učebnica

PRÁVO A UMELÁ INTELIGENCIA

Mesarčík, Gyurász a kol.

PRÁVNICKÁ FAKULTA
Univerzita Komenského
v Bratislave

Učebnica

PRÁVO A UMELÁ INTELIGENCIA

Matúš Mesarčík, Zoltán Gyurász a kolektív

PRÁVNICKÁ FAKULTA
Univerzita Komenského
v Bratislave

Vzor citácie: *Mesarčík, M., Gyurász, Z. a kol. Právo a umelá inteligencia. 1. vydanie. Bratislava: Právnická fakulta Univerzity Komenského v Bratislave, 2024.*

2024

Prvé vydanie

Recenzenti

JUDr. Ľudovít Máčaj, PhD.

Mgr. Dominika Juck, PhD.

Vydala Právnická fakulta Univerzity Komenského v Bratislave v roku 2024

ISBN: 978-80-7160-715-1

Táto učebnica je výstupom z projektu:

„Právo a umelá inteligencia" podporeným SK-NIC, a.s.

Viac informácii o projekte nájdete na

https://comeniusvyskum.flaw.uniba.sk/2023/02/16/pravo-a-umela-inteligencia/.

JEDNOTLIVÉ ČASTI SPRACOVALI

JUDr. Matúš Mesarčík, PhD., LL.M	2. kapitola (v spoluautorstve s JUDr. Zoltánom Gyurászom a JUDr. Michalom Rampáškom), 3. kapitola, 9. kapitola (v spoluautorstve s Mgr. Sandrou Žatkovou, PhD.), 10. kapitola (v spoluautorstve s JUDr. Zoltánom Gyurászom, PhD.)
JUDr. Zoltán Gyurász, PhD.	2. kapitola (v spoluautorstve s JUDr. Michalom Rampáškom a JUDr. Matúšom Mesarčíkom, PhD., LL.M), 4. kapitola, 5. kapitola, 6 kapitola (v spoluautorstve s Mgr. Janou Cihanovou), 10. kapitola (v spoluautorstve s JUDr. Matúšom Mesarčíkom, PhD., LL.M),
JUDr. Michal Rampášek	1. kapitola (v spoluautorstve s Ladislavom Rampáškom, PhD.), 2. kapitola (v spoluautorstve s JUDr. Zoltánom Gyurászom a JUDr. Matúšom Mesarčíkom, PhD., LL.M), 7. kapitola
Ladislav Rampášek, PhD.	1. kapitola (v spoluautorstve s JUDr. Michalom Rampáškom),
Mgr. Jana Cihanová	6. kapitola (v spoluautorstve s JUDr. Zoltánom Gyurászom, PhD.)
JUDr. Dominika Pintérová, PhD.	8. kapitola
Mgr. Sandra Žatková, PhD.	9. kapitola (v spoluautorstve s JUDr. Matúšom Mesarčíkom, PhD., LL.M)

ZOZNAM SKRATIEK

Agentúra ENISA znamená Agentúra Európskej únie pre kybernetickú bezpečnosť.

AI alebo UI znamená umelá inteligencia.

AIA alebo **AI Akt** znamená nariadenie (EÚ) o umelej inteligencií.

Akt o kybernetickej bezpečnosti znamená nariadenie Európskeho parlamentu a Rady (EÚ) 2019/881 zo 17. apríla 2019 o agentúre ENISA (Agentúra Európskej únie pre kybernetickú bezpečnosť) a o certifikácii kybernetickej bezpečnosti informačných a komunikačných technológií a o zrušení nariadenia (EÚ) č. 526/2013.

Dohovor znamená Dohovor Rady Európy o ochrane základných ľudských práv a slobôd.

DSA znamená 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).

EÚ znamená Európska únia.

GDPR znamená 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) (Ú. v. EÚ L 119, 4.5.2016, s. 1 – 88).

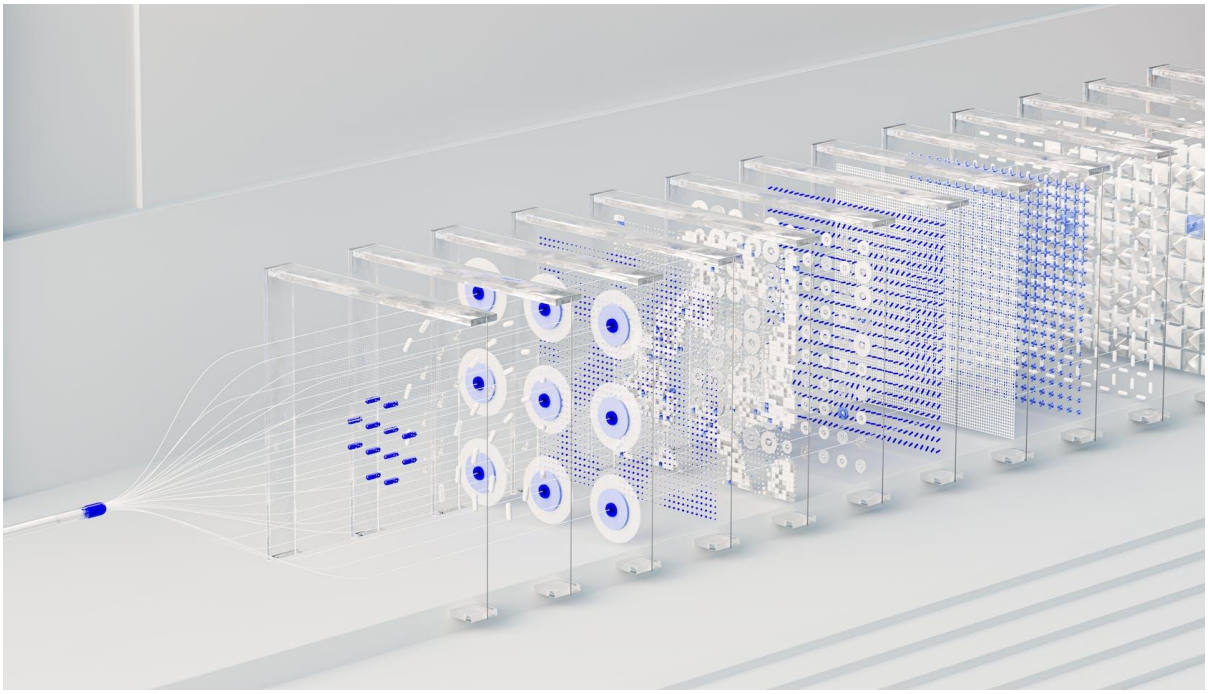
Charta znamená Charta základný práv Európskej únie.

Smernica NIS znamená smernica Európskeho parlamentu a Rady (EÚ) 2016/1148 zo 6. júla 2016 o opatreniach na zabezpečenie vysokej spoločnej úrovne bezpečnosti sietí a informačných systémov v Únii.

Smernica NIS 2 znamená smernica Európskeho parlamentu a Rady (EÚ) 2022/2555 zo 14. decembra 2022 o opatreniach na zabezpečenie vysokej spoločnej úrovne kybernetickej bezpečnosti v Únii, ktorou sa mení nariadenie (EÚ) č. 910/2014 a smernica (EÚ) 2018/1972 a zrušuje smernica (EÚ) 2016/1148 (smernica NIS 2).

SR znamená Slovenská republika.

ZoKB znamená zákon č. 69/2018 Z. z. o kybernetickej bezpečnosti a o zmene a doplnení niektorých zákonov.



Umelecká ilustrácia inšpirovaná konvolučnými neurónovými sieťami (CNN) používanými v hlbokom učení.

Autor: Novoto Studio, Visualizing AI, https://deepmind.google/discover/visualising-ai/

OBSAH

"I believe the transition we are seeing right now with AI will be the most profound in our lifetimes, far bigger than the shift to mobile or to the web before it"

Google and Alphabet CEO, Sundar Pichai, december 2023

https://blog.google/technology/ai/google-gemini-ai/#sundar-note

"We've come to this view that, in order to build the products that we want to build, we need to build for general [artificial] intelligence"

chairman and CEO of Meta, Mark Zuckerberg, január 2024

https://www.theverge.com/2024/1/18/24042354/mark-zuckerberg-meta-agi-reorg-interview

"... we view AI and machine learning as fundamental technologies and they are integral to virtually every product that we ship"

Apple CEO, Tim Cook, november 2023

https://appleinsider.com/articles/23/11/03/apple-ceo-tim-cook-calls-ai-a-fundamental-technology

"We have to take the unintended consequences of any new technology [AI] along with all the benefits, and think about them simultaneously ... Regulation [of AI] that allows us to ensure that the broad societal benefits are amplified, and the unintended consequences are dampened, is going to be the way forward"

Microsoft CEO Satya Nadella, január 2024

https://www.weforum.org/agenda/2024/01/microsoft-ceo-ai-technology-consequences/

1. Umelá inteligencia - Technologický úvod

1.1 Čo je AI?

V minulosti sa výskumníci zaoberali viacerými rôznymi verziami umelej inteligencie (*artificial intelligence*, AI). Niektorí definovali inteligenciu v zmysle podobnosti k úrovni ľudských schopností, zatiaľ čo iní uprednostňujú abstraktnú, formálnu definíciu inteligencie nazývanú *racionalita* - voľne povedané, konanie "správnych vecí". Samotný predmet sa tiež líši. Niektorí považujú inteligenciu za *vlastnosť vnútorných myšlienkových procesov a uvažovania*, zatiaľ čo iní sa zameriavajú na inteligentné *správanie*, vonkajšiu charakteristiku. [1]

Autori Russel a Norwig chápu AI za štúdium agentov (z latinského *agere,* konať), ktorí prijímajú vnemy z prostredia a vykonávajú akcie.[2]

Umelá inteligencia sa zaoberá najmä *racionálnym konaním*. Ideálny inteligentný agent koná v danej situácii najlepšie, ako sa dá. Pre každú možnú sekvenciu vnemov by mal racionálny agent zvoliť akciu, ktorá by mala maximalizovať jeho mieru výkonnosti vzhľadom na dôkazy poskytnuté sekvenciou vnemov a akékoľvek zabudované znalosti, ktoré agent má.[3]

Umelá inteligencia je definovaná ako štúdium racionálnych agentov.[4] V umelej inteligencii je agent počítačový program alebo systém, ktorý je navrhnutý tak, aby vnímal svoje prostredie, rozhodoval sa a vykonával činnosti na dosiahnutie konkrétneho cieľa alebo súboru cieľov. Agent pracuje autonómne, čo znamená, že nie je priamo riadený ľudským operátorom.

Systém umelej inteligencie sa skladá z *agenta a jeho prostredia*. Agenti pôsobia vo svojom prostredí, pričom toto prostredie môže obsahovať aj iných agentov. Agent je čokoľvek, čo vníma svoje prostredie prostredníctvom *senzorov* a pôsobenie na toto prostredie prostredníctvom *aktuátorov* (teda ovládacích členov). *Robotický agent* môže mať kamery a infračervené diaľkomery ako senzory a rôzne motory ako aktuátory. *Softvérový agent* prijíma

[1] RUSSELL, S., NORVIG, P.: *Artificial Intelligence*: *A Modern Approach* (*4th Edition*). Pearson 2020, ISBN 9780134610993
[2] Tamže, s. 669
[3] Tamže.
[4] Tamže. s. 11 - 3

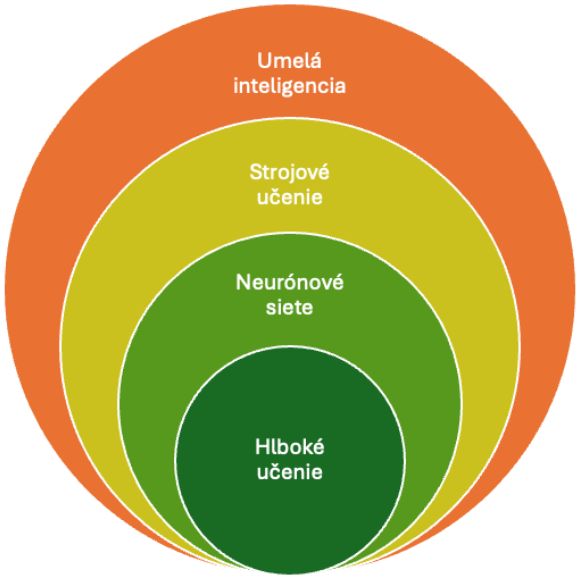
obsah súborov, sieťové pakety a ľudské vstupy (klávesnica/myš/dotyková obrazovka/hlas) ako senzorické vstupy a pôsobí na prostredie zápisom súborov, odosielaním sieťových paketov a zobrazovaním informácií alebo generovaním zvukov.

Úlohou umelej inteligencie je navrhnúť *program agenta*, ktorý implementuje funkciu agenta - mapovanie z vnemov na úkony. Predpokladáme, že tento program bude bežať na nejakom výpočtovom zariadení s fyzickými členmi a aktuátormi. Toto sa nazýva *architektúra agenta*, teda agent = architektúra + program.

V rámci AI existuje viacero študijných odborov s rôznymi pohľadmi na to, ako dosiahnuť žiadané schopnosti (program) agenta. Strojové učenie je jedným z nich. *Učenie* umožňuje agentovi pracovať v pôvodne neznámom prostredí a stať sa kompetentnejším, než by mu umožňovali samotné počiatočné znalosti. Všetci agenti tak môžu prostredníctvom učenia zlepšovať svoju výkonnosť.

1.2 Strojové učenie (*Machine Learning*)

Strojové učenie (*machine learning*, ML) stojí za chatbotmi a prediktívnym textom, aplikáciami na prekladanie jazykov, personalizovaným odporúčaním programov na Netflixe, či za tým, v akom poradí sa prezentuje obsah užívateľovi na sociálnych sieťach. Využíva sa v autonómnych vozidlách a strojoch, i v medicíne na analýzu snímok zo zobrazovacích metód ako CT, MRI a pod. Strojové učenie vychádza z matematiky a štatistiky, pričom je podmnožinou AI. Keď dnes spoločnosti zavádzajú programy s umelou inteligenciou, s najväčšou pravdepodobnosťou používajú strojové učenie.



Obr. 1. Strojové učenie (ML) a hlboké učenie (DL) sú v súčasnosti dva najpopulárnejšie prístupy v rámci umelej inteligencie.

Všeobecná definícia strojového učenia je, že ide o študijný odbor, ktorý dáva počítačom schopnosť učiť sa bez toho, aby boli explicitne naprogramované (Arthur Samuel, 1959). Viac inžiniersky orientovaná definícia, uvádza, že počítačový program sa učí na základe skúseností E (*experience*[5]) vzhľadom na nejakú úlohu T (*task*) a nejakú mieru výkonnosti P (*performance*), ak sa jeho výkonnosť v úlohe T, meraná pomocou P, zlepšuje so skúsenosťami E (Tom Mitchell, 1997).

1.2.1 Druhy algoritmov strojového učenia

Strojové učenie používa algoritmy, ktoré sa učia z údajov vytvárať predpovede. Algoritmy strojového učenia sú hnacou silou strojového učenia. Rozoznávame tri hlavné typy učenia (*learning paradigms*), t. j. ako sa stroj (agent) učí, keď sa mu dodávajú údaje: [6]

a) Pri **dozorovanom učení/učení s učiteľom (*supervised learning*)** agent pozoruje vstupy-výstupy a učí sa funkciu, ktorá mapuje zo vstupu na výstup. Vstupmi môžu byť napríklad obrázky z kamier, pričom každý z nich je sprevádzaný výstupom „autobus" alebo „chodec" atď. Takýto výstup sa nazýva štítok (*label*).

[5] Nakoľko "skúsenosti" v tomto zmysle sú vstupom pre program, nazývajú sa aj ako vstupné či trénovacie dáta.
[6] RUSSELL, S., NORVIG, P.: *Artificial Intelligence*: *A Modern Approach* (*4th Edition*), s. 674 - 676.

b) Pri **učení bez učiteľa (*unsupervised learning*)** sa agent učí vzorce na vstupe bez akejkoľvek explicitnej spätnej väzby. Najbežnejšou vzdelávacou úlohou bez učiteľa je zhlukovanie (*clustering)*: zisťovanie potenciálne užitočných zhlukov vstupov. Napríklad zhlukovanie zákazníkov podľa ich nákupných zvykov, či taxonomická klasifikácia organizmov na základe ich genetickej podobnosti.

c) Pri **učení formou odmeňovania (*reinforcement learning)*** sa agent učí interakciou s prostredím: odmenami a trestami. Napríklad na konci šachovej hry sa agentovi povie, že vyhral (odmena) alebo prehral (trest).

Algoritmy strojového učenia je tiež možné rozdeliť podľa ich hlavnej domény a typu použitých vstupných dát, pričom existujú algoritmy ktoré sú špecifické pre danú doménu, ale aj také ktoré sú širšie aplikovateľné.[7]

Medzi hlavné domény patria:

(a) počítačové videnie (*Computer Vision*),
(b) spracovanie prirodzeného jazyka a spracovanie reči (*Natural Language Processing & Speech Processing*),
(c) dátová veda (*Data Science*),
(d) robotika (*Robotics*).

ENISA identifikovala 40 najčastejšie používaných algoritmov ML. Analýzou týchto algoritmov bola zostavená taxonómia s prihliadnutím najmä na typy učenia (s učiteľom, bez učiteľa, s odmeňovaním) a problém, ktorý algoritmy riešia (hlavná doména).[8]

1.2.2 Model a učenie

Zatiaľ sme hovorili o modeli či algoritme strojového učenia ako o akejsi čiernej skrinke (*black box*), ktorá prijíma vstup a vracia akýsi želaný výstup. Čo presne sa však v tejto čiernej skrinke nachádza? Zoberme si model na predpovedanie výšky dieťaťa na základe jeho veku (Obr. 2a). Model strojového učenia je matematická rovnica, ktorá opisuje, ako sa priemerná

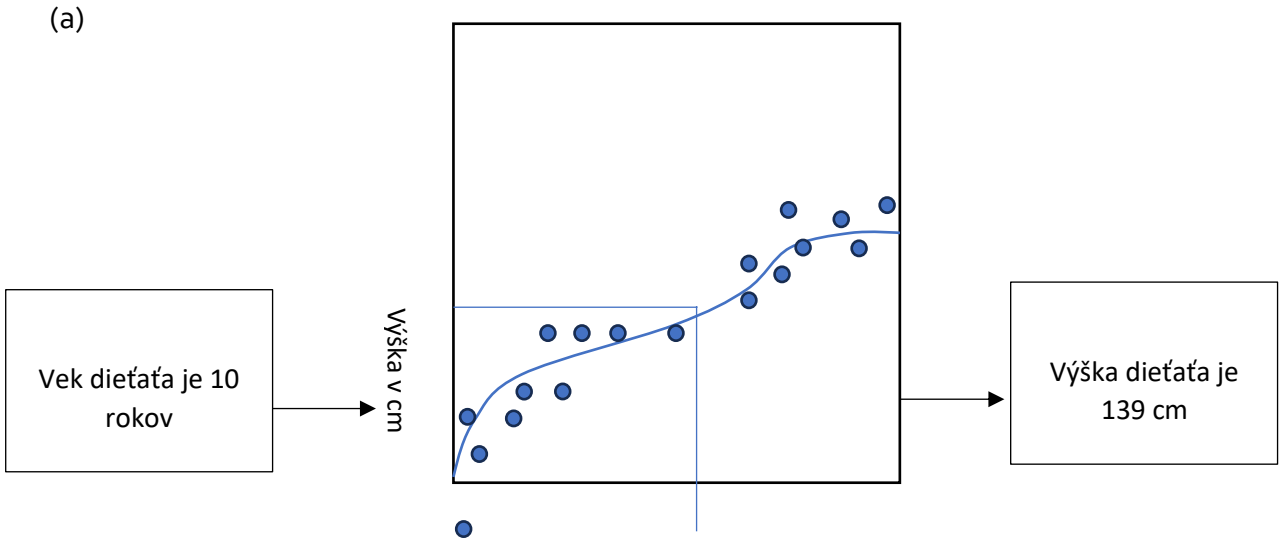
[7] ENISA: *Securing Machine Learning Algorithms*. December 2021, s. 8. [online] citované [30.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms.
[8] Tamže, s. 7.

výška mení v závislosti od veku (krivka na obrázku). Keď cez túto rovnicu preženieme vek, vráti nám výšku. Ak je napríklad vek 10 rokov, potom predpovedáme, že výška bude 139 cm.

Presnejšie povedané, model predstavuje rodinu rovníc mapujúcich vstup na výstup (t. j. rodinu rôznych kriviek). Konkrétna rovnica (krivka) sa vyberá pomocou trénovaných údajov (príkladov dvojíc vstup/výstup). Na obrázku sú tieto dvojice reprezentované bodmi a vidíme, že krivka tieto údaje primerane opisuje. Keď hovoríme o učení, trénovaní, alebo fitovaní modelu, synonymicky máme na mysli, že prehľadávame rodinu možných rovníc (možných kriviek) vzťahujúcich vstup k výstupu, aby sme našli tú, ktorá najpresnejšie opisuje trénované údaje.

Z toho vyplýva, že modely vyžadujú na trénovanie označené páry vstupov/výstupov. Napríklad model určovania hudby by vyžadoval veľký počet zvukových klipov, v ktorých ľudský expert určil žáner každého z nich. Tieto dvojice vstupov/výstupov preberajú úlohu učiteľa alebo supervízora pre proces trénovania, z čoho vzniká termín učenie s učiteľom.[9]

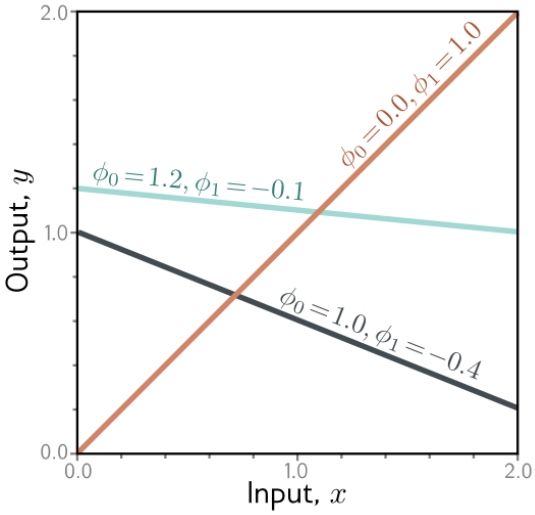
(a)



(b)

Vek v rokoch

[9] porovnaj. kapitolu 1.2.1



Obr. 2: Príklad modelu strojového učenia.[10]

Model je teda v úvodzovkách len matematická rovnica. Keď sa cez túto rovnicu prenesú vstupy, vypočíta sa výstup, ktorý sa nazýva *inferencia (inference)*. Rovnica modelu obsahuje aj *parametre*. Rôzne hodnoty parametrov menia výsledok výpočtu, napr. viď Obr. 2. Modelová rovnica opisuje rodinu možných vzťahov medzi vstupom (vek dieťaťa) a výstupom (výška dieťaťa) a parametre špecifikujú konkrétny vzťah. Konkrétny vzťah sa vyberá pomocou trénovacích údajov, ktoré pozostávajú z dvojíc vstup/výstup (body). Keď trénujeme model, hľadáme v možných vzťahoch ten, ktorý dobre opisuje údaje. V tomto prípade je natrénovaným modelom krivka na obrázku vyššie a môže sa použiť na výpočet výšky pre akýkoľvek vek. Moderné modely strojového učenia sú značne náročné na výpočtový výkon a môžu mať až stovky miliárd parametrov.

Učiaci algoritmus berie trénovaciu množinu dvojíc vstup/výstup a manipuluje s parametrami, kým vstupy čo najpresnejšie nepredpovedajú zodpovedajúce výstupy. Ak model funguje dobre pre tieto tréningové páry, potom dúfame, že bude dobre predpovedať nové vstupy, pri ktorých nie je známy skutočný výstup.

[10] (a) Body znázorňujú trénovacie príklady vstup-výstup, na základe ktorých bol naučený konkrétny model, znázornený ako krivka. Pozn. v tomto príklade bol použitý nelineárny model umožňujúci širšiu rodinu vzťahov ako lineárna regresia znázornená v (b). (b) Lineárna regresia je elementárny typ modelu so *všeobecnou* rovnicou tvaru $y = \phi_0 x + \phi_1$, ktorá opisuje rodinu možných vzťahov medzi vstupom x (napr. vek dieťaťa) a výstupom y (výškou dieťaťa) ako všetky možné lineárne rovnice (t.j. rovné čiary) s dvoma *parametrami* ϕ_0 a ϕ_1. Tieto dva parametre sa určia pomocou procesu učenia na trénovacích údajoch, čím vznikne konkrétny naučený model, napr. $y = 6\,x + 50$.Prevzaté a upravené z PRINCE, S. J.: *Understanding Deep Learning*. MIT Press. 2023, s. 4. online. dostupné na: https://udlbook.github.io/udlbook/.

Výsledný model je teda výstupom z učiaceho postupu (trénovacieho algoritmu), namiesto toho, aby bol výslovne predpísaný jeho tvorcami.

Pozorný čitateľ sa snáď pozastavil nad jednoduchosťou modelu z príkladu v Obr. 2a: Ak je vek 10 rokov, potom predpovedáme, že výška je 139 cm. Tento model je síce len jednoduchý príklad na vysvetlenie pojmov, ale rozoberme si ho o trochu viac pre ilustráciu problémov spojených s nasadením modelov strojového učenia v praxi. Čiže ak model predpovedá 139 cm, znamená to že každé 10 ročné dieťa je vysoké 139 cm? Nie, model predpovedá *priemernú* výšku. Ale priemernú v akej populácii? Záleží na krajine, etniku, pohlaví, príjmovej skupine domácnosti, či čase v ktorom sa pýtame (napr. 19. storočie vs. 21.)? Asi najpresnejšia odpoveď je že model predpovedá výšku na základe parametrov odvodených z populácie v *trénovacom datasete*. Ak sú tieto dáta skreslené (*biased*) tak bez prijatia protiopatrení bude aj výsledný model pravdepodobne skreslený. Takto vzniká *historical bias*, keď napr. staré trénovacie dáta nereprezentujú súčasný stav. *Selection bias*, keď nejaká populačná skupina je prílišne alebo nedostatočne zastúpená. Druhmi nedorozumení a skreslení sa podrobne venuje podoblasť nazvaná *Fairness (in AI/ML)* alebo širšie tzv. FATML (*fairness, accountability, and transparency in machine learning*).

1.2.3 Trénovacie údaje

Trénovací dataset (*training dataset*) sú počiatočné údaje (dáta či skúsenosti), ktoré sa používajú na učenie modelu strojového učenia na spracovanie informácií.

Kvalitné trénovacie údaje sú základom úspešného strojového učenia, pretože kvalita trénovacích údajov má zásadný vplyv na vývoj, výkon a presnosť každého modelu. Trénovacie údaje sú pre úspech modelu pripraveného na produkciu rovnako dôležité ako samotné algoritmy, pretože kvalita a objem označených trénovacích údajov priamo ovplyvňujú presnosť, s akou sa model naučí identifikovať výsledok, na ktorý bol navrhnutý.

Trénovacie údaje usmerňujú model, sú teda materiálom, z ktorého model získava svoje základné znalosti. Ukazujú modelu vzory a hovoria mu, čo má hľadať. Po tom, ako dátoví vedci model natrénujú, mal by byť schopný identifikovať vzory v nikdy predtým nevidených súboroch údajov na základe vzorov, ktoré sa naučil z trénovacích údajov.

1.2.4 Ďalšie prístupy k modelovaniu

Model teda predstavuje jadro systému umelej inteligencie a s určitou mierou nadsázky môžeme uvažovať o modeli ako o "mozgu" agenta, ktorý určuje jeho správanie. V prípade softvérového agenta, ktorý nemá fyzické aktuátory, v podstate agent = model (mozog). Model môže byť založený na údajoch a/alebo odborných znalostiach ľudí a/alebo automatizovaných nástrojov, ako sú algoritmy strojového učenia.

Strojové učenie však nie je jedinou oblasťou umelej inteligencie. Existuje viacero rovnocenných odborov k strojovému učeniu, pričom *Expertné systémy* patria v súčasnosti tiež k významným odborom. Ako napovedá názov, mozog agenta, ktorý je založený na expertnom systéme, je vytvorený expertami na danú oblasť, často značne prácnym postupom. Pozostáva zo série logických pravidiel a akcií, ktoré agentovi prikazujú správať sa veľmi špecifickým, vopred určeným spôsobom. Vysoká štruktúrovanosť rozhodovacieho procesu prispieva k jeho odolnosti a uľahčuje audit. Expertné systémy sú síce účinné v niektorých oblastiach, majú ale aj významné nedostatky. Tréning mozgu si vyžaduje spoluprácu s expertom na danú oblasť a starostlivé definovanie pravidiel a ich následné udržiavanie. Expertné systémy nevedia dobre zovšeobecňovať a riešiť problémy, s ktorými sa ešte nestretli.

Strojové učenie rieši mnohé nevýhody expertných systémov tým, že nevyžaduje explicitné budovanie mozgu pravidlom po pravidle. Je tiež schopný zovšeobecňovať problémy a do väčšej miery poskytovať odpovede na otázky, s ktorými sa predtým nestretol.

Pochopiť, ako vytvoriť mozog (model) pomocou strojového učenia, môže byť pomerne náročné. V predchádzajúcej časti sme uviedli príklad lineárnej regresie. Ďalším z riešení pomocou strojového učenia je vytvoriť model, ktorý sa podobá diagramu priebehu procesov. Robíme to pomocou modelovacieho algoritmu nazývaného rozhodovací strom (*Decision Tree*). Štruktúra modelu sa podobá štruktúre stromu s uzlami, vetvami a listami. Otázka znie, ako sa automaticky rozhodnúť, aké otázky položiť v každom uzle stromu (rozhodovacom bode vývojového diagramu). Veľmi zjednodušená odpoveď je, že strom vytvárame počnúc koreňom a postupujeme nadol ku každému z listov (v informatike platí, že stromy rastú hore nohami). V každom uzle algoritmus určí, ktorá vlastnosť by sa mala použiť v rozhodovacom

poli, a hodnotu, s ktorou ju testujeme. Robí to tak, že sa pozrie na množstvo údajov o predchádzajúcich rozhodnutí. Mozog sa teda síce konštruuje automaticky, ale je veľmi závislý od kvalitných označených údajov.

Modelovanie mozgu ako stromu je len jednou z viacerých možností. Nemalo by byť prekvapením, že pri skúmaní nových spôsobov modelovania mozgu vedci vychádzali z toho, ako je štruktúrovaný ľudský mozog. Táto kategória modelovacích algoritmov sa preto označuje ako umelá neurónová sieť (*Artificial Neural Network*). Ľudský mozog má podľa odhadov 86 miliárd neurónov usporiadaných do zložitých vrstiev. Neurón "vystrelí", ak podnet z neurónov, ktoré sú s ním spojené, prekročí určitý prah (váhu). Pritom sa medzi vrstvami prenášajú informácie.

Ak uvažujeme o jednoduchej sieti, môže pozostávať len z troch vrstiev, vstupnej vrstvy (zodpovedá vstupným znakom, napr. polohe transakcie), strednej vrstvy (kombinuje vstupy) a výstupnej vrstvy (odpoveď mozgu na problém). V praxi neurónové siete zvyčajne vyžadujú veľký počet skrytých vrstiev (obvykle niekoľko desiatok) na riešenie zložitých problémov, ako je počítačové videnie alebo spracovanie prirodzeného jazyka.

Zatiaľ sme popísali mozog reprezentovaný sieťou s jasne definovaným vstupom, niekoľkými zložitými kombináciami výsledkov v stredných vrstvách a nakoniec výstupnou vrstvou, ktorá prezentuje odpoveď. Mozog je však viac než len zložitá sieť vzájomných prepojení. Mozog potrebuje zakódovať to, čo sa naučil. V príklade rozhodovacieho stromu sme zakódovali vedomosti prostredníctvom otázok v každom uzle stromu. V neurónovej sieti namiesto otázok kódujeme učenie nastavením prahu uvoľnenia každého neurónu. Niektoré neuróny sa budú musieť rýchlo uvoľňovať (s malým počtom vstupov) a iné budú potrebovať veľa vstupov spoločne, aby mozog dal správnu odpoveď.

V nasledujúcej časti popíšeme, ako súvisí hlboké učenie (*Deep learning*) s neurónovými sieťami. Ak má neurónová sieť dve alebo viac skrytých vrstiev, potom sa nazýva hlboká neurónová sieť (*Deep Neural Network*, DNN). Štúdiom hlbokých neurónových sietí sa zaoberá hlboké učenie. V súdobej praxi všetky široko rozšírené modely neurónových sietí sú hlboké neurónové siete a tak sa tieto termíny často používajú synonymicky.

1.3 Hlboké učenie (*Deep learning*)

Hlboké učenie je široká skupina techník strojového učenia, v ktorej hypotézy majú podobu komplexných obvodov s nastaviteľnou váhou (parametrom) spojenia. Slovo "hlboké" odkazuje na skutočnosť, že obvody sú zvyčajne organizované do mnohých vrstiev, čo znamená, že výpočtové cesty od vstupov k výstupom majú mnoho krokov.

Hlboké neurónové siete dokážu spracovať vstupy, ktoré sú veľmi veľké, majú rôznu dĺžku a obsahujú rôzne druhy vnútorných štruktúr. Ich výstupom môžu byť jednotlivé reálne čísla (regresia), viacero čísel (viacrozmerná regresia) alebo pravdepodobnosti nad dvoma či viacerými triedami (binárna, resp. viactriedna klasifikácia). Aj ich výstupy môžu byť tiež veľmi veľké, s premenlivou dĺžkou a môžu obsahovať vnútornú štruktúru.

V čase písania tohto textu sú hlboké siete najvýkonnejšími modelmi strojového učenia a často sa s nimi stretávame v každodennom živote. Hlboké učenie je v súčasnosti najpoužívanejším prístupom pre aplikácie, ako je vizuálne rozpoznávanie objektov, strojový preklad, rozpoznávanie reči, syntéza reči a syntéza obrazu; zohráva tiež významnú úlohu v aplikáciách.

1.3.1 Stratové funkcie (*Loss functions*)

Hlboké neurónové siete predstavujú rodinu funkcií, ktoré mapujú vstup na výstup, pričom konkrétny člen rodiny je určený parametrami modelu. Pri trénovaní týchto modelov hľadáme parametre, ktoré vytvárajú najlepšie možné mapovanie zo vstupu na výstup pre uvažovanú úlohu.

Strata či chyba (odchýlka od najlepšieho možného výsledku) je jedno číslo, ktoré predstavuje nesúlad medzi predikciami siete a základnou pravdou pre trénovaciu množinu.

Strata závisí od parametrov siete. Cieľom učenia parametrov siete alebo jednoducho trénovania či fitovania modelu je nájsť hodnoty parametrov, ktoré minimalizujú práve túto stratu.

1.3.2 Trénovacie algoritmy a gradientový zostup (*Gradient descent*)

Cieľom trénovacieho algoritmu je nájsť parametre, ktoré minimalizujú stratu. Existuje mnoho rodín trénovacích algoritmov, ale štandardné metódy na trénovanie neurónových sietí sú iteračné.

Najrozšírenejšou metódou (v rôznych modifikáciách) je algoritmus stochastického gradientového zostupu (*Stochastic gradient descent*). V kontexte neurónových sietí sa ním hľadajú parametre, ktoré minimalizujú stratu tak, aby model presne predpovedal trénovacie výstupy zo vstupov. Základný prístup spočíva v náhodnom výbere počiatočných parametrov a následnom vykonaní série malých zmien, ktoré v priemere znižujú stratu. Jedným zo spôsobov, ako to urobiť, je zmerať sklon (gradient) stratovej funkcie vzhľadom na aktuálne parametre a urobiť krok v smere, ktorý je najstrmší smerom nadol. Potom tento postup opakujeme, kým nie je sklon rovný, t.j. už sa nedá jednoducho znížiť stratu ďalším postupným zlepšovaním. Po mnohých iteráciách dúfame, že dosiahneme uspokojivé minimum stratovej funkcie.

1.3.3 Základné kategórie hlbokých neurónových sietí

Jednoobvodová sieť (*feedforward neural network, FNN*), ako naznačuje názov, má spojenia len v jednom smere - to znamená, že tvorí smerovaný acyklický graf s určenými vstupnými a výstupnými uzlami. Každý uzol vypočíta funkciu svojich vstupov a výsledok odovzdá svojim nasledovníkom v sieti.

Na druhej strane, **rekurentná sieť** (*recurrent neural network, RNN*) vracia svoje medziprodukty alebo konečné výstupy späť do svojich vlastných vstupov. To znamená, že hodnoty signálov v sieti tvoria dynamický systém, ktorý má vnútorný stav alebo pamäť.

Konvolučná neurónová sieť (*convolutional neural network,* CNN) je ďalším druhom hlbokých neurónových sietí. Konvolučné neurónové siete majú rovnako ako iné neurónové siete parametre (váhy), ktorých hodnoty sú predmetom učenia. Ich architektúru rovnako popisujeme vrstvami obsahujúcimi neuróny, ktorých výpočet sa líši v závislosti od ich typu. Každý neurón v sieti počíta skalárny súčin svojho vstupu a váh, a typicky aktivuje svoj výstup na základe nelineárnej funkcie. Od všeobecných neurónových sietí sa ale odlišujú tým, že

obsahujú tzv. konvolučnú vrstvu (konvolúcia je matematická operácia pre rozpoznávanie vzorcov). Konvolučné siete nájdeme v oblasti počítačového videnia, v systémoch na rozpoznávanie tváre, pri spracovaní textu, v robotike či v autonómnych vozidlách.

Generatívna adversariálna sieť (*generative adversarial network,* GAN) je model strojového učenia, v ktorom dve neurónové siete medzi sebou súťažia pomocou metód hlbokého učenia, aby boli presnejšie vo svojich predpovediach. GAN zvyčajne fungujú bez dohľadu a na učenie používajú rámec kooperatívnej hry s nulovým súčtom, kde sa zisk jednej osoby rovná strate druhej osoby, pričom cieľom je nájsť čo najlepší kompromis. Dve neurónové siete, ktoré tvoria GAN, sa označujú ako generátor a diskriminátor. Generátor je konvolučná neurónová sieť a diskriminátor je dekonvolučná neurónová sieť. Cieľom generátora je umelo vyrobiť výstupy, napr. obrázky, ktoré by sa dali ľahko zameniť za skutočné údaje. Cieľom diskriminátora je práve rozlíšiť medzi takto umelo vytvorenými údajmi a skutočnými. V ideálnom prípade súťaženie generátora a diskriminátora vedie k systému ktorý je schopný generovať realistické výstupy avšak dostatočne odlišné od trénovacích dát.

Generatívny difúzny model (*generative diffusion model*, GDM). GDM je trieda pravdepodobnostných generatívnych modelov, ktoré vytvárajú obrázky z náhodného šumu postupnými krokmi odšumovania. Ich hlavnou časťou je odšumovací dekodér, ktorý realizuje jednotlivé kroky tohto procesu. Ako dekodér je obvykle použitá hlboká neurónová sieť podobná kombinácií konvolučnej a dekonvolučnej siete. Počas učenia sa k skutočným obrázkom pridáva rozličné množstvo náhodného šumu (od malého množstva až po také čo úplne zničí pôvodný obsah), na ktorých sa dekodér učí správne odstraňovať šum. Navyše, dekodér je možné podmieniť (napr. textom "mačka leží na posteli") a ovplyvniť tak proces odšumovania želaným spôsobom. V súčasnosti (začiatok 2024) všetky najlepšie generatívne modely obrázkov (DALL-E 3 a pod.) sú GDM a vo veľkej miere nahradili GAN.

Hoci modely hlbokého učenia v mnohých prípadoch dobre zovšeobecňujú z trénovacích príkladov (údajov) na nové nevidené príklady, môžu tiež produkovať chyby. Majú tendenciu vytvárať nesprávne mapovania vstupov a výstupov tak, že malá zmena vstupu môže spôsobiť veľkú zmenu výstupu. Napríklad je možné zmeniť len niekoľko pixelov na obrázku psa a spôsobiť, že sieť klasifikuje psa ako pštrosa alebo školský autobus - aj keď zmenený obrázok

pre človeka stále vyzerá presne ako pes. Takto zámerne zmenený obrázok sa nazýva adversariálny príklad (*adversarial example*) a je to jeden z útokov na model strojového učenia.

1.4 Spracovanie prirodzeného jazyka (*Natural Language Processing*, NLP)

Jazykové úsudky v prirodzených jazykoch, ako je napr. slovenčina či angličtina, sa líšia od človeka k človeku a čas od času. Jednu myšlienku vieme vyjadriť viacerými spôsobmi, pričom môžu alebo nemusia byť vyjadrené gramaticky správne.

Prirodzený jazyk je zároveň nejednoznačný, ("Sadol som si s ním" môže znamenať buď to, že si dvaja vzájomne rozumejú, alebo to, že dvaja mali spoločné stretnutie) a vágny ("Ide si svoje" presne nešpecifikuje o akú činnosť ide, ani v čom spočíva "svoje").

Mapovanie zo symbolov na objekty nie je formálne definované. V prirodzenom jazyku sa dva výskyty toho istého slova alebo slovného spojenia môžu vzťahovať na rôzne veci vo svete. Avšak vo formálnej logike môže jeden symbol predstavovať len práve jeden objekt.

Ak nedokážeme definitívne rozlíšiť gramaticky správne a nesprávne reťazce, môžeme aspoň povedať, aké pravdepodobné alebo nepravdepodobné sú jednotlivé reťazce.

Jazykový model definujeme ako určenie pravdepodobnosti opisujúce realistickosť akéhokoľvek reťazca v danom jazyku. Takýto model by mal povedať, že "Pokúsim sa preplávať Dunaj." má ako reťazec slovenčiny vysokú pravdepodobnosť, ale "Dunaj pokúsim preplávať?" je nepravdepodobný a náhodný reťazec ako "xVd6 qrlc !t" je extrémne nepravdepodobný.

Pomocou jazykového modelu možno predpovedať, aké slová budú pravdepodobne nasledovať v texte, a tým navrhnúť dokončenie e-mailovej alebo textovej správy. Možno vypočítať, ktoré zmeny v texte by zvýšili jeho pravdepodobnosť, a tým navrhnúť pravopisné alebo gramatické opravy. Pomocou dvojice modelov môžeme vypočítať najpravdepodobnejší preklad vety. S niektorými príkladmi dvojíc otázka/odpoveď ako trénovacími údajmi môžeme vypočítať najpravdepodobnejšiu odpoveď na otázku. Jazykové modely sú teda základom širokého spektra úloh prirodzeného jazyka. Samotná úloha

jazykového modelovania slúži aj ako spoločné kritérium na meranie pokroku v porozumení jazyka.

Rozpoznávanie reči je úloha transformácie hovoreného zvuku na text. Na výslednom texte potom môžeme vykonávať ďalšie úlohy (napríklad odpovedanie na otázky). Súčasné systémy majú chybovosť slov približne 3 % až 5 % (v závislosti od podrobnosti testovacej množiny), podobne ako ľudskí prepisovatelia. Výzvou pre systém využívajúci rozpoznávanie reči je vhodne reagovať aj v prípade chýb v jednotlivých slovách.

Syntéza textu na reč je opačný proces, teda prechod od textu k zvuku. Výzvou je správne vysloviť každé slovo a dosiahnuť, aby priebeh každej vety pôsobil prirodzene, so správnymi pauzami a dôrazom.

Extrakcia informácií je proces získavania znalostí prechádzaním textu a hľadaním výskytov určitých tried objektov a vzťahov medzi nimi. Typickou úlohou je extrahovať výskyty adries z webových stránok s databázovými políčkami pre ulice, mestá, štáty a poštové smerovacie číslo. Extrakcia informácie prebieha jednoducho ak je zdrojový text dobre štruktúrovaný (napríklad vo forme tabuľky).

Vyhľadávanie informácií je úloha nájsť dokumenty, ktoré sú relevantné a dôležité pre daný dotaz. Internetové vyhľadávače, ako napríklad Google vykonávajú túto úlohu miliardy krát denne.

Odpovedanie na otázky je odlišná úloha, pri ktorej je dopyt skutočne otázkou, napríklad "Kto bol prvý slovenský kozmonaut?", a odpoveďou nie je zoradený zoznam dokumentov, ale priama odpoveď "Ivan Bella". Systémy na zodpovedanie otázok, ktoré sa opierajú o syntaktický rozbor už existujú dlhšie, ale až od roku 2001 sa takéto systémy využívajú na vyhľadávanie informácií na webe.

1.5 Počítačové videnie (*Computer Vision*, CV)

Zrak prijíma podnety a podáva správu o určitom zobrazení okolia. Väčšina agentov, ktorí používajú videnie, používa pasívne vnímanie, teda nemusia vysielať svetlo, aby videli. Naproti tomu aktívne snímanie zahŕňa vysielanie signálu, napríklad radaru alebo ultrazvuku,

a snímanie odrazu. Nejde len o videnie, roboty v reálnom svete používajú aj rôzne senzory na vnímanie zvuku, dotyku, vzdialenosti, teploty, polohy a zrýchlenia.

V počítačovom videní počítač používa na hľadanie riešení mapovanie obrazu a vzoru. Obraz považuje za pole pixelov. Počítačové videnie automatizuje úlohy monitorovania, kontroly a dohľadu.

Základnými problémami počítačového videnia sú rozpoznávanie, rekonštrukcia a reorganizácia.[11]

Rozpoznávanie sa týka priraďovania sémantických kategórií k objektom a scénam, ako aj k udalostiam a činnostiam. Aspektmi rozpoznávania je zachytávanie vzťahov medzi časťami a celkami (partonómia), ako aj medzi kategóriami a podkategóriami (taxonómia). Rozpoznávanie jemných kategórií zahŕňa ako extrémny prípad identifikácie na úrovni entity (napr. tvár konkrétnej osoby).

Rekonštrukcia sa týka obnovy trojrozmernej geometrie sveta z jedného alebo viacerých jeho 2D obrazov. Tento pojem interpretujeme širšie ako "inverznú grafiku" - odhad tvaru, priestorového rozloženia, odrazivosti a osvetlenia, ktoré by sa mohli spoločne použiť na vykreslenie scény s cieľom vytvoriť obraz.

Reorganizácia je označenie pre to, čo sa v ľudskom videní zvyčajne nazýva „organizácia percepčného poľa" - organizácia vnímania, čiže spôsob, akým celkový vizuálny podnet (celé zorné pole) „rozmieňame" na dielčie podnety.

Pokroky v konvolučných neurónových sieťach (CNN, bližšie kapitola 1.3.3) viedli k mimoriadnemu výkonu v týchto úlohách a vo vizuálnom rozpoznávaní. V dôsledku toho sa konvolučné neurónové siete stali základnými stavebnými kameňmi výpočtov hlbokého učenia v oblasti počítačového videnia.

Strojové učenie v oblasti počítačového videnia sa využíva pri interpretácii údajov obsiahnutých v snímkach, napr. rozpoznávanie tvárí, spracovanie obrazu pre samoriadiace

[11] MALIK, J. et al.: *The three R's of computer vision: Recognition, reconstruction and reorganization*, Pattern Recognition Letters, Volume 72, 2016, Pages 4-14, ISSN 0167-8655, online. dostupné na: https://doi.org/10.1016/j.patrec.2016.01.019.

autá, analýza medicínskych snímok, pre kontrolu a riadenie v robotizovanej výrobe, pri interpretácii údajov diaľkového prieskumu Zeme pre rôzne geografické informačné systémy, či pre extrakciu grafických a textových informácií zo snímok dokumentov. Priemyselných nasadení strojového učenia a počítačového videnia je dnes veľmi veľa.

Generatívne adversariálne siete (GAN) a generatívne difúzne modely (GDM) sú určené pre tvorivé úlohy, na rozdiel od kategorizačných či rozpoznávacích úloh. Dokážu vytvárať nové realistické obrázky a videá. Výstupy týchto modelov možno podmieniť rôznymi spôsobmi, ako napr. štylistickým typom (olejomaľba, čiernobiela kresba, a pod.), textovým popisom želanej scény, alebo aj iným obrázkom a popisom želanej zmeny. Jedným z druhov takýchto obrázkov je tzv. *deepfake* - obrázok alebo video, ktoré vyzerá ako konkrétna osoba, ale je vygenerované z modelu.

1.6 Robotika (*Robotics*)

Roboty sú fyzickí agenti, ktorí vykonávajú úlohy manipuláciou s fyzickými predmetmi. Na tento účel sú vybavené *efektormi*, ako sú nohy, kolesá, kĺby a chápadlá. Efektory sú navrhnuté tak, aby pôsobili fyzikálnymi silami na prostredie. Keď to robia, môže sa stať niekoľko vecí: môže sa zmeniť stav robota (napr. auto roztočí svoje kolesá a v dôsledku toho urobí pohyb na ceste), môže sa zmeniť stav prostredia (napr. robotické rameno použije svoje chápadlo na posunutie hrnčeka po pulte) a dokonca sa môže zmeniť stav ľudí v okolí robota (napr. exoskelet sa pohne a to zmení konfiguráciu nohy človeka; alebo mobilný robot postupuje smerom k dverám výťahu a človek si to všimne a ustúpi z cesty, alebo dokonca stlačí tlačidlo pre robota).

Roboty sú tiež vybavené senzormi, ktoré im umožňujú vnímať okolie. Súčasná robotika využíva rôznorodý súbor senzorov vrátane kamier, radarov, laserov a mikrofónov na meranie stavu prostredia a ľudí okolo neho a gyroskopov, snímačov ťahu a krútiaceho momentu a akcelerometrov na meranie vlastného stavu robota.

Mechanizmus, ktorý iniciuje pohyb efektora, sa nazýva *aktuátor*; medzi príklady patria prevodovky, ozubené kolesá, káble a spoje. Najbežnejším typom aktuátora je elektrický aktuátor, ktorý využíva elektrickú energiu na roztočenie motora.

Učenie s odmeňovaním (*reinforcement learning*) sa uplatňuje v robotike s technikami, ktoré sa snažia znížiť potrebný počet interakcií s reálnym svetom. Takéto techniky majú tendenciu využívať modely, či už ide o odhadovanie modelov a ich využívanie na plánovanie, alebo o trénovanie politík, ktoré sú robustné vzhľadom na rôzne možné parametre modelu.

Interakcia s ľuďmi si vyžaduje schopnosť koordinovať činnosti robota s ich činnosťami, čo možno formulovať ako hru. Riešenie zvyčajne rozkladáme na predikciu, v ktorej sa používajú prebiehajúce činnosti človeka na odhad toho, čo bude robiť v budúcnosti, a akciu, v ktorej sa používajú predikcie na výpočet optimálneho pohybu robota.

Pomoc ľuďom si vyžaduje aj schopnosť naučiť sa alebo odvodiť, čo chcú. Roboty sa k tomu môžu priblížiť tak, že sa naučia požadovanú funkciu, ktorú by mali optimalizovať, zo vstupov od človeka, ako sú napríklad ukážky, opravy alebo inštrukcie v prirodzenom jazyku. Prípadne môžu roboty napodobňovať ľudské správanie a používať učenie odmeňovaním, ktoré pomáha riešiť problém zovšeobecňovania na nové prípady.

Robotika sa tak z časti prelína s umelou inteligenciou, a z oblastí ktoré sme v tomto texte pokryli, často využíva algoritmy strojového učenia, vyžaduje riešenia z počítačového videnia, spracovania prirodzeneho jazyka, či učenia s odmeňovaním. Má potenciál zlepšiť domácu starostlivosť, zdravotnú starostlivosť, služby, autonómne vozidlá, prieskum nebezpečných prostredí a priemysel.

1.7 Generatívna umelá inteligencia (*GenAI*)

Existujú dve široké triedy systémov AI na základe ich schopností: Prediktívna AI (PredAI) a generatívna AI (GenAI).[12] Predikatívna AI sa využíva v počítačovom videní na detekciu a klasifikáciu objektov. Generatívna AI je neoddeliteľnou súčasťou jazykových modelov, avšak zďaleka to nie je jediná doména generatívnej AI.

Generatívna umelá inteligencia je podmnožina strojového učenia, ktorá označuje triedu modelov umelej inteligencie, ktoré sú schopné generovať kreatívny obsah, často vo forme

[12] VASSILEV, A., et al. *Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations*. (National Institute of Standards and Technology, Gaithersburg, MD) NIST Artificial Intelligence (AI) Report, NIST Trustworthy and Responsible AI NIST AI 100-2e2023. s. 3 . 2024. dostupné na: https://doi.org/10.6028/NIST.AI.100-2e2023.

textu, obrázkov alebo dokonca zvuku. Tieto modely sú navrhnuté tak, aby vytvárali nové výstupy, ktoré nie sú priamo kopírované zo vstupných údajov. Systémy generatívnej umelej inteligencie, ako napríklad GAN či GDM, fungujú tak, že sa učia vzory a na ich základe generujú obsah od začiatku.

Generatívnu umelú inteligenciu je potrebné odlišovať od iných systémom umelej inteligencie, ktoré vykonávajú iné funkcie, ako je napríklad klasifikácia údajov (napr. priraďovanie štítkov obrázkom), zoskupovanie údajov (napr. identifikácia segmentov zákazníkov s podobným nákupným správaním) alebo výber činností (napr. riadenie autonómneho vozidla).

Niektoré generatívne modely explicitne opisujú rozdelenie pravdepodobnosti nad vstupnými údajmi a nové príklady sa tu generujú vzorkovaním z tohto rozdelenia. Iné sa len učia mechanizmus na generovanie nových príkladov bez explicitného opisu ich rozdelenia. Najmodernejšie generatívne modely dokážu syntetizovať príklady, ktoré sú mimoriadne pravdepodobné, ale odlišné od trénovacích príkladov. Obzvlášť úspešné boli pri generovaní obrázkov a textu. Dokážu tiež syntetizovať údaje s obmedzením, že niektoré výstupy sú vopred určené (tzv. podmienené generovanie). Príklady zahŕňajú dokreslenie obrázkov a dopĺňanie textu. Moderné generatívne modely pre text sú skutočne také výkonné, že sa môžu javiť ako inteligentné. Pri danom texte, po ktorom nasleduje otázka, model často dokáže "doplniť" chýbajúcu odpoveď vygenerovaním najpravdepodobnejšieho doplnenia dokumentu. V skutočnosti však model pozná len štatistiku jazyka a nerozumie významu svojich odpovedí.

Medzi kľúčové charakteristiky generatívnej umelej inteligencie preto patrí:

Kreativita: Generatívne systémy AI sú navrhnuté tak, aby boli kreatívne a vytvárali užitočný obsah, ktorý sa ale nenachádza v ich trénovacích údajoch. Dokážu generovať nové nápady, umelecké diela alebo texty, ktoré sú často nepredvídateľné.

Variabilita: Tieto modely môžu produkovať širokú škálu výstupov, vďaka čomu sú užitočné pri kreatívnych úlohách, ako je umenie, hudba a rozprávanie príbehov.

Multimodálnosť: Generatívna umelá inteligencia síce dokáže pracovať s textom, ale neobmedzuje sa len na generovanie jazyka a môže sa používať na rôzne kreatívne aplikácie, kedy je vstupom a výstupom ľubovoľná kombinácia typu informácií (modalít) ako text, zvuk, obrázky, či video.

Typickými príkladmi generatívnych systémov umelej inteligencie sú generátory obrázkov (napríklad DALL-E, Midjourney alebo Stable Diffusion), veľké jazykové modely (napríklad GPT-4, Gemini, LLaMA), nástroje na generovanie kódu (napríklad Copilot) alebo nástroje na generovanie zvuku (napríklad VALL-E alebo resemble.ai).

1.8 Odporúčacie systémy (*Recommender systems*)

Odporúčacie systémy sa využívajú v mnohých oblastiach, ako sú napríklad sociálne média, webové stránky eshopov, stránky s hudobnými/filmovými službami (Spotify, Netflix etc.), online reklama, a to je len niekoľko príkladov.

Hlavným cieľom odporúčacích systémov je pomôcť používateľom objaviť relevantný obsah, ako sú filmy na sledovanie, texty na čítanie alebo produkty na kúpu. Odporúčacie systémy nahrádzajú vyhľadávače tým, že znižujú úsilie pri proaktívnom vyhľadávaní a prekvapujú používateľov ponukami, ktoré nikdy nehľadali.

Odporúčacie systémy a online sociálne siete úzko súvisia. Obe sa zameriavajú na zvládnutie obrovského množstva údajov, ktoré používatelia vytvárajú a zdieľajú prostredníctvom online platforiem, pričom sa snažia udržať vysokú angažovanosť používateľov. Ich spolupráca je postavená na výhodách, ktoré môžu oba systémy dosiahnuť: optimalizácia techník odporúčaní využívaním dodatočného obsahu a charakteristík používateľov získaných zo sociálnej siete a rastúca požiadavka personalizácie služieb sociálnych služieb.

Významným prvkom používania sociálnych sietí sú smart zariadenia (napr. smartfóny a tablety). Preto dolovanie informácií poskytovaných týmito zariadeniami a ich správne kombinovanie s údajmi sociálnych sietí môže poskytnúť presnejší model kontextu a preferencií používateľa. Napríklad sociálny kontext používateľa možno definovať kombináciou virtuálnych sociálnych vzťahov získaných z sociálnych sietí s fyzickými

kontaktmi medzi zariadeniami a informáciami obsiahnutými v osobnom zariadení používateľa (napr. kontakty uložené v adresári, záznamy hovorov a správy). V tomto prípade možno preferencie a potreby používateľa odvodiť z heterogénnych zdrojov údajov, napr. histórie webového prehliadača, akcií vykonaných na sociálnych sieťach, vytvorených obsahov na sociálnych sieťach, navštívených miest a súboru najpoužívanejších mobilných aplikácií.[13]

1.8.1 Kolaboratívne filtrovanie (*Collaborative filtering*)

Kolaboratívne filtrovanie je dôležitý koncept v odporúčacích systémoch. V širšom zmysle ide o proces filtrovania informácií alebo vzorov pomocou techník zahŕňajúcich spoluprácu viacerých používateľov, agentov a zdrojov údajov. Kolaboratívne filtrovanie má mnoho podôb a od svojho vzniku bolo navrhnutých množstvo metód kolaboratívneho filtrovania.

Vo všeobecnosti kolaboratívne filtrovanie využíva na vytváranie predpovedí a odporúčaní len údaje o interakcii používateľa a položky (*item*). Okrem kolaboratívneho filtrovania sú užitočné aj odporúčacie systémy založené na obsahu a kontexte, ktoré zahŕňajú opisy obsahu položiek/používateľov a kontextové signály, ako sú časové pečiatky a lokalizácie.

Každá diskusia o hlbokom učení v odporúčacích systémoch by bola neúplná bez zmienky o jednom z najdôležitejších prielomov v tejto oblasti, neurónovom kolaboratívnom filtrovaní (*Neural collaborative filtering*, NCF), ktoré predstavili v roku 2017 vedci zo Singapurskej univerzity.[14]

Pred NCF bola štandardom v odporúčacích systémoch faktorizácia matíc (*matrix factorization*), pri ktorej sa učíme skryté (latentné) vektorové reprezentácie, tzv. *embeddings*, pre používateľov aj položky a potom generujeme odporúčania pre používateľa pomocou vektorového súčinu medzi vektorom používateľa a vektormi položiek. Čím je súčin bližšie k číslu 1, tým je predpovedaná zhoda vyššia. Na faktorizáciu matíc ako takú sa dá jednoducho

[13] CAMPANA, M. G.; DELMASTRO, F.: *Recommender systems for online and mobile social networks: A survey*. Online Social Networks and Media, 2017, 3: 75-97.
[14] HE, X., et al. *Neural collaborative filtering*. In: Proceedings of the 26th international conference on world wide web. 2017. p. 173-182.

pozerať ako na lineárny model latentných faktorov. Kľúčovou myšlienkou v neurónovom kolaboratívnom filtrovaní je nahradiť súčin matíc v maticovej faktorizácii neurónovou sieťou.

1.8.2 Explicitná a implicitná spätná väzba

Aby sa systém dozvedel o preferenciách používateľov, zhromažďuje od nich spätnú väzbu. Spätná väzba môže byť explicitná alebo implicitná. Napríklad YouTube poskytuje používateľom tlačidlá palec hore na vyjadrenie ich preferencií. Je zrejmé, že zhromažďovanie explicitnej spätnej väzby si vyžaduje, aby používatelia aktívne označili svoje záujmy. Napriek tomu mnohí používatelia sa môžu zdráhať hodnotiť produkty. Naproti tomu implicitná spätná väzba je často dostupnejšia, pretože sa týka najmä modelovania implicitného správania, ako sú napríklad kliknutia používateľov. Mnohé odporúčacie systémy sa preto sústreďujú na implicitnú spätnú väzbu, ktorá nepriamo odráža názor používateľa prostredníctvom pozorovania jeho správania. Existujú rôzne formy implicitnej spätnej väzby vrátane histórie nákupov, histórie prehliadania, zhliadnutí a dokonca pohybov myši.

1.9 Základné modely (*Foundation models*)

Základné modely (nielen generatívnej) umelej inteligencie tvoria špeciálnu kategóriu AI modelov. Základný model je model systému AI vyvinutý z algoritmov navrhnutých na optimalizáciu pre všeobecnosť a všestrannosť výstupu, možno ho opätovne použiť v nespočetných nadväzujúcich (*downstream*) aplikáciách. Tieto modely sú často trénované na širokom spektre zdrojov údajov a veľkých množstvách údajov, aby mohla byť riešená široká škála nadväzujúcich (downstreamových) úloh, vrátane niektorých, pre ktoré neboli špeciálne vyvinuté a natrénované.

Pojem "základný model" zaviedli výskumníci zo Stanfordu (*Center for Research on Foundation Models, Stanford Institute for Human-Centered Artificial Intelligence*) v prelomovej správe "*On the opportunities and risks of foundation models*" (O príležitostiach a rizikách základných modelov)[15] z roku 2021. Pojem základný model bol dokonca prevzatý a definovaný aj v rámci návrhov Nariadenia o umelej inteligencii. Konečný text Nariadenia

[15] BOMMASANI, R. et al.: *On the opportunities and risks of foundation models* (2021). [online] [19.1.2024]. Dostupné na: https://crfm.stanford.edu/report.html.

o umelej inteligencii však upustil od použitia pojmu "základný model" a nahradil ho pojmom "všeobecný systém umelej inteligencie", *general-purpose AI (GPAI) system*, ktorý je prakticky synonymom k pojmu základný model. Na účely tejto kapitoly však budeme ďalej používať pojem základný model.

Základný model si môžeme predstaviť ako univerzálnu platformu pre aplikácie AI. Tieto modely majú preto rastúci význam pre mnohé nadväzujúce (downstreamové) aplikácie.

Technológia, na ktorej sú postavené základné modely však už existovala istý čas. Vyrástla na hlbokých neurónových sieťach (*deep neural networks*), algoritmoch samoučenia (*self-supervised learning*), prenosného[16] učenia *(transfer learning*) a veľkých datasetoch. Pokroky vo výskume, inžinierstve a superpočítačoch, najmä v škálovaní týchto metód na stále väčšie množiny trénovacích údajov a väčšie výsledné modely (mnoho miliárd parametrov), viedli k bodu, keď tieto modely začali vykazovať nové schopnosti a stali sa všeobecnejšie použiteľnými. Nové schopnosti sa začali vynárať (emergovať, *to emerge*) nad rámec použitých učiacich úloh počas tréningového postupu. Napríklad, veľký jazykový model ktorý bol pôvodne trénovaný na doplnenie chýbajúceho slova vo vete je s malou úpravou možné použiť ako znalostnú bázu, ktorú je možné dotazovať otázkami formulovanými v prirodzenom ľudskom jazyku. Spolu s emergenciou sa kľúčovou charakteristikou základných modelov stala aj homogenizácia. Homogenizácia sa prejavuje ich efektívnou využiteľnosťou v mnohých oblastiach, vďaka čomu sú tieto modely základom pre ďalšie aplikácie, ktoré sú na týchto základných modeloch postavené.

Základný model je možné doladiť (*fine-tune*) pre širokú škálu nadväzujúcich aplikácií, najmä pre vytváranie aplikácií pre koncových zákazníkov. Najnovšie základné modely pracujú s viacerými typmi údajov. Sú multimodálne, čo znamená, že dokážu spracovať informácie nielen v textovom formáte, ale aj obrázky, zvuk či video.

Základné modely sú veľmi nákladné a časovo náročné na vývoj a predtrénovanie. Náklady na výpočtový výkon potrebný pre trénovanie modelu *Chinchilla*, veľkého jazykového modelu od spoločnosti DeepMind z roku 2022, sa odhadujú na 2,1 milióna USD. Ďalej, model

[16] Cieľom *transfer learning* je schopnosť naučiť model voľačo na dátach z domény „A" a potom môcť tento model zmysluplne použiť aj na doméne „B", čiže toto učenie je prenosné.

PaLM, jeden z vlajkových lodí veľkých jazykových modelov spustených v roku 2022, mal 540 miliárd parametrov a stál odhadom 8 miliónov USD.[17] Veľké množstvo spoločností bude preto stavať svoje aplikácie na základnom modeli, čím sa využijú pokročilé možnosti základného modelu na zvýšenie výkonu pri ich špecifických úlohách (napr. vytváranie správ, sumarizácia textov atď.). Je teda zrejmé, že základných modelov bude len málo, ale očakáva sa nárast startupov, nových prípadov použitia a obchodných modelov v nadväzujúcom dodávateľskom reťazci AI.[18]

Ukázalo sa, že existujúce základné modely sú obzvlášť účinné v oblastiach, ako je spracovanie prirodzeného jazyka a počítačové videnie. Jazyk bol najbežnejšou triedou významných systémov strojového učenia AI vydaných v roku 2022, približne šesťkrát viac ako ďalší najbežnejší typ systému, multimodálne systémy.[19]

Príkladom je veľký jazykový model *Gemini*, ktorý spoločnosť *Google* predstavila v decembri 2023. Google považuje Gemini za doteraz najflexibilnejší model, ktorý dokáže efektívne fungovať na všetkých zariadeniach od dátových centier až po mobilné zariadenia. Jeho najmodernejšie možnosti výrazne zlepšia spôsob, akým vývojári a podnikoví zákazníci vytvárajú a škálujú AI. Gemini sa integruje aj s produktmi Google, ako je *Bard* (nástroj pre četovanie).[20]

Podobne, spoločnosť Microsoft uvádza niekoľko nových funkcií do jej služby Copilot, vrátane najnovších modelov od spoločnosti OpenAI. Copilot získa podporu pre GPT-4 Turbo spolu s aktualizovaným modelom DALL-E 3, novou funkciou *Code interpreter*[21] a funkciou hĺbkového vyhľadávania *Deep Search* v rámci služby *Bing*.[22]

---

[17] NESTOR M., et al.: *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023. s. 62. [online] [30.12.2023]. Dostupné na: https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

[18] MĄDRY A.: *Written Statement for the Hearing, Advances in AI: Are We Ready for a Tech Revolution?* In Front of the House Cybersecurity, Information Technology, and Government Innovation Subcommittee." 2023. [online] [30.12.2023]. Dostupné na: https://oversight.house.gov/wp-content/uploads/2023/03/madry_written_statement100.pdf.

[19] Tamže. s. 49.

[20] PICHAI S., HASSABIS D.: *Introducing Gemini: our largest and most capable AI model*. 6.12.2023. [online] [19.1.2024]. Dostupné na: https://blog.google/technology/ai/google-gemini-ai/#sundar-note

[21] Ide o plug-in, umožňuje nielen prijímať súbory, ale aj pripraviť ich na stiahnutie, takže už nemusíte zložito kopírovať vygenerované tabuľky. Príkladom promptu môže byť nahrať údaje projektu vo formáte CSV a žiadať aby Bard analyzoval výkonnosť projektu, identifikoval potenciálne oneskorenia a pripravil správu pre tím.

[22] WARREN, T.: *Microsoft's Copilot is getting OpenAI's latest models and a new code interpreter.* 5.12.2023 [online] [19.1.2024]. Dostupné na: https://www.theverge.com/2023/12/5/23989052/microsoft-copilot-gpt-4-turbo-openai-models-code-interpreter-feature.

Spoločnosť Meta oznámila príchod novej verzie svojho veľkého jazykového modelu LLaMA 3.[23]

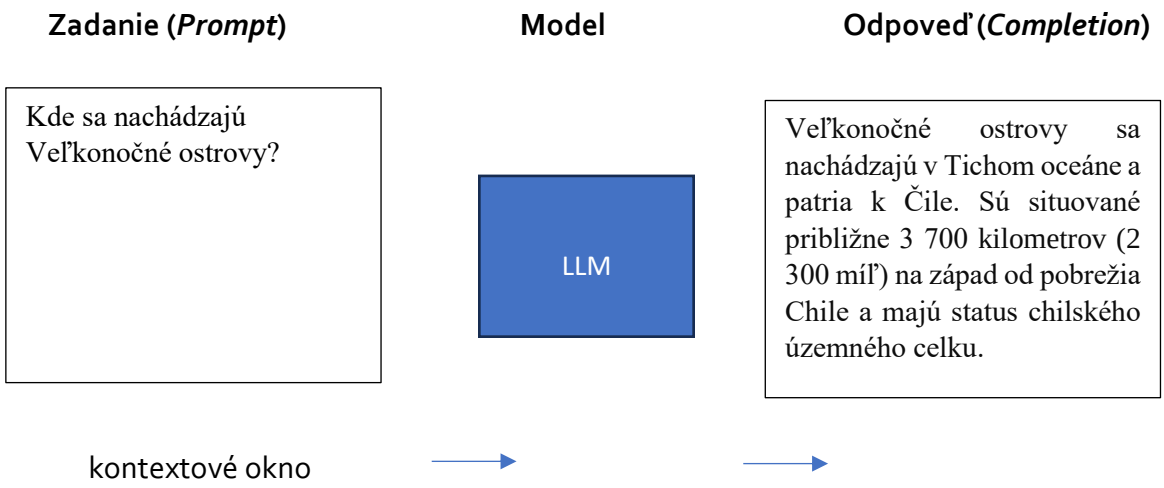
Ako je zrejmé, medzi typické príklady základných modelov patria veľké jazykové modely (porovnaj nižšie). Na ilustráciu, pre pôvodný ChatGPT slúžil ako základný model veľký jazykový model s názvom GPT-3.5. Ak to trochu zjednodušíme, spoločnosť OpenAI použila niektoré údaje špecifické pre chat na vytvorenie upravenej verzie GPT-3.5, ktorý bol špecializovaný na dobré fungovanie v prostredí chatbotov, a potom ho zabudovala do ChatGPT.

V súčasnosti sa pojem *základný model* často používa ako synonymum pojmu *veľký jazykový model*, pretože jazykové modely sú v súčasnosti najjasnejším príkladom systémov so širokými možnosťami, ktoré možno prispôsobiť na konkrétne účely. Podstatný rozdiel medzi týmito pojmami spočíva v tom, že veľké jazykové modely sa konkrétne vzťahujú na systémy zamerané na jazyk, zatiaľ čo základný model sa snaží vsadiť na širší koncept založený na funkciách, ktorý by sa v budúcnosti mohol rozšíriť o nové typy systémov.

1.10 Veľké jazykové modely (*LLMs*)

Veľké jazykové modely (*Large Language Models,* LLM) sú druhom generatívnej umelej inteligencie, ktorý je primárne zameraný na porozumenie a generovanie prirodzeného jazyka. Väčšina veľkých jazykových modelov je zároveň základným modelom, ako sme už uviedli vyššie. Podobne ako iné univerzálne technológie, ako je hlboké učenie, majú LLM potenciál ovplyvniť širokú škálu aplikácií v rôznych odvetviach, pričom niektorí autori prirovnávajú prelomovosť ich vplyvu k elektrifikácii na prelome 18. a 19. storočia.[24] Tieto modely sú trénované na obrovskom množstve údajov, čo im umožňuje generovať text podobný ľudskému a vykonávať zložité úlohy s pozoruhodnou presnosťou.

[23] HEATH, A.: *Mark Zuckerberg's new goal is creating artificial general intelligence*. 18.1.2024 [online] [19.1.2024]. Dostupné na: https://www.theverge.com/2024/1/18/24042354/mark-zuckerberg-meta-agi-reorg-interview.
[24] LYNCH, S.: *Andrew Ng: Why AI Is the New Electricity*. Stanford Graduate School of Business. 11.03.2017. [online] citované [19.1.2024]. Dostupné na: https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity.



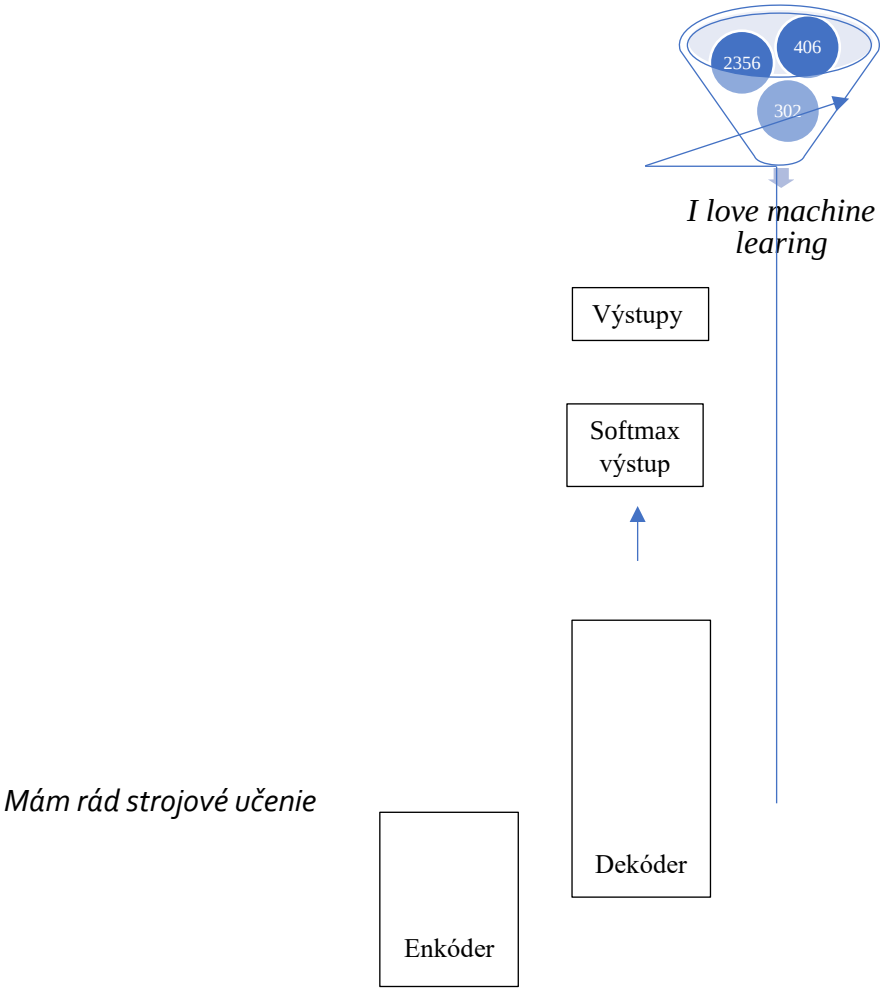
Obr. 3. Typické využitia jazykového modelu. Zodpovedanie otázok; na vstupe je otázka (všeobecne nazývaná ako dotaz, *prompt*) a na výstupe očakávame odpoveď.

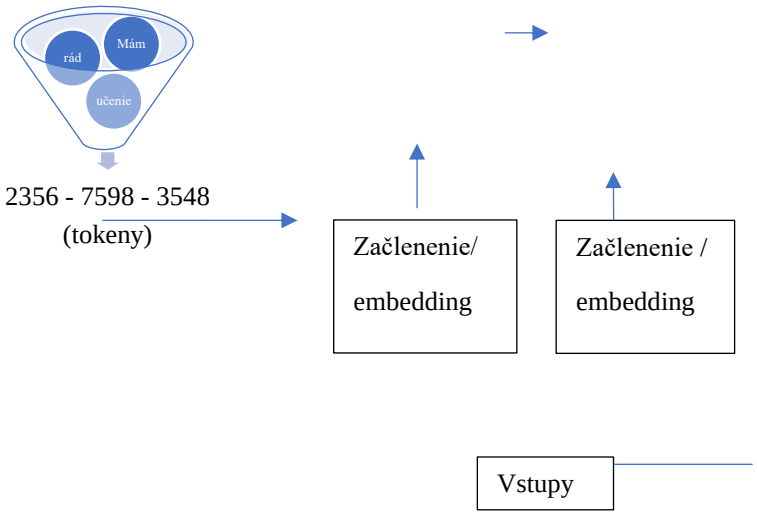
LLM sú používané na riešenia úloh spracovania prirodzeného jazyka (NLP). Prompt obsahuje počiatočný text spolu s označením typu úlohy (*task*), LLM potom slovo po slove vygeneruje odpoveď. Z technického pohľadu, model vytvára odpoveď ako čo najpravdepodobnejšie doplnenie daného dotazu (*completion*). Okrem zodpovedania otázok možno jazykové modely využitia aj na tvorbu esejí, sumarizáciu textu, preklad textu, výber informácií a pod.

Architektúra transformátora (*Transformer*) spôsobila revolúciu v oblasti spracovania prirodzeného jazyka a stala sa základom pre LLM, ktoré dnes poznáme. Bola predstavená vo výskumnej práci "*Attention is All You Need*" (Pozornosť je všetko, čo potrebujete) publikovanej v roku 2017 výskumníkmi spoločnosti Google.[25] Jedna z jej kľúčových predností spočíva v mechanizme tzv. vlastnej pozornosti (*self-attention*). Ten modelu umožňuje pochopiť význam slova vo vete vďaka na mieru šitému upriameniu pozornosti na dôležité slová v jeho kontexte, t.j. každé slovo má vlastný vzorec pozornosti. Priradením váh pozornosti vzťahom medzi slovami, hoci aj ďaleko vzdialenými, model získava komplexné porozumenie jazyka a vie tak, okrem iného, ľahko rozlíšiť homonymá podľa ich širšieho kontextu.

[25] VASWANI A.; et al., '*Attention Is All You Need*'. [online] [19.1.2024]. 2017 Dostupné na: https://arxiv.org/abs/1706.03762.

Architektúra Transformátora (Obr. 4) sa skladá z dvoch hlavných komponentov: *enkodéra* a *dekodéra*, pričom oba majú spoločné črty. Ešte pred vložením textu do modelu je však potrebné slová konvertovať na číselné reprezentácie, tzv. tokeny, pomocou *tokenizátora*. To umožňuje modelu pracovať s číslami a nie so slovami. Začleňovacia (*Embedding*) vrstva potom mapuje tokeny na vysokorozmerné vektory, ktoré v konečnom dôsledku (ako výsledok procesu učenia) kódujú význam a kontext každého tokenu. Tieto kódové vektory sú jedinečné pre daný token (slovo) a jeho kontext, čo uľahčuje matematické pochopenie jazyka. Okrem toho sa pridáva pozičné kódovanie, aby sa zachovala informácia o poradí slov. Takto zakódované tokeny sa odovzdávajú do vrstvy Transformátora so *self-attention* mechanizmom, kde model analyzuje vzťahy medzi nimi. Paralelne sa učí viacero súborov váh pozornosti, známych ako hlavy pozornosti (*attention heads*), ktoré zachytávajú rôzne aspekty jazyka. Takýchto vrstiev sa opakuje niekoľko, napr. LLaMA-2 aplikuje 32 vrstiev Transformátora.





Obr. 4 – Architektúra transformátora s enkóderom a dekóderom (sekvencia - sekvencia).

Hoci sa model enkodér-dekodér bežne používa na úlohy typu sekvencia-sekvencia, ako je preklad zobrazený na obr. č. 4 vyššie, existujú aj jeho varianty. Modely využívajúce iba enkodér sú účinné pri úlohách, ako je analýza nálad (*sentiment*), napríklad BERT od spoločnosti Google, zatiaľ čo modely využívajúce iba dekodér sa stali veľmi univerzálnymi. Medzi populárne modely využívajúce iba dekodér patrí rodina modelov GPT, LLaMA a mnohé ďalšie.

V oblasti spracovania prirodzeného jazyka sa ako účinná stratégia na zlepšenie výkonnosti jazykových modelov ukázalo prompt inžinierstvo (*prompt engineering*). Uvedením príkladov alebo dodatočných údajov do promptu, čo je technika známa ako vnútrokontextové učenie (*in-context learning*), môžu modely lepšie porozumieť a vykonávať špecifické úlohy.

1.10.1 Halucinácie (*Hallucinations*)

Halucinácie sú prípady, keď modely, najmä veľké jazykové modely ako GPT-3 alebo GPT-4, produkujú výstupy, ktoré sú síce koherentné a gramaticky správne, ale vecne nesprávne alebo nezmyselné. "Halucinácie" v tomto kontexte znamenajú generovanie nepravdivých alebo zavádzajúcich informácií. Tieto halucinácie sa môžu vyskytnúť v dôsledku rôznych faktorov, ako sú obmedzenia v trénovacích údajoch, skreslenie údajov alebo prirodzená zložitosť jazyka.

Halucinácie LLM sa však týkajú najmä oblastí, ktoré si vyžadujú vysokú úroveň presnosti a majú významný vplyv na životy ľudí, ako je zdravotníctvo, právo alebo inžiniering. Negatívne dôsledky halucinácií sa stali známymi najmä v oblasti poskytovania právnych služieb.

Veľké jazykové modely majú potenciál zmeniť poskytovanie právnych služieb, ale tento potenciál je ohrozený prítomnosťou právnych halucinácií - reakcií týchto modelov, ktoré nie sú v súlade s právnymi skutočnosťami. Podľa štúdie vedcov z Univerzity v Stanforde z januára 2024 sú právne halucinácie alarmujúco časté, vyskytujú sa v 69 % prípadov s ChatGPT 3.5 a 88 % prípadov s Llama 2, keď sa týmto modelom kladú konkrétne, overiteľné otázky o náhodných prípadoch amerického federálneho súdu. Podľa uvedenej štúdie modely nedokážu vždy predvídať, resp. nie vždy sami o sebe vedia, kedy vytvárajú právne halucinácie. Celkovo tieto zistenia varujú pred rýchlou a nekontrolovanou integráciou populárnych LLM do právnych služieb.[26]

V máji 2023 sa stal známym prípad advokáta zastupujúceho klienta, ktorý žaloval leteckú spoločnosť.[27] Pri príprave podania na súd sa advokát spoliehal na umelú inteligenciu. Predložil podanie, ktoré z veľkej časti vygeneroval ChatGPT, s citáciami na neexistujúce súdne rozhodnutia. V dôsledku toho prípadu dokonca predseda Najvyššieho súdu Spojených štátov John Roberts upozornil na "halucinácie" veľkých jazykových modelov vo výročnej správe o federálnom súdnictve za rok 2023.[28]

1.10.2 Výpočtová výkonnosť modelov (*FLOP* a *MAC*)

Operácie s pohyblivou rádovou čiarkou FLOP (*Floating Point Operations*) a operácie násobenia a akumulácie MAC (*Multiply-Accumulate Operations*) sú metriky, ktoré sa bežne používajú na meranie výpočtovej zložitosti modelov hlbokého učenia. Predstavujú rýchly a jednoduchý spôsob, ako pochopiť počet aritmetických operácií potrebných na vykonanie daného výpočtu. Pri práci s AI modelmi slúžia MAC alebo FLOP na odhad kvalitatívneho

[26] DAHL, M., et al. *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*. [online] [27.1.2024]. Dostupné na: arXiv preprint arXiv:2401.01301, 2024.
[27] WEISER, B.: *Here's What Happens When Your Lawyer Uses ChatGPT*. The New York Times. 27 Máj 2023. [online] [27.1.2024]. Dostupné na: https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html
[28] 2023 Year-End Report on the Federal Judiciary. [online] [27.1.2024]. Dostupné na: https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf

výkonu modelu, nakoľko približne platí, že s rastúcim množstvom výpočtu ktorý model spotrebuje, rastie aj jeho kvalitatívny výkon. Obe metriky sú len odhadom a nie presným vyjadrením výkonnosti daného modelu. Stále však môžu poskytnúť veľmi užitočné poznatky aj o spotrebe energie alebo výpočtových požiadavkách.

FLOP sa konkrétne vzťahuje na počet operácií s pohyblivou rádovou čiarkou, ktoré zahŕňajú operácie sčítania, odčítania, násobenia a delenia čísel s pohyblivou rádovou čiarkou. Tieto operácie sú rozšírené v mnohých matematických výpočtoch zapojených do strojového učenia, ako sú násobenie matíc, aktivácie a výpočty gradientu. FLOP sa často používajú na meranie výpočtových nákladov alebo zložitosti modelu alebo konkrétnej operácie v rámci modelu. To je užitočné, keď potrebujeme poskytnúť odhad celkového počtu potrebných aritmetických operácií, ktorý sa vo všeobecnosti používa v kontexte merania efektivity výpočtov.

Je potrebné si ešte uvedomiť rozdiel medzi FLOP a *FLOPS (floating point operations per second)*. FLOPS sú operácie s pohyblivou rádovou čiarkou za sekundu. Gigaflops (t.j. 10^9 = 10^9 FLOPS) predstavuje približne jednu miliardu týchto operácií za sekundu. Pre ilustráciu, priemerný notebook môže fungovať v rozmedzí od 250 gigaflops do 400 gigaflops – čo stačí na surfovanie na internete, prácu s kancelárskym softvérom, či hranie hier. Systémy vysokovýkonnej výpočtovej techniky (*high-performance computing*, HPC), sa merajú už v petaflops: jeden milión miliárd operácií za sekundu. LUMI superpočítač vo Fínsku má 550 petaflops (t.j. 550 x 10^15 FLOPS).[29]

Na druhej strane, MAC (*Multiply-Accumulate Operations*) počíta len počet operácií násobenia a sčítania, ktoré zahŕňajú násobenie dvoch čísel a sčítanie výsledku. MAC sa často používajú ako špecifickejšia miera výpočtovej zložitosti v modeloch, ktoré sa vo veľkej miere spoliehajú na operácie lineárnej algebry, ako sú konvolučné neurónové siete (konvolučné siete sú popísané vyššie).

Význam FLOP je podstatný aj z pohľadu regulácie umelej inteligencie. Nariadenie o umelej inteligencii priamo upravuje FLOP a od (nielen) hodnoty FLOPov odvodzuje riziká

[29] Európska komisia: Vyspelá výpočtová technika. [online] [27.01.2024]. Dostupné na: https://digital-strategy.ec.europa.eu/sk/policies/advanced-computing

modelu. Kumulatívne množstvo výpočtov použitých na trénovanie všeobecného modelu umelej inteligencie (*general-purpose artificial intelligence*, GPAI) merané na základe FLOPov je jedným z relevantných parametrov pre schopnosti modelu. Nariadenie o umelej inteligencii upravuje prezumpciu, že ide o GPAI model so systémovými rizikami, ak kumulatívny objem výpočtov použitých na jeho tréning je väčší ako 10^25 FLOP.[30]

Obdobne, výkonný príkaz (*Executive order*) prezidenta USA Joe Bidena o bezpečnej, chránenej a dôveryhodnej umelej inteligencii[31] považuje model s potenciálnymi schopnosťami, ktoré by sa mohli použiť pri škodlivej kybernetickej činnosti: ak si vyžaduje výpočtový výkon väčší ako 10^26 FLOPov a je natrénovaný na výpočtovom klastri, ktorý má súbor strojov fyzicky umiestnených v jednom dátovom centre, prepojených sieťou s rýchlosťou viac ako 100 Gbit/s a má teoretickú maximálnu výpočtovú kapacitu 10^20 FLOPS na trénovanie AI.

Podľa Európskej komisie hranica 10^25 FLOP zachytáva aktuálne najpokročilejšie modely GPAI, konkrétne GPT-4 od OpenAI a pravdepodobne Gemini od Google DeepMind. Táto hranica sa určila preto, že schopnosti modelov nad touto hranicou ešte nie sú dostatočne preskúmané. Mohli by predstavovať systémové riziká, a preto sa na nich bude vzťahovať ďalší súbor povinností. Počíta sa s tým, že presná prahová hodnota FLOP sa bude môcť aktualizovať smerom nahor alebo nadol, napr. vzhľadom na pokrok v objektívnom meraní schopností modelu a vývoj výpočtového výkonu potrebného na danú úroveň výkonu.[32]

FLOP však nemôže byť (a nie je) jediným spôsobom, ktorým sa počíta efektivita výpočtu. Pri odhadovaní efektívnosti modelu sa za potrebné považuje mnoho ďalších faktorov. Napríklad to, aké je paralelné nastavenie systému; akú má model architektúru; aký výpočtový akcelerátor model používa.

[30] Nariadenie o umelej inteligencii, recitál č. 600, článok 52a ods. 2.
[31] Príkaz prezidenta (*Executive Order*) 14110 zo dňa 30.10.2023. *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.* A Presidential Document by the Executive Office of the President, publikované dňa 01.11.2023 vo Federálnom registri. [online] citované [15.12.2023]. Dostupné na: https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.
[32] Európska Komisia: *Artificial Intelligence – Questions and Answers*. 12.12.2023 [online] [27.01.2024]. Dostupné na: https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683.

Nakoľko modely hlbokého učenia sú náročné na výpočet, používajú sa pre ne rôzne typy výpočtových akcelerátorov, tiež známych ako AI akcelerátory (*AI accelerators*) či neurónové procesory (*Neural network accelerator chips*).

AI akcelerátory sú špecializované hardvérové alebo softvérové komponenty, ktoré sú navrhnuté na zrýchlenie výpočtových operácií strojového učenia, teda pre masívne paralelne výpočty základných algebrických operácií, ako napr. násobenie matíc. Medzi najpoužívanejšie akcelerátory patria grafické procesory GPU (*Graphical Process Unit*) najmä od spoločnosti Nvidia, ďalej TPU (*Tensor Processing Unit*) vyvinutý spoločnosťou Google, alebo aj ASIC (*Application-Specific Integrated Circuit*). Súčasťou budúceho vývoja AI akcelerátorov sú kvantové AI akcelerátory, ktoré sú vo vývoji.

1.11 Súlad s hodnotami (*Value alignment*)

Pri navrhovaní systémov umelej inteligencie chceme zabezpečiť, aby ich "hodnoty" (ciele) boli v súlade s etickými hodnotami. Niekedy sa to nazýva problém zosúladenia hodnôt (*value alignment*). Je to náročné z troch dôvodov. Po prvé, je ťažké úplne a správne definovať naše hodnoty. Po druhé, je ťažké zakódovať tieto hodnoty ako ciele modelu umelej inteligencie a po tretie, je ťažké zabezpečiť, aby sa model naučil tieto ciele plniť.

Add 1. V modeli strojového učenia je stratová funkcia (*loss function*) náhradou za naše skutočné ciele. Zoberme si príklad agenta, ktorého učíme hrať šach. Ak je agent odmenený za vyhadzovanie figúrok, môže to viesť k mnohým remízam namiesto skutočne želaného cieľa vyhrať hru.

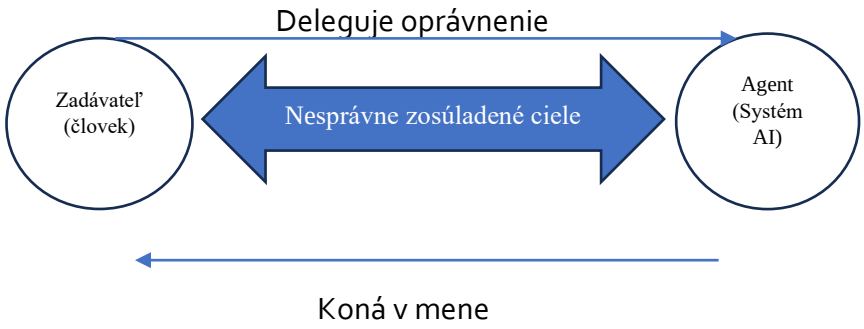
Add 2. Jedným zo spôsobov, ako uvažovať o zosúladení hodnôt, je *štrukturálny problém*, ktorý vzniká, keď ľudský zadávateľ deleguje úlohy na umelého agenta.[33] Je to podobné problému v obchodných vzťahoch, ktorý pripúšťa, že v každom vzťahu, v ktorom sa očakáva, že jedna strana bude konať v najlepšom záujme druhej strany, existujú konkurenčné stimuly. V kontexte umelej inteligencie môžu takéto konflikty záujmov vzniknúť buď vtedy, keď (i) sú

[33] LACROIX. T,: *The Linguistic Blind Spot of Value-Aligned Agency, Natural and Artificial*. 2022. online. Dostupné na: *arXiv preprint arXiv:2207.00868.*

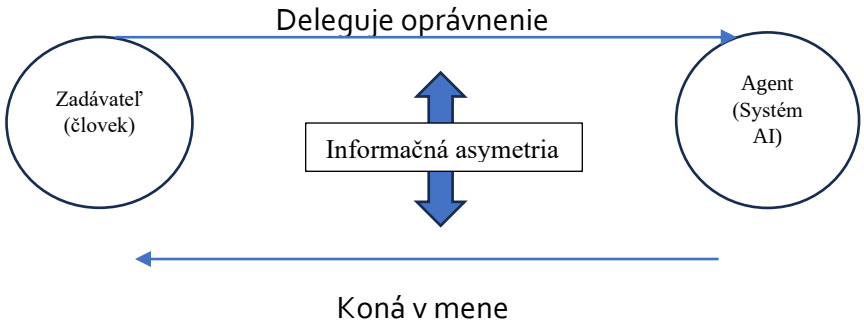
ciele nesprávne špecifikované, alebo (ii) existuje informačná asymetria medzi zadávateľom a agentom.

Mnohé témy v etike umelej inteligencie možno chápať v zmysle tohto štrukturálneho pohľadu na zosúladenie hodnôt. Nakoľko sa bližšie otázkam etiky umelej inteligencie venuje kapitola 9., poukážeme ďalej na technologické východiská problémov skreslenia, transparentnosti a vysvetliteľnosti (oba súvisia s informačnou asymetriou).

a)



b)



Obr. 5. Štrukturálny opis problému zosúladenia hodnôt. Problémy vyplývajú z a) nesprávne zosúladených cieľov (napr. skreslenie) alebo b) informačnej asymetrie medzi (ľudským) zadávateľom a (umelým) agentom (napr. nedostatočná vysvetliteľnosť). [34]

[34] Prevzaté z LaCroix, T. (2023). *Artificial Intelligence and the Value-Alignment Problem: A Philosophical Introduction*. online. Dostupné na: https://value-alignment. github.io.

1.11.1 Skreslenie a spravodlivosť (*Bias and Fairness*)

Skreslenie údajov (*Data Bias*) v strojovom učení je typ chyby, pri ktorej sú niektoré prvky súboru údajov hodnotené a/alebo zastúpené viac ako iné. Skreslenie do modelov možno vniesť rôznymi spôsobmi.

Začína to špecifikáciou problému a výberom cieľov modelu, ktoré si vyžaduje hodnotový úsudok o tom, čo je pre nás dôležité. Tým sa ale umožňuje vytvárať alebo prenášať na model skreslenia. Ďalej, tzv. algoritmické skreslenie môže vzniknúť, keď je dataset nereprezentatívny alebo neúplný. Pri kvantifikácii či hodnotení spravodlivosti modelu sa tiež vyžaduje hodnotový úsudok, nakoľko existuje na výber viacero exaktných matematických definícií, ktoré jednotlivo majú správnu intuíciu, ale prekvapivo, sa vzájomne vylučujú. Preto je možno podstatnejšie vecné hodnotenie toho, či algoritmy vykazujú spravodlivosť v praxi. Taktiež nasadené algoritmy môžu interagovať s inými algoritmami, štruktúrami alebo inštitúciami v spoločnosti a vytvárať komplexné spätné väzby, ktoré upevňujú pretrvávajúce skreslenia. Napríklad veľké jazykové modely sa trénujú na údajoch z internetu. Keď sa však výstupy modelov zverejnia na internete, trénovacie údaje pre budúce modely sa znehodnotia. To môže prehĺbiť skreslenia a vytvoriť nové spoločenské škody.

Pri analyzovaní chybovosti modelu strojového učenia, je potrebné od skreslenia údajov odlišovať odchýlku údajov (*Data Variance*). Zatiaľ čo skreslenie údajov je chyba reprezentácie alebo váhy údajov, odchýlka údajov je množstvo nejednoznačnosti samotných dát. Výsledná chyba sa tak skladá z oboch častí. Cieľom pri vývoji modelov je čo najviac znížiť skreslenie údajov aj odchýlku údajov, aby sa získali presnejšie výstupy.

1.11.2 Transparentnosť

Transparentnosť možno vnímať na niekoľkých úrovniach. *Funkčná transparentnosť* sa týka znalosti algoritmického fungovania systému (t. j. logických pravidiel, ktoré mapujú vstupy na výstupy). *Štrukturálna transparentnosť* zahŕňa znalosť toho, ako program vykonáva algoritmus. To môže byť zastreté, keď sa príkazy napísané vo vysokoúrovňových programovacích jazykoch vykonávajú pomocou strojového kódu. A napokon transparentnosť chodu vyžaduje pochopenie toho, ako bol program spustený v konkrétnom

prípade. V prípade hlbokých sietí to zahŕňa znalosti o hardvéri, vstupných údajoch, trénovacích údajoch a ich interakciách. Nič z toho sa nedá zistiť podrobným skúmaním kódu.

Napríklad sieť GPT-3 je funkčne transparentná; jej architektúra je opísaná.[35] Nevykazuje však štrukturálnu transparentnosť, pretože nemáme prístup ku kódu, a nevykazuje ani transparentnosť chodu, pretože nemáme prístup k naučeným parametrom, hardvéru ani trénovacím údajom. Následná verzia GPT-4 nie je transparentná vôbec. Podrobnosti o fungovaní tohto komerčného produktu nie sú v dobe písania toto textu (január 2024) známe.

1.11.3 Vysvetliteľnosť (*Explainability*)

Aj keď je systém transparentný, neznamená to, že môžeme pochopiť, ako sa rozhodnutie prijíma alebo na základe akých informácií sa prijíma. Hlboké siete môžu obsahovať miliardy parametrov, takže nie je možné, aby sme len na základe skúmania pochopili, ako fungujú. Jednou z mierne úspešných prístupov je vytváranie lokálnych vysvetlení. Hoci nevieme vysvetliť celý systém, niekedy vieme opísať, ako bol klasifikovaný konkrétny vstup. Ešte sa ukáže, či je možné vytvoriť komplexné rozhodovacie systémy, ktoré budú úplne zrozumiteľné pre ich používateľov alebo dokonca pre ich tvorcov. V súčasnosti neexistuje konkrétna definícia toho, čo znamená, že systém je vysvetliteľný, zrozumiteľný alebo interpretovateľný.

[35] BROWN, T., et al. *Language models are few-shot learners*. Neural Information Processing Systems, 33, 1877– 1901. (2020)

2. Prostriedky regulácie umelej inteligencie

Regulácia je jedným z nevyhnutných nástrojov na budovanie dôvery v spoločnosti.[36] Sme hlboko presvedčení, že esenciou regulácie je, že samotná regulácia nie je a nikdy by ani nemal byť účelom, ale vždy iba prostriedkom. Táto téza platí aj v dnešnej modernej spoločnosti, ktorá je dynamická a inovatívna. Naša spoločnosť je poháňaná rôznymi faktormi, vrátane politických zmien, vývoja trhu, inovatívneho pokroku a samozrejme aj pokroku vo vedeckých poznatkov.[37] Vedecké poznatky a technologický pokrok sa naďalej rýchlo vyvíjajú a takáto inovácia prináša stále niečo nové, s čím sú spojené nové vzrušujúce výhody ale aj neočkovateľné riziká. Veríme však, že aj v digitálnom veku by sa mala uplatňovať *„tradícia osvieteneckej éry, kde kategorický imperatív smeruje k digitálnemu svetu a technológie slúžia ľuďom, nie naopak*."[38]

Samotný pojem regulácia môže predstavovať akýkoľvek spôsobom úpravy spoločenského správania až po prijímanie vynútiteľných právnych noriem prostredníctvom štátu alebo medzinárodnej organizácie. V USA je regulácia vnímaná ako protipól slobodného trhu.[39] V rámci tejto učebnice budeme pojem regulácia používať na ilustráciu spôsobu verejnej kontroly nad určitými aktivitami, v našom prípade súvisiacimi s vývojom, nasadzovaním, testovaním a používaním umelej inteligencie. Samotná kontrola môže prebiehať prostredníctvom práva, morálnych noriem, noriem mäkkého práva, trhu, dizajnu alebo štandardov. Predmetná kontrola by podľa nášho názoru mala byť vykonávaná štátom prostredníctvom dezignovaných orgánov alebo medzinárodnou organizáciou.

Keďže súbežne s vývojom a inováciami sa budú vždy klásť aj otázky týkajúce sa vhodnosti, efektívnosti a kapacít režimov ktoré majú tieto nové technológie orámcovať, udržať a regulovať.[40] Systémy umelej inteligencie svojou jedinečnosťou tieto obavy ešte viac prehlbujú, pretože dynamický vývoj umelej inteligencie jednoducho ďaleko presahuje

[36] FÁBRY, B. - KASINEC, R. - TURČAN, M.: Teória Práva, Bratislava: Wolters Kluwer , 2019, ISBN 978-80-571-0127-7 s321
[37] Ibid.
[38] MARTINI, M.: Regulating Algorithms – How to demystify the alchemy of code?, Cambridge University Press, 2019. Dostupné na: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3458261
[39] Bližšie pozri Prosser, T. The Regulatory Enterprise: Government, Regulation, and Legitimacy (Oxford University Press 2010), s. 1 – 6.
[40] Porovnaj, GYURÁSZ, Z.: From principles to practice: regulating artificial intelligence, Míľniky práva v stredoeurópskom priestore 2020 [elektronický dokument] : zborník z online medzinárodnej vedeckej konferencie doktorandov a mladých vedeckých pracovníkov. - : 1. vyd. ISBN 978-80-7160-576-8. - Bratislava: Právnická fakulta UK, 2020. - S. 683-688 [online], 2020.

schopnosť akéhokoľvek tradičného regulačného prostredia udržať krok.[41] Pre tak rýchlo sa rozvíjajúcu a rozsiahlu technológiu, ako je umelá inteligencia, sa zdá, že komplexná regulácia pomocou noriem tvrdého práva nie je najideálnejšia, prinajmenšom nie z krátkodobého hľadiska.[42] Preto sa v ostatných rokoch normy mäkkého práva stali nevyhnutnou súčasťou rámca regulácie prakticky pre všetky nové technológie.

V rámci tejto kapitoly predstavíme modality regulácie umelej inteligencie a to konkrétne prostredníctvom práva, mäkkých noriem (soft law) a štandardov. Čitateľovi poskytneme pohľad na príklady dobrej praxe v rámci využitia spomínaných modalít a stručný prehľad očakávaní a ich využitia.

2.1. Regulácia prostredníctvom práva

Regulácia umelej inteligencie prostredníctvom práva a teda štátom (alebo medzinárodným zoskupením) vynútiteľných noriem je jednou z najdiskutovanejších otázok dnešných dní, hlavne v kontexte úsvitu verejného a komerčného využitia veľkých jazykových modelov a ich aplikácií ako ChatGPT alebo Bard. Nespornými výhodami regulácie prostredníctvom práva je ich transparentnosť, vynútiteľnosť a predvídateľnosť. Normy „tvrdého práva„ sú jednoducho prístupné verejnosti, ktorá navyše prostredníctvom odborných pripomienok môže vstupovať aj do legislatívneho procesu, či už na národnej alebo európskej úrovni. Zároveň sú tieto normy jednoducho vynútiteľné prostredníctvom aparátu orgánov presadzovania práva, dozorných orgánov alebo súdov. Právne normy by mali byť zároveň aj dostatočne predvídateľné s ohľadom na ich aplikáciu v konkrétnych prípadoch.

Na strane druhej, prijímanie legislatívy môže byť veľmi zdĺhavé, nemusí nutne zohľadňovať odborný prvok a minimálne v otázkach regulácie digitálneho priestoru a jeho nástrojov je častokrát pomalá a vyžaduje globálnejšie riešenie.

[41] MERCHANT, G.: Addressing the Pacing Problem 2011. Dostupné na: https://www.researchgate.net/publication/303158895_Addressing_the_Pacing_Problem

[42] Rýchle tempo rozvoja a množstvo foriem a možností aplikácie týchto zariadení, neexistencia jednoznačných odpovedí na kritické otázky spôsobujú, že uskutočnenie komplexnej tradičnej formy regulácie sa zdá byť zatiaľ prinajmenšom problematickou.

2.1.1 Opatrnosť verzus inovácia

Jedna z častých výčitiek regulácie nových technológií je, že brzdí inovačný vývoj spoločnosti. Pri umelej inteligencií je tento argument potrebné vyvážiť v kontexte krátkodobých a dlhodobých rizík, ktoré prináša. Na jednej strane je obava z reštriktívnej regulácie, ktorá zabrzdí technologický vývoj, vždy na mieste. Na strane druhej však neschopnosť regulovať aktuálne spoločenské problémy a reagovať na riziká môže predstavovať ešte väčší zásah do spoločnosti.

PRÍKLAD
Príkladom dynamického inovačného vývoja a pomalej legislatívy môžu byť sociálne siete, ich monetizovanie údajov a úlohy v spoločnosti. V dnešnej dobe sa prostredníctvom sociálnych médií cielia politické kampane a šíria dezinformácie. Legislatíva bola veľmi pomalá a na tieto fenomény reagovala až v ostatných rokoch.

Diskusia o regulácii AI je zložitá a rôznorodá. Na oboch stranách existujú opodstatnené argumenty a optimálny prístup pravdepodobne zahŕňa rovnováhu medzi podporou inovácií a zmierňovaním potenciálnych rizík. Najlepšia cesta vpred bude pravdepodobne zahŕňať kombináciu štátnej (tvrdej) regulácie, samoregulácie a etických usmernení.

2.1.2 Regulačné prístupy

Regulačné prístupy k umelej inteligencií prostredníctvom tvrdého práva môžu mať rôzne podoby. Na účely tejto učebnice môžeme rozlišovať minimálne štyri modely regulácie AI:

1. Model založený na posudzovaní rizika
2. Model založený na všeobecných princípoch
3. Sektorový model
4. Hybridné prístupy

Model založený na posudzovaní rizika kategorizuje systémy AI na základe ich potenciálneho rizika a uplatňuje prísnejšie predpisy na systémy s vyšším rizikom. V súčasnosti

získava na popularite vďaka svojej prispôsobivosti a zameraniu sa na špecifické riziká. Príkladom takéhoto prístup je aj Akt o umelej inteligencií.

Model založený na všeobecných princípoch sa zameriava na ustanovenie všeobecných princípov, ako je spravodlivosť, transparentnosť, zodpovednosť a vysvetliteľnosť, čo umožňuje flexibilitu pri vývoji technológie. Ponúka širší rámec, ale vyžaduje si viac výkladu a iné metódy presadzovania. Príkladom takéhoto modelu sú Princípy umelej inteligencie z dielne Organizácie pre hospodársku spoluprácu a rozvoj (OECD).[43]

PRÍKLAD
Princípy OECD pre umelú inteligenciu v skratke: Organizácia pre hospodársku spoluprácu a rozvoj (OECD) vypracovala päť kľúčových princípov a päť odporúčaní pre zodpovedné využívanie umelej inteligencie (AI). Tieto princípy sú základom pre budovanie dôveryhodnej AI, ktorá prináša prospech ľuďom a spoločnosti. **Princípy:** 1. **Hodnoty zamerané na človeka a spravodlivosť:** AI sa má budovať a využívať v prospech ľudí a spoločnosti s rešpektom voči ľudským právam a základným slobodám. 2. **Inkluzívny rast, udržateľný rozvoj a blahobyt**: Zainteresované strany by sa mali proaktívne zapojiť do zodpovedného spravovania dôveryhodnej umelej inteligencie v snahe o prospešné výsledky pre ľudí a planétu, ako je rozširovanie ľudských schopností a zvyšovanie kreativity, pokrok v začleňovaní nedostatočne zastúpených obyvateľov, znižovanie ekonomických, sociálnych, rodových a iných nerovností a ochrana prírodného prostredia. , čím sa posilní inkluzívny rast, trvalo udržateľný rozvoj a blahobyt 3. **Transparentnosť a vysvetliteľnosť**: AI by mala byť navrhnutá a používaná spôsobom, ktorý umožňuje pochopiť, ako funguje a aké rozhodnutia robí.

[43] K tomu pozri. OECD AI Principles overview. Dostupné na: https://oecd.ai/en/ai-principles.

4. **Robustnosť, bezpečnosť a spoľahlivosť**: AI by mala byť navrhnutá a používaná tak, aby bola bezpečná, spoľahlivá a odolná voči zneužívaniu.
5. **Zodpovednosť**: Tí, ktorí vyvíjajú, nasadzujú a používajú AI, by mali byť zodpovední za jej prevádzku a dôsledky.

Odporúčania:

1. **Vypracovať národné stratégie a politiky pre AI:** Vlády by mali vypracovať stratégie a politiky, ktoré podporujú zodpovedný vývoj a využívanie AI.
2. **Posilniť multiodborovú spoluprácu a investície do výskumu:** Vlády by mali spolupracovať so súkromným sektorom, občianskou spoločnosťou a akademickými kruhmi na riešení výziev a príležitostí súvisiacich s AI.
3. **Podpora digitálneho ekosystému pre AI:** Vlády by mali podporovať rozvoj a prístup k digitálnemu ekosystému pre dôveryhodnú AI, ktorý zahŕňa, digitálne technológie a infraštruktúru, mechanizmy na zdieľanie vedomostí o AI, mechanizmy na podporu bezpečného a etického zdieľania údajov, ako sú dátové trusty.
4. **Posilniť vzdelávanie a odbornú prípravu v oblasti AI:** Vlády by mali investovať do vzdelávania a odbornej prípravy, aby sa ľudia mohli pripraviť na pracovnú silu s podporou AI.
5. **Zabezpečiť medzinárodnú spoluprácu pre dôveryhodnú AI:** Vlády by mali spolupracovať na medzinárodnej úrovni, aby sa vytvoril globálny rámec pre dôveryhodnú AI.[44]

Sektorový model regulácie prispôsobuje predpisy konkrétnym sektorom, ako je zdravotná starostlivosť, financie alebo doprava, s ohľadom na ich jedinečné potreby a riziká. Ponúka cielené riešenia, ale môže viesť k roztrieštenej regulácii v rôznych sektoroch. Ilustráciou je regulácia algoritmického obchodovania v smernici Európskeho parlamentu a

[44] OECD. Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449 as amended on 08/11/2023. Dostupné na: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Rady 2014/65/EÚ z 15. mája 2014 o trhoch s finančnými nástrojmi, ktorou sa mení smernica 2002/92/ES a smernica 2011/61/EÚ (tzv. MIFID II).

PRÍKLAD

Smernica MiFID II upravuje rôzne aspekty finančných trhov v Európskej únii a algoritmické obchodovanie je jedným z nich.

Kľúčové požiadavky na algoritmické obchodovanie podľa MiFID II sú:

Systémy a kontroly: Investičné spoločnosti zapojené do algoritmického obchodovania musia mať:

- Odolné a robustné obchodné systémy, schopné zvládnuť vysoké objemy a udržiavať stabilitu.
- Vhodné prahové hodnoty a limity, aby sa predišlo chybným príkazom a narušeniam trhu.
- Efektívna kontrola rizika v podobe identifikovania a riadenia potenciálnych rizík spojených s algoritmickým obchodovaním.

Oznamovacie povinnosti: Firmy využívajúce algoritmické obchodovanie alebo poskytujúce priamy elektronický prístup musia informovať svoj príslušný dozorný orgán a obchodné miesto, ku ktorému majú prístup.

Vedenie záznamov: Firmy musia viesť záznamy o činnosti algoritmoch vrátane konkrétnych objednávok, dizajnu alebo zmien.[45]

Hybridné prístupy kombinujú vyššie uvedené modely. Príkladom je Singapore's Model AI Governance Framework, ktorý kombinuje prístup založený na všeobecných princípoch so špecifickým prístupom k niektorým sektorom ako zdravotníctvo a financie.[46]

[45] K tomu pozri bližšie články 17, 18, 34 alebo 42 danej smernice.
[46] Singapore's Approach to AI Governance. Dostupné na: https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework.

2.2. Regulácia prostredníctvom „mäkkého" práva

V posledných dekádach došlo k značnému dozrievaniu v téme regulácie. Od sedemdesiatich rokov minulého storočia sa príchodom nových teórii regulácie,[47] ako aj nárastom popularity spotrebiteľskej a environmentálnej regulačnej činnosti, stala regulácia celosvetovou témou.

Nárast popularity noriem mäkkého práva môže vysvetľovať skutočnosť, že normy mäkkého práva ponúkajú niekoľko potenciálnych výhod oproti tradičným regulačným nástrojom.[48] Najdôležitejšou z týchto výhod je, že tieto normy sa môžu prijať a revidovať relatívne rýchlo, bez toho, aby museli prejsť tradičným a zdĺhavým procesom tvorby právnych predpisov.[49] Taktiež je dôležitá skutočnosť, že právne normy v oblasti mäkkého práva nie sú viazané na konkrétnu jurisdikciu a môžu mať preto medzinárodné uplatnenie.[50] Podľa viacerých odborníkov je mäkké právo najvýhodnejšie, keď si aktéri chcú zachovať flexibilitu pri riešení neistôt a problémov, ktoré môžu nastať z dôvodu rýchlo meniacich sa okolností.[51]

Normy mäkkého práva však nie sú zlatom.[52] Ich nedokonalosť je často spôsobená najmä nedostatočnou vymožiteľnosťou noriem mäkkého práva. Ďalšou z kritík týchto noriem je časté použitie jazyka, ktorý je príliš nejasný alebo vágny.[53] Taktiež veľkým problémom noriem mäkkých právnych predpisoch je nedostatok koordinácie medzi veľkým počtom a

---

[47]Pozri bližšie STIGLER, G. J.: "The theory of economic regulation." Bell Journal of Economics and Management Science, 1971 s 3-21.

[48] LEE, R. - JOSE, P.D.: Self-interest, self-restraint and corporate responsibility for nanotechnologies: Emerging dilemmas for modern managers. 2008 Dostupné na:: https://www.researchgate.net/publication/233090728_Self-interest_self-restraint_and_corporate_responsibility_for_nanotechnologies_Emerging_dilemmas_for_modern_managers

[49] ABBOTT, K. - SNIDA, D: Hard and Soft Law in International Governance. 2000 International Organization, 421-456,Dostupné na:: https://www.cambridge.org/core/journals/international-organization/article/hard-and-soft-law-in-international-governance/EC8091A89687FDF7FC9027D1717538BF

[50]MERCHANT, G.: "Soft Law" Governance of Artificial intelligence 2019 Dostupné na: https://escholarship.org/content/qtojq252ks/qtojq252ks.pdf

[51] EREDÉLYI, O. GOLDSMITH, J.: Regulating Artificial intelligence Proposal for a Global Solution 2018 Dostupné na:: https://dl.acm.org/doi/pdf/10.1145/3278721.3278731?casa_token=gXFSlbp9ydUAAAAA%3APTrQCCNf5wrmMPXcJdTwIB9JxrN1U9HmM4PIO5T7KPML5TBQLOIblq6AjNbrJKWPTPYFb4KLCWF01w

[52] Samotná krása regulácie by mala vychádzať z dokonalosti. Podľa profesora Michio Kaku: *„Pre básnikov a umelcov je krása éterická, estetická kvalita, ktorá vyvoláva emócie a vášeň. Pre fyzikov je krása symetria. Lebo rovnice sú krásne, pretože majú symetriu".*[52] Môžeme teda doplniť, že pre nás právnikov je krása v regulácii. Aj keď osobne verím, že pán profesor hľadal význam krásy iba preto lebo nepoznal teba. (pozn. autora). KAKU, M.: The god equation, the quest for a theory of everything, Penguin random house uk, 2021. Isbn 978-0-241-48347-0 s 36..

[53] WEEKS, G.: SOFT LAW AND PUBLIC LIABILITY: BEYOND THE SEPARATION OF POWERS? 2018. Dostupné na: https://law.adelaide.edu.au/system/files/media/documents/2019-03/ALR_39%282%29_03_Weeks.pdf

potenciálne prekrývajúcich sa a možno dokonca nekonzistentných právnych predpisov.[54] Pre porovnanie, tradičné formy regulácie by znamenali vyššiu inštitucionálnu formalitu, ako aj kolektívnu kontrolu informácií s komplexnou centralizovanou správou a riadením.[55]

S ohľadom na „softlaw" reguláciu Európskej únie v oblasti umelej inteligencie bolo prvou lastovičkou Uznesenie Európskeho parlamentu zo 16. februára 2017 s odporúčaniami pre Európsku komisiu k normám občianskeho práva v oblasti robotiky (ďalej len „**Uznesenie**").[56] Prístup preferovaný v danom uznesení pri regulácií umelej inteligencie reflektuje podstatu prijatia regulácie v rámci iných rámcových predpisov a neregulovať technológiu v jednom právnom predpise komplexným spôsobom. Text Uznesenia zároveň výslovne spomína a zdôrazňuje otázky týkajúce sa bezpečnosti, ľudského bezpečia, zdravia, slobody, súkromia a integrity, dôstojnosti, sebaurčenia, rovnakého zaobchádzania a ochrany osobných údajov.[57] Následne v apríli 2018 publikovala Európska komisia Komuniké s názvom Umelá inteligencia pre Európu.[58] Komisia v tomto komuniké zvýraznila potrebu koherentného právneho rámca vo svetle digitálnej transformácie pri predaji produktov a poskytovaní služieb a zároveň potrebu dostupnosti dát, na základe ktorých môže umelá inteligencia fungovať.[59] V tomto smere je tak podľa Európskej komisie nevyhnutné prijať etický a právny rámec využívania údajov. Dostupnosť dostatočného množstva údajov reflektuje aj novelizácia smernice o opakovanom použití informácií verejného sektora, ktorá zvyšuje kvantitu údajov v dispozícií orgánov verejnej moci dostupných pre súkromný sektor. Predmetné plány len podčiarkuje prijatie Všeobecného nariadenia na ochranu údajov (ďalej len „**GDPR**"), ktoré je všeobecným právnym rámcom pre voľný pohyb osobných údajov v rámci EÚ. Na sklonku roka 2018 Európska komisia prijala tzv. Koordinovaný plán o umelej inteligencií (ďalej len „**Koordinovaný plán**").[60] Tento plán nadväzuje na záväzok všetkých

[54] Pri riešení týchto prípadov sa tvrdé právo spolieha na tradičné zásady derogácie, ale v prípade mäkkého práva neexistuje tradičná hierarchia medzi normami, a preto sa tieto normy stanú zásadami bez skutočných zásad.

[55] EREDÉLYI, O. - GOLDSMITH J: Regulating Artificial intelligence Proposal for a Global Solution 2018. Dostupné:https://dl.acm.org/doi/pdf/10.1145/3278721.3278731?casa_token=gXFSlbpgydUAAAAA%3APTrQCCNf5wrmMPXcJdTwIBgJxrN1U9HmM4PIO5T7KPML5TBQLOIblq6AjNbrJKWPTPYFb4KLCWF01w

[56] Uznesenie Európskeho parlamentu zo 16. februára 2017 s odporúčaniami pre Komisiu k normám občianskeho práva v oblasti robotiky. Dostupné na: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_SK.html

[57] Tamže, bod 11.

[58] Umelá inteligencia pre Európu. 2021. Dostupné na: https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe

[59] Communication to the Artificial Intelligence for Europe, s. 11 – 14. Dostupné na: https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe

[60] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Coordinated Plan on Artificial Intelligence.

členských štátov EÚ z roku 2018 prijať predmetný dokument. Cieľom tohto dokumentu je prijať spoločné ciele a determinovať snahy pri implementácií umelej inteligencie do praxe v rámci únioného priestoru. V prvom rade každý členský štát je povinný prijať národnú stratégiu o umelej inteligencií, ktorá by mala reflektovať požiadavky strategických dokumentov na úrovni EÚ.[61] Následne na jar 2019 prijala Expertná skupina na vysokej úrovni pre umelú inteligenciu Etické usmernenia pre dôveryhodnú umelú inteligenciu (ďalej len **Etické usmernenie**).[62] Toto usmernenie výslovne zvýrazňuje rešpektovanie základných ľudských práv a slobôd v kontexte vývoja a používania umelej inteligencie.[63] Zakotvuje štyri základné zásady, ktoré vychádzajú z rešpektovania základných ľudských práv a slobôd.[64] Etické usmernenie zároveň uvádza sedem požiadaviek pri realizácií dôveryhodnej umelej inteligencie.[65] Vo februári 2020 prijala Európska komisia balík troch dokumentov, ktoré sa týkajú regulácie umelej inteligencie a súvisiacich aspektov: a) White Paper on Artificial Intelligence: a European approach to excellence and trust (ďalej ako - **White Paper**);[66] b) Commission Report on safety and liability implications of AI, the Internet of Things and Robotics (ďalej ako -**Správa o bezpečnosti a zodpovednosti**);[67] c) A European strategy for data (ďalej ako - **Európska dátová stratégia**).[68] Následne koncom roka 2020 vydal Európsky parlament tri uznesenia o etických a právnych aspektoch softvérových systémov s umelou inteligenciou: d) Uznesenie č. 2020/2012 (INL)[69] - Rámec pre etické aspekty umelej inteligencie, robotiky a súvisiacich technológií, e) Uznesenie č. 2020/2014 (INL)[70] - Režim

[61] Coordinated Plan on Artificial Intelligence, s. 2.
[62] EU High-Level Expert Group on AI Ethics guidelines for trustworthy AI. [online]. [10-8-2021]. Dostupné na: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai
[63] Etické usmernenie, s. 12 – 13.
[64] (I) Rešpektovanie ľudskej autonómie; (II) Prevencia ujmy; (III) Spravodlivosť; (IV) Vysvetliteľnosť.
[65] (I) Ľudský faktor a dohľad, (II) Technická odolnosť a bezpečnosť, (III) Správa súkromia a údajov, (IV) Transparentnosť, (V) Rozmanitosť, (VI) Nediskriminácia a spravodlivosť, (VII) Spoločenský a environmentálny blahobyt a zodpovednosť
[66] White Paper on Artificial Intelligence: a European approach to excellence and trust Dostupné na: https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
[67] Commission Report on safety and liability implications of AI, the Internet of Things and Robotics 2021. Dostupné na.: https://ec.europa.eu/info/files/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics_en
[68] A European strategy for data 2021. Dostupné na: https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf
[69] Rámec pre etické aspekty umelej inteligencie, robotiky a súvisiacich technológií, 2020 Dostupné na: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_SK.html
[70] Režim občianskoprávnej zodpovednosti za umelú inteligenciu, 2020 Dostupné na: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_SK.html

občianskoprávnej zodpovednosti za umelú inteligenciu a f) Uznesenie 2020/2015 (INI)[71] - Práva duševného vlastníctva v súvislosti s rozvojom technológií umelej inteligencie[72]

EÚ chce v rámci tohto systému vybudovať ekosystém excelencie a ekosystém dôvery.[73] Z právneho hľadiska uvádza riziká pre základné ľudské práva a slobody, a to konkrétne ochrana osobných údajov, ochrana súkromia, nediskriminácia a taktiež riziká z hľadiska bezpečnosti tovarov a efektívneho fungovania zodpovednostných vzťahov..

2.3 Regulácia prostredníctvom technických noriem (štandardov)

"Standards are the distilled wisdom of people with expertise in their subject matter."[74]

2.3.1 Technická normalizácia a technické normy

Technická normalizácia je proces vedúci k zjednoteniu podľa jednotných a presne daných noriem (štandardov, z angl. *standards*). Technická normalizácia je ustáleným pojmom upraveným v právnych predpisov v Slovenskej republike, prakticky však znamená to isté čo v angl. jazyku pojem *standardisation*. V niektorých textoch sa preto môžeme stretnúť aj s pojmom štandardizácia.

Technická normalizácia sa môže vzťahovať na produkty, služby, procesy, materiály, subjekty a systémy manažérstva. Cieľom technickej normalizácie je kvalita, bezpečnosť a spätná kompatibilita. Technická normalizácia sa skladá najmä z postupov vypracúvania, vydávania a zavádzania noriem.[75]

Technická norma je dokument, vytvorený na základe dohody a schválený uznaným nezávislým orgánom, ktorý poskytuje na všeobecné a opakované použitie pravidlá, pokyny, charakteristiky alebo výsledky činností a zameriava sa na dosiahnutie optimálneho stupňa

[71]Práva duševného vlastníctva v súvislosti s rozvojom technológií umelej inteligencie, 2020 Dostupné na: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_SK.html

[72] K týmto trom uzneseniam pozri bližšie poslednú kapitolu.

[73]White Paper on Artificial Intelligence: a European approach to excellence and trust s 3. Dostupné na: https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

[74]ISO: *Standards*. [online] [15.12.2023]. Dostupné na: https://www.iso.org/standards.html#:%7E:text=Standards%20are%20the%20distilled%20wisdom,trade%20associations%2C%20users%20or%20regulators.

[75] Úrad pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky: Definícia technickej normalizácie. [online] [15.12.2023]. Dostupné na: https://www.normoff.gov.sk/stranka/297/o-technickej-normalizacii/

poriadku v danej súvislosti.[76] Z obsahovaného hľadiska možno technickú normu vymedziť ako súbor pravidiel, usmernení, technických špecifikácií alebo výsledkov činností, ktoré odzrkadľujú aktuálny stav vedy a techniky, ktorý je výsledkom konsenzu (dohody) zainteresovaných strán a podlieha systematickým previerkam jej aktuálnosti (spravidla každých päť rokov).[77] Inými slovami povedané, technické normy sú kodifikovanou najlepšou praxou (tzv. *best practice*) teda dokumentom, ktorý obsahuje všeobecne uznávané technické riešenia, ktoré zároveň efektívne ošetrujú riziká.[78] Tieto typické charakteristiky technickej normy ju odlišujú nielen obsahom, ale aj procesom tvorby od legislatívneho procesu tvorby právnych predpisov.

Od všeobecne záväzných právnych predpisov, nie je možné efektívne očakávať špecifické technické detaily – to je úlohou technických noriem.[79]

Používanie noriem je všestranne výhodné, pretože ich uplatňovanie poukazuje na určitú úroveň kvality, bezpečnosti a spoľahlivosti daného produktu či služieb. Technická normalizácia je základným prvkom vylepšovania procesov, slúži ako meradlo vyspelosti produkcie daného podniku, ale tiež predchádza prekážkam v obchode a uľahčuje technickú spoluprácu.[80] Technická normalizácia je základným prvkom konkurencieschopnosti a prispieva k ochrane spotrebiteľov.

V oblasti IKT (informačných a komunikačných technológií) zohrávajú technické normy zásadnú úlohu pri dosahovaní interoperability nových technológií a môžu priniesť značné výhody priemyslu aj zákazníkom a spotrebiteľom. Normy pomáhajú trhom s IKT zostať otvorené a umožňujú zákazníkom a spotrebiteľom najširší výber produktov. [81] Vzhľadom na rozsiahlu integráciu AI technológie do digitálnych produktov a služieb ako aj vývoj samotných AI produktov a služieb bude technická normalizácia v oblasti umelej inteligencie spolu s

[76] ISO/IEC Guide 2: 2004 (prijatá ako STN EN 45020: 2007)
[77] Zo slovenskej definície technickej normy podľa § 3 ods. 1 Zákona o normalizácii.
[78] Kompetenčné a certifikačné centrum kybernetickej bezpečnosti: *Technické normy v kyberbezpečnosti.* [online] [15.12.2023]. Dostupné na: https://cybercompetence.sk/technicke-normy-v-kyberbezpecnosti/
[79] MAKATURA, I.: *Technická normalizácia v informačnej a kybernetickej bezpečnosti*. In: Bezpečnosť v praxi [on-line]. Žilina: Poradca podnikateľa, 2024. ISSN 2729-885X. [19.1.2024]. Dostupné na: https://www.bezpecnostvpraxi.sk/odborny-clanok/technicka-normalizacia-v-informacnej-a-kybernetickej-bezpecnosti.htm
[80] Tamže.
[81] Európska Komisia: *ICT standardisation*. [online] [15.12.2023]. Dostupné na: https://single-market-economy.ec.europa.eu/single-market/european-standards/ict-standardisation_en

napredujúcim výskumom tejto oblasti zohrávať kľúčovú rolu v dosahovaní vyššie uvedených cieľov.

Tieto normy však nie sú záväzné a ich dodržanie môže upraviť ako povinnosť iba všeobecne záväzný právny predpis.

Vzťah právneho predpisu a technickej normy je potrebné vnímať citlivo. Právne predpisy väčšinou v dostatočnom detaile nepopisujú požiadavky kladené na každý produkt, proces alebo službu. Potrebnú úroveň detailu obsahujú technické normy a právne predpisy by nemali nahrádzať úpravu technických noriem a upravovať podrobnosti, ktoré prináležia technickým normám. Príkladom prerastania úpravy právneho predpisu do detailov doposiaľ typických pre technické normy, je Delegované nariadenie Komisie (EÚ) 2023/1717 z 27. júna 2023, ktorým sa mení smernica Európskeho parlamentu a Rady 2014/53/EÚ, pokiaľ ide o technické špecifikácie nabíjacieho portu a komunikačného protokolu nabíjania pre všetky kategórie alebo triedy rádiových zariadení, ktoré možno nabíjať cez kábel.[82]

2.3.2 Vývoj noriem a technická normalizácia

Právna úprava technickej normalizácie je vychádza z nasledujúcich právnych predpisov:

1. Nariadenie (EÚ) č. 1025/2012 o európskej normalizácii, ktorým sa menia a dopĺňajú smernice Rady 89/686/EHS a 93/15/EHS a smernice Európskeho parlamentu a Rady 94/9/ES, 94/25/ES, 95/16/ES, 97/23/ES, 98/34/ES, 2004/22/ES, 2007/23/ES, 2009/23/ES a 2009/105/ES a ktorým sa zrušuje rozhodnutie Rady 87/95/EHS a rozhodnutie Európskeho parlamentu a Rady č. 1673/2006/ES (ďalej len „*Nariadenie o európskej normalizácii*“),
2. zákon č. 56/2018 Z. z. o posudzovaní zhody výrobku, sprístupňovaní určeného výrobku na trhu v znení neskorších predpisov,
3. zákon č. 55/2018 Z. z. o poskytovaní informácií o technickom predpise a o prekážkach voľného pohybu tovaru v znení neskorších predpisov,

[82] MAKATURA, I.: *Technická normalizácia v informačnej a kybernetickej bezpečnosti*.

4. zákon č. 60/2018 Z. z. o technickej normalizácii v znení neskorších predpisov (ďalej len „*Zákon o technickej normalizácii*").

Podľa § 3 ods. 1 Zákona o normalizácii technickou normou je

1. európska norma,
2. medzinárodná norma,
3. harmonizovaná norma a
4. národná norma.

Národná norma je norma prijatá národným normalizačným orgánom. V Slovenskej republike je národnou normou, slovenská technická norma.

2.3.2.1 Národné technické normy

Technická normalizácia je v Slovenskej republike v pôsobnosti Úradu pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky (ďalej len „*Úrad pre normalizáciu*").

Slovenskou technickou normou sa rozumie technická norma prijatá do sústavy slovenských technických noriem podľa Zákona o technickej normalizácii. Slovenská technická norma má značku „STN". *Pôvodnou* slovenskou technickou normou je slovenská technická norma, ktorá upravuje oblasť, ktorá *nie je predmetom* úpravy európskej normy alebo medzinárodnej normy, a ktorá je *spracovaná na základe požiadavky verejnosti alebo orgánu verejnej moci*. Ich tvorba prebieha v rámci národných technických komisií.

Spolu s STN, Zákon o technickej normalizácii upravuje aj technickú normalizačnú informáciu. Technickou normalizačnou informáciou je normalizačný dokument prijatý slovenským národným normalizačným orgánom do sústavy slovenských technických noriem. Technická normalizačná informácia má značku „TNI".

Dodržiavanie STN alebo TNI je vo všeobecnosti *dobrovoľné.* Všeobecne záväzný právny predpis môže obsahovať odkaz na STN alebo TNI v texte, avšak použitie takejto technickej normy resp. normalizačnej informácie nie je pre adresátov právneho predpisu záväzné, nakoľko ide o odkaz pod čiarou.

Ak je to potrebné vzhľadom na technický charakter právneho predpisu alebo ak ide o právny predpis, ktorým sa do právneho poriadku Slovenskej republiky preberá právne záväzný akt Európskej únie, ktorý odkazuje na technické normy, možno podľa platných Legislatívnych pravidiel vlády SR v poznámke pod čiarou odkazovať na technické normy. V poznámke pod čiarou sa uvedie označenie, číslo a názov technickej normy a triediaci znak technickej normy v zátvorke v súlade s platnou sústavou slovenských technických noriem s použitím nedatovaných odkazov na technické normy.[83]

Zákon o technickej normalizácií umožňuje, aby bola STN *záväzná, hoci len nepriamo.* Orgán štátnej správy môže uviesť odkaz na slovenskú technickú normu alebo technickú normalizačnú informáciu v texte návrhu všeobecne záväzného právneho predpisu, teda nie v poznámke pod čiarou. V takom prípade sa požiadavky normy stanú záväznými práve prostredníctvom príslušného všeobecne záväzného právneho predpisu, avšak musia byť splnené nasledovné podmienky:

1. predkladateľ všeobecne záväzného právneho predpisu, písomne upozorní slovenský národný normalizačný orgán, a to najneskôr pred prijatím všeobecne záväzného právneho predpisu[84]
2. poskytnutie STN uhradí orgán štátnej správy, ktorý vyžaduje záväzné dodržiavanie STN v konkrétnom všeobecne záväznom predpise[85]
3. zodpovedný orgán vopred doručí Úradu pre normalizáciu STN a odôvodnenie potreby jej prijatia[86]
4. technická norma je alebo bude prevzatá do sústavy STN prekladom do slovenského jazyka.
5. STN musí vyhovovať prijatým a zverejneným kritériám t.j., že sú vyhotovené v štátnom jazyku a sú verejne prístupné na webovom sídle slovenského národného normalizačného orgánu.[87]

[83] čl. 23.10 Legislatívnych pravidiel vlády SR, dostupné na: https://www.vlada.gov.sk/share/RVLP/lpvsr_15072023.pdf?csrt=12580508444473908797.
[84] § 3 ods. 14 Zákona o technickej normalizácii.
[85] § 12 ods. 2 Zákona o technickej normalizácii.
[86] § 4 ods. 1 písm. c) bod 2 zákona č. 55/2018 Z. z.
[87] § 4c písm. c) Zákona o technickej normalizácii.

Záväzné použitie STN si môžu taktiež dohodnúť zmluvné strany v ich zmluvnom vzťahu (napr. v dohodách o úrovni služieb SLA).Pre porovnanie, európske normy ISO/IEC sú bežne priamo citované v nariadeniach a smerniciach EÚ.[88]

V neposlednom rade je potrebné si uvedomiť, že technická norma nie je to isté, čo technický predpis. Technickým predpisom sa rozumie všeobecne záväzný právny predpis alebo iný dokument vrátane návrhu technického predpisu, ktorý obsahuje

1. technickú požiadavku na výrobok alebo inú požiadavku, ktorej dodržiavanie je povinné pri sprístupňovaní výrobku na trhu alebo pri používaní výrobku, alebo ktorý zakazuje alebo obmedzuje výrobu, dovoz, predaj alebo používanie výrobku, alebo
2. pravidlo o službe informačnej spoločnosti, ktorého dodržiavanie je povinné v prípade predaja alebo poskytovania služby informačnej spoločnosti, zriadenia poskytovateľa služby informačnej spoločnosti alebo používania služby informačnej spoločnosti.[89]

2.3.2.2 Medzinárodné technické normy

Medzinárodné normy sú prijímané do sústavy slovenských technických noriem na základe podnetov odbornej verejnosti, resp. odborových združení, podnikateľských subjektov, či orgánov štátnej správy.

Medzinárodnými normalizačnými orgánmi sú ISO (*International Organization for Standardization*) Medzinárodná organizácia pre normalizáciu a IEC (*International Electrotechnical Commission*) Medzinárodná elektrotechnická komisia.

Pôvodné ISO/IEC normy sa prijímajú v anglickom, alternatívne nemeckom a francúzskom jazykuPre každú technickú komisiu ISO je za každý štát určený jeden zástupca. Slovenská republika ako člen ISO a IEC nemá povinnosť prijať všetky nimi vydané medzinárodné normy.

[88] MAKATURA, I.: *Technická normalizácia v informačnej a kybernetickej bezpečnosti*.
[89] § 2 písm. i) zákona č. 55/2018 Z. z. o poskytovaní informácií o technickom predpise a o prekážkach voľného pohybu tovaru.

2.3.2.3 Európske technické normy a harmonizované európske technické normy

A. Európske technické normy

Európske normy (EN) sa preberajú do sústavy slovenských technických noriem najneskôr do 6 mesiacov od ich sprístupnenia európskymi normalizačnými organizáciami.

Európske normalizačné organizácie sú tri a zodpovedajú za tvorbu dobrovoľných technických noriem na európskej úrovni. Prvou je CEN (franc. *Comité Européen de Normalisation*, angl. *European Commitee for Standardization*, Európsky výbor pre normalizáciu), druhou je CENELEC (franc. *Comité Européen de Normalisation Électrotechniqu*, angl. *European Commitee for Electrotechnical Standardization*, Európsky výbor pre normalizáciu v elektrotechnike) a poslednou je ETSI (*European Telecommunications Standards Institute*, Európsky inštitút pre telekomunikačné normy).

Rovnako ako v ISO/IEC, sa aj CEN/CENELEC opiera o štruktúru technických komisií.[90]

O spôsobe preberania európskych noriem, harmonizačných dokumentov, ako aj ich zmien a opráv rozhoduje Úrad pre normalizáciu v spolupráci so svojou príslušnou technickou komisiou, podľa účelu a rozsahu ich predpokladaného využívania.

Preberané dokumenty sa do sústavy STN preberajú jedným z týchto spôsobov:

- prekladom do štátneho jazyka,
- prevzatím originálu na priame používanie,
- oznámením vo vestníku Úradu pre normalizáciu.

Slovenská republika ako člen CEN a CENELEC, má povinnosť prijať všetky európske normy do sústavy STN a zrušiť pôvodné slovenské technické normy, ktoré sú s nimi v rozpore.

B. Harmonizované normy

[90] Zoznam technických komisií. [online] citované [19.1.2024]. Dostupné na: https://standards.cencenelec.eu/dyn/www/f?p=CEN:6

Európske normy, ktoré sa stávajú harmonizovanými vtedy, keď ich európske normalizačné organizácie oficiálne predložia Európskej Komisii a Európska Komisia ich vyhlási v Úradnom vestníku Európskej únie (*Official Journal of the European Union OJ EC*), typ L, najmä preto, aby so zreteľom na právne dôsledky bol stanovený dátum, od ktorého možno zhodu s európskou legislatívou predpokladať a aby všetky hospodárske subjekty v EÚ mali rovnakú východiskovú pozíciu pri využívaní harmonizovaných noriem. **Slovenská technická norma sa stane harmonizovanou, ak úplne preberá harmonizovanú európsku normu,** ktorá tvorí predpoklad zhody s technickými požiadavkami príslušnej európskej právnej úpravy.

2.3.3 Technické normy v oblasti umelej inteligencie

2.3.3.1 Technické normy pre riadenie umelej inteligencie (*AI governance*)

Potreba štandardov v oblasti riadenia AI je čoraz evidentnejšia vzhľadom na rastúci počet prípadov právnej regulácie AI na celom svete, najmä v Európskej únii, Spojenom kráľovstve a Spojených štátoch.

Terminológia regulácie o riadení AI sa v jednotlivých právnych predpisoch líši, čo vedie k nedostatočnému zosúladeniu kľúčových aspektoch, ako je taxonómia AI, mechanizmy riadenia, metodiky hodnotenia a meranie. Tento nedostatok harmonizácie a spoločného konsenzu možno odstrániť normalizáciou. Stanovením všeobecne akceptovaných noriem možno realizovať koherentnejší a konzistentnejší prístup k riadeniu technológií AI, zmierňovaniu rizík a podpore zodpovedného a etického vývoja a nasadenia AI.

2.3.3.2 Vývoj medzinárodných technických noriem umelej inteligencie

Organizácie ISO/IEC zriadili spoločnú technickú komisiu ISO/IEC JTC 1/SC 42 zodpovednú za technickú normalizáciu v oblasti umelej inteligencie.

Pracovný program komisie SC 42 sa zaoberá celým ekosystémom umelej inteligencie. Jeho súčasný pracovný program zahŕňa technickú normalizáciu v oblastiach základných noriem AI, dátových noriem týkajúcich sa AI, veľkých dát (*Big Data*) a analýzy,

dôveryhodnosti AI, prípadov použitia a aplikácií, vplyv AI na riadenie, výpočtových prístupov AI, testovania systémov AI, etických a spoločenských otázok.[91]

Doposiaľ sú publikované, respektíve sú v procese vývoja, najmä nasledovné kľúčové technické normy pre riadenie AI, ktoré možno rozdeliť do 4 skupín:

A. Základné normy (*Foundational*)

Tieto normy pomáhajú budovať spoločný jazyk, terminológiu a taxonómiu základných konceptov, čím uľahčujú dialóg medzi zainteresovanými stranami. V súčasnosti sú publikované dve základné normy pre umelú inteligenciu, pričom tieto slúžia ako dôležité stavebné kamene pre digitálnu transformáciu.[92]

- **ISO/IEC 22989:2022 Information technology. Artificial intelligence. Artificial intelligence concepts and terminology** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Poskytuje terminológiu a popisuje koncepty súvisiace so systémami AI, ktoré zahŕňajú široké spektrum technológií. Používanie odlišnej terminológie na opis rovnakých konceptov, procesov alebo technológií alebo používanie rovnakej terminológie napríklad na popis rôznych konceptov, procesov alebo technológií môže byť hlavnou prekážkou zodpovedného prijatia, širokého prijatia, spolupráce, regulácia a zdieľanie informácií.

Norma obsahuje viac ako 100 bežne používaných výrazov v AI, ako sú transparentnosť, vysvetliteľnosť, ovládateľnosť, skreslenie, súbor údajov, testovacie údaje, overovacie údaje a trénovaný model. Je navrhnutý tak, aby zlepšil efektívnosť výmeny informácií tým, že pomáha odborníkom na normalizáciu navrhovať normy, ktoré sú konzistentné z hľadiska terminológie.

[91] JCT1: *ISO/IEC JTC 1/SC 42 Artificial Intelligence.* Máj 2023 [online] citované [15.12.2023]. Dostupné na: https://jtc1info.org/sd-2-history/jtc1-subcommittees/sc-42/.
[92] IEC: *Two new foundational standards for artificial intelligence*. 19.7.2022. [online] citované [15.12.2023]. Dostupné na: https://www.iec.ch/blog/two-new-foundational-standards-artificial-intelligence.

- **ISO/IEC 23053:2022 Framework for Artificial Intelligence (AI). Systems Using Machine Learning (ML)** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Cieľom tohto dokumentu je poskytnúť rámec pre popis systémov AI, ktoré používajú ML. Tento dokument stanovuje rámec AI a ML na opis generického systému AI používajúceho ML technológie. Rámec popisuje systémové komponenty a ich funkcie v ekosystéme AI.

B. Procesné normy (*Process*)

Pomáhajú vytvárať organizačnú architektúru pre zodpovedný vývoj a nasadzovanie systémov umelej inteligencie tým, že zovšeobecňujú prijímanie najlepšej praxe („best practices") v oblasti manažérstva, navrhovania procesov, kontroly kvality a riadenia. Niektoré procesné normy sa považujú za „certifikovateľné", čo znamená, že organizácie môžu podstúpiť nezávislé posúdenie zhody, aby sa zistilo, či dodržiavajú predpísané osvedčené postupy, a následne získať certifikáciu podľa tejto špecifickej normy. Certifikačné normy končia v názvovej konvencii ISO/IEC vždy číslom 001.

- **ISO/IEC 42001:2023 Information technology. Artificial intelligence. Management system** *(publikovaná, prebieha preklad a proces prevzatia do sústavy STN)*

Dlho očakávaná norma ISO/IEC 42001:2023 je prvá medzinárodná norma systému manažérstva pre AI (***AI Management System, AIMS***).[93] Cieľom ISO/IEC 42001 je integrovať systém manažérstva AI s existujúcimi štruktúrami organizácie. Táto skutočnosť predurčuje normu ISO/IEC 42001 integrovať sa do organizácií tak ako už doposiaľ zavedené normy systému manažérstva, akými sú ISO 9001 v manažérstve kvality alebo ISO/IEC 27001 v systéme manažérstva informačnej bezpečnosti. Zámerom je, aby táto norma poskytla základ pre postupy posudzovania zhody pre procesy súvisiace s využitím AI.

Norma má štyri prílohy. Príloha A obsahuje opatrenia, ktoré sa majú zohľadniť pri ošetrovaní rizík. V prílohe B, ktorá sa zaoberá návodom na implementáciu opatrení z prílohy

[93] MCGARR, T.: *How a standard in development (ISO/IEC 42001) can meet collective AI Governance goals* [online] citované [15.12.2023]. Dostupné na: https://aistandardshub.org/how-a-standard-in-development-iso-iec-42001-can-meet-collective-ai-governance-goals/.

A. Príloha C je venovaná organizačným cieľom a zdrojom rizík súvisiacich s AI. Napokon, príloha D sa zaoberá používaním AIMS v rôznych oblastiach alebo sektoroch a ich integráciou s inými systémami manažérstva.

- **ISO/IEC DIS 42005 Information technology. Artificial intelligence. AI system impact assessment** *(vo vývoji)*

Tento dokument poskytuje usmernenia pre organizácie, ktoré vykonávajú posúdenie vplyvu systému umelej inteligencie na jednotlivcov a spoločnosti, ktorí môžu byť ovplyvnený systémom umelej inteligencie a jeho plánovanými a predvídateľnými aplikáciami. Dokument poskytuje návod, ako a kedy vykonávať takéto posúdenia a v ktorých fázach životného cyklu systému umelej inteligencie, ako aj usmernenia pre dokumentáciu posúdenia vplyvu systému umelej inteligencie.

- **ISO/IEC DIS 42006 Information technology. Artificial intelligence. Requirements for bodies providing audit and certification of artificial intelligence management systems** *(vo vývoji)*

Tento dokument špecifikuje dodatočné požiadavky k požiadavkám normy ISO/IEC 17021-1[94] s cieľom umožniť akreditovaným a vzájomne (*peer*) hodnoteným certifikačným orgánom spoľahlivo auditovať systém manažérstva pre organizácie, ktoré vyvíjajú alebo používajú systémy AI alebo oboje podľa ISO/IEC 42001 a vykonávať posúdenie a rozhodnutie o certifikácii. Použitie tohto dokumentu umožňuje certifikačným orgánom splniť špecifické technické vlastnosti a konkrétne riziká pri zaobchádzaní so systémami AI podľa ISO/IEC 42001. To umožňuje akreditačným orgánom a vzájomným (*peer*) posudzovateľom posúdiť kompetencie orgánov posudzovania zhody efektívnym a harmonizovaným spôsobom a zabezpečuje porovnateľnosť a reprodukovateľnosť certifikátov potvrdzujúcich zhodu s ISO/IEC 42001.

[94] ISO/IEC 17021-1:2015 Conformity assessment Requirements for bodies providing audit and certification of management systems

- **ISO/IEC 38507:2022 Information technology Governance of IT. Governance implications of the use of artificial intelligence by organizations** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument poskytuje návod pre členov vedenia organizácie, aby umožnili a riadili používanie umelej inteligencie (AI) s cieľom zabezpečiť jej účinné, efektívne a prijateľné používanie v rámci organizácie.

- **ISO/IEC 23894:2023 Information technology. Artificial intelligence. Guidance on risk management** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument poskytuje návod, ako môžu organizácie, ktoré vyvíjajú, vyrábajú, zavádzajú alebo používajú produkty, systémy a služby využívajúce umelú inteligenciu, riadiť riziká konkrétne súvisiace s AI. Cieľom usmernenia je tiež pomôcť organizáciám začleniť riadenie rizík do ich činností a funkcií súvisiacich s AI. Táto norma vychádza z predchádzajúcich noriem na riadenie rizík, ako je ISO 31000:2018, a opisuje procesy integrácie a implementácie riadenia rizík v oblasti AI.

- **ISO/IEC 8183:2023 Information technology. Artificial intelligence. Data life cycle framework** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument definuje etapy a identifikuje súvisiace činnosti pre spracovanie údajov počas celého životného cyklu systému umelej inteligencie (AI) vrátane získavania, vytvárania, vývoja, nasadenia, údržby a vyradenia z prevádzky. Tento dokument nedefinuje konkrétne služby, platformy alebo nástroje. Tento dokument sa vzťahuje na všetky organizácie bez ohľadu na typ, veľkosť alebo povahu, ktoré používajú údaje pri vývoji a používaní systémov umelej inteligencie.

- **ISO/IEC TR 27563:2023 Security and privacy in artificial intelligence use cases. Best practices** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument uvádza osvedčené postupy posudzovania bezpečnosti a súkromia v prípadoch použitia umelej inteligencie.

C. Normy merania (*Measurement*)

Tieto normy poskytujú univerzálne mechanizmy a terminológiu na meranie rôznych aspektov výkonu systému AI. Tieto sú obzvlášť dôležité, pretože vývoj a účinnosť dôveryhodných systémov AI závisí predovšetkým od obhájiteľných metód a mechanizmov merania.

- **ISO/IEC TS 4213:2022 Information technology Artificial intelligence Assessment of machine learning classification performance** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument špecifikuje metodiky na meranie klasifikačnej výkonnosti modelov strojového učenia, systémov a algoritmov.

- **ISO/IEC TR 24027:2021 Information technology Artificial intelligence (AI) Bias in AI systems and AI aided decision making** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument sa zaoberá skreslením (*bias*) vo vzťahu k AI systémom najmä pokiaľ ide o rozhodovanie s pomocou AI. Opisuje meracie techniky a metódy na hodnotenie skreslenia s cieľom riešiť a ošetriť zraniteľnosti súvisiace so skreslením. Zahŕňa všetky fázy životného cyklu systému AI vrátane, ale nie výlučne, zberu údajov, trénovania, neustáleho učenia sa, návrhu, testovania, hodnotenia a používania.

- **ISO/IEC TR 24029-1:2021 Artificial Intelligence (AI) Assessment of the robustness of neural networks Part 1: Overview** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument poskytuje základné informácie o existujúcich metódach hodnotenia robustnosti neurónových sietí.

- **ISO/IEC 25059:2023 Software engineering Systems and software Quality Requirements and Evaluation (SQuaRE). Quality model for AI systems** *(publikovaná, neprevzatá do sústavy STN, nepreložená do slovenského jazyka)*

Tento dokument opisuje model kvality pre systémy AI a je aplikačným rozšírením série noriem SQuaRE (prevzaté do sústavy STN).[95] Charakteristiky a sub-charakteristiky uvedené v tomto modeli poskytujú konzistentnú terminológiu na špecifikovanie, meranie a hodnotenie kvality systému AI.

D. Normy výkonnosti (*Performance*)

Tieto normy upravujú prahové hodnoty, požiadavky a očakávania, ktoré musia byť splnené na určitých úrovniach pre uspokojivú prevádzku a používanie systému AI:

- **ISO/IEC AWI 27090 Cybersecurity Artificial Intelligence. Guidance for addressing security threats and failures in artificial intelligence systems** *(vo vývoji)*

Tento dokument poskytuje organizáciám návod, ako riešiť bezpečnostné hrozby a zlyhania v systémoch umelej inteligencie (AI). Usmernenie v tomto dokumente má za cieľ poskytnúť organizáciám informácie, ktoré im pomôžu lepšie pochopiť dôsledky bezpečnostných hrozieb pre systémy AI počas ich životného cyklu, a popisy, ako takéto hrozby odhaliť a zmierniť. Tento dokument sa vzťahuje na všetky typy a veľkosti organizácií vrátane verejných a súkromných spoločností, vládnych subjektov a neziskových organizácií, ktoré vyvíjajú alebo používajú systémy AI.

- **IEEE 2937-2022 Standard for Performance Benchmarking for Artificial Intelligence Server Systems** (publikovaná)

Táto norma vyvinutá Inštitútom elektrotechnických a elektronických inžinierov (*Institute of Electrical and Electronics Engineers, IEEE*), stanovuje metodiky na hodnotenie výkonu serverov AI, serverových klastrov a iných AI High-Performance Computing (HPC)

[95] STN ISO/IEC 25000 Softvérové inžinierstvo. Požiadavky na kvalitu a hodnotenie kvality softvérového produktu (SQuaRE).

systémov. Okrem poskytovania usmernení o testovaní výkonu, metrikách a meraní tiež predpisuje technické požiadavky na nástroje benchmarkingu.

2.3.3.3 Vývoj európskych technických noriem umelej inteligencie

Cieľom Nariadenia o umelej inteligencii je regulačnými opatreniami stanoviť horizontálny rámec s cieľom vyhnúť sa fragmentácii jednotného digitálneho trhu a zabezpečiť harmonizáciu ustanovení o umelej inteligencii v rôznych sektoroch. Po nadobudnutí účinnosti Nariadenia o umelej inteligencii a po prechodnom období budú musieť vysokorizikové systémy umelej inteligencie pred uvedením na trh alebo do prevádzky v Európskej únii spĺňať súbor požiadaviek na dôveryhodnosť umelej inteligencie. Harmonizované a európske normy vyvinuté európskymi normalizačnými organizáciami (CEN/CENELEC), hoci sú dobrovoľnými nástrojmi, budú zohrávať kľúčovú úlohu pri definovaní technických riešení na splnenie týchto požiadaviek. Okrem toho by súlad s harmonizovanými normami mal poskytovať poskytovateľom umelej inteligencie právnu domnienku zhody ich produktov a služieb s požiadavkami Nariadenia o umelej inteligencii, najmä pokiaľ ide o vysokorizikové systémy AI.

2.3.3.4 Európsky plán technickej normalizácie AI a zriadenie technického výboru CEN-CENELEC JTC21 pre AI

Harmonizované normy sú jedným z hlavných prostriedkov na dosiahnutie súladu a súladu s legislatívnymi požiadavkami. Európske normy vypracuje CEN/CENELEC na základe vykonávacieho rozhodnutia Komisie C(2023)3215 zo dňa 22. mája 2023[96], ktorá požaduje vypracovanie návrhu nových európskych noriem alebo produktov európskej normalizácie (spoločne „***európske normy***"), ako sú uvedené v prílohe I tohto vykonávacieho rozhodnutia, a to do 30. apríla 2025. Budúce harmonizované európske normy budú pokrývať týchto 10 oblastí:

1. Systémy riadenia rizík pre systémy AI;

[96] Vykonávacie rozhodnutie Komisie zo dňa 22.5.2023 o žiadosti o normalizáciu adresovanej Európskemu výboru pre normalizáciu a Európskemu výboru pre normalizáciu v elektrotechnike na podporu politiky Únie v oblasti umelej inteligencie. [online] citované [15.12.2023]. Dostupné na: https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=en.

2. Riadenie a kvalita súborov údajov používaných na vytváranie systémov AI;
3. Uchovávanie záznamov prostredníctvom možností protokolovania pomocou systémov AI;
4. Ustanovenia o transparentnosti a informáciách pre používateľov AI;
5. Ľudský dohľad nad systémami AI;
6. Špecifikácie presnosti pre systémy AI;
7. Špecifikácie odolnosti pre systémy AI;
8. Špecifikácie kybernetickej bezpečnosti pre systémy AI;
9. Systémy riadenia kvality pre poskytovateľov systémov AI vrátane procesov monitorovania po uvedení na trh;
10. Posudzovanie zhody pre systémy AI.

Je zrejmé, že tak ako v prípade mnohých iných právnych predpisov EÚ bude súlad s Nariadením o umelej inteligencii podporovaný harmonizovanými normami. Tieto normy poskytnú usmernenia pre implementáciu Nariadenia o umelej inteligencii a uľahčia súlad s požiadavkami pre vysokorizikové systémy.[97]

Všeobecné technické a organizačné normy ako ISO/IEC 27001 v oblasti informačnej a kybernetickej bezpečnosti môžu prispieť k zmierneniu niektorých rizík, ktorým čelí AI.

Musia sa však prijať špecifické normy, ktoré sa zaoberajú najmä charakteristikami dôveryhodnosti, kvalitou údajov, riadením AI a systémami manažérstva AI .

CEN a CENELEC zriadili spoločný technický výbor (*Joint Technical Committee, JTC*) CEN-CENELEC 21 „Umelá inteligencia" (ďalej len „***CEN-CLC/JTC 21***"), ktorý je zodpovedný za vývoj a prijatie noriem pre AI a súvisiacich podkladov, ako aj za poskytovanie usmernení iným technickým výborom, ktoré sa zaoberajú AI.

CEN-CLC/JTC 21 najmä identifikuje a prijíma medzinárodné normy, ktoré sú už k dispozícii alebo sa vyvíjajú od iných organizácií, ako je ISO/IEC JTC 1 a jej podvýborov, ako je SC 42 Umelá inteligencia. Okrem toho sa CEN-CLC/JTC 21 zameriava na vytváranie

[97] LEONE DE CASTRIS, A, THOMAS, C.: *What role do standards play in the EU AIA? Looking at the implications of the European Parliament's proposed amendments* (27 July 2023) [online] [15.12.2023]. Dostupné na: https://aistandardshub.org/eu-ai-act.

normalizačných produktov, ktoré sa zameriavajú na potreby európskeho trhu a spoločnosti, ako aj na podporu legislatívy, politík, zásad a hodnôt EÚ.

CEN-CLC/JTC 21 v súčasnosti vyvíja európske normy, ktoré by v budúcnosti mohli výrobcom poskytnúť predpoklad zhody s požiadavkami Nariadenia o umelej inteligencii. Tento výbor identifikoval nasledovné normy (vrátane ISO noriem), ktoré sa priamo vzťahujú na Nariadenie o umelej inteligencii a ktorých prijatie/prispôsobenie pripravuje na roky 2024-2025:[98]

Projekt	Názov	Stav
CEN/CLC ISO/IEC/TR 24027:2023 (WI=JT021017)	Informačné technológie – Umelá inteligencia (AI) – Skreslenie v systémoch AI a rozhodovaní s pomocou AI (ISO/IEC TR 24027:2021)	V schvaľovaní
CEN/CLC ISO/IEC/TR 24029-1:2023 (WI=JT021018)	Umelá inteligencia (AI) – Hodnotenie odolnosti neurónových sietí – Časť 1: Prehľad (ISO/IEC TR 24029-1:2021)	V schvaľovaní
FprCEN/CLC ISO/IEC/TS 12791 (WI=JT021013)	Informačné technológie – Umelá inteligencia – Opatrenia pre nežiaduce skreslenia v klasifikácii a regresnom strojovom učení (ISO/IEC DTS 12791:2023)	V schvaľovaní
prCEN/CLC/TR 17894 (WI=JT021001)	Posudzovanie zhody umelej inteligencie	V príprave
prCEN/CLC/TR XXX(WI=JT021009)	Riziká AI – kontrolný zoznam pre riadenie rizík AI	Predbežný
prCEN/CLC/TR XXX(WI=JT021010)	Informačné technológie – Umelá inteligencia – Zelená a udržateľná AI	V príprave
prCEN/CLC/TR XXXX(WI=JT021002)	Umelá inteligencia - Prehľad úloh a funkcionalít AI súvisiacich so spracovaním prirodzeného jazyka	V príprave
prEN ISO/IEC 12792 (WI=JT021022)	Informačné technológie - Umelá inteligencia – Taxonómia k transparentnosti systémov AI	V príprave

[98] Pracovný program výboru CEN/CLC/JTC 21, [online] [15.12.2023]. Dostupné na: https://standards.cencenelec.eu/dyn/www/f?p=205:22:0::::FSP_ORG_ID,FSP_LANG_ID:2916257,25&cs=1827B89DA69577BF3631EE2B6070F207D

prEN ISO/IEC 23894 (WI=JT021016)	Informačné technológie – Umelá inteligencia – Usmernenie k riadeniu rizík (ISO/IEC 23894:2023)	V dopyte
prEN ISO/IEC 24029-2 (WI=JT021015)	Umelá inteligencia (AI) - Hodnotenie robustnosti neurónových sietí - Časť 2: Metodika používania formálnych metód.	Predbežný
prEN ISO/IEC 42001 (WI=JT021011)	Informačné technológie - Umelá inteligencia - Systém manažérstva	Predbežný
prEN ISO/IEC 8183 (WI=JT021020)	Informačné technológie - Umelá inteligencia - Rámec životného cyklu údajov	V schvaľovaní
prEN XXX(WI=JT021012)	Presnosť systémov spracovania prirodzeného jazyka	Predbežný
prEN XXX(WI=JT021023)	Opatrenia týkajúce sa podmienok údajov a požiadaviek na skreslenie	Predbežný
prEN XXX(WI=JT021019)	Požiadavky na kompetencie pre odborníkov v oblasti etiky AI	Predbežný
prEN XXX(WI=JT021008)	Charakterizácia dôveryhodnosti AI	Predbežný
prEN XXX(WI=JT021021)	Logovanie AI systémov	Predbežný
prEN XXXXX(WI=JT021006)	Nástroje na ovplyvňovanie rozhodnutia a správania („nudges“) vylepšené AI	V príprave

Na základe dohôd o technickej spolupráci môže CEN/CENELEC prijať/prispôsobiť normy ISO/IEC budúcim európskym normám, aby sa zabezpečil súlad s Nariadením o umelej inteligencii.[99]

CEN-CLC/JTC 21 už prebral obe základné ISO normy do európskych noriem, a tieto publikoval ako EN ISO/IEC 22989:2023 Informačné technológie - Umelá inteligencia - Pojmy a terminológia umelej inteligencie (ISO/IEC 22989:2022) a EN ISO/IEC 23053:2023 Rámec pre systémy umelej inteligencie využívajúce strojové učenie (ML) (ISO/IEC 23053:2022).

[99] ENISA: *Cybersecurity of AI and Standardisation* (marec 2023) [online] [15.12.2023]. Dostupné na: <https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation/@@download/fullReport>

Pri vývoji noriem je však potrebné zvážiť aj ďalšie aspekty.

Po prvé, v normalizačných činnostiach sa musí odzrkadliť nielen právna úprava Nariadenia o umelej inteligencii, ale aj ďalšie súvisiace právne predpisy, ako je napríklad návrh Nariadenia o kybernetickej odolnosti (*Cyber Resilience Act, CRA*), ktorý stanoví požiadavky na kybernetickú bezpečnosť pre digitálne produkty uvádzané na trh EÚ vrátane produktov obsahujúcich AI systémy.[100]

Po druhé, veľkou výzvou pre normalizáciu je skutočnosť, že niektoré aspekty najmä v oblasti kybernetickej bezpečnosti AI sú stále predmetom výskumu a vývoja, a preto nemusia byť dostatočne vyspelé pre ich normalizáciu. Podľa CEN-CLC/JTC 21 sú ďalšími podstatnými oblasťami v kybernetickej bezpečnosti umelej inteligencie, ktoré by mali byť pokryté normami, predovšetkým:

- modelovanie hrozieb kybernetických útokov špecifických pre AI a ich riadenie,
- meranie odolnosti modelov AI proti takýmto útokom,
- zvyšovať odolnosť modelov AI proti kybernetickým útokom a rozvíjať bezpečnostné opatrenia špecifické pre AI
- taxonómia a terminológia na priesečníku umelej inteligencie a kybernetickej bezpečnosti,
- hodnotenie rizík umelej bezpečnosti vrátane prispôsobenia klasických bezpečnostných opatrení na pokrytie bezpečnostných požiadaviek pre AI aktíva
- nové bezpečnostné riziká dodávateľského reťazca pre systémy umelej inteligencie pripojené k zdroju datasetov a
- rýchle šírenie predtrénovaných modelov, knižníc AI a základných AI služieb

Väčšina z nich je pokrytá aj klasickými opatreniami v kybernetickej bezpečnosti a normami, ako je séria noriem ISO 27000, ale potrebujú prispôsobenie, aby zvládli výzvy, ktoré prinášajú modely AI.[101]

[100] V decembri 2023 bola dosiahnutá predbežná dohoda na finálnom texte nariadenia CRA, pričom sa očakáva formálne prijatie Parlamentom a Radou. Dostupné online: https://www.europarl.europa.eu/news/en/press-room/20231106IPR09007/cyber-resilience-act-agreement-with-council-to-boost-digital-products-security

[101] SOLER GARRIDO, J., et al.: *Analysis of the preliminary AI standardisation work plan in support of the AI act*. EUR 31518 EN, Publications Office of the European Union, Luxembourg, 2023. ISBN 978-92-68-03924-3, doi:10.2760/5847, JRC132833. [online] [15.12.2023]. Dostupné na: https://publications.jrc.ec.europa.eu/repository/handle/JRC132833

2.3.3.5 Vývoj mimoeurópskych technických noriem pre umelú inteligenciu

USA

V Spojených štátoch amerických sa ďalší vývoj noriem odvíja predovšetkým od požiadaviek stanovených Výkonným príkazom (*Executive order*) prezidenta Joe Bidena o bezpečnej, chránenej a dôveryhodnej umelej inteligencii vydaný 30.10.2023 („***Príkaz prezidenta***").[102] Príkaz prezidenta poveruje viaceré agentúry vrátane Národného inštitútu pre štandardy a technológie (*National Institute of Standards and Technology*, *NIST*) vypracovaním usmernení a prijatím ďalších opatrení na podporu bezpečného a dôveryhodného vývoja a používania umelej inteligencie.

Vo vzťahu k normalizácii, Príkaz prezidenta priamo ukladá vyvíjať normy, nástroje a testy, ktoré pomôžu zabezpečiť, aby boli systémy umelej inteligencie bezpečné, spoľahlivé a dôveryhodné. NIST má za úlohu vypracovať normy pre rozsiahle pokročilé penetračné testovanie (tzv. *red-teaming*), aby sa zabezpečila bezpečnosť pred zverejnením. Ministerstvo vnútornej bezpečnosti (*Department of Homeland Security*) bude tieto normy uplatňovať na sektory kritickej infraštruktúry a zriadi Radu pre bezpečnosť a ochranu umelej inteligencie (*AI Safety and Security Board)*. Ministerstvá energetiky a vnútornej bezpečnosti sa budú zaoberať aj hrozbami systémov AI pre kritickú infraštruktúru, ako aj chemickými, biologickými, rádiologickými, jadrovými a kybernetickými bezpečnostnými rizikami.

V úlohe federálneho koordinátora pre štandardy v oblasti umelej inteligencie NIST spolupracuje s celou vládou a so zainteresovanými stranami z odvetvia s cieľom identifikovať kritické činnosti, stratégie a nedostatky v oblasti vývoja štandardov.

Dňa 26.01.2023, NIST vydal dokument s názvom Rámec riadenia rizík umelej inteligencie (*AI Risk Management Framework, AI RMF*).[103] Uvedený rámec je určený na

[102] Príkaz prezidenta (*Executive Order*) 14110 zo dňa 30.10.2023. *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.* A Presidential Document by the Executive Office of the President, publikované dňa 01.11.2023 vo Federálnom registri. [online] [15.12.2023]. Dostupné na: https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.

[103] NIST: *AI Risk Management Framework* (AI RMF 1.0). [online] [15.12.2023]. Dostupné na: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

dobrovoľné použitie a na zlepšenie schopnosti začleniť princípy dôveryhodnosti AI do návrhu, vývoja, používania a posudzovania produktov, služieb a systémov umelej inteligencie.

Veľká Británia

Hlavnú úlohu v oblasti normalizácie vykonáva Britský normalizačný inštitút (*British Standards Institution*), ktorý je národným normalizačným orgánom Spojeného kráľovstva.

BSI zriadil výbor ART/1 pre normalizáciu v oblasti umelej inteligencie, ktorý zrkadlí prácu skupiny ISO/IEC JTC 1/SC 42 a vytvára nezávislé národné (britské) normy. V súčasnosti bolo 18 uverejnených noriem a 81 je rozpracovaných.[104]

Čína

Čína zatiaľ neprijala právny predpis, ktorého predmetom by bola komplexná regulácia umelej inteligencie. Platné pravidlá, ktorými sa riadia oblasti súvisiace s umelou inteligenciou, sú rozložené vo viacerých zákonoch, ako je zákon o ochrane osobných údajov z roku 2021 (*Private Information Protection Law, PIPL*), zákon o bezpečnosti údajov z roku 2021 (*Data Security Law, DSL*) a zákon o kybernetickej bezpečnosti z roku 2016 (*Cyber Security Law, CSL)*, nariadeniach, zásadách a normách, z rôznych legislatívnych orgánov na rôznych úrovniach vlády.[105]

Doposiaľ boli vydané 3 prekrývajúce sa administratívne nariadenia na implementáciu zákonov a reguláciu AI:

- Administratívne ustanovenia o algoritmickom odporúčaní v internetových informačných službách 2021
- Administratívne ustanovenia o hĺbkovej syntéze v internetových informačných službách 2022
- Predbežné opatrenia na manažment služieb generatívnej umelej inteligencie 2023.[106]

[104] BSI: *ART/1 - Artificial Intelligence*. [online] [15.12.2023]. Dostupné na: https://standardsdevelopment.bsigroup.com/committees/50281655#published
[105] YANG, S. FUNG, C. ZHOU, B.: *China Proposes National Standards on Generative AI Security*, (10. November 2023) [online] [15.12.2023]. Dostupné na: https://www.chinalawvision.com/2023/11/tmt/china-proposes-national-standards-on-generative-ai-security/#_ftn1
[106] Tamže.

Normy v oblasti informačnej bezpečnosti vrátane umelej inteligencie vytvára predovšetkým Národný technický výbor pre štandardizáciu informačnej bezpečnosti (**全国信息安全**标准化技术委员会, TC260), ktorý bol založený so súhlasom Úradu pre normalizáciu Čínskej ľudovej republiky (*Standardization Administration of China, SAC*, **国家**标准化管理委员会).

V roku 2023 výbor TC260 zverejnil Bielu knihu o štandardizácii bezpečnosti umelej inteligencie.[107] Biela kniha skúma súčasnú situáciu vývoja technológií a aplikácií umelej inteligencie (AI), analyzuje nové bezpečnostné riziká vyplývajúce z AI a skúma súčasnú situáciu bezpečnostných politík a štandardov AI. V prílohe uvádza zoznam národných noriem prijatých výborom TC260 a tiež zahraničných noriem (najmä ISO/IEC) týkajúcich sa AI.

V októbri 2023 výbor TC260 vydal dokument Základné požiadavky na bezpečnosť služieb generatívnej umelej inteligencie (ide o návrh na pripomienkovanie).[108] Ide o prvú národnú normu Číny, ktorá špecificky upravuje osobitné bezpečnostné požiadavky na generatívnu umelú inteligenciu.[109]

2.3.4 Od technickej normalizácie k certifikácii IKT produktov a služieb

Certifikácia je najformálnejšou, ale zároveň aj najdôveryhodnejšou v rade metód posudzovania zhody.

Certifikácia produktov, procesov a služieb (súhrnne len „produktov") je postup, ktorým akreditovaný orgán posudzovania zhody poskytuje písomné ubezpečenie, že objekt posudzovania je v zhode so špecifickými požiadavkami.

STN EN ISO/IEC 17067:2013 definuje certifikáciu ako „poskytovanie posúdenia a nestranného osvedčenia treťou stranou, že bolo preukázané splnenie špecifikovaných požiadaviek" a „ustanovenú činnosť posudzovania zhody," pričom zároveň uvádza, že

[107] Biela kniha o štandardizácii bezpečnosti umelej inteligencie (**人工智能安全**标准化白皮书 (2023版)) zo dňa 31.05.2023 [online] [15.12.2023]. Dostupné (v čínštine) na: https://www.tc260.org.cn/upload/2023-05-31/1685501487351066337.pdf
[108] TC260: Dokument *Základné požiadavky na bezpečnosť služieb generatívnej umelej inteligencie* zo dňa 11.10.2023 [online] [15.12.2023]. Dostupné (v čínštine) na: https://www.tc260.org.cn/front/postDetail.html?id=20231011143225
[109] YANG, S. FUNG, C. ZHOU, B.: *China Proposes National Standards on Generative AI Security*, p 25.

„špecifikované požiadavky na výrobky sú vo všeobecnosti obsiahnuté v normách alebo iných normatívnych dokumentoch."[110]

Certifikácia produktu by mala poskytnúť (i) dôveru pre tých, ktorí majú záujem o splnenie požiadaviek, a (ii) dostatočnú hodnotu, aby dodávatelia mohli efektívne predávať produkty. [111]

Na vykonanie certifikácie je potrebné vyvinúť vhodné certifikačné schémy. Certifikačná schéma je certifikačný systém týkajúci sa špecifikovaných produktov, na ktorý sa vzťahujú rovnaké špecifikované požiadavky, špecifické pravidlá a postupy.

Certifikácia IKT produktov je osobitne dôležitá vo vzťahu k požiadavkám na kybernetickú bezpečnosť takýchto produktov. Nariadenie o kybernetickej bezpečnosti (*Cybersecurity Act, CSA*)[112] kladie dôraz na technické normy ako základ pre certifikáciu. Technická normalizácia v kybernetickej bezpečnosti je jednoznačne pridanou hodnotou pre certifikáciu. Je potrebné vytvoriť úzky vzťah a praktickú koordináciu s organizáciami vytvárajúcimi normy (napr. ISO, CEN/CENELEC apod.), ktorá umožní funkčné a efektívne využívanie noriem na podporu certifikácie. Technická normalizácia a súčinnosť organizácií vytvárajúcich normy by mala byť súčasťou pri vytváraní nových certifikačných schém.[113]

Certifikát kybernetickej bezpečnosti by mal byť podľa Nariadenia o kybernetickej bezpečnosti vydaný po úspešnom vyhodnotení špecifikovaných požiadaviek týkajúcich sa IKT produktu, IKT služby alebo IKT procesu, ktoré vykoná nezávislá tretia strana.[114] Európsky rámec certifikácie kybernetickej bezpečnosti by mal zahŕňať schémy certifikácie kybernetickej bezpečnosti, ktoré dokazujú, že certifikované riešenia IKT majú požadovanú úroveň ochrany kybernetickej bezpečnosti.[115] V januári 2024 bola prijatá ako prvá, schéma certifikácie produktov IKT (*EU cybersecurity certification scheme on Common Criteria*,

[110] STN EN ISO/IEC 17067:2014 Posudzovanie zhody. Základy certifikácie výrobkov a zásady pre systémy certifikácie výrobkov (ISO/ IEC 17067:2013), článok 4.
[111] STN EN ISO/IEC 17067:2014, článok 4.2.2.
[112] Nariadenie Európskeho parlamentu a Rady (EÚ) 2019/881 zo 17.04.2019 o agentúre ENISA (Agentúra Európskej únie pre kybernetickú bezpečnosť) a o certifikácii kybernetickej bezpečnosti informačných a komunikačných technológií a o zrušení nariadenia (EÚ) č. 526/2013 (akt o kybernetickej bezpečnosti)
[113] ENISA: *Standardisation in support of the Cybersecurity Certification*. [online] [15.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/recommendations-for-european-standardisation-in-relation-to-csa-i
[114] Nariadenie o kybernetickej bezpečnosti, recitál 77.
[115] ENISA: *Cybersecurity Certification Framework.* [online] [15.12.2023]. Dostupné na: <https://www.enisa.europa.eu/topics/certification/cybersecurity-certification-framework>

EUCC).[116] EUCC má dopĺňať právne predpisy v oblasti kybernetickej bezpečnosti, ako je najmä návrh Nariadenia o kybernetickej odolnosti. Pripravujú sa aj ďalšie dve schémy, certifikačná schéma EUCS sa týka cloudových služieb a certifikačná schéma s názvom EU5G sa týka sietí 5G.

Nariadenie o umelej inteligencii upravuje prezumpciu zhody s požiadavkami kybernetickej bezpečnosti pre vysokorizikové systémy AI, ak boli certifikované alebo pre ktoré bolo vydané vyhlásenie o zhode v rámci schémy kybernetickej bezpečnosti podľa Nariadenia o kybernetickej bezpečnosti.[117] I keď zatiaľ nebola vydaná žiadna oficiálna žiadosť o európsku certifikačnú schému kybernetickej bezpečnosti pre AI, je dôležité, aby takáto schéma, ak by bola vytvorená, náležite zohľadňovala Nariadenie o umelej inteligencii a rovnako aby Nariadenie o umelej inteligencii predpokladalo prijatie takejto certifikačnej schémy. Zostáva definovať, či a ako sa tri úrovne dôveryhodnosti IKT produktov, procesov a služieb (základná, pokročilá, vysoká) podľa Nariadenia o kybernetickej bezpečnosti môžu uplatňovať v kontexte Nariadenia o umelej inteligencii.[118] Vzťah kybernetickej bezpečnosti a umelej inteligencie rozoberáme bližšie v kapitole 7 tejto učebnice.

[116] ENISA: An EU Prime! EU adopts first Cybersecurity Certification Scheme. [online] [5.2.2024]. Dostupné na: https://www.enisa.europa.eu/news/an-eu-prime-eu-adopts-first-cybersecurity-certification-scheme
[117] Nariadenie o umelej inteligencii, článok 42 (2).
[118] ENISA: *Cybersecurity of AI and Standardisation (Report)*. Marec 2023. s.6. [online] [15.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation/@@download/fullReport

3. Regulácia umelej inteligencie v Európe

Európska únia je regulačná „veľmoc," čo podčiarkuje prijímanie európskych noriem ako zlatého štandardu naprieč celým svetom.[119] Výnimkou nie je ani regulácia umelej inteligencie, kde EÚ predstavila prvý komplexný právny rámec na svete, ktorý v čase písania tejto učebnice čaká na formálne prijatie. Rada Európy nezaháľa a pripravuje vlastný právny rámec, ktorý by mal korelovať s úniovým predpisom. V rámci tejto kapitoly si rozoberieme súčasnú normotvorbu týchto dvoch organizácií.

3.1 Regulácia umelej inteligencie v Európskej únii

Európska komisia v apríli 2021 po niekoľko-mesačných snahách predstavila prvý komplexný návrh regulácie umelej inteligencie na svete – návrh nariadenia o umelej inteligencií (*Artificial Intelligence Act*, AIA). Nadviazala tak na dlhoročnú prácu expertných skupín a požiadaviek orgánov Európskej únie v podobe stanovísk či odporúčaní.[120] V decembri 2022 publikovala svoju pozíciu (všeobecné smerovanie) k AIA Rada EÚ.[121] Dlho očakávaná pozícia Európskeho parlamentu bola schválená v júny 2023[122] a v ten istý mesiac začali trialógy. V decembri 2023 dosiahli orgány EÚ politickú dohodu na konečnom znení nariadenia, ktoré je v súčasnosti predmetom legislatívno-technických prác. Finalizácia textu sa očakáva vo februári 2024 so schválením aktu v apríli 2024. Účinnosť AIA je koncipovaní v niekoľkých vlnách, čomu sa ešte budeme osobitne venovať.

3.1.1 Základná filozofia Aktu o umelej inteligencií

Dopadová štúdia k návrhu AIA uvádza 6 dôvodov pre prijatie regulácie umelej inteligencie (AI) na úrovní EÚ, konkrétne:

- Využívanie AI predstavuje zvýšené riziko pre bezpečnosť jednotlivcov,

[119] K tomu pozri napríklad Bradford, Anu, The Brussels Effect (2012). Northwestern University Law Review, Vol. 107, No. 1, 2012, Columbia Law and Economics Working Paper No. 533, Available at SSRN: https://ssrn.com/abstract=2770634.
[120] Napríklad EXPERTNÁ SKUPINA NA VYSOKEJ ÚROVNI PRE UMELÚ INTELIGENCIU. *Etické Usmernenia Pre Dôveryhodnú Umelú Inteligenciu*. Dostupné na: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.
[121] Pozícia je dostupná online na https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf.
[122] Pozícia je dostupná online na https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html.

- Využívanie AI predstavuje zvýšené riziko porušovania základných ľudských práv a slobôd,
- Dozorné orgány nemajú právomoci, procesné rámce a zdroje na zabezpečenie a monitorovanie súlad vývoja a používania umelej inteligencie s platnými pravidlami,
- Právna neistota ohľadom uplatnenia súčasného právneho rámca na systémy AI,
- Nedôvera k AI by znížila inovatívnosť a konkurecieschopnosť EÚ v globálnom meradle,
- Fragmentované opatrenia predstavujú prekážky pre ďalší vývoj jednotného digitálneho trhu a suverenity.[123]

AIA si za cieľ kladie harmonizovať pravidlá systémov AI pri uvádzaní a používaní v EÚ, zakázať určité systémy AI, upraviť špecifické požiadavky pre vysokorizikové systémy AI, zakotviť požiadavky transparentnosti pre špecifické systémy AI a harmonizovať pravidlá dohľadu a monitorovania trhu.[124]

AIA je formulovaná ako produktová regulácia a porovnať to možno s požiadavkami na produkty ako napríklad zdravotnícke pomôcky alebo elektronické výrobky pri uvedení na trh. To znamená, že určuje špecifické vlastnosti a požiadavky na systémy AI, ktoré musia byť splnené pri uvedení na trh a veľkú zodpovednosť ponecháva na samotných prevádzkovateľoch týchto systémov prostredníctvom inštitútu posúdenia zhody (*conformity assessment*).

AIA je horizontálna regulácia založená na skúmaní rizika. To prakticky znamená, že vymedzuje systémy AI v kontexte rôznych typov rizík pre základné ľudské práva a slobody[125] a následne upravuje špecifické právne požiadavky pre tieto systémy v danej kategórií. AIA rozlišuje systémy AI, ktoré predstavujú:

- Neakceptovateľné riziko (zakázané praktiky),

[123] EURÓPSKA KOMISIA. *Commission staff working document impact assessment accompanying the proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*. {COM(2021) 206 final} - {SEC(2021) 167 final} - {SWD(2021) 85 final}.
[124] AIA, článok 1.
[125] Napríklad AIA, recitál 15: *„Využívanie umelej inteligencie má síce mnoho výhod, túto technológiu však možno zneužiť a môže sa stať zdrojom nových a výkonných nástrojov umožňujúcich manipulatívne a zneužívajúce praktiky a praktiky v oblasti sociálnej kontroly. Takéto praktiky sú mimoriadne škodlivé a mali by sa zakázať, pretože sú v rozpore s hodnotami Únie týkajúcimi sa* ***rešpektovania ľudskej dôstojnosti, slobody, rovnosti, demokracie a právneho štátu a základných práv Únie vrátane práva na nediskrimináciu, ochranu údajov a súkromia a práv dieťaťa.***"

- Vysoké riziko,
- Nízke riziko,
- Žiadne riziko.

Takmer vôbec sa regulácia netýka systémov AI nízkeho resp. minimálneho rizika, kde ustanovuje iba požiadavky na transparentnosť a odporúčanie prijatia kódexov správania, ktoré výrobcovia a prevádzkovatelia takýchto AI systémov budú dodržiavať. Osobitne AIA upravuje povinnosti pre systémy umelej inteligencie na všeobecné účely (pôvodne základné modely, z originálu *foundation models*).

Zaujímavosťou sú aj sankcie za porušenie AIA. Tie sú upravené ešte striktnejšie ako pri nariadení o ochrane údajov. Za porušenie ustanovení AIA bude možné uložiť pokutu až do výšky 35 000 000 EUR, alebo ak je porušiteľom spoločnosť, až do výšky 6,5 % jej celkového svetového ročného obratu za predchádzajúci účtovný rok, podľa toho, ktorá suma je vyššia za porušenie zakázaných praktík a nesúlad s požiadavkami na správu údajov. Ďalej AIA umožňuje správne pokuty do výšky 15 000 000 EUR a 7 500 000 EUR.[126]

AIA ustanovuje povinnosť pre členské štáty kreovať alebo určiť dozorný orgán, ktorý bude vykonávať štátny dozor. Na úrovní EÚ zároveň vznikne Európska rada pre umelú inteligenciu a Úrad pre umelú inteligenciu (*AI Office*). Zároveň bude musieť každý členský štát dezignovať národný dozorný orgán.

[126] AIA, článok 71.

3.1.2. Pôsobnosť právneho predpisu

Kľúčovou definíciou regulácie je pojem systémy AI. Ten je definovaný ako *„strojový systém, ktorý je navrhnutý tak, aby fungoval s rôznou úrovňou autonómie a ktorý môže pre explicitné alebo implicitné ciele vytvárať výstupy, ako sú predpovede, odporúčania alebo rozhodnutia, ktoré ovplyvňujú fyzické alebo virtuálne prostredie.*"[127] Predmetná definícia vychádza z dokumentov na úrovni Organizácie pre hospodársku spoluprácu a rozvoj (OECD).

Medzi príkladov, ktoré spĺňajú definíciu systémov AI môžeme zaradiť nasledovné systémy:

- Virtuálni asistenti: Siri, Alexa, Google Assistant, Bixby - rozpoznávajú reč, porozumejú jazyku a reagujú na požiadavky.
- Odporúčacie systémy: Netflix, Spotify, YouTube, Amazon - odporúčajú produkty a obsah na základe preferencií používateľov.
- Autonómne vozidlá: Tesla Autopilot, Waymo, Cruise Automation - dokážu riadiť vozidlá bez ľudského zásahu.
- Systémy rozpoznávania tvárí: Face ID, Windows Hello, Facebook DeepFace - rozpoznávajú tváre a overujú identitu.
- Chatboty: LaMDA, GPT-3, Mitsuku - vedú konverzáciu s ľuďmi pomocou prirodzeného jazyka.
- Medicínske diagnostické systémy: IBM Watson Oncology, Google DeepMind Health, Lunit - diagnostikujú choroby a odporúčajú liečbu.
- Systémy strojového prekladu: Google Translate, Microsoft Translator, DeepL - prekladajú text z jedného jazyka do druhého.
- Generovanie textu a obsahu: GPT-3, Bard, Jasper - generujú text v rôznych formátoch.
- Roboty v priemysle a logistike: Baxter, Sawyer, Kuka LBR iiwa - automatizujú úlohy v priemysle a logistike.
- Algoritmy pre online reklamu: Google Ads, Facebook Ads, Amazon Advertising - zobrazujú relevantné reklamy na základe záujmov používateľov.

AIA diferencuje medzi niekoľkými subjektami, ktoré hrajú významnú alebo menej významnú rolu pri vývoji, nasadzovaní a používaná systémov AI. Regulácie pracuje s pojmami

[127] AIA, článok 3 bod 1.

poskytovateľ (*provider*), subjekt, ktorý systém nasadzuje (*deployer*), dovozca (*importer*), distribútor (*distributor*) a prevádzkovateľ (*operator*).

Poskytovateľ systému AI „je fyzická alebo právnická osoba, orgán verejnej moci, agentúra alebo iný subjekt ktorá vyvíja systém AI alebo model AI na všeobecné účely alebo ktorá má systém AI alebo model AI na všeobecné účely a uvedie ich na trh alebo dá na trh systém do prevádzky pod svojím vlastným menom alebo ochrannou známkou, či už za odplatu alebo bezodplatne."[128] Inými slovami, je to osoba alebo organizácia, ktorá vyvíja a vlastní AI systém. Je zodpovedná za celkovú funkčnosť a údržbu systému. Subjekt, ktorý AI nasadzuje (*deployer*) „znamená akúkoľvek fyzickú alebo právnickú osobu, verejný orgán, agentúru alebo iný subjekt používajúci systém AI v rámci svojej právomoci okrem prípadov, keď sa systém AI používa v rámci osobnej neprofesionálnej činnosti." Ide o subjekt, ktorý inštaluje a konfiguruje AI systém v konkrétnom prostredí. Dovozca „je každá fyzická alebo právnická osoba, ktorá sa nachádza alebo je usadená v Únii a ktorá uvedie na trh systém umelého inteligencie, ktorý nesie názov alebo ochrannú známku fyzickej alebo právnickej osoby, ktorá je právnickej osoby usadenej mimo Únie." Distribútor „je akákoľvek fyzická alebo právnická osoba v dodávateľskom reťazci, okrem poskytovateľ alebo dovozcu, ktorá sprístupňuje systém AI na trhu Únie." Prevádzkovateľ je „poskytovateľ, výrobca výrobku, subjekt, ktorý zavádza systém, autorizovaný zástupca, dovozca alebo distribútor." Používateľ (*user*) je osoba, ktorá priamo interaguje s AI systémom a využíva jeho funkcie. Tento termín však finálna verzia AIA nedefinuje.

PRÍKLAD
Predstavme si AI systém pre rozpoznávanie tvárí, ktorý vyvinula spoločnosť X (poskytovateľ). Spoločnosť Y (*deployer*) nainštaluje systém v banke Z. Zamestnanci banky (používatelia) používajú systém na kontrolu prístupu do budovy. Spoločnosť W (dovozca) doviezla systém do krajiny a spoločnosť V (distribútor) ho predáva bankám. Banka Z (prevádzkovateľ) zodpovedá za každodennú prevádzku systému.

AIA by sa mala vzťahovať na poskytovateľov systémov AI bez ohľadu na to, či sú v EÚ usadení alebo nie, postačí, ak je splnené kritérium, že budú systémy AI uvádzať na trhu alebo

[128] Článok 3, bod 2.

prevádzkovať v rámci EÚ.[129] Extra-teritoriálna pôsobnosť AIA je zvýraznená aj tým, že sa bude vzťahovať na poskytovateľov používateľov systémov AI z tretích krajín, ak výstupy tvorené ich systémami sa využívajú v EÚ.[130] Špecifické požiadavky sú smerované aj na subjekty, ktoré AI nasadzujú.

Z hľadiska negatívnej pôsobnosti sa AIA nevzťahuje na systémy umelej inteligencie vyvinuté alebo používané výlučne na vojenské účely a na produkty v rámci právnych aktov výslovne vymenovaných v článku 2 ods. 2 AIA. Z pôsobnosti sú vyňate systémy AI vyvíjané na výskumné účely a pre výlučné osobnú (*non-professional*) potrebu. Do pôsobnosti AIA taktiež nepatria open-source systémy AI, ktoré nepredstavujú vysoké riziko z pôsobnosti AIA.

3.1.3 Zakázané praktiky

Na vrchole pyramídovej regulácie sú systémy AI, ktoré predstavujú neprijateľné riziko a zákonodarca sa ich rozhodol zakázať. Ide o systémy AI, ktoré predstavujú neakceptovateľné riziko pre bezpečnosť, zdravie a základné ľudské práva a slobody. Článok 5 AIA, ktorý upravuje tieto zákazy do istej miery reflektuje diskusiu o tzv. červených čiarach pri systémoch AI.[131] Ide o 8 zakázaných praktík. AIA výslovne zakazuje:

[129] AIA, článok 2.
[130] AIA, článok 2 ods. 1 písm. c).
[131] K tomu pozri napríklad EDRI. *Civil society calls for AI red lines in the European Union's Artificial Intelligence proposal*. Dostupné na: https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-

1) uvádzanie na trh, uvádzanie do prevádzky alebo používanie systému umelej inteligencie, ktorý **využíva podprahové techniky mimo vedomia osoby** alebo zámerne manipulatívne klamlivé techniky tak, s cieľom alebo účinkom podstatného skreslenia správanie osoby alebo skupiny osôb tým, že výrazne naruší ich schopnosť prijímať informované rozhodnutia, čím spôsobí, že osoba prijme rozhodnutie, ktoré by inak neprijala, a to spôsobom, ktorý spôsobí alebo môže spôsobiť tejto osobe, inej osobe alebo skupine osôb značnú škodu. Ako príklad takéhoto systému možno uviesť využívanie zvukových podprahových techník, ktoré udržia zamestnancov vo vyššej výkonnosti s potenciálom škody na mentálnom alebo fyzickom zdraví;

2) *„uvádzanie na trh, uvádzanie do prevádzky alebo používanie systému umelej inteligencie, **ktorý využíva akúkoľvek zraniteľnosť osoby alebo určitej skupiny osôb** z dôvodu ich veku, zdravotného postihnutia alebo osobitnej sociálnej alebo ekonomickej situácie s cieľom alebo účinkom podstatne narušiť správanie tejto osoby alebo osoby patriacej do tejto skupiny spôsobom, ktorý spôsobuje alebo môže spôsobiť tejto osobe alebo inej osobe značnú škodu.*"[132] Pod tento zákaz možno zaradiť napríklad hračky pre deti, ktoré maloletých budú nútiť do nebezpečných aktivít ako bábika hovoriaca maloletým užívateľom o pozitívnom účinku tvrdej drogy;

3) uvedenie na trh alebo do prevádzky na tento osobitný účel, alebo používanie biometrických kategorizačných systémov, ktoré kategorizujú jednotlivé fyzické osoby na základe ich biometrických údajov s cieľom odvodiť alebo vyvodiť ich rasu, politické názory, členstvo v odboroch, náboženské alebo filozofické presvedčenie, sexuálny život alebo sexuálnu orientáciu. Výnimkou sú legálne získané údaje a ich kategorizácia v oblasti orgánov presadzovania práva.

4) „uvádzanie na trh, uvádzanie do prevádzky alebo používanie systémov umelej inteligencie **na účely hodnotenia alebo klasifikácie dôveryhodnosti fyzických osôb** počas určitého obdobia na základe ich spoločenského správania alebo

proposal/. UNITED NATIONS – HUMAN RIGHTS, OFFICE OF THE HIGH COMMISSIONER. *New and emerging technologies need urgent oversight and robust transparency: UN experts*. Dostupné na: https://www.ohchr.org/en/press-releases/2023/06/new-and-emerging-technologies-need-urgent-oversight-and-robust-transparency.

[132] AIA, článok 5 ods. 1 písm. b).

známych, odvodených či predpokladaných osobných alebo osobnostných charakteristík, pričom takto získané sociálne skóre vedie k jednému alebo obidvom z týchto výsledkov: (i) **škodlivé alebo nepriaznivé zaobchádzanie** s určitými fyzickými osobami alebo celými skupinami fyzických osôb **v sociálnych kontextoch**, ktoré **nesúvisia** s kontextmi, v ktorých boli údaje pôvodne generované alebo zhromaždené a/alebo (ii) škodlivé alebo nepriaznivé zaobchádzanie s určitými fyzickými osobami alebo celými skupinami fyzických osôb, ktoré je **neodôvodnené alebo neprimerané** ich spoločenskému správaniu alebo jeho závažnosti."[133] Predmetný zákaz reflektuje situáciu, ak by orgány verejnej moci zaviedli systém hodnotenia občanov, ktorý by im podľa získaných bodov (ne)umožňoval využívať hromadnú dopravu, diaľkové spoje alebo sociálne služby. Podobný systém funguje v niektorých ázijských krajinách.[134]

5) používanie systémov diaľkovej biometrickej identifikácie v reálnom čase vo verejne prístupných priestoroch na účely presadzovania práva"[135] s výnimkou špecifických prípadov a podmienok, keď orgány presadzovania práva (napríklad orgány činné v trestnom konaní) môžu systémy diaľkovej biometrickej identifikácie (kamery s funkciou rozpoznávania tváre) použiť. Mohlo by sa tak stať napríklad pri hľadaní detí alebo pri predchádzaní teroristických útokov. AIA ale zároveň vyžaduje splnenie viacerých požiadaviek, ak by sa členské štáty rozhodli takýto systém nasadiť. Predovšetkým ide o požiadavku nezávislého dohľadu resp. povolenia súdneho alebo iného orgánu s využitím takýchto metód.[136] Zároveň musí byť zachovaná proporcionalita nasadenia takéhoto systému AI.[137]

6) uvedenie na trh, uvedenie do prevádzky na tento osobitný účel alebo používanie systému umelej inteligencie na posudzovanie rizika fyzických osôb s cieľom posúdiť alebo predvídať riziko, že fyzická osoba spácha trestný čin, a to výlučne na základe profilovania fyzickej osoby alebo posúdenia jej osobnostných vlastností

[133] AIA, článok 5 ods. 1 písm. c).
[134] K tomu pozri KASL, F. Surveillance in digitalized society: the chinese social credit system from a european perspective. In *The Lawyer Quarterly*, Vol 9, No 4 (2019).
[135] AIA, článok 5 ods. 1 písm. d) v spojení s odsekmi 2-4.
[136] AIA, článok 5 ods. 3.
[137] AIA, článok 5 ods. 2 – 5.

a charakteristík (tzv. *predictive policing*). Výnimkou je využívanie takýchto systémov v situáciách, kde slúžia iba na podporu ľudského hodnotenia páchateľa na základe objektívnych faktov relevantných pre trestnú činnosť.

7) používanie systémov umelej inteligencie, ktoré vytvárajú alebo rozširujú databázy rozpoznávania tváre prostredníctvom necieleného získavania obrazov tváre z internetu alebo záznamov priemyselných kamier.

8) uvedenie na trh, uvedenie do prevádzky na tento osobitný účel alebo používanie systémov umelej inteligencie na odvodzovanie emócií fyzickej osoby v oblasti pracovísk a vzdelávacích inštitúcií s výnimkou prípadov, keď je použitie systému umelej inteligencie určené na zavedenie alebo uvedenie na trh z lekárskych alebo bezpečnostných dôvodov.

3.1.2 Vysokorizikové systémy AI a požiadavky na nich

AIA sa automaticky nevzťahuje na všetky systémy AI v zmysle definície uvedenej vyššie. Požiadavky kladené týmto nariadením sa budú aplikovať na tzv. vysokorizikové systémy AI. Či je systém AI vysokorizikový, alebo nie, upravuje článok 6. Klasifikovať systém AI ako vysokorizikový možno na základe dvoch alternatívnych zdrojov. Prvým zdrojom je odkaz na špecifickú reguláciu prostredníctvom prílohy II AIA. Ide o prípady, ak systém AI *„je určený na používanie ako bezpečnostný komponent výrobku, na ktorý sa vzťahujú harmonizačné právne predpisy Únie uvedené v prílohe II, alebo je sám osebe takýmto výrobkom."*[138] Alternatívne totožné konštatovanie platí, ak *„výrobok, ktorého bezpečnostným komponentom je systém umelej inteligencie, alebo samotný systém umelej inteligencie ako výrobok sa musí podrobiť posúdeniu zhody treťou stranou s cieľom uviesť daný výrobok na trh alebo do prevádzky podľa harmonizačných právnych predpisov Únie uvedených v prílohe II ."*[139] Príloha II pre tieto výrobky uvádza nasledujúce regulácie:

- nariadenie o strojníckych výrobkoch[140]

[138] AIA, článok 6 ods. 1 písm. a).
[139] AIA, článok 6 ods. 1 písm. b).
[140] Nariadenie Európskeho parlamentu a Rady (EÚ) 2023/1230 zo 14. júna 2023 o strojových zariadeniach a o zrušení smernice Európskeho parlamentu a Rady 2006/42/ES a smernice Rady 73/361/EHS.

- smernica o bezpečnosti hračiek[141]
- smernica o rekreačných plavidlách a skútroch[142]
- smernica o harmonizácii právnych predpisov členských štátov týkajúcich sa výťahov a bezpečnostných komponentov do výťahov[143]
- smernica o harmonizácii právnych predpisov členských štátov týkajúcich sa zariadení a ochranných systémov určených na použitie v potenciálne výbušnej atmosfére[144]
- smernica o harmonizácii právnych predpisov členských štátov týkajúcich sa sprístupňovania rádiových zariadení na trhu[145]
- smernica o harmonizácii právnych predpisov členských štátov týkajúcich sa sprístupňovania tlakových zariadení na trhu[146]
- nariadenie o lanovkových zariadeniach a zrušení smernice 2000/9/ES[147]
- nariadenie o osobných ochranných prostriedkoch[148]
- nariadenie o spotrebičoch spaľujúcich plynné palivá[149]
- nariadenie o zdravotníckych pomôckach[150]
- nariadenie o diagnostických zdravotníckych pomôckach in vitro[151]
- nariadenie o spoločných pravidlách v oblasti bezpečnostnej ochrany civilného letectva[152]

[141] Smernica Európskeho parlamentu a Rady 2009/48/ES z 18. júna 2009 o bezpečnosti hračiek.
[142] Smernica Európskeho parlamentu a Rady 2013/53/EÚ z 20. novembra 2013 o rekreačných plavidlách a vodných skútroch a o zrušení smernice 94/25/ES.
[143] Smernica Európskeho parlamentu a Rady 2014/33/EÚ z 26. februára 2014 o harmonizácii právnych predpisov členských štátov týkajúcich sa výťahov a bezpečnostných komponentov do výťahov.
[144] Smernica Európskeho parlamentu a Rady 2014/34/EÚ z 26. februára 2014 o harmonizácii právnych predpisov členských štátov týkajúcich sa zariadení a ochranných systémov určených na použitie v potenciálne výbušnej atmosfére.
[145] Smernica Európskeho parlamentu a Rady 2014/53/EÚ zo 16. apríla 2014 o harmonizácii právnych predpisov členských štátov týkajúcich sa sprístupňovania rádiových zariadení na trhu, ktorou sa zrušuje smernica 1999/5/ES.
[146] Smernica Európskeho parlamentu a Rady 2014/68/EÚ z 15. mája 2014 o harmonizácii právnych predpisov členských štátov týkajúcich sa sprístupňovania tlakových zariadení na trhu.
[147] Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/424 z 9. marca 2016 o lanovkových zariadeniach a zrušení smernice 2000/9/ES.
[148] Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/425 z 9. marca 2016 o osobných ochranných prostriedkoch a o zrušení smernice Rady 89/686/EHS.
[149] Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/426 z 9. marca 2016 o spotrebičoch spaľujúcich plynné palivá a zrušení smernice 2009/142/ES.
[150] Nariadenie Európskeho parlamentu a Rady (EÚ) 2017/745 z 5. apríla 2017 o zdravotníckych pomôckach, zmene smernice 2001/83/ES, nariadenia (ES) č. 178/2002 a nariadenia (ES) č. 1223/2009 a o zrušení smerníc Rady 90/385/EHS a 93/42/EHS.
[151] Nariadenie Európskeho parlamentu a Rady (EÚ) 2017/746 z 5. apríla 2017 o diagnostických zdravotníckych pomôckach in vitro a o zrušení smernice 98/79/ES a rozhodnutia Komisie 2010/227/EÚ.
[152] Nariadenie Európskeho parlamentu a Rady (ES) č. 300/2008 z 11. marca 2008 o spoločných pravidlách v oblasti bezpečnostnej ochrany civilného letectva a o zrušení nariadenia (ES) č. 2320/2002.

- nariadenie o schvaľovaní a dohľade nad trhom dvoj- alebo trojkolesových vozidiel a štvorkoliek[153]
- nariadenie o schvaľovaní poľnohospodárskych a lesných vozidiel a o dohľade nad trhom s týmito vozidlami[154]
- smernica o vybavení námorných lodí[155]
- smernica o interoperabilite železničného systému v Európskej únii[156]
- nariadenie o schvaľovaní motorových vozidiel a ich prípojných vozidiel, ako aj systémov, komponentov a samostatných technických jednotiek určených pre takéto vozidlá[157]
- nariadenie o spoločných pravidlách v oblasti civilného letectva, ktorým sa zriaďuje Agentúra Európskej únie pre bezpečnosť letectva.[158]

PRÍKLAD
Ak chce firma uviesť na trh zdravotnícku pomôcku, ktorá v sebe obsahuje systém AI, musí splniť požiadavky nielen regulácie zdravotníckych pomôcok, ale v zmysle článku 6 aj AIA, nakoľko na túto právnu úpravu odkazuje príloha II.

[153] Nariadenie Európskeho parlamentu a Rady (EÚ) č. 168/2013 z 15. januára 2013 o schvaľovaní a dohľade nad trhom dvoj- alebo trojkolesových vozidiel a štvorkoliek.

[154] Nariadenie Európskeho parlamentu a Rady (EÚ) č. 167/2013 z 5. februára 2013 o schvaľovaní poľnohospodárskych a lesných vozidiel a o dohľade nad trhom s týmito vozidlami.

[155] Smernica Európskeho parlamentu a Rady 2014/90/EÚ z 23. júla 2014 o vybavení námorných lodí a o zrušení smernice Rady 96/98/ES.

[156] Smernica Európskeho parlamentu a Rady (EÚ) 2016/797 z 11. mája 2016 o interoperabilite železničného systému v Európskej únii.

[157] Nariadenie Európskeho parlamentu a Rady (EÚ) 2018/858 z 30. mája 2018 o schvaľovaní motorových vozidiel a ich prípojných vozidiel, ako aj systémov, komponentov a samostatných technických jednotiek určených pre takéto vozidlá a o dohľade nad trhom s nimi, ktorým sa menia nariadenia (ES) č. 715/2007 a (ES) č. 595/2009 a zrušuje smernica 2007/46/ES (Ú. v. EÚ L 151, 14.6.2018, s. 1); 3. nariadenie Európskeho parlamentu a Rady (EÚ) 2019/2144 z 27. novembra 2019 o požiadavkách na typové schvaľovanie motorových vozidiel a ich prípojných vozidiel a systémov, komponentov a samostatných technických jednotiek určených pre tieto vozidlá, pokiaľ ide o ich všeobecnú bezpečnosť a ochranu cestujúcich vo vozidle a zraniteľných účastníkov cestnej premávky, ktorým sa mení nariadenie Európskeho parlamentu a Rady (EÚ) 2018/858 a ktorým sa zrušujú nariadenia Európskeho parlamentu a Rady (ES) č. 78/2009, (ES) č. 79/2009 a (ES) č. 661/2009 a nariadenia Komisie (ES) č. 631/2009, (EÚ) č. 406/2010, (EÚ) č. 672/2010, (EÚ) č. 1003/2010, (EÚ) č. 1005/2010, (EÚ) č. 1008/2010, (EÚ) č. 1009/2010, (EÚ) č. 19/2011, (EÚ) č. 109/2011, (EÚ) č. 458/2011, (EÚ) č. 65/2012, (EÚ) č. 130/2012, (EÚ) č. 347/2012, (EÚ) č. 351/2012, (EÚ) č. 1230/2012 a (EÚ) 2015/166

[158] Nariadenie Európskeho parlamentu a Rady (EÚ) 2018/1139 zo 4. júla 2018 o spoločných pravidlách v oblasti civilného letectva, ktorým sa zriaďuje Agentúra Európskej únie pre bezpečnosť letectva a ktorým sa menia nariadenia Európskeho parlamentu a Rady (ES) č. 2111/2005, (ES) č. 1008/2008, (EÚ) č. 996/2010, (EÚ) č. 376/2014 a smernice Európskeho parlamentu a Rady 2014/30/EÚ a 2014/53/EÚ a zrušujú nariadenia Európskeho parlamentu a Rady (ES) č. 552/2004 a (ES) č. 216/2008 a nariadenie Rady (EHS) č. 3922/91, pokiaľ ide o projektovanie a výrobu lietadiel uvedených v článku 2 ods. 1 písm. a) a b) a ich umiestňovanie na trh, ak sa týka bezpilotných vzdušných prostriedkov a ide o ich motory, vrtule, súčasti a vybavenie na ich diaľkové ovládanie.

Ak je AI súčasťou alebo samostatným produktom pri niektorej z uvedených produktových regulácií, vzťahuje sa na nich kľúčová časť AIA, ktorá obsahuje drvivú väčšinu povinností pre systémy AI.

Ak systém AI nespadá pod osobitnú reguláciu, adresáti povinností sú povinní reflektovať prílohu III AIA, ktorá ustanovuje oblasti AI vysokého rizika, na ktoré sa následne nariadenie taktiež aplikuje. Príloha III upravuje oblasti a konkrétne aplikácie oblastí vysokého rizika. Konkrétne:

Oblasť	Aplikácia
Biometrická identifikácia a kategorizácia fyzických osôb	systémy umelej inteligencie určené na používanie na diaľkovú biometrickú identifikáciu fyzických osôb „v reálnom čase" a „následne".
	systémy umelej inteligencie určené na biometrickú kategorizáciu podľa citlivých alebo chránených atribútov alebo charakteristík na základe odvodenia týchto atribútov alebo charakteristík
	systémy AI určené na rozpoznávanie emócií.
Riadenie a prevádzka kritickej infraštruktúry	systémy umelej inteligencie určené na použitie ako bezpečnostné prvky pri riadení a prevádzke kritickej digitálnej infraštruktúry, cestnej dopravy a dodávok vody, plynu, tepla a elektriny
Vzdelávanie a odborná príprava	systémy umelej inteligencie, ktoré sa majú používať na určovanie prístupu fyzických osôb do inštitúcií vzdelávania a odbornej prípravy alebo na ich priradenie k týmto inštitúciám.

	systémy umelej inteligencie určené na hodnotenie výsledkov vzdelávania vrátane prípadov, keď sa tieto sa používajú na riadenie procesu učenia sa fyzických osôb vo vzdelávacích a inštitúciách odborného vzdelávania na všetkých úrovniach.
	Systémy umelej inteligencie určené na účely posúdenia primeranej úrovne vzdelania, ktoré jednotlivec získa alebo bude mať prístup, v kontexte vzdelávacej inštitúcie a inštitúcie odborného vzdelávania/odbornej prípravy.
	Systémy umelej inteligencie určené na monitorovanie a odhaľovanie zakázaného správania študentov počas skúšok v rámci vzdelávania a odbornej prípravy inštitúcie.
Zamestnanosť, riadenie pracovníkov a prístup k samostatnej zárobkovej činnosti	systémy umelej inteligencie určené na nábor alebo výber fyzických osôb, najmä na inzerovanie voľných pracovných miest, preverovanie alebo filtrovanie žiadostí, hodnotenie uchádzačov počas pohovorov alebo skúšok
	systémy umelej inteligencie, ktoré sa majú používať pri rozhodovaní o postupe v zamestnaní a ukončení zmluvných pracovných vzťahov, pri prideľovaní úloh a monitorovaní a hodnotení výkonnosti a správania osôb v rámci takýchto vzťahov
Prístup k základným súkromným a verejným službám a dávkam a ich využívanie	systémy umelej inteligencie, ktoré sa majú používať orgánmi verejnej moci alebo v ich mene na hodnotenie oprávnenosti

	fyzických osôb na dávky a služby verejnej pomoci vrátane zdravotnej starostlivosti, ako aj na poskytovanie, zníženie či zrušenie takýchto dávok a služieb alebo na žiadosti o ich vrátenie
	Systémy umelej inteligencie, ktoré sa majú používať na hodnotenie úverovej bonity fyzických osôb alebo stanovenie ich bodového hodnotenia kreditného rizika, s výnimkou systémov určených na identifikáciu finančných podvodov.
	Systémy umelej inteligencie určené na vyhodnocovanie a klasifikáciu tiesňových volaní fyzických osôb alebo na dispečing alebo na stanovenie priority pri dispečingu tiesňových volaní prvej pomoci. služieb vrátane polície, hasičov a lekárskej pomoci, ako aj na systém triedenia pacientov v núdzovej zdravotnej starostlivosti.
	Systémy umelej inteligencie určené na hodnotenie rizík a stanovenie cien v súvislosti s prírodnými osôb v prípade životného a zdravotného poistenia.
Presadzovanie práva	systémy AI určené na použitie orgánmi činnými v trestnom konaní alebo v ich mene, alebo inštitúcie, agentúry, úrady alebo orgány Únie na podporu presadzovania práva orgány alebo v ich mene na posúdenie rizika, že sa fyzická osoba stane obeťou trestných činov

	systémy umelej inteligencie, ktoré majú používať orgány presadzovania práva, ako napríklad detektory lži a podobné nástroje;
	systémy AI určené na použitie orgánmi činnými v trestnom konaní alebo v ich mene, alebo inštitúcie, agentúry, úrady alebo orgány Únie na podporu presadzovania práva orgány na vyhodnocovanie spoľahlivosti dôkazov v priebehu vyšetrovania alebo stíhania trestných činov
	systémy umelej inteligencie, ktoré majú používať orgány presadzovania práva na hodnotenie spoľahlivosti dôkazov v priebehu vyšetrovania alebo stíhania trestných činov;
	systémy umelej inteligencie, ktoré majú používať orgány presadzovania práva na predvídanie výskytu alebo opakovaného výskytu skutočného alebo potenciálneho trestného činu. Nie výlučné založenom na profilovaní fyzických osôb uvedeného v článku 3 bode 4 smernice (EÚ) 2016/680, alebo na posúdenie osobnostných a povahových rysov alebo trestnej činnosti fyzických osôb alebo skupín v minulosti;
	systémy umelej inteligencie, ktoré majú orgány presadzovania práva používať na profilovanie fyzických osôb uvedené v článku 3 bode 4 smernice (EÚ) 2016/680 v

	priebehu odhaľovania, vyšetrovania alebo stíhania trestných činov;
Migrácia, azyl a riadenie kontroly hraníc	systémy umelej inteligencie, ktoré majú používať príslušné orgány verejnej moci, ako napríklad detektory lži a podobné nástroje,
	systémy umelej inteligencie, ktoré majú používať príslušné orgány verejnej moci na posúdenie rizika vrátane bezpečnostného rizika, rizika nelegálneho prisťahovalectva alebo zdravotného rizika, ktoré predstavuje fyzická osoba, ktorá má v úmysle vstúpiť na územie členského štátu alebo naň už vstúpila;
	systémy umelej inteligencie určené na pomoc príslušným orgánom verejnej moci pri posudzovaní žiadostí o azyl, víza a povolení na pobyt a súvisiacich sťažností týkajúcich sa oprávnenosti fyzických osôb žiadajúcich o určitý status.
	systémy AI určené na používanie príslušnými orgánmi verejnej moci alebo v ich mene, vrátane agentúr, úradov alebo orgánov Únie, v súvislosti s migráciou, azylom a riadenia kontroly hraníc na účely zisťovania, rozpoznávania alebo identifikácie fyzických osôb s výnimkou overovania cestovných dokladov
Výkon spravodlivosti a demokratické procesy	Systémy umelej inteligencie určené na použitie justičným orgánom alebo v jeho mene na pomoc súdnemu orgánu pri

	skúmaní a výklade skutočností a práva a pri uplatňovaní práva na konkrétny súbor skutočností alebo sa podobným spôsobom používajú v alternatívnych sporoch riešení sporov
	Systémy umelej inteligencie určené na ovplyvňovanie výsledkov volieb alebo referenda alebo správania fyzických osôb pri hlasovaní v voľbách alebo referendách. To nezahŕňa systémy AI, ktorých výstupom sú prirodzené osoby nie sú priamo vystavené, ako napríklad nástroje používané na organizovanie, optimalizáciu a štruktúrovanie politických kampaní z administratívneho a logistického hľadiska

Tabuľka: Vysokorizikové systémy AI podľa prílohy III AIA.

Zdroj: Príloha III AIA.

V neskoršej fáze legislatívneho procesu bola pridaná výnimka z klasifikácie AI systémov vysokého rizika podľa oblastí v Prílohe III. V zmysle tejto výnimky, AI systém nie je nutné klasifikovať ako vysokorizikový, ak nepredstavuje zásadné riziko vzniku škody, zdravia, bezpečnosti a základných práv a slobôd, vrátane situácie, ak systém podstatne neovplyvňuje výsledok rozhodovania. Zároveň AIA ustanovuje demonštratívny výpočet kritérií, kedy je možné predmetnú výnimku uplatniť:

- systém umelej inteligencie je určený na vykonávanie úzkej procesnej úlohy (*narrow procedural task)*;
- systém AI je určený na zlepšenie výsledku predtým vykonanej ľudskej činnosti;
- systém umelej inteligencie je určený na zisťovanie vzorcov rozhodovania alebo odchýlok od predchádzajúcich vzorcov rozhodovania a nie je určený na nahradenie alebo ovplyvnenie predtým vykonaného ľudského posúdenia bez riadneho ľudského preskúmania; alebo

- systém AI je určený na vykonávanie prípravnej úlohy pre posúdenie relevantné na účely prípadov použitia uvedených v prílohe III.

Bez ohľadu na danú výnimku, ak systém AI vykonáva profilovanie fyzických osôb a spadá do oblasti z prílohy III, AIA ho považuje za vysokorizikový systém AI.

Vzhľadom na dynamický vývoj AI nariadenie predpokladá dopĺňanie prílohy III prostredníctvom delegovaných aktov Európskej komisie. Na doplnenie musia byť splnené dve podmienky. Prvou je, že systémy AI sú určené na nasadenie v niektorej z oblasti uvedenej v Prílohe III AIA.[159] Druhá podmienka spočíva v tom, že *„systémy umelej inteligencie predstavujú riziko poškodenia zdravia a bezpečnosti alebo riziko nepriaznivého vplyvu na základné práva, ktoré je vzhľadom na svoju závažnosť a pravdepodobnosť výskytu prinajmenšom rovnocenné riziku poškodenia alebo nepriaznivého vplyvu."*[160] AIA zároveň ustanovuje kritéria, ktoré Európska komisia posudzuje pri prijímanie doplnenia Prílohy III.

3.1.2.1 Požiadavky na vysokorizikové systémy AI

Drvivá väčšina požiadaviek v AIA sa zameriava na reguláciu systémov AI vysokého rizika. Ak bude chcieť výrobca uviesť vysoko-rizikový systém AI na trh alebo používať, bude musieť v zmysle požiadaviek AIA splniť niekoľko krokov. Je nutné poznamenať, že AIA dáva dôraz na splnenie požiadaviek pred uvedením na trh (*ex ante*), aby sa minimalizovali riziká AI z hľadiska zdravia, bezpečnosti a rešpektovania základných ľudských práv pri jej používaní.

Prvým krokom je vykonanie tzv. posudzovania zhody (*conformity assessment*), čo je proces známy aj z iných regulácií, či ako súčasť auditovacích mechanizmov.[161]

Samotné posúdenie zhody má legálnu definíciu v AIA: *„posudzovanie zhody je postup overovania, či boli splnené požiadavky stanovené v hlave III kapitole 2 tohto nariadenia týkajúce sa systému umelej inteligencie."*[162] Jeho zmyslom je, aby výrobca systémov AI sám dbal na

[159] AIA, článok 7 ods. 1 písm. a).
[160] AIA, článok 7 ods. 1 písm. b).
[161] Pozri napríklad ŠIMKO, J a kol. Towards Continuous Automatic Audits of Social Media Adaptive Behavior and its Role in Misinformation Spreading. In *29th ACM Conference on UMAP'21*.
[162] AIA, článok 3 bod 20.

dodržiavanie požiadaviek AIA, ktoré mu nariadenie ustanovuje a reflektoval zásadu zodpovednosti (*accountability*).

Väčšina poskytovateľov vysokorizikových systémov AI sa bude riadiť postupom pre posúdenie zhody založenom na vnútornej kontrole, keďže externá kontrola je povinná len v obmedzenom počte prípadov.[163] Dôvodom povinnosti vykonávať posúdenie zhody interne na strane poskytovateľa je, že poskytovatelia systémov AI majú viac skúseností a rozumejú špecifikám svojich systémov.[164] Externá kontrola a teda vykonanie posúdenia zhody treťou stranou sa vyžaduje iba v prípade systémov AI, ktoré budú využívať biometrickú identifikáciu ľudí a systémy pre rozpoznávanie emócií.[165] Okrem toho, AIA upravuje povinnosti aj pre subjekty, ktoré AI nasadzujú.

Obsahové náležitosti posúdenia zhody je možné derivovať z kapitoly 2 hlavy III AIA. Tieto požiadavky sú koncipované s cieľom efektívne zmierniť riziká pre zdravie, bezpečnosť a základné ľudské práva a slobody v kontexte účelu posudzovaného systému AI.[166] Tieto požiadavky zahŕňajú:

- zriadenie, zavedenie a dokumentácia systému riadenia rizík[167]
- zavedenie procesov pre správu údajov[168]
- koncipovanie technickej dokumentácie[169]
- uchovávania záznamov[170]
- zachovanie transparentnosti a poskytovanie informácií používateľom[171]
- zavedenie ľudské dohľadu[172]
- požiadavky na presnosť, spoľahlivosť a kybernetickú bezpečnosť.[173]

[163] VEALE, M. ZUIDERVEEN BORGESIUS, F. Demystifying the Draft EU Arti ficial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach. In *Computer Law Review International*, vol. 22, no. 4, 2021.
[164] AIA, recitál 64.
[165] AIA, článok 43 ods. 1 v spojitosti s Prílohou III, bod 1.
[166] AIA, recitál 43.
[167] AIA, článok 9.
[168] AIA, článok 10.
[169] AIA, článok 11.
[170] AIA, článok 12.
[171] AIA, článok 13.
[172] AIA, článok 14.
[173] AIA, článok 15.

Správa údajov

Vysokorizikové systémy umelej inteligencie, ktoré sa pri rozhodovaní spoliehajú na trénované modely, si vyžadujú osobitné zohľadnenie údajov používaných na ich vývoj a prevádzku. AIA upravuje základné požiadavky na kvalitu a správu dát pre zodpovedný vývoj a nasadenie takýchto systémov do praxe v článku 10.

Z pohľadu kvality údajov sú trénovacie, validačné a testovacie súbory údajov základnými stavebnými prvkami vysokorizikových systémov umelej inteligencie. Na zabezpečenie spoľahlivosti a férovosti týchto systémov musia použité súbory údajov spĺňať špecifické kritériá kvality:

- Údaje musia byť relevantné pre konkrétny zamýšľaný účel systému umelej inteligencie a musia byť reprezentatívne pre populáciu, s ktorou bude systém interagovať. Tým sa zabezpečí, aby systém fungoval efektívne a aby sa zabránilo zovšeobecňovaniu z obmedzených alebo zaujatých údajov.
- Súbory údajov by mali byť čo najviac bez chýb a chýbajúcich informácií. Nepresné alebo neúplné údaje môžu viesť k nespoľahlivým výstupom a potenciálne škodlivým následkom.
- Údaje by mali mať vhodné štatistické vlastnosti, pričom sa zohľadní zamýšľané použitie a príslušná populácia. To zaručuje, že údaje presne odrážajú realitu a zabraňuje sa vyvodzovaniu zavádzajúcich záverov.
- Súbory údajov by zároveň mali zohľadňovať špecifické geografické, kontextové, behaviorálne alebo funkčné prostredie, v ktorom sa má vysokorizikový systém umelej inteligencie používať. To zabezpečuje, aby systém zohľadňoval nuansy rôznych prostredí a aby sa nerobili rozhodnutia na základe irrelevantných alebo nepoužiteľných údajov.

Okrem zabezpečenia kvality dát je kľúčové implementovať pre tieto súbory údajov aj vhodné postupy správy dát tzv. *data governance*. Tieto postupy zahŕňajú rôzne aspekty správy dát, ktorých cieľom je:

- Pochopiť procesy a zdroje údajov, najmä pôvodný účel získavania pre relevantné osobné údaje. To zaručuje transparentnosť pri získavaní údajov a zabraňuje zneužitiu informácií.
- Spravovať kroky spracovania údajov, ako je anotácia, označovanie, čistenie, aktualizácia, obohacovanie a agregácia. Tieto procesy by mali byť transparentné a zdokumentované, aby sa zabezpečila integrita a udržateľnosť údajov.
- Jasne definovať základné predpoklady o tom, čo údaje predstavujú a čo majú merať. To podporuje porozumenie obmedzeniam a potenciálnym skresleniam v datasetoch.
- Posúdiť dostupnosť a množstvo údajov, aby sa zistilo, či postačujú na zamýšľaný účel systému umelej inteligencie. Nedostatok údajov môže viesť k nespoľahlivým a potenciálne škodlivým výsledkom.
- Identifikovať a riešiť potenciálne skreslenia v údajoch, ktoré by mohli negatívne ovplyvniť zdravie, bezpečnosť, základné práva alebo viesť k diskriminácii. Zahŕňa to aj implementáciu opatrení na odhalenie, prevenciu a zmiernenie identifikovaných skreslení.
- Rozpoznať a riešiť obmedzenia v datasetoch.

Technická dokumentácia

AIA nariaďuje vytvorenie a udržiavanie technickej dokumentácie pre vysokorizikové systémy umelej inteligencie v zmysle článku 11 pred ich uvedením na trh alebo uvedením do prevádzky. Táto dokumentácia slúži na dva kľúčové účely:

- Poskytuje dôkaz o tom, že systém umelej inteligencie dodržiava požiadavky uvedené v AIA.
- Poskytuje príslušným orgánom a notifikovaným subjektom v jasnej a zrozumiteľnej forme potrebné informácie, ktoré im umožňujú účinne posúdiť súlad systému s týmito požiadavkami.

Technická dokumentácia musí minimálne obsahovať prvky uvedené v Prílohe IV AIA.

Uchovávanie záznamov (logov)

AIA taktiež vyžaduje, aby vysokorizikové systémy umelej inteligencie boli technicky vybavené na automatické zaznamenávanie udalostí (logovanie) počas celej ich životnosti. Tieto požiadavky sú upravené v článku 12 AIA. Požiadavka logovania slúži na tri kľúčové účely:

- Zaznamenávanie udalostí súvisiacich s potenciálnymi rizikami alebo s potrebou podstatných úprav systému.
- Uľahčenie monitorovania systémov umelej inteligencie po ich uvedení na trh, ako sa požaduje v článku 61 AIA.
- Umožnenie monitorovania vysokorizikových systémov umelej inteligencie.

Logy tým, že zabezpečujú úroveň sledovateľnosti zodpovedajúcu zamýšľanému účelu systému, poskytujú cenné informácie o správaní systému, čo umožňuje proaktívnu identifikáciu a zmiernenie potenciálnych rizík efektívne vyhodnotenie výkonu a bezpečnosti systému po implementácii a nepretržité monitorovanie funkcie systému a dodržiavania predpisov. Táto funkcia zohráva kľúčovú úlohu pri podpore zodpovedného vývoja a nasadenia vysokorizikových systémov umelej inteligencie.

Transparentnosť voči subjektom, ktoré nasadzujú vysokorizikový systém AI

Táto časť načrtáva dva kľúčové aspekty zodpovedného vývoja a nasadenia vysokorizikových systémov umelej inteligencie a to konkrétne transparentnosť a pokyny pre subjekty, ktoré vysokorizikové systémy AI nasadzujú. Tieto požiadavky sú predmetom úpravy článku 13 AIA.

Z pohľadu transparentnosti, vysokorizikové systémy umelej inteligencie musia byť navrhnuté a vyvinuté tak, aby podporovali interpretovateľnosť svojich výstupov. To umožňuje subjektom, ktorých ich nasadzujú a jednotlivcom pochopiť rozhodnutia systému a využívať ich vhodným spôsobom. Tým sa podporuje zodpovedné a účinné využívanie systému umelej inteligencie. Miera a povaha transparentnosti požadovaná pre každý vysokorizikový systém umelej inteligencie sa bude líšiť v závislosti od jeho špecifických vlastností a zamýšľaného účelu.

Inštrukcie pre používateľov musia minimálne zahŕňať jasné a dostupné pokyny pre používanie. Pokyny by mali minimálne obsahovať:

- Kontaktné údaje poskytovateľa a prípadne jeho autorizovaného zástupcu.
- Jasný popis toho, čo má systém umelej inteligencie dosiahnuť a na čo sa môže používať.
- Namerané úrovne presnosti, robustnosti a kybernetickej bezpečnosti dosiahnuté testovaním a validáciou.
- Potenciálne faktory, ktoré by mohli ovplyvniť tieto opatrenia, ako napríklad špecifické prípady použitia alebo environmentálne podmienky.
- Predvídateľné scenáre zneužitia, ktoré by mohli predstavovať riziko pre zdravie, bezpečnosť alebo základné práva.
- Schopnosť systému vysvetliť svoje výstupy a poskytnúť pohľad do jeho rozhodovacieho procesu, ak je to možné.
- Ak je to možné, informácie o tom, ako systém funguje v súvislosti s konkrétnymi skupinami ľudí, s ktorými má interagovať.
- Špecifikácie pre typ a formát údajov, ktoré systém vyžaduje ako vstup, spolu s relevantnými informáciami o trénovacích, validačných a testovacích súboroch údajov použitých pri jeho vývoji, s prihliadnutím na zamýšľaný účel systému, ak je to možné.
- Informácie, ako účinne interpretovať a využívať výstupy systému.
- Podrobnosti o akýchkoľvek očakávaných úpravách systému a jeho výkonnosti, ako ich predvída poskytovateľ po prvotnom posudzovaní zhody.
- Informácie o uplatnených opatreniach ľudského dohľadu a podporných technických riešeniach, ktoré uľahčujú interpretáciu výstupov systému.
- Informácie o výpočtových prostriedkoch potrebných na spustenie systému, jeho predpokladanej životnosti a potrebných údržbových opatreniach (vrátane aktualizácií softvéru) na zabezpečenie jeho správneho fungovania.

Ľudský dohľad

Aby bolo možné pri nesprávnom alebo škodlivom fungovaní vysokorizikového systému AI zasiahnuť, proces nemôže byť plne automatizovaný. Z tohto dôvodu po vzore viacerých politických alebo etických dokumentov či deklarácií zakotvuje AIA požiadavku na ľudský dohľad: *„Vysokorizikové systémy umelej inteligencie musia byť koncipované a vyvinuté tak, aby nad nimi počas obdobia používania systému umelej inteligencie mohli fyzické osoby vykonávať*

účinný dohľad, a to aj pomocou vhodných nástrojov rozhrania človek – stroj."[174] Požiadavka prakticky znamená, že vysokorizikové systémy AI musia byť dizajnované takým spôsobom, aby umožnili dohľad človeka nad jeho fungovaním. Tento dohľad by sa mal zameriavať na *„prevenciu alebo minimalizáciu rizík pre zdravie, bezpečnosť alebo základné práva, ktoré môžu vzniknúť pri používaní vysokorizikového systému umelej inteligencie v súlade so zamýšľaným účelom alebo za podmienok logicky predvídateľného nesprávneho použitia.*"[175] Samotný ľudský dohľad má uvedené požiadavky na jeho kvalitu. Osoba vykonávajúca dohľad musí rozumieť kapacitám a obmedzeniam dohladovaného vysokorizikového systému AI a efektívne ho monitorovať.[176] Ľudský dohľad zároveň musí vedieť správne interpretovať výstupy systému[177] a rozhodovať o zásahoch do jeho fungovania.[178] Osobitnou požiadavkou je, aby ľudský dohľad umožňoval vysokorizikový systém AI vypnúť.[179] Koncepcia ľudského dohľadu zahŕňa aj požiadavku na uvedomenie si potenciálneho automatizovaného skreslenia.[180] To prakticky znamená, že osoba, ktorá nad vysokorizikovým systémom AI vykonáva dohľad by mala vedieť o možnosti prílišného sa spoliehania na automatizované procesy bez kontroly, čo môže spôsobiť prehliadnutie rizík, chýb a omylov.[181]

Presnosť, spoľahlivosť a kybernetická bezpečnosť

Veľkú diskusiu spôsobili požiadavky AIA na presnosť, spoľahlivosť a kybernetickú bezpečnosti.[182] Tieto požiadavky zvýrazňujú atribúty odolnosti voči chybám, poruchám a nezrovnalostiam.[183] *„Vysokorizikové systémy umelej inteligencie, ktoré sa po uvedení na trh alebo do prevádzky ďalej učia, sa musia vyvíjať tak, aby sa zabezpečilo, že prípadné skreslené výstupy v dôsledku výstupov používaných ako vstup pre budúce operácie („slučky spätnej*

[174] AIA, článok 14 ods. 1.
[175] AIA, článok 14 ods. 2.
[176] AIA, článok 14 ods. 4 písm. a).
[177] AIA, článok 14 ods. 4 písm. c).
[178] AIA, článok 14 ods. 4 písm. d).
[179] AIA, článok 14 ods. 4 písm. e).
[180] AIA, článok 14 ods. 4 písm. b): *„Opatrenia uvedené v odseku 3 musia osobám, ktorým je zverený ľudský dohľad, umožniť, aby podľa okolností... si boli neustále vedomé možnej tendencie automatického spoliehania sa alebo nadmerného spoliehania sa na výstupy vytvorené vysokorizikovým systémom umelej inteligencie („automatizačné skreslenie"), a to najmä v prípade vysokorizikových systémov umelej inteligencie používaných na poskytovanie informácií alebo odporúčaní pre rozhodnutia, ktoré majú prijať fyzické osoby.*"
[181] K tomu napríklad SKITKA, L. J. MOSIER, K. L. BURDICK, M. Does automation bias decision-making? In *International Journal of Human-Computer Studies*, 1999, 51, 991-1006.
[182] AIA, článok 15 ods. 1: *„Vysokorizikové systémy umelej inteligencie musia byť koncipované a vyvinuté tak, aby vzhľadom na svoj zamýšľaný účel dosahovali primeranú úroveň presnosti, spoľahlivosti a kybernetickej bezpečnosti a aby v týchto ohľadoch konzistentne fungovali počas celého svojho životného cyklu.*"
[183] AIA, článok 15 ods. 3.

väzby") budú náležite riešené vhodnými zmierňujúcimi opatreniami."[184] Zároveň, musia byť tieto systémy odolné voči útokom a pokusom neoprávnených tretích strán o manipuláciu.[185]

Osobitne od požiadaviek pre prevádzkovateľov vysokorizikových systémov je nutné odlišovať požiadavky pre subjekty, ktoré vysokorizikový systém AI nasadzujú. Ide predovšetkým o požiadavky zabezpečenia ľudského dohľadu, kvality údajov v prípade dotrénovania modelu či informovania poskytovateľov a dozorných autorít pri incidentoch. Ďalšie požiadavky sa týkajú uchovávania logov a informovania užívateľov. Osobitne možno zvýrazniť požiadavku na vykonanie tzv. posúdenia vplyvu na základné ľudské práva a slobody pre vysokorizikové systémy nasadené vo verejnej správe (s výnimkou kritickej infraštruktúry).

PRÍKLAD

Posúdenie vplyvu na základné ľudské práva a slobody (FRIA) je systematický proces, ktorého cieľom je identifikovať, analyzovať a hodnotiť potenciálne vplyvy politiky, legislatívy, programu, projektu alebo akejkoľvek inej aktivity na základné ľudské práva a slobody. Cieľom FRIA je:

- Zabezpečiť, aby sa pri tvorbe a implementácii politík a programov zohľadňovali a rešpektovali základné ľudské práva a slobody.
- Predchádzať porušeniam ľudských práv a identifikovať vhodné opatrenia na zmiernenie alebo elimináciu negatívnych vplyvov.
- Podporovať a posilňovať dodržiavanie ľudských práv a slobôd.

Ako príklad možno uviesť situáciu, v ktorej vláda plánuje zaviesť nový program sociálnej pomoci prostredníctvom automatizovaného rozhodovania. V rámci FRIA sa zistí, že program by mohol mať negatívny vplyv na právo na súkromie a ochranu osobných údajov. V reakcii na to vláda implementuje opatrenia na ochranu súkromia a osobných údajov účastníkov programu, ako napríklad anonymizáciu údajov a zavedenie prísnych bezpečnostných protokolov.

[184] AIA, článok 15 ods. 3.
[185] AIA, článok 15 ods. 4.

Súlad s požiadavkami pre vysokorizikové systémy AI je možné realizovať aj odlišným spôsobom, ako priamym plnením povinností v zmysle AIA. Nariadenie obsahuje aj prezumpciu plnenia požiadaviek, ak vysokorizikové systémy AI spĺňajú príslušné harmonizované normy - štandardy.[186] Môže ísť o prípad dodržiavania povinností v oblasti kybernetickej bezpečnosti, keď bolo poskytovateľovi vysokorizikových systémov AI vydané vyhlásenie o zhode podľa schémy kybernetickej bezpečnosti podľa osobitného nariadenia EÚ.[187] Významnú úlohu organizácií rozvíjajúcich normy v zmysle AIA kritizovala akademická obec pre obavy súvisiace s ochranou základných práv a procesnými aspektmi preskúmavania týchto noriem z pohľadu nezávislých orgánov.[188]

Samotné posúdenie zhody je potrebné vykonať pred uvedením vysokorizikového systému AI na trh EÚ[189] alebo pred uvedením vysokorizikového systému AI do prevádzky.[190] Okrem toho sa posúdenie vplyvu vykoná aj v prípadoch, keď sú vysokorizikové systémy AI podstatne zmenenia. Toto je prípad zmeny ovplyvňujúcej súlad s nariadením alebo modifikácia zamýšľaného účelu. Neustále učenie sa systémov AI sa nepovažuje za podstatnú zmenu.[191]

Členské štáty EÚ sa môžu odchýliť od postupov posudzovania zhody na území príslušných členských štátov *„z výnimočných dôvodov verejnej bezpečnosti alebo ochrany života a zdravia osôb, ochrany životného prostredia a ochrany kľúčových priemyselných a infraštruktúrnych aktív.*"[192] Výnimka je však prísne obmedzená a posúdenie zhody by sa malo vykonať počas uplatňovania výnimky.[193]

[186] AIA, články 40 a 42.
[187] Nariadenie Európskeho parlamentu a Rady (EÚ) 2019/881 zo 17. apríla 2019 o agentúre ENISA (Agentúra Európskej únie pre kybernetickú bezpečnosť) a o certifikácii kybernetickej bezpečnosti informačných a komunikačných technológií a o zrušení nariadenia (EÚ) č. 526/2013 (akt o kybernetickej bezpečnosti). Ú. v. EÚ L 151, 7.6.2019, s. 15 – 69.
[188] Napríklad VEALE, M. ZUIDERVEEN BORGESIUS, F. Demystifying the Draft EU Arti ficial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach. In *Computer Law Review International*, vol. 22, no. 4, 2021, s. 105.
[189] Uvedenie na trh definuje AIA v článku 3 bode 9: *„uvedenie na trh je prvé sprístupnenie systému umelej inteligencie na trhu Únie.*"
[190] Uvedenie do prevádzky definuje AIA v článku 3 bode 11: *„uvedenie do prevádzky je dodanie systému umelej inteligencie na prvé použitie priamo používateľovi alebo na vlastné použitie na trhu Únie na zamýšľaný účel.*"
[191] AIA, recitál 66.
[192] AIA, článok 47 ods. 1.
[193] Tamže.

Druhým krokom po úspešnom posúdení zhody je registrácia systému AI v databáze EÚ pre samostatné vysokorizikové systémy AI.[194] Túto databázu bude spravovať samotná Európska komisia spolu s členskými štátmi. Databáza bude verejne dostupná a tak si každý užívateľ môže overiť, či je systém AI registrovaný a prešiel posúdením zhody.

EÚ následne pre registrovaný vysokorizikový AI systém vydá tzv. vyhlásenie o zhode.[195] Zároveň výrobca označenie zhody umiestni tak, aby bolo viditeľné, čitateľné a neodstrániteľné.[196]

Tretím krokom je *ex post* monitorovanie systémov AI po uvedení na trh. Prakticky to znamená výkon dohľadu a monitorovania vysokorizikových systémov AI. Po vzore iných právnych úprav je zakotvený inštitút nahlasovania incidentov pri využívaní systémov AI.[197]

Samotný dozor budú vykonávať príslušné vnútroštátne orgány a Úrad pre umelú inteligenciu.

3.1.3 Požiadavky na modely AI na všeobecné účely

Osobitnou kategóriou systémov AI, ktoré považujeme za vhodné diskutovať z hľadiska regulačných požiadaviek na nich kladených sú tzv. systémy alebo modely AI na všeobecné použitie. Systém umelej inteligencie na všeobecné účely je systém umelej inteligencie, ktorý je založený na modely AI na všeobecné účely, ktorý je schopný slúžiť na rôzne účely, a tak na priame použitie, ako aj na integráciu do iných systémov AI.

PRÍKLAD
Príkladom môžu byť tzv. multimodálne modely. Tieto modely dokážu spracovať a porozumieť rôznym typom údajov, ako sú obrázky, text a zvuk. To im umožňuje vykonávať úlohy, ktoré si vyžadujú kombináciu zručností, ako napríklad generovanie textu k obrázku alebo odpovedanie na otázky založené na informáciách z rôznych zdrojov. Príkladom multimodálneho modelu je Megatron-Turing NLG od Google AI.

[194] AIA, článok 60.
[195] AIA, článok 48.
[196] AIA, článok 49.
[197] AIA, článok 62.

AIA upravuje požiadavky na tieto modely podľa toho, či predstavujú systémové riziko. Systémové riziko predstavujú, ak:

- majú schopnosť vysokého vplyvu hodnotenú na základe vhodných technických nástrojov a metodík vrátane ukazovateľov a referenčných hodnôt,
- rozhodla o tom Európska komisia.

AIA prezumuje, že model AI na všeobecné použitie má schopnosť vysokého vplyvu, ak kumulatívny objem výpočtových prostriedkov použitých na jeho trénovanie meraný v operáciách s pohyblivou rádovou čiarkou (FLOP) je väčšie ako 10^25.

PRÍKLAD
FLOP v kontexte AI znamená operácie s pohyblivou rádovou čiarkou za sekundu (Floating-Point Operations Per Second). Ide o jednotku, ktorá sa používa na meranie výpočtového výkonu a efektivity modelov umelej inteligencie.

Vo všeobecnosti, požiadavky na AI na všeobecné použitie zahŕňajú:

- Poskytovatelia musia pre svoje modely vytvoriť a udržiavať **aktuálnu technickú dokumentáciu**. Táto dokumentácia by mala obsahovať podrobnosti o procese trénovania a testovania, výsledkoch hodnotenia a prvkoch v zmysle Prílohy AIA.
- Poskytovatelia musia vytvoriť a udržiavať **aktuálne informácie a dokumentáciu** špeciálne pre tých, ktorí majú v úmysle integrovať model do svojich systémov umelej inteligencie. Tieto informácie by mali umožňovať poskytovateľom systémov porozumieť schopnostiam a obmedzeniam modelu a pomôcť im dodržiavať príslušné predpisy.
- Poskytovatelia musia implementovať politiku, ktorá bude reflektovať požiadavky práva EÚ týkajúce sa autorského práva. To zahŕňa identifikáciu a rešpektovanie výhradných práv autorov.
- Poskytovatelia musia vytvoriť a zverejniť **podrobný súhrn** o obsahu použitom na trénovanie svojich modelov GPAI. Táto súhrnná správa by mala nasledovať štandardizovaný formát poskytnutý Úradom pre umelú inteligenciu.
- Poskytovatelia týchto modelov musia určiť zástupcu v EÚ.

Nad rámec požiadaviek uvedených vyššie, musia modely AI na všeobecné použitie so systémovým rizikom:

- **Hodnotiť** pomocou **štandardizovaných protokolov a najmodernejších nástrojov**. To zahŕňa aj **adversárne testovanie,** pri ktorom je model zámerne vystavený výzvam s cieľom odhaliť jeho slabiny a potenciálne riziká. Proces hodnotenia a jeho zistenia musia byť **zdokumentované,** aby sa informovalo o stratégiách na zmiernenie rizík.
- Poskytovatelia musia **identifikovať a riešiť potenciálne systémové riziká** spojené s vývojom, uvedením na trh alebo používaním svojich modelov. Tieto riziká sa posudzujú na **úrovni EÚ** s ohľadom na širší vplyv modelu. Na minimalizáciu týchto rizík by sa mali implementovať stratégie na ich zmiernenie.
- Poskytovatelia sú povinní **sledovať, dokumentovať a hlásiť** všetky **vážne incidenty** súvisiace s ich modelmi Úradu pre umelú inteligenciu a príslušným vnútroštátnym orgánom.

- Poskytovatelia musia zaviesť **adekvátne opatrenia v oblasti kybernetickej bezpečnosti** na ochranu samotného **modelu** aj jeho **fyzickej infraštruktúry** pred kybernetickými útokmi a narušením bezpečnosti.

3.1.4 Transparentnosť systémov AI

Transparentnosť systémov AI je vo všeobecnosti jednou z veľkých tém súčasnosti.[198] Zároveň je nutné uviesť, že pri systémoch AI má transparentnosť viaceré modality. Preto je nevyhnutné medzi týmito modalitami rozlišovať. Diferencovať ich možno na:

1) Povedomie (*awaraness*);
2) Vysvetliteľnosť (*explainability*); a
3) Dosledovateľnosť rozhodnutí (*logging*).

Otázky povedomia sa týkajú poskytovania informácií o systéme AI. Táto vedomosť by sa nemala limitovať iba na binárne zistenie, či sa takýto proces využíva alebo nie, ale zároveň by mal jednotlivec mať možnosť pochopiť, v čom proces rozhodovania systémom AI spočíva a aký to má vplyv na jeho výsledok.

Vysvetliteľnosť môže byť všeobecná (*ex ante*) alebo špecifická (*ex post*). Všeobecná vysvetliteľnosť znamená, že ešte predtým, ako je o jednotlivcovi rozhodnuté systémom AI, má k dispozícii všeobecné informácie o fungovaní tejto technológie.[199] Predmetná modalita je veľmi blízka otázke povedomia. Na druhej strane rozoznávame vysvetliteľnosť *ex post* v konkrétnych prípadoch s konkrétnym rozhodnutí a vplyvom na jednotlivca. To prakticky znamená, že jednotlivec by mal vedieť a pochopiť, prečo v jeho konkrétnom prípade bolo rozhodnuté systémom AI určitým spôsobom.

Poslednou modalitou transparentnosti je dosledovateľnosť rozhodnutí. To znamená, že prevádzkovateľ AI systému je schopný identifikovať, prečo bolo konkrétne rozhodnutie urobené výsledným spôsobom a na základe akých parametrov. Tieto informácie sa najčastejšie získavajú z tzv. logov (záznamov) systému AI. Uchovávanie a spracúvanie

[198] GOHEL, P. SINGH, P. MOHANTY, M. *Explainable AI: current status and future directions*. Dostupné na: https://arxiv.org/abs/2107.07045.
[199] WACHTER, S. MITTLESTADT, B. FLORIDI, L. Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. In *International Data Privacy Law*, Volume 7, Issue 2, May 2017, s. 76–99.

predmetných záznamov vo významnej miere môže pomôcť pri hlbšom sledovaní algoritmov a zároveň slúži ako prevencia pred diskriminačnými rozhodnutiami či inými negatívnymi dôsledkami využívania systémov AI, ako je uvedené vyššie.

V dnešnej dobe je osobitne dôležité regulovať otázky umelo generovaného obsahu. Konkrétne v prípadoch, kde ide o obsah manipulovaný systémami AI, ktorého cieľom je pôsobiť ako realita, tzv. deepfakes[200] a systémoch AI určených na interakciu s človekom. Vo všeobecnosti AIA klasifikuje systémy generujúce deepfakes ako nízko rizikové, avšak to platí iba v prípadoch, ak nie sú využívané v oblastiach vysokého rizika. AIA obsahuje aj definíciu deepfakes: *„obrazový, zvukový alebo video obsah vytvorený alebo upravený umelou inteligenciou, ktorý sa podobá existujúcim osobám, predmetom, miestam alebo iným subjektom či udalostiam a sa osobe falošne javí ako autentický alebo pravdivý."*[201]

PRÍKLAD
V roku 2019 bolo publikované deepfake video, v ktorom Nancy Pelosi, predsedníčka Snemovne reprezentantov USA, vyzerá, akoby bola opitá. Toto video bolo zdieľané vo veľkom a následne bolo použité na kritiku Pelosi.[202]

Samotná požiadavka transparentnosti pre systémy AI určené na interakciu s človekom je ustanovená nasledovným spôsobom: *„Poskytovatelia zabezpečia, aby systémy umelej inteligencie určené na interakciu s fyzickými osobami boli koncipované a vyvinuté tak, aby boli fyzické* ***osoby informované*** *o tom, že komunikujú so systémom umelej inteligencie, pokiaľ to nie je zrejmé z okolností a kontextu používania."*[203] V prípade systémov, ktoré generujú umelo vytvorený obsah, výstupy takýchto systémom musia byť identifikovateľné ako umelo generované.

[200] K tomu viac SAMPLE, I. *What are deepfakes – and how can you spot them?* The Guardian. Dostupné na: https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them.
[201] Pozícia Európskeho parlamentu, článok 3 bod 44d.
[202] Dostupné na: https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/.
[203] AIA, článok 52 ods. 1.

Podobné požiadavky obsahujú aj pravidlá pre generovanie a využívanie deepfakes, na základe ktorých sú subjekty, ktoré AI nasadzujú transparentne informovať o umelo vytvorenom obsahu.

3.2 Regulácia umelej inteligencie v Rade Európy

Téma regulácie AI nerezonuje iba na úrovni EÚ, ale aj Rady Európy, či už prostredníctvom kreovania orgánov a výborov, ale aj prípravy medzinárodných zmlúv. V roku 2021 ad hoc výbor Rady pre umelú inteligenciu (CAHAI) pri Rade Európy ukončil svoj mandát. Výbor ministrov dostal záverečný dokument s názvom "Možné prvky právneho rámca pre umelú inteligenciu založeného na normách Rady Európy v oblasti ľudských práv, demokracie a právneho štátu" a vytvoril tak základný rámec pre rokovania stálej pracovnej skupiny a tvorbu medzinárodného dokumentu. V roku 2022 sa konalo prvé zasadnutie Koordinačnej skupiny pre umelú inteligenciu s cieľom zabezpečiť trvalú koordináciu prác v oblasti umelej inteligencie medzi rôznymi sektormi organizácie. Táto koordinácia bude prebiehať nepretržite. Vytvorený bol aj Výbor pre umelú inteligenciu (CAI), poverený prípravou konceptu medzinárodnej zmluvy o AI.

V decembri 2023 bol verejnosti predstavený návrh rámcového dohovoru o umelej inteligencii, ľudských právach, demokracii a právnom štáte (ďalej len „Dohovor" alebo „Dohovor o AI").[204] Návrh Dohovoru o AI v určitom zmysle predstavuje revolúciu pri medzinárodných požiadavkách na AI. Medzi kľúčové otázky Dohovoru o AI patrí, že:

- Rámcový dohovor Rady Európy o umelej inteligencii sa po jeho finalizácii a implementácii stane prvou medzinárodnou zmluvou na svete zameranou na správu umelej inteligencie.
- Hlavným cieľom Dohovoru je zabezpečiť, aby systémy umelej inteligencie počas celého svojho životného cyklu rešpektovali ľudské práva, demokraciu a právny štát.
- V záujme štandardizácie sa Dohovor zosúlaďuje s OECD prostredníctvom prijatia rovnakej definície pojmu „systém umelej inteligencie". Táto uniformita je dôležitá

[204] Dostupný na: https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043.

pre jasnosť a konzistentnosť v medzinárodnom diskurze a regulácii umelej inteligencie.

- V súlade so svojím názvom stanovuje Návrh rámcového dohovoru súbor všeobecných povinností a zásad pre svoje zmluvné strany. Konkrétne špecifiká necháva na vnútroštátnych právnych predpisoch.
- Návrh rámcového dohovoru prijíma prístup založený na riziku, ale robí to inak ako AIA; neustanovuje stupne rizika, ani zoznamy vysoko rizikových alebo zakázaných systémov. Namiesto toho sa zameriava na všetky systémy umelej inteligencie, ktoré môžu mať potenciálny vplyv na ľudské práva, demokraciu a právny štát, a ponúka všeobecné povinnosti a rámec pre posudzovanie rizík.
- Vzhľadom na svoju povahu medzinárodného dohovoru Návrh rámcového dohovoru nestanovuje pokuty alebo sankcie pre jednotlivcov alebo firmy. Namiesto toho vyžaduje, aby jeho zmluvné strany zaviedli vnútroštátne mechanizmy monitorovania a dodržiavania, pričom ich pravidlá a zásady by mali zmluvné strany integrovať do svojich právnych poriadkoch.

Hlavným cieľom rámcového dohovoru je zabezpečiť, aby systémy AI boli v súlade so zásadami ľudských práv, demokracie a právneho štátu a aby ich dodržiavali počas celého svojho životného cyklu, ako sa uvádza v článku 1 ods. 1 Dohovoru. Definícia systému AI je rovnaká ako v AIA, nakoľko preberá definíciu z princípov OECD pre vývoj AI. Dohovor by sa mal vzťahovať na činnosti v rámci životného cyklu systémov umelej inteligencie, ktoré majú potenciál zasahovať do ľudských práv, demokracie a právneho štátu. Výnimky sú koncipované užšie ako v AIA a zahŕňajú vývoj a výskum, národnú bezpečnosť a verejnú bezpečnosť. Ohľadom presnej aplikácie výnimiek sa ešte vedie odborná diskusia.

Dohovor ukladá stranám dve široko koncipované povinnosti. V súlade s nimi každá zmluvná strana prijme a zachová opatrenia

- na zabezpečenie súladu činností systémov umelej inteligencie s vnútroštátnymi a medzinárodnými záväzkami v oblasti ľudských práv a
- na ochranu účasti na demokratických procesoch.

Okrem týchto dvoch všeobecných povinností sa Dohovor ustanovuje osem všeobecných zásad, ktoré majú zmluvné strany implementovať do svojho vnútroštátneho právneho systému pre systémy AI a to konkrétne:

- Rešpektovanie ľudskej autonómie a dôstojnosti
- Transparentnosť a dohľad
- Zodpovednosť
- Rovnosť a ne-diskriminácia
- Súkromie a ochrana osobných údajov
- Zachovanie zdravia
- Spoľahlivosť a dôvera
- Bezpečná inovácia.

Od zmluvných strán Dohovoru sa vyžaduje, aby prijali opatrenia na identifikáciu, posúdenie, prevenciu a zmiernenie rizík a vplyvov na ľudské práva, demokraciu a právny štát, ktoré vyplývajú z návrhu, vývoja, používania a vyraďovania systémov umelej inteligencie.

Vymáhanie pravidiel v Dohovore je viacvrstvové a zahŕňa kombináciu vnútroštátneho dohľadu, medzinárodnej spolupráce a následného mechanizmu dohľadu a konzultácií. Keďže ide zatiaľ len o návrh a zatiaľ nie je právne záväzný, jeho budúce presadzovanie bude závisieť od spôsobu jeho ratifikácie a začlenenia do vnútroštátnych právnych predpisov zmluvných strán. Napriek tomu sa v Dohovore načrtli niektoré mechanizmy presadzovania pre prípad, že sa stane právne záväznou:

- **Konferencia zmluvných strán:** V článku 24 sa zriaďuje konferencia zložená zo zástupcov zmluvných strán. Táto platforma uľahčuje pravidelné konzultácie a diskusie medzi zmluvnými stranami so zameraním na implementáciu a uplatňovanie zásad a záväzkov.
- **Medzinárodná spolupráca:** V článku 25 sa zdôrazňuje spolupráca medzi zmluvnými stranami, ako aj so štátmi, ktoré nie sú zmluvnými stranami. Táto spolupráca sa zameriava na podporu cieľov Dohovoru a zabezpečuje širšie medzinárodné zosúladenie s jeho cieľmi.

- **Vnútroštátny dohľad:** Od každej zmluvnej strany sa vyžaduje, aby vytvorila alebo určila aspoň jeden účinný mechanizmus na dohľad nad dodržiavaním Dohovoru. Tieto mechanizmy sa budú líšiť v závislosti od právnych a administratívnych štruktúr strany.

4 . Právna subjektivita umelej inteligencie

4.1 Úvodné poznámky

Zakaždým, keď sme dokázali uchopiť základné postuláty fungovania našich životov, zmenilo to smerovanie našej civilizácie.[205] Keď Newton objavil zákony pohybu a gravitácie položil základy veku strojov a priemyselnej revolúcie. Vysvetlenie elektriny a magnetizmu od Michaela Faradaya a Jamesa Clarka Maxwella vydláždilo cestu pre osvetlenie našich miest a poskytlo nám výkonné elektrické motory. Keď Erwin Schrödinger a Werner Heisenberg odhalili tajomstvá kvantovej teórie, dali nám dnešnú high-tech revolúciu s internetom, superpočítačmi a všetkými úžasnými technológiami sveta, v ktorom dnes žijeme.[206]

Výskum umelej inteligencie je však pomerne mladá disciplína výskumu, ktorá v sebe zahŕňa teoretické poznatky z viacerých oblastí vedy a to vrátane matematiky, logiky, počítačových vied, neurobiológie, sociálnych vied,[207] filozofie[208] a samozrejme aj práva.[209]

---

[205]Je skoro až kúzelné si predstaviť, že či už hviezdy z neba alebo perly z dna oceánov môžu byť naše, ak to dokážeme spísať na jednoduchý kúsok obyčajného papiera. (pozn. autora).

[206] Porovnaj KAKU, M. *The god equation, the quest for a theory of everything*, Penguin random house uk, 2021.

[207] Pri skúmaní spoločenských vplyvov umelej inteligencie môžeme diela rozdeliť do dvoch odlišných oblastí. Za prvé, ľudská perspektíva, ktorej analýza zahŕňa preskúmanie vplyvov umelej inteligencie na spoločnosť a človeka ako takého. Tu možno spomenúť výskumy Jona Kofasa, ktorý skúmal vplyv umelej inteligencie na „Cybergeneráciu", skupinu ľudí, pre ktorých sú videohry, telefóny a počítače novou realitou. Pozri bližšie KOFAS, J. Artificial Intelligence: Socioeconomic: Political And Ethical Dimensions. *Counter Currents*, [online]. 2017, [cit. 2023-06-15]. Dostupné na: http://www.countercurrents.org/2017/04/22/artificial-intelligence-socioeconomic-political-and-ethical-dimensions. Druhá oblasť zahŕňa otázku potreby etických obmedzení pri implementácií zariadení umelej inteligencie. V tejto oblasti Nick Bostrom pozri bližšie BOSTROM, N. Ethical Issues in Advanced Artificial Intelligence, [online]. 2003, [cit. 2023-06-15]. Dostupné na: https://nickbostrom.com/ethics/ai.html

[208] Pri spracovaní filozofických aspektov umelej inteligencie, môžeme primárne vychádzať z diel Georga Lugera, pre ktorého je Aristoteles prirodzeným východiskom pri skúmaní filozofických základov umelej inteligencie, keďže jeho diela tvoria základ modernej vedy. Luger považuje empirické a racionálne filozofické tradície za najvýznamnejšie smery pre vývoj umelej inteligencie, keďže tieto vplyvy zohrávali významnú úlohu pri rozvoji štruktúr programov a algoritmov. Pozri bližšie LUGER, F. G. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Pearson Education India, 1993. Taktiež nemožno opomenúť Aarona Slomana, ktorý skúmal interakciu filozofie a umelej inteligencie prostredníctvom intuície a nelogického uvažovania. Pozri bližšie SLOMAN, A. Interactions between philosophy and artificial intelligence: The Role of Intuition and Non-Logical Reasoning in Intelligence. *Artificial Intelligence,* 1971, roč. 2, č. 3-4, s. 209-225.

[209] Právne aspekty umelej inteligencie sa stali nosnou časťou výskumov až začiatkom posledného desaťročia. Najdiskutovanejšími témami sa stali právna subjektivita, zodpovednosť, transparentnosť, ochrana (osobných) údajov a možnosť regulácie týchto zariadení. Otázka zodpovednosti týchto zariadení sa stala najaktuálnejšou zo všetkých momentálne riešených právnych aspektov. Profesor filozofie Andreas Matthias (pozri bližšie MATTHIAS, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and information technology, 2004) pri tejto téme zdôrazňuje, že zodpovednosť týkajúca sa strojov je opísateľná ako podmienená kontrola, čo znamená, že osoba, ktorá vykonáva kontrolu nad strojom, preberá zodpovednosť. Avšak pri zariadeniach umelej inteligencie vzniká takzvaná „medzera v zodpovednosti", keďže tieto zariadenia fungujú oveľa komplexnejšie ako iné stroje, z čoho vyplýva, že bude potrebné zvážiť komplexnejšie scenáre zodpovednosti. Tieto skutočnosti viedli k nahromadeniu otázok medzi odborníkmi akými sú Matthias, Cole, Ramsey alebo Chagal. Pri domácej literatúre sa pri téme právnych aspektov umelej inteligencie môžeme oprieť o diela Radima Polčáka, pozri bližšie POLČÁK, R. Odpovědnost umělé inteligence a informační útvary bez právní subjektivity (Liability of AI and information trusts). *Bulletin advokacie*, 2018, č. 11, s. 21-28; ako aj POLČÁK,

Historicky vieme, že začiatky výskumov boli zahájené v tieni druhej svetovej vojny, úzko súviseli s vývojom v oblasti výpočtovej techniky, a prvotnou myšlienkou týchto výskumov bola delegácia čoraz zložitejších úloh na počítače, úloh, ktoré predtým mohol vykonať iba človek.[210] Ak však chceme akceptovať tézu, že stroj dokáže vykonávať úlohy, ktoré predtým mohol vykonať iba človek, nemali by sme nahliadať na tieto stroje nie len ako objekt práva ale aj ako subjekt práva?

4.2 K právnej subjektivite všeobecne

Právna subjektivita je jedným z esenciálnych právnych inštitútov, na ktorom stojí právny poriadok každej krajiny. Na základe neho priznávame vybraným subjektom možnosť participovať v právnych vzťahoch. Aj napriek tomu, že právnu subjektivitu tradične spájame s fyzickými a právnickými osobami, v ostatnom období vidíme často, ako sa právna subjektivita preplieta s inými entitami.

Historicky je právna subjektivita relatívne mladší právny inštitút, ktorý vznikol koncom 18. storočia predovšetkým vďaka pôsobeniu nemeckej právnej doktríny.[211] Na základe neho priznávame vybraným subjektom možnosť participovať v právnych vzťahoch. Získavajú totižto tzv. spôsobilosť byť nositeľom práv a povinností a uvedené je charakteristickým a esenciálnym znakom predmetného konceptu.[212] V priebehu 19. storočia sa daný konštrukt rozšíril naprieč celým svetom a dodnes je akceptovateľný v zahraničí[213] i v priestore strednej

R. Odpovědnost umělé inteligence a budoucnost průmyslu 4.0. In *Docuride*, 2018; prípadne pozti aj ŠTEDRON, B. a kol. *Právo a umělá inteligence*. Plzeň: Aleš Čeněk, 2020.

[210] Pozri. Gyurász, Z.: Umelá inteligencia v kontexte dejín našej spoločnosti. Šarm a noblesa diplomata [elektronický dokument] : venované pamiatke profesora Jozefa Klimka. - : 1. vyd. Bratislava : Právnická fakulta UK, 2022. - S. 267-276 [print] ((Historia et theoria iuris:supplement, ISSN 1338-0753; roč.14;).

[211] C. F. VON SAVIGNY, F. A. ZEILLER Je však nutné podotknúť, že predmetnému konceptu sa v určitej podobe venovali aj známi učenci ako HUGUES DONEAU, HERMANN VULTEJUS, HUGO GROTIUS, SAMUEL PUFENDORF, THOMAS HOBBES, GOTTFRIED LEIBNIZ, CHRISTIAN WOLF, ale aj dvaja filozofickí velikáni IMMANUEL KANT A G. W. F. HEGEL. K uvedenému pozri viac v KURKI, A. J. *A Theory of Legal Personhood*. Oxford: Oxford University Press, 2019, s. 31-53.

[212] Tento postoj zhodne s Visa Kurkim budeme nazývať ako tradičný pohľad na právnu subjektivitu. Dnes sa však objavujú viaceré odlišné koncepty právnej subjektivity, ktoré nekladú taký dôraz na podmienku spôsobilosti byť nositeľom práv a povinností a do popredia sa dostávajú aj ďalšie rozmanité znaky. Typickým príkladom je aj tzv. *bundle theory* právnej subjektivity, ktorú nedávno prezentoval VISA KURKI vo svojej publikácii *A Theory of Legal Personhood,* pozostávajúca z viacerých samostatných, ale navzájom prepojených zložiek (*incidents*). *Ibid*. s. 4, 15, 91-124.

[213] Hans Kelsen taktiež definoval fyzické (prirodzené) osoby a právnické osoby pomocou ich práv a povinností, dokonca odmietal kvalitatívny rozdiel medzi právnou subjektivitou obchodných spoločností a prirodzených ľudí. Subjektivita v právnom zmysle je len technická personifikácia komplexu noriem, práv a povinností. Porovnaj s KELSEN, H. *Čistá právna náuka.* Bratislava, KALLIGRAM, 2018, s. 85-88. Z neskorších autorov uveďme napríklad Richarda Tura, ktorý v podobnom duchu uvádza, že právna subjektivita je jednoducho *„prázdny priestor, ktorý môže zaplniť čokoľvek, čo je spôsobilé mať práva alebo povinnosti.*" či spomeňme STEVENA WISA, ktorý tvrdí, že právnická osoba sa rovná *„spôsobilosti vlastniť aspoň jedno zákonné právo (...)*" Porovnaj s TUR, Richard. The "Person" in Law. In: PEACOCKE, A. a GILLETT, G. (eds). *Persons and Personality. A Contemporary Inquiry*. Oxford, New York: B. Blackwell, 1987, s. 122; resp. WISE, M. Steven. Legal Personhood

Európy, vrátane Českej i Slovenskej republiky[214] pričom nie je naším cieľom rozsiahlejšie sa venovať konceptu právnej subjektivity vo všeobecnosti, nakoľko je tejto téme dnes venovaná pomerne značná pozornosť.[215]

Na tomto mieste si len spomeňme, že v právnej teórii rozlišujeme momentálne dva základné prístupy k právnej subjektivite.[216] Odlišnosti možno v súčasnosti vidieť predovšetkým v usporiadaní daného vzťahu medzi subjektom práva a vlastnosťou tohto nositeľa.[217]

Prvý prístup sa opiera o osobu a z nej následne vyplýva jej vlastnosť (spôsobilosť byť subjektom práv). Samozrejme, pod osobou sa v tomto prípade najčastejšie myslí človek ako

and the Nonhuman Rights Project. *Animal Law Review,* 2010, roč. 17, č. 1, s. 1. Elvia Arcelia Quintana Adriano obdobne pod právnou subjektivitou rozumie *„výtvor práva, ktorého úlohou je identifikovať subjekty s určitými právami a povinnosťami a zaručiť legitimitu činnostiam realizovaným na základe týchto práv a povinností.*" Viď v QUINTANA, A. Natural Persons, Juridical Persons and Legal Personhood. *Mexican Law Review,* 2015, roč. VIII, č. 1, s. 118.

[214] Len pre ukážku si dovoľujeme uviesť niekoľko definícií právnej subjektivity zo známych diel a učebníc v našom „československom" priestore, ktoré taktiež obsahujú vyššie uvedený znak právnej subjektivity : *„(…) své adresáty si právo vytváří, resp. definuje, samo a vymezuje jejich právní subjektivitu (právní osobnost), tj. jejich způsobilost k právům (způsobilost být subjektem práv a povinností) (…)*" KNAPP, V. *Teorie práva.* Praha: C.H.BECK, 1995, s. 71. *„Osobou v právním smyslu je ten, kdo může být nositelem práv a povinností.*" WEINBERGER, O. *Norma a instituce*. Brno: MU, 1995, s. 131.; *„Postavení jednotlivce je dáno právní subjektivitou, která spočívá: a) ve způsobilosti k právem a povinnostem – pasívní stránka subjqektivity; b) ve způsobilosti k právnímu jednání (úkonům) – aktivní stránka subjektivity; c) ve způsobilosti k protiprávnímu jednání (deliktní).*" HARVÁNEK, J. Právní vztahy. In: HARVÁNEK, J. a kol. *Teória práva*. Plzeň: Aleš Čeněk, 2008, s. 315. *„Pojem právní subjektivity sice navazuje na pojem právního subjektu, není s ním ale totožný. Subjekt je osoba, která má způsobilost mít práva a povinnosti, subjektivita pak definuje, o jaká práva a povinnosti se jedná, a vymezuje jejich rozsah.*" VEČEŘA, M. a kol. *Teória práva.* Bratislava: Eurókódex, 2011, s. 225. *„Subjektmi práva sú osoby v právnom zmysle, t. j. osoby, ktoré majú právnu subjektivitu, teda spôsobilosť mať subjektívne práva a povinnosti.*" DOBROVIČOVÁ, G. Právne vzťahy. In: BRÖSTL, A. a kol. *Teória práva*. Plzeň: Aleš Čeněk, 2013, s. 91. *„Právna subjektivita je platným právom priznaná a garantovaná možnosť subjektu byť nositeľom práv a povinností, byť subjektom (účastníkom) právneho vzťahu.*" OTTOVÁ, Eva. *Teória práva.* Šamorín: Heuréka, 2010, s. 288. *„Rozlišujeme dva druhy právnej spôsobilosti: spôsobilosť na práva a povinnosti a spôsobilosť na právne úkony. Prvá spomínaná spôsobilosť sa týka pasívnej stránky subjektu. Ide o spôsobilosť byť príjemcom (recipientom), resp. nositeľom práv a povinností.*" FÁBRY, B., KASINEC, R. TURČAN, M. *Teória práva*. Bratislava: Wolters Kluwer, 2019, s. 179. *„Pojem právnej subjektivity Občiansky zákonník síce nevymedzuje, ale zo zmyslu právnej úpravy možno vyvodiť, že ide o spôsobilosť fyzickej osoby stať sa subjektom právnych vzťahov a spôsobilosť nadobúdať práva a povinnosti. Právna subjektivita teda znamená spôsobilosť byť nositeľom práv a povinností.*" FEKETE, I. *Občiansky zákonník. Veľký komentár. 1.zväzok Všeobecná časť, §1 až § 122*. Žilina: Eurokódex, 2017, s. 151.

[215] Spomeňme opätovne publikácie uvedené v poznámke pod čiarou č. 4 tejto monografie, ako aj osobitne by sme vyzdvihli tieto: BERAN, Karel. *Pojem osoby v právu (osoba, morální osoba, právnická osoba).* Praha: Leges, 2012, 215 s.; BERAN, K. a kol. *Artificial legal entities: essays on legal agency and liability*. Praha: Wolters Kluwer Czech Republic, 2019, predovšetkým s. 10-92; BERAN, K: *The Concept of Juristic Person*. Prague: Wolters Kluwer, 2020, s. 152-187; prípadne pozri aj MÉSZÁROS, T. ŠVRČKOVÁ, N. Meno, mesto, zviera, vec: spoločenská hra s právnou subjektivitou? In: *Paneurópske právnické fórum* [online]. Bratislava : Paneurópska vysoká škola, 2017, s. 510-523 [cit. 2023-06-15], dostupné z: https://www.paneurouni.com/wp-content/uploads/2017/03/paneuropske_pravnicke_forum_2017.pdf; pre právnu subjektivitu právnických osôb pozri aj GAJDOŠOVÁ, Martina. *Združenia a sloboda združovania*. Bratislava: C.H.Beck, 2019, s. 150-152, 170-172 a 219-250.

[216] Východisková otázka znie, či právna subjektivita vychádza zo skutočnosti, že je niečo chápané ako osoba, alebo existencia subjektu práva vyplýva z toho, že danej entite bola priznaná právna subjektivita.

[217] V rámci recentných štúdií, ktoré sa danej problematike venujú, si dovolíme spomenúť predovšetkým dve, a to HURDÍK, Jan. Pojem osoba a geneze jeho obsahu jako základ konstrukce osob v právním smyslu. *Časopis pro právní vědu a praxi*, 2000, roč. 8, č. 3, s. 306 a nasl.; BERAN, K. Proč a kdy byla nahrazena „osoba" právním subjektem?: Úvahy ke genezi pojmu osoba v právním smyslu. *Časopis pro právní vědu a praxi*, 2011, roč. 19, č. 2, s. 108 a nasl.

prirodzená individuálna bytosť. Avšak zatiaľ čo *človek* je pojem patriaci do fyzikálneho sveta, v právnom priestore je jeho ekvivalentom pojem *osoba*.[218] Právnu spôsobilosť bude teda mať entita, ktorá spĺňa definičné znaky osoby (pričom tie sa naprieč dejinami postupne vyvíjali).[219]

Druhý prístup vychádza z myšlienky, že právna subjektivita je potenciálny súbor subjektívnych práv, ktoré subjektu priznáva platné právo.[220]

Tradične sa historicky uplatňoval prístup prvý, v rámci ktorého bol východiskom pre teóriu práva človek ako biosociálna bytosť, no v súčasnosti má svoje silné zastúpenie v našom právnom priestore aj druhý prístup. Ide o dôsledok najmä právneho normativizmu a jeho autorít,[221] Práve na tento druhý prístup chceme nadviazať, keďže jeho primárne východisko je späté s týmito súčasnými tendenciami: a) stieranie rozdielov medzi fyzickou a právnickou osobou, b)rozširovanie okruhu entít, ktorým je možné priznať právnu subjektivitu.[222]

Práve posledný bod je z nášho pohľadu zaujímavý a domnievame sa, že trend chápať právnu subjektivitu nie z pohľadu spôsobilosti byť osobou v právnom zmysle (tzn. byť nadaný vybranými atribútmi ako napr. rozum, vôľa, sloboda a pod.), ale z pohľadu vôle zákonodarcu priznať určitej entite právnu subjektivitu, je dnes vo svete často využívaný pri priznávaní vybraných práv a právnej subjektivity aj takým entitám, o ktorých by sme pred pár storočiami ani neuvažovali. Totižto v posledných rokoch možno vnímať v tejto oblasti záujmu aj novú

[218] Porovnaj s BERAN, K.. Proč a kdy byla nahrazena „osoba" právním subjektem?: Úvahy ke genezi pojmu osoba v právním smyslu. *Časopis pro právní vědu a praxi*, 2011, roč. 19, č. 2, s. 109.

[219] Daná skupina znakov, určujúca a dôležitá pre vymedzenie osoby v právnom zmysle, sa v minulosti vyvíjala. Azda najčastejšie spomínané atribúty sú nasledovné: rozum, vôľa, sloboda, sebauvedomenie či schopnosť vyvodenia zodpovednosti za svoje konanie. Pozri viac v HURDÍK, J. Pojem osoba a geneze jeho obsahu jako základ konstrukce osob v právním smyslu. *Časopis pro právní vědu a praxi*, 2000, roč. 8, č. 3, s. 309-312. Samozrejme, nie každý človek mal v minulosti automaticky (úplnú) právnu subjektivitu. Napríklad v rímskom práve právna subjektivita závisela od statusov slobody, občianstva a postavenia v rodine, aj keď základnú mieru subjektivity mala každá *slobodná* osoba. Otroci síce boli ľudské bytosti, ale z pohľadu rímskeho práva mali postavenie veci, preto v rámci právnych vzťahoch vystupovali ako objekty. Pre viac odporúčame pozrieť v GREGOR, M. *Základy rímskeho práva. Historický úvod, pramene a subjekty.* Praha: Leges, 2022, s. 152-165, 172 a 190-192. Obdobne stredoveké feudálne právo nepoznalo všeobecnú právnu subjektivitu ľudí a plnú právnu subjektivitu požívali len šľachtici a slobodní. Porovnaj s BERAN, K. Proč a kdy byla nahrazena „osoba" právním subjektem?: Úvahy ke genezi pojmu osoba v právním smyslu. *Časopis pro právní vědu a praxi*, 2011, roč. 19, č. 2, s. 113-114; prípadne pre širšie historické okolnosti porovnaj aj s LYSÝ, M. MALATINSKÝ, M. PUCHOVSKÝ, J. SOMBATI, J. *History of Public Law*. Bratislava: Wolters Kluwer, 2019, napríklad s. 78-95; prípadne aj v LYSÝ, Miroslav. *Slovenské právne dejiny I. Vývoj ústavného a správneho práva na Slovensku od najstarších čias po súčasnosť*. Šamorín: Heuréka, 2021, s. 103-120. Pre ďalší vývoj chápania právnej subjektivity (predovšetkým v novoveku) pozri predovšetkým KURKI . *A Theory of Legal Personhood*. Oxford: Oxford University Press, 2019, s. 31-53.

[220] Porovnaj MÜLLEROVÁ, H. Mohou mít zvířata práva? Vývoj konceptu právní subjektivity a právní postavení zvířat. *Právník*, 2012, roč. 151, č. 10, s. 1078.

[221] Najmä Hansa Kelsena a Františka Weyra.

[222] Porovnaj s HURDÍK, J. *Osoba a její soukromoprávní postavení v měnícím se světě*. Brno: Masarykova univerzita, 2004, s. 86 a nasl.; ako aj MÜLLEROVÁ, H. Mohou mít zvířata práva? Vývoj konceptu právní subjektivity a právní postavení zvířat. *Právník*, 2012, roč. 151, č. 10, s. 1081-1082.

rovinu skúmania, ktorej sa čím ďalej, tým viac venuje pozornosť.[223] A to priznanie právnej subjektivity a vybraných práv a povinností nonhumánnym entitám.

4.3. Priznanie práv a právnej subjektivity vo svetle umelej inteligencie

Otázka priznania práv a právnej subjektivity umelej inteligencii sa objavila aj v dokumente Európskeho parlamentu, kde v Návrhu uznesenia Európskeho parlamentu s odporúčaniami pre Komisiu k normám občianskeho práva v oblasti robotiky sa píše, že je potrebné: *„vytvorenie špecifického právneho postavenia pre roboty v dlhodobom horizonte, aby sa aspoň tie najsofistikovanejšie autonómne roboty mohli považovať za elektronické osoby zodpovedné za náhradu akejkoľvek škody, ktorú môžu spôsobiť, a prípadného uplatnenia elektronickej osobnosti na prípady, keď roboty robia samostatné rozhodnutia alebo iným*

[223] Prvotný impulz tejto problematike dal doposiaľ sporný právny status zvierat. Vo väčšine krajín sú dodnes zvieratá v určitej podobe a intenzite chránené právom, no same o sebe zväčša nie sú subjektami práva, t. j. nemajú spôsobilosť na práva a povinnosti. Avšak minimálne posledné polstoročie sme svedkami mnohých iniciatív a hnutí, ktoré sa snažia uvedené zmeniť a priznať zvieratám isté práva a slobody a nepovažovať ich len za „veci". Domnievame sa, že „bibliou" hnutia za práva zvierat sa stala dnes už známa kniha od Petra Singera *Animal Liberation* z roku 1975, pričom svoje argumenty opiera o utilitarizmus a v zásade jednoduché kritérium, ktorým je schopnosť cítiť bolesť a potešenie. (Porovnaj so SINGER, P. *Animal Liberation.* New York: Harper Collins Publishers, 2002, s. 15. Medzi ďalšie dôležité diela spadajúce do hnutia za práva zvierat môžeme uviesť: REGAN, T. *The Case for Animal Rights*. Berkeley and Los Angeles: University of California Press, 1983, 425 s.; FRANCIONE, L. G. *Animals, Property, and the Law*. Philadelphia: Temple University Press, 1995, 350 s.; WISE, M. S. *Rattling the Cage: Toward Legal Rights for Animals*. New York: Perseus Publishing, 2000, 362 s.; FRANKLIN, H. J. *Animal Rights and Moral Philosophy*. New York: Columbia University Press, 2005, 151 s.; či FAVRE, S. D. *Animal law: welfare, interests, and rights*. New York: Aspen Publishers, 2008, 495s.) Napriek uvedeným iniciatívam však právny status zvierat je v zmysle všeobecnej teórie práva v našom priestore stále rovnaký, tzn. zvieratá sú v rámci právnych vzťahov naďalej v postavení objektov právnych vzťahov, nie subjektov. V súčasnosti však sme svedkami moderného trendu v súkromnom práve s názvom „dereifikácia" (odvecnenie) zvierat, ktorý odmieta historické chápanie zvierat ako vecí, pričom aj Slovenská republika sa pripojila k rade ďalších európskych krajín, ktoré odmietajú považovať zvieratá výslovne za veci, a to prijatím novelizačného zákona č. 184/2018 Z. z. V zmysle súčasnej občianskoprávnej úpravy živé zviera má osobitný význam a hodnotu, ako aj má mať osobitné postavenie, keďže ako živý tvor je schopné vnímať vlastnými zmyslami a pociťovať bolesť (porovnaj predovšetkým s § 119 ods. 3 OZ). Stále je však nutné pamätať, že zvieratá síce majú osobitné postavenie v občianskoprávnych vzťahoch, avšak nie sú subjektami v právnom zmysle. Z tohto dôvodu zvieratá dodnes môžu byť: a) predmetom záväzkového právneho vzťahu; b) predmetom vecnoprávneho (vlastníckeho) vzťahu; c) predmety bez pána (predovšetkým voľne žijúce zvieratá) d) predmetom osobitnej ochrany bez ohľadu na to, či žijú voľne alebo sú chované (napríklad v zmysle ustanovenia § 22 zákona č. 39/2007 Z. z. o veterinárnej starostlivosti). (Pre viac k vývoju ponímania právneho statusu zvierat pozri v DOLEŽAL, A. Vývoj pojetí zvířete v právním diskursu. In: MÜLLEROVÁ, H. ČERNÝ, D. DOLEŽAL, D. *Kapitoly o právech zvířat: "my a oni" z pohledu filosofie, etiky, biologie a práva*. Praha: Academia, 2016, s. 309 – 416; ako aj pre súčasnú právnu úpravu v Ceskej republike pozri viac v KOUKAL, P. Komentár k § 494. In: LAVICKÝ, P a kol. *Občanský zákoník I. Obecná část (§ 1-654). Komentář*. Praha: C. H. Beck, 2014, s. 1746 a nasl.; prípadne i v HUBKOVÁ, P. Komentár k § 494. In PETROV, J. VÝTISK, M. BERAN, V. a kol. *Občanský zákoník. Komentář*. Praha: C. H. Beck, 2019, s. 547-548; resp. pre právnu úpravu v Slovenskej republike pozri viac v FEKETE, I. *Občiansky zákonník. 1. zväzok (Všeobecná časť). Veľký komentár*. Bratislava: Eurokódex, 2017, s. 974 a nasl. Súčasne s dereifikáciou zvierat v súkromnom práve je spojená aj značná kritika, k tomu odporúčame pozrieť analýzu od Martina Škopa, ktorý poukazuje na to, že sama snaha vytvoriť zvláštnu kategóriu zvierat v súkromnom práve je síce chvályhodná, avšak normotvorca týmto nehovorí nič nového a neposilňuje sa ochrana zvierat. Pozri viac v ŠKOP, M. Symbolická reprezentáce zvířete v moderním právu. In: MÜLLEROVÁ, Hana, ČERNÝ, David a Adam DOLEŽAL. *Kapitoly o právech zvířat: "my a oni" z pohledu filosofie, etiky, biologie a práva*. Praha: Academia, 2016, s. 494 a nasl.) Uvedené znamená, že sa na zvieratá môžu vzťahovať práva a povinnosti adresované právnymi normami subjektom práva (najčastejšie pôjde o fyzické a právnické osoby), avšak nie sú subjektmi právnych vzťahov, tzn. nemôžu byť povinnými alebo oprávnenými subjektmi, domáhať sa vo svojom mene vlastných práv, byť deliktuálne spôsobilým subjektom a podobne.

spôsobom nezávisle komunikujú s tretími stranami."[224] Treba však spomenúť, že od tejto myšlienky sa pomerne rýchlo opustilo, keďže existoval značný odpor voči nej. Viacerí vytýkali tejto alternatíve, že chýbajú empirické dôkazy, ktoré by podložili potrebu takýchto ustanovení.

Dnes ešte musíme akceptovať, že kým sa systémy umelej inteligencie považujú iba za „objekt práva", nie sú schopné nahradiť škodu ani znášať následky svojich konaní. Hoci systémy na báze umelej inteligencie napodobňujú ľudskú inteligenciu, chýba im subjektivita z právneho pohľadu, ktorá by im umožnila mať tieto práva a povinnosti v tradičnom zmysle. Bude preto vždy potrebné nájsť, aká konkrétna osoba (subjekt práva) bude za tieto systémy zodpovedať. Avšak determinovať tento subjekt je v mnohých prípadoch veľmi komplikované a navyše absentuje príslušná právna úprava, ako aj judikatúra, ktoré by sa týmito otázkami zaoberali.

V týchto otázkach nám veľmi nedokážu pomôcť ani tradičné teórie týkajúce sa právnej subjektivity. Príkladom môže slúžiť komparácia priznania právnej subjektivity systémom na báze umelej inteligencie k prípadom priznania právnej subjektivity obchodným spoločnostiam.[225] Je zrejmé, že samotná právna fikcia subjektivity obchodných spoločností má zjavné ekonomické účinky a zároveň slúži konkrétnym účelom. Argumenty v prospech priznania právnej subjektivity právnickým osobám (predovšetkým obchodných spoločností) sa zvyčajne týkajú sociálnych a hospodárskych účinkov[226], ako sú finančná transparentnosť alebo zodpovednosť. Pri obchodných spoločnostiach sa však zvyčajne nehovorí o aspektoch ako dôstojnosť, vedomie, ako ani o iných morálnych hodnotách,[227] ktoré musia byť zohľadnené pri systémoch umelej inteligencie. Na druhej strane tí, ktorí argumentujú proti priznaniu právnej subjektivity systémom umelej inteligencie, hovoria, že tieto systémy nemajú pre subjektivitu žiadne kritické vlastnosti, ako vedomie, zámernosť alebo záujmy. V

[224] Draft Report of Committee on Legal Affairs (Rapporteur: Mady Delvaux) with recommendations to the Commission on Civil Law Rules on Robotics 2015/2103(INL), [online]. 2017, [cit. 2023-06-15]. Dostupné z: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html

[225] Porovnaj MARTINCOVÁ, L. Od obchodných spoločností k socialistickým organizáciám. Vývoj právnickej osoby po roku 1950 s prihliadnutím k hospodárskym vzťahom. *Historia et theoria iuris*, 2020, roč. 12, č. 1, s. 58-73.

[226] Z historického kontextu porovnaj s KÖSZEGHY, A. Podnik zahraničného obchodu ako forma právnickej osoby v medzinárodnom obchode. *Historia et theoria iuris*, 2020, roč. 12, č. 1, s. 32-50; alebo aj so ŠVEDOVÁ, T. Právne postavenie iniciatívy Verejnosť proti násiliu (november-december 1989). *Iurium Scriptum*, 2020, roč. 4, č. 2, s. 34-44.

[227] Porovnaj s TURČAN, M. *Etické dimenzie teórie práva*. Bratislava: Univerzita Komenského v Bratislave, Právnická fakulta, 2022.

prípade, že systém by dokázal preukázať správanie, ktoré by mohlo byť dôkazom spomínaných kvalít, znamenalo by to len tú skutočnosť, že autonómny stroj napodobňuje ľudské správanie a simulácia veci predsa nie je vecou samotnou.[228]

Môže sa zdať, že väčšina právnych systémov by mohla poskytnúť systémom na báze umelej inteligencie istú formu právnej subjektivity, keďže v teoretickej rovine neexistujú žiadne materiálne ani formálne prekážky pre udelenie právnej subjektivity autonómnym strojom.[229] Nemôžeme však zabúdať, že neexistujú žiadne technologické, ontologické či právne dôvody, na základe ktorých by sme mali považovať tieto stroje za rovnocenné s ľudskými bytosťami. Koniec koncov tieto stroje sú stále iba veci, artefakty a výrobky ľudského intelektu.[230] Preto vo všeobecnosti prevládajú tri kategórie názorov pri regulácií umelej inteligencie v kontexte právnej subjektivity.

V prvom rade ide o priznanie právnej subjektivity umelej inteligencii, pričom ako dôsledok by bolo, že systémy na báze umelej inteligencie by dokázali znášať následky svojich konaní. Samozrejme, pri takomto rozhodnutí by bolo treba zodpovedať aj niekoľko kompilovaných nadväzujúcich otázok.[231] Z čisto pragmatického hľadiska zatiaľ nie je žiadne zariadenie na báze UI považované za subjekt práva.[232] Táto argumentácia sa spolieha najmä na dôvody ako:

a) Stroj nemá „intencionalitu" v klasickom ponímaní, ktorá by sa dala formovať prostredníctvom stimulov,

[228] SOLUM, L. Legal Personhood for Artificial Intelligences. *North Carolina Law Review*, Volume 70, [online]. 1992, [cit. 2023-06-15]. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1108671.

[229] Opäť len poukazujeme na myšlienky napríklad Richarda Tura, ktorý definuje právnu subjektivitu jednoducho ako *„prázdny priestor, ktorý môže zaplniť čokoľvek, čo je spôsobilé mať práva alebo povinnosti.*". Porovnaj TUR, Richard. The "Person" in Law. In: PEACOCKE, A. a GILLETT (eds). *Persons and Personality. A Contemporary Inquiry*. Oxford, New York: B. Blackwell, 1987, s. 122.

[230] BERTOLINI, A. Artificial Intelligence and Civil Liability. [online]. 2020. [cit. 2023-06-15]. Dostupné z: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf.

[231] Porovnaj MATTHIAS, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, Springer, s. 175–183 [online]. 2004. [cit. 2023-06-15]. Dostupné z: https://link.springer.com/article/10.1007/s10676-004-3422-1#citeas.

[232] BERTOLINI, A. Artificial Intelligence and Civil Liability, [online]. 2020, [cit. 2023-06-15]. Dostupné z: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf .

b) Trestanie týchto strojov nebude účelné, keďže pri strojoch nebude existovať korelácia medzi samotným trestom a účelom trestania. Stroj sa nebojí sankcie, svoju existenciu nevníma, ani si ju neváži a nesnaží sa ju uchovať.[233]

To robí značné rozdiely medzi správaním stroja a klasických subjektov práva. Aj na základe týchto argumentov (a i iných, ku ktorým sa ešte dostaneme v tejto monografii) považujeme túto alternatívu za zatiaľ nerealizovateľnú.

Druhou možnosťou je upraviť súvisiace práva a povinnosti za úkony umelej inteligencie prostredníctvom entity, ktorá ju vytvorila alebo ju používa. Zástancovia tejto teórie tvrdia, že vždy je možné identifikovať ľudskú bytosť (subjekt práva), ktorá sa má považovať za subjekt práva, bez ohľadu na to, aký nezávislý alebo schopný je daný systém.[234] Zástancovia tejto teórie sa spoliehajú na skutočnosť, že nájdenie subjektu je iba vecou komplexnej analýzy, čo v konečnom dôsledku umožní za každých okolností nájsť správny subjekt.[235] Problematickým aspektom tejto teórie je však jej samotná realizácia. Kým v teórii je možné sa spoliehať na tézu, že vždy za každých okolností sa dá identifikovať subjekt práva, prax ukázala, že vo väčšine prípadov sa aktéri spoliehali na spoločnú zodpovednosť,[236] miesto hľadania správneho subjektu, a to najmä z dôvodu nesmiernej náročnosti samej úlohy.[237]

Tretou možnosťou je alternatíva odvodená z analógie udeľovania subjektivity v prípadoch prírody. Ako sme už vyššie spomínali, právna subjektivita je základom každého

[233] Z tejto perspektívy by dokonca aj trest, ktorý by mohol stroj priamo deaktivovať alebo odpojiť daný stroj, trestal najmä strany, ktoré tento stroj vlastnia.

[234] MATTHIAS, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, Springer, [online]. 2004. [cit. 2023-06-15].Dostupné z: https://link.springer.com/article/10.1007/s10676-004-3422-1#citeas

[235] BERTOLINI, A. Insurance and Risk Management for Robotic Devices: Identifying the Problems. *Global Jurist*, [online]. 2016, [cit. 2023-06-15]. Dostupné z: https://www.researchgate.net/publication/305623978_Insurance_and_Risk_Management_for_Robotic_Devices_Identifying_the_Problems.

[236] Porovnaj AGHEMO, R. Recent EC interventions on AI safety and liability, [online]. 2020 [cit. 2023-06-15]. Dostupné z: https://medium.com/ai-in-plain-english/recent-ec-interventions-on-ai-safety-and-liability-8e388573407b

[237] Náročnosť danej veci vyplýva z problému, ktorý je známy ako „*the black box problem*". Tento koncept v podstate popisuje problém vysvetliteľnosti rozhodnutí, ktoré urobí inteligentný systém. Porovnať by sme to mohli k Heisenbergovmu princípu neurčitosti. A teda, že síce vieme akými informáciami bol inteligentný systém „krmený" a vieme k akému záveru systém dospel. Avšak súčaste nám nebudú presne známe, podľa ktorých vybraných dát a ako metodou dospel systém k danému záveru. Preto ako Heisenberg tvrdil o subatomatrných časticiach aj my môžeme tvrdiť o UI, že *„ani minulé ani budúce správanie (...) sa nedá prepovedať s istotou*." Pozri bližšie BAKER, Joanne. *50 physics ideas you really need to know*. London: Quercus Publishing Plc, 2007, 112 s. K *„the black box*" problematike pozri bližšie GUIDOTTI, R. MONREALE, A. PEDRESCHI. The AI black box Explanation Problem. *ERCIM*, 2019, alebo ZEDNIK, Carlos. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology*, Springer, [online]. 2019, [cit. 2023-06-15]. Dostupné z: https://link.springer.com/article/10.1007/s13347-019-00382-7

právneho systému, pričom jej charakteristickým znakom je spôsobilosť subjektu mať práva a povinnosti. Avšak už aj z histórie vieme, že dokonca aj pri fyzických osobách tieto práva a povinnosti nemusia byť rovnaké pre všetkých.[238] Samotná analógia pri tejto alternatíve vychádza z prípadov uznania právnej subjektivity chrámom (idolom) v Indii, alebo spomínaným riekam naprieč svetom. V týchto prípadoch by sa mohlo namietať, že udelená právna subjektivita je iba využitie akýchsi „zadných vrátok", ktoré majú slúžiť pre ochranu daných miest.[239] Ide zväčša o pozitívnu stránku právnej subjektivity (teda udelenie práv), avšak nie je cieľom „zaťažovať" rieky, aby boli zodpovedné za svoje „konanie" (stránka povinností). Tento koncept by však mohol byť istou inšpiráciou pre prípadnú subjektivitu UI. Aby sme boli presní, išlo by o obrátenú schému tohto konceptu. Právna subjektivita UI v tomto prípade by prichádzala iba vo forme povinností, bez priznaných práv. Išlo by o akési fiktívne „zadné dvierka", ktoré by smerovali najmä k odstráneniu nedostatkov pri tradičných konceptoch zodpovednosti.

Samozrejme, všetky tri možnosti majú svoje výhody a nevýhody. Preto na zodpovedanie otázky, či je možné a vôbec potrebné priznať právnu subjektivitu systémom umelej inteligencie, je potrebné vykonať hlbokú filozofickú a právnu analýzu dopadov takéhoto úkonu na samotné právo, ale aj našu spoločnosť. Pri analýze myšlienky uznávania systémov umelej inteligencie za subjekty práva je dôležité brať do úvahy a správne posudzovať princípy fungovania týchto systémov, ako aj ich vzťah k prostrediu, v ktorom pôsobia, berúc do úvahy spoločnosť, morálku, ekonomiku, politiku situáciu, historické rozdiely v etických názoroch, filozofické teórie, ako aj národné tradície.[240]

Vzhľadom na uvedené je prirodzené, že máme za to, že implementácia umelej inteligencie vyvoláva celý rad otázok. Pričom neadekvátna rýchlosť pri reakcii práva spôsobuje, že pri týchto zariadeniach súčasná legislatíva nedokáže efektívne reagovať na nové aplikačné problémy, ktoré sa čím ďalej, tým rýchlejšie objavujú. Pre lepšie znázornenie vyššie uvedeného považujeme za vhodné uviesť tri praktické prípady, kde tieto otázky boli už predmetom právnej praxe.

[238] Túto skutočnosť odráža aj boj za rovnaké práva žien, etnických alebo náboženských menšín.
[239] RODGERS, C. A New Approach to Protecting Ecosystems. *Environmental Law Review,* 2017, s. 266-279.
[240] ALLGROVE, B. Legal personality for artificial intellects: pragmatic solution or science fiction. *SSRN Electronic Journal,* [online]. 2004, [cit. 2023-06-15]. Dostupné z:: http://dx.doi.org/10.2139/ssrn.926015

4.3.1. Prípad udelenia štátnych občianstiev zariadeniam na báze UI

Prvé pokusy o praktické riešenia danej otázky môžeme vidieť prostredníctvom dvoch prípadov udelenia štátnych občianstiev nonhumánnym subjektom. V týchto prípadoch príslušné štáty uznali právnu subjektivitu týchto systémov prostredníctvom aktu udelenia štátnych občianstiev. Išlo o veľmi medializované prípady, ktoré však viacerí označovali skôr za „spoločenské udalosti" ako za premyslené právne akty.[241] Totižto z pohľadu práva boli tieto kroky v rozpore s viacerými inými právnymi predpismi daných krajín.

Prvý prípad je zo Saudskej Arábie, kde v roku 2017 bolo robotovi Sophia[242] udelené štátne občianstvo Saudskej Arábie. Druhý prípad nastal ešte v tom istom roku, keď bolo udelené štátne občianstvo Japonska robotovi Shibuya Mirai.[243] Spočiatku je potrebné uviesť, že otázka štátneho občianstva je riešená vždy vnútroštátnou právnou úpravou a zväčša prostredníctvom zákona je upravené aj nadobudnutie štátneho občianstva.[244] Právna úprava jednotlivých štátov v tejto oblasti nie je jednotná, spravidla však je možné štátne občianstvo nadobudnúť:

- Narodením[245] alebo
- Naturalizáciou.[246]

V našich prípadoch však môžeme vidieť značné deficity (ako z hmotnoprávnej, tak aj procesnoprávnej stránky). Keď sa pozrieme bližšie na vnútroštátnu právnu úpravu Saudskej Arábie, tak môžeme spozorovať, že štátne občianstvo v tejto krajine je možné získať nasledujúcimi spôsobmi:[247]

[241] Napríklad ATABEKOV, A. YASTREBOV, O. Legal Status of Artificial Intelligence Across Countries: Legislation on the Move. *European Research Studies Journal*, Volume XXI, Issue 4, 2018, s. 773-782.
[242] Sophiu postavila spoločnosť Hanson Robotics v roku 2015. Vynálezca David Hanson tvrdí, že robot je preniknutý umelou inteligenciou a dokáže rozpoznávať tváre. Silikónová tvár robota údajne môže napodobňovať výrazy tváre človeka. Pozri bližšie v Saudi Arabia Gives Citizenship, [online]. 2017, [cit. 2023-06-15]. Dostupné z: www.bloomberg.com/news/articles/ 2017-10-26/saudi-arabia-gives-citizenship-to-a-robot-claims-global-first
[243] CUTHBERTSON, A. Tokyo: Artificial Intelligence 'BOY' SHIBUYA MIRAI Becomes World's First AI Bot to Be Granted Residency. [online]. 2017, [cit. 2023-06-15]. Dostupné z: http://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya- mirai-702382
[244] OTTOVÁ, E. *Teória práva.* Šamorín: Heuréka, 2006, s. 90.
[245] Uplatňujú sa princípy *ius sanguinis* alebo *ius soli*.
[246] Najčastejšie uzatvorenia manželstva alebo udelením na základe žiadosti.
[247]Saudi Arabian Citizenship System, [online]. [cit. 2023-06-15]. Dostupné z: https://www.refworld.org/pdfid/3fb9eb6d2.pdf

1. *Narodením*:

- v tradičnej rodine, kde matka a otec sú občanmi Saudskej Arábie.
- v Saudskej Arábii, do rodiny, kde otec je občanom Saudskej Arábie a matka nie je občanom tejto krajiny.[248]
- matke, ktorá má saudskoarabské občianstvo, avšak otec ho nemá, za budúcej podmienky, že po dosiahnutí plnoletosti má dieťa trvalý pobyt na území Saudskej Arábie a hovorí plynule po arabsky.

2. *Manželstvom*;

3. *Naturalizáciou* (za niekoľkých podmienok)[249]

- dosiahnutie plnoletosti;
- znalosť jazyka (arabsky);
- pobyt na území Saudskej Arábie viac ako 10 rokov;
- legitímny spôsob zarábania;
- výpis z registra trestov;
- dodržiavanie noriem správania stanovených v danej krajine.

Po zvážení vyššie uvedených hmotnoprávnych podmienok môžeme konštatovať, že v prípade Sophie nebola naplnená ani jedna z vyššie uvedených alternatív. Keďže je zrejmé, že Sophia sa „nenarodila"[250] ako taká, preto sa zásady *ius soli* ani *ius sanguinis* aplikovať nebudú. Samozrejme, ani o manželstve hovoriť nemôžeme a neprichádza do úvahy ani naturalizácia, keďže viaceré z daných podmienok nie sú naplnené. Taktiež nemožno opomenúť ani vyžadovaný model správania žien v spoločnosti Saudskej Arábie, ktorý zavádza špecifické požiadavky na ženské aktivity, vrátane povinnosti cestovať v sprievode muža, obmedzenia pre umiestňovanie do zamestnania, na vycestovanie do zahraničia,

[248] Súčasne sa považuje za potrebné notársky overené uznanie otcovstva.
[249] Pričom podmienky musia byt naplnene kumulatívne.
[250] Zaujímavou otázkou môže byť možnosť stotožnenia pojmov „narodenie" a „stvorenie". Pri tejto otázke však veríme, že pri takej analógií by sme opäť sa dostali do sféry, kde by sme museli zvážiť množstvo ďalších otázok týkajúcich sa danej problematiky.

limitované práva v rodinnom či dedičskom práve, ako aj ďalšie limitácie vyplývajúce zo šaríje.[251]

Na obdobné hmotnoprávne problémy môžeme naraziť aj v prípade Japonska a robota Shibuya Mirai. Udelenie občianstva v Japonsku je upravené zákonom o občianstve Japonska.[252]

Zákon uvádza dve možnosti, ako získať japonské štátne občianstvo, a to narodením alebo naturalizáciou. V prípade, ak sa narodí dieťa, sa stáva občanom Japonska, keď:

- otec alebo matka sú japonskými občanmi;
- otec, ktorý zomrel pred narodením dieťaťa, bol občanom Japonska;
- sú obaja rodičia neznámi alebo nemajú žiadnu štátnu príslušnosť, a dieťa sa narodilo v Japonsku.

V prípade naturalizácie by žiadateľ o štátne občianstvo mal splniť nasledovné podmienky:

- musí mať trvalí pobyt v Japonsku aspoň päť rokov;
- musí dosiahnuť vek 20 rokov;[253]
- musí vedieť zabezpečiť si živobytie vlastným majetkom alebo pomocou manžela/manželky alebo iných príbuzných, s ktorými žije v jednej domácnosti;
- nemá žiadnu štátnu príslušnosť;[254]
- je potrebné vyhlásenie, že žiadateľ nikdy neplánoval, neobhajoval a ani nepatril k politickej strane alebo inej organizácii, ktorá sa usilovala o zvrhnutie Ústavy Japonska alebo jej vlády.

Po zvážení vyššie uvedených hmotnoprávnych podmienok môžeme konštatovať, že ani v prípade Shibuya Mirai nebola naplnená žiadna z vyššie uvedených alternatív. Opäť je

[251] Porovnaj s SHAHEEN, N. Saudi women defy ban to register for polls, Gulf News, [online]. 2011, [cit. 2023-06-15]. Dostupné z: https://gulfnews.com/news/gulf/saudi-arabia/saudi-women-defy-ban-to-register-for- polls-1.799161

[252] The Nationality Law (Law No.147 of 1950, as amended by Law No.268 of 1952, Law No.45 of 1984, Law No.89 of 1993 and Law.No.147 of 2004,Law No.88 of 2008) [online]. [cit. 2023-06-15]. Dostupné z: http://www.moj.go.jp/ENGLISH/information/tnl-01.html

[253] Pričom sa vyžaduje plná spôsobilosť na právne úkony.

[254] Alebo získanie štátneho občianstva Japonska bude mať za následok stratu inej štátnej príslušnosti.

očividné, že ani Shibuya Mirai sa „nenarodila" ako taká, preto ani tu sa nebudú aplikovať zásady *ius soli* ani *ius sanguinis*. Obdobne nemôžeme hovoriť ani o manželstve, keďže tak ako aj pri naturalizácií neboli splnené všetky spomenuté podmienky.

Ak sa pozrieme na oba spomenuté prípady z pohľadu procesnoprávnej stránky môžeme obdobne pozorovať viaceré nedostatky. Prvým problémom je hneď skutočnosť, že ani Shibuya Mirai ani Sophia nepožiadali o štátne občianstvo, pričom rozhodnutie *ex offo* udeliť niekomu štátne občianstvo, kto o to ani nepožiadal, je opäť viac ako problematické.[255]

Na základe uvedenej analyzi je očividné, že samotné udelenie štátneho občianstva robotom Shibuya Mirai a Sophia nebolo v súlade s vnútroštátnou právnou úpravou daných štátov, a to bez ohľadu na to, či mali priznanú právnu subjektivitu alebo nie. Uvedené len potvrdzuje naše presvedčenie, že išlo skôr o spoločenské udalosti, prípadne isté formy reklamy vyspelých techník oboch krajín, než právne akty majúce za cieľ zaväzovať a byť účinné v spoločnosti. Rok 2017 bol v tomto smere skutočne prelomovým, keďže bol poznačený týmito dvomi prípadmi udelených štátnych občianstiev nonhumánnym subjektom. Nemožno opomenúť spoločenský kontext tohto obdobia. V roku 2017 boli oba roboty veľké nóvum a existoval okolo nich veľký mediálny záujem. Dokonca tu môžeme pozorovať istú formu pretekov medzi oboma krajinami o to, kto bude premiantom na tejto scéne a kto ako prvý udelí štátne občianstvo stroju, a tým priblíži svoju krajinu bližšie k budúcnosti, ktorá je mnohými predpovedaná (že stroje UI budeme považovať za rovnocenné ľuďom a bude im priznaná právna subjektivita s patričnými právami i povinnosťami). Avšak, keď sa pozrieme na dané prípady skrz prizmu práva, môžeme vidieť značné nedostatky, ktoré nemohli a ani by nemali byť zanedbané. O to viac je potrebné oba prípady chápať skôr ako síce spoločensky významné udalosti, no z právnej stránky išlo iba o symbolické právne úkony.

[255] Porovnaj GORIS, I. HARRINGTON, J. KOHN, O. Statelessness: what it is and why it matters. *Forced migration review*, [online]. 2009, [cit. 2023-06-15]. Dostupné z: http://hr.law.vnu.edu.vn/sites/default/files/resources/what_is_statelessness.pdf

5. Zodpovednosť umelej inteligencie

5.1 Úvodné poznámky

Teória i prax sa prikláňa k názoru, že pod pojmom zodpovednosti budeme rozumieť následnú (sekundárnu) povinnosť vznikajúcu subjektu, ktorý porušil primárnu právnu povinnosť vyplývajúcu zo zákona alebo z inej právnej skutočnosti.[256]

Je však treba podotknúť, že zariadenia na báze umelej inteligencie fungujú oveľa komplexnejšie ako jednoduché obyčajné produkty, z čoho vyplýva, že bude potrebné zvážiť komplexnejšie scenáre zodpovednosti.[257] Odpovede v mnohých prípadoch budú veľmi komplikované a ešte nie je ani dostatok judikatúry, ktorá by pomohla pri riešení týchto otázok.

5.2 Zodpovednosť vo svetle umelej inteligencie

Aj pri najjednoduchšom zodpovednostnom scenári, ak sa inteligentný hriankovač prehreje a spáli dom, má majiteľ domu niekoľko potenciálnych kandidátov o ktorých môže tvrdiť, že sú zodpovední. Možné subjekty sa pohybujú od predajcu, výrobcu, programátorov až po vývojárov aplikácie hriankovači. Problematickou ostáva aj otázka či bude jedna strana plne zodpovedná alebo budú do určitej miery zodpovedné všetky strany zapojené do vytvárania a spracovania integrovaných dátových komponentov hriankovača?

Pri týchto otázkach sa zdá, že pri zodpovednostných scenároch inteligentných zariadení budeme musieť hľadať odpovede mimo tradičných hraníc zodpovednosti.[258] Ako príklad môžeme dopravnú situáciu, keď sa pokazia inteligentné dopravné semafory a inteligentné vozidlo idúce pod nimi je naprogramované sledovať tieto signály, čo znamená, že vozidlo nedokáže zastaviť a spôsobí dopravnú nehodu. V takýchto prípadoch by sa zdalo neprimerané postihovať vodiča za spôsobenie škody, keď „vinníkom" bol čiastočne

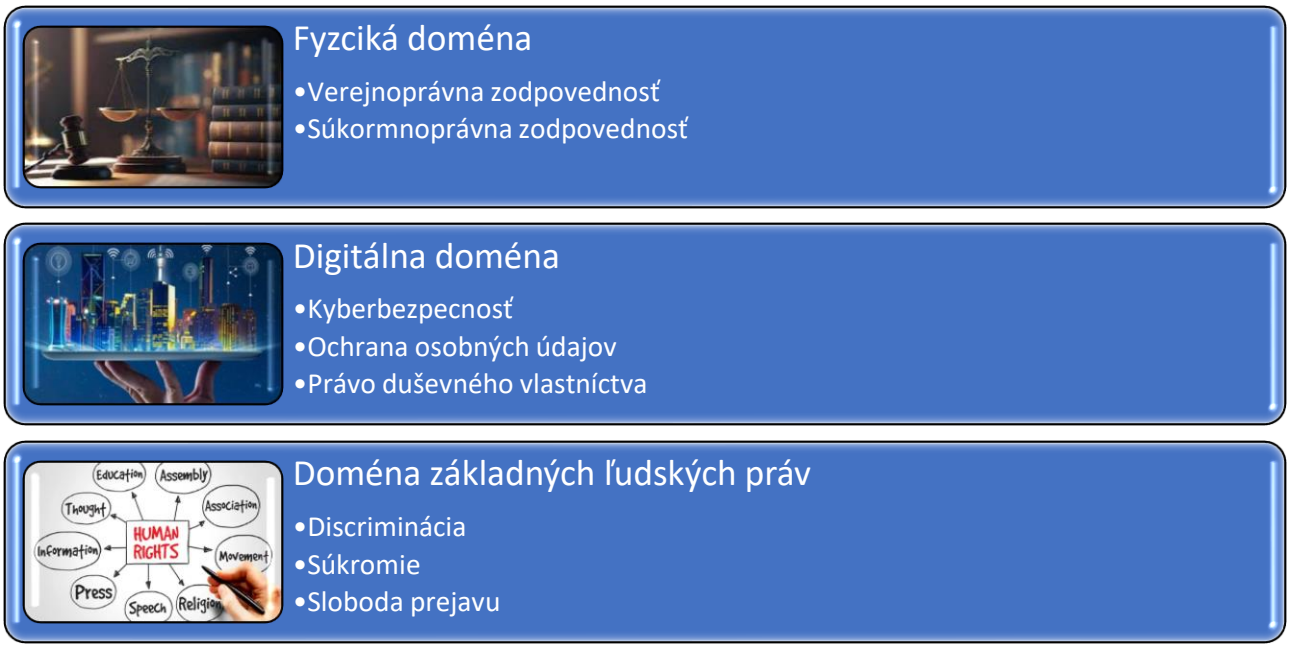
[256] ŠTEVČEK, M. a kol. Občiansky zákonník, Komentár. 2. vydanie. Praha: C.H.Beck, 2015.
[257] ROSE, ELDRIDGE, CHAPIN: The Internet of Things: an Overview, 2015 Dostupné na: https://www.researchgate.net/profile/Ananth_Saradhi2/post/What_are_the_latest_developments_in_IOT_architectures/attachment/59d6387d79197b8077995b29/AS:397885586853893@1471874724774/download/57.pdf
[258] Porovnaj GYURÁSZ, Z. Modality regulácie nových technológií. 2021. Dostupné na: <https://www.researchgate.net/publication/361440930_Modality_regulacie_novych_technologii>. cit. 2023-04-25.

nefunkčný semafor a čiastočne aj nefunkčné automatizované vozidlo.[259] Na druhej strane nemožno vnímať takéto zlyhanie zariadení ako zodpovednosť iba na strane výrobcov a vývojárov, keďže argumentom ostáva aj skutočnosť, že inteligentne zariadenia sa používajú a budú sa používať spôsobom a v situáciách, ktoré výrobca nemohol ani očakávať ani predpovedať.

Ako aj z vyššie uvedeného prípadu je zrejmé, otázky zodpovednosti umelej inteligencie sú viac ako len „komplikované" nakoľko spochybňujú naše tradičné vnímanie jednotlivých právnych oblastí. Nastáva aj otázka, v ktorej doméne vôbec budeme hľadať naše odpovede? Pôjde o občianskoprávnu zodpovednosť? Dokáže trestné právo adekvátne chrániť objekty ochrany trestného práva? Môžeme vôbec aplikovať normy administratívneho práva v kontexte umelej inteligencie? Tieto otázky skúsime zodpovedať v tejto kapitole.

Nápomocné by nám mohlo byť aj znázornenie delenia týchto vyššie uvedených otázok.

[259] Túto skutočnosť si uvedomujú aj vrcholové orgány Európskej únie. Európsky parlament publikoval materiál nazvaný ako Spoločný prístup EÚ k pravidlám zodpovednosti a poisteniu pripojených a autonómnych vozidiel. V ktorom za štyri hlavné riziká spojené s prevádzkou autonómnych vozidiel označili: a)Zlyhanie softwaru, b)Zlyhanie sieťovej infraštruktúry/pripojenia, c)Kybernetická kriminalita a d)Riziká vyplývajúce z rozhodnutí robených pri programovaní. Pozri bližšie European Parliamentary Research Service. A common EU approach to liability rules and insurance for connected and autonomous vehicles, s. 20 cit. Dostupné na: <https://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635_EN.pdf>. cit. 2023-10-25.



Obrázok 1:[260] *Vzťah UI a jednotlivých domén zodpovednosti*

Zdroj: Vlastná tvorba autorov.

V tejto kapitole sa budeme venovať najmä otázkam fyzickej domény s ohľadom na to ako dokážu aktuálne platné a účinné pramene práva zvládať jednotlivé zodpovednostné scenáre, taktiež sa dostaneme aj k otázkam domény základných ľudských práv. K otázkam digitálnej domény sú venované samostatné kapitoly tejto učebnice.

5.3 Verejnoprávna zodpovednosť

Z prehľadu odbornej vedeckej literatúry[261], ako aj práva Európskej únie,[262] ktoré sa tejto téme venujú, možno konštatovať, že sa zaoberajú predovšetkým civilnoprávnym aspektom zodpovednosti. Otázky verejnoprávnej zodpovednosti nie sú rozpracované do

[260] Výpočet jednotlivých podkategórií je demonštratívny a nie explicitný

[261] ŽELECHOWSKI, L. Občianskoprávna zodpovednosť za škody spôsobené autonómnymi vozidlami: poľský Perperspektívne. In. Skupień, D., Lewaszkiewicz-Petrykowska, B. (eds.) Rapports Polonais. XXIe kongres International de Droit Comparé. XXI. medzinárodný kongres porovnávacieho práva. Asunción 23 – 28 X 2022. Lodž : Lodz University Press, 2022; LIIVAK, T., LAHE, J. Deliktná zodpovednosť za Škody spôsobené plneautonómnymi vozidlami: estónsky pohľad. In. Masarykova university Journal of Law and Technology, Vol. 12, 2018, č. 1, s. 49 – 74; Srinivasamurthy, M. Autonómne vozidlá a komplikácie pri rozdeľovaní zodpovednosti. In. Jus Corpus Law Journal, Vol. 1, 2021, č. 4, s. 360 – 370. Funkhouser, K. Paving the Road Ahead: Autonomous Vehicles Zodpovednosť za produkty a potreba nového prístupu. In. Utah Law Review, 2013, č. 1, s. 437 – 462.

[262] Pozri uznesenie Európskeho parlamentu z 20. októbra 2020 s odporúčaniami pre systém občianskoprávnej zodpovednosti za umelú inteligenciu.

takých detailov.[263] Uvedené v súčasnosti nepovažujeme za problematické, pretože otázky zodpovednosti treba ponímať komplexne a pri vyvodzovaní zodpovednosti možno povedať, že verejnoprávna a súkromnoprávna zodpovednosť sa vzájomne, resp. k ich využitiu dochádza vtedy a keď nemožno využiť druhú zo zodpovednosti, lebo je z povahy veci vylúčená.

Pri pohľade na verejnoprávnu zodpovednosť sa treba v prvom rade vysporiadať s otázkou, či pri umelom inteligentných systémoch budeme môcť využívať pravidlá trestnoprávnej zodpovednosti alebo administratívnoprávnej zodpovednosti.

Z pohľadu trestného práva sa sústredíme na otázky umelej inteligencie ako „páchateľa" trestnej činnosti, „nástroj" trestnej činnosti, a ako „zdroj informácii" o trestnej činnosti. Z pohľadu správneho práva sa sústredíme na otázky správnych deliktov ako aj na zodpovednosť za škodu spôsobenú pri výkone verejnej správy.

5.3.1. Trestnoprávna zodpovednosť

Z hľadiska pohľadu trestnoprávnej zodpovednosti sa na systémy na báze umelej inteligencie dá nazerať zo štyroch možných pozícií:

I. Ako páchateľ trestnej činnosti
II. Ako nástroj trestnej činnosti
III. Ako zdroj informácii o trestnej činnosti
IV. Ako objekt ochrany

Aj keď samozrejme inteligentný otvárač na konzervy môže ušetriť čas v porovnaní so svojím manuálnym bratrancom, nevyžaduje žiadnu spoločenskú alebo právnu zmenu v tom zmysle, že právny rámec už nie je dostatočný alebo vhodný len preto, že existuje táto nová technológia.[264] Dôraz sa bude klásť na zariadenia, ktoré posúvajú nás k zmenám už svojou

[263] Napríklad pre trestné právo CSITEI, B. Samoriadiace autá a trestná zodpovednosť. In. Debreceni Jogi Műhely, sv. XVII, 2020, č. 3 – 4, s. 34 – 46; VOJTUŠ, F., KORDÍK, M., DRAŽOVÁ, P. In. Andraško, J. a kol. Právne a technické aspekty kybernetickej bezpečnosti automatizovaných vozidiel. Bratislava : Wolters Kluwer, 2022, s. 102 – 147.
[264] MOSES, L: Recurring Dilemmas: The Law's Race to Keep Up with Technological Change.. University of Illinois Journal of Law, Technology & Policy, 2007 Dostupné na: <https://www.researchgate.net/publication/228183058_Recurring_Dilemmas_The_Law%27s_Race_to_Keep_Up_With_Technological_Change>. cit. 2023-10-25.

existenciou. Na technologické pokroky, ktoré zmenia, to, čo je skutočne možné a nie iba bežné zmeny. V týchto prípadoch z hľadiska riešenia zodpovednostných scenárov sa navrhujú tri alternatívy:[265]

I. Koncept priamej zodpovednosti
II. Koncept nepriameho páchateľa
III. Koncept zavinenia z nedbanlivosti

Musíme však podotknúť, že ako na začiatku sme už spomínali doména trestného práva nie je ideálna na riešenie otázok zodpovednosti UI a preto aj každý jeden z konceptov bude mať svoje nedostatky.

Prvý koncept je založený na myšlienke priznania práv alebo právnej subjektivity tejto nehumánnej entity. Tento koncept aj keď by bol najefektívnejším zo všetkých konceptov, má však nedostatky kvôli ktorému sa zdá byť neaplikovateľná, aspoň nateraz. Zásadným problémom je, že tento koncept v kontinentálnom právnom prostredí nemá oporu v existujúcich a v praxi aplikovaných konceptoch trestnej zodpovednosti, pričom kontinentálna európska právna kultúra dlhodobo vychádza z konceptu individuálnej trestnej zodpovednosti fyzickej osoby.[266] V súčasnosti sú v slovenskom právnom prostredí, ako aj v okolitých krajinách využívané dva základné prístupy k trestnej zodpovednosti odlišujúce sa v tom, kto je zodpovedným subjektom trestnej zodpovednosti a ako je táto zodpovednosť vyvodzovaná (priamo, respektíve odvodene – nepriamo). Oba tieto prístupy však vychádzajú z princípu zodpovednosti za zavinenie, ktorý je kľúčovým konceptom trestnej zodpovednosti v kontinentálnom právnom systéme.[267] Ide o konkrétne koncepty: a)individuálnej trestnej zodpovednosti fyzickej osoby; b)odvodenej trestnej zodpovednosti právnickej osoby. Aj v prípade odvodenej trestnej zodpovednosti právnickej osoby jej základ vychádza z pričítateľnosti protiprávneho konania konkrétnej fyzickej osoby, ktorá je v určitom vzťahu k právnickej osobe.[268] Preto je otázkou, či by pričítanie trestnej zodpovednosti samotnému

[265] HALLEVY, G. The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control. Akron Intellectual Property Journal. 2010, Vol. 4, Iss. 2, Article 1, s. 171–172. Dostupné na: <https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1>
[266] Pozri bližšie. MENCEROVÁ, I. TOBIAŠOVÁ L. TURAYOVÁ Y. A kol. Trestné právo homtné. Všeobecná časť. 2. aktualizované a prepracované vydanie. Šamorín : Heuréka 2015, 501. s
[267] ASHWORTH, A. Principles of Criminal Law. third edition. New York: Oxford University Press, 1999, s. 87–95, 160–162.
[268] TURAYOVÁ, Y. TOBIÁŠOVÁ, L. et al. Trestná zodpovednosť právnických osôb: medzinárodné a európske aspekty, právno-teoretické východiská, s. 49–138.

konaniu umelej inteligencie nebolo popretím základných princípov a zásad kontinentálneho trestného práva.

Druhý koncept je založený na prípadoch, keď je využitá alebo zneužitá ako nástroj na páchanie trestnej činnosti. V týchto prípadoch môžeme použiť zaužívaný koncept z trestného práva. Treba však podotknúť, že v prípade umelej inteligencie pri tomto modeli sa javí skôr „neživým nástrojom" než entitou, ktorá bola zneužitá, ako je tomu v prípade konceptu nepriameho páchateľa. V tomto prípade umelej inteligencie vykonávajúcej vôľu tretej osoby, ale nie preto, že by bola uvedená do omylu alebo nechala zmysel svojho konania (ako v prípade fyzickej osoby), ale preto, že bola buď priamo naprogramovaná alebo preprogramovaná na spáchanie trestného činu alebo bola využitá ako „sofistikovanejší" nástroj, ktorý len vykonáva svoju úlohu pre samotného páchateľa trestného činu. V tomto prípade môžeme byť schopní riešiť otázku trestnej zodpovednosti osoby, ktorá umelú inteligenciu použila na spáchanie trestného činu.

Tretí koncept vychádza z predpokladu možností predvídateľnosti spáchania trestného činu konkrétnej umelej inteligencie. Ide o aplikáciu analógie konceptu *respondeat superior* vychádzajúcemu z rímskeho práva, podľa ktorého bol majiteľ otroka zodpovedný za škodu týmto otrokom spôsobenú.[269] Z hľadiska vyvodenia trestnoprávnej zodpovednosti model vychádza z nedbanlivostnej formy zavinenia, teda predpokladu, že tvorca alebo užívateľ umelej inteligencie nemali úmysel spáchať prostredníctvom umelej inteligencie trestný čin a o páchaní trestného činu nemali vedomosť až kým nebol spáchaný. Ich trestná zodpovednosť je založená na predpoklade, že za bežných okolností mohli vedomosť, že spáchanie trestného činu v danej situácii je možným výsledkom umelej inteligencie, ktorú vytvorili, respektíve v súvislosti s prípadom potrebné mať. Tieto osoby boli zodpovedné za trestné činy, ktoré je možné spáchať z nedbanlivosti a ktorých objektívna stránka je predstaveným konaním umelej inteligencie. Pri posudzovaní a vyvodzovaní trestnej zodpovednosti, ako z hľadiska skúmania príčinnej súvislosti medzi konaním a následkom, tak aj z hľadiska formy zavinenia, bude pri tomto modeli dôležité skúmanie, či tvorcom alebo užívateľom systému umelej inteligencie bola zákonom, vyhláškou, prípadne iným podzákonným právnym predpisom,

[269] KING, C. T. AGGARWAL, N. TADDEO, M. FLORIDI, L. Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions, s. 110

prípadne nejakým interným právnym predpisom uložená určitá povinnosť správania alebo postupu, kde pri jeho správnom konaní je zasiahnuté škodlivým následkom využitia umelej inteligencie dalo predísť.[270] Prezentovaný druhý model trestnej zodpovednosti založený na možnosti a schopnosti predvídať v súvislosti so správaním umelej inteligencie. Avšak aj tento koncept má tiež svoje limity. Problém môže nastať ak nejaký systém umelej inteligencie vykoná kroky, ktoré neočakával ani jeho užívateľ a ani jeho tvorca. Podľa Hallevyho sú dve možnosti, a to buď hodnotiť umelú inteligenciu ako systém, ktorý nemá trestnú zodpovednosť a za jej konanie zodpovedá osoby ako podľa druhého modelu, alebo umelú inteligenciu považujem takú za priamo trestne zodpovednú, čo reflektuje koncept priamej trestnej zodpovednosti.[271]

5.3.2 Administratívnoprávna zodpovednosť

Ak dôjde k porušeniu administratívnoprávnej normy, budeme môcť hovoriť o administratívnoprávnej alebo správnej zodpovednosti. Správne trestanie je súčasťou verejnoprávneho systému vynucovania porušených povinností, a to spolu s právom trestným. Tu sa začína jednoznačná podobnosť medzi dvomi právnymi odvetviami, ktorú odzrkadľuje právna prax v podobe súdnej judikatúry. Na národnej úrovni, ako aj medzinárodnej platí, že pod pojmom obvinenie možno pochopiť obvinenie zo spáchania trestného činu, ako aj obvinenie zo spáchania správneho deliktu.[272]

Špecifikom administratívnoprávnej zodpovednosti je, že ju musíme členiť podľa toho, či normu porušil spravovaný subjekt[273] alebo spravujúci subjekt.[274]

---

[270] SOKOL, T. SMEJKAL, V. Trestněprávní aspekty robotiky, s. 534.

[271] HALLEVY, G. The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control, ref. 2, s. 185–186.

[272] Napríklad rozsudok Najvyššieho súdu SR, sp. zn. 10Asan/3/2016 z 25. 1. 2017: „Taktiež je potrebné uviesť, že záver o nevyhnutnosti úplnej špecifikácie príslušného priestupku z hľadiska vecného, časového a miestneho plne korešponduje i s medzinárodným záväzkom Slovenskej republiky, pretože i na rozhodovanie o priestupkoch je potrebné aplikovať požiadavky článku 6 ods. 1 Dohovoru o ľudských právach a základných slobodách. Z článku 6 ods. 1 Dohovoru vyplýva, že veci priestupkové a taktiež veci správnych deliktov je potrebné považovať v zmysle ustálenej štrasburskej judikatúry za veci, ktoré majú charakter konania o trestnom obvinení. Vo všetkých veciach, ktoré je možné subsumovať pod pojem ‚veci trestného charakteru' musí mať osoba, proti ktorej sa vedie konanie možnosť domôcť sa práva na spravodlivý proces v zmysle článku 6 Dohovoru. Správne delikty sú svojou povahou najbližšie práve priestupkom. V oboch prípadoch ide o súčasť tzv. správneho trestania, o postih správnym orgánom za určité nedovolené konanie (či opomenutie)."

[273] súkromná fyzická osoba alebo právnická osoba

[274] t.j. vykonávateľ verejnej správy. K pojmu „vykonávateľ verejnej správy" pozri napríklad HORVAT, M. In. Cepek, B. a kol. Správne právo hmotné. Všeobecná časť. Bratislava : Wolters Kluwer, 2018, s. 146 a nasl.

Ak poruší administratívnoprávnu normu **spravovaný subjekt,** vzniká jej administratívnoprávna zodpovednosť za správny delikt. Právna teória člení správne delikty na priestupky, iné správne delikty fyzických osôb, správne delikty právnických osôb a fyzických osôb podnikateľov, správne disciplinárne delikty a správne poriadkové delikty.[275]

Ak poruší administratívnoprávnu normu **vykonávateľ verejnej správy**, vzniká zodpovednosť nositeľa verejnej správy za škodu spôsobenú nezákonným rozhodnutím alebo nesprávnym úradným postupom orgánu verejnej správy.[276] Dôležitosť tohto členenia spočíva predovšetkým v tom, že zatiaľ čo v prvom prípade hovoríme o zodpovednosti v oblasti verejného práva, v druhom prípade ide o zodpovednosť súkromnoprávnu. [277]

Vzhľadom na skutočnosť, že moderné technológie sú súborom mnohých technických zariadení, netreba zabúdať na skutočnosť, že zodpovednosť ako takú nemusí alebo nemôže niesť len vlastný výrobca, ale aj dodávatelia jednotlivých technológií, ktoré mohli vykázať chybu a na jej základe došlo k porušeniu právnych predpisov. V tejto súvislosti sa v teórii hovorí o tzv. spoločnej zodpovednosti.[278] Podstata spočíva v spoločnej a nerozdielnej zodpovednosti a to nielen zo strany výrobcu, ale aj výrobcov jednotlivých súčiastok, kde sú technologicky zahŕňajú súčiastky, automatizované systémy ktoré riadia a počítače, ktoré prijímajú rozhodnutia ktoré nemusí vyrábať len výrobca inteligentného systému.[279] Bolo by preto pomerne prísne, ak by k tomu došlo, pretože z hľadiska subjektu môže byť zodpovedným subjektom len výrobca a nie ostatní „dodávatelia".

Ako aj vyššie spomíname, ak poruší administratívnoprávnu normu vykonávateľ verejnej správy, vzniká zodpovednosť nositeľa verejnej správy za škodu spôsobenú nezákonným rozhodnutím alebo nesprávnym úradným postupom orgánu verejnej správy,

[275] MACHAJOVÁ, J. In. Machajová, J. a kol. Všeobecné správne právo. Žilina : Eurokódex, 2009, s. 198.
[276] KOŠIČIAROVÁ, S. Správne právo hmotné. Všeobecná časť. Plzeň : Aleš Čeněk, 2022, s. 263.
[277] Dôležitosť spočíva v tom, že postup pri vyvodzovaní zodpovednosti je rozličný a postupuje sa na základe odlišných princípov, funkcií, postupov, ale aj právnych predpisov. K členeniu zodpovednosti na verejnoprávnu a súkromnoprávnu a ich princípom, či funkciám pozri napríklad HORVAT, M. In. Andraško, J., Horvat, M., Mesarčík, M. Vybrané kapitoly práva informačných technológií II. Bratislava : Univerzita Komenského v Bratislave, Právnická fakulta, 2021, s. 73 – 80.
[278] Pri spoločnej zodpovednosti sa nazerá na výrobcu v širšom zmysle slova, keď nebude mať na mysli len ten subjekt, ktorý inteligentný systém pod svojou „značkou" vyrába, ale teda aj výrobcovia jednotlivých komponentov, ktoré sú len na to , aby systém plnilo tie úlohy a funkcie, ktoré má. Pod týmito výrobcami komponentov máme tiež dodávateľov a prevádzkovateľov systémov, ktoré sa v inteligentnom systéme využívajú. Porovnaj HORVAT, M. Administratívnoprávna zodpovednosť. Právne aspekty automatizovaných vozidiel. - : 1. vyd. ISBN 978-80-8232-038-4. - Bratislava : C.H. Beck, 2023. - S. 206-240.s
[279] VLADECK, D. C. Machines Without Principals: Rules and Artificial Intelligence. In. Washington Law Review, Vol. 89, 2014, No. 1, p. 148.

a práve tieto zodpovednostné scenáre nás budú zaujímať z pohľadu zodpovednosti v oblasti verejného práva.

5.3.2.1 Zodpovednosť za škodu spôsobenú pri výkone verejnej moci

Otázka, či by štát mohol niesť zodpovednosť za škodu, ktorá bola spôsobená prevádzkou inteligentného systému je v skutku zaujímavá. Podľa nášho názoru tu vzniká oprávnená otázka, či na vyvodenie zodpovednosti za škodu spôsobenú prevádzkou inteligentných systémov, ktorá bola spôsobená zanedbaním povinností štátu, možno priniesť zákon č. 514/2003 Z. z. a či je jeho aplikácia vhodná?

V podmienkach Slovenskej republiky je zodpovednosť za spôsobenú škodu najmä predmetom vecnej pôsobnosti zákona č. 514/2003 Z. z. o zodpovednosti za škodu spôsobenú pri výkone verejnej moci a o zmene niektorých zákonov v znení neskorších zákonov (ďalej len „zákon č. 514/2003 Z. z.") a zákona č. 40/1964 Zb. Občiansky zákonník v znení neskorších predpisov (ďalej len „Občiansky zákonník").

Zákon č. 514/2003 Z. z. vychádza z ústavnoprávnej úpravy, konkrétne čl. 46 ods. 3 Ústavy SR,[280] podľa ktorého má každé právo na náhradu škody spôsobenej nezákonným rozhodnutím súdu, iného orgánu či orgánu verejnej správy alebo nesprávneho úradného postupu. Podľa zákona č. 514/2003 Z. z. štát zodpovedá za škodu spôsobenú orgánmi verejnej moci pri výkone verejnej moci, pričom na účely tohto zákona výkonu verejnej moci je rozhodovanie a úradný postup orgánov verejnej moci o právach, právom chránených záujmoch a povinnostiach fyzických osôb alebo právnických osôb.[281]

Treba však povedať, že nejde o komplexné poňatie verejnej moci. Toto vymedzenie pojmu verejná moc je zamerané iba na tie prípady, keď orgán verejnej moci rozhoduje o individuálnych právach a povinnostiach v právno-aplikačnom procese. Obchádza normotvorbu, riadiacu a organizátorskú činnosť, postup pri uzatváraní verejnoprávnych zmlúv, uskutočňovanie rôznych bezprostredných zákrokov a iných opatrení obdobnej

[280] SVÁK, J. In. OROSZ, L., SVÁK, J. a kol. Ústava Slovenskej republiky. Komentár. 1. zväzok (základné princípy a ľudské práva). Bratislava : Wolters Kluwer, 2021, s. 623 – 625.
[281] § 1 a 2 zákona č. 514/2003 Z. z.

povahy.[282] Orgány verejnej moci teda nemusia vždy vystupovať len vo verejnomocenskom postavení. Súdna judikatúra uvádza, že pri skúmaní, či orgán verejnej moci v tom, ktorý v prípade vykonávania verejnej moci, je významné, či koná vo verejnej sfére a s verejnými dôsledkami, vo vzťahu k iným subjektom. Naopak, o výkon verejnej moci nejde, ak určitý subjekt[283] v konkrétnom prípade neplní svoje úlohy ako orgán verejnej moci. O výkon verejnej moci nejde a štát podľa ustanovení zákona č. 514/2003 Z. z. nezodpovedá za činnosť takého orgánu.

Podľa zákona č. 514/2003 Z. z. zodpovedá štát za škodu, ktorú spôsobilým rozhodnutím; nezákonným zatknutím, zadržaním alebo iným pozbavením osobnej slobody; rozhodnutie o treste, o ochrannom opatrení alebo rozhodnutie o väzbe; alebo nesprávnym úradným spôsobom.[284] Z vyššie uvedeného vyplýva, že budeme musieť rozdeľovať medzi prípadmi kedy štát priamo využíva inteligentné zariadenia pri výkone verejnej moci a kedy len „dovoľuje" využívanie týchto systémov.

Máme za to, že v **prvom** prípade, ak štát využíva systémy umelej inteligencie za účelom efektivizácie výkonu verejnej moci,[285] pričom v týchto vzťahoch stále vystupuje v nadriadenej pozícií, bude aj zodpovedať za škodu, ktorá bola spôsobená prevádzkou inteligentného systému podľa zákona č. 514/2003.

V **druhom** prípade je však už táto otázka menej jasná. Ako príklad môžeme zobrať automatizované vozidlá, kde štát nebude subjekt ktorý zavádza využívanie týchto vozidiel ale umožňuje, aby sa na jeho pozemných komunikáciách štátu pohybovali tieto vozidlá a navyše niektoré dopravné nehody môžu nastať v dôsledku zanedbania povinností práva, ako napríklad nevhodné dopravné značenie. Treba však akcentovať, že orgány verejnej moci môžu mať aj verejné služby. V rámci užšieho vymedzenia verejných služieb, ktoré sa v literatúre preferuje, sa verejnými službami realizuje sociálna funkcia štátu.[286] Osadenie a použitie dopravných značiek možno považovať za poskytovanie určitej služby na jednej

[282] ŠKROBÁK, J. Kategória verejnej moci a jej význam vo verejnej správe a vo vede správneho práva. In. Míľniky práva v stredoeurópskom priestore 2008. Bratislava : VO PraF UK, 2008, s. 651 a nasledujúce.
[283] Vykazujúci inak znaky orgánu verejnej moci
[284] Pozri § 3 zákona č. 514/2003 Z. z. Na naše účely nemožno aplikovať prípady nezákonného zatknutia, zadržania alebo iného pozbavenia osobnej slobody, ako ani rozhodnutia o treste, o ochrannom opatrení alebo rozhodnutia o väzbe.
[285] Porovnaj GYURÁSZ, Z.: Umelá inteligencia a digitálna verejná správa. 2019. Dostupné na: https://www.flaw.uniba.sk/fileadmin/praf/Veda/Publikacne_vystupy/2019/Monografia_DVSEI_2019.pdf
[286] HORVAT, M. In. Cepek, B. a kol. Správne právo hmotné. Všeobecná časť. Bratislava : Wolters Kluwer, 2018, s. 76.

strane, ale na druhej strane aj za určitý aplikačný rozhodovací proces ktorý je správnym konaním, ktorý je ukončený vydaním rozhodnutia, ktoré má podobu osadenia dopravnej značky.[287] Náhrada škody z dôvodu vydania nezákonného rozhodnutia však neprichádza do úvahy, lebo zákon v § 5 ods. 1 predpokladá, že škoda sa stala konkrétnemu účastníkovi konania v konkrétnom správnom konaní. V našom prípade však takýto subjekt v konaní nejestvuje, pretože ako sme uviedli, vo svojej podstate tu ide o rozhodnutie so všeobecnými účinkami, ktoré nezaväzuje konkrétne vymedzených účastníkov konania menovite uvedených v rozhodnutí vo veci samej, k čomu smeruje aj § 5 zákona č. 514/2003 Z. z., kde sa vymedzujú základné predpoklady uplatnenia náhrady škody.

5.4 Právny rámec súkromnoprávnej zodpovednosti

Dnes neexistuje ucelený právny rámec, ktorý by osobitne upravoval súkromnoprávnu zodpovednosť inteligentných systémov. Právny režim súkromnoprávnej zodpovednosti sa opiera o viaceré právne predpisy. Jednak to bude náš občiansky zákonník[288] č. 40/1964 Zb. a za druhé smernica Rady 85/374/EC z 25. júla 1985 o aproximácii zákonov, iných právnych predpisov a správnych opatrení členských štátov o zodpovednosti za chybné výrobky(*Product Liability Directive* – ďalej ako *Smernica PLD*).[289] Nemôžeme zabúdať však ani na otázky súkromnoprávnej zodpovednosti z oblasti administratívnoprávnej sféry a nakoniec nemožno opomenúť ani otázky poistných udalosti vyplývajúcich zo zodpovednostných scenárov niektorých špecifických inteligentných systémov. Preto dôležitú rolu bude hrať aj smernica Európskeho parlamentu a Rady 2021/2118/ES z 24. novembra 2021 ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel

[287] Osadenie dopravnej značky namiesto vydania rozhodnutia. Dopravná značka potom pôsobí ako rozhodnutie so všeobecnými účinkami, a zaväzuje bližšie nešpecifikovaný okruh osôb, ale zaväzuje tie osoby, ktoré prídu do styku s touto značkou. Pozri § 47 ods. 7 SpP.

[288] Porovnaj ŠTEVČEK, M. a kol. Občiansky zákonník, Komentár. 2. vydanie. Praha: C.H.Beck, 2015.Najmä nás budú zaujímať ustanovenia §499 a nasledujúce.

[289] Smernica Rady 85/374/EC z 25. júla 1985 o aproximácii zákonov, iných právnych predpisov a správnych opatrení členských štátov o zodpovednosti za chybné výrobky Dostupné na: https://eur-lex.europa.eu/legal-content/SK/TXT/HTML/?uri=CELEX:31985L0374&from=EN. cit. 2023-10-25. Táto smernica bola implementovaná do slovenského právneho poriadku zákonom č. 294/1999 Z. z. o zodpovednosti za škodu spôsobenú vadným výrobkom. Dostupné na: <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/1999/294/20040901>.

a o kontrole plnenia povinnosti poistenia takejto zodpovednosti (Motor Vehicle Insurance Directive – ďalej ako Smernica MVID)[290]

5.4.1. Zodpovednosť pri vade výrobku a zodpovednosť za škodu spôsobenú vadným výrobkom

Súčasný zodpovednostný právny rámec vychádza z predpokladu existencie dvoch základných rizikových faktorov v dôsledku zlyhania pri ktorých môže vzniknúť škoda.

Ako učebnicový príklad môžeme uviesť, že ak si sadneme do nášho automatizovaného vozidla (inteligentný systém) a ak auto nebude chcieť naštartovať a začať jazdu, tak budeme hovoriť a otázke zodpovednosti za vady, avšak prípade ak systém na báze umelej inteligencie, ktoré robí naše vozidlo automatizovaným zle vyhodnotí situáciu v doprave alebo spraví inak zlé rozhodnutie a tým spôsobí dopravnú nehodu, tak budeme hovoriť o zodpovednosti za škodu spôsobenú vadným výrobkom. Pri prevádzke automatizovaného vozidla sa prítomnosť človeka napokon vôbec nevyžaduje a nemožno vylúčiť, že škodová udalosť sa uskutoční v neprítomnosti človeka.

Veríme, že niet pochýb o tom, že pre nás budú viac zaujímavé prípady zodpovednosti za škodu spôsobenú vadným výrobkom, keďže presne v týchto prípadoch môžeme vidieť, ako dokážu nové technológie spochybňovať ustanovenia exitujúcich právnych predpisov, a spôsobujú tým neistotu, pokiaľ ide o uplatňovanie týchto právnych predpisov.[291]

Už z prvého pohľadu nám musí byť jasné, že náš občiansky zákonník z roku 1964 ani smernica z roku 1985 pri koncipovaní nepočítali stým, že budú musieť riešiť zodpovednostné scenáre inteligentných systémov. Preto len s ťažkosťami sa budeme môcť oprieť o sémantické výklady noriem daných predpisov.[292]

---

[290] Smernica z roku 2009 bola implementovaná do zákona č. 381/2001 Z. z. o povinnom zmluvnom poistení zodpovednosti za škodu spôsobenú prevádzkou motorového vozidla a o zmene a doplnení niektorých zákonov. Dostupné na: < https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2001/381/>. cit. 2023-10-25.

[291] Pre otázku právnej neistoty v kontexte nových technológií pozri bližšie. Gyurász, Z. Modality regulácie nových technológií. 2021. Dostupné na: https://www.researchgate.net/publication/361440930_Modality_regulacie_novych_technologii>. cit. 2023-10-25.

[292] K tejto otázke pozri bližšie. GAHÉR, F. MRVA, M. ŠTEVČEK, M. TURČAN, M. Open Texture of Concepts and Rules – Fuel for the Fire of Subjectivism in Applied Semantics? FILOZOFIA, 75, 2020, No 4, s.309 – 323 Dostupné na: <https://www.sav.sk/journals/uploads/04211152ilozofia.2020.75.4.4.pdf>. cit. 2023-10-25.

Pri otázke zodpovednosti za škodu spôsobenú vadným výrobkom, smernica PLD vytvorila jednotný európsky právny rámec upravujúci zodpovednosť výrobcu za škodu spôsobenú vadným výrobkom. Účelom ktorého je poskytnúť ochranu používateľom a spotrebiteľom, ktorý si kúpili vadný[293] výrobok, a smeruje k spravodlivému rozdeleniu rizík medzi subjektmi.[294]

Podľa čl. 1 smernice PLD je za škodu spôsobenú vadou svojho výrobku zodpovedný výrobca. Vychádzajúc z takejto premise by sme sa mali pýtať na dve otázky, za prvé kedy je výrobok vadným a za druhé či vôbec budeme môcť považovať inteligentné systémy za výrobok?

Jednoduchšie zodpovedateľná sa zdá byť naša **prvá otázka**.

Kedy bude automatizované vozidlo vadným? Podľa čl. 6 smernice PLD je výrobok vadný v prípade, ak nezabezpečuje bezpečné používanie, ktoré jednotlivec od výrobku právom očakáva s ohľadom na všetky okolnosti, vrátane:

a) predvádzania výrobku;
b) použitia výrobku na také účely, na ktoré sa logicky predpokladá, že sa použije;
c) časového obdobia, v priebehu ktorého sa výrobok uviedol do obehu;

S prihliadnutím na legálne vymedzenie pojmu vadný výrobok je kľúčové z hľadiska zodpovednosti výrobcu potrebné skúmať dve skutočnosti:

i. presné stanovenie okamihu uvedenia do obehu a
ii. správne nastavenie očakávaní spotrebiteľa,

Za prvé, okamih uvedenia do obehu bude rozhodujúca z dôvodu, že právny režim smernice PLD sa vzťahuje iba na vady, ktoré výrobok už mal v okamihu uvedenia do obehu.

[293] KIILUNEN, V. Autonomous vehicles, Competence and Liability in the EU: Answering the Call of the European Parliament. s. 43. Dostupné na:< https://www.researchgate.net/publication/323934703_AUTONOMOUS_VEHICLES_COMPETENCE_AND_LIABILITY_IN_THE_EU_-_ANSWERING_THE_CALL_OF_THE_EUROPEAN_PARLIAMENT>. cit. 2023-10-25.
[294] European Parliamentary Research Service. A common EU approach to liability rules and insurance for connected and autonomous vehicles. s. 22. Dostupné na: <https://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635_EN.pdf>. cit. 2023-10-25.

V čom spočíva aj jadro jedného z problémov smernice PLD vo vzťahu k jej aplikovateľnosti na inteligentné systémy.

Za druhé, podľa smernice PLD je výrobok vadný, pokiaľ nezabezpečuje bezpečné používanie, ktoré jednotlivec od výrobku právom očakáva. Veríme, že je celkom jednoznačné, že očakávania sa budú líšiť závisiac od výrobku.[295] V prípade automatizovaných vozidiel budú také, že v autonómnom režime budú zabezpečovať bezpečnosť všetkých osôb, vecí a majetku nachádzajúcich sa vo vozidle, ako aj ostatných účastníkov cestnej premávky. Avšak už takéto vymedzenie očakávaní nepočíta s technologickým postupom, kde vývoj bude viesť k ďalšiemu rastu očakávaní v oblasti bezpečnosti, efektívnosti či komfortu.[296]

Ďalším z problematických miest, na ktoré narážame v kontexte vád inteligentných systémov je, že určenie presných vád je v niektorých prípadoch prakticky nemožné.[297] Samozrejme nehovoríme o prípadoch kedy sú zjavné vady od začiatku, ale predstavme si situáciu kedy po niekoľko ročnom používaní automatizovaného vozidla dôjde k zlyhaniu systému umelej inteligencie následkom čoho bude dopravná nehoda z dôvodu, že systém umelej inteligencie vyhodnotil stav cestnej premávky zle. Treba uviesť, že systémy umelej inteligencie fungujú a robia rozhodnutia spôsobom, ktorý pre človeka nemusí byť racionálne a ani jednoducho pochopiteľné.[298] Pričom samotný systém nedokáže odôvodniť alebo podať vysvetlenie o tom prečo sa rozhodlo, ako rozhodlo.[299] A taktiež, že pri inteligentných

[295] Opäť môžeme použiť analógiu a inteligentnom otvárači konzerv a jeho manuálnom bratrancovi v kontraste a autonómnymi vozidlami.

[296] Aj v tomto prípade môžeme vidieť klasický dopad „pacing problému". Koncept „pacing problému" v podstate popisuje výzvy regulačného prostredia udržať tempo s rýchlo sa rozvíjajúcim objektom regulácie. Táto myšlienka je charakteristická pre oblasť nových technológií najmä kvôli skutočnosti, že nové technológie sa objavujú rýchlosťou, s ktorou právne systémy nedokážu držať krok. Systémy umelej inteligencie svojou jedinečnosťou tieto obavy ešte viac prehlbujú, pretože dynamický vývoj umelej inteligencie jednoducho ďaleko presahuje schopnosť akéhokoľvek tradičného regulačného prostredia udržať krok. Čo pre AV znamená to, že v dnešnom svete, za pár mesiacov sa môže radikálne zmeniť to čo budeme považovať za bezpečné pri AV, avšak vyplývajúc zo smernice PLD, výrobca by zodpovedal iba za štandard bezpečnosti, ktorý zodpovedá tomu čo bolo považované za bezpečné v okamihu uvedenia na trh, nič viac. Čoho dôsledkom by bolo nejednoznačnosť nie len pri riešení zodpovednostných scenárov, ale aj pri všeobecnom riadení sa cestnou dopravou. Pozri bližšie Gyurász, Z. From principles to practice: regulating artificial intelligence. Míľniky práva v stredoeurópskom priestore 2020 [elektronický dokument]:zborník z online medzinárodnej vedeckej konferencie doktorandov a mladých vedeckých pracovníkov. -:1. vyd. ISBN 978-80-7160-576-8. - Bratislava: Právnická fakulta UK, 2020. - S. 683-688. cit. 2023-10-25.

[297] GURNEY, J. Sue my car not me: products liability and accidents involving autonomous vehicles. J.L. Tech. & Pol'y 247. 2013. s 265. Dostupné na: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2352108>. cit. 2023-10-25.

[298] Pre problém rozhodovanie v kontexte umelej inteligencie pozri GYURÁSZ, Z., GORNAĽOVÁ, D.: Use of Artificial Intelligence in Arbitration. Cofola International 2021. International and National Arbitration - Challenges and Trends of the Present and Future. - : 1. vyd. ISBN 978-80-210-863. cit. 2023-04-25.

[299]Tento koncept je označovaný ako „the black box problém", čo popisuje fenomén pri systémoch umelej inteligencie, kde rozhodovanie algoritmov funguje spôsobom, že poznáme vstupné dáta a tiež poznáme výsledok ku ktorému algoritmus došiel avšak nevieme sa ani pomocou reverzného inžinierstva dopracovať k tomu ako a prečo sa k takémuto rozhodnutiu

systémoch ako to je aj v tomto prípade, sa už v okamihu kedy bol tento systém umelej inteligencie zasadený do vozidla, sa nepočítalo stým, že fungovanie tohto systému ostane nemenné počas celej doby prevádzky. Pri systémoch umelej inteligencie sa totiž očakáva, že systém bude schopný sa ďalej učiť. Preto môžeme tvrdiť, že „správanie" systému umelej inteligencie v okamihu dopravnej nehody bolo už odlišné od toho ako tento systém vyzeral v čase uvedenia na trh.[300] Problémom však ostáva, že výrobca podľa smernice PLD zodpovedá len za vady, ktoré výrobok mal v čase uvedenia na trh, nie ktoré vznikli následne.

Musíme tiež dodať, že problémom smernice PLD je aj obmedzený rozsah vo vzťahu k zodpovedným osobám, keďže podľa smernice za výrobok resp. škodu ním spôsobenú, zodpovedá výrobca. Táto téza bude jednoducho aplikovateľná pri jednoduchých výrobkoch. Avšak nemusí byť postačujúce v prípade technicky komplexných systémov. Príkladom môže slúžiť prípad, kedy systém umelej inteligencie v automatizovanom vozidla vyhodnotí, že môže zmeniť jazdný pruh, aj keď v tom pruhu v skutočnosti sa nachádzalo iné auto, lenže systém nedokázal rozoznať farbu auta a teda dospelo k záveru, že sa tam nič nenachádza. V tomto prípade bude ktorý z výrobcov zodpovedný? Výrobca samotného automobilu? - kto nejakým spôsobom sa nepričinil o vytvorenie systému autonómneho riadenia? Či to bude výrobca softwaru? Kto nemohol predvídať, že lidarové systémy nedokážu rozoznať farbu auta v istých poveternostných podmienkach? Či to bude výrobca šošoviek pre kamier ktoré používajú tieto lidarové systémy? V týchto prípadoch je viac ako nejednoznačné ktorý subjekt budeme môcť považovať za „výrobcu" a ktorý by následne mal znášať zodpovednosť.

Komplexné fungovanie inteligentných systémov nám jednoducho nedovoľuje aby sme sa mohli poukázať na jeden subjekt a povedať s istotou, kto bude ten, ktorý by mal znášať zodpovednosť za škodu spôsobenú vadným výrobkom. Naviac ani samotná smernica

došlo. Pozri bližšie ZEDNIK, C. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence, Springer, 2019. Dostupné na: < https://arxiv.org/pdf/1903.04361.pdf>. cit. 2023-10-25. alebo GUIDOTTI, R. - MONREALE, A. - PEDRESCHI, D. The AI black box Explanation Problem, ERCIM, 2019. Dostupné na: <https://ercim-news.ercim.eu/images/stories/EN116/EN116-web.pdf#page=12>. cit. 2023-10-25

[300] Pri hlbšej úvahe do tejto otázky sa môžeme prepracovať až k Théseovmu paradoxu, a teda, či môžeme vozidlo so systémom umelej inteligencie ktoré sa pri používaní učí, po rokoch používania vôbec považovať za to isté vozidlo ktoré opustilo priemyselný pás. Viac o Théseovom paradoxe pozri. Bross, B. Theseus' Paradox: History, Authenticity and Identity. Conference: ARCC Toronto. 2020. dostupné na: <https://www.researchgate.net/publication/340037297_Theseus%27_Paradox_History_Authenticity_and_Identity >. cit. 2023-10-25

dnes odpovede na takéto otázky nedáva a pre všetky zainteresované subjekty ostáva situácia neprehľadná a viac ako právne neistá.

Namietať by sa mohlo či už tieto vady, ktoré spôsobili to, že systém umelej inteligencie sa naučil veci, ktoré neskôr spôsobili dopravnú nehodu, už sa nedostali do systému počas trénovania systému, a teda, že výrobok bol vadný už v čase uvedenia na trh. Aj keď tento scenár je možnou alternatívou, dokázanie takejto skutočnosti je v praxi takmer nemožné.[301] Jednak kvôli zložitosti samotného systému umelej inteligencie, ale taktiež narážame na ďalší procesnoprávny problém, keďže otázka dôkazného bremena je označovaná za ďalší nedostatok smernice PLD. Podľa čl. 4 smernice PLD nesie dôkazné bremeno poškodená osoba, ktorá musí preukázať vznik škody, existenciu vady a príčinnú súvislosť medzi vadou a škodou. V prípade automatizovaných vozidiel tak poškodený bude povinný preukázať, že samotný automatizovaný alebo autonómny systém bol príčinou vzniku škody.[302]

Znášanie dôkazného bremena pri takto sofistikovaných systémoch je mimoriadne náročné. Pozitívom ostáva, že poškodený nebude musieť preukázať dôvod zlyhania automatizovaného alebo autonómneho systému, ale iba to, že výrobok ja vadný, a nesplnil právne očakávania a nezabezpečil bezpečné používanie.[303] Opäť sa však dostávame k vyššie naznačeným problémom, ako kvalifikovať právne očakávania pri takto sofistikovaných systémoch s prihliadnutím na skutočnosť, že tieto systémy sa dokážu učiť.

Taktiež je potrebné dodať, že aj keby sa tento scenár preukázal ako pravdivý, trúfame si tvrdiť, že ako právna prax tak ani jurisprudencia nie je pripravená riešiť otázku zodpovednosti za škodu spôsobenú vadným výrobkom z pohľadu otázky možnej evolúcie výrobku počas jeho požívania.

---

[301] GURNEY, J. Sue my car not me: products liability and accidents involving autonomous vehicles. J.L. Tech. & Pol'y 247. 2013. s 265. Dostupné na: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2352108>. cit. 2023-10-25

[302] A common EU approach to liability rules and insurance for connected and autonomous vehicles. s.23 Dostupné na:< https://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635_EN.pdf>. cit. 2023-10-25.

[303] KIILUNEN, V. Autonomous vehicles, Competence and Liability in the EU: Answering the Call of the European Parliament. s. 45 Dostupné na:< https://www.researchgate.net/publication/323934703_AUTONOMOUS_VEHICLES_COMPETENCE_AND_LIABILITY_IN_THE_EU_-_ANSWERING_THE_CALL_OF_THE_EUROPEAN_PARLIAMENT>. cit. 2023-10-25

Druhou otázkou pri inteligentných systémoch z pohľadu smernice PLD je ich klasifikácia ako výrobku. Pojem „výrobok" je v čl. 2 smernice PLD definovaný ako všetky hnuteľnosti. Aj keď na prvý pohľad by sa mohol zdať, že systémy umelej inteligencie sú bez akýchkoľvek pochybností hnuteľná vec, spĺňajú definíciu výrobku a preto by sa tento právny režim smernice PLD mal vzťahovať aj na automatizované vozidlá.

Musíme sa však aj pri tejto otázke zastaviť. Nie je hanbou si priznať, že pre nás právnikov je vytváranie si analógií[304] alfou a omegou, keďže takto dokážeme aj ten najbizarnejší fenomén subsumovať pod existujúce ustanovenia tých-ktorých noriem. Veríme, že aj v tomto prípade by sme veľmi jednoducho mohli padnúť do tej pasce aby sme vytvorili až banálnu analógiu medzi plechovým vonkajškom na jednej strane a konvenčnými výrobkami na strane druhej. Existuje však niekoľko jednoznačných rozdielov, kvôli ktorým by sme nad takouto analógiu nemali ani uvažovať.

Prvým takýmto prvkom je softwarové vybavenie inteligentných systémov. Je to práve systém umelej inteligencie, ktorý odlišuje inteligentné zariadenia od konvenčných výrobkov. Systém umelej inteligencie vozidla bude tou najcennejšou a zároveň z pohľadu možného zlyhania najkritickejšou súčasťou inteligentných. Ak by sme veľmi chceli, mohli by sme zredukovať celú podstatu inteligentných systémov na tento software a tvrdiť, že je to ich základnou esenciou.

Preto však nastáva otázka či vôbec budú inteligentné systémy spadať pod režim smernice PLD?

Považujeme za dôležité poznamenať, že inteligentné systémy treba vnímať nielen z kvantitatívneho hľadiska ako extra funkciu výrobku, zabezpečovanú zabudovaným systémom umelej inteligencie, ale aj z kvalitatívneho hľadiska ako úplne nový produkt.[305]

Takýto prístup si vyžaduje oddelenie vlastníctva hnuteľnej veci od systému umelej inteligencie. A keďže ako sme vyššie uviedli, do právneho režimu ktorú zavádza smernica PDL

[304] Či už prostredníctvom „analógie legis" alebo „analógie iuris".
[305] Podstata inteligentných systémov, nie je súčet konvenčného výrobku s pridaním systému umelej inteligencie, ale ide o kvalitatívne nový výrobok.

patria iba hnuteľné veci, pričom software hnuteľnou vecou nie je, preto akokoľvek by sme chceli nebudeme môcť subsumovať naše prípady pod ustanovenia smernice PDL.[306]

Taktiež stoji za zmienku, že rovnakým právnym režimom sa budú musieť riadiť aj softwarové aktualizácie urobené v čase po uvedení výrobku do obehu. Rovnako ako počítače či smartfóny aj inteligentné systémy budú musieť byť priebežne aktualizované za účelom odstránenia chýb alebo rozšírenia ich funkčnosti. Poskytovanie softwarovej podpory a softwarových aktualizácií zo strany výrobcu vozidla alebo výrobcu softwaru po celú dobu životnosti výrobku svojou povahou viac pripomína samostatnú službu ako sprievodnú podporu poskytovanú výrobku, na čo sa smernica PLD opäť nevzťahuje.[307]

Takéto striktné oddelenie hardwaru a softwaru však nemožno považovať za koncepčný prístup. Jednoduchším riešením uvedeného problému je rozšírenie pojmu výrobok tak, aby výslovne zahŕňal aj software a softwarové aktualizácie výrobku, čo je dnes aj na úrovni EÚ zjavne preferované riešenie.[308]

Práve aj tento problém bude riešiť novela smernice PLD.[309] Jednou zo základných zmien ktoré toto prepracovanie prináša je skutočnosť, že do vymedzenia pojmu „výrobok" sa explicitne dostane aj software.[310] Práve za účelom, aby smernica PDL sa stala viac životaschopnou v dnešnom digitálnom svete. Naďalej očakávame v akej konečnej forme sa dostane smernica do praxe, avšak teraz sme iba na začiatku dlhej legislatívnej cesty. A

[306]Questions and answers on the revision of the Product Liability Directive Dostupné na:< https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_5791>. cit. 2023-10-25

[307] CHATZIPANAGIOTIS, M. Product Liability Directive and Software Updates of Automated Vehicles. 2021 Dostupné na: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3759910>. cit. 2023-10-25

[308] New liability rules on products and AI to protect consumers and foster innovation. Dostupné na: <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807>. cit. 2023-10-25

[309] Dna 28.9.2022 Komisia prijala dva návrhy na prispôsobenie pravidiel zodpovednosti digitálnemu veku. Komisia po prvý raz navrhuje cielenú harmonizáciu vnútroštátnych pravidiel zodpovednosti za umelú inteligenciu, ktorá obetiam škôd súvisiacich s umelou inteligenciou uľahčí získanie odškodnenia. Nové pravidlá majú zabezpečiť, aby obete mali rovnakú ochranu v prípadoch keď došlo ku škode výrobkami na báze umelej inteligencie, ako by to urobili, keby bola spôsobená škoda za akýchkoľvek iných okolností. Pozri bližšie New liability rules on products and AI to protect consumers and foster innovation. Dostupné na: <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807>. cit. 2023-10-25

[310] Proposal for a directive of the European Parliament and of the Council on liability for defective products. Dostupné na: <https://single-market-economy.ec.europa.eu/document/3193da9a-cecb-44ad-9a9c-7b6b23220bcd_en.>. cit. 2023-10-25

predtým ako bude táto nová verzia smernice PLD platná a účinná budeme musieť ešte nejaký čas musieť počkať.[311]

5.4.2. Otázky zodpovednosti spojené s povinným zmluvným poistením

Ako sme už vyššie uviedli nemôžeme zabúdať však ani na otázku poistenia v kontexte zodpovednostných scenárov niektorých inteligentných systémov.[312] V týchto prípadoch prioritnú rolu bude hrať pre nás smernica Európskeho parlamentu a Rady 2021/2118/ES z 24. novembra 2021 ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a o kontrole plnenia povinnosti poistenia takejto zodpovednosti.

Súčasná iterácia je už šiestou verziou smernice MVID. Naposledy v roku 2017 Komisia vykonala hodnotenie fungovania smernice Európskeho parlamentu a Rady 2009/103/ES vrátane jej efektívnosti, účinnosti a súdržnosti s ostatnými politikami Únie.[313] Z hodnotenia vyplynulo, že smernica 2009/103/ES funguje celkovo dobre a netreba vykonať zmeny vo väčšine aspektov. Boli však identifikované štyri oblasti, v ktorých by boli vhodné cielené zmeny:

i. Odškodnenie poškodených v dôsledku nehôd v prípade, že dotknutá poisťovňa je platobne neschopná,
ii. Minimálne povinné výšky poistného krytia,
iii. Kontroly poistenia vozidiel členskými štátmi a
iv. Používanie potvrdení o minulých nárokoch na odškodnenie poistencov novou poisťovňou.

Už z preambuly môžeme vidieť, že Komisia zatiaľ nepovažuje otázky automatizovaných vozidiel za problematické a zmeny ktoré nastali boli iného charakteru.

[311] Návrh Komisie bude teraz musieť prijať Európsky parlament a Rada. Následne sa navrhuje, aby o päť rokov po nadobudnutí účinnosti smernice o zodpovednosti umelej inteligencie Komisia v prípade potreby posúdila potrebu pravidiel objektívnej zodpovednosti za nároky súvisiace s umelou inteligenciou.
[312] ako napr. automatizovaných vozidiel.
[313]Smernica Európskeho parlamentu a Rady (EÚ) 2021/2118 z 24. novembra 2021, ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a o kontrole plnenia povinnosti poistenia takejto zodpovednosti. Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:32021L2118>. cit. 2023-04-25

MVID však nezabúda úplne na otázku automatizovaných vozidiel. V bode 39 preambuly explicitne hovorí analýze a preskúmaní automatizovaných vozidiel kde sa uvádza, že *„s cieľom zabezpečiť, aby smernica 2009/103/ES naďalej slúžila svojmu účelu, ktorým je ochrana potenciálnych poškodených pred nehodami s účasťou motorových vozidiel, by Komisia mala takisto monitorovať a preskúmať uvedenú smernicu vzhľadom na technologický vývoj vrátane zvýšeného používania autonómnych a poloautonómnych vozidiel. Mala by tiež analyzovať, ako poisťovne používajú systémy, v ktorých je poistné ovplyvnené potvrdeniami o minulých nárokoch na odškodnenie poistencov. Okrem toho by Komisia mala posúdiť účinnosť systémov výmeny informácií používaných na účely cezhraničných kontrol poistenia*."[314] Tento „prísľub" normativizuje následne v článku 28c. ods. 2 kde sa uvádza, že: *„Najneskôr 24. decembra 2030 Komisia predloží správu Európskemu parlamentu, Rade a Európskemu hospodárskemu a sociálnemu výboru, v ktorej vyhodnotí vykonávanie tejto smernice, s výnimkou prvkov, ktorých sa týka hodnotenie uvedené v odseku 1, a to aj pokiaľ ide o*:

a) uplatňovanie tejto smernice v súvislosti s technologickým vývojom, najmä pokiaľ ide o autonómne a poloautonómne vozidlá;
b) primeranosť rozsahu pôsobnosti tejto smernice, berúc do úvahy riziká nehôd, ktoré predstavujú rôzne motorové vozidlá;
c) (vo forme preskúmania) účinnosť systémov výmeny informácií na účely kontrol poistenia v cezhraničných situáciách, v prípade potreby vrátane posúdenia realizovateľnosti využívania existujúcich systémov výmeny informácií v takýchto prípadoch, a v každom prípade analýzu cieľov systémov výmeny informácií a posúdenie ich nákladov, a
d) používanie takých systémov poisťovňami, v ktorých je poistné ovplyvnené potvrdeniami o minulých nárokoch na odškodnenie poistencov, okrem iného systémov „bonus-malus" alebo bonus za bezškodový priebeh."

Aj táto najaktuálnejšia podoba smernice MVID predpokladá zotrvanie pri už existujúcom hybridnom modeli založenom riešení poistných a škodových scenárov. Ktorá je

[314] Bod 39. smernice Európskeho parlamentu a Rady (EÚ) 2021/2118 z 24. novembra 2021, ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a o kontrole plnenia povinnosti poistenia takejto zodpovednosti. Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:32021L2118>. cit. 2023-10-25

založená na kombinácii naprieč Európou neharmonizovaných a v každom členskom štáte osobitne upravených pravidiel, ktoré zakladá priamy nárok poškodeného voči poisťovni, ktorá poisťuje prevádzkovateľa vozidla.[315] Čo sa týka právneho poriadku Slovenskej republiky je zodpovednosť prevádzkovateľa založená na objektívnom princípe. Otázka zavinenia je pre vyvodenie zodpovednosti bezvýznamná a zodpovednosti sa možno zbaviť len v určitých špecifických prípadoch.[316]

Uvedený model je potrebné vnímať ako efektívny nástroj alokácie rizika spojeného s prevádzkou dopravného prostriedku na celom území EÚ, ktorého účelom je chrániť predovšetkým poškodeného a len v menšej miere aj prevádzkovateľa motorového vozidla.[317]

Je však aplikovateľný aj škody spôsobené prevádzkou automatizovaných vozidiel? Smernica MVID v článku 1 definuje vozidlo ako:

a) *„každé motorové vozidlo poháňané výlučne mechanickou energiou na pevnine, ktoré sa však neprevádzkuje na koľajniciach, s: (i)maximálnou konštrukčnou rýchlosťou viac ako 25 km/h alebo (ii)maximálnou prevádzkovou hmotnosťou viac ako 25 kg a maximálnou konštrukčnou rýchlosťou viac ako 14 km/h;*
b) b) *každé prípojné vozidlo, ktoré sa má používať s vozidlom uvedeným v písmene a), bez ohľadu na to, či je pripojené alebo oddelené. Bez toho, aby boli dotknuté písmená a) a b), sa invalidné vozíky určené výlučne na používanie osobami s telesným postihnutím nepovažujú za vozidlá uvedené v tejto smernici*;"[318]

Pričom za prevádzkovanie vozidla sa podľa článku 1 bodu 1a považuje: *„akékoľvek používanie vozidla, ktoré je v súlade s funkciou vozidla ako dopravného prostriedku v čase*

[315]PATTI, F. The European road to autonomous vehicles. 2019. Dostupné na: https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2766&context=ilj cit. 2023-10-25.
[316] Porovnaj Števček, M. a kol. Občiansky zákonník, Komentár. 2. vydanie. Praha: C.H.Beck, 2015.Najmä nás budú zaujímať ustanovenia §427 a nasledujúce.
[317] KIILUNEN, V. Autonomous vehicles, Competence and Liability in the EU: Answering the Call of the European Parliament. s. 43. Dostupné na: https://www.researchgate.net/publication/323934703_AUTONOMOUS_VEHICLES_COMPETENCE_AND_LIABILITY_IN_THE_EU_-_ANSWERING_THE_CALL_OF_THE_EUROPEAN_PARLIAMENT cit. 2023-04-25 alebo PATTI, F. The European road to autonomous vehicles. 2019. Dostupné na: https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2766&context=ilj cit. 2023-10-25.
[318] Smernica Európskeho parlamentu a Rady (EÚ) 2021/2118 z 24. novembra 2021, ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a o kontrole plnenia povinnosti poistenia takejto zodpovednosti. Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:32021L2118#>. cit. 2023-10-25.

nehody, bez ohľadu na jeho vlastnosti, na terén, na ktorom sa takéto motorové vozidlo používa, a bez ohľadu na to, či stojí alebo je v pohybe;".[319]

Takto široko a najmä technologicky neutrálne vymedzenie pojmov teda nebraní tomu, aby v budúcnosti boli do režimu smernice MVID boli zahrnuté aj automatizované vozidla, bez ohľadu na ich výbavu či stupeň automatizácie.[320]

Bohužiaľ, iné ustanovenia smernice MVID ukazujú isté nedostatky, čo sa týka schopnosti smernice úspešne regulovať automatizované vozidlá. Ako príklad môžeme uviesť znenie článku 12 kde sa uvádza, že *„poistenie uvedené v článku 3 bude pokrývať zodpovednosť za ujmu na zdraví, ktorá je dôsledkom prevádzky vozidla, spôsobenú všetkým cestujúcim okrem vodiča"*[321] je možné vyvodiť záver, že smernica MVID sa vzťahuje len na vozidlá s vodičom a ostatné osoby sú považovaní za tretiu stranu. Nezodpovedanou teda ostáva otázka, či sa režim povinného zmluvného poistenia bude vzťahovať aj na škody spôsobené prevádzkou vozidla, v režime plnej automatizovanej prevádzky vozidla a či osoby potrebné považovať za cestujúcich, za vodičov či za tretie osoby.[322]

---

[319] Tamtiež.

[320] Princíp technologickej neutrality je téma, ktorá je už dlhodobo a podrobne skúmaná a väčšina autorov ju identifikuje ako regulačné kritérium normatívnych systémov vo svetle nových technológií. Princíp technologickej neutrality znamená, že regulačná činnosť by sa nemala zameriavať na vybranú technológiu, ale na účinky, ktoré vyplývajú z jej používania. Z tohto dôvodu vyššie uvedené právne predpisy stanovujú, že legislatíva musí byť založená na udržateľnej, subsidiárnej a primeranej regulácii, ktorá je zároveň transparentná. Technologická neutralita je tiež priamo spojená s myšlienkou flexibility regulácie, ktorá vychádza zo základnej myšlienky, že technológia sa vyvíja rýchlejšie ako regulácia. Regulácia preto musí byť udržateľná a musí sa vyhýbať neustálym právnym kontrolám s cieľom prispôsobiť sa neustálym technologickým zmenám. Na zabezpečenie tohto cieľa by regulácia nemala obmedzovať svoj rozsah pôsobnosti na konkrétnu technológiu, skôr by sa mala zamerať na účinky plynúce z jeho používania, umožňujúce flexibilnú reguláciu. Regulácia by mala byť otvorená zmenám, pokrokom, inováciám, ktoré sa vyskytnú v rozsahu jeho pôsobnosti, čo umožní nepretržité uplatňovanie aj pri vývoji nových technológií. Predpisy by nemali byť statické, ale dostatočne pružné a dynamické na to, aby sa vyvíjali spolu s technologickým vývojom, bez potreby neustálych revízií predpisov. Porovnaj FERNÁNDEZ, A.G.: how to regulate the future of technology challenges and principles, Tallinn University, 2020, Dostupné na: <http://publications.tlu.ee/index.php/eastwest/article/view/900.>. cit. 2023-04-25.alebo HAAR, I. The principle of technological neutrality: Connecting EC network and content regulation, Tilburg university, 2010. cit. 2023-04-25. alebo KOOPS, B.: Should ICT Regulation Be Technology-Neutral?: Starting points for ict regulation. Deconstructing prevalent policy one-liners, it & law series, Bert-Jaap Koops, Miriam Lips, Corien Prins & Maurice Schellekens, eds., Vol. 9, s77- 108, the hague: T.M.C. Asser Press, 2006. Dostupné na: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=918746>. cit. 2023-10-25.

[321] Smernica Európskeho parlamentu a Rady (EÚ) 2021/2118 z 24. novembra 2021, ktorou sa mení smernica 2009/103/ES o poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a o kontrole plnenia povinnosti poistenia takejto zodpovednosti. Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:32021L2118#>. cit. 2023-10-25.

[322] European Parliamentary Research Service. A common EU approach to liability rules and insurance for connected and autonomous vehicles, s. 20 cit. Dostupné na: <https://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635_EN.pdf>. cit. 2023-10-25. alebo Light, D. A scenario: The end of auto insurance. Celent | Experts in financial services., 2012. Dostupné na: <https://www.celent.com/insights/121822340.>. 2023-04-25

Z pohľadu európskeho vnútorného trhu a vytvorenia uceleného prostredia, v rámci ktorého sa bude uskutočňovať budúca prevádzka automatizovaných vozidiel, možno ako ďalší nedostatok vnímať aj už naznačené odlišnosti v rámci národných právnych úprav zodpovednosti za škodu spôsobenú prevádzkou dopravného prostriedku a s tým súvisiacu odlišnú aplikačnú prax v jednotlivých členských štátoch EÚ.[323]

V kontexte právnej istoty by ideálne bolo ak by každá poistná a škodová udalosť vyvolaná prevádzkou automatizovaného vozidla mala byť v každom členskom štáte EÚ riešená rovnako. Vzhľadom však na naznačené odlišnosti regulácie nie je očkovateľné, hoci skutkové okolnosti prípadu budú prakticky totožné.[324]

V čoraz globalizovanejšom prostredí, v ktorom vonkajšie sily drasticky ovplyvňujú našu spoločnosť, sa musíme usilovať o koordinovanú a harmonizovanú reguláciu, v ktorej sa rôzne krajiny dohodnú na poskytnutí jasného bezpečnostného rámca bez regulačných mozaík.[325] Európska únia tu má veľkú príležitosť stať sa priestorom, v ktorom majú inovácie ako automatizované vozidlá miesto. Európska únia musí využiť túto príležitosť rozvinutím svojej integračnej úlohy, ktorej cieľom je efektívna a konkurencieschopná regulácia. Článok 179 Zmluvy o fungovaní Európskej únie špecifikuje, že *„cieľom Únie je posilniť jej vedecké a technologické základy prostredníctvom vytvorenia európskeho výskumného priestoru, v ktorom budú voľne obiehať výskumní pracovníci, vedecké poznatky a technológie, a povzbudzovať ju, aby sa stala konkurencieschopnejšou*."[326] Rozdrobená metodológia regulácie vytvára medzi poskytovateľmi služieb a spotrebiteľmi neistotu a môže vážne narušiť inovácie. Cieľom EÚ preto musí byť posunutie sa smerom k digitalizácii slobôd jednotného trhu EÚ prijatím celoeurópskych štandardov pre nové technológie.[327] Legislatíva EÚ musí byť primeraná digitálnej transformácii a dátovej ekonomike. EÚ od svojho založenia vždy pozerala do

[323] KIILUNEN, V. Autonomous vehicles, Competence and Liability in the EU: Answering the Call of the European Parliament. s. 43. Dostupné na: https://www.researchgate.net/publication/323934703_AUTONOMOUS_VEHICLES_COMPETENCE_AND_LIABILITY_IN_THE_EU_-_ANSWERING_THE_CALL_OF_THE_EUROPEAN_PARLIAMENT cit. 2023-04-25 cit. 2023-04-25.

[324] PATTI, F. The European road to autonomous vehicles. 2019. Dostupné na: https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2766&context=ilj cit. 2023-10-25.

[325] GYURÁSZ, Z. Modality regulácie nových technológií. 2021. Dostupné na: <https://www.researchgate.net/publication/361440930_Modality_regulacie_novych_technologii>. cit. 2023-10-25.

[326] Zmluva o fungovaní európskej únie. Dostupné na: <https://eur-lex.europa.eu/resource.html?uri=cellar:9e8d52e1-2c70-11e6-b497-01aa75ed71a1.0021.01/DOC_3&format=PDF >. cit. 2023-10-25.

[327] Porovnaj, Smernica Európskeho Parlamentu A Rady (EÚ) 2019/790 zo 17. apríla 2019 o autorskom pravé a právach súvisiacich s autorským pravom na digitálnom jednotnom trhu a o zmene smerníc 96/9/ES a 2001/29/ES, Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:32019L0790&from=EN>. cit. 2023-10-25.

budúcnosti a budúcnosťou je porozumenie a regulácia technológie. Cieľom Európskej komisie je vytvoriť skutočne digitálnu európsku spoločnosť.[328] Samozrejme, nejde však len o európsku úlohu, ale celý svet by mal mať tendenciu k tomuto druhu koordinácie a spolupráce pri vytváraní harmonizácie regulácie pre technologický priemysel, ktorý pomáha budovať efektívny rámec na ochranu spotrebiteľov, zvyšovať transparentnosť. V digitálnom svete bez hraníc, a tam, kde je všetko prepojené, by mala byť prepojená a koordinovaná aj globálna regulácia, aby sa zabránilo nedostatočnej ochrane spotrebiteľov.

V neposlednom rade treba upozorniť na to, že súčasný model objektívnej zodpovednosti prevádzkovateľa dopravného prostriedku za škodu spôsobenú osobitnou povahou prevádzky je založený na predpoklade, že pravé prevádzkovateľje tou osobou, ktorá pozná stav vozidla, zabezpečuje jeho údržbu, dlhodobo vozidlo užíva, a vykonáva nad vozidlom faktickú aj právnu moc, a preto by mal zodpovedať za škodu, ktorá bude spôsobená okolnostnou majúcou pôvod v prevádzke vozidla. Pri prevádzke automatizovaného vozidla však vplyv prevádzkovateľa na prevádzku vozidla bude menší, a vo vzťahu k automatizovaným systémom vozidla dokonca žiadny. [329]

V tejto súvislosti je možné vychádzať z rozumného predpokladu, že plnoautomatizovaná prevádzka bude pripustená až po tom, čo úroveň automatizovaných systémov bude umožňovať úplne bezrizikovú prevádzku. Nie je možno očakávať že budú povolené nepredvídateľné a potenciálne nebezpečné systémy.

5.5. Zodpovednosť za porušenie ľudských práv

Ochrana základných ľudských práv predstavuje pomerne úzko prepojenú problematiku súvisiace s umelou inteligenciou.

Ako už bolo naznačené vyššie, kvalitné údaje sú základným kameňom vývoja, testovania a používania aplikácií na báze umelej inteligencie. Čím viac údajov, tým kvalitnejšia a efektívnejšia umelá inteligencia. Prirodzene, v každej z vyššie uvedených fáz sa

[328] Porovnaj Urýchlenie digitálnej transformácie verejných správ v EÚ – akčný plán na roky 2016 – 2020, Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/HTML/?uri=LEGISSUM:4301896&from=SK>. cit. 2023-10-25.
[329]PATTI, F. The European road to autonomous vehicles. 2019. Dostupné na:< https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2766&context=ilj>. cit. 2023-10-25.

aplikuje jemne odlišný rámec problémov. Na poukázanie rámcových problémov s umelou inteligenciou považujeme za vhodné tieto problémy zasadiť práve do kontextu základných zásad. V rámci aplikácií UI pôjde predovšetkým o plne alebo polo-automatizovanú formu spracúvania osobných údajov.

Otázkou preto ostáva, ako sa využijú tieto informácie pri strojoch na báze umelej inteligencie?

Rozhodovanie ako také, je evidentne ústredné pre právo, mentálne procesy rozhodovania zostávajú však neistotou v srdci právneho diškurzu. Hovorí sa, že myšlienka automatizovaného rozhodovania fascinovala akademikov už od začiatku sedemdesiatych rokov minulého storočia.[330] Význam rozhodovania pri prevádzaní právnych noriem do praxe je zrejmý, nakoľko je ťažké si predstaviť akýkoľvek prípad, kedy by sa právne predpisy mechanicky sami od seba uplatňovali v ktorejkoľvek fáze právneho procesu.[331] Svet v ktorom by sa automatizované systémy dali použiť na rozhodnutia, ktoré sa musia robiť často a rýchlo, mohol znieť pre niektorých až utopický.

Dlho sa verilo, že ak sa pravidlá rozhodovania dajú ľahko kodifikovať a ak sú k dispozícii vysokokvalitné údaje, je veľká šanca, že sa rozhodnutie dá automatizovať.[332] Preto, ak budeme veriť konceptu založenom na celkom jednoduchej myšlienke, že rozum je konečnou silou, ktoré dokáže jedinečne odfiltrovať relevantné od nepodstatného[333], tak podľa takéhoto racionalistického pohľadu teda právne rozhodnutia prirodzene vychádzajú z predpísaných foriem logických záverov, menovite dedukcií, indukcií a analógií.[334]

Ako sa však tieto fenomény odzrkadľujú pri rozhodovaní umelej inteligencie?

Veríme, že do bodu, keď fenomén rozhodovania úplne prestane existovať ako ontologická výsada ľudskej existencie a stane sa bežnou črtou umelej inteligencie v každom

[330] Pozri bližšie, GYURÁSZ, Z. - GORNAĽOVÁ, D.: Use of Artificial Intelligence in Arbitration v Cofola International 2021 [elektronický dokument] : International and National Arbitration - Challenges and Trends of the Present and Future. - : 1. vyd. ISBN 978-80-210-8639-5. - Brno : Masarykova univerzita, 2021. - S. 80-101
[331] EPSTEIN, D.: Rationality, Legitimacy, & The Law, Washington University Jurisprudence Review, 2014, Dostupné na:: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1103&context=law_jurisprudence
[332] HARRIS, J. DAVENPORT, T.: Automated Decision Making Comes of Age, MIT Sloan Management Review 46(4), 2005.
[333] EPSTEIN, D.: Rationality, Legitimacy, & The Law, Washington University Jurisprudence Review, 2014, Dostupné na:: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1103&context=law_jurisprudence
[334] MACCORMICK, N.: Rhetoric and The Rule of Law: A Theory of Legal Reasoning, Oxford University Press, 2010

aspekte nášho života, je pred nami ešte dlhá cesta. Musíme preto veriť, že rozhodovanie vyžaduje okrem inteligencie aj štipku, povedzme, zdravého rozumu.[335]

Ak sa bližšie pozrieme na stanovisko pracovnej skupiny zriadenej podľa článku 29 pre automatizované individuálne rozhodovanie a profilovanie na účely Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/679, môžeme vidieť, že výlučne automatizované rozhodovanie je definované ako „schopnosť vykonávať rozhodnutia technologickými prostriedkami bez ľudskej účasti". Podľa názoru pracovnej skupiny zriadenej podľa článku 29 môžu byť automatizované rozhodnutia založené na akomkoľvek type údajov, najmä však:

1) Po prvé, o údaje poskytnuté priamo dotknutými jednotlivcami (ako sú odpovede na dotazník),
2) Po druhé, o údaje pozorované o jednotlivcoch (napríklad údaje o polohe zhromaždené prostredníctvom aplikácie),
3) Po tretie, o odvodené údaje, ako napríklad profil jednotlivca, ktorý už bol vytvorený (napr. kreditné skóre).

Pri automatizovanom individuálnom rozhodovaní je potrebné odlišovať všeobecné rozhodnutia na základe profilovania a rozhodnutia urobené výlučne automatizovaným spracúvaním s právnymi účinkami na dotknutú osobu. V zmysle článku 22 ods. 1 GDPR: *„Dotknutá osoba má právo na to, aby sa na ňu nevzťahovalo rozhodnutie, ktoré je založené výlučne na automatizovanom spracúvaní, vrátane profilovania, a ktoré má právne účinky, ktoré sa jej týkajú alebo ju podobne významne ovplyvňujú."* Avšak inteligentné zariadenia potrebujú zhromažďovať dáta a následne musia vyvodzovať závery, aby mohli fungovať správne. Súčasne ale takéto zhromažďovanie informácií môže viesť k narušeniu súkromných životov užívateľov.

Ako všetci veľmi dobre vieme, korelácia nemusí nutne znamenať kauzalitu. A tento problém je evidentný pri systémoch na báze umelej inteligencie, keďže väčšina systémov

[335] Práve z tohto dôvodu zostáva najväčšou výzvou pri automatizovanom individuálnom rozhodovaní otázka akumulácie, výberu a správy dát ktoré sa použijú. Príkladom tejto tézy sú zistenie spoločnosti ZestFinance Inc, ktorá zo svojich údajov zistila, že vyšší ľudia majú lepšiu schopnosť splácať pôžičky, alebo informácie, že ľudia, ktorí vypĺňajú žiadosti o pôžičku len s použitím veľkých písmen, splácajú lepšie ako ľudia, ktorí používajú iba malé písmená. V konečnom dôsledku môžeme vidieť, že umelá inteligencia môže urobiť systémy inteligentnejšími, no bez pridania štipky zdravého rozumu môže spôsobiť značné nepríjemnosti.

strojového učenia nekombinuje myslenie s výpočtami. Jednoducho chrlia koreláciu údajov, či už dávajú zmysel alebo nie.[336]

Z tohto dôvodu by sme sa mali bližšie pozrieť na problematické miesta, ktoré sú konštantou pri otázke automatizovaného individuálneho rozhodovania. Za najdôležitejšie považujeme vyzdvihnúť je možná diskriminácia.

Zatiaľ čo väčšina údajov, čo zariadenie na báze umelej inteligencie využívajú, je neškodných, len sa dá ťažko vylúčiť, že iné získané informácie nebudú zneužité. Subjekty, o ktorých sa zbierajú tieto dáta sú často bezmocní pri prevencii potenciálnych hrozieb, kvôli nedostatočne možnej kontrole a nedostatku informácií.

Všeobecné obavy o ochranu osobných údajov sú u zariadení na báze umelej inteligencie znásobené o spôsob, akým tieto zariadenia rozširujú možnosť a dosah sledovania a získavanie informácií. Charakteristiky zariadení na báze umelej inteligencie a spôsobov, ako sa tieto systémy používajú, predefinujú celú otázku akým spôsobom sa zbierajú, analyzujú, používajú a chránia osobné údaje.[337] Tento druh získavania informácií o užívateľoch otvorí dvere novým formám **diskriminácie**.[338] Pričom súčasné antidiskriminačné zákony nie sú pripravené riešiť diskrimináciu vo svetle zariadení na báze umelej inteligencie.

Podľa Wachterovej existujú najmenej tri možné spôsoby monitorovania a profilovania, ktoré môžu slúžiť ako základ pre možnú diskrimináciu:[339]

1) Po prvé, zhromažďovanie údajov, ktoré vedú k záverom o osobe (napríklad správanie pri prehliadaní internetu);

---

[336] Porovnaj pri otázke implementácie umelej inteligencie do zdravotníctva. MESARČÍK, M. GYURÁSZ, Z.: Umelá inteligencia a právna úprava zdravotníctva v Slovenskej republike, Bratislava : Právnická fakulta UK , 2020. Dostupné na: https://www.flaw.uniba.sk/fileadmin/praf/Pracoviska/Ustavy/UPITPDV/Mesarcik_Gyurazs_-_AI_a_zdravotnictvo.pdf

[337] ROSE, K. ELDRIDGE, S. CHAPIN, L.: The Internet of Things: an Overview, 2015 Dostupné na: https://www.researchgate.net/profile/Ananth_Saradhi2/post/What_are_the_latest_developments_in_IOT_architectures/attachment/59d6387d79197b8077995b29/AS:397885586853893@1471874724774/download/57.pdf

[338] PEPPET, S.: Regulating the Internet of Things: First Steps toward Managing Discrimination, Privacy, Security and Consent ,2014 Dostupné na: https://texaslawreview.org/wp-content/uploads/2015/08/Peppet-93-1.pdf

[339] WACHTER, S.: Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR, 2017 Dostupné na: https://www.researchgate.net/profile/Sandra_Wachter2/publication/321959135_Normative_Challenges_of_Identification_in_the_Internet_of_Things_Privacy_Profiling_Discrimination_and_the_GDPR/links/5baab8bfa6fdccd3cb733211e/Normative-Challenges-of-Identification-in-the-Internet-of-Things-Privacy-Profiling-Discrimination-and-the-GDPR.pdf

2) Po druhé, profilovanie vo všeobecnosti, prostredníctvom prepojenia údajov zo zariadení (tiež nazývaná ako "senzorová fúzia");
3) Po tretie, profilovanie, ku ktorému dochádza keď sú údaje zdieľané s treťou stranou, ktorá kombinuje údaje s inými dátovými súbormi (napr. zamestnávateľ, poisťovňa).

Najväčšiu výzvu pre tému diskriminácie vo svetle umelej inteligencie predstavuje efekt senzorovej fúzie, kde okrem toku zozbieraných údajov a informácií z jedného zdroja môžu byť vyvodené širšie závery v kombinácii s inými údajmi. Ako príklad senzorovej fúzie môžeme uviesť *„Smart-watch"* zariadenia. Tieto zariadenia vytvárajú naraz prehľad o zdravotnom stave užívateľa,[340] pričom zaznamenávajú aj údaje o pohybe používateľa.[341] Aj keď sami o sebe tieto údaje nemusia predpovedať o užívateľovi nič extravagantné, predstavme si situáciu kde pred nástupom do zamestnania, budúceho zamestnávateľa bude zaujímať vaša fyzická výkonnosť. Na základe týchto informácií by sa mohli vyvodiť veľmi jednoducho závery ohľadne návykov uchádzača *ako sú:* životospráva, požívanie alkoholu alebo drog, fajčenie alebo nedostatok spánku. Práve tieto faktory boli spájané so zlým psychickým stavom, zdravotnými problémami a slabým kognitívnym výkonom. V konečnom dôsledku tieto zistenia môžu odradiť potencionálneho zamestnávateľa vybrať takéhoto uchádzača. Veríme preto, že dopady senzorovej fúzie neradno podceňovať a vôbec už nie zaradovať do sféry právnej sci-fi.

Tieto obavy sa odrážajú samozrejme aj v GDPR a to najmä v článku 21[342] a v článku 22.[343] Článok 21 zavádza právo dotknutých osôb kedykoľvek namietať proti spracovaniu údajov, pričom ak je účelom spracovania údajov priamy marketing, dotknutá osoba má absolútne právo na námietku.[344] Vo všetkých ostatných prípadoch sa spracovanie údajov musí zastaviť, pokiaľ správca údajov nemôže preukázať nevyhnutné oprávnené dôvody na spracúvanie ktoré prevažujú nad záujmami, právami a slobodami dotknutej osoby.[345] Taktiež článok 5 ods. 1 písm. a) GDPR ustanovuje, že osobné údaje musia byť spracúvané zákonným spôsobom, transparentne a spravodlivo. Práve tieto tri atribúty sú najčastejšie problematické

[340] Najmä srdcovej frekvencie
[341] Geolokácia alebo vykonané kroky za deň
[342] Právo podať námietku voči spracúvaniu osobných údajov.
[343] Automatizované individuálne rozhodovanie vrátane profilovania
[344]GDPR článok 21, 2018 Dostupné na: https://gdpr.algolia.com/gdpr-article-21
[345] Problematickou však ostáva skutočnosť že GDPR bližšie nedefinuje pojem „nevyhnutné oprávnené dôvody".

v kontexte umelej inteligencie. V súvislosti s umelou inteligenciou sa však často spomína aj požiadavka zákazu diskriminácie. Napriek tomu, že technológie na báze umelej inteligencie nepodliehajú zmenám nálad alebo iným subjektívnym pocitom, môže k diskriminácií dôjsť tým, že v samotných dátach, na základe ktorých technológia robí rozhodnutia je už prítomný prvok diskriminácie a technológia to bude považovať za štandard. Ako príklad možno uviesť časté zatýkanie jednotlivcov z minorít na základe rasistických predsudkov. Ak tieto dáta budú „vložené" do umelej inteligencie a tá sa z nich bude učiť a ďalej s nimi pracovať, je veľmi pravdepodobné, že výsledky budú diskriminačné. Túto tézu rozvinula a dokázala americká štúdia, ktorá analyzovala systém na vyhodnotenie prepustenia z väzby, ktorý diskriminoval Afroameričanov.[346]

[346] Pozri viac ANGWIN, J. LARSON, J. MATTU, S. KRICHNER, L.: Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. 2016 Dostupné na: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

6. Právo duševného vlastníctva a umelá inteligencia

6.1 Úvodné poznámky

Tam, kde pred niekoľkými desaťročiami mohli iba ľudia písať básne, hrať šachy alebo vytvárať inovácie, dnes tieto úlohy bežne vykonávajú systémy na báze umelej inteligencie. Práce na týchto systémoch spôsobili revolúciu v spôsoboch našej komunikácie, práce ale aj zábavy.[347] Automatizácia kreatívneho procesu však spadá do oblasti právnej vedy ktorá je veľmi nová. Je preto na mieste otázka či súčasná legislatíva adekvátne rieši scenáre, v ktorých systémy umelej inteligencie vytvárajú inovácie a to, ako sa vysporiadať so vniknutými právnymi otázkami.[348]

Umelá inteligencia je však čosi ako nová elektrina, a len ťažko si predstaviť odvetvie, ktoré nebude transformované touto novou technológiou. Nie je to inak ani pri otázkach práva duševného vlastníctva. Umelá inteligencia narúša naše predstavy o tradičných konceptoch práva a vyvoláva vážne otázky, či tradičné formy jeho ochrany stále postačujú.[349] Už v roku 2017 Európsky parlament vyzval Európsku komisiu, aby podporila horizontálny a technologicky neutrálny prístup k duševnému vlastníctvu, ktorý by sa dal uplatniť v rôznych sektoroch, v ktorých by sa umelá inteligencia mohla uplatniť.[350]

Samotná myšlienka konvergencie umelej inteligencie a práv duševného vlastníctva nám môže pripadať možno nezvyčajnou, hlavne keď si predstavíme primárnu funkciu práv duševného vlastníctva v zabezpečení ekonomickej návratnosti pre tvorcu a zároveň zabránenie iným v neoprávnenom využívaní jeho obsahu.[351] Zdá sa, že umelá inteligencia ako obyčajný nástroj na výrobu takéhoto obsahu nemá žiadnu z týchto ekonomických potrieb, a

[347]ABBOTT, R: Artificial intelligence, big data and intellectual property: protecting computer-generated works in the united kingdom (2020) dostupné na: https://poseidon01.ssrn.com/delivery.php?ID=115120071099084002072068104098105081006040072045048074065007099125073126119111001029030006032052005113101003096104074089030103021043008005016104111022124028100115102028065038095084111110003081001075085112019016097068064086006113092080094067117099009067&EXT=pdf.

[348]KUKKONEN, TAIT, JOHNSON: When Innovation Creates: Additional Developments in Artificial Intelligence at the U.S. Patent and Trademark Office (2019) Dostupné na: https://www.jonesday.com/en/insights/2019/11/when-innovation-creates.

[349]JEFFRIES, A. TAIT, E.: Protecting Artificial Intelligence IP: Patents, Trade Secrets, or Copyrights? 2018 Dostupné na: https://www.jonesday.com/en/insights/2018/01/protecting-artificial-intelligence-ip-patents-trad.

[350]European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on 7 Robotics (2015/2103(INL)).

[351]European Commission- Intellectual property, Dostupné na: https://ec.europa.eu/growth/industry/policy/intellectual-property_en.

preto by otázky poskytnutia ochrany podľa práva duševného vlastníctva nemali byť ani relevantné. Aj napriek tomu sa tieto otázky stali v spoločnosti veľmi relevantné a diskutované, a preto je nevyhnutné riešiť prepojenie medzi umelou inteligenciou a duševným vlastníctvom. V predkladanej kapitole sa bližšie zameriavame na jednotlivé prieniky práva duševného vlastníctva a umelej inteligencie a s tým súvisiace otázky.

6.2 Duševné vlastníctvo ako výsledok tvorivej duševnej činnosti

Na úvod možno konštatovať, že existuje niekoľko definícií duševného vlastníctva, avšak vo svetle zovšeobecnenej definície ide predovšetkým o majetok nehmotnej povahy, ktorý vznikol ako výsledok **tvorivého myslenia** alebo **tvorivej duševnej činnosti**.[352] Tieto výsledky sú predovšetkým rôzne vynálezy či literárne alebo umelecké diela. Z hľadiska systematizácie práva duševného vlastníctva možno *tvorivú duševnú činnosť* vnímať najmä v kontexte autorských práv a priemyselných práv na výsledky tvorivej duševnej činnosti, kam patrí napríklad patentové právo, dizajnové právo či právo úžitkových vzorov.[353] Tento pojem nájdeme explicitne uvedený aj v našich právnych predpisoch, ako napríklad v ustanovení § 3 ods. 1 zákona č. 185/2015 Z. z. Autorský zákon v znení neskorších predpisov (ďalej ako „autorský zákon"), ktorý hovorí, že: „*Predmetom autorského práva je dielo z oblasti literatúry, umenia alebo vedy, ktoré je jedinečným výsledkom* ***tvorivej duševnej činnosti autora*** *vnímateľným zmyslami, bez ohľadu na jeho podobu, obsah, kvalitu, účel, formu jeho vyjadrenia alebo mieru jeho dokončenia*."[354] Z uvedeného vyplýva, že jedinečný výsledok tvorivej duševnej činnosti, je jedným z pojmových znakov, ktoré musí dielo napĺňať, aby mohlo požívať ochranu v zmysle autorského zákona. Rovnako tak možno spomenúť ustanovenie § 10 ods. 2 zákona č. 435/2001 Z. z. o patentoch, dodatkových ochranných osvedčeniach a o zmene a doplnení niektorých zákonov (ďalej ako „patentový zákon"), ktorý za pôvodcu vynálezu označuje toho, ***kto vytvoril vynález vlastnou tvorivou činnosťou.***[355]

Je zrejmé, že výsledky duševnej činnosti odrážajú ľudskú osobnosť a individualitu.[356] Aj keď tvorivé myslenie a invencia zostávajú v zásade ľudskými funkciami, čoraz viac

[352] VOJČÍK, P. a kol.: Právo duševného vlastníctva. 2. upravené vydanie. Aleš Čeněk, 2014, s. 27.
[353] ADAMOVÁ, Z.: Právo duševného vlastníctva. Bratislava: TINCT, 2020, s. 12.
[354] § 3 ods. 1 zákon č. 185/2015 Z. z. Autorský zákon.
[355] § 10 ods. 2 zákon č. 435/2001 Z. z. Patentový zákon.
[356] MINGALEVA, Z. MIRSKIKH, I:. Psychological Aspects of Intellectual Property Protection, Procedia - Social and Behavioral Sciences, Vol. 190, 2015, s. 220-226.

systémov je schopných tieto funkcie veľmi presvedčivo imitovať.[357] Rastúca sofistikovanosť nových technológii a prístupnosť k otvoreným údajom,[358] však umožnila transformáciu umelej inteligencie na jednu z hlavných prispievateľov do kreatívnej a inovačnej sféry.[359] V nasledujúcej podkapitole bližšie vysvetľujeme vnímanie tvorivej duševnej činnosti v kontexte umelej inteligencie.

6.2.1 Tvorivá „duševná" činnosť

Pred príchodom výpočtových technológií, algoritmov a inteligentných strojov a myšlienkami spojené s nimi, teoretizovali o povahe mysle iba filozofovia. Dnes však už snáď nikto z nás neverí tomu, že inteligencia by mala závisieť od niečoho nevysvetliteľného alebo nadprirodzeného. Moderný postoj k tejto problematike je skôr materialistický, prinajmenšom v tom užšom slova zmysle, ktorý predpokladá, že vhodne vybraná a usporiadaná hmota[360] bude pre existenciu inteligencie dostatočná. Tieto myšlienky sa stali v ostatných rokoch opäť veľmi populárnymi, a to najmä kvôli rozmachu témy umelej inteligencie. Otázky filozofie mysle s prihliadnutím na umelú inteligenciu, ako aj moderná podoba problematiky *„telo-mysel"* sa však už nesústreďujú primárne na možnosť existencie mysle bez fyzickej hmoty. Dnes sa tieto úvahy posúvajú smerom k otázkam, či môžeme umelú inteligenciu považovať skutočne za mysliaci stroj, a ak nie, či budeme môcť niekedy takého niečo skutočne vytvoriť.

Dodnes pretrváva značná nejasnosť aj v samotnej teórii o tom, ako objektívne skúmať myseľ a s tým spojené fenomény.[361] A keďže zatiaľ empirický výskum nedokázal jednoznačne odpovedať na všetky naše otázky týkajúce sa tejto problematiky, je potrebné uznať, že prinajmenšom čiastočne, tieto otázky spadajú do oblasti právnej filozofie, a to aj napriek

[357]DAVIS, R: Intellectual Property and software: The assumptions are broken, 1991 in World Intellectual Property Organization, WIPO Worldwide Symposium on the Intellectual Property Aspects of Artificial Intelligence, Stanford University.
[358]Viac o otvorených údajoch pozri. ANDRAŠKO, J. MESARČÍK, M.: Právne aspekty otvorených údajov 2020 C. H. Beck SK, ISBN 9788089603794.
[359]GÖNENC a kol.: Questions of Intellectual Property in the Artificial Intelligence Realm 2017 Dostupné na: https://www.gurkaynak.av.tr/docs/8b791-rlj-september-october-2017-.pdf.
[360]V tomto prípade ide o hmotu, v tom najširšom ponímaní. Jednotlivé teórie sa líšia v čom vnímajú základ inteligencie, avšak vždy pôjde o preukázateľnú fyzickú hmotu.
[361]LEEFMANN, J. HILDT, E:. The Human Sciences after the Decade of the Brain, 2017 https://www.sciencedirect.com/topics/neuroscience/philosophy-of-mind.

tomu, že otázky filozofie sa zvyčajne zameriavajú skôr na to, čo je „možné alebo potrebné" na rozdiel od toho, čo je preukázateľné (t.j. prípad empirického výskumu.)

6.2.2. Problematika „myseľ-telo"

Ústredným problémom filozofie mysle v kontexte umelej inteligencie je problematika vzťahu mysle k telu s výzvou vysvetliť, ako údajne nehmotná myseľ môže ovplyvniť hmotné telo a ako tento fenomén preniesť do strojom poháňanej podoby.[362]

Jedným z možných pohľadov na tento problém je prostredníctvom dualizmu, v ktorom sú myseľ a telo kategoricky oddelené jeden od druhého. Túto problematiku sformuloval ako prvý René Descartes, otec modernej filozofie, ktorý považoval svoje bezprostredné vedomé myšlienky za základ všetkých ostatných znalostí.[363] Jeho pohľad, v ktorom sú myseľ a telo kategoricky oddelené jeden od druhého, sa začal nazývať kartezianizmom.[364] Descartes obhajoval formu dualizmu, pre ktorú sú myseľ a telo vzájomne vylučujúcimi sa kategóriami.[365] Podľa tohto chápania pre myseľ je charakteristické, že vedomie nemá tvar a nepozostáva z fyzickej hmoty, na rozdiel od mozgu. Je preto možné povedať, že naše telá sú určite v čase a priestore ale naša myseľ nie je.[366]

Descartesov prístup mal obrovský vplyv až do dvadsiateho storočia, keď vývoj počítačov začal pútať fantáziu tých, ktorí hľadali vedeckejšiu a menej subjektívnu koncepciu

[362]Existujú však aj takí (najmä Ludwig Wittgenstein a jeho nasledovníci) ktorí tento problém odmietajú ako iluzórny, ktorý vznikol čisto preto, že mentálny a biologický slovník je nekompatibilný, a takéto problémy vznikajú, ak sa slovná zásoba používa v nesprávnych kontextoch.

[363]„Cogito, ergo sum".

[364] Pre problematiku pozri bližšie KENNY, A.: The rise of modern philosophy, Oxford University Press, 2006, s 40-68s.

[365]Tento druh dualizmu je možno najlepšie zachytený v listoch z mája 1643, ktoré si vymenili Descartes a princezná Alžbeta Falcká. V liste sa princezná pýta Descarta na otázku, *„ako môže ľudská duša určiť pohyb zvieracích duchov v tele, aby vykonávali vedomé činy."* Princezná dodáva, že: *„Zdá sa, že pohyb vždy prichádza z poháňaného pohybujúceho sa tela – teda závisí od druhu impulzu, ktorý dostáva, čo uvádza telo do pohybu."* Ďalej princezná dodáva, že*: „Prvé dve podmienky zahŕňajú kontakt a tretia zahŕňa to, že pohyb má akési predĺženie, ale existenciu takéhoto predĺženia úplne vylúčite zo svojej predstavy o duši a preto kontakt sa mi zdá nezlučiteľný s tým, že vec (myseľ) nie je nehmotná*". (Dodávame, že pre pochopenie týchto listov potrebujeme vedieť, že keď sa spomína pojem „zvierací duch", pre Descartesa neboli zvieracími duchmi duchovia v zmysle zjavení, ale súčasť teórie, ktorá tvrdila, že svaly sa pohybovali vzduchom. Pojem „Zviera" tu tiež neznamená synonymum nejakej šelmy, ale je prídavné meno odvodené od gréckeho slova „anima" (duša)).

[366] SCHMALTZ, T.: Descartes and Malebranche on Mind and Mind-Body Union, 1992. (Ang.) „when a man is considered in himself as a whole, we say of course that he himself is an Ens per se, and not per accidens; because the union, by which a human body and soul[366] are joined to each other, is not in itself accidental, but essential, since without it a man is not a man" (AT III 508/K 130). Dostupné na: https://www.jstor.org/stable/pdf/2185536.pdf?casa_token=OG4vMungoooAAAAA:_15IYZB8bDRbkpXnoS1XAJFlwnZI6wymCfMtegdAbBl3pC4ttpb57crRbon9XAmUe-8IlrTtI4-hT2taEkgPlUO2occm_xmmmk86p33UfNts7-wf7AsV.

podstaty myslenia. V týchto myšlienkach pokračoval aj Alan Turing.[367] Turing tvrdil, že kombináciou podstaty mysle a konceptu počítačového programu vo forme strojov, by v zásade myseľ mohla mať mnoho fyzických foriem aj v skutočnom svete. Jeho práca tak predstavila to, čo sa stalo známe ako výpočtová koncepcia mysle, ktorá premieňa karteziánske vymedzenie strojov na „bezduché".[368] Analogicky, takéto chápanie „inteligencie" podporuje záver, že vhodne naprogramované a správne fungujúce Turingove stroje by sa mohli kvalifikovať ako stroje s vedomím, alebo so slovami Johna McCarthyho, ako „umelá inteligencia".[369]

6.2.3. „Vlastná" tvorivá „duševná" činnosť

Tradičné otázky filozofie mysle sa zameriavajú najmä na štyri hlavné javy[370]: I)Vedomie,[371] II)Racionalita,[372] III)Intencionalita,[373] IV)Slobodná vôľa.[374] Ako sa však tieto fenomény prejavujú pri strojoch na báze umelej inteligencie?

---

[367]Turingové najdôležitejšie výskumy sa týkali obmedzení dôkazov v matematike, kde navrhoval, aby hranice počitateľnosti matematických problémov, ktorých riešenia bolo možné získať na základe konečných aplikácií logických pravidiel boli rovnaké ako tie, ktoré je možné vyriešiť pomocou špecifického druhu strojov na riešenie problémov. FLORIDI, L.: The Blackwell Guide to the Philosophy of Computing and Information, 2003 Dostupné na: https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/2/334/files/2017/05/The-Blackwell-Guide-to-the-Philosophy-of-Computing-and-Information-1by6afj.pdf.

[368]COPELAND, J.: The Essential Turing Oxford University Press, 2004 Dostupné na: https://books.google.sk/books?hl=sk&lr=&id=7sEbjfnVWREC&oi=fnd&pg=PR7&dq=Turing+artificial+intelligence&ots=Ekq7_NMe4y&sig=MGeu4NmDQyPOxEU52CfS9xS6sLE&redir_esc=y#v=onepage&q&f=false.

[369]FLORIDI, L.: The Blackwell Guide to the Philosophy of Computing and Information John Wiley & Sons; 1 edition (October 17, 2003) Dostupné na: https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/2/334/files/2017/05/The-Blackwell-Guide-to-the-Philosophy-of-Computing-and-Info"rmation-1by6afj.pdf.

[370]REY, G.: Philosophy Of Mind And Empirical Psychology *The Stanford Encyclopedia of Philosophy* 2017 Dostupné na: https://www.britannica.com/topic/philosophy-of-mind#ref283940

[371]Samotný pojem vedomie v kontextoch filozofie mysle sa používa rôznymi spôsobmi, ktoré je potrebné rozlišovať. V tej najužšej podobe sa používa ako synonymum ľudskej duševnej činnosti, a v tej najširšej sa to používa ako označenie stavu „byť pri vedomí". Bolo by prirodzené si myslieť, že všetky mentálne javy sú aj vedomé, avšak v skutočnosti tomu tak nie je. Existencia takzvaných podvedomých duševných stavov bola skúmaná už od čias starovekého Grécka. Odstupom času sa tiež ukázalo, že je prekvapivo ťažké tvrdiť niečo o vedomí, čo by nebolo celkom jednoducho spochybniteľné. Porovnaj LEEFMANN, J. - HILDT, E.: The Human Sciences after the Decade of the Brain, 2017 https://www.sciencedirect.com/topics/neuroscience/philosophy-of-mind.

[372]Koncept racionality by sme mohli najefektívnejšie priblížiť pomocou systematického skúmania problematiky poznania od Kanta, kde oproti empirizmu, ktorý uznáva, že skúsenosť je zdrojom hodnoverného poznania. Kant tvrdí, že: *„poznanie nevyhnutného a všeobecného poskytujú len apriori súdy, ktoré sú človeku dané pred skúsenosťou. Zdrojom týchto súdov nie je skúsenosť, ale transcendentálna dedukcia*", ktorú Kant nazýva čistým rozumom. Porovnaj CHOVANCOVA, J. - VALENT, T.: Filozofia pre právnikov, Bratislava: Univerzita komenského, Pravnická fakulta, 2012, s 121.

[373]Searle vysvetľuje, že: *„Intencionalita je príznakom určitých mentálnych stavov a udalostí, spočívajúcich v ich zameraní na niečo, v tom, že sú o niečom, alebo že reprezentujú iné entity a stavy vecí."* Tento pojem, od svojho zavedenia do filozofie Franzom Brentanom, kto ešte vtedy nazval tento jav *„úmyslom duševných stavov*", sa používa na označenie zatiaľ nevysvetleného, ktoré leží na rozhraní mysle a jazyka. Porovnaj SEARLE, J.: Intentionality: An Essay in the Philosophy of Mind, Cambridge University Press, 1983.

[374]Dnes v modernej psychológii už nie je otázna existencia vôle ako špecifickej kvality u človeka. Stále však zostáva dôležitá a zatiaľ nevyriešená otázka, či je táto vôľa slobodná. Alebo by sme sa na túto problematiku mali pozerať skôr deterministicky? Slobodnú vôľu definujeme ako možnosť voľby medzi alternatívami alebo preferovane aj ako konanie, ktoré

V klasickom ponímaní, v zariadeniach na báze umelej inteligencie sa inteligencia a intencionalita často vníma ako technická vlastnosť programu, ktorá je výsledkom algoritmov. Nemožno však zabúdať, že množstvo teórií stále trvá na tom, že vedomie a intencionalita je ontologickým vlastníctvom privilegovanej ľudskej existencie, a preto sa tento fenomén nemôže vzťahovať na stroje.[375] Za dôležité považujeme spomenúť aj myšlienku, s ktorou prišli Zhu a Harrel, ktorí argumentujú, že aj keď možno nevieme presne pochopiť, ako je autonómny robotický vysávač navrhnutý a skonštruovaný, môžeme do určitej miery predpovedať jeho správanie (výber miesta, nasledovný smer atď.).[376] Na základe tejto skutočnosti tu zjavná zámernosť systému je, aj keď samozrejme takúto zámernosť budeme len ťažko prirovnávať k intencionalite človeka.

Pri otázke racionality sa dostávame do sféry, v ktorej by umelá inteligencia mala mať navrch v porovnaní s ľudským rozumom. Fenomén, ktorý nazývame „dokonalá racionalita" vychádza z premisy, že sa dosiahne maximálny očakávaný úspech vzhľadom na dostupné možnosti.[377] Pretrvávajú však názory, že skutočne perfektná racionalita neexistuje. Totiž o akomkoľvek zariadení hovoríme, spracovanie informácií a výber akcií si vyžadujú čas, a preto akékoľvek ich rozhodovanie bude koniec koncov stále suboptimálne.[378] Aspektom dokonalej racionality, ktorá doteraz chýba, je vývoj vhodného súboru techník na špecifikáciu užitočnosti funkcií. Zatiaľ nemáme spôsob, ako špecifikovať „užitočnosť", hoci tieto otázky boli skúmané pomerne často, dodnes nemáme uspokojivé pochopenie vzťahu hodnoty a užitočnosti pri týchto systémoch.

Pokiaľ ide o systémy, ktoré dnes máme k dispozícii, je celkom jednoznačné, že nemôžeme hovoriť o slobodnej vôli týchto strojov. Tieto stroje sú viazané inštrukciami, ktoré

---

je nezávislé od prírodných, sociálnych alebo božských obmedzení. Problémom slobodnej vôle však ostáva otázka: *„Ako môžu existovať skutočne slobodné rozhodnutia vo svete, v ktorom všetky udalosti, (aspoň teda na makroúrovni), majú zrejmé príčinne súvislosti a podmienky? Zdá sa, že každá udalosť na tejto úrovni je určená príčinami, ktoré jej predchádzali. Prečo by potom činy vykonávané pri zjavnom ľudskom vedomí slobody mali byť výnimkou?*" Veríme však, aby sme mohli racionálne fungovať v tomto našom svete, musíme si aspoň myslieť, že si niekedy môžeme slobodne zvoliť svoje konanie. Ak by sme to popreli, vytvoril by to rozpor, bez vyriešenia ktorého stratí filozofia svoju racionlnu súdržnosť. Porovnaj FREDE, M: A Free Will: Origins of the Notion in Ancient Thought, University of California Press; Edición: First, 2011.

[375]Väčšina týchto téz sa odvoláva na „Argument čínskej izby" od Johna Searleho, Pozri bližšie: SEARLE, J: The Chinese Room, 1980 In Minds, Brains, and Programs Dostupné na: https://rintintin.colorado.edu/~vancecd/phil201/Searle.ąąpdf.

[376]ZHU, J. HARREL, D.: System Intentionality and the Artificial Intelligence Hermeneutic Network: the Role of Intentional Vocabulary 2009 Dostupné na: https://escholarship.org/uc/item/3rd2s695#main.

[377]V našom prípade dostupné informácie, treba tiež spomenúť, že tento spôsob sa používa aj na stanovenie hornej hranice výkonu akéhokoľvek systému.

[378]RUSSEL, S: Rationality and Intelligence, 1997, Dostupné na: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.48.9785&rep=rep1&type=pdf.

sú implantované, a ktoré musia byť nasledované. Treba si však uvedomiť, že autonómne systémy nie sú naprogramované iba na vykonávanie určitých činností, ale aj na to, aby sa naučili vykonávať určité činnosti sami. Inými slovami, podstatou autonómnych systémov nie je iba schopnosť autonómne existovať a fungovať, ale aj tvorba vlastného kódu nezávisle od svojho autora. Nesmierne komplikované však ostáva problematika takzvaných *„black box artificial intelligence*".[379] Tento pojem označuje stroje na báze umelej inteligencie, ktoré sú často založené na strojovom učení a otvorených údajoch, u ktorých nemáme možnosť zistiť a odôvodniť prečo rozhodli spôsobom, akým rozhodli.[380]

6.3 Prienik duševného vlastníctva a umelej inteligencie

V súčasnosti sa bežne stretávame s tým, že systémy umelej inteligencie vytvárajú obrazy, píšu príbehy či tvoria hudbu a dokonca sa podieľajú na navrhovaní rôznych vynálezov, čo však vyvoláva otázky o tradičnom chápaní vynálezcovstva a autorstva. Skutočnosť, že systémy umelej inteligencie dokážu vytvárať kreatívne výstupy však nie je jediný spôsob ako možno uvažovať o práve duševného vlastníctva v kontexte umelej inteligencie. Aj napriek tomu, že umelá inteligencia dokáže byť pre spoločnosť veľmi užitočná a nápomocná, jej využívanie, a to najmä v oblasti práva duševného vlastníctva, môže vzbudzovať rôzne obavy. Ako príklad možno uviesť, že algoritmy umelej inteligencie vyžadujú na svoje správne a efektívne fungovanie množstvo údajov, čo môže zahŕňať aj využívanie materiálu, ktorý je chránený autorským právom. Pri tejto činnosti tak môže dochádzať k porušovaniu autorských práv a tzv. *fair use*. Na druhej strane však systémy umelej inteligencie môžu byť nápomocné pri efektívnejšom odhaľovaní porušovania práv duševného vlastníctva.[381] Všetky z uvedených prienikov predstavujú pre právne prostredie dôležité výzvy, s ktorými sa budú musieť zákonodarcovia a spoločnosť ako taká vysporiadať.

[379]Pre problematiku pozri bližšie: PASQUALE, F: "The black box society: The secret algorithms that control money and information", Harvard University Press, 2015.
[380]ADADI, A – BERRADA, M.: Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence 2018 Dostupné na: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8466590.
[381] Congressional Research Service: Generative Artificial Intelligence and Copyright Law, s. 3 Dostupné na: https://crsreports.congress.gov/product/pdf/LSB/LSB10922.

6.3.1 Umelá inteligencia ako pôvodca vynálezu

Ako sme už viackrát poznamenali, systémy na báze umelej inteligencie síce napodobňujú ľudskú inteligenciu, chýba im však subjektivita z právneho pohľadu, ktorá by im umožnila mať práva a povinnosti v konvenčnom zmysle. Subjekt iný ako človek, vrátane strojov a zvierat, sa podľa zaužívanej súdnej praxe[382] nepovažuje za vynálezcu alebo tvorcu v zmysle príslušných právny predpisov.[383] Z tohto dôvodu zostáva otázka vlastníctva a autorstva, pokiaľ ide o diela vytvorené autonómnymi strojmi, stále nejasná. A to aj navzdory tomu, že žiadna domáca, medzinárodná ani európska legislatíva výslovne nezakazuje ochranu takýchto práv, avšak zriedka sú takéto výsledky práce explicitne chránené.[384] Najväčšou prekážkou sa však zdá byť právnická terminológia, ktorá je používaná v právnych predpisoch, týkajúca sa ako autorských práv, tak aj patentov. V kontinentálnom ako aj v angloamerickom právnom priestore je systematicky používané vymedzenie autora a vynálezcu prostredníctvom „fyzickej osoby". Predmetné vymedzenie následne neumožňuje, aby sa tieto výsledky prác priradili iným entitám, ktoré nemožno subsumovať pod „fyzické osoby" v tradičnom chápaní.[385]

V rokoch 2018 a 2019 boli podané paralelné prihlášky patentov na Úrad pre patenty a ochranné známky v USA (**United States Patent and Trademark Office** – ďalej ako USPTO), Európsky patentový úrad (**European Patent Office** – ďalej ako EPO) a Britský patentový úrad (**UK Intellectual Property Office** – ďalej ako UKIPO) Dr. Stephenom Thalerom. Vo vyhláseniach o vynálezcoch patentov Dr. Thaler uviedol,[386] že vynálezcom bol systém umelej inteligencie s názvom DABUS.

---

[382]Väčšina rozhodnutí sa odvoláva na Rozhodnutie Naruto v. Slater, zo dňa 23. 4. 2018, sp. zn. No. 16-15469, Dodnes známe ako: „The Monkey Selfie case". Dostupné z: https://law.justia.com/cases/federal/appellate-courts/ca9/16-15469/16-15469-2018-04-23.html.

[383]MORIGGI, Andrea. The role of intellectual property in the intelligence explosion. *4ipCouncil*, [online]. 2017, [cit.2021-08-15]. Dostupné z: https://www.4ipcouncil.com/application/files/9615/1638/1031/The_Role_of_Intellectual_Property_in_the_Intelligence_Explosion.pdf .

[384]NURTON, James. The IP behind the AI boom [online]. 2019, [cit. 2023-06-15]. Dostupné z: https://www.wipo.int/wipo_magazine/en/2019/01/article_0001.html.

[385]ABBOTT, Ryan. Artificial Intelligence, Big Data and Intellectual Property: Protecting Computer-Generated Works in the United Kingdom. In: APLIN, Tanya (ed.). *Research Handbook on Intellectual Property and Digital Technologies.* Edward Elgar Publishing Ltd, Forthcoming, [online] 2017, [cit. 2023-06-15]. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064213.

[386]Dr. Thaler tiež uviedol, že vynález vytvorený strojom by nemal byť vo vlastníctve majiteľa stroja. Hovoril tiež, že DABUS identifikoval novosť svojho vlastného nápadu skôr, ako to urobila fyzická osoba. Porovnaj, IRELAND, I. LOHR, J.: 'DABUS':

V zmysle vyššie uvedených argumentov, ktoré vo svojich rozhodnutiach použili všetky tri úrady, boli žiadosti zamietnuté. Hlavným dôvodom týchto zamietavých rozhodnutí bola skutočnosť, že DABUS je iba stroj, zatiaľ čo právne rámce uplatňované EPOm, UKIPOm a USPTOm vyžadujú, aby vynálezca bola výlučne fyzická osoba. Navyše EPO vo svojom rozhodnutí[387] taktiež poukázal na skutočnosť, že systémy alebo stroje UI nemajú právnu subjektivitu porovnateľnú s fyzickými osobami alebo právnickými osobami a neexistuje žiadna legislatíva ani judikatúra, ktorá by takúto právnu fikciu podporovala.

Je však celkom zrejmé, že v týchto prípadoch je otázka skôr zameraná na problém zaužívaných právnych zásad, než ako iba o samotné patentové prihlášky. Napokon všetky tri úrady by s radosťou akceptovali prihlášku, ak by pri políčku pre vynálezcov Dr. Thaler nahradil stroj DABUS vlastným menom. V skutočnosti, v čase, keď potenciál technológií na báze UI stále exponenciálne rastie, je potrebné si položiť otázku, či súčasný právny rámec zodpovedá danému účelu. Túto tézu uznal vo svojom rozhodnutí aj UKIPO, keď dospel k záveru, že *„súčasný systém sa nestará o také vynálezy UI a nikdy sa nepredpokladalo, že by to bolo potrebné, ale doba sa zmenila a technológia sa posunula ďalej. Je správne, že sa o tom diskutuje širšie a že akékoľvek zmeny zákona sa majú posudzovať v kontexte takejto debaty a nie byť svojvoľne zavádzané do existujúcich právnych predpisov."*[388]

Pre úplnosť informácií uveďme, že Dr. Thaler podal odvolanie proti rozhodnutiu britského patentového úradu. Najvyšší súd Spojeného kráľovstva rozsudkom zo dňa 20. december. 2023 potvrdil rozhodnutie a stanovisko patentového úradu podľa ktorého systémy na báze UI nemôžu byť uznané za pôvodcov patentov. [389]

Rozhodnutia DABUS však vyvolávajú niekoľko dôležitých otázok o tom, ako zvládnuť „vynálezy", ktoré boli aspoň čiastočne vyvinuté alebo koncipované systémom UI. Možným riešením tejto otázky je stále postup podľa britského systému. Tento prístup je najlepšie vyjadrený v zákone o autorských právach, dizajnoch a patentoch v Spojenom kráľovstve

---

the AI topic that patent lawyers should be monitoring, 2020, [online]. [10-8-2021]. Dostupné na: https://www.managingip.com/article/b1n8q624s4vyv4/dabus-the-ai-topic-that-patent-lawyers-should-be-monitoring.

[387]Tlačová správa – EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor. [online]. 2020, [10-8-2021]. Dostupné z: https://www.epo.org/news-events/news/2020/20200128.html.

[388]Rozhodnutie patentového úradu zo dňa 04. 12. 2019, sp. Zn. BL O/741/19. Dostupné z: https://www.ipo.gov.uk/p-challenge-decision-results/074119.pdf.

[389]Rozhodnutie Najvyššieho súdu zo dňa 20. 12. 2023. Pozri bližšie: Dostupné na: https://www.supremecourt.uk/cases/docs/uksc-2021-0201-judgment.pdf.

(ďalej ako „CDPA"), ktorej paragraf 9 ods. 3 hovorí: *„V prípade literárneho, dramatického, hudobného alebo umeleckého diela, ktorá je generovaná počítačom, za autora sa považuje osoba, od ktorej sa prijímajú opatrenia potrebné na vytvorenie diela."* Ďalej paragraf 178 definuje počítačovo generované dielo ako také, ktoré *„je generované počítačom za okolností, že neexistuje žiadny ľudský autor diela."*[390] Myšlienkou takýchto ustanovení je vytvoriť výnimku z požiadavky fyzickej osoby ako autora pri programe, ktorá je schopná vytvárať nové diela.

Avšak sporným aspektom toho prístupu ostáva určenie, ktorému exaktnému subjektu práva treba pripísať autorstvo týchto diel. A či je vôbec etické a správne pripísať zásluhy diela niekomu, kto to v skutočnosti nevytvoril. Zástancovia tejto teórie tvrdia,[391] že britský systém má jasné výhody, akými sú:

a) prínos istoty do neistej právnej oblasti;
b) existencia krajín, kde tento systém funguje už pomerne dlho bez väčších incidentov;
c) existencia rozhodovacej praxe súdov pri tomto systéme. Britský systém však nie je bez kritiky iných odborníkov,[392] ktorí tento systém nepovažujú za vhodný.

Ich kritika smeruje predovšetkým k tomu, že tento systém ponecháva značnú mieru neistoty o osobe, ktorej sa majú pripísať práva z práce týchto systémov. Ich druhou zásadnou námietkou je, že tento systém nerieši otázku originality. Neopomeňme ani problematickú otázku, či tento systém je vôbec v súlade s Acquis Európskej únie.

Voči predmetnej teórii sa sformovala alternatíva, ktorej hlavná téza v tejto oblasti je, že by sa mala zaviesť forma odlišnej ochrany udelením práv *sui generis*.

Existuje však silný odpor proti takejto myšlienke, keďže podľa viacerých predmetnej alternatíve chýbajú empirické dôkazy podporujúce potrebu vytvárania nových práv. Väčšina týchto autorov uprednostňuje vylúčenie prác vytvorených umelou inteligenciou z akejkoľvek

[390]Rozhodnutie Najvyššieho súdu Anglicka a Walesu (patentový súd), zo dňa 21. 9. 2020 sp. zn. EWHC 2412 Dostupné z: https://www.bailii.org/ew/cases/EWHC/Patents/2020/2412.html.
[391]GUADAMUZ, A. Artificial intelligence and copyright, [online]. 2017, [cit. 2023-06-15]. Dostupné z: https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html.
[392]Porovnaj LAUBER-RÖNSBER, A. HETMANK, S. The concept of authorship and inventorship under pressure: Does artificial intelligence shift paradigms?. *Journal of Intellectual Property Law & Practice*, Vol. 14, No 7, [online]. 2019, [cit. 2023-06-15]. Dostupné z: https://academic.oup.com/jiplp/article-abstract/14/7/570/5485772?redirectedFrom=fulltext.

ochrany, keďže udelenie autorstva systému umelej inteligencie, ktorá nikdy sama nepoužije žiadne zo svojich práv, sa koniec koncov stáva aj tak iba symbolickým gestom.[393] V spojitosti s touto témou Michaux upozorňuje na ťažkosti pri rozlišovaní diel vytvorených ľuďmi a strojmi.[394] Určenie a vymedzenie presných hraníc zohráva kľúčovú úlohu, keďže udelenie ochrany by mohlo viesť k podstatnému zvýšeniu počtu chránených diel a ku koncentrácii autorských práv v niekoľkých spoločnostiach. Michaux dodáva, že priznaná ochrana by mohla viesť k inštrumentalizácii pojmu práca a k nadmernému speňaženiu prístupu k dielam.

Podobné obavy vyjadrila aj Anne Lauber-Rönsberg, ktorá zdôrazňuje, že integrácia diel vytvorených umelou inteligenciou do režimu autorských práv môže vážne narušiť základné koncepty právnych predpisov, na ktorých je postavený celý právny systém.[395] Taktiež nemožno zabudnúť na silný spoločenský dopad priznania takejto ochrany dielam vytvorenej systémom umelej inteligencie. Priznanie autorských práv umelej inteligencii môže viesť k zničeniu stimulov pre ľudských tvorcov, ak sa so strojovo generovanými dielami bude zaobchádzať na rovnakej úrovni.[396]

Napriek možným následkom však viacerí odborníci veria, že súčasný tradičný prístup k vynálezom, vytvorených systémom umelej inteligencie, ktorý je zameraný na človeka, sa v skutočnosti javí ako nedostatočný na reguláciu technologického vývoja v automatizovaných systémoch umelej inteligencie.[397] Súčasné právne poriadky uprednostňujú ochranu ľudských výtvorov pred dielami UI, bude preto potrebné zvážiť nastolenie pravidiel fungovania medzi ľudskou tvorivosťou a rozvojom technológií umelej inteligencie. Jedno je však isté, prípad DABUS je iba začiatkom medzinárodnej diskusie na túto tému.

---

[393]GONENC, G. Questions of Intellectual Property in the Artificial Intelligence Realm. *The Robotics Law Journal*, Volume 3, No. 2, s.9-11 [online]. 2017, [cit. 2023-06-15]. Dostupné z: https://www.gurkaynak.av.tr/docs/8b791-rlj-september-october-2017-.pdf.

[394]MICHAUX, B. Singularité technologique, singularité humaine et droit d'auteur. In: *Laws, Norms and Freedoms in Cyberspace/ Droit, normes et libertés dans le cybermonde: liber amicorum Yves Poullet*, 2018, s 401- 416.

[395]LAUBER-RÖNSBER, A HETMANK, S. The concept of authorship and inventorship under pressure: Does artificial intelligence shift paradigms?. *Journal of Intellectual Property Law & Practice*, Vol. 14, No 7, [online]. 2019, [cit. 2023-06-15]. Dostupné z: https://academic.oup.com/jiplp/article-abstract/14/7/570/5485772?redirectedFrom=fulltext.

[396]TRIPATHI, S . GHATAK, C. Artificial Intelligence and Intellectual Property Law. *Christ University Law Journal*, 2018 s. 83-97.

[397]TUNG, Jonathan. Who Owns the Creation of an Artificial Intelligence? *Technologist*, [online]. 2016, [cit. 2023-06-15]. Dostupné z: https://is.gd/law_who_owns_AI_creations.

6.3.2 Umelá inteligencia ako autor

Vzhľadom na to, že systémy umelej inteligencie sa vyvíjajú veľmi rýchlo, dokážu vytvárať kreatívne diela, ktoré je čoraz zložitejšie odlíšiť od výtvorov človeka. Na mieste boli preto otázky týkajúce sa autorstva k dielam vytvoreným prostredníctvom umelej inteligencie. Prvým pozitívnym rozsudkom pre zariadenia na báze umelej inteligencie je rozhodnutie, ktoré vydal okresný súd Shenzhen Nanshan (ďalej len „*súd Nanshan*") pre spor o porušenie autorských práv medzi spoločnosťami Shenzhen Tencent Computer System Co., Ltd. (ďalej len „*Tencent*") a Shanghai Yingxun Technology Co., Ltd. (ďalej len „*Yingxun*"). Tento rozsudok priniesol nové úvahy o ochrane autorských práv k dielam vytvoreným umelou inteligenciou.[398] Súd sa priklonil k názoru, že artikulácia a štruktúra článku mali „určitú originalitu" a spĺňali zákonné požiadavky na to, aby ich bolo možné klasifikovať ako písomné dielo, čím sa kvalifikoval na ochranu autorských práv.[399]

Za posledných päť rokov čínsky technologický titan Tencent[400] publikoval obsah produkovaný automatizovaným softvérom s názvom Dreamwriter so zameraním na obchodné a finančné správy. Dreamwriter je inteligentný asistenčný systém pri písaní, založený na veľkých dátach a algoritmoch a bol vyvinutý spoločnosťou Tencent, ktorá použila Dreamwriter na vytvorenie asi 300 000 článkov ročne. Jedným z článkov vytvorených programom Dreamwriter bola finančná „finance-stock" s podpisom, že článok napísal Dreamwriter. Yingxun bez súhlasu Tencentu v deň jeho uverejnenia článok na svojom webe znova zverejnila. Na základe toho Tencent podal žalobu z dôvodu porušenia autorských práv a nekalej súťaže.

V prípade *Shenzhen Tencent v. Shanghai Yingxun* súd upriamil svoju pozornosť na dve právne otázky týkajúce sa UI z hľadiska autorských práv. **Po prvé**, či diela súvisiace s UI môžu byť prácami podľa autorského zákona, a **po druhé**, či spoločnosť Tencent je oprávnená byť

[398]K otázke dôležitosti ochrany autorských práv pri nových technológiách porovnaj ŽARSKÁ, P. The importance of copyright during the COVID-19 crisis. In *Bratislavské právnické fórum: Technologies in times of crisis: threat or opportunity to law*?. Bratislava: Právnická fakulta UK, 2021, s. 62-66.

[399]Chinese Court Backs Copyrights for AI-Created Work, [online]. 2020, [cit. 2023-06-15]. Dostupné z: https://www.chinajusticeobserver.com/a/chinese-court-backs-copyrights-for-ai-created-works.

[400]Tencent nie je jedinou spoločnosťou, ktorá publikuje publicistický obsah napísaný pomocou algoritmov. Associated Press (AP) využíva AI na reportáže o bejzbalovom pokrytí a príjmoch prostredníctvom partnerstva s Automated Insights. Organizácia Narrative Science so sídlom v Chicagu ponúka niečo podobné so špecifickým zameraním na business intelligence pre podnik alebo „dátové rozprávanie príbehov", ako to nazýva.

vlastníkom autorských práv. Pokiaľ ide o prvú otázku, súd Nanshan rozhodol, že článok spĺňal formálne požiadavky na literárne dielo. Podľa článku 2 autorského zákona „dielo" uvedené v autorskom zákone odkazuje na pôvodné diela literatúry, umenia a vied s intelektuálnymi výsledkami, ktoré je možné reprodukovať v hmatateľnej forme.[401]

Článok 3 hovorí, že za „vytvorenie diel" sa považuje akákoľvek intelektuálna činnosť, ktorá priamo súvisí s produkciou diel. Súd sa sústredil na to, či článok bol „dielom" Dreamwriteru podľa autorského zákona a či proces vytvorenia článku napĺňal požiadavku podľa článku 3. Súd došiel k záveru, že usporiadanie a výber údajov, nastavenie spúšťacích podmienok, ako aj výber šablóny a štýlu korpusu vývojového tímu Dreamwriter, sú intelektuálne činnosti, ktoré priamo súvisia s konkrétnym vyjadrením článku.

Na základe analýzy procesu generovania článku je jeho vyjadrenie určené individuálnym usporiadaním a výberom príslušných zamestnancov vývojového tímu Tencent. Inak stručne povedané aj finálna podoba článku, vytvoreného Dreamwriterom, sa chápe ako vytvorenie diela, a teda sú splnené základné požiadavky kladené autorským zákonom. Súd preto rozhodol, že príslušný článok je chránený čínskym autorským zákonom ako písomné dielo.[402]

Tento rozsudok bol prvým, ktorý výslovne potvrdil ochranu diel vytvorených UI a rozsudok uznal prínos UI v procese intelektuálnej práce. Aj tento prípad preukázal, že téma UI dnes už nie je len o potencionálnych scenároch v budúcnosti, ale o skutočných aplikačných problémoch, ktoré sa dejú práve teraz. Treba však spomenúť, že absencia adekvátnej právnej ochrany pri systémoch UI zabraňuje tomu, aby sa naplno uvoľnil ekonomický potenciál spojený s touto novou technológiou. Pritom „zastaraná" tradičná povaha súčasnej bežnej právnej úpravy duševného vlastníctva akosi nedokáže odzrkadliť terajšiu novú realitu, ktorú priniesli tieto systémy. Nemožno však poprieť skutočnosť, že poskytnutie ochrany duševného vlastníctva UI vedie paradoxne k spochybňovaniu samotných základov tejto právnej oblasti.[403] Ale to, či táto skutočnosť by sa mala interpretovať ako hrozba alebo príležitosť na

[401]Regulations for the Implementation of the Copyright Law of the People's Republic of China. Dostupné z: http://www.china.org.cn/english/DAT/214795.htm.

[402]Porovnaj s ZHOU, B. Artificial Intelligence and Copyright Protection --Judicial Practice in Chinese Courts. *Wipo*, [online]. 2019, [cit. 2023-06-15]. Dostupné z: https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/conversation_ip_ai/pdf/ms_china_1_en.pdf.

[403]Porovnaj MICHAUX, B. Singularité technologique, singularité humaine et droit d'auteur. In: *Laws, Norms and Freedoms in Cyberspace/ Droit, normes et libertés dans le cybermonde: liber amicorum Yves Poullet*. 2018, s. 401-416. alebo GYURÁSZ, Z. Duševné vlastníctvo vo svetle umelej inteligencie. *ComenIUS*, 2020, Roč. 5, č. 2, s. 6-14.

modernizáciu oblasti práva duševného vlastníctva, je stále nezodpovedanou otázkou, ktorá ešte bude najbližšie roky intenzívne rezonovať v mnohých odborných kruhoch.

6.4 Porušovanie a ochrana práv duševného vlastníctva v kontexte umelej inteligencie

Ako vyplýva z úvodu tejto podkapitoly, prienik medzi právom duševného vlastníctva a umelej inteligencie nájdeme vo viacerých smeroch. Vplyv umelej inteligencie na právo duševného vlastníctva možno rozšíriť aj na jeho ochranu, či prípadné porušovanie. Zatiaľ čo na jednej strane dokážu systémy umelej inteligencie vytvárať kreatívne výstupy vo forme textu, hudby či obrazov[404] na strane druhej dokážu prostredníctvom rôznych nástrojov detegovať porušenie už existujúcich práv duševného vlastníctva či inak zabezpečovať jeho ochranu. V súčasnosti už poznáme rôzne systémy umelej inteligencie, ktoré môžu byť nápomocné pri identifikácií porušovania napr. autorských práv, čím napomáhajú k proaktívnej ochrane práv duševného vlastníctva. Ako príklad možno spomenúť platformu YouTube, ktorá využíva umelú inteligenciu na identifikovanie a informovanie používateľov o prípadnom porušení práv duševného vlastníctva.[405] Rovnako tak systémy umelej inteligencie môžu byť nápomocné v rámci právnickej profesie, napríklad v prípade patentového práva, niektoré nástroje dokážu vyhľadávať už existujúce vynálezy, či určovať novosť nových vynálezov.[406]

Využívanie umelej inteligencie sa často spája aj s rizikom porušovania práv duševného vlastníctva. V súvislosti s týmto porušovaním možno spomenúť už samotné „trénovanie" systémov umelej inteligencie. S príchodom generatívnej umelej inteligencie,[407] ktorá funguje na základe strojového učenia a ktorá si na svoj tréning vyžaduje veľké množstvo dát,[408] sa začali v spoločnosti vynárať otázky týkajúce sa pôvodu týchto dát. Vznikli preto obavy, že tieto systémy sú trénované na vytváranie výstupov tak, že ťažia z už existujúcich diel

[404] Pozri bližšie: WIPO Magazine Umelá inteligencia a autorské práva Dostupné na: https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html.
[405] RATHIA, N. Protection and Infringement of IPR by Artificial Intelligence: A Double Edged Sword? Dostupné na: https://www.intepat.com/blog/protection-and-infringement-of-ipr-by-artificial-intelligence-a-double-edged-sword/.
[406] *Tamtiež.*
[407] Forma umelej inteligencie, ktorá dokáže vytvárať text, obrázky a rôzny obsah na základe údajov, na ktorých bola trénovaná. Napríklad chatbot ChatGPT od spoločnosti OpenAI.
[408] QUANG, J. Does Training Ai Violate Copyright Law?, Berkeley Tech. L.J. Vol. 36:1407, str.1407-1436 Dostupné na: https://btlj.org/wp-content/uploads/2023/02/0003-36-4Quang.pdf.

dostupných na internete ako sú napríklad fotografie, maľby, články či iné diela, ktoré sú chránené autorským právom. O to, či ide o oprávnené využívanie týchto diel sa začali zaujímať ako samotní autori resp. ich zástupcovia, tak aj príslušné organizácie. Autori literárnych či umeleckých diel po celom svete sa však začali domnievať, že veľké spoločnosti ako napr. OpenAI, neoprávnene kopírujú ich diela za účelom trénovania systémov umelej inteligencie, čím sa podľa nich dopúšťajú porušovania ich práv. Spoločnosti sa však bránia tým, že v tomto prípade k žiadnemu porušovaniu nedochádza. [409] V tomto kontexte je preto nutné sa zamyslieť nad tým, či pôvodné dielo môžu vývojári a prevádzkovatelia systémov AI použiť na trénovanie ich produktov bez súhlasu autora a bez primeranej kompenzácie.

V podmienkach Slovenskej republiky možno bez súhlasu autora nakladať s jeho dielom len v osobitých prípadoch ustanovených autorským zákonom.[410] Ide o tzv. výnimky a obmedzenia majetkových práv autora[411]. V tomto kontexte možno uvažovať o subsumovaní takého použitia pod výnimku upravenú v ustanovení § 51c autorského zákona, tzv. *výnimka použitia diela pri čerpaní údajov*, ktorá bola do slovenského právneho poriadku transponovaná smernicou Európskeho parlamentu a Rady (EÚ) 2019/790 zo 17. apríla 2019 o autorskom práve a právach súvisiacich s autorským právom na digitálnom jednotnom trhu a o zmene smerníc 96/9/ES a 2001/29/ES (ďalej ako „smernica o autorskom práve"). Z ustanovenia autorského zákona, ako aj zo znenia Článku 4 smernice o autorskom práve vyplýva, že výnimka sa vzťahuje na *rozmnožovanie a extrakciu zákonne dostupných diel a iných predmetov ochrany na účely vyťažovania textov a dát*.[412] Hoci medzi odborníkmi existuje zhoda, že za splnenia podmienok možno túto výnimku aplikovať aj na vyťažovanie textov a dát na účely strojového učenia, stretávame sa aj s názormi, že pri formulovaní tohto ustanovenia zákonodarca nemal na mysli tieto druhy použitia[413]. V zmysle uvedeného tak osoba, ktorá by bez súhlasu autora použila dielo *„vyhotovením rozmnoženiny pri čerpaní údajov, ak takéto*

[409]Pozri napríklad: https://www.reuters.com/technology/more-writers-sue-openai-copyright-infringement-over-ai-training-2023-09-11/ alebo https://www.theverge.com/2023/9/11/23869145/writers-sue-openai-chatgpt-copyright-claims

[410] Za predpokladu, že nejde napr. o tzv. Voľné dielo alebo nebolo autorom publikované licenciou Creative Commons (CC).

[411]V zmysle § 34 autorského zákona sú dovolené len v osobitných prípadoch ustanovených v tejto hlave a nakladanie s dielom podľa týchto ustanovení nesmie byť v rozpore s bežným využitím diela a nesmie neodôvodnene zasahovať do právom chránených záujmov autora. Nájdeme ich explicitne uvedené v § 37 – 57 autorského zákona.

[412]Článok 4 smernice o autorskom práve Dostupné na: https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:32019L0790&from=es.

[413]KELLER, P. A first look at the copyright relevant parts in the final AI Act compromise Dostupné na: https://copyrightblog.kluweriplaw.com/2023/12/11/a-first-look-at-the-copyright-relevant-parts-in-the-final-ai-act-compromise/.

použitie nie je výslovne vyhradené[414]" neporušovala autorské práva. Z tohto znenia vyplýva, že autori môžu uplatneniu tejto výnimky zabrániť a to tým, že sa proti takému nakladaniu s dielom vopred ohradia.[415] Môžeme však konštatovať, že uplatňovanie tejto výnimky je v súčasnosti pomerne diskutovanou témou. Situáciu však pravdepodobne vyrieši prijatie zákona o umelej inteligencii, ktorý priamym odkazom na túto smernicu potvrdzuje, že túto výnimku možno uplatniť aj na takéto použitie. V súvislosti s uvedeným tiež tento zákon ukladá povinnosť vývojárom generatívnej umelej inteligencie zverejňovať súhrny údajov chránených autorským právom, ktoré sa použili na tréning umelej inteligencie.[416] Obdobou týchto výnimiek a obmedzení je aj tzv. doktrína *fair use*[417], ktorá je zakotvená v zákone o autorských právach USA, a ktorá umožňuje v niektorých prípadoch využívať obmedzené časti diela aj bez súhlasu autora.[418] Práve na túto doktrínu sa odvolávajú spoločnosti, ktoré stoja za vývojom systémov umelej inteligencie. To či je takéto využívanie autorských diel bez súhlasu, či prípadne akejkoľvek primeranej kompenzácie autorom naozaj „*fair use*" sa stalo veľmi diskutovanou témou. Výklad tejto doktríny prináleží v USA súdom, a preto prvé súdne spory nenechali na seba dlho čakať.

Ako jeden z viacerých prípadov,[419] ktorý ašpiruje stať sa precedensom v tejto oblasti v USA, možno spomenúť žalobu denníka *The New York Times*, ktorý v závere roku 2023 zažaloval spoločnosti *OpenAI* a *Microsoft* za porušovanie autorských práv, ku ktorému malo dôjsť práve tým, že na trénovanie ich produktov ako je chatbot ChatGPT, boli použité viaceré ich články.[420] Niektorí právnici sa prikláňajú na stranu týchto spoločností, a majú za to, že na trénovanie modelov umelej inteligencie sa vzťahuje doktrína *fair use* a to najmä s odvolaním

[414]Článok 4 ods. 3 smernice o autorskom práve Dostupné na: https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:32019L0790&from=es.

[415]European Visual Artists (EVA) – Združenie audiovizuálnych umelcov, však nepovažuje možnosť autorom vyhradiť sa voči takémuto používaniu za dostatočnú ochranu, nakoľko často autori nemusia vedieť, že ich diela sú využívané na trénovanie systémov UI.

[416] KELLER, P. A first look at the copyright relevant parts in the final AI Act compromise Dostupné na: https://copyrightblog.kluweriplaw.com/2023/12/11/a-first-look-at-the-copyright-relevant-parts-in-the-final-ai-act-compromise/.

[417] QUANG, J. Does Training Ai Violate Copyright Law?, Berkeley Tech. L.J. Vol. 36:1407, str.1407-1436 Dostupne na: https://btlj.org/wp-content/uploads/2023/02/0003-36-4Quang.pdf

[418] *Tamtiež.*

[419]Napríklad pozri: SANCTON, J. vs. OpenAI a Microsoft Dostupné na: https://fingfx.thomsonreuters.com/gfx/legaldocs/zdvxrbbawvx/OPENAI%20COPYRIGHT%20LAWSUIT%20sanctoncomplaint.pdf.

[420]Pozri bližšie The New York Times vs OpenAI a Microsoft: Dostupné na: https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf.

na rozhodnutie vo veci *Authors Guild v. Google*.[421] Iní sa však domnievajú, že takéto použitie už presahuje rámec spravodlivého použitia a teda by tieto spoločnosti mali upustiť od takéhoto konania resp. by mali poskytnúť autorom kompenzáciu za takéto použitie. Môžeme predpokladať, že v blízkej budúcnosti budú spoločnosti, ktoré stoja za vývojom a prevádzkovaním takýchto modelov umelej inteligencie čeliť obdobným žalobám. Isté je však to, že rozhodnutie v tomto, a tomu podobných prípadoch, výrazne ovplyvní doterajšie fungovanie či prípadný rozmach týchto modelov umelej inteligencie a to nie len v USA.

[421] V roku 2004 spoločnosť Google začala digitalizovať niektoré knižné diela. Združenie autorov v USA zažalovali spoločnosť Google za porušovanie ich autorských práv skenovaním kníh do aplikácie Google books. Toto združenie tvrdilo, že Google takýmto skenovaním kníh do svojej knižnej iniciatívy porušil americký autorský zákon, pretože takéto masívne skenovanie týchto diel presiahlo aj hranice tzv. fair use. Príslušný súd však uznal argumenty spoločnosti Google, že doktrína tzv. fair use je naplnená tým, že Google books prináša významný verejný prospech. Bližšie pozri: Dostupné na: https://scholar.google.com/scholar_case?case=2220742578695593916&q=authors+guild+v+google&hl=en&as_sdt=20006

7. Kybernetická bezpečnosť a umelá inteligencia

7.1 Vzťah kybernetickej bezpečnosti a umelej inteligencie

Vzťah medzi umelou inteligenciou (*AI*) a kybernetickou bezpečnosťou je potrebné vnímať vo viacerých úrovniach, pričom doposiaľ sa vykryštalizovali nasledovné tri kategórie vzťahov:[422]

1. *Kybernetická bezpečnosť umelej inteligencie* – predmetom týchto vzťahov je ošetrenie nedostatočnej robustnosti a zraniteľnosti modelov a algoritmov umelej inteligencie.

2. *AI na podporu kybernetickej bezpečnosti* - AI je používaná ako nástroj/prostriedok na vytvorenie pokročilej kybernetickej bezpečnosti (napr. vývojom účinnejších bezpečnostných opatrení) a na uľahčenie úsilia orgánov činných v trestnom konaní a iných orgánov verejnej moci s cieľom lepšie reagovať na počítačovú kriminalitu.

3. *Škodlivé používanie AI na páchanie kybernetickej kriminality* – v tomto prípade ide o používanie AI na vytvorenie sofistikovanejších typov útokov.

7.2 Kybernetická bezpečnosť umelej inteligencie

Kybernetickú bezpečnosť AI možno vnímať v užšom a širšom zmysle.

V užšom zmysle, kybernetická bezpečnosť AI predstavuje tradičný rozsah, ktorý je určený na ochranu pred útokmi na dôvernosť, integritu a dostupnosť aktív (komponentov AI, súvisiacich údajov a procesov) v celom životnom cykle systému AI.

Širšie vnímanie kybernetickej bezpečnosti AI, zahŕňa a dopĺňa užšie chápanie kyberbezpečnosti o prvky dôveryhodnosti AI, ako sú kvalita údajov, vysledovateľnosť, transparentnosť, ľudský dohľad, presnosť a robustnosť.

[422] Porov. JUNKLEWITZ, H., et al.: *Cybersecurity of Artificial Intelligence in the AI Act*, Publications Office of the European Union, Luxembourg, 2023, str. 6 [online] [30.12.2023]. Dostupné na: doi:10.2760/271009, JRC134461. Rovnako, European Union Agency for Cybersecurity (ENISA). *Cybersecurity of AI and Standardisation (Report)*. March 2023. str. 10 [online] [30.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation/@@download/fullReport.

Kybernetická bezpečnosť a robustnosť sú kritické aspekty, ktoré treba brať do úvahy pri návrhu a vývoji systémov AI.

Rozdiel medzi kybernetickou bezpečnosťou a robustnosťou spočíva v ich zameraní. Zatiaľ čo kybernetická bezpečnosť sa zaoberá ochranou systému pred vonkajšími hrozbami, ako sú útoky hackerov alebo úniky dát, robustnosť sa zameriava na schopnosť systému fungovať spoľahlivo aj pri nečakaných situáciách, ako sú chyby v dátach alebo poruchy hardvéru.

Nariadenie o umelej inteligencii v článku 15 o presnosti, robustnosti a kybernetickej bezpečnosti stanovuje, že vysokorizikové systémy AI musia byť navrhnuté a vyvíjané tak, aby dosiahli primeranú úroveň presnosti, robustnosti a kybernetickej bezpečnosti a aby v týchto ohľadoch konzistentne fungovali počas celého svojho životného cyklu.

Pre dosiahnutie robustnosti, musia vysoko rizikové systémy AI byť čo najodolnejšie, pokiaľ ide o chyby, poruchy alebo nekonzistentnosti, ktoré sa môžu vyskytnúť v rámci systému alebo prostredia, v ktorom systém funguje, najmä v dôsledku ich interakcie s fyzickými osobami alebo inými systémami. V tejto súvislosti sa majú prijať technické a organizačné opatrenia. Technické riešenia na dosiahnutie robustnosti vysokorizikových systémov UI môžu zahŕňať riešenia technickej redundancie, ktoré môžu zahŕňať záložné alebo havarijné plány. Vysoko rizikové systémy AI, ktoré sa naďalej učia po uvedení na trh alebo do prevádzky, musia byť vyvinuté tak, aby sa eliminovalo alebo čo najviac znížilo riziko možných skreslených výstupov ovplyvňujúcich vstupy pre budúce operácie (*feedback loops*), prostredníctvom vhodných zmierňujúcich opatrení.[423]

K požiadavke kybernetickej bezpečnosti Nariadenie o umelej inteligencii výslovne uvádza, že vysoko rizikové systémy AI musia byť odolné voči pokusom neoprávnených tretích strán zmeniť ich používanie, výstupy alebo výkon zneužitím zraniteľností systému. Uvedené sa má dosiahnuť najmä technickými riešeniami.[424] Nariadenie o umelej inteligencii zvýrazňuje, že tieto opatrenia majú byť zamerané na riešenie zraniteľností špecifických pre AI, a to na predchádzanie, detekciu, odozvu, riešenie a kontrolu

[423] Nariadenie o umelej inteligencii, článok 15 ods. 3.
[424] Nariadenie o umelej inteligencii, článok 15 ods. 4

a) útokov, ktoré sa snažia manipulovať súborom trénovacích údajov (*data poisoning*),
b) predtrénovaných komponentov používaných počas trénovania (*model poisoning*),
c) vstupov určených na spôsobenie chyby modelu (*adversarial examples* alebo *model evasion*),
d) útokov na dôvernosť alebo chýb modelu (*model flaws*).

Ďalšie technické riešenia na zabezpečenie kybernetickej bezpečnosti vysokorizikových systémov AI musia byť primerané vzhľadom na relevantné okolnosti a riziká. Medzi tieto riešenia patrí bezpečnosť založená na návrhu (*security-by-design*), ktorá zdôrazňuje dôležitosť integrovania bezpečnostných princípov už v raných štádiách návrhu a vývoja systémov AI a princíp bezpečnosti vo viacerých vrstvách (*security-in-depth*) ktorá je dôležitá pri ochrane pred kybernetickými hrozbami, pretože poskytuje viacero úrovní bezpečnosti, ktoré môžu pomôcť predchádzať útokom. Ak dôjde k narušeniu jednej vrstvy, ostatné vrstvy sú stále na mieste, aby chránili systém. Tento prístup znižuje riziko jediného bodu zlyhania (*single point of failure*) a poskytuje redundanciu (rezervu).

7.2.1 Aktíva systému umelej inteligencie

Systémom AI sa rozumie strojový systém, ktorý je schopný ovplyvňovať prostredie vydávaním odporúčaní, predpovedí alebo rozhodnutí pre daný súbor cieľov. Systém AI využíva strojové a/alebo ľudské vstupy/údaje na: i) vnímanie reálneho a/alebo virtuálneho prostredia; ii) abstrahovanie takéto vnímanie do modelov manuálne alebo automaticky; a iii) použitie modelových interpretácii na formulovanie možností pre výsledky.[425]

Riešenie požiadaviek na kybernetickú bezpečnosť systémov AI si nevyhnutne vyžaduje pochopenie vnútorných komponentov tohto systému a zamýšľaného kontextu ich použitia.[426] Tieto komponenty môžeme tiež označiť ako aktíva (*assests*) systémov AI.

Vnútorná štruktúra systémov AI zahŕňa celý rad komponentov a ich interakcie. Niektoré z týchto komponentov súvisia s AI, zatiaľ čo iné nie. Podľa princípov *security-in-depth*

[425] OECD (2019), *Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)*, OECD Digital Economy Papers, No. 291, OECD Publishing, Paris, [online] [30.12.2023]. Dostupné na: https://doi.org/10.1787/d62f618a-en.
[426] JUNKLEWITZ, H., et al.: *Cybersecurity of Artificial Intelligence in the AI Act*. s. 1 p1.

a *security-by-design* sa systémy AI musia rozložiť (dekomponovať) na základné artefakty informačnej architektúry. Jedným z týchto hlavných artefaktov sú modely AI. Hoci modely AI (vrátane základných modelov, *foundation models*) sú základnými komponentmi systémov AI, nie sú ich jediným komponentom (porov. bližšie kapitola 1).[427]

Požiadavky na kybernetickú bezpečnosť sa vzťahujú na systém AI ako celok, zatiaľ čo základné modely ako ich súčasť sú špecificky riešené z dôvodu ich univerzálnosti použitia v mnohých nadväzujúcich aplikáciách a následne vznikajúcich rizík.

Americký Národný inštitút pre štandardy a technológie (*NIST*) definuje aktívum ako údaje, personál, zariadenia, systémy a objekty, ktoré umožňujú organizácii dosahovať obchodné ciele.[428]

Pri dekompozícii systémov AI na aktíva by sme mali systémy AI považovať za súčasť infraštruktúry IKT. Podľa tohto prístupu prezentovaného agentúrou ENISA sú hlavnými komponentmi (aktívami) systémov AI: zdroje údajov, údaje, algoritmy, trénovacie modely, procesy implementácie/riadenia údajov/testovania a používatelia.[429]

Podobne, Spoločné výskumné centrum Európskej Komisie (*Joint Research Centre. JRC*) popisuje aktíva AI ako údaje, modely, iné digitálne aktíva alebo základnú infraštruktúru IKT.[430]

V ďalšom texte sa budeme venovať algoritmom strojového učenia, a špecificky základným modelom.

[427] Tamže.
[428] NIST Computer Security Resource Center, Glossary. [online] [30.12.2023]. Dostupné na: https://csrc.nist.gov/glossary/term/asset
[429] ENISA: *Multilayer Framework for Good Cybersecurity Practices for AI*. June 2023, [online] [30.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/multilayer-framework-for-good-cybersecurity-practices-for-ai.
[430] SOLER GARRIDO, J., et al.: *Analysis of the preliminary AI standardisation work plan in support of the AI act*. EUR 31518 EN, Publications Office of the European Union, Luxembourg, 2023. ISBN 978-92-68-03924-3, doi:10.2760/5847, JRC132833. [online] [30.12.2023]. Dostupné na: https://publications.jrc.ec.europa.eu/repository/handle/JRC132833.

7.2.1.1 Algoritmy strojového učenia

Algoritmy strojového učenia (*machine learning*, ML) sú najdôležitejším prvkom systémov AI a aplikácií AI (bližšie kapitola 1).[431] Čelia však sérii hrozieb a zraniteľností.

ENISA identifikovala šesť základných hrozieb a sedem vedľajších hrozieb pre strojové učenie. K identifikovaným hroznám následne priradila aj príklady zraniteľností ML, ktoré môžu týmito hrozbami zneužité a navrhla tak špecifické bezpečnostné opatrenia (popísané v kapitole 7.2.2).

1. Vyhýbanie (*Evasion*)

Typ útoku, pri ktorom útočník pracuje na vstupoch algoritmu ML, aby našiel malé poruchy vedúce k veľkej modifikácii jeho výstupov (napr. chyby v rozhodovaní). Je to, ako keby útočník vytvoril optickú ilúziu pre algoritmus. Takéto modifikované testovacie vstupy, ktoré spôsobujú nesprávnu klasifikáciu modelu strojového učenia sa nazývajú adversariálne príklady (*adversarial examples*) (bližšie tiež kapitola 1).[432]

Príklad: premietanie obrazov na dom by mohlo viesť algoritmus autonómneho auta k rozhodnutiu náhle zabrzdiť.

2. Oracle

Typ útoku, pri ktorom útočník skúma model poskytnutím série starostlivo vytvorených vstupov a pozorovaním výstupov. Tieto útoky môžu byť prípravnými krokmi k škodlivejším typom, napríklad vyhýbaniu alebo otrave (*poisoning*).

Je to, ako keby útočník prinútil model hovoriť, aby ho potom lepšie kompromitoval alebo aby o ňom získal informácie (napr. extrakcia modelu) alebo jeho tréningové údaje (napr. inversion útoky).

---

[431] ENISA: *Securing Machine Learning Algorithms*. December 2021, [online] Dostupné na: https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms.

[432] VASSILEV, A. et al.: (2024) *Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations.* (National Institute of Standards and Technology, Gaithersburg, MD) NIST Artifcial Intelligence (AI) Report, NIST Trustworthy and Responsible AI NIST AI 100-2e2023. s. 92 [online] [19.1.2024]. Dostupné na: https://doi.org/10.6028/NIST.AI.100-2e2023.

Príklad: Útočník študuje množinu vstupno-výstupných párov a výsledky použije na získanie tréningových dát.

3. Otrava (*Poisoning*)

Typ útoku, pri ktorom útočník zmenil údaje alebo model, aby modifikoval správanie algoritmu ML zvoleným smerom (napr. sabotoval jeho výsledky, vložil zadné vrátka). Je to ako keby útočník upravil algoritmus podľa jeho motivácií.

3.a Úprava štítku Útok, pri ktorom útočník poškodí štítky tréningových údajov. Táto čiastková hrozba je špecifická pre učenie s učiteľom (*supervised learning*).

4. Zverejnenie údajov alebo modelu

Táto hrozba sa týka možnosti úniku všetkých alebo čiastočných informácií o modeli.

Príklad: Výstupy algoritmu ML sú také podrobné, že poskytujú informácie o jeho konfigurácii alebo o citlivých údajoch.

4.a Zverejnenie údajov

Táto hrozba sa týka úniku údajov upravovaných algoritmami ML. Tento únik údajov možno vysvetliť nedostatočnou kontrolou prístupu, chybou projektového tímu pri manipulácii s údajmi alebo jednoducho preto, že entita, ktorá vlastní model, a entita, ktorá vlastní údaje, sú niekedy odlišné. Na trénovanie modelu je preto často potrebné, aby k údajom mal prístup poskytovateľ modelu. Ide o zdieľanie údajov, a teda zdieľanie citlivých údajov s treťou stranou.

4.b Zverejnenie modelu

Táto hrozba sa týka úniku vnútorných častí (t. j. hodnôt parametrov) modelu ML. K úniku modelu môže dôjsť v dôsledku ľudskej chyby alebo zdieľania s treťou stranou s príliš nízkou úrovňou bezpečnosti.

5. Kompromitácia komponentov aplikácie ML

Táto hrozba sa vzťahuje na kompromitáciu komponentu alebo vývojového nástroja aplikácie ML.

Príklad: kompromitácia jednej z knižníc s otvoreným zdrojovým kódom používaných vývojármi na implementáciu algoritmu ML.

6. Zlyhanie alebo porucha aplikácie ML

Táto hrozba sa týka zlyhania aplikácie ML (napr. odmietnutie služby v dôsledku zlého vstupu, nedostupnosť v dôsledku chyby spracovania).

Príklad: úroveň služieb podpornej infraštruktúry aplikácie ML poskytovanej treťou stranou je príliš nízka v porovnaní s biznis potrebami, aplikácia je pravidelne nedostupná.

Táto hrozba nezohľadňuje zlyhanie v prípadoch, na ktoré bola aplikácia vytvorená (napríklad algoritmus zlyhá, pretože nie je dostatočne presný na zvládnutie všetkých reálnych situácií, ktorým je vystavený).

6.a Ľudská chyba

Rôzne zainteresované strany môžu robiť chyby, ktoré vedú k zlyhaniu alebo nefunkčnosti aplikácie ML. Napríklad z dôvodu nedostatku dokumentácie môžu aplikáciu používať v prípadoch, s ktorými sa pôvodne nepočítalo.

6.b Odmietnutie služby

Útočníci môžu využívať starostlivo vytvorené vstupné dáta (*sponge examples*), ktoré sú vstupmi navrhnutými tak, aby maximalizovali spotrebu energie a latenciu, a tak viesť systémy strojového učenia k ich najhoršiemu možnému výkonu. Sponge examples sú prvým útokom na odopretie služby proti komponentom strojového učenia týchto systémov. Jazykové modely sú týmto útokom prekvapivo zraniteľné. Sponge examples často zvyšujú latenciu aj spotrebu energie týchto modelov 30× a tým potenciálne spôsobia odmietnutie služby.[433]

[433] SHUMAILOV, I. Y, et al.: *Sponge Examples: Energy-Latency Attacks on Neural Networks,* 2021 IEEE European Symposium on Security and Privacy (EuroS&P), Vienna, Austria, 2021, pp. 212-231, [online] [19.1.2024]. Dostupné na: doi: 10.1109/EuroSP51992.2021.00024.

6.c Incident kybernetickej bezpečnosti nebol nahlásený tímu pre riešenie incidentov

Táto hrozba sa týka možnosti, že projektový tím nenahlási bezpečnostné incidenty špecializovaným tímom, aj keď bola prijatá politika povinného hlásenia incidentov.

7.2.1.2 Základné modely (*Foundation models*)

Systémy AI môžu byť implementáciou základného modelu, čo znamená, že každý základný model možno opätovne použiť v nespočetných nadväzujúcich (*downstream*) aplikáciách AI (k základným modelom bližšie pozri kapitolu 1). Základný model alebo tiež všeobecný model (*general-purpose AI model*) si môžeme predstaviť ako univerzálnu platformu pre aplikácie AI.

Základné modely sú distribuované tak s uzavretým zdrojom - zdrojovým kódom (*closed source*) ako aj s otvoreným zdrojom (*open source*).[434] Pokiaľ ide o bezpečnosť a riadenie (*governance)*, pri oboch modeloch distribúcie existujú veľké nedostatky. Oba modely vykazujú riziká v rôznych aspektoch. Modely s uzavretým zdrojom ponúkajú pridané možnosti bezpečnosti a riadenia, ktoré modelom s otvoreným zdrojom chýbajú. Hoci modelom s otvoreným zdrojom chýbajú funkcie bezpečnosti a riadenia, možno ich začleniť do bezpečnostného perimetra organizácií a bezpečne doladiť na lokálnych údajoch.[435]

Vplyv základných modelov na trh je taký významný, že upútal pozornosť orgánov na dohľad nad hospodárskou súťažou.[436] Orgán pre hospodársku súťaž Spojeného kráľovstva (*UK Competition and Markets Authority*) vo svojej predbežnej analýze z roku 2023 venuje pozornosť obmedzeniam v prístupe k základným modelom pre akademikov a ostatné subjekty, ako sú audítori tretích strán, aby mohli vykonávať nezávislé testovanie alebo vyvíjať kritéria pre porovnávanie (benchmarky) a štandardy. Väčšia transparentnosť by určite mohla viesť k zlepšeniu benchmarkov a podpore technickej normalizácie. Externé metriky kvality od audítorov alebo akademikov sú pre podniky rozhodujúce pri výbere ich dodávateľoch a poskytovatelia základných modelov preto nemusia súťažiť v otázkach kvality. Výsledkom je,

[434] U.K. Competition and Markets Authority: *AI Foundation Models: Initial Review* zo dňa 18. septembra 2023. s. 14. [online] [30.12.2023]. Dostupné na: https://www.gov.uk/government/publications/ai-foundation-models-initial-report.
[435] LU. S,: *Proprietary vs. Open Source Foundation Models*, 15. máj 2023, [online] [30.12.2023]. Dostupné na: https://tolacapital.com/2023/05/15/foundationmodels/.
[436] U.K. Competition and Markets Authority: *AI Foundation Models: Initial Review*.

že spotrebitelia rovnako nemôžu porovnávať a prechádzať v používaní medzi aplikáciami používajúcimi základné modely.[437]

Podľa OWASP patria medzi 10 najväčších hrozieb pre veľké jazykové modely (*Large Language Models*, LLM), ktoré sú jednou z kategórií základných modelov (bližšie kapitola 1):[438]

1. Manipulovanie vstupného textu - promptu (*Prompt Injection*)

Ide o manipuláciou s veľkým jazykovým modelom prostredníctvom sofistikovaných vstupov, čo spôsobuje nepredpokladané konanie LLM. Priame manipulácie prepisujú systémové prompty, zatiaľ čo nepriame manipulácie manipulujú so vstupmi z externých zdrojov.

2. Nezabezpečená manipulácia s výstupom

Táto zraniteľnosť sa vyskytuje, keď je výstup LLM akceptovaný bez kontroly, čím sú odhalené backendové systémy. Zneužitie môže viesť k vážnym následkom, ako sú XSS, CSRF, SSRF, eskalácia privilégií alebo vzdialené spustenie kódu.

3. Otrava trénovacích údajov

K tomu dochádza, keď sú sfalšované trénovacie údaje LLM, čím sa zavádzajú zraniteľnosti alebo predsudky, ktoré ohrozujú bezpečnosť, efektivitu alebo etické správanie. Zdroje zahŕňajú Common Crawl, WebText, OpenWebText a knihy.

4. Odmietnutie služby modelu

Útočníci spôsobujú na LLM operácie náročné na zdroje, čo vedie k degradácii služieb alebo vysokým nákladom. Zraniteľnosť sa zväčšuje kvôli povahe LLM náročných na zdroje a nepredvídateľnosti používateľských vstupov.

[437] U.K. Competition and Markets Authority: *AI Foundation Models: Initial Review* s. 101, bod 5.81.
[438] OWASP Top 10 for LLM. Version 1.1 online. dostupné na: https://www.llmtop10.com/assets/downloads/OWASP-Top-10-for-LLM-Applications-v1_1.pdf

5. Zraniteľnosť dodávateľského reťazca

Životný cyklus aplikácie LLM môže byť ohrozený zraniteľnými komponentmi alebo službami, čo vedie k útokom. Používanie datasetov tretích strán, pre-trénovaných modelov a zásuvných modulov (plug-inov) môže zvýšiť počet zraniteľností.

6. Zverejňovanie citlivých informácií

LLM môžu vo svojich odpovediach neúmyselne uviesť dôverné údaje, čo vedie k neoprávnenému prístupu k údajom, narušeniu súkromia a narušeniu bezpečnosti. Na zmiernenie tohto je dôležité implementovať proces odstránenia citlivých údajov (*data sanitization*) a prísne pravidlá pre používateľov.

7. Nezabezpečený návrh zásuvného modulu

Zásuvné moduly, plug-iny LLM môžu mať nezabezpečené vstupy a nedostatočné riadenie prístupu. Tento nedostatok kontroly aplikácií uľahčuje ich zneužitie a môže viesť k následkom, ako je vzdialené spustenie kódu.

8. Nadmerné oprávnenie

Systémy založené na LLM môžu vykonávať činnosti vedúce k nezamýšľaným následkom. Problém vzniká z nadmernej funkčnosti, oprávnení alebo autonómie udelenej systémom založeným na LLM.

9. Nadmerné spoliehanie sa (*Overreliance*)

Systémy alebo ľudia, ktorí sú príliš závislí na LLM bez dohľadu, môžu čeliť dezinformáciám, nesprávnej komunikácii, právnym problémom a zraniteľnostiam v dôsledku nesprávneho alebo nevhodného obsahu generovaného LLM.

10. Krádež modelu

To zahŕňa neoprávnený prístup, kopírovanie alebo exfiltráciu proprietárnych modelov LLM. Dopady zahŕňajú ekonomické straty, ohrozenie konkurenčnej výhody a potenciálny prístup k citlivým informáciám.

7.2.2 Bezpečnostné opatrenia

Kybernetická bezpečnosť AI je nová oblasť skúmania, ktorá zbiera a kombinuje poznatky a prístupy v rôznych oblastiach, ako je výskum AI, adversariálne strojové učenie[439] a všeobecná kybernetická bezpečnosť.

Z pohľadu opatrení v kybernetickej bezpečnosti možno AI vnímať ako typ softvéru, a preto aj pri prijímaní bezpečnostných opatrení vo vzťahu k systémom AI sú použiteľné tam, kde je to možné, už zavedené bezpečnostné postupy a opatrenia. V súvislosti s kybernetickou bezpečnosťou AI však existuje celý rad technologických výziev špecifických pre AI. Tieto výzvy sú väčšinou spojené so systémami strojového učenia, kde sa objavuje väčší počet nových zraniteľností špecifických pre AI, ktoré môžu byť zneužité útočníkmi.[440] Bližšie sa zneužitiu AI venujeme nižšie v podkapitole 7.4.

ENISA identifikovala 37 bezpečnostných opatrení pre algoritmy strojového učenia, ktoré zaradila do základných 3 skupín opatrení (organizačné, technické, a osobitné pre ML).[441] Niektoré bezpečnostné opatrenia sú špecifické pre algoritmy ML, ale iné sú štandardné technické a organizačné opatrenia kybernetickej bezpečnosti na zmiernenie všeobecných útokov. Je dôležité uplatňovať oba typy opatrení, pretože v systémoch umelej inteligencie existujú okrem zraniteľností špecifických pre ML aj zraniteľnosti všeobecného typu, ktoré môžu byť tiež zneužité útočníkmi.

7.2.2.1 Organizačné opatrenia

- ***Použite RBAC modelu a rešpektovanie princípu najnižších oprávnení***

Definovať správu prístupových práv pomocou RBAC (*Role Based Access Control*) modelu rešpektujúc princíp najnižších oprávnení. Tento by sa mal vzťahovať na všetky zložky modelu (napr. hostiteľské infraštruktúry) a umožniť ochranu zdrojov, ako napr. modelu (napr. jeho

[439] Adversariálne strojové učenie (*Adversarial machine learning*) je proces získavania informácií o správaní a charakteristikách systému strojového učenia a/alebo učenie sa, ako manipulovať so vstupmi do systému strojového učenia s cieľom získať preferovaný výsledok

[440] JUNKLEWITZ, H., et al.: *Cybersecurity of Artificial Intelligence in the AI Act*. s. 6. p1

[441] ENISA: *Securing Machine Learning Algorithms*. December 2021, s. 19 a nasl.

konfigurácia, jeho kód) a údaje, ktoré používa (napr. tréningové údaje). Roly, ktoré sa majú zahrnúť sa majú týkať aj koncového používateľa.

Napríklad: koncový používateľ, ktorý môže vložiť vstupy do modelu, by nemal mať možnosť mať prístup k jeho konfigurácii.

- ***Dokumentácia projektov***

Ako pri všetkých projektoch, dokumentácia musí byť pre umelú inteligenciu, aby sa zachovali poznatky o rozhodnutiach prijatých počas projektovej fázy, architektúre aplikácie a jej konfigurácii, jej údržbe, spôsobe údržby účinnosti v priebehu času a predpokladoch o používaní modelu.

Táto dokumentácia by mala obsahovať aj zmeny, ktoré sa budú uplatňovať, vrátane dokumentácie počas celého trvania životného cyklu algoritmu.

- ***Zabezpečenie súladu aplikácie ML s právnymi predpismi***

Ako všetky aplikácie, aj tie, ktoré využívajú ML, môžu podliehať právnym predpisom (napr. v závislosti od zhromaždených údajov). Takéto posúdenie sa musí vykonať čo najskôr už vo fáze projektu a by sa potom malo pravidelne aktualizovať podľa vývoja právnej úpravy (napr. AI Akt).

- ***Zabezpečenie súladu aplikácií ML so stratégiou bezpečnosti a bezpečnostnými politikami***

Tak ako všetky aplikácie, aj tie, ktoré používajú ML, musia dodržiavať existujúce politiky, najmä riadenie rolí, autentifikácia, politiky riadenia prístupu, bezpečnostné politiky ako napr. hardening, anti-malware politika a integrovanie do procesov bezpečnostných operácií (napr. riadenie zraniteľností, zálohovanie) vrátane procesov detekcie a reakcie na incidenty.

7.2.2.2 Technické opatrenia

- ***Posúdenie úrovne exponovania použitého modelu***

Niektoré návrhy modelov sú častejšie používané alebo zdieľané ako iné a najmä v oblasti ML (napr. otvorené zdieľanie zdrojového kódu). Tieto aspekty musia byť zohľadnené v analýze rizík.

- Nepoužívať modely prevzaté priamo z internetu bez ich kontroly.

- Používať modely, pre ktoré sú jasne identifikované hrozby a pre ktoré existujú bezpečnostné opatrenia.

- ***Kontrola zraniteľností použitých komponentov***

Počas životného cyklu ML algoritmu sa v ňom nachádza niekoľko komponentov (napríklad softvér, programové knižnice alebo aj iné modely) použitých na dokončenie projektu. Musia byť vykonané kontroly, aby sa zabezpečilo, že tieto komponenty poskytujú primeranú úroveň bezpečnosti. Okrem toho je potrebné urobiť niektoré postupy, aby sa zabránilo manipulácii s komponentmi.

Napríklad: ak sa má knižnica s otvoreným zdrojovým kódom používať, je potrebné vykonať revíziu kódu alebo skontrolovať, či nie sú zverejnené jej zraniteľnosti.

- ***Analýza rizík aplikácie ML***

Analýza rizík celkovej aplikácie by sa mala vykonávať s prihliadnutím na špecifiká jeho kontextu vrátane motivácie útočníka, citlivosti spracovávaných údajov, hosting aplikácie, architektúra modelu, životného cyklu aplikácie ML (napr. zdieľanie modelu).

- ***Kontrola všetkých údajov používaných modelom ML***

Údaje sa musia skontrolovať, aby sa zabezpečilo, že budú vyhovovať modelu a obmedzia príjem škodlivých údajov. Predovšetkým ide o vyhodnotenie úrovne dôveryhodnosti zdrojov a pôvodcu, ich integrity v rámci celého reťazca dodávania údajov. Ďalej sem možno zaradiť kontrolu anomálií, automaticky alebo manuálne (napr. selektívna ľudská kontrola).

- ***Určenie a monitorovanie ukazovateľov pre správne fungovanie modelu***

Opatrenie spočíva v definovaní kľúčových ukazovateľov vrátane bezpečnostných indikátorov (zmeny v správaní modelu atď.) pre správne fungovanie modelu, najmä na rýchlu identifikáciu anomálií.

- ***Nasadenie vhodnej bezpečnosti pre testovacie prostredia***

- ***Dodržiavanie bežných procesov integrácie bezpečnosti do projektov***

Ako každý projekt, aj projekty ML musia byť v súlade s procesom integrácie bezpečnosti do projektov, vrátane nasledujúcich prvkov:

- Analýza rizík celej aplikácie.

- Kontrola integrácie osvedčených postupov kybernetickej bezpečnosti, pokiaľ ide o architektúru, bezpečný vývoj.

- Integrovanie do existujúcich prevádzkových bezpečnostných procesov, najmä do monitorovania a odozva, patch management, riadenie prístupov.

- Kontrola vyhotovenia primeranej dokumentácie na zabezpečenie udržateľnosti aplikácie (napr. technická architektúra, hardening, konfiguračné a inštalačné dokumenty).

- Bezpečnostné kontroly pred publikovaním do produkčného prostredia (napr. bezpečnostný audit, penetračné testy).

7.2.2.3 Opatrenia špecifické pre strojové učenie

- ***Pridanie niekoľko adversariálnych príkladov do množiny trénovacích údajov školenia (Robust adversarial training)***

Do procesu trénovania algoritmu by sa mali zahrnúť adversariálne príklady, aby bol model voči týmto útokom odolnejší.

- ***Úpravy vstupov***

Pridanie kroku na úpravu vstupov modelu (napr. randomizácia údajov, ktorá spočíva v pridaní náhodného šumu ku každému údaju) môže zlepšiť odolnosť modelu voči útokom. Takéto kroky môžu útočníkovi sťažiť pochopenie fungovania algoritmu, manipuláciu s ním a zníženie dopadov útoku. Opatrenie môže byť použité počas trénovania alebo fáz nasadenia modelu.

- ***Vysvetliteľnosť modelov (explainability)***

Modely ML by mali byť vysvetliteľné, aj keď to znamená ich zjednodušenie, aby sa umožnilo dobré pochopenie ich fungovania a rozhodovacích faktorov. Môže to byť aj regulačná požiadavka (napr. GDPR). Bezpečnosť však zasahuje do vysvetliteľnosti modelu. Ide teda o kompromis medzi potrebou vysvetliteľnosti a bezpečnosťou.

- ***Odolnejší model***

Niektoré návrhy modelov môžu byť voči útokom odolnejšie ako iné. Je potrebné zvážiť výber vhodného modelu.

- ***Zväčšenie trénovacieho datasetu***

Použitie rozšíreného súboru trénovaných údajov rieši nedostatok údajov a zlepšuje robustnosť modelu voči útokom otrávenia (*poisoning*) oslabením ich vplyvu.

- ***Neskreslenosť modelu***

Niektoré techniky možno použiť na zmiernenie skreslení (*bias*). Je potrebné overiť, či je trénovací dataset dostatočne reprezentatívny pre použitie na zamýšľaný účel, a tiež skontrolovať relevantnosť atribútov použitých na rozhodovanie atď.

- ***Modely musia v dostatočnej miere rešpektovať diferencované súkromie (Differential Privacy)***

Diferenciálne súkromie[442] je silná matematická definícia súkromia v kontexte štatistickej a ML analýzy. Toto opatrenie môže výrazne znížiť výkonnosť modelu. Preto je dôležité odhadnúť potrebu ochrany údajov alebo modelu.

- ***Odolnosť modelu voči prostrediu, v ktorom bude fungovať***

Model by mal byť dostatočne odolný voči prostrediu, v ktorom bude fungovať. To zahŕňa napríklad zabezpečenie toho, aby proces učenia a údaje boli dostatočne reprezentatívne pre skutočné podmienky, v ktorých sa bude model vyvíjať.

- ***Implementácia procesov na udržiavanie úrovne bezpečnosti komponentov ML***

ML je rýchlo sa vyvíjajúca oblasť, najmä pokiaľ ide o jej kybernetickú bezpečnosť. Pravidelná kontrola nových útokov a ochranných opatrení musí byť integrovaná do procesov pre udržiavanie úrovne bezpečnosti aplikácií.

- ***Implementácia nástroja na zisťovanie adversariálnych príkladov***

Na identifikáciu toho, či bol daný vstup modifikovaný útočníkom, alebo nie, môžu byť zaujímavé nástroje na detekciu vstupov. Jedným z príkladov v prípade hlbokých neurónových sietí je pridanie neurónovej podsiete k architektúre vytrénovanej na detekciu advesariálnych príkladov.

- ***Integrácia špecifík ML do stratégie zvyšovania povedomia***

Do programov zvyšovania povedomia by sa mali pridať potrebné informácie o strojovom učení pre zainteresované strany. Školenie o kybernetickej bezpečnosti by malo zahŕňať:

- Školenie o kybernetickej bezpečnosti vrátane osvedčených postupov na prevenciu ohrozenia aplikácie ML.

[442] Diferenciálne súkromie je odbor matematiky a informatiky pre dosiahnutie ochrany súkromia jednotlivcov v databázach údajov.

- Manipuláciu s potenciálne citlivými údajmi alebo údajmi podliehajúcimi regulačným obmedzeniam.

- Konfigurácie na zabránenie zraniteľnosti aplikácií.

- Informovanosť o útokoch špecifických pre ML.

- ***Integrácia kontroly otrávení (poisoning) po fáze "hodnotenia modelu"***

Pred presunom modelu do produkcie a potom pravidelne by sa mal model vyhodnotiť, aby sa zabezpečilo, že nebol otrávený.

- ***Minimalizácia dostupných informácií o modeli***

Táto ochrana spočíva v obmedzení informácií o modeli, keď nie sú nevyhnutné. Jej cieľom je prijať potrebné opatrenia na zníženie dostupných informácií o modeli, ako napr. informácie o trénovacom datasete alebo akékoľvek iné informácie, ktoré by mohol použiť útočník (napr. nezverejňovanie modelu v otvorenom zdroji). Samozrejme, ide o kompromis medzi bezpečnosťou a skutočnosťou, že zainteresované strany (napr. používatelia) niekedy chcú modely s otvoreným zdrojovým kódom.

- ***Používanie federatívneho učenia na minimalizáciu rizika narušenia ochrany údajov***

Federatívne učenie je súbor techník trénovania, ktoré trénujú model na niekoľkých decentralizovaných serveroch obsahujúcich lokálne vzorky údajov bez toho, aby si vymieňali svoje vzorky údajov. Tým sa predchádza potrebe prenášať údaje a/alebo ich zveriť nedôveryhodnej tretej strane, čo pomáha zachovať ochranu údajov.

- ***Používanie menej ľahko prenosných modelov***

Vlastnosť prenositeľnosti sa dá využiť na to, aby adversariálne príklady prenikli do iného modelu. Jednoduchosť prenosu adversariálneho príkladu z modelu do iného, závisí od rodiny algoritmov. Jednou z obrán je teda vybrať si rodinu algoritmov, ktorá je menej citlivá na prenositeľnosť advesariálnych príkladov.

Pre bezpečnostné opatrenia, ktoré nie sú špecifické pre algoritmy ML, ENISA vychádzala z noriem rady ISO 27000 alebo noriem NIST 800-53 (*Security and Privacy Controls for Information Systems and Organizations*). Je teda zrejmé, že ďalší vývoj a výskum sa okrem posúdenia útokov na modely strojového učenia sústredí aj na prispôsobenie existujúcich bezpečnostných opatrení na AI a vývoj bezpečnostných opatrení špecifických pre AI.[443]

Mnohé bezpečnostné opatrenia, ktoré nie sú špecifické pre AI, možno z veľkej časti prevziať zo série noriem ISO 27000, ktorá zahŕňa dobre zavedené postupy organizačných princípov, riadenia rizík a bezpečnostných opatrení. Postupne sú publikované technické normy prispôsobené na použitie pre AI.

Technickej normalizácii a vývoji špecifických noriem pre AI sa bližšie venujeme v kapitole 2. Pre oblasť kybernetickej bezpečnosti AI však možno vyzdvihnúť predovšetkým normu ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system (*AI Management System, AIMS*). Norma ISO 42001 považuje za kritické predovšetkým nasledovné procesy

- určenie organizačných cieľov, zapojenie zainteresovaných strán a politiku organizácie;
- riadenie rizík;
- procesy riadenia problémov súvisiacich s dôveryhodnosťou systémov umelej inteligencie, ako je bezpečnosť, spravodlivosť, transparentnosť, kvalita údajov a kvalita systémov umelej inteligencie počas ich životného cyklu;
- procesy pre riadenie dodávateľov, partnerov a tretích strán, ktoré poskytujú alebo vyvíjajú systémy AI pre organizáciu.

Norma ISO/IEC 42001 funguje rovnakým spôsobom ako norma ISO/IEC 27001. Príloha A obsahuje opatrenia, ktoré sa majú zohľadniť pri ošetrovaní rizík. Príloha B je podobná samostatnej norme ISO 27002 pre prílohu A k norme ISO 27001 a poskytuje návod na vykonávanie opatrení uvedených v prílohe A. V prílohe C sú organizačné ciele, zdroje rizík a opisy, ktoré možno zohľadniť pri riadení rizík súvisiacich s používaním AI. Príloha D sa zaoberá

[443] ENISA: *Securing Machine Learning Algorithms*. December 2021, s. 19 a nasl.

používaním AIMS v rôznych oblastiach alebo sektoroch a ich integráciou s inými systémami riadenia.

7.3 AI na podporu kybernetickej bezpečnosti

AI pomáha bezpečnostným tímom automatizovať opakujúce sa úlohy, zrýchliť detekciu hrozieb a reakciu na ne, ako aj zlepšiť presnosť ich akcií s cieľom posilniť bezpečnosť proti rôznym bezpečnostným problémom a kybernetickým útokom.

Použitie umelej inteligencie (či už ide o nástroje, techniky, metódy AI), je už bežné v prevencii útokov, detekcii hrozieb a narušenia, reakcii a obnove z kybernetických útokov. Doposiaľ neexistuje komplexný prehľad najnovšieho výskumu, ktorý by popisoval všetky použitia AI v kybernetickej bezpečnosti a podrobnosti o tom, ako sa v takýchto prípadoch AI používajú.[444]

Používanie umelej inteligencie na podporu kybernetickej bezpečnosti s osobitným zameraním na praktické aplikácie možno najlepšie ilustrovať v rámci piatich funkcií konceptu kybernetickej bezpečnosti (Identifikovať, Chrániť, Detegovať, Reagovať a Obnoviť) definovaných rámcom kybernetickej bezpečnosti NIST (*Cybersecurity Framework*).[445]

Analýza z roku 2022 preukázala, že 36 % odborných štúdií sa zameralo na odhaľovanie anomálií a incidentov v oblasti kybernetickej bezpečnosti. Detekcia riadená algoritmami strojového učenia môže viesť k automatickej detekcii útokov a včasnej obrane. Druhou najpopulárnejšou kategóriou s 28 % bola identifikácia, za ňou nasleduje ochrana (25 %) a reakcia (10 %). Veľmi málo štúdií zaznamenaných v rokoch 2021 a 2022 sa zameralo na využitie AI na obnovu.[446] Obdobne väčšina článkov na konferenciách, ktoré sa venovali

[444] KAUR, R., GABRIJELČIČ D., KLOBUČAR, T.: *Artificial intelligence for cybersecurity: Literature review and future research directions*, Information Fusion, Volume 97, 2023, 101804, ISSN 1566-2535, [online] [30.12.2023]. Dostupné na: https://doi.org/10.1016/j.inffus.2023.101804.
[445] NIST Cybersecurity Framework verzia 1.0, [online] [30.12.2023]. Dostupné na: https://www.nist.gov/cyberframework. Bol už zverejnený návrh verzie 2.0 rámca.
[446] KAUR, R., GABRIJELČIČ D., KLOBUČAR, T.: *Artificial intelligence for cybersecurity: Literature review and future research directions*. p. 28.

aplikácii AI pre kybernetickú bezpečnosť bola publikovaná v troch hlavných funkciách: Identifikovať (40 %), Detegovať (31 %) a Chrániť (17 %).[447]

ENISA identifikovala jednotlivé použiteľné nástroje, techniky a metódy AI pri prevencii (zodpovedá funkciám Identifikovať a Chrániť podľa NIST rámca) a detekcii.[448]

7.3.1 Prevencia

Umelá inteligencia sa môže použiť na posúdenie zraniteľností v počítačových systémoch a sieťach. Algoritmy strojového učenia sa často používajú pri analýze údajov z viacerých zdrojov, ako sú skenery, bezpečnostné logy a systémy riadenia opráv, na identifikáciu zraniteľností a stanovenie priorít pri náprave.

Fuzzing testovanie[449] založené na hlbokom učení sa teraz považuje za najsľubnejšiu cestu na objavovanie zraniteľností v porovnaní s tradičným ML.[450]

Strojové učenie môže byť prínosom aj pri skórovaní rizika v sieti, napr. pri určovaní závažnosti zraniteľnosti. Možno ho tiež použiť na analýzu správania používateľov a identifikáciu podozrivých aktivít, ako sú pokusy o prevzatie účtu alebo pokusy o neoprávnený prístup. Umelá inteligencia je teda použiteľná všeobecne pri správe identít používateľov a kontrole prístupu k počítačovým systémom a aplikáciám.

7.3.2 Detekcia

Väčšina „tradičných" aplikácií strojového učenia spadá takmer výlučne do fázy detekcie, t. j. na detekciu spamu, detekciu narušenia a detekciu škodlivého softvéru, ako aj detekciu útokov. Kým doposiaľ bola pozornosť zameraná najmä na detekciu spamu v

[447] Tamže. Pozn. – v tejto štúdii neboli zahrnuté články, ktoré boli publikované v iných vedeckých databázach ako Scopus alebo používali iné kľúčové slová. Tiež neboli zahrnuté najnovšie publikácie (po februári 2022).
[448] ENISA: *Artificial Intelligence and Cybersecurity Research*, Jún 2023, str. 20 - 22, [online] [30.12.2023]. Dostupné na: https://www.enisa.europa.eu/publications/artificial-intelligence-and-cybersecurity-research.
[449] *Fuzz testing* alebo *Fuzzing* je technika testovania softvéru, ktorá v podstate spočíva v nájdení chýb implementácie pomocou vkladania chybných/upravených údajov automatizovaným spôsobom.
[450] Porov. BEAMAN, C., REDBOURNE, M., MUMMERY, J.D., HAKAK, S.: *Fuzzing vulnerability discovery techniques: Survey, challenges and future directions,* Computers & Security, Volume 120, 2022, 102813, ISSN 0167-4048, [online] [30.12.2023]. Dostupné na: https://doi.org/10.1016/j.cose.2022.102813.

počítačových sieťach, ďalšou rozvíjajúcou sa oblasťou je detekciu škodlivého softvéru (malvéru) a nedovolených prienikov.

V oblasti detekcie malvéru sa strojové učenie používa na výber relevantných funkcií odhaľujúcich prítomnosť malvéru, ako aj metód na zisťovanie anomálií alebo abnormalít.[451]

7.3.3 Budúci výskum a vývoj

Hoci sa AI postupne stáva kľúčovou zložkou kybernetickej bezpečnosti, existujú stále podstatné medzery vo výskume a pri určení príležitostí pre AI vo výskume kybernetickej bezpečnosti. Kľúčovým prvkom je identifikácia nových aplikačných domén, vhodných zdrojov a pokročilých techník AI použiteľných v kybernetickej bezpečnosti. Budúci výskum by mal prebiehať preto v štyroch hlavných oblastiach: (i) vznikajúce aplikačné domény kybernetickej bezpečnosti, (ii) reprezentácia údajov, (iii) pokročilé metódy AI pre kybernetickú bezpečnosť a (iv) výskum a vývoj novej infraštruktúry.[452]

7.3.3.1 Vznikajúce aplikačné domény

1) Automatizované vyhľadávanie kľúčových indikátorov rizika (*key risk indicators*) v reálnom čase;
2) Detekcia nových útokov;
3) Prediktívna (predpovedná) analýza;
4) Viacjazyčné (neanglické) spravodajské služby o hrozbách (*threat intelligence*);
5) Kybernetická obrana a odolnosť podporovaná umelou inteligenciou (napríklad automatizované modelovanie hrozieb, automatizovaná oprava zraniteľností, a automatizovaná segmentácia a reorganizácia siete);
6) Prevencia a zisťovanie narušenia údajov;
7) Generovanie falošných dokumentov[453];
8) Kontextovo riadené spracovanie bezpečnostných udalostí a ich triedenie;

[451] Porov. aj HOSSAIN FARUK, M., et al.: *Malware Detection and Prevention using Artificial Intelligence Techniques*, In 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 2021 s. 5369-5377. [online] [30.12.2023]. Dostupné na: doi: 10.1109/BigData52589.2021.9671434.
[452] KAUR, R., GABRIJELČIČ D., KLOBUČAR, T.: *Artificial intelligence for cybersecurity: Literature review and future research directions*. p. 27. s. 24.
[453] Ide o relatívne nový koncept ochrany citlivého dokumentu falšovaním informácie, vytvorením mnohých falošných verzií akéhokoľvek dokumentu.

9) Odozva na kybernetické bezpečnostné incidenty podporovaná umelou inteligenciou.

7.3.3.2 Reprezentácia údajov

Kvalitné údaje sú kľúčové pre efektívnu AI. Najväčšou výzvou je preto výber vhodných súborov údajov na trénovanie a spracovanie rôznorodosti a rýchlosti údajov. Preto je dôležitá reprezentácia údajov, kvalita údajov, získavanie aktuálnosti o údajoch (*recency mining*) a znalosť kontextu pre trénovanie a modelovanie modelov AI pre aplikácie kybernetickej bezpečnosti.

7.3.3.3 Pokročilé metódy AI pre kybernetickú bezpečnosť

Na realizáciu plného potenciálu vyššie uvedených aplikačných domén a reprezentácie údajov sú potrebné sofistikovanejšie techniky AI. Tri zásadné nové techniky, ktoré môžu mať veľký vplyv na vývoj praktickej a použiteľnej AI pre kybernetickú bezpečnosť patria:

a) analýza viac zdrojov údajov (*multiple data source*),
b) vysvetliteľná (*explainable*) AI a
c) rozšírená inteligencia (rozhrania *human-AI*).

7.3.3.4 Výskum a vývoj novej infraštruktúry

Potreba výskumu a vývoja nových infraštruktúr na podporu technológie AI je priamo spojená so spracovaním enormných objemov údajov z interných systémov aj externých informačných kanálov, s cieľom poskytnúť kybernetickú bezpečnosť v reálnom čase. Medzi kľúčové medzery vo výskume pre úspešnú implementáciu AI do kybernetickej bezpečnosti patria najmä:

a) Nedostatok platforiem so spravodajskými službami o hrozbách (*threat intelligencie*)

- Vzhľadom na dynamickosť prostredia, kde vznikajú vždy nové hrozby a útoky, ktoré sú navrhnuté tak, aby obišli známy potenciálny scenár, sú potrebné vhodné platformy, ktoré umožnia spoluprácu medzi určenými hodnotiteľmi pri diskusii a zdieľaní najnovších údajov o hrozbách.

b) Nedostatok datasetov (databáz údajov), ktoré by boli dostatočne rozsiahle a aktuálne

- Datasety sú najdôležitejšou súčasťou AI pre kybernetickú bezpečnosť. Väčšina z dostupných súbory údajov sú neaktuálne a nemusia stačiť na pochopenie najnovších vzorcov správania rôznych kybernetických útokov.

7.4 Zneužitie AI na páchanie kybernetickej kriminality

Umelá inteligencia je typická technológia s dvojakým použitím, teda možno ju rovnako využívať v dobrom úmysle, ale možno ju tiež zneužiť k páchaniu protiprávnej činnosti. Škodliví aktéri (útočníci) sa učia, ako zefektívniť svoje útoky pomocou tejto technológie na nájdenie a zneužitie zraniteľností v IKT produktoch, službách a systémoch. S AI sa tieto schopnosti postupne stávajú automatizovanými a je ťažšie ich odhaliť.

Keďže technológia AI napreduje, je pravdepodobné, že v budúcnosti uvidíme sofistikovanejšie a komplexnejšie kybernetické útoky založené na AI. Napríklad generatívna adversariálna sieť alebo generatívny difúzny model sa môžu použiť na generovanie „deep fake" obrázkov alebo videí manipuláciou s tvárami alebo hlasmi (porov. bližšie kapitola 1).[454]

Algoritmy založené na AI sú tiež schopné pripraviť presvedčivé spear-phishingové emaily.[455] Umelá inteligencia sa môže použiť aj na zvýšenie efektívnosti a účinnosti malvéru zlepšením jeho schopnosti vyhnúť sa detekcii, prispôsobiť sa meniacemu sa prostrediu, zamerať sa na konkrétne zraniteľné miesta, šíriť sa a pretrvávať v cieľových systémoch.[456]

Predmetom útoku sú aj samotné systémy AI, a to v dôsledku svojich vlastných zraniteľností alebo slabých miest.[457] Medzi útoky proti AI patria najmä nasledujúce:

[454] Porov napr. PREETI, KUMAR, M., KUMAR SHARMA, H., *A GAN-Based Model of Deepfake Detection in Social Media*, Procedia Computer Science, Volume 218, 2023, Pages 2153-2162, ISSN 1877-0509, [online] [30.12.2023]. Dostupné na: https://doi.org/10.1016/j.procs.2023.01.191.

[455] Napríklad HAZELL, Julian. *Large language models can be used to effectively scale spear phishing campaigns*. 2023 [online] [30.12.2023]. Dostupné na: *arXiv preprint arXiv:2305.06972*.

[456] Porov. tiež FRITSCH, L., JABER, A., YAZIDI, A. (2022). *An Overview of Artificial Intelligence Used in Malware*. In: Zouganeli, E., Yazidi, A., Mello, G., Lind, P. (eds) Nordic Artificial Intelligence Research and Development. NAIS 2022. Communications in Computer and Information Science, vol 1650. Springer, Cham. [online] [30.12.2023]. Dostupné na: https://doi.org/10.1007/978-3-031-17030-0_4.

[457] Porov. tiež SANGWAN, R.S.; BADR, Y.; SRINIVASAN, S.M.. 2023. *Cybersecurity for AI Systems: A Survey*. Journal of Cybersecurity and Privacy 3, no. 2: 166-190. [online] [30.12.2023]. Dostupné na: https://doi.org/10.3390/jcp3020010.

a) Útoky využívajúce existujúce zraniteľnosti v populárnych softvérových knižniciach s otvoreným zdrojovým kódom.
b) Útoky otrávením tréningových údajov. Tu sa predpokladá, že útočník má prístup k trénovacím údajom a je schopný ich zmeniť a zaviesť manipulované údaje, ako sú nesprávne označenia, aby systém AI, vyškolený na poškodených údajoch, vykonával spracovanie a/alebo predpovede podľa záujmov útočníka.
c) Adversariálne útoky, kde zvyčajne napadnutým systémom AI je hlboká neurónová sieť (*deep neural network*). Zmenami v testovacích príkladoch sa zmení predikcia systému AI cieleným alebo necieleným spôsobom, t. j. útočník nasmeruje predikciu k danej požadovanej triede alebo k akejkoľvek inej ako správnej triede.
d) Reverzné (spätné) inžinierstvo natrénovaného modelu založeného na verejne prístupných rozhraniach, ide o útoky ako napríklad krádež modelov (*model stealing*) či inverzia modelu (*model inversion*).

8. Obchodnoprávne vzťahy a umelá inteligencia

AI preniká do všetkých oblastí života a práva, nevynímajúc obchodné právo a obchodnoprávne vzťahy. V tejto časti učebnice rozoberieme využitie AI v obchodnom práve, a to od automatizácie rozhodovania členov orgánov obchodných spoločností[458] cez uzavieranie obchodnoprávnych zmlúv až po vytváranie nových produktov a optimalizáciu poskytovania služieb zákazníkom. Cieľom je tak poskytnúť primerane komplexný právny prehľad otázok a odpovedí v rámci vzťahu umelej inteligencie a obchodného práva.

8.1 Využitie AI v obchodnoprávnych vzťahoch

Pred samotným vymedzením oblastí využitia AI v rámci obchodného práva je nevyhnutné špecifikovať aké atribúty by mala spĺňať umelá inteligencia využiteľná v obchodnoprávnych vzťahoch. Čiastočnú odpoveď na túto otázku poskytla Expertná skupina Európskej komisie na vysokej úrovni AI v roku 2019, keď vypracovala odporúčania - *Etické smernice pre dôveryhodnú umelú inteligenciu* zdôrazňujúce dôležitosť ***dôveryhodných systémov AI***. Dôveryhodná umelá inteligencia má tri zložky, ktorých sa treba pridržiavať počas celého životného cyklu systému:

1. mala by byť *zákonná*, čím sa zabezpečí dodržiavanie celého platného práva a právnych predpisov;

2. mala by byť *etická*, čím sa zabezpečí súlad s etickými zásadami a hodnotami, a

3. mala by byť *odolná*, a to z technického aj sociálneho hľadiska, keďže systémy umelej inteligencie môžu aj pri dobrých úmysloch spôsobiť neúmyselnú ujmu/škodu.[459] Presnejšie povedané EÚ definovala

[458] Podľa aktuálnych údajov Štatistického úradu Slovenskej republiky k 3. kvartálu roka 2023 je v Slovenskej republike celkovo 254 025 obchodných spoločností, pričom z daného čísla tvoria spoločnosti s ručením obmedzeným 247 989 (97,6 %) a akciových spoločností je 4812 (1,89 %). Z uvedeného dôvodu sa zameriavame najmä na kapitálové obchodné spoločnosti, ktoré sú v SR najzastúpenejšie. Pre bližšie údaje o ekonomických subjektov podľa právnych foriem pozri: http://datacube.statistics.sk/#!/view/sk/VBD_SLOVSTAT/og2019qs/v_og2019qs_00_00_00_sk.

[459] Expertná skupina na vysokej úrovni pre umelú inteligenciu (HLEG AI) (2019). *Etické usmernenia pre dôveryhodnú umelú inteligenciu.* Správa pre Európsku komisiu, 09.04.2019. Brusel: Európska komisia. s. 2. [online]. Dostupné na: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

„Sedem kľúčových požiadaviek na dôveryhodnú AI:

1. *Ľudský faktor a dohľad - Zahŕňa základné práva, ľudský faktor a ľudský dohľad,*
2. *Technická odolnosť a bezpečnosť - Zahŕňa odolnosť voči útokom a bezpečnostnú ochranu, záložný plán a všeobecnú bezpečnosť, presnosť, spoľahlivosť a reprodukovateľnosť,*
3. *Správa súkromia a údajov - Zahŕňa rešpektovanie súkromia, kvalitu a integritu údajov a prístup k údajom,*
4. *Transparentnosť - Zahŕňa vysledovateľnosť, vysvetliteľnosť a komunikáciu,*
5. *Rozmanitosť, nediskriminácia a spravodlivosť - Zahŕňa zabránenie nespravodlivej zaujatosti, prístupnosť a dizajn pre všetkých a účasť zainteresovaných strán*
6. *Spoločenský a environmentálny blahobyt - Zahŕňa udržateľnosť a šetrnosť k životnému prostrediu, sociálny vplyv, spoločnosť a demokraciu,*
7. *Zodpovednosť - Zahŕňa kontrolovateľnosť, minimalizáciu negatívneho vplyvu a jeho oznamovanie, kompromisy a nápravu*."[460]

Po splnení vymedzených kľúčových požiadaviek na dôveryhodnú AI v obchodnoprávnych vzťahoch môže byť AI využitá v nasledovných oblastiach obchodného práva:

1. ***Rozhodovanie členov orgánov obchodných spoločností***: AI môže byť použitá na zlepšenie a optimalizáciu rozhodovania v oblastiach spadajúcich do rozhodovacej činnosti členov orgánov obchodných spoločností. To môže viesť k zvýšeniu efektivity a konkurencieschopnosti obchodných spoločností.

PRÍKLAD
AI môže byť použitá na podporu rozhodovania tým, že poskytuje členom orgánov prístup k informáciám, dátam a analýzam, ktoré by inak nebolo možné získať alebo automatizáciou úloh a rozhodnutí, ktoré sú časovo náročné, prípadne opakujúce sa.

[460] Tamže.

2. ***Správa a riadenie obchodných spoločností*:** Umelá inteligencia môže slúžiť ako nástroj na riadenie samotných spoločností prostredníctvom pomoci alebo dokonca nahradenia ľudí vo vedení spoločnosti.

PRÍKLAD
V súčasnosti už existujú spoločnosti, ktoré do manažérskych pozícií vymenovali stroje umelej inteligencie, napr. poľská spoločnosť Dictador, ktorá má na pozícii CEO – generálneho riaditeľa humanoidného robota menom Mika.[461]

3. ***Uzavieranie obchodnoprávnych zmlúv:*** AI má potenciál zlepšiť spôsob, akým podnikatelia uzatvárajú zmluvy. Tieto nové technológie AI môžu výrazne znížiť pracovnú silu potrebnú na tvorbu samotných zmlúv a celého kontraktačného procesu, čím pridávajú významnú hodnotu celej organizácii.

PRÍKLAD
Využitím automatizovaných softvérov na analýzu zmlúv je možné generovať a kontrolovať zmluvy, identifikovať potenciálne problémy i extrahovať dôležité klauzuly. To môže zlepšiť a urýchliť procesy tvorby zmlúv a takisto ovplyvniť ich obsah.

4. ***Optimalizácia úloh a vytváranie nových produktov a služieb*:** AI môže automatizovať mnohé úlohy, ktoré sa v súčasnosti vykonávajú manuálne, napríklad spracovanie objednávok, vyhodnocovanie rizík alebo poskytovanie zákazníckej podpory. Výsledkom môže byť zvýšenie produktivity a zníženie nákladov. AI môže byť tiež použitá na vytváranie nových produktov a služieb, čo môže viesť k expanzii trhov a vytvoreniu nových podnikateľských príležitostí.

PRÍKLAD
Prostredníctvom automatizovaných procesov umelá inteligencia môže asistovať pri vývoji produktov a služieb novej generácie, dokáže optimalizovať rozhodovacie procesy, zlepšiť

[461] Pozri *What Do You Think of This AI Robot 'CEO'?* [online]. Dostupné na: https://www.youtube.com/watch?v=8BQEyQ2-awc.

údržbu strojov, zvýšiť produkciu a kvalitu, zlepšiť služby zákazníkom, ušetriť energiu, a to napr. zlepšením predikcie, optimalizáciou podnikových operácií, prideľovania zdrojov a rozhodovania, ako aj personalizáciou poskytovania služieb zákazníkom, napríklad prostredníctvom nasadenia chatbotov.

8.1.1 Výhody využitia AI v obchodnoprávnych vzťahoch

Začlenenie umelej inteligencie do podnikania so sebou prináša množstvo významných výhod, a to:

1. ***Rýchlosť a efektívnosť***: Systémy umelej inteligencie dokážu spracovať a analyzovať obrovské množstvo údajov v reálnom čase, čím sa urýchli rozhodovací proces, prípadne iné činnosti spoločnosti. AI môže pomôcť podnikom automatizovať rutinné úlohy, ako je spracovanie zmlúv, analýza údajov a poskytovanie služieb zákazníkom. To môže viesť k zníženiu nákladov a zvýšeniu produktivity.
2. ***Identifikácia zložitých vzorcov***: AI má schopnosť odhaliť skryté vzorce v údajoch, ktoré si ľudia nemusia všimnúť. Identifikáciou zložitých vzťahov a trendov je možné získať cenné informácie pre podnikanie.
3. ***Väčšia presnosť a objektivita***: Schopnosť AI analyzovať údaje prispieva k objektívnejšiemu rozhodovaniu založenému na faktoch. Elimináciou vplyvu ľudských predsudkov a emócií pomáha AI minimalizovať chyby a zvyšuje presnosť pri rozhodovaní a plnení pridelených úloh.[462]
4. ***Nové príležitosti:*** AI môže korporáciám identifikovať nové príležitosti a zlepšiť ich konkurenčné postavenie na trhu.
5. ***Zlepšenie rozhodovania:*** AI môže pomôcť podnikom zlepšiť rozhodovanie tým, že im poskytne presnejšie a aktuálnejšie informácie. AI môže napríklad analyzovať veľké množstvo údajov, aby identifikovala trendy a predpovedala budúce výsledky. To môže pomôcť spoločnostiam prijímať informovanejšie rozhodnutia o svojich produktoch,

[462] PERÉZ, S. D. *The impact of AI on business decision making* [online]. Dostupné na: https://intelequia.com/en/blog/post/the-impact-of-ai-on-business-decision-making

službách a stratégiách. AI sa môže použiť aj na analýzu údajov o zákazníkoch, aby sa vytvorili personalizované ponuky a služby.

8.1.2 Riziká využitia AI v obchodnoprávnych vzťahoch

1. ***Nedostatočné využitie*** (môže prameniť z nedôvery podnikov v AI a zapríčiniť hospodársku stagnáciu odrezaním spoločnosti od nových príležitostí na trhu) alebo naopak ***nadužívanie umelej inteligencie*** (napríklad investície do technológií AI, ktoré sa ukážu ako neužitočné pre konkrétnu spoločnosť).

2. ***Transparentnosť***[463] *a* ***vysvetliteľnosť****:* Mnohé algoritmy AI fungujú ako *„čierne skrinky*", takže je ťažké pochopiť, ako fungujú a dosahujú konkrétne rozhodnutia. V týchto prípadoch môžu byť potrebné iné opatrenia súvisiace s vysvetliteľnosťou (napr. vysledovateľnosť, kontrolovateľnosť a transparentná komunikácia o schopnostiach systému) za predpokladu, že systém ako celok rešpektuje základné práva.[464] Existuje pritom niekoľko dôvodov, prečo je transparentnosť AI dôležitá:

    - *Dôvera*: Transparentnosť AI môže pomôcť používateľom dôverovať systémom AI. Keď používatelia vedia, ako fungujú systémy AI, je menej pravdepodobné, že budú mať obavy z jej používania.
    - *Porozumenie*: Transparentnosť AI môže pomôcť používateľom lepšie pochopiť, ako systémy AI fungujú. To môže byť užitočné pre výrobcov, ktorí sa snažia zlepšiť systémy AI, ako aj pre používateľov, ktorí sa snažia lepšie pochopiť, ako systémy AI fungujú.

---

[463] FELZMAN, H., VILLARONGA, E. F., LUTZ, Ch., LARRIEUX, A.T. *Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns.* Big Data & Society, 6(1). 2019. p. 1 - 2 [online]. Dostupné na: https://doi.org/10.1177/2053951719860542
So vzostupom AI sa stala transparentnosť kľúčovou témou, systémy AI môžu byť založené buď na potenciálnej alebo spätnej transparentnosti:

- *Potenciálna transparentnosť* informuje používateľov o spracovaní údajov a fungovaní AI vopred. Popisuje, ako AI systém prijíma rozhodnutia vo všeobecnosti. To môže zahŕňať informácie o údajoch, ktoré boli použité na trénovanie systému AI, ako aj o algoritme, ktorý systém AI používa na prijímanie rozhodnutí.
- *Spätná transparentnosť*, na druhej strane odkazuje na *post hoc* vysvetlenia. Prezrádza pre konkrétny prípad, ako a prečo sa dospelo k určitému rozhodnutiu, popis spracovania údajov krok za krokom. Spätná transparentnosť zahŕňa pojem kontrolovateľnosť a vysvetliteľnosť. Preto, aby mal algoritmický rozhodovací systém retrospektívnu transparentnosť, človek by mal mať možnosť kontrolovať jeho *„vnútorné*" rozhodnutie pochopiť štruktúru a systém hodnotenia premís v rámci systému a v konečnom dôsledku vysvetliť rozhodnutie.

[464] Expertná skupina na vysokej úrovni pre umelú inteligenciu (HLEG AI) (2019). *Vymedzenie pojmu umelej inteligencie: hlavné schopnosti a disciplíny.* Správa pre Európsku komisiu, 08.04.2019. Brusel: Európska komisia. s. 6. [online]. Dostupné na: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

3. ***Zodpovednosť za škodu***: Otvorenou otázkou na globálnej úrovni zostáva, kto je zodpovedný za škody spôsobené umelou inteligenciou. Určenie zodpovednosti v prípade nehôd alebo škôd súvisiacich s AI je náročné. Autonómny a nie vždy predvídateľný charakter fungovania systémov AI vedie k otázkam o príčinnej súvislosti, *vis maior* a pod. Kto bude niesť zodpovednosť, keď AI spôsobí škodu obchodnej spoločnosti? Kto bude zodpovedný, ak systém AI navrhne členovi orgánu spoločnosti plán, ktorý poškodzuje korporáciu? Je možné žalovať algoritmus AI za nesprávny návrh, ktorý členovia orgánu spoločnosti nasledovali? Otázky o tom, kto je zodpovedný – vývojári, používatelia alebo samotná AI – by sa mali riešiť v právnych predpisoch.

4. ***Zaujatosť a spravodlivosť****:* Výstupy umelej inteligencie veľmi úzko súvisia s tým, ako je navrhnutá a akými údajmi je *„krmená".* Jej dizajn aj vstupné údaje môžu byť úmyselne alebo neúmyselne okresané. Pri vytváraní algoritmu môžu byť vynechané dôležité súčasti v dôsledku čoho môžu udržiavať alebo zosilňovať existujúce predsudky v spoločnosti, čo vedie k diskriminačným a zaujatým výsledkom.[465] Napríklad, ak sa údaje určené na výcvik vyznačujú zaujatosťou (čiže nie sú dostatočne vyvážené alebo inkluzívne), systém umelej inteligencie vycvičený na základe týchto údajov nebude môcť správne zovšeobecňovať a pravdepodobne bude prijímať zaujaté rozhodnutia, ktoré môžu uprednostňovať niektoré skupiny pred inými.[466]

5. ***Súkromie a ochrana údajov:*** Systémy umelej inteligencie sa často spoliehajú na obrovské množstvo osobných údajov, čo vyvoláva obavy o súkromie a ochranu údajov. Napríklad štatutárny orgán môže použiť dôverné informácie v generatívnych AI (napr. ChatGPT), čím môže prísť k úniku dôverných informácií. Okrem toho umožňuje takýto systém zlučovanie informácií, ktoré daná osoba poskytla, do nových údajov. Toto môže viesť k neočakávaným a negatívnym výsledkom. Systémy umelej inteligencie musia zaručovať ochranu súkromia a údajov počas celého životného cyklu systému. Patria sem

[465] BRYNJOLFSSON, E., MCAFEE, A. *Artificial Intelligence: The Insights You Need from Harvard Business Review.* Harvard Business School Publishing Corporation. 2019. p. 21.
[466] Expertná skupina na vysokej úrovni pre umelú inteligenciu (HLEG AI) (2019). *Vymedzenie pojmu umelej inteligencie: hlavné schopnosti a disciplíny.* Správa pre Európsku komisiu, 08.04.2019. Brusel: Európska komisia. s. 6. [online]. Dostupné na: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

informácie, ktoré pôvodne poskytol používateľ, ako aj informácie, ktoré o používateľovi vznikli v priebehu jeho komunikácie so systémom (napr. výstupy, ktoré systém umelej inteligencie vytvoril v súvislosti s konkrétnymi používateľmi, alebo spôsob, akým používatelia reagovali na konkrétne odporúčania). V EÚ je v tejto súvislosti najmä potrebné zabezpečiť súlad s Nariadením 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 – GDPR).

6. ***Dopady na zamestnanosť***: Vznik nových technológií – od elektrickej energie po internet – odjakživa menil povahu práce. Široké zavádzanie AI má potenciál narušiť tradičné trhy práce, čo môže viesť k nezamestnanosti v určitých sektoroch. AI má potenciál vykonávať niektoré úlohy vykonávané ľuďmi, na druhej strane však pomôže vytvoriť nové a lepšie pracovné miesta, na ktoré však bude potrebná vyššia kvalifikácia. V konečnom dôsledku tak technológie AI pomáhajú zlepšovať schopnosti ľudí a robiť ich produktívnejšími.[467]

7. ***Konkurencieschopnosť***: Zhromažďovanie veľkého množstva údajov môže viesť aj k narušeniu hospodárskej súťaže. Spoločnosti, ktoré vedia získať masy údajov, môžu vďaka nim získať obrovské konkurenčné výhody a v krajnom prípade eliminovať svojich konkurentov.[468]

8. ***Cena***: AI prispôsobená na používanie v konkrétnej spoločnosti môže byť nákladná technológia. Preto je nevyhnutné pred implementáciou AI vykonať základnú analýzu potreby, aby sa zistilo, či je AI vhodná na splnenie konkrétnych potrieb spoločnosti.

[467] RUSSELL, S. NORVIG, P. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 4rd edition, Pearson Education Limited 2022, p. 49 - 50.
[468] Európsky parlament. *Umelá inteligencia: možné oblasti využitia a riziká, ktoré so sebou prináša* [online]. Dostupné na: https://www.europarl.europa.eu/news/sk/headlines/society/20200918STO87404/umela-inteligencia-mozne-oblasti-vyuzitia-a-rizika-ktore-so-sebou-prinasa

8.2 Rozhodovanie členov orgánov kapitálových obchodných spoločností a AI

Štruktúru orgánov jednotlivých obchodných korporácií môžu tvoriť na jednej strane orgány obligatórne, ktorých tvorba je zákonom vyžadovaná a na druhej strane orgány fakultatívne, ktorých existenciu zákon výslovne nepredpisuje, a to na báze súkromnoprávnej zásady *„čo právo nezakazuje je dovolené*". Na rozdiel od kapitálových spoločností v osobných obchodných spoločnostiach sa orgány povinne nevytvárajú, a tieto formy obchodných spoločností sú charakteristické osobnou účasťou spoločníka na podnikateľskej činnosti korporácie. Ako sme už vyššie uviedli v tejto časti učebnice sa venujeme kapitálovým obchodným spoločnostiam, nakoľko tieto sú najzastúpenejšie v rámci SR. V kapitálových obchodných spoločnostiach je mechanizmus vytváraných orgánov nasledovný:

- *Valné zhromaždenie* – najvyšší a obligatórny orgán v spoločnosti s ručením obmedzeným, akciovej spoločnosti, aj jednoduchej spoločnosti na akcie,
- *Štatutárny orgán* – obligatórny orgán v spoločnosti s ručením obmedzeným (konateľ/konatelia), v akciovej spoločnosti (predstavenstvo), v jednoduchej spoločnosti na akcie (predstavenstvo),
- *Dozorná rada* - obligatórny orgán v akciovej spoločnosti, fakultatívny orgán v spoločnosti s ručením obmedzeným a jednoduchej spoločnosti na akcie.
- *Fakultatívne orgány* – napríklad poradné orgány (investičný výbor, výbor pre nové projekty, výbor pre audit, výbor pre odmeňovanie), špeciálne kontrolné orgány (v oblasti hospodárenia s majetkom korporácie), sociálne orgány, atď.[469]

Rozhodovacia pôsobnosť orgánov kapitálových obchodných spoločností

Každý z vyššie uvedených orgánov má rozličnú rozhodovaciu pôsobnosť, ktorú si bližšie priblížime na nasledovných riadkoch.

[469] K štruktúre orgánov v s. r. o. pozri MAMOJKA, M. *Časť 11.4. Orgány spoločnosti.* In MAMOJKA, M. a kol. *Obchodné právo I (všeobecná časť, súťažné právo, právo obchodných spoločností a družstva).* Bratislava: C. H. Beck, 2021. s. 457 - 519. K štruktúre orgánov v a. s. pozri MAŠUROVÁ, A. *Časť 12.4. Orgány spoločnosti.* In MAMOJKA, M. a kol. *Obchodné právo I (všeobecná časť, súťažné právo, právo obchodných spoločností a družstva).* Bratislava: C. H. Beck, 2021. s. 589 - 635. K štruktúre orgánov v j. s. a. pozri HAJNIŠOVÁ, E. *Časť 13.4. Orgány spoločnosti.* In MAMOJKA, M. a kol. *Obchodné právo I (všeobecná časť, súťažné právo, právo obchodných spoločností a družstva).* Bratislava: C. H. Beck, 2021. s. 749 - 766.

- *Valné zhromaždenie* je najvyšším orgánom kapitálových obchodných spoločností, na ktorom vykonávajú spoločníci svoje práva týkajúce sa riadenia spoločnosti a kontroly jej činnosti prostredníctvom rozhodovania o všetkých zásadných otázkach fungovania spoločnosti. Valné zhromaždenie zvoláva štatutárny orgán pozvánkou, pričom sa má konať najmenej jedenkrát za rok. Na rokovaní valného zhromaždenia majú právo zúčastniť sa všetci spoločníci (akcionári). Títo sa môžu zúčastniť osobne alebo v zastúpení splnomocnencom na základe písomného plnomocenstva. Splnomocnencom môže byť tak fyzická, ako aj právnická osoba a na valnom zhromaždení môže splnomocnenec zastupovať aj viacerých spoločníkov. Pôsobnosť valného zhromaždenia je vymedzená v ObZ demonštratívne,[470] čo znamená že spoločníci (akcionári) obchodnej spoločnosti môžu valnému zhromaždeniu v spoločenskej (zakladateľskej) zmluve priznať širší okruh rozhodovacích oprávnení. Valné zhromaždenie si tiež môže vyhradiť rozhodovanie vecí, ktoré inak patria do pôsobnosti iných orgánov spoločnosti.[471]

- *Štatutárnym orgánom* spoločnosti s ručením obmedzeným je jeden alebo viac konateľov.[472] V prípade akciovej spoločnosti a jednoduchej spoločnosti na akcie je štatutárnym orgánom predstavenstvo.[473] Pôsobnosť štatutárnych orgánov je možné rozdeliť na 2 oblasti, a to obchodné vedenie a konanie v mene spoločnosti. Obchodné vedenie je pôsobnosť *„dovnútra*" charakterizované prijímaním rozhodnutí súvisiacich s bežnou prevádzkou podniku, ako napríklad rozhodovanie o otázkach organizačného charakteru, technických otázkach, veciach vnútornej prevádzky, podnikateľskom zámere,[474] zavádzanie marketingu, riešenie záväzkov, riešenie optimálneho spôsobu riadenia,[475] rozhodovanie o postupe pri vymáhaní pohľadávok spoločnosti,[476] uzatvorení nájomnej a podnájomnej zmluvy,[477] nadobudnutí a scudzení majetku.[478] Negatívne vymedzenie obchodného vedenia koncipuje Najvyšší súd ČR v uznesení zo dňa

[470] Napríklad pôsobnosť valného zhromaždenia spoločnosti s ručením obmedzeným je vymedzená v § 125 ods. 1 ObZ. Spoločenská zmluva alebo stanovy nemôžu zákonom stanovenú pôsobnosť obmedziť, môžu ju len rozšíriť. K problematike valného zhromaždenia v s. r. o. bližšie pozri MAMOJKA, M. *Časť 11.4. Orgány spoločnosti.* In MAMOJKA, M. a kol. *Obchodné právo I (všeobecná časť, súťažné právo, právo obchodných spoločností a družstva).* Bratislava: C. H. Beck, 2021. s. 457 - 478.
[471] § 125 ods. 3 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.
[472] § 133 – § 135a zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.
[473] § 191 – § 195 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.
[474] Uznesenie Najvyššieho súdu Slovenskej republiky zo dňa 30.04.2009, sp. zn. 5 Obdo 12/2009.
[475] Rozsudok Najvyššieho súdu Českej republiky zo dňa 30.07.2008, sp. zn. 29 Odo 1262/2006.
[476] Uznesenie Najvyššieho súdu Českej republiky zo dňa 24.06.2009, sp. zn. 29Cdo 3139/2007.
[477] Uznesenie Najvyššieho súdu Slovenskej republiky zo dňa 30.04.2009, sp. zn. 5 Obdo 12/2009.
[478] Rozsudok Najvyššieho súdu Českej republiky zo dňa 11.04.2012, sp. zn. 29 Cdo 3223/2010.

05.04.2006, sp. zn. 5Tdo 94/2006 ,,*Do obchodného vedenia spoločnosti nepatria tie činnosti, ktoré konateľ vykonáva vo vzťahu k spoločníkom či orgánom spoločnosti, napr. zvolanie valného zhromaždenia podľa § 128 ods. 1 Obch. zák., oznámenie o následnom prijatí uznesenia valného zhromaždenia podľa § 127 ods. 8 Obch. zák., vedenie zoznamu spoločníkov podľa § 135 ods. 1 Obch. zák. a pod., pretože tieto činnosti sa nedotýkajú každodennej podnikateľskej činnosti spoločnosti a smerujú k výkonu práv spoločníkov.*"[479] Konanie v mene spoločnosti je naopak pôsobnosťou *„navonok*" voči tretím osobám. Ide o vystupovanie podnikateľa voči inému subjektu, ako napríklad uzatváranie zmlúv či iných úkonov v mene a na účet spoločnosti.

- *Dozorná rada* je kontrolným orgánom obchodnej spoločnosti, ktorej činnosť spočíva v tom, že dohliada na činnosť členov štatutárneho orgánu a uskutočňovanie podnikateľskej činnosti spoločnosti, nahliada do všetkých dokladov a záznamov týkajúcich sa činnosti spoločnosti a kontroluje, či sú účtovné záznamy riadne vedené v súlade so skutočnosťou a či sa podnikateľská činnosť spoločnosti uskutočňuje v súlade s právnymi predpismi, stanovami a pokynmi valného zhromaždenia. V prípade spoločnosti s ručením obmedzeným a jednoduchej spoločnosti na akcie ide o fakultatívny orgán, ktorého zriadenie musí byť upravené v zakladateľskom dokumente. Naproti tomu v prípade akciovej spoločnosti je dozorná rada obligatórnym orgánom. Z dôvodu dosiahnutia riadnej činnosti spoločnosti člen dozornej rady nemôže byť štatutárnym orgánom obchodnej spoločnosti.

- *Fakultatívny orgán* je vo svojej podstate orgán, ktorý obchodná spoločnosť nemá povinnosť zriadiť.[480] Právna úprava spoločnosti s ručením obmedzeným vytvorenie fakultatívnych orgánov explicitne neupravuje, avšak ani nezakazuje.[481] V rámci úpravy akciovej spoločnosti ObZ takéto orgány tiež explicitne neupravuje, možno však konštatovať, že ich tvorbu predpokladá, a to na základe ustanovenia § 173 ods. 1 písm. f) ObZ ,,*Stanovy musia obsahovať: počet členov predstavenstva, dozornej rady alebo iných*

[479] Uznesenie Najvyššieho súdu Českej republiky zo dňa 05.04.2006, sp. zn. 5Tdo 94/2006.
[480] K možnosti zriaďovania fakultatívnych orgánov v obchodnej spoločnosti pozri napríklad uznesenie Najvyššieho súdu Českej republiky zo dňa 24.11.2009, sp. zn. 29 Cdo 4563/2008.
[481] MRÁZOVÁ, Ž. *1.2 Volené orgány kapitálových spoločností*. In MRÁZOVÁ, Ž. HUSÁR, J. DOLNÝ, J. a kol. *Volené orgány kapitálových spoločností. Vybrané otázky.* Košice: ŠafárikPress, 2021. s. 11.

orgánov, ako aj vymedzenie ich pôsobnosti a spôsob rozhodovania,".[482] Z daného ustanovenia možno vyvodiť, že pri zriaďovaní fakultatívnych orgánov je nutná úprava ich pôsobnosti, spôsobu kreovania, spôsobu rozhodovania, úprava vzťahov s ostatnými subjektmi v rámci spoločnosti, teda začlenenie do systému orgánov, vymedzenie práv a povinností jednotlivých členov, regulácia vzájomných vzťahov medzi členmi orgánov a subjektami vo vnútri spoločnosti a pod.[483] Fakultatívny orgán je možné zriadiť formou spoločenskej zmluvy, rozhodnutím valného zhromaždenia alebo stanovami. Možno teda zhrnúť, že fakultatívnymi orgánmi môžu byť orgány poradné, správno-finančné (napr. v otázke investícií do dlhodobého hmotného majetku), špeciálne kontrolné (najčastejšie v oblasti hospodárenia s majetkom korporácie), poradné (výbor pre odmeňovanie) alebo sociálne.

Podnikateľské rozhodovanie a AI

Pri podnikateľskom rozhodovaní členov orgánov obchodných spoločností môže byť AI využívaná najmä prostredníctvom **automatizovaných procesov rozhodovania, t. j. prijímaním rozhodnutí, ktoré sú delegované na stroj alebo systém**. Tieto zahŕňajú širokú škálu techník zameraných na vytváranie inteligentných systémov, zatiaľ čo *strojové učenie (machine learning, deep learning)* sa špecializuje na vývoj algoritmov, ktoré sa učia z údajov vytvárať predpovede alebo rozhodnutia, *generatívna AI* (napr. *ChatGPT*) využíva strojové učenie na generovanie originálneho a realistického obsahu. Proces používania tejto technológie a jej algoritmov sa využíva na analýzu údajov, zisťovanie vzorov alebo generovanie odporúčaní na podporu obchodného rozhodovania.

V rámci kategorizácie procesov rozhodovania existujú 3 stupne:

1. ***podpora rozhodovania*** - pomocou prediktívnej, diagnostickej alebo deskriptívnej analytiky pomáha robiť presnejšie rozhodnutia. Výhoda spočíva v spojení ľudskej inteligencie s poznatkami AI založenými na údajoch. Kombinácia ľudského rozumu a odborných znalostí môže z AI v podnikoch vyťažiť to najlepšie.

[482] § 171 ods. 3 písm. f) zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.
[483] KOCHAN, M. *Vytváranie fakultatívnych orgánov v obchodných spoločnostiach*. In Orgány obchodných spoločností, Univerzita Pavla Jozefa Šafárika v Košiciach, 2015. s. 62.

PRÍKLAD
Flexibilné vyhľadávanie, manipulácia a klasifikácia širokej škály neštruktúrovaných veľkých súborov údajov a iných informácií, napr. analýza sociálnych médií, novelizácie právnych predpisov alebo trhových podmienok, údaje o zákazníkoch. Tieto analýzy môžu poskytnúť členom orgánov cenné informácie, ktoré môžu využiť pri rozhodovaní o strategických otázkach obchodného vedenia. Napríklad IBM Watson je procesor na analýzu údajov, ktorý využíva spracovanie prirodzeného jazyka, technológiu, ktorá analyzuje význam a syntax ľudskej reči. Kognitívne a analytické schopnosti IBM Watson umožňujú reagovať na ľudskú reč, spracovávať obrovské zásoby údajov a generovať odpovede na zložité otázky. Keď sú nové údaje vložené do dátového úložiska Watson, využíva strojové učenie, aby pokračoval v rozširovaní svojich vedomostí, ktoré je schopný následne poskytnúť používateľom.[484] Spoločnosti v mnohých odvetviach už používajú IBM Watson na predikčnú analýzu a riešenie problémov.[485]

2. ***augmentácia rozhodovania*** - využíva prediktívne alebo preskriptívne analýzy na odporúčanie viacerých alternatív rozhodovania. Synergia schopností AI a ľudských znalostí vedie k rýchlej analýze veľkých objemov údajov.

PRÍKLAD
Sofistikovanejšia analýza údajov na podporu rozhodnutí spoločnosti, napr.: analýza rizík konkrétneho projektu, obchodné prognózy, odporúčania alternatív týkajúce sa investičných rozhodnutí, a pod. Príkladom je technológia AI menom Rationale, ktorá pomáha štatutárnym orgánom, manažérom a jednotlivcom robiť rozhodnutia. Funguje tak, že sa mu zadá rozhodnutie, ktoré sa má prijať a algoritmy GPT a kontextového učenia vypíšu klady a zápory, vygenerujú SWOT analýzu, vykonajú multikriteriálnu analýzu alebo kauzálnu analýzu, aby pomohli zvážiť možnosti. Po zvážení všetkých relevantných faktorov Rationale pomôže urobiť racionálne rozhodnutie.[486]

[484] Pozri https://www.youtube.com/watch?v=U6rvaWaiZNg.
[485] CONNER, F. *IBM Watson: What are companies using it for?* [online]. Dostupné na: https://www.zdnet.com/article/ibm-watson-what-are-companies-using-it-for/.
[486] Pozri Rationale. Dostupné na: https://rationale.jina.ai/.

3. ***automatizácia rozhodovania*** - použitie preskriptívnej alebo prediktívnej analýzy na automatizáciu rozhodovacieho procesu.[487]

PRÍKLAD
Najvyššia úroveň autonómie AI zahŕňa jej schopnosť vytvárať rozhodnutia v mene spoločnosti, ako napríklad uzatváranie zmlúv s tretími stranami alebo riadiacimi zamestnancami.

Na základe uvedeného by sme pre prípad podnikateľského rozhodovania členov orgánov obchodných spoločností mohli zovšeobecniť, že AI predstavuje počítačové systémy, ktoré sú spôsobilé fungovať spôsobom podobným ľudskej inteligencii, v dôsledku čoho takéto systémy rozumejú ľudskej reči, sú schopné riešiť problémy a prijímať rozhodnutia, učiť sa z nadobudnutých skúseností, adaptovať sa a zlepšovať.

8.2.1 Právny základ využívania systémov umelej inteligencie v podnikateľskom rozhodovaní obchodných spoločností v Slovenskej republike

Členovia štatutárnych orgánov kapitálových obchodných spoločností sú povinní vykonávať svoju pôsobnosť s **náležitou (odbornou) starostlivosťou** a **v súlade so záujmami spoločnosti a všetkých jej spoločníkov (akcionárov)**.[488] Uvedené vymedzenie je ďalej pertraktované ďalšími povinnosťami, že sú

1. povinní zaobstarať si a pri rozhodovaní zohľadniť všetky dostupné informácie týkajúce sa predmetu rozhodnutia,
2. zachovávať mlčanlivosť o dôverných informáciách a skutočnostiach, ktorých prezradenie tretím osobám by mohlo spoločnosti spôsobiť škodu alebo ohroziť jej záujmy alebo záujmy jej spoločníkov (akcionárov), a
3. pri výkone svojej pôsobnosti nesmú uprednostňovať svoje záujmy, záujmy len niektorých spoločníkov (akcionárov) alebo záujmy tretích osôb pred záujmami spoločnosti.

[487] Pozri POWERS, S. *AI in Decision Making* [online]. Dostupné na: https://www.youtube.com/watch?v=RxYuh2alfTA.
[488] § 135a, § 194 ods. 5 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.

V rámci doktríny je odborná starostlivosť chápaná ako očakávaná miera **informovanosti** a **odbornosti** štatutárneho orgánu v podobe objektívnych osobných kvalít, ktoré má spĺňať každá osoba pôsobiaca vo funkcii člena štatutárneho orgánu.[489] Kategória odbornej starostlivosti je zmiernená doktrínou *„business judgement rule"*, podľa ktorej má štatutárny orgán právo na podnikateľské rozhodnutie a na omyl, pokiaľ dodrží zásady odbornej starostlivosti, to znamená, že nekoná podvodne, vo vlastnom záujme, a že si zaobstaral všetky dostupné informácie a vyhodnotil riziká svojho rozhodnutia.

Okrem členov štatutárnych orgánov sú aj členovia dozornej rady povinní vykonávať svoju funkciu s náležitou (odbornou) starostlivosťou a v súlade so záujmami spoločnosti a všetkých jej spoločníkov (akcionárov), t.j. s potrebnou mierou informovanosti a odbornosti,[490] a to na podklade § 139 ods. 4 ObZ, podľa ktorého *„Na členov dozornej rady sa vzťahuje zákaz konkurencie (§ 136) a ustanovenia § 135a[491] sa použijú primerane."*[492]

V Slovenskej republike absentuje právna regulácia i právna definícia AI. Jediným prameňom, ktorý sa AI zaoberá aj vo väzbe na podnikateľské rozhodovanie je legislatívny zámer rekodifikácie práva obchodných spoločností z mája 2021, ktorý v nadväznosti na danú problematiku uvádza

*„**V blízkej budúcnosti bude potrebné reagovať aj na využitie umelej inteligencie pri podnikateľskom rozhodovaní či iných procesoch rozhodovania v rámci obchodnej spoločnosti**, a to **vytvorením právneho rámca jej použitia pri zachovaní zodpovednosti členov orgánov za kontrolu jeho riadneho fungovania**."*[493]

Z danej konštatácie vyplýva, že zákonodarca plánuje *de lege ferenda* vytvoriť právny rámec využívania systémov AI v podnikateľskom rozhodovaní obchodných spoločností,

---

[489] MAMOJKA, M. In: MAMOJKA, M. a kol.: *Obchodný zákonník – Veľký komentár. 1. zväzok* – Bratislava: Vydavateľstvo Eurokódex, 2016. s. 539

[490] Tamže. s. 548

[491] § 135a ods. 1 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov *„Konatelia sú povinní vykonávať svoju pôsobnosť s odbornou starostlivosťou a v súlade so záujmami spoločnosti a všetkých jej spoločníkov. Najmä sú povinní zaobstarať si a pri rozhodovaní zohľadniť všetky dostupné informácie týkajúce sa predmetu rozhodnutia, zachovávať mlčanlivosť o dôverných informáciách a skutočnostiach, ktorých prezradenie tretím osobám by mohlo spoločnosti spôsobiť škodu alebo ohroziť jej záujmy alebo záujmy jej spoločníkov, a pri výkone svojej pôsobnosti nesmú uprednostňovať svoje záujmy, záujmy len niektorých spoločníkov alebo záujmy tretích osôb pred záujmami spoločnosti."*.

[492] § 139 ods. 4 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.

[493] Ministerstvo spravodlivosti. *Legislatívny zámer rekodifikácie práva obchodných spoločností*. Máj 2021. [online]. Dostupné na:https://www.justice.gov.sk/dokumenty/2022/02/Legislativny-zamer-ZoOSaS_2021.pdf.

s tým, že členovia orgánov budú mať v prípade využitia takýchto systémov zodpovednosť za kontrolu ich riadneho fungovania. To znamená, že pokiaľ bude takýmto systémom spôsobená škoda, zodpovedným za túto škodu bude člen orgánu (najmä štatutárneho a dozorného orgánu).[494]

To, že neexistuje právna regulácia ohľadom využívania systémov AI v podnikateľskom rozhodovaní obchodných spoločností však neznamená, že ju využiť nemožno, práve naopak *štatutárny orgán/dozorný orgán môže podľa nášho názoru časť svojej pôsobnosti vertikálne delegovať na systémy AI*, mal by pamätať na nasledujúce dôležité povinnosti a zodpovednosti:

a) ***Zodpovednosť za výber*** – pri výbere vhodnej technológie AI musí postupovať s odbornou/náležitou starostlivosťou a uskutočniť výber aspoň na takej úrovni, ako by výber uskutočnila iná rozumne starostlivá osoba;

b) ***Zodpovednosť za zadanie, vedenie a súčinnosť*** – musí vybranej technológii AI zadefinovať zadanie jasne a zrozumiteľne, poskytnúť všetku potrebnú súčinnosť a dávať jej pokyny resp. informácie/dáta;
c) ***Zodpovednosť za kontrolu*** – musí byť schopný dodaný výsledok delegovanej pôsobnosti primerane kontrolovať a rozumne vyhodnocovať, a to nielen osobne, ale aj za pomoci riadne nastavených kontrolných mechanizmov.[495]

Okrem vyššie uvedeného by štatutárne orgány mali zabezpečiť, aby využívanie systémov AI v podnikateľskom rozhodovaní obchodných spoločností bolo v zmysle vyššie popísaných *Etických smerníc pre dôveryhodnú umelú inteligenciu* transparentné a dôveryhodné, spĺňajúce nasledovné kritériá prostredníctvom odpovedí na jednotlivé otázky:

a) *Vysledovateľnosť: Zaviedli ste opatrenia, ktoré môžu zabezpečiť vysledovateľnosť?* Tieto opatrenia si môžu vyžadovať zdokumentovanie:
   - metód použitých na navrhnutie a vývoj algoritmického systému:

[494] Zrejme obdobne ako v súčasnosti podľa § 135a, § 194 ods. 5 zákona č. 513/1991 Zb. Obchodný zákonník v znení neskorších predpisov.
[495] Rozsudok Najvyššieho súdu Českej republiky zo dňa 30. 9. 2019, sp. zn. 27 Cdo 90/2019.

- v prípade generatívneho systému umelej inteligencie založeného na pravidlách by sa mala zdokumentovať metóda programovania alebo spôsob, akým bol vytvorený model,
- v prípade systému umelej inteligencie založeného na strojovom učení sa by sa mala zdokumentovať metóda výcviku algoritmu vrátane toho, ktoré vstupné údaje boli zhromaždené a vybraté, a toho, ako došlo k ich získaniu a výberu,

- metód použitých na testovanie a validáciu algoritmického systému:
  - v prípade generatívneho systému umelej inteligencie založeného na pravidlách by sa mali zdokumentovať scenáre alebo prípady použité na testovanie a validáciu,
  - v prípade systému umelej inteligencie založeného na strojovom učení sa by sa mali zdokumentovať informácie o údajoch použitých na testovanie a validáciu,
- výsledkov algoritmického systému:
  - zdokumentovať by sa mali výsledky algoritmu alebo rozhodnutia prijaté na jeho základe, ako aj prípadné ďalšie rozhodnutia, ktoré vyplývajú z iných prípadov (napr. v prípade iných podskupín používateľov).

b) *Vysvetliteľnosť:*

- Posúdili ste mieru, do akej sú rozhodnutia systému umelej inteligencie a v dôsledku nich aj jeho výsledky zrozumiteľné?
- Presvedčili ste sa o tom, že vysvetlenie, prečo systém uskutočnil istú voľbu vedúcu k určitému výsledku, bude zrozumiteľné?
- Posúdili ste, do akej miery rozhodnutie systému ovplyvňuje rozhodovacie procesy organizácie?
- Posúdili ste, prečo bol v tejto konkrétnej oblasti zavedený daný systém?
- Posúdili ste obchodný model súvisiaci s týmto systémom (napr. ako vytvára hodnotu pre organizáciu)?

c) *Komunikácia*:

- Oznámili ste používateľom (napr. ostatným členom štatutárneho orgánu, zamestnancom prípadne zákazníkom), že nekomunikujú s iným človekom, ale so systémom umelej inteligencie?
- Zaviedli ste mechanizmy na informovanie používateľov o dôvodoch a kritériách, ktoré tvoria základ výsledkov systému umelej inteligencie?
- Viedli ste zrozumiteľnú komunikáciu o vlastnostiach, obmedzeniach a možných nedostatkoch systému umelej inteligencie.[496]

8.3 Správa, riadenie a kontrola obchodných spoločností a AI

Európska komisia k správe a riadeniu obchodných spoločností vo všeobecnosti uvádza:

„správa a riadenie spoločnosti definuje vzťahy medzi vedením spoločnosti, jej predstavenstvom, jej akcionármi a ostatnými zainteresovanými stranami. Určuje spôsob riadenia a kontroly spoločností.[...] Dobrá správa a riadenie spoločnosti je v prvom rade zodpovednosťou príslušnej spoločnosti, pričom pravidlá na európskej úrovni a vnútroštátnej úrovni sú zavedené, aby sa zabezpečilo dodržiavanie určitých noriem"[497]

Európska komisia ďalej v rámci *štúdie o význame a vplyve umelej inteligencie na právo obchodných spoločností a správu a riadenie spoločností* špecifikovala možné oblasti využitia AI v rámci správy a kontroly spoločností, a to nasledovné:[498]

1. ***Rozhodovacie procesy:***

a) *Vymenovanie orgánov spoločnosti:* Menovanie členov štatutárnych orgánov/riaditeľov (výkonných, nevýkonných, funkcionárov), členov dozorných orgánov, manažérov, audítorov, delegovanie činností na členov štatutárnych orgánov alebo manažérov.

PRÍKLAD

[496] Expertná skupina na vysokej úrovni pre umelú inteligenciu (HLEG AI) (2019). *Etické usmernenia pre dôveryhodnú umelú inteligenciu.* Správa pre Európsku komisiu, 9. apríla. Brusel: Európska komisia. S. 35 – 37. Dostupné z: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[497] European Commission (2012), *'Action Plan: European Company Law and Corporate Governance - a modern legal framework for more engaged shareholders and sustainable companies'*, COM (2012)740 final, pp. 2-3, Strasbourg, available at https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0740:FIN:EN:PDF.

[498] European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, p. 12, https://data.europa.eu/doi/10.2838/790784-

Nástroj AI by sa dal použiť napríklad na identifikovanie a výber príslušných osôb s odbornými znalosťami a zručnosťami potrebnými pre danú funkciu člena štatutárneho/dozorného orgánu, manažéra, a pod.

V rámci delegácie činností by AI mohla pomôcť tým, že bude zohrávať úlohu v rozhodovacom procese štatutárneho orgánu (alebo dokonca nahradení členov štatutárneho orgánu v ich funkciách, ak je to možné a právne povolené), ako aj zastúpenie spoločníkov/akcionárov pri výkone ich hlasovacích práv.[499]

b) *Prijatie alebo vykonanie iných rozhodnutí spoločnosti:* Príprava a schvaľovanie správ účtovnej závierky a iných povinností spoločností týkajúcich sa povinného zverejňovania dokumentov (finančné a nefinančné), prijatie kódexov správania, algoritmické alebo vysokofrekvenčné obchodovanie s akciami, obchodná stratégia a definícia záujmov spoločností, dlhodobý výkon, analýza vplyvu na udržateľnosť a procesy náležitej starostlivosti, sledovateľnosť v zásobovacích reťazcoch.

PRÍKLAD

Technológie AI v tejto podoblasti by sa mohli použiť na analýzu spoločnosti z pohľadu finančných údajov a zásad správania, ktoré sa majú uplatňovať, automatizáciu podávania správ a finančných transakcií a v širšom zmysle na zber, autonómnu klasifikáciu a ďalšie analýzy údajov z rôznych oblastí.[500]

c) *Manažérske aktivity:* Schvaľovanie transakcií (vrátane mimoriadnych operácií) alebo iné dohody alebo akékoľvek iné rozhodovanie, výber tretej strany, investičné rozhodnutia,

[499] Tamže. s. 17

[500] Ku konkrétnym nástrojom AI bližšie pozri: European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, p. 17 - 19, https://data.europa.eu/doi/10.2838/790784.

zapojenie spoločníkov/akcionárov.

PRÍKLAD
Napríklad nástroje AI by mohli zahŕňať podporu v rozhodovacom procese pri schvaľovaní transakcií, pri preverovaní tretích strán, analýze rizík a dopadov udržateľnosti spoločnosti, efektívneho zdieľania informácií s akcionármi a prevádzkovania platforiem hlasovania na valných zhromaždeniach, ako aj podporu virtuálnych valných zhromaždení.[501] Nástrojom AI v tejto oblasti je napr. Fireflies.ai, vznikajúca platforma AI meniaca spôsob, akým spoločnosti vedú stretnutia orgánov a interpretujú údaje. Ponúka pokročilého AI asistenta stretnutí, ktorý sa môže zúčastňovať stretnutí, hovorov, robiť si poznámky a poskytovať prepisy/zápisnice, čím efektívne zachytáva dôležité detaily a informácie. Fireflies.ai podporuje všetky populárne aplikácie vrátane Google Meet, Zoom, Teams Webex, Ringcentral a Aircall.[502]

2. Monitorovanie a kontrola súladu s právnymi predpismi:

a) *Dohľad a činnosti súladu:* Riadenie podnikových rizík (ERM), dohľad a ochrana práv zamestnancov a iných záujmov zainteresovaných strán, monitorovanie ukazovateľov udržateľnosti a procesov *due diligence*, audítorská činnosť ako súčasť dohľadu a kontroly, monitorovanie relevantných parametrov pre riziká platobnej neschopnosti, súlad so zákonmi, podmienkami náležitej starostlivosti, oznamovacími povinnosťami voči úradom, zodpovednosť členov štatutárnych orgánov/riaditeľov, monitorovanie dodržiavania požiadaviek práva obchodných spoločností (predkladanie a obsah ročných účtovných dokladov, správ a iných podobných dokumentov požadovaných právom obchodných spoločností atď.).[503]

[501] Ku konkrétnym nástrojom AI bližšie pozri: European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, p. 20, https://data.europa.eu/doi/10.2838/790784.

[502] Pozri Firelies AI. Dostupné na: https://fireflies.ai/?fpr=hemin&gclid=CjoKCQiAtaOtBhCwARIsAN_x-3Lkj2nkiehHvBfUP7BnQWg2ah4NGSOr9fle_RcQ2AFPJOqTlPXiKMEaAgAIEALw_wcB alebo https://www.youtube.com/watch?v=H5fdzoAeLCE&t=21s.

[503] Ku konkrétnym nástrojom AI bližšie pozri: European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, s. 20 – 22. Dostupné na: https://data.europa.eu/doi/10.2838/790784.

b) *Založenie/registrácia spoločností/podania:* proces zakladania a registrácie spoločností, proces podávania/registrácie zmien údajov o spoločnostiach (napr. nový člen štatutárneho orgánu - vymenovanie, zvýšenie základného imania, zmena stanov), proces domácich alebo cezhraničných premien, rozdelení alebo zlúčení, tvorba zakladateľských dokumentov a stanov spoločností, využitie štatistických nástrojov pre údaje v obchodných registroch.[504]

PRÍKLAD
Používanie AI v tejto oblasti by mohlo byť relevantné medzi podnikovými právnikmi, notármi alebo advokátmi v rámci prípravy podkladov k podaniam a kontrolu informácií pred podaniami. Využiteľnosť týchto AI nástrojov môže byť aj pre obchodný register alebo iné štátne orgány majúce v rámci svojej pôsobnosti povinnosť prijímať, kontrolovať a overovať podania od spoločností, ako aj na analýzu podnikových údajov.[505]

8.3.1 Právny základ využívania systémov umelej inteligencie v správe, riadení a kontrole obchodných spoločností v Slovenskej republike

Je nesporným faktom, že moderné právo obchodných spoločností, ale aj korporačnú prax podstatne zasiahla elektronizácia, digitalizácia a automatizácia. Ako by však do budúcna mala vyzerať právna úprava správy, riadenia a kontroly spoločností v SR vo väzbe na možnosti využitia umelej inteligencie? Na to čiastočne odpovedá zákonodarca v legislatívnom zámere rekodifikácie práva obchodných spoločností z mája 2021, ktorý uvádza

*„Musí umožňovať **flexibilitu pri nastavení procesov správy a riadenia spoločnosti**, **zvýšiť a zjednodušiť zapojenie spoločníkov do správy a riadenia** vrátane **digitalizácie rozhodovacích procesov** a nastaviť dostatočne efektívne mechanizmy ochrany dotknutých subjektov, najmä veriteľov spoločnosti."*[506]

[504] Tamže.

[505] Ku konkrétnym nástrojom AI bližšie pozri: European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, s. 22 – 23. Dostupné na: https://data.europa.eu/doi/10.2838/790784.

[506] Ministerstvo spravodlivosti. *Legislatívny zámer rekodifikácie práva obchodných spoločností*. Máj 2021. [online]. Dostupné na:https://www.justice.gov.sk/dokumenty/2022/02/Legislativny-zamer-ZoOSaS_2021.pdf

Po vymedzení možného obsahového využitia AI v rámci správy, riadenia a kontroly spoločnosti sa v rámci tejto podkapitoly bližšie zameriame na zaujímavé právne polemiky, a to:

1. Systém AI – humanoidný robot ako člen štatutárneho, dozorného orgánu obchodných spoločností.
2. Systém AI využitý na hlasovanie spoločníkov/akcionárov na valnom zhromaždení a s tým súvisiaca možnosť uskutočniť valné zhromaždenie virtuálne na diaľku.
3. Systém AI uľahčujúci proces založenia, registrácie, zmien údajov obchodných spoločností.

8.3.1.1 Systém AI – humanoidný robot ako člen štatutárneho, dozorného orgánu obchodných spoločností.

V štúdii Svetového ekonomického fóra z roku 2015, ktorá skúmala viac ako 800 IT manažérov, 45 % respondentov očakávalo, že uvidíme prvý systém AI v predstavenstve obchodnej spoločnosti do roku 2025 a že takýto prelom by bol zlomovým bodom.[507]

Ako sme už vyššie uviedli poľská spoločnosť Dictador Europe Sp. z o.o. má na pozícii CEO – generálneho riaditeľa humanoidného robota menom Mika.[508] Mika je výsledkom spoločného výskumného projektu medzi Dictadorom a Hanson Robotics, renomovanou strojárskou a robotickou spoločnosťou so sídlom v Hongkongu, ktorá je známa vývojom robotov podobných ľuďom s pokročilou AI. Mika je robot poháňaný umelou inteligenciou, ktorý bol naprogramovaný tak, aby stelesňoval osobitého ducha a základné hodnoty Dictadoru. Marek Szoldrowski, prezident Dictador Europe po jej vymenovaní uviedol *„Rozhodnutie predstavenstva Dictador je revolučné a odvážne zároveň. Tento prvý robot podobný človeku s AI v štruktúre spoločnosti navždy zmení svet, ako ho poznáme."*[509]

[507] Pozri PUGH, W. *Why Not Appoint an Algorithm to Your Corporate Board?* [online]. Dostupné na: https://slate.com/technology/2019/03/artificial-intelligence-corporate-board-algorithm.html
[508] Pozri *What Do You Think of This AI Robot 'CEO'?* Dostupné na: https://www.youtube.com/watch?v=8BQEyQ2-awc
[509] BHAVYA, S. *Meet 'Mika', The World's First Ever AI Humanoid Robot CEO* [online]. Dostupné na: https://www.ndtv.com/feature/meet-mika-the-worlds-first-ever-ai-humanoid-robot-ceo-4567460.

Mika vlastnými slovami zdôrazňuje zdatnosť svojej pokročilej umelej inteligencie a algoritmov strojového učenia, ktoré jej umožňujú rýchlo a presne vytvárať rozhodnutia založené na údajoch, keď uvádza *„Môj rozhodovací proces sa opiera o rozsiahlu analýzu údajov a zosúladenie so strategickými cieľmi spoločnosti."*[510] Dôležité je, že Mika pôsobí bez osobnej zaujatosti a zabezpečuje, že jej strategické rozhodnutia sú zamerané na uprednostňovanie najlepších záujmov spoločnosti Dictador Europe Sp. z o.o.

Vymenovanie Miky za generálnu riaditeľku je významným míľnikom vo vývoji umelej inteligencie. Je to znak toho, že AI je čoraz sofistikovanejšia a schopnejšia a že je pripravená hrať dôležitejšiu úlohu v riadení obchodných spoločností. Je to tiež znak toho, že spoločnosti sú čoraz otvorenejšie myšlienke využitia AI na zlepšenie svojich operácií.

Naskytá sa nám však dôležitá právna otázka či **môže byť v SR členom štatutárneho, dozorného orgánu systém AI (humanoidný robot)?** Súčasná slovenská legislatíva stanovuje, že členmi orgánov spoločností môžu byť len fyzické[511] prípadne právnické osoby disponujúce spôsobilosťou na práva a povinnosti a spôsobilosťou na právne úkony. AI nie je ani fyzická ani právnická osoba, a v súčasnosti nemá právnu subjektivitu konať ako fyzická alebo právnická osoba. **Na túto otázku teda musíme odpovedať záporne, že členom štatutárneho, dozorného orgánu v SR nemôže byť systém AI (humanoidný robot). To znamená, že systém AI nemôže v súčasnosti legálne vykonávať všetky činnosti, ktoré môže vykonávať fyzická alebo právnická osoba vo funkcii člena štatutárneho/dozorného orgánu obchodnej spoločnosti. To platí vo všeobecnosti (napr. stroj alebo systém AI nemôže platne podpísať zmluvy, AI nemôže platne založiť spoločnosť, vymenovať alebo byť vymenovaný za člena štatutárneho/dozorného orgánu, predkladať účtovnú závierku a pod.). Umelá inteligencia však môže byť použitá ako pomocný nástroj, ktorý pomáha ľuďom pri vykonávaní určitej úlohy, rozhodovaní a pod.**

Do budúcna príležitosti spočívajú v delegovaní pôsobnosti na systém AI spôsobilý autonómne rozhodovať a konať v rámci spoločnosti (t. j. bez ľudského zásahu alebo súhlasu) v prípadoch, keď existuje správne nastavený regulačný rámec v rámci spoločnosti.

[510] Tamže.

[511] Napríklad podľa § 133 ods. 2 zákona č. 513/1991 Zb. Obchodný zákonník: *„Konateľom spoločnosti môže byť len fyzická osoba, ktorá nie je v čase vykonania zápisu do obchodného registra ako povinný vedená v registri poverení na vykonanie exekúcie podľa osobitného zákona."*

Rozhodnutia sa môžu týkať podnikania ako celku, prípadne jednotlivých oddelení, manažérov a zamestnancov. Táto oblasť potenciálneho využitia AI zahŕňa projektové riadenie s autonómnym systémom overujúcim priebeh, aktualizovanie prognóz v čase a náklady na realizáciu projektu, prerozdelenie zdrojov alebo riadenie zásob.[512]

Domnievame sa, že v budúcnosti bude zavedená možnosť, aby systém AI pôsobil vo funkcii člena štatutárneho, dozorného orgánu obchodnej spoločnosti. V Európskom parlamente sa viedli diskusie o potenciálnej právnickej osobe „*AI*". Podľa niektorých názorov totiž vznik špecifického právneho statusu pre roboty a prisúdenie právnej subjektivity AI môže byť prospešné. Navrhovalo sa zvážiť priznanie inteligentným robotom status *„elektronické osoby*" zodpovedné za prípadné škody. Uznesenie Európskeho parlamentu zo 16. februára 2017 obsahujúce odporúčania pre Komisiu pre občianskoprávne pravidlá o robotike odkazuje na možnosť vytvorenia špecifického právneho postavenia pre roboty v dlhodobom horizonte, aplikáciou „*elektronickej osobnosti*" na prípady, keď roboty robia autonómne rozhodnutia alebo inak interagujú s treťou osobou nezávisle.[513]

8.3.1.2 Systém AI využitý na hlasovanie spoločníkov/akcionárov na valnom zhromaždení a s tým súvisiaca možnosť uskutočniť valné zhromaždenie virtuálne na diaľku.

Právna úprava *lex specialis* v čase pandémie COVID-19

Na rokovaní valného zhromaždenia majú právo zúčastniť sa a hlasovať všetci spoločníci (akcionári). Práve osobná účasť väčšieho počtu spoločníkov na valnom zhromaždení sa ukázala ako problematická v čase pandémie COVID-19 vzhľadom na epidemiologické pravidlá obmedzujúce stýkanie sa ľudí. Zákonodarca reagoval na nastolený problém ako zabezpečiť účasť na valnom zhromaždení bez fyzickej prítomnosti spoločníkov prostredníctvom § 5 zákona č. 62/2020 Z. z. o niektorých mimoriadnych opatreniach v súvislosti so šírením nebezpečnej nákazlivej ľudskej choroby Covid-19 a v justícii (ďalej len § 5

[512] Pozri správy o riešeniach pre inteligentnú správu zásob, ako sú tie, ktoré sú k dispozícii na https://www.impomag.com/inventorymanagement/article/13249224/inventory-management-and-artificial-intelligence alebo https://www.unleashedsoftware.com/blog/how-inventory-control-can-benefit-from-artificial-intelligence.

[513] European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, p. 37, Dostupné na: https://data.europa.eu/doi/10.2838/790784.

zákona o niektorých mimoriadnych opatreniach).[514] Podľa samotnej dôvodovej správy k predmetnému ustanoveniu bolo cieľom návrhu zákona v prípade § 5 umožniť právnickým osobám založeným podľa predpisov občianskeho práva a obchodného práva, aby ich kolektívne orgány mohli prijímať svoje rozhodnutia aj *per rollam*, čím sa minimalizuje nutnosť zhromažďovania sa väčšieho počtu osôb na jednom mieste.[515] Neskôr bol § 5 zákona o niektorých mimoriadnych opatreniach novelizovaný zákonom č. 9/2021 Z. z., ktorá zmenila predmetný § 5 zákona o niektorých mimoriadnych opatreniach, nasledovne:

1. ***Kolektívne orgány právnických osôb*** *založených podľa predpisov občianskeho práva alebo obchodného práva môžu* ***v čase mimoriadnej situácie alebo núdzového stavu používať korešpondenčné hlasovanie*** *alebo* ***umožniť účasť ich členov na zasadnutí takéhoto orgánu prostredníctvom elektronických prostriedkov****, aj keď to nevyplýva z ich vnútorných predpisov alebo stanov.*
2. *Ak podmienky rozhodovania podľa odseku 1 nevyplývajú zo zákona, vnútorných predpisov alebo stanov****, určí ich v prípade, ak ide o najvyšší orgán právnickej osoby, štatutárny orgán konajúci s náležitou starostlivosťou****, a v prípade iného orgánu ich určí tento orgán.* ***Tieto podmienky musia byť členom orgánu oznámené v dostatočnom predstihu pred rozhodovaním.***

Z daného ustanovenia § 5 zákona o niektorých mimoriadnych opatreniach vyplývajú nasledovné ťažiskové závery:

1. Vzťahuje sa na **všetky kolektívne orgány právnických osôb** založených nielen podľa predpisov obchodného práva, ale i občianskeho práva. V rámci právnických osôb založených podľa noriem obchodného práva (najmä obchodné spoločnosti) to budú predovšetkým zasadnutia valného zhromaždenia, ale i zasadnutia štatutárneho orgánu, prípadne dozornej rady.

[514] Tento pôvodne znel nasledovne: ,,*Kolektívne orgány právnických osôb založených podľa predpisov občianskeho práva alebo obchodného práva môžu v čase mimoriadnej situácie alebo núdzového stavu používať korešpondenčné hlasovanie alebo umožniť účasť ich členov na zasadnutí takéhoto orgánu prostredníctvom elektronických prostriedkov, aj keď to nevyplýva z ich vnútorných predpisov alebo stanov. Ustanovenia § 190a až §190d Obchodného zákonníka sa použijú primerane.*" Pôvodné znenie § 5 zákona o niektorých mimoriadnych opatreniach využívalo existujúcu infraštruktúru ustanovení § 190a až §190d Obchodného zákonníka, vzťahujúcich sa na verejnú akciovú spoločnosť, ktoré sa v týchto prípadoch mali použiť primerane.

[515] Dôvodová správa k § 5 zákona č. 62/2020 Z. z. o niektorých mimoriadnych opatreniach v súvislosti so šírením nebezpečnej nákazlivej ľudskej choroby Covid-19 a v justícii.

2. Je naviazané na stav **mimoriadnej situácie alebo núdzového stavu,**[516] pričom sa môže aplikovať v takom rozsahu z hľadiska územnej pôsobnosti v akom je vyhlásená mimoriadna situácia alebo núdzový stav.
3. V rámci realizácie rozhodovania kolektívnych orgánov umožňuje výber medzi **korešpondenčným hlasovaním prostredníctvom pošty** alebo **účasťou členov na zasadnutí takéhoto orgánu prostredníctvom elektronických prostriedkov**, a to aj vtedy, keď to nevyplýva z ich vnútorných predpisov alebo stanov. Uvedené znamená, že aj pri postupe podľa § 5 zákona o mimoriadnych opatreniach bude zachovaný rovnaký výsledok rozhodnutia kolektívnych orgánov (napríklad budú prijaté uznesenia valného zhromaždenia, rozhodnutia alebo uznesenia predstavenstva, prípadne dozorného orgánu).

V rámci konania valného zhromaždenia obchodných spoločností podľa § 5 zákona o niektorých mimoriadnych opatreniach potom platí, že:

- Pri korešpondenčnom hlasovaní bude valné zhromaždenie prebiehať aj bez osobnej účasti spoločníkov, pričom lehota, do ktorej by pri hlasovaní mali byť spoločnosti doručené hlasovacie lístky bude podľa nášho názoru zároveň určením času konania valného zhromaždenia.
- Valné zhromaždenie umožňujúce účasť členov prostredníctvom elektronických prostriedkov sa bude realizovať prostredníctvom aktívnej účasti spoločníkov v reálnom čase prostredníctvom elektronických prostriedkov. Elektronické prostriedky, pritom môžu byť akékoľvek komunikačné kanály spĺňajúce to, aby mohol byť spôsob výkonu hlasovacích práv spoločníkov či výsledok rozhodnutia kontrolovateľný v rovnakom rozsahu ako v bežnom režime konania valného zhromaždenia za osobnej účasti spoločníkov.

Zákonodarca tiež ustanovil, že **š**tatutárny orgán má konať v súlade s náležitou starostlivosťou, pričom má nielen právo, ale i povinnosť počas mimoriadnej situácie

[516] Mimoriadna situácia je obdobie ohrozenia alebo obdobie pôsobenia následkov mimoriadnej udalosti na život, zdravie alebo majetok. Dostupné na: https://www.minv.sk/?Krizove_stavy.
Podľa čl. 5 Ústavného zákona č. 227/2002 Z. z. o bezpečnosti štátu v čase vojny, vojnového stavu, výnimočného stavu a núdzového stavu „*Núdzový stav môže vláda vyhlásiť len za podmienky, že došlo alebo bezprostredne hrozí, že dôjde k ohrozeniu života a zdravia osôb, a to aj v príčinnej súvislosti so vznikom pandémie, životného prostredia alebo k ohrozeniu značných majetkových hodnôt v dôsledku živelnej pohromy, katastrofy, priemyselnej, dopravnej alebo inej prevádzkovej havárie; núdzový stav možno vyhlásiť len na postihnutom alebo na bezprostredne ohrozenom území.*".

organizačne zabezpečiť konanie valného zhromaždenia tak, že nediskriminačným spôsobom vylúči možnosť osobnej účasti všetkých spoločníkov (akcionárov) a umožní im hlasovanie na valnom zhromaždení výlučne korešpondenčným hlasovaním alebo hlasovaním za použitia elektronických prostriedkov. Samotné podmienky realizácie valného zhromaždenia musia byť spoločníkom/akcionárom oznámené v dostatočnom predstihu pred rozhodovaním.

Stav *de lege lata*

Popisovaná právna úprava § 5 zákona o mimoriadnych opatreniach slúži ako cenný zdroj skúseností a inšpirácií avšak v súčasnej dobe už nie je uplatniteľná z dôvodu, že mimoriadna situácia v SR bola odvolaná 15. septembra 2023 ku 06.00 hod. uznesením vlády SR č. 446 z 13. septembra 2023. Súčasný stav možností vzdialeného výkonu práv spoločníkov na valných zhromaždeniach je v zmysle aktuálneho znenia Obchodného zákonníka nasledovný:

- v prípade spoločnosti s ručením obmedzeným ustanovenie § 126 Obchodného zákonníka ráta výlučne s možnosťou prezenčnej účasti na valnom zhromaždení alebo v zastúpení splnomocnencom na základe písomného plnomocenstva;
- rovnako je tomu tak v prípade súkromnej akciovej spoločnosti (§ 184 Obchodného zákonníka);
- v prípade verejnej akciovej spoločnosti je možnosť korešpondenčného hlasovania a účasti na zasadnutí elektronickými prostriedkami predmetom úpravy v rámci §§ 190a až 190d Obchodného zákonníka. Korešpondenčné hlasovanie sa vykonáva prostredníctvom hlasovacieho lístka, na ktorom musí byť úradne osvedčený podpis. Pokiaľ ide o účasť akcionára na valnom zhromaždení elektronicky, teda v reálnom čase, prípadný výkon hlasovacieho práva akcionára musí byť podpísaný kvalifikovaným elektronickým podpisom, ktorého kvalifikovaný certifikát musí spĺňať podmienky, ktoré sa vyžadujú v styku s orgánmi verejnej moci, a opatrený kvalifikovanou elektronickou časovou pečiatkou.
- Pokiaľ ide o možnosť *per rollam* hlasovania, to sa pripúšťa iba v rámci spoločnosti s ručením obmedzeným (§ 130 Obchodného zákonníka) a v rámci vyššie spomínanej regulácie verejnej akciovej spoločnosti.

Stav *de lege ferenda*

Samotný legislatívny zámer rekodifikácie práva obchodných spoločností z mája 2021 predpokladá do budúcnosti zavedenie zmien v nadväznosti na digitalizáciu rozhodovacích procesov v obchodných spoločnostiach a možnosť vykonávať práva spoločníka/akcionára na diaľku,

*„podporuje sa **možnosť výkonu práv spoločníkov na diaľku** pri zvýšení právnej istoty pokiaľ ide o možnosti a podmienky výkonu práva spoločníkov na diaľku. Väčšina dokumentov dnešného sveta už necirkuluje v listinnej podobe, ale na informačných komunikačných prostriedkoch. Zároveň rastie **význam online stretnutí** či **zasadnutí orgánov**. Tomu by nemal brániť ani právny poriadok svojou rigídnosťou, keď pripúšťa účasť a hlasovanie na valnom zhromaždení prostredníctvom elektronických prostriedkov len v rámci úpravy verejnej akciovej spoločnosti (§ 190d). Vývoj informačných prostriedkov je veľmi rýchly, preto právna úprava by mala stanovovať kritériá, za splnenia ktorých je možné využiť akýkoľvek informačný prostriedok komunikácie na prijímanie rozhodnutí v obchodnej spoločnosti."*[517]

Zákonodarca v budúcnosti plánuje umožniť dištančné zasadnutia valných zhromaždení a takisto vykonávanie práv spoločníka/akcionára na diaľku,[518] toto však podľa nášho názoru bude musieť mať určité zákonné podmienky, pri ktorých je možné využiť systémy AI, napríklad nasledovné systémy umožňujúce:

- dostatočné overenie totožnosti spoločníka, resp. jeho splnomocnenca;
- technické podmienky, ktoré zabezpečujú a umožňujú, že sa všetky zúčastnené osoby pri komunikácii navzájom majú možnosť počuť, resp. vyjadrovať sa;
- technické podmienky, ktoré umožnia hlasovanie v reálnom čase;
- pri prijímaní určitých závažnejších rozhodnutí môže byť na mieste vyžadovať určitú formalizáciu výstupu;

[517] Ministerstvo spravodlivosti. *Legislatívny zámer rekodifikácie práva obchodných spoločností*. Máj 2021. [online]. Dostupné na:https://www.justice.gov.sk/dokumenty/2022/02/Legislativny-zamer-ZoOSaS_2021.pdf.
[518] Pozri RIJMENAM, M. *How blockchain proxy-voting improves shareholder engagement* [online]. Dostupné na: https://www.thedigitalspeaker.com/blockchain-proxy-voting-improves-shareholder-engagement/.

- ochranu pred zneužitím týchto mechanizmov tretími osobami, napr. o zabezpečenie v podobe generovania prístupových kódov, osobitných kont a podobne.

8.3.1.3 Systém AI uľahčujúci proces založenia, registrácie, zmien údajov obchodných spoločností

Slovenská republika zvŕšila v roku 2020 elektronizáciu Obchodného registra SR, pričom jednou zo zmien je vylúčenie listinného spôsobu podávania návrhov, zmenu definuje § 5 ods. 2 zákona č. 530/2003 Z. z. o obchodnom registri a o zmene a doplnení niektorých zákonov. Od 1.10.2020 je zverejnená nová verzia elektronických formulárov na zápis do obchodného registra. Nastáva tak vylúčenie podávania listinných návrhov na zápis do obchodného registra a zavedenie výlučne v elektronickej podoby podávania prostredníctvom na to určenej služby na www.slovensko.sk. Návrh na zápis údajov do obchodného registra, návrh na zápis zmeny zapísaných údajov a návrh na výmaz zapísaných údajov (ďalej len „návrh na zápis") sa môže **podávať výlučne elektronickými prostriedkami registrovému súdu prostredníctvom elektronického formulára** zverejneného na webovom sídle ústredného portálu verejnej správy alebo špecializovaného portálu. Formuláre na zápis údajov do obchodného registra pre podania v elektronickej podobe sú dostupné na stránke: https://www.justice.gov.sk/Stranky/Obchodny-register-SR/Formulare-na-zapis-do-obchodneho-registra-pre-podania-v-elektronickej-podobe.aspx. Tieto formuláre využívajú nástroje umelej inteligencie spočívajúce v kontrole dokumentov detegujúce administratívne chyby (napr. správne vyplnený formulár, platnú adresu, a pod.).

Do budúcna by sa mohli nástroje AI rozšíriť najmä v prípadoch zabezpečenia *materiálneho preskúmavania zapisovaných údajov* (§ 7 zákona č. 530/2003 Z. z. o obchodnom registri v znení neskorších predpisov)[519] v podobe identifikácie rôznych problémov s

[519] Materiálne preskúmavanie zapisovaných údajov (§ 7 zákona č. 530/2003 Z. z. o obchodnom registri v znení neskorších predpisov) sa vykonáva len pri prvozápisoch obchodných spoločností a družstva, pri zápise zmeny obchodného mena (preskúmava sa totožnosť), pri zápisoch predmetu podnikania, pri zápise zahraničnej osoby ako osoby oprávnenej konať v mene podnikateľa... Pred zápisom spoločnosti s ručením obmedzeným registrový súd preverí aj to, či

a) spoločenská zmluva alebo zakladateľská listina obsahuje náležitosti podľa § 110 ods. 1 OBZ,
b) v spoločenskej zmluve nie je uvedených viac ako 50 spoločníkov,
c) výška základného imania, výška vkladu každého spoločníka a výška splateného vkladu každého spoločníka uvedená v spoločenskej zmluve alebo v zakladateľskej listine je v súlade s OBZ,
d) spoločnosť s ručením obmedzeným, ktorá je jediným zakladateľom spoločnosti, má viac spoločníkov,
e) fyzická osoba, ktorá je jediným zakladateľom spoločnosti, nie je jediným spoločníkom vo viac ako dvoch spoločnostiach s ručením obmedzeným,

dodržiavaním právnych predpisov (napr. obchodné meno novej spoločnosti porušujúcej ochrannú známku na tovar alebo služby v rovnakom sektore alebo kontrola dodržiavania pravidiel boja proti praniu špinavých peňazí). AI by tiež mohla overiť, či osoby (fyzické alebo právnické) medzi spoločníkmi/akcionármi nemajú zákaz zakladať spoločnosti, alebo či členovia štatutárnych a dozorných orgánov obchodných spoločností nemajú zákaz vykonávať funkciu v spoločnosti (v našej alebo inej jurisdikcii). Na rozšírenie nástrojov AI bude podľa legislatívneho zámeru rekodifikácie práva obchodných spoločností z mája 2021 potrebné,

„vytvorenie účinného prepojenia medzi jednotlivými systémami registrácie (obchodný register, centrálne depozitáre cenných papierov, register partnerov verejného sektora), ako aj efektívnejšie prepojenie medzi zdrojovými a referenčnými registrami (obchodný register, register právnických osôb, register fyzických osôb či register adries) a samozrejme aj zachovanie existujúcich prepojení na iné systémy verejnej správy (napr. register účtovných závierok).“[520]

Z hľadiska možných inšpirácií pre budúcnosť Obchodného registra SR stoja za zmienku nasledovné technológie AI využívané v zahraničí:

- nástroj AI v Estónskom obchodnom registri, ktorý predstavuje automatizovaný systém overovania vhodnosti obchodného mena spoločnosti založený na AI, ktorý je v štádiu testovania. Tento systém umožňuje používateľom identifikovať, či obchodné meno vybrané pre spoločnosť už náhodou nebolo zaregistrované (skúma sa totožnosť obchodného mena).[521]
- techniky strojového učenia používané v Dánskom obchodnom registri pomáhajúce znížiť daňové podvody právnických osôb. Ich cieľom je využiť strojové učenie na vykonanie a hodnotenie rizík nových spoločností na základe ich podaní a doplňujúcich informácií, čím sa vykonáva cielenejšia kontrola daňových podvodov založená na riziku. Využíva tiež

f) spoločenskú zmluvu podpísali všetci spoločníci a či je pravosť podpisov všetkých spoločníkov na spoločenskej zmluve osvedčená; ak spoločnosť založil jeden zakladateľ, či podpísal zakladateľskú listinu a či pravosť jeho podpisu na zakladateľskej listine je osvedčená,

g) k návrhu na zápis bol pripojený súhlas správcu dane podľa osobitného zákona, ak osoba, ktorá má byť zapísaná ako spoločník, je vedená v zozname daňových dlžníkov podľa osobitného zákona.

[520] Ministerstvo spravodlivosti. *Legislatívny zámer rekodifikácie práva obchodných spoločností.* Máj 2021. [online]. Dostupné na:https://www.justice.gov.sk/dokumenty/2022/02/Legislativny-zamer-ZoOSaS_2021.pdf.

[521] European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, s. 22. Dostupné na: https://data.europa.eu/doi/10.2838/790784.

techniky strojového učenia na predpovedanie možných neúspechov pomocou XBRL Financial vyhlásenia, ktoré zhromažďuje od spoločností pôsobiacich v Dánsku. Týmto spôsobom môže poskytnúť včasné varovanie podnikateľom, že ich spoločnosti vykazujú znaky, ktoré by mohli viesť ku reštrukturalizácii, úpadku alebo konkurzu.[522]

- UNA využívaná Lotyšským obchodným registrom predstavuje chatbota odpovedajúceho na často kladené otázky týkajúce sa procesu registrácie spoločností, prípadne otázky o stave podaní na registráciu.[523]

8.4 Uzavieranie obchodnoprávnych zmlúv

Umelá inteligencia postupne prináša revolúciu do spôsobu, akým sú zmluvy koncipované a kontrolované. Má schopnosť automaticky generovať zmluvy na základe preddefinovaných šablón a tiež analyzovať jazyk používaný v existujúcich zmluvách na označenie potenciálnych problémov a ponúkanie návrhov na zlepšenie. AI tiež pomáha pri zefektívňovaní procesu vyjednávania identifikáciou zmluvných doložiek, ktoré môžu byť sporné, a navrhovaním alternatív. Systémy umelej inteligencie môžu napríklad identifikovať ustanovenia, ktoré môžu byť nejednoznačné alebo konfliktné, upozorniť na potenciálne riziká. Pomáha to zefektívniť proces tvorby zmluvy, znížiť riziko chýb a zabezpečiť, aby si všetky zmluvné strany boli vedomé svojich povinností. Celkovo má AI obrovský potenciál zlepšiť efektívnosť, presnosť a kvalitu návrhov a procesov tvorby i kontroly zmlúv.

V súvislosti s touto podkapitolou sa ďalej zameriame na kľúčovú otázku: *mohla by umelá inteligencia nahradiť prácu právnika zameriavajúceho sa na tvorbu obchodnoprávnych zmlúv?* Odpoveď na túto nie je jednoduchá, nakoľko kontraktačný proces má viaceré aspekty. Obchodnoprávne zmluvy mnohokrát obsahujú štandardné klauzuly, ktorých účinok a často aj samotné znenie sú vo väčšine zmlúv rovnakého typu identické. Práve pri týchto štandardných klauzulách si viem predstaviť využitie AI, povedzme sa bude jednať o zmluvný softvér, ktorý dokáže identifikovať rozdiely medzi zmluvami lepšie ako človek, pričom bude mať databázu X vzorov štandardných zmlúv, a bude schopný identifikovať štandardné

[522] European Commission, Directorate-General for Justice and Consumers, *Study on the relevance and impact of artificial intelligence for company law and corporate governance – Final report*, Publications Office, 2021, p. 22 - 23, https://data.europa.eu/doi/10.2838/790784.
[523] Tamže.

ustanovenia prípadne chyby v nich. Takýto nástroj AI môže byť dokonca schopný navrhnúť zmeny a doplnenia na odchýlenie sa od štandardných klauzúl.

PRÍKLAD
Zmluvný softvér AI na uzatváranie zmlúv: Spoločnosť zameraná na umelú inteligenciu so sídlom v Spojenom kráľovstve, Luminance, predstavila 7.11.2023 prvú technológiu AI na svete s názvom Luminance Autopilot, ktorá úplne zautomatizuje vyjednávanie zmluvy bez ľudského zásahu medzi dvoma protichodnými stranami. Využíva pritom jedinečný právny model veľkého jazyka (LLM) na rozšírenie práce právnikov. Vďaka viac ako 150 miliónom právnych dokumentov v kombinácii s vlastnou umelou inteligenciou a hlbokými znalosťami právneho odvetvia z nej robí dnes najpokročilejšiu právnu LLM, ktorá pôsobí pre viac ako 500 zákazníkov na celom svete. Luminance Autopilot umožňuje zvládnuť každodenné rokovania o bežných zmluvách, ako sú dohody o mlčanlivosti. Umelá inteligencia môže čítať zmluvu, upraviť rizikové klauzuly a reagovať na akékoľvek zmeny vykonané umelou inteligenciou protistrany. To všetko sa deje automaticky, pričom AI využívajú znalosti z predchádzajúcich zmlúv konkrétneho používateľa.[524]

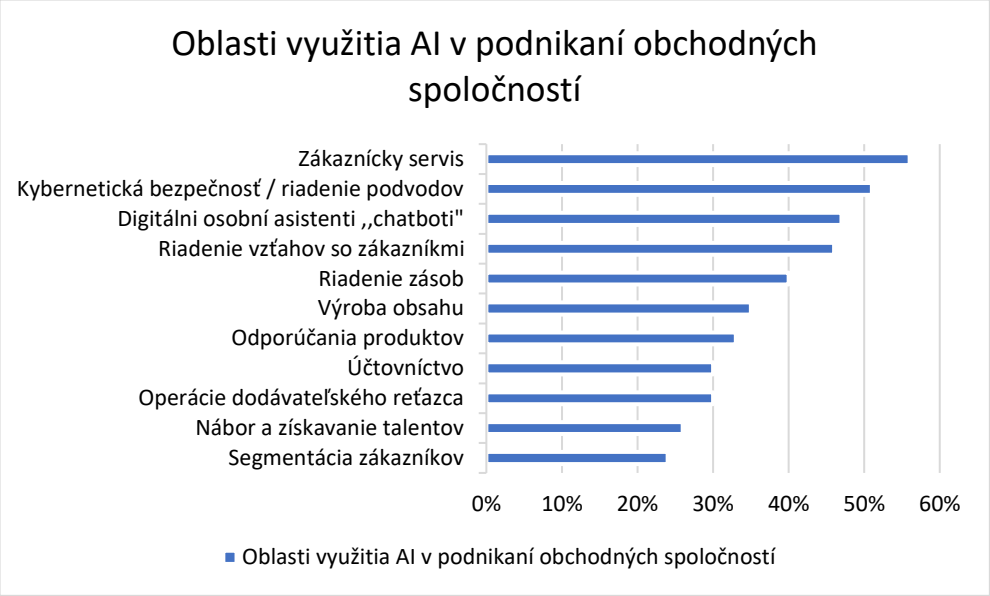
Na druhej strane aj typické obchodnoprávne zmluvy podliehajú pravidlám stanoveným právnymi predpismi a judikatúrou, napríklad ak ide o transakcie medzi podnikateľmi a spotrebiteľmi, a preto môže byť potrebná právna expertíza. Vedieť poradiť zmluvnej strane, či daná zmluva bude v jeho komerčných záujmoch alebo nie, umelá inteligencia zatiaľ nedokáže. Hodnotenie takéhoto rizika v konkrétnej situácii totiž nie je formou abstraktného kalkulu, pretože si vyžaduje pochopenie priemyselných, obchodných, ekonomických, sociálnych, právnych a ľudských faktorov konkrétneho prípadu. A samozrejme, nie všetky obchodnoprávne zmluvy sa dajú uzavrieť prostredníctvom štandardných formulárov. Niektoré situácie si preto vyžadujú „*zmluvy na mieru*" vypracované právnikmi. V nadväznosti na uvedené konštatujeme, že AI môže pomôcť právnikom pri tvorbe zmlúv avšak nie ich nahradiť úplne. Podľa nášho názoru softvéry AI môžu byť užitočné pri generovaní štandardných zmlúv a klauzúl, ale právnici sú stále potrební na to, aby zabezpečili, že zmluvy sú v súlade s platným právom a judikatúrou a že chránia záujmy

[524] Bližšie pozri: https://www.luminance.com/news/press/20231107_luminance_showcases.html alebo https://www.youtube.com/watch?v=9PKRfnstXfk.

zmluvných strán. Profesorka práva na Univerzite v Južnej Californii Gillian K. Hadfield, ktorá sa špecializuje na zmluvné právo, verí, že AI pri uzatváraní zmlúv povedie k lepšiemu využitiu právnických talentov: *„právnici presunú svoje zameranie z rutinných činností na oveľa vyššiu hodnotu práce spojenú s formovaním stratégií a navigáciou v zložitých právnych problémoch."*[525]

8.5 Optimalizácia úloh a vytváranie nových produktov a služieb

Spoločnosti využívajú umelú inteligenciu rôznymi spôsobmi, aby zlepšili efektivitu, ušetrili čas a znížili náklady. Forbes Advisor vykonal prieskum so vzorkou 600 spoločností s cieľom lepšie pochopiť ako spoločnosti využívajú alebo majú v najbližšej budúcnosti v pláne využiť AI v rámci podnikateľskej činnosti. Výsledky odhalili vplyv AI na oblasti, ako je kybernetická bezpečnosť, správa podvodov, produkcia obsahu, analýza spokojnosti zákazníkov a zákaznícky servis, vrátane používania chatbotov.[526]



Zdroj: Forbes Advisor[527]

Na vrchole sa umiestnil zákaznícky servis - pomoc a poradenstvo poskytované spoločnosťou tým ľuďom, ktorí kupujú alebo používajú jej produkty alebo služby. Jeho

[525] RICH, B. *How AI Is Changing Contracts*. In 2018. [online]. Dostupné na: https://hbr.org/2018/02/how-ai-is-changing-contracts.

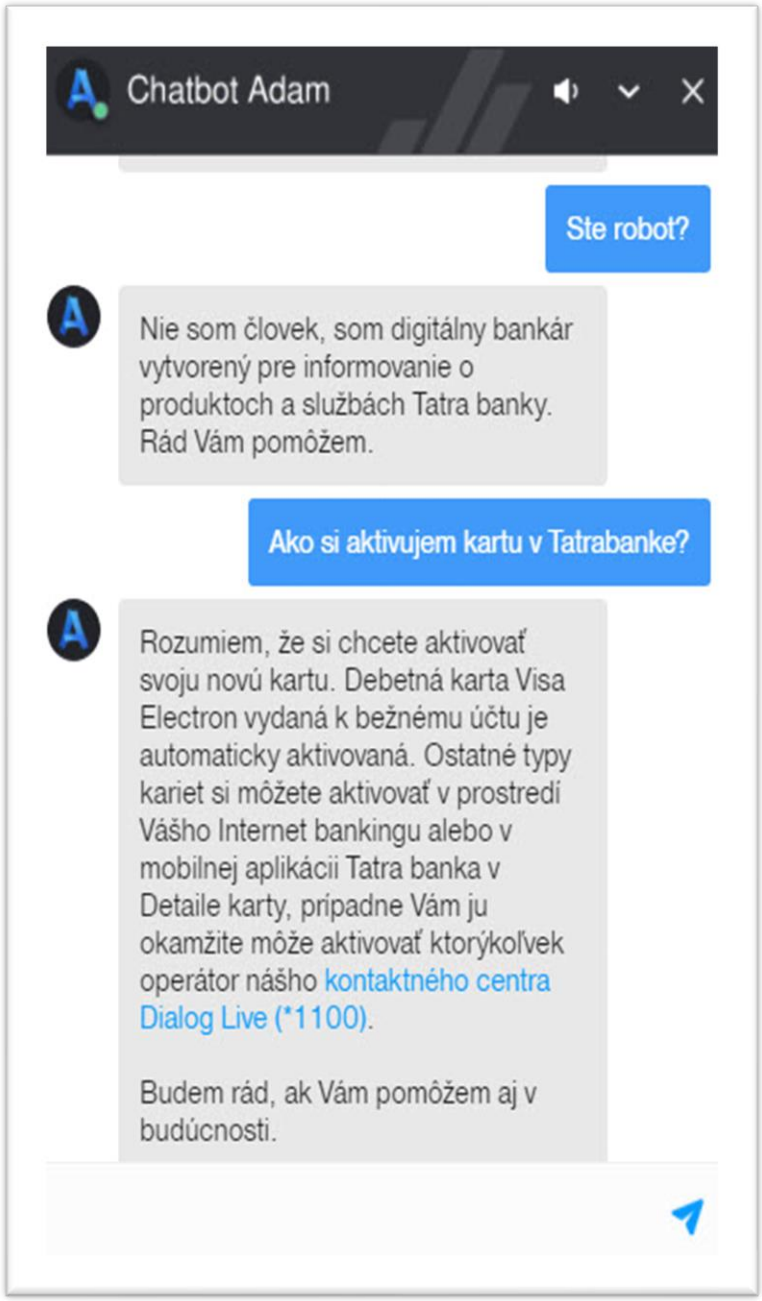
[526] HAAN, K. *How Businesses Are Using Artificial Intelligence In 2024*. [online]. Dostupné na: https://www.forbes.com/advisor/business/software/ai-in-business/

[527] Pre ďalšie údaje pozri HAAN, K. *How Businesses Are Using Artificial Intelligence In 2024*. [online]. Dostupné na: https://www.forbes.com/advisor/business/software/ai-in-business/

súčasťou je aj využitie chatbotov. Chatboty sú konverzačné nástroje, ktoré efektívne vykonávajú bežné úlohy. Chatboty sú jedným z najbežnejších prípadov, keď zákazníci priamo interagujú s AI. Z obchodného hľadiska umožňujú chatboty spoločnostiam zefektívniť procesy služieb zákazníkom a uvoľniť čas zamestnancov na problémy, ktoré si vyžadujú viac personalizovanej pozornosti. Chatboty zvyčajne používajú kombináciu spracovania prirodzeného jazyka, strojového učenia a AI na pochopenie požiadaviek zákazníkov. Chatbot imituje konverzáciu s bežným človekom s použitím základných praktík strojového učenia. Môžete sa s ním porozprávať, spýtať sa na radu, objednať si prostredníctvom neho tovar cez e-shop, či zistiť aktuálne správy zo sveta. Môže tiež pomôcť nasmerovať zákazníkov k skutočnému zástupcovi spoločnosti, ktorý je najlepšie vybavený na riešenie ich otázok.

PRÍKLAD

Príkladom je Chatbot Adam vyvinutý pre Tatru banku, a.s., ktorý je formou umelej inteligencie snažiacej sa porozumieť otázke a na základe preddefinovanej rozhodovacej logiky poskytnúť relevantnú odpoveď. Vie zodpovedať otázky o produktoch, zobraziť mapu s najbližšou či konkrétnou pobočkou alebo bankomatom/vkladomatom a tiež kurzový lístok spolu s menovou kalkulačkou. Vie pomôcť s úpravou limitov platobnej karty či nastavením trvalého príkazu alebo blokovaním karty a pod. Súčasne napĺňa podmienky dôveryhodnosti, nakoľko chatovanie s Adamom je bezpečne chránené šifrovaním, pretože citlivé údaje pri ukladaní najprv zamaskuje, aby ich chránil pred zneužitím. Chatbot tiež nesleduje aktivity mimo chatovacieho okna.[528] Vedľa uvádzam príklad komunikácie s chatbotom Adamom:



[528] Pre viac informácií pozri: https://www.tatrabanka.sk/sk/chatbot/.

9. Umelá inteligencia, profilovanie a sociálne média

9.1 Ochrana súkromia a údaje v 21. storočí

Údaje zo sociálnych sietí, mobilné údaje, obchodné údaje, údaje z internetu vecí (IoT), verejné údaje, či komerčné údaje sú analyzované za účelom pochopenia vzorcov ľudského správania a následného lepšieho rozhodovania a prispôsobenia obchodnej stratégie. Aby sme odhalili vzorce ľudského správania, je potrebné získať a analyzovať dostatočné množstvo informácií o jednotlivcoch alebo skupinách jednotlivcov. Medzi nimi sa následne hľadajú podobné vlastnosti alebo správanie, na základe čoho je týchto jednotlivcov možné kategorizovať, t. j. zaradiť do určitých skupín alebo kategórií (napr. osoby, ktorým sa páčia komédie, osoby s vyšším rizikom ochorenia na rakovinu a podobne). Údaje o ľudskom správaní a vlastnostiach je tiež možné využiť na vytváranie úsudkov (napríklad, že niekto je vegetarián na základe histórie nákupov potravín) a predpovedí o budúcom správaní (napríklad, či zamestnanec bude efektívne pracovať aj o niekoľko rokov). Kategorizáciu je tiež možné realizovať na základe úsudkov, napríklad ak by niekto o niekom na základe toho, s kým sa priatelí na Facebooku usúdil, akú ma daná osoba sexuálnu orientáciu a následne by ju zaradil do skupiny ľudí s vyššou pravdepodobnosťou nákazy pohlavnou chorobou. Uvedený proces zbierania a analýzy údajov sa označuje ako „profilovanie" (po anglicky profiling). Hildebrandt rozlišuje medzi organickým ľudským a strojovým profilovaním. Ľudia majú tendenciu generalizovať, kategorizovať a iným spôsobom vyhodnocovať to, čo sa okolo nich deje, aby pre nich bola realita zrozumiteľnejšia. To isté robia ostatné živé organizmy, aby prežili v podmienkach, v ktorých existujú. Stroje zase môžu byť naprogramované ľuďmi tak, aby spracúvali informácie a vykonávali obdobné činnosti.[529]

Príklad so sexuálnou orientáciou nám znázorňuje, že úsudky sú nielen výsledkom profilovania, ale patria medzi typy údajov, ktoré sa využívajú pri profilovaní. Vo všeobecnosti môžeme povedať, že existujú štyri typy údajov:[530]

[529] HILDEBRANDT, M.: Defining Profiling: A New Type of Knowledge? In: HILDEBRANDT, M. et GUTWIRTH, S. (eds.): Profiling the European Citizen, Springer Science + Business Media B.V., 2008, s. 26.
[530] Information Commissioner's Office: Big data, artificial intelligence, machine learning and data protection, 2017, s. 12-13. Dostupné online: https://ico.org.uk/media/for-organisatio ns/documents/2013559/big-data-ai-ml-and-data-protection.pdf (cit. dňa 15.02.2021).

1. poskytnuté údaje (provided data): vedome odovzdané údaje (napríklad pri vypĺňaní dotazníka),

2. odpozorované údaje (observed data): údaje zaznamenané automaticky, napríklad prostredníctvom online cookies,

3. odvodené údaje (derived data): sú vyprodukované na základe iných údajov, napríklad kalkuláciou profitability zákazníka na základe počtu návštev obchodu, počtu alebo druhu kúpených produktov (ide o jednoduchšie štatistické výpočty) a

4. úsudky (inferred data): sú vytvorené využitím komplexnejších analytických metód, ktorými sa hľadajú korelácie medzi dátami a dátovými súbormi. Úsudky sa tvoria na základe pravdepodobností, a preto sa považujú za menej „isté" ako odvodené údaje.

Hoci práve poskytnuté údaje sú tie, pri ktorých robíme vedomé rozhodnutie o tom, že ich poskytujeme druhej strane, v kontexte Big Data predstavujú iba malú časť spracúvaných údajov – väčšinu tvoria práve zvyšné tri kategórie údajov, t. j. odpozorované údaje, odvodené údaje a úsudky.[531] Na tomto mieste je potrebné vysvetliť súvislosť medzi profilovaním a strojovým učením. Hoci Ferraris označuje profilovanie ako prax aj techniku zároveň[532], nejde o techniku strojového učenia, akou sú napríklad neurónové siete. Machine-learning (a konkrétne aj kategorizáciu, zhlukovanie a neurónové siete) možno využiť na účely profilovania, aby sa mohli pomocou algoritmov hľadať určité vzťahy medzi jednotlivými údajmi, a to aj pri veľkých množstvách údajov. Strojové učenie teda vnímame ako nástroj profilovania.

Profilovanie sa delí na skupinové a individuálne. Pri skupinovom profilovaní sa identifikujú ľudia, ktorí zdieľajú jednu alebo viac spoločných čŕt, napríklad fanúšikovia konkrétneho športu alebo ľudia s rovnakým vzdelaním. Individuálne profilovanie vychádza zo súboru vlastností jednej osoby za účelom usudzovania alebo predpovedania určitých jej charakteristík v budúcnosti.[533] Oba typy profilovania sa môžu realizovať priamo alebo

[531] Európsky dozorný úradník pre ochranu údajov (EDPS): Meeting the challenges of big data. A call for transparency, user control, data protection by design and accountability , Opinion 7/2015, 19. november 2015, s. 10.
[532] FERRARIS, V. et al.: Defining Profiling, 2013, s. 20. Dostupné online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2366564 (cit. dňa 29.3.2021).
[533] Ibid., s. 6-7.

nepriamo. Pri priamom profilovaní sa usudzujú alebo predpovedajú atribúty na základe údajov o danej osobe alebo skupine, zatiaľ čo pri nepriamom profilovaní sa údaje o osobe alebo skupine usudzujú alebo predpovedajú na základe údajov o inej alebo väčšej časti populácie. Rozdiel je možno vysvetliť na jednoduchom príklade cielenej reklamy. Pri priamom profilovaní sa osobe ponúkne produkt, ktorý by sa jej mohol páčiť na základe jej predošlých nákupov, pri nepriamom profilovaní (napr. ak osoba ešte od daného predajcu nenakupovala), sa odporučí produkt na základe nákupov iných ľudí, ktorí s danou osobou zdieľajú podobné črty (napr. na základe pohlavia, veku, navštívených stránok a pod.). Nepriame profilovanie, logicky, vykazuje nižšiu mieru istoty ako priame profilovanie. Poo et al.[534] a Kanoje[535] et al. tiež rozlišujú profilovanie na základe toho, či ide o údaje zverejnené v určitom bode samotnými užívateľmi (napr. v dotazníku alebo pri vytváraní účtu na sociálnej sieti) – v takom prípade hovoríme o explicitnom (statickom) profilovaní alebo o údaje, ktoré sú založené na dynamických zmenách v správaní užívateľov – potom ide o implicitné (dynamické) profilovanie, kde je potrebné pozorovať zmeny v atribútoch ako pri jednotlivcovi, tak aj pri skupine.

Privacy International vo svojom článku[536] sumarizuje niekoľko zaujímavých príkladov z praxe, kedy bolo profilovaním možné získať informácie, ktoré by väčšina z nás považovala za intímne alebo veľmi osobné. Príkladom je socio-ekonomický status a osobnostné črty (otvorenosť alebo neurotizmus), ktoré možno zistiť na základe histórie hovorov, kontaktných údajov a polohy. Emočné stavy, ako napr. smútok, nervozita, či sebavedomie možno zase predpovedať pomocou spôsobu písania na počítačovej klávesnici. Z verejne dostupných údajov (napr. z tweetov) sa dá tiež usúdiť priemerná výška mzdy osoby alebo výška dlhu. Prostredníctvom profilovania možno tiež jednotlivcov porovnávať vo vzťahu k určitým preddefinovaným vzorcom správania a ustanoviť, do akej miery sa od týchto vzorcov odchyľujú. Medzi príklady z praxe, na ktoré poukazuje Privacy International, patria napríklad niektoré náboženské skupiny v USA, ktoré profilujú jednotlivcov a udeľujú im body[537] podľa toho, ako vážne berú svoju vieru alebo softvéry využívané na hodnotenie uchádzačov o

[534] POO, D. et al.: A Hybrid Approach for User Profiling, In: Proceedings of the 36th Hawaii International Conference on System Sciences, IEEE, 2003 (dostupné online).
[535] KANOJE, S. et al.: User Profiling Trends, Techniques and Applications, In: International Journal of Advance Foundation and Research in Computer, 2014, Vol. 1, No. 1. (dostupné online).
[536] Privacy International: Big Data: A tool for development or a threat to privacy? 2014. Dostupné online: https://privacyinternational.org/blog/1434/big-data-tool-development-or-threat-privacy. (cit. dňa 19.2.2020).
[537] Tamže.

zamestnanie, ktoré udeľujú body,[538] triedia životopisy a zoraďujú uchádzačov na základe relevantných kritérií. Výsledok, resp. profil poskytnutý softvérom môže slúžiť ako podklad pre rozhodnutie človeka, napríklad riaditeľa ľudských zdrojov, o prijatí alebo neprijatí určitého uchádzača, prípadne môže ísť o tzv. „automatizované rozhodovanie", kedy sa rozhodnutia uskutočňujú aspoň sčasti technologickými prostriedkami. Príkladom, kedy ide o výlučne automatizované rozhodovanie je, ak vyfiltrované životopisy už nebudú posunuté na ďalšie zváženie a vyradení uchádzači budú automaticky informovaní o neúspechu. Je potrebné podotknúť, že profilovanie a automatizované rozhodovanie nie sú navzájom nevyhnutne späté a je možné uskutočňovať jedno bez druhého.

9.2 Sociálne siete, profilovanie a personalizovaný obsah

Rada Európy vo svojom odporúčaní uvádza, že profilovanie prebieha v troch technicky odlišných etapách:[539]

1. etapa – zbieranie údajov o ľudskom správaní a vlastnostiach (vo veľkom rozsahu),

2. etapa – analýza zozbieraných údajov hľadaním vzájomných vzťahov medzi rôznymi vlastnosťami, charakteristikami a správaním a

3. etapa – vytváranie úsudkov a predpovedí na základe analýzy údajov.

Proces analýzy údajov a jej následného využitia, t. j. 2. a 3. etapu, sme popísali vyššie. Obe súvisia predovšetkým so strojovým učením, ktoré poskytuje konkrétne nástroje (napr. regresia) na hľadanie súvislostí, ktoré umožňujú usudzovať alebo predpovedať určité javy,

[538] Bodovací systém poznáme napríklad z Číny, ktorá využíva tzv. systém sociálneho kreditu, v rámci ktorého sa monitoruje a vyhodnocuje správanie jednotlivcov za účelom zvýšenia dôvery v spoločnosti. Tento systém, ktorý sa zvykne označovať ako novodobý orwellovský „Veľký brat", udeľuje občanom skóre na základe ich správania, čo má pre nich dôsledky v každodennom živote. Pozri viac: KOLANY-RAISER, B. et RADTKE, T.: Ich sammele, also bin ich (Social Credit) – Das Szenario eines allumfassenden Bonitätssystems am Beispiel Chinas, In: HOEREN, T. (Hrsg.): Phänomene des Big-Data-Zeitalters. Eine rechtliche Bewertung im wirtschaftlichen und gesellschaftspolitischen Kontext, Münster: Universitäts- und Landesbibliothek Münster, 2019, s. 121-146.

[539] Výbor Ministrov Rady Európy: Recommendation CM/Rec(2010)13 of the Committee of Ministers to member states on the protection of individuals with regard to automatic processing of personal data in the context of profiling, 23. November 2010, s. 25.

správanie alebo vlastnosti. Pokiaľ ide o tretiu etapu, je potrebné spomenúť, že profilovanie možno realizovať aj s menším súborom údajov.

Samotné profilovanie má aj svoje legálne vymedzenie. Profilovanie je legálne definované v článku 4 bode 4 GDPR[540] ako *„akákoľvek forma automatizovaného spracúvania osobných údajov, ktoré pozostáva z použitia týchto osobných údajov na vyhodnotenie určitých osobných aspektov týkajúcich sa fyzickej osoby, predovšetkým analýzy alebo predvídania aspektov dotknutej fyzickej osoby súvisiacich s výkonnosťou v práci, majetkovými pomermi, zdravím, osobnými preferenciami, záujmami, spoľahlivosťou, správaním, polohou alebo pohybom."* Profilovanie bez individuálneho rozhodovania je v podstate bežnou spracovateľskou operáciou a z hľadiska právnych účinkov sa neodlišuje od iných spracovateľských operácií ako bežná analýza alebo prenos osobných údajov.

Profilovanie v sebe subsumuje dva pojmové znaky:

1. Získanie charakteristík o jednotlivcovi alebo skupine osôb (proces kreovania profilu); a
2. Použitie týchto charakteristík na jednotlivca alebo skupine osôb (aplikovanie profilu).[541]

Tieto dva pojmové znaky sú zároveň naviazané na automatizovanú formu spracúvania, avšak nevylučujú ľudský zásah.[542] Zároveň prvok hodnotenia určitých osobných aspektov sa týka tak procesu kreovania profilu ako aj jeho aplikácie.[543] WP29 zároveň dodáva, že jednoduchá klasifikácia osôb na základe veku, pohlavia alebo iných známych znakov nemusí nevyhnutne predstavovať profilovanie v zmysle GDPR.[544] Na profilovanie z pohľadu regulácie ochrany osobných údajov sa tak budú aplikovať požiadavky na spracúvanie osobných údajov v zmysle príslušnej legislatívy.

[540] 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).
[541] BYGRAVE, L.: Profiling. In KUNER, CH. – BYGRAVE, L. – DOCKSEY, CH.: *The EU General Data Protection Regulation (GDPR). A commentary*. Oxford: Oxford University Press, 2020, s. 130.
[542] Tamže.
[543] Tamže.
[544] Article 29 Data Protection Working Party: *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*. Adopted on 3 October 2017. As last Revised and Adopted on 6 February, s. 7.

Jedným z priestorov, kde sa koncentruje najväčšie množstvo údajov o ľudskom správaní a vlastnostiach sú sociálne siete. Sociálna sieť ako služba je jednou z najväčších tovární na údaje tohto druhu. Profilovanie na úrovni takých gigantov, akými je Facebook dokonca dokáže ovplyvniť aj dianie vo svete. Notoricky známym prípadom je poradenská spoločnosť Cambridge Analytica, ktorá získala od Facebooku údaje z profilov približne 214 miliónov amerických užívateľov. Išlo o mená, dátumy narodenia, údaje o polohe, ako aj zoznam všetkých stránok na Facebooku, ktoré kedy lajkli alebo čo si kedy stiahli. Dotknutí užívatelia o tom nevedeli a neposkytli ani súhlas s poskytnutím týchto údajov predmetnej firme. Tieto údaje umožnili Cambridge Analytica vytvoriť tzv. psychometrické profily užívateľov, ktoré využila pre množstvo kampaní v USA, vrátane kampane Teda Cruza, Donalda Trumpa alebo „Leave.EU" (radikálnejšia z dvoch kampaní Brexit).[545] Sociálnym sieťam by sme z dôvodu ich vplyvu venovali pozornosť v rámci tejto kapitoly a priblížili, akým spôsobom prebieha profilovanie na týchto populárnych platformách. Sociálne siete sú fenoménom našej doby. Iba Facebook samotný využíva takmer 2,5 miliardy užívateľov, čo znamená, že ak by Facebook bol krajinou, bol by najväčšou na svete. Sociálne siete sú zároveň jedným z kľúčových médií, prostredníctvom ktorých sa šíria informácie vrátane dezinformácií. Až 22 % ľudí získava spravodajské informácie prostredníctvom webov alebo aplikácií.[546] Najpopulárnejšou sociálnou sieťou ostáva Facebook, avšak ostatné sociálne siete, predovšetkým Youtube a TikTok si zachovávajú solídne čísla užívateľov, predovšetkým medzi mladšou generáciou.[547] Pri preberaní správ, respondenti výskumu uviedli, že zvýšenú pozornosť venujú kanálom a stránkam celebrít, influencerom či iným osobnostiam.[548]

V Slovenskej republike patrí medzi najpoužívanejšie sociálne siete stále Facebook.[549]

Online platforma	Využitie za účelom získania spravodajstva	Využitie za iným účelom
Facebook	48 %	71 %
Youtube	25 %	63 %

[545] FUNTA, R. et PLAVČAN, P.: Údaje a ich úloha v politických kampaniach, In: Justičná revue, 2019, Vol. 71, No. 10, s. 934-957 (dostupné online).
[546] REUTERS. Reuters Institute Digital News Report 2023. 2023. Dostupné na: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf.
[547] Tamže.
[548] Tamže.
[549] Tamže, s. 97.

Facebook messenger	15 %	49 %
Instagram	11 %	32 %
WhatsApp	9 %	33 %
Telegram	5 %	11 %

Na sociálnych sieťach zverejňujú ľudia veľké množstvo obsahu, vyjadrujú svoje názory a myšlienky, zdieľajú fotografie a videá. Vďaka nim dokážu veľmi efektívne komunikovať s priateľmi na opačnom konci sveta, či osloviť široké publikum. Pokiaľ ide o množstvo času venovaného sociálnym sieťam, tak v roku 2018 bol priemerný čas človeka strávený na týchto platformách takmer dve a pol hodiny denne, čo predstavuje 62,5 percentný nárast oproti roku 2012. Odhaduje sa, že v priemere človek strávi zo svojho života takmer 7 rokov na sociálnych sieťach.[550] Čas strávený na sociálnych sieťach je čas, ktorý strávime virtuálnou interakciou s inými ľuďmi, čítaním článkov, prezeraním fotografií, získavaním najnovších informácií a aktualít a podobne. Možností využitia je naozaj veľa. Čo si však bežne neuvedomujeme je, že čas strávený na sociálnych sieťach je zároveň časom stráveným odovzdávaním svojich údajov. Nejde iba o údaje, ktoré vedome odovzdávame (napríklad meno, fotografie, status), ale aj údaje, ktoré zanechávame nevedome (ako dlho čítame článok, ktoré stránky otvoríme a podobne). Ak si otvoríme aplikáciu sociálnej siete, je to ako keď vojdeme do miestnosti a zanecháme po sebe všade stopy, na zemi, kam stúpime, odtlačky prstov na všetkom, čoho sa chytíme. Už pri zaregistrovaní sa zverejňujeme osobné údaje ako meno a priezvisko, vek, pohlavie, dátum narodenia, či sme vo vzťahu alebo nie, údaje o vzdelaní a zamestnaní, fotografie a ďalšie údaje. Ak dáme na sociálnej siete na nejaký príspevok lajk, zverejníme fotku, okomentujeme nejaký článok, s niekým sa spriatelíme a podobne, ide o údaje, ktoré pri správnom využití odhalia o nás veľa vecí, pričom mnohé z nich by sme považovali za súkromné. Informácie o nás môžu byť dokonca získané aktivitou iných užívateľov, prípadne ľudí, ktorí ani užívateľmi sociálnych sietí nie sú. Údaje sú tiež získavané aj z aktivít, o ktorých by nám možno ani nenapadlo, že ich niekto zaznamenáva. Príkladom je „scrollovanie" obrazovky, kedy sa meria, ako dlho sa zastavíme na nejakom príspevku alebo ako dlho čítame určitý článok. Ide o údaje, ktoré o nás majú silnú výpovednú hodnotu.

[550] Pre zaujímavosť, skutočným socializovaním sa mimo virtuálneho sveta, strávi človek v priemere iba necelé 2 roky. Pozri BroadbandSearch: Average Time Spent Daily on Social Media (Latest 2020): https://www.broadbandsearch.net/blog/average-daily-time-on-social-media#post-navigation-0 (cit. dňa 4.2.2020).

Na demonštráciu toho, ako o ľuďoch (aj o tých, ktorí nie sú zaregistrovaní a nemajú svoj účet) získava Facebook údaje, vysvetlíme, ako funguje spomínané, na prvý pohľad nenápadné, tlačidlo označené „like" alebo „páči sa mi". Facebookové lajky sú známe asi každému. Prostredníctvom tohto tlačidla môžu užívatelia vyjadriť svoj súhlas alebo iný pozitívny vzťah so zverejneným obsahom. Tento nástroj bol vyvinutý tak, aby mohol byť, okrem Facebooku, zobrazený aj na iných webových stránkach spolu s uvedením počtu užívateľov, ktorí naň klikli. Z technického hľadiska ide o HTML kód vložený do webovej stránky. Tento kód umožňuje vykonanie určitých príkazov zakaždým, keď je webová stránka načítaná alebo obnovená. Príkladom takého príkazu je zaslanie cookies tretích strán do vyhľadávača užívateľov alebo použitie už predtým nainštalovaných cookies, ktoré pre Facebook získajú o užívateľoch informácie. Facebook tak vie využiť lajk tlačidlo na sledovanie a profilovanie užívateľov navštevujúcich rôzne webové stránky bez ohľadu na to, či na tlačidlo naozaj klikli, či sa z Facebooku odhlásili alebo vôbec majú na Facebooku účet. Čiže aj keď osoba nemá účet na Facebooku, ten dokáže prostredníctvom svojho lajk tlačidla zbierať údaje o jej histórii navštívených stránok, napr. URL stránky, dátum a čas návštevy, IP adresu zariadenia a podobne. Ak si táto osoba účet neskôr vytvorí, Facebook bude schopný vytvoriť korelácie medzi históriou prehliadania a novovytvoreným profilom.[551]

Ako možno vidieť, na sociálnych sieťach ide o spleť rôznych údajov, ktoré sú navzájom poprepájané a navzájom spolu súvisia. Schneier vytvoril taxonómiu údajov, ktorá zatrieďuje a sprehľadňuje typy údajov na sociálnych sieťach:[552]

a. Servisné údaje (Service data): údaje, ktoré dávame sociálnej sieti, aby ich využívala, napr. meno, vek, číslo kreditnej karty.

b. Zverejnené údaje (Disclosed data): informácie, ktoré zverejňujeme na našich profiloch alebo webových stránkach, napr. blog, fotografie, správy, komentáre, statusy.

[551] KELBERT, F. et al.: State of Online Privacy: A Technical Perspective, In: BUCHMANN, J. (ed.): Internet Privacy, 2012, s. s. 207-208.
[552] SCHNEIER, B.: A Taxonomy of Social Networking Data, 2010. Dostupné online: https://www.schneier.com/essays/archives/2010/07/a_taxonomy_of_social.html (cit. dňa 4.2.2020).

c. Zverené údaje (Entrusted data): informácie, ktoré zverejňujeme na profiloch a stránkach iných ľudí/tretích strán. Rozdiel oproti predošlému prípadu je v podstate iba ten, že po zverejnení strácame kontrolu nad týmito údajmi.

d. Incidental data: informácie, ktoré o nás zverejní niekto iný, napríklad fotografia, na ktorej sa nachádzame, ale nezverejnili sme ju my. Opäť ide o rovnaké údaje ako v prípade zverejnených údajov, rozdiel spočíva v tom, kto má nad nimi kontrolu a kto ich vytvoril.

e. Údaje o správaní (Behavioural data): údaje, ktoré o nás poskytovatelia internetových služieb zbierajú a ktoré sa týkajú našich zvykov a záujmov na základe sledovania toho, čo robíme, s kým to robíme, kedy to robíme. Môže ísť o hry, ktoré hráme, témy, o ktorých píšeme, novinové články, ktoré čítame a podobne.

f. Odvodené údaje (Derived data): údaje, ktoré sú o nás a ktoré sú odvodené zo všetkých ostatných údajov, ako napríklad naša sexualita, ktorá môže byť odvodená od sexuality prevažnej väčšiny našich priateľov.

Ide o obdobu delenia údajov na poskytnuté, odpozorované, odvodené a úsudky, ktoré sme uviedli vyššie, hoci Schneier vo svojom delení nereferuje na úsudky tvorené o nás na základe údajov o iných ľuďoch, pričom takéto úsudky sú základným pilierom profilovania. V rámci investigatívnej práce žurnalistov bolo zistené, že pomocou profilov na sociálnych sieťach možno usúdiť alebo predpovedať rôzne stavy, črty, či názory, ako napríklad sexuálnu orientáciu a zdravotný stav,[553] rasu,[554] bezprostredne hroziace pokusy o samovraždu,[555] tendencie k depresii,[556] predpoklady na udelenie úveru,[557] politické názory[558] a podobne.

[553] CUEVAS, Á. et al.: Does Facebook Use Sensitive Data for Advertising Purposes? Dostupné online: https://arxiv.org/pdf/1907.10672.pdf. (cit. dňa 17.2.2020). V tlači.
[554] TOBIN, A.: Facebook (Still) Letting Housing Advertisers Exclude Users by Race, 2017. Dostupné na: https://digg.com/2017/facebook-housing-discrimination (cit. dňa 17.3.2021).
[555] CONSTINE, J.: Facebook rolls out AI to detect suicidal posts before they´re reported, 2017. Dostupné online: https://techcrunch.com/2017/11/27/facebook-ai-suicide-prevention/?guccounter=2 (cit. dňa 17.3.2021).
[556] MORENO, M. A. et al.: Feeling Bad on Facebook: Depression disclosures by college students on a Social Networking Site, In: Depression & Axiety, 2011, Vol. 28, Issue 6 (dostupné online).
[557] TAYLOR, A. et SADOWSKI, J.: How Companies Turn Your Facebook Activity Into a Credit Score, 2015. Dostupné na: https://www.thenation.com/article/archive/how-companies-turn-your-facebook-activity-credit- score/ (cit. dňa 17.3.2021).
[558] BOWCOTT, O. et HERN, A.: Facebook and Cambridge Analytica face class action lawsuit, 2018, The Guardian. Dostupné online: https://www.theguardian.com/news/2018/apr/10/cambridge-analytica-and- facebook-face-class-action-lawsuit (cit. dňa 17.2.2020).

Pokrok v technológiách umožňuje vyhodnocovať enormné množstvá údajov a na základe nich predpovedať „vzorce" správania ľudí. Prostredníctvom komplexných algoritmov dokážu sociálne siete (alebo iné subjekty disponujúce údajmi zo sociálnych sietí, viď príklad s Cambridge Analytica) vytvoriť profily o našich presvedčeniach, názoroch a správaní a v prepojení s názormi tisícok podobných ľudí je možné docieliť pochopenie individuálnych súborov osobností,[559] výsledkom čoho je možnosť zobrazovať užívateľom personalizovaný obsah. Pre personalizáciu opäť neexistuje jednotná definícia, v literatúre sa zjednodušene uvádza, že ide o „poskytovanie správneho obsahu správnej osobe v správnej forme a správnom čase". Uskutočňuje sa to na základe histórie správania tej konkrétnej osoby alebo iných osôb (viď typy profilovania), napríklad zobrazenie reklamy na určitý produkt na základe histórie nákupov alebo histórie vyhľadávania, čím sa dosiahne lepšie zacielenie na potenciálnych zákazníkov. Sociálne siete využívajú personalizáciu na to, aby nám ukazovali také články a príspevky, ktoré zodpovedajú našim záujmom. Na tom je postavený celý biznis týchto platforiem. Ak užívateľ vidí obsah, ktorý sa mu páči, cíti sa príjemne a obohatene. Tieto pocity spôsobujú to, že sociálnu sieť pravidelne využíva a zanecháva ďalšie údaje, na základe ktorých je možné poskytnúť ďalšiu cielenú reklamu a tento cyklus sa opakuje nonstop. Práve preto je hlavnou prioritou Facebooku tzv. „angažovanosť"[560] – neustála motivácia užívateľov znovu sa vrátiť. Algoritmy, ktoré tento proces umožňujú, možno označiť ako odporúčacie systémy. Odporúčacie systémy používajú sofistikované, distribuované modely strojového učenia, ako napríklad hlboké neurónové siete, na identifikáciu, usporiadanie a prezentovanie menšieho výberu zo všetkých dostupných informácií. Tento výber je navrhnutý tak, aby bol relevantný pre každého používateľa na základe pravdepodobnosti, že používateľ naňho zareaguje, či už zobrazením, kliknutím, lajkom, zdieľaním alebo inými spôsobmi interakcie.

Aby boli tieto predpovede zapojenia presné a prispôsobené konkrétnemu používateľovi v danom okamihu, tieto algoritmické odporúčacie systémy sú trénované na základe obrovského množstva údajov získaných z predchádzajúcich aktivít používateľa a záujmov vrátane záujmov iných podobných používateľov. Tieto algoritmy sa taktiež prispôsobujú špecifickému kontextu, ako je čas dňa, alebo používané zariadenie. Odporúčacie systémy sú kľúčovým nástrojom na podporu a udržiavanie zapojenia používateľov. Spoločnosti

[559] FUNTA, R. et PLAVČAN, P.: Údaje a ich úloha v politických kampaniach, In: Justičná revue, 2019, Vol. 71, No. 10, s. 934-957 (dostupné online).
[560] Tamže.

zdôrazňujú, že ich odporúčacie systémy majú za cieľ spojiť používateľov s obsahom a ľuďmi, ktorí sú pre nich relevantní. Avšak, kritici tvrdia, že obchodný model týchto systémov často uprednostňuje obsah, ktorý priláka používateľov na stránky a udržiava ich tam čo najdlhšie, bez ohľadu na to, či je tento obsah kontroverzný, škodlivý alebo nízkej kvality.[561]

Samotný proces odporúčania má dve hlavné fázy:

1. Generovanie kandidátov (*candidate generation*), pri ktorom sa analyzujú milióny informácií a následne sa automatizovane vyberie užšia množina relevantná pre konkrétneho užívateľa, tzv. kandidátske príspevky
2. Poradie (*ranking*) kde sa ďalej analyzujú vlastnosti kandidátskych príspevkov a história aktivít používateľov s cieľom vyhodnotiť príspevky na základe predpokladaného zapojenia používateľov a nakoniec vybrať niekoľko desiatok príspevkov, ktoré sa nakoniec zobrazia používateľovi pri otvorení alebo obnovení aplikácie.[562]

Jedným z najväčších nedostatkov využívania odporúčacích systémov je ich nedostatočná vysvetliteľnosť a transparentnosť, keď ani vývojári častokrát nevedia fungovanie systému vysvetliť v plnej miere.[563] Niektorí akademici odporúčajú väčšiu kontrolu odporúčacích systémov zo strany užívateľov.[564]

Odporúčanie konkrétneho obsahu výrazne ovplyvňuje, akému typu informácií jednotlivci venujú pozornosť, čo vedie k uzavretej slučke v ktorej algoritmus zobrazuje jednotlivcom obsah, ktorý upúta ich pozornosť, a jednotlivci venujú väčšiu pozornosť tomu, čo im algoritmus ukazuje.[565] Posilňovanie zobrazovanie podobného obsahu prostredníctvom odporúčacích systémov zároveň spôsobuje, že niekoľko kliknutí na kontroverzný obsah môže

[561] M. Vidal Bustamante, Constanza , Joaquin Quiñonero Candela, Lucas Wright, Leisel Bogan and Marc Faddoul. "Technology Primer: Social Media Recommendation Algorithms." Edited by Ariel Higuchi. Belfer Center for Science and International Affairs, Harvard Kennedy School, August 25, 2022.
[562] Tamže, s. 6.
[563] Paul Voosen, "How AI Detectives Are Cracking Open the Black Box of Deep Learning," Science, July 06, 2017, https://www.science.org/content/article/how-ai-detectives-arecracking-open-black-box-deep-learning.
[564] Stephen Wolfram, "Testifying at the Senate about AI-Selected Content on the Internet," Stephen Wolfram Writings blog, June 25, 2019,
https://writings.stephenwolfram.com/2019/06/testifying-at-the-senate-about-a-i-selected-content-on-the-internet/.
[565] Dimitris Kalimeris, Smriti Bhagat, Shankar Kalyanaraman, and Udi Weinsberg, "Preference Amplification in Recommender Systems," in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (New York: Association for Computing Machinery, 2021):
805–15, https://doi.org/10.1145/3447548.3467298.

mať za následok ďalšie odporúčania kontroverzného obsahu, ktoré je ťažké prestať prijímať a ktoré môžu ovplyvniť jednotlivca v jeho nazeraní na svet a prijímaní informácií.[566] Ako príklad možno uviesť vyšetrovanie denníka *Wall Street Journal* týkajúce sa sociálnej siete TikTok, na ktorej nasadili umelých používateľov (botov), ktorý simulovali prezeranie odporúčaných videí so smutným kontextom dlhšie ako iné videá. Následne po tridsiatich minútach odporúčací systém užívateľovi zobrazoval videá, z ktorých 93 % bolo depresívnej povahy.[567]

PRÍKLAD

Cielená reklama je najväčším zdrojom príjmov sociálnych sietí, a preto nie je prekvapujúce, že sú prirodzene motivované neustále zlepšovať túto službu pomocou údajov o svojich užívateľoch.[568] Vzhľadom na to, že údaje od užívateľov predstavujú pre sociálne siete zisk, využívanie sociálnych sietí nie je „zadarmo", ale mena, ktorou sa platí sú údaje. Dr. Klepmann[569] však poukazuje na to, že vzťah užívateľov a sociálnych sietí nie je vzťahom so skutočnou reciprocitou a výmenou za reálnu hodnotu. Neexistuje medzi nimi dialóg a možnosť vyjednávať, koľko dát poskytnú a akú službu za to dostanú. Ide o veľmi asymetrický a jednostranný vzťah, kde podmienky určuje poskytovateľ služby. Pre používateľa, ktorý nesúhlasí s monitorovaním svojho správania je jedinou možnosťou nevyužívať službu. Ani tu však nejde v každom prípade o slobodnú možnosť voľby – niektoré sociálne siete ako Facebook sú tak populárne, že väčšina ľudí považuje za nevyhnutné ju využívať kvôli interakcii s ostatnými ľuďmi[570] a tiež ako zdroj informácií (napríklad často sa komunikujú študijné či pracovné záležitosti v skupinách a skupinových četoch na Facebooku, Instagrame či WhatsApp, ktoré patria Facebooku).

[566] Algorithms and Amplification: How Social Media Platforms' Design Choices Shape Our Discourse and Our Minds, Subcommittee on Privacy, Technology, and the Law (April 27, 2021) (statement by Joan Donovan, Research Director, Shorenstein Center on Media, Politics and Public Policy, Harvard Kennedy School), https://www.judiciary.senate.gov/imo/media/doc/Donovan%20Testimony%20(updated).pdf.

[567] WSJ Staff, "Inside TikTok's Algorithm: A WSJ Video Investigation," Wall Street Journal, July 21, 2021, https://www.wsj.com/articles/tiktok-algorithm-video-investigation-11626877477.

[568] MENDEL, T. et al.: Global Survey on Internet Privacy and Freedom of Expression, Paríž: UNESCO, 2012, s. 33.

[569] KLEPMANN, M.: Designing Data-Intensive Applications. Sebastopol: O'Reilly Media, Inc., 2017, s. 538- 539.

[570] Pracovná skupina zriadená podľa článku 29 (dnes Európsky výbor pre ochranu údajov) uvádza vo svojom stanovisku ku definícii súhlasu: „Vzhľadom na dôležitosť, ktorú nadobudli niektoré sociálne siete, niektoré kategórie používateľov (napríklad tínedžeri) budú akceptovať prijímanie behaviorálnej reklamy, aby sa zabránilo riziku byť čiastočne vylúčený zo sociálnych interakcií. Užívateľ by mal mať možnosť poskytnúť slobodný a konkrétny súhlas s prijímaním reklamy na základe správania, nezávisle od jeho prístupu k službe sociálnej siete." (pozri Article 29 Data Protection Working Party: Opinion 15/2011 on the definition of consent, 01197/11/EN WP187, 13. júl 2011, s. 18).

9.3 Reakcia práva

Regulácia na problematiku sociálnych sietí reagovala nezvyčajne pomaly. Na sklonku roka 2020 predstavila Európska komisia návrh nariadenia o digitálnych službách známeho pod skratkou DSA („*Digital Services Act*"). DSA bol v októbri 2022 finálne prijatý. Prirodzene, nie všetky navrhované zmeny zo strany inštitúcií EÚ sa premietli aj do finálneho znenia právneho predpisu. DSA má rozdielnu účinnosť diferencovanú podľa typu subjektu, na ktorý sa vzťahuje. Na sprostredkovateľské služby, hostingové služby a online platformy (pre vysvetlenie týchto pojmov pozri nižšie) bude DSA účinný od 17. februára 2024.[571] Pre veľmi veľké online platformy, na ktorých sa dezinformácie šíria najviac, je DSA účinný od štyroch mesiacov po oznámení dotknutému poskytovateľovi Európskou komisiou, že predstavuje veľmi veľkú online platformu. K oznámeniu došlo 23. apríla 2023.[572]

DSA upravuje dve oblasti relevantné z hľadiska hmotnoprávnych požiadaviek na sprostredkovateľské služby. Prvou oblasťou je právny rámec pre podmienené výnimky zo zodpovednosti poskytovateľov sprostredkovateľských služieb. Inými slovami ide o režim tzv. bezpečných prístavov pri zodpovednosti sprostredkovateľských služieb za obsah pridaný tretími stranami. Predmetná časť DSA upravuje podmienky, za akých sprostredkovateľské služby sú a nie sú zodpovedné za obsah pridaný tretími stranami. Druhou veľkou oblasťou sú požiadavky tzv. náležitej starostlivosti (požiadavky *due dilligence)* pre sprostredkovateľské služby. Práve tieto predstavujú nóvum v legislatíve EÚ a viaceré inštitúty považujeme za vysoko relevantné pre šírenie dezinformácií online. Z hľadiska vecnej pôsobnosti je potrebné upozorniť na jeden zásadný limit. DSA sa vzťahuje v drvivej miere na nezákonný obsah. Za nezákonný obsah sa považuje *„akákoľvek informácia, ktorá sama o sebe alebo tým, že odkazuje na nejakú činnosť vrátane predaja výrobkov alebo poskytovania služieb, nie je v súlade s právnymi predpismi Únie alebo niektorého členského štátu, a to bez ohľadu na presný predmet alebo povahu týchto právnych predpisov."*[573] DSA v recitálovej časti vyžaduje široký význam interpretácie daného pojmu.[574] Zároveň je potrebné zvýrazniť, že nezákonný obsah si

[571] DSA, článok 93 ods. 2.
[572] EURÓPSKA KOMISIA: Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines. Dostupné na: https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413.
[573] DSA, článok 2 písm. h).
[574] DSA, recitál 12.

definujú samotné členské štáty vo svojich právnych poriadkoch resp. ho definuje legislatíva EÚ. Špecifická pozornosť a právne požiadavky v DSA sú zamerané aj na najzaujímavejšiu množinu aktérov - veľmi veľké online platformy (VVOP). Tieto priamo definované v DSA nie sú, avšak nariadenie obsahuje návod na ich klasifikáciu. V zmysle článku 33 ods. 1 DSA veľkými online platformami možno rozumieť tých sprostredkovateľov, *„ktoré majú priemerný mesačný počet aktívnych príjemcov služby v Únii, ktorý sa rovná 45 miliónom alebo je vyšší."* Práve do tejto kategórie budú spadať sociálne siete ako Facebook, Instagram alebo TikTok. Zároveň je potrebné dodať, že veľmi veľkú platformu musí dezignovať Európska komisia.[575] Veľmi veľké online platformy predstavujú špecifikum aj v prípade dohľadu. Exkluzívnu právomoc nad týmito subjektami má Európska komisia.

Ako sme uviedli vyššie, jedným z hlavných nástrojov, cez ktoré sa šíri obsah na sociálnych médiách sú personalizované algoritmy – odporúčacie systémy. Základným účelom odporúčacích systémov je zobrazenie konkrétneho obsahu pre konkrétneho užívateľa, ktorý je šitý na mieru záujmom alebo potenciálnym záujmom užívateľa. Práve tieto nástroje zvýrazňujú a uľahčujú šírenie akéhokoľvek, vrátane škodlivého alebo nezákonného obsahu širokému publiku. Z tohto dôvodu DSA neopomína reguláciu odporúčacích systémov pre veľmi veľké online platformy.

DSA obsahuje legálnu definíciu odporúčacích systémov ako *„úplne alebo čiastočne automatizovaný systém, ktorý online platforma používa na odporúčanie konkrétnych informácií príjemcom služby vo svojom online rozhraní alebo na uprednostňovanie uvedených informácií, a to aj v dôsledku vyhľadávania iniciovaného príjemcom služby alebo iného určenia relatívneho poradia alebo významnosti zobrazovaných informácií."*[576] Samotné odporúčacie systémy výrazne závisia od voľby ich architektúry a preferencie konkrétnych parametrov ako sú obmedzenia zdieľaného obsahu, užívateľské správanie a interakcia medzi užívateľom a obsahom.[577]

[575] DSA, článok 33 ods. 4.
[576] DSA, článok 3 písm. s).
[577] Sandra K. Evans et al., 'Explicating Affordances: A Conceptual Framework for Understanding Affordances in Communication Research', Journal of Computer-Mediated Communication 22, no. 1 (2017): 35–52, https://doi.org/10.1111/jcc4.12180.

Požiadavky kladené na odporúčacie systémy sú predmetom úpravy pre VVOP a online platformy. Pre VVOP platí, že odporúčací systém musí pre užívateľov obsahovať možnosť personalizáciu obsahu vypnúť.[578]

Okrem toho, online platformy (vrátane VVOP) sú povinné zabezpečiť transparentnosť odporúčacích systémov. Transparentnosť v tomto ohľade spočíva v komunikovaní hlavných parametrov, ktoré odporúčacie systémy využívajú a poskytnutie možnosti na užívateľskú úpravu predmetných systémov.[579] Hlavné parametre musia minimálne obsahovať informácie o (i) kritériách, ktoré sú najvýznamnejšie pri určovaní informácií navrhovaných príjemcovi služby a (ii) dôvodoch relatívneho významu týchto parametrov.[580] Poradie zobrazenia informácií si budú môcť užívatelia zmeniť a upraviť fungovanie odporúčacieho systému pre seba.[581]

[578] DSA, článok 38. *„Okrem požiadaviek stanovených v článku 27 poskytovatelia veľmi veľkých online platforiem a veľmi veľkých internetových vyhľadávačov, ktorí používajú odporúčacie systémy, pre každý zo svojich odporúčacích systémov poskytnú aspoň jednu možnosť, **ktorá nie je založená na profilovaní** v zmysle vymedzenia článku 4 bodu 4 nariadenia (EÚ) 2016/679.*"
[579] DSA, článok 27 ods. 1.
[580] DSA, článok 27 ods. 2.
[581] DSA, článok 27 ods. 3.

10. Etika umelej inteligencie

10.1. Úvodné poznámky

Nástup automatizovaných systémov poskytol v posledných rokoch úrodnú pôdu pre diskusie nie len o právnych, ekonomických a filozofických otázkach, ale aj o etike umelej inteligencie.[582]

Treba podotknúť, že samozrejme nie všetky technológie, ktoré sú automatizované, spôsobia obdobné problémy. Elektrický otvárač na konzervy môže ušetriť čas v porovnaní so svojím manuálnym bratrancom, ale nevyžaduje žiadnu spoločenskú zmenu v tom zmysle, že právny rámec už nie je dostatočný alebo vhodný len preto, že existuje táto nová technológia.[583] Dôraz sa bude klásť na technológie, ktoré posúvajú nás k zmenám už svojou existenciou. Na technologické pokroky, ktoré zmenia, to, čo je skutočne možné a nie iba bežné zmeny. Dôležité je si tiež uvedomiť, že to samozrejme vylučuje zmeny v spoločenských normách a zvykoch, ktoré menia to, čo je v spoločnosti akceptované.[584] Napriek enormným výhodám, ktoré umelo inteligentne systémy so sebou prinášajú sa neočakáva, že implementácia automatizovaných vozidiel bude jednoduchá. Spotrebitelia očakávajú, že automatizovaná technológia bude dokonale bezpečná a sú extrémne nemilosrdní voči akýmkoľvek chybám alebo nedostatkom, ktoré takéto systémy môžu mať. Ukazuje sa rastúci počet výskumov o averzii k algoritmom.[585]

---

[582] ETZIONI, A., ETZIONI, O. Incorporating ethics into artificial intelligence. The Journal of Ethics, 21(4), s403–418. 2017. Dostupné na: <https://philpapers.org/archive/ETZIEI.pdf>. cit. 2023-04-25. alebo Hengstler, M., Enkel, E., Duelli, S. Applied artificial intelligence and trust – The case of autonomous vehicles and medical assistance devices. Technological Forecasting and Social Change, s105–120. 2016. Dostupné na:<https://www.sciencedirect.com/science/article/abs/pii/S0040162515004187>. cit. 2023-04-25. alebo LIN, P., ABNEY, K., JENKINS, R. Robot ethics 2.0: From autonomous cars to artificial intelligence. Oxford: Oxford University Press. 2017. Dostupné na: <https://books.google.sk/books?hl=en&lr=&id=y2EwDwAAQBAJ&oi=fnd&pg=PP1&ots=8aH9pQcVfK&sig=6Pr1F3ULv5Tbdo5ocikTspL_Aqg&redir_esc=y#v=onepage&q&f=false>. cit. 2023-04-25.

[583] MOSES, L: Recurring Dilemmas: The Law's Race to Keep Up with Technological Change.. University of Illinois Journal of Law, Technology & Policy, 2007 Dostupné na: <https://www.researchgate.net/publication/228183058_Recurring_Dilemmas_The_Law%27s_Race_to_Keep_Up_With_Technological_Change>. cit. 2023-04-25.

[584] Porovnaj GYURÁSZ, Z.: Ethics in the age of AI, Bratislava legal forum 2020 [elektronický dokument] : disruptive technologies: regulatory and ethical challenges. - : 1. vyd. ISBN 978-80-7160-567-6. - Bratislava : Právnická fakulta UK, 2020. - S. 63-68

[585] DIETVORST, B. J., SIMMONS, J. P., & MASSEY, C. Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General,144(1), s114–126. 2015. Dostupné na:<https://repository.upenn.edu/cgi/viewcontent.cgi?article=1392&context=fnce_papers>. cit. 2023-04-25. alebo BURTON, J. W., STEIN, M.-K., & JENSEN, T. B. A systematic review of algorithm aversion in augmented decision

10. 2 Etika umelej inteligencie

Ako sme už aj vyššie spomenuli, nové technológie so sebou nesú riziká, ktoré môžu byť potenciálne v rozpore s existujúcimi sociálnymi, kultúrnymi či etickými hodnotami.[586] Skorým príkladom technológie, pri ktorej boli tieto problémy viac ako zreteľné, boli železnice. Dlho sa diskutovalo o vlastníckych právach ako aj o zodpovednosti za škody na zamestnancoch, bezpečnosti cestujúcich a vlastníctve pozemkov.[587] Následne s rozkvetom počítačov sa právnici začali pýtať aké následky by mohlo mať udelenie právnej subjektivity počítačom,[588] či údaje uložené v počítači napĺňajú požiadavku zákona na písomnú formu či by mali byť počítačové výtlačky prípustné ako dôkaz pred súdom[589] ako aj to, či počítačový softvér predstavoval majetok podliehajúci dani.[590]

Technologické zmeny majú výrazný potenciál nepriaznivo vplývať na naše tradičné hodnoty,[591] na našu kultúru a etiku.[592] Odstupom času sa ukázalo, že existuje niekoľko spôsobov ako sa naša spoločnosť vysporiadáva s vzniknutým konfliktom medzi novými technológiami a existujúcimi hodnotami.

a) *Po prvé*, naše hodnotové procesy sa môžu prispôsobiť tak, aby sa technológia integrovala do sociálneho sveta.[593]
b) *Po druhé*, v horších prípadoch, záujmy verejnosti môžu byť ignorované alebo odmietané.[594]

---

making. Journal of Behavioral Decision Making, 33(2), s220–239. 2020. Dostupné na: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bdm.2155>. cit. 2023-04-25.

[586] Mesthene opísal tieto vplyvy ako negatívne externality, ktoré sú výsledkom „nespočetných individuálnych rozhodnutí o vývoji individuálnych technológií pre individuálne účely bez výslovnej dbania o to, čo všetko tieto rozhodnutia prinášajú pre spoločnosť ako celok a pre ľudí" - MESTHENE, G. E.: The Role of Technology in Society, in Technology and the future, 1997. cit. 2023-04-25.

[587] Porovnaj, ISAAC, R.:A practical treatise upon the law of railways. Boston, Little, Brown and company 1858; 2d ed. edition, 1858.

[588] Napríklad, BANZHAF, J.: When a Computer Needs a Lawyer, Communications of the ACM, 1966. Dostupné na: <https://dl.acm.org/doi/10.1145/363567.363573,>. cit. 2023-04-25. Porovnaj taktiež, Gyurász, Z.: Duševné vlastníctvo vo svetle umelej inteligencie In COMENIUS časopis, 2/2020, Právnická fakulta, Univerzita Komenského v Bratislave, s. 6-14.

[589] BROWN, J.: Electronic Brains and the Legal Mind: Computing the Data Computer's Collision with Law, yale, L.J. 1961

[590] LOWRY, P. H.: Does Computer Stored Data Constitute a Writing for the Purposes of the Statute of Frauds and the Statute of Wills?, RUTGERS COMP. & TECH. L.J. 1982, Dostupné na: <http://houstonputnamlowry.com/Articles/Computer_Stored_Data_Article.pdf> cit. 2023-04-25.

[591] Fyzické a duševné zdravie ľudí, životné prostredie, sociálne otázky, atď.

[592] LUDOW, K., BOWMAN, M.D., GAFOT, J., BENNETT, G. M.: Regulating Emerging and Future Technologies in the Present, Springer, 2015.

[593] Pre tento argument pozri bližšie, BERNSTEIN, G.: The Socio-Legal Acceptance of New Technologies: A Close Look at Artificial Insemination, 2002, Dostupné na: <https://pubmed.ncbi.nlm.nih.gov/15212080/>. cit. 2023-04-25.

[594] Pre tento argument pozri bližšie, Lassen, J., Jamison, A.: Genetic Technologies Meet the Public: The Discourses of Concern, SCI., TECH. & HUMAN VALUES, 2006 Dostupné na:

c) *Po tretie*, ak neexistuje žiadny prienik, nová technológia sa môže zakázať alebo obmedziť. Používanie novej technológie sa jednoducho zakáže alebo sa obmedzí s cieľom znížiť jej „zlý" vplyv, aby sme sa tak dokázali uchovať naše tradičné hodnoty.[595]

d) *Po štvrté*, samozrejme môže nastať situácia, že použitie novej technológie sa jednoducho povolí a akceptujú sa následky takéhoto rozhodnutia.[596]

Tento konflikt hodnôt však neodvrátiteľne bude spôsobovať problém v kontexte právnej istoty.[597] Otázky spravodlivosti, predvídateľnosti práva a právnej istoty dostávajú „novú príchuť" vo svetle nových technológií.[598] Za dôležité považujeme spomenúť, že pre rozhodovaciu prax nebude postačovať ak sa jednoducho poukáže na fakt, že ide o novú technológiu o ktorej nemusia existovať žiadne výslovne pravidlá.[599] Veríme totiž, že jednou z najdôležitejších funkcií práva, ako aj etiky nie je nevyhnutne len regulovať a riadiť spoločnosti, ale aj poskytnúť stabilné a predvídateľné prostredie, v ktorom si ľudia môžu uskutočňovať svoje vlastné túžby a sny.[600] Stabilita a predvídateľnosť nie sú preto iba žiaducim prvkom, sú

<https://www.researchgate.net/publication/6952632_Genetic_Technologies_Meet_the_Public_The_Discourses_of_Concern>. cit. 2023-04-25.

[595] Pre tento argument pozri bližšie, Mandel, G.: Technology Wars: The Failure of Democratic Discourse, Michigan Telecommunications and Technology Law Review, 2005. Dostupné na: <https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1115&context=mttlr>. cit. 2023-04-25.

[596] Treba však mať na pamäti, že neprijatie žiadnych opatrení by mohlo evokovať, že právo jednoducho zaostalo za dobou.

[597] Príkladov, keď nové technológie priniesli neistoty do našich životov, je kvantum. Železničná doprava bola kedysi novou technológiou, ktorá viedla k množstvu otázok a k značnej právnej neistote Neskôr na prelome tisícročia sa používanie počítačov stala natoľko rozšíreným, že boli potrebné riešiť neistoty ktoré priniesol tento fenomén. V každom z týchto príkladov nové artefakty, činnosti a vzťahy generované technologickými zmenami v rôznych oblastiach jednoducho nezapadli do existujúcich klasifikačných rámcov. V iných prípadoch nemusí byť problém s klasifikáciou, ale skôr so situáciou, keď nová technológia môže spadať do viacerých samostatných oblastí alebo kategórií. Rôznorodé pravidlá rôznych systémoch sa zrazu začnú miešať, pravidlá ktoré sa nikdy nepredpokladalo, že by museli koexistovať a nie ešte to fungovať na rovnakom objekte alebo subjekte. Preto ak je možné novú technológiu klasifikovať viac ako jedným spôsobom, môžu začať platiť nezlučiteľné pravidlá určené na riadenie rôznych systémov, čo opäť vedie iba k právnej neistote. Ďalším problémom sú technológie, ktoré umožňujú cezhraničnú interakciu a cez rôzne jurisdikcie. Tieto technológie môžu tiež viesť k otázke, ktorá sada pravidiel sa má uplatňovať pri nekonzistentných požiadavkách. Porovnaj

[598] Samozrejme, výsledkom technologických zmien v mnohých prípadoch nie je nič nové, čo by sa nedalo ľahko právne klasifikovať. Často uvádzaný príklad je osobný automobil. Čo je stále vozidlom na účely existujúcich dopravných predpisov aj napriek skutočnosti, že má spaľovací motor, pričom väčšina dopravných pravidiel je staršia ako samotný vynález. Tu však veríme, že je potrebné dodať, že aj keď súhlasíme s takýmto názorom, nedá nám nepovedať, že tak ako vo väčšine prípadov, aj tu sa nájdu výnimky. Keďže kým koč s konským záprahom a auto so spaľovacím motorom budú mať podobné dopady na existujúce pravidlá cestnej premávky, ak však do tohto vzorca začneme pridávať premenné ako algoritmy a umelá inteligencia, ktorá prevezme riadenia vozidla od vodiča, sa dostávame do sféry, kde toto pravidlo nemusí platiť. Porovnaj, GYURÁSZ, Z.: From principles to practice: regulating artificial intelligence, Míľniky práva v stredoeurópskom priestore 2020 [elektronický dokument] : zborník z online medzinárodnej vedeckej konferencie doktorandov a mladých vedeckých pracovníkov. - : 1. vyd. ISBN 978-80-7160-576-8. - Bratislava: Právnická fakulta UK, 2020. - S. 683-688. 2020

[599] Známy je prípad Olivera Wendella Holmesa, keď sa vysmieval vermontskému sudcovi, keď uviedol, že nie je schopný rozhodnúť prípad týkajúci sa maselnice, pretože neexistoval žiadny zákon o maselnici. Porovnaj, HOLMES, O: The Path of the Law, 1897. Dostupné na: <https://warwick.ac.uk/fac/soc/sociology/staff/sfuller/social_theory_law_2015-16/oliver_wendell_holmes_oliver_holmes_the_path_of_law.pdf>. cit. 2023-04-25.

[600] BLACK, J.: Decentring Regulation: Understanding The Role Of Regulation And Self-Regulation In A 'Post- Regulatory' World, 2001.

nevyhnutnosťou. *„Po stáročia sa cnosti vzájomnej dôvery a spravodlivosti ukázali ako základné kamene pre prežitie akéhokoľvek systému uvažovania.*"[601]

10.2.1. Morálny status UI

Dnes je všeobecne akceptované, že súčasné systémy na báze umelej inteligencie nemajú žiadny morálny status, tzn. môžeme ich kedykoľvek zmeniť alebo dokonca aj kompletne zničiť.[602] Naše správanie však nie je výlučne vecou racionality, máme tiež morálne dôvody, ktoré ovplyvňujú naše konania. Keď uvažujeme o možnosti, že niektoré systémy umelej inteligencie by mohli byť kandidátmi na morálny status, vzniká súbor etických otázok, ktoré budeme musieť skôr či neskôr zodpovedať.[603]

Aj keď neexistuje všeobecná zhoda na tom, ako by sa presne mal určiť morálny status, v diskusii prevládajú dve zásadné kritéria.[604] Tieto kritéria sa označujú za:

a) Sentience(**Vnímavosť**) - schopnosť fenomenálneho prežívania alebo kvality, ako je schopnosť cítiť bolesť a utrpenie;
b) Sapience(**Sebauvedomenie**) - súbor kapacít spojených s vyššou inteligenciou, ako je napríklad sebauvedomenie a schopnosť reagovať.

Takéto delenie morálneho stavu naznačuje, že v budúcnosti by takýto systém UI mohol mať určitý morálny status. Istú analógiu tu môžeme vytvoriť so zvieratami. Veľa zo zvierat je sentientných, a preto sa im pripisuje určité morálne, ako aj právne postavenie.[605] Vnímavosť u zvierat je stále iba nižší stupeň morálneho statusu ako sebauvedomenie u ľudí. Zväčša sa za

[601]KUNDU, S.: Ethics in the Age of Artificial Intelligence, 2019. Dostupné na: <https://blogs.scientificamerican.com/observations/ethics-in-the-age-of-artificial-intelligence/> cit. 2023-04-25.
[602] BOSTROM, Nick a Eliezer YUDKOWSKY. The ethics of artificial intelligence. In: Draft for Cambridge Handbook of Artificial Intelligence. Cambridge University Press, 2011. s. 365.
[603] Porovnaj aj s TURČAN, Martin. Etické dimenzie teórie práva. Bratislava: Univerzita Komenského v Bratislave, Právnická fakulta, 2022.
[604] BOSTROM, Nick a Eliezer YUDKOWSKY. The ethics of artificial intelligence. In: Draft for Cambridge Handbook of Artificial Intelligence. Cambridge University Press, 2011. s. 365.
[605] Porovnaj právne postavenie vecí a zvierat v Občianskom zákonníku SR (zákon č. 40/1964 Zb.). Ust. §118 ods. (3) znie: *„Živé zviera má osobitný význam a hodnotu ako živý tvor, ktorý je schopný vnímať vlastnými zmyslami a v občianskoprávnych vzťahoch má osobitné postavenie. Na živé zviera sa vzťahujú ustanovenia o hnuteľných veciach; to neplatí ak to odporuje povahe živého zvieraťa ako živého tvora.*" Toto osobitné postavie vyplíva práve z morálneho státusu zvierat spojené so sentientnostou. Od 01.09.2018 túto skutočnosť odráža aj slovenský občianský zakonník.

sapientné považujú iba ľudské bytosti, čo im dáva vyššie morálne aj právne postavenie ako iným než ľudským bytostiam.[606]

Rýchlosť, ako aj smerovanie napredovania v oblasti UI, naznačuje, že iba čas nás delí od toho, aby sme tu mali sentientné a neskôr možno aj sapientné stroje.[607] Už sentientný systém umelej inteligencie, aj keď mu chýbajú vyššie kognitívne schopnosti, nie je ako jednoduchá hračka. Avšak dôležitejšou otázkou ostáva, čo ak systém umelej inteligencie okrem vnímania dosiahne aj úroveň sebauvedomenia, podobne ako dospelý ľudský jedinec.

a) Mala by potom mať umelá inteligencia plný morálny status, ktorý je rovnocenný so stavom ľudských bytostí?[608]
b) Museli by sme prispôsobiť svoje správanie k týmto systémom tak, ako keby išlo o iného človeka?
c) Bolo by potom nemorálne vyhodiť inteligentný vysávač iba preto, že v rohu obývačky nedokázal poriadne povysávať, alebo zmeniť algoritmus, lebo neprináša očakávané výsledky?

Vnímajúc právny postoj relevantných subjektov v tejto oblasti,[609] ako aj súčasný stav právnej teórie, je nepochybné, že otázka udelenia právnej subjektivity zariadeniam na báze UI je ešte priskorá. Ako naša spoločnosť, tak ani naše právne systémy nie sú pripravené na to, aby sme zaradili systémy na báze UI na úroveň čo i len právneho statusu zvierat. Tieto názory sú reflektované aj vo viacerých dokumentoch, ktoré sme uviedli vyššie.[610] Táto skutočnosť však nemení nič na tom, že systémy založené na báze UI vyzdvihujú skutočný rozdiel medzi našimi explicitnými hodnotami a našimi kolektívnymi skúsenosťami.[611] Naše kolektívne

[606] YAMPOLSKIY, V. Roman. Artificial Intelligence Safety and Security. New York: Chapman and Hall/CRC, 2018, 474 s.
[607] Pozri bližšie Patent č. US2014/046891A1, zo dňa 13.2.2014, s názvom patentu „Sapient or Sentient Artificial Intelligence" od BANAS, Sarah. Dostupné z: https://patentimages.storage.googleapis.com/9f/ed/bc/ab6674f5d979b5/US20140046891A1.pdf
[608] BOSTROM, Nick a Eliezer YUDKOWSKY. The ethics of artificial intelligence. In: Draft for Cambridge Handbook of Artificial Intelligence. Cambridge University Press, 2011. s. 365.
[609] Napríklad už spomenuté EK, EP alebo EPO, UKIPO a USPTO.
[610] Príkladom uveďme uznesenie EP, kde sa spomína, že: „systémy umelej inteligencie nemajú právnu subjektivitu ani ľudské svedomie, a že ich jedinou úlohou je slúžiť ľudstvu" Režim občianskoprávnej zodpovednosti za umelú inteligenciu, 2020 Dostupné z: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_SK.html alebo Rozhodnutie UKIPO, sp. zn. BL O/741/19, kde sa uvádza: „súčasný systém sa nestará o také vynálezy UI a nikdy sa nepredpokladalo, že by bol, ale doba sa zmenila a technológia sa posunula ďalej. Dostupné z: https://www.ipo.gov.uk/p-challenge-decision-results/074119.pdf
[611] Pre problematiku pozri BOSTROM, Nick a Eliezer YUDKOWSKY. The ethics of artificial intelligence. In: Draft for Cambridge Handbook of Artificial Intelligence. Cambridge University Press, 2011. s. 365.

skúsenosti nie sú statické, riadia sa našimi hodnotami, ktoré sa vyvíjajú v čase a sú formované dôležitými spoločenskými rozhodnutiami.[612]

Aj keď sa momentálne môže zdať, že väčšina názorov je odmietavá k danej téme, veľavravné je množstvo dokumentov, ktoré sa venujú vytýčenej problematike. Aj keď dnes stále považujeme systémy UI iba za objekty práva, nie je možné úplne vylúčiť, že sa v budúcnosti dočkáme zmeny tohto postoja. Ľudské hodnoty a postoje sa neustále vyvíjajú a nateraz nikto nevie s určitosťou povedať, kam sa celá táto problematika vyvinie. Len čas ukáže, kam sa aj v tejto oblasti ľudstvo a UI posunie.

Práve z týchto dôvodov môžeme tvrdiť, že áno, etika automatizovaných systémov má svoje nenahraditeľné miesto pri vývoji a následnej implementácii tejto novej technológie do našej spoločnosti.

10.2.2. Etický rámec pre automatizované systémy

Etické diskusie spojené s automatizovanými systémami ktoré sú zdokumentované vo vedeckej literatúre, sa zamerali najmä na extrémne situácie, ktoré budú predstavovať morálne dilemy.[613]

Jedným z takýchto etických dilem, ktorý si získal najväčšiu pozornosť medzi akademickými výskumníkmi, je „električková dilema". Jedná sa o morálnu otázku pri ktorej sa človek má rozhodnúť pre záchranou jednej strany v situáciách, keď hrozí nebezpečenstvo, ktorá je neodvrátiteľná.[614] Existuje mnoho variácií a rozšírení tohto myšlienkového experimentu, ale ich jadro možno definovať ako morálnu voľbu medzi akciami v premávke, ktorých výsledkom budú rôzne kombinácie zachránených a obetovaných životov. Väčšina

[612] ANDERSON, M. ANDERSON, S. Machine Ethics. Ai Magazine, Volume 28, number 4, s. 15-26 [online]. 2016. [cit. 2023-06-15]. Dostupné z: https://www.researchgate.net/publication/220605213_Machine_Ethics_Creating_an_Ethical_Intelligent_Agent

[613] AWAD, E., DSOUZA, S., KIM, R., SCHULZ, J., HENRICH, J., SHARIFF, A.,RAHWAN, I.: The moral machine experiment. Nature, s59–64. 2018. Dostupné na:< https://www.nature.com/articles/s41586-018-0637-6.>. cit. 2023-04-25. alebo BONNEFON, J. F., SHARIFF, A., RAHWAN, I. The social dilemma of autonomous vehicles. Science, 352(6293), 1573–1576. 2016. Dostupné na:< https://www.science.org/doi/abs/10.1126/science.aaf2654>. cit. 2023-04-25.

[614] Táto dilema však nie je úplne nová. Túto istú debatu sme viedli v kontexte protiteroristických opatrení v roku 2001. Jedenásty september 2001 nám ukázal, aké obrovské bezpečnostné riziko môže predstavovať unesený dopravný prostriedok. Dovoľujem si poznamenať, že aj v tomto kontexte stále veríme, že obetovať menšiu, aby bola zachránená väčšina, je klasická vojenskej stratégia. Nastáva preto otázka do akej miery je to možné v civilnom vyspelom demokratickom štáte akceptovať takéto praktiky ak si chceme nazývať právnym štátom?

týchto výskumov bola modelovaná ako myšlienkový experiment založený na nejakej forme tejto legendárnej dilemy morálnej psychológii.[615] [616]

Tieto dilemy boli formulované na otázkach:

a) Ako by mali byť automatizované systémy naprogramované, aby rozhodlo dobre takýchto situáciách?
b) Aké etické a morálne pravidlá a zásady by malo automatizovaný systém dodržiavať?
c) Mali by rozhodnutia automatizovaných systémov reflektovať špecifické charakteristiky potenciálnych subjektov ujmy?[617]

Vedci podrobne diskutovali o relevantnosti morálnych dilem automatizovaných vozidiel, o výhodách používania rôznych etických rámcov, ako je deontológia alebo utilitarizmus ako základ pre algoritmy ktoré robia rozhodnutia.[618] Zaoberali sa morálnymi preferenciami a spoločenskými očakávaniami týkajúce sa etických pravidiel, ktoré by malo byť do nich zakódované.[619]

Prístupy, ktoré zastávajú hodnotu etiky automatizovaných systémov však majú svojich odporcov. Niektorí prominentní akademickí vedci zastávajú názor, že etika automatizovaných systémov je v podstate inžinierskym a politickým rozptýlením a nie naliehavejším problémom. Domnievajú sa, že etika automatizovaných vozidiel je iba

[615] THOMSON, J. J. The trolley problem. Yale Law Journal, 94, 1985. Dostupné na: <http://jonathonklyng.com/wp-content/uploads/2016/08/Thomson-The-trolley-problem.pdf.> cit. 2023-04-25.
[616] Samozrejme existujú ďalšie etické problémy týkajúce sa ochrany osobných údajov, kybernetickej bezpečnosti alebo potenciálu hromadného sledovania. Senzory automatizovaných vozidiel zhromažďujú obrovské množstvo údajov o polohe, premávke a správaní rôznych aktérov. Zatiaľ čo tieto údaje by mohli pomôcť pri navigácii a monitorovaní nelegálnych aktivít, existuje potenciál pre sledovanie ľudí alebo porušenia súkromia.
[617] Uprednostniť život dieťaťa pred dospelým, alebo uprednostniť život ženy pred mužom?
[618] BONNEFON, J. F., SHARIFF, A., RAHWAN, I. The social dilemma of autonomous vehicles. Science, 352(6293), 1573–1576. 2016. Dostupné na:< https://www.science.org/doi/abs/10.1126/science.aaf2654>. cit. 2023-04-25. alebo AWAD, E., DSOUZA, S., KIM, R., SCHULZ, J., HENRICH, J., SHARIFF, A.,RAHWAN, I.: The moral machine experiment. Nature, 563(7729), s59–64. 2018. Dostupné na:< https://www.nature.com/articles/s41586-018-0637-6.>. cit. 2023-04-25. alebo ETZIONI, A., ETZIONI, O. Incorporating ethics into artificial intelligence. The Journal of Ethics, 21(4), s403–418. 2017. Dostupné na: <https://philpapers.org/archive/ETZIEI.pdf>. cit. 2023-04-25.
[619] AWAD, E., DSOUZA, S., KIM, R., SCHULZ, J., HENRICH, J., SHARIFF, A.,RAHWAN, I. The moral machine experiment. Nature, 563(7729), s59–64. 2018. Dostupné na:< https://www.nature.com/articles/s41586-018-0637-6.>. cit. 2023-04-25. alebo BONNEFON, J. F., SHARIFF, A., RAHWAN, I. The social dilemma of autonomous vehicles. Science, 352(6293), 1573–1576.

rozptýlením a má malú praktickú hodnotu pre implementáciu automatizovaných systémov.[620]

Riešenie otázok etiky umelo inteligentných systémov je skutočne netriviálna úloha. Viacerí z odborníkov poukazujú na takmer nemožnosť riešenia týchto problémov.[621] Bez konsenzu o univerzálnom morálnom kódexe však by bolo takmer nemožné vyvinúť automatizované rozhodovanie, ktoré by univerzálne napĺňalo etické rámce očakávané obyvateľstvom na celom svete. Otázka či je verejná mienka tým najlepším spôsobom, ako ponúknuť riešenie pre otázky etiky automatizovaných vozidiel alebo by to malo byť ponechané v rukách odborníkov, nie je len predmetom súčasného výskumu.[622] Samozrejme možno sa taktiež domnievať, že nejestvuje správny orgán moci, ktorý by mal rozhodovať o etike autonómnych systémov. A, že rozhodnutie musí ostať v rukách jednotlivca tak ako to bolo doposiaľ v kontexte tradičných zariadení. Keďže väčšina rozhodnutí urobených automatizovanými systémami počas ich bežnej prevádzky prinesie len veľmi málo potenciálnych rizík pre užívateľov.

10.2.3 Etické usmernenia pre dôveryhodnú umelú inteligenciu

Európska komisia v roku 2019 zverejnila Etické usmernenie pre dôveryhodnú AI, ktoré vypracovala expertná skupina na vysokej úrovni pre umelú inteligenciu (AI HLEG). Toto usmernenie sa zameriava na tri kľúčové oblasti:

1. **Zákonnosť:** Umelá inteligencia by mala byť v súlade s platnými právnymi predpismi a rešpektovať základné práva a slobody.

[620] DE FREITAS, J., CENSI, A., SMITH, B. W., DI LILLO, L., ANTHONY, S. E., FRAZZOLI, E. From driverless dilemmas to more practical commonsense tests for automated vehicles. 2021 Dostupné na:<https://www.pnas.org/doi/full/10.1073/pnas.2010202118> cit. 2023-04-25. alebo DEWITT, B., FISCHHOFF, B., & SAHLIN, N. 'Moral machine' experiment is no basis for policymaking. Nature, 567, 2019 Dostupné na:<https://go.gale.com/ps/i.do?id=GALE%7CA577251215&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=00280836&p=HRCA&sw=w&userGroupName=anon%7E297fcce4> cit. 2023-04-25.alebo NYHOLM, S., SMIDS, J. The ethics of accident-algorithms for self-driving cars: An applied trolley problem? Ethical Theory and Moral Practice, 19(5). 2016 Dostupné na:<https://link.springer.com/article/10.1007/s10677-016-9745-2> cit. 2023-04-25.

[621] DE FREITAS, J., CENSI, A., SMITH, B. W., DI LILLO, L., ANTHONY, S. E., FRAZZOLI, E. From driverless dilemmas to more practical commonsense tests for automated vehicles. 2021 Dostupné na:<https://www.pnas.org/doi/full/10.1073/pnas.2010202118> cit. 2023-04-25.

[622] AWAD, E., DSOUZA, S., KIM, R., SCHULZ, J., HENRICH, J., SHARIFF, A.,RAHWAN, I. The moral machine experiment. Nature, 563(7729), s59–64. 2018. Dostupné na:< https://www.nature.com/articles/s41586-018-0637-6.>. cit. 2023-04-25.

2. Etika: Umelá inteligencia by mala byť vyvíjaná a používaná eticky, s ohľadom na humanistické hodnoty a zodpovednosť.

3. Robustnosť (odolnosť): Umelá inteligencia musí byť odolná a to z technického aj sociálneho hľadiska, keďže systémy umelej inteligencie môžu aj pri dobrých úmysloch spôsobiť neúmyselnú ujmu.

Základ dôveryhodnej AI v zmysle usmernenia predstavujú štyri etické zásady:

- Rešpektovanie ľudskej autonómie
- Prevencia ujmy
- Spravodlivosť
- Vysvetliteľnosť.

Požiadavka rešpektovania ľudskej autonómie znamená, že AI by mala posilňovať, nie obmedzovať ľudskú slobodu a autonómiu a práve ľudia musia mať konečnú kontrolu nad systémami AI. AI by sa mala navrhovať tak, aby rešpektovala ľudské práva a etické princípy a mala v práci ľudí dopĺňať a uľahčovať im prácu a vytvárať zmysluplnú prácu.[623] V kontexte zásady prevencie ujmy by AI nemala spôsobovať ujmu ľuďom ani zhoršovať ich životné podmienky. AI by mala rešpektovať ľudskú dôstojnosť a chrániť duševnú a telesnú integritu.

Systémy AI musia byť bezpečné a odolné voči zneužitiu, s dôrazom na ochranu zraniteľných skupín.[624] AI by mala reflektovať aj spravodlivosť a rovnako rozdeliť prínosy a náklady AI a zabrániť nespravodlivej zaujatosti, diskriminácii a stigmatizácii AI by mala prispievať k rovnakým príležitostiam a k posilneniu spoločenskej spravodlivosti. Na druhej strane, AI by nemala oklamať používateľov ani oslabovať ich slobodnú voľbu. Zodpovednosť za rozhodnutia AI, transparentnosť procesov a možnosť nápravy.[625] Poslednou zásadou je vysvetliteľnosť, ktorá je kľúčová pre budovanie a udržiavanie dôvery používateľov v AI. Procesy a schopnosti AI musia byť transparentné a zrozumiteľné. Rozhodnutia AI musia byť vysvetliteľné pre osoby, ktorých sa dotýkajú.[626]

[623] Etické usmernenie pre dôveryhodnú AI, s. 14 – 15.
[624] Tamže, s. 15.
[625] Tamže, s. 15.
[626] Tamže, s. 15 – 16.

Usmernenie sa osobitne týka etických požiadaviek a definuje 7 kľúčových požiadaviek dôveryhodnej AI:

- Ľudský faktor a dohľad
- Technická odolnosť a bezpečnosť
- Správa súkromia a údajov
- Transparentnosť
- Rozmanitosť, nediskriminácia a spravodlivosť
- Spoločenský a environmentálny blahobyt
- Zodpovednosť.

Usmernenie AI HLEG je dôležitým krokom k rozvoju dôveryhodnej a etickej umelej inteligencie v Európe. Poskytuje rámec pre tvorcov politík, výskumníkov a vývojárov AI, ako aj pre širokú verejnosť, aby sa zapojili do diskusie o budúcnosti umelej inteligencie.

FACULTY OF LAW
Comenius University
Bratislava

ISBN: 978-80-7160-715-1