

# Právo na odôvodnené rozhodnutie a algoritmizácia rozhodovacích systémov

Andrej Krištofik\*

**Abstrakt:** Jednou z podstatných náležitostí, nielen súdnych, ale v podstate akýchkoľvek verejnoprávných rozhodnutí je ich odôvodnenie. Takéto odôvodnenie musí spĺňať určité štandardy a ich nenaplnenie môže viesť minimálne k vytvoreniu priestoru pre napadnutie takéhoto rozhodnutia, v horšom prípade k porušeniu práva na spravodlivý proces. Táto práca sa bude zaoberať otázkou, aký minimálny štandard a v akých rozhodnutiach existuje pre odôvodnenia a či je možné ho naplniť aj pri algoritmickom rozhodovaní bez zásahu človeka, s ohľadom, mimo iné, na technickú komplikáciu predstavovanú problémom algoritmickej čiernej skrinky. Článok vychádza z Nálezu Ústavného súdu II. ÚS 2522/19, respektíve rozsudku ESLP vo veci *Paradiso a Campinelli proti Taliansku* – rozsudky na seba vzájomne nadväzujúce, ktoré spomínajú aj automatizované rozhodovanie ako neprípustné, čo je záver ktorý sa tento článok snaží vyvrátiť. V neposlednej rade bude výsledok konfrontovaný s článkom 22 GDPR, respektíve primárne s francúzskou a maďarskou úpravou, ktorá rozširuje právo na vysvetlenie (právo pochopiteľnosti) subjektu.

**Kľúčové slová:** automatizácia, právo na spravodlivý proces, odôvodnené rozhodnutie, GDPR, algoritmizácia spravodlivosti

## Úvod

Všetky úvahy o inovovaní súdnych procesov sú viazané niekoľkými podmienkami. Tie môžeme v podstate rozdeliť na tie, od ktorých je možné nejakým spôsobom upustiť, no jedná sa o nejakú formu „*best practice*“ a obecné to nedáva zmysel.<sup>1</sup> Tú druhú skupinu potom tvoria tie, ktorým sa jednoducho nemožno vyhnúť, teda také podmienky, ktoré predpokladá zákon. V tejto podkategórii sa potom ešte nachádzajú aj (základné) ľudské práva, ktoré do určitej miery predstavujú jadro týchto podmienok, nakoľko je možné si predstaviť určitý zásah, napr. do rôznych procesných pravidiel, ktoré svojou zmenou umožnia fungovanie nejakej inovácie, táto zmena má však mantinely stanovené práve týmito základnými právami.

Zrejme najrelevantnejším základným právom v tomto ohľade, ktoré nám predstavuje práve toto pomyselné vymedzenie mantinelov implementácie, bude právo na spravodlivý proces, ktoré vo svojej podstate nie je právom ale súhrnným označením pre celý výpočet práv, ktoré sa pod túto hlavičku bežne radia. Posúdenie dopadu akejkoľvek inovácie, v našom prípade sa bude jednať o automatizáciu rozhodovacieho procesu,<sup>2</sup> je potrebné

\* Mgr. Andrej Krištofik, autor je doktorandom na Ústave práva a technológií Právnickej fakulty Masarykovej Univerzity. E-mail: andrej.kristofik@gmail.com. ORCID: <https://orcid.org/0000-0001-8150-0362>. Tento článok vznikol na Masarykovej univerzite v rámci projektu *Právo a technológia X* číslo MUNI/A/1484/2021 podporeného z účelovej podpory na špecifický vysokoškolský výskum poskytovanej Ministerstvom školstva ČR v roku 2022.

<sup>1</sup> Príkladom môže byť, pri zavádzaní technických inovácií, „Susskindova maxima“ že akýkoľvek takýto postup musí byť, a viesť, k väčšej miere efektivity, inak sa jedná o „inováciu pre inováciu“. Pozri: SUSSKIND, R. *Online courts and the future of justice*. Oxford: Oxford University Press, 2019.

<sup>2</sup> Automatizácia rozhodovacieho procesu je v kontexte tejto práce chápaná ako prenesenie rozhodovacej činnosti výhradne na neľudských aktérov a to v celom jeho procese, s výnimkou nejakej asistenčnej funkcie v podobe zadávania dát. Celé rozhodnutie je v tomto systéme vykonané bez akéhokoľvek zásahu človeka.

vykonať voči všetkým jednotlivým zložkovým právam. V tejto práci sa však budeme venovať analýze len jedného z nich a tým je právo na odôvodnené rozhodnutie, ktoré je ešte ako také súčasťou ďalšieho, nadradeného, zložkového aspektu práva na spravodlivý proces, označované najčastejšie ako aspekt „spravodlivého konania“.<sup>3</sup>

## 1. Odôvodnené rozhodnutie

Neodôvodnenie rozhodnutia je teda závažnou procesnou vadou (a porušením práva na spravodlivý proces), ako také teda je relatívne často namietaným právom pred Ústavným súdom, čo nám dáva relatívne dobrú predstavu o obsahu tohto práva. Aplikácia tohto práva sa opiera nielen o jeho zakotvenie v medzinárodných dokumentoch, ale aj na národnej úrovni v čl. 36 Listiny, a následne na zákonnej úrovni v kontexte civilného konania v § 157 zákona č.99/1963 Sb., prípadne v trestnom konaní v § 134 zákona č. 141/1961 Sb., pričom určité rozhodnutia si podľa zákona vyžadujú ďalšie špecifické body odôvodnenia, ako je tomu napríklad pri rozhodnutí o vzatí do väzby podľa § 73c, eventuálne v správnom konaní sa náležitosti odôvodnenia riadia podľa § 68 zákona č. 500/2004 Sb.<sup>4</sup>

Právo na spravodlivý proces je okrem národnej úpravy zakotvené napríklad aj v čl. 6 Európskeho dohovoru o ľudských právach („Dohovor“), pod hlavičkou ktorého sa tiež skrýva právo na odôvodnené rozhodnutie.<sup>5</sup> To má za účel chrániť záujmy jednotlivcov a umožňovať verejnosti kontrolu výkonu súdnej moci, respektíve kontrolu rozsudkov vynešených jej menom.<sup>6</sup> Takéto ponímanie je v kontraste s historickým vznikom odôvodnení, v ktorých pôvodne svoje rozhodnutie nižší súd vysvetľoval nadradenej inštancii či orgánu dozoru.<sup>7</sup> Okrem toho však už ESLP, respektíve Dohovor, na národné súdy nekladie nejaké širšie povinnosti, respektíve iba také, ktoré v porovnaní s požiadavkami čl. 36 Listiny môžeme považovať za minimálne.<sup>8</sup> To sa odráža potom samozrejme v rôznych národných prístupoch, od napríklad toho českého, ktorý dôsledne dbá na úmerné vysporiadanie sa s každou vznesenou námietkou, cez ten anglosaský, ktorý si zakladá na budovaní naratívu až po sylogistické odôvodnenia typické pre francúzske súdy,<sup>9</sup> ktorý je práve pri diskusii o automatizácii odôvodnenia najzaujímavejší.<sup>10</sup>

<sup>3</sup> Pozri napr. MOLEK, P. *Právo na spravodlivý proces*. Praha: Wolters Kluwer, 2012, s. 274–281.

<sup>4</sup> S. ř. spr. hlavne však preberá podmienky stanovené o. s. ř. pre odôvodnenia, a to v § 64. Pozri napr. KÜHN, Z. et al. *Soudní řád správní. Komentář*. Praha: Wolters Kluwer ČR, 2019, s. 508–510.

<sup>5</sup> KMEC, J. et al. *Evropská úmluva o lidských právech: komentář*. Praha: C. H. Beck, 2012, s. 757–762.

<sup>6</sup> Rozsudok č. 35289/11 vo veci *Tatishvili proti Rusku* zo dňa 22. 2. 2007.

<sup>7</sup> MOLEK, P. *Právo na spravodlivý proces*, s. 276.

<sup>8</sup> Tu zrovna s národnou úpravou a Nálezmi ÚS napr. rozhodnutie vo veci *Perez proti Francúzsku*, najmä § 81, kedy ESLP tvrdí, že od súdu nemôžeme očakávať že sa vyjadrí ku každému vznesenému bodu.

<sup>9</sup> LASSER, M. *Judicial deliberations – A comparative analysis od judicial transparency nad legitimacy*. Oxford University Press, 2014, s. 156.

<sup>10</sup> Lasser dokonca francúzskych sudcov označuje za „stroje“, ktoré mechanicky uplatňujú platné právo a odôvodnenie je často konštruované formou „Zákon hovorí... a preto...“. Tento efekt potom pripisuje najmä silnej formalistickej tradícii vo francúzskom súdnictve, kedy sudca „jednoducho aplikuje“ relevantné číslo ustanovenia *Code civil* „pasívne, mechanicky, až matematicky“. V Lesserovom ponímaní potom tento text nielenže „sám“ generuje rozhodnutie, respektíve priamo z jeho textácie vychádza záver, ale zároveň tým aj určuje štruktúru rozhodnutia (odôvodnenia). Tu je potom zaujímavé sledovať široké možnosti automatizácie sudcovského rozhodovania ako takého, kedy je v súčasnosti za „strojový“ či „automatický“ považovaný striktno formalistický prístup, napr. Francúzov, ktorý sa zásadne vyhýbajú akémukoľvek precedentnému rozhodovaniu, ktoré je založené výhradne na presných pravidlách. Oproti tomu sa technicky javia byť vhodným kandidátom do budúcnosti prístupy strojového učenia, ktoré vychádzajú z podstatnej miery práve z predchádzajúcich rozhodnutí.

Z vyššie uvedených právnych predpisov vyplýva niekoľko základných podmienok, ktoré by malo odôvodnenie rozhodnutia spĺňať. Za prvé by sa malo jednať o určité zhrnutie žaloby, teda čoho a na základe akých dôvodov sa žalobca domáhal, súčasťou čoho by mala byť aj replika. Následne by malo byť v rozhodnutí uvedené aké skutočnosti boli považované za preukázané, dôkazy, ktorými sa súd pri svojom rozhodnutí riadil ako aj opis týchto úvah ktoré viedli k rozhodnutiu, spolu s odôvodnením vynechania niektorého z dôkazov, a v poslednej rade aj záver, ktorý z toho vyvodil. I keď s drobnými variáciami pri rôznych druhoch rozhodnutí či konaní, toto možno považovať za obecnú kostru obsahu odôvodnenia.

Tieto zákonom vymedzené základné podmienky potom svojím výkladom rozširuje, či ozrejmuje Ústavný súd vo svojej judikatúre. Môžeme sa tak dozvedieť, že porušením práva na spravodlivý proces, respektíve nedostatočným odôvodnením, je aj také odôvodnenie, ktoré sa nevypradálo so všetkými vznesenými námietkami, respektíve ako tvrdí Ústavný súd v Náleze I. ÚS 2449/20 zo dňa 9. 2. 2021, súd sa musí (riadne) vysporiadať so všetkými *relevantnými* námietkami, vznesenými v priebehu procesu, a to s každou z nich práve v takom rozsahu v akom sú relevantné.<sup>11</sup> Obdobne sa k tomu postavil aj v Náleze II. ÚS 2522/19 zo dňa 10. 12. 2020, kde Ústavný súd (skutočne len obdobne) tvrdí, že porušením práva na spravodlivý proces je explicitné nereagovanie na také tvrdenie, ktoré si to pre svoju rozhodnú povahu pre proces vyžaduje.<sup>12</sup>

Oba tieto Nálezy sú pre túto prácu zvlášť zaujímavé, a to jednako z dôvodu, že sa opierajú o judikatúru Európskeho Súdu pre Ľudské Práva („ESLP“), a tiež preto, že oba spomínajú aj automatizované rozhodovanie.<sup>13</sup> Odkaz na judikatúru je tak trochu dvojsečná zbraň, na jednej strane nám totiž poskytuje ďalšie vhlad do tejto problematiky, na druhej strane to tiež znamená, že sa od Ústavného súdu nedočkáme žiadnej argumentácie, či aspoň odôvodnenia aplikácie predmetnej judikatúry, ktorá bude rozobratá v ďalšej časti článku.

## 1.1 Judikatúra ESLP a automatizácia

Prvým z rozsudkov ESLP, na ktoré sa Ústavný súd pri svojom opakovaní mantry „*Porušením práva na spravodlivý proces môže byť zejména stereotypní či automatizované odůvodnění [...]*“<sup>14</sup> odkazuje, je rozsudok vo veci *Paradiso a Campanelli proti Taliansku*. Rozsudok sa venuje možnosti cezhraničného osvojovania detí a svojím obsahom obecné nie je nejak zvlášť podstatný, či zaujímavý, pre otázky odôvodňovania súdnych rozhodnutí. Pre túto prácu, a zjavne aj pre Ústavný súd, je tento rozsudok zaujímavý najmä v odstavci 210 (a souv.). V ňom práve súd hovorí, že k porušeniu práv nedošlo, nakoľko odôvodnenie zá-sahu do práv nebolo vykonané stereotypne či automaticky, respektíve bola zvážená individuálna situácia dieťaťa, o ktorom bolo rozhodované. Je zvláštne, že sa Ústavný súd tak chytil práve tohto odstavca v predmetnom rozsudku, nakoľko sa celý rozsudok venuje

<sup>11</sup> Nález I. ÚS 2449/20 zo dňa 9. 2. 2021, najmä potom § 31.

<sup>12</sup> Nález II. ÚS 2522/19 zo dňa 10. 12. 2020, najmä potom § 30.

<sup>13</sup> Ústavný súd má samozrejme mnoho viac Nálezov zaoberajúcich sa odôvodňovaním rozhodnutia ako súčasť práva na spravodlivý proces, ako napríklad Nález III. ÚS 84/94 zo dňa 20. 6. 1995, kde právo opisuje vzťah týchto dvoch práv, respektíve pripisuje právu na odôvodnenie úlohu poistky voči sudcovskej svojvoľe, takáto analýza, akokoľvek zaujímavá, však neprináša nič do užšej problematiky tejto práce, Nálezy spomínané v tejto práci boli teda volené podobne, akoby súdy mali voliť body ku ktorým sa vyjadrujú – podľa kritéria relevantnosti.

<sup>14</sup> Pozri napr. I. ÚS 2449/20 zo dňa 9. 2. 2021 § 34.

úplne inej záležitosti, do určitej miery to ani nie je hlavná otázka toho odstavca, no najmä je z autorovho pohľadu problematická nejasnosť významu „stereotypizovaného a automatického“ rozhodovania. Rozsudok totiž používa termín „automatic“ a nie „automated“, ktoré sa bežne používa s algoritmickými rozhodovacími systémami, a hlavne potom z kontextu rozsudku, ako aj z kontextu druhého odkazovaného rozsudku *ESLP X proti Litve*, na ktoré sa odkazuje aj odstavec 210, vyvstáva otázka, či pod pojmom automatický nebolo myslené také konanie, ktoré prebieha bez rozmyslu, jednoduchým prepisovaním nejakého vzoru. Nad to oba rozsudky považujú takéto rozhodovanie za potencionálne problematické výhradne pri rozhodovaní o otázkach osobného štatútu detí.

Z autorovho pohľadu je argumentácia týmto ustanovením ako odmietnutím automatizovaného odôvodňovania problematická, nakoľko je veľmi nejasné, prakticky neodôvodnené a špecifické pre zvlášť zraniteľnú skupinu. Zaujímavými sú však dovetky v týchto odstavcoch, ktoré nám často poskytujú nejakú bližšiu informáciu o obsahu odôvodnenia ako takého – teda povinnosť zvážiť a vysporiadať sa so špecifickou individuálnou situáciou, v dostatočne podrobnej miere. Odhliadnuc teda od možnosti neprípustnosti automatizovaného odôvodnenia v určitých prípadoch, bolo by vhodné sa ďalej pozrieť na reálne možnosti konštruovania odôvodnení nejakým automatizovaným systémom, a to najmä s ohľadom na technické limitácie, akými je najmä problém algoritmickej čiernej skrinky („*algorithmic black box*“).

## 2. Čierna skrinka

Jednou z častých námietok, ktoré sa vyskytujú pri algoritmickej (nielen) rozhodovacej procesov, je problém tzv. čiernej skrinky. Tento problém, respektíve stav, referuje k použití takého algoritmu, ktorého vnútorná logika je skrytá používateľovi, a v podstate aj jeho tvorcovi. Takýto stav veľmi často nastáva pri využití algoritmov strojového učenia, nakoľko programátorom je pri mnohých aplikáciách programovaný „učiaci“ algoritmus, nie výsledný rozhodovateľ.<sup>15</sup> Takýto stav čiernej skrinky býva vo výskume umelej inteligencie často označovaný nielen ako praktický problém, ale aj etický.<sup>16</sup> Jednou z otázok tejto práce je, či ho bude nutné označiť aj za právny problém.

Na prvý pohľad sa môže zdať, že problém algoritmickej čiernej skrinky vytvára absolútnu prekážku automatizovanému odôvodňovaniu, nakoľko z pohľadu človeka tu jednoducho nie je čo odôvodniť, poťažmo potom by táto situácia mala vytvárať absolútnu prekážku pre automatizované odôvodňovanie ako také, nakoľko nám tento stav zabráňuje poznať jeho logiku, čo je z pohľadu rozoberaného práva na odôvodnenie vo väčšine prípadov<sup>17</sup> neprípustný stav.

V skutočnosti by sa však nemuselo jednať až o tak páľčivý problém, jednako preto, že existuje rozsiahla odborná diskusia, s už v tomto momente obstojnými výsledkami, o tom,

<sup>15</sup> Ako poukazuje Dymitruk, modely ktoré najspoľahlivejšie reprodujú sudcovské rozhodovanie sú tie, ktoré sú prakticky nevyvetliteľné a naopak. Pozri: DYMITRUK, M. The Right to a Fair Trial in Automated Civil Proceedings. *Masaryk University Journal of Law and Technology*. 2019, č. 1.

<sup>16</sup> GUIDOTTI, R. et al. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*. 2018, č. 5, s. 93:1–93:42.

<sup>17</sup> Predchádzajúca kapitola opomenula situácie kedy je prípustné rozhodnutie neodôvodniť, no to je prípustné práve z dôvodu, že odôvodňovanie je v takých situáciách neopodstatnené, či proti logike veci, teda sa na ne rozoberané právo v podstate nevzťahuje. Pozri: MOLEK, P. *Právo na spravедlivý proces*. Praha: Wolters Kluwer, 2012, s. 274–281.

ako tento problém odstrániť, prípadne preto, že je vôbec otázne či tento problém vôbec existuje.<sup>18</sup>

Ak by sme chceli hľadať ekvivalenciu tohto právneho problému na poli počítačovej vedy, museli by sme sa pozrieť na výskum v oblasti tzv. vysvetliteľnej AI („*explainable AI*“). Táto oblasť však nerieši len odôvodnenie, zaoberá sa viac transparentnosťou a interpretovateľnosťou modelov a ich výsledkov, a práve pokroky v oblasti interpretovateľnosti by mohli vyriešiť aj (právnu) otázku odôvodnenia. Mittelstadt et al. identifikujú niekoľko takýchto modelov, ktoré by mali byť schopné vytvárať *post hoc*<sup>19</sup> odôvodnenie rozhodnutia vykonaného algoritmom, ktorý by bol inak označený za čiernu skrinku.<sup>20</sup> V zásade by sa malo jednať o taký spätné propagovaný proces, ktorý by prideliť váhu jednotlivým uzlom, ktoré tvoria (neurónovú) sieť, teda na konci tohto procesu môžeme vidieť, ktoré body v akej miere zavážili pri formovaní finálneho rozhodnutia. Takúto analýzu potom možno vnímať vo svetle zákonného požiadavku na rozhodnutie ako na vysvetlenie, respektíve určenie toho, ktorými dôkazmi sa súd riadil a ktoré skutočnosti považoval za preukázané.

Takýto postup sa zdá byť ako ten najvhodnejší práve s ohľadom na zákonné požiadavky na odôvodnené rozhodnutie, no nie je jediný. Skupina odborníkov na vysokej úrovni pre umelú inteligenciu zriadená Európskou komisiou<sup>21</sup> vydala v rámci svojej činnosti Pokyny k dôveryhodnosti AI,<sup>22</sup> v ktorých uznáva, že niečo ako čierna skrinka existuje a treba s tým náležite narábať. Skupina je toho názoru, že v závislosti na kontexte rozhodnutia by mala postačovať aj len jednoduchá vedomosť mechanizmu, na ktorom algoritmus spočíva. Ako sama Skupina ďalej uvádza, čiernu skrinku často vytvára práve to, že nie sme schopní poznať ani ten, eventuálne takéto odôvodnenie je zrejme nedostačujúce s ohľadom na právo na spravodlivý proces, respektíve v kontexte rozhodovania o právach a povinnostiach jedincov.

Ako druhý možný spôsob mitigácie dopadu čiernej skrinky potom navrhuje „*rule of law by design*“, teda nastaviť systém, nie len algoritmického rozhodovania, ale obecne kontroly a narábania s rozhodnutím, takým spôsobom, aby nasledoval podmienky, ktoré by na takéto rozhodnutie kládli zásady vlády práva. V našom prípade by to ale znamenalo vyriešiť problém čiernej skrinky algoritmu tým, že by nikdy nevznikla, teda nastavíť

<sup>18</sup> Na jednej strane je nutné poznamenať, že existujú rozhodovacie algoritmy, ktoré nevedú k vytvoreniu čiernej skrinky, tie sú však značne simplistické a skôr nevhodné pre naše účely, a tiež preto, že časť odborníkov zaoberajúcich sa vysvetliteľnosťou algoritmov (m. j. označované ako „*Explainable AI*“) často opakuje názor, že čierna skrinka vzniká vtedy, keď programátor „nedonútil“ svoj program urobiť všetko čo má.

<sup>19</sup> Tu je nutné zdôrazniť, že je rozdiel medzi ľudským *post hoc* odôvodňovaním a tým algoritmickým. U ľudí tento proces prebieha ako formulácia rozhodnutia a následné nasadenie argumentov na toto rozhodnutie, miesto toho aby daného človeka k rozhodnutiu viedli práve tieto argumenty. U algoritmov je *post hoc* odôvodnenie skôr akýmsi „vytiahnutím“ týchto argumentov zo systému, teda existujú ešte pred vytvorením odôvodnenia. V tomto ohľade sa dá tvrdiť, že je takéto rozhodnutie ešte transparentnejšie než to ľudské, ktoré často funguje v opačnom poradí, prípadne je odôvodnenie písané inou osobou (asistentom) ako rozhodnutie. K otázkam prehadzovania poradia rozhodnutia a odôvodnenia pozri napr. aj ŽUFFOVÁ-KUNČOVÁ, T. Bylo dříve vejce nebo slepice? Lze připustit, aby soudce o případu nejdříve rozhodl, a až poté hledal pro své rozhodnutí oporu v právu? In: *Iurium* [online]. 15. 1. 2022 [cit 2022-06-09]. Dostupné zo: <<https://www.iurium.cz/2022/01/15/bylo-drive-vejce-nebo-slepice-lze-pripustit-aby-soudce-o-pripadu-nejdříve-rozhodl-a-az-pote-hledal-pro-sve-rozhodnutí-oporu-v-právu/>>.

<sup>20</sup> MITTELSTADT, B. – RUSSELL, C. – WACHTER, S. Explaining Explanations in AI. *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 2019.

<sup>21</sup> Pozri: <<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>>.

<sup>22</sup> EURÓPSKA KOMISIA. High-level expert group on artificial intelligence. Ethics guidelines for trustworthy AI. In: *ec.europa.eu* [online]. 2019. Dostupné zo: <<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>>.

rozhodovací proces tak aby bola zachovaná kontrola, respektíve nejaký ľudský dohľad. To však s ohľadom na spomínané technické limitácie jednoduchých rozhodovacích algoritmov v súčasnosti nie je schodná cesta.

## 2.1 Právo na porozumenie

Ako Mittelstadt rozoberá v úvode do spomínanej práce, rozsah pojmu vysvetliteľná AI zahŕňa viacero aspektov transparentnosti a pochopiteľnosti algoritmickeho rozhodovania, jednou z nich je aj práve porozumiteľnosť (štruktúry algoritmu).<sup>23</sup> Toto je prvý z odkazov na ustanovenia GDPR ako možnej inšpirácie, respektíve ako doposiaľ v zásade jedinej existujúcej právnej úpravy, ktorá sa aspoň okrajovo dotýka problematiky tejto práce a mimo iné tiež uvádza právo na porozumenie vnútornej logiky procesu, ktorý o dátovom subjekte automaticky rozhodoval, a preto sa budeme týmto podobnosťami venovať v ďalšej časti. Ďalším existujúcim právnym predpisom, ktorý to vidí zhodne s Mittelstadtom, je aj kanadské Nariadenia o automatizovanom rozhodovaní<sup>24</sup>, ktoré potrebu porozumiteľnosti rozhodovacieho procesu zdôrazňuje najmä s ohľadom na spätnú kontrolu funkčnosti systému a správnosti algoritmu, porozumiteľnosť teda v tomto prípade nie je adresovaná osobe, o ktorej bolo rozhodované ale orgánu dozoru.

Táto úvaha nás však privádza k ďalšej komplikácii automatizovaného odôvodňovania rozhodnutí, a to k otázke jeho formy. Ostatne, ako uvádza Mittelstadt, vysvetliteľná AI by mala mať aj sociologicko-filozofické aspekty, medzi ktoré sa radí aj otázka, či je produkovateľné vysvetlenie algoritmickej čiernej skrinky „užitočné“ pre adresáta. Obecne potom je potrebné mať na akékoľvek automatizované vysvetlenie požiadavku toho, aby bolo dostatočne porozumiteľné.<sup>25</sup> Čo do konkrétnej formy odôvodnenia súdneho rozhodnutia nám zákon ani komentárová literatúra toho veľmi nezdelujú a v zásade sa obmedzujú len na výpočet bodov povinného obsahu, ktorý bol už uvedený v úvodnej časti.<sup>26</sup> Bližšie informácie sa potom nevieme dozvedieť ani na nadnárodnej, respektíve európskej úrovni, kedy sa ESLP vo svojich nálezhoch obdobne obmedzuje na výpočet jednotlivých bodov, ktoré v odôvodnení musia zaznieť. Z pohľadu ESLP je to najmä reflexia *hlavného* argumentu, špecifických a dôležitých bodov, a takých bodov, ktoré sa odkazujú na Dohovor.<sup>27</sup> Súd už však bližšie nešpecifikuje, ako sa má s týmito bodmi narábať. Ako poukazuje Kühn, prísť s nejakým štandardom kvality u súdnych odôvodnení je prakticky nemožné, nakoľko štýl a spôsob odôvodňovania rozhodnutí sa líši nielen naprieč odvetviami súdnictva, ale aj súdmi, či dokonca na úrovni jednotlivých sudcov.<sup>28</sup> Jediným, a neúspešným, pokusom o štandardizáciu kvality sudcovského odôvodňovania mal byť návrh kódexu disciplinárnych priestupkov, ktorý za nedostatočne odôvodnené rozhodnutie považoval také, ktoré

<sup>23</sup> MITTELSTADT, B. – RUSSELL, C. – WACHTER, S. *Explaining Explanation in AI*.

<sup>24</sup> Pozri: <<https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>>.

<sup>25</sup> Mnohé aproximatívne modely algoritmickeho vysvetlenia majú totiž ako svoj výstup len váhu jednotlivých uzlov v rozhodovacej sieti, čo z pohľadu bežného občana možno označiť za skôr nepostačujúce odôvodnenie.

<sup>26</sup> Pozri napr. SCHELLEOVÁ, I. a kol. *Civilní Proces*. EUROLEX-BOHEMIA, 2016, s. 564.

<sup>27</sup> Pozri napr. rozsudok č. 11161/08 vo veci *Mont Blanc Trading Ltd a Antares Titanium Trading Ltd proti Ukrajine*, najmä § 82 a 84, alebo rozsudok č. 76240/01 vo veci *Wagner a J. M. W. L. proti Luxembursku*, najmä § 96.

<sup>28</sup> To Kühn vidí ako dôvod, prečo sa kontrola činnosti súdov obmedzuje na príliš dlho trvajúce rozhodovanie („pomalú prácu“) a počet rozhodnutí zmenených druhostupňovou inštanciou. KÜHN, Z. *The Quality of Justice and of Judicial Reasoning in the Czech Republic*. In: BENCZE, M. – NG, Y. (eds). *How to Measure the Quality of Judicial Reasoning*. Springer. 2018, s. 174–176.

porušovalo platné právo alebo jeho zaužívanú interpretáciu, nepoužívalo racionálnu argumentáciu alebo bolo arbitrárne, či porušovalo *nároky na kvalitné rozhodovanie*.<sup>29</sup>

Takéto bodové vymedzenie obsahu nás potom logicky navádza na úvahy o možnostiach formulárového odôvodnenia. To by z pohľadu automatického generovania textu malo byť tou najjednoduchšou možnosťou, nakoľko by len na vopred predurčená miesta dopĺňalo predpokladané hodnoty, zrejme v podobe bodov, ktoré zákon predpokladá ako obsah rozhodnutia, teda mimo iné výpočet skutočností, ktoré boli pre rozhodnutie považované za preukázané, a dôkazy, ktorými sa riadil, prípadne mieru, do akej pre dané rozhodnutie, ktorý konkrétny dôkaz zavážil, nakoľko práve táto hodnota je jedným z výstupov aproximatívnych vysvetľovacích modelov opísaných v predchádzajúcej časti. Otázkou však zostáva, či práve takéto „jednoduché“ doplnenie predpísaného odôvodnenia nebude konstituovať práve stereotypizované odôvodnenie, ktoré v *určitých* prípadoch môže zakladať porušenie práva na spravodlivý proces. Kvôli zvrchu opísanej pojmovej nejasnosti by sme podľa všetkého zatiaľ mohli na túto otázku povedať, že sa bude jednať o problém pri rozhodovaní o otázkach osobného štatútu mladistvých, prípadne inej zvlášť zraniteľnej skupiny obyvateľstva – jedná sa o také rozhodnutia, u ktorých ESLP, s ohľadom na dodržanie práva na spravodlivý proces, v spomínanej judikatúre kladie dôraz na konkrétne zváženie a zhodnotenie každej jednej situácie. Možno sa domnievať, že ESLP volí takéto prípady práve pre potrebu ich silnej individualizácie, teda povinnosť sudcu priamo sa oboznámiť a zvážiť buďto želanie a názor (ako mimo iné ukladá napríklad aj článok 12 Dohovoru o právach dieťaťa), prípadne relevantný stav danej osoby (ako napríklad pri obmedzení svojprávnosti), a ako je uvedené v *Paradiso a Campanelli*, rozhodnutie musí reflektovať konkrétne dopady týchto pozorovaní na rozhodnutie.

Pri úvahách o prípustnosti formulárového odôvodňovania pri iných druhoch rozhodnutí sa ponúka trochu cynická otázka, akým spôsobom sa pracuje s textami rozhodnutí v súčasnosti a či už teraz neexistuje niečo ako formulárový rozsudok, ktorý vzniká prepisovaním relevantných údajov v predchádzajúcich rozhodnutiach rovnakého druhu. Sú to potom najmä rozhodnutia vo veciach, v ktorých nie je potrebné nariadiť ústne konanie, ktoré sú pre svoju jednoduchosť a strohosť prakticky identické.

Druhá varianta je založená na predpoklade, že pre absenciu akýchkoľvek sťažností, či náleзов, môžeme predpokladať, že až na nejaké odchýlky je súčasný stav odôvodnení vyhovujúci, alebo minimálne v zhode s minimálnymi požiadavkami práva na spravodlivý proces. Je teda možné hľadať také automatizované technické riešenie, ktoré by v určitej miere kopirovalo stávajúce odôvodnenia. Už z takéhoto opisu sa núka riešenie založené na strojovom učení, ktoré by vychádzalo z existujúceho rozhodnutia, respektíve by ho ako dataset využilo. Vhodným kandidátom je potom systém založený na jazykovom modeli GPT-3.<sup>30</sup> Tento systém generuje text na základe parametrov (čo do jeho obsahu) a kopíruje pritom štylistiku vzorových textov. Pri správnom nastavení je teda možné si predstaviť automatizované generovanie originálneho súvislého textu, ktorý nahrádza „ľudské“ odôvodnenie rozhodnutia.<sup>31</sup>

<sup>29</sup> Ibidem, s. 174–185.

<sup>30</sup> Pozri: <<https://openai.com/blog/openai-api/>>.

<sup>31</sup> K schopnosti GPT-3 generovať dlhé súvislé texty pozri článok v *The Guardian*, ktorý ním bol vygenerovaný: GPT-3. A robot wrote this entire article. Are you scared yet, human? In: *the Guardian* [online]. 8. 9. 2020 [cit.]. Dostupné zo: <<http://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>>.

Takáto predstava nás však privádza k ďalšej limitácii, respektíve zrejme bližšie k okruhu prípadov a rozhodnutí u ktorých by bolo možné potenciálne uvažovať o automatizácii rozhodovania, respektíve aj len odôvodňovania. Stav, kedy je každé jedno rozhodnutie (alebo aj odôvodnenie) naviazané na všetky predchádzajúce, v zmysle tréningového datasetu, síce značne zvyšuje právnu istotu a konzistentnosť rozhodovacej praxe a jej predvídateľnosť, no zároveň je prakticky nemožné produkovanie rozhodnutí v „hard case“, teda prípady, ktoré sa nedajú rozhodnúť výhradne prostredníctvom jednoduchej interpretácie platného práva.<sup>32, 33</sup> Takéto prípady by potom podľa Ústavného súdu mali vo svojom odôvodnení o to podrobnejšie reflektovať hodnoty a zásady, na základe ktorých bolo rozhodnuté,<sup>34</sup> musí sa teda prakticky jednať o jedinečné rozhodnutie, ktoré nie je generovateľné. Minimálne takéto prípady predstavujú absolútnu limitáciu automatizovaného odôvodnenia.

### 3. Otázka porozumiteľnosti a súčasná právna úprava

Táto práca sa zaoberá teoretickou situáciou, ktorej realizácia by si zrejme vyžiadala vlastnú právnu úpravu. Najbližšie k úprave automatizovaného rozhodovania, a teda aj jeho odôvodňovania, má Všeobecné nariadenie o ochrane údajov (ďalej len „GDPR“). Článok 22 hovoriaci o automatizovanom rozhodovaní, respektíve 13–15 ustanovujúce informačnú povinnosť, tohto nariadenia zakladá subjektom pri automatizovanom rozhodovaní (profilovaní) právo na jednako informácie o tom, že k tomu došlo, a čo je pre nás zaujímavejšie, nárok na poznanie „logiky“ rozhodovania, a ako vyplýva m. j. z recitálov, mať nárok na následné odôvodnenie rozhodnutia.

V kontexte týchto ustanovení potom vznikla diskusia o tom, či je týmto založené nové právo na vysvetlenie,<sup>35</sup> a ak áno, čo je jeho obsahom, prípadne či je *ex ante* informácia ekvivalentnom *ex post* vysvetlenia.<sup>36</sup> V kontexte rozhodovania moci verejnej o právach a povinnostiach samozrejme takéto nie je dostačujúce,<sup>37</sup> tieto ustanovenia sú zaujímavé

<sup>32</sup> *Hard cases* a ich rozhodovanie v Dworkinovskom ponímaní, kedy Dworkin v diele *Taking Rights Seriously* uvádza tézu, že existuje set prípadov, ktoré nejdú rozhodnúť výhradne pozitivistickou interpretáciou práva ale je nutná nejaká miera diskrecie. Pozri: DWORKIN, R. *Taking Rights Seriously*. Harvard University Press, 1978, s. 81–130.

<sup>33</sup> Využívanie historických datasetov a s nimi spojených štatistických metód (strojového učenia) so sebou prináša celý rad ďalších komplikácií pre právo, respektíve spravodlivosť, ako napríklad dopad na otvorenú štruktúru práva a jeho ďalší vývoj, či problém „znovu nalézání“ práva v každom novom prípade ako poistku individualizácie potrebnej pre rozdielnosť konceptov práva a spravodlivosti, k tomu pozri: DERRIDA, J. *Síla zákona*. Oikoymenth, 2002, s. 21–28. Tieto a ďalšie súvisiace problémy si však zaslúžia vlastné, hlbšie pojednanie.

<sup>34</sup> Pozri nálež I. ÚS 403/03 z 17. 8. 2005.

<sup>35</sup> Zrovňaj GOODMAN, B. – FLAXMAN, S. European Union regulations on algorithmic decision-making and a “right to explanation”. *AI Magazine*. 2017, č. 3, s. 50–57, ktorý ako prvý argumentovali jeho existenciu s opačným názorom mimo iné aj Mittelstadta v: WACHTER, S. – MITTELSTADT, B. – FLORIDI, L. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*. 2017, Vol. 7, Iss. 2, s. 76.

<sup>36</sup> MALGIERI, G. Automated decision-making in the EU Member States: The right to explanation and other “suitable safeguards” in the national legislations. *Computer Law & Security Review*. 2019, č. 5, s. 105327. Dostupné zo: <<https://www.sciencedirect.com/science/article/pii/S0267364918303753>>.

<sup>37</sup> Tu pozri napr. URÍČAŘ, M. et al. *Obecné nařízení o ochraně osobních údajů: komentář*. Praha: C. H. Beck, 2021, s. 608, kde je dovedené, že správca musí subjektu poskytnúť okrem iného informácie, že k takému druhu rozhodovania dochádza a poskytnúť zmysluplné informácie o použitom postupe, význame a predpokladaných dôsledkoch takého rozhodovania. Tieto podmienky sú s ohľadom na vyššie uvedené nároky pre odôvodnené rozhodnutie zjavne nedostačujúce a príliš abstraktné, než aby mohli byť považované za odôvodnenie v kontexte verejnoprávneho rozhodnutia o právach a povinnostiach.



najmä z pohľadu otázky, čo to znamená vysvetliť a či je nejaký rozdiel oproti pojmu pochopiť, čo do formálnej stránky odôvodnenia. V prvom rade je nutné podotknúť, že už samotná požiadavka vyslovená v článkoch 13–15, teda poznať vnútornú logiku, ide proti konceptu algoritmickej čiernej skrinky, odhliadnuc od toho, nám ďalej hovorí, že táto logika musí byť poznateľná „*zmysluplným spôsobom*“, <sup>38</sup> čo nás vracia do rozoberaného problému práhu pochopiteľnosti, nakoľko ani Pracovná skupina 29 vo svojich Pokynoch tento pojem bližšie neidentifikuje, hovorí len, že sa tým nemyslí odhalenie algoritmu ako takého, prípadne že v závislosti od kontextu by sa mohlo jednať o zdelenie rozhodovacích kritérií. Jarolímková v komentári uvádza, že správca nie je povinný vysvetľovať subjektu (rozhodovaci) algoritmus dopodrobna, či mu ho celý odhaliť, z čoho potom dovodzuje aj „riešenie“ pre problém čiernej skrinky, teda že nakoľko neexistuje povinnosť komplexného vysvetlenia nie je problém, že správca v podstate ani takéto vysvetlenie nikdy podať nemôže. <sup>39</sup> Tento nedostatok potom, ako uvádza v komentári, je možné preklenúť nejakou formou certifikácie („*informace o pravidelném testování*“), <sup>40</sup> pri probléme čiernej skrinky, teda inherentne netransparentného systému, sa núka otázka, čo vlastne má takéto certifikácia osvedčovať.

V otázke pochopiteľnosti, respektíve vhodnej formy podávaného vysvetlenia Jarolímková, respektíve WP 260 rev.01 uvádzajú, že informácie musia byť poskytnuté podľa zásady transparentnosti, teda stručne, zrozumiteľne, ľahko prístupným spôsobom za použitia jasných a jednoduchých jazykových prostriedkov. <sup>41</sup>

Určitý náznak, že medzi právom na vysvetlenie a právom na pochopenie bude nejaký rozdiel, potom plynie z národných zapracovaní GDPR, ktoré takéto právo zavádzajú. Konkrétne sa jedná o Maďarsko a Francúzsko, <sup>42</sup> ktoré zvolili podľa Malgieriho tzv. proaktívny prístup. <sup>43</sup> V porovnaní s ostatnými národnými úpravami, respektíve aj samotným Nariadením ako takým nám poskytujú informácie o tom, kedy je logika algoritmickeho rozhodnutia pochopiteľná, nakoľko implementácie priamo stanovujú, že subjektu musia byť známe nielen jednotlivé faktory rozhodovania ale aj ich váha a to v „pochopiteľnej“ forme. Takýto opis je prakticky opisom proximatívnych modelov vysvetliteľnej AI opísaných v tretej kapitole, takýto názor, respektíve takéto možnosť implementácie je však v rozpore s nálezom francúzskeho ústavného súdu, ktorý za neprípustnú považuje implementáciu takéto rozhodovacieho algoritmu, u ktorého pri vývoji každej jednotlivéj časti nehral hlavnú úlohu človek. <sup>44</sup>

Táto krátka exkurzia mala dva ciele, jednako priblížiť spôsob, akým k tejto problematike pristupuje zatiaľ prakticky jediná úprava, aspoň nejakej formy, automatizovaného

<sup>38</sup> DATA PROTECTION WORKING PARTY 29. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: *ec.europa.eu* [online]. Dostupné zo: <<https://ec.europa.eu/newsroom/article29/items/612053>>.

<sup>39</sup> URÍČAŘ, M. et al. *Obecné nařízení o ochraně osobních údajů: komentář*, s. 608–609.

<sup>40</sup> *Ibidem*, s. 609.

<sup>41</sup> *Ibidem*.

<sup>42</sup> V kontexte súdnej automatizácie (í keď už priamo rozhodovania a nie odôvodnení) volí Francúzsko obdobie prístup ako ESLLP, teda zhladáva ho neprípustné v otázkach osobnostného štatútu. Art. 10(1), Loi n° 78-17 du 6 janvier 1978. Dostupné zo: <<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037085952>>.

<sup>43</sup> MALGIERI, G. *Automated decision-making in the EU Member States: The right to explanation and other “suitable safeguards” in the national legislations*.

<sup>44</sup> CONSEIL CONSTITUTIONNEL. *Décision n° 2018-765 DC du 12 juin 2018*. Dostupné zo: <<https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm>>.

rozhodovania, a po druhé, aj v nadväznosti na nedostačujúci charakter tejto úpravy, poukázať na komplexnosť tejto problematiky,<sup>45</sup> ktorú s ohľadom na rozoberaný druh verejnoprávnych rozhodnutí nemožno odbiť len jednoduchým odstavcom.

## Záver

Vyslovenie záveru, či je z pohľadu práva na spravodlivý proces možné automaticky generovať odôvodnenia rozsudkov, v zásade stojí na zodpovedaní dvoch podotázok. Tou prvou je otázka, čo všetko by malo materiálne odôvodnenie splňovať na to, aby bolo v súlade s právom na odôvodnené rozhodnutie,<sup>46</sup> a tiež aké musí naplňovať formálne požiadavky, teda napríklad či sa môže jednať o formulárové odôvodňovanie. V nadväznosti nato je potom potrebné si položiť otázku technického charakteru, teda či existuje také algoritmicke riešenie, ktoré by týmto požiadavkám vyhovovalo a zároveň neprinášalo ďalšie komplikácie, teda práve typicky nejaký druh nevysvetliteľnosti algoritmickej činnosti, ako upozorňujú Slebst a Powles, jedná sa však skutočne o netriviálny technický aspekt tohto práva.<sup>47</sup>

Nezávislou otázkou sú možné limitácie takéhoto „základného“ modelu. Tou prvou, ku ktorej máme už teraz, i keď mierne nejednoznačnú, judikatúru, je problém rozhodovania o osobnostných otázkach mladistvých, čo je predmet, ktorý zrejme možno zobecniť na takéto rozhodnutia u zvlášť zraniteľných osôb, u ktorých je vyžadovaná taká miera individualizácie rozhodnutia a priameho sa oboznamovania s konkrétnymi, osobnými aspektami situácie, že ju technicky nemožno zaručiť v takej miere, aby nedochádzalo k porušeniu práva na spravodlivý proces či iných práv relevantných v danej situácii. Ďalšou predpokladanou limitáciou bude potom odôvodňovanie rozhodnutí, ktoré si vyžadujú originálnu, v zásade mimoprávnu (teda hodnotovú, či princípovú), a rozsiahlu argumentáciu, teda v prípade *hard cases*.

Na opačnej strane týchto limitácií potom stoja rozhodnutia, o ktorých možno rozhodnúť bez nariadenia ústneho jednania, ktoré sú pre svoju jednoduchosť a repetitívnosť danú vecou samou, ktorá sudcovi neposkytuje priestor pre inováciu, kreativnosť a vo veľkej miere ani pre vlastné uváženie, ideálnym adeptom na takúto automatizáciu.<sup>48</sup>

<sup>45</sup> Tu mimo iné napríklad aj s ohľadom na rozdielnosť úrovne na ktorej sa vysvetlenie poskytuje podľa GDPR, kde stačí vysvetliť spôsob akým bude rozhodnutie vykonané a úpravu napr. o. s. r., kde je nutné vysvetliť spôsob, akým bolo dospeť ku konkrétnemu rozhodnutiu.

<sup>46</sup> Poľažmo s tým potom aj m. j. s Vládou práva, pozri: VENICE COMMISSION. *Rule of Law Checklist*. Strassbourg: COE, 2016. Dostupné zo: <[https://www.venice.coe.int/images/SITE%20IMAGES/Publications/Rule\\_of\\_Law\\_Check\\_List.pdf](https://www.venice.coe.int/images/SITE%20IMAGES/Publications/Rule_of_Law_Check_List.pdf)>.

<sup>47</sup> SELBST, A. D. – POWLES, J. Meaningful information and the right to explanation. *International Data Privacy Law*. 2017, Vol. 7, Iss. 4, s. 233–242.

<sup>48</sup> Tento článok vychádza z modelu, kedy rozhodnutie, ktoré je automatizované odôvodňované je tiež samo vynesené bez zásahu človeka. Odhliadnuc od časti, ktorá sa výslovne venuje problému *black boxu*, teda úvod časti 3, mal by tento článok slúžiť aj ako základný úvod do automatizovaného odôvodnenia rozhodnutí ľudských sudcov.

## **The Right to a Reasoned Decision and the Algorithmization of Decision-Making Systems**

Andrej Krištofík (<https://orcid.org/0000-0001-8150-0362>)

**Abstract:** One of the essential elements, not only of judicial decisions, but in fact of any public-law decision, is the statement of reasons. Such reasoning must meet certain standards, and failure to do so may lead to at least the creation of scope for appeal, and at worse to a violation of the right to a fair trial. This paper will address the question of what minimum standard exists for reasoning and in what decisions, and whether it can be met in algorithmic decision-making without human intervention, considering, among other things, the technical complication presented by the algorithmic black-box problem. The article is based on the Constitutional Court's ruling II Constitutional Court 2522/19 and the ECtHR's ruling in *Paradiso and Campinelli v. Italy* – judgments which are mutually related and which also mention automated decision-making as inadmissible, a conclusion which this article seeks to refute. Last but not least, the result will be confronted with Article 22 GDPR, or primarily with the French and Hungarian regulation, which extends the right of explanation (right of intelligibility) to the subject.

**Keywords:** automation, right to a fair trial, reasoned judgement, GDPR, algorithmic justice