu a aktualizácie

  Používateľ za zneužitie, nesprávne použitie a interpretáciu, správnosť vstupných dát, human-in-the-loop prístup vrátane kontroly výstupov, označenie výstupu, že bol generovaný AI

- Uplatňovanie objektívnej zodpovednosti (postupné úplné opustenie konceptu subjektívnej zodpovednosti), prezumpcia zavinenia a právo obetí na prístup k dôkazom (systém musí byť transparentný)

- Kategorizácia systémov podľa miery rizika: vysoko rizikové systémy vyžadujú splnenie prísnejších pravidiel.

  Pred vyvodením zodpovednosti za škodu je nutná správna analýza prípadu a zohľadnenie technických aspektov rozobraných vyššie, až následne sa aplikuje správny právny režim.

14.4 Návrh právnej úpravy od budúcna

Právna úprava umelej inteligencie je vo vzťahu k zodpovednosti za škodu ešte neúplná, čo na jednej strane predstavuje neistotu, na druhej strane však poskytuje priestor reagovať na inovatívne iniciatívy a ďalej hľadať čo

najuniverzálnejšie riešenia problematických oblastí[7]. V závere tejto kapitoly je prehľadná tabuľka, ktorá ukazuje, kde sa nachádzame dnes a navrhuje predpokladané riešenia do budúcna. Medzi kategórie, ktoré pravdepodobne budú v budúcnosti zmenené, patria:

- Objektívna zodpovednosť za škodu spôsobenú umelou inteligenciou - zaviesť režim objektívnej zodpovednosti za vysoko rizikové systémy umelej inteligencie, aby sa komplexne pokryli fyzické, virtuálne a ekonomické škody.
- Obrat dôkazného bremena - zaviesť vyvrátiteľné domnienky zavinenia, aby sa uľahčilo súdne konanie pre obete v prípade porušenia.
- Rozšírenie rozsahu zodpovednosti - zahrnúť nehmotné škody, ako sú zaujatosť a diskriminácia, narušenie súkromia a ekonomické straty, do rozsahu zodpovednosti za systémy umelej inteligencie.
- Harmonizácia definícií - zosúladenie kategórií rizík a definícií v pravidlách zodpovednosti a nariadení o umelej inteligencii, prechod na klasifikáciu systémov s „vysokým dopadom".
- Posilnenie vymožiteľnosti - posilnenie vymožiteľnosti prostredníctvom nezávislých auditov a dozorného orgánu na úrovni EÚ s cieľom zlepšiť vyvodzovanie zodpovednosti.

[7] Porov. FORGÓ, N., NECATI PEHLIVAN, C., VALCKE, P. (eds.): *AI Governance and Liability in Europe: A Primer*. (1 ed.) Wolters Kluwer, 2025, s. 472.

	Pôvodné nariadenie	**Aktuálna úprava**	**Návrhy do budúcna**
Cieľ	Harmonizácia procesných pravidiel pre nároky na náhradu škody súvisiacej s AI	Uplatňovať nariadenie o umelej inteligencii + smernicu o zodpovednosti za chybné výrobky + vnútroštátne zákony o deliktoch	soft law, špecifické odvetvové pravidlá alebo nový legislatívny balík
Kľúčové aspekty	Predpoklad príčinnej súvislosti (objektívna zodpovednosť)	- Nariadenie o AI: povinnosti v oblasti dodržiavania predpisov - Smernica o zodpovednosti za výrobky: chybné výrobky - Vnútroštátne právne predpisy: zodpovednosť za škodu	Kódexy správania Usmernenia pre súdy
Typ zodpovednosti	Na základe zavinenia (bez novej objektívnej zodpovednosti)	Zmiešané: objektívna zodpovednosť za chybné výrobky + subjektívna zodpovednosť za ostatné	V budúcnosti by sa mohla zaviesť hybridná alebo objektívna zodpovednosť
Úroveň harmonizácie	Čiastočná (len procesné)	Nízka (nejednotné vnútroštátne predpisy)	Stredná až vysoká, ak sa prijmú nové smernice alebo sektorové pravidlá
Dopad na spoločnosti (business)	Zvýšené riziko súdnych sporov, povinnosti týkajúce sa transparentnosti	Dodržiavanie nariadenia o AI + pravidiel bezpečnosti výrobkov	Závisí od prístupu: poistenie alebo soft law = nižšia záťaž
Status	Stiahnuté (február 2025)	Platné PP: Nariadenie o AI (2024), Smernica o zodpovednosti za chybné výrobky (2024)	Diskusia; pravdepodobne najskôr nezáväzné usmernenia; opätovné prerokovanie stiahnutého nariadenia

Záver

Právna úprava zodpovednosti za škodu prešla za posledné roky veľkým vývojom. Viaceré medzinárodné organizácie pripravili odporúčania či nezáväzné predpisy, ako AI regulovať a pristupovať k jej regulácii eticky, s ohľadom na ochranu ľudských práv i práva na náhradu škody. Na úrovni EÚ boli prijaté dva záväzné právne akty, ktoré čiastočne zodpovednosť za škodu spôsobenú AI regulujú. Nie je možné s istotou tvrdiť, že bude EÚ či medzinárodná úprava pre celú oblasť zodpovednosti za škodu spôsobenú AI v najbližšom období úplne harmonizovaná. Najpravdepodobnejšie budeme sledovať vývoj, skutočnú potrebu regulovať AI a do akej miery, príp. ako prísne, riadiť sa soft law, súdnymi rozhodnutiami či usmerneniami, aby vznikla dôvera k AI - ak k systému nebude dôvera, nebude sa používať. Súčasťou nadviazania dôvery je právna istota pre používateľov, že v prípade vzniku škody bude vyvodzovaná zodpovednosť a zjednaná adekvátna náprava škody.

Zoznam použitej literatúry a zdrojov

1. ALMADA, M., PETIT, N.: *The EU AI Act: Between the rock of product safety and the hard place of fundamental rights.* In: Common Market Law Review, Vol. 62, No. 1, 2025, pp. 85–120, Online: https://kluwerlawonline.com/journalarticle/Common+Market+Law+Review/62.1/COLA2025004 .
2. BUITEN, M.C.: *Product liability for defective AI*. In: European Journal of Law and Economics, Vol. 57, 2024, pp. 239–273. Online: https://doi.org/10.1007/s10657-024-09794-z.
3. EUROPEAN UNION: *Chybné výrobky – zodpovednosť*. EUR-Lex, naposledy aktualizované 19.8.2025. Online: https://eur-lex.europa.eu/SK/legal-content/summary/defective-products-liability.html.
4. FORGÓ, N., NECATI PEHLIVAN, C., VALCKE, P. (eds.): *AI Governance and Liability in Europe: A Primer*. (1 ed.) Wolters Kluwer, 2025. Online: https://ucrisportal.univie.ac.at/en/publications/ai-governance-and-liability-in-europe-a-primer/.
5. Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 z 13. júna 2024, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii), OJ L, 12.7.2024.
6. Návrh Smernica Európskeho parlamentu a Rady o prispôsobení pravidiel mimozmluvnej občianskoprávnej zodpovednosti umelej inteligencii (smernica o zodpovednosti za umelú inteligenciu) 2022/0303 (COD)).
7. Smernica Európskeho parlamentu a Rady (EÚ) 2024/2853 z 23. októbra 2024 o zodpovednosti za chybné výrobky a o zrušení smernice Rady 85/374/EHS, OJ L, 18.11.2024.
8. TACHE, C. E. P., VÂLCU, E. N.: *Artificial intelligence and corporate liability towards a new legal-ethical contract in the dynamics of emerging global human rights convergences.* In: Juridical Tribune - Review of Comparative and International Law, Vol. 15, No. 2, 2025, pp. 281–305. DOI: https://doi.org/10.62768/TBJ/2025/15/2/04.

Informácie o autorovi
JUDr. Natália KOBULSKÁ, LL.M.
Univerzita Pavla Jozefa Šafárika v Košiciach
Filozofická fakulta
Dekanát
Email: natalia.kobulska@upjs.sk

15 Digitálny konštitucionalizmus: cesta k rešpektovaniu ľudských práv v dobe algoritmov?

Soňa Juhásová

https://doi.org/10.33542/HRF-0541-2-15

Abstrakt

S rastúcim vplyvom digitálnych platforiem na verejný diskurz sa limity ľudských práv prestávajú formovať výlučne v rámci štátnych inštitúcií a súdov. Koncept digitálneho konštitucionalizmu, inak nazývaného aj súkromný konštitucionalizmus, vystihuje tento posun tým, že opisuje, ako súkromní aktéri, najmä digitálne platformy, vytvárajú a presadzujú vlastné normatívne systémy, ktoré svojou štruktúrou pripomínajú ústavné poriadky. Tieto systémy určujú, aké prejavy sú prípustné, ako sa chráni súkromie a aké mechanizmy nápravy sú dostupné, a to všetko mimo rámca demokratickej legitimity a verejnoprávnej zodpovednosti. Príspevok ponúka teoretickú reflexiu otázky, či sa tento model môže vyvinúť do digitálnej správy založenej na ľudskoprávnych štandardoch, najmä v kontexte algoritmického rozhodovania. Zároveň analyzuje napätie medzi zmluvnou slobodou spoločností, algoritmickou správou a univerzalitou ľudských práv. Ako ilustráciu uvádza prípad Meta Oversight Board, ktorý predstavuje pokus o implementáciu princípov proporcionality, transparentnosti a riadneho procesu do vnútorného rozhodovania platformy. Príspevok zároveň poskytuje kritickú analýzu limitov tohto druhu konštitucionalizmu, poukazujúc najmä na nedostatok demokratickej legitimity, obmedzenú transparentnosť a nedostatočné mechanizmy zodpovednosti. Diskusia je zasadená do rámca formujúceho sa digitálneho konštitucionalizmu Európskej únie, ktorý sa prostredníctvom Aktu o digitálnych službách (DSA) usiluje o reflektovanie ústavných princípov aj do súkromnej správy online priestoru.

Úvod

Písal sa 6. január 2021. Vo Washingtone D.C. sa mali spočítať hlasy voliteľov v súlade s výsledkami prezidentských volieb. Vtedajší americký prezident Donald Trump zverejnil na svojom Facebookovom účte sériu príspevkov, v ktorých spochybňoval legitimitu volieb a nepriamo povzbudzoval svojich priaznivcov, aby sa im vzopreli . Niekoľko hodín nato došlo k útoku na Kapitol, na čo Facebook zareagoval bezprecedentným krokom – zablokoval účet

úradujúceho prezidenta Spojených štátov na neurčito.[1] Rozhodnutie, ktoré v demokratickom štáte prináleží súdu či zákonodarnej moci, sa tak odohralo v rámci interných procesov súkromnej spoločnosti.

O niekoľko mesiacov neskôr o opodstatnenosti tohto kroku rozhodol Meta Oversight Board (ďalej ako „MOB"), samoregulačný orgán spoločnosti Meta, ktorý svojou štruktúrou, postupmi a argumentáciou pripomína ústavný súd. MOB preskúmala, či bol zásah v súlade s princípmi proporcionality, nevyhnutnosti a ochrany slobody prejavu, a odporučila spoločnosti Meta jasnejšie pravidlá pre obdobné situácie.[2]

Tento proces sa odohral mimo rámca verejného práva, no v duchu ústavných princípov, čím ilustruje presun rozhodovacej moci o základných právach z verejných inštitúcií na súkromné platformy. Prípad s americkým prezidentom tak upriamil pozornosť na nastupujúci fenomén rozhodovania v digitálnom priestore, kedy súkromné subjekty – digitálne platformy vykonávajú normotvornú, rozhodovaciu i sankčnú právomoc a ich rozhodnutia majú dopad na základné práva používateľov.

Cieľom tohto článku je analyzovať, ako sa prostredníctvom tohto posunu formuje koncept digitálneho konštitucionalizmu, teda prenos ústavných princípov a hodnôt do súkromnej správy online priestoru. Zámerom je preskúmať, či sa tento model môže vyvinúť do formy digitálnej správy založenej na ľudskoprávnych štandardoch, a predovšetkým aké limity predstavujú absencia demokratickej legitimity, nedostatočná transparentnosť a zodpovednosť.

Prínosom predkladaného článku je aj to, že v slovenskej právnickej akademickej literatúre doposiaľ chýba systematické spracovanie fenoménu digitálneho konštitucionalizmu. Cieľom textu je preto čiastočne zaplniť túto medzeru a otvoriť diskusiu o tom, ako sa ústavné princípy premietajú do správy digitálneho priestoru, najmä v kontexte európskeho práva.

Metodologicky sa tento článok opiera o kvalitatívnu analýzu prameňov digitálneho konštitucionalizmu s komparáciou prístupov Európskej únie (najmä Aktu o digitálnych službách, ďalej ako „DSA") a modelu samoregulácie regulácie platforiem, reprezentovaného prípadom MOB. Analýza vychádza z interpretácie

---

1 HERN, A.; PAUL, K.: *Donald Trump Twitter ban comes to end amid calls for tougher action*. The Guardian, 7. január 2021. Online: https://www.theguardian.com/us-news/2021/jan/07/donald-trump-twitter-ban-comes-to-end-amid-calls-for-tougher-action.

2 META OVERSIGHT BOARD: *Decision on the indefinite suspension of Donald J. Trump's accounts* (Case No. FB-691QAMHJ). Oversight Board, 5. máj 2021. Online: https://www.oversightboard.com/decision/fb-691qamhj/.

akademickej literatúry o digitálnom konštitucionalizme a skúma, ako a či sa princípy proporcionality, transparentnosti a riadneho procesu premietajú do praxe platforiem.

Po úvode sa v prvej časti článku pozornosť venuje teoretickým východiskám digitálneho konštitucionalizmu a otázke, ako môže automatizované rozhodovanie ovplyvňovať ľudské práva. Druhá časť ponúka prehľad v rámci EÚ, sledujúci vývoj od obdobia „digitálneho divokého západu" k formujúcemu sa modelu digitálneho konštitucionalizmu. Tretia časť analyzuje digitálny konštitucionalizmus v praxi na príklade MOB ako pokusu o implementáciu ústavných princípov v súkromnej správe online priestoru. Štvrtá časť identifikuje kritické momenty MOB, limity súčasného vývoja digitálneho konštitucionalizmu. Záverečná časť článku prináša zhrnutie a odporúčania pre tvorcov právnych rámcov.

15.1 Teoretické východiská digitálneho konštitucionalizmu – ako algoritmy rozhodujú o ľudských právach?

Pojem digitálny konštitucionalizmus vznikol ako reakcia na rastúcu moc súkromných digitálnych platforiem, ktoré svojím rozhodovaním čoraz viac zasahujú do priestoru tradične vyhradeného verejnej moci. Kým klasický konštitucionalizmus predpokladal, že obmedzovanie práv vychádza zo štátnej moci, v digitálnej ére túto úlohu čoraz viac preberajú súkromné digitálne platformy. Tie tak určujú hranice slobody prejavu, súkromia a prístupu k informáciám, čím sa stávajú novými kvázi-ústavnými aktérmi bez formálnej demokratickej legitimity.[3]

V tradičnom ponímaní je primárnou úlohou konštitucionalizmu a ústavného práva predovšetkým ochrana základných práv a vzájomné vyvažovanie jednotlivých zložiek mocí, čo vedie k nastoleniu ústavnej rovnováhy.[4] Možno argumentovať, že rozmach digitálnych technológií tento stav zásadne ovplyvnil.

Čo sa týka ochrany základných ľudských práv, digitálne platformy v tomto smere predstavujú dvojsečnú zbraň. Na jednej strane totiž rozširujú možnosti jednotlivcov pri výkone základných práv, keďže výrazne zvyšujú kapacitu

---

[3] Pozri DE GREGORIO, G.: *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge: Cambridge University Press, 2022; CELESTE, E.: *Digital constitutionalism: The role of internet bills of rights*. London – New York: Taylor & Francis, 2022.

[4] PETERS, A.: *Compensatory constitutionalism: the function and potential of fundamental international norms and structures.* In: Leiden Journal of International Law, roč. 19, č. 3, 2006, s. 579–610.

prenášať a prijímať informácie, čím z ústavného hľadiska posilňujú výkon práv založených na výmene informácií – najmä slobodu prejavu, slobodu náboženského vyznania, slobodu zhromažďovania či slobodu podnikať.[5]

Na druhej strane, rozvoj digitálnych technológií prináša aj nové riziká pre výkon základných práv. Mechanizmy, ktoré umožňujú slobodnú výmenu informácií, sa môžu zároveň stať nástrojom ich zneužívania, napríklad prostredníctvom šírenia nenávistných prejavov, kyberšikany, ohovárania či detskej pornografie. Digitálne platformy navyše disponujú technickými prostriedkami na obmedzovanie, monitorovanie a zaznamenávanie komunikácie používateľov. Takéto zásahy môžu ovplyvniť nielen slobodu prejavu, informácií či združovania, ale v prípadoch spracúvania dôverných údajov alebo osobných dát aj právo na súkromie a ochranu osobných údajov.[6]

Rovnováha jednotlivých zložiek mocí bola v tomto ohľade značným spôsobom modifikovaná. Kým tradične sme predpokladali dichotómiu moci medzi štátom a jednotlivcom, digitálna éra priniesla na scénu tretieho hráča – veľké digitálne platformy. Tieto spoločnosti vytvárajú vlastné kvázi normatívne poriadky, ktoré určujú rozsah a limity práv používateľov, keďže prostredníctvom svojich zmluvných podmienok, algoritmov a procesov na moderovanie obsahu *de facto* uplatňujú výkon verejnej moci.[7] Tento výkon moci súkromnými aktérmi, ktorí na globálnej úrovni vykonávajú funkcie podobné verejnej správe, je jednou z ústredných tém, ktorými sa zaoberá koncept digitálneho konštitucionalizmu pri zmene ústavnej rovnováhy.

Online platformy disponujú osobitnou formou súkromnej moci najmä vďaka svojej schopnosti zhromažďovať a monetizovať obrovské množstvo obsahu a spracúvať dáta. Pôsobia tak ako digitálni strážcovia prístupu[8], ktorí rozhodujú o tom, aké informácie sa k používateľom dostanú a ako môžu jednotlivci uplatňovať svoje práva v online prostredí.

---

5 CELESTE, E.: *Digital constitutionalism: a new systematic theorisation*. In: International Review of Law, Computers & Technology, roč. 33, č. 1, 2019, s. 79.

6 *Ibid*.

7 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 57; BELLI, L., VENTURINI, J.: *Private ordering and the rise of terms of service as cyber-regulation*. In: Internet Policy Review, roč. 5, č. 4, 2016; CELESTE, E.: *Terms of Service and Bills of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment?* In: International Review of Law, Computers & Technology, roč. 33, č. 2, 2018.

8 LAIDLAW, E. B.: *Private Power, Public Interest: An Examination of Search Engine Accountability*. In: International Journal of Law and Information Technology, roč. 17, č. 1, 2009; LAIDLAW, E. B.: *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility*. Cambridge: Cambridge University Press, 2015.

Je nevyhnutné zdôrazniť, že platformy rozhodujú o právach používateľov predovšetkým prostredníctvom moderovania obsahu, ktorý predstavuje jeden z najviditeľnejších prejavov ich moci.[9] Moderovanie obsahu možno charakterizovať ako „*všetky situácie, kedy platforma zasahuje do komunikácie medzi používateľmi, definujúc pri tom pravidlá komunikácie, formy zapojenia sa a spôsoby šírenia obsahu*.“[10] Takto definované moderovanie obsahu nemožno vnímať len ako technickú operáciu, ale predstavuje vlastný normotvorný proces, v rámci ktorého súkromné spoločnosti rozhodujú o rozsahu a podobe základných práv v digitálnom priestore, najmä o slobode prejavu.

Možno zhrnúť, že digitálny konštitucionalizmus sa snaží o zodpovedanie otázky, ako stanoviť hranice výkonu moci platformami v algoritmickej spoločnosti.[11] Predstavuje snahu nanovo premyslieť, akým spôsobom má byť výkon moci v digitálnej ére obmedzovaný a čo predstavuje jeho zdroj legitimity. Ide o súbor princípov a hodnôt, ktoré usmerňujú a formujú normatívne reakcie na výzvy, ktoré prináša rozvoj digitálnych technológií. V nasledujúcej časti článok stručne opíše vývoj normatívnych reakcií na území Európskej únie počas posledných desaťročí.

15.2 Európsky prehľad: od digitálneho divokého západu k digitálnemu konštitucionalizmu

Politika Európskej únie v oblasti digitálnych technológií prešla za posledné dve desaťročia zásadnou premenou – od liberálneho ekonomického prístupu k ústavne orientovanému modelu správy digitálneho priestoru. Tento posun možno rozdeliť do troch vývojových etáp: digitálny liberalizmus, súdny aktivizmus, ktoré vyústili do súčasne sa formujúceho európskeho digitálneho konštitucionalizmu.[12]

15.2.1 Prvá fáza: Digitálny liberalizmus

Počiatočné obdobie bolo formované logikou vnútorného trhu a ekonomickej integrácie členských štátov. Digitálne prostredie sa vnímalo ako

---

9 DENARDIS, L., HACKL, A. M.: *Internet governance by social media platforms*. In: Telecommunications Policy, roč. 39, č. 9, 2015.

10 DA CONCEIÇÃO, L. H. M.: *A constitutional reflector? Assessing societal and digital constitutionalism in Meta's Oversight Board.* In: Global Constitutionalism, roč. 13, č. 3, 2024, s. 585.

11 PADOVANI, C., SANTANIELLO, M.: *Digital constitutionalism: Fundamental rights and power limitation in the Internet eco-system*. In: International Communication Gazette, roč. 80, č. 4, 2018.

12 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: *International Journal of Constitutional Law*, roč. 19, č. 1, 2021.

priestor, do ktorého by verejná moc mala zasahovať len minimálne, pričom sa predpokladalo, že regulácia zdola od samotných spoločností poskytne efektívnejší rámec ako štátna centralizovaná regulácia prostredníctvom právnych noriem, ktorá sa vnímala ako prekážka inovácií.[13] Technológie boli chápané primárne ako prostriedok hospodárskeho rastu a technologických inovácií a nie ako nástroj výkonu moci.

Prístup EÚ bol preto založený na štyroch základných slobodách vnútorného trhu. Typickým prejavom tohto prístupu bola Smernica o elektronickom obchode[14], ktorej cieľom bolo zabezpečiť voľný pohyb služieb informačnej spoločnosti a fungovanie vnútorného trhu. Zaviedla režim výnimiek zo zodpovednosti pre poskytovateľov internetových služieb a tzv. online sprostredkovateľov, ktorým priznala pasívnu úlohu – teda výnimku zo zodpovednosti za obsah tretích strán, pokiaľ si neboli vedomí jeho nezákonnosti.[15] Tento rámec však nepriamo umožnil rýchly rast online platforiem, ktoré získali značný priestor pri určovaní rozsahu ochrany základných práv prostredníctvom vlastných mechanizmov moderovania obsahu.

15.2.2 Druhá fáza: Súdny aktivizmus

Odklon od digitálneho liberalizmu a začiatok novej éry v európskom vývoji možno pozorovať v dôsledku dvoch primárnych javov: čiastočná premena povahy online platforiem a zvyšovanie významu Charty základných práv EÚ po prijatí Lisabonskej zmluvy.[16]

V prvom rade sa v tomto čase začala výraznejšie meniť povaha online platforiem. Kým na konci 20. storočia boli hostingoví poskytovatelia vnímaní ako pasívni aktéri, ktorí len technicky sprístupňujú alebo ukladajú obsah tretích strán, v tom čase sa zvýraznila rola sociálnych sietí, ktoré začali aktívne organizovať digitálny priestor. Ich dátovo orientovaný biznis model stojí na profilovaní používateľov a cielení reklamy, čo si vyžaduje intenzívne spracúvanie údajov a automatizované moderovanie obsahu.[17] Platformy už preto nebolo

---

13 HARDY, I. T.: *The proper legal regime for cyberspace*. In: University of Pittsburgh Law Review, roč. 55, 1993.

14 Smernica 2000/31/ES Európskeho parlamentu a Rady z 8. júna 2000 o určitých právnych aspektoch služieb informačnej spoločnosti na vnútornom trhu, najmä o elektronickom obchode (smernica o elektronickom obchode),

15 Smernica o elektronickom obchode, čl. 14.

16 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 50

17 SRNICEK, N.: *Platform Capitalism*. Cambridge: Polity Press, 2017; ESTEVE, A.: *The Business of Personal Data: Google, Facebook, and Privacy Issues in the EU and the USA*. In: International Data Privacy Law, roč. 7, 2017.

možné považovať len za neutrálnych poskytovateľov, keďže prostredníctvom algoritmov a moderovania obsahu začali selektovať, odporúčať, pretláčať a potláčať určité typy prejavov, čím aktívne formovali online diskurz.[18]

Taktiež, v roku 2017 došlo k podpisu Lisabonskej zmluvy, čo znamenalo, že základné ľudské práva obsiahnuté v Charte nadobudli rovnakú právnu silu ako zakladajúce zmluvy a ich dodržiavanie sa stalo záväzné pre členské štáty ako aj pre inštitúcie EÚ.[19] V takto posilnenom ústavnom prostredí začal Súdny dvor EÚ systematicky používať Chartu ako základný nástroj pri riešení výziev digitálnej éry. Keďže smernica o elektronickom obchode neprešla dlhé zmenou, práve súdna aktivita Súdneho dvora upozornila na konflikty medzi ekonomickými záujmami a ochranou základných práv v informačnej spoločnosti. Súdny dvor sa v rámci svojej rozhodovacej činnosti posunul od čisto ekonomického pohľadu, typického pre predchádzajúcu liberálnu fázu, k prístupu založenému na ochrane základných práv.[20]

V tejto súvislosti zohrali dôležitú úlohu viaceré rozhodnutia Súdneho dvora. Pri online obsahu sa Súdny dvor odklonil od pôvodného chápania sprostredkovateľov ako pasívnych aktérov (napr. Google France[21]) a uznal, že niektoré platformy vykonávajú aktívnu úlohu prostredníctvom automatizovaných technológií ako napríklad optimalizácia, prezentácia alebo propagácia ponúk, ako to potvrdil vo veci L'Oréal.[22] V prípadoch Scarlet Extended a Netlog[23] následne vyvažoval ekonomické záujmy platforiem, najmä slobodu podnikať, s právom používateľov na ochranu osobných údajov a slobodu prejavu pri posudzovaní zákazu všeobecného monitorovania.[24]

Možno konštatovať, že hoci súdny aktivizmus úplne nevyriešil otázky týkajúce sa zodpovednosti platforiem alebo transparentnosti pri moderovaní obsahu, poukázal na možné ohrozenie základných práv, ktoré predstavuje nárast

---

18 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 50

19 Charta základných práv Európskej únie, čl. 51.

20 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 51.

21 Rozsudok Súdneho dvora EÚ z 23. marca 2010, v spojených veciach C-236/08 až C-238/08, Google France SARL a Google Inc. proti Louis Vuitton Malletier SA.

22 Rozsudok Súdneho dvora EÚ z 12 júla 2011, vo veci C-324/09 L'Oréal SA and ostatní proti eBay International AG and ostatní,§ 124.

23 Rozsudok Súdneho dvora EÚ z 24. novembra 2011, vo veci C-70/10, Scarlet Extended proti SABAM; Rozsudok Súdneho dvora EÚ z 16. februára 2012 vo veci C-360/10, SABAM proti Netlog NV.

24 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021.

nových síl platforiem v informačnej spoločnosti, a tým pripravil pôdu pre ďalšiu fázu – európsky digitálny konštitucionalizmus.

15.2.3 Tretia fáza: Digitálny konštitucionalizmus

Rýchly technologický vývoj v kombinácii s liberálnym ústavným prístupom prispel k tomu, že online platformy získali dominantné postavenie v digitálnom prostredí. Ich rozsiahle využívanie algoritmických nástrojov na moderovanie obsahu a spracovanie údajov výrazne znížilo transparentnosť ich rozhodovania a začali vyvstávať otázky ohľadom ich zodpovednosti.[25] V neposlednom rade, vzostup online platforiem otvoril zásadné otázky týkajúce sa rešpektovania základných práv[26], predovšetkým slobody prejavu a súkromia, ale taktiež negatívne dopady na samotné fungovanie demokratických procesov v digitálnom prostredí.[27] Európska únia preto začala systematicky prijímať právne mechanizmy na obmedzenie tejto súkromnej moci.

Jeho prvým výrazným legislatívnym prejavom je Všeobecné nariadenie o ochrane údajov[28] (ďalej ako „GDPR"), ktoré presadilo chápanie ochrany údajov ako základného práva a zároveň ako nástroja stanovenia limitov súkromných subjektov, ktoré spracúvajú dáta v masovom rozsahu.[29] Charakteristickým znakom tohto prístupu je dôraz na zodpovednosť a reguláciu založenú na miere rizika. GDPR nepracuje len s následnou kontrolou porušení, ale zavádza proaktívne povinnosti pre prevádzkovateľov, ako sú zásady „privacy by design"[30] či povinnosť vykonávať posúdenia vplyvu na ochranu údajov v prípadoch zvýšeného rizika pre práva jednotlivcov.[31] Zároveň posilňuje procesné postavenie dotknutých osôb, najmä prostredníctvom ochrany pred výlučne automatizovaným rozhodovaním, ktoré má právne účinky.[32] Všetky

---

25 BURRELL, J.: *How the machine 'thinks': Understanding opacity in machine learning algorithms.* In: Big Data & Society, roč. 3, č. 1, 2016.

26 BALKIN, J.: *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*. In: U.C. Davis Law Review, roč. 51, 2018.

27 NEMITZ, P.: *Constitutional democracy and technology in the age of artificial intelligence*. In: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, Vol. 376, Issue 2133, 2018.

28 Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/679 z 27. apríla 2016 o ochrane fyzických osôb pri spracúvaní osobných údajov a o voľnom pohybe takýchto údajov, ktorým sa zrušuje smernica 95/46/ES (všeobecné nariadenie o ochrane údajov). OJ L 119/1, 4.5.2016. Ďalej len „GDPR".

29 DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 64.

30 GDPR, čl. 25.

31 GDPR, čl.35.

32 GDPR, čl. 21.

tieto aspekty možné chápať ako zhmotnenie ústavných princípov transparentnosti, zodpovednosti a spravodlivého procesu v digitálnom prostredí.

Na tento vývoj nadväzuje DSA, ktorý je označovaný za najvýraznejší prejav európskeho digitálneho konštitucionalizmu[33], ba dokonca za „*digitálnu ústavu Európy*".[34] DSA posilňuje procesné záruky, zvyšuje transparentnosť moderovania obsahu a zavádza mechanizmy posudzovania a zmierňovania systémových rizík vyplývajúcich z fungovania a dizajnu platforiem. Tým sa digitálny konštitucionalizmus v EÚ posúva od ochrany individuálnych práv k širšiemu rámcu správy digitálneho priestoru, ktorý sa snaží reagovať aj na štrukturálne a systémové dôsledky platformovej moci.

Cieľom DSA je reagovať na moc online platforiem, najmä v oblasti moderovania obsahu, a aktualizovať regulačný model založený na smernici o elektronickom obchode. Hoci DSA zachováva výnimky zo zodpovednosti a zákaz všeobecnej povinnosti monitorovania, výrazne posilňuje transparentnosť, zodpovednosť a procesné záruky pri moderovaní obsahu. V tomto smere ukladá platformám povinnosti odôvodnenia rozhodnutí[35], zohľadňovania základných práv pri moderovaní obsahu[36] a zabezpečenia vnútorných aj vonkajších mechanizmov nápravy. Osobitný režim sa vzťahuje na veľmi veľké online platformy, ktoré musia identifikovať a zmierňovať systémové riziká vyplývajúce z dizajnu ich služieb, vrátane algoritmických systémov. Možno konštatovať, že ide v tomto smere ide aj o kodifikáciu rozhodnutí Súdneho dvora, ktoré boli spomínané v predchádzajúcej časti. Zavedenie vysokých sankcií za porušenie povinností podčiarkuje ambíciu EÚ vytvoriť horizontálny a vynútiteľný rámec, ktorý posilňuje ochranu slobody prejavu a ďalších základných práv a zároveň reflektuje ústavné hodnoty EÚ. Tým sa digitálny konštitucionalizmus v EÚ posúva od ochrany individuálnych práv k širšiemu rámcu správy digitálneho priestoru, ktorý sa snaží reagovať aj na štrukturálne a systémové dôsledky moci platforiem.

---

33 DE GREGORIO, G.: *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge: Cambridge University Press, 2022, s. 211.

34 POLLICINO, O.: *The Quadrangular Shape of the Geometry of Digital Power(s) and the Move Towards a Procedural Digital Constitutionalism*. In: European Law Journal, roč. 29, č. 1–2, 2023, s. 29; BRADFORD, A.: *Europe's Digital Constitution*. In: Virginia Journal of International Law, roč. 64, 2023, s. 1–62.

35 Nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 z 19. októbra 2022 o jednotnom trhu s digitálnymi službami a o zmene smernice 2000/31/ES (akt o digitálnych službách), OJ L 277, 27.10.2022, čl. 17.

36 *Ibid*, čl. 14 ods. 4

15.3 Meta Oversight Board: digitálny konštitucionalizmus v praxi

Popri verejnoprávnej regulácii predstavujú reakciu na rastúcu moc platforiem aj samoregulačné mechanizmy, ktorými sa snažia preniesť ústavné princípy do vlastnej správy online priestoru. Príkladom takéhoto prístupu je MOB. Predstavuje *kvázi* samoregulačný a mimosúdny orgán zriadený spoločnosťou Meta, ktorého cieľom je dohliadať na rozhodnutia o moderovaní obsahu na platformách Facebook a Instagram. Občas býva zjednodušene označovaný ako „*Najvyšší súd Facebooku a Instagramu*“[37], keďže zhmotňuje pokus o zavedenie prvkov ústavného rozhodovania do súkromnej správy online priestoru.

Možno konštatovať, že digitálny konštitucionalizmus sa v modeli MOB prejavuje predovšetkým v inštitucionálnej a normatívnej rovine.

Z inštitucionálneho hľadiska má proces rozhodovania MOB štruktúrovaný a kvázi-súdny charakter.[38] Každý prípad môže byť nahlásený samotným používateľom platformy alebo priamo vybratý spoločnosťou Meta. Používatelia musia spĺňať určité procesné kritériá – napríklad musia byť aktívnymi používateľmi a v lehote 15 dní od rozhodnutia Mety sa musia odvolať.[39] MOB má však plnú diskrečnú právomoc vybrať si prípady, ktorým sa bude venovať. Zameriava sa na emblematické prípady s potenciálom ovplyvniť budúce politiky a prax – najmä tie, ktoré sa dotýkajú slobody prejavu, ľudských práv či otázok verejného diskurzu.[40]

Vybrané prípady preberajú päťčlenné panely, ktorých zloženie by malo reflektovať regionálnu a rodovú rozmanitosť.[41] MOB zhromažďuje podklady od dotknutých strán[42], verejnosti a odborníkov[43], pričom konečné rozhodnutie prijíma väčšinovým hlasovaním. Rozhodnutia MOB majú záväzný účinok, Meta

---

37 DOUEK, E.: *Facebook's "Oversight Board": Move Fast with Stable Infrastructure and Humility*. In: North Carolina Journal of Law & Technology, roč. 21, č. 1, 2019; GRADONI, L.: *Constitutional Review via Facebook's Oversight Board: How Platform Governance Had Its Marbury v Madison*, Verfassungsblog, 10. február 2021. Online: https://verfassungsblog.de/fob-marbury-v-madison/.

38 JAFFE, L. L.: *The right to judicial review I*. In: Harvard Law Review, roč. 71, 1957, s. 401

39 Oversight Board Bylaws, Article 3 sec. 1.1.

40 MAZÚR, J., GRAMBLIČKOVÁ, B.: *New Regulatory Force of Cyberspace: The Case of Meta's Oversight Board*. In: Masaryk University Journal of Law and Technology, roč. 17, č. 1, 2023, s. 8.

41 OVERSIGHT BOARD: Oversight Board Bylaws, 2025, Art. 1, sec. 3.1.3. Online: https://www.oversightboard.com/wp-content/uploads/2025/07/Oversight-Board-Bylaws-June-2025.pdf.

42 *Ibid*, Article 3 Sec. 3.

43 *Ibid*, Article 1, Sec. 3.1.4.

je povinná ich implementovať a zabezpečiť, aby podobné situácie boli posudzované v súlade s vytvoreným precedensom.[44] Teda možno konštatovať, že rozhodovací proces MOB napĺňa určité verejne dostupné procesné štandardy, ktoré formálne evokujú princípy riadneho procesu, transparentnosti a právnej istoty známe zo súdneho rozhodovania.

Z normatívneho hľadiska MOB integruje základné hodnoty a ciele digitálneho konštitucionalizmu priamo do svojho rozhodovania. Pri posudzovaní jednotlivých prípadov sa opiera nielen o interné zásady spoločnosti Meta, ale aj o medzinárodné ľudskoprávne štandardy, najmä článok 19 Medzinárodného paktu o občianskych a politických právach (ICCPR). V prípadoch konfliktu medzi zásadami spoločnosti a ľudskoprávnymi normami sa MOB zaväzuje uprednostniť ochranu ľudských práv.[45]

Napokon, možno pozorovať snahu o formovanie **určitých** nadnárodných štandardov moderovania obsahu. Jej rozhodnutia majú ambíciu presiahnuť národné ústavné rámce a formovať vlastný prístup k otázkam slobody prejavu a politického diskurzu. Tým sa Meta čiastočne dištancuje od regionálnych ústavných tradícií a vytvára svoj vlastný model digitálnej ústavnosti.[46] Niektorí autori dokonca MOB pripisujú potenciál stať sa efektívnym novým typom medzinárodného orgánu ochrany ľudských práv, pričom MOB prirovnávajú k medzinárodným ľudskoprávnym tribunálom.[47]

Pri bližšom pohľade na najznámejšie rozhodnutie MOB, ktoré bolo načrtnuté už v úvode príspevku – rozhodnutie o pozastavení účtu Donalda Trumpa, možno sledovať, ako sa vyššie spomenuté princípy digitálneho konštitucionalizmu premietli do tohto rozhodnutia, najmä v podobe obmedzovania súkromnej moci, posilňovania transparentnosti, aplikácie ľudskoprávnych štandardov, ako aj medzinárodného rozmeru rozhodnutia.

MOB potvrdila rozhodnutie spoločnosti Meta pozastaviť Trumpove účty na Facebooku a Instagrame po útoku na Kapitol 6. januára 2021, keďže jeho príspevky mohli predstavovať riziko podnecovania násilia.[48] Zároveň však MOB

---

44 *Ibid*, Article 2, sec. 2.3.

45 META OVERSIGHT BOARD**:** *Our Work*. Online: https://www.oversightboard.com/our-work/.

46 DA CONCEIÇÃO, L. H. M.: *A constitutional reflector? Assessing societal and digital constitutionalism in Meta's Oversight Board*. In: Global Constitutionalism, roč. 13, č. 3, 2024, s. 570.

47 HELFER, L. R., LAND, M. K.: *The Meta Oversight Board's Human Rights Future*. In: Cardozo Law Review, roč. 44, 2022, s. 2233.

48 META OVERSIGHT BOARD: *Decision on the indefinite suspension of Donald J. Trump's accounts* (Case No. FB-691QAMHJ). Oversight Board, 5. máj 2021. Online: https://www.oversightboard.com/decision/fb-691qamhj/.

konštatovala, že neurčitá a časovo neobmedzená povaha pozastavenia účtu je v rozpore s pravidlami spoločnosti aj s požiadavkami predvídateľnosti, ktoré sú vlastné princípom právneho štátu. „Neurčitý zákaz" nebol v komunitných štandardoch nikde upravený, a preto MOB nariadila, aby Meta určila jasnú a primeranú sankciu v súlade s vlastnými pravidlami a s medzinárodnými normami ochrany ľudských práv.[49] V tomto smere možno ilustrovať obmedzenie súkromnej moci a taktiež apel na transparentnosť, predvídateľnosť a spravodlivý proces.

Ďalej sa v rozhodnutí prejavil ľudskoprávny rámec digitálneho konštitucionalizmu. MOB hodnotila Trumpove príspevky nielen podľa interných pravidiel spoločnosti Meta, ale aj v kontexte medzinárodných ľudských práv, najmä článku 19 ICCPR – práva na slobodu prejavu. Pri hodnotení zásahu do slobody prejavu bol zároveň aplikovaný aj tradičný test proporcionality, bežne používaný pri vyvažovaní ľudských práv, v rámci ktorého skúmala legitímnosť sledovaného cieľa, nevyhnutnosť zásahu a jeho primeranosť vo vzťahu k ochrane verejného záujmu.[50] Týmto spôsobom došlo k horizontálnemu pôsobeniu ľudských práv voči súkromným subjektom, čo tvorí jednu z ústredných myšlienok digitálneho konštitucionalizmu.

V neposlednom rade, rozhodnutie zároveň ilustruje nadnárodný charakter digitálneho konštitucionalizmu, keďže MOB výslovne odmietla aplikovať americký absolútny výklad slobody prejavu.[51] Namiesto toho sa prihlásila k univerzálnym ľudskoprávnym štandardom a svojim rozhodnutím pripustila zásah do slobody prejavu, ak takýto zásah prejde vyššie spomínaným testom proporcionality.

15.4 Kritické pohľady

Napriek tomu, že MOB býva označovaná za „*bezprecedentný experiment v samospráve platforiem*"[52] a pokus o implementovanie princípov digitálneho konštitucionalizmu do praxe, jej fungovanie sprevádza množstvo zásadných kritík, ktoré spochybňujú jej efektivitu, nezávislosť a schopnosť reálne obmedziť moc spoločnosti Meta a chrániť práva používateľov.

---

49 *Ibid*.

50 *Ibid*.

51 DA CONCEIÇÃO, L. H. M.: *A constitutional reflector? Assessing societal and digital constitutionalism in Meta's Oversight Board*. In: Global Constitutionalism, roč. 13, č. 3, 2024, s. 570.

52 KLONICK, K.: *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*. In: Yale Law Journal, roč. 129, č. 8, 202, s. 2449.

Jednou z hlavných výhrad sú funkčné deficity a neefektívnosť MOB ako mechanizmu nápravy pre používateľov. Dáta poukazujú na výraznú nedostupnosť tohto mechanizmu pre bežných používateľov platforiem. Zatiaľ čo približne 15 % prípadov predložených samotnou spoločnosťou Meta je prijatých na preskúmanie, v prípade podaní zo strany používateľov ide len o 0,002 % prípadov.[53] Na túto nerovnováhu nadväzuje aj rozdiel v samotných výsledkoch rozhodovania. Prípady iniciované Metou boli potvrdené v približne 53 % rozhodnutí, zatiaľ čo v prípade používateľských podaní bola úspešnosť len okolo 12,5 %.[54] Takáto disproporcia spochybňuje charakter MOB ako efektívneho a dostupného opravného prostriedku a naznačuje, že mechanizmus je štrukturálne nastavený v prospech záujmov platformy.

Ďalšia línia kritiky smeruje k demokratickým deficitom a používaniu konštitucionálnych analógií, teda že je MOB často prirovnávaná k najvyššiemu súdu[55] alebo ľudskoprávnym tribunálom.[56] Avšak, v skutočnosti MOB nespĺňa potrebné ústavné kritériá, aby mohla figurovať ako skutočne nezávislý kvázi-súdny orgán. V tomto smere tu absentujú klasické ústavné mechanizmy bŕzd a protiváh, keďže Meta si zachováva rozhodujúce slovo pri akejkoľvek zmene Charty MOB aj pri implementácii jej odporúčaní.[57] Rovnako tak je MOB nepriamo finančne závislá od samotnej spoločnosti.[58] V neposlednom rade MOB trpí zásadným deficitom demokratickej legitimity – jej členovia nie sú volení používateľmi ani verejnosťou a možné mechanizmy participácie zostávajú bez rozhodujúceho vplyvu. Tým pádom *„prinajlepšom, slúži architektúra MOB na to, aby výsledok získal väčšiu vnútornú legitimitu, čím sa kvalifikuje ako samoregulačné rozhodnutie, ktoré prešlo špeciálnym postupom. Tieto črty*

---

53 PONGÓ, T., KÁLMÁN, R., SZANISZLÓ, R.B. AND TRIBL, N.: *Oversight Board–An Effective Tool to Protect Free Speech?* In: Masaryk University Journal of Law and Technology, Vol. 19, No. 2, 2025, s. 171.

54 *Ibid*, s. 165.

55 DOUEK, E.: *Facebook's "Oversight Board": Move Fast with Stable Infrastructure and Humility*. In: North Carolina Journal of Law & Technology, roč. 21, č. 1, 2019; GRADONI, L.: *Constitutional Review via Facebook's Oversight Board: How Platform Governance Had Its Marbury v Madison*. *Verfassungsblog*, 10. február 2021. Online: https://verfassungsblog.de/fob-marbury-v-madison/.

56 HELFER, L. R., LAND, M. K.: *The Meta Oversight Board's Human Rights Future*. In: Cardozo Law Review, roč. 44, 2022, s. 2233.

57 KLONICK, K.: *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*. In: Yale Law Journal, roč. 129, č. 8, 202, s. 2457–2466.

58 *Ibid*, s. 2469.

nezodpovedajú demokratickej ústavnej architektúre, ktorá by oprávňovala operovať s pojmami „súd" a „ústava".[59]

Napokon je potrebné zdôrazniť, že MOB disponuje len veľmi obmedzenou schopnosťou reálne zasiahnuť do výkonu moci spoločnosti Meta na systémovej úrovni. Hoci jej rozhodnutia môžu mať precedenčný charakter pri posudzovaní jednotlivých zásahov do obsahu, MOB nemá právomoc ani mandát zasahovať do tvorby pravidiel platformy. Jej pôsobenie sa tak obmedzuje na dodatočné preskúmavanie individuálnych prípadov, čo sa v kontexte komplexných digitálnych ekosystémov javí ako nedostatočný nástroj kontroly.[60] Samotné interné analýzy spoločnosti Meta pritom poukazujú na to, že riešenie jednotlivých príspevkov či komentárov má len marginálny vplyv na šírenie dezinformácií a iných škodlivých javov, ktoré často vznikajú v dôsledku fungovania algoritmov a optimalizácie obsahu na angažovanosť používateľov, nie v dôsledku porušenia obsahových pravidiel.[61]

Hoci úloha MOB v oblasti moderovania obsahu môže pôsobiť ako prejav zodpovedného správania, v skutočnosti nenahrádza potrebu nezávislej kontroly najzásadnejších prvkov moci platforiem, teda algoritmických systémov a celkovej architektúry digitálneho priestoru. Absencia právomoci kontrolovať algoritmy a architektúru platformy tak predstavuje zásadný limit jej efektívnosti. V dôsledku toho môže MOB skôr prispievať k legitimizácii existujúcej koncentrácie súkromnej moci než k jej reálnemu obmedzeniu.[62]

Kritika MOB však nepredstavuje izolovaný problém viazaný výlučne na inštitucionálny dizajn jednej platformy, ale otvára širšiu otázku samotných limitov digitálneho konštitucionalizmu ako normatívneho a teoretického rámca.

Jedna z najzásadnejších výhrad voči digitálnemu konštitucionalizmu spočíva v tom, že samotné prepojenie digitálnych platforiem na princípy

---

59 KELLER, C. I., PEREIRA, J. R. G.: *Digital constitutionalism as an online speech governance framework: A critical approach*. In: SPIECKER gen. DÖHMANN, I., SCHERTEL MENDES, L., CAMPOS, R. R. (eds.): Digital Constitutionalism. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2025, s. 305.

60 MAZÚR, J., GRAMBLIČKOVÁ, B.: *New Regulatory Force of Cyberspace: The Case of Meta's Oversight Board*. In: Masaryk University Journal of Law and Technology, roč. 17, č. 1, 2023, s. 12.

61 O'SULLIVAN, D., SUBRAMANIAM, T., DUFFY, C.: *Not Stopping "Stop the Steal:" Facebook Papers Paint Damning Picture of Company's Role in Insurrection*, CNN, 22. október 2021. Online: https://edition.cnn.com/2021/10/22/business/january-6-insurrection-facebook-papers/index.html.

62 DOUEK, E.: *Facebook's "Oversight Board": Move Fast with Stable Infrastructure and Humility*. In: North Carolina Journal of Law & Technology, roč. 21, č. 1, 2019. s. 74; KLONICK, K.: *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*. In: Yale Law Journal, roč. 129, č. 8, 2020, s. 2475.

právneho štátu ešte neznamená skutočné obmedzenie ich moci. Zavedenie pravidiel transparentnosti, proporcionality či procesných záruk totiž implicitne potvrdzuje fakt, že práve tieto súkromné subjekty sú oprávnené rozhodovať o hraniciach online prejavu. Pokiaľ sa regulačné zásahy obmedzujú len na normatívne rámcovanie výkonu tejto moci, bez toho, aby spochybňovali samotnú architektúru a inštitucionálne mechanizmy, ktoré ju umožňujú, dochádza skôr k jej formalizácii než k jej narušeniu.[63]

V tomto zmysle digitálny konštitucionalizmus často pôsobí paradoxne: namiesto oslabovania asymetrických mocenských vzťahov im poskytuje nový jazyk legitimity. Ústavné metafory, procesné štandardy a ľudskoprávna argumentácia môžu vytvárať dojem, že výkon moci platforiem je dostatočne kontrolovaný a legitimizovaný, hoci jeho materiálne základy – algoritmická architektúra, ekonomika pozornosti či centralizované rozhodovanie zostávajú nedotknuté. Kritici preto upozorňujú, že digitálny konštitucionalizmus v mnohých podobách nefunguje ako nástroj redistribúcie moci, ale skôr ako teoretický rámec, ktorý prekrýva nerovnováhu moci „*plášťom legitimity*“.[64]

Ďalší problém spočíva v idealizovanom chápaní subjektov, na ktoré sa digitálny konštitucionalizmus odvoláva. Koncept často pracuje s predstavou používateľa ako obdobou ústavodarného subjektu – ľudu , bez toho, aby dostatočne reflektoval materiálne podmienky digitálnych platforiem, ktoré fungujú ako globálne, komerčné infraštruktúry. Analógie so spoločenskou zmluvou či verejnou sférou sú preto kritizované ako zjednodušené, keďže digitálny priestor je v skutočnosti organizovaný prostredníctvom súkromných zmluvných vzťahov s cieľom čo najvyššieho zisku, nie na demokratickú participáciu.[65] Tieto výhrady spoločne naznačujú, že digitálny konštitucionalizmus stojí na napätí medzi normatívnou ambíciou ochrany základných práv a materiálnou realitou súkromnej moci platforiem a platformovej ekonomiky.

---

63 KELLER, C. I., PEREIRA, J. R. G.: *Digital constitutionalism as an online speech governance framework: A critical approach*. In: SPIECKER gen. DÖHMANN, I., SCHERTEL MENDES, L., CAMPOS, R. R. (eds.): Digital Constitutionalism. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2025, s. 309.

64 HELBERGER, N.: *The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power*. In: Digital Journalism, roč. 8, č. 6, 2020, s. 842–854.

65 TERZIS, P.: *Against digital constitutionalism*. In: European Law Open, roč. 3, č. 2, 2024, s. 341.

Záver

V závere možno konštatovať, že cieľom predloženého príspevku bolo analyzovať presun rozhodovacej moci o základných právach z verejných inštitúcií na súkromné digitálne platformy a preskúmať, ako sa v tomto kontexte formuje koncept digitálneho konštitucionalizmu. Možno zhrnúť, Digitálny konštitucionalizmus predstavuje reakciu na zásadnú premenu výkonu moci v digitálnom priestore, kde súkromné platformy nadobudli schopnosť systematicky ovplyvňovať výkon základných práv.

Analýza MOB ukázala, že hoci samoregulačné mechanizmy môžu do správy digitálnych platforiem vnášať jazyk ústavných princípov, ako sú proporcionalita, transparentnosť a riadny proces, ich schopnosť reálne obmedziť koncentráciu súkromnej moci zostáva obmedzená. Možno konštatovať, že samoregulačné modely nezasahujú do samotných štrukturálnych a technických základov platformovej moci, prostredníctvom ktorých sa výkon tejto moci uskutočňuje, a preto nedokážu adresovať jej systémové dôsledky.

Z uvedených zistení vyplýva, že ochranu základných práv v digitálnom priestore nemožno založiť výlučne na samoregulačných mechanizmoch platforiem. Ťažisko normatívnej odpovede sa v európskom kontexte presúva k záväzným verejnoprávnym rámcom, predovšetkým k DSA, ktorý predstavuje pokus o pretavenie princípov digitálneho konštitucionalizmu do právne vynútiteľných povinností. Osobitný význam má v tomto smere článok 14 ods. 4 DSA, ktorý požaduje, aby zmluvné podmienky platforiem boli formulované a uplatňované v súlade so základnými právami používateľov.

Zároveň však bolo poukázané na to, že samotné procesné záruky a individuálne opravné prostriedky nepostačujú na účinné riešenie výziev algoritmickej spoločnosti. Regulačné úsilie je preto potrebné rozšíriť aj na systémové riziká vyplývajúce z dizajnu platforiem, ich obchodných modelov a algoritmických systémov, ktoré zásadným spôsobom formujú verejný diskurz ešte pred samotným rozhodovaním o obsahu. Len kombinácia právne záväznej regulácie, zvýšenej transparentnosti algoritmického fungovania a účinnej verejnej kontroly môže zabezpečiť, aby digitálny konštitucionalizmus neostal len metaforou, ale stal sa funkčným nástrojom ochrany základných práv v digitálnom priestore.

Zoznam použitej literatúry a zdrojov

1. BALKIN, J.: *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*. In: U.C. Davis Law Review, roč. 51, 2018, s. 1151–1210.
2. BELLI, L., VENTURINI, J.: *Private ordering and the rise of terms of service as cyber-regulation*. In: Internet Policy Review, roč. 5, č. 4, 2016, s. 1–17.
3. BRADFORD, A.: *Europe's Digital Constitution*. In: Virginia Journal of International Law, roč. 64, 2023, s. 1–62.
4. BURRELL, J.: *How the machine 'thinks': Understanding opacity in machine learning algorithms.* In: Big Data & Society, roč. 3, č. 1, 2016, s. 1–12.
5. CELESTE, E.: *Digital constitutionalism: a new systematic theorisation*. In: International Review of Law, Computers & Technology, roč. 33, č. 1, 2019, s. 76–99.
6. CELESTE, E.: *Digital constitutionalism: The role of internet bills of rights*. London – New York: Taylor & Francis, 2022, 254 s.
7. CELESTE, E.: *Terms of Service and Bills of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment?* In*:* International Review of Law, Computers & Technology, roč. 33, č. 2, 2018, s. 122–138.
8. DA CONCEIÇÃO, L. H. M.: *A constitutional reflector? Assessing societal and digital constitutionalism in Meta's Oversight Board*. In: Global Constitutionalism*,* roč. 13, č. 3, 2024, s. 557–590.
9. DE GREGORIO, G.: *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge: Cambridge University Press, 2022, 366 s.
10. DE GREGORIO, G.: *The rise of digital constitutionalism in the European Union*. In: International Journal of Constitutional Law, roč. 19, č. 1, 2021, s. 41–70.
11. DENARDIS, L., HACKL, A. M.: *Internet governance by social media platforms*. In: Telecommunications Policy, roč. 39, č. 9, 2015, s. 761–770.
12. DOUEK, E.: *Facebook's "Oversight Board": Move Fast with Stable Infrastructure and Humility*. In: North Carolina Journal of Law & Technology, roč. 21, č. 1, 2019, 1–78.
13. ESTEVE, A.: *The Business of Personal Data: Google, Facebook, and Privacy Issues in the EU and the USA.* In: International Data Privacy Law, roč. 7, 2017, s. 36–47.
14. GRADONI, L.: *Constitutional Review via Facebook's Oversight Board: How Platform Governance Had Its Marbury v Madison*, Verfassungsblog*,* 10. február 2021. Online: https://verfassungsblog.de/fob-marbury-v-madison/.

15. HARDY, I. T.: *The proper legal regime for cyberspace*. In: University of Pittsburgh Law Review, roč. 55, 1993, s. 993–1038.
16. HELBERGER, N.: *The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power*. In: Digital Journalism, roč. 8, č. 6, 2020, s. 842–854.
17. HELFER, L. R., LAND, M. K.: *The Meta Oversight Board's Human Rights Future*. In: Cardozo Law Review, roč. 44, 2022, s. 2233–2280.
18. HERN, A.; PAUL, K.: Donald Trump Twitter ban comes to end amid calls for tougher action, The Guardian, 7. január 2021. Online: https://www.theguardian.com/us-news/2021/jan/07/donald-trump-twitter-ban-comes-to-end-amid-calls-for-tougher-action.
19. Charta základných práv Európskej únie, Európska únia, 2000/C 364/01, zo dňa 7. decembra 2000.
20. JAFFE, L. L.: *The right to judicial review* I. In: Harvard Law Review, roč. 71, 1957, s. 401–422.
21. KELLER, C. I., PEREIRA, J. R. G.: *Digital constitutionalism as an online speech governance framework: A critical approach*. In: SPIECKER gen. DÖHMANN, I., SCHERTEL MENDES, L., CAMPOS, R. R. (eds.): *Digital Constitutionalism*. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2025, s. 291–310.
22. KLONICK, K.: *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*. In: Yale Law Journal, roč. 129, č. 8, 2020, s. 2449–2499.
23. LAIDLAW, E. B.: *Private Power, Public Interest: An Examination of Search Engine Accountability*. In: International Journal of Law and Information Technology, roč. 17, č. 1, 2009, s. 113–145.
24. LAIDLAW, E. B.: *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility*. Cambridge: Cambridge University Press, 2015.
25. MAZUR, J., GRAMBLIČKOVÁ, B.: *New Regulatory Force of Cyberspace: The Case of Meta's Oversight Board*. In: Masaryk University Journal of Law and Technology, roč. 17, č. 1, 2023, s. 3–32.
26. META OVERSIGHT BOARD: *Decision on the indefinite suspension of Donald J. Trump's accounts (Case No. FB-691QAMHJ).* Oversight Board, 5. máj 2021. Online: https://www.oversightboard.com/decision/fb-691qamhj/.
27. META OVERSIGHT BOARD**:** *Our Work*. Online: https://www.oversightboard.com/our-work/.
28. META OVERSIGHT BOARD**:** *Oversight Board Bylaws*, marec 2024. Online: https://www.oversightboard.com/wp-content/uploads/2024/03/Oversight-Board-Bylaws.pdf.

29. Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/679 z 27. apríla 2016 o ochrane fyzických osôb pri spracúvaní osobných údajov a o voľnom pohybe takýchto údajov, ktorým sa zrušuje smernica 95/46/ES (všeobecné nariadenie o ochrane údajov). OJ L 119/1, 4.5.2016.
30. Nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 z 19. októbra 2022 o jednotnom trhu s digitálnymi službami a o zmene smernice 2000/31/ES (akt o digitálnych službách). OJ L 277/1, 27.10.2022.
31. NEMITZ, P.: *Constitutional democracy and technology in the age of artificial intelligence*. In: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, Vol. 376, Issue 2133, 2018, s. 1–14.
32. O'SULLIVAN, D., SUBRAMANIAM, T., DUFFY, C.: *Not Stopping "Stop the Steal:" Facebook Papers Paint Damning Picture of Company's Role in Insurrection*. CNN, 22. október 2021. Online: https://edition.cnn.com/2021/10/22/business/january-6-insurrection-facebook-papers/index.html.
33. PADOVANI, C., SANTANIELLO, M.: *Digital constitutionalism: Fundamental rights and power limitation in the Internet eco-system*. In: International Communication Gazette, roč. 80, č. 4, 2018, s. 295–301.
34. PETERS, A.: *Compensatory constitutionalism: the function and potential of fundamental international norms and structures*. In: Leiden Journal of International Law, roč. 19, č. 3, 2006, s. 579–610.
35. POLLICINO, O.: *The Quadrangular Shape of the Geometry of Digital Power(s) and the Move Towards a Procedural Digital Constitutionalism*. In: European Law Journal, roč. 29, č. 1–2, 2023, s. 29–47.
36. PONGÓ, T., KÁLMÁN, R., SZANISZLÓ, R. B., TRIBL, N.: *Oversight Board – An Effective Tool to Protect Free Speech?* In: Masaryk University Journal of Law and Technology, roč. 19, č. 2, 2025, s. 151–178.
37. Rozsudok Súdneho dvora EÚ z 12 júla 2011, vo veci C-324/09 L'Oréal SA and ostatní proti eBay International AG and ostatní. ECLI:EU:C:2011:474.
38. Rozsudok Súdneho dvora EÚ z 16. februára 2012 vo veci C-360/10 Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) proti Netlog NV. ECLI:EU:C:2012:85.
39. Rozsudok Súdneho dvora EÚ z 23. marca 2010, v spojených veciach C-236/08 až C-238/08, Google France SARL a Google Inc. proti Louis Vuitton Malletier SA, Google France SARL proti Viaticum SA a Luteciel SARL, Google France SARL proti Centre national de recherche en relations humaines (CNRRH) SARL a iní. ECLI:EU:C:2010:159.

40. Rozsudok Súdneho dvora z 24. novembra 2011, vo veci C-70/10, Scarlet Extended SA proti Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM). ECLI:EU:C:2011:771.
41. Smernica 2000/31/ES Európskeho parlamentu a Rady z 8. júna 2000 o určitých právnych aspektoch služieb informačnej spoločnosti na vnútornom trhu, najmä o elektronickom obchode (smernica o elektronickom obchode). OJ L 178/1, 17.7.2000.
42. SRNICEK, N.: *Platform Capitalism*. Cambridge: Polity Press, 2017, 120 s.
43. TERZIS, P.: *Against digital constitutionalism*. In: European Law Open, roč. 3, č. 2, 2024, s. 336–352.

Informácie o autorovi

Mgr. Soňa JUHÁSOVÁ
Univerzita Komenského v Bratislave
Právnická fakulta
Ústav práva informačných technológií a práva duševného vlastníctva
Email: sona.juhasova@flaw.uniba.sk

16 Čítanie mysle, prepisovanie pravidiel? Ľudské práva v ére neurotechnológií

Laura Gazdagová

https://doi.org/10.33542/HRF-0541-2-16

Abstrakt

"Freedom of thought includes the right to develop ideas autonomously and free from impermissible external influence, making protection against neurotechnological interventions essential. However, neither the Human Rights Committee (HRC) nor the European Court of Human Rights (ECtHR) has clearly defined what freedom of thought entails. Despite the rapid development of neurotechnologies, this right remains often overlooked; even the Council of Europe's Guide on Artificial Intelligence does not mention it. Some authors, such as Alegre, propose a broad interpretation of "thought," encompassing emotional states, political opinions, and even ordinary mental processes. A precise definition of the concept of "thought" is necessary to assess whether the protection of freedom of thought is sufficient, or whether a new right to mental self-determination is needed to safeguard the entire mind, including emotions and non-rational processes. Neurotechnologies devices that interact with the human nervous system have revolutionized medicine and neuroscience, enabling the treatment of Parkinson's disease and drug-resistant depression, improving cognitive and motor abilities, and even altering perception and memory. These technologies include brain-computer interfaces (BCIs), deep brain stimulation (DBS), and transcranial direct current stimulation (tDCS). Advances in neurotechnology and artificial intelligence provide unprecedented insights into the human brain, with increasing possibilities to influence and measure brain activity. These developments raise complex ethical and legal questions. Proponents of neurorights argue for the introduction of new human rights to protect mental processes and brain data. This study focuses on the research question: which existing explicitly recognized rights may be threatened by the development of neurotechnologies, and is there a need to introduce new neurorights?"

Úvod

Pokroky v oblasti neurotechnológií a umelej inteligencie (angl. Artifical Inteligence (ďalej len „AI“)) otvárajú nové horizonty v porozumení a modulácii mozgových procesov. Predmetné moderné technológie dokážu monitorovať mozgovú aktivitu s vysokou presnosťou a umožňujú nielen diagnostiku či zmiernenie príznakov neurologických a psychiatrických ochorení, ale aj asistenciu osobám so zdravotným postihnutím pri obnove mobility a

komunikácie. Takýto technologický vývoj prináša však aj nové etické a predovšetkým právne výzvy. V súčasnosti sa diskutuje, do akej miery dokážu existujúce právne rámce zabezpečiť ochranu jednotlivcov voči potenciálnym rizikám spojeným s monitorovaním a ovplyvňovaním mozgových funkcií. Na medzinárodnej aj regionálnej úrovni sa preto vytvárajú odporúčania a regulačné mechanizmy, ktoré majú za cieľ minimalizovať riziká zneužitia neurotechnológií a zároveň podporiť ich bezpečné a eticky zodpovedné využitie. Praktické skúsenosti z adaptácie existujúcich práv v kontexte nových technológií, poukazujú na skutočnosť, že progresívna interpretácia už existujúcich a explicitne zakotvených ľudských práv môže byť účinnejšia a udržateľnejšia než vytváranie úplne nových ľudských práv. Súčasne je dôležité, aby regulačné a etické rámce boli flexibilné a dokázali reagovať na rýchle technologické zmeny, pričom zachovávajú ochranu základných práv, dôstojnosti a autonómie jednotlivca. Cieľom článku je preto preskúmať a zhodnotiť možnosti a obmedzenia existujúcich ľudsko-právnych mechanizmov pri regulácii neurotechnológií a ich integrácie s AI a zároveň identifikovať potenciálne prístupy, ktoré by mohli podporiť ich bezpečné a eticky prijateľné využitie v budúcnosti.

16.1 Integrácia neurotechnológií a umelej inteligencie v modernej medicíne a technike

Moderné technológie možno považovať za jeden z pilierov súčasnej vedecko-technickej transformácie, ktoré majú výrazne pozitívny dopad na rôzne aspekty ľudskej existencie a ktorých vplyv bude v nasledujúcich rokoch s najväčšou pravdepodobnosťou naďalej narastať. Avšak rýchle a rozsiahle zmeny sú často komplexné, dynamické a nie vždy transparentné či ľahko predvídateľné. Tieto skutočnosti preto vyvolávajú závažné otázky a obavy týkajúce sa širších spoločenských, etických a právnych dôsledkov využívania moderných technológií v súčasnej spoločnosti. Moderné technológie zároveň rýchlo menia naše možnosti porozumieť fungovaniu ľudského mozgu, ako aj zasahovať do jeho funkcií a dysfunkcií, najmä keď sa integrujú s neurotechnológiou. Niektoré neinvazívne neurotechnológie sú známe a využívané už celé desaťročia. Typickým príkladom je elektroencefalografia (ďalej ako „EEG"), ktorá sa široko používa na meranie mozgovej aktivity, alebo transkraniálna magnetická stimulácia (ďalej ako „TMS"), umožňujúca externú stimuláciu určitých oblastí mozgu.[1] TMS je zároveň schválenou terapeutickou

[1] HENSING, J., FAIKA, A.: *Neurotechnology: A New Frontier for Prosperity and Security in Germany and Europe?* In: *GPPi,* 2025. Online:

metódou pri liečbe niektorých duševných porúch, ako napríklad depresie. V posledných rokoch sa však výrazne rozvíjajú aj invazívne neurotechnológie, ktoré si vyžadujú implantáciu zariadení priamo do mozgu (a teda prostredníctvom neurochirurgie) alebo iné fyzické zásahy do nervového systému. Ide o novšiu, rýchlo sa rozvíjajúcu oblasť, úzko spätú s pokrokom v biotechnológii a nanotechnológii. Medzi tieto techniky patrí napríklad elektrokortikografia, ktorá umožňuje presnejší záznam nervovej aktivity, a hlboká mozgová stimulácia (angl. Deep Brain Stimulation ďalej ako „DBS"), ktorá je dnes dostupnou a klinicky overenou metódou. DBS sa používa najmä pri liečbe Parkinsonovej choroby, kde pomáha znižovať tremor a ďalšie symptómy choroby. [2] Mnohé z týchto technológií fungujú na princípe uzavretých spätnoväzbových systémov, v ktorých algoritmy AI priebežne analyzujú „neurálne" signály a v reálnom čase prispôsobujú stimulačné alebo dekódovacie parametre.[3] V takýchto prípadoch moderné technológie preukazujú (predovšetkým AI) svoj potenciál, napríklad pri DBS umožňujú presnejšiu identifikáciu cieľových oblastí mozgu a individualizáciu stimulačných protokolov na základe rozpoznávania komplexných vzorcov ochorení vo veľkoobjemových neurálnych dátach.

Neurotechnológie teda zahŕňajú širokú škálu invazívnych (resp. semiinvazívnych) a neinvazívnych zariadení schopných dekódovať a modulovať nervovú aktivitu. Významný pokrok sa dosiahol najmä v rozhraniach schémy *„mozog-počítač"* (angl. Brain-Computer Interfaces, ďalej ako „BCI"), ktoré spájajú mozgovú aktivitu s externým počítačom alebo strojom, a to buď v otvorenej, alebo uzavretej slučke. Napríklad intrakortikálne elektródové implantáty dokážu dekódovať požadované pohyby paží a umožniť ovládanie robotických ramien alebo kurzorov aj u úplne paralyzovaných pacientov.[4] Nedávne experimenty zamerané na motorickú kontrolu preukázali presnosť dekódovania pohybov presahujúcu 85 percent.[5] Ešte väčší pokrok sa zaznamenal pri dekódovaní reči pričom v roku 2024 vedci pomocou implantovaných senzorov dokázali rekonštruovať pokusy o reč u osôb s

https://gppi.net/2025/03/27/neurotechnology-a-new-frontier-for-prosperity-and-security-in-germany-and-europe.

2 *Ibid.*

3 *Ibid.*

4 Bližšie pozri: WILLETT, F.R. et al*.: A high-performance speech neuroprosthesis*. In: *Nature*, roč. 620, 2023, s. 1031–1036. Online: https://doi.org/10.1038/s41586-023-06377-x.

5 PATRICK-KRUEGER, K.M., BURKHART, I. a CONTRERAS-VIDAL, J.L.: *The state of clinical trials of implantable brain–computer interfaces*. In: Nature Reviews Bioengineering, 2025, s. 50–67. Online: https://doi.org/10.1038/s44222-024-00239-5.

amyotrofickou laterálnou sklerózou s presnosťou až 97,5 percenta.[6] Predmetné štúdie preukazujú, že invazívne BCI dokážu premietať mentálnu aktivitu do konkrétnych akcií, čím umožňujú obnovu komunikácie a mobility u osôb so zdravotným postihnutím.[7] Kľúčovým prvkom umožňujúcim dnešné fungovanie BCI je z veľkej časti samozrejme aj AI. Moderné rozhrania BCI sú postavené na algoritmoch strojového a hlbokého učenia, ktoré zabezpečujú presné spracovanie, interpretáciu a predikciu neurálnych signálov. Bez týchto algoritmických nástrojov by spoľahlivé dekódovanie komplexných mozgových dát nebolo možné. V oblasti BCI plní umelá inteligencia najmä tieto funkcie:

- rozpoznávanie komplexných vzorcov v mozgových signáloch, čo výrazne zvyšuje schopnosť systémov identifikovať používateľove mentálne príkazy či zámery,
- filtrovanie šumu v reálnom čase, ktoré umožňuje stabilné a presné dekódovanie „neurálnej" aktivity aj v dynamických podmienkach,
- predikcia používateľových zámerov a emocionálnych stavov, a to na základe jemných, pre ľudského pozorovateľa takmer nerozoznateľných zmien v neurónovej aktivite,
- adaptácia používateľského rozhrania podľa kontextu a spätnej väzby, čo vedie k vyššej miere personalizácie, efektívnejšiemu ovládaniu a dlhodobej spoľahlivosti systému.[8]

Uvedené schopnosti AI výrazne rozširujú výkonové možnosti BCI, ktoré tak postupne nadobúdajú charakter vysoko adaptívnych, prediktívnych a používateľsky orientovaných neurotechnológií. Integrácia AI tak nepredstavuje iba technologické rozšírenie, ale stáva sa základnou architektonickou súčasťou súčasných neurotechnologických systémov.

16.1.1 Krátky exkurz do regulácie umelej inteligencie v EÚ a jej dôsledky pre neurotechnológie

Z regionálneho hľadiska, najmä z pohľadu právotvorby EÚ, je s ohľadom na problematiku neurotechnológií v korelácii s AI relevantná práva úprava podľa

---

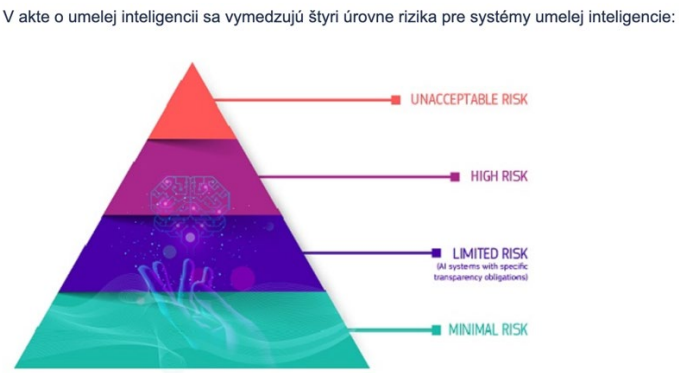
6 CARD, N.S. et al.: *An accurate and rapidly calibrating speech neuroprosthesis*. In: *New England Journal of Medicine*, roč. 391, 2024, s. 609–618.

7 Informácie a primárne zdroje získané: SZOSZKIEWICZ, Ł., YUSTE, R.: *Mental privacy: navigating risks, rights and regulation: Advances in neuroscience challenge contemporary legal frameworks to protect mental privacy.* In: EMBO Reports, č. 14, 2025, s. 3469–3473. Online: https://doi.org/10.1038/s44319-025-00505-6.

8 ZHANG, H. et al.: *Brain-computer interfaces: the innovative key to unlocking neurological conditions.* In: *International Journal of Surgery*, roč. 110, č. 9, 2024, s. 5745–5762. Online: https://doi.org/10.1097/JS9.0000000000002022.

nariadenia Európskeho parlamentu a Rady (EÚ) 2024/1689 (známeho ako Akt o umelej inteligencii) (ďalej len ako „AIA"), ktorá v čl. 3 ods. 1 definuje AI ako *„strojový systém, ktorý je dizajnovaný na prevádzku s rôznymi úrovňami autonómnosti, ktorý môže po nasadení prejavovať adaptabilitu a ktorý pre explicitné alebo implicitné ciele odvodzuje zo vstupov, ktoré dostáva, spôsob generovania výstupov, ako sú predpovede, obsah, odporúčania alebo rozhodnutia, ktoré môžu ovplyvniť fyzické alebo virtuálne prostredie."* Predmetná definícia podľa tvrdení Santalu zahŕňa aj pokročilé algoritmy strojového učenia, bežne využívané aj v neurovede, napríklad konvolučné neurónové siete[9] dekódujúce mozgové signály z EEG či generatívne siete rekonštruujúce vizuálne podnety z mozgovej aktivity. AI systém môže existovať ako samostatný softvér alebo byť súčasťou iného produktu a pritom nemusí byť fyzicky integrovaný do hardvéru (napr. softvér pre spracovanie signálov z EEG-headsetu). *Z toho však nevyplýva skutočnosť, že každý systém AI automaticky podlieha regulácii AIA. Podľa AIA sa regulačný rámec uplatní len vtedy, ak systém patrí do jednej z definovaných rizikových kategórií.*

***Ilustrácia č. 1 rizikových kategórií využívania AI podľa AIA*[10]**



AIA, ako je možné vyvodiť z Ilustrácie č. 1, rozlišuje viacero úrovní rizika systémov AI, pričom každá z nich je spojená so špecifickým regulačným režimom. Na najvyššej úrovni pyramídy sa nachádzajú systémy považované za

---

[9] Pre odbornejšie vysvetlenie pozri: XU, Y., VAZIRI-PASHKAM, M.: *Limits to visual representational correspondence between convolutional neural networks and the human brain*. In: Nature Communications, roč. 12, č. 1, 2021, s. 2065. Online: https://doi.org/10.1038/s41467-021-22244-7.

[10] EURÓPSKA KOMISIA: *AI Act*, posledná aktualizácia 11. mája 2026. Online: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

„neprijateľné riziko“, ktoré predstavujú jednoznačnú hrozbu pre bezpečnosť, základné práva a spoločenské fungovanie. Tieto systémy sú podľa AIA normatívne zakázané. Do tejto kategórie patria najmä postupy využívajúce manipulatívne a klamlivé techniky založené na umelej inteligencii, využívanie zraniteľností jednotlivcov, systémy sociálneho skórovania, modely predikcie individuálneho rizika spáchania trestnej činnosti, plošné zhromažďovanie údajov z internetu alebo kamerových systémov na účely tvorby databáz na rozpoznávanie tvárí, ako aj systémy na rozpoznávanie emócií v pracovnom prostredí alebo vo vzdelávacích inštitúciách. Rovnako je zakázaná biometrická kategorizácia umožňujúca odvodenie chránených charakteristík a používanie vzdialeného, v reálnom čase fungujúceho biometrického identifikačného systému na účely presadzovania práva na verejne prístupných miestach.[11]

Nižšiu úroveň predstavujú systémy označené ako *„vysoko rizikové“.* Ide o prípady, pri ktorých môže nesprávne fungovanie umelej inteligencie spôsobiť vážne ohrozenie zdravia, bezpečnosti alebo základných práv jednotlivca. Sem patrí používanie AI ako bezpečnostných komponentov kritickej infraštruktúry, systémy ovplyvňujúce prístup k vzdelaniu, produkty zdravotníckej povahy využívajúce AI, nástroje v oblasti zamestnávania a riadenia pracovnej sily, systémy podmieňujúce prístup k nevyhnutným službám, technológie určené na diaľkovú biometrickú identifikáciu či kategorizáciu, ako aj riešenia používané v oblasti presadzovania práva, migrácie, azylu, správy hraníc, justičnej činnosti a demokratických procesov. Vysokorizikové systémy podliehajú prísnemu regulačnému režimu, ktorý vyžaduje zavedenie primeraných mechanizmov hodnotenia a zmierňovania rizík, používanie kvalitných databáz minimalizujúcich diskriminačné výsledky, povinnosť viesť podrobné záznamy umožňujúce sledovateľnosť výsledkov, vypracovanie rozsiahlej technickej dokumentácie a zabezpečenie dostatočného ľudského dohľadu. Dôraz sa kladie aj na odolnosť, kybernetickú bezpečnosť a presnosť systémov. Samostatnú kategóriu predstavujú systémy spojené s požiadavkou transparentnosti. Ide o situácie, v ktorých je potrebné, aby používateľ vedel, že interaguje so systémom umelej inteligencie. Napríklad pri využívaní chatbotov musí byť jasne uvedené, že osoba komunikuje so strojom. Poskytovatelia generatívnej AI sú zároveň povinní zabezpečiť identifikovateľnosť výstupov, pričom deepfake obsah a texty určené na informovanie verejnosti v oblastiach verejného záujmu musia byť jasne a viditeľne označené. Najnižšiu kategóriu tvoria systémy s minimálnym alebo zanedbateľným rizikom, ktoré AIA priamo neupravuje. Väčšina súčasných aplikácií AI používaných v rámci EÚ ako napríklad AI prvky vo videohrách či

11 *Ibid.*

filtre nevyžiadanej pošty, patrí práve do tejto skupiny, keďže nepredstavujú významné riziko pre bezpečnosť alebo základné práva jednotlivcov.

Používanie systémov na rozpoznávanie emócií a mentálnych stavov (angl.emotion recognition systems (ďalej ako „ERS")), ktoré predstavuje jednu z charakteristických funkcií neurotechnológií, je podľa AIA klasifikované ako vysoko rizikové. V určitých kontextoch najmä na pracovisku či vo vzdelávaní je zároveň normatívne vylúčené, s výnimkou ich využitia na medicínske alebo bezpečnostné účely. Je potrebné poznamenať, že systémy rozpoznávania emócií zahŕňajú inferenciu alebo identifikáciu emócií aj „zámerov". AIA pritom rozlišuje medzi emóciami (napr. šťastie, hnev, smútok) či úmyslami a fyzickými stavmi (ako únava), ktoré nemusia automaticky spadať pod ERS, pokiaľ sa nepoužívajú na odvodenie psychických stavov. Tieto rozdiely môžu významne sťažiť praktickú implementáciu regulačných požiadaviek zo strany poskytovateľov neurotechnológií. Kombinované využívanie neurotechnológií s ďalšími senzorickými vstupmi, napríklad s technológiami sledovania očí, môže viesť k inferencii informácií o emocionálnych či mentálnych stavoch, čo je podľa AIA neprístupné. Naopak, kategorizácia osôb na základe zdravotných alebo genetických charakteristík spadá medzi systémy označené ako „vysoko rizikové". Tieto síce nie sú explicitne zakázané, avšak podliehajú sprísnenému regulačnému režimu, vrátane rozsiahlych požiadaviek na posudzovanie bezpečnosti a súladu.[12] Predmetná situácia by bola relevantnou napríklad, ak by sa EEG v kombinácii s inými biometrickými údajmi použilo na odvodenie pravdepodobnosti vzniku Parkinsonovej choroby, epileptických záchvatov alebo duševného stavu osoby.[13] V praxi tak jeden a ten istý AI systém môže spĺňať kritériá pre viacero kategórií napríklad vysokorizikový alebo zakázaný systém v závislosti od jeho použitia. Preto je zásadné, aby vývojári a poskytovatelia neurotechnológií dôkladne analyzovali zamýšľané aj predvídateľné scenáre použitia a zvážili ich právnu klasifikáciu podľa AIA.[14]

16.2 Ľudsko-právne aspekty využívania moderných neurotechnológií

Jednou z hlavných prekážok postupného rozvoja neurotechnológií v spojitosti s vývojom AI je istá miera *„netransparentnosti"* mnohých pokročilých systémov a modelov, ktorú so sebou prináša rýchly technologický pokrok. Tento

---

12 SANTALU, N.: *Neurotechnologies under the EU AI Act: Where law meets science*. In: *IPPA*. Online: https://iapp.org/news/a/neurotechnologies-under-the-eu-ai-act-where-law-meets-science.

13 *Ibid.*

14 *Ibid.*

nedostatok predstavuje zásadnú výzvu pre ich uplatnenie, obzvlášť v prípadoch, keď sa systémy priamo interagujú s ľudským mozgom. Neurónové dáta sú jedinečne citlivé, nakoľko odhaľujú najintímnejšie procesy jednotlivca ako myšlienky, spomienky, duševné stavy, emócie, správanie, osobnosť a zdravotné problémy, a dokonca dokážu predpovedať budúce správanie, zdravotné riziká alebo kognitívny výkon. Rozhodujúce je, že tieto signály odrážajú aj podvedomú a nedobrovoľnú aktivitu a odhaľujú informácie, ktoré si jednotlivci nemusia vedome uvedomovať. V danej situácii je potrebné, okrem iného, *ex ante* zvážiť aj možné ľudsko-právne dôsledky využívania neurotechnológií, spojené predovšetkým so zásahom do mentálneho súkromia či integrity jednotlivca, ako aj nastavenie primeraných mechanizmov zodpovednosti, ktoré by zabránili ich zneužitiu a zabezpečili súlad s ľudskoprávnymi štandardmi. Neurotechnológie otvárajú teda nové horizonty medicíny (a nie len), no zároveň nastoľujú zásadnú otázku: Ako sa prudký technologický rozvoj neurotechnológií zlučuje s ochranou ľudských práv? Správa Poradného výboru Rady OSN pre ľudské práva (A/HRC/57/61) dospela okrem iného k záveru, že neurotechnológie zasahujú do oblasti ľudských práv mimoriadne osobitým spôsobom. Priame prepojenie ľudského mozgu s digitálnymi sieťami prináša závažné etické dôsledky pre základné hodnoty, na ktorých stojí systém ľudských práv, predovšetkým ľudskú dôstojnosť, súkromie a slobodnú vôľu jednotlivca, a zároveň môže predstavovať prostriedok, ktorým sa mení samotná ľudská podstata. Výbor zároveň zdôraznil, že uplatňovanie prístupu založeného na ľudských právach vo všetkých národných aj medzinárodných politikách musí byť jednou z hlavných priorít. Na základe uvedeného správa odporúča:

- pokračovať v diskusiách o vhodnosti vytvorenia osobitného mandátu pre oblasť nových technológií, ktorý by poskytoval usmernenia, ako zabezpečiť, aby sa neurotechnológie vyvíjali a používali v plnom súlade s ľudskými právami;
- uplatňovať zásadu náležitej starostlivosti (*due diligence*) pri regulácii, monitorovaní a sankcionovaní činností subjektov, ktoré neurotechnológie vyvíjajú, komercializujú alebo vyžadujú ich používanie, s cieľom predchádzať ohrozeniu uplatňovania ľudských práv a prijímať opatrenia na nápravu ich porušení;
- **vytvoriť ochranný regulačný rámec**, ktorý zohľadní osobitosti neurotechnológií vrátane ich súčasných aj potenciálnych dopadov na ľudské práva;
- prispôsobiť vnútroštátne právne poriadky predovšetkým v rámci občianskeho, trestného a pracovného právo, a to novým výzvam, ktoré

neurotechnológie prinášajú, a taktiež zároveň rozvíjať inštitucionálne mechanizmy, schopné predvídať a predchádzať porušovaniu ľudských práv;

- **posilniť kompetencie národných inštitúcií pre ľudské práva**, aby dokázali účinnejšie reagovať na nové technologické výzvy;
- aktívne sa zapájať do medzinárodných diskusií o riadení neurotechnológií a súvisiacich oblastí, ako je napríklad umelá inteligencia, a zvážiť prijatie medzinárodných nástrojov, ktoré by umožnili zavedenie moratória alebo zákazu používania technológií predstavujúcich riziko zneužitia – najmä v oblasti armády, presadzovania práva a trestnej justície;
- zabezpečiť, aby osoby so zdravotným postihnutím, staršie osoby a ďalšie zraniteľné skupiny mali prístup k bezpečným, spoľahlivým a cenovo dostupným neurotechnológiám v súlade s ľudskými právami, a aby boli efektívne chránené pred ich negatívnymi dopadmi a zneužitím počas vývoja a implementácie;
- zároveň zaručiť dostupnosť neurotechnológií pre osoby, ktoré z nich môžu reálne profitovať na zdravotné a medicínske účely.[15]

V súvislosti s uvedeným je dôležité sledovať, ako sa k problematike neurotechnológií bude v priebehu času stavať medzinárodné spoločenstvo (*societas maxima)*, či už prostredníctvom vyššie spomínaných aktivít OSN a EÚ, alebo činností iných regionálnych organizácií.

16.2.1 Potreba diskusií o nových „neuro“ - právach?

Niektorí vedci v súčasnosti rozvíjajú koncepty nových ľudských práv, ktoré by reagovali na etické a právne výzvy spojené s rozvojom neurotechnológií. Medzi najznámejšie prístupy patrí návrh Marcella Iencu a Roberta Andorna, ktorí formulovali ideu tzv. „*neuro - práv*“ (neurorights) štyroch nových práv: práva na mentálne súkromie**,** mentálnu integritu**,** psychologickú kontinuitu **a** kognitívnu slobodu.[16] Tieto práva podľa autorov čiastočne nadväzujú na štyri etické priority definované Yuste et al.,[17] no zameriavajú sa výslovne na právnu ochranu mentálnych procesov jednotlivca. *Právo na mentálne súkromie* má

---

15 HUMAN RIGHTS COUNCIL: *Impact, opportunities and challenges of neurotechnology with regard to the promotion and protection of all human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/57/61, 8 August 2024.

16 MARCELLO, I., ANDORNO, R.: *Towards new human rights in the age of neuroscience and neurotechnology*. In: Life Sciences Society and Policy, roč. 13, č. 1, 2017, s. 5. Online: https://doi.org/10.1186/s40504-017-0050-1.

17 YUSTE, R. et al.: *Four ethical priorities for. neurotechnologies and AI*. In: Nature, roč. 551, 2017, s. 159. Online: https://doi.org/10.1038/551159a.

chrániť citlivé informácie uložené v mysli pred neoprávneným zberom, používaním alebo vymazaním, a to ešte pred ich vonkajším prejavením. *Právo na mentálnu integritu* má poskytovať ochranu pred neautorizovanými zásahmi do mozgu. *Právo na psychologickú kontinuitu* sa sústreďuje na zachovanie stability osobnej identity a má chrániť pred zásahmi, ktoré by mohli meniť neurálne procesy či nevedome manipulovať jednotlivca, napríklad prostredníctvom neuromarketingu.[18] Napokon, *právo na kognitívnu slobodu* zdôrazňuje ochranu pred donucovaním či nevedomým využívaním neurotechnológií proti vôli jednotlivca. Alternatívny prístup predstavuje Andrea Lavazza, ktorý namiesto viacerých osobitných práv navrhuje zavedenie jedného komplexného práva na mentálnu integritu.[19] To by podľa neho zahŕňalo nielen ochranu pred zásahmi, ale aj právo na súkromie a sebaurčenie, čím by poskytovalo širší rámec ochrany ako návrh Iencu a Andorna. Na druhej strane, vedci ako Nora Hertz tvrdia, že existujúce právo na slobodu myslenia možno konzistentne interpretovať ako poskytujúce komplexnú ochranu mentálnych procesov a údajov z mozgu, čím sa vytvára normatívny základ pre používanie neurotechnológií. Podľa jej názoru je pritom vyvíjajúca sa interpretácia práva na slobodu myslenia presvedčivejšia než zavedenie nového ľudského práva na mentálnu sebauzavretosť. Právo na slobodu myslenia je zakotvené vo väčšine univerzálnych ako aj regionálnych ľudskoprávnych zmluvách:

[18] HERTZ, N.: *Neurorights – Do we Need New Human Rights? A Reconsideration of the Right to Freedom of Thought*. In: Neuroethics, roč. 16, č. 5, 2023. Online: https://doi.org/10.1007/s12152-022-09511-0.

[19] LAVAZZA, A.: *Freedom of thought and mental integrity: The moral requirements for any neural prosthesi*. In: Frontiers in Neuroscience, roč. 12, 2018, s. 82. Online: https://doi.org/10.3389/fnins.2018.00082.

Ilustrácia č. 2 - Prehľad univerzálnych a regionálnych ľudskoprávnych dohovorov upravujúcich právo na slobodu myslenia

Všeobecná deklarácia ľudských práv (OSN, 1948)	**Čl. 18**	**„Každý má právo na slobodu myslenia, svedomia a náboženstva;** toto právo zahŕňa slobodu zmeniť svoje náboženstvo alebo presvedčenie a slobodu prejavovať svoje náboženstvo alebo presvedčenie sám alebo spoločne s inými, verejne alebo súkromne, vyučovaním, vykonávaním, bohoslužbou a zachovávaním obradov.“
Medzinárodný pakt o občianskych a politických právach (ICCPR, 1966)	**Čl. 18 ods. 1**	**„Každý má právo na slobodu myslenia, svedomia a náboženstva.** Toto právo zahŕňa slobodu mať alebo prijať náboženstvo alebo presvedčenie podľa vlastnej voľby a slobodu prejavovať svoje náboženstvo alebo presvedčenie...“
Európsky dohovor o ľudských právach (EDLP, 1950)	**Čl. 9 ods. 1**	**„Každý má právo na slobodu myslenia, svedomia a náboženstva.** Toto právo zahŕňa slobodu zmeniť svoje náboženstvo alebo vieru a slobodu prejavovať svoje náboženstvo alebo vieru...“
Charta základných práv Európskej únie (2000)	**Čl. 10 ods. 1**	**„Každý má právo na slobodu myslenia, svedomia a náboženstva.** Toto právo zahŕňa slobodu zmeniť náboženstvo alebo vieru a slobodu prejavovať náboženstvo alebo vieru...“
Americký dohovor o ľudských právach (Pakt San José, 1969)	**Čl. 13 ods. 1**	**„Každý má právo na slobodu myslenia a prejavu.** Toto právo zahŕňa slobodu vyhľadávať, prijímať a rozširovať informácie a myšlienky akéhokoľvek druhu, bez ohľadu na hranice, ústne, písomne alebo tlačou, vo forme umenia, alebo akýmikoľvek inými prostriedkami podľa vlastného výberu.“
Arabská charta o ľudských právach (2004)	**Čl. 30 ods. 1**	**„Každá osoba má právo na slobodu myslenia, svedomia a náboženstva.** Toto právo zahŕňa slobodu zachovať si alebo zmeniť náboženstvo alebo presvedčenie a slobodu prejavovať svoje náboženstvo alebo presvedčenie...“

William Schabas však uvádza, že článok 18 Medzinárodného paktu o občianskych a politických právach *„vyžaduje, aby sa štát zdržal zásahov do duchovnej a morálnej existencie jednotlivca.*“ V kontexte neurotechnológií, napríklad pri aplikácii hlbokého mozgového stimulátora (DBS), by ochrana práva na slobodu myslenia platila len v prípade ovplyvnenia postojov, názorov alebo presvedčení jednotlivca pričom manipulácia s emóciami by pravdepodobne pod ochranu nespadala.[20] Z nášho pohľadu predstavuje ďalšie relevantné a explicitne zakotvené právo, ktoré je aplikovateľné na skúmanú problematiku, právo na súkromie. V tejto súvislosti v roku 2025 prostredníctvom svojho reportu A/HRC/58/58, vyzval Špeciálny spravodajca OSN pre právo na súkromie všetky štáty, aby implementovali špecifické regulačné rámce týkajúce sa neurotechnológií a neurónových údajov, vzhľadom na ich zásadný vplyv na ochranu súkromia, ľudskú dôstojnosť a základné ľudské práva. Ruud v rámci na diskusií, ktoré prebehli počas okrúhleho stola, spoločne organizovaného Radou Európy (ďalej len „CoE“) a Organizáciou pre

[20] Vychádzajúc z: BUBLITZ, C.: *Neurotechnológie a ľudské práva: opätovné stanovenie a potvrdenie viacvrstvovej ochrany osoby*. In: The International Journal of Human Rights, roč. 28, č. 5, 2024, s. 782–807. Online: https://doi.org/10.1080/13642987.2024.2310830.

hospodársku spoluprácu a rozvoj (ďalej len „OECD"), na základe ktorých bol vypracovaný Rapporteur Report s názvom *„Neurotechnologies and Human Rights Framework"* prezentoval svoj názor,[21] že namiesto vytvárania nových kategórií ľudských práv môže byť z hľadiska právnej systematiky a technickej realizovateľnosti efektívnejšie adaptovať existujúce práva tak, aby sa vzťahovali aj na kontext neurotechnológií. Argumentoval tým, že Európsky dohovor o ľudských právach už obsahuje ustanovenia schopné reagovať na nové neurotechnologické riziká, pokiaľ budú interpretované širšie práve v otázke súhlasu so zásahmi do neurálnych procesov, ktorý by nemal zahŕňať len súhlas s liečbou, ale aj so samotným monitorovaním mozgu, a to aj v prípade neinvazívnych postupov.

Martinho da Silva nadviazala na túto líniu argumentácie a uznala potrebu chrániť tzv. neuro-práva, avšak zdôraznila, že z právnoteoretického hľadiska je vhodnejšie vnímať neurotechnológie ako zdroj nových hrozieb pre existujúce práva, než ako dôvod na vytváranie úplne nových právnych kategórií. Upozornila na to, že súčasná spoločnosť je výrazne presýtená reguláciami a každý nový typ „práva" musí byť uplatniteľný, vykonateľný a zrozumiteľný. Poukázala na už existujúce právne piliere ako je ľudskú dôstojnosť, fyzická a mentálna integritu, právo na súkromie (vrátane mentálneho súkromia), sloboda myslenia, rovnosť a zásadu nediskriminácie, ktoré sú aplikovateľné aj na sféru neurotechnológií. Ako ilustráciu uviedla adaptáciu práva na súkromie v oblasti genetických technológií, ktorá môže slúžiť ako model aj pre neurotechnológie. Existujúcim príkladom ochrany implicitne zakotveného ľudského práva prostredníctvom širšieho výkladu explicitne zakotveného práva je aj prax ESĽP, ktorý v kontexte ochrany environmentálnych práv interpretuje právo na súkromný a rodinný život podľa čl. 8 Európskeho dohovoru o ľudských právach široko. Je dôležité poznamenať, že judikatúra ESĽP predstavuje jedinečný zdroj poznania aplikácie tzv. „evolutívneho výkladu".[22] Ako uvádza Dörr: *„Keď hľadáte dynamické prístupy k výkladu, judikatúra Európskeho súdu pre ľudské práva v Štrasburgu je*

[21] RADA EURÓPY (COE) a OECD: *Neurotechnologies and Human Rights Framework: Do We Need New Rights? (Round Table on the Human Rights Issues Raised by the Applications of Neurotechnologies),* 2021, Online: https://www.coe.int/sk/web/human-rights-and-biomedicine/round-table-on-the-human-rights-issues-raised-by-the-applications-of-neurotechnologies.

[22] BELLA, D.: Evolutívny výklad v praxi medzinárodných súdnych orgánov. In: PAĽUCHOVÁ, P. (zost.): *Vybrané otázky medzinárodného práva súkromného a verejného: zborník študentských príspevkov.* Bratislava: Univerzita Komenského v Bratislave, Právnická fakulta, 2024, 51 s.

zvyčajne jedným z prvých kandidátov, ktorí vám prídu na um.“[23] Vôbec prvým prípadom, v ktorom ESĽP použil evolutívny výklad, bol prípad Tyrer proti Spojenému kráľovstvu, v ktorom uviedol: *„Súd tiež musí pripomenúť, že Dohovor (pozn. autora: Dohovor o základných ľudských právach a slobodách) je živý nástroj, ktorý, ako Komisia správne zdôraznila, sa musí vykladať vo svetle súčasných podmienok. V súčasne prejedávanej veci súdu neostáva iné, ako byť ovplyvnený vývojom a všeobecne uznávanými normami v trestnej politike členských štátov Rady Európy.“*[24] Tento prístup ESĽP naznačuje, že podobná metodológia výkladu by mohla v budúcnosti slúžiť ako základ pre rozšírenú ochranu práv aj v súvislosti s využívaním neurotechnológií**.** Vzhľadom na vyššie spomínaný koncept vývoja nových „neuro-práv“ je dôležité spomenúť aj samotné požiadavky na rozšírenie rámca ľudských práv, ktoré často predstavujú trend vychádzajúci z analýz praktických výziev a technologických zmien, ktoré odhaľujú medzery v existujúcich ľudsko-právnych normách. Prvým krokom k právnemu uznaniu nového ľudského práva môže byť jeho začlenenie do prameňov povahy *soft law.. Soft law* síce nie je právne záväzný prameň *strictu sensu*, avšak disponuje špecifickými vlastnosťami, a to predovšetkým tým, že poskytuje usmernenia pre interpretáciu medzinárodného práva, pričom ho možno definovať ako flexibilný a ľahko prijateľný, čo ho predurčuje pre oblasti, ktoré sa rýchlo vyvíjajú alebo menia, alebo kde existuje konsenzus medzi štátmi, no sú zdráhavé prijať záväzné záväzky. *Soft law* navyše disponuje morálnou autoritou a vytvára štandard očakávania, často sa stáva základom pre vznik obyčajového pravidla alebo prijatie zmluvy. V rámcoch vyššie spomínaného reportu CoE a OECD, Roure zdôraznila potrebu mechanizmov na ochranu integrity, dôstojnosti, autonómie, súkromia a zásady nediskriminácie a varovala pred zásahmi, ktoré by mohli obmedziť slobodu a sebaurčenie jednotlivcov, napríklad prostredníctvom invazívnych dohľadových aktivít. Pričom ako realistickú cestu v predmetnej problematike taktiež označila *soft-law* nástroje.

Koncept nových ľudských práv môže vzniknúť aj ako obyčajové pravidlo, avšak dokázať dve obligatórne podmienky obyčajového pravidla, a to *usus longaevus* a *opinio juris sive necessitatis* je z povahy veci náročné. Preto je bežnejšie zavádzať nové ľudské práva prostredníctvom medzinárodnej zmluvy alebo dodatkového protokolu, resp. úpravou existujúcich zmlúv o ľudských právach. Právne uznanie nových ľudských práv môže byť tiež podporené rozhodnutiami medzinárodných súdnych a kvázi-súdnych orgánov, ako je ESĽP

---

[23] ABI-SAAB, G. et al.: *Evolutionary Interpretation and International Law*, 1. vydanie, Oxford: Hart Publishing, 2019, s. 116.

[24] Rozhodnutie ESĽP vo veci *Tyrer proti Spojenému kráľovstvu*, sťažnosť č. 5856/72, rozsudok z 25. apríla 1978.

alebo ľudskoprávnych výboroch OSN (angl. Treaty bodies) OSN, napríklad prostredníctvom ich všeobecných komentárov alebo širšej interpretácie existujúcich, explicitne zakotvených ľudských práv. Je však potrebné brať do úvahy, že v prípade rozhodnutí ľudsko-právnych Výborov OSN ide o rozhodnutia s čisto odporúčajúcim charakterom s právne nezáväznou povahou. Hoci sa často predpokladá, že vytvorenie nového ľudského práva by mohlo odstrániť istú dávku neistoty a vyplniť právne medzery v oblasti ochrany mentálnych procesov, nie je zrejmé, že takýto krok by viedol k väčšej právnej istote alebo silnejšej ochrane. Existujúce a explicitne zakotvené práva ako napríklad už viackrát spomínané právo na slobodu myslenia alebo právo na súkromie, už v súčasnosti poskytujú normatívny základ pre reguláciu neurotechnológií a ich dynamická interpretácia môže ponúknuť dostatočný rámec pre riešenie nových technologických výziev. Hertz tvrdí, že právo na slobodu myslenia možno koherentne vykladať tak, aby zahŕňalo široké spektrum mentálnych procesov a foriem kognitívnej autonómie.[25] Treba tiež zdôrazniť, že samotný nárast nových ľudských práv môže mať kontraproduktívne účinky, a to tým, že zvyšuje nároky na dohľad a monitorovanie a môže oslabovať účinnosť ochrany tých práv, ktoré sú skutočne fundamentálne. Argument, že navrhované „neuro- rights" sú už implicitne obsiahnuté v existujúcich právach, skôr podporuje potrebu ich progresívnej interpretácie než zavedenie úplne nových kategórií práv. Z týchto dôvodov sa zdá presvedčivejšie sústrediť sa na „evolutívny výklad" existujúcich ľudských práv než usilovať o tvorbu nových práv, ktorých pridanie by nemuselo automaticky posilniť ľudskoprávny rámec a mohlo by ho dokonca fragmentovať. Takýto prístup zároveň umožňuje reagovať na výzvy neurotechnológií spôsobom, ktorý je právne konzistentný, normatívne stabilný a medzinárodne akceptovateľný.

Záver

V závere možno konštatovať, že otázka ochrany ľudských práv v kontexte neurotechnológií nevyžaduje nevyhnutne vytváranie úplne nových ľudských práv. Existujúce práva, ako právo na súkromie a právo na slobodu myslenia, poskytujú dostatočný právny rámec na reguláciu nových technologických výziev, ak sú interpretované dynamicky a progresívne. Evolutívny súdny výklad umožňuje medzinárodným súdnym aj kvázi-súdnym orgánom prispôsobiť význam týchto práv aktuálnym spoločenským, technologickým a etickým

[25] HERTZ, N.: *Neurorights – Do we Need New Human Rights? A Reconsideration of the Right to Freedom of Thought*. In: Neuroethics, roč. 16, č. 5, 2023. Online: https://doi.org/10.1007/s12152-022-09511-0.

podmienkam, čím sa zabezpečuje ich účinná ochrana bez fragmentácie právneho systému. Použitie *soft-law* mechanizmov a odporúčacích nástrojov medzinárodných orgánov podporuje implementáciu existujúcich práv a poskytuje spoľahlivý základ pre budúcnosť otázky týkajúcej sa ľudských práv a neurotechnológií. Zavádzanie nových práv by mohlo viesť k zvýšeniu administratívnej a dohľadovej záťaže a oslabiť účinnosť ochrany práv, ktoré sú fundamentálne. Z právneho hľadiska sa preto ukazuje ako efektívnejší prístup sústrediť sa na adaptáciu a evolutívny výklad existujúcich ľudských práv, čím sa dosahuje právna istota, koherencia a medzinárodná akceptovateľnosť pri regulácii neurotechnológií. Tento prístup zároveň umožňuje adresovať nové technologické riziká spôsobom, ktorý rešpektuje integritu, dôstojnosť a autonómiu jednotlivca. Celkovo progresívna interpretácia existujúcich práv predstavuje teoreticky aj prakticky presvedčivejšiu cestu k ochrane mentálnej autonómie a „kognitívnych práv", pričom zabezpečuje stabilitu a konzistenciu ľudsko-právneho rámca v meniacom sa technologickom prostredí.

Zoznam použitej literatúry a zdrojov

1. ABI-SAAB, G. et al.: *Evolutionary Interpretation and International Law*, 1. vydanie, Oxford: Hart Publishing, 2019, s. 116.
2. BELLA, D.: *Evolutívny výklad v praxi medzinárodných súdnych orgánov*. In: PAĽUCHOVÁ, P. (zost.): Vybrané otázky medzinárodného práva súkromného a verejného: zborník študentských príspevkov. Bratislava: Univerzita Komenského v Bratislave, Právnická fakulta, 2024, 51 s.
3. BUBLITZ, C.: *Neurotechnológie a ľudské práva: opätovné stanovenie a potvrdenie viacvrstvovej ochrany osoby*. In: The International Journal of Human Rights, roč. 28, č. 5, 2024, s. 782–807. Online: https://doi.org/10.1080/13642987.2024.2310830.
4. CARD, N.S. et al.: *An accurate and rapidly calibrating speech neuroprosthesis*. In: New England Journal of Medicine, roč. 391, 2024, s. 609–618.
5. EURÓPSKA KOMISIA: *AI Act*, posledná aktualizácia 11. mája 2026. Online: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.
6. HENSING, J., FAIKA, A.: *Neurotechnology: A New Frontier for Prosperity and Security in Germany and Europe?* In: GPPi, 2025. Online: https://gppi.net/2025/03/27/neurotechnology-a-new-frontier-for-prosperity-and-security-in-germany-and-europe.
7. HERTZ, N.: *Neurorights – Do we Need New Human Rights? A Reconsideration of the Right to Freedom of Thought*. In: Neuroethics, roč. 16, č. 5, 2023. Online: https://doi.org/10.1007/s12152-022-09511-0.
8. HUMAN RIGHTS COUNCIL: *Impact, opportunities and challenges of neurotechnology with regard to the promotion and protection of all human rights – Report of the Human Rights Council Advisory Committee*, A/HRC/57/61, 8 August 2024.
9. LAVAZZA, A.: *Freedom of thought and mental integrity: The moral requirements for any neural prosthesi*. In: Frontiers in Neuroscience, roč. 12, 2018, s. 1–10. Online: https://doi.org/10.3389/fnins.2018.00082.
10. MARCELLO, I., ANDORNO, R.: *Towards new human rights in the age of neuroscience and neurotechnology*. In: Life Sciences Society and Policy, roč. 13, č. 1, 2017, s. 1–27. Online: https://doi.org/10.1186/s40504-017-0050-1.
11. PATRICK-KRUEGER, K.M., BURKHART, I. a CONTRERAS-VIDAL, J.L.: *The state of clinical trials of implantable brain–computer interfaces*. In: Nature Review Bioengineering, 2025, s. 50–67. Online: https://doi.org/10.1038/s44222-024-00239-5.

12. RADA EURÓPY (COE) a OECD: *Neurotechnologies and Human Rights Framework: Do We Need New Rights? (Round Table on the Human Rights Issues Raised by the Applications of Neurotechnologies),* 2021. Online: https://www.coe.int/sk/web/human-rights-and-biomedicine/round-table-on-the-human-rights-issues-raised-by-the-applications-of-neurotechnologies.
13. Rozhodnutie ESĽP vo veci *Tyrer proti Spojenému kráľovstvu*, sťažnosť č. 5856/72, rozsudok z 25. apríla 1978.
14. SANTALU, N.: *Neurotechnologies under the EU AI Act: Where law meets science*. In: IPPA, 2025. Online: https://iapp.org/news/a/neurotechnologies-under-the-eu-ai-act-where-law-meets-science.
15. SZOSZKIEWICZ, Ł., YUSTE, R.: *Mental privacy: navigating risks, rights and regulation : Advances in neuroscience challenge contemporary legal frameworks to protect mental privacy.* In: EMBO Reports, 2025, č. 14, s. 3469–3473. Online: https://doi.org/10.1038/s44319-025-00505-6.
16. WILLETT, F.R. et al*.: A high-performance speech neuroprosthesis*. In: *Nature*, roč. 620, 2023, s. 1031–1036. Online: https://doi.org/10.1038/s41586-023-06377-x.
17. XU, Y., VAZIRI-PASHKAM, M.: *Limits to visual representational correspondence between convolutional neural networks and the human brain*. In: Nature Communications, roč. 12, č. 1, Article 2065, 2021, s. 1–16. Online: https://doi.org/10.1038/s41467-021-22244-7.
18. YUSTE, R. et al.: *Four ethical priorities for. neurotechnologies and AI*. In: Nature, roč. 551, 2017, s. 159–163. Online: https://doi.org/10.1038/551159a.
19. ZHANG, H. et al.: Brain-computer interfaces: the innovative key to unlocking neurological conditions. In: *Int J Surg*, roč. 110, č. 9, 2024, s. 5745–5762. Online: https://doi.org/10.1097/JS9.0000000000002022.

Informácie o autorovi

Mgr. Laura GAZDAGOVÁ
Univerzita Pavla Jozefa Šafárika v Košiciach
Právnická fakulta
Ústav medzinárodného práva a európskeho práva
Email: laura.gazdagova@student.upjs.sk

17 Žurnalistický prejav v ére sociálnych sietí a AI – úsvit alebo súmrak slobody prejavu?

Jakub Vanacký

https://doi.org/10.33542/HRF-0541-2-17

Abstrakt

Príspevok sa venuje skúmaniu súhry dvoch aktuálnych javov – straty primátu tradičných médií a nárastu využívania umelej inteligencie – a ich spoločenským dôsledkom. Cieľom je pomocou empirickej a normatívnej analýzy objasniť vplyv týchto procesov na žurnalistickú prax, právne prostredie a etické štandardy mediálnej komunikácie. Príspevok analyzuje súčasný stav právnej úpravy žurnalistického prejavu a umelej inteligencie, ako aj recentný vývoj v oblasti žurnalistiky. Identifikuje materiálne pramene práva a ponúka alternatívy regulácie, najmä prostredníctvom profesijného etického kódexu. Na tento účel zas využíva axiologickú analýzu. Osobitná pozornosť sa venuje významu umelej inteligencie a nových technológií v kontexte „vojny" prebiehajúcej v informačnom priestore. Záver príspevku prináša úvahy de lege ferenda o možnostiach budúcej regulácie a o ďalšom vývoji slobody prejavu v oblasti žurnalistiky a nových médií. Prínos príspevku vidíme vo význame skúmania tejto problematiky, v záujme zdravého vývoja demokratickej spoločnosti.

Úvod

Súčasný stav mediálneho priestoru môže v spoločnosti vyvolávať vášne.[1] Zároveň je nepochybné, že aj pohľad na to, čo sa rozumie pod „kvalitnou" žurnalistikou sa môže líšiť, dokonca až tak výrazne, že pohľady možno považovať až za „výrazne protikladné".[2] Niekomu sa bude zdať mediálny priestor absolútne v poriadku. Pre niekoho bude príliš zahltený negativizmom. Ďalší bude tvrdiť, že média v súčasnosti neinformujú o dôležitých veciach a radšej sa zameriavajú na neutrálne informácie o bezvýznamných udalostiach, iný povie, že informujú jednostranne, alebo prípadne, že všetky „tradičné" médiá klamú, či pretláčajú

[1] Autor príspevkom nadväzuje na svoje štúdium skúmanej problematiky, predovšetkým na svoju diplomovú prácu. Na niektorých miestach využíva poznatky nadobudnuté a prezentované už skôr v spomínanej diplomovej práci. Porov.: VANACKÝ, J.: *Sloboda prejavu v digitálnom priestore*. školiteľ: Michal Mrva. Bratislava: Univerzita Komenského v Bratislave, 2025. 123 s.

[2] MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, s. 28.

jeden názor a je nutné hľadať „pronárodnejšie“ alternatívy, ako uvádza napríklad Slovenská asociácia novinárov.[3]

Tento príspevok sa nebude zaoberať riešením komplexného problému žurnalistiky, ktorý môže pre každého spočívať v inom jej aspekte. Zároveň jeho účelom nemá byť iba upriamiť pozornosť na potrebu tvorby etických pravidiel pre oblasť AI. Jeho primárnym účelom je zdôvodniť, že pre zachovanie významu žurnalistiky vo verejnom priestore a zároveň zachovanie slobody prejavu sa vyžaduje, aby sa okrem právnej úpravy v kontexte tejto problematiky prispôsobili aj profesijne kódexy.

Sekundárne, je účelom tohto príspevku apelovať na profesijné organizácie žurnalistov aby prijali také etické kódexy, ktoré budú reflektovať meniacu sa dobu a budú poskytovať praktické nástroje pre vývojárov AI. Vedľajším cieľom, či prínosom je v tomto príspevku preskúmanie oblasti vzájomného vzťahu medzi právom, AI, žurnalistikou a spoločnosťou, ktoré môže prispieť k zdravému vývoju v demokratickej spoločnosti.

Hypotézy, ktoré budeme v tomto príspevku overovať sú dve. Prvá: Ústup tradičných médií mení mediálne prostredie a odhaľuje nové možnosti deformácie informačného priestoru.

A druhá: Nové právne a etické mechanizmy, vrátane využitia umelej inteligencie, sú nevyhnutné na ochranu slobody prejavu a práva na informácie pred spomínanými deformáciami.

Základne výskumné otázky, ktoré si v tomto príspevku za účelom pokladáme sú: Ako sa ústup tradičných médií prejavuje a aké spoločenské dôsledky tento jav prináša? Ako reagujú mechanizmy etickej samoregulácie a právna úprava v Slovenskej republike na nové formy mediálnej komunikácie a šírenia informácií? Aká je budúcnosť využitia umelej inteligencie pri žurnalistike a pri regulácii mediálneho priestoru? Je možná bez narušenia základných práv a slobôd? Akú úlohu môže zohrať odborná a etická samoregulácia (napr. prostredníctvom etických kódexov) pri formovaní zodpovedného mediálneho prostredia?

V nadväznosti na uvedené sa pokúsime zodpovedať na nasledujúce pridružené otázky: Čo sú to tradičné médiá a ako sa mení ich postavenie? Ako[4] je oblasť žurnalistiky regulovaná na Slovensku? Aký je význam profesijnej etiky v

---

[3] SLOVENSKÁ ASOCIÁCIA NOVINÁROV: Vyhlásenie Slovenskej asociácie novinárov zo dňa 17.5.2024. Slovenská asociácia novinárov, 17. máj 2024. Online: https://www.novinarisan.sk/2024/05/17/vyhlasenie-san-k-atentatu-na-roberta-fica-2/.

[4] V zmysle výraznejšieho vplyvu (právne nezáväzných) etických kódexov alebo záväznej právnej úpravy.

žurnalistike dneška? Má sa prispôsobiť ona novému rozmeru mediálnej sféry alebo sa má snažiť nový rozmer mediálnej sféry prispôsobiť zaužívaným štandardom? Aké zmeny, či hrozby možno očakávať v masmediálnom priestore v súvislosti s príchodom nových technológií? Je prípustné využitie umelej inteligencie v záujme zachovania pluralitného a demokratického verejného diškurzu? Ak áno, akým spôsobom?

Na uvedené otázky sa pokúsime odpovedať pomocou viacerých metód, predovšetkým však prostredníctvom deskriptívnej normatívnej analýzy právnych predpisov a axiologickej analýzy žurnalistickej profesie. V nasledujúcej časti sa však najprv zamerajme na objasnenie niekoľkých základných pojmov.

17.1 Tradičné média v kontexte nových technológii

Prioritou tejto časti príspevku bude ozrejmiť čo máme na mysli pod pojmami tradičné/nové médiá a aký je vzťah medzi tradičnými médiami a novými technológiami. Logická stavba, ako napovedá názov kapitoly, teda sleduje najprv zmeny v oblasti médií a následne ilustruje najvýznamnejšiu zmenu v oblasti technológií v súčasnosti – stále väčšiu popularitu umelej inteligencie. Nakoľko si úvahy o skúmanej problematike vyžadujú aj istú mieru teoretických poznatkov zo skúmanej oblasti, v úvode ponúkneme aj krátky historický exkurz a teoretickú úvahu o povahe žurnalistiky. Neodškriepiteľnou súčasťou tejto kapitoly bude aj odpoveď na otázku, aké javy môžeme pozorovať v súvislosti so zmenou pozície tradičných médií ako primárneho zdroja informácií pre majoritu v populácii. Zároveň v tejto časti nášho príspevku formulujeme základné východiská pre naše využívanie pojmu „AI“ v druhej časti príspevku.

Otázkou, ktorú však považujeme za podstatnú zodpovedať skôr než prejdeme k „jadru veci“ je: Akým spôsobom súvisí žurnalistika s ľudskými právami? Myslíme si, že táto otázka už bola zodpovedaná v množstve učebníc žurnalistiky. Preto je možno vhodnejšie pýtať sa, prečo je oprávnené zameriavať sa na jej zmeny v súvislosti s umelou inteligenciou? Odpoveď znie, že žurnalistika, ako inštitúcia dohliadajúca na politiku, je „lakmusovým papierom“ ohrozenia demokracie a ľudských práv. Pokiaľ v akejkoľvek krajine má hroziť, že dôjde k úpadku týchto ideálov, sloboda prejavu a právo na (pravdivé) informácie, ako imanentné nástroje žurnalistiky, budú prvé, ktoré budú signalizovať hrozbu. V súvislosti s týmto tvrdením i pokiaľ by nemalo dôjsť k ohrozeniu spomínaných ideálov, predpokladáme značný súvis, blížiaci sa k úmernosti, medzi kvalitou žurnalistickej produkcie a kultivovaním verejnej diskusie a teda aj medzi žurnalistikou a demokraciou.

Niekto by mohol namietať, že ide o vágne východisko – veď práve v neslobode totalít, v cenzúrou jasne definovaných hraniciach medzi propagandou, teda štátom schvaľovaných médiách a reálnou žurnalistikou, teda disidentskými samizdatmi, je možné vidieť nepriamu úmeru! Čím nižšia sloboda, tým kvalitnejšia žurnalistika! Avšak takúto argumentáciu nepovažujeme za relevantnú hneď z viacerých dôvodov.

Prvý – mediálnu oblasť je potrebné hodnotiť komplexne. Nedá sa povedať, že žurnalistický produkt je kvalitnejší v nedemokratickom zriadení, nakoľko aj štátom schvaľovaná produkcia je súčasťou a tá, skrze nerešpektovanie základného princípu žurnalistickej práce – pravdivosti – nemusí vôbec napĺňať požiadavky[5] na označenie „kvalitná". V súčte potom v takomto zriadení máme jediné – jasnejšie definovanú hranicu medzi kvalitou a nekvalitou.

Druhým dôvodom, prečo odmietame vyššie uvedené tvrdenie je, že spomínané by znamenalo, že v demokratickom zriadení kvalita žurnalistického produktu, oproti nedemokratickému zriadeniu klesá. To už je ale nutné zahrnúť aj ekonomický rozmer problematiky. V demokratickom zriadení, v trhovo orientovanej ekonomike sa bude kvalita meniť podľa požiadaviek trhu. Podľa zmien na strane ponuky a dopytu. V demokratickej spoločnosti, teda v spoločnosti, ktorá si hneď po práve na život a osobnú integritu a po volebných právach, najväčšmi musí ceniť slobodu prejavu a právo na informácie, potom platí naše tvrdenie. Ak budeme podporovať hodnoty pravdivosti – teda informovanosti a slobody médií, bude rásť aj demokratické zmýšľanie v spoločnosti. Inak povedané, ak predpokladáme, že demokratickú spoločnosť konštituujú jednotlivci, ktorých hodnoty sú sloboda a pravda, potom možno predpokladať, že sa priamo úmerne tomuto bude chovať aj mediálny priestor. V konečnom dôsledku tak potom stále väčšia časť jednotlivcov bude považovať tieto hodnoty za dôležité, čo povedie k zvýšeniu občianskej angažovanosti a povedomia teda následne k demokratickejšiemu spoločenstvu.

Znamená to snáď, že v demokratickom spoločenstve bude kvalita konštantne rásť? Určite nie. Tu by bolo možné do hĺbky rozoberať nuansy vzťahu medzi dopytom a ponukou a problematiku trhového hospodárstva, to však nemáme v úmysle. Ekonomické hľadisko nemá byť gro tohto príspevku. Na čo sa však zameriame je otázka spoločenskej objednávky smerom k žurnalistike. Táto objednávka bude musieť najprv nájsť odpoveď na otázku – čo od žurnalistiky chceme? Kvalitu, či kvantitu? Odporuje jedno druhému? A ak sa rozhodneme pre niektorú odpoveď, máme očakávať, že riešenie jednoducho sprostredkuje trh,

[5] Predovšetkým požiadavku aby bola objektívna, založená na faktoch a nie podriadená štátnej moci.

alebo je namieste požadovať reakciu tak povediac „z hora“, či možno skôr „z vnútra“?

17.1.1 O povahe žurnalistiky

Ak v právnej teórii zaznamenávame rôzne pohľady na vzťah medzi právom, politikou a etikou (či morálkou), potom možno konštatovať, že rovnako tak sa môže líšiť percepcia vzťahu práva a žurnalistiky. Žurnalistika predstavuje koncept, ktorý svojím spoločenským dosahom vplýva na politiku a skrze ekonomickú oblasť sú tieto oblasti úzko prepojené,[6] dalo by sa povedať vzťahom podobným tretiemu Newtonovmu zákonu pohybu.[7] Spolu s ďalšími inštitúciami (ako cirkvi, či vzdelávacie a kultúrne inštitúcie) žurnalistika formuje obraz spoločnosti, v spoločnosti.

Nemožno predpokladať, že žurnalistika verejné dianie, verejnú mienku, či morálku spoločnosti iba objektívne reflektuje alebo nezúčastnene komentuje, hoci sa niekto môže nazdávať, že taký by mal byť jej prístup. Otázkou, ktorú však na tomto mieste presvedčivo nezodpovieme, môže byť, či niekedy existovala skutočne objektívna žurnalistika, alebo ide iba o ideál, ku ktorému sa upína, respektíve by sa mala podľa niekoho názoru upínať. Hoci, vyslovme názor, že reálnejšie sa zdá to druhé. Žurnalistika je teda jedným zo zásadných aktérov v politickom súboji. V tomto smere je dobré upozorniť, že poznáme aj pojem verejnoprávnych médií, ktorých prínos by mal byť práve v spomínanej nezávislej žurnalistickej práci. Otázne je, do akej miery je takýto ideál prakticky naplniteľný a ak áno, akým spôsobom sa k nemu vieme dopracovať. V kontexte posledných rokov je diskusia o verejnoprávnych médiách a predovšetkým o snahách o ich politizáciu na Slovensku no v princípe naprieč celou Európskou úniou stále pomerne živá.[8]

Žurnalistika v známej funkcii „strážneho psa“ dohliada na vývoj práva a prácu politikov no má aj mnoho ďalších funkcií. McQuail spomína predovšetkým funkciu významného aktéra v socio-politickej sfére ako aj

---

6 REMIŠOVÁ, A.: *Etika médií*. Bratislava: Kalligram, 2010, s. 32–37. Pozn.: Zároveň je na žurnalistiku vplývané zo strany politiky, ako aj z ekonomickej sféry – z hľadiska obsahu, no často i formy. Myslieť si iné sa zdá naivné, hoci by sme sa mohli nazdávať, že v ideálnej, hoci možno príliš utopickej, spoločnosti by nemalo ísť o priamy vplyv.

7 Ktorý znie: *„Každá akcia vyvoláva rovnako veľkú a opačne orientovanú reakciu. Sily pôsobenia sú vždy vzájomné.“* Porov.: THORNTON, S. T.; MARION, J. B.. *Classical Dynamics of Particles and Systems*. 5. vydanie. Belmont: Brooks/Cole, 2004, s. 49.

8 REUTERS INSTITUTE FOR THE STUDY OF JOURNALISM: *Digital News Report 2025 – Slovakia.* 2025. Online: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025/slovakia.

občianskej spoločnosti, národ zjednocujúcu funkciu, či funkciu „advokácie“[9] za záujmy istej časti publika, či strany istej kauzy.[10] V takom prípade je istotne otázkou aj existencia novinárskej objektivity, ktorú sme spomínali vyššie. Každopádne, tieto funkcie napĺňajú rôzne médiá v rôznej miere, čo častokrát vyplýva aj z názvu.[11] V súvislosti s takýmto posudzovaním médií (na základe názvu) pôsobia zaujímavo názvy niektorých komerčných televízií na Slovensku – „Markíza“ a „JOJ“, viac pochopiteľne, avšak stále celkom nevýpovedne[12] potom vyznieva názov „TA3“. Túto krátku polemiku, uvádzame skôr na odľahčenie.

Ak sme však skôr uviedli, že žurnalistika je zásadným aktérom v politickom boji, potom je nutné spomenúť, že skrze svoje početné funkcie a skrz vplyv na verejnú mienku, významný bude aj dopad žurnalistiky na, v súčasnosti prebiehajúce, konflikty na medzinárodnej, no i vnútroštátnej, politickej úrovni. Na Slovensku môže aj menej pozorný divák zaregistrovať, aké médiá považuje vládnuca časť politického spektra za „problémové“, čo súvisí prevažne s názorovými rozpormi. To značí že pokiaľ ide o povahu žurnalistiky, tak z pohľadu politiky ide o nástroj, ktorého povaha bude v podstatnej miere relativistická a bude závisieť od politickej afiliácie toho ktorého média.

Právo zas využíva regulačné prostriedky na reguláciu žurnalistického prejavu a formáciu mediálneho priestoru. Nemožno zamlčať ani vplyv žurnalistiky na utváranie obrazu o práve a právnikoch, súdoch a súdnych rozhodnutiach. Pre právo predstavuje žurnalistika prostriedok prezentácie pluralitného názorového spektra a najmä realizácie základných práv, totiž práva na informácie a slobody prejavu. V tomto ohľade je povaha žurnalistiky objektívnejšia, pretože tieto funkcie by sa mal v demokratickom zriadení, v nejakej miere, snažiť napĺňať každý publikujúci subjekt.

Z historického hľadiska - sme boli svedkami mediálnej revolúcie v čase, keď orálne zdieľanie aktuálnych informácií urýchlila kníhtlač, ktorá umožnila masovú produkciu tlačovín, od letákov s každodennými[13] informáciami, až po

---

9 V zmysle verejnej podpory myšlienky, plánu či spôsobu konania; nie ako právnické povolanie „advokát“.

10 MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, s. 102.

11 *Ibid*, s. 40.

12 Nevýpovedne z hľadiska informovanie o tom, ktorú funkciu napĺňa toto médium. Spomeňme, že zaujímavé by bolo reflektovať aj názvy ďalších médií, napr.: SME, Hlavné správy, (Denník)N, Pravda, +1deň, či iné...

13 Tu možno vidieť prepojenie s etymológiou slova – žurnal(-istika), zo staro-francúzskeho „*jornel*“- t.j. „deň“, prípadne latinského „diurnalis“, t.j. „denný“. Porov.: MERRIAM-WEBSTER: *Journal*. In: Merriam-Webster Dictionary. Online: https://www.merriam-webster.com/dictionary/journal#word-history.

komplexnejšie, viacstranové spisy o obchode, politike či spoločnosti.[14] Neskôr ďalšie zmeny priniesli možnosti telekomunikácie prostredníctvom rádiových vĺn rozhlasu. Nepochybnou inováciou v zdieľaní informácií boli televízne stanice. V súčasnosti sa do popredia dostávajú sociálne siete a osobitne aj internetové vyhľadávače. Popri všetkých spomínaných metódach masovej komunikácie si printové médiá zachovali svoju pozíciu ako hlavného zdroja informácií pre širokú populáciu.[15] Uvedené platilo pre väčšiu časť elektrifikovaného sveta, až donedávna.

Na základe štatistických údajov vieme konštatovať, že v súčasnosti už tento primát stratili. Pre reálie Slovenskej republiky sú relevantné vo významnej miere údaje Českého štatistického úradu, ktoré vypovedajú aj o občanoch Slovenska: Internet využíva ako zdroj spravodajských informácií na Slovensku až 63 % obyvateľov vo veku od 16 po 74 rokov a až 57 % tejto časti populácie na tento účel využíva sociálne siete.[16][17] Zároveň aj v publikáciách zahraničných odborníkov z oblasti výskumu no i praxe nachádzame úvahy o budúcnosti tradičných médií.[18]

Čo na základe uvedeného môžeme tvrdiť o povahe žurnalistiky dnes? Domnievame sa, že predovšetkým to, že sa mení. Osobitné zmeny prinášajú do žurnalistiky nové médiá a s nimi aj nové technológie. Našim úmyslom bude v nasledujúcej časti, venovať sa týmto fenoménom a čiastočne rozpracovať akým spôsobom definujeme ich obsah, s ohľadom na tému nášho príspevku.

---

14 MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, s. 23.

15 STATISTA: *Most trusted sources of news and information worldwide*, 2025. Online: https://www.statista.com/statistics/381455/most-trusted-sources-of-news-and-info-worldwide/.

16 ČESKÝ STATISTICKÝ ÚŘAD: *Informační společnost v číslech – 2024*, s. 11. Online: https://csu.gov.cz/docs/107508/76df3534-a565-79c7-41c7-ce6cf2f44737/06100424c.pdf?version=1.0.

17 VANACKÝ, J.: *Sloboda prejavu v digitálnom priestore*. školiteľ: Michal Mrva. Bratislava: Univerzita Komenského v Bratislave, 2025. s. 21.

18 Porov.: NGUYEN, M. N.: JOURNALISM AND SOCIAL MEDIA: *The Transformation of Journalism in the Age of Social Media and Online News*. In: European Journal of Social Sciences Studies, Vol. 7, No. 6, 2022. Online: https://www.researchgate.net/publication/365132142_JOURNALISM_AND_SOCIAL_MEDIA_THE_TRANSFORMATION_OF_JOURNALISM_IN_THE_AGE_OF_SOCIAL_MEDIA_AND_ONLINE_NEWS; alebo aj https://rsf.org/en/rsf-publishes-new-report-protect-europe-s-public-media ;

17.1.2 Zmeny v mediálnom prostredí v kontexte základných práv a slobôd

V tejto časti príspevku považujeme za podstatné špecifikovať niektoré pojmy, ktoré budú viac figurovať v našom príspevku. Ide predovšetkým o pojmy, ktoré sa môžu svojim obsahom zdať priširoké a preto je vhodné ich pre účely nášho príspevku, zúžiť. Predovšetkým ide o pojmy „tradičné médiá", „nové médiá" a „AI". Samotný pojem „médiá" budeme na účely tohto príspevku považovať za notorietu, ktorú spôsobom s ktorým sa stotožňujeme my i Remišová, definuje už Schellmann.[19]

Na požiadavku definície pojmu tradičných médií je možné odpovedať z viacerých hľadísk. Z chronologického hľadiska by sme pod pojmom „tradičné médiá" rozumeli tie, ktoré vznikli pred digitálnou érou, teda na škále od tradičných po nové médiá, by sa bližšie k tradičným nachádzali printové a rozhlasové médiá, niekde v prostriedku televízne médiá a na opačnej strane internet a sociálne siete. Toto kritérium samotné je však stále nepostačujúce, pretože napríklad televízne médiá nezaradzuje jednoznačne. V prípade využitia iného kritéria ako chronologického, napríklad z hľadiska miery využitia nových technológií, sa nezdá že by sa rozloženie na spektre výrazne odlišovalo. Preto je nevyhnutné vychádzať pri definícii pojmu „tradičné médiá" predovšetkým z toho, čím sa odlišujú od „nových médií". Takýto prístup nám umožní jednoznačnejšie oddeliť nové médiá od tých tradičných.

Teda jednoznačnejšie kritérium, ktoré nám zároveň pomáha zadefinovať aj pojem „nové médiá" sleduje odlišnosti medzi novými a tradičnými médiami. Pre tento účel sledujeme kľúčové atribúty, ktoré dané médium prináša do mediálneho priestoru. S novými médiami sa podľa McQuaila spájajú nasledovné javy – „znížené bariéry pri publikácii informácií[20], individualizácia a mobilita publika, vysoká diverzita na strane ponuky (obsahu), interaktivita médií, menšia (obmedzená) kontrola a regulácia,...,zvýšená globalizácia".[21]

---

19 REMIŠOVÁ, A.: *Etika médií*. Bratislava: Kalligram, 2010, s. 19.

20 Parafrázované. Orig.: „Menší (nebo žádné) překážky v přístupu do veřejného prostoru."

21 MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, s. 185. Pozn.: McQuail spomína aj ďalšie, ktoré sme v našom texte vynechali, ide o „obchádzanie kedysi dominantných médií" a „multimedialitu" – v texte vyššie ich neuvádzame, nakoľko máme za to, že ide až o druhotný jav spojený s príchodom nových technológií, viac než nových médií a zároveň, že v princípe ide o totožný jav. Teda domnievame sa, že k obchádzaniu printových médií dochádza vďaka zvýšenej miere využívania smartfónov a počítačov, nie ako sprievodný jav zvyšovania popularity online žurnalistiky. Inými slovami – nekupujeme si smartfóny, ani laptopy aby sme si mohli prečítať správy prostredníctvom týchto zariadení. Ak ich máme –

Na základe prítomnosti týchto javov môžeme považovať uvedené za pojmové znaky nových médií a na základe nich potom jasne zadefinovať, že ide o „nové médiá“. Pod novými médiami teda rozumieme internet a sociálne siete, ktoré ako jediné vieme spojiť s týmito spomínanými javmi. Tradičné médiá potom, za pomoci spojenia známych kritérií budú tie médiá, ktoré vznikli pred érou digitalizácie, fungujú prevažne v analógovom prostredí, vyznačujú sa centralizovanou kontrolou obsahu (prostredníctvom redakčnej činnosti), jednosmerným modelom komunikácie (od redakcie k publiku, teda adresátom) a teda aj obmedzenou spätnou väzbou od publika.

Sekundárne potom možno pozorovať aj ďalšie javy, ktoré so sebou prináša prechod od tradičných médií k novým médiám. Hoci digitalizácia médií so sebou istotne nesie aj množstvo pozitívnych javov, vymenujeme na tomto mieste, v záujme témy príspevku prednostne tie javy, ktoré sú považované za problematické. Obzvlášť pri pohľade na krízové alebo mimoriadne situácie, ktorým musí spoločnosť čeliť sa stávajú tieto javy viditeľnejšie. Medzi takéto javy patrí napríklad nárast konšpiračného myslenia v spoločnosti, zvýšená miera šírenia dezinformácií, polarizácia spoločnosti a napokon nárast preferencií populistických politických strán a znižovanie dôvery v inštitúcie.[22] Slovenská republika nie je ojedinelou krajinou, ktorá sa potýka s týmito javmi.[23] Na základe uvedeného by sa mohlo zdať, že nová éra médií je skôr súmrakom slobody prejavu. Na tomto mieste nášho príspevku si ale už dovolíme vysloviť domnienku, že možno práve naopak. Otázne je, akým spôsobom dokážeme v tejto oblasti využiť nové technológie, ako napríklad nástroje AI. Tým sa budeme venovať v nasledujúcej časti.

využívame ich na to, no nepredpokladáme, že je to primárny dôvod kúpy týchto zariadení a teda, že tu je možné vidieť akúsi významnejšiu súvislosť so zmenami mediálneho prostredia.

22 ŠKRIPEK, M., et al.: *Čo nás stoja dezinformácie? Hlavní aktéri, strategická komunikácia a ekonomické aspekty dezinformácií.* Bratislava: Výskumné centrum Slovenskej spoločnosti pre zahraničnú politiku (RC SFPA), Friedrich Ebert Stiftung, 2022, s. 5–6 a ďalšie. Online: https://library.fes.de/pdf-files/bueros/slowakei/20643.pdf.

23 COUNCIL OF EUROPE: *Europe Press Freedom Report 2024: Confronting Political Pressure, Disinformation, and the Erosion of Media Independence. Annual assessment of press freedom in Europe by the partner organisations of the Safety of Journalists Platform*. 2025, s. 16 a 31. Online: https://rm.coe.int/prems-013425-gbr-2519-annual-report-2025-correction-cartooning/1680b48f7b.

17.1.3 Nové technológie v informačnom priestore

V tomto bode sa môže zdať potrebné zadefinovať ešte, v súčasnosti už takmer triviálny pojem, „AI“[24]. Ak je tento pojem, ako píšeme, na hranici triviálnosti, vyžaduje si skutočne bližšie rozpracovanie? V súčasnosti už je nutné pojem „AI“ považovať za tak všeobecný, ako povedzme pojem „telefón“ v čase nástupu smartfónov. Iným spôsobom povedané – pojem „AI“ nedokáže úplne uspokojivo vyjadriť špecifiká denotátu. V bežnej reči a pri pedagogickej praxi nepredstavuje tento fakt problém, avšak v odborných a vedeckých publikáciách je nedostačujúce, ak sa nepristupuje k hlbšej špecifikácii toho, čo má autor na mysli, keď hovorí o „AI“. Keď sa teda na tomto mieste pýtame: „Čo teda máme rozumieť pod pojmom „AI“?“ zároveň sa vždy pýtajme: „Ktorá/aká „AI“ je pre nás relevantná v danom kontexte. Čo nás zaujíma v oblasti slobody prejavu a žurnalistiky? Je pochopiteľné, že pre žurnalistiku budú s veľkou pravdepodobnosťou relevantné takzvané LLM, teda „veľké jazykové modely“[25], schopné analyzovať a spracúvať text, s čím ďalej, tým väčšou presnosťou a podobnosťou ľudskej reči, avšak využitie iných foriem AI v oblasti žurnalistiky, v budúcnosti nie je vylúčené.

Ak otázka znie, „Čo je AI?“ jednoduchá odpoveď bude znieť, že ide o *„počítačový systém, ktorý disponuje podobnými kvalitami ako ľudský mozog“* širšie potom „ disponuje schopnosťou interpretovať a produkovať reč spôsobom, ktorý sa zdá byť ľudským, dokáže rozoznať alebo vytvoriť vizuálne diela, riešiť problémy a učiť sa z dát ktoré sú mu ponúknuté“.[26] Takto znie lexikálna definícia pojmu AI. Po kritickej úprave takejto definície by sme mohli dospieť k takejto jej skrátenej verzii: počítačový systém, ktorý dokáže analyzovať, schématizovať a kategorizovať poskytnuté dáta a následne sa dokáže z týchto dát učiť a následne riešiť problémy[27].

---

[24] V príspevku sa využíva pojem „AI“ a „umelá inteligencia“ ako synonymá, nakoľko ide o skratku z anglického „artificial intelligence“, čo sa zaužíva prekladať ako „umelá inteligencia“.

[25] Z anglického „large language models“; Porov.: CAMBRIDGE DICTIONARY: *LLM.* In: Cambridge Dictionary. Online: https://dictionary.cambridge.org/dictionary/english/llm.

[26] „artificial intelligence : the use or study of computer systems or machines that have some of the qualities that the human brain has, such as the ability to interpret and produce language in a way that seems human, recognize or create images, solve problems, and learn from data supplied to them“; Porov.: CAMBRIDGE DICTIONARY: *Artificial intelligence*. In: Cambridge Dictionary. Online: https://dictionary.cambridge.org/dictionary/english/artificial-intelligence.

[27] V tomto bode máme na mysli požiadavky používateľa (na riešenie), *ergo* tzv. „prompty“.

V praxi však takáto definícia stále nevystihuje špecifiká rôznorodých systémov umelej inteligencie. V tomto ohľade potom pojem bude môcť nabrať rôzne tvary podľa kritérií, ktoré zvolíme na rozdelenie. Pokiaľ pre nás bude kritériom priehľadnosť procesu riešenia – budeme rozlišovať tzv. „white-box trénované", „grey-box trénované" a „black-box trénované" AI.[28] Pokiaľ kritériom bude ľudská intervencia, či ľudský faktor pri procese trénovania – budeme rozlišovať „učenie s učiteľom"[29], „učenie bez učiteľa"[30], „učenie sa zo spätnej väzby"[31] a tzv. „učenie s čiastočným zapojením učiteľa"[32].[33]

V oblasti umelej inteligencie sa zmnožujú nové pojmy, či už je to v sfére vývoja, v podobe pojmov ako „strojové učenie", „hĺbkové učenie", „neurálne siete", alebo v sfére užívania, prejavujúc sa v pojmoch ako „veľké jazykové modely" či „generatívny model" alebo „diskriminačný model". Pre účely tohto príspevku nie je praktické a ani žiadúce, aby sme sa pokúšali všetky vysvetliť, to by si mala klásť za účel obšírnejšia, v súčasnosti vznikajúca, odborná literatúra. Postačí ak sa zameriame na už spomínanú otázku – čo bude znamenať AI pre zmeny v oblasti žurnalistiky?

Pokiaľ prihliadame na to, ako je v súčasnosti regulovaná oblasť AI – najrelevantnejšie sa javí Nariadenie Európskeho parlamentu, známe taktiež ako AI Act. Toto nariadenie sa vo svojom texte stavia rezervovane k definovaniu či kategorizácii typov AI. Vyslovene špecifikuje tieto kategórie ako „model/systém AI na všeobecné účely"[34] alebo „model/systém AI na všeobecné účely so

---

28 BOOS, E., MÄLZER, M., CONRAD, F. et al.: *An Engineer-Friendly Terminology of White, Black and Grey-Box Models.* In: MODELSWARD 2025: 13th International Conference on Model-Based Software and Systems Engineering. Proceedings, 2025, s. 313–320. Online: https://www.researchgate.net/publication/389542353_An_Engineer-Friendly_Terminology_of_White_Black_and_Grey-Box_Models; Zaujímavý je však aj prehľad z pred 13 rokov. - Porov.: KHAN, M. E., KHAN, F.: *A Comparative Study of White Box, Black Box and Grey Box Testing Techniques*. In: International Journal of Advanced Computer Science and Applications, roč. 3, č. 6, 2012, s. 12–15. Online: https://www.researchgate.net/publication/270554162_A_Comparative_Study_of_White_Box_Black_Box_and_Grey_Box_Testing_Techniques.

29 Voľný preklad z anglického: „Suprevised learning";

30 Voľný preklad z anglického: „Unsuprevised learning";

31 Voľný preklad z anglického: „Reinforcement learning";

32 Voľný preklad z anglického: „Semi-suprevised learning";

33 SAP: *What is machine learning*? Online: https://www.sap.com/products/artificial-intelligence/what-is-machine-learning.html.

34 Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 zo dňa 13. júna 2024 ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828, OJ L, 12.7.2024, článok 3, bod 63.

systémovým rizikom“[35]. Ak by sme sa však v tejto práci rozhodli aplikovať len takéto delenie neposlúžilo by to jej účelu, nakoľko ide len o delenie na „nebezpečné“ a negáciu tohto atribútu. Pre prax a vývoj AI sa môže zdať toto delenie vhodné, pre vedecké/odborné skúmanie, sa javí nevhodný, či prinajmenšom obmedzujúci.

Zároveň je vhodné spomenúť Rámcový dohovor Rady Európy o umelej inteligencii a ľudských právach, demokracii a právnom štáte, ktorého signatárom je Európska únia a ako jej súčasť aj Slovensko. Tento dohovor však ešte nebol ratifikovaný žiadnym zo signatárskych štátov, preto opomenieme bližší rozbor jeho obsahu. Je však nutné poznamenať, že ide o prvú medzinárodne právne záväznú zmluvu v oblasti AI.[36] Na vnútroštátnej úrovni nemožno hovoriť o zákonnej regulácii. Tak je akákoľvek regulácia tvorená buď zhora (EÚ/Rada Európy) alebo zospodu – vnútorné poriadky inštitúcií, združení či spoločností.

S príchodom AI sa v rôznych odvetviach spájajú rôzne očakávania. Tieto majú pozitívny aj negatívny charakter. Pokiaľ ide o žurnalistiku, budeme vychádzať z už publikovanej štúdie na túto tému, od autorského kolektívu Dodds, Zamith, Lewis.[37] Ako uvádza Dodds, prínosy, ktorý môže prinášať pre žurnalistiku AI sú nasledovné: lepšia dostupnosť informačného obsahu, ľahšie preklady, prispôsobenie obsahu, lacnejšia produkcia a v neposlednom rade, viac priestoru pre žurnalistov na angažovanie sa v občianskej spoločnosti či lokálnej politike.[38] Spomína sa však zároveň aj niekoľko hrozieb, či negatív, ktoré AI pre žurnalistiku môže v budúcnosti predstavovať. Predovšetkým sa tým myslia tieto: masová produkcia dezinformácií a odpadového, marketingu prispôsobeného obsahu; kríza autenticity – teda strata povedomia o tom, čo je tvorené človekom, čo strojom a čomu sa dá veriť; posilnenie filtrovania obsahu, v súvislosti s predošlými dvoma bodmi; následné „vyčerpanie“ z novín a strata záujmu o žurnalistickú tvorbu zo strany adresátov a následný pokles záujmu o verejne dianie.[39] Otázkou, ktorú sme si teda v súvislosti s touto problematikou položili už skôr a ktorú teraz opäť musíme klásť je – čo chceme pre žurnalistiku do budúcnosti? Ako dokážeme zabezpečiť, že systémy AI budú napomáhať

---

35 *Tamže*, článok 51; Príloha XIII.

36 Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Strasbourg, 5 September 2024, Council of Europe Treaty Series – No. 225.

37 V texte ďalej uvádzame len „Dodds“.

38 DODDS, T., ZAMITH, R., LEWIS, S. C.: *The AI turn in journalism: Disruption, adaptation, and democratic futures*. In: Journalism, Vol. 27, No. 4, 2025, s. 7. Online: https://www.researchgate.net/publication/391770691_The_AI_turn_in_journalism_Disruption_adaptation_and_democratic_futures

39 *Tamže*, s. 8.

žiadúcemu rozvoju žurnalistiky v podobe nových médií? Ak odpoveď znie, že regulácia – odkiaľ má prísť a v akej forme?

17.2 K diskusii o potrebe regulácie informačného prostredia

Timothy Snyder, v knihe *O tyranii: dvadsať ponaučení z 20. storočia,* formuluje ťažiskové odporúčania a varovania, ktoré majú napomôcť tomu, aby sa predišlo opakovaniu totalitných omylov minulého storočia. Napriek tomu, že všetky možno považovať za relevantné, pre náš príspevok uvedieme ilustratívne najmä dve, konkrétne tieto: ponaučenie č. 2 – „*Bráňte inštitúcie*" a č.5 – „*Nezabúdajte na profesionálnu etiku*".[40] Týmito dvoma direktívami by sme sa chceli riadiť v nasledujúcej časti. Prečo? Brániť inštitúcie síce Snyder požaduje pred totalitnými silami v štáte, [41] avšak nám sa javí, že pokiaľ sa rozhodneme brániť nejakú inštitúciu, znamená to starať sa o ňu, jej rozvoj, o jej dobro. A profesionálna etika? O nej hovorí Snyder že bola tým, čo mohlo spomaliť, ak nie efektívne brániť nacistickému Nemecku pri Holokauste.[42] Píše: „Ak nám povedia, že situácia je výnimočná, musíme sa riadiť prísne profesnou etikou. Potom neexistuje nič ako „len som plnil rozkazy"."[43] My si dovolíme tvrdiť, že profesionálna etika je preto kľúčovým rozmerom, nielen v bránení extrému, ale v zabezpečovaní správneho chodu akéhokoľvek systému. Vytvára požiadavku na kvalitu.

Revolúcia mediálneho priestoru, ktorá ako sme si uviedli skôr, bude s príchodom AI stále výraznejšia si vyžaduje dôslednú ochranu demokratických princípov. Pri výrazných spoločenských zmenách môžu často trpieť práve ľudské práva,[44] ktoré považujeme v súčasnosti za materiálnu súčasť demokracie. Zachovanie slobody prejavu, no zároveň práva na pravdivé informácie, v algoritmom definovanom online-priestore, si bude podľa nášho názoru vyžadovať ochranu inštitúcie - žurnalistiky.

Hoci sú snahy dosiahnuť spomínané pomocou legislatívy, želaný efekt týmto spôsobom nemôže byť dosiahnutý, nakoľko žurnalistická obec je, pochopiteľne, veľmi citlivá na akékoľvek vonkajšie regulácie slobody prejavu. Na Slovensku nami skúmanú oblasť upravuje zákon o mediálnych službách a zákon

---

40 SNYDER, T.: *O tyranii: dvadsať ponaučení z 20. storočia*. Bratislava: Premedia, 2017, s. 6.

41 *Tamže*, s. 19–21.

42 *Tamže*, s. 32–34.

43 *Tamže*, s. 34.

44 Príkladom možno uviesť najmä industriálnu revolúciu. Táto je s takzvanou „AI revolúciou" častokrát prirovnávaná. Deje sa tak nakoľko ich sprevádzajú podobné obavy – zo straty pracovných miest, zmeny spoločenských štruktúr a tak podobne.

o publikáciách. Z hľadiska európskej úpravy, nesmieme zabudnúť známy Akt o digitálnych službách. Žiadny z nich sa však nevenuje etickému hľadisku povolania novinára a ani tento pojem samotný sa v nich nevyskytuje. Je možné domnievať sa, že ak by niektorý z týchto zákonov tento pojem používal, bolo by ho nutné definovať a tomu sa zákonodarca evidentne vyhýba.[45]

Ako riešenie sa ponúka preferencia profesionálnych etických kódexov ako nástroja stanovenia požiadaviek na riadny výkon profesie. Upravovať oblasť AI formou etických pravidiel navrhujú aj niektorí odborníci, dokonca aj v našom právnom prostredí.[46] Uvedené chápe aj Európsky parlament, keď v samotnom texte DSA, čl.45, ods.1 vyjadruje podporu tvorbe „Kódexov správania".[47] Požiadavka na etickú kontrolu vykonávania žurnalistickej profesie je teda legitímna, na čom sa už dlhšie zhoduje aj veľká časť odbornej verejnosti z oblasti žurnalistiky.[48] To sa potvrdzuje skutočnosťou, že azda najväčšia profesijná organizácia novinárov, *Slovenský syndikát novinárov,*[49] prijala svoj Etický kódex novinára (ďalej len „Etický kódex"). Otázkou je, či Etický kódex, v stave v akom je, dokáže efektívne zohľadniť aj nové hrozby a možnosti, ktoré ako spomíname vyššie, so sebou prináša AI. Tomu sa budeme venovať v prvej podkapitole.

V oblasti žurnalistiky prirodzene nachádzame implicitné očakávania na napĺňanie istých spoločnosťou uznávaných hodnôt. Ide o spoločenskú požiadavku vyplývajúcu z charakteru žurnalistiky samotnej. Tieto očakávania môžu mať rôznu formu a ich bližšiu špecifikáciu rozpracúva odborná literatúra. Napríklad McQuail uvádza ako najvšeobecnejšie kategórie tieto: „pravdivosť, slobodu, solidaritu, rovnosť, poriadok a súdržnosť".[50] Podľa nášho názoru je možné tvrdiť nasledovné: Napĺňanie týchto hodnôt je merným

---

45 Niektorí iste povedia – „Našťastie!"

46 KLUČKA, J.: *Úloha a dôležitosť etických pravidiel v systémoch umelej inteligencie*. In: Právny obzor, Roč. 108, č.3, 2025, s. 240–254. Online: https://doi.org/10.31577/pravnyobzor.2025.3.02.

47 Nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 zo dňa 19. októbra 2022 o jednotnom trhu s digitálnymi službami a o zmene smernice 2000/31/ES (akt o digitálnych službách; tiež DSA), L 277/1, 27.10.2022, článok 45 ods.1: „S cieľom prispieť k riadnemu uplatňovaniu tohto nariadenia Komisia a výbor podporujú a uľahčujú vypracovanie dobrovoľných kódexov správania na úrovni Únie, pričom zohľadňujú najmä osobitné výzvy spojené s bojom proti rôznym druhom nezákonného obsahu a riešením systémových rizík podľa práva Únie, najmä s právnymi predpismi pre hospodársku súťaž a ochranu osobných údajov.";

48 REMIŠOVÁ, A.: *Etika médií*. Bratislava: Kalligram, 2010, s. 104.

49 *Tamže*, s. 90.

50 MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, s. 72–73. Pozn.: Autor si dovoľuje poznamenať, že jednotlivé hodnoty sa môžu do istej miery prekrývať, niektoré z vymenovaných až po hranicu totožnosti. Napríklad medzi hodnotou „solidarity" a „súdržnosti" sa zdá byť rozdielnosť až zanedbateľná.

faktorom kvality. V druhej podkapitole sa preto budeme venovať demonštrácii toho, akým spôsobom môže byť otázka napĺňania týchto hodnôt (konkrétne hodnoty pravdivosti) problematická s príchodom AI. Zameriame sa pri tom na rozporuplné úvahy o práve na pravdivé informácie zo strany Ústavného súdu Slovenskej republiky a Európskeho súdu pre ľudské práva. V druhej kapitole priebežne ponúkame úvahy *de lege ferenda*, ako formu apelu na odbornú verejnosť.

V závere uvedieme zistenia ku ktorým sme dospeli pri vypracovaní príspevku, s prihliadnutím na tézy, východiská a ciele, ktoré sme si vytýčili v úvode nášho príspevku.

Než prejdeme k materiálnej časti tejto kapitoly, musíme rozpracovať otázku, ktorú sme vystavali v predošlej kapitole. Táto otázka znie: „Čo chceme od žurnalistiky?“. Odpoveď sa môže líšiť, no po eliminácii zjednodušení a zanedbaní niektorých ekonomických aspektov[51] žurnalistického prostredia bude možné zúžiť výber na binárnu úroveň: kvalita *versus* kvantita.

Tieto dve odpovede sa nutne nemusia vylučovať. Je však dôležité poznať naše preferencie, pokiaľ tvoríme víziu žurnalistiky do budúcnosti. Podľa nás je v súčasnosti vhodné zamerať sa na kvalitatívny rozmer a sústrediť sa na novú rolu, ktorú žurnalistika môže v spoločnosti mať. Máme tým na mysli, ako píše Dodds: *„journalism's less-well-understood role (yet one that is more pressing than ever) as a media-and-technology educator for society.“.*[52]

---

[51] Je pochopiteľné, že na našu otázku je možné odpovedať aj z hľadiska toho, že pokiaľ chceme zachovať inštitúcie, musíme zabezpečiť ich ekonomickú udržateľnosť. Sme však toho názoru, že v racionálnej spoločnosti si médiá, ktoré ponúkajú to čo spoločnosť žiada na seba zarobia. V našej polemike nebudeme namietať, že rozmery ako: aktuálnosť, prominentné osobnosti, šokujúce udalosti, dramatické príbehy – sú neoddeliteľnými súčasťami žurnalistickej tvorby. Ale keďže náš príspevok čiastočne adresujem celej žurnalistickej obci, považujeme za irelevantné riešiť túto otázku. Ekonomické hľadisko si bude sledovať každé médium osobitne, avšak etické pravidlá si stanovuje profesia a to bez ohľadu na príslušnosť ku konkrétnemu médiu. Týmto tvrdením zároveň nevylučujeme skutočnosť, že aj redakcie môžu prijať osobitné etické kódexy, pokiaľ neodporujú všeobecnému Etickému kódexu novinára. Dokonalým potvrdením našich tvrdení bude kariérny vývoj Michala Kovačiča v uplynulých rokoch, keď po vyjadrení morálneho postoja síce ako jednotlivec musel opustiť významnú pozíciu moderátora v televízii Markíza, avšak pomocou crowd-fundingovej kampane bol schopný vyzbierať viac ako pol milióna na fungovanie vlastného projektu, online spravodajského kanálu 360°.

[52] DODDS, T., ZAMITH, R., LEWIS, S. C.: *The AI turn in journalism: Disruption, adaptation, and democratic futures*. In: Journalism, Vol. 27, No. 4, 2025, s. 11. Poznámka (voľný preklad autora): „menej známu rolu žurnalistiky, (hoci pálčivejšiu než kedykoľvek predtým) ako učiteľa v oblasti médií a technológií.“

Média sa budú musieť časom čoraz častejšie snažiť preraziť epistemické bubliny v ktorých sa uzatvárame vďaka filtrácii obsahu zo strany sociálnych sietí a tak sa vynára požiadavka, aby nezáväzne bola táto oblasť regulovaná skôr – ideálne aby sa so zmenami začalo už teraz. To môže byť smerodajné pri úvodných fázach, kedy ešte nedošlo k stavom, v ktorých je nutné oblasť regulovať záväzne a prostredníctvom právnych noriem, čo je nežiadúce, zo strany žurnalistov (kvôli hrozbe zneužitia zo strany mocenských štruktúr, pod zámienkou ochrany právneho štátu)[53] a kedy je už nemožné kontrolovať technológie kvôli ich masovej popularite a rýchlemu šíreniu.[54] Zachytenie tohto *momenta* môže byť kľúčovým krokom pri zmenách v mnohých profesiách, nie len v žurnalistike. Je zároveň na mieste pýtať sa, či sme takýto zlomový bod už nezmeškali. Optimisticky povieme, že zatiaľ nie. Pesimisticky a možno so štipkou realistického pohľadu, od toho nemáme ďaleko.

17.2.1 Etický kódex novinára

Otázkou ako sme spomenuli vyššie pre nás v tejto časti je, akým spôsobom Etický kódex novinára upravuje problematiku využívania AI v žurnalistike. Po zbežnej analýze pomocou vyhľadávania kľúčových výrazov ako „AI", „UI" či „umelá inteligencia" zistíme, že v texte sa nenachádza explicitná úprava problematiky ktorou sa v našom príspevku zaoberáme. To môže už samo o sebe byť problémom. Dôležité je však zamerať sa aj na implikované usmernenia, ktoré Etický kódex ponúka. Naša diskusia sa preto v tejto časti bude uberať dvoma smermi. Najprv analyzujeme problémy, ktoré môžu nastať pri rôznej interpretácii implicitných etických noriem v texte Etického kódexu a následne budeme rozoberať problematiku absencie tých explicitných.

V Etickom kódexe nájdeme v zásade tri pasáže, od ktorých by sme mohli vyvodzovať etické normy pre novinára pracujúceho s AI: Čl. 2, bod 2 a 4; Čl.3 bod 1,2,3 a 5; Čl. 7 bod 1 a 3; V prvom z uvedených článkov ide o základné princípy práce novinára. V druhom zas ide o vzťah novinára a verejnosti. V tretej časti zas o vzťah novinára a jeho kolegov. Kvôli rozsahu tohto príspevku nebudeme rozoberať text jednotlivých článkov z každého použiteľného interpretačného rámca, nakoľko v záujme problematizácie postačuje, ak viacero subjektov si môže niektorú normu vyložiť odlišne. Už len implicitná povaha posudzovaných noriem predpokladá, že tento problém nastane a z uvedeného dôvodu možno

---

53 THOMSON REUTERS FOUNDATION: *Weaponizing the Law: Attacks on Media Freedom 2023*, s. 13. Online: https://www.trust.org/wp-content/uploads/legacy/weaponizing-law-attacks-media-freedom-report-2023.pdf.

54 Tu máme na mysli známy koncept – tzv. „Collingridge dilemma".

konštatovať, že je v prvom rade problematické, že Etický kódex sa touto problematikou explicitne nezaoberá.

V Čl.2 sa dozvedáme, že: „Hlavnými zásadami, ktorými sa novinár riadi vo svojej práci, sú nestrannosť, vyváženosť, objektivita, čestnosť, pravdivosť, zodpovednosť a dôsledné overovanie faktov.“[55] Zároveň sa dozvedáme aké sú pravidlá preberania obsahu pre novinára publikujúceho na internete. Tento článok vo svojej všeobecnosti má najväčší potenciál zahrnúť do svojho obsahu aj problematiku AI, avšak práve táto všeobecnosť nám bráni akceptovať tento článok ako efektívny nástroj etickej regulácie. Spomeňme hodnoty objektivity, pravdivosti, čestnosti a dôsledného overovania faktov – pri všetkých spomínaných je možné od AI požadovať nejakú mieru reflektovania týchto hodnôt, ale je nepravdepodobné, že morálne zdanlivo ambivalentná AI bude schopná reflektovať ich pri svojich výstupoch. Všetka zodpovednosť za napĺňanie týchto hodnôt je na osobe novinára.

Problémom je, pokiaľ si uvedomíme, že už v súčasnosti môže množstvo novinárov využívať AI pomerne nekriticky, bez toho aby bolo na prvý pohľad evidentné, že napríklad publikovaný článok nenapĺňa tieto hodnoty, respektíve – že môže byť kompletne vytvorený AI. Dovolíme si predpokladať, že žurnalistická obec časom akceptuje pokojne aj úplné využitie AI pri tvorbe informačného obsahu ako eticky neproblematické, dôraz sa však bude klásť na overovanie faktov.[56] Do budúcnosti bude preto podľa nás potrebné vytvárať nástroje AI schopné spolupracovať so žurnalistami pri ich práci a zároveň schopné dodržiavať vyššie spomínané hodnoty. To však samo o sebe nebude postačovať, nakoľko stále absentuje vyjadrenie, či je samotné použitie AI pri žurnalistickej práci etické.

Čl. 3 sa venuje zdrojovej stránke informácií poskytovaných novinármi, vo vzťahu k verejnosti. Dôraz sa kladie na overovanie informácií aspoň z dvoch nezávislých zdrojov, na zdieľanie zdroja informácií či na neupravovanie textového, obrazového a zvukového zdrojového materiálu. Osobitne zaujímavý je pre nás však bod 3, ktorý znie: „Novinár alebo médium nezverejňuje žiadne informácie, o ktorých vie, že sú nepravdivé. Pravdivé informácie vychádzajú z faktov a ich objektívnej interpretácie, s prihliadnutím na skutočný kontext popisovaného deja, bez skresľovania objektívnej reality a zamlčovania (opomenutia) časti existujúcich faktov a súvislostí.“ Umelá inteligencia môže v budúcnosti byť veľmi efektívnym nástrojom na zabezpečenie zdieľania

---

55 SLOVENSKÝ SYNDIKÁT NOVINÁROV: *Etický kódex novinára*, článok 2, bod 2. Online: https://www.ssn.sk/eticky-kodex-novinara/.

56 V odbornej literatúre sa často používa výraz z anglického jazyka „*fact-checking*“.

informácii formou „objektívnej interpretácie“ a „bez skresľovania objektívnej reality“, ale je dôležité, aby ju jej vývojári vedeli týmto schopnostiam naučiť. Ako je možné toto realizovať, ak ani naším súdom nie je úplne jasné, čo je možné považovať za naozaj objektívne a čo nie? Tomuto problému sa budeme viac venovať v druhej podkapitole.

Autorstvo je problémom, ktorý reguluje čl. 7 v bode 1. V rozsahu tohto bodu by mohlo byť najproblematickejšie považovať za etické využitie AI, nakoľko pokiaľ novinár nesmie vydávať cudzí text za vlastný – za čí treba považovať text vytvorený AI? A podľa čoho možno rozlišovať pomer autorstva na článku, ktorý vypracuje AI podľa dôsledného „promptu“ novinára? Odpovede na túto problematiku a na problematiku práva duševného vlastníctva pri výsledkoch práce za použitia AI rozpracúva aj súčasná odborná spisba.[57] Hoci základné princípy už sa v tejto oblasti začalo formovať, všetko stojí v závere na morálnej integrite novinára – či ak vydáva článok vytvorený AI za svoj, túto skutočnosť prizná. Vzhľadom na tento rozmer môže byť potom problematické vyvodzovanie morálnej zodpovednosti. Etické rady nemajú nástroj, ktorým by bolo možné autorstvo overiť.

Zároveň tu v čl. 7 nájdeme bod č. 3, ktorý káže novinárovi vytvárať nátlak na spolupracujúcich novinárov, aby nekonali neeticky a protiprávne. O dosahoch tohto bodu pri využití AI sa domnievame, že vo svete, v ktorom je umelá inteligencia dostupná takmer každému a skoro kedykoľvek a kdekoľvek, je toto obsolentná požiadavka. Ak AI má k dispozícii každý, je pravdepodobnejšie, že ju každý ticho využije, než aby na využitie zo strany kolegov upozorňoval. Aj z tohto dôvodu považujeme za pravdepodobné, že využitie AI v novinárskej profesii bude do budúcna možno považovať za normu.

Ako sme uviedli vyššie – najproblematickejšie sa nám javí absencia explicitnej úpravy AI v Etickom kódexe. Táto absencia spôsobuje nerovnováhu, v ktorej autori, ktorí sa Etickým kódexom neriadia vôbec, budú AI využívať bez okolkov a autori, ktorí sa ním riadia sú ponechaní bez vysloveného súhlasu, či jasnej a jednoznačnej interpretácie. Sme preto toho názoru, že ak má byť do budúcnosti Etický kódex efektívnym nástrojom, ktorý bude novinármi akceptovateľný a využiteľný, potom sa musí prispôsobiť dobe. Nakoniec, je to práve jedna z výhod etických kódexov, že si nevyžadujú celospoločenskú zhodu, prostredníctvom parlamentných zástupcov, na to aby sa zmenili. Postačuje zhoda v rámci toho spoločenstva, ktoré si tento kódex prijalo. V súvislosti

---

57 FRITZ, J.: *Understanding authorship in artificial intelligence-assisted works*. In: Journal of Intellectual Property Law & Practice, roč. 20, č. 5, 2025, s. 354–364. Online: https://academic.oup.com/jiplp/article/20/5/354/7965768.

s týmto tvrdením si dovoľujeme vysloviť predpoklad, že ak má nestať efektívna zmena v oblasti, v ktorej dochádza k signifikantnému technologickému pokroku, vyžaduje sa aktívna účasť a zapojenie aj na úrovni nižšej než je parlamentná. Vyžaduje sa dôkladná komunikácia a spolupráca medzi žurnalistami[58] a vývojármi AI. Pozrime sa teraz na jednu z problematických oblastí v ktorých sa využitie AI v žurnalistike môže stretnúť s absenciu jasnej odpovede zo strany súdov.

17.2.2 Nejasnosti v kontexte s právom na pravdivé informácie

V tejto časti sa zameriame špeciálne na prípad Európskeho súdu pre ľudské práva, Hertel proti Švajčiarsku.[59] V stručnosti prípad sumarizujeme a problematizujeme ho v kontexte AI.

Pán Hertel, (ďalej len „pán H“) bol nezávislým výskumníkom[60] v oblasti technológii. Z hľadiska svojej osobnostnej kvalifikácii išlo o osobu s titulom „Dr.“ z oblasti technických vied, z Federálneho technického inštitútu v Zürichu, ktorý však už bol na dôchodku a svoj čas na dôchodku si krátil skúmaním vplyvov technológií na životné prostredie a ľudský organizmus. Svoje výskumy realizoval vo vlastnom amatérskom laboratóriu. Pán H sa v jednom zo svojich výskumov rozhodol zaoberať vplyvom mikrovlnných rúr na ľudský organizmus. Šokujúce zistenia prameniace zo spomínaného výskumu sa rozhodol zaslať na publikáciu do environmentalisticky ladeného, no v podstate tematicky veľmi rozmanitého časopisu - Franz Webber, v ktorom bol článok napokon aj skutočne publikovaný, v 19. čísle tohto časopisu. Tento článok bol titulným a centrálnym článkom tohto časopisu, ktorý sa predával v dvoch jazykových verziách francúzskej a nemeckej.

V reakcii na publikáciu spomínaného výskumu sa však ohradilo niekoľko vedcov a prebehla diskusia na tému faktickej problematickosti tohto výskumu,[61] v ktorej vysvitlo, že zistenia ktoré boli uvedené v článku pána H boli nepravdivé a nesprávne, postupy neštandardné a *de facto* odporovali zaužívaným vedeckým faktom – totiž, že samotné používanie mikrovlnnej rúry nespôsobuje rakovinu. V reakcii na článok i kritiku pána H, zo strany odbornej verejnosti, sa svojich práv začalo domáhať Švajčiarske združenie predajcov elektrospotrebičov. Evidentným dôvodom, sformulovaným aj v ich žalobe proti pánovi H, pre ktoré je potrebné, aby pán H nesmel pokračovať v šírení svojich „nezmyslov“, bola

---

58 Ako, nepochybne, aj inými profesiami.

59 Rozhodnutie ESĽP vo veci *Hertel proti Švajčiarsku*, sťažnosť č. 59/1997/843/1049, rozsudok zo dňa 25. augusta 1998. Online: https://hudoc.echr.coe.int/eng?i=001-59366.

60 Rozumej, nepôsobil na žiadnej inštitúcii.

61 Podobne ako na Slovensku v prípade vládneho splnomocnenca Petra Kotlára.

potreba zabránenia nekalej súťaži a celkovo ochrana hospodárskej súťaže. Predajcovia vnímali, že by článok mohol poškodiť ich hospodárske záujmy, obzvlášť aj preto, že v časopise Franz Webber bol článok zverejnený s obrázkami smrtky s kosou, držiacej mikrovlnnú rúru. Obrázok smrtky s mikrovlnnou rúrou sa dostal aj na titulnú stranu a v princípe aj na cca. osem nasledujúcich strán, teda na každú stranu, na ktorej sa nachádzal článok pána H. V tomto ohľade však pán H uviedol, že o obrazovej prílohe svojej štúdie nevedel a dištancuje sa od nej.

Švajčiarske súdy rozhodli v súlade s názorom Združenia švajčiarskych predajcov elektrospotrebičov a zakázali pánovi H tvrdiť isté veci týkajúce sa jeho štúdie. V súvislosti s tým sa pán H rozhodol obrátiť na Federálny súd, ktorý jeho sťažnosť zamietol. Táto skutočnosť aktívne legitimovala pána H na podanie podnetu na ESĽP. Ten následne rozhodol v jeho prospech, teda, že rozhodnutím Švajčiarskych súdov došlo k porušeniu čl. 10, so zdôraznením niektorých skutočností, vyberme si však len tie, ktoré môžu, podľa nás byť v budúcnosti sporné v kontexte AI. Predovšetkým budú problematické kvôli negatívnej odpovedi na otázku, či existuje právo na pravdivé informácie.

„a) Švajčiarske orgány mali takto možnosť voľnej úvahy v rozhodovaní, či tu bola „nevyhnutná

spoločenská potreba“, aby bolo vydané predmetné súdne nariadenie proti sťažovateľovi. Je však potrebné obmedziť rozsah voľnej úvahy, keď to, o čo ide, nie je len „obchodné“ tvrdenie jednotlivca, ale jeho účasť na diskusii ovplyvňuje všeobecný záujem napríklad na verejnom zdraví; v danom prípade nemôžeme poprieť, že takáto diskusia existovala. Týkala sa vplyvov mikrovlnných rúr na ľudské zdravie.

...

b) Nie je podstatné, že jeho názor je v menšine, alebo že sa zdá byť neopodstatneným, pretože v oblasti, v ktorej je nepravdepodobnosť existencie akejkoľvek istoty, by bolo obzvlášť nezmyselné obmedziť slobodu prejavu len na všeobecne akceptované myšlienky.“[62]

Rozpor s vyššie uvedeným vnímame napríklad v tvrdeniach ústavného právnika Jána Drgonca, v jeho komentári k Ústave SR. Ten tvrdí, že: *„lož, klamstvo, je frekventovanou pracovnou metódou politikov aj masmédií, ale to z nej nerobí prejav pod ústavnou ochranou, lebo osobitným záujmom, ktorý podľa čl. 26 ods. 4 dovoľuje obmedziť slobodu prejavu, je aj mravnosť, s ktorou nie je lož*

[62] Rozhodnutie ESĽP vo veci *Hertel proti Švajčiarsku*. Pozn.: Uvedené možno nájsť v rozhodnutí, v slovenskom preklade v časti s najvýznamnejšími právnymi vetami rozhodnutia.

zlučiteľná."[63] V súlade s týmto tvrdením znie rozhodnutie Ústavného súdu SR so spisovou značkou PL. ÚS 12/09, 11. 7. 2012. V ňom tvrdí Ústavný súd nasledovné: *„Ústavný súd v prvom rade nenachádza presvedčivý dôvod, pre ktorý by šírenie nepravdivých informácií malo požívať ústavnú ochranu a pre ktorý by sa teda na obmedzenie šírenia takýchto informácií mali nevyhnutne vzťahovať podmienky ustanovené v čl. 26 ods. 4 ústavy."*[64] Je teda zjavné, že v istom ohľade sa na tomto názore zhoduje súdna autorita s právnou doktrínou.

Jedným problémom je, že Ústavný súd v tom prípade rozhoduje iným spôsobom, než ako argumentuje Európsky súd pre ľudské práva. Tento problém má však nadstavbu. Nami zmienené rozhodnutie je z hľadiska konzistentnosti judikatúry Ústavného súdu tiež sporné, nakoľko práve podobnú argumentáciu ako Európsky súd pre ľudské práva v prípade pána H, využil Ústavný súd o niekoľko rokov neskôr, v rozhodnutí II.ÚS 307/2014. Tam spomína nasledovné: *„zákonné obmedzenie slobody prejavu médií [nemožno odôvodniť] požiadavkou na objektívne a nestranné informovanie v rámci spravodajských programov a politicko-publicistických programov... nie je možné vyvodiť z ústavného práva jednotlivcov na prijímanie informácií (čl. 26 ods. 2 ústavy a čl. 10 ods. 1 dohovoru), keďže právo prijímať informácie zahŕňa právo prijímať informácie tak pravdivé, overiteľné a neutrálne, ako aj fiktívne, neoveriteľné, nepravdivé alebo pravdu skresľujúce či prezentujúce určitý svetonázor."*[65] Je preto evidentné, že v odbornej verejnosti ako ani medzi súdmi neexistuje jednoznačná zhoda na existencii práva na pravdivé informácie.

My si myslíme, že so vzrastajúcim vplyvom AI bude toto právo stále relevantnejšie. Preto očakávame, že v nasledujúcich rokoch bude musieť Európsky súd prehodnotiť svoj názor na túto problematiku. Rozhodnutie obhajujúce právo na pravdivé informácie by bolo spoločensky veľmi prospešné. Napomohlo by v prispôsobení Etického kódexu, ako aj pri vývoji eticky tvoriacej AI. Išlo by zároveň o právo, ktoré by umožnilo zachovať istú mieru kvalitatívnej úrovne publikovaných vedeckých názorov v žurnalistickej oblasti. Obzvlášť dôležité sa to javí v tejto dobe, kedy sme konfrontovaní s problematikou post-faktuality.

V závere tejto časti je ešte nutné spomenúť, že pokiaľ nastanú okolnosti za ktorých sa bude musieť Európsky súd pre ľudské práva vyrovnať s týmto

---

63 DRGONEC, J.: Čl. 26. In: DRGONEC, J.: Ústava Slovenskej republiky. 2. vydanie. Bratislava: C. H. Beck, 2019, s. 671.

64 Nález Ústavného súdu Slovenskej republiky zo dňa 11. 7. 2012, sp. zn. PL. ÚS 12/09; článok 6, bod 2, ods.4.

65 Nález Ústavného súdu Slovenskej republiky zo dňa 18. 12. 2014, sp. zn. II. ÚS 307/2014; článok 61.

problémom, pôjde o zaujímavú a veľmi zásadnú zmenu. Pokiaľ takéto okolnosti nenastanú, bude sa s touto problematikou skôr, alebo neskôr musieť vyrovnať odborná verejnosť sama. Je predovšetkým v ich záujme, aby prijali novú úpravu etických pravidiel skôr, než bude v niektorom problematickom prípade musieť rozhodovať spomínaný súd, nakoľko jeho záväzné kritériá môžu byť v konečnom dôsledku prísnejšie, než by si stanovili stavovské organizácie novinárov sami.

Zároveň sa domnievame, že vzrastajúci záujem o AI bude viesť k potrebe využitia samotných nástrojov AI na vyhodnocovanie kvality jednotlivých žurnalistických výstupov. Ak majú mať nástroje AI vytvorený rámec, v ktorom budú vedieť rozoznávať medzi objektívnymi informáciami, dezinformáciami a halucináciami, potom budeme musieť akceptovať, že sú názory a sú fakty. Podľa nášho názoru, si fakty vyžadujú vyššiu mieru ochrany. Tým máme na mysli napríklad aj využitie AI ako nástroja mäkkej etiky, teda napríklad ako fact-chceckera. Samozrejme, v takom prípade si takýto systém bude stále vyžadovať osobu, ktorá bude zodpovedať za prácu „AI fact-checkera". Tu vidíme jednu z možných metamorfóz žurnalistickej profesie do budúcnosti. Využitie AI v tejto funkcii si však, s prihliadnutím na súčasný stav, bude vyžadovať ešte hlbšiu implementáciu jej jednoduchších uplatnení. Preto považujeme za korektné uviesť tu v závere, že akékoľvek kroky, ktoré odporúčame v našom príspevku majú povahu prevencie a vyžadujú si hlbšie rozpracovanie zodpovednými subjektmi, hoci ako uvádza Dodds, môže ísť už o veľmi súčasný problém a preto sa nám javí, že je potrebné sa o túto problematiku zaujímať.[66]

Záver

V závere sa môžeme vrátiť k východiskám, ktoré sme si stanovili v úvode. Ukázali sme, že ústup tradičných médií – chápaných ako centralizované, redakčne kontrolované a prevažne analógové kanály komunikácie – zásadne mení podobu mediálneho prostredia. Internet a sociálne siete, spojené so zníženými bariérami publikovania, vysokou diverzitou obsahu, interaktivitou a obmedzenou reguláciou, vytvárajú nové možnosti deformácie informačného priestoru. Tieto zmeny sa v spoločnosti prejavujú nárastom konšpiračného myslenia, šírením dezinformácií, polarizáciou a oslabovaním dôvery v inštitúcie, čo potvrdzuje prvú z našich hypotéz. Zároveň však nevyhnutne neznamenajú súmrak slobody prejavu, ale skôr test odolnosti demokratických inštitúcií a kvality žurnalistiky ako „lakmusového papiera" stavu demokracie.

---

66 DODDS, T., ZAMITH, R., LEWIS, S. C.: *The AI turn in journalism: Disruption, adaptation, and democratic futures*. In: Journalism, Vol. 27, No. 4, 2025, s. 9.

Druhá hypotéza – potreba nových právnych a etických mechanizmov, vrátane využitia umelej inteligencie – sa v priebehu príspevku ukázala ako opodstatnená. Na právnej úrovni sme poukázali na to, že existujúce rámce (či už na úrovni EÚ alebo Rady Európy) sa zameriavajú najmä na kategorizáciu rizík a základnú reguláciu systémov AI, no len okrajovo reflektujú špecifiká žurnalistickej profesie. Zároveň sme ukázali, že právna regulácia sama osebe naráža na citlivosť žurnalistickej obce na zásahy do slobody prejavu a na obavy z „regulácie zvonka". Aj preto sa javí ako kľúčové doplniť právne nástroje o vnútroprofesijné mechanizmy – najmä o aktualizované etické kódexy, ktoré budú schopné pružnejšie reagovať na technický vývoj a zároveň ostanú prijateľné pre samotných novinárov.

Analýza Etického kódexu novinára ukázala, že napriek tomu, že v sebe implicitne obsahuje viaceré relevantné hodnoty (pravdivosť, objektivita, nestrannosť, dôsledné overovanie faktov, rešpekt k autorstvu), chýba mu explicitná reflexia využívania AI v žurnalistickej praxi. To vedie k interpretačnej neistote – novinári, ktorí sa kódexom riadia, nemajú jasnú odpoveď na otázku, v akom rozsahu a za akých podmienok je eticky prípustné využívať AI pri tvorbe textu, pri overovaní informácií či pri práci so zdrojmi. V prostredí, kde je AI *de facto* dostupná „každému a vždy", sa zároveň ukazuje ako iluzórne očakávať, že etická kontrola bude fungovať len prostredníctvom vzájomného „dozerania" medzi kolegami. Práve preto sme zdôraznili potrebu jasnejšie formulovať pravidlá týkajúce sa autorstva, transparentnosti využitia AI a zodpovednosti za obsah, ktorý je publikovaný pod menom novinára.

Osobitnú pozornosť sme venovali otázke práva na pravdivé informácie a jeho vzťahu k slobode prejavu v kontexte AI. Na príklade judikatúry Európskeho súdu pre ľudské práva a Ústavného súdu SR sme ukázali, že v európskom právnom priestore neexistuje ustálená zhoda na tom, či – a v akej podobe – možno hovoriť o práve na pravdivé informácie. Kolísanie medzi dôrazom na ochranu pluralitnej diskusie, zahŕňajúcej aj menšinové či nepravdepodobné názory, a tendenciou vylúčiť vedome nepravdivé informácie z ústavnej ochrany, vytvára neistotu, ktorá sa s príchodom systémov AI ešte zvýrazňuje. Domnievame sa, že práve masová schopnosť AI generovať „pravdivo pôsobiace" nepravdy a skreslenia bude do budúcna nútiť právnu aj odbornú verejnosť k precíznejšiemu vymedzeniu vzťahu medzi faktmi, názormi a ochranou pravdy ako hodnoty.

V tomto kontexte sme vyjadrili presvedčenie, že posilnenie konceptu práva na pravdivé informácie – minimálne ako interpretačnej zásady – by mohlo významne prispieť k etickému usmerneniu vývoja a používania AI v žurnalistike. Takéto právo by poskytlo hodnotový rámec jednak pre podobu profesijných

kódexov, jednak pre samotný dizajn AI nástrojov (napríklad fact-checkerov či systémov na odhaľovanie dezinformácií). Zároveň by mohlo fungovať ako protiváha post-faktuálnym tendenciám, v ktorých sa rozdiel medzi názorom a faktom nekontrolovane rozostruje.

Napokon, na základe východísk z úvodu môžeme konštatovať, že budúcnosť žurnalistického prejavu v ére sociálnych sietí a AI nie je vopred daná technológiou samotnou. Nepochybne bude nutné prehodnocovať jej využitie vzhľadom na ekologickú a energetickú záťaž, ktorú môže AI predstavovať, v oblasti žurnalistiky však už AI s najväčšou pravdepodobnosťou ostane. Rozhodujúce bude pre systém to, či sa nám podarí súbežne „brániť inštitúcie" a „nezabudnúť na profesionálnu etiku" – teda posilniť žurnalistiku ako demokratickú inštitúciu a zároveň aktualizovať jej profesijné pravidlá tak, aby boli zrozumiteľné aj v digitálnom a algoritmickom prostredí. Ak sa podarí nastaviť živý dialóg medzi právom, profesijnou etikou a vývojármi AI, potom môže byť AI skôr nástrojom kultivácie verejnej diskusie než jej rozvratu. Odpoveď na otázku z názvu – či ide o úsvit alebo súmrak slobody prejavu – tak bude závisieť predovšetkým od kvality našich inštitúcií, etických noriem, schopnosti vyvodzovať z nich zodpovednosť a profesionálnej zodpovednosti, nie od samotnej existencie novej, perspektívnej technológie.

Zoznam použitej literatúry a zdrojov

1. BOOS, E., MÄLZER, M., CONRAD, F. et al.: *An Engineer-Friendly Terminology of White, Black and Grey-Box Models*. In: MODELSWARD: 13th International Conference on Model-Based Software and Systems Engineering. Proceedings, 2025, s. 313–320. Online: https://www.researchgate.net/publication/389542353_An_Engineer-Friendly_Terminology_of_White_Black_and_Grey-Box_Models.
2. CAMBRIDGE DICTIONARY: *Artificial Intelligence*. In: Cambridge Dictionary. 2025. Online: https://dictionary.cambridge.org/dictionary/english/artificial-intelligence.
3. CAMBRIDGE DICTIONARY: *LLM.* In: Cambridge Dictionary. 2025. Online: https://dictionary.cambridge.org/dictionary/english/llm.
4. ČESKÝ STATISTICKÝ ÚŘAD: *Informační společnost v číslech* – 2024. Online: https://csu.gov.cz/docs/107508/76df3534-a565-79c7-41c7-ce6cf2f44737/06100424c.pdf?version=1.0.
5. DODDS, T., ZAMITH, R., LEWIS, S. C.: *The AI turn in journalism: Disruption, adaptation, and democratic futures*. In: Journalism, Vol. 27, No. 4, 2025, s. 1–15. Online: https://www.researchgate.net/publication/391770691_The_AI_turn_in_journalism_Disruption_adaptation_and_democratic_futures.
6. DRGONEC, J.: *Ústava Slovenskej republiky-Komentár*. 2. vydanie. Bratislava: C. H. Beck, 2019, 1792 s.
7. *Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, Strasbourg, 5 September 2024, Council of Europe Treaty Series – No. 225.
8. FRITZ, J.: *Understanding authorship in artificial intelligence-assisted works*. In: Journal of Intellectual Property Law & Practice, roč. 20, č. 5, 2025, s. 354–364. Online: https://academic.oup.com/jiplp/article/20/5/354/7965768.
9. KLUČKA, J.: *Úloha a dôležitosť etických pravidiel v systémoch umelej inteligencie*. In: Právny obzor, roč. 108, 2025, č.3, s. 240–254. Online: https://doi.org/10.31577/pravnyobzor.2025.3.02.
10. MCQUAIL, D.: *Žurnalistika a společnost*, Praha: Univerzita Karlova, Karolinum, 2016, 254 s.
11. MERRIAM-WEBSTER: *Journal.* In: Merriam-Webster Dictionary. 2025. Online: https://www.merriam-webster.com/dictionary/journal#word-history.

12. Nález Ústavného súdu Slovenskej republiky zo dňa 11. 7. 2012, sp. zn. PL. ÚS 12/09.
13. Nález Ústavného súdu Slovenskej republiky zo dňa 18. 12. 2014, sp. zn. II. ÚS 307/2014.
14. Nariadenie Európskeho parlamentu a Rady (EÚ) 2022/2065 zo dňa 19. októbra 2022 o jednotnom trhu s digitálnymi službami a o zmene smernice 2000/31/ES (Akt o digitálnych službách), L 277/1, 27.10.2022.
15. Nariadenie Európskeho parlamentu a Rady (EÚ) 2024/1689 zo dňa 13. júna 2024 ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie a ktorým sa menia nariadenia (ES) č. 300/2008, (EÚ) č. 167/2013, (EÚ) č. 168/2013, (EÚ) 2018/858, (EÚ) 2018/1139 a (EÚ) 2019/2144 a smernice 2014/90/EÚ, (EÚ) 2016/797 a (EÚ) 2020/1828 (akt o umelej inteligencii), L, 12.7.2024.
16. RADA EURÓPY: *Europe Press Freedom Report 2024: Confronting Political Pressure, Disinformation, and the Erosion of Media Independence. Annual assessment of press freedom in Europe by the partner organisations of the Safety of Journalists Platform*, March 2025. Online: https://rm.coe.int/prems-013425-gbr-2519-annual-report-2025-correction-cartooning/1680b48f7b.
17. REMIŠOVÁ, A.: *Etika médií*, Bratislava: Kalligram, 2010, 312 s.
18. REUTERS INSTITUTE FOR THE STUDY OF JOURNALISM: *Digital News Report 2025 – Slovakia*. 10. December 2025. Online: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025/slovakia.
19. Rozhodnutie ESĽP vo veci *Hertel proti Švajčiarsku*, sťažnosť č. 59/1997/843/1049, rozsudok zo dňa 25. augusta 1998. Online: https://hudoc.echr.coe.int/eng?i=001-59366.
20. SAP: *What is machine learning?* Online: https://www.sap.com/products/artificial-intelligence/what-is-machine-learning.html.
21. SLOVENSKÁ ASOCIÁCIA NOVINÁROV: *Vyhlásenie Slovenskej asociácie novinárov zo dňa 17.5.2024*. Slovenská asociácia novinárov. Online: https://www.novinarisan.sk/2024/05/17/vyhlasenie-san-k-atentatu-na-roberta-fica-2/.
22. SLOVENSKÝ SYNDIKÁT NOVINÁROV: *Etický kódex .novinára.* 17. jún 2017. Online: https://www.ssn.sk/eticky-kodex-novinara/.
23. SNYDER, T.: *O tyranii: dvadsať ponaučení z 20. storočia.* Bratislava: Premedia, 2017, 103 s.

24. STATISTA: *Most trusted sources of news and information worldwide*. 2025. Online: https://www.statista.com/statistics/381455/most-trusted-sources-of-news-and-info-worldwide/
25. THORNTON, S. T.; MARION, J. B.: *Classical Dynamics of Particles and Systems*. 5. vydanie. Belmont: Brooks/Cole, 2004, 656 s.
26. ŠKRIPEK, M., et al.: *Čo nás stoja dezinformácie? Hlavní aktéri, strategická komunikácia a ekonomické aspekty dezinformácií*. Bratislava: Výskumné centrum Slovenskej spoločnosti pre zahraničnú politiku (RC SFPA), Friedrich Ebert Stiftung, 2022. Online: https://library.fes.de/pdf-files/bueros/slowakei/20643.pdf.
27. THOMSON REUTERS FOUNDATION: *Weaponizing the Law: Attacks on Media Freedom 2023*. Online: https://www.trust.org/wp-content/uploads/legacy/weaponizing-law-attacks-media-freedom-report-2023.pdf.
28. VANACKÝ, J.: *Sloboda prejavu v digitálnom priestore*. školiteľ: Michal Mrva. Bratislava: Univerzita Komenského v Bratislave, 2025, 123 s.

Informácie o autorovi

Mgr. Jakub Vanacký
Univerzita Komenského v Bratislave
Právnická fakulta
Katedra teórie práva a filozofie práva
Email: jakub.vanacky@flaw.uniba.sk

Human Rights and Freedoms under Pressure from Algorithms: International and National AI Challenges

Peer-reviewed proceedings of an international scientific conference

Edited by: Mgr. Lucia Bakošová, PhD.

Publisher: Pavol Jozef Šafárik University in Košice
by ŠafárikPress Publishing

Year: 2026
Pages: 333
Author's sheets: 22,45
Edition: first

DOI: https://doi.org/10.33542/HRF-0541-2
ISBN 978-80-574-0541-2 (e-publication)