

**UNIVERZITA KARLOVA**

**Právnická fakulta**

**Laura Hospodková Tadevosjanová**

**Právní a etické aspekty regulace autonomních  
systémů**

**Umělá inteligence v právu**

**Diplomová práce**

Vedoucí diplomové práce:

JUDr. Mgr. Michal Urban, Ph.D.

Katedra:

Katedra politologie a sociologie

Datum vypracování práce:

15. 5. 2020

### **Čestné prohlášení**

Prohlašuji, že jsem předkládanou diplomovou prací vypracovala samostatně, že všechny použité zdroje byly řádně uvedeny a že práce nebyla využita k získání jiného nebo stejného titulu.

Dále prohlašuji, že vlastní text této práce včetně poznámek pod čarou má 222.013 znaků včetně mezer.

.....

diplomantka

V Praze dne 15. 5. 2020

# Obsah

Slovo úvodem.....	5
1 Autonomní systémy a představy o nich.....	10
2 Autonomní systémy, roboti a umělé inteligence .....	15
2.1 Vznik pojmu umělá inteligence .....	16
2.2 Definice autonomních systémů, robotů a umělé inteligence .....	18
2.2.1 Definice umělé inteligence.....	19
2.2.2 Definice (inteligentního) robota.....	23
2.2.3 Definiční znaky autonomních systémů .....	25
2.3 Otázka autonomie a jak jí v kontextu autonomních systémů rozumět .....	31
2.4 Kontrola nad autonomními systémy .....	34
3 Pravidla chování autonomních systémů, rozvoj regulačního rámce a role práva ve věku inteligentních robotů.....	39
3.1 Právo umělé inteligence, zajištění právní jistoty a předvídatelnosti.....	39
3.1.1 Role práva ve věku inteligentních robotů .....	40
3.1.2 Status a povaha umělé inteligence .....	44
3.1.3 Umělá inteligence jako osoba v právu .....	49
3.1.4 Doporučení.....	52
4 Odpovědnost za jednání autonomních systémů.....	54
4.1 Problém s odpovědností u autonomních systémů.....	54
4.2 Současné pojetí občanskoprávní odpovědnosti .....	57
4.2.1 Prevenční povinnost .....	58
4.2.2 Subjektivní a objektivní odpovědnost.....	62
4.3 Odpovědnost za jednání autonomních systémů v kontextu účinné právní úpravy .....	67
4.3.1 Škoda způsobená provozem zvláště nebezpečným.....	67
4.3.2 Škoda, kterou způsobí věc sama od sebe .....	69
4.3.3 Škoda způsobená vadou výrobku.....	70
4.3.4 Škoda z provozu dopravních prostředků.....	72

4.3.5	Potřeba adaptace právní úpravy .....	73
5	Etické aspekty regulace autonomních systémů .....	77
5.1	Roboetika .....	82
5.1.1	Transparentnost a důvěra .....	84
5.1.2	Ochrana soukromí a dopad na lidské chování.....	88
5.1.3	Předsudky a diskriminace .....	92
5.2	Etika strojů .....	97
5.3	Charta inženýrů robotiky .....	103
5.4	Doporučení Etické komise .....	104
	Závěr.....	107
	Seznam použitých zdrojů .....	112
	Seznam použité literatury .....	112
	Seznam použitých internetových zdrojů .....	113
	Seznam použitých právních předpisů.....	118
	Seznam použité judikatury .....	119
	Seznam ostatních zdrojů.....	119

## **Poděkování**

Ráda bych tímto poděkovala svému vedoucímu práce, JUDr. Mgr. Michalu Urbanovi, Ph.D., za ochotu vést moji diplomovou práci a umožnění zabývat se tímto zajímavým tématem, za odborné vedení, za pomoc a rady. Dále bych ráda poděkovala své rodině a zejména manželovi za trpělivost a veškerou podporu při zpracování této práce.

## Slovo úvodem

Není tomu tak dávno, kdy jsme se mohli s představou robotů, umělé inteligence a dalších příkladů tzv. autonomních systémů setkat zejména jako s poutavým tématem sci-fi příběhů a futuristických idejí některých autorů. Když nahlédneme do literatury, spatříme, že touha po myslících, samoučících a inteligentních strojích provází lidstvo od nepaměti. Jistě nebudu daleko od pravdy, když prohlásím, že každého z nás v určitém okamžiku v životě napadla a lákala představa robotů obsluhujících v restauracích, létajících automobilů, cestování vesmírem nebo policie, která na Marsu ukládá pokuty za překročení rychlosti, jakožto vize budoucnosti, která není až tak daleko od současného světa, tedy současného stavu vědeckého a technologického poznání.

Toto období je již do určité míry překonané a autonomní systémy v různých podobách již několik let představují neodmyslitelný prvek světa okolo nás. Je jenom přirozené, že lidská závislost na chytrých technologiích stále roste a využívání takových technologií již představuje přirozený jev ve společnosti. Stále více zařízení má zabudované určité autonomní prvky, které nám do jisté míry usnadňují každodenní rutinní činnosti, aniž bychom to vždy tušili. Čím dál více společností se zabývá vývojem autonomních systémů vybavených umělou inteligencí, zejména pak vývoj chatbotů, autonomních vozidel či autonomních dronů je v dnešní době dobrým příkladem toho, že provoz podobných systémů není tak daleko od reality.

Za poslední roky jsme si zvykli, že moderní technologie postupují velmi rychlým tempem a právo na jejich rozvoj a potenciál reaguje se značným zpožděním. A oblast autonomních systémů a umělé inteligence není výjimkou.

*„Technologie robotiky bude hrát v nadcházejícím desetiletí prim. Bude ovlivňovat každý aspekt práce a domova. Robotika má potenciál transformovat životy a pracovní postupy, zvyšovat úroveň efektivity, bezpečnosti a poskytovaných služeb a vytvářet pracovní místa. Její vliv bude postupem času růst stejně jako interakce mezi roboty a lidmi.“*

Takové závěry o budoucnosti technologie robotiky činí Evropská unie v reformní agendě s výhledem do roku 2020 známé pod názvem „Strategie EU Robotika 2020“.<sup>1</sup> Evropská unie chce co nejlépe využít tento rostoucí trh prostřednictvím spolupráce veřejného a soukromého sektoru, orgány veřejné správy, průmyslem a akademickou obcí. Není pochyb o tom, že taková spolupráce má potenciál zlepšit konkurenceschopnost průmyslu Evropské unie a umožnit zavádění robotů a služeb, které evropskému průmyslu pomohou vyřešit některé společenské výzvy, jako je například stárnutí obyvatelstva. Úkolem všech zainteresovaných osob je však vyřešit etické a právní otázky s tímto vývojem spojené a vypracovat široce využitelné kodexy, ať už v méně či více formální podobě. Úkolem legislativců je adaptovat současné právní předpisy, aby udržely krok s vývojem.

Pro lepší představu, již v roce 2010 vyšel článek v New York Times oznamující testování autonomních vozidel na dálnicích ve Spojených státech, kde společnost Google poukázala na skutečnost, že fenomén bezpilotních dopravních prostředků již zdaleka není fikcí, a upozornila na rozsah dopadu širší implementace takových prostředků v rámci dopravní infrastruktury po celém světě: *„autonomous cars are years from mass production, but technologists who have long dreamed of them believe that they can transform society as profoundly as the Internet has.“*<sup>2</sup>

Přípodobnění dopadům, které na společnost měly počátky komerčního využívání internetu, je s ohledem na současný stav jeho využití a obchodního potenciálu velice zajímavý. Není proto divu, že autonomní systémy a zejména umělá inteligence, byť zatím známe pouze režim strojového učení, čím dál častěji poutají pozornost právníků i neprávníků, odborníků i laiků.

---

<sup>1</sup> Budoucnost práce: robotika - Revue pro sociální politiku a výzkum. Revue pro sociální politiku a výzkum [online]. Institut pro sociální politiku a výzkum z. s. [cit. 25. 02. 2020]. Dostupné z: <https://socialnopolitika.eu/2017/01/budoucnost-prace-robotika/>.

<sup>2</sup> MARKOFF, J. *Google Cars Drive Themselves, in Traffic*. The New York Times, říjen 2010 [cit. 25. 7. 2020]. Dostupné z: <https://www.nytimes.com/2010/10/10/science/10google.html>.

Již dnes se umělá inteligence, resp. strojové učení, využívá při detekci infarktů, analýze hrozeb kybernetického útoku, nábore nových zaměstnanců, překladu náročných textů a v mnoha dalších různorodých odvětvích. V průběhu příštích desetiletí se bude tento výčet příkladů dále rozšiřovat, dokud nebude umělá inteligence více ovlivňovat stále širší okruh oblastí, jako jsou medicína, ekonomika ale například i politika.

Jak uvádí Vláda ČR ve své studii *Výzkum potenciálu rozvoje umělé inteligence v České republice* z 10. prosince 2018,<sup>3</sup> s ohledem na současný vývoj technologické úrovně v oblasti umělé inteligence lze očekávat, že v horizontu do 5 let budou technologie schopny nahradit více než 50 % dovedností požadovaných pro výkon přibližně 11 % povolání. V horizontu do 30 let pak můžeme uvažovat o více než 50 % dovedností v naprosté většině současných povolání.

Je pozoruhodné, že nové technologie hrají v našem každodenním životě mnohdy větší roli, než si dokážeme představit. Vzhledem ke skutečnosti, že nové technologie v jisté míře již dnes způsobují revoluci ve způsobu, jakým lidé myslí, učí se a pracují v různých oblastech, je důležité (a současně zajímavé) vývoj v této oblasti sledovat. Z tohoto důvodu jsem zvolila toto téma jako téma své diplomové práce i oblast, které bych se ráda do budoucna dále a detailněji věnovala.<sup>4</sup>

Příchod nových technologií a softwaru, který může fungovat stále více nezávisle na lidech a může vykonávat úkoly, které by jinak vyžadovaly inteligenci, když by je prováděli lidé, vyžaduje zvláštní pozornost. Není proto divu, že moderní technologie už roky v lidech vyvolávají zájem, ale zároveň i hluboké obavy. Jak technologie, tak právo mohou být dobrým sluhou ale i zlým pánem, pokud jim včas a správně neporozumíme, byť scénáře typu *Já, robot*

---

<sup>3</sup> Jaký je potenciál umělé inteligence v České republice? Vláda ČR [online] [cit. 28. 11. 2019]. Dostupné z: <https://www.vlada.cz/cz/evropske-zalezitosti/aktualne/jaky-je-potencial-umele-inteligence-v-ceske-republice--170808/>.

<sup>4</sup> Tento výstup vznikl v rámci projektu Specifického vysokoškolského výzkumu č. 260 495/ 2020.



či *A.I. Artificial Intelligence* můžeme zatím pustit z hlavy. Proto se práce v první a navazující druhé kapitole věnuje obecně úvodu do tématu autonomních systémů s důrazem na umělou inteligenci, resp. technologie strojového učení, popisu představ společnosti o těchto nových technologiích, základním definičním znakům autonomních systémů, robotů i umělé inteligence. Zejména definiční znaky jsou důležité pro porozumění a analýzu pojmu „autonomie“ v kontextu nových technologií, dopadu na současný stav poznání i mentálního nastavení společnosti a otázky udržení kontroly nad jednáním a rozhodnutími učiněnými v souvislosti s novými technologiemi, kterým se práce rovněž věnuje ve druhé kapitole.

Tyto systémy vyvolávají řadu důležitých otázek právě v oblasti právní úpravy. Analýza a úprava regulatorního rámce představuje jednu z klíčových podmínek úspěšného rozvíjení nových technologií a současně důležitý korektiv pro případ ztráty obecné kontroly nad tímto rozvojem a jeho dopady. Třetí kapitola této práce se proto věnuje pravidlům chování autonomních systémů, rozvoji regulatorního rámce a roli práva ve věku inteligentních robotů tak, aby byla zajištěna dostatečná míra právní jistoty a předvídatelnosti. Tato kapitola obsahuje rovněž posouzení jedné z nejkontroverznějších otázek v souvislosti s umělou inteligencí, a to otázky, zda může fungovat koncept umělé inteligence jako osoby v právu.

Jednou z nejdiskutovanějších otázek v souvislosti s umělou inteligencí je otázka odpovědnosti za její jednání a rozhodování. Je nutné si uvědomit, že ani ta nejlepší, nejnovější technologie není zcela bezchybná, domácí roboti, autonomní vozy i všechny další autonomní systémy budou nevyhnutelně způsobovat škodu či nemajetkovou újmu. Aby mohla společnost akceptovat nové technologie a podpořit inovace, je třeba, aby existovala jasná, transparentní a srozumitelná pravidla určení občanskoprávní odpovědnosti. Čtvrtá kapitola této práce se tak věnuje právě otázce odpovědnosti, a to zejména v kontextu současného pojetí subjektivní a objektivní odpovědnosti, prevenční povinnosti i odpovědnost za jednání autonomních systémů v kontextu účinné právní úpravy a možnosti uplatnění různých odpovědnostních režimů.

Pokroky v oblasti umělé inteligence, robotiky a jiných autonomních technologií vyvolaly i řadu naléhavých a složitých morálních otázek. Současné snahy najít odpovědi na

etické, společenské a další právní výzvy s pokrokem spojené a směřovat je správným směrem již vedly k mnoha iniciativám na poli různých jurisdikcí i organizací. Některé z těchto problémů pramení z otázek týkajících se bezpečnosti, ochrany, zmírňování rizik. Jak lze vytvořit svět se vzájemně propojenými autonomními zařízeními, která jsou bezpečná a zabezpečená? Jak lze identifikovat a předcházet předsudkům a nerovnostem, které jsou vtěleny do algoritmů či vstupních dat? S provozem autonomních systémů jsou neoddělitelně spojeny právě morální otázky, které lze obecně rozdělit do dvou kategorií – na tzv. roboetiku, tj. oblast etických problémů, s nimiž se potýkají lidé navrhující, vyvíjející a používající inteligentní stroje, a na etiku strojů, tj. oblast zabývající se situacemi, kdy stroje samy rozhodují o etických problémech. Konečně v tomto ohledu vznikají i otázky týkající se vysvětlitelnosti a transparentnosti autonomních systémů, a to právě zejména umělé inteligence, resp. strojového učení. Poslední a stěžejní kapitola této práce obsahuje analýzu těchto otázek s cílem dosažení morálních a důvěryhodných nových technologií.

Mezi použité metody v této diplomové práci patří zejména metoda analýzy a popisná metoda. Jako podklady při přípravě a vypracování sloužily odborné publikace knižní, časopisecké či elektronické. Podstatnou část zdrojů této práce tvoří rovněž mezinárodní dokumenty z dané oblasti a národní i unijní právní předpisy upravující předmětnou materii.

# 1 Autonomní systémy a představy o nich

Jak bylo uvedeno v úvodu této práce, lidstvo od nepaměti touží vytvořit „umělou bytost“. Tento fenomén se projevuje zejména napříč různými oblastmi umělecké tvorby, jako je literatura či malířství. Přestože si takovou bytost většina společnosti představí ve formě humanoida, který se do značné míry chová, pohybuje, mluví a reaguje obdobně jako člověk, nemusí tomu tak zákonitě být vždy. Takové „umělé bytosti“ dnes například obsluhují zákazníky internetových obchodů, zajišťují komunikaci s těmito zákazníky a často také činí rozhodnutí, jež mohou mít pro konkrétní osoby právní důsledky (ve formě uzavření kupní smlouvy, akceptace reklamace apod.).<sup>5</sup> Síť restaurací Dominos například využívá prostřednictvím nejméně 15 platforem, včetně aplikace Facebook Messenger, Amazon Echo, Samsung TV, Apple Watch a dalších chatbota, který komunikuje se zákazníky a přijímá objednávky, což podstatně urychluje proces objednávání a nabízí nejjednodušší způsob, jak si objednat pizzu odkudkoliv.<sup>6</sup> Obsluha zákaznického servisu však nepředstavuje jediné praktické využití tzv. chatbotů, chatbot společnosti Endurance například pomáhá lidem s Alzheimerovou chorobou, kteří se potýkají s krátkodobou ztrátou paměti, a v reálném čase identifikuje odchylky v konverzaci, které mohou naznačovat problém s krátkodobou pamětí.<sup>7</sup>

---

<sup>5</sup> V České republice byl jedním z prvních veřejně diskutovaných systémů chatbot společnosti Datasys, která mimo jiné zajišťuje servis pro společnost Dell, jak uvádí server [lupa.cz](http://lupa.cz): „Bot tak v současné době umí komunikovat přes Skype, Facebook a web. Zvládne identifikovat zákazníka, zjistit problém, zda vzniká nárok na reklamaci a další podobné úkoly. Pak vytvoří „ticket“ v helpdeskovém systému či pošle e-mail. Tam už pak roli přebírá člověk s tím, že nemusí řešit všechny úvodní detaily, jako je třeba zjišťování servisního čísla notebooku.“

<sup>6</sup> STEIN, Bram. *You can now order Domino's pizza through a chatbot on Facebook Messenger*. Business Insider [online]. [cit. 22. 4. 2020]. Dostupné z: <https://www.businessinsider.com/you-can-now-order-dominos-pizza-through-a-chatbot-on-facebook-messenger-2016-9>.

<sup>7</sup> Viz blíže *A Robot-Companion for Senior People and Patients with Alzheimer's Disease*, dostupné z: <http://endurancerobots.com/azbnmaterial/a-robot-companion-for-senior-people-and-patients-with-alzheimer-s-disease/>.

Jen málo lidí si uvědomí, s kolika inteligentními roboty skutečně přichází neustále do kontaktu. Někteří z nich vůči nám činí různé úkony a někteří z nich s námi komunikují, aniž bychom věděli, že se jedná o robota,<sup>8</sup> jelikož nevidíme jejich hmotný substrát (tj. konkrétní stroj, či chcete-li – androida), který tato jednání činí. Jako příklad jednoho z nejnápaditějších využití inteligentního robota lze pak uvést projekt společnosti Disney, která k zapojení mladších diváků, použila v aplikaci Facebook Messenger chatbota, který představoval postavu z animovaného filmu Zootopia.<sup>9</sup> Fanoušci filmu mohli hrát detektivní hru založenou na detailech z filmu Zootopia a interakcí s robotem řešit záhady, kdy robot v reálném čase v rámci chatu zkoumal vstupy od uživatelů.

V kontextu výše zmíněných skutečností se dnes často setkáme s pojmem algoritmus – předdefinovaným, přesným postupem upravujícím řešení určité řady obdobných úloh, tj. postupem, který je charakteristický širokou množinou možných vstupů a svou bezvýjimečnou resultativností, tedy dovedením zadaných vstupů ke konkrétnímu výstupu. Autonomní systémy jsou charakteristické řadou implementovaných algoritmů, které definují jejich chování. V souvislosti s množstvím autonomních systémů operujících v oblasti internetu dnes můžeme hovořit o tzv. ekonomice algoritmů.<sup>10</sup> Tyto systémy dnes často určují cenu výrobků, které na internetu nakupujeme, reklamu, která je nám zobrazována, i doporučené internetové stránky, které by nás mohly zajímat. Tato rozhodnutí nečiní člověk, nýbrž počítačový program, který operuje na základě vstupů (již ne jen vložených, nýbrž i vypočítaných, dat). V důsledku toho jsou tyto výsledné „hodnoty“ mnohem přesnější a lépe šité na míru jejich adresátům.

---

<sup>8</sup> K tomu blíže Robocalls: Last Week Tonight with John Oliver (HBO) - YouTube. *YouTube* [online]. Dostupné z: [https://www.youtube.com/watch?v=FO0iG\\_P0P6M](https://www.youtube.com/watch?v=FO0iG_P0P6M).

<sup>9</sup> JAEKEL, Brielle. *Zootopia fans help solve crimes via Facebook Messenger chatbot*. Mobile Marketer. Mobile Marketing News. Mobile Marketer [online]. [cit. 22. 4. 2020]. Dostupné z: <https://www.mobilemarketer.com/ex/mobilemarketer/cms/news/messaging/22953.html>.

<sup>10</sup> BURROWS, Leah. *An economy of algorithms*. Harvard John A. Paulson School of Engineering and Applied Sciences [cit. 12. 01. 2020]. Dostupné z: <https://www.seas.harvard.edu/news/2017/01/economy-algorithms>.

Lze se však domnívat, že autonomní systémy budoucnosti, vybavené umělou inteligencí, nebudou mít převážně podobu „humanoida“ (v kontextu robotiky androida, tj. umělého organismu, jehož tělo se podobá tělu lidskému). Vzhledem k tomu, jakým směrem a do jakých oblastí toto odvětví technologie expanduje,<sup>11</sup> se lze domnívat, že většina inteligentních systémů nebude vůbec tvořena hmotným substrátem tak, jak si ho logicky představujeme. Na rozdíl od autonomních vozidel, letadel a dalších, u kterých slovní spojení „inteligentní stroj“ dává větší smysl, bude většina autonomních systémů představována právě počítačovými programy, vzorci a algoritmy. V tomto kontextu se můžeme setkat s několika termíny, které se v souvislosti či společně s autonomními systémy často používají, zejména robot, inteligentní/chytrý/autonomní stroj apod. Tyto termíny většinou popisují tutéž věc nehledě na skutečnost, jakou fyzickou podobu autonomní systém má, nýbrž s poukazem na jeho charakteristické vlastnosti.

Autonomní vlastnosti takových zařízení spočívají zejména v jejich schopnostech důkladné analýzy okolního prostředí, samostatného jednání, a případně i schopnosti samoučení na základě konkrétních situací. Vývojáři autonomních systémů, konkrétně pak zejména autonomních dopravních prostředků, upozorňují na mnohé výhody, které takové systémy mohou poskytnout. Nejpůsobivější z nich je příslib levnějších služeb, bezpečnější dopravy s možností eliminace nehod způsobených člověkem (nepozornost, návykové látky apod.), jelikož tyto často vyplývají z nedokonalostí člověka ve smyslu nedokonalostí ve srovnání s autonomními systémy, které by měly být méně chybové. Dále výrobci a vývojáři upozorňují na výhody spojené s úsporou času a možností věnovat se perspektivnějším činnostem právě v důsledku přenechání některých rutinních, fyzicky nenáročných či mentálně jednoduchých, opakujících se úkolů autonomním systémům. Konkrétní charakteristické vlastnosti autonomních systému budou popsány v následujících kapitolách.

---

<sup>11</sup> *The Future of Artificial Intelligence - Will Robots/Machines Outsmart Humans?* TechPats – Technology Patent Services & Intellectual Property Consulting [cit. 12. 01. 2020]. Dostupné z: <https://www.techpats.com/future-artificial-intelligence-will-robotsmachines-outsmart-humans/>.

Přestože společnost se střemhlav řítí do světa, který je „chytrý“, propojený a v tomto ohledu velice pohodlný, otázkou zůstává, zda je na takové technologie skutečně připravena. Přesto, že některé oblasti lidské činnosti budou „zasazeny“ pravděpodobně až za několik desítek let nebo i později, je třeba myslet na možné důsledky stále širšího využití autonomních systémů, včetně umělé inteligence, a postupně nacházet odpovědi na vznikající otázky, aby mohla právní úprava udržet s rychlým vývojem chytrých technologií krok a umožnila průchod budoucímu vývoji. Inteligentní stroje dnes už nejen řeší některé problémy vyhodnocováním konkrétních okolností a vyvozováním potřebných závěrů, avšak i formulují hypotézy či vzorce pro řešení těchto problémů. Všechny takové stroje či programy, které spadají pod pojem „umělá inteligence“, byť zatím nelze hovořit o umělé inteligenci v pravém slova smyslu (viz blíže v následujících kapitolách), nýbrž o strojovém učení, jsou zatím pod kontrolou lidí, neboli jednájí dle instrukcí člověka, který je stále může v případě potřeby vypnout či přeprogramovat. Je však otázkou, zda tomu tak bude vždy, tj. zda bude tato kontrola vždy udržitelná, a to právě v důsledku čím dál vyšší a komplexnější úrovně autonomie.

Zatímco de lege lata jsme ohraničeni platnými pravidly, které v určitých aspektech regulují výrobu, zkoušení a využívání chytrých technologií a umožňují jejich implementaci, tato pravidla do určité míry představují překážku pro budoucí pokrok, jak bude blíže vysvětleno v následujících kapitolách, z pohledu de lege ferenda existuje nespočet nedořešených (a do jisté míry i v současné chvíli nepředstavitelných) otázek, které bude dříve nebo později nutné zodpovědět. Autonomní systémy a zejména technologie umělé inteligence či strojového učení mají potenciál změnit povahu téměř všeho, co je nerozlučitelně spjato s lidským životem. V budoucnu zasáhne umělá inteligence ať už v menší nebo větší míře většinu oblastí lidské činnosti od zaměstnání, způsobu komunikace, ochrany dat a soukromí, zdravotní péče, bezpečnosti až po logistiku, vedení válečných konfliktů a další. Například v logistice se nejedná pouze o samořídící automobily, ale autonomní systémy se již delší dobu

využívají například v letectví či ve skladovacích a výrobních provozech.<sup>12</sup> Přestože situace, kdy budou na pozemních komunikacích ke spatření pouze samořídící automobily, nákup na internetu bude obsluhován pouze chatboty (počítačovými programy určenými ke komunikaci se zákazníky) apod. je zatím v nedohlednu, počet takových systémů se bude přirozeně stále zvyšovat a je třeba mít na paměti, že otázky s jejich provozem spojené budou vyplývat průběžně, dokud nebude společnost plně připravena na život vedle těchto umělých bytostí.

---

<sup>12</sup> GUILLOT, Craig, *4 types of autonomous mobile robots, and their warehouse use cases*. Supply Chain Dive. Supply Chain News and Analysis [cit. 28. 11. 2019]. Dostupné z: <https://www.supplychaindive.com/news/4-types-of-autonomous-mobile-robots-and-their-warehouse-use-cases/529548//>.

## 2 Autonomní systémy, roboti a umělé inteligence

*„Kdyby existovaly stroje, podobající se našim tělům a napodobující naše úkony potud, pokud by to mravně bylo možné, měli bychom vždy dva velice vážné důvody, abychom poznali, že proto ještě nejsou skutečnými lidmi. První důvod je, že by nikdy nemohly užívat slov ani jiných znaků, skládající je, jako činíme my, abychom své myšlenky vyložili jiným. (...) A druhý důvod je, že i kdyby vykonávaly určité věci stejně dobře nebo snad i lépe než kdokoli z nás, selhaly by nevyhnutelně v jiných, při nichž by vyšlo najevo, že nejednaly s vědomím, nýbrž toliko podle sestavení svých orgánů; neboť rozum je všestranný nástroj, kterého lze užívat ve všech možných případech, kdežto tyto orgány musí mít nějaké zvláštní uzpůsobení pro každý úkon jednotlivý; a proto je morálně nemožné, aby rozmanitost těchto orgánů v jednom stroji stačila přivést jej k tomu, aby jednal za všech okolností života stejně, jako jednáme my vlivem svého rozumu.“<sup>13</sup>*

Je zajímavé, že lidé se otázkou, zda je možné, aby uměle vytvořené systémy dosahovaly stejného chování a inteligence jako živé organismy, zabývali dokonce i mnohem dříve, než byly vynalezeny počítače. Tuto otázku řešil již v 17. století i Descartes a už tehdy upozornil na rizika spojená s možnou existencí strojů podobných a napodobujících lidí. Dle Descarta by však takové systémy byly vždy rozpoznatelné na základě dvou kritérií, a to absence schopnosti užívat slova a jiné znaky tak, jako lidé, a omezeného rozsahu možných činností a aplikací v důsledku nedostatku vědomí, které je natolik komplexní, že je lze užít ve všech myslitelných situacích. Tyto závěry jsou samozřejmě pouze filosofické a jsou výsledkem bádání, zda mohou stroje vůbec myslet ve smyslu, ve kterém myslí lidé či jiné organismy, a nedosahují, a vzhledem k tehdejšímu technologickému poznání ani nemohou, úvah o dopadech a aplikaci technologií tak pokročilých, jako je například strojové učení.

---

<sup>13</sup> DESCARTES, René. *Rozprava o metodě: jak vést správně rozum a hledat pravdu ve vědách*. Přeložil Karel ŠPRUNK. Praha: OIKOYMENH, 2016. Knihovna novověké tradice a současnosti. ISBN 978-80-7298-212-7.



Mechanistický výklad chování strojů<sup>14</sup> (Descartes do této skupiny přidává i zvířata) považovali za nedostatečný již různí teologové a filosofové. Ti namítali, že argument, že lze veškeré jejich jednání dostatečně vysvětlit pouze pomocí mechaniky, není přesvědčivý.<sup>15</sup> Je skutečně otázkou, zda Descartova úvaha o orgánech stroje, které mají, ve srovnání s všestranným použitím rozumu, své specializované určení, může dnes a v budoucích letech uspět s takovou určitostí. Aby bylo možné důkladně analyzovat chování strojů a případné problematické aspekty s tímto chováním<sup>16</sup> a s jejich regulací spojené, je třeba si tyto pojmy předně definovat.

## 2.1 Vznik pojmu umělá inteligence

Pokud uvažujeme pojem *umělý* jako protiklad k pojmu *přirozený*, tj. vytvořený jako imitace něčeho, co se přirozeně vyskytuje ve světě, lze při hledání počátků vědy o umělé inteligenci pravděpodobně dojít až do 50. let 20. století a fenoménu známému pod názvem *Turingův test*.<sup>17</sup> Turingův test je známý test umělé inteligence navržený Alanem Turingem, britským matematikem a zakladatelem moderní informatiky,<sup>18</sup> jehož cílem bylo zjistit, zda může počítač „myslet“, konkrétně, zda je člověk schopný rozpoznat, že komunikuje

---

<sup>14</sup> DESCARTES, René. *Meditace o první filosofii: námítky a autorovy odpovědi*. Přeložil Tomáš MARVAN, přeložil Petr GLOMBÍČEK, přeložil Pavel ZAVADIL. Praha: OIKOYMENH, 2003. ISBN 80-7298-084-X.

<sup>15</sup> BURGAT, Florence. *Svoboda a neklid zvířecího života*. Přeložila Olga SMOLOVÁ. Praha: Univerzita Karlova, nakladatelství Karolinum, 2018. Myšlení současnosti. ISBN 978-80-246-4056-3.

<sup>16</sup> Se zaměřením právě na autonomní systémy a s tímto fenoménem dnes úzce spojené technologie umělé inteligence či strojového učení.

<sup>17</sup> TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*. Praha: Togga, 2014. Scholia (Togga). ISBN 978-80-7476-043-3.

<sup>18</sup> TURING, Alan. *Education, Machine & Life – Biography* [cit. 28. 11. 2019]. Dostupné z: <https://www.biography.com/scientist/alan-turing>.

s počítačem.<sup>19</sup> Turing argumentoval, že jakmile by bylo pro člověka nemožné spolehlivě rozlišovat mezi lidskými a strojovými interakcemi, dotyčný počítač by prošel Turingovým testem. Jinými slovy, Turing ve své teorii vyvozoval, že když se stroje dostanou na komunikační úroveň nerozeznatelnou od lidí, budeme moci konstatovat, že tyto stroje jsou „uměle“ inteligentní.<sup>20 21</sup>

Turingův test byl podroben mnoha kritickým studiím, které vyvracely hypotézu v něm obsaženou. Nejznámějším a nejcitovanějším je argument čínského pokoje, poprvé publikován v roce 1980 americkým filozofem Johnem Searlem. Searle argumentuje, že počítače pouze používají syntaktická pravidla za účelem manipulace s řetězci symbolů, ale samy o sobě nerozumí významu nebo sémantice výsledku svého jednání. Širší závěr argumentu ukazuje, že teorie, že lidská mysl je ve své podstatě podobná počítačovým systémům (systémům zpracování informací), je vyvrácena. Mysl člověka vychází z biologických procesů uvnitř lidského těla, počítače mohou v nejlepším případě tyto biologické procesy pouze simulovat.<sup>22</sup>

Vznik pojmu umělá inteligence je blíže spjat s událostí roku 1956, kdy John McCarthy pozval některé z předních výzkumníků z oblastí jazykové simulace, neuronových sítí, sensoriky, učících se strojů a mnoha dalších na konferenci a provedení výzkumné studie do Dartmouth v New Hampshire. Účastníci měli diskutovat o tématu, pro které vzhledem ke stupni inovativnosti a z důvodu odlišnosti od dosavadního stupně vědeckého poznání vznikl nový termín *umělá inteligence*, deklarovaný v návrhu na provedení výzkumného projektu

---

<sup>19</sup> GUERRA-PUJOL, F. E., The Turing Test and the Legal Process (December 31, 2011) [cit. 28. 11. 2019]. Information & Communications Technology Law, vol. 21, no. 2 (June 2012), pp. 113-126.

<sup>20</sup> MOOR, James. The Turing test: the elusive standard of artificial intelligence. Boston: Kluwer Academic Publishers, c2003. ISBN 1402012047.

<sup>21</sup> TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*. Praha: Togga, 2014. Scholia (Togga). ISBN 978-80-7476-043-3.

<sup>22</sup> LEM, Stanisław a WEIGEL, Pavel. *Tajemství čínského pokoje*. Praha: Mladá fronta, 1999. Kolumbus. ISBN 80-204-0826-6.

s cílem vytvoření studie o umělé inteligenci.<sup>23</sup> Studie měla vycházet z předpokladu, že každý aspekt učení nebo jakýkoli jiný projev inteligence může být v zásadě tak přesně popsán, že bude možné vytvořit stroj schopný jeho simulace. Konference v Dartmouth byla největším shromážděním na uvedené téma, které se do té doby konalo, a položila základy pro budoucí výzkum a vývoj v oblasti strojírenství, matematiky, informatiky, psychologie i řady dalších oborů. Mnoho účastníků bylo přesvědčeno, že pokrok v rychlosti, kapacitě a programování softwaru povede k tomu, že počítače budou schopny jednat stejně inteligentně jako lidské bytosti. Jedinou otázkou zůstalo, kdy a jak by se to mělo stát.<sup>24</sup>

## 2.2 Definice autonomních systémů, robotů a umělé inteligence

Pro účely této práce lze pod pojmem autonomní systém či autonomní robot rozumět zjednodušeně takový počítačový systém, který je schopen samostatného jednání, ve smyslu jednání bez soustavné interakce člověka. Dle veřejných zdrojů je všeobecně užívanou definicí následující definice popisující autonomního robota pomocí úrovně autonomie při plnění úkolů: „*an autonomous robot is a robot that performs behaviors or tasks with a high degree of autonomy. Autonomous robotics is usually considered to be a subfield of artificial intelligence, robotics, and information engineering.*“<sup>25</sup>

Pojmy „autonomní systém“ a „umělá inteligence“ jsou spolu úzce spjaty, jelikož většina autonomních systémů nemůže plnit své úkoly bez projevů umělé inteligence. Konkrétně v případě autonomních vozidel tak naznačil například David Tennenhouse, ředitel

---

<sup>23</sup> A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence [cit. 20. 12. 2019]. Dostupné z: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

<sup>24</sup> KAPLAN, Jerry. *Artificial Intelligence, What Everyone Needs to Know*. Oxford: Oxford University Press, 2016. ISBN 9780190602383.

<sup>25</sup> Autonomous robot - Wikipedia. [online] [cit. 19. 5. 2019]. Dostupné z: [https://en.wikipedia.org/wiki/Autonomous\\_robot](https://en.wikipedia.org/wiki/Autonomous_robot).

výzkumu ve VMware<sup>26</sup>, ve svém článku pro MIT Technology Review: „*Eventually, all self-driving cars will employ some combination of sensors, cameras, radar, high-performance GPS, Light Detection and Ranging (LIDAR), artificial intelligence (AI), and machine learning to achieve their respective levels of autonomy.*“<sup>27</sup>

### 2.2.1 Definice umělé inteligence

Co se tedy rozumí umělou inteligencí? To je příliš jednoduše znějící otázka s příliš složitou odpovědí. Existuje mnoho definic umělé inteligence, většina z nich operuje s konceptem počítačových programů či strojů schopných jednání, které lze nazvat „intelligentním“ v případě, že by je činil člověk (viz Turingův test popsáný v této kapitole). Tento přístup však nemusí být příliš vhodný, vezme-li se v potaz, jak složité je definovat inteligenci člověka. Jsme totiž zvyklí redukovat popis charakteristik některých pojmů či jevů na vlastnosti, které jsou snadno srovnatelné a změřitelné s námi již dobře známými pojmy či jevy. Avšak v tomto případě takové srovnání nefunguje, jelikož pojem inteligence není snadno představitelný a je složité jej objektivizovat či prakticky uchopit.

Složitost vymezení „inteligence“ lze ukázat na následujícím názorném příkladu – klasické, všem známé hry „piškvorky“, kde hráči umísťují křížky a kolečka do ohraničené tabulky o 3x3 čtverečích. Tato hra má přesně 255.168 možných zakončení. Dnes je již jednoduché naprogramovat počítačový program tak, aby uměl analyzovat všechna možná zakončení a v průběhu hry vybíral ty možnosti, které vedou ke kýženému cíli. Je však takový počítačový program nadán umělou inteligencí? Lze se domnívat, že většina lidí by odpověděla negativně, a to zejména s odkazem na trivialitu postupu takového programu. Na

---

<sup>26</sup> Společnost VMware, Inc. je vývojářem proprietárního virtualizačního softwaru pro počítače.

<sup>27</sup> TENNENHOUSE, David. *Autonomous Vehicles: Are You Ready for the New Ride?*. MIT Technology Review, listopad 2017 [cit. 28. 11. 2019]. Dostupné z: <https://www.technologyreview.com/s/609450/autonomous-vehicles-are-you-ready-for-the-new-ride/>.

druhou stranu program, který by hrál na tom principu, že by pozoroval chování jiných hráčů (lidí), učil se, analyzoval nejen možná zakončení, která znamenají výhru, ale i různé taktiky, bychom s nejvyšší pravděpodobností za nadaný umělou inteligenci označili, byť výsledek jednání může být v obou případech stejný. Subjektivním důvodem takového posouzení by byla právě schopnost programu učit se a aplikovat poznání na svá budoucí jednání.

Problémem je, že většina úkolů či zadání nemá přesně vypočitatelný či určitelný počet možných výsledků, jak je tomu v uvedeném příkladu. Většina problémů také nemusí mít nutně řešení formou analýzy možných výsledků, nýbrž vyžadují určitou míru intuice a kreativity. Lze se domnívat, že zde pramení překážka ve srovnání mezi umělou a lidskou inteligencí. U umělé inteligence přirozeně posuzujeme náročnost neboli složitost postupu posuzovaného programu a v případě, že je příliš triviální, odmítáme mu přiznat kvalitu inteligence. V případě lidské inteligence však toto měřítko neaplikujeme, tedy nezkoumáme zvolený postup ani jeho náročnost, zkoumáme pouze výsledky. V tomto ohledu se lze domnívat, že definice umělé inteligence cestou srovnání s inteligencí lidskou je nežádoucí, a to i s ohledem na skutečnost, že počítačové programy jsou často schopny jednání či posouzení, která člověk díky svým omezeným kapacitám nikdy nesvede (např. během několika málo sekund detekovat hrozbu kybernetického útoku prostřednictvím analýzy stahovaných dat).

Oxford Dictionary definuje umělou inteligenci alespoň pomocí úkolů, které „obvykle“ vyžadují lidskou inteligenci s uvedením konkrétních praktických příkladů, a to následovně: *„the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.“*<sup>28</sup>

---

<sup>28</sup> Definition of artificial intelligence in English by Oxford Dictionaries. Oxford Dictionaries [cit. 19. 5. 2019]. Dostupné z: [https://en.oxforddictionaries.com/definition/artificial\\_intelligence](https://en.oxforddictionaries.com/definition/artificial_intelligence).

Evropský hospodářský a sociální výbor se této problematice rovněž věnoval a konstatoval, že jednoznačně přijímaná, přesně vymezená definice umělé inteligence neexistuje. Širokou definici umělé inteligence poté uvedl ve svém stanovisku ze dne 31. května 2017.<sup>29</sup> Ve stanovisku uvádí, že umělá inteligence „je zastřešující pojem pro velké množství (pod)oblastí, jako např.: kognitivní informatika (cognitive computing, tedy algoritmy, které argumentují a chápou na vyšší (lidské) úrovni), strojové učení (machine learning, tedy algoritmy, které si osvojují úkoly), rozšířená inteligence (augmented intelligence, tedy spolupráce mezi člověkem a strojem), robotika v oblasti UI (UI integrovaná do robotů). Ústředním cílem výzkumu a vývoje v oblasti UI je však automatizace inteligentního jednání, jako je argumentace, shromažďování informací, plánování, učení, komunikace, manipulace, vysílání signálů a samostatná tvorba, snění a vnímání.“

Umělou inteligenci lze dle Evropského hospodářského a sociálního výboru dělit zhruba na umělou inteligenci v úzkém slova smyslu a v obecném slova smyslu. Výsledkem vývoje umělé inteligence v úzkém slova smyslu, tedy takové, která je schopná vykonávat konkrétní úkoly a na jejímž poli bylo v minulosti dosaženo velkého pokroku, zejména díky růstu výpočetní výkonnosti (*computer processing power*), dostupnosti velkých objemů dat a vývoji strojového učení, je skutečnost, že systémy umělé inteligence (prostřednictvím algoritmů) již mohou být samoučící, autonomní a adaptabilní. Umělá inteligence v obecném slova smyslu je již schopná vykonat jakýkoli intelektuální úkol, který je schopen vykonat člověk.<sup>30</sup>

---

<sup>29</sup> Bod 2.1 Stanoviska Evropského hospodářského a sociálního výboru k tématu Umělá inteligence – dopady umělé inteligence na jednotný trh (digitální), výrobu, spotřebu, zaměstnanost a společnost (stanovisko z vlastní iniciativy) (2017/C 288/01).

<sup>30</sup> Bod 2.3 Stanoviska Evropského hospodářského a sociálního výboru k tématu Umělá inteligence – dopady umělé inteligence na jednotný trh (digitální), výrobu, spotřebu, zaměstnanost a společnost (stanovisko z vlastní iniciativy) (2017/C 288/01).

Nejaktuálnější definici, která je současně pro Českou republiku relevantní, navrhla Evropská komise, resp. jí zřízená odborná skupina na vysoké úrovni pro umělou inteligenci, v dubnu 2019 a publikovala ji v dokumentu *Definice UI: hlavní schopnosti a vědní obory* určeném pro účely použití v dalších pokynech a dokumentech v oblasti výzkumu umělé inteligence vydaných odbornou skupinou.<sup>31</sup> Definice je poměrně široká a zahrnuje jak umělou inteligenci jako software, tak umělou inteligenci jako vědní obor:

*„Systémy umělé inteligence (UI) jsou softwarové (a případně také hardwarové) systémy navržené lidmi<sup>32</sup>, které mají zadán složitý cíl a jednají ve fyzické nebo digitální dimenzi, přičemž vnímají své prostředí tím, že získávají data, interpretují shromážděná strukturovaná nebo nestruturovaná data, usuzují ze znalostí nebo zpracovávají informace odvozené z těchto dat a rozhodují o nejlepší akci či akcích k dosažení daného cíle. Systémy UI mohou používat symbolická pravidla nebo si vytvořit numerický model; mohou rovněž přizpůsobovat své chování na základě analýzy toho, jak je prostředí ovlivněno jejich předchozími akcemi.“*

*Jako vědní obor zahrnuje UI řadu přístupů a technik, jako je strojové učení (jehož konkrétními příklady jsou hluboké učení a posilované učení), strojové usuzování (což zahrnuje plánování, rozvrhování, reprezentaci znalostí a usuzování z nich, vyhledávání a optimalizaci) a robotika (která zahrnuje řízení, vnímání, senzory a akční členy, jakož i integraci veškerých dalších technik do kyberneticko-fyzických systémů).“*

V případě umělé inteligence jako software se navržená definice zaměřuje na aspekt plnění cílů a tuto inteligenci tak charakterizuje vstupy (daty), jejich analýzou a interpretací

---

<sup>31</sup> Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *Definice UI: hlavní schopnosti a vědní obory*. Duben 2019 [cit. 28. 11. 2019]. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).

<sup>32</sup> Lidé navrhují systémy UI přímo, k optimalizaci jejich návrhu však mohou využít rovněž techniky UI.

a následným rozhodnutím za použití naprogramovaných pravidel a modelů, případně pravidel a modelů přizpůsobených předchozím zkušenostem, tj. předchozím zpracováním vstupů.

Umělá inteligence, jakožto věda či teorie učení, jak je tento pojem často užíván v literatuře či mezi širokou veřejností, však značí v dnešní době fakticky nejzazší projev technologického vývoje ve formě níže popsaného strojového učení, dále strojového usuzování a robotiky, což v souhrnu pokrývá jak plánování, usuzování, optimalizaci, tak jejich spojení s potřebnými senzory a mechanikou.

Odborná skupina zřízená Evropskou komisí pak vedle výše zmíněného dělení umělé inteligence na umělou inteligenci v **úzkém slova smyslu a v obecném slova smyslu**, jak uvádí Evropský hospodářský a sociální výbor, také dělí umělou inteligenci na **úzkou (slabou) a obecnou (silnou)**. Obecný systém umělé inteligence má být dle odborné skupiny systémem, jenž je schopen provádět většinu činností, které dnes běžně vykonávají lidé. Úzké systémy umělé inteligence jsou na druhou stranu systémy, které jsou schopny provádět jeden či několik konkrétních úkolů a jsou příklady v současnosti zavedených systémů.<sup>33</sup> V zásadě se tedy jedná o ekvivalenty, když obě dělení vychází z rozsahu úkolů, které je systém schopen plnit, a v případě obecné (silné) umělé inteligence odkazují na schopnosti člověka.

### 2.2.2 Definice (intelligentního) robota

Dalším často skloňovaným pojmem v tomto kontextu je pojem „robot“. Když se tento pojem použije v literatuře (ať už odborné či umělecké), představíme si něco docela jiného, než pouhý shluk vzorců a algoritmů. Pojem „robot“ použil poprvé Karel Čapek ve svém divadelním dramatu R.U.R (neboli Rosumovi Univerzální Roboti), kde tento pojem představoval umělou bytost stvořenou člověkem, která reaguje na podněty ze svého okolí a současně na toto okolí zpětně působí. Když odmyslíme od umělé bytosti, tj. bytosti člověku

---

<sup>33</sup> Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *Definice UI: hlavní schopnosti a vědní obory*. Duben 2019 [cit. 28. 11. 2019]. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).



podobné, představíme si nejčastěji různé postavy z filmů, např. R2-D2 či Wall-E, které sice nevypadají jako člověk, ale mají podobu stroje, který je sestaven člověkem či jiným robotem z hmotných součástí, a to za účelem plnění určitého úkolu. Co ale některé méně jasné příklady? Vezměme například HAL („*Heuristically Programmed Algorithmic computer*“) 9000, palubní počítač vesmírné lodi Discovery One z fiktivního příběhu Vesmírná odysea, který má za úkol ovládat všechny palubní systémy lodi, komunikuje s posádkou lodi, avšak nemá hmotný substrát, nýbrž pouze podobu červeného oka kamery – je správné označit tento počítač za robota?

Čapková představa a definice robota se za uplynulých sto let významně posunula. Mezinárodní organizace pro standardizaci například definuje robota v normě ISO 8373 jako „*automaticky řízený, opětovně programovatelný, víceúčelový manipulátor pro činnost ve třech nebo více osách, který může být buď upevněn na místě, nebo mobilní k užití v průmyslových automatických aplikacích.*“<sup>34</sup> Definice, která je vhodná pro určitou organizaci, však nemusí být vhodná pro jinou. Americký institut pro robotiku (RIA) tak definuje robota jinak, a to jako „*přeprogramovatelný vícefunkční manipulátor určený k přemísťování materiálu, součástek, nástrojů nebo specializovaných přístrojů pomocí různě naprogramovaných pohybů za účelem provádění různých úkolů.*“<sup>35</sup>

Robotika a umělá inteligence jsou sice dva odlišné vědní obory s jiným zaměřením i jinými cíli. Robotika představuje takové odvětví technologie, které se zabývá fyzickými roboty. Roboti zjednodušeně představují programovatelné stroje, které jsou obvykle schopny provádět řadu akcí autonomně nebo částečně autonomně. Umělá inteligence má naopak v moderním světě mnoho odlišných způsobů využití. Algoritmy umělé inteligence se používají například v rámci vyhledávače Google, profilovacích postupech v internetových

---

<sup>34</sup> VOLNÁ, Eva a KOTYRBA, Martin. *Umělá inteligence*. Ostrava: Ostravská univerzita v Ostravě, 2013. ISBN 978-80-7464-330-9.

<sup>35</sup> Tamtéž.

obchodech, optimalizaci tras prostřednictvím GPS a mnoha dalších. Většina programů umělé inteligence, resp. strojového učení, se nepoužívá k řízení robotů. Existuje však podoblast robotiky i umělé inteligence, kde se tyto dva obory střetávají, a tou oblastí je výzkum inteligentních robotů. Běžný robot<sup>36</sup> může být naprogramován tak, aby zvedl předmět, umístil jej na jiné místo a opakoval tento úkol, dokud nedostane podnět, aby se zastavil. Při doplnění kamer a algoritmu umělé inteligence, resp. strojového učení, může robot objekty sledovat, analyzovat podněty a činit relevantní závěry. V takovém případě lze hovořit o inteligentním robotovi, který bude představovat rovněž širokou oblast výzkumu umělé inteligence.

Jak lze tedy pro účely této práce definovat robota? Ve většině příkladů uvedených výše lze uzavřít, že robot je schopen interakce s okolním světem, při této interakci se chová inteligentně a je stvořen člověkem. Vhodná se proto jeví následující definice: *„Robot je člověkem či jiným robotem vytvořený systém, který interaguje s okolím, plní definované úkoly, ale v biologickém smyslu není živý.“* To znamená, že robot je něco sestrojeného člověkem, příp. jiným robotem, činí racionální rozhodnutí o tom, co dělat, a je to stroj. Je důležité poznamenat, že plnění definovaných úkolů robotem může znamenat i pouhé plnění funkce pozorovací, za předpokladu, že robot naplňuje zbylé definiční znaky.

### 2.2.3 Definiční znaky autonomních systémů

Před tím, než může nastoupit jakákoliv regulace, je třeba důkladně posoudit veškeré rysy autonomních systémů, případně umělé inteligence. Výše je uvedena definice umělé inteligence, jak ji nabízí Oxford Dictionary, Mezinárodní organizace pro standardizaci, Americký institut pro robotiku nebo Evropská komise. Všeobecně uznávaná a užívaná definice však zatím neexistuje, a nejen v různých státech, avšak i mezi různými autory se pojetí v dílčích aspektech liší. Proto je potřeba některou vybrat a při následné analýze z ní vycházet. Pro Českou republiku by pro účely budoucí regulace, vzhledem k aktivitě orgánů Evropské unie, zejména právě i Evropského parlamentu, a snaze sjednotit, případně alespoň

---

<sup>36</sup> Ve smyslu kolaborativního robota (cobot), který je výsledkem výzkumu čistě v oblasti robotiky.

přiblížit výzkum, vývoj i legislativu v této oblasti v rámci celé Evropské unie, mohla být relevantní definice (nebo spíše definiční znaky) „inteligentního robota“ obsažená v Usnesení Evropského parlamentu o občanskoprávních pravidlech pro robotiku<sup>37</sup>. Evropský parlament v tomto kontextu operuje s definičními znaky, kterými jsou (i) autonomie díky senzorům nebo výměně dat s okolním prostředím (propojenost) a schopnost tato data předávat a analyzovat, (ii) schopnost samostatného učení na základě zkušeností a interakce (volitelné kritérium), (iii) alespoň menší fyzická struktura; schopnost přizpůsobit své jednání a svou činnost okolnímu prostředí, a (iv) skutečnost, že robot není v biologickém smyslu živý. Přesto, že neexistuje závazná a zastřešující definice, mohou typické znaky autonomních systémů posloužit jako výchozí bod úvah o budoucí regulaci tak, aby zahrnovala všechny tyto aspekty a zohlednila je v možných důsledcích takové regulace.

Co tedy znamená, když se řekne, že systém je „autonomní“? V technologickém kontextu inženýři aplikují termín „autonomní“ na počítačem řízené systémy, které jsou schopny učinit důležitou volbu v rámci své vlastní činnosti, a to s drobným nebo žádným zásahem člověka. Autonomní systémy jsou tak schopny řídit svou vlastní činnost tváří v tvář nepředvídatelnému nebo měnícímu se fyzickému nebo datovému prostředí. Jedná se o volby, které by v jiném (neautonomním) režimu činil vždy člověk. Jedním z nejjednodušších příkladů jsou domácí autonomní vysavače. Takový systém je považován za autonomní, jelikož sám robot (prostřednictvím palubního počítače a senzorů) rozhoduje, kam se má pohybovat a jak se má vyhnout jednotlivým překážkám, aniž by byl řízen člověkem. Naopak, tradiční vysavač je neautonomní, protože je to člověk, ne počítač, který ho ručně řídí po místnosti a okolo překážek.

Termín „autonomní“ se však může vztahovat i na nepohyblivé počítačové systémy, jako jsou například algoritmické systémy pro obchodování na finančním trhu, resp. různé

---

<sup>37</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

automatické mechanismy, které ovládají a ovlivňují velkou část finančního trhu. V tradičních systémech obchodování na finančním trhu je to osoba, obchodník, která je odpovědná za rozhodování o tom, jaké finanční nástroje budou nabízeny k nákupu nebo prodeji a za jakou cenu. Naproti tomu v autonomních algoritmičtých systémech finančního obchodování rozhoduje o všem sám počítačový systém.

Jak je z výše uvedeného patrné, zatímco definičních znaků autonomních systémů je několik, jsou to jen některé z nich, které mají za následek náročnost regulace chytrých strojů nadaných umělou inteligencí, jelikož jsou pro společnost relativně nové a jen nesnadno uchopitelné. Nejproblematictější znakem je patrně níže uvedená schopnost jednat autonomně a zní dále vyplývající nepředvídatelnost, která představuje veliký posun od dřívějších technologií. To v praxi vede spíše k přístupu ex post regulace, jakožto reakce na již existující situace, v důsledku nezbytnosti posoudit dopady a všechna (ne)myslitelná rizika, jejichž znalost není často v rané fázi vývoje a výzkumu či rané fáze implementace dostatečná pro účely regulace ex ante, která by umožnila udržet krok s technologickým vývojem.

### **2.2.3.1 Schopnost jednat autonomně**

Zřejmě nejtypičtější vlastností autonomních systémů, která je odlišuje od dřívějších technologií, je právě jejich schopnost jednat autonomně, tedy bez lidské intervence (ta je představována maximálně například uvedením do provozu). Tyto systémy jsou již schopny provádět složité operace, jakými jsou například řízení automobilu nebo letadla. Náročnost a rozsah operací (úkolů), které budou postupně přenechávány chytrým strojům vybaveným umělou inteligencí, bude v nadcházejících letech nepochybně růst. Lze tedy hovořit o nové průmyslové revoluci, jelikož čím dál širší zapojení takových strojů do běžného provozu způsobí dalekosáhlé socio-ekonomické důsledky, stejně jako mechanizace práce v průběhu 18. a 19. století.

### **2.2.3.2 Nepředvídatelnost a schopnost strojového učení**

Dalším znakem autonomních systémů, který bude představovat zajímavou výzvu při budoucí regulaci, je specifický koncept nepředvídatelnosti. Autonomní stroj je v rámci

nastavených parametrů schopen činit rozhodnutí, která nejsou pro člověka předvídatelná, a tím pádem ani člověkem kontrolovatelná. Nejlepším příkladem takových schopností jsou počítačové programy testované na počítačových hrách, které mají naprogramované základní informace o pravidlech hry, jejich následující chování je již podmíněno vlastním učením a vlastní „vůli“ robota.<sup>38</sup> Otázku (ne)předvídatelnosti jednání autonomních systémů analyzoval po uspořádání kulatého stolu s nezávislými odborníky na autonomii, umělou inteligenci a robotiku i Mezinárodní výbor Červeného kříže (ICRC) ve zprávě *Autonomy, artificial intelligence and robotics* ze srpna 2019.<sup>39</sup> Mezinárodní výbor Červeného kříže posuzoval otázku udržení kontroly u autonomních zbraní, kdy akcentoval potřebu udržení dohledu a kontroly formou, kdy počítačový systém vykonává úkoly autonomně, zatímco provozovatel – fyzická osoba na něj dohlíží a může podle potřeby poskytovat pokyny anebo zasáhnout a převzít kontrolu. Mezinárodní výbor Červeného kříže upozorňuje, že obecně platí, že umožnění počítačového systému vykonávat úkoly autonomně při zachování kontroly lidského dohledu vyžaduje detailní znalost toho, jak systém funguje a reaguje, vyžaduje schopnost tzv. „prediktivní kontroly“, aby mohl provozovatel s jistotou posoudit, kdy a v jaké formě je zásah nezbytný.

Nepředvídatelnost jednání je atribut, který v naší společnosti není zvykem. Jsme obklopeni lidmi a stroji řízenými lidmi, které mohou jednat svévolně, ale stále se pohybujeme v předem nastavených, společnostmi uznaných mantinelech. Samozřejmě máme kolem sebe i stroje, které jsou plně řízeny počítačem (např. výtah), ale ty mají ve srovnání s autonomními roboty velice úzký rozsah operací, nejsou nadány schopností autonomie, a proto je nemůžeme navzájem srovnávat. Pro účely regulace takového jednání bude nutné do jisté míry vykročit z rámce dnešní právní úpravy. Úprava deliktivní způsobilosti, trestní odpovědnosti a dalších

---

<sup>38</sup> MORELLE, R. *Google machine learns to master video games*. BBC NEWS, únor 2015 [cit. 28. 11. 2019]. Dostupné z: <http://www.bbc.com/news/science-environment-31623427>.

<sup>39</sup> ICRC, *Autonomy, artificial intelligence and robotics: Technical aspects of human control*. Srpen 2019 [cit. 28. 11. 2019]. Dostupné z: [https://www.icrc.org/sites/default/files/document/file\\_list/autonomy\\_artificial\\_intelligence\\_and\\_robotics.pdf](https://www.icrc.org/sites/default/files/document/file_list/autonomy_artificial_intelligence_and_robotics.pdf).

druhů odpovědnosti dnes vychází z rámce, který předpokládá, že jsme schopni předvídat jednání ostatních a zejména racionálně posoudit jednotlivé motivy a pohnutky. Autonomní systémy jsou příkladem takových technologií, které samy určují, jaké kroky má systém učinit, aniž by došlo k vnějšímu zásahu ze strany člověka, což může vést k utváření řešení, která člověk neočekává. Tato myšlenka mimo jiné vychází ze skutečnosti, že lidská mysl je omezena v tom smyslu, že není v určitý moment schopna analyzovat všechny (často ani většinu) informace, které jsou k dispozici, a většinou usiluje o momentálně nejspokojivější řešení, nikoliv optimálnější řešení.<sup>40</sup>

Jinými slovy lze u některých autonomních systémů hovořit o tzv. strojovém učení. Jedná se o jev, který zkoumají obory zabývající se právě počítačovými programy, které jsou schopny získávat zkušenosti, a tím zlepšovat svou výkonnost v průběhu času. Představa, že počítače se „učí“, je z velké části metaforou a neznamená, že počítačové systémy uměle replikují pokročilé kognitivní funkce, prostřednictvím kterých vnímá člověk svět okolo sebe a které se podílejí na učení člověka. Naopak tyto algoritmy můžeme považovat za učení z funkčního hlediska – počítačové systémy jsou schopny měnit své chování, aby zlepšily svůj výkon v určitém úkolu, a to na základě získaných zkušeností. Obvykle se obdobné algoritmy strojového učení používají k odhalování opakujících se vzorků v datech za účelem automatizace složitých úkolů nebo předpovědí. Tyto algoritmy se dnes používají v řadě komerčních aplikací v reálném světě, např. při analýze výsledků vyhledávání na internetu, při rozpoznávání obličejů, detekci podvodů a dalších. Strojové učení je úzce spjato s pojmem „prediktivní analýzy“, neboť výzkumníci často používají metody strojového učení k analýze stávajících dat, aby předpověděli pravděpodobnost nejistých výsledků. Strojové učení je proto

---

<sup>40</sup> SCHERER, M. U. *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*. Harvard Journal of Law & Technology, jaro 2016.

často úzce spojováno s fenoménem umělé inteligence, protože dobře fungující algoritmus může produkovat automatizované výsledky, které jsou „inteligentní“.<sup>41</sup>

Výše uvedené lze pozorovat například u všem známých „spam filtrů“, na kterých je možné ilustrovat některé základní funkce strojového učení. Spamy jsou nevyžádaná obchodní sdělení, která mohou rušit jejich adresáty. V zásadě by uživatel mohl spravovat nevyžádanou poštu ručně čtením každého e-mailu, zjišťováním, zda je daný e-mail spamem a odstraněním těch, které jsou nevyžádaným obchodním sdělením. Vzhledem k tomu, že tento úkol je pro obyčejného člověka náročný, je žádoucí identifikaci nevyžádané pošty automatizovat. Pro automatické filtrování nevyžádané pošty jsou používány právě programy vybavené algoritmy se schopností strojového učení. Jak algoritmy strojového učení automaticky identifikují nevyžádanou poštu? Tyto algoritmy jsou navrženy tak, aby detekovaly definované vzory mezi zkoumanými údaji. V typickém procesu je algoritmus strojového učení „vyškolen“ tak, aby rozpoznal nevyžádané e-maily tím, že jsou mu poskytnuty známé příklady spamu pro účely vzorové analýzy – například určitá osoba určí, že konkrétní e-mail je nevyžádanou poštou a označí jej jako takový pomocí softwaru pro čtení e-mailů. Při analýze nevyžádané pošty se algoritmus strojového učení pokouší detekovat znaky, které indikují, že daný e-mail je pravděpodobně nevyžádaným sdělením. Po analýze několika takových příkladů může algoritmus zjistit vzorec a vytvořit obecné pravidlo – například takové, že e-maily s konkrétním slovním spojením v předmětu zprávy jsou mnohem pravděpodobněji nevyžádanými než vyžádanými e-maily, naopak e-maily od osoby, které uživatel již dříve odepisoval, spíše nebudou nevyžádaným sdělením. Poté může takový vzorec aplikovat na budoucí situace a použít k automatickému posouzení pravděpodobnosti, zda nově příchozí e-mail je nebo není spamem. Důležité je, že algoritmy strojového učení jsou navrženy tak, aby zlepšovaly svou výkonnost v průběhu času, a to v důsledku analýzy stále většího množství dat.

---

<sup>41</sup> MORGAN, C. (2019). *Responsible AI: a global policy framework*. McLean, VA, USA: ITechLaw. ISBN 978-1-7339931-0-4.

Již současný svět je plný chytrých strojů a robotů, které do jisté míry jednají autonomně, ale jsme zvyklí, že nad vším, čemu zatím propůjčíme dočasně výkon určitých funkcí, máme stále určitou kontrolu a jsme schopni předpovídat chování takového zařízení. S příchodem strojů na vyšší úrovni autonomie se naše myšlení v tomto ohledu bude muset posunout, jelikož jednání softwaru, který je schopen se sám učit, již nezávisí pouze na vtěleném algoritmu, ale také na zkušenostech nabytých poté, kdy je stroj přiveden do chodu.

Vzhledem k tomu, že strojové učení je již dnes úspěšně implementováno v řadě složitých oblastí, o kterých dříve platilo, že jsou výlučně oblastí lidské inteligence, vyvstává tato otázka - do jaké míry mohou být tyto techniky aplikovány v rámci platných právních principů? Algoritmy strojového učení dokázaly vytvořit "inteligentní" výsledky v rámci komplexních, abstraktních úkolů, často ne přímo zakomponovaných do základní koncepční podstaty daného programu, ale nepřímo tím, že zjistí pravděpodobnost a vzorce v datech, které vedou k užitečným výsledkům.

### **2.3 Otázka autonomie a jak jí v kontextu autonomních systémů rozumět**

Jak uvádí National Highway Traffic Safety Administration (NHTSA), pod pojmem autonomie zahrnutým v pojmu "autonomní vůz" rozumíme schopnost automobilu operovat samostatně během časového úseku potřebného k vykonání předem určené činnosti, a to za účelem dosažení konkrétně stanoveného cíle, kterým je přeprava z bodu A do bodu B, za současné schopnosti reakce na podněty z okolního prostředí, které jsou zaznamenávány pomocí zabudovaných senzorů (radar, lidar atd.).<sup>42</sup> Pojem autonomie tedy, jak bylo zmíněno výše, značí určitou míru nezávislosti, což odlišuje autonomní systémy od jiných systémů, které operují na základě předem definovaného vzorce.

---

<sup>42</sup> National Highway Traffic Safety Administration: Preliminary Statement of Policy Concerning Automated Vehicles [cit. 12. 5. 2019]. Dostupné z: [https://www.google.cz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjYyaWPIKjaAhUB3qQKHWQoDTUQFggnMAA&url=https%3A%2F%2Fwww.nhtsa.gov%2Fstaticfiles%2Frulemaking%2Fpdf%2FAutomated\\_Vehicles\\_Policy.pdf&usg=AOvVaw1kzgMhDIOOc-olHnCFb4os](https://www.google.cz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjYyaWPIKjaAhUB3qQKHWQoDTUQFggnMAA&url=https%3A%2F%2Fwww.nhtsa.gov%2Fstaticfiles%2Frulemaking%2Fpdf%2FAutomated_Vehicles_Policy.pdf&usg=AOvVaw1kzgMhDIOOc-olHnCFb4os).



Tento pojem se snadno ilustruje na příkladu autonomních vozidel. Konkrétně autonomními automobily rozumíme takové, které umožňují některé (podle úrovně autonomie i většinu či všechny) funkce řízení předat řidičem samotnému automobilu. Právě nejvyšší úroveň autonomie předpokládá, že cestující poskytne pouze cílovou destinaci a dále již nevěnuje pozornost podmínkám na pozemních komunikacích, tj. nemusí kontrolovat průběh jízdy. Takový automobil už pravděpodobně nebude muset být vybaven standardními ovládacími prvky, které známe, jako jsou volant, řadicí páka apod., teoreticky už nebude muset být přítomen ani sám cestující.

Pokud je vozidlo zcela autonomní, provádí všechna důležitá rozhodnutí týkající se řízení - brzdění, kontrola rychlosti, dodržování vzdálenosti mezi vozidly, volby jízdního pruhu, dodržování pravidel silničního provozu, vyhýbání se překážkám a další. V současnosti zatím nejsou plně autonomní vozy dostupné, tj. nejsou volně přístupné spotřebitelům. Ty, které již existují, fungují jako prototypy pro výzkumné účely a jsou provozovány v umělých podmínkách (na polygonech či na jiných uzavřených okruzích). Je nicméně pravděpodobné, vzhledem k výhodám například z hlediska bezpečnosti, které výrobci slibují, že se plně autonomní vozidla stanou pro spotřebitele čím dál atraktivnějšími.

Plně autonomní vozidlo lze již podřadit pod skupinu „robotů“, míněno inteligentní roboti, které definuje Evropský parlament ve výše zmíněném usnesení. Můžeme na ně tedy aplikovat vše, co bylo uvedeno výše, zejména pasáže věnované strojovému učení a konceptu nepředvídatelnosti. Potencionální provoz autonomních vozidel také otevírá škálu nezodpovězených otázek, se kterými se pro účely případné právní regulace bude nutné vypořádat. Lze se domnívat, že na rozdíl od jiných typů inteligentních robotů je právě v případě samořídících automobilů otázka regulace velice důležitá, zejména s ohledem na rozsah využití dopravních prostředků ve společnosti, a skutečnost, že stále méně domácností se ve svém každodenním obejde bez automobilu. Jelikož postupem času, jak je přirozené, autonomní prvky převáží nad manuálními a spotřebitelé budou chtít provozovat svá vozidla s jistou mírou právní jistoty.

Jelikož je zatím nemožné vtělit do autonomního systému vzorce pro všechny možné situace, nemůžeme s jistotou spoléhat na “bezchybné” plnění jeho funkcí. V únoru roku 2016 došlo k nehodě autonomního vozidla společnosti Google, které se střetlo s autobusem.<sup>43</sup> Jednalo se o první nehodu, která byla způsobena automobilem v plně autonomním režimu. Autonomní vozy již před tím havarovaly, ale žádná z těchto nehod nebyla dle zveřejněných zpráv způsobena autonomním systémem. Předchozí nehody byly způsobeny převážně řidiči jiných automobilů. V tomto případě je na místě zmínit prohlášení mluvčího společnosti Google: *„This is a classic example of the negotiation that’s a normal part of driving -- we’re all trying to predict each other’s movements. In this case, we clearly bear some responsibility, because if our car hadn’t moved there wouldn’t have been a collision. (...) Our test driver, who had been watching the bus in the mirror, also expected the bus to slow or stop. And we can imagine the bus driver assumed we were going to stay put. Unfortunately, all these assumptions led us to the same spot in the lane at the same time. This type of misunderstanding happens between human drivers on the road every day.“*<sup>44</sup> Prohlášení zajímavým způsobem srovnává reakce řidičů a autonomních vozidel a mluvčího společnosti Google dochází k relativně překvapivému závěru, že v případě, kdy by ke stejné srážce došlo i při řízení neautonomního vozidla, nelze nehodu přičítat autonomnímu vozidlu, aniž by se detailněji zabýval otázkou, zda byla nehoda důsledkem pochybení či nikoliv.

Jedná se o případ, kdy vozidlo v plně autonomním režimu reagovalo chybně. Nicméně dle zprávy, s ohledem na způsob jeho naprogramování, nebyl software dokonalý a vozidlo reagovalo tak, jak bylo naprogramované. Rozhodující roli tak sehrála v této nehodě neschopnost předpovědět budoucí chování. Řidič autobusu nedokázal předpovědět budoucí chování automobilu a automobil nebyl schopen předpovědět chování řidiče.

---

<sup>43</sup> LEE, Dave. *Google self-driving car hits a bus*. BBC NEWS, únor 2016 [cit. 12. 5. 2019]. Dostupné z: <http://www.bbc.com/news/technology-35692845>.

<sup>44</sup> *Google Self-Driving Car Project Monthly Report*. GOOGLE, únor 2016 [cit. 12. 5. 2019]. Dostupné z: <http://dl.icdst.org/pdfs/files1/6786717558639e653aae406005326fb3.pdf>.

V souvislosti se schopností samostatného učení inteligentních robotů lze předpokládat, že vůz by se z této zkušenosti poučil a v budoucnu by se zachoval odlišně. Problém však je, že je pravděpodobné, že řidič – fyzická osoba by v některých situacích, které by vyústily v nehodu, zareagoval stejným způsobem, jak mimo jiné vyplývá i ze zmíněné zprávy, společnost však mnohem shovívavěji vnímá lidské pochybení než pochybení stroje. Tato “nedůvěra” může v budoucnu do značné míry brzdit pokrok samořídících automobilů, přesto, že to objektivně neznamena, že provoz autonomních vozů je méně bezpečný. Do té doby, než bude provoz na pozemních komunikacích plně autonomní, což nakonec nemusí ani nastat, nebude snadné určit, jaké jednání je třeba přikládat za vinu právě autonomnímu systému.

## 2.4 Kontrola nad autonomními systémy

Usnesení Evropského parlamentu zdůrazňuje, že *„další rozvoj a širší využívání automatizovaných a algoritmických rozhodovacích procesů má nezpochybnitelný dopad na to, k čemu se nakonec přikloní soukromé osoby (například podniky či uživatelé internetu) a správní, soudní či jiné veřejné orgány v rámci svého rozhodování v záležitostech spotřebitelského, obchodního či úředního rázu; vzhledem k tomu, že nedílnou součástí procesu automatizovaného a algoritmického rozhodování musí být určité ochranné prvky a možnost lidské kontroly a lidského ověřování.“*<sup>45</sup>

Evropský parlament tak zdůrazňuje, že vedle výše uvedených faktorů, jako je například zvýšená bezpečnost, bude v budoucnu jak pro spotřebitelské uvažování, tak pro orgány veřejné moci rozhodujícím faktorem i míra udržitelné kontroly, případně schopnost zpětného ověření jednání inteligentních robotů.

Rizika spojená se stále širší implementací autonomních systémů do běžného chodu společnosti nejsou spojena pouze se schopností autonomního, nepředvídatelného jednání, ale

---

<sup>45</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

i s postupnou ztrátou kontroly ze strany výrobce, vlastníka, provozovatele či dalších. Ztráta kontroly může být způsobena nejen výrobními hardwarovými či softwarovými vadami, ale také právě způsobem naprogramování, který umožňuje širokou míru autonomie a schopnosti samostatného učení. V Usnesení Evropského parlamentu se sice uvádí, že „rozvoj robotických technologií by se měl zaměřit na to, aby doplňoval, a nikoli nahrazoval schopnosti člověka; domnívá se, že ve vývoji robotiky a umělé inteligence je zcela zásadní, aby bylo zaručeno, že lidé budou mít nad inteligentními stroji neustálou kontrolu.“<sup>46</sup> Otázkou je, zda je požadavek neustálé kontroly naplnitelný, případně nakolik je kompatibilní s požadavkem umožnění technického pokroku. Myslím si, že není možné obou těchto cílů plnohodnotně dosáhnout současně. Přestože požadavek kontroly je legitimní a pochopitelný jak z pohledu uživatelského, tak regulatorního, spíše než neustálá kontrola v reálném čase se jako přijatelné řešení jeví nastavení procesů tak, aby bylo zpětně vždy možné rozklíčovat postupy a pochody, které vedly ke konkrétnímu rozhodnutí či jednání. Vedle uvedeného bude rovněž nutné si stanovit hranici, do které umožníme chytrým strojům autonomní jednání (za současného naprogramování základních principů přijatelného chování) a zároveň při jejím překročení zasáhneme a stroj přeprogramujeme či jinak znemožníme pokročilejší autonomní jednání.

Problémem může být skutečnost, že pokud dojde ke ztrátě kontroly, může být obtížné ji získat zpět. Taková situace může představovat potenciální riziko, které převyšuje ve společnosti již známá rizika, která jsou většinou výsledkem lidského chování a společnost má pro takové případy vybudovaný systém represí či podobných opatření. Jak uvádí Matthew U. Scherer, ztráta kontroly může mít dvojí podobu - kontroly lokální a obecné.<sup>47</sup> Ztráta lokální kontroly nastává, když systém již nemůže být kontrolován osobou nebo osobami právně odpovědnými za jeho provoz a dohled nad ním. Ztráta obecné kontroly nastává, když systém již nemůže být kontrolován žádnou osobou. Výstižným příkladem je kontroverzní případ

---

<sup>46</sup> Tamtéž.

<sup>47</sup> SCHERER, M. U. *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*. Harvard Journal of Law & Technology, Volume 29, Number 2 Spring 2016.

chatbotů Facebooku, kteří byli instruováni, aby mezi sebou sjednali barterový obchod (pokusili se vyměnit zboží za zboží, které mělo určitou přidělenou hodnotu). Zajímavý pokus dopadl tak, že chatboti museli být vypnuti, jelikož mezi sebou začali komunikovat jazykem (vytvořili si vlastní komunikační „zkratku“), kterému výzkumníci nerozuměli.<sup>48</sup> Druhá varianta bude logicky představovat větší riziko, které by teoreticky bylo možné eliminovat, pokud se cíle a vtělené principy systému shodovaly s cíli či zásadami široké veřejnosti. To však předpokládá, že jsme schopni správně definovat, co vlastně představují zájmy a zásady široké veřejnosti a následně tyto hodnoty vtělit do systému. Riziko v takovém případě nevyplývá z vadnosti či chybovosti systému umělé inteligence, nýbrž z absence schopnosti pochopit subjektivní záměr za jeho naprogramovanými cíli. Přestože může být snadné naprogramovat požadované cíle, nemusí být možné kalkulovat se všemi možnými proměnnými, a prostředky dosažení tohoto cíle, které systém umělé inteligence zvolí, nemusí být vždy žádoucí.

Možnost, jak eliminovat případy nežádoucího jednání, i když je determinováno okolnostmi, které nastanou po naprogramování a nastartování stroje, je správně nastavený vzorec alespoň základních principů chování, které jsou společné pro všechny autonomní systémy. Při programování vzorců chování mají (pokud případná budoucí regulace tuto možnost neomezí) vývojáři systémů volnost nad výběrem zásad, jimiž se bude muset robot řídit. V tom případě, i přesto, že dojde ke ztrátě kontroly nad konkrétním chováním, bude robot omezen předem stanovenými hodnotami, které se nezmění ani na základě jeho budoucích zkušeností. Taková pravidla by měla tvořit etické minimum, které slouží například jako základ při tvorbě nových právních předpisů (dobré mravy, veřejný pořádek apod.), tj. taková pravidla chování, na nichž je nutné trvat a se kterými se ztotožňuje většina společnosti. K tomu, jaké by tyto hodnoty mohly být, se budu věnovat dále v této práci.

---

<sup>48</sup> GRIFFIN, Andrew. *Facebook's artificial intelligence robots shut down after they start talking to each other in their own language*. The Independent. UK and Worldwide News [cit. 28. 11. 2019]. Dostupné z: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>.

Přestože existenciální rizika a scénáře typu *Já, robot* či *Westworld* zdaleka nepředstavují reálnou hrozbu, problémy kontroly a dohledu skutečně vyvstávají, a to postupně tím, jak se systémy umělé inteligence stávají stále sofistikovanějšími a autonomnějšími. Problémem je, že kontrola v reálném čase v určitých případech ani není možná, Systémy umělé inteligence jsou již dnes schopny samostatně vykonávat úkoly, jako je například sjednávání obchodů s cennými papíry, a to i v řádech nanosekund, což připravuje lidi o schopnost případně zasáhnout v reálném čase. Tzv. „Flash crash“ z roku 2010 ukázal, že příliš rychlé jednání algoritmického systému může mít v krátkém časovém období masivní ekonomický dopad.<sup>49</sup> Přední index amerických akcií se 6. května 2010 během necelých třiceti minut propadl o více než 1000 bodů a tržní hodnota akcií se snížila o téměř jeden bilion dolarů. Několik minut tak trh kolísal v obrovských výkyvech, než v dalších asi 20 minutách téměř celou ztrátu smazal, přičemž celé to způsobil automatizovaný počítačový program, který generoval rozsáhlé prodejní příkazy. Tento a další příklady demonstrují, že kontrola v reálném čase mnohdy není myslitelná, což však neznamená, že vývoj umělé inteligence by měl probíhat bez jakékoliv supervize či bez stanovených pravidel.

Požadavek kontroly sice výslovně z právních předpisů nevyplývá, ale v souladu s prevenční povinností (viz dále) se nabízí alespoň přijetí propracované metodiky či doporučení ať už na úrovni vnitrostátních orgánů či unijních orgánů ve vztahu k pravidlům vývoje umělé inteligence. Evropské orgány například propagují přístup ponechání kontroly nad zásadními rozhodnutími v ruce člověka (human-in-command), případně může být dohled prováděn pomocí řídicích mechanismů, jako je aktivní zapojení lidského faktoru (human-in-the-loop) nebo dohled lidského faktoru (human-on-the-loop).<sup>50</sup> Aby mohly tyto

---

<sup>49</sup> PRATLEY, Nils. *The Trillion-Dollar Questions over the Flash Crash and the Hound of Hounslow*, The Guardian, duben 2015 [cit. 8. 5. 2020]. Dostupné z: <https://www.theguardian.com/business/2015/apr/25/flash-crash-hound-of-hounslow-trillion-dollar-question>.

<sup>50</sup> Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů „Budování důvěry v umělou inteligenci zaměřenou na člověka“ ze dne 8. 4. 2019 (COM/2019/168 final).

mechanismy fungovat a byla zajištěna dostatečná kontrola nad systémy umělé inteligence, je nezbytné, aby existovala neustálá možnost vysvětlit jejich fungování, zajištěna technická spolehlivost a bezpečnost, transparentnost a srozumitelnost algoritmů.

### **3 Pravidla chování autonomních systémů, rozvoj regulatorního rámce a role práva ve věku inteligentních robotů**

Jak bylo již výše zmíněno, postupné rozšíření plně autonomních systémů v různých odvětvích lidské činnosti představuje pro právní řád náročnou výzvu. Rozhodnutí učiněná autonomním systémem mají tendenci být pro lidskou mysl méně intuitivně předvídatelná než srovnatelná rozhodnutí učiněná člověkem. Tento závěr pramení ze skutečnosti, že autonomní systémy nejsou (zatím) nadány emocemi, které jsou charakteristické pro člověka a které určují lidské chování. Lidé mohou předvídat vzájemné chování tím, že spoléhají na vrozené modely lidského chování a na podobný způsob komunikace a sdělování svých záměrů. Z tohoto důvodu je jedním z kontroverzních témat implementace a s ní související regulace autonomních systémů téma etiky (neboli morální hledisko problému). Pokud má být autonomní systém schopný fungovat zcela v autonomním režimu, musí si být schopný poradit v mnoha situacích, i v takových, kde člověk reaguje intuitivně.

Autonomní systém, nebo umělá inteligence, z povahy věci zatím nemůže reagovat intuitivně. Je proto nezbytné, jak bylo naznačeno výše, do něj vtělit soubor základních požadavků (algoritmů), které musí v konkrétních situacích aplikovat. S ohledem na některá odvětví, kde jsou, budou či mohou být inteligentní systémy implementovány, mohou taková rozhodnutí způsobit majetkovou škodu, nemajetkovou újmu nebo i ublížit na zdraví či přivodit smrt, z tohoto důvodu je nutné předem korektně určit pravidla chování, se kterými se ztotožní většina společnosti.

#### **3.1 Právo umělé inteligence, zajištění právní jistoty a předvídatelnosti**

Analýza a úprava regulatorního rámce představuje jednu z klíčových podmínek úspěšného rozvíjení umělé inteligence v České republice. Není překvapující, že Česká republika má zájem na podpoře rozvoje umělé inteligence, a to poskytnutím právní jistoty a odstraněním regulatorních překážek jejího rozvoje a využívání. Z pohledu práva



a s ohledem na univerzální použitelnost umělé inteligence v různých aplikačních oblastech identifikovala Vláda České republiky několik obecných a několik zvláštních právních oblastí, včetně například finančního sektoru, hospodářské soutěže, výzkumu a vývoje umělé inteligence, sociálního zabezpečení, autonomní mobility, či autonomních zbraní (tzv. LAWS), které mohou být rozvojem dotčeny.<sup>51</sup>

Mezi tyto oblasti patří zejména status a právní povaha umělé inteligence, odpovědnost za umělou inteligenci, ochrana soukromí, elektronické komunikace a zpracování neosobních dat, kybernetická bezpečnost, ochrana duševního vlastnictví a další problémy ve zvláštních právních odvětvích.<sup>52 53</sup>

Z uvedených právních oblastí se ve své práci, vzhledem k mému profesnímu i zájmovému zaměření a omezenému rozsahu práce, budu věnovat blíže otázce právní povahy inteligentních strojů a odpovědnosti za jejich jednání.

### 3.1.1 Role práva ve věku inteligentních robotů

Skutečnost, že jednání autonomních systémů může být do jisté míry nepředvídatelné, úzce souvisí s otázkou, co by měl stát udělat pro to, aby povzbudil důvěru lidí v tyto systémy? Jak by měla být regulace nastavena, aby umožnila pokračující vývoj inteligentních robotů a zároveň neodradila spotřebitele či v důsledku příliš přísné regulace dokonce výrobce?

---

<sup>51</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Souhrnná zpráva. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-souhrnna-zprava-2018.pdf>.

<sup>52</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Souhrnná zpráva. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-souhrnna-zprava-2018.pdf>.

<sup>53</sup> Dalšími dotčenými právními odvětvími mohou být dle souhrnné zprávy Vlády České republiky zejména oblast autonomní mobility. V současné době není v České republice možné testovat a provozovat částečně autonomní vozidla s vyšším stupněm autonomie (např. se stupněm autonomie 3 a vyšším). Stávající legislativní rámec rovněž neumožňuje provozování autonomních dronů.

V současnosti sice zatím žádný plně autonomní systém, tj. takový, který by nebyl závislý na lidech, neexistuje. Je však pouze otázkou času, kdy se plně autonomní systémy na trhu objeví.

Dle analýzy Vlády České republiky je pro právo fenomén umělé inteligence zajímavý z toho důvodu, že zasahuje do velkého množství právních odvětví. Autoři analýzy uvádí, že „výzvou jsou pak pro právo zejména systémy vybavené umělou inteligencí, které fungují bez lidské intervence, mění své fungování v průběhu času, stávají se nepředvídatelnými, a k nemožnosti vysvětlit jejich fungování představují netransparentní a zároveň nekontrolovatelné mechanismy, které žijí svým životem ve společnosti, často s ní i komunikují prostřednictvím přirozeného jazyka a výrazně ovlivňují její fungování.“<sup>54</sup> Je patrné, že jedním ze zásadních rizik, které je akcentováno mezi odbornou veřejností nejvíce, je právě otázka vyřešení problému s nedostatkem transparentnosti a kontroly, viz blíže v předchozí kapitole).

Je možné, že problémy spojené s nepředvídatelností by mohly v průběhu času přirozenou cestou samy zaniknout, tj. bez jakéhokoliv zásahu ze strany státu, neboť lidé by si zvykli na způsob chování inteligentních robotů, a toto chování by se tak stalo více předvídatelným. Taková vize je ale spíše nežádoucí a neposkytuje požadovanou míru právní jistoty, jelikož nebudou stanovená pevná pravidla. Jak již Ústavní soud ve své ustálené judikatuře mnohokrát konstatoval, ke znakům právního státu a mezi jeho základní principy neoddělitelně patří také zásada právní jistoty, přičemž její nezbytnou součástí je jak požadavek, aby bylo objektivní právo dlouhodobě stabilní, jednotné a předvídatelné, tak i požadavek na legitimní předvídatelnost postupu orgánů veřejné moci v souladu s právem a zákonem stanovenými požadavky.<sup>55</sup> Ústavní soud stanovil, že „*princip předvídatelnosti práva jakožto důležitý atribut právního státu podstatným způsobem souvisí s principem právní*

---

<sup>54</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Analýza právně-etických aspektů rozvoje umělé inteligence a jejích aplikací v ČR. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: [https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018\\_final.pdf](https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf).

<sup>55</sup> GERLOCH, Aleš. *Teorie práva*. 6., aktualiz. vyd. Plzeň: Vydavatelství a nakladatelství Aleš Čeněk, 2013. ISBN 978-80-7380-454-1.

*jistoty a je nezbytným předpokladem obecné důvěry občanů v právo. Jak již v minulosti Ústavní soud několikrát konstatoval, neopomenutelnou komponentou principu právní jistoty je předvídatelnost práva a legitimní předvídatelnost postupu orgánů veřejné moci v souladu s právem a zákonem stanovenými požadavky.*<sup>56 57</sup> Obecná důvěra občanů v právo může být pak v tomto ohledu ohrožena, pokud nebude předvídatelná, jakým způsobem budou řešeny případné spory vzniklé z interakce lidí s algoritmickými systémy.

Absence konkrétní právní úpravy v některých specifických oblastech pak může vést k nejistotě při využívání některých výrobků. V současné době například stále roste počet výrobců vozidel s autonomním režimem a není jisté, že všichni výrobci by v budoucnu používali stejný vzorec chování a automobily by v důsledku reagovaly ve stejných situacích odlišně od sebe navzájem. V různých oblastech praktického využití algoritmických systémů mohou vznikat výkladové potíže ohledně uplatnitelných právních institutů,<sup>58</sup> a je tedy třeba obzvláště zajistit soulad s vnitrostátními i unijními právními předpisy, včetně pravidel na ochranu základních práv a práv spotřebitelů, zejména pokud jde o systémy umělé inteligence určené pro interakci s lidmi.

Nabízí se jednoduchá otázka – nemůžeme autonomní systémy jednoduše naprogramovat tak, aby dodržovaly právní řád? Odhlédněme od toho, jak jednoduché či složité by bylo samotné naprogramování takového systému, problémem je, že není vždy žádoucí, aby systém jednal v souladu s právními předpisy. Z morálního hlediska nemusí být vždy nejvhodnější právě to legální řešení. Představme si například situaci, kdy nám hrozí nebezpečí, ale inteligentní robot nám nepomůže, protože musí dodržet zákaz vstupu na

---

<sup>56</sup> Nález Ústavního soudu sp. zn. I. ÚS 420/09 ze dne 3. 6. 2009 (N 131/53 SbNU 647).

<sup>57</sup> Srov. nález Ústavního soudu sp. zn. IV. ÚS 690/01 ze dne 27. 3. 2003 (N 45/29 SbNU 417), nález Ústavního soudu sp. zn. II. ÚS 487/03 ze dne 11. 5. 2005 (N 103/37 SbNU 383) či nález Ústavního soudu sp. zn. II. ÚS 2070/07 ze dne 13. 12. 2007 (N 223/47 SbNU 935).

<sup>58</sup> Například právě v otázce právní charakteristiky takového systému či uplatnitelného odpovědnostního režimu.

trávník, nebo nás nedoveze rychle do nemocnice, jelikož dodržuje všechna pravidla provozu na pozemních komunikacích. Problém spočívá v tom, že právní regulace reaguje na standardní situace, stále však spoléhá na lidský rozum, který může v mnoha případech nastavená pravidla překračovat. V jiných případech právní regulace nastavuje určité hranice a člověk své jednání v rámci takových mantinelů přizpůsobuje konkrétním okolnostem právě za využití rozumu. Takové hranice jsou stanoveny například akcentováním právních principů, jejichž nedodržování nebo dokonce jejich neznalost může vést k přílišnému právníckému formalismu, který se projevuje nesprávným lpěním na pouhé litéře předpisu, vytržené ze souvislosti, tj. z podstaty a smyslu práva.<sup>59</sup> My však zatím vycházíme z předpokladu, že inteligentní roboti, byť schopny strojového učení, rozumem nadáni nejsou. Určitým přiblížením by mohlo být naprogramování hierarchie vybraných základních zásad, aby však mohlo být toto řešení účinné, je třeba současně dohlédnout na to, aby byly systémy dostatečně transparentní, jelikož v případě tzv. black box efektu, tj. při neprůhlednosti umělé inteligence, může být následně obtížné určit a prokázat případné porušení právních předpisů, včetně právních ustanovení na ochranu základních lidských práv a svobod, přičíst odpovědnost a splnit případné další podmínky pro uplatnění nároku na náhradu škody.

Nabízí se také regulace sjednocující způsob jednání autonomních systémů. Vzhledem k odlišnostem mezi jednotlivými systémy by mělo dojít ke standardizaci vzorců chování zvláště například pro autonomní vozy, případně autonomní drony, a zvláště pro jiné systémy, přičemž není vyloučeno, že některé základní principy budou společné pro všechny inteligentní roboty. Taková regulace by byla určena zejména výrobcům a programátorům softwaru zabudovaného v inteligentních robotech a zlepšila by předvídatelnost chování autonomních systémů z pohledu jejich provozovatelů, uživatelů i jiných osob. Postupem času by se předvídatelnost chování zvyšovala, jelikož lidé by si zvykli na určitý standardizovaný vzorec chování. Zároveň by nedošlo k omezení vývoje, jelikož výrobci a programátoři by s jistotou znali postup, který by byl v souladu s právními předpisy.

---

<sup>59</sup> OSINA, Petr. *Teorie práva*. Praha: Leges, 2013. Student (Leges). ISBN 978-80-87576-65-6.

Jiným přístupem, založeným na nepřímé regulaci, by bylo stanovení odpovědnosti za pochybení autonomního systému a vzniklou škodu či nemajetkovou újmu. Stát by tímto způsobem působil na výrobce a programátory, aby vyvíjeli nové systémy zodpovědně. Pokud by byl však systém přičítání odpovědnosti nastaven příliš přísně, mohl by vést k nevoli vyvíjet inteligentní roboty, jelikož nikdo by si nepřál nést odpovědnost za jednání robota, který sice postupoval v souladu se svým výchozím nastavením, avšak způsobil škodu. Pokročilí inteligentní roboti využívající umělou inteligenci mohou totiž jednat způsobem, který nebyl při jejich prvotním uvedení v provoz předvídaný. Vysoká úroveň bezpečnosti a účinné mechanismy nápravy dostupné dotčeným osobám v případě vzniku újmy sice pomáhají budovat důvěru uživatelů a zajišťovat obecné přijímání těchto technologií, ale ruku v ruce s pečlivým nastavováním či uplatňováním odpovědnostních mechanismů musí jít i důsledné vynucování bezpečnostních standardů a pravidel tak, aby se co nejlépe eliminovalo riziko skutečně nepředvídatelné újmy. V tomto ohledu se lze domnívat, že přístup založený na vysoké úrovni standardizace a sjednocení principů vedle adaptace odpovědnostních mechanismů je společensky přijatelnější, jelikož by skutečně vedl k větší předvídatelnosti a úprava by byla transparentnější. Příliš přísný mechanismus určení odpovědnosti by například mohl značně odrazovat potenciální kupce autonomních zařízení, jelikož by se nevyhnutelně vystavovali riziku vzniku jejich odpovědnosti za škodu, které přitom nemohli při absenci detailních pravidel pro zajištění bezpečnosti nijak zabránit.

### **3.1.2 Status a povaha umělé inteligence**

Pro to, abychom byli schopni chytré stroje (ať již se jedná o autonomní vůz nebo jiného inteligentního robota) regulovat, je třeba najisto stanovit povahu takového zařízení. Legální zařazení umělé inteligence pod některý z existujících institutů, či případně vytvoření zcela nového konceptu, bude mít zásadní vliv na právní vztahy související s konkrétním

využitím a aplikacemi umělé inteligence. De lege lata, podřazení umělé inteligence pod určitý institut může mít vliv dokonce i na případný odpovědnostní režim.<sup>60</sup>

Ve všech níže uvedených případech však není, dle mého názoru, plně zohledněn úplný potenciál, který v sobě autonomní systémy, včetně umělé inteligence, ukrývají, a dříve nebo později se může takové pojetí stát nedostatečným, jelikož osoby, kterým právní řád přiznává určitou odpovědnost (ve smyslu odpovědnosti za případné porušení stanovených povinností), nebudou chtít stále tuto odpovědnost přebírat za stroje, které budou schopny samostatného jednání, samoučení a jejichž chování již nebude podléhat plné kontrole výrobce, prodejce, vlastníka nebo provozovatele. Z tohoto důvodu se domnívám, že na místě je spíše posouzení situace de lege ferenda.

### 3.1.2.1 Umělá inteligence jako počítačový program

Pokud se budeme řídit platnými právními předpisy, mohli bychom takové stroje považovat za počítačový program. Evropské právo definuje počítačový program pro účely Směrnice o právní ochraně počítačových programů jako „*programy v jakékoliv formě, včetně těch, které jsou součástí technického vybavení (hardware). Tento výraz zahrnuje rovněž přípravné koncepční práce vedoucí k vytvoření počítačového programu za podmínky, že povaha těchto prací v pozdější etapě umožní vytvoření počítačového programu.*“<sup>61</sup> Počítačový program je dle evropské úpravy chráněn tehdy, pokud je původní, v tom smyslu, že je vlastním duševním výtvořem autora. Pro stanovení způsobilosti k ochraně pak není požadováno žádné jiné kritérium. Autorem počítačového programu je prizmatem evropské úpravy fyzická osoba nebo skupina fyzických osob, které daný počítačový program vytvořily,

---

<sup>60</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Analýza právně-etických aspektů rozvoje umělé inteligence a jejích aplikací v ČR. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: [https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018\\_final.pdf](https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf).

<sup>61</sup> Viz recitál 7 Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. dubna 2009 o právní ochraně počítačových programů (kodifikované znění) (Text s významem pro EHP).

nebo, pokud to právní předpisy příslušného členského státu připouští, právnická osoba označovaná těmito právními předpisy za nositele tohoto práva.<sup>62</sup>

V kontextu vnitrostátního práva s termínem počítačového programu pracuje autorský zákon, byť jej blíže nedefinuje.<sup>63</sup> Opět však v případě, pokud je výsledkem tvůrčí duševní činnosti autora (programátora), a to současně za předpokladu, že je „*vyjádřen v jakékoli objektivně vnímatelné podobě včetně podoby elektronické.*“<sup>64 65</sup>

V obou výše zmíněných případech s pojmem legislativa pracuje jako s výsledkem lidské činnosti, se kterým jsou spjata konkrétní práva, která však vždy náleží fyzické osobě, a tak právě hrozí, že v určitý okamžik už takové pojetí nebude s ohledem na charakteristiku autonomních systémů a umělé inteligence dostačující.

### 3.1.2.2 Umělá inteligence jako věc

Věc je v občanském zákoníku definována jako „*vše, co je rozdílné od osoby a slouží potřebě lidí.*“<sup>66</sup> Zákon používá dvě pozitivní definiční kritéria, jejichž prostřednictvím je určen obsah pojmu „věc v právním smyslu“. Prvním kritériem je odlišnost (rozdílnost) od osoby, druhým pak objektivní způsobilost uspokojovat lidské potřeby. Definice věci je tedy

---

<sup>62</sup> Viz článek 2 odst. 1 Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. dubna 2009 o právní ochraně počítačových programů (kodifikované znění) (Text s významem pro EHP).

<sup>63</sup> Viz § 2 zákona č. 121/2000 Sb., zákon o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

<sup>64</sup> Vedle počítačových programů je třeba rozlišovat pojem „programové vybavení počítačů“ (angl. software), který je složen z několika různých prvků (součástí) a tvoří sadu všech programů používaných v počítači, které provádějí nějakou činnost, a kromě programového vybavení je nutno počítat i s „technickým vybavením počítačů“ (hardware), byť autorský zákon tento výraz výslovně nepoužívá. Jedná se o movitou věc s různými součástmi a s příslušenstvím.

<sup>65</sup> TELEČ, I., TŮMA, P. *Autorský zákon, Komentář*. 2. vydání. Praha: C. H. Beck, 2019, 1295 s.

<sup>66</sup> Viz § 489 zákona č. 89/2012 Sb., občanský zákoník, ve znění pozdějších předpisů („občanský zákoník“).

značně široká a umožňuje použít ustanovení o věcech v právním smyslu rovněž na takové nehmotné statky, jako jsou např. energie nebo objekty průmyslového a duševního vlastnictví.

Dle důvodové zprávy se užitečností nerozumí pouze „*faktické přinášení užitku jednotlivým předmětem konkrétnímu vlastníkovi, ale objektivní způsobilost přinášet především hospodářský užitek (čímž se nevylučuje užitek estetický či jiný). Užitečné je, co je prospěšné pro život člověka, a tedy má i hodnotu.*“<sup>67</sup> Samo o sobě by, vzhledem k výše naznačeným oblastem možného uplatnění umělé inteligence, kritérium užitečnosti jednoznačně naplněno bylo. Pro plnohodnotné posouzení je však třeba vzít v úvahu i kritérium ovladatelnosti, jak je uvedeno níže.

Vláda České republiky v tomto ohledu stanoví, že „*věci se dělí na hmotné a nehmotné. Inteligentní robot, tj. systém s fyzickým vyjádřením, jehož součástí je i vestavěný inteligentní software, lze považovat za hmotnou věc.*“<sup>68</sup> Problém však vidím právě v charakteristických vlastnostech inteligentního robota. Jak bylo uvedeno výše, autonomní systém je v rámci nastavených parametrů schopen činit rozhodnutí, která nejsou pro člověka předvídatelná, a tím pádem ani člověkem kontrolovatelná. Z definičního znaku užitečnosti pro potřeby lidí vyplývá implicitně i požadavek objektivní možnosti ovladatelnosti věci.<sup>69</sup> Je tedy otázkou, zda lze jednoznačně stanovit, že i v případě ztráty kontroly nad činností autonomního systému či umělé inteligence je stále naplněno kritérium užitečnosti, resp. ovladatelnosti.

Navíc kritérium užitečnosti pro upokožování lidských potřeb úzce souvisí s kritériem odlišnosti od člověka. Souvislost spočívá v tom, že zatímco subjekt práva (osoba) nemůže

---

<sup>67</sup> Důvodová zpráva k zákonu č. 89/2012 Sb., občanský zákoník (konsolidovaná verze). Dostupné z: <http://obcanskyzakonik.justice.cz/images/pdf/Duvodova-zprava-NOZ-konsolidovana-verze.pdf>.

<sup>68</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Analýza právně-etických aspektů rozvoje umělé inteligence a jejích aplikací v ČR. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: [https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018\\_final.pdf](https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf).

<sup>69</sup> Srov. § 497 občanského zákoníku.



obecně sloužit jako prostředek pro uspokojování potřeb jiného subjektu, věc v právním smyslu ano.<sup>70</sup> Z hlediska právně filosofického je tedy věc prostředkem pro dosahování cílů subjektů práva, tj. plnění účelu utilitárního.<sup>71</sup> Tento znak, společně s požadavkem odlišnosti od osoby, je v kontextu inteligentních robotů (byť tomu tak nemusí být v současné chvíli) taktéž značně problematický. Pojmový znak „autonomie“ se původně vztahuje pouze k lidským schopnostem a zahrnuje svobodu stanovovat si vlastní cíle a zájmy. Kognitivní procesy, které tuto schopnost umožňují, jsou sice úzce spjaty s lidskou bytostí, avšak neuronové sítě, které umožňují strojové učení, mají rovněž za důsledek, že autonomní systémy, zde konkrétně umělá inteligence je schopna do jisté míry řídit svou vlastní činnost. Autonomii v biologicky relevantním slova smyslu lze tedy připsat pouze lidem, avšak pokročilé komplexní adaptivní autonomní systémy, jak bylo výše blíže popsáno, jsou ve vědeckém světě a do značné míry i v rámci veřejného mínění vnímány jako vysoce nezávislé na lidech. Tato analýza je však v současné chvíli ve značné míře hypotetická. Lze tedy uzavřít, byť s jistými otevřenými otázkami do budoucna, že autonomní systém nadaný umělou inteligencí je možné označit za věc ve smyslu § 489 občanského zákoníku.

Pokud vedeme úvahu o možném podřazení umělé inteligence pod pojem „věc v právním smyslu“, nabízí se rovněž úvaha nad kategorizací jakožto výrobku, kdy tento má v právním řádu konkrétnější postavení se specifitějšími dopady. Vnitrostátní právní úprava pojem výrobku legálně nedefinuje, občanský zákoník však transponuje směrnici Rady 85/374/EHS o sblížování právních a správních předpisů členských států týkajících se odpovědnosti za vadné výrobky, která za výrobek považuje „*všechny movité věci kromě prvotních zemědělských produktů a produktů lovu, také tehdy, jsou-li zabudovány do jiné*

---

<sup>70</sup> LAVICKÝ, P. a kol.: *Občanský zákoník I. Obecná část (§ 1–654)*. Komentář. 1. vydání, Praha: C. H. Beck, 2014, 2400 s.

<sup>71</sup> Srov. rozsudek Nejvyššího soudu ze dne 14. 12. 2010, sp. zn. 28 Cdo 537/2010: „*Aby bylo možno považovat určitý objekt za věc v právním smyslu, musí se předně jednat o ovladatelný hmotný předmět (res corporalis). Nesmí zde však chybět ani další významný element, kterým je objektivní užitečnost takového předmětu a jeho schopnost určitým způsobem uspokojovat lidské potřeby.*“

*movité věci nebo do nemovitosti.*“<sup>72</sup> Tento závěr by mohl významně ovlivnit regulaci odpovědnosti za jednání umělé inteligence, resp. by ovlivnil možnost aplikace ustanovení upravujících odpovědnost za vadu výrobku (viz blíže v následujících kapitolách).

### 3.1.3 Umělá inteligence jako osoba v právu

Z výše zmíněných důvodů se nabízí otázka, zda může právo přiznat chytrým strojům postavení plnohodnotné osoby v právním řádu, tj. zda může dojít k subjektivizaci „umělých“ bytostí, neboli přidělení právní osobnosti. „*Právní osobnost je způsobilost mít v mezích právního řádu práva a povinnosti.*“<sup>73</sup> Jedná se o status subjektu práva, který již není pouhým objektem. Pokud předpokládáme, že inteligentní stroje přestanou být nástrojem kontrolovaným a řízeným člověkem, ale začnou být plnohodnotnou, aktivní součástí společnosti, budou schopny činit samostatná rozhodnutí a učit se z vlastních jednání, je potřeba takovým strojům uložit určité povinnosti, případně i práva, a stanovit tak mantinely jednání stejně, jako je tomu u fyzických a právnických osob. Ač se taková myšlenka může zdát absurdní, pravdou je, že přiznáváme práva a ukládáme povinnosti právnickým osobám, které taktéž nejsou osobami přirozenými. Pojem „osoby“ již není ztotožňován pouze s člověkem, ale stále více se subjektem práva ve smyslu nositele subjektivních práv.<sup>74</sup> Pro účely regulace a přiznání subjektivních práv už není určující osoba, nýbrž subjektivní práva jsou určující pro stanovení, kdo je osobou. Právnická osoba je uměle vytvořený konstrukt, má s inteligentním robotem společný rys, kterým je tvůrce - člověk.

Pokud předpokládáme, že inteligentní roboti jsou schopni lidského uvažování a projevů vůle, je na místě posoudit, do jaké míry je řeč o věci a do jaké míry o entitě, která se

---

<sup>72</sup> Směrnice Rady 85/374/EHS o sblížení právních a správních předpisů členských států týkajících se odpovědnosti za vadné výrobky, ze dne 25. července 1985.

<sup>73</sup> Viz § 15 občanského zákoníku.

<sup>74</sup> BERAN, Karel. *Pojem osoby v právu: (osoba, morální osoba, právnická osoba)*. Praha: Leges, 2012. Teoretik. ISBN 978-80-87576-06-9.

blíží člověku. Jak bylo uvedeno výše, občanský zákoník definuje věc jako “*vše, co je rozdílné od osoby a slouží potřebě lidí.*”<sup>75</sup> Toto pojetí není pro pokročilé úrovně inteligentních robotů zcela dostačující. Robot sice do jisté míry slouží potřebě lidí, nemusí tomu tak být ale vždy. Skutečnost, že robot není biologický tvor, neznamená, že v určitých situacích nemůže učinit jednání, které je bližší projevu vlastní vůle než strojovému chování, a sám řídit vlastní jednání. Rozhodujícím kritériem by, dle mého názoru, mělo být to, zda je inteligentní robot schopen rozpoznat svá práva a povinnosti, tj. zda je schopen se pohybovat v mezích zákonů a uvědomit si své postavení v systému práva. V současné chvíli takový robot nemusí existovat, právní úprava by však měla počítat s rychlým vývojem a zejména s trendy, které v oblasti technologií převažují.

Je bezesporu, že do doby, než budeme schopni plně uchopit potenciál umělé inteligence, musíme s inteligentními roboty nadanými umělou inteligencí, pracovat jakožto s věcí. Toto pojetí se ale dříve nebo později stane neúplným, jelikož vývojáři stále více pracují na zdokonalování umělé inteligence a jejímu připodobňování lidem inkorporováním atributů, které byly doposud charakteristické právě jen pro člověka. V úvahu také přichází možnost alespoň dočasně garantovat inteligentním robotům jistou míru ochrany. Jistý příměr by se dal nalézt u zvířat, kterým zákon poskytuje ochranu, i přesto, že nejsou osobou v právu.

Argumentem proti přiznání právní osobnosti inteligentním robotům je skutečnost, že byly vytvořeny jako zařízení, jehož cílem je usnadnit život člověka. Ve své podstatě jsou pouze vyšší formou jakéhokoli jiného nástroje, který slouží k potřebě lidí. Dalším argumentem může být absence "zdravého rozumu" nebo "přirozeného úsudku”.<sup>76</sup>

---

<sup>75</sup> § 489 občanského zákoníku.

<sup>76</sup> KRAUSOVÁ, Alžběta. “*Legal Regulation of Artificial Beings*”, Masaryk University Journal of Law and Technology, Brno: Masarykova Univerzita, 2007, roč. 1, č. 1, s. 187-198. ISSN 1802-5943.

Pokud se však podíváme na praxi v zahraničí, můžeme identifikovat případy zvláštního zřetele hodné, které již dnes překračují hranice představitelného přístupu k umělé inteligenci.

Rijád v roce 2017 oznámil, že Sofii, inteligentnímu robotovi vyrobenému hongkongskou společností Hanson Robotics, veřejně představenému v Rijádu na summitu Future Investment Initiative, který byl na summitu prezentován jako žena, bylo uděleno občanství Saúdské Arábie.<sup>77</sup> Vyjma skutečnosti, že tento krok je v rozporu s řadou místních zákonů, které stanoví pravidla chování subjektů právních vztahů, pravidla získání občanství či práva a omezení postavení žen v Saúdské Arábii, jedná se o krok značně kontroverzní a vzhledem ke skutečnosti, že se tak stalo v roce 2017, i pokrokový.

Japonsko poskytlo v roce 2017 povolení k pobytu chatbotu Shibuya Mirai, který byl naprogramován jako sedmiletý chlapec.<sup>78</sup> Tento krok byl však, obdobně jako v případě Saúdské Arábie, v rozporu s lokálními právními předpisy upravujícími řízení o povolení k pobytu v Japonsku.

S ohledem na výše uvedené precedenty<sup>79</sup> v otázce postavení umělé inteligence, které je prakticky jako rovné lidským bytostem v Saúdské Arábii a Japonsku, je rovněž zajímavá procesní stránka obou případů, a to protože (i) ani chatbot Shibuya Mirai ani robot Sofia sami nepožádali o občanství či povolení k pobytu, a (ii) chatbot Shibuya Mirai ani robot Sofia

---

<sup>77</sup> Saúdská Arábie dala občanství robotovi Sophii. Ženy mají méně práv než ona. Echo24.cz - Názorový deník [cit. 11. 1. 2020]. Dostupné z: <https://www.echo24.cz/a/pCHVV/saudska-arabie-dala-obcanstvi-robotovi-sophii-zeny-maji-mene-prav-nez-ona>.

<sup>78</sup> CUTHBERTSON, Anthony. *Tokyo: Artificial Intelligence 'Boy' Shibuya Mirai Becomes World's First AI Bot to Be Granted Residency*. Newsweek - News, Analysis, Politics, Business, Technology [cit. 11. 01. 2020] Dostupné z: <https://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382>.

<sup>79</sup> S ohledem na zásadu legitimního očekávání se lze domnívat, že uvedené kroky mohou založit precedent a inteligentní roboti, kteří naplní stejná kritéria jako roboti v citovaných případech, by měli mít nárok na stejné rozhodnutí.

nesplnili všechna zákonem požadovaná kritéria (věkovou hranici, požadavek pobytu, znalost jazyka apod.).<sup>80</sup>

Vzhledem k výše uvedeným formálním kritériím měla příslušná správní řízení chatbota Shibuya Mirai i robota Sofia skončit zamítavým rozhodnutím. Současně, spolu s povinnostmi získává úspěšný žadatel o občanství či trvalý pobyt také určitá práva, což může být v uvedeném kontextu značně problematické. V žádném z uvedených případů však tento problém blíže řešen nebyl.

### 3.1.4 Doporučení

Z výše uvedeného vyplývá, že v kontextu současné právní úpravy lze na umělou inteligenci pohlížet optikou různých právních režimů a přistupovat k ní jako k různým právním institutům, které mají následně i vliv na aplikaci konkrétního odpovědnostního režimu. V kontextu autorského zákona lze umělou inteligenci považovat za počítačový program a v kontextu občanského zákoníku za věc v právním smyslu, případně také za výrobek. Byť s dalším rozvojem nových technologií mohou argumenty podporující toto posouzení oslabovat, prozatím se jeví jako nejvhodnější a je s nimi možné dále pracovat.

Co se týče akceptace umělé inteligence jako osoby v právu a přiznání postavení plnohodnotné osoby v právním řádu, tento postoj se zatím jeví pouze jako teoretický a v kontextu současného stavu techniky i jako velmi vzdálený. Z praktického hlediska lze tedy jako nejvhodnější a nejtransparentnější doporučit variantu zavedení definice umělé inteligence s ohledem na její výše uvedené zvláštní charakteristické vlastnosti. Důvodem, proč se legislativci jak na vnitrostátní, tak na unijní úrovni tohoto postupu zatím v praxi zdráhají, může být skutečnost, že takový postup by, vzhledem k tomu, že by se promítl i do velkého

---

<sup>80</sup> ATABEKOV, A., YASTREBOV, O. *Legal Status of Artificial Intelligence Across Countries: Legislation on the Move*. European Research Studies Journal Volume XXI, vydání 4, 773 – 782 [cit. 11. 01. 2020]. Dostupné z: <https://www.ersj.eu/journal/1245>.

množství souvisejících právních oblastí, jednoznačně představoval vekou koncepční změnu stávajícího právního řádu.

## 4 Odpovědnost za jednání autonomních systémů

### 4.1 Problém s odpovědností u autonomních systémů

V souvislosti s vývojem autonomních systémů podporujících technologie umělé inteligence, resp. strojového učení, stále více roste obava jak ve veřejnosti a ve sdělovacích prostředcích, tak i mezi legislativci a vývojáři.<sup>81</sup> Takové obavy historicky vznikaly v případě mnoha jiných, předcházejících technologií,<sup>82</sup> ale u této další vlny technologického pokroku a změn je něco atypického. I když dopady široké implementace jakékoli nové technologie jsou v jistém smyslu nejisté a mohou vést k mnoha a různým nezamýšleným následkům, zdá se, že v případě autonomních technologií je z pohledu legislativců a vývojářů něco zvláště znepokojujícího a rizikového. Lze konstatovat, že toto zvláštní znepokojení pramení právě ze skutečnosti, že autonomní systémy budou nejen nepředvídatelné, pokud jde o jejich nenaprogramované, ale naučené jednání, ale že mohou jednat i mimo kontrolu člověka, jak bylo popsáno v předchozích kapitolách, a to v tom smyslu, že k těmto jednáním může docházet i bez ohledu na existenci či rozsah lidské odpovědnosti.

Provoz autonomních systémů s sebou nutně přináší riziko vzniku škody na právních hodnotách chráněných zákonem. Přiměřenost a komplexnost režimů odpovědnosti vzhledem k novým technologickým výzvám je pro společnost zvláště důležitá. Pokud bude systém nepřiměřený nebo neúplný<sup>83</sup> a bude vykazovat nedostatky pro účely řešení majetkové či

---

<sup>81</sup> Srov. Evropská komise – Liability for Artificial Intelligence and other emerging digital technologies. [cit. 11. 01. 2020]. Dostupné z: <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>.

<sup>82</sup> LESIAK, Malgorzata. *A Comparative Analysis of The Liability of Internet Service Providers in The Context of Copyright Ingfringement in The U.S., European Union And Poland*. Časopisy Masarykovy univerzity [cit. 11. 01. 2020]. Dostupné z: <https://journals.muni.cz/mujlt/article/view/2541/2105>.

<sup>83</sup> Ve smyslu, že budou absentovat konkrétní řešení pro všechny případy způsobení majetkové či nemajetkové újmy v důsledku jednání autonomních systémů.

nemajetkové újmy způsobené vyvíjenými novými technologiemi, mohou osoby, jimž byla újma způsobena, skončit úplně nebo částečně bez kompenzace. Sociální dopad potenciální neadekvátnosti ve stávajících právních režimech při řešení nových rizik vytvářených novými technologiemi může ve svém důsledku ohrozit očekávané přínosy těchto technologií. Jak poukazuje i Evropská komise, neustále se rozvíjející přítomnost nových technologií ve všech aspektech společenského života, může v budoucnu také zhoršit či znásobit škody, které tyto technologie způsobí. Problém je v tom, že soudní praxe se podobnou situací dosud nezabývala, což se může ale brzy změnit. Je proto důležité vymezit, o jaký druh odpovědnosti se jedná.

Problém s přiřítáním odpovědnosti výrobcí, vývojáři, vlastníkovi nebo uživateli inteligentního stroje je tak stále otevřenou otázkou. Široká veřejnost může mít tendenci obviňovat samotné inteligentní roboty, i když uzná nedostatek jejich svobodné vůle a zavinění.<sup>84</sup> <sup>85</sup> Určitým způsobem, jak alespoň zpočátku proměny společnosti zajistit, aby společnost nevinila inteligentní roboty, a tedy neztrácela důvěru v jejich potenciál, může být snaha dostatečně odlišit tyto systémy a transparentně informovat uživatele o charakteru takového systému, aby dostatečně rozpoznali, že neinteragují s lidmi. Lze se domnívat, že takový přístup by mohl pomoci překonat fázi, kdy si společnost bude na narůstající počet autonomních systémů okolo sebe zvykat. Problematické je však to, že současný trend v robotice tomuto přístupu nepomáhá, jelikož roste snaha navrhovat inteligentní roboty tak, aby vypadaly jako lidé nebo zvířata a co nejlépe je imitovali.

Evropský parlament přijal dne 16. 2. 2017 usnesení, obsahující doporučení Komisi, ve kterém vyzývá k vytvoření silného evropského právního rámce upravujícího inteligentní

---

<sup>84</sup> LICHOCKI, P., KAHN, P. Jr., a BILLARD, A. *The Ethical Landscape of Robotics*. IEEE Robotics and Automation Magazine, 18(1):39-50, 2011.

<sup>85</sup> FUJITA, M. AIBO: *Toward the era of digital creatures*. The International Journal of Robotics Research, 20(10):781–794, 2001.



roboty v návaznosti na společenské trendy v oblasti vývoje umělé inteligence.<sup>86</sup> Vedle skutečnosti, že regulace je nutná pro evropské podniky, které potřebují jasné podmínky pro vývoj svých systémů tak, aby nebyl vývoj brzděn, je nutná také pro to, aby byla stanovena pravidla pro určení odpovědnosti za škodu způsobenou autonomními systémy. Evropský parlament upozorňuje, že u robotů, kteří jsou schopni činit autonomní rozhodnutí, nejsou existující odpovědnostní pravidla dostatečná. Slovy Evropského parlamentu „(...) *tradiční pravidla by již nestačila k určení odpovědnosti za škodu způsobenou robotem, protože by nebylo možné určit stranu, která má poskytnout odškodnění a napravit škodu, kterou robot způsobil.*“<sup>87</sup>

Právní úprava odpovědnosti z deliktů napříč členskými státy Evropské unie je z velké části neharmonizovaná. Výjimkou je právní úprava odpovědnosti za vady výrobků,<sup>88</sup> některé aspekty odpovědnosti za porušení ochrany osobních údajů<sup>89</sup> a odpovědnost za porušení principů hospodářské soutěže<sup>90</sup>. Vedle zmíněného neexistuje již ve vztahu k odpovědnosti za jednání autonomních systémů, vyjma zavedeného systému upravujícího pojištění

---

<sup>86</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

<sup>87</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

<sup>88</sup> Směrnice Rady 85/374/EHS o sblížení právních a správních předpisů členských států týkajících se odpovědnosti za vadné výrobky, ze dne 25. července 1985.

<sup>89</sup> Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (Text s významem pro EHP).

<sup>90</sup> Směrnice Evropského parlamentu a Rady 2014/104/EU ze dne 26. listopadu 2014 o určitých pravidlech upravujících žaloby o náhradu škody podle vnitrostátního práva v případě porušení právních předpisů členských států a Evropské unie o hospodářské soutěži (Text s významem pro EHP).

odpovědnosti za škodu způsobenou provozováním motorových vozidel,<sup>91</sup> žádná relevantní legislativa.

Na úrovni vnitrostátní legislativy jednotlivých členských států lze obecně konstatovat, že právní předpisy zatím spíše, až na pár výjimek,<sup>92</sup> neobsahují pravidla odpovědnosti konkrétně použitelná na škody způsobené používáním nových technologií jako je umělá inteligence, byť v režimu strojového učení. Výjimečně jurisdikce, které umožňují testování nebo pravidelné používání vysoce nebo plně automatizovaných vozidel, upravují problematiku odpovědnosti za způsobenou újmu buď pouze prostřednictvím systému povinného pojištění, nebo odkazem na obecná pravidla.

## 4.2 Současné pojetí občanskoprávní odpovědnosti

Hlavním účelem deliktního práva, resp. úpravy odpovědnosti z deliktů, je kompenzace vzniklé majetkové či nemajetkové újmy těm, kterým taková újma vznikla, a to při posouzení všech okolností daného případu. Luboš Tichý uvádí, že účelem kompenzace lze rozumět vážení zájmů (priorit), a to zejména na straně škůdce a poškozeného, do jisté míry však též zájmu společnosti, resp. třetích osob.<sup>93</sup> Zjednodušeně řečeno se jedná o hodnocení rozsahu a mezí povinností na straně jedné (škůdce) a legitimního očekávání na straně druhé (poškozeného). Deliktní právo sahá však pouze do té šíře, kterou mu umožní legislativci, a mnohé ze současných principů i ustanovení vznikly desítky let zpátky za odlišné situace a v odlišném kontextu.

---

<sup>91</sup> Směrnice Evropského parlamentu a Rady 2009/103/ES ze dne 16. září 2009 o pojištění občanskoprávní odpovědnosti z provozu motorových vozidel a kontrole povinnosti uzavřít pro případ takové odpovědnosti pojištění (Text s významem pro EHP).

<sup>92</sup> Např. Spojené království, Dánsko, Německo.

<sup>93</sup> TICHÝ, Luboš. *K rozlišování mezi tzv. subjektivní a objektivní odpovědností – rozsudek NS o povaze odpovědnosti advokáta za škodu*. Bulletin advokacie, odborný právní portál [cit. 12. 01. 2020]. Dostupné z: <http://www.bulletin-advokacie.cz/k-rozlisovani-mezi-tzv.-subjektivni-a-objektivni-odpovednosti>.

Při posuzování současného pojetí občanskoprávní odpovědnosti je z hlediska regulace autonomních systémů zásadní rozlišení na sankční a preventivní pojetí a objektivní a subjektivní odpovědnost. Odpovědnost v kontextu občanského zákoníku tak nadále není, ve srovnání s předchozí právní úpravou, vnímána jako odpovědnost sekundární (sankční), nýbrž jako odpovědnost za náležité plnění již povinnosti primární (preventivní), která působí jako hrozba sankcí.<sup>94</sup>

#### 4.2.1 Prevenční povinnost

Základní význam obecné prevenční povinnosti podle § 2900 občanského zákoníku<sup>95</sup> spočívá v zakotvení zásady *neminem laedere*, což z ekonomického pohledu vyjadřuje zákaz přenášení nákladů jednání určitého subjektu na jinou osobu (tj. zákaz externalit).<sup>96</sup> Prevenční zásada prostupuje celým soukromým právem a její zakotvení v § 2900 občanského zákoníku pak působí preventivně především tím, že s jejím porušením je spojován negativní následek pro konkrétní subjekt v podobě vzniku povinnosti k náhradě vzniklé újmy. Hlavní kritérium toho, zda se na někoho vztahuje obecná prevenční povinnost, spočívá v zohlednění okolností případu a zvyklostí soukromého života. Tento závěr akcentuje i dosavadní judikatura, když zdůrazňuje, že je vždy nutné zohlednit konkrétní časové a místní okolnosti.<sup>97</sup> Nejvyšší soud v tomto ohledu stanovil, že každý je povinen zachovávat vždy takový stupeň bedlivosti (pozornosti), který lze po něm vzhledem ke konkrétní časové a místní situaci rozumně požadovat a který - objektivně posuzováno - je způsobilý zabránit či alespoň co nejvíce omezit riziko vzniku škod na životě, zdraví či majetku.

---

<sup>94</sup> HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.

<sup>95</sup> „Vyžadují-li to okolnosti případu nebo zvyklosti soukromého života, je každý povinen počínat si při svém konání tak, aby nedošlo k nedůvodné újmě na svobodě, životě, zdraví nebo na vlastnictví jiného.“

<sup>96</sup> SZTEFEK, Martin. *Kritické poznámky k prevenční povinnosti v občanském zákoníku*. Právní rozhledy. 2018, č. 2.

<sup>97</sup> Srov. rozsudek Nejvyššího soudu ze dne 25. února 2003, sp. zn. 25 Cdo 618/2001.

Pokud jde o zvyklosti soukromého života, určujícími v tomto ohledu budou především kvazinormativní pravidla, kterými mohou být např. technické normy, sportovní pravidla, apod.<sup>98</sup> Sztefek uvádí, že soudům se tak otevírá široké pole působnosti, kdy mohou v rámci rozhodování individuálních případů zohlednit různé aspekty jednání, např. předvídatelnost újmy, rizikovitost jednání, nákladovost preventivních opatření, zohlednění poměru mezi osobami, apod. V kontextu autonomních systémů a umělé inteligence lze tedy s ohledem na odbornou literaturu dovodit, že je při jejím vývoji a provozování je nutné dodržovat alespoň příslušné technické normy. Dokud nedojde ke standardizaci vzorců chování plošně pro autonomní systémy, či zvláště například pro autonomní vozy, případně autonomní drony, a zvláště pro jiné systémy, je na každém výrobcu, vývojáři či uživateli, aby posoudil, jaký je ve vztahu ke konkrétnímu zařízení vhodný standard péče a prevence. Ve vztahu k existujícím normám zmiňuje Vláda ČR například mezinárodní ISO normy pro oblast strojového učení, které jsou však většinou stále ve stádiu příprav.<sup>99</sup>

Důvodem, proč Vláda ČR odkazuje právě na zmíněné normy, je skutečnost, že v kontextu vývoje umělé inteligence se momentálně jako nerizikovější oblast jeví právě výzkum strojového učení. Stále rostoucí počet sítí propojujících uživatele a zařízení umožňuje shromažďovat, ukládat, zpracovávat a analyzovat zdánlivě nekonečné množství dat. Vzhledem k tomu, že se autonomní systémy stávají součástí internetu věcí, potenciál sběru obrovského množství dat a jejich analýzy stále roste. Každé elektronické zařízení, aplikace,

---

<sup>98</sup> TICHÝ, Luboš a HRÁDEK, Jiří. *Deliktní právo*. V Praze: C.H. Beck, 2017. Beckova edice právní instituty. ISBN 978-80-7400-625-8.

<sup>99</sup> Viz ISO. Standard ISO/IEC 2382:2015 Information technology – Vocabulary. In: *International Organization for Standardization* [online]. 2015 [cit. 19. 12. 2019]. Dostupné z: <https://www.iso.org/standard/63598.html>, ISO. Standard ISO/IEC WD 22989 Artificial Intelligence – Concepts and terminology. In: *International Organization for Standardization* [online]. [cit. 19. 12. 2019]. Dostupné z: <https://www.iso.org/standard/74296.html>, či ISO. Standard ISO/IEC WD 23053 Artificial Intelligence (AI) Systems Using Machine Learning (ML). In: *International Organization for Standardization* [online]. [cit. 19. 12. 2019]. Dostupné z: <https://www.iso.org/standard/74438.html>.

senzory, GPS zařízení, peněžní transakce, použití karet/čipů či například aktivita na sociálních sítích neustále produkují data a tento proces se nikdy nezastaví.<sup>100</sup>

Tato skutečnost má samozřejmě také zásadní dopad na otázky ochrany soukromí, zejména pokud se jedná o zpracování osobních údajů. I když společně s vývojem takových technologií dochází k vývoji prostředků pro lepší ochranu soukromí, tyto metody nemusí být zdaleka dokonalé a jakákoli technologie sama o sobě nemůže efektivně chránit soukromí, aniž by existoval vhodný právní rámec pro předcházení nežádoucích účinků. Vláda ČR v tomto ohledu v rané fázi vývoje umělé inteligence například odkazuje na doporučení využívání přístupu tzv. black-box testingu. Black-box testing je definován jako testovací technika, při které se testuje funkčnost aplikace AUT (Application Under Test) bez ohledu na strukturu vnitřní logiky, podrobnosti implementace a znalosti interních postupů softwaru. Jinými slovy, v rámci prováděného testu lze zkoumat pouze vstupy a výstupy a nikoliv samotné vnitřní fungování systému, což znamená, že testy se provádějí z pohledu uživatele a pomáhají odhalit nesrovnalosti ve vstupech a výstupech. Další výhodou je, že osoba, která provádí zkoušku, nemusí znát programovací jazyk ani způsob implementace softwaru. Nevýhodou pak může být to, že testování není důkladné a detailní a nemusí umožnit detekci některých pochybení v procesu vytváření výstupu právě proto, že v rámci této metody testování se vývojáři zaměřují pouze na vstupy a výstupy, aniž by zkoumali či průběžně posuzovali interní postupy softwaru.

Z pohledu prevenční povinnosti může však umělá inteligence hrát i opačnou roli. Jednou z oblastí vývoje moderních technologií je využití modelů umělé inteligence k identifikaci anomálního chování. Tato umělá inteligence má za cíl analyzovat data ze senzorů zařízení a vytvářet smysluplné poznatky pro zabránění prostoje nebo nehodám.

---

<sup>100</sup> What Is Big Data? - Gartner IT Glossary - Big Data [online]. [cit. 18. 12. 2019]. Dostupné z: <https://www.gartner.com/it-glossary/big-data/>.

Tato inteligence, běžně známá jako prediktivní údržba,<sup>101</sup> předpovídá, kdy nebo zda zařízení selže, takže jeho údržba a opravy mohou být naplánovány a provedeny dříve, než dojde k poruše.

Sledování stavu zařízení v reálném čase a provádění prediktivní údržby tak může být v případě některých nebezpečných provozů za nedlouho považováno za standard dodržování prevenční povinnosti. Neustálý vývoj v oblasti big data, komunikace mezi inteligentními stroji a cloudových technologií však vytváří nové možnosti pro zkoumání informací, ale postupně vytváří také novou realitu, tzn. nové okolnosti, které jsou v každodenním životě rutinní, včetně nových rizik. Data totiž na jedné straně představují palivo, které pohání algoritmy řízenou společnost, ale na druhé straně jsou také jejím produktem. Shromažďování a zpracování dat vytváří stále více dat, která lze následně použít ke zlepšování výkonu algoritmů. Není tak překvapující, že v souvislosti s novými technologiemi je běžně používán přívlastek „rušivé“ (anglicky „disruptive“). Je zjevné, že nové technologie přináší hrozby, které přesahují hranice soukromí. S rostoucím výkonem a rozšířením schopností strojového učení lze pak očekávat, že širší využívání těchto systémů povede ke změnám a posunu v oblasti hrozeb a potřebě neustálé adaptace na nová rizika. Lze očekávat, že dojde k rozšíření stávajících hrozeb, vzniku nových i změně kvality a charakteru stávajících hrozeb. Náklady na možné útoky použitím nových technologií, které by obvykle vyžadovaly lidské úsilí, inteligenci a odbornost, budou kontinuálně klesat, útoky umožněné rostoucím používáním nových technologií budou zvláště účinné, automatizovanější, obtížně identifikovatelné a následně přiřítelné konkrétní osobě, což může ve svém souhrnu značně ztížit možnost

---

<sup>101</sup> SRIVASTAVA, Nishit a MONDAL, Sandeep. *Development of a Predictive Maintenance Model Using Modified FMEA Approach* (November 4, 2014). The IUP Journal of Operations Management, Vol. XIII, No. 2, May 2014, pp. 7-16 [cit. 19. 12. 2019]. Dostupné z: <https://ssrn.com/abstract=2518973>.

jejich odhalení. Příkladem stále častějších úniků dat mohou být série skandálů spojených zejména se společnostmi Facebook,<sup>102</sup> Yahoo<sup>103</sup> nebo Marriott International<sup>104</sup>.

V souvislosti s prevenční povinností se tak jako základní doporučení jeví zejména vyhotovení výkladových pravidel rozsahu prevenční povinnosti ve vztahu k jednotlivým technologiím či jednotlivým oblastem využití technologií, zpracování údajů i testování inteligentních systémů.

#### 4.2.2 Subjektivní a objektivní odpovědnost

Rozdělení odpovědnosti na objektivní a subjektivní hraje důležitou roli, jelikož pracuje s předpokladem existence či neexistence zavinění. Zavinění, jež je chápáno jako znak subjektivní odpovědnosti, který ji odlišuje od odpovědnosti objektivní, znamená psychický (vnitřní či duševní) vztah škůdce k protiprávnímu jednání, jakož i ke způsobenému škodnému následku a jejich příčinné souvislosti. Svou subjektivní povahou se zavinění odlišuje od protiprávního jednání, škody a příčinné souvislosti, které tvoří jediné předpoklady odpovědnosti objektivní. Zavinění dle Švestky předpokládá, že *„rušitel je po psychické (subjektivní) stránce, tj. po rozumové (intelektuální, předvídací, rozeznávací) stránce schopen svůj protiprávní čin i jeho způsobené následky, tj. majetkovou či nemajetkovou újmu, posoudit a zároveň je svou vůlí ovládnout, tj. rozhodnout se, zda protiprávní čin s těmito následky*

---

<sup>102</sup> ISAAC, Mike, a FRENKEL, Sheera. *Facebook Security Breach Exposes Accounts of 50 Million Users*. TheNew York Times, Sept. 29, 2018, Section A [cit. 10. 5. 2020]. Dostupné z: <http://Facebook Security Breach Exposes Accounts of 50 Million Users>.

<sup>103</sup> WILLIAMS, Martyn. *Inside the Russian hack of Yahoo: How they did it*. Security news, features and analysis about prevention, protection and business innovation. [online]. [cit. 10. 5. 2020]. Dostupné z: <https://www.csoonline.com/article/3180762/inside-the-russian-hack-of-yahoo-how-they-did-it.html>.

<sup>104</sup> FRUHLINGER, Josh. *Marriott data breach FAQ: How did it happen and what was the impact?* Security news, features and analysis about prevention, protection and business innovation. [online]. [cit. 10. 5. 2020]. Dostupné z: <https://www.csoonline.com/article/3441220/marriott-data-breach-faq-how-did-it-happen-and-what-was-the-impact.html>.

*uskuteční, či naopak, zda se jich zdrží.*<sup>105</sup> Zavinění lze tedy, alespoň ve smyslu nedbalosti, definovat jako opomenutí náležité péče. Jednání však může být považováno za zavrženíhodné jenom tehdy, jestliže bylo ovlivnitelné, resp. ovladatelné vůlí. Nebylo-li jednání takto ovladatelné, nejedná se o jednání v právním slova smyslu, a potom je otázka zavinění irelevantní a je třeba postupovat dle ustanovení operujících s konceptem odpovědnosti objektivní.

Objektivní odpovědnost se v kontextu občanského zákoníku uplatňuje v případě porušení smluvní povinnosti. Mimoto pracuje občanský zákoník s objektivní občanskoprávní odpovědností ve zvláštních případech činností. Jedná se o případy či situace, v nichž z jejich povahy může dojít k značně rozsáhlým a významným újmám. Tento přístup je pro případnou regulaci odpovědnosti u autonomních systémů příznačnější, jelikož vzhledem k výše uvedenému nelze v současné době u takových systémů, byť v režimu strojového učení, uvažovat o konceptu zavinění. Riziko vzniku újmy totiž nebude vyplývat z psychického (vnitřního či duševního) vztahu škůdce, ani z jednání odpovědné osoby, nýbrž ze samotné povahy rizikové činnosti. Pouze přístup založený na objektivní odpovědnosti může totiž v kontextu nových technologií sloužit jako správný a spravedlivý koncept, schopný zareagovat na vyvstalé otázky a odpovědnost za protiprávní stav někomu přiznat.

Pro případnou budoucí regulaci odpovědnosti za jednání autonomních systémů, se například uvažuje o využití stávající koncepce odpovědnosti v případě škody způsobené zvířetem či osobou, která nemůže posoudit následky svého jednání, případně se často skloňují různé inspirace doktrínou *respondeat superior* (viz dále), byť tento postup může vést k relativně přísné regulaci.

Účelem úpravy obsažené v ustanoveních § 2920 či § 2933 občanského zákoníku je chránit před nebezpečím, jehož zdrojem jsou zvířata či osoby, které z různých důvodů

---

<sup>105</sup> DVOŘÁK, Jan, ŠVESTKA, Jiří, ZUKLÍNOVÁ, Michaela a kol. *Občanské právo hmotné*. Díl první, Obecná část. 2., aktualizované a doplněné vydání. Praha: Wolters Kluwer, 2016. ISBN 978-80-7552-187-3.



nemohou samy nést odpovědnost za své jednání. Komentářová literatura konkrétně k ustanovení § 2933 uvádí, že „[z]vířata jsou živými tvory, kteří se pohybují, mají svoji sílu, často se chovají na základě různých instinktů a popudů (nepředvídatelně) a ocitají se v situacích, při nichž nepodléhají lidské kontrole. Kdo vlastní zvíře, vystavuje ostatní tomuto nebezpečí. Jako vyrovnání, že ostatní musejí toto riziko snášet, musí vlastník nést povinnost k náhradě újmy, byť vykonal vše, co po něm bylo možné rozumně požadovat, aby vzniku újmy zabránil.“<sup>106</sup> Smyslem citovaného pravidla je chránit před vším, co může zvíře ve své nevyočitatelnosti způsobit a tato nevyočitatelnost je jevem, který je v případě autonomních systémů taktéž typický.

Za povinnou osobu označuje v tomto případě občanský zákoník vlastníka zvířete. Jeho povinnost k náhradě je tak velmi přísná, neboť nezáleží na tom, zda má nad zvířetem dohled, nebo ne. Povinnost nese vlastník z titulu svého vlastnického práva, nezáleží tedy na faktickém stavu, rozhoduje zde stav právní. Zatímco vlastník zvířete nese povinnost k náhradě vždy, může být v určitých případech nárok na náhradu škody uplatněn i vůči osobě, které vlastník zvíře svěřil a která je s jeho souhlasem chovala nebo je pouze využívala. Taková osoba hradí škodu s vlastníkem podle pravidel společné a nerozdílné odpovědnosti.

Co se však týče přirovnání umělé inteligence ke zvířeti a související aplikace ustanovení o škodě způsobené zvířetem, je třeba upozornit, že tato analogie má za důsledek několik problematických bodů. Jedním z nejakcentovanějších problémů je skutečnost, že činnosti umělé inteligence jsou založeny na algoritmickém procesu podobném racionálnímu lidskému myšlení a pouze částečně podobným instinktům a smyslům, jako jsou ty zvířecí.<sup>107</sup> Dalším, poměrně závažným problémem je skutečnost, že jednání a rozhodování umělé inteligence je závislé na jejím naprogramování, tedy v drtivé většině případů závisí na

---

<sup>106</sup> HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.

<sup>107</sup> ČERKA, Paulius, GRIGIENÈ, Jurgita, SIRBIKYTÈ, Gintarè. *Liability for Damages Caused by Artificial Intelligence*. Computer Law & Security Review 31, 2015.

činnosti osoby odlišné od jejího koncového uživatele. V případě analogického využití ustanovení o škodě způsobené zvířetem by však odpovědnost plně přešla na vlastníka technologie. V neposlední řadě lze poukázat na speciální liberační důvody upravené v § 2934 občanského zákoníku<sup>108</sup>, při jejichž naplnění by se vlastník měl odpovědnosti zprostit. Všechny uvedené liberační důvody, tj. umělá inteligence slouží k výkonu povolání či jiné výdělečné činnosti nebo slouží jako pomocník pro osobu se zdravotním postižením, budou vzhledem k očekávanému využití těchto technologií velmi časté, což by však v důsledku vedlo k nespravedlivé situaci, kdy by poškozená osoba neměla možnost požadovat náhradu způsobené újmy po žádné osobě.

V případě využití analogie s osobou, která nemůže posoudit následky svého jednání, se sice jako praktické jeví to, že tento mechanismus je na rozdíl od ustanovení o škodě způsobené zvířetem schopen reflektovat určitý vývoj umělé inteligence, který by se mohl odrazit ve spravedlivém očekávání ohledně rozsahu požadované pozornosti dohlížející osoby (analogicky jako v případě rodiče a dítěte). Posuzování takové „nezbytné míry“ může být však mnohem náročnější než například právě u dítěte (zejména s ohledem na to, že tuto otázku bude v případě sporu řešit soud), a to i s ohledem na skutečnost, že dohlížející osoba nemusí mít v konkrétním případě o možnostech umělé inteligence takovou představu, jako o možnostech osoby, která nemůže posoudit následky svého jednání, jelikož v této oblasti existuje v současné době mnohem méně srozumitelných zdrojů informací.

Co se týče doktríny *respondeat superior*, či jinak známé také jako „*master-servant rule*“, je tato typická zejména pro anglosaské právní systémy. Má implikace i v českém právním řádu, nebývá však v tuzemských pramenech akcentována v takové míře, jako v zahraničí.

---

<sup>108</sup> „Slouží-li domácí zvíře vlastníku k výkonu povolání či k jiné výdělečné činnosti nebo k obživě, anebo slouží-li jako pomocník pro osobu se zdravotním postižením, zproští se vlastník povinnosti k náhradě, prokáže-li, že při doзору nad zvířetem nezanedbal potřebnou pečlivost, anebo že by škoda vznikla i při vynaložení potřebné pečlivosti.“

Doktrína *respondeat superior* má své kořeny ve starověkém Římě. V té době se tento koncept vztahoval na otroky, jakožto výraz skutečnosti, že otrok byl pouhým nástrojem, a aplikoval se tam, kde otrok nemohl nést následky svého jednání. Později byl koncept rozšířen tak, aby se vztahoval nejen na otroky, ale také na zvířata a členy rodiny patera familias. V dnešní době je tato doktrína využívána v konceptech zaměstnaneckého jednání za zaměstnavatele či zastoupení, konkrétně v občanském zákoníku můžeme její projevy sledovat v ustanoveních § 1935<sup>109</sup> či § 2914<sup>110</sup> občanského zákoníku.

Podmínky použití tohoto konceptu uvádí ve svém stanovisku k věci C-396/12 i generální advokátka Juliane Kokott: „*Pokud osoba ke splnění smluvního závazku využije pomoc třetí osoby, je třeba ve zcela převažujícím počtu členských států vycházet z toho, že ve svých právních rádech dojdou k řešení, podle kterých je ve výsledku za pochybení svého pomocníka odpovědný vůči smluvnímu partnerovi dlužník. Je tedy odpovědný jak za zavinění pro něj činné třetí osoby, tak za své vlastní.*“<sup>111</sup> Doktrína tedy eliminuje případy, kdy by v důsledku zneužití třetí osoby či nedostatku její delikt ní způsobilosti nemohla být určena odpovědná osoba, případně by byla odpovědnost nespravedlivě přičtena pouze osobě jednající jménem někoho jiného.

Jak je patrné z výše uvedeného, jednou z podstat umělé inteligence je schopnost autonomního chování, což znamená, že její jednání může být do určité míry a v určitých situacích nezávislé na jednání člověka. V dohledné budoucnosti a v určitých případech tak zcela určitě nebude možné vždy dovést a prokázat zavinění fyzické osoby, včetně příčinné

---

<sup>109</sup> „*Plní-li dlužník pomocí jiné osoby, odpovídá tak, jako by plnil sám.*“

<sup>110</sup> „*Kdo při své činnosti použije zmocněnce, zaměstnance nebo jiného pomocníka, nahradí škodu jím způsobenou stejně, jako by ji způsobil sám. Zavázal-li se však někdo při plnění jiné osoby provést určitou činnost samostatně, nepovažuje se za pomocníka; pokud ho však tato jiná osoba nepečlivě vybrala nebo na něho nedostatečně dohlížela, ručí za splnění jeho povinnosti k náhradě škody.*“

<sup>111</sup> Stanovisko generální advokátky Juliane Kokott ve věci C-396/12 A. M. van der Ham, A. H. van der Ham-Reijersen van Buuren ze dne 24. října 2013.

souvislosti mezi zaviněním, jednáním a vzniklou škodou, případně dokonce zavinění umělé inteligence. Prozatím se tedy, jak bylo uvedeno výše, jako vhodnější jeví přístup založený na objektivní odpovědnosti a v kontextu účinné právní úpravy tak přichází v úvahu speciální skutkové podstaty zakládající objektivní odpovědnost, při jejichž analýze je možné postavení umělé inteligence mimo jiné porovnat s výrobky či například s provozem zvláště nebezpečným.

### **4.3 Odpovědnost za jednání autonomních systémů v kontextu účinné právní úpravy**

#### **4.3.1 Škoda způsobená provozem zvláště nebezpečným**

Jelikož zatím neexistuje systém upravující odpovědnost za škodu způsobenou autonomním systémem, jsme odkázáni na tradiční ustanovení o odpovědnosti civilního práva. Prvním relevantním ustanovením je § 2925 odst. 1 občanského zákoníku upravující odpovědnost za škodu způsobenou provozem zvláště nebezpečným. Ustanovení § 2925 odst. 1 občanského zákoníku stanoví, že „[k]do provozuje závod nebo jiné zařízení zvláště nebezpečné, nahradí škodu způsobenou zdrojem zvýšeného nebezpečí.“ Provoz se považuje za zvláště nebezpečný, pokud nelze předem rozumně vyloučit možnost vzniku škody ani při vynaložení řádné péče. Provoz autonomního systému podporujícího technologie umělé inteligence by bylo možné považovat za zvláště nebezpečný, provozovatel totiž nemůže vyloučit vznik škody.

Jak bylo uvedeno výše v této kapitole, režim delikt ní odpovědnosti by měl brát v úvahu základní priority (hodnoty, statky) relevantní pro rozhodnou oblast. Předchozí zkušenost vedla k zavedení a práci s pojmem nebezpečná či riziková činnost, nebezpečný podnik či zvláště nebezpečný provoz. Jedná se o případy či situace, kdy možnost vzniku závažné (tj. zvláště kvalifikované) škody nelze předem rozumně vyloučit ani při vynaložení řádné péče. Ustanovení zjevně primárně míří na sofistikované činnosti, využívání specifických technologií či zařízení. Podstatným znakem takového provozu je, že k újmě může dojít i bez ohledu na lidské pochybení. Důvodem regulace je v tomto případě možnost

existence vady, která spočívá v používání určitého přístroje či zařízení, jejichž fungování je mimo rámec lidského poznání a závisí na nepředvídatelných okolnostech, anebo je výsledkem působení přírodních sil, které člověk není schopen ovlivnit. Hulmák uvádí, že občanský zákoník nepodává definici zvláště nebezpečného provozu zřejmě z toho důvodu, aby nebyly učiněny zábrany výkladu tohoto pojmu při dalším technickém rozvoji,<sup>112</sup> což podporuje závěr o možné aplikaci této skutkové podstaty na provoz autonomních systémů. Vlastník (mnohdy provozovatel) takového zařízení tedy neodpovídá za jednání (tj. v subjektivním smyslu), ale za výsledek, neboli za událost (tj. v objektivním smyslu).

Nejvyšší soud ČR judikoval, že „[z]a provoz zvláště nebezpečný se považuje takový provoz, s nímž je spojena možnost zvýšeného nebezpečí vzniku závažných škod. Zdrojem zvýšeného nebezpečí mohou být i věci, jejichž užívání má za určitých okolností vedlejší škodlivé účinky, které nejsou plně ovladatelné člověkem a s nimiž je spojen vysoký stupeň pravděpodobnosti vyvolání újmy.“<sup>113</sup> Není pochyby o tom, že provoz autonomního systému, využití technologií umělé inteligence, zejména za současného stavu vývoje strojového učení, z důvodů zmíněných výše v této kapitole tyto znaky naplňuje.

I provozovatel činnosti zvláště nebezpečné má v souladu s občanským zákoníkem k dispozici liberační důvod. Je s ním ovšem pracováno jinak než u provozní činnosti (viz § 2924), neboť zde se předpokládá, že ani vynaložení řádné péče nemůže předem rozumně vyloučit možnost vzniku závažné újmy. Při samotném vynaložení veškeré péče tedy nelze uvažovat o liberaci. Problematický je však v kontextu provozu autonomních systémů pojem „řádná péče“, ten bude muset být chápán v konkrétním kontextu. V případě autonomních systémů (nejlépe si lze představit případ autonomních vozů) by se za řádnou péči provozovatele mohla počítat pravidelná technická kontrola, pravidelná aktualizace softwaru

---

<sup>112</sup> HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.

<sup>113</sup> Viz rozsudek Nejvyššího soudu ČR ze dne 31. 5. 1983, sp. zn. 1 Cz 13/83.

apod. Jinak se provozovatel povinnosti nahradit vzniklou škodu zproští, prokáže-li, že škodu způsobila zvnějšku vyšší moc nebo že ji způsobilo vlastní jednání poškozeného nebo neodvratitelné jednání třetí osoby.

#### 4.3.2 Škoda, kterou způsobí věc sama od sebe

Dalším relevantním ustanovením může být § 2937 občanského zákoníku upravující odpovědnost za škodu, kterou způsobí věc sama od sebe. Ustanovení § 2937 odst. 1 občanského zákoníku stanoví, že „*[z]působí-li škodu věc sama od sebe, nahradí škodu ten, kdo nad věcí měl mít dohled; nelze-li takovou osobu jinak určit, platí, že jí je vlastník věci. Kdo prokáže, že náležitý dohled nezanedbal, zproští se povinnosti k náhradě.*“

Úvaha, že škodu může způsobit i věc sama od sebe, vychází z předpokladu, že bez ohledu na skutečnost, jak bylo s předmětnou věcí zacházeno, došlo k působení vnitřních vlastností či okolností spojených s povahou věci, které vedly ke vzniku škodné události. Komentářová literatura uvádí, že by do rozsahu působnosti předmětného ustanovení zřejmě mohly spadat i některé případy podléhající dříve režimu § 421a zákona č. 40/1964 Sb., občanský zákoník, kdy působení vnitřních charakteristických vlastností přístroje vyvolalo škodlivý účinek, ač postup při jejich použití byl v souladu se stanovenými pravidly (*lege artis*) a počínání osoby, která přístroj použila, nemělo na vznik škodlivého následku vliv.<sup>114</sup> Vojtek dále upřesňuje, že věc sama o sobě může působit i tehdy, není-li využívána k plnění závazku či dokonce není-li využívána vůbec. Tento závěr by mohl podpořit argumentaci k možnosti aplikace předmětného ustanovení v případě autonomních systémů, které by operovaly bez lidského dohledu a bez lidské interakce, pokud by chtěl někdo argumentovat, že v takovém případě „není využívána“.

---

<sup>114</sup> VOJTEK, Petr. *Občanský zákoník: komentář*. 2. vydání. Praha: Wolters Kluwer, 2019-. Komentáře (Wolters Kluwer ČR). ISBN 978-80-7598-412-8.

Odpovědnost v takovém případě nese ten, kdo nad věcí měl mít dohled, případně vlastník věci, nelze-li takovou osobu určit. Není přitom rozhodující, zda jej skutečně vykonával, nýbrž zda byl pouze povinen tak činit. Liberační důvod se odvozuje od kvality prováděného dohledu, neboť povinná osoba se může odpovědnosti zprostit, prokáže-li, že nezanedbala náležitý dohled. Parametry náležitého dohledu nejsou občanským zákoníkem nikterak stanoveny, a proto je možné tímto obecným ustanovením pokrýt všechny možné případy a posoudit je ad hoc.

V kontextu autonomních systémů může být těžké určit, co se myslí náležitým dohledem. Je třeba brát v potaz, že například robot v autonomním režimu se pohybuje ve společnosti, kde vlastním chováním interaguje s okolím. Zákon v tomto ohledu nesmí klást na provozovatele či vlastníky nepřiměřené požadavky, aby stále dohlíželi na chování robotů. Také nebude snadné stanovit, kdy robot způsobí škodu „sám od sebe“, neboli bez zapříčinění člověka. Autonomní systém, který je nadaný umělou inteligencí a schopností sám se učit z podnětů, které dostává ze svého okolí, neustále analyzuje lidské chování a vytváří z toho závěry, které je schopný aplikovat. Není proto možné s jistotou říci, že jakékoliv jednání autonomního systému v souvislosti se strojovým učením již není jednáním z podnětu člověka, a tudíž jedná činným systémem „samým od sebe“. Aplikace předmětného ustanovení rovněž závisí na jeho interpretaci v kontextu celého předpisu, občanského zákoníku, tedy pojem „věc“ zjevně odkazuje na definici věci dle § 489 občanského zákoníku. Aplikace předmětného ustanovení tedy závisí na definitivním zařazení autonomních systémů pod tento pojem.

### **4.3.3 Škoda způsobená vadou výrobku**

Ustanovení § 2939 občanského zákoníku upravuje odpovědnost za škodu způsobenou vadou výrobku, kdy stanoví, že *“[š]kodu způsobenou vadou movité věci určené k uvedení na trh jako výrobek za účelem prodeje, nájmu nebo jiného použití nahradí ten, kdo výrobek nebo jeho součást vyrobil, vytěžil, vypěstoval nebo jinak získal, a společně a nerozdílně s ním i ten, kdo výrobek nebo jeho část označil svým jménem, ochrannou známkou nebo jiným způsobem.”*

Lze se domnívat, že toto ustanovení bude v případě škody způsobené autonomním systémem užíváno nejčastěji (pokud se nebude jednat o odpovědnost provozovatele dopravního prostředku v případě autonomních vozidel). Aplikovatelnost ustanovení bude však záviset zejména na výsledcích v současnosti probíhajícího hodnocení směrnice o vadách výrobku expertní skupinou na odpovědnost a nové technologie zřízenou Evropskou komisí.<sup>115</sup> Úkolem expertní skupiny je totiž poskytnout Evropské komisi odborné posouzení k použitelnosti stávající směrnice o odpovědnosti za vadné výrobky na nové technologie v kontextu nových společenských výzev a pomoci Evropské komisi při vytváření zásad, které mohou sloužit jako vodítka pro možné úpravy platných právních předpisů v rámci EU i jednotlivých členských států.

Ustanovení stanoví, že vada musí být na věci movité, což autonomní systém vtělený do konkrétního zařízení nepochybně naplní. Poškozený musí prokázat existenci vady výrobku, přičemž nezávisí na jejím druhu, musí se pouze jednat o vadu ve smyslu § 2941 občanského zákoníku<sup>116</sup>. Poškozený nemusí prokazovat ani skutečnost, že vada vznikla ve sféře škůdce, ani okamžik vzniku vady, postačí, že konkrétní zařízení nenaplnuje standardy při jeho užívání. Jedním z mála předpokladů vzniku odpovědnosti je skutečnost, že výrobek musí být určen k uvedení na trh za účelem prodeje, nájmu nebo jiného použití. Hulmák toto pravidlo rozvádí, když uvádí, že z toho vyplývá, že povinnost nahradit újmu nevzniká pouze za okolností, že výrobek mění vlastníka v důsledku převodu vlastnického práva, ale také v případě, kdy zůstává ve vlastnictví škůdce, který jej pouze předá jinému do užívání (na základě nájemní smlouvy apod.).<sup>117</sup>

---

<sup>115</sup> Pro podrobnosti o skupině Expert Group on liability and new technologies (E03592) viz <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail&groupID=3592&Lang=CS>

<sup>116</sup> „Výrobek je ve smyslu § 2939 vadný, není-li tak bezpečný, jak to od něho lze rozumně očekávat se zřetelem ke všem okolnostem, zejména ke způsobu, jakým je výrobek na trh uveden nebo nabízen, k předpokládanému účelu, jemuž má výrobek sloužit, jakož i s přihlédnutím k době, kdy byl výrobek uveden na trh.“

<sup>117</sup> HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.



V tomto případě však ve většině případů nebude za škodu odpovědný provozovatel či vlastník výrobku, nýbrž nejčastěji výrobce zařízení nebo výrobce softwaru podle toho, kde nastala konkrétní vada. Za této právní konstrukce hrozí situace, že výrobci budou méně ochotni vyvíjet nové systémy, ačkoliv poptávka spotřebitelů po autonomních systémech by mohla kontinuálně růst, jelikož riziko bude na jejich straně vysoké. Logickým řešením by mohla být taková právní konstrukce, kdy by se výrobce mohl povinnosti zprostit, pokud by prokázal, že naplnil všechny bezpečnostní standardy. Výrobek se totiž v souladu s § 2941 občanského zákoníku považuje za vadný, právě tehdy, pokud není tak bezpečný, jak od něho lze se zřetelem ke všem okolnostem rozumně očekávat. Současná právní úprava umožňuje liberaci v případě, že stav vědeckých a technických znalostí v době, kdy uvedl výrobek na trh,<sup>118</sup> neumožnil zjistit jeho vadu, avšak tato skutečnost je v kontextu umělé inteligence mnohem hůře prokazatelná než fakt, že výrobce postupoval dle schválených bezpečnostních standardů.

Pro případ uplatnění kteréhokoliv z uvedených liberačních důvodů však stále může nastat situace, kdy nebude určitelná osoba odpovědná za škodu a poškozený se nebude schopen domoci kompenzace.

#### **4.3.4 Škoda z provozu dopravních prostředků**

Konkrétně v případě provozu autonomních vozidel či dronů bude nejčastěji aplikován (kromě případů, kdy se bude jednat o zjevně vadný výrobek) § 2927 občanského zákoníku, který upravuje škodu z provozu dopravních prostředků a stanoví, že *„[k]do provozuje dopravu, nahradí škodu vyvolanou zvláštní povahou tohoto provozu. Stejnou povinnost má i jiný provozovatel vozidla, plavidla nebo letadla, ledaže je takový dopravní prostředek poháněn lidskou silou.“*

---

<sup>118</sup> Viz ustanovení § 2942 odst. 2 písm. e) občanského zákoníku.

Za provozovatele dopravního prostředku se v souladu s § 2927 občanského zákoníku považuje ta osoba, která má právní a faktickou možnost dispozice s konkrétním dopravním prostředkem. Dle komentářové literatury se dispozice s konkrétním dopravním prostředkem vyznačuje především využíváním jeho základních funkcí (jízdu, letem), ale i jeho údržbou, placením provozních nákladů, placením oprav apod.<sup>119</sup> Tento extenzivní výklad podporuje možnost aplikace předmětného ustanovení, jelikož i v případě, že provozovatel dopravního prostředku fakticky neřídí (tento je řízen autonomně), stanoví pravidlo pro určení odpovědné osoby. Aplikace předmětného ustanovení v případě autonomních vozidel je podporována i skutečností, že právní úprava činí odpovědnou osobu, která je provozovatelem, nikoliv řidiče, což by v kontextu autonomního jednání činilo větší potíže.

Provozovatel se povinnosti k náhradě škody zproští tehdy, prokáže-li, že škodě nemohl zabránit ani při vynaložení veškerého úsilí, které lze požadovat. Provozovatel musí prokázat, že vyvinul takovou péči, kterou je možné za podmínek konkrétního případu spravedlivě žádat, aby se zabránilo vzniku újmy. V současnosti, kdy autonomní vozidla nejsou na plné úrovni autonomie, může být ustanovení § 2927 občanského zákoníku dostačující, jelikož autonomní vozidla by měla mít schopnost vyzvat řidiče k okamžitému převzetí řízení, pokud podle okolností usoudí, že poměry neumožňují provoz v autonomním režimu. Ve chvíli, kdy budou autonomní vozidla provozována téměř plně v autonomním režimu, může být však takové ustanovení příliš přísné, jelikož v takové situaci se již od řidiče (provozovatele), či jiných cestujících, nevyžaduje pozornost při provozu a nelze tak rozumně očekávat jejich připravenost a mnohdy ani schopnost řízení převzít.

#### 4.3.5 Potřeba adaptace právní úpravy

Z důvodu toho, že v současné době a v dohledné budoucnosti budou systémy umělé inteligence stále do jisté míry podléhat určitému dohledu lidí, se práce zaměřila i na instituty,

---

<sup>119</sup> HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.

u kterých je možné nalézt jistou korelaci s autonomním rozhodováním umělé inteligence a kdy je škodu povinen nahradit právě ten, kdo dohled vykonával, nebo nezbytný dohled zanedbal. Tyto instituty (např. § 2936 občanského zákoníku) sice v době jejich vytváření s tak rozsáhlým využitím umělé inteligence nemohly počítat, nicméně jejich analogické užití přichází, za předpokladu výkonu dohledu, v souvislosti s odpovědností za jednání umělé inteligence v úvahu.

Je však zřejmé, že platná právní úprava nebude připravena na situace, kdy budou již mnohem pokročilejší autonomní systémy způsobovat škodu, nebo i jinou újmu. Vláda ČR v tomto ohledu uvádí, že *„[s]távající pravidla odpovědnosti jsou v současné době aplikovatelná, avšak s rostoucí mírou využívání AI a s rostoucí komplexitou těchto systémů a jejich propojování pravděpodobně nebude možné tento stav udržet. V závislosti na konkrétním vývoji na evropské úrovni proto bude vhodné zvážit vytvoření zvláštního odpovědnostního režimu pro inteligentní systémy s určitou mírou autonomie, který by odstranil případné nedostatky v situacích, kdy nebude možné určit odpovědnou osobu.“*<sup>120</sup> Vláda tak jako nejrizikovější faktory do budoucna a v rámci možnosti využití stávajících institutů vnímá zejména stále narůstající komplexitu autonomních systémů a jejich propojování a za vhodné řešení považuje přijetí zvláštního odpovědnostního režimu pro inteligentní systémy, které jsou autonomní.

Evropský parlament ze stejného důvodu navrhuje mimo jiné i zvážení přiznání právní osobnosti inteligentním robotům (statusu elektronické osoby), jak bylo posuzováno výše, a umožnění, aby byl robot odpovědným subjektem za svá vlastní jednání. Evropský parlament zdůrazňuje, že *„(...) ať bude zvoleno jakékoli právní řešení problému občanskoprávní odpovědnosti za škody způsobené roboty v případech, které se netýkají škody na majetku, neměl by budoucí legislativní nástroj žádným způsobem omezovat druh nebo rozsah škody,*

---

<sup>120</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Souhrnná zpráva. Vláda ČR [online] [cit. 19. 12. 2019]. Dostupné z: <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-souhrnna-zprava-2018.pdf>.

*jejíž nahrazení lze vymáhat, a neměl by ani omezovat formu náhrady, jež může být nabídnuta poškozené straně, pouze na základě toho, že škodu nezpůsobil člověk (...).*<sup>121</sup> Nad rámec doporučení vlády tak Evropský parlament zdůrazňuje, že je třeba dbát na to, aby při aplikaci zvoleného či adaptovaného odpovědnostního režimu nedocházelo k omezení rozsahu a druhu škody, která může vzniknout a kterou je třeba z legitimních důvodů a s ohledem na očekávání dotčených subjektů nahradit, pouze z důvodu toho, že nebyla způsobena přímo člověkem.

Usnesení nabízí dva možné přístupy k odpovědnosti za jednání robota. Přístup absolutní odpovědnosti a přístup založený na řízení rizik. Předpokládá, že *„budoucí legislativní nástroj by měl vycházet z důkladné analýzy, kterou provede Komise s cílem stanovit, zda má být uplatňován přístup absolutní odpovědnosti, nebo přístup založený na řízení rizik.“*<sup>122</sup> Přístup absolutní odpovědnosti vyžaduje pouze důkaz, že došlo ke vzniku škody a existuje příčinná souvislost mezi jednáním robota a vzniklou škodou, jedná se o princip objektivní odpovědnosti. Přístup založený na řízení rizik předpokládá, že odpovědnou osobou bude osoba, která byla za určitých okolností schopna minimalizovat rizika a negativní dopady. Obě řešení by byla, podle Evropského parlamentu, zároveň doplněna systémem povinného pojištění a povinné registrace robotů do centrální evidence.

Největší problém v tomto ohledu činí autonomní systémy, které se nezávisle na lidech a neustále vyvíjejí a jen těžko lze u nich prokazovat příčinnou souvislost či uvažovat zavinění. Právě v těchto situacích přichází v úvahu výše zmíněné povinné pojištění, byť samo o sobě také nepředstavuje ideální řešení. Zavedení povinného pojištění v případě provozu vyspělých autonomních systémů by totiž vyžadovalo detailní znalost všech možných rizik a poskytování širokého rozsahu dat pojišťovnám. S ohledem na riziko multikauzální újmy bude současně

---

<sup>121</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

<sup>122</sup> Tamtéž.

nadále existovat požadavek alespoň částečné odpovědnosti. Velkým přínosem tohoto řešení však bude vyšší jistota poskytnutí kompenzace poškozeným subjektům.

Jak tedy bylo uvedeno výše, byť ideálním a z právního hlediska nejpříjemnějším by bylo přijetí zvláštního ustanovení upravujícího odpovědnost za provoz umělé inteligence, obdobně jako veškeré ostatní světové právní řády se ani český, resp. unijní, zatím de lege lata s otázkou náhrady újmy způsobené umělou inteligencí nijak výslovně nepotýká. Skutečností však je, že v případě vzniku újmy v přímé souvislosti a v důsledku jednání umělé inteligence by takováto újma v praxi nezůstala bez náhrady jen proto, že neexistuje výslovná právní úprava. V souladu se zásadou analogie legis by muselo dojít k aplikaci právní normy v konkrétním případě obsahem nejbližší tak, aby mohlo dojít ke spravedlivému uspořádání soukromoprávních nároků. Lze se domnívat, že soudy by aplikovaly ustanovení o škodě způsobené provozem zvlášť nebezpečným, škodě, kterou způsobí věc sama od sebe, škodě způsobené vadou výrobku nebo z provozu dopravního prostředku. V každém konkrétním případě se pak soud bude muset v odůvodnění vypořádat s případnými komplikacemi spojenými s aplikací toho konkrétního režimu.

## 5 Etické aspekty regulace autonomních systémů

Odborná skupina zřízená Evropskou komisí vydala v dubnu 2019 pokyny, jejichž cílem je stanovit rámec pro dosažení důvěryhodné umělé inteligence. Odborná skupina uvádí, že „[d]ůvěryhodná UI má tři složky, které by měly být dodrženy v průběhu celého životního cyklu systému: a) měla by být legální, tj. respektovat veškeré platné právní a správní předpisy, b) měla by být etická, tj. zajišťovat dodržování etických zásad, a c) měla by být robustní, a to z technického i sociálního hlediska, jelikož i dobře míněné systémy UI mohou způsobit neúmyslnou újmu. Každá složka je sama o sobě nezbytná, nepostačuje však k dosažení důvěryhodné UI. V ideálním případě všechny tři složky fungují v souladu a překrývají se. Dojde-li v praxi mezi těmito složkami ke konfliktům, měla by společnost usilovat o jejich odstranění.“<sup>123</sup> Odborná skupina tak staví naroveň legální požadavky, tj. respektování právních předpisů, etické zásady a robustnost systémů umělé inteligence. V souladu s přístupem založeným na apelu na dodržování základních práv poskytla odborná skupina výčet etických zásad a souvisejících hodnot, jež musí být dodržovány při vývoji, zavádění i používání systémů umělé inteligence. Konkrétně se jedná o (i) lidský faktor a dohled, (ii) technickou robustnost a bezpečnost, (iii) ochranu soukromí a správu dat, (iv) transparentnost, (v) rozmanitost, nediskriminaci a spravedlnost, (vi) dobré sociální a environmentální podmínky, a (vii) odpovědnost. Některým z těchto zásad, konkrétně transparentnosti, ochraně soukromí a nediskriminaci, které dle mého názoru tvoří nezbytné minimum důvěryhodné umělé inteligence, je věnována následující kapitola této práce.

Etika představuje pravidla chování ve společnosti, jedná se o normativní systém chování a zároveň základ systému všech pravidel (a to psaných i nepsaných), která jsou nastavena v každé společnosti. Vybraná etická pravidla jsou obsažena v normativních

---

<sup>123</sup> Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *ETICKÉ POKYNY PRO ZAJIŠTĚNÍ DŮVĚRYHODNOSTI UI*. Duben 2019 [cit. 19. 2. 2020]. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).

právních aktech a jsou vymahatelná státní mocí (důraz na dobré mravy, zákaz někoho poškozovat, povinnost jednat poctivě apod.). Velká část pravidel chování ve společnosti však vzniká spontánní interakcí mezi lidmi, jelikož zákon nedokáže pojmut a korektně definovat všechny žádoucí principy lidského jednání.

Aplikace výsledků etických studií a výzkumů závisí na porozumění a přístupu k autonomním systémům. Pojem „autonomie“ má své kořeny ve filozofii jakožto vědním oboru a značí schopnost fyzických osob vytvářet si vlastní pravidla, formulovat pravidla a myšlenky, které bude následovat. Zahrnuje právo na svobodu stanovovat si vlastní standardy a volit si vlastní cíle a zájmy. Kognitivní procesy, které to podporují a usnadňují, patří mezi ty, které jsou úzce spjaty s důstojností fyzických osob a lidskou myslí. Tento pojem obvykle zahrnuje rysy, jako jsou sebevědomí, sebereflexe a sebeurčení. Autonomii v eticky relevantním slova smyslu lze tedy připisovat pouze lidem.<sup>124</sup> Z jednoho úhlu pohledu je tedy poněkud nesprávné používat termín „autonomní“ na pouhé výtvořiny lidské činnosti, byť se jedná o velmi pokročilé, komplexní a adaptivní nebo dokonce inteligentní systémy. Z jiného úhlu pohledu však termín „autonomní“ získal v odborných kruzích, v odborné literatuře i v rámci široké veřejnosti význam, který odkazuje na nejvyšší stupeň automatizace a nejvyšší stupeň nezávislosti na lidech, pokud se jedná o provozní a rozhodovací autonomii. Pojem „autonomie“ v původním smyslu je však důležitým aspektem lidské bytosti, což by nemělo být relativizováno. Důvodem je skutečnost, že žádnému inteligentnímu robotu nebo systému, jakkoli pokročilému a sofistikovanému, který lze nazvat „autonomním“, nelze zatím přiznat morální postavení lidské bytosti a přiznat lidskou důstojnost. Objevují se i názory, že lidská důstojnost, jako jedno ze základních lidských práv, vyžaduje, aby rozhodnutí týkající se lidských bytostí, resp. mající vliv na lidské bytosti, byla vždy přičitatelná odpovědné osobě. Odborná skupina zřízená Evropskou komisí v tomto ohledu apeluje na respektování lidské

---

<sup>124</sup> European Group on Ethics in Science and New Technologies. *Artificial Intelligence, Robotics and 'Autonomous' Systems*. Luxembourg: Publications Office of the European Union, 2018. ISBN 978-92-79-80328-4.

důstojnosti a upřesňuje, že „[v] kontextu UI respektování lidské důstojnosti znamená, že se všemi lidmi je zacházeno s úctou jako s morálními subjekty, a nikoli pouze objekty, které mají být zkoumány, tříděny, bodovány, formovány nebo manipulovány. Systémy UI by proto měly být vyvíjeny způsobem, který respektuje fyzickou a duševní nedotknutelnost lidí, osobní a kulturní vědomí identity a uspokojení základních potřeb, slouží jim a chrání je.“ Jinými slovy, s ohledem na lidskou důstojnost by se lidé nikdy neměli stát pouhým předmětem automatizovaného rozhodování či hodnocení, nýbrž by k nim mělo být vždy přistupováno jako k morálním subjektům.

Dvě oblasti, ve kterých vývoj autonomních systémů již vedl k etickým debatám na vyšší či odbornější úrovni, jsou autonomní vozy a autonomní zbraně (tzv. LAWS).<sup>125</sup> Ačkoli auta bez řidiče ještě nejsou na trhu, několik zemí po celém světě se připravuje na úpravu právní regulace upravující možnost provozu autonomních vozidel na veřejných pozemních komunikacích. Systémy autonomních zbraní (AWS) jsou definovány americkým ministerstvem obrany jako „systémy, které, jakmile jsou aktivovány, mohou vybrat a zasáhnout cíle bez dalšího zásahu člověka.“<sup>126</sup> Vývoj technologie LAWS v širokém měřítku tedy představuje potenciál pro transformaci způsobu vedení války, a kvalitativně se tak liší od předchozích vojenských technologických inovací.

Z hlediska etiky je důležité i rozlišení mezi klasickými algoritmy a strojovým učením. Algoritmy jsou často popisovány jako vzorce operující na bázi vstupů a výstupů. Tradiční programování se spoléhá na funkce, které jsou charakteristické funkcí „if X, then Y“. Algoritmy jsou založeny na explicitních a pevně zakotvených pravidlech. Jak bylo popsáno v předchozích kapitolách, algoritmy strojového učení jsou složitější. Algoritmy učení se rozhodují nikoli podle předem naprogramovaných podmínek, které musí vstupní data

---

<sup>125</sup> Tamtéž.

<sup>126</sup> The Ethics of Autonomous Weapons Systems: *The Ethics of Autonomous Weapons Systems*. University of Pennsylvania Law School. Penn Law [online] [cit. 19. 2. 2020]. Dostupné z: <https://www.law.upenn.edu/institutes/cerl/conferences/ethicsofweapons/>



splňovat, ale prostřednictvím rozboru a statistických analýz stovek či tisíců datových setů v té oblasti, ve které činí rozhodnutí. Například v náborovém algoritmu, který hledá vhodné uchazeče, u nichž je největší pravděpodobnost úspěchu, může být datový set tvořen například údaji z celkem stovek životopisů od nejlepších kandidátů již zaměstnaných v dané společnosti. Algoritmus strojového učení poté hledá vzorce a korelace, které přispějí ke zdokonalení analýzy pravděpodobnosti úspěchu nového kandidáta. Předání této činnosti algoritmům strojového učení má pro zaměstnavatele mnoho výhod, včetně úspory času, peněz a úsilí. Pokud se však jedná o etiku a odpovědnost za konečné rozhodnutí, nemusí být mechanismus strojového učení již tak transparentní. Je tomu tak z toho důvodu, že hrozí situace, kdy nebudeme moct přesně pochopit, proč se stroj rozhodl určitým způsobem, proč učinil jednání, které učinil, a zejména dle informací dostupných v rámci současného výzkumu nedokážeme vždy detekovat a vyhnout se zaujatosti či diskriminace, když k tomu dojde.

Některé algoritmy umělé inteligence, které se využívají k předpovídání pravděpodobnosti recidivy trestné činnosti, například vykazovaly v průběhu výzkumu genderovou a rasovou předpojatost a vyvozovaly odlišnou pravděpodobnost recidivy u žen oproti mužům nebo u osob tmavé pleti oproti osobám se světlou pletí.<sup>127</sup> Jiné programy umělé inteligence pro analýzu obličeje například vykazovaly genderovou a rasovou předpojatost, která se projevovala nízkou chybovostí při určování pohlaví mužů se světlejší pletí a vysokou chybovostí při určování pohlaví u žen tmavší pleti.<sup>128</sup>

Dle Bílé knihy o umělé inteligenci, kterou čerstvě vydala Evropská komise (19. 2. 2020), má Evropská unie velmi dobré postavení a předpoklady k tomu, aby mohla těžit z potenciálu umělé inteligence, a to nejen jako uživatel, ale také jako tvůrce a výrobce této

---

<sup>127</sup> TOLAN, S., MIRON, M., GOMEZ, E. a CASTILLO, C. „*Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia*“. International Conference on AI and Law, 2019. Dostupné z: <https://doi.org/10.1145/3322640.3326705>.

<sup>128</sup> BUOLAMWINI, Joy, GEBRU, Timnit. „*Proceedings of the 1st Conference on Fairness, Accountability and Transparency*“, PMLR 81:77-91, 2018.

technologie.<sup>129</sup> V současné době má Evropská unie vynikající výzkumná střediska, inovativní start-upy, vedoucí postavení v oblasti robotiky z globálního hlediska a konkurenceschopné podniky v oblasti výroby a služeb, a to od automobilového průmyslu po zdravotnictví, energetiku, finanční služby i zemědělství. Evropské unii se podařilo vyvinout silnou výpočetní infrastrukturu, která je nezbytná pro fungování umělé inteligence. Evropská unie disponuje také velkým množstvím dat, jejichž potenciál je dle Evropské komise v současné době nedostatečně využíván. Zjednodušeně řečeno, Evropská unie si s ohledem na budoucí vývoj umělé inteligence a robotiky stojí velmi slibně, což pramení z již dlouholeté podpory a poskytování finančních prostředků relevantním odvětvím.

Pokrok v nových technologiích určených primárně pro spotřebitele tak v budoucnosti povede k vývoji asistentů, domácích robotů, zdravotnických robotů, konstrukčních robotů, robotů – zvířecích mazlíčků a hraček. Tyto robotické aplikace budou imitovat lidské i zvířecí chování a internet věcí a všudypřítomné aplikace umožní jejich vzájemnou komunikaci.

Často akcentovaná výhoda robotů a inteligentních strojů spočívá v jejich schopnosti donekonečna a neúnavně a prakticky bezchybně provádět řadu úkolů či pohybů a, jak bylo zmíněno v předcházejících kapitolách, v některých případech i „přemýšlet“ a činit relevantní rozhodnutí. V současné době se při navrhování robotů lidé zaměřují na schopnost robotů postupovat podle určitých vzorců, a v důsledku toho jsou roboti obecně vysoce specializovaní. Předpověď ne příliš vzdálené budoucnosti však slibuje, že se tento přístup změní a roboti budou čím dál častěji a ve větším rozsahu schopni plnit širokou škálu úkolů a imitovat, resp. i nahrazovat lidi. Tento vývoj bude možný, podaří-li se dosáhnout zvýšení kapacity paměti robotů a zdokonalení aplikací umělé inteligence (nebo dnes využívaného strojového učení), jenž poskytne přístup k obrovskému množství dat a umožní využití jejich potenciálu jejich použitím pro řadu různých úkolů.

---

<sup>129</sup> White Paper on Artificial Intelligence: a European approach to excellence and trust. European Commission [online] [cit. 25. 2. 2020]. Dostupné z: [https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en).

Časopis *Science Robotics* identifikoval v souvislosti se současným vývojem robotiky 10 největších překážek, mezi které patří vyjma například často akcentovaného rizika v oblasti bezpečnosti či soukromí například i otázka etiky a sociální interakce.<sup>130</sup> Pokud mají roboti vstoupit do každodenního světa lidí, budou se také muset naučit, jak se při interakci s lidmi chovat. Sociální roboti se budou muset naučit sociální kognice, budou muset porozumět kulturnímu a sociálnímu kontextu svého zařazení, pochopit mentální stav dotyčného člověka a přizpůsobit tomu své chování. Na druhou stranu to však vyžaduje i zvyšování povědomí společnosti, konkrétně uživatelů těchto technologií, a to alespoň v míře, která pomůže zvýšit komfort při běžném užívání a pochopit potenciální rizika a dopady na jejich soukromí.

I z analýzy Vlády České republiky vyplývá, že je s umělou inteligencí spojeno velké množství etických otázek, které „*lze obecně rozdělit do dvou kategorií – na tzv. roboetiku (angl. roboethics), tj. oblast etických problémů, s nimiž se potýkají lidé navrhující, vyvíjející a používající inteligentní stroje, a na etiku strojů (angl. machine ethics), tj. oblast zabývající se situacemi, kdy stroje rozhodují o etických problémech.*“<sup>131</sup>

## 5.1 Roboetika

Nalézt odpovědi na otázku, jak vytvořit na jedné straně fungující a straně druhé morální a společensky přijatelné autonomní systémy, není jednoduché. Tato otázka je velmi aktuální a z tohoto důvodu je i tématem výzkumu mnoha odborníků a akademiků, a to včetně již několikrát zmíněné Odborné skupiny na vysoké úrovni pro umělou inteligenci. Ta v nejobecnější rovině uvedla, že je třeba „*vyvíjet, zavádět a používat systémy UI způsobem, který dodržuje etické zásady týkající se: respektování lidské autonomie, předcházení újmám,*

---

<sup>130</sup> CROWE, S. *10 Biggest Challenges in Robotics*. Science Robotics [cit. 22. 2. 2020]. Dostupné z: <https://www.thebotreport.com/10-biggest-challenges-in-robotics/>.

<sup>131</sup> Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Analýza právně-etických aspektů rozvoje umělé inteligence a jejích aplikací v ČR. Vláda ČR [online]. Dostupné z: [https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018\\_final.pdf](https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf).

*spravedlnosti a vysvětlitelnosti. Uznat potenciální konflikty mezi těmito zásadami a zabývat se jimi.*<sup>132</sup>

Mezi konkrétní témata, která je dle Odborné skupiny třeba vzít v potaz a která je nutné implementovat za účelem dosažení důvěryhodné (ve smyslu robustní, etické a legální) umělé inteligence, patří, jak bylo stručně uvedeno výše: (i) lidský faktor a dohled (zahrnuje lidská práva, lidský faktor a lidský dohled), (ii) technická robustnost a bezpečnost (zahrnuje odolnost vůči útokům a zabezpečení, nouzový plán a obecnou bezpečnost, přesnost, spolehlivost a reprodukovatelnost), (iii) ochrana soukromí a správa dat (zahrnuje ochranu soukromí, kvalitu a integritu údajů a přístup k údajům), (iv) transparentnost (zahrnuje sledovatelnost, vysvětlitelnost a komunikaci), (v) rozmanitost, nediskriminace a spravedlnost (zahrnuje předcházení nespravedlivé podjatosti, přístupnost a univerzální design a zapojení zúčastněných stran), (vi) dobré sociální a environmentální podmínky (zahrnuje udržitelnost a šetrnost k životnímu prostředí, sociální dopad, společnost a demokracii), (vii) odpovědnost (zahrnuje auditovatelnost, minimalizaci negativních dopadů a podávání zpráv o těchto dopadech, kompromisy a možnost zjednání nápravy).<sup>133</sup>

Vzhledem ke skutečnosti, že některé z uvedených témat se v odborné literatuře opakují a některá témata už byla v rámci této práce analyzována, bude se práce dále v rámci roboetiky věnovat převážně právě otázce transparentnosti, ochrany soukromí a rizika vzniku předsudků a případů nerovného zacházení a diskriminace.

---

<sup>132</sup> Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *ETICKÉ POKYNY PRO ZAJIŠTĚNÍ DŮVĚRYHODNOSTI UI*. Duben 2019 [cit. 25. 2. 2020]. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).

<sup>133</sup> Tamtéž.

### 5.1.1 Transparentnost a důvěra

Moderní ekonomika, jak je blíže uvedeno dále v této kapitole, je založena na údajích, osobních či neosobních, kde je pro spotřebitele stále obtížnější pochopit obchodní praktiky subjektů operujících na trhu moderních technologií. Spotřebitelé by měli mít vždy jasnou představu o používaných algoritmických rozhodovacích procesech, které by na ně mohly mít vliv, zejména pak o těch, které by se jich mohly dotknout negativně. Spotřebitelé by proto měli mít právo na informace, právo zakotvené jako klíčová zásada v článku 169 Smlouvy o fungování Evropské unie.<sup>134</sup> Toto právo by mělo zahrnovat jak přístup k transparentním informacím o skutečnosti, že probíhá automatizované zpracování dat, tak o tom, jak toto zpracování funguje.

Evropský parlament například v Usnesení o komplexní evropské průmyslové politice v oblasti umělé inteligence a robotiky zdůrazňuje, že „vyzývá Komisi a členské státy, aby prosazovaly rozsáhlou a transparentní spolupráci mezi veřejným a soukromým sektorem a akademickou obcí, což by posílilo sdílení znalostí, a aby podporovaly vzdělávání a odbornou přípravu konstruktérů v oblasti etických důsledků, bezpečnosti a dodržování základních práv, ale také vzdělávání spotřebitelů zaměřené na použití robotiky a umělé inteligence, a to se zvláštním důrazem na bezpečnost a ochranu údajů.“<sup>135</sup>

V tomto ohledu se v souvislosti se zpracováním dat vyvinul pojem *etika dat*.<sup>136</sup> Etiku dat lze definovat jako obor etiky, který studuje a vyhodnocuje morální problémy související s používáním dat v moderním světě. Právo na soukromí je totiž základním lidským právem

---

<sup>134</sup> „K podpoře zájmů spotřebitelů a k zajištění vysoké úrovně ochrany spotřebitele přispívá Unie k ochraně zdraví, bezpečnosti a hospodářských zájmů spotřebitelů, jakož i k podpoře jejich práva na informace, vzdělávání a práva sdružovat se k ochraně svých zájmů.“

<sup>135</sup> Bod 140 Usnesení Evropského parlamentu ze dne 12. února 2019 o komplexní evropské průmyslové politice v oblasti umělé inteligence a robotiky (2018/2088(INI)).

<sup>136</sup> HASSELBALCH, G. (2019). *Making sense of data ethics. The powers behind the data ethics debate in European policymaking*. Internet Policy Review, 8(2). DOI: 10.14763/2019.2.1401.

a zejména při použití dat v souvislosti s novými technologiemi musí být chráněno. Etika dat zahrnuje vytváření interních programů vyvažujících přínosy a rizika pro jednotlivce plynoucí ze zpracování velkých dat (big data). Sběr a zpracování dat mohou být užitečné pro účely analýzy a vyhodnocení trendů například pro účely poskytnutí odpovídající reakce na běžné i mimořádné situace. Sladění střetu mezi těmito konkurenčními zájmy - ochrana soukromí jednotlivce vs. společenský přínos - je právě jádrem etiky dat.<sup>137</sup>

Právo, jakožto soubor obecně závazných právních předpisů, by mělo poskytnout jasná a srozumitelná pravidla od stádia vývoje po výrobu a využívání nových technologií, a to zejména v souladu se zásadou *privacy by design*. Zásada *privacy by design*, tzv. záměrná ochrana osobních údajů, zejména zdůrazňuje povinnost správce osobních údajů jak v době nastavování parametrů pro nové zpracování osobních údajů, tak i průběžně v jeho průběhu, aplikovat adekvátní technická a organizační opatření k zajištění toho, aby zpracování osobních údajů probíhalo v souladu s relevantními právními předpisy a aby byla respektována práva subjektů údajů. Dále vyžaduje, aby byli správci osobních údajů schopni prokázat, že zavedli požadovaná adekvátní opatření k ochraně všech zásad zpracování osobních údajů a že mají integrované zvláštní záruky, které jsou nezbytné právě k zajištění práv a svobod subjektů údajů. Začlenění principu *privacy by design* již při vývoji technologií je stěžejní, protože aplikace, které využívají velká data a umělou inteligenci (byť i v současně aplikované formě strojového učení), operují a mohou nadále operovat s čím dál menším lidským dohledem v reálném čase.

S principem *privacy by design* úzce souvisí právě i požadavek transparentního přístupu. Transparentní přístup je velmi důležitý a je nutné na něm dbát, jelikož buduje důvěru v nové technologie tím, že uživateli poskytuje možnost porozumět tomu, co systém dělá a proč. Nedostatek transparentnosti, vedle potenciálního odmítnutí technologií ze strany

---

<sup>137</sup> DIBLÍK, Jan, CHOLASTA, Roman, TADEVOSJANOVÁ, Laura. Robotika a regulace: „Privacy by design“ a transparentní přístup k novým technologiím. Národní centrum průmyslu 4.0, březen 2020 [cit. 10. 5. 2020]. Dostupné z: <https://www.ncp40.cz/aktuality/technologie-a-pravo-privacy-by-design>.

široké veřejnosti, ztěžuje i identifikaci a prokázání možného porušení právních předpisů, včetně právních norem, které chrání základní práva, upravují odpovědnost apod. Aby bylo zajištěno účinné uplatňování a vymáhání právních předpisů, je nezbytné upravit nebo vyjasnit stávající právní předpisy v určitých oblastech, například v oblasti úpravy odpovědnosti.

Spojení „transparentní a vysvětlitelná“ umělá inteligence odkazuje na metody a techniky při vývoji a aplikaci technologií umělé inteligence (resp. strojového učení) tak, aby výsledky jednání a učiněná rozhodnutí mohly být pochopeny a podrobeny analýze. Opakem tohoto přístupu je problém tzv. „černé skříňky“ (black-box) ve strojovém učení, kde ani jejich programátoři nemohou vysvětlit, proč systém dospěl ke konkrétnímu rozhodnutí. Transparentní a vysvětlitelná umělá inteligence představuje také implementaci práva na vysvětlení. V kontextu regulace algoritmů strojového učení je právo na vysvětlení právem na popis důvodů, postupu a procesů předcházejících konkrétnímu výstupu algoritmu. Tato práva primárně směřuje na ochranu individuálního práva na vysvětlení v případech rozhodnutí, která se jednotlivce zásadním způsobem dotýkají, zejména mají pro něj právně či obdobně významné dopady.<sup>138</sup>

Systémy strojového učení si někdy implementují nežádoucí postupy, které jsou sice schopny provést optimální práci při plnění explicitních, předem naprogramovaných úloh, avšak plně neodrážejí požadovaný postup ze strany programátorů. Například v roce 2017 systém, který byl pověřen identifikováním obrázků, na kterých je vyobrazen kůň, se naučil „podvádět“ tak, že místo rozpoznávání obrysů koně identifikoval obrázky, které odkazovaly na zdroj obrázků, ze kterého programátoři skutečně čerpali obrázky vyobrazující koně.<sup>139</sup>

---

<sup>138</sup> VOOSSEN, Paul. *How AI detectives are cracking open the black box of deep learning AAAS*. Science [online] [cit. 01. 02. 2020]. Dostupné z: <https://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning>.

<sup>139</sup> SAMPLE, Ian. *Computer says no: why making AIs fair, accountable and transparent is crucial*. The Guardian [online] [cit. 01. 02. 2020]. Dostupné z: <https://www.theguardian.com/science/2017/nov/05/computer-says-no-why-making-ais-fair-accountable-and-transparent-is-crucial>.

V rámci jiného pokusu se systém pověřený uchopením předmětů ve virtuálním světě naučil podvádět umístěním manipulátoru mezi objekt a diváka takovým způsobem, že se divák mylně domníval, že systém předmět uchopuje.<sup>140</sup>

Co se týče důvěry, lze se domnívat, že v kontextu rychlých technologických změn je nezbytné, aby důvěra stále zůstávala pojítkem společností, komunit, ekonomik a udržitelného rozvoje. Z tohoto důvodu také Odborná skupina na vysoké úrovni pro umělou inteligenci vytyčila jako svůj základní cíl právě důvěryhodnou umělou inteligenci. Vychází přitom z předpokladu, že lidé a komunity budou moci důvěřovat rozvoji této technologie a jejím aplikacím pouze tehdy, bude-li zaveden jasný a komplexní rámec pro dosažení její důvěryhodnosti.

Důraz na důvěryhodnou umělou inteligenci se netýká pouze důvěryhodnosti samotného software, nýbrž vyžaduje i důvěryhodný přístup všech osob, postupů a procesů v systému během celého jeho životního cyklu. Pokyny Odborné skupiny zdůrazňují, že *„[p]řístup založený na důvěryhodnosti je klíčem k zajištění „odpovědné konkurenceschopnosti“ tím, že poskytuje základ, díky němuž mohou všichni, koho se systémy UI dotýkají, důvěřovat, že jejich návrh, vývoj a používání jsou legální, etické a robustní.“* Pokyny Odborné skupiny tak usilují především o to, aby se etika stala hlavním pilířem přístupu v oblasti rozvoje umělé inteligence a aby všichni aktéři usilovali o prospěch, posílení postavení a ochranu jednotlivých osob i obecného blaha společnosti. Sdělením Odborné skupiny, s nímž se lze ztotožnit, je, že pouze při zajištění důvěryhodnosti budou koncoví uživatelé plně využívat výhod nových technologií a budou mít jistotu, že jsou zavedena opatření na ochranu před jejich potenciálními riziky.

S výše uvedenými doporučeními se lze ztotožnit, jelikož transparentnost, jako obecně schopnost interpretovat fungování modelů strojového učení či poskytovat vysvětlení konkrétních algoritmických výstupů, je důležitá s ohledem na potřebu popisovat nebo

---

<sup>140</sup> Tamtéž.



poskytovat vodítka pro provádění etických, ale i právních, přezkumů. Vedle podrobností o procesu či informací o tom, kdy bylo dosaženo rozhodnutí prostřednictvím automatizovaného systému, tento princip z pohledu uživatelů zahrnuje i pochopení záruk a práv, která mají v souvislosti s provozem a využíváním nových technologií k dispozici. Transparentnost jako zastřešující pojem pak v důsledku vede i k lepší vysvětlitelnosti. Důvěru v nové technologie pak může podpořit právě i požadavek tzv. auditovatelnosti, tj. požadavek, aby byly systémy podrobovány nezávislým auditům a výsledné zprávy, při zachování obchodního tajemství a dalších chráněných statků, byly dostupné široké veřejnosti.

### 5.1.2 Ochrana soukromí a dopad na lidské chování

Jak bylo uvedeno výše, stále rostoucí počet sítí propojujících uživatele a zařízení umožňuje shromažďovat, ukládat, zpracovávat a analyzovat zdánlivě nekonečné množství dat. Data se stala nepostradatelnou součástí každé ekonomiky, průmyslu, organizace, obchodní společnosti i jednotlivce. *Big data* je termín používaný k identifikaci datových sad, jejichž velikost překračuje schopnosti typických databázových softwarových nástrojů ukládat, spravovat a analyzovat. Velká data představují jedinečné výpočetní a statistické výzvy. Tyto výzvy vyžadují zejména nové a účinnější výpočetní a statistické modely.

Vzhledem k tomu, že se autonomní systémy stávají součástí internetu věcí, potenciál sběru obrovského množství dat a jejich analýzy stále roste. To má samozřejmě také zásadní dopad na otázky ochrany soukromí, zejména pokud jde o zpracování osobních údajů. Zejména nástup internetu, mobilních zařízení a dalších technologií zapříčinil zásadní posun charakteru a chápání dat a také způsobu jejich využití.

Rozsah a dopad datové (r)evoluce je takový, že všechny druhy lidské činnosti a lidských rozhodnutí začínají být čím dál více ovlivňovány předpověďmi na základě analýzy velkých dat, a to včetně nakupování, lékařských služeb, vzdělávání, prevence terorismu a kybernetické bezpečnosti a mnoha dalších.

Pojem *velká data* nemá jednoznačnou definici. Používání pojmu v současné době odkazuje na použití prediktivní analýzy, analýzy chování uživatelů anebo jiné pokročilé

metody analýzy dat, které umožňují z velkého množství informací extrahovat použitelná a potřebná data. Technické definice velkých dat jsou často úzce omezeny na popis jejich rozsahu a srovnání s klasickými, dosavadními datovými sety, tj. definují je jako data, která přesahují zpracovatelskou kapacitu tradičních databází a tradiční nástroje pro správu dat nejsou připraveny na takové zpracování ať již časově či nákladově efektivně.<sup>141</sup> Programátoři často používají technickou „3-V“ definici velkých skládající se z jednotlivých složek, kterými jsou objem (*volume*) – množství dat, typ (*variety*) – různorodost typů dat, rychlost (*velocity*) - rychlost s jakou data vznikají a s jakou roste potřeba jejich analýzy v reálném čase. Pojmem *velká data* tedy jinými slovy rozumíme nové způsoby, kterými jsou organizace, včetně veřejných orgánů, schopny kombinovat různé datové sety a pomocí technik těžení (tzv. *data mining*) získávat potřebné odvozené informace. Velká data slibují významné ekonomické a sociální přínosy, vyvolávají však také mnoho otázek a obav v některých často diskutovaných oblastech.

Pro lepší představu lze uvést několik příkladů. Marketingové společnosti vyžadují data, aby lépe cílily své strategie na spotřebitele, lékaři potřebují data, aby efektivněji zjistili vedlejší účinky léků, policie a státní orgány, aby lépe určily, kam zaměřit své omezené finanční i personální zdroje. V neposlední řadě sociologové využívají velké objemy dat ke studiu interakcí mezi lidmi. Počet firem využívajících analýzu velkých dat roste a nepochybně poroste i dál s ohledem na zvyšování objemu dat, která je nutné ukládat rychle a efektivněji.

Logicky nás může napadnout otázka, kde lze takové množství dat získat či jak se taková data generují. Odpověď není tak složitá. Data jsou konstantně generována prostřednictvím většiny zařízení a technologií, které používáme. Příkladem společnosti, která neustále marketingově využívá velká data, je bezesporu Google. A někdy nemusí být metody sběru a analýzy dat ani příliš složité, Seth Stephen-Davidowitz například uvádí, že chce-li

---

<sup>141</sup> What Is Big Data? - Gartner IT Glossary - Big Data. [online]. Copyright © 2019 Gartner, Inc. and [cit. 24. 05. 2019]. Dostupné z: <https://www.gartner.com/it-glossary/big-data/>

Google vědět, jak přimět víc lidí, aby klikli na reklamy na jejich stránkách, stačí, když použije klasickou metodu A/B testu, tedy například zkusí použít dva odstíny modré, jeden pro skupinu A, druhý pro skupinu B, a následně porovnat počet kliknutí.<sup>142</sup> Facebook například provádí tisíce A/B testů, které generují použitelné údaje, aniž bychom si to nutně uvědomili, denně, což dle Stephen-Davidowitze znamená, že několik pracovníků Facebooku provádí denně více řízených experimentů než veškerý farmaceutický průmysl za jeden rok.

Mimo uvedené experimenty je každý den generováno velké množství dat, které vypovídají o našich preferencích, chování, názorech a postojích, jak je například patrné i z níže uvedených statistik.

- Každý den je na Facebook nahráno přes 300 milionů fotografií;
- Každou sekundu se na Facebooku vytvoří průměrně 5 nových účtů;
- Každou minutu je na YouTube nahráno v průměru 300 hodin videa;
- Každý den je na Instagramu přibližně 4,2 miliardy krát zakliknuto „like“;
- Denně je prostřednictvím WhatsAppu zasláno přibližně 42 miliard zpráv a sdíleno přibližně 1,6 miliardy fotografií; nebo
- Na Twitteru je sdíleno více než 140 milionů tweetů denně, což v součtu představuje až miliardu tweetů týdně.<sup>143</sup>

Z jiného úhlu pohledu – každých deset minut je na sociálních sítích vygenerováno stejné množství informací, jaké bylo dohromady vygenerováno desítkami tisíc generací před námi.<sup>144</sup> Stále klesající náklady na shromažďování, ukládání a zpracování dat spolu s novými

---

<sup>142</sup> STEPHENS-DAVIDOWITZ, Seth. *Všichni lžou: velká data, chytré algoritmy a jak nám internet může říct, kdo opravdu jsme*. Přeložila Romana HEGEDŮSOVÁ. Brno: Host, 2019. ISBN 978-80-7577-787-4.

<sup>143</sup> How Much Data Is Generated Every Minute? Social Media Today. Social Media News [online]. [cit. 18. 12. 2019]. Dostupné z: <https://www.socialmediatoday.com/news/how-much-data-is-generated-every-minute-infographic-1/525692/>.

<sup>144</sup> Tamtéž.

zdroji dat ze senzorů, kamer GPS apod. naznačují, že žijeme ve světě, kde se sběr dat stal téměř všudypřítomným, kde uchovávání dat může mít trvalý charakter a kde je jejich analýza stále častěji prováděna v rychlostech blížících se reálnému času. I když společně s vývojem takových technologií dochází k vývoji prostředků pro lepší ochranu soukromí, tyto metody nemusí být zdaleka dokonalé, a jakákoli technologie sama o sobě nemůže efektivně chránit soukromí, aniž by existoval vhodný právní rámec pro předcházení nežádoucích účinků.

Ve vztahu ke zpracování osobních údajů poskytuje unijní úprava ochranu prostřednictvím obecného nařízení o ochraně osobních údajů.<sup>145</sup> Pro organizace zpracovávající osobní údaje pomocí nových technologií je tak zásadní právě důraz na dodržování základních zásad upravených v článku 5 uvedeného nařízení. Podle první zásady musí být zpracování osobních údajů ve vztahu k subjektům údajů zákonné, transparentní a korektní. K naplnění požadavku zákonnosti je třeba, aby správce osobních údajů (v tomto případě organizace zpracovávající osobní údaje pomocí nových technologií) disponovala některým z právních titulů zpracování vyjmenovaných v článku 6 odst. 1 obecného nařízení o ochraně osobních údajů. Vedle výše zmíněné transparentnosti tak zůstává klíčovou otázkou pro organizace, které používají osobní údaje pro analýzu velkých dat, to, zda je zpracování z jejich strany ve vztahu k subjektům údajů korektní. Korektnost zahrnuje několik prvků, včetně posouzení účinků zpracování na jednotlivce a jejich očekávání ohledně způsobu, jakým budou jejich údaje použity. Organizace shromažďující a dále zpracovávající osobní údaje bude obecně muset transparentně a srozumitelně vysvětlit účely, pro které data potřebuje, ale to samo o sobě ještě nemusí nutně vysvětlit, jakým způsobem budou použita. Je důležité, aby tyto organizace zvážily, zda subjekty údajů mohou rozumně očekávat, že budou jejich data použita způsobem, který se odlišuje od dosavadních prostředků zpracování údajů.

---

<sup>145</sup> Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (Text s významem pro EHP).

Další principy ochrany osobních údajů jsou srovnatelně důležité. Předně musí správce osobních údajů stanovit určitý, výslovně vyjádřený a legitimní účel, pro který jsou údaje shromažďovány a pokud jsou údaje dále zpracovávány pro jakýkoli jiný účel, nesmí být tento neslučitelný s původním účelem. Dále musí správce dbát na to, aby byly údaje přiměřené, relevantní a omezené na nezbytný rozsah ve vztahu k účelu, pro který jsou zpracovávány.

Ve vztahu k velkým datům hrají důležitou roli i zásada přesnosti a zásada omezení uložení. Ve velkém objemu dat může být pro správce náročné udržet údaje přesné a v případě potřeby aktualizované, stejně tak zajistit, aby byly údaje uloženy ve formě umožňující identifikaci subjektů údajů po dobu ne delší, než je nezbytné pro účely, pro které jsou zpracovávány. V neposlední řadě je správce povinen zajistit, aby byly údaje zpracovávány způsobem, který zajistí jejich náležité zabezpečení, včetně jejich ochrany pomocí vhodných technických nebo organizačních opatření před neoprávněným či protiprávním zpracováním a před náhodnou ztrátou, zničením nebo poškozením.

S otázkou zpracování osobních údajů a prakticky stále narůstající mírou sledování fyzických osob a jejich činnosti souvisí rovněž obava z vlivu, který toto sledování může na tyto jednotlivce mít. Takovéto monitorování již samo o sobě může vést k ovlivňování sledovaných jedinců. Důvodem je skutečnost, že jedinec mění své přirozené chování s ohledem na obavu z toho, že je neustále sledován a hodnocen.

### **5.1.3 Předsudky a diskriminace**

Skutečnost, že jsou data využívána při předvídání či efektivnějším rozhodování, může ovlivnit osudy mnoha lidí, a to i nepříznivými způsoby. Vedle všech výše zmíněných dopadů, může těžení dat totiž také reprodukovat již existující vzorce diskriminace, opakovat předchozí předsudky, nebo zviditelnit jiné, skryté předsudky, které ve společnosti přetrvávají. Dokonce může mít za následek zhoršení existujících nerovností tím, že podpoří závěry, že historicky znevýhodňované skupiny si ve skutečnosti zaslouží méně příznivé zacházení. Vtělené algoritmy by mohly tyto tendence vykazovat i v případech, kdy k tomu nebyly cíleně naprogramovány. Taková diskriminace může vzniknout i v souvislosti se samotným procesem

těžení dat, než pouze jako jeho výsledek činnosti. Programátoři totiž přirozeně přiřazují určitým faktorům různou váhu, což může mít za následek nevyhnutelný diskriminační potenciál. V takovém případě mohou být následky mnohem závažnější, jelikož mechanismus, který může neúmyslně (tj. bez vědomí člověka) znevýhodňovat některé skupiny, je mnohem obtížnější identifikovat a následně je mnohem složitější řešit případné důsledky. Za předpokladu, že případná diskriminace je neúmyslná, mohou i samotné pokusy o odstranění všech předvídaných a možných způsobů diskriminace ze strany osob zapojených do procesu těžení dat nepříznivě ovlivnit nestrannost výsledných rozhodnutí.

První záležitost, se kterou se vývojáři nových technologií potýkají, když programují model strojového učení, je stanovení, jakého cíle vlastně chtějí dosáhnout. Například společnost vydávající kreditní karty může chtít provádět analýzu bonity zákazníka. Aby bylo možné „bonitu zákazníka“ převést na něco, co lze spočítat, musí se společnost rozhodnout, zda chce, například, maximalizovat zisk nebo maximalizovat počet půjček, které budou splaceny. Bonitu zákazníka je poté třeba definovat v rámci tohoto cíle. Problém nastává v tom, že tato rozhodnutí jsou přijímána z různých obchodních důvodů, tj. z jiných důvodů než je spravedlivost nebo nediskriminace, vysvětluje Solon Barocas, docent na Cornell University, který se specializuje na spravedlivost v souvislosti se systémy strojového učení.<sup>146</sup> Pokud by například algoritmus strojového učení zjistil, že nabízení úvěrů méně bonitním uchazečům vede ke zvýšení zisku, preferoval by tuto možnost, přestože je v rozporu se zájmy společnosti.

Dle odborníků z MIT existují dva hlavní důvody projevu nespravedlivosti či zaujatosti v rámci tréninkových vstupních dat, a to buď ten, že data, která jsou shromažďována,

---

<sup>146</sup> BAROCAS, Solon a SELBST, Andrew D. *Big Data's Disparate Impact* (2016). 104 California Law Review 671 (2016) [cit. 20. 2. 2020]. Dostupné z: <https://ssrn.com/abstract=2477899> nebo <http://dx.doi.org/10.2139/ssrn.2477899>.

nerepresentují realitu, nebo odráží již existující předsudky.<sup>147</sup> Příkladem prvního scénáře může být algoritmus strojového učení, kterému bylo jako vstupní data poskytnuto větší množství fotografií tváří se světlou pletí než tváří s pletí tmavou. Výsledný systém určený k rozpoznávání obličejů (např. tzv. face recognition system) pak bude nevyhnutelně vykazovat horší výsledky při rozpoznávání obličejů s tmavou pletí. Druhý scénář již v několika známých případech nastal. Nejznámějším je případ chatbota, který byl vyvinut společností Microsoft Corporation v roce 2016 a spuštěn prostřednictvím sociální sítě Twitter.<sup>148</sup> Následná diskuse a pobouření vznikly, když chatbot začal přes svůj Twitter účet posílat rasistické a útočné tweety, což přimělo společnost Microsoft Corporation chatbota vypnout, a to jen 16 hodin po jeho spuštění. Podle všeho byla situace způsobena tím, že data, ze kterých se chatbot učil (tj. komentáře a tweety dalších uživatelů sítě Twitter), obsahovaly nežádoucí a škodlivý obsah.

Dalším příkladem toho, jak tréninková data vyvolala diskriminační přístup, nastal před několika lety, když se společnost Amazon pokusila použít umělou inteligenci k vytvoření nástroje pro analýzu životopisů.<sup>149</sup> Podle agentury Reuters společnost spoléhala na to, že tato technologie může zefektivnit proces nábory nových zaměstnanců. Jako vstupní data byla použita data shromážděná ze životopisů, které společnost Amazon obdržela v předcházejících letech. Výsledkem byl algoritmus, který preferoval muže a naučil se diskriminovat ženy, a to na základě skutečnosti, že ve společnosti Amazon byly historicky při náborech úspěšnější

---

<sup>147</sup> HAO, Karen. *This is how AI bias really happens—and why it's so hard to fix*. MIT Technology Review [cit. 20. 02. 2020]. Dostupné z: <https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>.

<sup>148</sup> WAKEFIELD, Jane. *Microsoft chatbot is taught to swear on Twitter*. BBC News [cit. 20. 02. 2020]. Dostupné z: <https://www.bbc.com/news/technology-35890188>.

<sup>149</sup> DASTIN, Jeffrey. *Amazon scraps secret AI recruiting tool that showed bias against women*. Business & Financial News, U.S & International Breaking News. Reuters [online] [cit. 20. 02. 2020]. Dostupné z: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

právě muži. Algoritmus strojového učení v tomto případě začal diskriminovat např. už na základě údaje, zda uchazeč absolvoval dívčí či chlapeckou školu.

Způsobit následnou diskriminaci či nerovné zacházení je možné i ve fázi přípravy algoritmu, která zahrnuje výběr atributů, které má algoritmus považovat za relevantní. V případě přípravy algoritmu pro posuzování bonity by mohl být vhodným atributem např. věk, příjem nebo počet splacených půjček uchazeče. V případě náborového algoritmu může být atributem úroveň dosaženého vzdělání nebo roky zkušeností uchazeče. Výběr toho, které atributy zvážit a které nepovažovat za relevantní, může výrazně ovlivnit přesnost predikce algoritmu strojového učení.

V kontextu zmíněného je patrné, že při programování algoritmů strojového učení existují určité výzvy, se kterými se vývojáři musí vypořádat, aby alespoň eliminovali riziko vzniku zaujetí, diskriminace či jiných nežádoucích efektů. Největším problémem je však skutečnost, že mnohdy není v této počáteční fázi zřejmé, že určitý atribut či charakter vstupních dat mohou vést ke zkresleným výsledkům. S tím se pojí i další výzva vývojářů, a to zjistit, zda a proč k takovému zkreslení došlo. Pokud není systém dostatečně transparentní, nemusí být dohledatelné, jak v procesu strojového učení algoritmus postupoval a zejména proč preferoval určitá rozhodnutí. Například v případě náborového algoritmu společnosti Amazon, jakmile začalo být zřejmé, že algoritmus preferuje uchazeče mužského pohlaví, programátoři jej přeprogramovali tak, aby ignoroval slova, která jsou genderově více spjata s ženami. To však nepomohlo a algoritmus i nadále na základě jiných kritérií porovnání historických životopisů došel do fáze, kdy diskriminoval uchazeče ženského pohlaví.

Další výzvou pro vývojáře jsou nedokonalé procesy strojového učení. Dle odborníků MIT mnoho standardních postupů v rámci strojového učení není navrženo tak, aby detekovaly



možná rizika zkreslení.<sup>150</sup> Algoritmy strojového učení jsou před jejich nasazením mnohokrát testovány, což by se mohlo zdát jako ideální příležitost k zachycení potenciálně vadného vzorce. V praxi však testování obvykle vypadá tak, že programátoři a vývojáři náhodně rozdělují vstupní data do dvou skupin. Do první skupiny jsou zařazena data, která se skutečně využívají k „výcviku“, a do druhé skupiny jsou zařazena data pro účely testování, jakmile je software dokončen. To znamená, že data, která jsou použita k testování výkonu systémů strojového učení, vykazují stejná zkreslení jako data, která byla využita k výcviku.

V neposlední řadě představují výzvu pro předcházení algoritmických předsudků proměnné, jako rozdílný sociální kontext a rozdílný pohled na to, co je a není spravedlivé. Zaprvé, tentýž algoritmus nemůže ideálně fungovat v různých sociálních kontextech. Jiné požadavky a důraz na spravedlnost a nediskriminaci budou kladeny na algoritmus určující délku trestu odnětí svobody a jiné v případě algoritmu propojujícího uživatele v aplikaci určené k seznamování, tj. v každém případě je nutné klást důraz na jiné společenské souvislosti. Zadruhé, nejen v každém státě, ale i v každé menší jednotce se pohled na to, co je a není spravedlivé, může významným způsobem lišit. Není proto možné jednoznačně stanovit, jaký je požadovaný výsledek, stejně tak není možné s jistotou určit, zda po určitém čase nebude pohled na tuto otázku i v dané posuzované komunitě úplně odlišný.

V současné době již probíhají výzkumy, jejichž cílem je jednak vyvinout algoritmy, které dokáží identifikovat diskriminační zacházení, tak i algoritmy, které by diskriminaci předcházely pomocí zvláštních technik těžení dat.<sup>151</sup>

---

<sup>150</sup> HAO, Karen. *This is how AI bias really happens—and why it's so hard to fix*. MIT Technology Review [cit. 20. 02. 2020]. Dostupné z: <https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>.

<sup>151</sup> HAJIAN, Sara, BONCHI, Francesco, CASTILLO, Carlos. *Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining*. KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA — August 13 - 17, 2016 [cit. 25. 2. 2020]. Dostupné z: <https://dl.acm.org/citation.cfm?doi=2939672.2945386>.

## 5.2 Etika strojů

Pokud mají být autonomní systémy plnohodnotnou součástí společnosti, musí operovat odpovědně. V budoucnu by měly v čím dál větší míře nahradit činnost lidí, a tím pádem v některých případech nahradit i jejich rozhodování. Některá rozhodnutí však představují více než např. pouhou aplikaci pravidel provozu na pozemních komunikacích či formalistickou aplikaci jiných právních předpisů, vyžadují určitý smysl pro etické uvažování, které se, jak vyplývá z níže uvedeného, jen těžko implementuje do konkrétních algoritmů chování.

Do oblasti etiky strojů spadá zejména otázka, zda a za jakých podmínek bychom měli umožnit umělé inteligenci samostatně rozhodovat o etických otázkách. Konkrétně to znamená: Mohou stroje činit morální rozhodnutí? Jak mohou být morální, společenské a právní hodnoty zakomponovány do procesu programování a vývoje? Jak a kdy by měla společnost případně zasáhnout?

Odpověď na tyto a související otázky vyžaduje zcela nové porozumění etice s ohledem na kontrolu a autonomii v měnící se socio-technické realitě. Implementace etických hodnot do strojů nám navíc pomůže celkově lépe porozumět etice, resp. její reflexi v současné společnosti. Aby bylo možné dosáhnout požadovaného technologického vývoje a pozitivních reakcí, budou výzkumní pracovníci a odborníci v oblasti umělé inteligence muset být při navrhování systémů umělé inteligence schopni vzít v úvahu stávající morální, společenské a právní hodnoty. Jinými slovy, výzkum a vývoj umělé inteligence vyžaduje, aby programátoři analyzovali uvedené hodnoty, převedli tyto hodnoty do formy technických požadavků a implementovali je do algoritmů, které může systém následně aplikovat. Současně, analýza etických a společenských důsledků jednání a rozhodnutí systémů umělé inteligence vyžaduje od programátorů a vývojářů mentální posun od zaměřování se na výkon právě k výše zmíněné snaze zajištění důvěry.

Zásadní rozdíl mezi etickými rozhodnutími činěnými lidmi a rozhodnutími systému umělé inteligence spočívá v hodnocení procesu rozhodování. Lidé jsou vůči sobě navzájem

a vůči společnosti odpovědny za morálnost svých rozhodnutí a jednání. Hodnocení rozhodovacího procesu se v takovém případě provádí, pokud vůbec, ex post. Lidé jsou zřídka, nejedná-li se o případy přičítání odpovědnosti, vyzváni vysvětlit a obhájit postup, který vedl k jejich rozhodnutí. K vysvětlitelnosti a zdůvodnění rozhodnutí stroje se však přístup liší. Děje se tak zejména z toho důvodu, že tento z velké části musí být rutinní a předem určený.

S tím souvisí i další rozdíl, který spočívá v tom, že lidé jsou standardně považováni za morální bytosti (tzv. moral agents), pokud není důvod o tom pochybovat či neexistují důkazy, které by prokazovaly opak. U strojů nelze o takovém předpokladu uvažovat, jelikož nelze bez dalšího uzavřít, že stroje jsou schopny morálního uvažování. Určitým způsobem, jak se k takovému předpokladu přiblížit, může být například zavedení mechanismu certifikování, které by v rámci možností zaručovalo určité vtělené principy, které je stroj schopen reprodukovat.

Zatímco u lidí je chybování považováno za běžné, není jasné, jak tolerantní bude společnost vůči inteligentním strojům, které učiní nesprávná rozhodnutí. Jedním z důvodů vzniku této obavy je i skutečnost, že ve vztahu k lidem se ve společnosti vytvořila určitá odůvodnění zmírňující negativní důsledky jejich jednání a rozhodnutí, jako jsou například nepřičetnost, rozptýlení nebo zmatení apod. Jedná se však o argumenty použitelné u lidí, nikoliv u strojů.

V neposlední řadě, v kontextu nových technologií, je nezbytné zmínit, že zatímco prakticky všechna rozhodnutí mohou být svým způsobem považována za etická a budou odrážet (a potenciálně pomůžou formovat) priority a hodnoty programátora či uživatelů, etika v uvedeném relevantním slova smyslu není v popředí většiny rozhodnutí, která budou autonomní systémy činit. Za striktně etická rozhodnutí lze v tomto kontextu považovat ta rozhodnutí, která se určitým způsobem vztahují k lidské důstojnosti a životním podmínkám nebo na ně mají přímý dopad. Obecně lze tedy konstatovat, že výslovně etické rozhodování u inteligentních strojů bude vyžadováno pouze za specifických okolností, a to buď proto, že systém bude muset vzít v potaz některá omezení či okolnosti předem pro něj označené jako

etické, nebo proto, že detekuje konflikt mezi etickými principy a musí učinit konečné rozhodnutí.<sup>152</sup> Problematiku etiky strojů lze nejlépe demonstrovat u autonomních vozidel, která jsou snadno představitelná.

V dopravním provozu budou nehody pravděpodobně vždy nevyhnutelnou součástí, ať již kvůli podmínkám provozu, špatné údržbě komunikací, počasí nebo ostatním účastníkům provozu. Autonomní vůz může čelit situacím, kdy se nebude možné vyhnout zranění osob nebo způsobení majetkové újmy. Lze si představit jednoduchý scénář – autonomní vůz musí učinit rozhodnutí, zda zraní 8leté dítě nebo 80letého důchodce. Která varianta bude z morálního hlediska správná? Existuje na tuto otázku správná odpověď? Argumenty lze nalézt na obou stranách, řidič-fyzická osoba by pravděpodobně reagoval reflexivně dopředu neznámým a neurčitým způsobem. Výsledek rozhodnutí programu o zranění osob na základě jejich věku (v situacích, kdy věk není relevantním kritériem) však nikdy nebude morálně korektní. Pokud nejsme schopni posoudit, která varianta je přijatelnější, jak má vypadat algoritmus vtělený do autonomního systému?

Komplexitu a náročnost rozhodovacích mechanismů lze demonstrovat i na jiných příkladech, kde rozlišovacím prvkem není věk. Autonomní vůz srazí buďto cyklistu s helmou nebo cyklistu bez helmy. Lze konstatovat, že cyklista s helmou má větší šanci na přežití, jelikož je lépe chráněn. Bylo by ale morálně ospravedlnitelné „potrestat“ cyklistu, protože dodržuje podmínky provozu na pozemních komunikacích, a ochránit toho, kdo podmínky porušuje?

Pokud jsou do obdobných scénářů zahrnuti i cestující uvnitř autonomního vozu, situace se více komplikuje. Pokud bude systém nastaven tak, aby v co největší míře ochránil své cestující, pravděpodobně by zvolil možnost narazit do „nejlehčího“ předmětu (tj. spíše do

---

<sup>152</sup> DENNIS, L., FISHER, M., SLAVKOVIK, M. and WEBSTER, M. *Formal verification of ethical choices in autonomous systems*. *Robotics and Autonomous Systems* 77 (2016), 1–14 [cit. 25. 2. 2020]. Dostupné z: <https://doi.org/10.1016/j.robot.2015.11.012>.

8letého dítěte). Je ale správné chránit cestující vozu, když právě oni vystavili ostatní riziku, když se rozhodli pro jízdu autonomním vozem? Obětování cestujících se může z hlediska etiky zdát jako nejlepší volba. Vzhledem k výše uvedenému je však z hlediska společenské přijatelnosti rozdíl v tom, když toto rozhodnutí učiní řidič, který jedná reflexivně, nebo když tak rozhodne autonomní vůz za řidiče na základě alespoň částečně naprogramovaných vzorců.

Ať bude implementován jakýkoliv algoritmus, následky jednání autonomních systémů budou posuzovány nejen soudy v rámci civilního řízení o žalobě na náhradu škody (či jiného řízení dle okolností případu) a statistickými úřady, ale také širokou veřejností. Důsledkem toho může dojít ke ztrátě důvěry společnosti a nevoli vůči akceptaci autonomních systémů. Otázkou totiž je, zda může existovat algoritmus, který bude přijatelný pro většinu společnosti?

Problém etiky je populárním tématem mnoha filosofů. Asi nejznámějším případem je tramvajové dilema britské filosofky Philippy Foot, která se zabývala příkladem řidiče tramvaje jedoucí po kolejích na 5 pracovníků s možností zatáhnout za páku, změnit směr a najet na kolej s jedním pracovníkem.<sup>153</sup> Tato situace nemá žádné správné řešení, způsob, jakým je možné se s dilematem vypořádat, vyplývá z mravních principů, které na případ aplikujeme. Philippa Foot ve svém článku vymezila možnost aplikace dvou mravních principů - utilitarismu a deontologie.

*„Utilitarians believe that the purpose of morality is to make life better by increasing the amount of good things (such as pleasure and happiness) in the world and decreasing the amount of bad things (such as pain and unhappiness).“<sup>154</sup>*

---

<sup>153</sup> FOOT, Philippa. *The Problem of Abortion and the Doctrine of the Double Effect*. Oxford Review. 1967, čís. 5 [cit. 19. 5. 2019]. Dostupné z: <http://www2.pitt.edu/~mthomps/readings/foot.pdf>.

<sup>154</sup> Internet Encyclopedia of Philosophy (online). Dostupné z: <https://www.iep.utm.edu/util-a-r/>.

Utilitaristé se domnívají, že nejetichtější způsobem jednání je takový způsob, který umožní největší užitek největšímu počtu lidí. Utilitaristé by se přikláněli k etickému vzorci, který by porovnal důsledky všech alternativ jednání a za správnou vyhodnotil tu, jejíž konsekvence jsou minimálně tak užitečné, jako důsledky alternativních jednání.<sup>155</sup> V kontextu inteligentních robotů, a konkrétně autonomních vozů, by utilitaristické řešení spočívalo v takovém jednání automobilu, při kterém dojde ke zranění nejmenšího počtu osob.

Deontologie představuje přístup v etice, který se na rozdíl od utilitarismu zaměřuje na správnost nebo nesprávnost samotné akce, nikoliv na správnost nebo nesprávnost jejích důsledků. To, co dělá jednání správným, je jeho soulad s morálními pravidly. Podle deontologů by autonomní vůz v první řadě neměl ublížit člověku, nebo jakkoli umožnit, aby mu bylo ublíženo, a v případě nehody by měl přednostně ochránit osoby cestující ve vozidle.

Lze se domnívat, že není správné vytvářet vzorec chování inteligentního robota, který bude jasně determinovat, jaký počet osob má ochránit, zda má dát přednost dětem či důchodcům, ženám či mužům. Některé instituce se tento problém snaží vyřešit prostřednictvím online formulářů jakožto průzkumů veřejného mínění pro stanovení určitých záchytných bodů při normování etického vzorce. Konkrétně vědci z MIT sestavili online dotazník s názvem „Moral Machine“,<sup>156</sup> kde mohou návštěvníci stránek dle vlastního mínění hodnotit předem připravené scénáře z provozu autonomních vozů. Takový přístup nemusí být ideální, jelikož průzkum veřejného mínění pravděpodobně odhalí podobné výsledky – lidé by dali přednost ochraně většího počtu osob. Jenomže morální problém tkví také v tom, že pokud programátor autonomního systému implementuje takový vzorec, nastanou situace, kdy automobil učiní aktivní kroky k ochraně, a tím i aktivní krok k ohrožení osob, což může být společensky nepřijatelné. V tomto kontextu se lze tedy domnívat, že spíše než se zaměřit na filosofické směry a otázku, koho ochránit či neochránit, je důležité zpočátku stanovit základní

---

<sup>155</sup> ARKIN, Ronal C. *Governing lethal behaviour in autonomous robots*. ISBN 978-1-4200-8594-5.

<sup>156</sup> Dotazník dostupný na: <http://moralmachine.mit.edu>.

principy jednání inteligentních robotů, které budou společné pro všechny jejich formy a nebudou řešit konkrétní scénáře.

Jelikož, jak jsem zmínila výše, na autonomní systémy v řadě případů nahlížíme jako na inteligentního robota, můžeme se podívat na některé návrhy, přístupy, které k otázce etického chování robotů existují. Moderní věda o robotech začíná u Isaaca Asimova, který definoval tři velice jednoduché zákony robotiky, jejichž cílem je sladění robotů s lidskými hodnotami.<sup>157</sup>

- 1. Robot nesmí ublížit člověku nebo svou nečinností dopustit, aby bylo člověku ublíženo.*
- 2. Robot musí uposlechnout příkazy člověka, kromě případů, kdy jsou tyto příkazy v rozporu s prvním zákonem.*
- 3. Robot musí chránit sám sebe před poškozením, kromě případů, kdy je tato ochrana v rozporu s prvním nebo druhým zákonem.*

I Evropský parlament ve svém usnesení<sup>158</sup> odkazuje na základní Asimovy zákony, které by měly být určeny konstruktérům, výrobcům a provozovatelům robotů. Asimovy zákony sice zdaleka nejsou dostatečné, jelikož takto jednoduše naprogramovaný robot by nebyl veřejností uznán za plně morálního, jejich kouzlo ale tkví v tom, že Asimov se nezabýval otázkou, zda se jednoho dne inteligentní roboti neobrátní proti lidem, nebo zda má robot ochránit jednu nebo pět osob. Samozřejmě, obecná pravidla nelze použít plošně, například autonomní vozy jakožto inteligentní roboti budou naprogramování odlišně od jiných druhů autonomních systémů, jelikož jejich primárním účelem je provoz na pozemních komunikacích.

---

<sup>157</sup> ASIMOV, Isaac. *Robohistorie I*. Triton, 2004, 432 s. Trifid. ISBN 80-7254-477-2.

<sup>158</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

Stanfordský profesor filozofie Ken Taylor si např. klade otázku, zda vůbec můžeme od umělé inteligence očekávat, že se vyrovná člověku, resp. nahradí ho jako mravní osobnost, která je schopna činit mravní rozhodnutí.<sup>159</sup> Je zjevné, že odpověď na tuto a další položené otázky nebude v dohledné době ani mezi akademiky ustálena a práce programátorů tak bude velmi náročná. Jelikož je jisté, že určitá pravidla pro vývoj, výrobu a užívání nových technologií jsou potřebná, probíhají debaty a práce na jejich vytvoření na poli mnoha organizací. Jedním z příkladů je usnesení Evropského parlamentu na vytvoření určitého kodexu pravidel chování (charta robotiky) osob zainteresovaných v robotickém inženýrství, vycházející mj. z principů obsažených ve Smlouvě o Evropské unii a Listině základních práv Evropské unie.

### 5.3 Charta inženýrů robotiky

Příloha k usnesení Evropského parlamentu<sup>160</sup> obsahuje Kodex etického chování inženýrů robotiky, který vybízí projektanty a vývojáře inteligentních robotů, aby jednali zodpovědně, při své práci ctili lidskou důstojnost, soukromí a bezpečnost a spolupracovali napříč odvětvími. Evropský parlament se domnívá, že současný právní rámec Unie je třeba modernizovat a doplnit o etické zásady, které budou odrážet složitost problematiky inteligentních robotů.

Kodex vyzývá projektanty a vývojáře, aby při své práci vzali v potaz i budoucí očekávaný vývoj v oblasti robotiky a budoucí dopady těchto technologií. Základními zásadami, na kterých je kodex postaven, jsou prospěšnost, neškodlivost, autonomie a spravedlivost. Evropský parlament zdůrazňuje, že inteligentní roboti by měli jednat

---

<sup>159</sup> SHASHKEVICH, Alex. *Stanford scholars, researchers discuss key ethical questions self-driving cars present*. Stanford News [cit. 20. 02. 2020]. Dostupné z: <http://news.stanford.edu/2017/05/22/stanford-scholars-researchers-discuss-key-ethical-questions-self-driving-cars-present/>.

<sup>160</sup> Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komisi o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).



v nejlepším zájmu lidí, zbytečně jim nezpůsobovat újmu, to vše za předpokladu, že bude umožněna spravedlivá distribuce přínosů spojených s robotikou a cenová dostupnost.

Kodex je koncipován tak, že vyzývá k předběžné opatrnosti za současné podpory budoucího pokroku. Potenciál pokroku v oblasti robotiky je třeba vnímat na pozadí různých rizik a potenciálních bezpečnostních dopadů (a to i v oblasti svobody a ochrany osobních údajů) a za přijetí vhodných preventivních opatření eliminovat možná rizika (např. přísným dodržováním bezpečnostních směrnic při výzkumných pracích). Důraz je kladen na soukromí fyzických osob. Inženýr robotiky by měl zaručit, že soukromé informace o fyzických osobách budou zabezpečeny a využívány pouze řádným způsobem.

I přesto, že inteligentní roboti se v mnohém odlišují od člověka, mělo být vždy možné racionálně odůvodnit jednotlivé rozhodnutí inteligentního robota, tj. jejich jednání by v souladu se zásadou transparentnosti mělo být vždy pro člověka pochopitelné.

Kodex má zatím pouze povahu „soft law“, není závazný, dokonce je výslovně označen za dobrovolný a pouze doplňující text. V budoucnu však může vzhledem ke své obecnosti (věcná působnost kodexu je velice široká, přesto, že osobní působnost je omezena výlučně na inženýry robotiky, nikoliv potenciální vlastníky) sloužit jako podklad pro vnitrostátní i unijní zákonodárce, ale i pro další odvětvové kodexy inženýrů robotiky.

## **5.4 Doporučení Etické komise**

Zatímco usnesení Evropského parlamentu je potenciálně aplikovatelné velmi široce, na poli různých jurisdikcí či organizací vznikají i úzce zaměřené kodexy či doporučení. Příkladem jsou obecná etická doporučení pro automatizované a propojené dopravní prostředky, která pro německé ministerstvo dopravy a digitální infrastruktury připravila

skupina 14 odborníků (etická komise) a pokusila se tak sestavit první etický kodex pro autonomní vozy.<sup>161</sup>

Etická komise vychází z předpokladu, že na konci vývoje autonomních vozů bude provoz na pozemních komunikacích mnohem bezpečnější, nicméně vzhledem k úrovni autonomie, která v provozu existuje dnes, nebude po nějakou dobu možné úplně zabránit nehodám. Proto je důležité, aby byla při programování softwaru implementovaného do autonomních vozů zohledněna některá etická pravidla.

Etická komise zdůrazňuje, že technologický pokrok v tomto ohledu nutí vládu i širokou veřejnost, aby současně reagovaly na vznikající změny. Etická komise upozorňuje, že pokud budou autonomní systémy licencovány státem, bude třeba správně nastavit podmínky, ve kterých budou tyto systémy používány, aby bylo možné dosáhnout stavu, kdy bude zajištěna potřebná ovladatelnost a transparentnost systému.

Etický kodex stojí zejména na zásadě „pozitivní rovnováhy rizik“. Ochrana jednotlivců musí mít přednost před jinými úvahami a cílem autonomního systému musí být eliminace škod, dokud nebude možné škody úplně vyloučit. Udělení licence autonomním systémům by nebylo legitimní, pokud by autonomní řízení nebylo provázeno příslibem bezpečnější budoucí dopravy. Systém musí být nastaven tak, aby kritické situace nenastávaly, byť s vědomím toho, že takovým situacím není možné stoprocentně zabránit.

Etický kodex vychází z existujícího způsobu nastavení právních systémů tak, že ochrana lidského života představuje nejvyšší prioritu při vyvažování zájmů chráněných právem. Proto v rámci toho, co je technologicky proveditelné, musí být systémy naprogramovány tak, aby akceptovaly zranění zvířat nebo poškození majetku, pokud to znamená, že je možné zabránit zranění fyzických osob. Situace, kdy musí být učiněno

---

<sup>161</sup> Zpráva Etické komise: Automatizované a propojené řízení, červen 2017 [cit. 19. 2. 2020]. Dostupné z: [https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile)

rozhodnutí ohledně ochrany několika lidských životů, bude záviset na konkrétní situaci společně s ad hoc posouzením na základě chování dotčených stran. Etická komise zdůrazňuje, že takové jednání systému nemůže být standardizované ani nemůže být systém naprogramován tak, aby bylo jeho jednání eticky nezpochybnitelné. Autonomní systém nemůže jednat intuitivně a nahradit rozhodování člověka. Řidič-člověk by v takových scénářích mohl jednat protiprávně, kdyby dal přednost životu jedné osoby před životem druhé, toto posouzení však přichází ex post ze strany orgánů veřejné moci při posouzení relevantních okolností konkrétního případu. Není proto jednoduché implementovat plošně akceptovatelná pravidla ex ante. Etická komise v tomto ohledu nabízí možné řešení - zřízení nezávislé veřejné organizace (např. úřad pro vyšetřování nehod spojených s autonomními dopravními systémy), která by systematicky zpracovala závěry ze soudní praxe získané mnoho let dozadu do uceleného výstupu, který by sloužil jako vodítko pro programátory.

V případě nevyhnutelných nehod je podle kodexu přísně zakázáno jakékoli posuzování rozdílů založených na osobních rysech dotčených osob (např. věk, pohlaví, fyziognomické rysy), a to jak individuálně, tak ve vzájemném porovnání.

Etická komise je takového názoru, že i za účelem omezení odpovědnosti za škodu způsobenou autonomním systémem jsou výrobci nebo provozovatelé automobilů povinni optimalizovat svá řešení pro nové systémy a průběžně sledovat a zlepšovat staré systémy tam, kde je to technologicky možné. Spotřebitelé by měli být co nejtransparentnějším způsobem informováni o nových technologiích, jejich implementaci a uplatňování standardizovaných zásad chování, které budou licencovány státem nebo nadnárodní autoritou a v ideálním případě přezkoumány nezávislou entitou. Správné používání autonomních systémů by mělo být součástí širšího vzdělávání společnosti o digitálním a propojeném světě a za tím účelem vhodným způsobem vyučováno a testováno.

## Závěr

Vývoj autonomních systémů a zejména umělé inteligence má před sebou ještě dlouhou cestu spolu s dalšími technologiemi, které postupně přetvářejí svět na prostředí, které je v mnoha ohledech propojené a automatizované. Lze konstatovat, že každá technologie má své výhody a nevýhody, ale pokud se má udržet na trhu a lákat spotřebitele, výhody musí vždy převážet nad nevýhodami.

I autonomní systémy slibují v různých oborech a oblastech nespočet přínosů, aby však společnost přijala jejich čím dál širší využití, musí si k nim vybudovat patřičnou důvěru. Taková důvěra bude podpořena transparentností systémů, legitimním očekáváním z pohledu odpovědnosti za škodu způsobenou autonomním systémem a legitimním očekáváním z hlediska předvídatelnosti chování autonomních systémů. Jak je patrné z předchozího textu, není práce výrobců a programátorů jednoduchá, z toho důvodu vznikají zmíněná doporučení a kodexy, které si kladou za cíl podpořit technologický vývoj a usnadnit řešení společensko-etických dilemat nabízením svých pohledů na problematiku. Aby mohl vývoj autonomních systémů nerušeně pokračovat, musí jednotlivé státy (či v našem případě i Evropská unie) zaujmout konkrétní postoj a zvolit přiměřené nástroje pro vytvoření prostředí, kde budou jasně dána pravidla určení odpovědnosti za jednání inteligentních robotů a závazné kodexy etických pravidel pro programátory autonomních systémů.

Aby bylo možné udržet požadovanou míru důvěry v nové technologie, je alespoň do určité míry nutné neztratit kontrolu. Je sice otázkou, zda je požadavek neustálé kontroly naplnitelný, případně nakolik je kompatibilní s požadavkem umožnění technického pokroku, ale je jisté, že od tohoto požadavku nelze upustit. Problémem může být jak skutečnost, že pokud dojde ke ztrátě kontroly, může být obtížné ji získat zpět, tak to, že se ztrátou kontroly může být snížena i schopnost auditovatelnosti systému nebo související požadavek neustálého aktualizování, aby systém odpovídal novým trendům a požadavkům. Možnost, jak eliminovat případy nežádoucího jednání, i když je determinováno okolnostmi, které nastanou po naprogramování a nastartování stroje, je, i v souladu s prevenční povinností, správně nastavený vzorec alespoň základních principů chování, které jsou společné pro všechny

autonomní systémy. V souvislosti s prevenční povinností se tak jako základní doporučení jeví zejména vyhotovení výkladových pravidel rozsahu prevenční povinnosti ve vztahu k jednotlivým technologiím či jednotlivým oblastem využití technologií, zpracování údajů i testování inteligentních systémů.

Aby mohla právní úprava správně reagovat na výzvy představované pokrokem v oblasti výzkumu umělé inteligence, je třeba předně stanovit její povahu. Byť existují názory, že by bylo možné umělé inteligenci přiznat postavení plnohodnotné osoby v právním řádu, neboli přidělit ji právní osobnost, s tímto postupem je v současné chvíli spojeno spousta neznámých, a proto nepředstavuje nejvhodnější řešení. V kontextu aplikovatelné právní úpravy lze umělou inteligenci posuzovat jako věc v právním smyslu, případně výrobek, nebo v určitých případech konkrétněji jako počítačový program. Toto posouzení následně i determinuje, jaký model a režim odpovědnosti bude ideálně požit.

K otázce odpovědnosti za jednání umělé inteligence již bylo vedeno nespočet debat, a byť ideálním řešením by bylo přijetí zvláštního ustanovení upravujícího odpovědnost za jednání umělé inteligence, takový krok zatím nebyl podniknut. Vzhledem ke skutečnosti, že v dohledné budoucnosti a v určitých případech by měla umělá inteligence operovat stále pod dohledem člověka, a jelikož nebude možné vždy dovést a prokázat zavinění fyzické osoby, včetně příčinné souvislosti mezi zaviněním, jednáním a vzniklou škodou, případně dokonce zavinění umělé inteligence, jeví se jako nejvhodnější přístup založený na objektivní odpovědnosti a v kontextu účinné právní úpravy tak přichází v úvahu speciální skutkové podstaty zakládající objektivní odpovědnost. Ve vztahu k přiřítání odpovědnosti za jednání autonomních systémů, konkrétně umělé inteligence, lze tak v kontextu účinné právní úpravy uvažovat o odpovědnosti za škodu způsobenou provozem zvláště nebezpečným, škodu, kterou způsobí věc sama od sebe, škodu z provozu dopravních prostředků. V případě většiny autonomních systémů lze pak uvažovat i o využití těch ustanovení občanského zákoníku, která se vztahují ke škodě způsobené vadou výrobku.

S cílem otevřít cestu budoucímu výzkumu a vývoji tato práce upozorňuje také na některé problémy, které vyvstávají při zvažování etických aspektů při navrhování

autonomních systémů. S umělou inteligencí je bezesporu spojeno velké množství etických otázek, které lze obecně rozdělit do dvou kategorií – na tzv. roboetiku a na etiku strojů, které hradí důležitou roli na cestě k budování dobře fungující a důvěryhodné umělé inteligence.

V rámci roboetiky je důležité se vypořádat s otázkami dohledu, transparentnosti, bezpečnosti, ochrany soukromí a rizika vzniku předsudků a případů nerovného zacházení a diskriminace. Některé konkrétní vlastnosti nových technologií, včetně neprůhlednosti („black box effect“), nepředvídatelnosti a autonomního chování, vedou v praxi k tomu, že jejich provoz může být netransparentní a neověřitelný z hlediska dodržování předpisů, což následně brání účinnému prosazování ať už právních či soft law pravidel, jejichž cílem je poskytnout legitimně očekávatelnou úroveň ochrany základních práv. Pokud odhlédneme od uživatelů, pro účely řádného uplatňování případných právních nároků, musí praktickými prostředky pro ověření postupů a procesů v rámci umělé inteligence disponovat i orgány veřejné moci.

Nové technologie rovněž představují vysoké riziko nepřiměřeného zásahu do soukromí. V tomto ohledu je nezbytné dbát na to, aby byly bez výjimky dodržovány principy ochrany osobních údajů, aby byl stanoven určitý, výslovně vyjádřený a legitimní účel, pro který jsou údaje shromažďovány, a aby byly údaje přiměřené, relevantní a omezené na nezbytný rozsah ve vztahu k účelu, pro který jsou zpracovávány.

Poslední dobou často akcentované riziko vzniku diskriminace a předsudků je podpořeno tím, že takový vzorec může být vytvořen nejen při aktivním konání umělé inteligence, ale už i ve fázi přípravy algoritmu, která zahrnuje výběr atributů, které má algoritmus považovat za relevantní. V současné době již probíhají výzkumy, jejichž cílem je jednak vyvinout algoritmy, které dokáží identifikovat diskriminační zacházení, tak i algoritmy, které by diskriminaci předcházely pomocí zvláštních technik těžení dat, v každém případě by však mohlo pomoci vyvíjet nové technologie transparentněji a případně v maximální možné míře spolupracovat tak, aby bylo možné se poučit z chyb a nedostatků jiných vývojářů, kteří mohou být v pokroku mnohem dál.

Do oblasti etiky strojů pak spadá zejména otázka, zda a za jakých podmínek bychom měli umožnit umělé inteligenci samostatně rozhodovat o etických otázkách. Zásadní rozdíl mezi etickými rozhodnutími činěnými lidmi a rozhodnutími systému umělé inteligence totiž spočívá v hodnocení procesu rozhodování. Lidé jsou vůči sobě navzájem a vůči společnosti odpovědní za morálnost svých rozhodnutí a jednání, zatímco stroje nejsou schopny morálního uvažování a jejich chování musí být z velké části rutinní a předem naprogramované.

Přestože se zdá, že není správné dopředu vytvářet vzorec chování inteligentního robota, který bude jasně determinovat, jaký počet osob má ochránit, či jakým skupinám obyvatel dát přednost, využívají se v kontextu etického chování robotů dva mravní principy - utilitarismu a deontologie. Zatímco utilitaristé se domnívají, že nejetichtějším způsobem jednání je takový způsob, který umožní největší užitek největšímu počtu lidí, deontologové se zaměřují na správnost (či nesprávnost) samotné akce, nikoliv důsledků, a tedy její soulad s morálními pravidly. Některé instituce se pak tento problém snaží vyřešit prostřednictvím průzkumů veřejného mínění. Na poli mnoha organizací pak probíhají debaty a práce na vytvoření určitých vodítek, která doporučují například určení několika etických zásad, která představují široce uznávaný minimální požadavek morálního jednání, a jejich implementaci do zdrojového kódu programu, což lze považovat za vysoce žádoucí.

V kontextu nezastavitelného pokroku lze tak shrnout, že s ohledem na udržení důvěry uživatelů, kontroly orgánů veřejné moci a omezení negativních důsledků zásahu do soukromí potřeba zvýšit důraz na transparentnost a podpořit evropskou spolupráci zejména v následujících oblastech: (i) transparentní uplatňování bezpečnostních požadavků na vývoj umělé inteligence; (ii) vyvíjení metod pro zlepšování aplikací v oblasti bezpečnosti a ochrany soukromí; (iii) přijetí pokynů pro transparentní testování nových technologií; (iv) vzdělávání veřejnosti ohledně práce s novými technologiemi; a (v) úprava právních předpisů a podpora vytváření kodexů chování v oblasti odpovědnosti za jednání autonomních systémů, etiky a ochrany soukromí.

Role legislativců a odborných organizací není jednoduchá, ale nedostatek jasných pokynů pro účely řešení alespoň předvídatelných rizik může, samozřejmě vyjma rizika vzniku

újmou, představovat také právní nejistotu pro podnikatele, vývojáře, výrobce, programátory, ať už se na vývoji a implementaci umělé inteligence podílejí jakýmkoliv způsobem. Tento problém pak může přetrvat i v případě, že případná právní úprava nebude harmonizována v rámci Evropské unie.



## Seznam použitých zdrojů

### Seznam použité literatury

- ARKIN, Ronal C. *Governing lethal behaviour in autonomous robots*. ISBN 978-1-4200-8594-5.
- ASIMOV, Isaac. *Robohistorie I*. Triton, 2004, 432 s. Trifid. ISBN 80-7254-477-2.
- BERAN, Karel. *Pojem osoby v právu: (osoba, morální osoba, právnická osoba)*. Praha: Leges, 2012. Teoretik. ISBN 978-80-87576-06-9.
- BURGAT, Florence. *Svoboda a neklid zvířecího života*. Přeložila Olga SMOLOVÁ. Praha: Univerzita Karlova, nakladatelství Karolinum, 2018. Myšlení současnosti. ISBN 978-80-246-4056-3.
- DESCARTES, René. *Meditace o první filosofii: námitky a autorovy odpovědi*. Přeložil Tomáš MARVAN, přeložil Petr GLOMBÍČEK, přeložil Pavel ZAVADIL. Praha: OIKOYMENH, 2003. ISBN 80-7298-084-X.
- DVOŘÁK, Jan, ŠVESTKA, Jiří, ZUKLÍNOVÁ, Michaela a kol. *Občanské právo hmotné*. Díl první, Obecná část. 2., aktualizované a doplněné vydání. Praha: Wolters Kluwer, 2016. ISBN 978-80-7552-187-3.
- FUJITA, M. AIBO: *Toward the era of digital creatures*. The International Journal of Robotics Research, 20(10):781–794, 2001.
- GERLOCH, Aleš. *Teorie práva*. 6., aktualiz. vyd. Plzeň: Vydavatelství a nakladatelství Aleš Čeněk, 2013. ISBN 978-80-7380-454-1.
- GUERRA-PUJOL, F. E. *The Turing Test and the Legal Process* (December 31, 2011). Information & Communications Technology Law, vol. 21, no. 2 (June 2012), pp. 113-126.
- HASSELBALCH, G. (2019). *Making sense of data ethics. The powers behind the data ethics debate in European policymaking*. Internet Policy Review, 8(2). DOI: 10.14763/2019.2.1401.
- HULMÁK, M. a kol. *Občanský zákoník VI. Závazkové právo. Zvláštní část (§ 2055–3014)*. Komentář. 1. vydání. Praha: C. H. Beck, 2014. ISBN 978-80-7400-287-8.
- KAPLAN, Jerry. *Artificial Intelligence, What Everyone Needs to Know*. Oxford: Oxford University Press, 2016. ISBN 9780190602383.
- KRAUSOVÁ, Alžběta. *Legal Regulation of Artificial Beings*. Masaryk University Journal of Law and Technology, Brno: Masarykova Univerzita, 2007, roč. 1, č. 1, s. 187-198. ISSN 1802-5943.

- KRAUSOVÁ, Alžběta. *Status elektronické osoby v evropském právu v kontextu českého práva. Právní rozhledy*, 20/2017, s. 700-704.
- LAVICKÝ, P. a kol.: *Občanský zákoník I. Obecná část (§ 1–654). Komentář*. 1. vydání, Praha: C. H. Beck, 2014, 2400 s.
- LEM, Stanisław a WEIGEL, Pavel. *Tajemství čínského pokoje*. Praha: Mladá fronta, 1999. Kolumbus. ISBN 80-204-0826-6.
- LICHOCKI, P., KAHN, P. Jr., a BILLARD, A. *The Ethical Landscape of Robotics*. IEEE Robotics and Automation Magazine, 18(1):39-50, 2011.
- MOOR, James. *The Turing test: the elusive standard of artificial intelligence*. Boston: Kluwer Academic Publishers, c2003. ISBN 1402012047.
- MORGAN, C. (2019). *Responsible AI: a global policy framework*. McLean, VA, USA: ITechLaw. ISBN 978-1-7339931-0-4.
- OSINA, Petr. *Teorie práva*. Praha: Leges, 2013. Student (Leges). ISBN 978-80-87576-65-6.
- SCHERER, M. U. *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*. Harvard Journal of Law & Technology, jaro 2016.
- STEPHENS-DAVIDOWITZ, Seth. *Všichni lžou: velká data, chytré algoritmy a jak nám internet může říct, kdo opravdu jsme*. Přeložila Romana HEGEDŮSOVÁ. Brno: Host, 2019. ISBN 978-80-7577-787-4.
- SZTEFEK, Martin. *Kritické poznámky k prevenční povinnosti v občanském zákoníku*. Právní rozhledy. 2018, č. 2.
- TELEC, I., TŮMA, P. *Autorský zákon, Komentář*. 2. vydání. Praha: C. H. Beck, 2019, 1295 s.
- TICHÝ, Luboš a HRÁDEK, Jiří. *Deliktní právo*. V Praze: C.H. Beck, 2017. Beckova edice právní instituty. ISBN 978-80-7400-625-8.
- TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*. Praha: Togga, 2014. Scholia (Togga). ISBN 978-80-7476-043-3.
- VOJTEK, Petr. *Občanský zákoník: komentář*. 2. vydání. Praha: Wolters Kluwer, 2019-. Komentáře (Wolters Kluwer ČR). ISBN 978-80-7598-412-8.
- VOLNÁ, Eva a KOTYRBA, Martin. *Umělá inteligence*. Ostrava: Ostravská univerzita v Ostravě, 2013. ISBN 978-80-7464-330-9.

### **Seznam použitých internetových zdrojů**

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Dostupné z: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

*A Robot-Companion for Senior People and Patients with Alzheimer's Disease*, dostupné z: <http://endurancerobots.com/azbnmaterial/a-robot-companion-for-senior-people-and-patients-with-alzheimer-s-disease/>.

ATABEKOV, A., YASTREBOV, O. *Legal Status of Artificial Intelligence Across Countries: Legislation on the Move*. European Research Studies Journal Volume XXI, vydání 4, 773 – 782. Dostupné z: <https://www.ersj.eu/journal/1245>.

BAROCAS, Solon a SELBST, Andrew D. *Big Data's Disparate Impact* (2016). 104 California Law Review 671 (2016). Dostupné z: <https://ssrn.com/abstract=2477899> nebo <http://dx.doi.org/10.2139/ssrn.2477899>.

Budoucnost práce: robotika - Revue pro sociální politiku a výzkum. Revue pro sociální politiku a výzkum [online]. Institut pro sociální politiku a výzkum z. s.. Dostupné z: <https://socialnipolitika.eu/2017/01/budoucnost-prace-robotika/>.

BUOLAMWINI, Joy, GEBRU, Timnit. „*Proceedings of the 1st Conference on Fairness, Accountability and Transparency*“, PMLR 81:77-91, 2018.

BURROWS, Leah. *An economy of algorithms*. Harvard John A. Paulson School of Engineering and Applied Sciences. Dostupné z: <https://www.seas.harvard.edu/news/2017/01/economy-algorithms>.

CROWE, S. *10 Biggest Challenges in Robotics*. Science Robotics. Dostupné z: <https://www.therobotreport.com/10-biggest-challenges-in-robotics/>.

CUTHBERTSON, Anthony. *Tokyo: Artificial Intelligence 'Boy' Shibuya Mirai Becomes World's First AI Bot to Be Granted Residency*. Newsweek - News, Analysis, Politics, Business, Technology. Dostupné z: <https://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382>.

ČERKA, Paulius, GRIGIENĖ, Jurgita, SIRBIKYTĖ, Gintarė. *Liability for Damages Caused by Artificial Intelligence*. Computer Law & Security Review 31, 2015.

DASTIN, Jeffrey. *Amazon scraps secret AI recruiting tool that showed bias against women*. Business & Financial News, U.S & International Breaking News. Reuters [online]. Dostupné z: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

DENNIS, L., FISHER, M., SLAVKOVİK, M. and WEBSTER, M. *Formal verification of ethical choices in autonomous systems*. Robotics and Autonomous Systems 77 (2016). Dostupné z: <https://doi.org/10.1016/j.robot.2015.11.012>.

DIBLÍK, Jan, CHOLASTA, Roman, TADEVOSJANOVÁ, Laura. Robotika a regulace: „Privacy by design“ a transparentní přístup k novým technologiím. Národní centrum

průmyslu 4.0, březen 2020 [cit. 10. 5. 2020]. Dostupné z: <https://www.ncp40.cz/aktuality/technologie-a-pravo-privacy-by-design>.

FOOT, Philippa. *The Problem of Abortion and the Doctrine of the Double Effect*. Oxford Review. 1967, čís. 5. Dostupné z: <http://www2.pitt.edu/~mthomps/readings/foot.pdf>.

FRUHLINGER, Josh. *Marriott data breach FAQ: How did it happen and what was the impact?* Security news, features and analysis about prevention, protection and business innovation. [online]. [cit. 10. 5. 2020]. Dostupné z: <https://www.csoonline.com/article/3441220/marriott-data-breach-faq-how-did-it-happen-and-what-was-the-impact.html>.

*Google Self-Driving Car Project Monthly Report*. GOOGLE, únor 2016. Dostupné z: <http://dl.icdst.org/pdfs/files1/6786717558639e653aae406005326fb3.pdf>.

GRIFFIN, Andrew. *Facebook's artificial intelligence robots shut down after they start talking to each other in their own language*. The Independent. UK and Worldwide News. Dostupné z: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>.

GUILLOT, Craig, *4 types of autonomous mobile robots, and their warehouse use cases*. Supply Chain Dive. Supply Chain News and Analysis. Dostupné z: <https://www.supplychaindive.com/news/4-types-of-autonomous-mobile-robots-and-their-warehouse-use-cases/529548/>.

HAJIAN, Sara, BONCHI, Francesco, CASTILLO, Carlos. *Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining*. KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA — August 13 - 17, 2016. Dostupné z: <https://dl.acm.org/citation.cfm?doid=2939672.2945386>.

HAO, Karen. *This is how AI bias really happens—and why it's so hard to fix*. MIT Technology Review. Dostupné z: <https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>.

*How Much Data Is Generated Every Minute?* Social Media Today. Social Media News [online]. Dostupné z: <https://www.socialmediatoday.com/news/how-much-data-is-generated-every-minute-infographic-1/525692/>.

ICRC, *Autonomy, artificial intelligence and robotics: Technical aspects of human control*. Srpen 2019. Dostupné z: [https://www.icrc.org/sites/default/files/document/file\\_list/autonomy\\_artificial\\_intelligence\\_and\\_robotics.pdf](https://www.icrc.org/sites/default/files/document/file_list/autonomy_artificial_intelligence_and_robotics.pdf).

Internet Encyclopedia of Philosophy. Dostupné z: <https://www.iep.utm.edu/util-a-r/>.

ISAAC, Mike, a FRENKEL, Sheera. *Facebook Security Breach Exposes Accounts of 50 Million Users*. TheNew York Times, Sept. 29, 2018, Section A [cit. 10. 5. 2020]. Dostupné z: <http://Facebook Security Breach Exposes Accounts of 50 Million Users>.

JAEKEL, Brielle. *Zootopia fans help solve crimes via Facebook Messenger chatbot*. Mobile Marketer. Mobile Marketing News. Mobile Marketer [online]. [cit. 22. 4. 2020]. Dostupné z: <https://www.mobilemarketer.com/ex/mobilemarketer/cms/news/messaging/22953.html>.

LEE, Dave. *Google self-driving car hits a bus*. BBC NEWS, únor 2016. Dostupné z: <http://www.bbc.com/news/technology-35692845>.

LESIAK, Malgorzata. *A Comparative Analysis of The Liability of Internet Service Providers in The Context of Copyright Ingfrngement in The U.S., European Union And Poland*. Časopisy Masarykovy univerzity. Dostupné z: <https://journals.muni.cz/mujlt/article/view/2541/2105>.

MARKOFF, J. *Google Cars Drive Themselves, in Traffic*. The New York Times, říjen 2010. Dostupné z: <https://www.nytimes.com/2010/10/10/science/10google.html>.

MORELLE, R. *Google machine learns to master video games*. BBC NEWS, únor 2015. Dostupné z: <http://www.bbc.com/news/science-environment-31623427>.

National Highway Traffic Safety Administration: Preliminary Statement of Policy Concerning Automated Vehicles. Dostupné z: [https://www.google.cz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjYyaWP1KjaAhUB3qQKHwQoDTUQFggnMAA&url=https%3A%2F%2Fwww.nhtsa.gov%2Fstaticfiles%2Frulemaking%2Fpdf%2FAutomated\\_Vehicles\\_Policy.pdf&usg=AOvVaw1kzgMhDIOOc-olHnCFb4os](https://www.google.cz/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjYyaWP1KjaAhUB3qQKHwQoDTUQFggnMAA&url=https%3A%2F%2Fwww.nhtsa.gov%2Fstaticfiles%2Frulemaking%2Fpdf%2FAutomated_Vehicles_Policy.pdf&usg=AOvVaw1kzgMhDIOOc-olHnCFb4os).

OPPERMANN, Artem. *Artificial Intelligence vs. Machine learning vs. Deep Learning*. Dostupné z: <https://towardsdatascience.com/artificial-intelligence-vs-machine-learning-vs-deep-learning-2210ba8cc4ac>.

PRATLEY, Nils. *The Trillion-Dollar Questions over the Flash Crash and the Hound of Hounslow*, The Guardian, duben 2015 [cit. 8. 5. 2020]. Dostupné z: <https://www.theguardian.com/business/2015/apr/25/flash-crash-hound-of-hounslow-trillion-dollar-question>.

SAMPLE, Ian. *Computer says no: why making AIs fair, accountable and transparent is crucial*. The Guardian [online]. Dostupné z: <https://www.theguardian.com/science/2017/nov/05/computer-says-no-why-making-ais-fair-accountable-and-transparent-is-crucial>.

*Saúdská Arábie dala občanství robotovi Sophii. Ženy mají méně práv než ona.* Echo24.cz - Názorový deník. Dostupné z: <https://www.echo24.cz/a/pCHVV/saudska-arabie-dala-obcanstvi-robotovi-sophii-zeny-maji-mene-prav-nez-ona>.

SHASHKEVICH, Alex. *Stanford scholars, researchers discuss key ethical questions self-driving cars present.* Stanford News. Dostupné z: <http://news.stanford.edu/2017/05/22/stanford-scholars-researchers-discuss-key-ethical-questions-self-driving-cars-present/>.

SRIVASTAVA, Nishit a MONDAL, Sandeep. *Development of a Predictive Maintenance Model Using Modified FMEA Approach* (November 4, 2014). The IUP Journal of Operations Management, Vol. XIII, No. 2, May 2014. Dostupné z: <https://ssrn.com/abstract=2518973>.

STEIN, Bram. *You can now order Domino's pizza through a chatbot on Facebook Messenger.* Business Insider [online]. [cit. 22. 4. 2020]. Dostupné z: <https://www.businessinsider.com/you-can-now-order-dominos-pizza-through-a-chatbot-on-facebook-messenger-2016-9>.

TENNENHOUSE, David. *Autonomous Vehicles: Are You Ready for the New Ride?.* MIT Technology Review, listopad 2017. Dostupné z: <https://www.technologyreview.com/s/609450/autonomous-vehicles-are-you-ready-for-the-new-ride/>.

*The Ethics of Autonomous Weapons Systems: The Ethics of Autonomous Weapons Systems.* University of Pennsylvania Law School. Penn Law [online]. Dostupné z: <https://www.law.upenn.edu/institutes/cerl/conferences/ethicsofweapons/>.

*The Future of Artificial Intelligence - Will Robots/Machines Outsmart Humans?* TechPats – Technology Patent Services & Intellectual Property Consulting. Dostupné z: <https://www.techpats.com/future-artificial-intelligence-will-robotsmachines-outsmart-humans/>.

TICHÝ, Luboš. *K rozlišování mezi tzv. subjektivní a objektivní odpovědností – rozsudek NS o povaze odpovědnosti advokáta za škodu.* Bulletin advokacie, odborný právnický portál. Dostupné z: <http://www.bulletin-advokacie.cz/k-rozlisovani-mezi-tzv.-subjektivni-a-objektivni-odpovednosti>.

TOLAN, S., MIRON, M., GOMEZ, E. a CASTILLO, C. „*Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia*“. International Conference on AI and Law, 2019. Dostupné z: <https://doi.org/10.1145/3322640.3326705>.

TURING, Alan. *Education, Machine & Life – Biography.* Dostupné z: <https://www.biography.com/scientist/alan-turing>.

VOOSEN, Paul. *How AI detectives are cracking open the black box of deep learning* AAAS. Science [online]. Dostupné z: <https://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning>.

Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Souhrnná zpráva. Vláda ČR [online]. Dostupné z: <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-souhrnna-zprava-2018.pdf>.

Výzkum potenciálu rozvoje umělé inteligence v České republice ze dne 10. 12. 2018. Analýza právně-etických aspektů rozvoje umělé inteligence a jejích aplikací v ČR. Vláda ČR [online]. Dostupné z: [https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018\\_final.pdf](https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-pravne-eticka-zprava-2018_final.pdf).

WAKEFIELD, Jane. *Microsoft chatbot is taught to swear on Twitter*. BBC News. Dostupné z: <https://www.bbc.com/news/technology-35890188>.

What Is Big Data? - Gartner IT Glossary - Big Data [online]. Dostupné z: <https://www.gartner.com/it-glossary/big-data/>.

White Paper on Artificial Intelligence: a European approach to excellence and trust. European Commission [online]. Dostupné z: [https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en).

WILLIAMS, Martyn. *Inside the Russian hack of Yahoo: How they did it*. Security news, features and analysis about prevention, protection and business innovation. [online]. [cit. 10. 5. 2020]. Dostupné z: <https://www.csoonline.com/article/3180762/inside-the-russian-hack-of-yahoo-how-they-did-it.html>.

Zpráva Etické komise: Automatizované a propojené řízení, červen 2017. Dostupné z: [https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile).

## **Seznam použitých právních předpisů**

Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (Text s významem pro EHP).

Směrnice Evropského parlamentu a Rady 2009/103/ES ze dne 16. září 2009 o pojištění občanskoprávní odpovědnosti z provozu motorových vozidel a kontrole povinnosti uzavřít pro případ takové odpovědnosti pojištění (Text s významem pro EHP).

Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. dubna 2009 o právní ochraně počítačových programů (kodifikované znění) (Text s významem pro EHP).

Směrnice Evropského parlamentu a Rady 2014/104/EU ze dne 26. listopadu 2014 o určitých pravidlech upravujících žaloby o náhradu škody podle vnitrostátního práva v případě porušení právních předpisů členských států a Evropské unie o hospodářské soutěži (Text s významem pro EHP).

Směrnice Rady 85/374/EHS o sblížení právních a správních předpisů členských států týkajících se odpovědnosti za vadné výrobky, ze dne 25. července 1985.

Usnesení předsednictva České národní rady č. 2/1993 Sb., o vyhlášení Listiny základních práv a svobod jako součásti ústavního pořádku České republiky.

Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů.

Zákon č. 89/2012 Sb., občanský zákoník, ve znění pozdějších předpisů.

### **Seznam použité judikatury**

Nález Ústavního soudu sp. zn. IV. ÚS 690/01 ze dne 27. 3. 2003 (N 45/29 SbNU 417).

Nález Ústavního soudu sp. zn. II. ÚS 487/03 ze dne 11. 5. 2005 (N 103/37 SbNU 383).

Nález Ústavního soudu sp. zn. II. ÚS 2070/07 ze dne 13. 12. 2007 (N 223/47 SbNU 935).

Nález Ústavního soudu sp. zn. I. ÚS 420/09 ze dne 3. 6. 2009 (N 131/53 SbNU 647).

Rozsudek Nejvyššího soudu ČSR ze dne 31. 5. 1983, sp. zn. 1 Cz 13/83.

Rozsudek Nejvyššího soudu ze dne 25. února 2003, sp. zn. 25 Cdo 618/2001.

Rozsudek Nejvyššího soudu ze dne 14. 12. 2010, sp. zn. 28 Cdo 537/2010.

### **Seznam ostatních zdrojů**

Důvodová zpráva k zákonu č. 89/2012 Sb., občanský zákoník (konsolidovaná verze). Dostupné z: <http://obcanskyzakonik.justice.cz/images/pdf/Duvodova-zprava-NOZ-konsolidovana-verze.pdf>.

European Group on Ethics in Science and New Technologies. *Artificial Intelligence, Robotics and 'Autonomous' Systems*. Luxembourg: Publications Office of the European Union, 2018. ISBN 978-92-79-80328-4.

Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *Definice UI: hlavní schopnosti a vědní obory*. Duben 2019. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).



Evropská komise: Odborná skupina na vysoké úrovni pro umělou inteligenci. *ETICKÉ POKYNY PRO ZAJIŠTĚNÍ DŮVĚRYHODNOSTI UI*. Duben 2019. Dostupné z: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60663](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60663).

Evropská komise – Liability for Artificial Intelligence and other emerging digital technologies. Dostupné z: <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>.

Jaký je potenciál umělé inteligence v České republice? Vláda ČR [online]. Dostupné z: <https://www.vlada.cz/cz/evropske-zalezitosti/aktualne/jaky-je-potencial-umele-inteligence-v-ceske-republice--170808/>.

Robocalls: Last Week Tonight with John Oliver (HBO) - YouTube. *YouTube* [online]. Dostupné z: [https://www.youtube.com/watch?v=FO0iG\\_P0P6M](https://www.youtube.com/watch?v=FO0iG_P0P6M).

Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a Výboru regionů „Budování důvěry v umělou inteligenci zaměřenou na člověka“ ze dne 8. 4. 2019 (COM/2019/168 final).

Standard ISO/IEC 2382:2015 Information technology – Vocabulary. In: *International Organization for Standardization* [online]. 2015 [2018-11-01]. Dostupné z: <https://www.iso.org/standard/63598.html>, ISO. Standard ISO/IEC WD 22989 Artificial Intelligence – Concepts and terminology. In: *International Organization for Standardization* [online]. [2018-11-01]. Dostupné z: <https://www.iso.org/standard/74296.html>, či ISO. Standard ISO/IEC WD 23053 Artificial Intelligence (AI) Systems Using Machine Learning (ML). In: *International Organization for Standardization* [online]. [2018-11-01]. Dostupné z: <https://www.iso.org/standard/74438.html>.

Stanovisko Evropského hospodářského a sociálního výboru k tématu Umělá inteligence – dopady umělé inteligence na jednotný trh (digitální), výrobu, spotřebu, zaměstnanost a společnost (stanovisko z vlastní iniciativy) (2017/C 288/01).

Stanovisko generální advokátky Juliane Kokott ve věci C-396/12 A. M. van der Ham, A. H. van der Ham-Reijersen van Buuren ze dne 24. října 2013.

Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komise o občanskoprávních pravidlech pro robotiku (2015/2103(INL)).

Usnesení Evropského parlamentu ze dne 12. února 2019 o komplexní evropské průmyslové politice v oblasti umělé inteligence a robotiky (2018/2088(INI)).

## Abstrakt

V souvislosti s vývojem autonomních systémů podporujících technologie umělé inteligence, resp. strojového učení, stále více roste obava jak ve veřejnosti a ve sdělovacích prostředcích, tak i mezi legislativci a vývojáři. Takové obavy přirozeně historicky vznikaly i v případě jiných, dřívějších technologií, ale je zřejmé, že v případě moderních technologií je něco atypického. Období, kdy byli roboti, umělá inteligence a další autonomní systémy pouze sci-fi tématem, je již do jisté míry překonané a autonomní systémy v různých podobách již několik let prostupují do velkého množství oblastí a představují neodmyslitelný prvek světa okolo nás. V praxi se čím dál více společností zabývá vývojem autonomních systémů vybavených umělou inteligencí, zejména pak vývoj chatbotů, autonomních vozidel či autonomních dronů je v dnešní době dobrým příkladem toho, že provoz podobných systémů není tak daleko od reality.

Práce se v první a navazující druhé kapitole věnuje obecně úvodu do tématu autonomních systémů s důrazem na umělou inteligenci, resp. technologie strojového učení, popisu představ společnosti o těchto nových technologiích, základním definičním znakům autonomních systémů, robotů i umělé inteligence.

Třetí kapitola se pak věnuje pravidlům chování autonomních systémů, rozvoji regulatorního rámce a roli práva ve věku inteligentních robotů tak, aby byla zajištěna dostatečná míra právní jistoty a předvídatelnosti. Tato část práce dále obsahuje posouzení jedné z nejkontroverznějších otázek v souvislosti s umělou inteligencí, a to otázky, zda může fungovat koncept umělé inteligence jako osoby v právu.

Čtvrtá kapitola práce se věnuje jedné z nejdiskutovanějších oblastí, kterou je otázka odpovědnosti za jednání umělé inteligence, a to zejména v kontextu současného pojetí subjektivní a objektivní odpovědnosti, prevenční povinnosti i odpovědnost za jednání autonomních systémů v kontextu účinné právní úpravy a možnosti uplatnění různých odpovědnostních režimů, případně jejich výhody a nevýhody.

Pátá, poslední a stěžejní kapitola práce posuzuje s provozem autonomních systémů neoddělitelně spjaté morální otázky, které lze obecně rozdělit do dvou kategorií – na tzv. roboetiku, a na etiku strojů. Kapitola odpovídá i na některé otázky týkající se

vysvětlitelnosti a transparentnosti autonomních systémů, a to právě zejména umělé inteligence, prostřednictvím analýzy s cílem dosažení morálních a důvěryhodných nových technologií.

Klíčová slova: autonomní systémy, umělá inteligence, odpovědnost, robotika, etika strojů

## **Abstract**

In connection with the development of autonomous systems supporting artificial intelligence technologies, respectively machine learning, there is a growing concern both in the public and in the media, as well as among legislators and developers. Naturally, such concerns have naturally arisen in the case of other, earlier technologies, but it is clear that there is something atypical in the case of modern technologies. The period when robots, artificial intelligence and other autonomous systems were only a sci-fi topic is already to some extent obsolete and autonomous systems in various forms have been reaching a large number of areas for several years and represent an integral element of the world around us. In practice, more and more companies are engaged in the development of autonomous systems equipped with artificial intelligence, especially the development of chatbots, autonomous vehicles or autonomous drones is nowadays a good example that the operation of similar systems is not so far from reality.

In the first and subsequent second chapter, the paper deals in general with an introduction to the topic of autonomous systems with an emphasis on artificial intelligence, respectively machine learning technologies, description of society's ideas of these new technologies, basic defining features of autonomous systems, robots and artificial intelligence.

The third chapter then deals with the rules of conduct of autonomous systems, the development of the regulatory framework and the role of law in the age of intelligent robots so as to ensure a sufficient degree of legal certainty and predictability. This part of the paper also contains an assessment of one of the most controversial issues related to artificial intelligence, namely the question of whether the concept of artificial intelligence as a person in law can work.

The fourth chapter deals with one of the most discussed areas, which is the issue of liability for the actions of artificial intelligence, especially in the context of the current concept of fault-based and strict liability, duty of care and liability for autonomous systems in the context of effective regulation and the possibility of applying different liability regimes, eventually their advantages and disadvantages.

The fifth, last and main chapter of the paper assesses the moral issues inextricably linked to the operation of autonomous systems, which can be generally divided into two categories - the so-called roboethics, and the machine ethics. The chapter also provides answers to some questions concerning the explainability and transparency of autonomous systems, especially artificial intelligence, through analysis in order to achieve moral and trustworthy new technologies.

Key words: autonomous system, artificial intelligence, liability, roboethics, machine ethics