AI as a judge and the problem of designation*

*Radoslav Pavlinský***

ABSTRACT

This article examines whether artificial intelligence (AI) could operate within a judicial system through the problem of designation, which poses the question of how society assigns roles and attaches certain expectations from them. Drawing on John Searle's theory of institutional facts, it argues that calling someone a judge is a collective act that along with certain constitutive rules creates a status carrying specific rights, duties, and symbolic authority. The article compares the designation of human judges with that of AI by analyzing three dimensions: formal, functional, and ontological. While AI could satisfy formal requirements, especially if we modify our existing legislation, and also functional requirements through technical sophistication, it cannot meet the ontological conditions of judging, such as self-awareness, conscience, and responsibility. These human qualities underpin genuine judgment and cannot be replicated by algorithmic systems. Thus, the barrier to AI becoming a judge is not merely procedural or functional, but ontological. Nevertheless, AI can serve as a valuable judicial tool, either by supporting legal research, pattern

* This article is a partial output of the following project: APVV-21-0336: Analýza súdnych rozhodnutí metódami umelej inteligencie (Analysis of the judicial decisions via methods of artificial intelligence)

** JUDr. Radoslav Pavlinský, PhD., Researcher at the Gustav Radbruch Institute of the Law Faculty - Universtity of Pavol Jozef Šafárik in Košice

recognition, and preliminary assessments, so long as its role remains distinct from that of a judge.

Key words: Designation, Artificial Intelligence, Judge, Institutional facts, Machine learning

AI kot sodnik in problem označitve

POVZETEK

Ta članek preučuje, ali bi lahko umetna inteligenca (UI) delovala v sodnem sistemu skozi problem označitve, ki odpira vprašanje, kako družba posameznikom in subjektom pripisuje določene vloge ter z njimi povezana pričakovanja. Izhajajoč iz teorije institucionalnih dejstev Johna Searla članek trdi, da je poimenovanje nekoga za sodnika kolektivno dejanje, ki skupaj z določenimi konstitutivnimi pravili ustvarja status, povezan s posebnimi pravicami, dolžnostmi in simbolno avtoriteto. Članek primerja označitev človeških sodnikov in umetne inteligence z analizo treh razsežnosti: formalne, funkcionalne in ontološke. Medtem ko bi umetna inteligenca lahko izpolnjevala formalne zahteve, zlasti ob ustreznih spremembah obstoječe zakonodaje, ter tudi funkcionalne zahteve zaradi svoje tehnične sofisticiranosti, ne more izpolniti ontoloških pogojev sojenja, kot so samozavedanje, vest in odgovornost. Te človeške lastnosti predstavljajo temelj resničnega sojenja in jih ni mogoče reproducirati z algoritmičnimi sistemi. Ovira, da bi umetna inteligenca postala sodnik, zato ni zgolj postopkovne ali funkcionalne narave, temveč ontološka. Kljub temu pa lahko umetna inteligenca služi kot dragoceno sodno orodje, bodisi pri podpori pravnim raziskavam, prepoznavanju vzorcev in predhodnih ocenah, dokler njena vloga ostaja jasno ločena od vloge sodnika.

Ključne besede: Označitev, Umetna inteligenca, Sodnik, Institucionalna dejstva, Strojno učenje.

1. Introduction

The emergence of what is broadly referred to as artificial intelligence (AI) has prompted our collective human imagination regarding its possible roles within society. This development is unsurprising. Humans, as a species, rely extensively on imagination and our social world is structured through an array of collectively constructed designations. Among the more provocative questions in our current times is whether AI could act as a judge, or more broadly, how AI might be integrated into a judicial system. This inquiry ultimately concerns whether AI can meaningfully participate in the human world of institutional constructs, symbolic roles, and socially embedded expectations. It is this question that the present article seeks to address.

The article consists of three parts. The first examines how people create what American philosopher John Rogers Searle terms institutional facts, of which law is a perfect example. This section links Searle's framework to the concept of designation and explores how each designation, especially that of a judge, carries specific expectations. These expectations are analyzed in three dimensions: formal, functional, and ontological. The second part considers whether an AI being designated as a judge could satisfy these expectations, and how such a designation differs from that of a human judge. This analysis aims to clarify the procedural, practical and conceptual limits of AI within judicial systems. The final part of the article addresses the other side of the debate, often neglected in the discourse on potential of the AI in our judicial systems. Human judges themselves are imperfect, raising the question of why AI should be evaluated against an idealized judge rather than a real one.

2. The act of designation and how a judge manifests in our society

„How can there be an objective world of money, property, marriage, governments, elections, football games, cocktail parties and law courts in a world that consists entirely of physical particles in fields of force, and in which some of these particles are organized into systems that are conscious biological beasts, such as ourselves?“ – John Rogers Searle (Searle, 1995, pp. xi-xii)

One of the most important differences between other animals and humans is the power of our imagination. This imagination shows in our everyday activity, often without us realizing. A simple buying a ticket to be transported from a place to place via a bus is anything but trivial and in fact is almost mindboggling in its complexity. At the same time, we imagine a person to be a „driver“, a metal carriage on circular shaped objects a „bus” and fellow humans inside this moving structure to be „passengers“. Things get even more complex when we add in the fact that bus arrives and departs at certain prerequisite times, we pretend that a ticket is anything more than a pressed wood matter with some dye on it and that everyone on the bus is convinced there are some kind of rules to follow, probably posted behind the backseat of the bus driver. Humans are a designatory species and our world is filled with designations. Each designation brings with itself certain expectations. We expect bus to take us somewhere, bus driver to be competent enough to drive the bus, ticket to be a proof of contract we made with the driver and the transportation company.

This brings me to John Searle and his theory of institutional facts. To Searle, for human mind complex ontology seems simple, while simple ontology seems difficult. (Searle, 1995, p. 4) In his book, *The Construction of Social Reality*, he seeks to explain how people work with designations and create institutional facts in order to make working with

true natural states of the objects easier. He connects the idea to intentionality and ontological subjectivity. Screwdriver is for example only a screwdriver based on its function of driving screws, thus ontologically subjective. (Ibid., pp. 7-10) Intentionality is connected to self-awareness and besides humans, all animals practice it to their best ability. It is through intentionality that institutions and institutional facts are created. Their existence is not tied to the material qualities of the objects, but to the people who consider them to be facts in their beliefs, desires, intentions, expectations and other intentional states. (Toráčová, 2012, pp. 83-84)

We all collectively impose status functions to objects. As such, the question is whether it is even important to us if the object in question is capable of fulfilling the conditions of satisfaction prescribed by the intentional states. By conditions of satisfaction, we mean the ontological ability of the objects to serve in their designated function. (Ibid., p. 90) In some cases, these are rather simple to fulfill. Screwdriver fulfills the condition of satisfaction rather easily when used for driving screws. But can we think of other ways to use the screwdriver? The answer to that question is yes. While we may not primarily think of screwdriver as a murder weapon, it could very well serve in that capacity. Each object thus can have many more functions imposed upon it, although this ability is limited by its ontological capabilities. In some cases, however, certain objects appear to satisfy the conditions of satisfaction in human minds, despite the object ontologically being incapable of fulfilling them. Such is also the perceived position of AI in our present society.

Aside from rather simple institutional facts connected to the objects, some institutional facts only come to fruition via so called constitutive rules. These not only regulate, but first of all they create the very possibility of a certain type of institutional facts, which then can be further

regulated. Designations thus have a following structure: „X counts as Y in C". (Moural, 2002, p. 277) To put it simply, existence of some institutional facts is regulated by certain rules and only after these rules apply can we as a society collectively accept the designation and assign functions. In other words, constitutive rules are „conditions of possibility" of an institutional fact. (Ricciardi, 2015, p. 1) This is where law comes into play. In many ways, legal world is built upon various designations, and its purpose is mainly regulation of institutional facts. As such, some designations are deemed harder to fulfill the conditions of satisfaction for. It is not enough for people to designate an object or subject to be a judge. For this institutional fact to become true, we first need to abide by a certain set of constitutive rules.

As stated above, every designation has certain expectations put on it. These are not necessarily connected to the conditions of satisfaction being fulfilled. We can have judges, for example, that do not fulfill the expectations put on them by society. The only requirement of a designation of judge as a role in a society is social acceptance. But this is not an agreeable state and thus in our own wisdom, we created additional constitutive rules in our legal systems, that both constitute a judge as an institutional fact, meaning it is not enough for designation to be socially accepted, and also put certain requirements on the designation, in an effort to fulfill the expectations of the society. We decided to split these requirements into three categories: formal, functional and ontological. Formal conditions are laid out in the law and vary from country to country. Functional conditions are connected to the ability of the judge to perform activities and tasks connected to his designation as a judge. Lastly, ontological conditions are deeper and represent some limitations on who or what can serve as a judge. We will start our analysis with the formal conditions.

3. How conditions are made to meet expectations

3.1. Formal and functional conditions and the AI

Each country has its own set of formal conditions to be met by an aspiring judge. These are often the subject of their own laws, to showcase the importance of the judge in the legal system. These conditions do vary, but most will contain the following requirements:

- age requirement
- formal education in law
- full legal capacity
- no criminal record
- medical fitness to perform the duties of a judge
- permanent residence within the country (and often also to be a citizen)
- passing some sort of judicial examination
- go through selection process
- consent to being appointed as a judge
- other prerequisites of judicial competence that guarantee they will perform the duties of a judge properly (most often connected to morality and integrity)[1]

Formal requirements are meant to fulfill certain expectations we as a society might place on the judge. We want judges to be wise, just, objective, knowledgeable, moral and one of us. None of these are obviously able to fulfill the expectations by themselves, but all of these need to be adhered to for a judge to be appointed. Just from the conditions alone, we see that these are clearly geared towards humans. AI cannot attend university and cannot age. It cannot give consent to be appointed as a judge. More pressingly, it lacks full legal capacity. Let us unpack now two of

[11] This summary of the requirements were pulled from the Slovak legal system, in particular § 5 of the Law no. 385/2000 Zz. on judges and lay judges: https://www.slov-lex.sk/ezbierky/pravne-predpisy/SK/ZZ/2000/385/

the most important formal requirements, that of education and selection process.

It would seem obvious that the requirement of education is meant to make sure that the judge is knowledgeable enough to perform his duties. But this requirement is much more complex than it may first seem. Historically, the stature of the legal expert, whether scholar, practitioner or a judge, rested not only on knowledge of content, but on the difficulty of acquiring and developing legal expertise. (5+) Years of study, training and cultivation were necessary, and it came at a rather large financial price. Not everyone could have become a judge. Thus, education is not merely an assurance of quality but also connected to the authority of the judge. People are more likely to respect the judge's decision, if he undertook rigorous legal training. The advent of AI use in legal fields means an erosion of that authority somewhat as the professional mystique gets lost, especially if we use AI in predicting the judicial outcomes. Furthermore, the result of extensive reliance on AI in generating judicial decisions may result in a generation of legal professionals who can work faster but not think deeper. (Skouteris, 2025, p. 6) In regard to whether AI can somehow supplement the requirement of education, it can. Through machine learning, AI can simulate the understanding of the legal landscape to a certain extent. It does however lack the ontological capacity to truly understand law and institutional facts.

Perhaps the most important prerequisite for becoming a judge is successfully completing the relevant selection process. Although selection criteria vary across jurisdictions, two principal models of judicial creation can be distinguished: judges who are elected and career judges. Each model carries distinct advantages and disadvantages. Elected judges face a frequently noted issue. Due to their need to obtain financial support for their election campaigns, their impartiality may be undermined their decisions are also more likely to be perceived as biased, as their decisions may appear

responsive to constituents or political donors, potentially weakening their authority. (Fine, Berthelot, Marsh, 2025, p. 6) At the same time, it is difficult to find a stronger source of democratic legitimacy than direct election.

Whereas elected judges must navigate the pressures of political accountability, career judges, by virtue of being appointed through a selection process and due to self-governance, are largely shielded from such concerns, thereby maintaining their legitimacy even when issuing unpopular decisions. (Ibid., p. 7) The German system, one of the most fully developed judicial career structures, exemplifies this model. The promotion process is formalized and closely aligned with civil service rules. Vacancies for judicial positions are publicly advertised, and candidates undergo a hierarchical evaluation involving court presidents and the Ministry of Justice, which ultimately determines the best-qualified applicant. Although the ministry or cabinet formally makes the final decision, advisory bodies such as presidential councils or judicial electoral committees often play a decisive role through their ranking of candidates. (Riedel, 2020, p. 2)

Of particular relevance here are the substantive criteria used to evaluate candidates. These criteria often extend far beyond the formal qualifications set out in legislation. Selection procedures aim to assess applicants in sufficient depth to allow a reliable prediction of their future judicial performance. (Riedel, 2013, p. 49) Commonly, the criteria are grouped into four categories: professional competence, personal competence, social competence, and leadership competence. Each category encompasses numerous specific expectations. While it is not possible to examine them exhaustively here, several examples present challenges if applied to the AI: willingness to assume responsibility for judicial decisions; sensitivity, patience, and understanding in dealing with parties, breadth of interests, natural authority, awareness of social responsibility, the ability to communicate (where AI faces inherent ontological limitations) and

acceptance as an authoritative figure.[2] It is noteworthy, however, that AI cannot be dismissed outright as incapable of meeting these categories. Through sophisticated machine-learning techniques, AI can increasingly simulate aspects of professional, social, and even personal competence at a functional level. Yet the deeper functional and ontological requirements of judging remain unresolved.

Judges are not merely appointed to office but are also entrusted with a set of core responsibilities integral to the exercise of judicial duties, most notably argumentation, interpretation, and decision-making. These functions are also the primary targets of often proposed AI-based replacements. Yet a fundamental obstacle arises. The formal requirements imposed by legislation do not exist solely to create the office of a judge, but to ensure that judicial duties are carried out to a prescribed professional and ethical standard. These conditions are meant to manifest in the judge's work. When tasks are delegated to AI systems, the link between these formal conditions and the resulting performance of judicial duties is under threat or even nullified. This raises a substantive argument against employing AI, whether as a judge or even as a judicial tool, because the statutory requirements permeate all aspects of judicial work. Nonetheless, potential benefits do exist. AI can streamline labor intensive or repetitive tasks, thereby allowing judges and lawyers to redirect their efforts toward more complex intellectual tasks. (Ruhil, 2025, p. 5551) Such efficiency gains could alleviate systemic burdens and improve access to justice. However, efficiency cannot be equated with justice itself.

Although AI can approximate certain judicial behaviors through machine learning techniques when trained on carefully curated datasets, it cannot meaningfully satisfy the formal criteria established in law. Further difficulty stems from

[2] These were pulled from the selection criteria for the position of a judge in the German federal state of Bavaria. Full list available at: https://www.justiz.bayern.de/media/pdf/anfprofilrus.pdf

a deeply rooted assumption about what constitutes the performance of judicial duties. Judging is seen as an inherently human activity. Even if AI can simulate many components of judicial reasoning, most people remain reluctant to accept a non-human judge. This reluctance reflects the importance of the ontological conditions traditionally associated with judicial office, conditions that AI, by its nature, cannot fulfill.

3.2. Ontological conditions and the AI

We now once again return to Searle and look at his famous Chinese room thought experiment. In it, we place person into a room and give him a set of instructions on how to respond to the questions asked by a Chinese correspondent from the outside. The trick is, that the person inside the room does not understand Chinese at all. The person who knows Chinese is asking our person inside the room some basic questions. The one inside the room consults his instructions and answers appropriately. In no shape or form can this be considered as him understanding Chinese. An argument could be made that while he himself does not understand Chinese, the entire system understands it. But that is also not possible. Whether it be the single person or the entire system, neither understands Chinese, only simulates the understanding. (Searle, 1980, p. 420)

The same can be said about the AI and its capability to serve as a judge. To be a good judge, the person needs the ability to work with institutional facts and understand the law, not merely simulate understanding of it. Especially using GenAI is problematic in this regard. Its responses are generated from patterns identified in the training data and refined through developer fine-tuning. However, these systems do not possess true understanding or the capacity for critical thinking. (Socol, Remolina, 2024, p. 5) A judge determining a criminal sentence must consider not only the facts but also the subtler dimensions of intent, redemption, and societal impact. Likewise, courts handling constitutional questions must often weigh individual freedoms against the

common good. These are all decisions rooted in history, human empathy, and contextual understanding that no machine, however advanced, can truly replicate. (Ruhil, 2025, p. 5551) Furthermore, there are other ontological conditions that AI fails to meet, such as responsibility and accountability.

These are often connected to lack of transparency in current GenAI supported judicial decision-making. Judges and other legal experts are not IT professionals and thus are not well versed in understanding how a specific decision was reached, and which biases or errors were introduced in the final rationale provided. (Socol, Remolina, 2024, p. 6) It is perfectly understandable and reasonable demand to seek to fully grasp how AI reasons before endorsing its use in judicial capacity. Even if we do not entirely understand human reasoning either. As humans, we also have a natural preference for the status quo, even when change could bring positive benefits overall. These tendencies may explain the widespread resistance to integrating AI into the judiciary. (Winter, 2022, pp. 198-199)

There are also ethical concerns connected to AI. Can AI be expected to give ethical dimensions of legal cases proper attention? (Ruhil, 2025, p. 5552) The law frequently governs issues with profound moral dimensions, such as abortion, homicide, and constitutional disputes surrounding values. Simultaneously, it pursues certain normative objectives, including the reduction of crime, the promotion of equality, and the redistribution of wealth. Moreover, law serves as an instrument for regulating state power, capable of being employed for both beneficial and harmful purposes on a scale beyond that of individual action. (Valvoda et al., 2024, p. 703) In other words, the number of considerations going into each individual case may prohibit the use of AI, deciding mainly on learned patterns. These patterns are of course the very problem with AI as a judge. When either the laws themselves or judicial decisions exhibit prejudice, and such

biases are not excluded from training datasets, generative AI systems risk replicating and amplifying them in their outputs or case predictions. Consequently, AI may inadvertently entrench existing societal biases, resulting in discriminatory outcomes and aggravating the very problems it seeks to address. (Fine, Berthelot, Marsh, 2025, p. 2)

It is also important to note that perceptions of AI and trust in its judicial application may vary significantly across social groups, each holding distinct expectations of what constitutes a just and legitimate system. While research conducted in the United States highlights differences across racial communities, such as Black, Hispanic, and White populations (Ibid., p. 7), these variations likely reflect broader social dynamics rather than race alone. Different groups interpret a judge's reliance on AI in different ways. For some, it may symbolize progress and equality of treatment, while for others, it may evoke skepticism or unease. Yet, the authority of the judiciary cannot rest solely on social acceptance or fluctuating public sentiment. Furthermore, public discourse frequently evaluates AI against an idealized conception of a flawless human judge, which is an unrealistic and logically flawed comparison that ignores the inherent imperfections of both human and artificial decision-making. In practice, judges differ widely in competence and integrity. Some act with exemplary professionalism, while others may be influenced by corruption, nepotism, or personal bias. The notion of a universally virtuous and infallible judge is therefore fiction. With this in mind, we now turn to the research conducted within our project and the insights it has yielded.

4. Judges are not perfect, why compare AI to a perfect judge?

The focus of our research was the application of artificial intelligence in the analysis of legal judgments. Specifically, we employed machine learning techniques to exam-

ine criminal verdicts issued in the Slovak Republic, with the dual objective of assessing their adequacy and establishing a baseline for predictive modeling of judicial outcomes. The AI system was trained on two datasets and subsequently tested on a separate dataset, tasked with predicting whether the verdict would be guilty or not guilty. (Varga et al., 2021, p. 1) Particular attention was devoted to evaluating adherence to the principle of *in dubio pro reo*, which stipulates that doubt should favor the defendant, who must be presumed innocent unless proven guilty.

The model's predictions were subsequently reviewed by legal experts to assess their accuracy. Through the optimization of machine learning filters by our IT specialists, the system achieved an accuracy rate exceeding 99%. (Ibid., p. 7) However, a deeper analysis revealed a critical discrepancy. Although the AI's predictions aligned with the legal experts' assessments, the actual judicial decisions made by human judges frequently diverged from the *in dubio pro reo* principle. In theory, one would expect stronger evidentiary reasoning to be associated with guilty verdicts, given the burden of proof lies with the prosecution. In practice, our findings suggest the opposite. Judges often appeared to have placed the burden of proof on defendants, effectively requiring them to demonstrate their innocence. (Ibid., p. 8) Notably, in such cases, the AI system consistently rendered a not-guilty verdict.

What do these findings reveal about human judges? The principle of *in dubio pro reo* served as a cornerstone of criminal justice since the Roman era and requires that doubt favor the accused. Yet our results indicate that this principle is not consistently upheld. Although further empirical investigation is necessary to fully explain this discrepancy, three preliminary explanations can be proposed. The first explanation is methodological. Since the AI based its predictions on the written reasoning contained within judicial decisions, any deficiencies in the completeness or clarity of that rea-

soning could have led the system to predict a not-guilty verdict. Variability in the quality of judicial reasoning is a well-documented phenomenon and could therefore account for part of the observed divergence. The second explanation relates to the evidentiary basis of judgments. In some cases, although the reasoning did not sufficiently substantiate a guilty verdict, judges may have relied on circumstantial or unrecorded evidence not explicitly reflected in the written decision. While such practice undermines transparency and procedural integrity, it nonetheless occurs in real-world adjudication. The third, and most concerning, possibility is normative. Some judges may consciously or unconsciously prioritize the perceived protection of society over the presumption of innocence, thereby lowering the evidentiary threshold required for conviction. This approach directly contravenes the rationale behind *in dubio pro reo*, which exists precisely to prevent the conviction of the innocent, which represents a miscarriage of justice that serves neither the individual nor society at large.

While artificial intelligence can, as demonstrated, serve as a valuable tool within judicial systems, revealing patterns and discrepancies in the performance of judicial duties, it faces a significant obstacle when considered as a potential decision-maker in legal proceedings. The core challenge lies in the fundamental legal principle that each case must be adjudicated independently, based solely on its unique facts and circumstances. Given that AI systems rely inherently on pattern learning and statistical generalization, they are ill-suited to assess individual cases in full isolation from prior examples. Furthermore, the process by which AI systems acquire knowledge presents additional concerns. As previously noted, careful attention must be paid to the composition of training datasets. Should datasets include cases reflecting systemic bias or discrimination, particularly racial bias, these patterns risk being replicated and amplified by the AI itself. Equally critical is the principle of ju-

dicial independence. In practice, it is the AI developers, rather than the judiciary, who determine how an AI system functions and from which datasets it learns. While this issue could be partially mitigated by allowing judges to select or curate the training datasets, ultimate control over programming and model design would still reside with the system's developers, raising questions about external influence and accountability.

There is, however, another side to the debate. While each case must indeed be adjudicated on its own merits, judges are simultaneously bound by the duty to uphold consistency in judicial decision-making. Maintaining legal certainty requires that judges remain aware of prior cases to ensure coherence and predictability within the legal system. To achieve this, judges must currently conduct manual searches through previous decisions and extract relevant keywords, a task that is both time consuming and methodologically complex. In the Slovak Republic, only one institution currently provides such a service, and it is limited to the judges of the Supreme Court. (Varga et al., 2022, p. 1)

Artificial intelligence could substantially improve this process. By developing a search tool based on machine learning, the benefits of automated case analysis could be extended beyond the Supreme Court to lower courts, thereby easing the workload of judges and enhancing access to relevant jurisprudence. Our IT team explored various techniques to optimize the tool's performance, testing multiple methods of key phrase extraction. The most promising approach proved to be the TF-IDF (Term Frequency–Inverse Document Frequency) method, which penalizes phrases that occur frequently across many documents while giving greater weight to less common, and therefore potentially more distinctive, phrases. (Ibid., p. 6) However, infrequent phrases are not necessarily semantically meaningful. Phrases that are semantically dissimilar from the dominant themes in a corpus are more likely to serve as informative key phrases useful for

clustering documents. When combined with autoencoders (unsupervised neural networks designed to generate efficient representations of data by isolating the most salient features), the method yielded the best results. We compared the automatically generated key phrases with both expert-written abstracts and manually extracted key phrases. The findings revealed that the algorithm's output corresponded more closely to expert abstracts than to manually extracted keyphrases. (Ibid., p. 7) This finding suggests that AI did overall a much better job than the people who used manual extraction methods.

Even in regard to the positives of the AI, however, several important caveats must be acknowledged. As previously discussed, the integration of AI into judicial functions may conflict with fundamental principles of the judiciary, particularly those concerning the nature and autonomy of judicial reasoning. If AI systems are permitted to assume a greater share of these interpretive or analytical tasks, these risks diminish the practical relevance of a judge's legal education and expertise. While AI may perform effectively when compared to an imperfect human judge, this comparison is inherently problematic. Training AI models on large datasets of judicial decisions would likely produce an "average" AI judge, one that replicates the statistical tendencies of the judiciary on the whole rather than embodying its aspirational standards in the idealized judge. Attempts to counteract this through curated datasets would raise serious ethical and legal concerns, as such curation could introduce bias or discrimination. Furthermore, granting a select group the authority to determine which judges' decisions qualify as exemplary would undermine the perceived legitimacy of the judiciary and erode public trust in the justice system.

Taken together, these considerations lead to the conclusion that, while AI may serve as a valuable tool in assisting judges and improving their efficiency, it should not, and indeed, cannot replace human judges. Despite its capacity to

outperform humans in certain narrow tasks, entrusting AI with judicial authority would jeopardize the foundational values upon which the integrity of the judiciary depends. What does this mean for its place in our judicial systems? Proposals for integrating artificial intelligence into judicial systems generally fall into two categories. We either seek the replacement of human judges with AI systems, or the augmentation of judges and lawyers through AI-based tools. (Valvoda et al., 2024, p. 700) The latter approach envisions only partial automation, typically through the use of Natural Language Processing (NLP) to assist with specific legal tasks, as demonstrated in our research.

When conceptualized as a tool to streamline procedures, minimize human error, and base decisions on data rather than subjective judgment, AI can be seen as promoting greater fairness and transparency within the judiciary. Conversely, if associated with bias, opacity, or the dehumanization of legal processes, AI risks eroding public trust in judicial outcomes. (Fine, Berthelot, Marsh, p. 4) The potential benefits of NLP are substantial, so much so that some consider rejecting its use could be viewed as analogous to refusing modern medical treatments that improve patient care. (Ruhil, 2025, p. 5553) Nevertheless, these advantages must be carefully balanced against the considerable ethical risks that accompany such technologies. (Valvoda et al., 2024, p. 705)

Equally important is the question of judicial authority and legitimacy. Empirical research conducted in the United States indicates that judges who rely solely on their professional expertise are perceived more favorably than those who employ AI, either entirely or in conjunction with human judgment. (Fine, Berthelot, Marsh, 2025, p. 14) This finding suggests that the public may equate AI use with a reduction in genuine human deliberation, particularly in duties of such importance as sentencing, thereby amplifying concerns regarding fairness and moral accountability. Even when AI serves only as a supportive tool, its influence on public trust

remains a critical factor in the broader discourse on the integration of AI within judicial systems.

5. Conclusion

In this paper, we have examined formal, functional, and ontological requirements a judge must satisfy and how these shape the role of a judge in our society. Through the lens of Searle's theory of institutional facts, we have shown that the judicial role is sustained by a complex interplay of designations, constitutive rules, and socially embedded expectations that define what it means when we designate someone to be a judge. While AI may credibly simulate functional aspects of judicial work and, with a rewrite or adding new sections to existing legislature, could even be formally designated as a judge, the ontological dimension, encompassing understanding, responsibility, conscience, and the capacity to work with institutional facts, remains beyond the reach of algorithmic systems. Importantly, this does not imply that human judges themselves are perfect arbiters of justice. They themselves are prone to error, and the tendency to measure AI against a hypothetical "perfect judge" is a common logical fallacy.

AI can in fact help identify patterns and inconsistencies within judicial reasoning with notable accuracy and expose both the strengths and the weaknesses of human adjudication. The divergence between algorithmic predictions and actual decisions reveals imperfections deeply embedded in legal practice, particularly in the application of principles such as *in dubio pro reo*. As a judicial tool, AI can clearly enhance judicial performance, improve access to justice, identify inconsistencies, and alleviate the burden of routine tasks. In this supportive capacity, AI may increase efficiency, transparency, and even the quality of our judicial systems. However, as argued throughout this paper, even this limited role raises issues that must be approached with care, par-

ticularly regarding accountability, bias, and over-reliance on algorithmic reasoning.

Overall, the narrative of AI as an "antidote to the legal chaos of the present" should also be rejected. Such framing reflects a growing reluctance to engage with the demands of justice and a troubling erosion of trust in the legal system. What does it say about us when we prefer automated justice to the collective effort of maintaining a functioning, accountable judiciary? If the judiciary is defined not only by performance of tasks but by a role embedded in formal, functional, and ontological requirements, then automating the tasks does not replicate the judicial function in society. It dissolves it.

Ultimately, the barrier to AI serving as a judge is not merely formal or functional, but ontological. Judging is not a task to be executed. It is a role to be inhabited, a role constituted by the very ontological features that make us human. AI can illuminate weaknesses in our legal system and support its operation, but it cannot be a judge. The institution of judging, with its ethical weight and societal significance, remains grounded in human agency, and it is within the human domain that justice must continue to be rendered, even if at times it may prove to be imperfect and messy.

LITERATURE AND SOURCES:

Fine, A., Berthelot, E.R., Marsh, S. (2025). Public perceptions of judges' use of AI tools in courtroom decision-making: An examination of legitimacy, fairness, trust, and procedural justice. Behavioral Sciences, 15/4, Article 476.

Moural, J. (2002). Searle's theory of institutional facts: A program of critical revision. In Speech Acts, Mind, and Social Reality – Discussions with John R. Searle./Grewendorf, G., Meggle, G. (eds.). Springer Dordrecht, pp. 271-286.

Ricciardi, M.. (2015). Constitutive rules and institutions. Available at: https://doi.org/10.13140/RG.2.1.4803.9841

Riedel, J. (2013). Training and Recruitment of Judges in Germany. International Journal For Court Administration, 5/2, pp. 42-53.

Riedel, J. (2020). Judicial review of judicial appointments in Germany. International Journal For Court Administration, 11/1, Article 2.

Ruhil, O. (2025). Justice in the age of algorithms: Can AI weigh morality? AI & Society, 40, pp. 5513-5551.

Searle, J.R. (1980). Minds, brains and programs. Behavioral and Brain Sciences, 3/3, pp. 417-424.

Searle, J.R. (1995). The construction of social reality. New York, United States: The Free Press.

Skouteris, T. (2025, Spring/Summer). What counts as legal knowledge in the age of AI? The Cairo Review of Global Affairs. Available at: https://www.thecairoreview.com/essays/what-counts-as-legal-knowledge-in-the-age-of-ai/

Socol, D.U. de la O., Remolina, N. (2024). Artificial intelligence at the bench: Legal and ethical challenges of informing—or misinforming—judicial decision-making through generative AI. Data & Policy, 6/e59. (e59).

Toráčová, P. (2012). Intentionality and what we can learn about it from Searle's theory of institu**ti**ons. Organon F, 19/SI2, pp. 83-92.

Valvoda, J. *et al.* (2024). The ethics of automating legal actors. Transactions of the Association for Computational Linguistics, 12, pp. 700-720.

Varga, D., *et al.* (2022). Keyphrase extraction from Slovak court decisions. Available at: https://ceur-ws.org/Vol-3226/paper16.pdf

VARGA, D., *et al.* (2021). Analysis and Prediction of Legal Judgements in the Slovak Criminal Proceedings. Available at: https://ceur-ws.org/Vol-2962/paper13.pdf

Winter, CH. K. (2022). The Challenges of artificial judicial decision-making for liberal democracy. In Judcial Decision-Making: Integrating Empirical and Theoretical Perspectives/ Bystranowski, P., Janik, B., Próchnicky, M. (eds.) Springer Cham, pp. 179-204.