

UNIVERZITA KARLOVA

Právnická fakulta

Jan Hořeňovský

Právo, společnost a nové technologie 21. století

Disertační práce

Školitel: Mgr. Petr Agha, LL.M., Ph.D.

Studijní program: Teoretické právní vědy – Teorie, filozofie a sociologie práva

Datum vypracování práce (uzavření rukopisu): 4. 2. 2024

Prohlašuji, že jsem předkládanou disertační práci vypracoval 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 384410 znaků včetně mezer.

JUDr. Jan Hořeňovský

V Praze dne 4. 2. 2024

Obsah

Poděkování.....	6
Úvod.....	7
OBEČNÁ ČÁST.....	10
1. Adaptabilita právních systémů v éře socio-technologických změn	10
1.1 Základní východiska: systémová teorie a relační ontologie	10
1.1.1. Relační ontologie	11
1.1.2. Koncept socio-technologické změny a policy driftu.....	12
1.2 Úvodní poznámky k regulaci nových technologií a problém relevance	13
1.3 Jak se mohou právní systémy vnitřně adaptovat (modely regulace).....	16
1.3.1. Nulový model.....	17
1.3.2. Law-centric model	17
1.3.3. Emergentní model.....	19
1.4 Praktické příklady	21
1.4.1. Desátá léta: Uber a Airbnb.....	21
1.4.2. Dvacátá léta: AI systémy a velká data	22
1.5 Závěr	24
2. Právní a kulturní aspekty teze o technologické stagnaci.....	27
2.1 Teze o technologické stagnaci v oblasti fyzického světa.....	28
2.2 Narůstající vnitřní komplexita práva a problém přeregulovanosti.....	31
2.3 Účel právní úpravy a narůstající rezistence společnosti vůči technologiím.....	32
2.4 Narůstající averze k riziku ve společnosti a právní regulaci.....	34
2.5 Měnící se socio-technologický kontext a vnitřně statický právní systém.....	38
2.6 Diskuse.....	40
2.7 Závěr	43
3. Právní systém a zvyšující se komplexita.....	45
3.1 Komplexita a entropie.....	45
3.2 Konsekvence pro právo a právní systém.....	46
3.3 Role moderních technologií ve vztahu k entropii a komplexitě.....	47
3.4 Závěr	50
ZVLÁŠTNÍ ČÁST	51
4. Umělá inteligence a právo v dlouhodobé perspektivě	51
4.1 Umělá inteligence jako subjekt práva	55
4.1.1. Právo jako svět subjektů a věcí.....	55
4.1.2. Co by umělá inteligence měla splňovat, aby byla subjekt práva?.....	57
4.1.3. Obecná umělá inteligence versus partikulární umělá inteligence	58
4.1.4. Obecná umělá inteligence jako subjekt práva.....	59

4.1.5.	Emulace lidského mozku jako subjekt práva	60
4.1.6.	Problém zastoupení umělou inteligencí v právních vztazích	61
4.1.7.	Umělá inteligence a postavení otroka v římském právu	62
4.1.8.	Závěr	64
4.2	Negativní lidská práva jako základ pro dlouhodobou bezpečnost AI	65
4.2.1.	Dosavadní práce	70
4.2.2.	Umělá inteligence	74
4.2.3.	Lidská práva	76
4.2.3.1.	Která práva by měla být chráněna	78
4.2.3.2.	Lidská práva jako relativní principy	80
4.2.3.3.	Nadřazenost lidských práv	82
4.2.3.4.	Proč lidská práva, a ne obecné lidské hodnoty?	83
4.2.3.5.	Lidská práva a etické teorie	84
4.2.4.	Vývoj umělé inteligence respektující lidská práva: technické aspekty	85
4.2.4.1.	Automatizované rozpoznávání rizika ohrožení práv	86
4.2.4.1.1.	VÝSTUPY: CO ODHADUJEME	87
4.2.4.1.2.	VSTUPY: REPREZENTACE ZÁMĚRŮ CHOVÁNÍ A KONTEXTU	88
4.2.4.1.3.	UČÍCÍ SE ZDROJE	91
4.2.4.2.	Vynucování omezení: problém integrace	94
4.2.4.3.	Konzervatismus s ohledem na nejistotu	96
4.2.4.4.	Prevence specification gaming	96
4.2.5.	Systém mezinárodní ochrany: explorativní návrh	98
4.2.5.1.	Úmluva	98
4.2.5.2.	Systém rozhodování/adjudikace	100
4.2.5.3.	Soukromý ekosystém	102
4.2.6.	Diskuse	103
4.2.6.1.	Práva v kontextu problému specifikace	104
4.2.6.2.	Rizika a omezení	105
4.2.7.	Závěr	107
4.3	Hrozba nových forem vlády	109
4.3.1.	AI totalitarismus	109
4.3.2.	AI technokracie	111
4.4	Závěr	116
5.	Sociální sítě, svoboda projevu a informační veřejný prostor	120
5.1	Význam svobody projevu v demokratické společnosti	120
5.2	Dělení projevů ve veřejném prostoru a éra úpadku svobody slova	122
5.3	Svoboda projevu a sociální sítě	127

5.4	Možnosti regulace sociálních sítí.....	129
5.4.1.	Regulace pomocí soutěžního práva.....	129
5.4.2.	Přímá regulace	130
5.4.3.	Platformy jako vydavatelé a otázka soukromoprávní odpovědnosti.....	133
5.4.4.	Horizontální účinek základních práv a pozitivní závazek státu	136
5.5	Boj s dezinformacemi	137
5.6	Problém zužujícího se mainstreamu	143
5.7	Závěr	146
6.	Technologie, právo a Covid-19.....	147
6.1	Debata o elektronickém sledování osob.....	147
6.2	Dobrovolné elektronické sledování osob.....	148
6.3	Nedobrovolné elektronické sledování osob	150
6.4	Debata o covid pasu	152
6.5	Možnost užití velkých dat a umělé inteligence při boji s covidem	153
6.6	Závěr	155
	Závěr	156
	Seznam zkratk	158
	Seznam použitých zdrojů.....	160

Poděkování

Předně bych chtěl poděkovat Elišce Mainclové, za neutuchající podporu i podnětné připomínky, nejen k této práci. Dále školiteli Petru Aghovi, za volnost při zpracování tématu. Právnické fakultě Univerzity Karlovy, že takto ambiciózní téma práce umožnila zkoumat. Janu Kyselovi, za příjemné prostředí katedry politologie a sociologie, na které jsem tuto práci měl čest psát. Adamu Růžičkovi a celému týmu Institutu H21, za plodné diskuse, které se v mnohém propsaly do této práce. Andree Prinzové, za pomoc při jazykové stránce textu. A v neposlední řadě kočkám Čuance a Daisy Von Han da Costě a kocourovi Chillimu Charlsburovi, za poctivé asistování při psaní.

Úvod

Práce, která se vám právě dostala do rukou, se věnuje poměrně široce vymezenému tématu vztahu práva, společnosti a nových technologií 21. století. Text je rozdělen na část obecnou, která se zaměřuje na oblasti tématu společné, a část zvláštní, která je úzce zaměřena na určité konkrétní problémy spojené s partikulárními technologiemi, jako jsou umělá inteligence nebo sociální sítě. Každá z těchto částí je dále rozdělena do tří kapitol, z nichž každou lze číst i samostatně. Texty vznikaly původně samostatně, avšak myslím, že se mi je podařilo seřadit a upravit způsobem, který poskytuje koherentní celek. I tak ale vnímám, že jde o text velmi široce rozkročený, který spoustu témat nevyhnutelně pouze lehce načrtne. Jde o daň za zpracovávání širokého tématu, kterou však v dnešní době stále rostoucí partikularizace dle mého mínění stálo za to zaplatit.

První, úvodní, kapitola vychází ze sociologie práva a jejím účelem je vytyčit základní terén, který je později hlouběji rozveden v dalších kapitolách ve vztahu ke konkrétním socio-technologickým změnám. Tato část vychází ze systémové teorie, která má podle mého názoru nejlepší předpoklady popsat fungování evoluce/adaptability práva v kontextu měnícího se socio-technologického prostředí. Jde o část, která, byť nedokonale, vytváří důležité uvedení tématu a rámec pro popsání základních jevů. Teorie zde představená přitom funguje obecně na studium jakýchkoli technologií v kontextu práva a společnosti.

Druhá kapitola se věnuje netradičnímu tématu technologické stagnace. Pojem „technologie“ totiž konceptuálně nezahrnuje pouze tzv. „informační technologie“, ale i ty ostatní. Varující je skutečnost, že i samotné předporozumění slova „technologie“ se v posledních letech proměňuje. Například „centrum práva a technologií“ by ještě před několika desítkami let znamenalo pracoviště, které řeší jaderné reaktory a jejich právní úpravu. Dnes bychom si takřka výlučně představili pracoviště práva informačních technologií. Jde o posun v jazyce, který nám říká něco důležitého o společnosti a prostředí, ve kterých žijeme. Technologie totiž stále znamenají i oblast energetiky, stavebnictví, léků, dopravy atd., přičemž v těchto oblastech stále více autorů dochází k názoru, že zažíváme technologickou stagnaci. Originalita mého přínosu v tomto podtématu spočívá v pokusu o hlubší propojení těchto hlasů s otázkami fungování a problémů současného práva.

Třetí kapitola, poslední v obecné části, je silně experimentální. Snaží se totiž v kontextu systémového pohledu na právo vysvětlit roli technologií ve vztahu k entropii a komplexitě jak těchto systémů, tak jejich prostředí. Věnuji se zde otázce, jak technologie ovlivňují praxi práva

a co mohou být benefity a omezení s tímto jevem spojená. Jde však pouze o prvotní náčrt tohoto dosud velmi málo probádaného tématu.

Čtvrtou kapitolu považuji za nejvýznamnější, a to jak z pohledu originality a inovativnosti, tak svojí důležitostí pro společnost a právo. Věnuje se totiž otázce dlouhodobých aspektů umělé inteligence, přičemž velmi silně akcentuje otázku bezpečnosti těchto systémů. Propojuji zde oblast teorie práva a teorie lidských práv s technickým pohledem, jak by se dala taková bezpečnost zajistit. Konkrétně je zde představen návrh, jak by se dalo inkorporovat dodržování negativních lidských práv do samotné technologie, včetně doprovodných regulativních mechanismů mezinárodního práva. Dále se zde věnuji důležitému tématu zapojování umělé inteligence do forem vládnutí a otázce právní subjektivity těchto strojů.

Pátá kapitola se věnuje aktuální otázce svobody projevu na veřejných internetových fórech. Snažím se zde hájit liberální pozici, která akcentuje důležitost této svobody a vyvrací argumenty pro její omezování spojené s morální panikou například ve vztahu k otázce šíření tzv. „dezinformací“. V této kapitole se snažím shrnout významné poznatky k tomuto tématu a představit pohled, který nabízí konstruktivní společenská a právní řešení.

Poslední kapitola reflektuje dobu pandemie Covid-19 a řeší některé otázky spojené s využíváním moderních technologií v boji s pandemií.

Tematicky jde o text, který by měl sloužit všem zájemcům o pochopení role práva v regulaci nových technologií. Je určen pro ty, kteří věří, že technologicky vyspělejší a lepší budoucnost je dosažitelná a že musíme aktivně hledat cesty, jak jí dosáhnout. Věřím, že partikulárně bude tato práce užitečná především všem právním akademikům, kteří se podobným tématům budou chtít věnovat. Pro ty by myšlenky zde obsažené mohly fungovat jako dobrý odrazový můstek pro vlastní výzkumné projekty. Přál bych si však, aby se tato práce propsala i do právní praxe, a to jak v cestě legislativní, tak judiciální, neboť jedině tak bude moci reálně ovlivnit společnost. V neposlední řadě si myslím, že je můj text psaný dostatečně srozumitelně, aby jej mohla číst i širší veřejnost. Věřím proto a doufám, že na moji práci navážou další autoři a že jim tento text může být jak inspirací, tak praktickým pomocníkem.

Je třeba také zmínit, že jde o práci, která je z většiny složena z již publikovaných textů – informace, kde byly které části publikovány, je uvedena na začátku každé z kapitol v první poznámce pod čarou. Nad rámec toho budou ještě části této disertace publikovány ve dvou knihách, které v době odevzdání této práce dosud nevyšly. V první řadě jde o knihu *Právní aspekty umělé inteligence – výzvy a příležitosti*, která vyjde v nakladatelství Wolters Kluwer

v průběhu roku 2024. V druhém případě jde o knihu *Právo na veřejný prostor*, která vyjde v tomtéž roce v nakladatelství Academia.

Části textu jsem psal na zahraničních univerzitách, především bych zmínil semestrální výzkumný pobyt na Tel Aviv University v Izraeli v zimním semestru 2021, kde jsem hostoval v pozici Cegla Visiting Research Fellow. Velmi užitečným byl také měsíční pobyt v listopadu 2022 na University of Hamburg v pozici Visiting scholar. Oba tyto pobyty byly nesmírně přínosné a chtěl bych poděkovat všem zapojeným, že bylo možné je uskutečnit. Části práce jsem také v průběhu studia prezentoval jak v Česku, tak v zahraničí, především v Paříži a Hamburku.

OBEČNÁ ČÁST

1. Adaptabilita právních systémů v éře socio-technologických změn¹

První kapitola se věnuje vztahu práva, společnosti a nových technologií z pohledu sociologie práva, konkrétně prizmatem aplikované systémové teorie² propojené s teorií regulace nových technologií. V první části představuji systémovou teorii a relační ontologii spolu s koncepty *socio-technologické změny a policy driftu*. V druhé části se zaměřuji na možné způsoby (modely), jak mohou právní systémy na tyto změny vnitřně reagovat, které lze dle mého názoru současně považovat za metodologické přístupy v oboru doktrinální právní vědy. V třetí části se na příkladech z desátých a dvacátých let tohoto století pokouším ukázat, jak může být tento konceptuální model uchopen v praxi.

Musím na úvod upozornit na jednu důležitou limitaci. Aplikace tohoto rámce je ze své podstaty omezena v tom smyslu, že předpokládá, že právní systém je relativně stálá praxe, která omezeným způsobem reaguje na dynamický socio-technologický kontext. Stále však musíme mít na paměti i možnost, že socio-technologické změny, k nimž dochází v okolí práva, mohou mít více než jen dopad na účinnost práva, slouží-li regulačním politikám. Tyto změny mohou nakonec zcela proměnit i samotnou povahu práva, například směrem ke stavu, který někteří autoři nazývají *právní singularitou*.³ Takový stav je však již mimo záběr této práce.

1.1 Základní východiska: systémová teorie a relační ontologie

Ve svém konceptuálním rámci vycházím ze systémové teorie, která považuje právní systém za *operativně uzavřený autopoietický systém*.⁴ Rozebírat celou historii systémové teorie by zde nebylo účelné a tyto rozbory lze v české právní vědě najít ve vysoké kvalitě na jiných místech.⁵ Alespoň ve zkratce bych zmínil, že teorie autopoietických systémů vychází původně

¹ Tato část vychází z a) HOŘEŇOVSKÝ, J. Modely regulace nových technologií v éře socio-technologických změn. *Obchodní právo*, 2022, roč. 31, č. 7–8, s. 18–27; b) HOŘEŇOVSKÝ, J. Modely přístupu práva k regulaci nových technologií a metodologie právní vědy. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022, s. 3–11.

² K systémové teorii v právu obecně – především viz ONDŘEJEK, P. *Koncepce práva jako systému*. Praha: Wolters Kluwer, 2020.

³ Viz např. AIDID, A.; ALARIE, B. *The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better*. Toronto: University of Toronto Press, 2023. Právní singularitu můžeme chápat jako podmnožinu sociální singularity.

⁴ LUHMANN, N. *Introduction to Systems Theory*. Cambridge: Polity Press, 2013, s. 63–83.

⁵ Především opět viz ONDŘEJEK, P. *Koncepce práva jako systému...*

z biologie a její základní podstatou jsou otázky fungování a reprodukce systémů.⁶ Luhmann takovýto druh systému popisuje následovně: „*Jako operativně uzavřené mají být označeny systémy, které jsou k vytváření vlastních operací odkázány na síť vlastních operací a v tomto smyslu se samy reprodukují.*“⁷

Pro právní systém je z tohoto pohledu klíčové jeho oddělení od prostředí.⁸ Buď něco je součástí právního systému, nebo je to prostředím, ve kterém právní systém působí. Uzavřenost v tomto kontextu však nelze považovat za úplnou separaci (odloučenost), pouze se tím chce říct, že se právní systém reprodukuje skrze „*vytváření vlastních operací*“.⁹ Autopoietická uzavřenost žádným způsobem nevylučuje kauzální působení systému se svým prostředím ani možnou závislost systému na prostředí (a naopak závislost prostředí na systému).¹⁰

1.1.1. Relační ontologie

Operativní oddělenost mezi právním systémem a jeho socio-technologickým prostředím přináší nutnost *relačního pohledu*.¹¹ Pochopení vztahu (relace) přitom vyžaduje zachycení vzájemné dynamiky. Analýza každého prvku zvláště samozřejmě může přinést cenné partikulární poznatky, nepomůže nám však s pochopením vysvětlujícího rámce relačních jevů. Jinými slovy, potřebujeme nejdříve pochopit dynamiku vztahu, abychom zjistili, jak a proč záleží na vlastnostech jednotlivých prvků relace.¹²

Abych tuto abstraktní myšlenku uvedl na jiném než právním příkladě – klasickou ukázkou vztahového fenoménu je například důvěra, kterou lze chápat jako určitý vztah mezi dvěma subjekty. Důvěra se může rozpadnout i tehdy, kdy oba její subjekty objektivně neudělaly nic špatně (tzn. chyba není na straně ani jednoho z nich). Problémem v takovém případě je nefunkčnost vzájemného vztahu, který se zde svým způsobem chová jako samostatná entita. Stejně platí pro vztah právního systému a jeho socio-technologického prostředí, ve kterém normativní systém působí. Také tento vztah může přestat fungovat i v situaci, kdy na samotné regulaci nebo prostředí (tedy na prvcích vztahu) není objektivně nic špatného.

⁶ Tamtéž, s. 63

⁷ LUHMANN, N. *Právo společnosti*. Praha: Academia, 2021, s. 41–42.

⁸ Tamtéž, s. 41.

⁹ Tamtéž, s. 41.

¹⁰ BAXTER, H. Niklas Luhmann's Theory of Autopoietic Legal Systems. *Annual Review of Law and Social Science*, 2013, vol. 9, no. 1, s. 170.

¹¹ K relačnímu přístupu v sociálních vědách obecně viz SELG, P.; VENTSEL, A. *Introducing Relational Political Analysis. Political Semiotics as a Theory and Method*. Palgrave Macmillan, 2020.

¹² Jde o myšlenku Adama Růžičky vyjádřenou ve vztahu ke krizím současných demokracií.

Vzhledem k relační povaze adaptace práva proto také nelze vytvořit stabilní produktovou teorii, jak by se měl právní systém ve vztahu k prostředí přizpůsobovat nebo dokonce jak by měl optimálně trvale vypadat. Ze své podstaty se totiž vůbec nejedná o jakkoli stabilní a tudíž předvídatelnou kategorii. To, co dobře funguje dnes, nemusí být vhodné zítra, a co bude vhodné zítra, může být zcela nepoužitelné v delším časovém horizontu, kdy se prostředí výrazně změní. Jde o stejný problém, jako bychom chtěli vymyslet optimálního biologického predátora. Žralok operuje výborně v moři, ale pokud toto prostředí vyschne a on se ocitne na poušti, jeho dříve vynikající vlastnosti budou zcela nepoužitelné.

Z těchto důvodů platí, že je možné přijít pouze s procedurálními přístupy vysvětlujícími, jak může být adaptabilita/evoluce právního systému ve vztahu k měnícímu se socio-technologickému prostředí kontextuálně realizována.¹³ Takový proces spočívá v rozbití vnitřní struktury systému a jeho opětovném začlenění do nové struktury. Jedná se o vývojový proces sebeorganizace systému.

1.1.2. Koncept socio-technologické změny a policy driftu

Jak vyplývá z výše uvedeného, právní systémy jsou vztahově provázané se svým socio-technologickým prostředím a vstupují do komplikovaných relací s různými dalšími společenskými sub-systémy (např. s ekonomikou či politikou). Může se přitom stát, že se socio-technologické prostředí, v němž právní systém působí, výrazně změní – toto nazýváme *socio-technologickou změnou* – přičemž právní systém tuto změnu po dlouhou dobu nemusí kognitivně vnímat a vnitřně reflektovat. V takovém případě právní systém, právě vzhledem ke své operativní autopoietické uzavřenosti, pokračuje v reprodukci závěrů (např. rozhodnutí), jako kdyby tato změna vůbec nenastala, což může narušovat dříve dobře fungující vztah. Toto může mít nepříznivý vliv jak na ostatní společenské sub-systémy, tak i na legitimitu samotného právního systému.

Jinými slovy, pokud právní systém zůstává vnitřně – operativně – nezměněn (což je otázkou jeho vnitřní struktury a vnitřních aktérů), zatímco vnější svět se výrazně proměňuje, nevyhnutelně dochází k posunům, jak právní systém působí na své prostředí. Společenský dopad práva není statický a nelze s ním zacházet staticky, neboť to vede k nezamýšleným jevům, které se v anglicko-jazyčné literatuře označují jako *policy drift*.¹⁴

¹³ Pozn. podobně jako u biologické teorie evoluce nebo teorie realizace relevance v kognitivní vědě.

¹⁴ RICCUCCI, M. N. *Policy Drift: Shared Powers and the Making of U.S. Law and Policy*. New York: NUP, 2018.

Policy drift znamená situaci, kdy „*politika nebo instituce není modernizována, aby odrazila měnící se vnější okolnosti, a tato nedostatečná modernizace způsobuje, že se výsledky politiky nebo instituce mění – někdy dramaticky*“.¹⁵ Klasickým příkladem tohoto jevu může být nominálně pevně stanovená výše státního starobního důchodu (politika uskutečněná skrze právní normu), jejíž reálná hodnota byla snížena v důsledku vysoké inflace (měnící se vnější okolnosti). Policy drift je přitom situace v mnoha případech nechtěná a nežádoucí, na kterou by právní systém měl reagovat, aby zachoval původní účel právní úpravy.

Je užitečné o policy driftech neuvažovat jako o něčem záměrně plánovaném, ale spíše jako o nezamýšleném a obvykle škodlivém následku často nepředvídatelné dynamiky mezi právním systémem a jeho měnícím se prostředím. Jejich účinek přitom může jít oběma směry: regulace může být v novém kontextu příliš restriktivní, nebo naopak nedostatečná. Obojí má potenciál být podobně problematické, neboť obojí neúmyslně negativně ovlivňuje vztah práva a společnosti.

I přesto, že je tento koncept obecně užitečný, ve zbytku textu používám mírně upravenou významovou verzi, která zdůrazňuje silnější vazbu na moderní disruptivní technologie. Dále tedy používám tento termín ve významu *technologicky zapříčiněného policy driftu*, čímž se rozumí *policy drift* způsobený významným (disruptivním) technologickým pokrokem. Role moderních technologií je totiž v kontextu *socio-technologických změn a policy driftů* zcela zásadní, neboť technologie jsou v současnosti jednou z nejvýznamnějších příčin změn prostředí, ve kterém právní systém působí.

1.2 Úvodní poznámky k regulaci nových technologií a problém relevance

Při obecných úvahách o vztahu práva, společnosti a nových technologií se nabízí několik kategorií výchozích přístupů. První kategorie zdůrazňuje nutnost regulace nových technologií jako takových a zaměřuje se na partikulární technologii a její technické parametry. Osobně se domnívám, že tento přístup není příliš perspektivní, neboť se – pohledem výše představeného rámce – zabývá pouze vlastností jednoho z prvků vztahu (technologemi), a nikoli neustále se měnící vztahovou realitou mezi právním systémem a jeho socio-technologickým prostředím. Právní regulace vycházející z tohoto přístupu také často bude příliš technicky specifická (což ji dělá pro většinu právníků nepřístupnou), náchylná k hypertrofii

¹⁵ GALVIN, D. J.; HACKER, J. S. The Political Effects of Policy Drift: Policy Stalemate and American Political Development. *Studies in American Political Development*, 2020, vol. 34, no. 2, s. 2.

(což neúměrně zvyšuje vnitřní komplexitu práva) a může rychle zastarat, tedy stát se obsoletní.¹⁶

Jak argumentuje Bennett Moses, „*technologie není ani zvláštním důvodem, ani zvláštním předmětem regulace*“.¹⁷ Lepším přístupem je zaměřit se na socio-technologické změny spojené s implementací technologií, nikoli na nové technologie *per se*.¹⁸ Tento alternativní přístup vychází z myšlenky, že většina nových problémů spojených s technologiemi nepochází z „*čehokoli technologického*“.¹⁹ Naopak, problémy v naprosté většině případů pramení ze „*skutečnosti, že se socio-technologický kontext, v němž zákony a regulace fungují, v důsledku proudu nových technologických možností vyvíjí*“.²⁰ Existuje zde však jedna významná výjimka. Pokud řešíme otázku, jak zajistit, že určitá technologie právní normy vnitřně dodržuje, je zaměření se na technologickou stránku věci zcela klíčové.²¹

Vzhledem k tomu, že je kapacita právních systémů měnit sebe sama (adaptovat se) ze své podstaty omezena²² a tempo, jakým se nové technologie objevují, se bude pravděpodobně zvyšovat, je potřeba zdůraznit také požadavek, aby se řešily pouze skutečně společensky *relevantní* problémy. Jak však poznat, co je relevantní? Podle Bennett Moses mohou být dobrým vodítkem následující kritéria: 1) dochází k selhání trhu; 2) dochází k zásahům do práv osob; 3) dopady technologie narušují sociální solidaritu; 4) je potřeba demokratická správa (democratic governance) nové technologie.²³

Jiní autoři používají lehce odlišná kritéria. Například Maas v souvislosti s AI systémy zdůrazňuje těchto šest oblastí: 1) selhání trhu; 2) rizika pro lidské zdraví, bezpečnost nebo životní prostředí; 3) rizika pro morální zájmy, práva nebo hodnoty; 4) narušení sociální solidarity; 5) rizika pro demokratický proces; 6) rizika pro soudržnost, účinnost nebo integritu

¹⁶ PETIT, N. Law and Regulation of Artificial Intelligence and Robots—Conceptual Framework and Normative Implications. *SSRN*, 2017, s. 16.

¹⁷ BENNETT MOSES, L. Regulating in the Face of Sociotechnical Change. In: BROWNSWORD, R. (eds.). *The Oxford Handbook of Law, Regulation and Technology*. Oxford: Oxford University Press, 2017, s. 1. Citováno dle Research Gate: <http://dx.doi.org/10.13140/RG.2.2.17805.67048>.

¹⁸ K tomuto konceptu viz MAAS, M. M. Aligning AI Regulation to Sociotechnical Change. In: BULLOCK, J. (eds.). *Oxford Handbook of AI Governance*. Oxford: Oxford University Press, 2022.

¹⁹ HEIDEGGER, M. *The Question Concerning Technology and Other Essays*. New York: Harper & Row Publishers, 1977.

²⁰ BENNETT MOSES, L. Regulating in the Face of Sociotechnical Change..., s. 2.

²¹ Viz pozdější část o umělé inteligenci respektující negativní lidská práva.

²² Pozn. omezená schopnost právního systému měnit sebe sama vyplývá z jeho účelu zajišťovat stabilitu ve společnosti. Kdyby se právní systém mohl měnit neomezeně, mohl by zítra vypadat úplně jinak než dnes. Takový systém by však jen stěží mohl plnit funkci zajišťování právní jistoty ve společnosti.

²³ BENNETT MOSES, L. Regulating in the Face of Sociotechnical Change..., s. 4–10.

stávajícího regulačního ekosystému, jehož úkolem je zmírňovat předchozí rizika.²⁴ Domnívám se, že oba tyto seznamy lze z pozice právního systému vnímat jako svého druhu *kritéria relevance*.

Riziko neuchopení problematiky relevance spočívá v tom, že bez (vědomého či nevědomého) využití *realizace relevance* se právní systémy mohou topit v irelevantních otázkách, což by mohlo vést k paralýze. Partikulární socio-technologická změna rovněž nemusí být nutně „*kvalitativně nová, dramatická, náhlá nebo špičková* [...] *aby vyvolala intenzivní a významnou změnu v rovnováze moci nebo ve společenských strukturách*“.²⁵ Ačkoli tedy může být realizace relevance ošemetný koncept, pokud by byl správně uchopen, může být také nesmírně cenný vzhledem ke své schopnosti poukázat na věci, kterých bychom si jinak nevšimli nebo jim nevěnovali dostatečnou pozornost.

Jak navíc argumentuje Maas: „*ne všechny vědecké průlomy, nové technologické možnosti nebo dokonce nové způsoby využití nutně vyvolají takovou socio-technologickou změnu, která vyžaduje pozornost regulačních orgánů*“.²⁶ Aby však bylo možné lépe rozpoznat, co regulační pozornost skutečně vyžaduje, měli by aktéři právního systému začít problém relevance více vnímat. Právní systém totiž nemůže mít ambici regulovat vše, nýbrž pouze to nejdůležitější, tedy to, co je nejvíce relevantní.

Vhodný postup tedy obecně spočívá v analýze socio-technologických změn spojených například se společenskou implementací AI systémů z pohledu výše zmíněných (nebo podobných) *kritérií relevance*. Například se tedy přednostně ptát, zda socio-technologická změna vede k ohrožení základních práv, a pokud ano, posléze hledat způsoby, jak by měl právní systém vnitřně reagovat. Pokud přitom není žádné z těchto kritérií naplněno, je otázka, zda má právní systém vyvíjet úsilí k náročné adaptaci, velmi důležitá. Pokud některé z kritérií naplněno naopak je, právní systém by reagovat spíše měl.

Tato kritéria relevance také do jisté míry odpovídají na základní otázku, kterou by Luhmann nebo Teubner položili – zda je právo se svou velmi specifickou funkční strukturou a binárním kódem vhodné pro řešení socio-technologických změn. Příkladně v případě některých problémů – například pokud jde o ochranu základních práv nebo zachování sociální

²⁴ MAAS, M. M. Aligning AI Regulation to Sociotechnical Change. In: BULLOCK, J. (eds.). *Oxford Handbook of AI Governance*. Oxford: Oxford University Press, 2022, s. 6–7. Citováno dle SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3871635.

²⁵ Tamtéž, s. 6.

²⁶ Tamtéž.

solidarity ve společnosti – pokládám za evidentní, že žádný jiný společenský sub-systém není schopen těchto cílů dosáhnout lépe.

1.3 Jak se mohou právní systémy vnitřně adaptovat (modely regulace)

Jakmile dokážeme identifikovat relevantní socio-technologické změny a jejich škodlivé společenské důsledky, musíme si položit následující otázku: Jak se může právní systém vnitřně přizpůsobit změněné realitě? Toto reflektuje skutečnost, že evoluce práva může být popsána jako „*souhra mezi endogenními evolučními mechanismy právního systému a exogenní evolucí společnosti jako celku*“.²⁷ A že pohledem Luhmanna: „*evoluce probíhá pouze tehdy, pokud jsou ve vztahu mezi systémem a prostředím zachovány jak odlišnosti, tak adaptace*“.²⁸ Právě vnitřní adaptace právního systému je předmětem této části.

V předchozí části představená východiska nyní propojím se třemi základními modely regulace nových technologií, které lze současně považovat za způsoby vnitřní změny právního systému, a ještě o krok dál za metodologii v oblasti doktrinální právní vědy. Pro další výzkum ponechávám otevřený sociologický pohled na tyto modely, který by nad rámec zde představeného musel vysvětlit také podmínky, jež způsobují, že se právní systém rozhoduje pro jeden z těchto přístupů. Takovou ambici nemám, tato část proto spíše zůstává na úrovni teorie regulace nových technologií.

První model logicky vyplývá z výše uvedené dynamiky mezi právním systémem a jeho socio-technologickým prostředím. Druhý a třetí model je inspirován prací Petita a De Coomana ve vztahu k regulaci umělé inteligence.²⁹ Čtenáře zde však musím upozornit, že definice a pojetí výše zmíněných autorů se od mého odlišuje. Petit a De Cooman se dle mého názoru dopouštějí jedné výrazné chyby, neboť striktně nerozlišují mezi situacemi, kdy jde o vnitřní změnu právního systému (resp. o to, jak taková změna probíhá), a kdy jde naopak o analýzu vztahu mezi systémem a jeho prostředím (z čehož lze dovodit relevanci pro právní systém).

Domnívám se, že tyto dvě věci je třeba striktně rozlišovat. V první situaci otázka totiž zní: bez ohledu na to, co je předmětem socio-technologické změny, jaké jsou prostředky, jimiž se právní systém může vnitřně přizpůsobit? Zatímco v druhém případě se ptáme, zda je vnější

²⁷ TEUBNER, G. Substantive and reflexive elements in modern law. *Law & Society Review*, 1983, vol. 17, no. 2, s. 263.

²⁸ LUHMANN, N. *Law as a Social System*. Oxford: OUP, 2004, s. 231.

²⁹ PETIT, N.; DE COOMAN, J. Models of Law and Regulation for AI. *RSCAS 2020/63 Robert Schuman Centre for Advanced Studies*, 2020, s. 1–5.

socio-technologická změna a její důsledky natolik relevantní, aby se právní systém vůbec snažil interně reagovat, a pokud ano, jakými jejími aspekty je třeba se primárně zabývat. V druhém případě půjde také o prvotní detekci potenciálně relevantních problémů.

1.3.1. Nulový model

První možností, jak může právní systém reagovat na měnící se socio-technologické prostředí, je *nedělat nic a tvářit se, jako by ke změně nedošlo*. Jedná se o pasivní model, jehož podstatou je zachování nezměněného vnitřního fungování a reprodukce právního systému. Tento přístup je z pohledu tvůrců práva, správních orgánů či soudců přímočarý a lákavý, ale v mnoha případech také velmi problematický vzhledem k výše diskutované možnosti technologicky zapříčiněných policy driftů.

V takovém případě tento model umožňuje narušit stávající společenskou smlouvu a někdy obchází zavedené regulace v daném odvětví. Tento pasivní přístup proto může být velmi problematický, pokud jde o skutečně disruptivní technologie. Dopad umělé inteligence nebo robotiky na společnost bude radikálně transformační a pasivita a staticita právních systémů může vést ke značným společenským problémům.

Mohou však existovat situace, kdy tento přístup dává smysl, a to za předpokladu, že není naplněno žádné z *kritérií relevance*. V takovém případě se posléze nabízí řešení nezatěžovat právní systém potřebou změny či proměny. Výhodou takového přístupu je, že nezatěžuje omezenou kapacitu právního systému měnit se a díky případné deregulaci vede ke snížení celkové úrovně složitosti právního systému.

1.3.2. Law-centric model

Druhým možným přístupem je tzv. *law-centric model*. V tomto modelu se na socio-technologickou změnu nahlíží optikou stávajícího právního rámce, případně jeho možných drobných modifikací. V rámci tohoto přístupu by právníci měli reinterpretovat stávající právo nebo provést skrze legislativu drobné změny ve stávající právní infrastruktuře tak, aby byl zachován účel právní úpravy, který je v důsledku socio-technologických změn a technologiemi způsobených policy driftů v nové situaci nenaplněván.

Cílem tedy není poskytnout radikálně nový pohled, ale provést zásah, který by pomohl lépe naplnit dříve stanovené cíle právního systému. Tento přístup je ze své podstaty reaktivní a dominují mu teleologické otázky, neboť v něm soudy, orgány veřejné moci a zákonodárci

musí zvažovat účel práva.³⁰ Vzhledem k relačnímu vztahu právního systému s dynamickým socio-technologickým prostředím se právní systém někdy musí změnit, aby jeho dopad na společnost zůstal stejný. Tento přístup je tedy primárně o analýze, zda a případně jak lze aplikovat již existující právní normy na novou situaci. Ve svém druhém kroku se posléze ptá, jak upravit stávající právní rámec, aby socio-technologickou změnu lépe reflektoval.

Ve vztahu k umělé inteligenci může první otázka například znít, zda současná právní úprava odpovědnosti dopadá i na tyto systémy, popřípadě jak. Ve druhém kroku se můžeme ptát, zda nepřidat do právního řádu speciální ustanovení, které by tyto případy specificky upravovalo. V oblasti evropského soutěžního práva budou zase předmětem *reinterpreta*ce, popřípadě úpravy či doplnění, například články 101 a násl. SFEU. Tento krok může učinit například orgán aplikující tuto právní úpravu či soud. V mnoha případech podobnou možnost však prvně nastíní akademická práce v oblasti doktrinální právní vědy.

V tomto kontextu je nezbytné zdůraznit, že se role a význam vědecké teorie v přírodních vědách a *normativních systémech* výrazně odlišují. V oblasti teoretické práce v přírodních vědách jde o artikulaci hypotézy, kterou by mělo být možno empirickým pozorováním buď vyvrátit, nebo potvrdit,³¹ přičemž názor teoretika v této oblasti chod samotného pozorovaného jevu příliš neovlivní. V oblasti normativních systémů naproti tomu doktrína systém skrze vnitřní argumentaci a logiku sama vytváří.³² Nejde tedy o pozici nezúčastněného třetího, ale naopak aktéra normativního systému.

Z toho, co je (facticita), přitom nelze odvodit, co by mělo být (normativita).³³ Normativní systémy jsou od objektivně pozorovatelného fyzického světa odděleny, přičemž jejich utváření, vnitřní strukturování a změna či proměna je věcí především samotného systému a jeho aktérů. Toto samozřejmě neznamená, že by se doktrinální práce neměly inspirovat v závěrech empirických zkoumání, či obecněji v tom, co se děje ve vnějším světě. Je-li však řeč o implementaci faktických pozorování do normativních systémů, popřípadě o otázku vnitřních vztahů v rámci normativních systémů, může být role empirie pouze nápomocná.

³⁰ PETIT, N.; DE COOMAN, J. Models of Law and Regulation for AI. *RSCAS 2020/63 Robert Schuman Centre for Advanced Studies*, 2020, s. 1–5.

³¹ POPPER, K. *The Logic of Scientific Discovery*. Routledge, 2002.

³² Tuto skutečnost dobře analyzuje G. Soros, byť tak nečiní v kontextu normativních systémů, který tento jev nazývá *reflexivita*. Viz SOROS, G. Fallibility, Reflexivity, and the Human Uncertainty Principle. *Journal of Economic Methodology*, 2014, vol. 20, no. 4, s. 309–329.

³³ BLACK, M. The Gap Between 'Is' and 'Should.' *The Philosophical Review*, 1964, vol. 73, no. 2, s. 165–181.

Změna právního systému musí vzejít z vnitřku právního systému, přičemž z toho, co empiricky pozorováno je, nelze bez dalšího odvodit, jak by právní systém vypadat *pro futuro* měl.

Zásadou by přitom mělo být, že pokud lze *socio-technologickou změnu* adekvátně vyřešit pomocí změny interpretace již existující právní úpravy, popřípadě drobnou změnou v legislativě, nemělo by se přistoupit k vymýšlení nové komplexní regulace. To souvisí s problémem, že každá nová regulace může mít nepříznivé vedlejší účinky a zvyšuje složitost právního systému. Musíme také vzít v úvahu, že vytváření nových rozsáhlých předpisů je finančně, časově i personálně nákladné.

Tento model bude fungovat především na úrovni *rozptýlených znalostí* (dispersed knowledge)³⁴ a spontánní tvorby práva³⁵, například prostřednictvím postupně se vyvíjející judikatury soudů nebo rozhodovací praxe správních orgánů (tzn. skrze soudní a správní precedenty). Nicméně v případě drobných úprav stávající právní infrastruktury je možný i více centralizovaný přístup skrze legislativní normotvorbu.

Pokud jde o omezení, jsou zde přinejmenším dvě důležitá. Zaprvé, lze pozorovat, že orgány veřejné moci nejsou příliš ochotny reinterpetace existujícího práva provádět. Toto bylo dobře vidět například na odporu podřadit pod veřejnoprávní regulaci hotelů Airbnb. Druhý problém spočívá v tradiční otázce výkladu v právu: co je ještě přijatelné změnit skrze interpretaci, a co již ne ani v rámci extenzivního výkladu. Jsem si vědom tohoto problému, proto do modelu zahrnuji i možnost drobných legislativních úprav stávajících pravidel a principů. Pokud tedy nelze provést nějakou změnu prostřednictvím reinterpetace, stále existuje způsob, jak normu nebo zásadu v rámci tohoto modelu legislativně upravit. Otázka limitů interpretace je však v právní vědě stále otevřená a dalece přesahuje rozsah a ambice této kapitoly.³⁶

1.3.3. Emergentní model

Třetím možným přístupem je tzv. *emergentní model*. Podstatou tohoto modelu je zavedení nové komplexní regulace, popřípadě dokonce nového odvětví práva. Do tohoto modelu bych také zahrnul vysoce inovativní přístupy, které jsou nyní nepředstavitelné, ale k nimž může socio-technologický a právní vývoj vést.

³⁴ SOWELL, T. *Knowledge And Decisions*. Basic books, 1979.

³⁵ HAYEK, F. A. *Law, Legislation and Liberty: A new statement of the liberal principles of justice and political economy*. Routledge Classics, 2012.

³⁶ Viz WINTR, J. *Metody a zásady interpretace práva*. Praha: Auditorium, 2019.

Tento přístup může být užitečný při řešení otázek, které současný regulační rámec řeší pouze okrajově, nebo je dokonce zcela pomíjí. Může jít také o případy, kdy je *socio-technologická změna* považována za příliš výraznou a specifickou na to, aby spadala pod stávající právní úpravu. V rámci tohoto modelu bychom se proto měli ptát, zda je *socio-technologická změna* ve svých důsledcích, které mohou být měřeny optikou *kritérií relevance*, tak závažná, že vyžaduje přístup *sui generis*.

Příklady užití tohoto modelu mohou být spatřovány v evropské regulaci, konkrétně v Aktu o digitálních službách (DSA) a v návrhu Aktu o umělé inteligenci (AIA). Tyto regulace také dobře ukazují na úskalí tohoto přístupu. Jde o pokusy o vysoce komplexní regulace, které mohou činit právní systém nepřístupný i vystudovaným právníkům. Zároveň taková regulace může být ve svých mnoha částech nadbytečná či může spadnout do zbytečného (technického) detailu. Tvorba podobné regulace je navíc nákladná a hledání konsensu mezi zapojenými účastníky komplikované. V případě české veřejné správy lze mít navíc opodstatněné pochyby o tom, zda by takovýto druh regulace vůbec dokázala vytvořit. Bude se proto pravděpodobně jednat spíše o pokusy na evropské úrovni. Pod tento přístup však lze teoreticky zahrnout i méně rozsáhlé české zákony, které by mohly plnit podobnou funkci.

Pro tento model je typické, že se na rozdíl od předchozích dvou občas zabývá socio-technologickými změnami, které mohou nastat až v budoucnosti. V takové situaci se právní systém snaží odhadnout rysy budoucího stavu socio-technologického prostředí. Někdy se takovému přístupu říká *proaktivní regulace* a je po ní často voláno například v mezinárodním právu.³⁷ Pokud se taková predikce podaří, situaci to výrazně zlepší, protože lze řešit kritické problémy ještě předtím, než nastanou. Na druhou stranu, pokud se budoucnost z jakéhokoli důvodu vymyká očekávání, mohou nám zůstat drahé, rozsáhlé a neužitečné regulace. Toto bude pravděpodobně častější případ.

Emergentní model přitom najde uplatnění také v oblasti autonomního smluvního práva. Dobrým příkladem je tzv. *Windfall Clause*³⁸, což je myšlenka dobrovolného závazku společností vyvíjejících umělou inteligenci, že darují významnou část svých zisků, pokud získají „*historicky bezprecedentní ekonomický neočekávaný zisk z vývoje pokročilé umělé*

³⁷ SHLOMO-AGON, S.; SALITERNIK, M. Proactive International Law Keeping Pace with an Accelerated World. In: *VerfBlog* [online]. 2022 [cit. 2022-11-15]. Dostupné na: <https://verfassungsblog.de/proactive-international-law/>.

³⁸ O'KEEFE, C.; CIHON, P.; FLYNN, C.; GARFINKEL, B.; LEUNG, J.; DAFOE, A. *The Windfall Clause: Distributing the Benefits of AI*. Centre for the Governance of AI Research Report, 2020. Future of Humanity Institute, University of Oxford. Dostupné z: <https://www.fhi.ox.ac.uk/windfallclause/>.

intelligence“.³⁹ Neočekávaný zisk se zde definuje jako úroveň příjmů vyšší než podstatná část (např. alespoň 1 %) celosvětového hospodářského výkonu. Mírou inovativnosti jde o právní řešení, které si zaslouží být v této kategorii, protože představuje zcela nový přístup v oblasti závazkového práva.⁴⁰

1.4 Praktické příklady

1.4.1. Desátá léta: Uber a Airbnb

Dobrym příkladem z posledního desetiletí je vztah platformy Uber ke dvěma různým regulacím ve dvou odlišných kontextech.

Za prvé, platforma Uber vedla k policy driftu na poli pracovního práva, neboť ve svém důsledku zhoršila ochranu práv zaměstnanců a ztížila jim možnost zakládat odbory.⁴¹ V tomto případě došlo z důvodu socio-technologické změny k zásadnímu policy driftu ve vztahu k pracovnímu právu, jenž zároveň naplňuje jedno z kritérií relevance – narušení sociální solidarity ve společnosti. Pracovní-právní ochrana z důvodu měnícího se společenského a technologického kontextu přestává plnit svůj původní účel, což ohrožuje sociální smír. Právní systém by tudíž na tuto změnu měl reagovat. I nečinnost je přitom reakcí *sui generis*, neboť i nečinnost vede k nevyhnutelné změně v působení právního systému na společnost.

Na druhou stranu, implementace těchto platforem sdílené ekonomiky vyvolala také policy drift ve vztahu k veřejnoprávní regulaci taxislužeb.⁴² V tomto případě však docházím k závěru, že žádné z *kritérií relevance* naplněno nebylo. Ve světě, kde existuje aplikace, která řidiče navádí po celou dobu jízdy, není třeba zkouška z cestopisu. Podobně aplikace fakticky znemožňuje okrádání zákazníků či nedůvodnou volbu delší trasy. Nová technologie a její společenská implementace proto učinila tuto regulaci do značné míry obsoletní, a tak by měla být zrušena. Nová technologie vedla ke snížení sociální complexity, jelikož učinila určité nežádoucí možnosti jednání fakticky nemožnými.

³⁹ Tamtéž, executive summary.

⁴⁰ Viz KNETL, Š. The windfall clause. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022, s. 135–142.

⁴¹ Viz např. ROGERS, B. The Social Costs of Uber. *University of Chicago Law Review Online*, 2015, vol. 82, no. 1, s. 85–102; CUNNINGHAM-PARMETER, K. From Amazon to Uber: Defining Employment in the Modern Economy. *B.U. L. REV.*, 2016, vol. 96, s. 1674-1727; THELEN, K. A. Regulating Uber: The Politics of The Platform Economy In Europe And The United States. *Perspectives on Politics*, 2018, vol. 16, no. 4, s. 938-953.

⁴² Viz zákon č. 111/1994 Sb., o silniční dopravě ve znění pozdějších předpisů.

Pokud jde o ubytovací platformy typu Airbnb, je tento přístup příznačný pro neschopnost orgánů veřejné moci uplatňovat na tyto poskytovatele služeb pravidla související s hotely nebo přijmout nové předpisy, které by se jimi zabývaly. Ačkoli by takové kroky byly možné, trvá roky než se k nim orgány veřejné správy, samosprávy či soudy odhodlají.⁴³ Taková nečinnost následně může zapříčinit zhoršení situace na trhu s bydlením, popřípadě neschopnost řešit přehlcení měst turisty.

V tomto případě se jedná o *socio-technologickou změnu*, která spočívá ve změně způsobu poskytování ubytování pro turisty. Původní regulace hotelů je opět předmětem technologicky zapříčiněného *policy driftu*, neboť její původní význam pro regulaci trhu s ubytováním se snížil. Na rozdíl od příkladu s platformou Uber ve vztahu k regulaci taxislužeb je zde však naplněno významné kritérium relevance – zajištění sociální solidarity.⁴⁴ Z tohoto důvodu není vhodná nulová varianta. V tomto případě by proto aplikace například law-centric modelu spočívala v analýze současné podoby veřejnoprávní regulace hotelů a v nalezení možností, jak tuto regulaci interpretovat způsobem, že zahrnuje také poskytovatele ubytování skrze tyto platformy.

1.4.2. Dvacátá léta: AI systémy a velká data

Nástup umělé inteligence v tomto desetiletí pravděpodobně přinese významné socio-technologické změny.⁴⁵ S touto technologií proto může být spojeno mnoho krátkodobých i dlouhodobých společenských výzev, na které bude potřeba reagovat. S rostoucí mírou autonomie těchto systémů může stále častěji docházet k selhání statických veřejných politik, k ohrožení občanských práv⁴⁶ a potenciálně i k existenčním rizikům pro lidstvo. Podstatná může být i provázanost s dalšími technologiemi, neboť pokrok v AI systémech může mít například významné dopady v oblasti robotiky, což by zase vyvolalo silné dopady na trh práce.

V Evropské unii přitom již v současnosti probíhají pokusy o regulaci dopadů umělé inteligence, které lze klasifikovat jako řešení v rámci emergentního modelu. Panuje shoda na tom, že rostoucí počet aplikací umělé inteligence vyžaduje regulační pozornost, což je aktuálně

⁴³ Toto řeší např. rozsudek Městského soudu v Praze ze dne 19. 8. 2021, č. j. 6 Af 20/2020-28. Zpoždění několika důležitých let je zde zcela očividné a bylo zapříčiněno neochotou orgánů samospráv tyto věci přehodnotit dříve.

⁴⁴ Viz např. SHABINA, Z. (eds.). Airbnb and its potential impact on the London housing market. *Urban Studies*, 2022, vol. 59, no. 1, s. 197–221.

⁴⁵ HARARI, Y. Reboot for the AI revolution. *Nature*, 2017, vol. 550, s. 324–327.

⁴⁶ COWGER, A. R. *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*. Lexington Books, 2020, s. 2.

přetaveno v návrh Aktu o umělé inteligenci. Toto nařízení rozděluje AI systémy do čtyř kategorií podle rizik, která s sebou nesou: a) nepřijatelné, a tudíž zakázané; b) vysoce rizikové pro zdraví, bezpečnost nebo základní práva, ale za určitých okolností ospravedlnitelné; c) s riziky ohledně transparentnosti a d) s nízkým nebo minimálním rizikem.

Ačkoli je návrh Aktu o umělé inteligenci novým předpisem, je zajímavé, že některá jeho ustanovení se již nyní zdají být nadbytečná. Příkladem z první kategorie nepřijatelného rizika je tak například zákaz AI systémů způsobujících psychickou nebo fyzickou újmu. Je však velmi obtížné si představit, jak by hypotetický systém způsobující psychickou nebo fyzickou újmu mohl být v souladu se stávajícími právními předpisy na ochranu spotřebitele⁴⁷ či se stávající ústavněprávní ochranou lidských práv, pokud by šlo o užívání této technologie veřejnou mocí. Jistě existují i méně jasné případy, například otázka, jak jsou před AI systémy chráněni podnikatelé. I tak mi ale přijde sporné, zda je nutné tuto věc řešit v novém komplexním nařízení, když by pravděpodobně stačilo několik výkladových změn (tzn. law-centric model) lépe reflektujících současnou podobu socio-technologické reality v národních právních řádech.

Vzhledem k tomu, že návrh Aktu o umělé inteligenci nepokrývá všechny oblasti, kterých se umělá inteligence dotkne, bude stejně nakonec nutné většinu problémů řešit prostřednictvím stávajícího právního rámce. Podle Petita patří mezi právní disciplíny, které se týkají AI systémů, zejména bezpečnost výrobků a odpovědnosti za ně, ochrana spotřebitele, duševní vlastnictví, medicínské právo, pracovní právo, ochrana soukromí, občanskoprávní odpovědnost, trestní odpovědnost, právní subjektivita, pojištění a daňové právo.⁴⁸ Je však jisté, že bychom zajímavé otázky našli i v ostatních částech právního řádu. V rámci korporálního práva bychom si například mohli představit právnickou osobu, kde úlohu zaměstnanců a statutárů může vykonávat AI systém – takový druh korporace však zákon o obchodních korporacích zatím nezná. Domnívám se proto, že problematika umělé inteligence brzy prostoupí – bez ohledu na Akt o umělé inteligenci – významnou část právního řádu.

Socio-technologická změna se bude často dotýkat trendů ve využívání AI systémů ve vztahu k velkým datům. Analýza těchto trendů může odhalit mnohé nedostatečnosti právní

⁴⁷ RILEY-SMITH, T.; MCCORMICK, L. Liability for Physical Damage. In: HERVEY, M.; LAVY, M. (eds.). *The Law of Artificial Intelligence*. Sweet & Maxwell, 2021, kapitola 5, s. 38–55.

⁴⁸ PETIT, N. Law and regulation of artificial intelligence and robots: conceptual framework and normative implications. Working paper. *SSRN*, 2017, s. 6.

regulace. Například obchodní model sociálních sítí poukázal na zásadní fenomén související s umělou inteligencí, který stávající právní úprava dobře nezachytila.

GDPR a další předpisy o ochraně osobních údajů totiž poměrně dobře chrání údaje lidí, které proudí od uživatelů k platformám.⁴⁹ Tyto předpisy však neřeší opačný tok informací: od platformy k uživatelům.⁵⁰ Tím se myslí situace, kdy doporučovací algoritmy aktivně poskytují pod maskou personalizace informace konkrétním uživatelům, což může negativně ovlivnit jejich autonomní postoje, blahobyt, a nakonec i demokratickou společnost jako celek.⁵¹ V této souvislosti existují v odborné literatuře názory, že doporučující AI systémy mohou uživatele tlačit do názorově extrémnějších pozic, protože pak jsou lépe předvídatelní. To pomáhá algoritmům plnit jejich partikulární cíl maximalizace času uživatelů stráveného na platformě.⁵²

Tento problém je vynikajícím příkladem toho, jak může i law-centric model pomoci řešit poměrně náročné problémy. Jedním z možných interpretačních řešení totiž může být zahrnout pod svobodu myšlení, jak je vyjádřena například v článku 9 EULP, také horizontální právo *nebýt obelháván* nebo *nebýt manipulován AI systémy*. O právu *nebýt obelháván AI systémy* v této souvislosti píše například Russell, který tento termín zavedl ve své knize *Human Compatible*.⁵³ Předpokládám, že by tato interpretační změna mohla sehrát zásadní roli nejen v otázkách manipulace na sítích sociálních médií, ale také pokud jde o pokročilejší budoucí algoritmy ovlivňující myšlení lidí obecně (například tzv. AI věštiny). K tomuto však více v kapitole přímo se věnující dlouhodobým aspektům působení umělé inteligence.

1.5 Závěr

Tato kapitola načrtla některá základní východiska vztahu práva, společnosti a nových technologií, z nichž některá budou podrobněji rozvedena v následujících kapitolách. Zároveň ukázala, proč přílišná konzervativnost práva může značně uškodit nejen společnosti, ale i samotné legitimitě právního systému. Věřím, že se tato kapitola může stát dobrým výchozím bodem pro uvažování o regulaci nových technologií obecně.

⁴⁹ TANNOCK, Q.; SUMROY, R. Data Protection and Privacy. In: HERVEY, M.; LAVY, M. (eds.). *The Law of Artificial Intelligence*. Sweet & Maxwell, 2021, kapitola 9.

⁵⁰ YEUNG, K. Why Worry about Decision-Making by Machine? In: YEUNG, K; LODGE, M. *Algorithmic Regulation*. Oxford: OUP, 2020, s. 34.

⁵¹ SUNSTEIN, C. R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2017, s. 1-97.

⁵² RUSSELL, S. *Human Compatible. AI and the Problem of Control*. London: Allen Lane, 2019, s. 9.

⁵³ Tamtéž.

Pro přehlednost přikládám modely a jejich charakteristiku v Tabulce 1:

Model	Nulový model	Law-centric model	Emergentní model
charakteristika	pasivní model zachovávající nezměněnou reprodukci právního systému	aktivní model usilující o zachování předchozího účelu již existující právní úpravy	radikální model zajišťující komplexní změnu prostřednictvím nového právního rámce
vztah k socio-technologickým změnám a technologicky podmíněným <i>policy driftům</i>	lhostejnost	zachovává stávající společenskou smlouvu řeší současné	může zachovat nebo změnit společenskou smlouvu řeší současné nebo budoucí
regulační techniky	žádné deregulace	reinterpretace drobné úpravy právního řádu	zavádění nových zákonů nebo institucí (významné legislativní změny)
vhodné pro	menší socio-technologické změny a technologicky podmíněné <i>policy drifty</i> s malými nebo žádnými riziky a negativními dopady	změny středního rozsahu vyžadující určitou regulační pozornost udržení rovných podmínek v odvětví	významné strukturální změny a hlubší nově vznikající problémy
existence nezbytnosti právního zásahu (naplnění kritéria relevance)	ne	ano	ano
nároky na regulační orgán a nezbytné odborné znalosti	žádné	nízké – střední	vysoké
příklady	veřejnoprávní regulace taxislužby	svoboda myšlení zahrnující <i>právo nebýt manipulován AI systémy</i>	Akt o umělé inteligenci Akt o digitálních službách Windfall Clause

výhody	<p>může poukázat na nadbytečnost právní úpravy</p> <p>nezvyšuje složitost právního systému</p> <p>nezatěžuje omezenou kapacitu pro změnu právního systému</p>	<p>je zvládnutelnější než tvorba nové komplexní regulace</p> <p>může probíhat na úrovni zákonodárců i vykladačů práva</p> <p>zachovává status quo a nemusí tedy řešit hodnotové otázky</p> <p>není tak nákladný jako Emergentní model</p> <p>nevede k hypertrofii práva</p>	<p>v případě významných změn může být jediným funkčním řešením</p> <p>může regulovat i silné nadnárodní korporace</p> <p>může řešit i jiné otázky než ty, které jsou nutně způsobeny zaváděním nových technologií (hodnotové otázky)</p>
slabiny	<p>vytváří právní nejistotu (s možným rdousivým efektem)</p> <p>umožňuje narušit stávající společenské smlouvy (obcházení zavedené regulace v daném odvětví)</p>	<p>neochota aktérů právních systémů provádět změny prostřednictvím výkladu</p> <p>nemusí být schopen řešit zásadní problémy</p>	<p>může vést k hypertrofii práva</p> <p>může neúměrně zvyšovat složitost právního systému</p> <p>nákladný</p> <p>riziko, že se normy brzy stanou nadbytečnými</p>

Tabulka 1 – Charakteristika modelů regulace

2. Právní a kulturní aspekty teze o technologické stagnaci⁵⁴

Dominantním narativem, převažujícím v dnešní západní společnosti, je přesvědčení, že žijeme v době nebývalého technologického pokroku, jenž historicky nemá obdoby. Máme sociální sítě, chytré telefony, internet a směřujeme k vyspělé umělé inteligenci.⁵⁵ Myslíme si, že se nacházíme v éře bezprecedentního technologického progresu, že se věci mění stále rychleji a že technologie mohou být velmi destruktivní záležitostmi, proto je lepší vůči těm novým zachovávat obezřetnost.

Tento narativ je však v posledních letech napadán řadou významných autorů jako Vijg nebo Thiel, jejichž perspektiva v českém a evropském veřejném prostoru schází.⁵⁶ Tito autoři poukazují na skutečnost, že progres byl v posledních zhruba 50 letech zaměřen především na informační technologie (a částečně biotechnologie), zatímco v ostatních oblastech, jako je energetika, doprava či stavebnictví, zažíváme výraznou stagnaci.⁵⁷ Jejich argument je komparativní a je založen na tvrzení, že se ve srovnání s první polovinou 20. století tempo technologických změn ve fyzickém světě výrazně snížilo.

Tato kapitola z tohoto paradigmatu vychází a klade si dva cíle. Prvním je seznámit čtenáře s tezí o technologické stagnaci. Druhým je analyzovat, jaké mohou být příčiny tohoto jevu a jakou roli v něm může sehrávat právní systém. Mluví-li se zde přitom o právním systému, moje pojetí má nejbližší, jak již bylo zmíněno dříve, k systémové teorii práva.⁵⁸ Toto pojetí dobře ilustruje, že mění-li se socio-technologické prostředí, ve kterém právo působí, mohou se společenské dopady vnitřně statického právního systému znatelně lišit od původně zamýšlených. Do pojmu právního systému přitom zahrnuji také faktickou institucionální aktivitu, nikoli pouze textaci právních norem a principů.

Ambicí této části není tento problém vyřešit ani čtenáře přesvědčit, že všechny argumenty v něm obsažené jsou nutně správné, nýbrž spíše na tuto problematiku upozornit

⁵⁴ Tato část vychází z článku HOŘEŇOVSKÝ, J. Právní a kulturní aspekty příčin a dopadů teze o technologické stagnaci. *Acta Universitatis Carolinae. Iuridica*, 2023, vol. 69, no. 1, s. 29–44.

⁵⁵ Viz např. KURZWEIL, R. *The Singularity Is Near*. New York: Penguin Books, 2006; HARARI, Y. N. *Homo Deus. A Brief History of Tomorrow*. Random House, 2017; BRYNJOLFSSON, E.; MCAFEE, A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. Norton, 2016.

⁵⁶ Viz např. VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century*. New York: Algora Publishing, 2011; GREF, L. *The Rise and Fall of American Technology*. New York: Algora Publishing, 2010; THIEL, P.; MASTERS, B. *Zero to One. Notes on Start Ups, or How to Build the Future*. London: Ebury Publishing, 2014; PHELPS, E. *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change*. Princeton University Press, 2016.

⁵⁷ BETZ, U. Is the force awakening? *Technological Forecasting & Social Change*, 2018, vol. 128, s. 299–300.

⁵⁸ LUHMANN, N. *Law as a Social System*. Oxford: Oxford University Press, 2004, s. 76–142.

českou a slovenskou odbornou veřejnost. Zároveň věřím, že tento text může otevřít dveře dalšímu výzkumu na poli jednotlivých oborů práva, včetně potenciální empirické analýzy. V závěru nabízím diskusi, ve které některé své hypotézy zpochybňuji.

2.1 Teze o technologické stagnaci v oblasti fyzického světa

V 60. a 70. letech 20. století byl progres v západním světě brán jako něco samozřejmého, o co se nebylo třeba příliš starat.⁵⁹ Převládající narativ ve společnosti byl, že technologie nemá žádné hranice.⁶⁰ Představy o budoucnosti byly velkolepé. Jaderná energie byla viděna jako příslib energie tak levné, že ji nebude dávat smysl ani měřit. Létající auta byla predikována jako blížká samozřejmost.⁶¹ Tento narativ přitom jistě nebyl vlastní pouze západním společnostem, nýbrž i tehdy komunistickému Československu.⁶²

Problém s přísliby budoucnosti druhé poloviny 20. století je však ten, že se jich ve skutečnosti příliš nenaplnilo. Nemáme energii v podstatě zadarmo a po našich ulicích se neprohánějí létající auta. Pořád však spalujeme uhlí, stejně jako před 150 lety a kvůli nedostatku ekologických a levných zdrojů energie čelíme závažné ekologické krizi. Pokud se podíváme, jak vypadají věci, které používáme v běžném životě (auta, domy, svoz odpadu, hromadná doprava atd.), a porovnáme je se stavem světa před 50 lety, uvidíme, že až na dílčí vylepšení žádné významné změny nenalezneme.⁶³ Toto se velmi liší od období mezi lety 1880–1930 a 1930–1980, kdy se svět změnil naopak velmi výrazně, a to ve všech aspektech lidských životů. Z tohoto úhlu pohledu lze argumentovat, že se v komparaci s druhou polovinou 19. století a první polovinou 20. století nacházíme v éře zpomalení technologického pokroku, tedy že se nacházíme v *éře nástupu relativní stagnace a zpomalení rychlosti pokroku*.

Nejde o to, že by se progres vůbec nekonal, ale že k němu zhruba od roku 1970 nedochází tak rychle jako v minulosti. Navíc se odehrává ve velmi limitovaných sektorech společnosti a ekonomiky.⁶⁴ Toto bývá dokládáno na empirických pozorováních založených na poklesu počtu nových makro-vynálezů, nově schválených léků atd.⁶⁵ Trend nástupu

⁵⁹ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 2.

⁶⁰ Tamtéž, s. 2.

⁶¹ THIEL, P. The End of the Future. In: *National Review* [online]. 3. 11. 2011 [cit. 2022-01-29]. Dostupné na: <https://www.nationalreview.com/2011/10/end-future-peter-thiel/>. [není stránkováno].

⁶² Viz např. text kapely Katapult *Až se bude psát rok 2006* nebo socialistická kinematografie té doby týkající se tehdy „nedaleké“ budoucnosti jako filmy „Zítřka vstanu a opařím se čajem“ nebo „Zabil jsem Einsteina, pánové...“.

⁶³ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 15–17.

⁶⁴ Tamtéž, s. 8.

⁶⁵ Tamtéž, s. 1–39.

technologické stagnace přelévající se do stagnace ekonomické je také empiricky podchycován některými ekonomy.⁶⁶ Není samozřejmě v možnostech tohoto textu přednést všechny argumenty podporující tuto tezi, přesto považuji za účelné seznámit čtenářstvo s některými údaji, na kterých bývá tento trend ilustrován. Pro začátek se zaměříme na dvě základní pozorování:

- Zatímco kumulativní počet vynálezů se od roku 1800 zvyšuje v podstatě exponenciálně, počet nových vynálezů za rok se od roku 1970 začal snižovat.⁶⁷
- Průměrný 10letý roční růst celkové produktivity výrobních faktorů v USA měl svůj vrchol v 50. letech 20. st. a od té doby setrvale klesá, s výjimkou 90. let, které přišly po takřka zcela stagnujících 80. letech.⁶⁸

Technologická stagnace se samozřejmě netýká všech oborů rovnoměrně, a některých dokonce vůbec. Existují oblasti, ve kterých je technologický rozmach velmi silný – svět internetu, softwaru, mobilních telefonů, počítačů a nově také umělé inteligence. Cowen k tomuto konstatuje: „*Nechci popírat nárůst produktivity tam, kde ho nacházíme, například v oblasti informačních technologií, ale obávám se, že tento nárůst je kompenzován poklesem produktivity v jiných oblastech ekonomiky.*“⁶⁹ Problémem je, že progres v informačních technologiích nemůže sám o sobě kompenzovat zpomalení v oblastech, jako je doprava, energetika či stavebnictví, které jsou pro životní úroveň lidí zcela klíčové. Betz v tomto kontextu pozoruje, že největší progres se v historii lidstva odehrál v oblastech výpočetní techniky, schopnosti ukládání informací a v oblasti komunikace, zatímco nejhůře se lidstvu daří posouvat zvyšování věku dožití, oblast stavebnictví a dopravy.⁷⁰

S tímto může úzce souviset pozorování o stagnaci růstu reálných příjmů amerických domácností.⁷¹ Reálný příjem amerických domácností zůstává po očištění inflací zhruba od 70.

⁶⁶ GORDON, R. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. New Jersey: Princetown University Press, 2016; PIKETTY, T. *Kapitál v 21. století*. Praha: Knižní klub, 2015.

⁶⁷ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 32.

⁶⁸ GORDON, R. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War...*, s. 547. [pozn. měřeno do roku 2014].

⁶⁹ COWEN, T. *The Great Stagnation: How America Ate All The Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*. New York: Penguin, 2010, ch. 2. [pozn. čteno jako e-kniha, přesnou stranu proto nelze identifikovat].

⁷⁰ BETZ, U. *Is the force awakening?...*, s. 299.

⁷¹ Pozn. využívat český příklad by bylo v této otázce nevyhovující vzhledem ke změně ekonomického režimu z centrálně plánovaného na tržní na přelomu 80. a 90. let. USA jsou vzhledem k nejlepším datům a největší kontinuitě nejlépe vypovídající příklad.

let 20. století stejný,⁷² výdaje na bydlení, zdravotní péči nebo vzdělávání přitom výrazně rostly.⁷³ Mezi lety 1948–1972 ročně rostl reálný příjem horních 10 % obyvatel USA v průměru o 2,46 % a u dolních 90 % obyvatel dokonce o 2,65 %. Mezi lety 1972–2013 toto číslo kleslo u horních 10 % na 1,42 % a u dolních 90 % nastal dokonce průměrný reálný pokles o 0,17 %⁷⁴

Nejen z těchto důvodů Thiel píše: „*Nejdůležitějším hospodářským jevem poslední doby je rozsáhlá stagnace reálných mezd a příjmů od roku 1973, kdy se ceny ropy zvýšily čtyřnásobně. V prvním přiblížení se zdá, že pokrok v oblasti počítačů a neúspěch v energetice se zhruba vzájemně vyrovnaly.*“⁷⁵ Jedno z potenciálních vysvětlení přitom spočívá právě v tezi o technologické stagnaci, která pokud je pravdivá, může znamenat závažný problém pro západní liberální demokracie, jež na příslibu růstu a zvyšování životní úrovně stojí. Někteří autoři dokonce ekonomické faktory, jako je stagnace příjmů, vnímají jako příčinu současného příklonu části veřejnosti k neliberálním politickým hnutím.⁷⁶

Technologický rozvoj může být samozřejmě nebezpečný. Je však důležité připustit možnost, že svět bez dostatečného technologického progresu může být ještě nebezpečnější. Progres v technologiích je zcela klíčový pro kontinuální nárůst bohatství a kvality života stále většího množství lidí, který dneska chápeme jako přirozeně daný.⁷⁷ Pokud chybí, ze společenského soužití se stává hra s nulovým součtem, ve které vždy někdo vyhraje a někdo prohraje.⁷⁸

Je zároveň účelné rozlišovat dva typy pokroku – horizontální a vertikální.⁷⁹ Podstatou horizontálního progresu je kopírování věcí, které fungují. Z hlediska posledních třiceti let je v tomto ohledu nejsilnějším prvkem *globalizace*.⁸⁰ Vertikální progres naopak spočívá ve vynalézání věcí zcela nových.⁸¹ Pokud je zde diskutována technologická stagnace, myslí se tím

⁷² VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 208; LEONHARDT, D. For many, a boom that wasn't. In: *The New York Times* [online]. 9. 4. 2008 [cit. 2022-30-1]. Dostupné na: <https://www.nytimes.com/2008/04/09/business/09leonhardt.html>.

⁷³ THIEL, P. The End of the Future... [není stránkováno].

⁷⁴ GORDON, R. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War...*, s. 609.

⁷⁵ THIEL, P. The End of the Future... [není stránkováno].

⁷⁶ GOODWIN, M.; HEATH, O. The 2016 Referendum, Brexit and the Left Behind: An Aggregate-level Analysis of the Result. *The Political Quarterly*, 2016, Vol. 87, no. 3, s. 323–332.

⁷⁷ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 20.

⁷⁸ THIEL, P. The End of the Future... [není stránkováno].

⁷⁹ THIEL, P.; MASTERS, B. *Od nuly k jedničce. Úvahy o startupech aneb jak tvořit budoucnost*. Brno: Jan Melvil Publishing, 2015, s. 14–16.

⁸⁰ Tamtéž, s. 14–16.

⁸¹ Tamtéž, s. 14–16.

tedy stagnace vertikální, která se projevuje v úpadku společenské inovativnosti ve fyzickém světě, nikoli absence pokroku skrze rozšiřování již existujících technologií. Nejde tedy o to, že bychom nedokázali rozšiřovat a vylepšovat již existující technologie, ale o skutečnost, že přestáváme být inovativní společností vyvíjející a implementující technologie zcela nové. Inovativní společnost, kterou bychom měli chtít být, proto můžeme definovat jako „společnost s vysokou mírou rozvoje, která adoptuje nové transformativní technologie“.⁸² Opět však musím zopakovat, že se tato teze netýká světa počítačů a softwaru, nýbrž výlučně tzv. fyzického světa.

2.2 Narůstající vnitřní komplexita práva a problém přeregulovanosti

Za prvé je nutno zmínit možnou souvislost tohoto jevu s problémy přeregulovanosti a narůstající vnitřní komplexity práva, které jsou důsledkem představy, že na každý lidský problém existuje vhodné právní řešení.⁸³ Takový legislativní optimismus je katalyzátorem nových regulací, které mohou představovat překážku pro implementaci nových technologií. Zároveň jsou nové technologie a na ně navazující společenské změny také důležitou příčinou hypertrofie práva.⁸⁴ Jde o kruh, který je postaven na představě právního řádu jako všeobjímajícího abstraktně-rationálního systému. V takové představě platí přímá úměra, že čím více externí svět nabývá na komplexitě, tím více komplexním se musí stávat i právo samotné.

Komplexita přitom samozřejmě nemusí být vždy špatnou věcí.⁸⁵ Co lze v současném právu spatřovat a kritizovat, bych proto spíše nazval *neřízená komplexita*, tzn. komplexita, která je „utržena z řetězů“. Jde o situace, kdy právní regulace upadá do kombinatorní exploze a za 1) si toho není vědoma, a za 2) neví, jak tuto situaci řešit jinak než prohlubováním tohoto problému. Problém neúměrného zvyšování komplexity v právu ve vztahu k inovativnosti spočívá ve skutečnosti, že nevyhnutelně vytváří prostory vnitřní rozpornosti, nekonzistence, nejistoty v tom, kterou normu užít, a nejistoty v tom, co který termín znamená.⁸⁶ Tento stav vede k právní nejistotě v základní dichotomii práva, tedy v rozlišení, zda je něco legální či nelegální.

⁸² VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 22.

⁸³ SUMPTION, J. *Trials of the State*. London: Profile Books, 2020, s. 3–19.

⁸⁴ HOLLÄNDER, P. Právnik v bludišti postmoderní dekonstrukce a Savignyho Ariadnina nit. In: GERLOCH, A.; KRZYŽANKOVÁ, K. Ž. (eds.). *Právo v měnícím se světě*. Plzeň: Aleš Čeněk, 2020, s. 44.

⁸⁵ SCHUCK, P. H. Legal Complexity: Some Causes, Consequences, and Cures. *Duke Law Journal*, 1992, vol. 42, no. 1, s. 8.

⁸⁶ HOLLÄNDER, P. Právnik v bludišti postmoderní dekonstrukce a Savignyho Ariadnina nit..., s. 41–56.

Důsledkem je právní regulace, která je často předimenzovaná, rozporuplná nebo často dokonce úplně zbytečná. Z práva se tak stává obraz dekonstrukce a fragmentace, namísto jednotného normativního systému.⁸⁷ Stav práva, vyznačující se příliš vysokou mírou komplexity a z ní vyvěrající právní nejistoty, jak argumentuje Schuck, příliš nemotivuje ke zkoušení nových věcí a vynalézání a společenské implementaci nových technologií. Naopak vede k pasivitě, kterou takový systém aktivně motivuje.⁸⁸ Zde můžeme spatřovat potenciálně první propojení teze o technologické stagnaci s právem.

Je-li toto propojení skutečné, tak pokud bychom chtěli podpořit technologický vývoj, potřebovali bychom snížit množství a komplikovanost regulace, a to právě v těch oblastech, které identifikujeme jako stagnující. Thiel takto například argumentuje, že jeden z důležitých důvodů rozmachu informačních technologií spočíval v prvotní absenci regulace,⁸⁹ která byla pomalu dotvářena až v momentě, kdy byly nejdůležitější vynálezy v této oblasti objeveny a infrastruktura vybudována. Nepoměr mezi přeregulovaným fyzickým světem a zprvu neregulovaným světem informačních technologií vedl k orientaci investic a talentu směrem ke světu počítačů, což zde posléze multiplikovalo úspěchy (a snižovalo je v jiných oblastech). Stojí přitom za upozornění, že v posledních letech se právní regulace snaží dosáhnout i svět informačních technologií.⁹⁰ Je však otevřenou otázkou, zda to ve výsledku nepovede ke snížení pokroku i tam, kde dosud fungoval velmi dobře.

2.3 Účel právní úpravy a narůstající rezistence společnosti vůči technologiím

Nejde samozřejmě pouze o množství regulace, nýbrž také o účel, který právní úprava sleduje. Určitá míra regulace je nezbytná, protože alternativa v podobě přenechání sporů pouze na rozhodovací činnost soudů a správních orgánů je velmi nákladná a nezaručuje *ex ante* právní jistotu.⁹¹ Otázkou však je, kolik regulace je ještě snesitelné mít a jaký účel má regulace plnit. Zda má být jejím primárním cílem ochrana bezpečí a eliminace rizika takřka za jakoukoli cenu, nebo naopak podpora inovativnosti, například skrze ochranu před potenciální odpovědností vynálezců. S tímto může souviset kulturní vnímání společnosti, které se nevyhnutelně promítá

⁸⁷ Tamtéž, s. 41–56; k tomuto také viz ONDŘEJEK, P. *Defragmentation of Law. Reconstruction of Contemporary Law as a System*. Cambridge: Intersentia, 2023.

⁸⁸ SCHUCK, P. H. *Legal Complexity: Some Causes, Consequences, and Cures...*, s. 19.

⁸⁹ THIEL, P.; MASTERS, B. *Od nuly k jedničce. Úvahy o startupech aneb jak tvořit budoucnost*, s. 75.

⁹⁰ Např. GDPR, DSA, AIA atd.

⁹¹ GLAESER, E.; SHLEIFER, A. The Rise of the Regulatory State. *J. Econ Literature*, 2003, vol. 41, no. 2, s. 402–403.

do práva. Právní systém je systémem kognitivně otevřeným⁹² a mnohé kulturní obavy společnosti proto nepřímo přenáší skrze své aktéry do svého vnitřního rozhodování a jazyka.

Trend technologické stagnace mimo oblast počítačů a softwaru může být proto také poháněn frustrací spojenou s narůstající resistencí společnosti k implementaci nových vynálezů v oblastech, jako je doprava, energetika nebo zemědělství.⁹³ Tuto resistenci lze sledovat například v několika desítek let trvajících protestech vůči jaderné energii nebo v nelibosti společnosti vůči akceptaci geneticky modifikovaných potravin, v obou případech v podstatě bez racionálního vědeckého základu.⁹⁴ Tento společenský tlak se samozřejmě přeneseně dotýká i činnosti regulátorů a justice.

V americkém kontextu uvádí Jaffe jako příklad tohoto fenoménu, že doba rozhodování amerických regulačních orgánů o geneticky modifikovaných plodinách se mezi lety 1994 a 2005 více než zdvojnásobila, a to bez jakéhokoli vysvětlení.⁹⁵ Mnohdy se také jedná o zásah soudu, který regulačnímu orgánu může například vyčíst, že dostatečně nepodložil bezpečnost nové technologie. Takto se tomu stalo například v USA, když jeden z federálních soudů rozhodl, že vláda před schválením pěstování geneticky modifikované cukrové řepy nedostatečně posoudila dopady této plodiny na životní prostředí.⁹⁶ Tento argument však lze užít v podstatě vždy a na cokoli.

Zajímavým příkladem tohoto jevu z práva EU je zase poměrně nový rozsudek SDEU C-528/16, ve kterém soud konstatoval, že směrnice Evropského parlamentu a Rady 2001/18/ES ze dne 12. března 2001 o záměrném uvolňování geneticky modifikovaných organismů do životního prostředí „*musí být vykládána v tom smyslu, že organismy získané technikami/metodami mutagenese představují geneticky modifikované organismy ve smyslu uvedeného ustanovení*“. Takový závěr vede k situaci, že je na metodu mutagenese aplikována striktní regulace, která omezuje technologický vývoj na území EU.⁹⁷ Tento závěr soudu přitom šel proti stanovisku generálního advokáta Michala Bobka a dle mnohých vědců nemá podklad

⁹² BAXTER, H. Niklas Luhmann's Theory of Autopoietic Legal Systems. *Annual Review of Law and Social Science*, 2013, vol. 9, s. 169.

⁹³ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 9.

⁹⁴ K jaderné energii viz PINKER, S. *Enlightenment Now*. London: Allen Lane, 2018, s. 145–152. Ke geneticky modifikovaným potravinám viz VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 104–105.

⁹⁵ JAFFE, G. Regulatory slowdown on GM crop decisions. *Nat Biotechnol*, 2006, vol. 24, no. 7, s. 748–749.

⁹⁶ Rozhodnutí The United States District Court for the Northern District of Carolina Center for food safety et al. v. Thomas J. Vilsack et al. ze dne 13. 10. 2010, No. C 08-00484 JSW.

⁹⁷ CALLAWAY, E. EU law deals blow to CRISPR crops. In: *Nature* [online]. 25. 7. 2018 [cit. 2022-01-29]. Dostupné na: <https://www.nature.com/articles/d41586-018-05814-6>.

ve vědeckých faktech. Zjednodušeně, na nebezpečnější techniku náhodné genetické modifikace se tato regulace neuplatní, ale na bezpečnější a člověkem řízenou mutagenezi ano. Z pohledu právního realismu zde lze dobře vidět, jak se resistance společnosti vůči GMO přenáší do evropského právního rozhodování a tím zpomaluje technologický rozvoj.

Vijg tento západní fenomén analyzuje následovně: „*Současné snadné podlehnutí úřadů tlaku veřejnosti (často proti veškeré vědecké a právní logice), je silně v neprospěch technologií při pokusech o jejich společenské uplatnění.*“⁹⁸ Vůle společnosti přijmout nové vynálezy je pro inovativnost společnosti stejně důležitá, jako samotná schopnost tyto vynálezy sestrojít.⁹⁹ Pokud bude společnost nové technologie odmítat, může to technologický progres silně limitovat. A vzhledem k tomu, že jsou soudci a regulátoři kulturně provázáni s prostředím, ve kterém žijí, dostávají se zprostředkovaně tyto obavy také do práva. Základním imperativem regulace je totiž v současnosti v mnoha oblastech důraz na bezpečnost, nikoli na inovativnost.¹⁰⁰ Mezi těmito hodnotami však existuje silný trade-off.

V této věci samozřejmě hrají velmi negativní roli medializované havárie, které obvykle vedou k další a restriktivnější regulaci.¹⁰¹ Dobrým příkladem tohoto jevu je oblast jaderné energie, ve které jsme se s požadavky na bezpečnost dostali tak daleko, že je dnes prakticky nemožné v západním světě postavit novou jadernou elektrárnu.¹⁰² Nepopírám přitom, že tyto technologie v sobě rizika a problémy obsahují. Například otázka jaderného odpadu je beze sporu velmi důležitá. Trvám však na tom, že se diskuse o těchto otázkách nevede v racionální rovině a že se dává přednost emocím a přílišnému důrazu na bezpečnost, a to i za cenu stagnace a úpadku těchto velmi důležitých technologických oblastí.

2.4 Narůstající averze k riziku ve společnosti a právní regulaci

Obecně jde v právu, mimo jiné, o otázku, jak vyvážit vztah mezi rizikem a příležitostmi. V západním světě je tato otázka ovlivněna jak kulturními, tak psychologickými faktory.¹⁰³

⁹⁸ VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 107.

⁹⁹ Tamtéž, s. 22.

¹⁰⁰ Např. FINDLAY, T. *Nuclear Energy and Global Governance: Ensuring Safety, Security and Non-proliferation*. Routledge, 2010.

¹⁰¹ BENZ, E. Lessons from Fukushima: Strengthening the International Regulation of Nuclear Energy. *Wm. & Mary Env'tl. L. & Pol'y Rev.* 845, 2013, vol. 37, issue 3/8, s. 845–883.

¹⁰² POTTER, B. Why Does Nuclear Power Plant Construction Cost So Much? In: *IFP* [online]. 1. 3. 2023 [cit. 2023-11-15]. Dostupné na: <https://ifp.org/nuclear-power-plant-construction-costs/>.

¹⁰³ GRUSZCZYNSKI, L. *Regulating Health and Environmental Risks under WTO Law: A Critical Analysis of the SPS Agreement*. Oxford: Oxford University Press, 2010, s. 20.

Silným kulturním faktorem je tzv. *Frankensteinův komplex*, tedy jakási západní¹⁰⁴ obava, že by se vlastní výtvor mohl jednoho dne obrátit se proti svému tvůrci a zvítězit nad ním.¹⁰⁵ Tento strach je přitom silně zakořeněn v západní kultuře až do té míry,¹⁰⁶ že se dostává přímo do oficiálních dokumentů Evropské unie.¹⁰⁷ Tento trend je spojený se širším jevem v právní regulaci, a to stále častějším preferování bezpečí na úkor jiných hodnot.¹⁰⁸ To má největší vliv na právní odvětví týkající se zaměstnanosti, ochrany osobních údajů, zdraví, bezpečnosti nebo ochrany spotřebitele.¹⁰⁹ Sumption tuto problematiku analyzuje následovně:

„V posledních letech se změnila míra rizika, kterou jsou lidé ochotni ve svém životě tolerovat. Na rozdíl od našich předků už nejsme ochotni přijmout kolo štěstěny jako běžnou událost lidské existence. Fyzické, finanční a emocionální bezpečí považujeme nejen za normální stav věcí, nýbrž i za nárok.“¹¹⁰

Zde se však dostáváme k paradoxnímu zjištění, na které upozorňuje Vijg. Svět se ve spoustě atributů kvality lidských životů zlepšuje.¹¹¹ Tento progres je přitom do značné míry zapříčiněn nárůstem regulace, silnějších vlád a větším důrazem na bezpečnost a snižování utrpení ve světě. Jde o nástroje, které do správy věcí veřejných přinášejí důraz na racionální stabilitu.¹¹² Tytéž mechanismy, které nám zapříčinily tento progres, nám ale také paradoxně zabraňují učinit stejné vědecko-technologické objevy, které šlo realizovat v naší méně racionální minulosti.¹¹³ Svět, který je bezpečný a relativně hojný, nám jednoduše může bránit podstupovat kroky, které mohou být pro náš budoucí úspěch zcela nezbytné.

Weimer v tomto kontextu ukazuje, že vůbec samotný pojem rizika v sobě obsahuje dvě roviny. Jedna rovina jej vnímá jako výdobytek modernity, který nám umožňuje se na

¹⁰⁴ Příkladem opačného vnímání této problematiky je např. kulturní oblast Japonska a obecně východní Asie.

¹⁰⁵ LEHMAN-WILZIG, S. Frankenstein unbound: Towards a legal definition of artificial intelligence. *Futures*, 1981, vol. 13, no. 6, s. 442–457.

¹⁰⁶ PETIT, N.; DE COOMAN, J. Models of Law and Regulation for AI. *European University Institute. Robert Schuman Centre for Advanced Studies*, 2020, RSCAS 2020/63, s. 8.

¹⁰⁷ COMMITTEE ON LEGAL AFFAIRS (EU). Report with recommendations to the Commission on Civil Law Rules on Robotics. 2017. Dostupné z: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html.

¹⁰⁸ SUMPTION, J. *Trials of the State...*, s. 3–19; VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 70–75.

¹⁰⁹ SUMPTION, J. *Trials of the State...*, s. 14.

¹¹⁰ Tamtéž, s. 16–17.

¹¹¹ Viz PINKER, S. *Enlightenment Now...*, s. 1–558; ROSLING, H.; ROSLING, O.; ROSLING RÖNNLUND, A. *Faktomluva: deset důvodů, proč se mýlím v pohledu na svět – a proč jsou věci lepší, než vypadají*. Brno: Jan Melvil Publishing, 2018.

¹¹² VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century...*, s. 10.

¹¹³ Tamtéž, s. 10.

racionálním základě vypořádávat s neurčitou budoucností.¹¹⁴ Z druhé roviny jde o moderní hazard, který pramení z přijímání nových technologií a jehož důsledky nemůžeme kontrolovat.¹¹⁵ Z této dichotomie vyplývá, že jakákoli snaha riziko zcela potlačit bude neúspěšná, neboť vždy povede k riziku jiného druhu. Pokud je společnosti zcela znemožněno vystavit se technologickým rizikům, může být neschopna přizpůsobit se novým – dosud nepředvídatelným – nebezpečím.¹¹⁶

Abych tuto tezi uvedl na příkladě, vraťme se k již zmíněné problematice jaderné energie. Spalování uhlí je konzervativní a na první pohled relativně bezpečná cesta výroby energie. Naproti tomu jaderná energie se může jevit na první pohled rizikovější. Pokud se ovšem zcela vyhneme rizikům spojeným s užitím jaderné energie, nemusíme být schopni reagovat na „nové nebezpečí“, například v podobě změny klimatu. Konzervativní averze k riziku se zde proto paradoxně může z pohledu nepředvídatelné budoucnosti ukázat jako více riziková než „nebezpečné hraní si s jádrem“. V oblasti komparace uhelné a jaderné energie je věc ještě komplikovanější, protože u uhelné energie většinou nepočítáme lidi, kteří zemřeli na znečištění ovzduší, což by celou statistiku bezpečnosti významně proměnilo.¹¹⁷

Wildavsky proto dochází k závěru, že bychom: „*Měli [...] zvolit kritérium čistého přínosu: více dobra než zla. Požadavek, aby výrobky nikomu neškodily, znamená nejen velké náklady, [...] rozsáhlou byrokracii a méně inovací, ale také horší zdraví.*“¹¹⁸ Toto kritérium mi přijde velmi rozumné. Že přehnaná kultura bezpečí může být nebezpečná, dobře ilustrují například nevyužité nově vyvinuté léky. Zde obvykle hrají roli dva faktory – některé léky jednoduše nejsou tržně profitabilní, což ve spojení s narůstající regulativní zátěží akcentující stále větší důraz na bezpečnost a etičnost výzkumu vede k tomu, že spousta z nich nakonec skončí nevyužita.¹¹⁹ Neříkám, že by vývoj léků měl být zcela neregulovaný, ale možná jsme již překročili únosnou míru. V tomto kontextu si vypůjčím ještě jednu delší citaci Sumptiona:

„Neštěstí, která se našim předkům zdála být nevyhnutelná, se nám zdají být snadno odstranitelná. Jakmile se na ně pohlíží jako na odvratitelné důsledky

¹¹⁴ WEIMER, M.; MARIN, L. The Role of Law in Managing the Tension between Risk and Innovation. *EJRR, Special Issue on Regulating New and Emerging Technologies*, 2016, vol. 7, issue 3, s. 12.

¹¹⁵ Tamtéž, s. 12.

¹¹⁶ WILDAVSKY, A. *Searching for Safety*. New Brunswick: Transaction, 1988, s. 5–6.

¹¹⁷ PINKER, S. *Enlightenment Now...*, s. 145–152.

¹¹⁸ WILDAVSKY, A. *Searching for Safety...*, s. 5–6.

¹¹⁹ STEWARD, D. (eds.). Equipose lost: ethics, costs, and the regulation of cancer clinical research. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 2010, vol. 28, s. 2925–2935.

lidského jednání, mají tendenci stát se vhodným předmětem pro přisouzení právní odpovědnosti. Po každém neštěstí tak máme sklon myslet si, že zákon buď musel být porušen, nebo není dostatečně přísný. Hledáme právní prostředky nápravy: žalobu, trestní stíhání nebo další legislativu. „Měl by na to být zákon“, zní všeobecné volání. Obvykle je, nebo brzy bude. Samozřejmě, že zákon ve skutečnosti neposkytuje řešení pro každé neštěstí. Očekává, že se lidé v rámci možností postarají o své vlastní zájmy. Předpokládá, že některá rizika bude možná nutné přijmout, protože společenské a ekonomické náklady na jejich odstranění jsou prostě příliš vysoké. Očekávání veřejnosti jsou však silným motorem právního vývoje. Soudci nerozhodují o případech podle stavu veřejného mínění. Je však jejich povinností zohlednit hodnoty společnosti, které slouží. Jednou z těchto hodnot se stala averze k riziku.“¹²⁰

Historickým příkladem regulace excesivně bránící jakémukoli riziku, která poškodila technologický progres své země na mnoho desítek let, je *Red Flag Act* z roku 1865 z Velké Británie, který se snažil regulovat zavádění automobilů ve městech. Tento zákon zavedl velice nízkou maximální rychlost v městských aglomeracích a také vyžadoval, aby chodec nesl červenou vlajku 60 yardů před vozidlem.¹²¹ Tato pravidla byla natolik restriktivní, že nikdo ve městě v autě jezdit nechtěl, což v důsledku poškodilo britský automobilový průmysl. Poučení z tohoto příkladu je, že můžeme v přehnané obavě z rizika přijmout regulaci, která razantně potlačí technologický rozvoj. To by přitom mohlo být pro naši civilizaci do budoucna fatální, neboť k překonání například klimatické krize nové vynálezy bezpodmínečně potřebujeme.¹²² Tento případ nám také ukazuje obecnější trend, že eliminace rizika nemůže být pro právo nikdy absolutní hodnotou, pouze otázkou míry.¹²³

Klíčovou roli zde také sehrávají regulátoři, kteří mají prevenci rizika zakódovanou ve své incentivní struktuře. Na případě léčiv však můžeme vidět i odvrácenou stranu této skutečnosti, a to za předpokladu, že je jakákoli újma dávana za vinu nedostatečnému kontrolnímu mechanismu. V takovém případě totiž není rozhodováno na základě principu

¹²⁰ SUMPTION, J. *Trials of the State...*, s. 16–17.

¹²¹ PETIT, N. *Law and Regulation of Artificial Intelligence and Robots – Conceptual Framework and Normative Implications*. SSRN, 2017, s. 12–13.

¹²² Viz GATES, B. *How to Avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need*. New York: Alfred A. Knopf, 2021.

¹²³ SUMPTION, J. *Trials of the State...*, s. 14–15.

poměrování „více dobra než zla“, nýbrž je požadována absolutní bezpečnost, která ale ve výsledku může stát životy mnohem většího množství lidí.

Epstein zajímavě upozorňuje, že v USA bylo nejvíce nových léků schváleno v době, kdy federální regulátor neměl pravomoc regulovat výrobu léků v jednotlivých státech.¹²⁴ Jedním z těchto případů je příklad objevení podávání inzulínu na diabetes, který byl předtím smrtelnou chorobou, na níž umíralo ohromné množství lidí. Od nápadu v roce 1921 do masové produkce v roce 1923 tehdy uběhly pouze tři roky, přičemž velkou roli sehrálo štěstí a absence legislativních a etických omezení, neboť celý proces výzkumu by dnes nebyl možný.¹²⁵ Do doby objevení inzulínu byla přítom běžná léčba hladovění. Skutečnost, že by v dnešní době léčba inzulínem nemohla vzniknout a již určitě ne v takto krátké době, by nás měla varovat, že můžeme v současnosti přicházet o podobně důležité objevy.

Vztah regulace a aktivity regulátorů je navíc v určitém rozporu se svobodou jednotlivce rozhodovat o tom, jak by měl být léčen.¹²⁶ V USA přítom rozhodovací praxe soudů málem judikovala ústavní právo jednotlivce sám se rozhodnout o využití potenciálně život zachraňující léčby, pokud prošla alespoň první fází schvalování.¹²⁷ Toto právo však nakonec judikováno nebylo.¹²⁸ Zatímco tak ve zdravotnickém právu platí *právo odmítnout léčbu*, to samé nelze říct o *právu léčbu přijmout bez ohledu na názor veřejné moci*.¹²⁹ To je přítom dle mého názoru v rozporu se základní zásadou medicínského práva – respektem vůči autonomii pacienta a jeho svobodnému a informovanému rozhodnutí.¹³⁰

2.5 Mění se socio-technologický kontext a vnitřně statický právní systém

Limitem pro inovativnost také může být právní regulace, která nereflektuje důležité změny ve společnosti (viz první část této práce). Že právní regulace a její dopad v čase není ve vztahu k inovativnosti neutrální, lze dobře vidět v oblastech práva duševního vlastnictví a soutěžního práva. Tyto oblasti také ukazují, že mění-li se sociální kontext, může být kdysi

¹²⁴ EPSTEIN, R. Can Technological Innovation Survive Government Regulation? *Harvard Journal of Law and Public Policy*, 2013, vol. 36, no. 1, s. 97.

¹²⁵ Tamtéž, s. 97–98.

¹²⁶ Tamtéž, s. 100.

¹²⁷ Tamtéž, s. 100.

¹²⁸ Viz *Abigail Alliance for Better Access to Developmental Drugs v. von Eschenbach*, 495 F.3d 695 (D.C. Cir. 2007), cert denied, 552 U.S. 1159 (2008).

¹²⁹ Viz EPSTEIN, R. The Erosion of Individual Autonomy in Medical Decision making: Of the FDA and IRBs. *Georgetown Law Journal*, 2007, vol. 96, no. 2, s. 559–582.

¹³⁰ BRAZIER, M.; CAVE, E. *Medicine, Patients and the Law*. London: Penguin Books, 2011, s. 64.

dobře, pro-inovativně, míněná právní úprava ve svém důsledku v nových poměrech kontraproduktivní.

De Beer takto například argumentuje k nutnosti změny paradigmatu v právu duševního vlastnictví a právu hospodářské soutěže, kdy na základě svého empirického výzkumu tvrdí, že „*panuje stále větší shoda na tom, že práva duševního vlastnictví mohou být často překážkou inovací, zejména sekvenčních a kumulativních inovací*“.¹³¹ Dnešní technologický svět je dle jeho názoru v oblasti inovací totiž mnohem více určován otevřeností platforem a spoluprací; je založen na filosofii, že pokud máme víc informací, které proudí mezi různými subjekty, tak máme lepší prostředí pro technologický progres.¹³² Tento společenský posun však není příliš reflektován ve výše zmíněných právních odvětvích.

Historicky, v době vzniku práva duševního vlastnictví a práva hospodářské soutěže, se totiž vývoj a výzkum odehrávaly uzavřeně v jednotlivých obchodních společnostech, které posléze nárokovaly patenty a dobývaly trh. V současnosti je však situace v nejvíce inovativních prostředích odlišná. Produkty jsou čím dál tím komplexnější, což vyžaduje spolupráci mezi společnostmi a výzkumnými týmy. Nelze zaměstnávat všechny odborníky na jednom místě; uzavřenost proto mnohdy vede k menšímu vědeckému výkonu. Zároveň je dnešní technologie často postavena na myšlence *platformy*.¹³³

Pokud je však pro inovativnost v dnešní době klíčová spolupráce (což je minimálně důležitá hypotéza), co nám to říká například o současné podobě evropského soutěžního práva a práva duševního vlastnictví? Půjčím si zde k odpovědi citaci de Beera:

„Ideologie, že hospodářská soutěž je nejlepším způsobem podpory inovací, je implicitně základem nejen duševního vlastnictví, ale i hospodářské soutěže a dalších rámcových politik trhu. Základním předpokladem politiky duševního vlastnictví je, že firmy nebudou investovat do inovací, pokud nebudou chráněny před konkurencí omezenými monopolními právy. Důkazy nyní ukazují, že tento

¹³¹ DE BEER, J. Intellectual Property and ‘Open’ Innovation: A Synthesis of Concepts. In: CALBOLI, I.; MONTAGNANI, M. R. (eds.). *Handbook of Intellectual Property Research*. Oxford: Oxford University Press, 2021, s. 733.

¹³² Tamtéž.

¹³³ Viz např. IZZO, M. Management Transition to Big Data Analytics: Exploratory Study on Airline Industry. *International Business Research*, 2019, vol. 12, no. 10, s. 48–56.

*předpoklad je příliš zjednodušený a nelze jej zobecnit; mnoho uživatelů a firem bude inovovat i bez práv duševního vlastnictví.*¹³⁴

Pokud se tedy společenské předpoklady, na kterých jsou právní odvětví postaveny, výrazně změnily, je nutné zahájit diskuse o vnitřní změně právního systému. Jinak může mít právní regulace, vzhledem k socio-technologické změně, opačný účinek, než jaký původně zamýšlela. Z podpory inovativnosti se stane její brzda.

Podobný problém „neaktualizace práva“ lze vnímat také v kontextu evropského soutěžního práva ve vztahu k technologickým společnostem. Na jedné straně chceme, aby byly evropské obchodní společnosti dostatečně velké, aby byly schopny soutěžit na globálním technologickém trhu. Na straně druhé nechceme, aby byly až moc velké, protože by to mohlo ohrozit konkurenci na vnitřním trhu EU. Problémem je, že se tyto cíle do jisté míry zároveň vylučují. Z teoretického pohledu je dobré si uvědomit, že jde o podobnou situaci, jaká nastala u změny způsobu, jak technologický progres probíhá. Pokud se změní společenské podmínky, může to ovlivnit dopad vnitřně statické právní regulace na společnost.

2.6 Diskuse

Roli práva a regulace v kontextu technologické stagnace je potřeba dát do souvislosti také s faktory kulturními, mezi které Vijg řadí například rostoucí vliv korporativismu v západních společnostech – tedy propojení státu a zájmových skupin s partikulárními zájmy, které jsou posléze implementovány do veřejných politik.¹³⁵ Dále identifikuje problém dysfunkčních firemních kultur a přílišného zaměření se na krátkodobou perspektivu s důrazem na plnění krátkodobých kvartálních plánů a na dosahování rychlých zisků na úkor dlouhodobého růstu a skutečné inovativnosti.¹³⁶ V neposlední řadě upozorňuje na problém „posedlosti penězi“, přílišné konzervativnosti mladých lidí ve volbě povolání a na fenomén nárůstu veřejného aktivismu, který v mnoha oblastech lobbuje proti jakýmkoli rizikům spojeným s implementací pokročilých technologií.¹³⁷ Pro Thiela je zase klíčovým kulturním faktorem naše současná filozofie ve vztahu k budoucnosti, kterou nazývá „neurčitým optimismem“ (budoucnost bude dobrá, ale nevíme proč a jak toho dosáhnout).¹³⁸ Argumentuje

¹³⁴ DE BEER, J. *Intellectual Property and 'Open' Innovation: A Synthesis of Concepts...*, s. 744.

¹³⁵ VIJG, J. *The Great Technology Slowdown: Looking for the Silver Lining*. [přednáška]. Praha: Institute H21, 18. září 2021. Dostupné na: <https://www.youtube.com/watch?v=eqsCNs8trTc>.

¹³⁶ Tamtéž.

¹³⁷ Tamtéž.

¹³⁸ THIEL, P.; MASTERS, B. *Od nuly k jedničce. Úvahy o startupech aneb jak tvořit budoucnost*, s. 70–80.

přítom, že neurčitý optimismus nebo neurčitý pesimismus nejsou vhodné modely pro technologický pokrok – tím je pouze určitý optimismus.¹³⁹

První otevřenou otázkou tedy je, do jaké míry jde o kulturní problém a jak velkou roli v něm skutečně sehrává právo. Může se totiž ukázat, že kulturní faktory převažují a že role práva a regulace je ve skutečnosti marginální. Zároveň však tato kapitola ilustrovala, že se kulturní trendy nevyhnutelně přelévají do praxe regulátorů a že ovlivňují podobu a interpretaci související právní regulace. Domnívám se proto, že půjde o propojené nádoby a že je třeba se v dalším výzkumu zaměřit na oba tyto aspekty společně.

Druhým důležitým protiargumentem může být, že základní teze této kapitoly jsou mylné. Tzn. že je zcela přirozené, že se různé druhy technologií vyvíjejí v různých dobách různým tempem.¹⁴⁰ Podobně argumentuje například také Kessler¹⁴¹ nebo v historii Kondratieff¹⁴², kteří tvrdí, že inovace probíhají ve vlnách a že tudíž můžeme být na pokraji dalšího rozmachu technologického průlomu. Mezi odvětví, která jsou na prahu obřího progresu, dle Kesslera patří vzdělávání, medicína, biotechnologie nebo robotika. Obecně lze pozorovat, že počet publikací predikující nástupy nových technologií (a jak změní svět) je ohromný. Je přitom samozřejmě nemožné odhadnout, co se stane v budoucnosti, a je možné, že tyto hlasy mají pravdu.

Osobně je však slýchám již velmi dlouho a myslím, že faktická realita a zkušenost naopak svědčí pesimistické tezi této kapitoly. Vezměme si například, kolik let se mluví o autonomních vozidlech a kolik různých etických či právních publikací na toto téma vzniká. A kolik autonomních vozidel reálně po silnicích jezdí. Stejně se již mnoho let diskutuje o ohromných objevech v biotechnologiích, které zcela transformují zdravotnictví. Avšak u investorů a pacientů lze prozatím spíše pozorovat zklamání.¹⁴³ Buď slibované technologie nejsou, nebo se je nedaří dostatečně rychle implementovat ve zdravotnických zařízeních. Podobných příkladů nepodloženého optimismu bychom našli mnoho (modulární jaderné reaktory, létající auta, nukleární fúze atd.).

¹³⁹ Tamtéž, s. 75.

¹⁴⁰ Za tuto myšlenku děkuji kritickému recenzentovi článku, ze kterého tato kapitola vychází.

¹⁴¹ KESSLER, A. *How We Got Here: A Slightly Irreverent History of Technology and Markets*. Harper Business, 2005.

¹⁴² KONDRATIEFF, N. *The Long Waves in Economic Life*. Martino Fine Books, 2014. [Pozn. jde o překlad do angličtiny z ruského originálu z roku 1926].

¹⁴³ THIEL, P.; MASTERS, B. *Od nuly k jedničce. Úvahy o startupech aneb jak tvořit budoucnost...*, s. 74.

Třetím potenciálním protiargumentem je, že vztah regulace a inovativnosti může fungovat i obráceně. Tedy že frekvence a kvalita právní regulace je určována spíše dynamikou technologického vývoje, nikoli že technologický vývoj je určován podobou právní regulace.¹⁴⁴ Rozmach práva informačních technologií by zde mohl být použit jako příklad, kdy příslušný vztah mohl takto fungovat. Je však otázkou, zda oblast informačních technologií v posledních 30 letech vlastně nebyla velmi specifická výjimka a že obecný trend leží jinde. Navíc nelze *pro futuro* vyloučit, že informační technologie čeká stejný osud jako starší technologické oblasti (a to v momentě, kdy je začne přílišná regulace dohánět – což se již pomalu děje).

Úspěch informačních technologií byl totiž postaven na několika faktorech, které jsou pro tuto oblast specifické a nemohou být jednoduše přeneseny jinam. Těmi byla alespoň z počátku kombinace malého množství regulace, malých startupů, absence odporu veřejnosti,¹⁴⁵ lehkého vstupu na trh a možnosti jednoduchého vytváření nových trhů. Tyto faktory o to více vyniknou v protikladu k fyzickému světu, kde je regulace naopak velmi rozsáhlá, odpor k riziku značný a který je ovládán velkými zkonstatitelnými korporacemi. Navíc ve fyzickém světě je velmi těžké vstoupit na existující trhy či vytvořit trhy zcela nové. Když se k tomu přidá skutečnost, že společnost přestala lehce přijímat nové vynálezy, a spirála neúspěchu v mnoha oblastech, stala se zde stagnace v podstatě nevyhnutelnou.

Navíc je logicky mnohem snazší pozorovat to, co vynalezeno bylo a jak na to právní systém ve své dynamice reagoval, oproti tomu, co vynalezeno kvůli podobě právního systému nebylo, popřípadě co nebylo společensky implementováno. Osobně mi navíc přijde hypotéza, že podoba právního systému nemá vliv na technologický vývoj, jako málo přesvědčivá. Spíše by mi proto přišlo účelnější debatovat míru, tzn. jak moc může podoba právního systému – vedle dalších faktorů – inovativnost ovlivňovat. A zároveň jak moc a v jakých oblastech je frekvence a kvalita právní regulace určována dynamikou technologického vývoje.

Čtvrtý faktor, který by mohl tezi této kapitoly do budoucna výrazně narušit, je potenciální dlouhodobý vliv výzkumu a vývoje, který proběhl v období pandemie Covid-19.¹⁴⁶ Uvádí se, že například progres v oblasti vakcín, který byl během pandemie vykonán, byl extrémně rychlý a za jiných okolností takřka nemožný. Podobně došlo díky epidemii k progresu

¹⁴⁴ Za tuto myšlenku děkuji kritickému recenzentovi článku, ze kterého tato kapitola vychází.

¹⁴⁵ Oblast informačních technologií je také specifická nižší mírou rizikovosti, výbuch počítače pochopitelně neohrozí tolik lidí jako výbuch jaderné elektrárny nebo nebezpečný lék.

¹⁴⁶ RENU, N. Technological advancement in the era of COVID-19. *SAGE Open Medicine*, 2021, vol. 9, s. 1–4.

na poli práce s velkými daty a umělou inteligencí, jejichž využití pomohlo tuto krizi v mnoha státech lépe zvládat.¹⁴⁷ Je proto do budoucna otevřenou otázkou, zda se tento progres vryje do dalšího tempa technologického vývoje, či zda šlo o ojedinělou situaci, která nebude mít dlouhodobé trvání.¹⁴⁸ I pokud by se však jednalo o první variantu, půjde opět o progres ve velmi limitovaných sektorech.

V neposlední řadě, jak bystře pozoruje Maas, lze v historii technologií nalézt mnoho případů, kdy došlo k tzv. *technologickému sebeomezení*, tedy situaci, kdy určité technologie mohly být vynalezeny, ale buď byl jejich vývoj úmyslně pozdržen, či úplně zastaven.¹⁴⁹ Šlo například o program sovětského internetu, meteorologických zbraní či desítek státních programů věnujících se jadernému zbrojení.¹⁵⁰ Pokud jsou některé technologie potenciálně velmi nebezpečné, může být vlastně dobře, že nejsou vynalezeny a že tomuto zpomalení pokroku právní řád napomáhá. V takové situaci plní právní systém cíl ochrany před technologickou disrupcí. Maas toto aktuálně diskutuje v kontextu umělé inteligence, ve kterém si jistě lze představit situace, kdy tento přístup dává, vzhledem k potenciální nebezpečnosti této technologie, smysl. Na druhou stranu si nejsem jistý, zda lze pozastavení vývoje konkrétní technologie přirovnávat k situaci, kdy nevznikají ani technologie velmi prospěšné, či takové, které by byly obecně bezpečné.

2.7 Závěr

V této kapitole jsem nejprve představil tezi o technologické stagnaci. Posléze jsem identifikoval kulturní a právní témata, která se tohoto jevu týkají. Mezi ně může patřit problém hypertrofie práva, skutečnost, že právní regulace klade příliš silný důraz na eliminaci rizika a potenciální problém některých právních odvětví v nerefluktování současného stavu společensko-technologických poměrů. Zároveň si uvědomuji, že mnohé premisy se mohou ukázat mylnými, což diskutuji v poslední části.

¹⁴⁷ GOULD, M.; JOSHI, I.; TANG, M. The power of data in a pandemic. In: *GOV. UK. Blog. Technology in the NHS* [online]. 28. 3. 2020. [cit. 2022-10-22]. Dostupné na: <https://healthtech.blog.gov.uk/2020/03/28/the-power-of-data-in-a-pandemic/>.

¹⁴⁸ Viz BETZ, U. et al. Game changers in science and technology – now and beyond. *Technological Forecasting & Social Change*, 2023, vol. 193, 122588.

¹⁴⁹ MAAS, M. Paths Untaken: The History, Epistemology and Strategy of Technological Restraint, and lessons for AI. In: *VerfBlog* [online]. 9. 8. 2022 [cit. 2022-10-22]. Dostupné na: <https://verfassungsblog.de/paths-untaken/>.

¹⁵⁰ Tamtéž.

Jak jsem psal v úvodu, tato část je teoretická a jejím účelem je toto téma otevřít kritice a dalším badatelům, především v kontextu českého a evropského práva. Toto téma je totiž mnohem lépe podchyceno v americké odborné literatuře, kde lze navíc díky ekonomické a politické kontinuálnosti tyto trendy lépe sledovat. Jde však o problém, který je zcela jistě relevantní i v Evropě. Porovnáme-li navíc technologický vývoj v Evropě a USA, i bez substantivní analýzy není těžké vidět, že Evropa v posledních letech nezažívá ani výrazný technologický progres v oblasti počítačů a softwaru.¹⁵¹ Pokud přitom chceme být skutečně inovativní společností, která bude konkurenceschopná USA, Izraeli a Číně, musíme hledat cesty, jak toho dosáhnout, a to i na poli právního systému.

¹⁵¹ SOLON, O. Peter Thiel: Europe is cracking down on Silicon Valley out of 'jealousy'. In: *The Guardian* [online]. 15. 3. 2018 [cit. 2022-10-22]. Dostupné na: <https://www.theguardian.com/technology/2018/mar/15/peter-thiel-silicon-valley-europe-regulation/>.

3. Právní systém a zvyšující se komplexita

Třetí kapitola obecné části se věnuje analýze, co pro právní systém znamená zvyšující se komplexita jak vnějšího světa, tak samotného právního systému a jakou roli v tomto jevu hrají moderní technologie. Jde o téma, které již bylo v předchozích kapitolách naznačeno, avšak až zde se mu dostává hlubšího rozpracování. Východiska jsou přitom pořád stejná: právní systém je operativně uzavřeným autopoietickým systémem, který vstupuje do složitých vztahů se svým okolím a jenž stojí na předávání informací.

3.1 Komplexita a entropie

Vzhledem k tomu, že systémová teorie považuje právní systém za uzavřený systém, domnívám se, že na něj lze aplikovat některé teorie uzavřených systémů pocházejících mimo právo. Však samotná myšlenka *autopoiesis* původně pochází z biologie, aby se později skrze sociologii dostala také do uvažování o právu. Je důležité si uvědomit, že tyto koncepty byly paralelně rozvíjeny i v jiných oborech, a to i poté, co je Luhmann přenesl do sociologie práva.¹⁵² Právě z tohoto důvodu považuji například kognitivní vědu a její poznatky o autopoietických systémech za relevantní, neboť ve skutečnosti vychází ze stejných základů.

Uvažujeme-li přitom o uzavřených systémech, které se skládají z informací a jejich předávání, hrají zásadní roli dva koncepty: entropie a komplexita. Entropii pro tento účel definujeme jako množství informačního obsahu, respektive jako otázku, kolik informací je v systému obsaženo.¹⁵³ Komplexitu zase definujeme jako kapacitu systému inkorporovat informace v určitý čas, případně jako měřítko obtížnosti kompletního popsání systému.¹⁵⁴ Z tohoto pohledu komplexita odráží množství informací potřebných k popisu „*všeho zajímavého v uzavřeném systému v daném časovém okamžiku*“.¹⁵⁵ Čím je komplexita vyšší, tím obtížnější je systém popsat způsobem, aniž by nám neuniklo něco důležitého.¹⁵⁶

Dle Modise platí, že v uzavřeném systému se entropie a komplexita zpočátku zvyšují společně, jinými slovy, čím větší je množství informací, tím obtížnější je určitý systém popsat.

¹⁵² MURRAY, J.; WEBB, T. E.; WHEATLEY, S. Encountering law's complexity. In: MURRAY, J. (eds.). *Complexity Theory and Law. Mapping An Emerging Jurisprudence*. London and New York: Routledge, 2019, s. 3–23.

¹⁵³ MODIS, T. Links between entropy, complexity, and the technological singularity. *Technological Forecasting and Social Change*, 2022, vol. 176, s. 1–3.

¹⁵⁴ Tamtéž, s. 2.

¹⁵⁵ Tamtéž, s. 2.

¹⁵⁶ Tamtéž, s. 2.

Na konci, poněkud paradoxně, platí opak, neboť když se entropie blíží svému maximu – v němž existuje absolutní neuspořádanost – komplexita systému zcela mizí. V okamžiku, kdy entropie dosáhne svého konečného stropu, je informační obsah maximální, ale také pro popsání systému zcela nezajímavý, protože se stal 100 % náhodnou informací.¹⁵⁷ Degradace informačního obsahu na směs nezajímavé náhodnosti začíná v okamžiku, kdy entropie dosáhne tzv. *inflexního bodu* své trajektorie, tj. když rychlost růstu dosáhne maxima.¹⁵⁸ V tomto bodě komplexita překročí své maximum a začne postupně opět klesat. Absolutní neuspořádanost jde posléze opět popsat dost jednoduše. Nakonec přestává jít o systém jako takový.

Za příklad tohoto jevu se ve světě fyziky udává káva s mlékem. Když nalejeme mléko do kávy, je entropie tohoto systému malá a totéž platí pro komplexitu, neboť takovou situaci lze velmi snadno popsat na tom, když si dáte café latté (dole káva, nahoře mléko). Situace se začne komplikovat v momentě, kdy vezmeme lžici a mícháme kávu s mlékem dohromady – zde platí, že se entropie systému zvyšuje a zároveň se zvyšuje komplexita. Právě v tomto bodě je komplexita systému nejvyšší. Úplně na konci, kdy už jsou mléko a káva zcela promíchány, je entropie maximální, ale situaci lze opět popsat velmi jednoduše, tzn. komplexita je opět nízká.¹⁵⁹ Platí tedy, že se zajímavost informačního obsahu v druhé polovině procesu růstu snižuje, stejně jako složitost systému. Na konci pak zůstává čistě náhodná informace a nulová schopnost systému začlenit novou zajímavou informací.¹⁶⁰

3.2 Konsekvence pro právo a právní systém

Byť může výše analyzované působit od práva odtažitě, jsem přesvědčen, že tento mechanismus ve skutečnosti platí i pro právní systém a umožňuje nám na teoretické úrovni pojmenovat moment, kdy v právním systému stále roste entropie, ale již se současně nezvyšuje komplexita (ta ve skutečnosti klesá). Právní systém by si proto měl být vědom svých přirozených limitů tváří v tvář rostoucí entropii a komplexitě jak vnějšího světa (s nímž je v relaci), tak i svého vlastního systému. Efektivní práce s entropií a komplexitou (resp. v současnosti především jejich snižování) a vyhýbání se překročení *inflexního bodu* je tak podmínkou zachování racionality právního systému.

¹⁵⁷ Tamtéž, s. 2.

¹⁵⁸ Tamtéž, s. 2.

¹⁵⁹ AARONSON, S.; CARROLL, S. M.; OUELLETTE, L. Quantifying the Rise and Fall of Complexity in Closed Systems: The Coffee Automaton. *arXiv:1405.6903v1*, 2014.

¹⁶⁰ Tamtéž.

Vyšší komplexita systému není samozřejmě vždy špatná věc a jednoduchost není vždy lepší řešení. Teprve určitá míra komplexity nám totiž vůbec umožňuje mluvit o existenci systému, tzn. určitá míra komplexity je vůbec podmínkou existence systému. Ani později nemusí být vyšší komplexita špatně. Toto vystihl Schuck, když napsal: „*Právní komplexita někdy vytváří spravedlivější, propracovanější, účinnější, ba dokonce jistější formy společenské kontroly.*“¹⁶¹ Toto je dobře vidět na příkladu trestního práva. Komplexnější úprava ukládání trestů sice obsahuje větší množství informací (větší entropii), avšak může také vytvářet spravedlivější trestní politiku, než kdyby za každý trestný čin hrozil trest smrti.

To, co však lze v současnosti kritizovat, by se dalo nazvat stavem nezvládnutého zvyšování entropie právního systému, který na začátku komplexitu zvyšuje (tzn. může i produkovat společenské dobro), ale od překročení *inflexního bodu* vede k jejímu snížení. Neschopnost vypořádat se s touto vlastností právního systému a pochopit tak jeho přirozené limity vede k hypertrofii práva¹⁶² se všemi známými problémy tohoto jevu.¹⁶³

Právnímu systému se tak teoreticky může stát, že v okamžiku, kdy jeho entropie dosáhne konečného stropu, bude informační obsah maximální, ale také zcela nezajímavý, protože půjde o 100% nahodilou směs informací. Nakonec tedy může existovat vysoká entropie a žádná komplexita. Tedy vůbec nepůjde o žádný právní systém. Jde samozřejmě o teoretické konstrukce, domnívám se však, že nám ukazují něco poměrně důležitého. Tyto postřehy rozvádím v následujících podkapitolách této části.

3.3 Role moderních technologií ve vztahu k entropii a komplexitě

Ve vztahu k potenciálnímu limitu maximální míry entropie v právním systému, při němž se komplexita ještě nesnižuje (tzn. nebyl překročen *inflexní bod*), je role moderních technologií zcela zásadní. Právní informační systémy a internet umožnily pracovat s mnohem větším množstvím informací v právu, než by kdykoli v minulosti bylo představitelné. Pokud lze zákony, judikaturu a odbornou literaturu vyhledávat prostřednictvím specializovaného

¹⁶¹ SCHUCK, P. H. Legal Complexity: Some Causes, Consequences, and Cures. *Duke Law Journal*, 1992, vol. 42, no. 1, s. 8.

¹⁶² Viz NOVOTNÝ, O. Hypertrofie práva, její aspekty a jak jí čelit. *Právník*, 2016, vol. 155, no. 7, s. 619–625; mezinárodní srovnání viz GERICKE, O. *Möglichkeiten und Grenzen eines Abbaus der Verrechtlichung*. Aachen: Shaker Verlag, 2003, s. 11.

¹⁶³ HOLLÄNDER, P. Právník v bludišti postmoderní dekonstrukce a Savignyho Ariadnina nit. In: GERLOCH, A.; KRZYŽANKOVÁ, K. Ž. (eds.). *Právo v měnícím se světě*. Plzeň: Aleš Čeněk, 2020, s. 44.

softwaru, umožňuje to vyšší míru komplexity právního systému, než byla v době knih a papíru. Stejně platí pro komparaci období před a po vynálezu a rozšíření knihtisku.

Pokud bychom čistě teoreticky přesunuli dnešní právní řád do roku 1900, tamní právníci by s ním bez našich technologických vymožeností vůbec nebyli schopni pracovat (tato teze určitě platí, vezmeme-li v potaz evropské právo). Množství informací by na ně bylo příliš vysoké, vedlo by to proto k mnohem vyšší míře nahodilosti při aplikaci práva a ve skutečnosti také ke snížení míry komplexity právního systému, neboť větší entropie automaticky neznamená větší komplexitu. Pokud bychom naopak přesunuli právní řád z roku 1900 do současnosti, přišel by nám pravděpodobně velmi jednoduchý, neboť bychom díky našim technologickým možnostem s ním mohli pracovat velmi jednoduše. Tyto příklady ukazují na skutečnost, že podoba technologií při aplikaci práva ovlivňuje, kde se *inflexní bod* nachází (nejde tedy o stálou kategorii).

Současně je třeba si uvědomit, že i nynější způsob fungování práva závislý na používání dnešního softwaru má své přirozené limity a že se k nim možná pomalu blížíme. Ty ale samozřejmě může dále výrazně prolomit například umělá inteligence, kupříkladu skrze nástroje založené na generativní AI. Již v současnosti se například objevují AI právní poradny, které mají potenciál změnit poskytování právních služeb poměrně výrazně. Zde je však na místě i otázka, jestli je tato cesta moudrá, neboť hrozí, že bychom se při poznávání a aplikaci práva mohli dostat mimo kognitivní schopnosti lidí a museli bychom zcela věřit strojům. Přílišným spoléháním se na technologie se také může snížit naše schopnost s právním řádem pracovat bez nich či obecně pracovitost právníku. Rostoucí míra komplexity a entropie právního systému se navíc logicky projevuje i v přetíženosti justičního systému, do kterého se moderní technologie jako AI v českém kontextu příliš nedaří implementovat, což v kontextu celkové proměny podoby právní praxe může být zásadní problém.

Také platí, že naše technologické možnosti za několik posledních desetiletí již vedly k výrazné změně podoby právní praxe. To se projevuje tím, že dnes se právníci mohou fakticky buď úzce specializovat, což vede k fragmentaci, nebo se stát tzv. „věčnými hledači“, což do systému vnáší nahodilost. Naše schopnost pracovat s větším množstvím informací a vyšší složitostí právního systému vedla k změně podoby a fungování právního systému a také jako vedlejší účinek umožnila obsáhlejší množství regulace. Možná i proto se vytvořila moderní (a v mnoha ohledech naivní) představa, že vše lze snadno regulovat a že na veškerý problém by

mělo existovat vhodné právní řešení.¹⁶⁴ Však se stačí podívat kolik nových oborů práva vzniklo za posledních 70 let.

Je zajímavé, že právní systémy historicky měly způsoby, jak míru entropie a komplexity snižovat. Například common law precedenční systém snižoval svoji složitost tím, že se zabýval pouze otázkami, které byly předmětem sporu, tedy tím, co se k němu dostalo prostřednictvím stran sporu. Neměl ambici abstraktně *ex ante* upravovat všechny potenciální situace, které mohou nastat. Úloha zákonů v common law v širším slova smyslu byla rovněž omezená. Historicky se zde zákony snažily upravit pouze určité aspekty vnějšího světa nebo provést konkrétní změny oproti precedenčnímu právu, spíše než aby byly všezahrnujícími racionalistickými pokusy o pokrytí mnoha oblastí v celé jejich šíři. Oba tyto přístupy jsou ve skutečnosti poměrně dobrými způsoby realizace relevance v právním systému a zároveň způsoby, jak snižovat míru entropie a komplexity systému.

Problém proto šlo lépe pozorovat v kontinentální právní kultuře, neboť ta vychází z osvětské racionalistické tradice a z přesvědčení, že je možné vytvořit koherentní normativní systém, který pokryje všechny myslitelné situace. Kontinentální právní systém se totiž snaží vytvořit deontologicky ideální normativní svět. Historickými příklady takových pokusů jsou kodifikace soukromého práva, jako je ABGB, BGB nebo Code civil.

Domnívám se, že tyto pokusy mohly fungovat jen díky tomu, že nebyly dovedeny k dokonalosti. Ve skutečnosti totiž také našly způsoby, jak entropii a komplexitu snižovat a neponořit systém práva do kombinatorní exploze, např. skrze roli právních principů či užívání velmi abstraktních, a tudíž i významově otevřených pojmů.¹⁶⁵ Zároveň je však třeba říci, že kontinentální právní kultura a její tradice jsou zásadní příčinou rostoucí složitosti práva v dnešním světě. Tento trend je nejzřetelněji patrný v právu Evropské unie, jejíž regulační snahy mohou brzy narazit na hranice toho, co je právní systém v současnosti schopen pojmut.

Rostoucí entropie a komplexita právního systému je také významný negativní důsledek většiny nových regulací. Existuje zde přitom, jak bylo ukázáno, nutnost chránit právní systém před rostoucí entropií a komplexitou a před překročením inflexního bodu. Rostoucí složitost vnějšího světa by se proto neměla vždy odrážet v právním systému, v něm by se mělo

¹⁶⁴ SUMPTION, J. *Trials of the state...*, kap. Law's Expanding Empire.

¹⁶⁵ K otázce, jak lze defragmentizovat právní systém viz ONDŘEJEK, P. *Defragmentation of Law. Reconstruction of Contemporary Law as a System...*

reflektovat pouze to dostatečně důležité. Toto bylo do značné míry diskutováno v první kapitole, která pracovala s tzv. *kritérii relevance*.

Zaměříme-li se na vztah právního systému a jeho prostředí, jedním z cílů práva je entropii ve svém okolí snižovat.¹⁶⁶ Pokud se však samotný právní systém nachází v příliš vysoké vlastní entropii a jeho rozhodování již není systémové, nýbrž nahodilé, pravděpodobně bude naopak entropii ve svém okolí zvyšovat. Entropii vnějšího světa přitom zvyšuje několik fenoménů. Jedním z nejvýznamnějších budou nové technologie, například sociální sítě, skrze rozšíření světa, na který by mělo právo reagovat. Svou roli však bude sehrávat také demokratický pluralismus, který je typický pro moderní liberální demokracie.

Důležitou roli sehrává právnické vzdělávání a obecně role právní vědy. Vzdělávání právníků by se nemělo utápět v podružných detailech, ale mělo by být zaměřené na principy a systém, což absolventům později umožní pochopit komplexitu a nepropadnout entropii. Podobně spatřuji integrační roli právní vědy, která může snížit míru entropie právního systému například sjednocením interpretačních metod práva, což je něco, co se dosud zcela nepodařilo.

3.4 Závěr

Tato kapitola ukázala, že existují hranice entropie a komplexity právního systému a že tyto limity mohou nové technologie posouvat. Existuje také fenomén zvyšování entropie a komplexity ve vnějším světě, na čemž mají technologie významný podíl. Podoba vývoje právního systému do budoucna je velmi nejasná a těžko předvídatelná. Domnívám se však, že zde pojmenované fenomény mohou pomoci při úvahách o proměně právního systému v současnosti a budoucnosti a že mohou ukázat na některé obecné jevy, které v evoluci práva nastávají.

¹⁶⁶ KNAPP, V. *Teorie práva*. Praha: C.H. Beck, 1995, s. 30–32.

ZVLÁŠTNÍ ČÁST

4. Umělá inteligence a právo v dlouhodobé perspektivě

Tato kapitola se zabývá oblastmi, které se tématu práva, společnosti a umělé inteligence týkají v horizontu nejbližších desetiletí. Na úvod proto diskutuji vůbec opodstatněnost takového přístupu v kontextu právní vědy. V první části se posléze věnuji otázce právní subjektivity umělé inteligence. V druhé části se zaměřuji na některé existenční hrozby spojené s umělou inteligencí a na způsoby, jak by jim právo mohlo pomoci zabránit. V poslední části se věnuji tématům AI totalitarismu a AI technokracie, což je diskuse týkající se nově se objevujících forem vlády v kontextu využití umělé inteligence ve správě věcí veřejných.

Když přemýšlíme o právu a umělé inteligenci, většinou se náš zrak upíná na problémy současnosti. Do této kategorie bude spadat například posuzování právní odpovědnosti v kontextu užívání AI systémů v medicíně¹⁶⁷ nebo třeba u policie. Občas se zamýšlíme také nad otázkami blízké budoucnosti, kdy například víme, že autonomní vozidla na ulicích pořád masivně nejezdí, ale doufáme, že by se to mohlo v nejbližších několika letech změnit – a tak o tom píšeme právnické texty a přemýšlíme nad možnostmi regulace.¹⁶⁸

Dlouhodobý pohled je jiný, neboť mu nejde o problémy nejbližších několika málo let. Jde mu o kontext nejbližších desetiletí (popř. v extrémnější podobě i staletí až tisíciletí).¹⁶⁹ Jeho základní premisa je, že životy lidí narozených například za sto let nejsou o nic méně důležité než životy lidí narozených v současnosti.¹⁷⁰ Jde o pohled na společnost, která představuje „*společenství živých, mrtvých a dosud nenarozených*“, jak by řekl slavný konzervativní myslitel Edmund Burke. Toto paradigma má výrazný přesah i do světa práva, v němž může hrát roli například při konceptualizaci otázky existence veřejných subjektivních práv budoucích generací.¹⁷¹

¹⁶⁷ Viz např. ZIKMUNDOVÁ, K. Artificial Intelligence and Medical Devices: Do We Need New Regulation? *Časopis pro právní vědu a praxi*, 2023, vol. 31, no. 2, s. 381–394.

¹⁶⁸ Viz např. JENKINS, R.; CERNY, D.; HRÍBEJ, T. *Autonomous Vehicle Ethics: The Trolley Problem and Beyond*. Oxford: OUP, 2022.

¹⁶⁹ GREAVES, H. (eds.). *A Research Agenda For The Global Priorities Institute*. Global Priorities Institute, 2020, verze 2.1, s. 7.

¹⁷⁰ WINTER, Ch. (eds.). *Legal Priorities Research: A Research Agenda*. Legal Priorities Project, 2021, s. 14.

¹⁷¹ Viz např. BEHRENDT, S. Grundrechtstheoretische Begründbarkeit einer Verantwortung gegenüber künftigen Generationen: Das grundrechtstheoretische Rechtsverhältnis in der Zeit. In: AMMANN, O. (eds.). *Verantwortung und Recht: 62. Junge Tagung Öffentliches Recht*. Baden-Baden: Nomos, 2022, s. 259–274.

V tomto paradigmatu se kupříkladu ptáme: Pokud je život lidí narozených v daleké budoucnosti stejně významný jako život naší generace, nemohou některé naše kroky ohrozit jejich právo na život nebo naději žít ve svobodné a demokratické společnosti? A pokud ano, mělo by to nějakým způsobem ovlivnit naše současné jednání? Podobný typ otázek se bude v těchto diskusích vyskytovat poměrně často.

Druhým klíčovým argumentem ve prospěch relevance tohoto přístupu je empirický pohled na časovou osu existence biologických druhů na této planetě. Pokud se podíváme, jak dlouho zde přetrvává průměrný savec, dostaneme se k číslu kolem milionu let.¹⁷² Z tohoto pohledu je lidský druh vlastně velmi mladý (cca 200 000 až 300 000 let) a statisticky by zde měl být ještě zhruba dalších 700 000 až 800 000 let. Toto číslo by však mohlo – při stálosti ostatních proměnných – být ještě výrazně vyšší vzhledem ke schopnostem lidí ovlivňovat planetární prostředí ve svůj prospěch a bezpečí.¹⁷³

Když se navíc zamyslíme nad pokrokem lidského druhu, uvidíme, že naše civilizace začala jen před několika málo tisíci lety a že exponenciální technologický vývoj zažíváme pouze posledních cca 300 let.¹⁷⁴ Vezmeme-li v potaz všechny potenciální budoucí generace za předpokladu, že by náš druh zůstal na této planetě až do jejího přirozeného zániku, jednalo by se až o 30 milionů generací.¹⁷⁵ To je velmi vysoké číslo a potenciál, který by jistě měl mít určitou hodnotu i v současnosti.

Jak ukazují Martínez a Winter, historicky právní systémy na zájmy budoucích generací příliš nehleděly. Nebyly zastoupeny v zákonodárných sborech, v jejich zájmu nešlo podávat žaloby a jejich zájmy nebyly zvažovány v analýze nákladů a přínosů na ministerstvech a v jiných správních orgánech v rámci moci výkonné.¹⁷⁶ Toto se v poslední době pomalu a intuitivně mění, což dokládají i empirická šetření zkoumající postoje právníků k těmto otázkám.¹⁷⁷ Jde také o paradigma, které se stále častěji dostává do právní argumentace například v ústavním právu.¹⁷⁸ Tyto postoje se přitom mnohdy opírají o principy a ustanovení,

¹⁷² WINTER, Ch. (eds.). *Legal Priorities Research: A Research Agenda*. Legal Priorities Project, 2021, s. 15.

¹⁷³ Tamtéž, s. 15.

¹⁷⁴ Tamtéž, s. 15.

¹⁷⁵ Tamtéž, s. 15.

¹⁷⁶ Viz WINTER, Ch.; MARTÍNEZ, E. *Experimental longtermist jurisprudence*. Legal Priorities Project. LPP working paper, 2021, n. 2.

¹⁷⁷ MARTÍNEZ, E.; WINTER, Ch. Is Legal Longtermism Common Sense? In: *VerfBlog* [online]. 9. 8. 2022 [cit. 2022-11-17]. Dostupné na: <https://verfassungsblog.de/is-legal-longtermism-common-sense/>.

¹⁷⁸ ARAÚJO, R.; KOESSLER, L. The Rise of the Constitutional Protection of Future Generations. LPP Working Paper, 2021, No. 7. Dostupné na SSRN.

kteřá nejsou nová. Takový přístup tedy často spadá do tzv. law-centric modelu. Araújo a Koessler k tomuto píší:

„Při bližším pohledu jsme identifikovali čtyři témata, v jejichž kontextu jsou v dnešních ústavách zmiňovány budoucí generace: životní prostředí, přírodní zdroje, společenské hodnoty a veřejné finance. Tato témata však nejsou v ústavách nová. Byla spíše přeformulována s ohledem na zájmy budoucích generací.“¹⁷⁹

V jiných otázkách bude třeba právní řády a právní vědu doplňovat, například o dosud nepřilíš rozšířený princip *In Dubio Pro Futura*.¹⁸⁰ Tento princip znamená, že v případě nejistoty by se mělo rozhodnout tak, aby to bylo v nejlepší zájmu budoucnosti nebo dlouhodobého udržitelného vývoje. Bublitz a Winter tento princip konkretizují následujícím způsobem:

„V případě pochybností, tj. když zákon neposkytuje jedinou správnou odpověď a několik odpovědí je stejně přijatelných, by soudy měly zvolit tu, která je z hlediska dlouhodobých zájmů lidstva nejpříznivější. To se týká soudních rozhodnutí v konkrétním případě nebo, pokud soudy vytvářejí nové normy, těchto nově vytvořených norem.“¹⁸¹

Existují samozřejmě také důležité námitky vůči tomuto přístupu. První a nejdůležitější je, že je nemožné ovlivnit takto vzdálenou budoucnost.¹⁸² Tato námitka stojí na představě, že krátkodobé problémy vždy převáží ty dlouhodobé a že právo by mělo mířit hlavně na současnost.¹⁸³ Podíváme-li se však na právo a právní instituty, některé z nich ovlivňují naši civilizaci po stovky, ne-li tisíce let.¹⁸⁴ Římské právo, byť tisíce let staré, dodnes ovlivňuje podobu civilního práva a například instituce parlamentu je koncept starý minimálně stovky let. V mnoha jurisdikcích nalezneme zákony, jejichž vznik se datuje do počátku 19. století,

¹⁷⁹ ARAÚJO, R.; KOESSLER, L. The Rise of the Constitutional Protection of Future Generations. In: *VerfBlog* [online]. 12. 8. 2022 [cit. 2022-11-17]. Dostupné na: <https://verfassungsblog.de/the-rise-of-the-constitutional-protection-of-future-generations/>.

¹⁸⁰ BUBLITZ, Ch.; WINTER, Ch. In Dubio Pro Futura. A Novel Decision Rule to Preserve the Long-Term Interests of Humanity. In: *VerfBlog* [online]. 19. 8. 2022 [cit. 2022-11-17]. Dostupné na: <https://verfassungsblog.de/in-dubio-pro-futura/>.

¹⁸¹ Tamtéž.

¹⁸² WINTER, Ch. (eds.). *Legal Priorities Research: A Research Agenda*. Legal Priorities Project, 2021, s. 17.

¹⁸³ Tamtéž, s. 17.

¹⁸⁴ Tamtéž, s. 17.

a přestože jistě nevznikaly s vidinou regulovat moderní svět 21. století, činí tak, a to většinou dokonce velmi obstojně.

Právní normy a instituce, pokud jsou dobře nastaveny ve svých základních principech a hodnotových východiscích, jednoduše mohou směřovat i do daleké budoucnosti a pro právníky by tedy mělo dávat smysl uvažovat v kontextu nikoli pouze současnosti, nýbrž i vzdálenější budoucnosti. Podobně mohou fungovat také kulturní normy, které si společnost buduje. Z empirického šetření přitom vyplývá, že až 70 % oslovených právníků se domnívá, že právo může ovlivnit vzdálenou budoucnost alespoň v horizontu 100 let.¹⁸⁵

Vzhledem k názvu této práce, která by měla být vázána na 21. století, se tohoto horizontu přidržím. Následující pasáže by proto měly reflektovat vývoj, který může nastat v průběhu tohoto století. Některé otázky jsou přitom velmi relevantní i v současnosti.

¹⁸⁵ WINTER, Ch.; MARTÍNEZ, E. *Experimental longtermist jurisprudence*. Legal Priorities Project. LPP working paper, 2021, n. 2, s. 15.

4.1 Umělá inteligence jako subjekt práva¹⁸⁶

Otázka dopadů umělé inteligence na právo a jeho budoucí vývoj je široké téma, které naskýtá mnoho zcela fundamentálních otázek. Tato kapitola se především zaměřuje na otázku právní subjektivity umělé inteligence, a to navíc z trochu specifického úhlu pohledu, neboť se zde věnuji především otázce právního jednání a zastoupení v právních vztazích.

Přestože jde o problematiku do jisté míry právně futuristickou, jde také o téma, které je již nyní velmi aktuální například s ohledem na algoritmické obchodování na finančních trzích. S technologickým postupem bude přitom otázka, zda by měla moci umělá inteligence právně jednat či zastupovat subjekty práva, čím dál tím více relevantní. Tato kapitola rozebírá teoretická východiska, včetně některých otázek etických, a načrtává určité možnosti směřování právní úpravy v budoucnosti.

4.1.1. Právo jako svět subjektů a věcí

Právo je společenský regulativní systém, který působí na autonomní lidskou vůli. Účelem práva je působit na vůli autonomních bytostí a tím ovlivňovat jejich jednání.¹⁸⁷ Právo je ale také zároveň produktem autonomní vůle, neboť jde o systém příkazů, zákazů a dovolení, které jsou výtvorem obecné vůle právotvorného subjektu. Jak lze na tomto schématu vidět, jedním ze základních komponentů práva je *autonomní vůle* – a kvůli autonomii vůle (tedy možnosti ovlivnit, co udělám) jsou právní subjekty nositeli práv (možnosti určitého jednání) a povinností (nutnosti určitého jednání pod hrozbou sankce).

V tomto kontextu existují subjekty práva, tedy nositelé konkrétních (procesních či hmotněprávních) vzájemně provázaných práv a povinností. V případě fyzických osob se (alespoň potenciálně) jedná o bytosti s biologicky vrozenou autonomií vůle, které svým jednáním aktivně ovlivňují a mění objektivní svět. V případě právnických osob jde o fiktivní konstrukty, které ale díky skutečnosti, že se většina lidského chování odehrává

¹⁸⁶ Tato část vychází z HOŘEŇOVSKÝ, J. Umělá inteligence jako hmotněprávní a procesní subjekt v civilním právu s přihlédnutím k zastoupení. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.) *Zastoupení. Specifika a kontext*. Praha: Wolters Kluwer, 2021, s. 259–269.

¹⁸⁷ KELSEN, H. *Reine Rechtslehre*. Vídeň: Deuticke, 1960, s. 31.

v institucionálním rámci,¹⁸⁸ působí na chování konkrétních fyzických osob alespoň zprostředkovaně.¹⁸⁹

V podstatě vše, co není subjektem, je, až na několik málo výjimek,¹⁹⁰ věcí. Věc se vyznačuje tím, že nemá autonomii vůle, nemůže se sama rozhodnout, že nějaké jednání uskuteční. Dům je jednoduše řečeno domem a nemůže si vybrat, zda si nechá dostavět své horní patro. Nemělo by proto valný smysl, aby byl subjektem práva. Vyhrožovat domu právní sankcí, pokud něco neuskuteční, by bylo stejně užitečné, jako rozkazovat slunci, aby přestalo svítit.

Výše zmíněné rozdělení v právu neexistuje od nepaměti. Pojetí subjektu práva se samozřejmě neustále vyvíjí. Ve starém Egyptě byly subjektem náboženské chrámy, protože plnily klíčovou hospodářskou úlohu.¹⁹¹ A přestože v římském právu neexistovaly klasické právnické osoby, které by při podnikání měly vlastní majetek nezávislý na majetku svých členů, přišli i tak Římané s širokou plejádou pohledů na korporace.¹⁹² Rozhodně nebylo vždy pravidlem, že každý člověk je automaticky plnoprávným subjektem; v římském právu bylo pro plnou subjektivitu klíčové až římské občanství. Nelze navíc zapomenout na odvěkou existenci otroků. Ve středověku se zase stávalo, že subjekty práva (povětšinou trestního) byla také zvířata. Čas od času tak byla stíhána prasata, koně nebo dokonce krysy, a to klasickým trestním procesem.¹⁹³

Jak tyto historické případy ukazují, samotný koncept a dichotomie subjektů a věcí jsou kulturně a historicky proměnlivé; nejedná se o věčné a neměnné se kategorie. Proměnlivost a nejednoznačnost na první pohled tak samozřejmých konceptů lze spatřovat i v § 494 OZ, který stanoví, že „živé zvíře má zvláštní význam a hodnotu již jako smysly nadaný živý tvor. Živé zvíře není věcí a ustanovení o věcech se na živé zvíře použijí obdobně jen v rozsahu, ve kterém to neodporuje jeho povaze“. Tímto nastalo další zmatení pojmů, neboť zvíře

¹⁸⁸ WEINBERGER, O. *Norma a instituce: úvod do teorie práva*. Plzeň: Aleš Čeněk, 2017, s. 23–26.

¹⁸⁹ Srov. o povaze právnických osob např. BERAN, K.; ČECH, P.; DVOŘÁK, B.; ELISCHER, D.; HRÁDEK, J., JANEČEK, V.; KÜHN, Z.; ONDŘEJEK, P. *Artificial legal entities: essays on legal agency and liability*. Praha: Wolters Kluwer, 2019.

¹⁹⁰ Věcí ani osobou není například Slunce.

¹⁹¹ BARFIELD, W.; PAGALLO, U. *Research Handbook on the Law of Artificial Intelligence*. Cheltenham: Edward Elgar Publishing, 2018, s. 228.

¹⁹² PATTERSON, J. L. The Development of the Concept of Corporation From Earliest Roman Times to A.D. 476. *The Accounting Historians Journal*, 1983, vol. 10, no. 1, s. 87–98.

¹⁹³ DINZELBACHER, P. Animal Trials: A Multidisciplinary Approach. *The Journal of Interdisciplinary History*, 2002, vol. 32, no. 3, s. 405–421.

v současnosti není subjektem ani věcí, takže lze mluvit o existenci další kategorie reality, kterou objektivní právo nějakým způsobem bere na zřetel.

Je důležité si uvědomit, že určité rozdělení smysl dávalo více než rozdělení jiné. Například přiznání subjektivity chrámu bylo ve své době zcela racionálním krokem. Starověký chrám představoval základní ekonomickou jednotku své doby, která pravidelně vstupovala do právních vztahů. Tím pádem bylo velmi praktické, aby byla samostatným subjektem podobně, jako jimi jsou dnešní právnické osoby. Naproti tomu souzení zvířat smysl nedávalo žádný. Otázka zní, do které z těchto kategorií spadá umělá inteligence.

4.1.2. Co by umělá inteligence měla splňovat, aby byla subjekt práva?

Dle Evropské komise je „*umělá inteligence* [...] *označení pro systémy, které vykazují inteligentní chování analyzováním jejich prostředí a s určitou mírou autonomie podnikají kroky k dosažení konkrétních cílů*“.¹⁹⁴ Bostrom o současném stavu umělé inteligence píše: „*Stroje dokážou komponovat ve stylu různých skladatelů vážné hudby, předčít v některých klinických diagnostických úkolech začínající lékaře, samostatně řídit automobily a přicházet s vynálezy, na které lze získat patent. Objevila se dokonce AI, která vypráví své vlastní vtipy*“.¹⁹⁵ Jedná se ale o dostatečné schopnosti, abychom uvažovali o vytvoření dalšího druhu subjektu práva?

Jedním z klíčových důvodů, proč je určité entitě přiznána právní subjektivita, je skutečnost, že jí tím také přiznáváme schopnost „*participovat v sociálně relevantních vztazích*“.¹⁹⁶ Spolu s klíčovým aspektem spočívajícím v *autonomii vůle* z tohoto logicky vyplývá následující úvaha: Aby dávalo smysl uvažovat o vytvoření třetího (například elektronického) subjektu práva, musela by taková entita:

- a) buď přímo autonomii vůle mít, nebo alespoň jednat způsobem, který je nerozpoznatelný od situace, jako by autonomii vůle měla; a zároveň
- b) by zde měly existovat relevantní společenské vztahy, ve kterých by hospodářsky či jinak pomohlo, pokud by umělá inteligence byla samostatným subjektem, a nikoli pouhou věcí.

¹⁹⁴ EVROPSKÁ KOMISE. Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému výboru a sociálnímu výboru a výboru regionů *Umělá inteligence pro Evropu*. COM/2018/237 final. 2018. Dostupné z: <https://eur-lex.europa.eu/legal-content/CS/ALL/?uri=COM:2018:237:FIN>.

¹⁹⁵ BOSTROM, N. *Superintelligence: až budou stroje chytřejší než lidé*. Praha: Prostor, 2017, s. 24.

¹⁹⁶ BARFIELD, W.; PAGALLO, U. *Research Handbook on the Law of Artificial Intelligence*. Cheltenham: Edward Elgar Publishing, 2018, s. 228.

4.1.3. Obecná umělá inteligence versus partikulární umělá inteligence

Otázka autonomie vůle (či alespoň jejího dostatečného napodobování) je u umělé inteligence komplikovanou záležitostí. Obecně ale lze v tomto kontextu rozdělit umělou inteligenci do dvou kategorií:

- 1) systémy s úzkým rozsahem kognitivních schopností; a
- 2) systémy, které mají předem nspecifikované širší uplatnění (tzv. obecná umělá inteligence).¹⁹⁷

Do první kategorie patří například samořídící auto. Nedomnívám se, že by v tomto prvním případě byla splněna některá z podmínek, kdy by dávalo smysl uvažovat o vzniku třetího subjektu práva. Samořídící auto nebude vstupovat do právních vztahů a jeho míra autonomie vůle je determinována pouze základním účelem – řízením na dopravních komunikacích. A přestože takové auto může například samo způsobit škodu, vždy lze snadno přičíst odpovědnost buď výrobcí (v případě vady výrobku),¹⁹⁸ nebo uživateli (pokud stroj nedal včas do opravy; či jednoduše na základě absolutní objektivní odpovědnosti za provoz dopravního prostředku).¹⁹⁹ Z tohoto důvodu mi zájem právníků o samořídící auta přijde v současnosti až přehnaný, neboť do práva (ani do etiky) ve skutečnosti žádné nové otázky nepřináší. Podobné bude přitom platit pro drtivou většinu partikulární umělé inteligence, která dnes představuje „*v podstatě všechny systémy, které se v současnosti používají*“.²⁰⁰ Vedle samořídících aut by se mohlo jednat například o autonomní sekačku na trávu.

Úvaha nad druhou kategorií je už o poznání zajímavější. Dle Bostroma musí *obecná umělá inteligence* splňovat následující charakteristiky:

- a) musí mít autonomní schopnost učit se;²⁰¹
- b) musí být schopna „*efektivně pracovat s neurčitostí a pravděpodobnostními informacemi*“²⁰²; a

¹⁹⁷ BOSTROM, N. *Superintelligence: až budou stroje chytřejší než lidé...*, s. 39.

¹⁹⁸ Viz § 2939 až 2943 OZ.

¹⁹⁹ Viz § 2927 a násl. OZ.

²⁰⁰ BOSTROM, N. *Superintelligence: až budou stroje chytřejší než lidé...*, s. 39.

²⁰¹ Tamtéž, s. 48.

²⁰² Tamtéž, s. 48.

- c) měla by mít „*schopnost získávat užitečné pojmy ze smyslových dat a svých vlastních vnitřních stavů a vkládat takto získané pojmy do flexibilních kombinatorických reprezentací, aby mohly být využity v logickém a intuitivním usuzování*“.²⁰³

Ve zkratce řečeno, musí se jednat o stroj, který by měl být schopen řešit předem neomezený okruh úkolů a autonomně vykonávat předem neomezené množství jednání ve vnějším světě. Pokud by se podařilo takový stroj vytvořit, důvody pro přiznání právní subjektivity (či obdobného postavení) takové entitě by již byly velmi silné, neboť by se alespoň z pohledu klíčového prvku autonomie vůle příliš neodlišovala od člověka samotného. Takový stroj však dosud vytvořen nebyl.

4.1.4. Obecná umělá inteligence jako subjekt práva

Obecná umělá inteligence by byla fakticky schopna autonomně vstupovat do právních vztahů, zavazovat se k povinnostem, být zastupována u soudu, nebo dokonce u soudu jinou osobu zastupovat. Je třeba si navíc uvědomit jednu důležitou věc. Obecná umělá inteligence by svou podstatou byla mnohem blíže osobě fyzické než osobě právnické. Právnická osoba je imaginární společenskou institucí, která nemá skutečnou autonomní vůli a nemůže sama od sebe jednat ve vnějším světě. Tomu dal za pravdu i občanský zákoník, když stanovil, že právnické osoby jsou nesvéprávné.²⁰⁴ Jedná se pouze o institucionální rámec, ve kterém se uskutečňuje chování konkrétních lidí v konkrétních situacích. Spolu s normami jde o struktury, které působí na vůli lidí, aby se nějakým způsobem chovali.²⁰⁵ Jde svým způsobem o únik od reality a odpovědnosti, který je veden praktickými důvody. Naproti tomu obecná umělá inteligence by měla vůli skutečně vlastní – vůli, na kterou by se působilo a která by také sama působila na okolní svět. Šlo by tedy o entitu velmi reálnou.

V této analýze je však jeden důležitý problém. Nikde není napsáno, že by taková entita přemýšlela stejným způsobem, jakým tak činí lidská mysl.²⁰⁶ Někteří autoři proto upozorňují na skutečnost, že naše současná na lidi orientovaná právní pravidla by se mohla ukázat až zoufale neúčinná, pokud by měla regulovat chování autonomních ne-lidských subjektů, založených například na strojovém učení.²⁰⁷ Technologický pokrok by z tohoto důvodu mohl

²⁰³ Tamtéž, s. 48.

²⁰⁴ Viz např. BERAN, K. Může být právnická osoba svéprávná? *Právník*, 2018, roč. 157, č. 8, s. 657–669.

²⁰⁵ WEINBERGER, O. *Norma a instituce: úvod do teorie práva...*

²⁰⁶ BOSTROM, N. *Superintelligence: až budou stroje chytrější než lidé...*, s. 58.

²⁰⁷ HARNED, Z.; WALLACH, H. Stretching Human Laws to Apply to Machines: The Dangers of a 'Colorblind' Computer. *Florida State University Law Review*, 2020, vol. 47, issue 3, s. 618–648.

vést k potřebě vytvořit zcela nový druh právních pravidel s omezenou osobní působností, která by cílila pouze na vůli umělých entit. Nejednalo by se však z pohledu práva o nic nového, neboť již v současném právním řádu máme mnoho norem, které se uplatní například pouze na některé konkrétní osoby, třeba na banky.²⁰⁸ V případě fyzických osob se pouze některé normy vztahují například na děti, spotřebitele atd.

4.1.5. Emulace lidského mozku jako subjekt práva

Velmi záleží, o jakém druhu umělé inteligence hovoříme. Jedním ze způsobů vytvoření obecné umělé inteligence, o kterém mluví například Kurzweil,²⁰⁹ je možnost emulace lidského mozku. Jde o metodu „*stvoření inteligentního softwaru, jež spočívá v nasnímání a následném přesném namodelování výpočetní struktury biologického mozku*“.²¹⁰ Takto vytvořená inteligence by mohla kopírovat konkrétní existující fyzickou osobu. Problémy, které takový scénář vyvolává, jsou poměrně evidentní. Měl by takový software mít již při svém vzniku právní osobnost?²¹¹ Změnila by se odpověď, pokud by již nežil původní člověk, od kterého je taková inteligence odvozena?²¹²

Jedná se o otázky, na které zatím neexistují adekvátní odpovědi. Jedním ze zajímavých akademických návrhů je, že by taková inteligence měla mít po určitou dobu podobný status nezletilého mezi patnácti a osmnácti lety života, po jehož dosažení by nabyla plné svéprávnosti.²¹³ Podstatou tohoto návrhu je, že by se během několika let, díky vlastním „životním“ zkušenostem, dostatečně odlišila od své „původní předlohy“, a mohla by proto jednat jako autonomní bytost, nikoli pouze jako kopie.²¹⁴ Pokud by byl tento návrh implementován, mohl by toho využít například advokát, který by tak mohl poskytovat právní služby i po své smrti. Navazující otázky jsou velmi zajímavé. Existují například nějaké důvody, ať již právní či etické, proč by nemohl zastupovat své klienty před soudy? Jedná se samozřejmě o úvahy trochu předběžné, neboť tato technologie nebude k dispozici v nejbližších letech. I tak

²⁰⁸ Viz zákon č. 21/1992 Sb., o bankách.

²⁰⁹ FRIEDERSDORF, C. Immortal but Damned to Hell on Earth

The danger of uploading one's consciousness to a computer without a suicide switch. In: *The Atlantic* [online]. 2015 [cit. 2024-01-07]. Dostupné z: <https://www.theatlantic.com/technology/archive/2015/05/immortal-but-damned-to-hell-on-earth/394160/>.

²¹⁰ BOSTROM, N. *Superintelligence: až budou stroje chytrější než lidé...*, s. 58.

²¹¹ MUZYKA, K. The Outline of Personhood Law Regarding Artificial Intelligences and Emulated Human Entities. *Journal of Artificial General Intelligence*, 2013, vol. 4, no. 3, s. 166.

²¹² Tamtéž, s. 166.

²¹³ Tamtéž, s. 166–167.

²¹⁴ Tamtéž, s. 167.

ale může dávat smysl se na takové otázky začít ptát, neboť technologický progres v této oblasti může být rychlejší, než očekáváme.

4.1.6. Problém zastoupení umělou inteligencí v právních vztazích

Představme si následující situaci: Podnikatel by se rozhodl přenechat investování a obhospodařování části svého majetku autonomní umělé inteligenci. Tato umělá inteligence by posílala nabídky na koupi podílů obchodních společností, které by byly akceptovány, odmítnuty nebo zaslány zpět s úpravami jako protinávrh. Umělá inteligence by všechny okolnosti daného obchodu důkladně zvážila a rozhodla, zde takovou smlouvu uzavře, nebo neuzavře. Podnikatel by o jednotlivých obchodech, ani o existenci druhých smluvních stran vůbec nevěděl.²¹⁵

Jaké problémy by taková situace v současnosti vyvolala? Jak upozorňuje Alexandre, problém je zde zcela evidentní: „*Předpokládáním konsensu mezi stranami, které si možná ani neuvědomují, že smlouva byla uzavřena, nebo že druhá strana vůbec existuje, tento přístup připravuje uzavření smlouvy o nejdůležitější prvek: setkání projevů vůle.*“²¹⁶ Tento problém platí i pro občanský zákoník, který v § 1724 stanoví, že „*smlouvou projevují strany vůli zříditi mezi sebou závazek a řídit se obsahem smluv*“. Pokud však ani netuším, že se nějaká smlouva uzavírá, nebo dokonce že druhá strana existuje, je tvrzení o uzavření smlouvy dosti problematické a ukazuje na alespoň teoretické problémy v současné právní úpravě. Jedná se o velmi dobrý argument pro zvážení přiznání právní subjektivity (a tedy i možnosti zastupovat někoho v právních vztazích) také umělé entitě, která disponuje dostatečnou autonomií vůle.

Na druhou stranu je zde otázka, jak se v tomto konkrétním případě liší obyčejný software obchodující na burze na základě primitivnějších algoritmů od obecné umělé inteligence. Jednou z odpovědí by mohla být *míra kontroly a předem nastavené pokyny a preference*, a to například vzhledem k následnému vztahu k odpovědnosti za škodu či ve vztahu k smluvní pokutě. Jednoduše řečeno, pokud software pouze vykonává pokyny a jedná za účelem maximalizace preferencí, jde o věc, která nějakým způsobem funguje (a za kterou jsem odpovědný). Pokud by však takový software dokázal jednat skutečně autonomně, tzn. že by mu byl pouze vyčleněn určitý majetek a on by s ním zcela samostatně nakládal, byla by situace odlišná. Toto dělení je však velmi schematické.

²¹⁵ ALEXANDRE, F. M. *The Legal Status of Artificially Intelligent Robots: Personhood, Taxation and Control*. Tilburg: Tilburg University, 2017, s. 15. Dissertation. Vedoucí práce Prof. Erik Vermeulen.

²¹⁶ Tamtéž, s. 15.

Tato otázka má ještě navazující konotace v oblasti korporálního práva, kdy lze uvažovat, že by podobné schéma šlo udělat například v rámci určité právnické osoby, kupříkladu společnosti s ručením omezeným. V takové situaci by umělá inteligence mohla nahradit zaměstnance a potenciálně i členy statutárního orgánu. Právní subjektivitu by v takovém případě měla právnická osoba, ta by však byla jen prázdou schránkou pro určitou umělou inteligenci, která by de facto právní subjektivitu této právnické osoby převzala. Vzhledem k tomu, že i dnes je spousta smluv uzavírána automaticky (tzv. smart contracts),²¹⁷ by taková entita řízená umělou inteligencí mohla dělat poměrně dost věcí. Zda je takové řešení možné dle současného korporálního práva, nebo by bylo třeba změnit právní úpravu, je v současnosti diskutovanou otázkou.²¹⁸ V odborné literatuře se dokonce diskutuje možnost, že by právní subjektivitu mohla získat i tzv. distribuovaná autonomní organizace, která je z technického pohledu počítačovým programem.²¹⁹

4.1.7. Umělá inteligence a postavení otroka v římském právu

Alternativní úvahou je připodobnění vztahu vlastníka a umělé inteligence ke vztahu pána a otroka z římského práva.²²⁰ Ve vztahu k hospodářské aktivitě by tato analogie do značné míry dávala smysl:

„Podobnosti jsou hojné: otroci byli nuceni pracovat, což umožnilo produktivitu a bohatství společnosti. Jako takoví obvykle neměli žádná práva, ale byli považováni za majetek svých pánů.“²²¹

Se souhlasem pána mohli otroci spravovat určitou vymezenou část majetku – tzv. *peculium*.²²² Otroek byl tak schopen nabývat vlastnické právo a držbu pro svého pána, který

²¹⁷ MARKOVÁ, J. Využití smart contract ve spotřebitelských vztazích, zejména ve smlouvě o zájezdu. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022, s. 89–101.

²¹⁸ Např. příspěvek Jany Guričové s názvem *Hledání právní osobnosti umělé inteligence* prezentovaný dne 27. 10. 2023 v rámci konference Právo Obchod Ekonomika, XII. ročník mezinárodního symposia na Štrbském plesu.

²¹⁹ ČERNÁ, M. K vybraným pojmům z oblasti řízení obchodních korporací v kontextu diskuze o právní osobnosti distribuovaných autonomních organizací. In: SEDLÁČEK, M.; STŘELEČEK, T. *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022, s. 125–134.

²²⁰ HUTTUNEN, A.; KULOVESI, J.; BRACE, W.; LECHNER, G.; SILVENNOINEN, K.; KANTOLA, V. Liberating Intelligent Machines with Financial Instruments. *Nordic Journal of Commercial Law*, 2010, no. 2, s. 4.

²²¹ Tamtéž, s. 4.

²²² KINCL, J.; SKŘEJPEK, M.; URFUS, V. *Římské právo*. Praha: C.H. Beck, 1997, s. 64.

byl ale odpovědný pouze do výše majetku vyčleněného v peculiu.²²³ Byl to elegantní způsob, který omezoval ručení pána podobně, jako tomu činí dnešní kapitálové obchodní společnosti. Otok tak byl schopen samostatně využívat svůj intelekt a samostatně vstupovat do právních vztahů, neměl ale právní subjektivitu a také měl omezený přístup k soudu.²²⁴ Otoku se tak přiznávala účast na *ius commercii*, neboli na hospodářském životu společnosti.²²⁵ Života politického se ale z povahy věci neúčastnil.

Když se zamyslíme nad povahou umělé inteligence a jejích možných jednání v hospodářském životě společnosti, jeví se římská inspirace velmi zajímavě. Pokud bych umělé inteligenci svěřil určitý majetek, aby s ním hospodařila dle svého uvážení, řešila by tato právní úprava situaci, kdy by umělá inteligence sama vstupovala do právních vztahů (aniž by však narušovala tradiční dualismus subjektů práva).

Na druhou stranu, v případě umělé inteligence musíme do úvahy zahrnout ještě osobu výrobce, což analogický vztah k postavení otroka – například v případě otázky odpovědnosti – trochu narušuje.²²⁶ Diskutujeme-li ale specificky otázku zastoupení, jde o průměr takřka dokonalý. Takovýto koncept by přitom mohl v omezeném rozsahu dávat smysl již při dnešních technických možnostech, tedy i pro partikulární (úzce zaměřenou) umělou inteligenci, a to například právě při obchodování na burze automatizovanými systémy.²²⁷

Důležitou výhodou tohoto řešení je, že by tento status vyřešil některé problémy, které se s umělou inteligencí jako subjektem práva obvykle pojí. Uvádí se, že pokud bychom z umělé inteligence udělali samostatný subjekt, mohla by mít či si činit nárok na lidská práva nebo právo volit.²²⁸ Otoki v římské říši z podstaty věci lidská práva neměli, včetně práva volebního, takže by takovýto právní status mohl řešit i otázky podobného ražení.

Otokový status ale samozřejmě může vyvolat také kontroverze a etické otázky. Je eticky přijatelné inspirovat se v právní úpravě postavení otroků? Podle mého názoru to přijatelné spíše je, neboť umění římského práva obecně tkví v tom, že právní konstrukty, které původně

²²³ GAMAUF, R. Slaves doing business: the role of Roman law in the economy of a Roman Household. *European Review of History: Revue europeenne d'histoire*, 2009, vol. 16, no. 3, s. 332.

²²⁴ Tamtéž, s. 332.

²²⁵ KINCL, J., SKŘEJPEK, M., URFUS, V. *Římské právo...*, s. 64.

²²⁶ HUTTUNEN, A.; KULOVESI, J.; BRACE, W.; LECHNER, G.; SILVENNOINEN, K.; KANTOLA, V. Liberating Intelligent Machines with Financial Instruments..., s. 4.

²²⁷ Tamtéž, s. 4.

²²⁸ YAMPOLSKIY, R. Could an artificial intelligence be considered a person under the law? In: *The Conversation* [online]. 5. 10. 2018 [cit. 2020-06-26]. Dostupné z: <https://theconversation.com/could-an-artificial-intelligence-be-considered-a-person-under-the-law-102865>.

sloužily pro potřeby antické říše založené na práci otroků, se osvědčily také mnoho tisíc let poté v úplně odlišných společnostech založených na odlišných hodnotách. Samozřejmě ale platí, že se jedná o právní konstrukt, který může být pro potřeby užití v 21. století pojmenován zcela jinak, což by bylo vzhledem k možným kontroverzím z pohledu veřejnosti vhodné učinit.

4.1.8. Závěr

V této kapitole byla probírána umělá inteligence a možné reakce práva *de lege ferenda* ve vztahu k otázkám, které může vyvolávat potřeba nechat se zastupovat umělou inteligencí v právních vztazích. Jedním ze zde prezentovaných řešení je učinit z umělé inteligence plnohodnotný subjekt práva, který by se zařadil vedle osob fyzických a právnických. Druhou probíranou variantou je užití právní konstrukce otroctví z římského práva, která se ve vztahu k umělé inteligenci jeví, při jiném pojmenování, vhodně (protože nejde o člověka). Třetí variantou je ponechat umělé inteligenci status věci; tato možnost ale může velmi brzo začít narážet na potřeby společnosti.

Všechny zde zmíněné možnosti jsou však na začátku diskuse v odborných kruzích a každá ze zde nastíněných cest se také může ukázat jako cesta nefunkční. Celá záležitost navíc samozřejmě závisí na tom, čeho bude v budoucnu technologicky dosaženo. Tato kapitola představuje pouze stručný obecný přehled, který ale může být zajímavým odrazovým můstkem pro další diskuse o této otázce.

4.2 Negativní lidská práva jako základ pro dlouhodobou bezpečnost AI²²⁹

Toby Ord ve své knize *Nad propastí: Existenční riziko a budoucnost lidstva* uvádí pravděpodobnost jedna ku deseti, že lidstvo do konce tohoto století vymře kvůli pokročilé AI.²³⁰ Existuje-li zde takto velké riziko, že naše jednání bude mít v horizontu několika málo generací zcela fatální následky, pravděpodobně jde o téma určité pozornosti hodné i v současnosti. I kdyby tato pravděpodobnost totiž ve skutečnosti byla o řád nižší, například na úrovni jedna ku stu, pořád by šlo o nezanedbatelný problém. A to především kvůli možným dopadům v podobě úmrtí miliard lidí žijících v současnosti a zabránění životu desítek až stovek miliard lidí žijících v budoucnosti.

Podobně varovný pohled nabízí například Bostrom, který situaci, ke které pomalu směřujeme, analogicky popisuje, jako by myši vyvíjely umělého člověka za účelem, aby jim poskytoval sýr.²³¹ Rozdíl v inteligenci mezi námi a pokročilou AI může být podobně významný jako mezi myši a člověkem. Je realistické předpokládat, že by myši byly v takové situaci schopné člověka plně ovládat a že by člověk dělal pouze to, co myši původně zamýšlely, aby dělal? Spíše nikoliv. V kontextu dlouhodobé bezpečnosti AI systémů jsou proto klíčové dva koncepty:

- 1) problém kontroly;²³²
- 2) sladění AI systémů s lidskými zájmy a hodnotami (tzv. AI alignment).²³³

V prvním případě jde o stále více naléhavou otázku, jak udržet kontrolu nad něčím, co bude každého z lidí převyšovat ve schopnostech i intelektu. V druhém případě jde o problém, jak zajistit, aby AI systémy jednaly v souladu s našimi nejlepšími zájmy a hodnotami, a to nejen úmyslně, ale i skrze nezamýšlené důsledky jejich jednání. Podíváme-li se na tento problém z pohledu práva, bude nás zajímat, jak může právní věda či regulace přispět v dlouhodobém horizontu k tomu, aby vývoj a vznik pokročilé umělé inteligence nevedl ke katastrofickým následkům. Obecně půjde o následující čtyři oblasti výzkumu, které by měly být prioritizovány:

- 1) snižování existenčních rizik spojených s umělou inteligencí;

²²⁹ Podstatná část této kapitoly vychází z článku HORENOVSKY, J.; BAJGAR, O. Negative Human Rights as a Basis for Long-term AI Safety and Regulation. *Journal of Artificial Intelligence Research*, 2023, vol. 76, s. 1043–1075. Oba autoři mají sdílené první autorství textu (shared first authorship). Tato informace je obsažena v části „Acknowledgment“ v samotném článku.

²³⁰ ORD, T. *Nad propastí: Existenční riziko a budoucnost lidstva*. Praha: Argo, 2022, s. 166 a násl.

²³¹ BOSTROM, N. *Superintelligence: až budou stroje chytřejší než lidé*. Praha: Prostor, 2017.

²³² RUSSELL, S. *Human Compatible. AI and the Problem of Control*. London: Allen Lane, 2019, s. 10–15.

²³³ GABRIEL, I. Artificial Intelligence, Values, and Alignment. *Minds & Machines*, 2020, vol. 30, s. 411–437.

- 2) snižování rizik spojených s utrpením v důsledku umělé inteligence;
- 3) společenské sdílení přínosů umělé inteligence;
- 4) meta-výzkum v oblasti umělé inteligence.²³⁴

V kontextu právní vědy zde přitom dochází k zajímavému větvení. Prvním směrem jdou otázky týkající se regulace výrobců, vývojářů, uživatelů atd. Toto je oblast, kterou jako právníci známe velmi dobře a umíme s ní pracovat. Druhý směr je komplikovanější, neboť se dotýká regulace samotné pokročilé umělé inteligence. Zahrnuje složité a interdisciplinární otázky typu, jak přenést dodržování právních pravidel a hodnotových pozic do samotné technologie nebo jak využít poznatky obsažené v právní vědě k ovlivnění jednání technologie, která ale v ničem nebude připomínat lidskou mysl a lidskou inteligenci. V této části se dále budu věnovat právě druhému směru.

Jak se systémy vytvořené člověkem stávají stále více autonomními – tzn. schopnými jednat bez lidského vedení ve stále širším spektru situací – budeme je muset vybavit obecnými principy, které jim umožní rozhodnout, které činnosti nebo výsledky jsou pro člověka žádoucí a kterým je třeba se naopak vyhnout. Při použití úzce vymezených pokynů totiž hrozí, že umělá inteligence bude směřovat k řešením, která tyto úzce specifikované pokyny sice splní, ale zároveň budou v rozporu s širšími lidskými hodnotami a zájmy.

Například jsme se již setkali s tvrzeními, že doporučovací systémy na sociálních sítích mohou uživatele tlačit k extrémnějším názorovým pozicím. To sice algoritmům pomáhá naplnit jejich úzký cíl maximalizace času, který uživatel stráví na webové stránce tím, že se uživatelé stanou předvídatelnějšími, ale zároveň jde o postup, který je v rozporu s širšími lidskými preferencemi ohledně toho, co by měl algoritmus dělat.^{235 236}

Vzhledem k tomu, že AI systémy budou do budoucna stále schopnějšími, pravděpodobně v čase významně porostou náklady na takzvaná *selhání specifikace* (specification failures). Z tohoto důvodu se již v současnosti objevují výzvy, aby se pokročilé AI systémy učily lidským hodnotám nebo preferencím, a to s cílem vyřešit tzv. *AI alignment*

²³⁴ WINTER, Ch. (eds.). *Legal Priorities Research: A Research Agenda*. Legal Priorities Project, 2021, s. 36.

²³⁵ RUSSELL, S. *Human Compatible: Artificial Intelligence and the Problem of Control*. London: Allen Lane, 2019.

²³⁶ Pro další příklady viz KRAKOVNA, V. et al. Specification gaming: the flip side of AI ingenuity. In: *Google Deep Mind* [online]. 21. 4. 2020. Dostupné z: <https://deepmind.google/discover/blog/specification-gaming-the-flip-side-of-ai-ingenuity/>.

problem.²³⁷ O jaké hodnoty by se však mělo jednat? Pojem *lidské hodnoty* vyvolává široké neshody mezi lidmi ohledně jejich vlastních hodnot a preferencí. To se stává vážným problémem, pokud uvažujeme o potřebě závazné regulace, která bude nezbytná pro globální řešení negativních externalit vyplývajících z umělé inteligence – včetně existenčních rizik pro lidstvo.²³⁸

Pokud má regulace vyžadovat dodržování souboru zásad nebo hodnot, musí alespoň nepřímou specifikovat, o jaké hodnoty se jedná. A má-li být takový regulační rámec přijat a prosazován na mezinárodní úrovni, potřebujeme soubor hodnot, které jsou samy o sobě široce uznávány mezinárodním společenstvím. Tedy zde existuje dvojí požadavek na hodnoty – ty by měly být zároveň obecné a široce přijímané.

Návrh, který jsem vypracoval společně s Ondřejem Bajgarem a který jsme publikovali v podobě článku *Negative Human Rights as a Basis for Long-term AI Safety and Regulation*,²³⁹ ze kterého tato část vychází, spočívá v aplikaci konceptu *negativních lidských práv*²⁴⁰ jako minimálního souboru principů, které by měly respektovat všechny AI systémy a které jsou už dnes široce přijímány a chráněny. Návrh omezuje na *negativní práva*,²⁴¹ jelikož se nám požadavek, aby všechny systémy umělé inteligence napomáhaly plnění pozitivních lidských práv, nezdá smysluplný (přirozeným nositelem povinností je v tomto případě stát). Mezi

²³⁷ YUDKOWSKY, E. Complex Value Systems in Friendly AI. In: SCHMIDHUBER, J.; THORISSON, K. R.; LOOKS, M. (eds.). *Artificial General Intelligence. AGI 2011. Lecture Notes in Computer Science, vol 6830*. Berlin, Heidelberg: Springer Publishing, 2011.

²³⁸ BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford: OUP, 2014; PETIT, N. Law and Regulation of Artificial Intelligence and Robots—Conceptual Framework and Normative Implications. *SSRN*, 2017; PETIT, N.; DE COOMAN, J. Models of Law and Regulation for AI. *EUI Working Paper RSCAS 2020/63*, 2020; CRITCH, A.; KRUEGER, D. AI Research Considerations for Human Existential Safety (ARCHES). *arXiv:2006.04948*, 2020; NGO, R. AGI Safety From First Principles. In: *AI Alignment Forum* [online]. 2020. Dostupné z: <https://www.alignmentforum.org/s/mzgtmmTKKn5MuCzFJ>; SAETRA, H. S. *Big Data's Threat to Liberty Surveillance, Nudging, and the Curation of Information*. London: Academic Press, 2021.

²³⁹ HORENOVSKY, J.; BAJGAR, O. Negative Human Rights as a Basis for Long-term AI Safety and Regulation. *Journal of Artificial Intelligence Research*, 2023, vol. 76, s. 1043–1075.

²⁴⁰ Ačkoli používáme termín lidská práva, do tohoto rámce by mohla být v budoucnu zahrnuta i práva dalších entit, například zvířat (viz OWE, A.; BAUM, S. D. Moral Consideration of Nonhumans in the Ethics of Artificial Intelligence. *AI and Ethics*, 2021, vol. 1, s. 517–528) nebo digitálních myslí (SHULMAN, C.; BOSTROM, N. Sharing the World with Digital Minds. In: CLARKE, S.; SAVULESCU, J. (eds.). *Rethinking Moral Status*. Oxford: OUP, 2021). To by bylo v souladu s trendem rozšiřování sféry subjektů základních práv a s tím souvisejícího rozšiřování morálního spektra (ANTHIS, J. R.; PAEZ, E. Moral circle expansion: A promising strategy to impact the far future. *Futures*, 2021, vol. 130, 102756.).

²⁴¹ Negativní práva, na která se zaměřuje náš návrh, chrání subjekt před škodlivými vnějšími zásahy, na rozdíl od pozitivních práv, která svého nositele k něčemu opravňují, např. ke vzdělání. Pokud se zmiňujeme o lidských právech, máme zpravidla na mysli pouze negativní práva, pokud není uvedeno jinak.

negativními lidskými právy se zpočátku navrhuje soustředit především na ty, na nichž již panuje elementární shoda – například právo na život nebo ochranu vlastnictví.²⁴²

Abychom tomuto zdánlivě abstraktnímu konceptu dali konkrétní význam, inspirujeme se soudním systémem, který je navržen tak, aby dával autoritativní odpověď, zda je určité jednání dovoleno. Tímto způsobem jsou již dnes lidská práva chráněna, a to prostřednictvím mezinárodního i vnitrostátního práva a soudních systémů. V tomto duchu náš návrh chápe určité jednání umělé inteligence jako souladné s lidskými právy pouze tehdy, kdy by je konkrétní soudní systém posoudil jako souladné na základě úmluvy vymezující, která konkrétní práva jsou chráněna.²⁴³

V praxi by se rozhodovací systém přímo uplatňoval pouze v malé množině případů – cílem je především navrhnout budoucí systémy umělé inteligence tak, aby samy předvíдалy, které chování může porušovat negativní lidská práva, a vyloučily takové strategie ještě před jejich provedením. Náš návrh také řeší související technický úkol klasifikovat jakékoli konkrétní (hypotetické) jednání jako souladné či nesouladné s negativními lidskými právy, přičemž rozhodovací systém by sloužil jako zdroj referenční pravdy. Toto je druh úkolu vhodný pro paradigma strojového učení. Náš návrh tedy řeší jak institucionální struktury, které mají dlouhodobě zajistit bezpečnost umělé inteligence, tak zarámování souvisejícího technického problému.

Náš návrh přispívá ke stavu poznání následujícími body:

- Poskytujeme nové rámování *AI alignment* problému, přičemž zdůrazňujeme potřebu oddělit pozitivní cíle a bezpečnostní omezení. V souladu s tím se zaměřujeme na to, jak by se systém neměl chovat, nikoli na to, jak by se chovat měl.

²⁴² Klíčovým principem, na němž stavíme, je konsenzus, který v současné době globálně nezahrnuje ochranu nelidských vnímajících bytostí, snad kromě základní ochrany proti týrání zvířat v některých zemích. Používáme tedy termín lidská práva, aby odpovídal současnému historickému kontextu, čímž se však nijak nesnažíme normativně prohlásit, že by práva nelidských bytostí měla být dlouhodobě vyloučena. Nelidské subjekty se již ostatně objevily v některých deklaracích souvisejících s umělou inteligencí, například v *Montrealské deklaraci o odpovědném rozvoji umělé inteligence* z roku 2018 (Montreal Declaration for a Responsible Development of Artificial Intelligence. 2018. Dostupné z: <https://montrealdeclaration-responsibleai.com/the-declaration/>). Podobně je tomu u práv dosud neexistujících subjektů, jako jsou budoucí generace.

²⁴³ Takový postoj navazuje na právní realismus (viz např. FISHER, W. W.; HORWITZ, M. J.; REED, T. A. *American Legal Realism*. Oxford: OUP, 1993); spíše než o teoretický význam a povahu práva se snažíme předvídat právo způsobem, jak bude skutečně v realitě praktikováno.

- Na základě poznatků z právní teorie navrhujeme koncept negativních lidských práv jakožto hodnotový základ bezpečnostních omezení a vysvětlujeme, proč jsou pro tento účel vhodná.
- Vysvětlujeme, jak by se dodržování negativních lidských práv mohlo využít při vytváření technických řešení pro bezpečnou umělou inteligenci, a nastiňujeme některé související výzvy jako inspiraci pro budoucí technickou práci. Rovněž uvádíme některé stávající oblasti, na nichž by taková práce mohla stavět.
- Oproti tomu, co považujeme za standardní způsob regulace – tedy sérii úzce zaměřených pravidel cílených na konkrétní problémy, které se již vyskytly – nabízí náš návrh obecnější pokyny pro systémy umělé inteligence s větší šancí uplatnění v nepředvídaných či nepředvídatelných budoucích situacích.
- Pro konkrétnější ilustraci předchozích bodů předkládáme návrh, jak by mohl fungovat mezinárodní regulační systém založený na našem návrhu.

Upozorňujeme také na následující tři omezení:

- Poskytujeme pouze částečný příspěvek k řešení problému specifikace při návrhu bezpečné umělé inteligence, který bude třeba doplnit vhodnými pozitivními cíli (ve směru *AI alignmentu*), jakož i dalšími bezpečnostními opatřeními, včetně právních.
- Neposkytujeme konkrétní technická řešení, ale spíše rámec, který vymezuje dodržování negativních lidských práv jako úlohu strojového učení, která je přístupná budoucí technické práci.
- V některých částech předpokládáme pokročilé schopnosti budoucích AI systémů, zejména pokud jde o předvídaní dopadů jejich činnosti. Zatímco schopnosti AI systémů se rychle rozvíjejí, jejich konkrétní podoba a míra zůstávají do určité míry předmětem spekulací, což znamená, že některé naše argumenty zůstávají odpovídajícím způsobem nejisté.

Naše práce je koncipována následovně: Po přehledu některých existujících studií nejprve stručně vysvětlujeme, na jaké systémy umělé inteligence se zaměřujeme. Poté nastiňujeme naše pojetí lidských práv pro tento účel a paralelu s jejich využitím v právu. Následně pokračujeme technickými úvahami o bezpečnosti umělé inteligence založené na negativních lidských právech. Na závěr popisujeme návrh, jak by bylo možné tuto bezpečnost začlenit do systému mezinárodních organizací.

4.2.1. Dosavadní práce

Naše analýza navazuje především na tři okruhy již existujících prací: (1) na akademické práce pojednávající o umělé inteligenci a lidských právech; (2) na seznamy obecných zásad, kterými se má řídit vývoj umělé inteligence; (3) na právní a politické dokumenty zabývající se otázkami lidských práv v souvislosti se systémy umělé inteligence.

V prvním okruhu se lidská práva již v současnosti objevují v diskusích o regulaci umělé inteligence. Články v této oblasti lze rozdělit jednak podle toho, zda uvažují pouze o krátkodobé perspektivě (takové většinou poukazují na problémy s aktuálně využívanými technologiemi), nebo zda uvažují i o dlouhodobé perspektivě, včetně pokročilejších budoucích systémů. Dále můžeme rozlišovat práce podle toho, zda zohledňují také technická řešení, nebo zda zůstávají pouze na abstraktní úrovni, či na úrovni interpretace reálných dopadů umělé inteligence optikou lidských práv.

Lze konstatovat, že se většina prací zaměřuje na již existující problémy – a v této skupině existují mnohé studie, které je nahlízejí optikou lidských práv.²⁴⁴ Tento směr je užitečný, neboť zdůrazňuje význam lidských práv pro umělou inteligenci a zahrnuje velkou část problémů spojených s umělou inteligencí do jednotného rámce. Mnoho práce bylo také odvedeno na řešení technických problémů souvisejících se současným porušováním lidských práv – nejvýznamnější se týká zamezování předsudků a diskriminace.²⁴⁵

Náš návrh se na rozdíl od výše zmíněných pokouší přejít od reakce na již existující problémy k přístupu více orientovanému na budoucnost. Také po technické stránce se snažíme posunout od mozaiky řešení konkrétních bezpečnostních opatření k obecnějšímu, komplexnímu řešení. Částečná řešení budou pravděpodobně po dlouhou dobu ve svých oblastech nadále dominovat. I z tohoto důvodu však existuje významné riziko, že se mezi partikulárními bezpečnostními přístupy objeví zásadní mezery a z nich vyloučí nové

²⁴⁴ Viz např. LATONERO, M. *Governing Artificial Intelligence: Upholding Human Rights & Dignity*. Data & Society, 2018; COWGER, A. R. *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*. Lexington Books, 2020; ANDERSEN, L. et al. *Human Rights in the Age of Artificial Intelligence*. Access Now. 2018. Dostupné z: <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>; RODRIGUES, R. Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 2020, vol. 4, 100005; AIZENBERG, E.; VAN DEN HOVEN, J. Designing for human rights in AI. *Big Data & Society*, 2020, vol. 7, no. 2, s. 1–14; RASO, F.; HILLIGOSS, H.; KRISHNAMURTHY, V.; BAVITZ, C.; KIM, L. *Artificial Intelligence & Human Rights: Opportunities & Risks*. The Berkman Klein Center, 2018.

²⁴⁵ MEHRABI, N.; MORSTATTER, F.; SAXENA, N.; LERMAN, K.; GALSTYAN, A. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 2021, vol. 54, no. 6, s. 1–35.

problémy – z tohoto důvodu věříme, že ambiciózní a náročný úkol najít obecnější řešení bezpečnosti je smysluplný.

Existují práce, které v kontextu umělé inteligence zohledňují lidská práva v jejich plné šíři. Donahoe a Metzger je například obhajují jako nástroje, které efektivně dosahují cílů, ke kterým směřují jiné etické rámce – staví člověka do středu pozornosti, pokrývají široké spektrum problémů a opírají se o globální konsenzus.²⁴⁶ I náš návrh staví především na posledních dvou důvodech. První důvod je však také zajímavý, a to zejména svým zaměřením na reálné důsledky jednání AI systémů, nikoli na konkrétní mechanismy, kterými je újma způsobena (například zkreslenými údaji) a které je v rámci regulace jen velmi obtížně možné vyčerpávajícím způsobem předvídat.

Mezi další autory, kteří zkoumají lidská práva v dlouhodobém kontextu, patří Risse²⁴⁷ a Gabriel²⁴⁸. I přesto, že uznávají hodnotu lidských práv jako vodítka v oblasti správy (governance) umělé inteligence, nevytvořili propojení s technickými řešeními, jež jsou předmětem našeho úsilí. Podobně máme dojem, že technickou stránku téměř zcela opomíjí Fjeld a Nagy²⁴⁹, Andersen²⁵⁰ i Rodrigues²⁵¹. Naším cílem je poskytnout rámec, který by pokrýval jak administrativní, tak technickou stránku problému způsobem, jenž by mohl ve svých principech v zásadě, pokud ne v detailu, dlouhodobě fungovat.

Skutečnost, že se lidská práva objevují jak v krátkodobém, tak v dlouhodobém kontextu, naznačuje, že by mohla přispět k propojení těchto dvou perspektiv. Pokud robustní ochrana lidských práv pomáhá jak při řešení krátkodobých, tak dlouhodobých hrozeb, mohla by lidská práva tvořit základ pro tzv. *neúplně teoretizovanou shodu* (incompletely theorized agreement), jak k tomu vyzývají například Stix a Maas²⁵². Díky tomu by obě tyto perspektivy nemusely

²⁴⁶ DONAHOE, E.; METZGER, M. M. Artificial Intelligence and Human Rights. *Journal of Democracy*, 2019, vol. 30, no. 2, s. 115–126.

²⁴⁷ RISSE, M. Human Rights and Artificial Intelligence: An Urgently Needed Agenda. *Human Rights Quarterly*, 2019, vol. 41, no. 1, s. 1-16.

²⁴⁸ GABRIEL, I. Artificial Intelligence, Values, and Alignment. *Minds & Machines*, 2020, vol. 30, s. 411–437.

²⁴⁹ FJELD, J.; NAGY, A. *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. The Berkman Klein Center for Internet & Society, Harvard University, 2020. Dostupné z: <https://cyber.harvard.edu/publication/2020/principled-ai>.

²⁵⁰ ANDERSEN, L. et al. *Human Rights in the Age of Artificial Intelligence*. Access Now. 2018. Dostupné z: <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>.

²⁵¹ RODRIGUES, R. Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 2020, vol. 4, 100005.

²⁵² STIX, C.; MAAS, M. M. Bridging the Gap: The Case for an ‘Incompletely Theorized Agreement’ on AI Policy. *AI and Ethics*, 2021, vol. 1, s. 261–71.

nutně konvergovat na filozofické úrovni, nýbrž alespoň by se mohly sbližovat na úrovni praktických řešení.

McGregor a kolektiv²⁵³ v článku, který je svým charakterem velmi blízký našemu návrhu, doporučují jako rámec pro algoritmickou odpovědnost především mezinárodní právo v oblasti lidských práv. Jejich text je nicméně psán s ohledem na transparentnost algoritmů, zatímco náš se zaměřuje na dlouhodobou bezpečnost umělé inteligence. I přesto se domníváme, že obsah našich článků je navzájem kompatibilní a poskytuje sice nepřiliš související, a přesto vzájemně se doplňující, materiál. Rovněž je třeba podotknout, že zatímco stávající mezinárodní právo v oblasti lidských práv se zaměřuje především na státy jako nositele povinností, náš návrh zdůrazňuje také povinnosti organizací vyvíjejících umělou inteligenci. Především nám však jde o mechanismy v samotných systémech umělé inteligence – což je potřebný posun, který zdůrazňuje například Smuha.²⁵⁴

Nejbližše našemu pohledu, avšak s větším filozofickým důrazem, je Gabriel²⁵⁵, který uvádí tři návrhy ve vztahu k problému *AI alignment*, které jsou relevantní pro naši tezi. Za prvé, podobně jako my, tvrdí, že normativní a technické aspekty jsou vzájemně propojeny. Za druhé upozorňuje na nezbytnost mít jasný cíl a rozlišuje mezi minimalistickým a maximalistickým pojetím. Náš přístup se v tomto dělení řadí k minimalistickému přístupu, protože se primárně snažíme vyhnout nebezpečným dopadům, nikoli maximalizovat pozitivní hodnotu umělé inteligence. Za třetí zdůrazňuje nutnost určit správné principy nebo hodnoty, na kterých lze *alignment* opřít. Jako jednu z možností uvádí lidská práva, jež však my rozpracováváme mnohem podrobněji.

Mezi maximalistickým a minimalistickým pojetím se nachází sféra *kooperativní umělé inteligence*²⁵⁶, která usiluje o vytvoření umělé inteligence, jež by napomáhala dosahování vzájemně prospěšných výsledků ve světě, kde se cíle jednotlivých aktérů liší. Minimalistický aspekt tohoto cíle – aby si aktéři navzájem nezpůsobovali újmu – identifikujeme jako faktor

²⁵³ MCGREGOR, L.; MURRAY, D.; NG, V. International Human Rights Law as a Framework for Algorithmic Accountability. *International and Comparative Law Quarterly*, 2019, vol. 68, no. 2, s. 309–343.

²⁵⁴ SMUHA, N. A. Beyond a Human Rights-Based Approach to AI Governance: Promise, Pitfalls, Plea. *Philosophy & Technology*, 2021, vol. 34, s. 91–104.

²⁵⁵ GABRIEL, I. Artificial Intelligence, Values, and Alignment. *Minds & Machines*, 2020, vol. 30, s. 411–437.

²⁵⁶ DAFOE, A.; BACHRACH, Y.; HADFIELD, G.; HORVITZ, E.; LARSON, K.; GRAEPEL, T. Cooperative AI: machines must learn to find common ground. *Nature*, 2021, 593, s. 33–36; DAFOE, A.; HUGHES, E.; BACHRACH, Y.; COLLINS, T.; MCKEE, K.R.; LEIBO, J.Z.; LARSON, K.; GRAEPEL, T. Open Problems in Cooperative AI. *arXiv:2012.08630*, 2020.

plně kompatibilní s přímou regulací. Také širší ambice kooperativní umělé inteligence však vnímáme jako užitečné rozšíření našeho návrhu, hodné dalšího následování.

V další oblasti literatury, vedle již zmíněných akademických prací, publikovala v posledních několika letech řada mezinárodních organizací, vlád a soukromých společností soubory obecných zásad umělé inteligence, přičemž mnoho z nich se týká lidských práv. Jobin a kolektiv²⁵⁷ poskytují dobrou rekapitulaci těchto publikovaných souborů. Mnohé z nich, například doporučení UNESCO v oblasti etiky umělé inteligence,²⁵⁸ lidská práva přímo zmiňují. Jiné obsahují principy, které lze snadno interpretovat jako spadající do oblasti negativních lidských práv. Jobin a kolektiv²⁵⁹ ve své studii zjistili, že 60 z 84 souborů principů obsahuje zásadu *non-maleficence* (která také zahrnuje například bezpečnost a ochranu před újmou) a 47 obsahuje ochranu soukromí. Obě tyto kategorie lze přitom interpretovat jako podmnožinu lidských práv. Navíc 60 souborů zahrnovalo zásadu odpovědnosti (zahrnující jak odpovědnost v anglickém smyslu *accountability*, tak odpovědnost v právním smyslu, *liability*), kterou také zdůrazňujeme v kontextu lidských práv.

Z našeho pohledu jsou mnohé z těchto souborů principů ve srovnání s naší prací cílené na kratší časové období a téměř žádné nevysvětlují význam uvedených pojmů v dostatečné podrobnosti. Také se příliš nevěnují tomu, jak by měly být uváděny do praxe jak z technického, tak z administrativního hlediska. Tato témata se snažíme zpracovat mnohem podrobněji, než je v dokumentech o zásadách obvyklé. V podstatě navrhuje nahradit mnohé z těchto zásad negativními lidskými právy, která poskytují konzistentnější, komplexnější a univerzálnější rámec připravený na výzvy, které mohou v budoucnu nastat.

Otázka možného ohrožení lidských práv ze strany systémů umělé inteligence se rovněž objevuje v strategických dokumentech a návrzích regulací. Například v *Aktu o umělé inteligenci* je ochrana lidských práv považována za jeden z hlavních důvodů, proč je třeba vytvořit novou regulaci. V kontextu USA je ohrožení lidských práv zmíněno například v *Závěrečné zprávě Komise pro národní bezpečnost o umělé inteligenci*²⁶⁰. Podobný přístup zaujímá i Australská

²⁵⁷ JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 2019, vol. 1, s. 389–399.

²⁵⁸ UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2021. Dostupné z: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

²⁵⁹ JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 2019, vol. 1, s. 389–399.

²⁶⁰ NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE. Final Report. 2021. Dostupné z: <https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>.

komise pro lidská práva ve své *Závěrečné zprávě o lidských právech a technologiích*²⁶¹, která se věnuje řadě potenciálních problémů souvisejících s umělou inteligencí a lidskými právy. Tyto práce však často opomíjejí úvahy o tom, jak by se ochrana lidských práv mohla v dlouhodobém horizontu rozvíjet, a nezabývají se tím, jak by se požadavek na respektování lidských práv mohl promítnout do konkrétních technických řešení – právě tyto mezery se v naší práci snažíme zaplnit.

4.2.2. Umělá inteligence

Naše práce se primárně zaměřuje na dlouhodobou bezpečnost umělé inteligence – pokrývá tak jakékoliv systémy vytvořené člověkem, jež autonomně provádí úkoly, které lidé běžně spojují s inteligencí. Umělá inteligence představuje širokou škálu technologií, a tak na jedné straně nastiňujeme, které technologie je náš rámec schopen zahrnout, a na druhé straně uvádíme, na které typy systémů se specificky zaměřuje a pro které by byl nejvíce přínosný.

Na nejobecnější úrovni se zdá být rozumné požadovat, aby všechny technologie, u nichž by to mohl být potenciálně problém, neporušovaly negativní lidská práva – například neubližovaly lidem, nepoškozovaly jejich majetek nebo neomezovaly jejich svobodu.²⁶² Obecný cíl našeho návrhu – aby umělá inteligence respektovala negativní lidská práva – je tedy relativně nezávislý na tom, kde definujeme hranici toho, co považujeme za systém umělé inteligence, a může zahrnovat současné systémy, jako jsou autonomní automobily, robotické vysavače nebo doporučovací algoritmy. Dokonce již existují výzvy k aplikaci lidskoprávního rámce k hodnocení rizik současných systémů.²⁶³

Výhodou tohoto obecného přístupu je, že by se administrativní prvky tohoto bezpečnostního rámce mohly začít rozvíjet relativně brzy, i když by se na technické úrovni v blízké době mohly nadále používat tradiční metody zajišťující bezpečnost. Také postupné změny současné právní infrastruktury (například v oblasti občanského práva) by mohly být dostačující. Ve skutečnosti mnohá porušení lidských práv prostřednictvím technologií jsou již

²⁶¹ AUSTRALIAN HUMAN RIGHTS COMMISSION. Human Rights and Technology Final Report. 2021. Dostupné z: <https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-final-report-2021>.

²⁶² To je v souladu s tím, že obecné zásady právní ochrany se zabývají prevencí jakékoli újmy – na takto obecné úrovni nemá umělá inteligence ani jiné nově vznikající technologie žádné zvláštní postavení.

²⁶³ LATONERO, M. *Governing Artificial Intelligence: Upholding Human Rights & Dignity*. Data & Society, 2018.

podle současných zákonů nelegální. Přesto může být vhodné implementovat zde navrhované obecnější mechanismy ještě před zavedením pokročilejších technologií.²⁶⁴

Předmětem našeho hlavního zájmu však jsou systémy ve vzdálenější budoucnosti, které by mohly operovat v širším spektru situací (a tak vyžadovat obecnější rozhodovací principy) a mít mnohem větší rozhodovací prostor (a být tak schopné přicházet s řešeními, která jsou pro člověka hůře předvídatelná). V důsledku toho budeme potřebovat obecnější opatření k zabránění nežádoucích důsledků jak na technické úrovni (rozhodování samotných systémů umělé inteligence), tak na úrovni správy/administrace (governance). Účelem je, abychom motivovali vlády a organizace vyvíjející umělou inteligenci k implementaci efektivních bezpečnostních opatření způsobem, který bude škálovatelný s rostoucí obecností AI systémů daleko nad rámec toho, co mohou pokrýt současná, často úzce zaměřená právní pravidla.²⁶⁵

Ačkoli obecné zásady mohou poskytnout užitečné vodítko, ne všechny systémy umělé inteligence budou mít schopnost naučit se rozpoznávat negativní lidská práva v plné šíři a u některých to ani nebude potřeba. Úroveň porozumění právům, která je pro konkrétní systém potřebná, by měla být přiměřená stupni jeho autonomie a konkrétním rizikům, která představuje. Předpokládáme však, že pokud bychom v budoucnu vytvořili systémy, které přesahují určitou úroveň obecnosti a autonomie, bylo by plné chápání a dodržování negativních lidských práv žádoucí. Náš návrh je koncipován s ohledem na takovou situaci. Když tedy hovoříme o umělé inteligenci nebo systémech s umělou inteligencí, čtenář by si především měl představit tuto situaci. Avšak mnohé prvky návrhu jsou prostřednictvím vhodných dílčích řešení použitelné i pro systémy menšího rozsahu.

Náš návrh se zaměřuje na civilní využití umělé inteligence se záměrem pokud možno zcela eliminovat porušování základních lidských práv. Vojenské a jiné bezpečnostní aplikace umělé inteligence, pokud jsou přípustné, představují složitější případ. Ten prozatím neřešíme, ale je diskutován jinde.²⁶⁶ Mnohé aspekty zde navržených rámcových řešení by však mohly být přizpůsobeny, aby byly užitečné i pro militární AI systémy. Mohlo by to být například realizováno nahrazením konceptu základních negativních lidských práv principy (vhodně

²⁶⁴ CAVE, S.; ÓHÉIGEARTAIGH, S. S. Bridging near-and long-term concerns about AI. *Nature Machine Intelligence*, 2019, vol. 1, no. 1, s. 5–6.

²⁶⁵ To může zahrnovat i otázky právní odpovědnosti organizací vyvíjejících umělou inteligenci.

²⁶⁶ PETIT, N. Artificial Intelligence and Automated Law Enforcement: A Review Paper. *SSRN*, 2018; MAAS, M. M.; LUCERO-MATTEUCCI, K.; COOKE, D. Military Artificial Intelligence as Contributor to Global Catastrophic Risk. In: BEARD, S. J. (eds.). *The Era of Global Risk*. Open Book Publishers, 2023; WARREN, A.; HILLAS, A. Lethal autonomous weapons systems: Adapting to the future of unmanned warfare and unaccountable robots. *Yale Journal of International Affairs*, 2017, vol. 12, no. 1, s. 71–85.

rozšířeného) mezinárodního humanitárního práva. Tuto možnost však ponecháváme nerozpracovanou.

4.2.3. Lidská práva

Nyní se zaměříme na vysvětlení, jaké pojetí lidských práv náš návrh přijímá a jak se inspiruje jejich používáním v právu. Z pohledu právní teorie lidské právo představuje formu normativní věty – jde tedy o výrok, který uvádí, že by něco „mělo být“.²⁶⁷ Například čl. 5 EULP stanoví, že „každý má právo na svobodu a osobní bezpečnost“. To v praxi znamená, že svoboda a osobní bezpečnost nemají být, až na právem povolené situace, omezovány. Toto ustanovení také vymezuje obecný základ pro rozhodování v konkrétních situacích. Jak však uvádí Bellamy, takové ustanovení je samo o sobě příliš neurčité a abstraktní:

„Charty jsou nezbytně formulovány na vysoké úrovni abstrakce, aby bylo dosaženo dohody. To ponechává neshody ohledně jejich obsahu, působnosti, sféry a subjektů práv, stejně jako ohledně nejlepších způsobů, jak je zabezpečit apod. k budoucímu rozhodnutí. [...] Dokonce i taková práva, která zrcadlí to, co prakticky všichni označují za velké morální zlo, například vražda nebo znásilnění, mohou nadále obsahovat neshody ohledně toho, jak by měla být vykládána a jak by měla působit v relaci s ostatními právy.“²⁶⁸

Abychom tedy pochopili skutečný obsah lidských práv, musíme se dívat nejen na abstraktní textaci slov, ale také na skutečný dopad jednotlivých práv na reálný svět: na způsoby chování, kterým se aktéři rozhodnou vyhnout s ohledem na nositele práv. Z právního hlediska by to mělo přibližně odpovídat tomu, co by rozhodovací systém – typicky prostřednictvím soudů – za určitých okolností posoudil jako protiprávní porušení lidského práva. Takže místo toho, abychom se příliš zaměřovali na definice lidských práv, zřejmě postačí určit, které chování je a není přípustné. Jelikož fungování rozhodovacích systémů je již zavedenou a prověřenou metodou interpretace abstraktních konceptů, jako jsou lidská práva v praktických situacích, v rámci našeho návrhu je význam lidských práv určen následujícím způsobem:

Určité jednání (skutečné nebo hypotetické) je v souladu s lidskými právy, pokud by jej tak posoudil konkrétní rozhodovací systém.

²⁶⁷ KELSEN, H. *General Theory of Norms*. Oxford: Clarendon Press, 1991.

²⁶⁸ BELLAMY, R. *Political Constitutionalism: A Republican Defence of the Constitutionality of Democracy*. Cambridge: Cambridge University Press, 2007, s. 50–51.

Detaily fungování tohoto procesu a jeho odůvodnění přesahují rámec našeho návrhu. Nicméně, při hodnocení, zda došlo k porušení lidských práv, je obvykle třeba odpovědět na tři otázky:

- 1) Která specifická lidská práva jsou v dané situaci relevantní?
- 2) Zasahuje konkrétní činnost umělé inteligence do těchto práv?
- 3) Pokud ano, lze tento zásah (s ohledem na interní pravidla normativního systému) odůvodnit právem přijatelnými důvody?

Z třetího uvedeného bodu vyplývá, že ne každý zásah do lidského práva je vždy protiprávní (jinými slovy, že jde o jeho porušení). Například právo na svobodu projevu může být v některých případech omezeno ve prospěch ochrany národní bezpečnosti atd.

Z této perspektivy je cílem právního systému být racionálním systémem poskytujícím jedinou autoritativní a definitivní odpověď. Právní systém přistupuje k jednání, ať už k činům, nebo k jejich následkům, binárně – a to v kódu legální/nelegální.²⁶⁹ Konkrétní činnost tedy buď představuje protiprávní zásah do lidských práv, nebo nepředstavuje. V právní praxi, minimálně při konečném verdiktu u soudu, neexistuje varianta „možná“. Právo, včetně lidských práv, lze tedy chápat jako souhrn systematizovaných předpovědí.²⁷⁰ Tyto předpovědi se týkají možných rozhodnutí v konkrétních situacích na základě zákonů nebo úmluv, dřívější rozhodovací praxe a dalších relevantních zdrojů poznání práva. Z tohoto pohledu aktéři, ať už se jedná o fyzické osoby, nebo právnické osoby, neustále posuzují, zda je něco legální, na základě předvídání (hypotetických) soudních rozhodnutí.

Tato struktura právního a soudního systému by mohla být vhodná pro strojové učení.²⁷¹ Na základě tisíců existujících případů a literatury z různých jurisdikcí by se umělá inteligence mohla naučit odhadovat pravděpodobnost, že její jednání poruší některé z ustanovení v konkrétní úmluvě (ačkoli by byly nutné i další zdroje učení – o tom více později).

²⁶⁹ Viz LUHMANN, N. *Právo společnosti*. Praha: Academia, 2021.

²⁷⁰ HOLMES, O. W. The Path of the Law. *Harvard Law Review*, 1897, vol. 10, no. 8, s. 457–478.

²⁷¹ BENJAMIN, A. The Path of the Law: Towards Legal Singularity. *University of Toronto Law Journal*, 2016, vol. 66, no. 4, s. 443–455; CASEY, A. J.; NIBLETT, A. The Death of Rules and Standards. *Indiana Law Journal*, 2017, vol. 92, no. 4, s. 1401–1447.

4.2.3.1. Která práva by měla být chráněna

Právní teorie rozlišuje dva druhy lidských práv: pozitivní a negativní.²⁷² Podstatou pozitivních práv je, že by se někomu mělo něco poskytnout – může tak například vzniknout povinnost státu vyplácet sociální dávky. Naopak, podstatou negativních práv je vytvoření sféry osobní autonomie, kterou ostatní aktéři nesmí narušovat.²⁷³ Jinými slovy, negativní práva vytvářejí jakousi ochrannou bariéru proti zásahům. Systém, který navrhujeme, je postaven na negativních lidských právech. Podle něj by měla být podmnožina negativních lidských práv globálně závazná pro všechny systémy umělé inteligence, aby byly ochráněny nejdůležitější aspekty našich životů.²⁷⁴

Zatímco u pozitivních práv je přirozeným nositelem povinností stát jako orgán veřejné moci, negativní práva ve svém nejširším výkladu mohou chránit každého jednotlivce před neoprávněnými zásahy jakéhokoli subjektu, včetně systémů umělé inteligence. Nepopíráme, že může být smysluplné hledat u umělé inteligence pomoc i při zajišťování pozitivních práv. Nezdá se nám však, že by takovou pomoc šlo nalézt u všech AI systémů, a domníváme se, že by bylo lepší je podporovat prostřednictvím jiných mechanismů. Neexistuje totiž žádný důvod, proč by všechny AI systémy měly být povinny poskytovat například zdravotní péči nebo starobní penze. Naším cílem je zajistit, aby systémy umělé inteligence nezasahovaly do autonomní sféry lidí, nikoliv určovat, jaké pozitivní funkce by měly plnit.

Navíc se zdá, že negativní práva jsou v mezinárodním společenství méně kontroverzní než mnohá práva pozitivní, jako jsou hospodářská a sociální práva.²⁷⁵ A i když v současném světě nemusí existovat úplná shoda ohledně negativních lidských práv jako ochrany obyvatel před státy,²⁷⁶ domníváme se, že nalezení shody ohledně negativních lidských práv ve vztahu k AI systémům by mohlo být výrazně dosažitelnější. Nikdo totiž nechce, aby nebezpečná umělá

²⁷² JELLINEK, G. *System der subjektiven öffentlichen Rechte*. Akademische Verlagsbuchhandlung von J. C. B. Mohr, 1892.

²⁷³ MELRO, A.; OLIVEIRA, L. *Constitutional Knowledge and Its Impact on Citizenship Exercise in a Networked Society*. IGI Global, 2019, s. 140.

²⁷⁴ Je třeba poznamenat, že mnoho prima facie pozitivních práv zahrnuje také negativní aspekt. Ačkoli tedy právo na zdravotní péči primárně znamená povinnost státu poskytovat zdravotní péči (pozitivní aspekt), zahrnuje také negativní aspekt spočívající v tom, že nikdo nesmí zasahovat do přístupu ostatních ke zdravotní péči. Tento negativní aspekt může být také přenesen do systémů umělé inteligence, ale opět bychom to navrhovali pouze v případech, kdy existuje konsenzus.

²⁷⁵ COURTNEY J.; HIRSCHL, R.; ROSEVEAR, E. Economic and Social Rights in National Constitutions. *The American Journal of Comparative Law*, 2014, vol. 62, no. 4, s. 1043–1094.

²⁷⁶ LAW, D. S.; VERSTEEG, M. The Evolution and Ideology of Global Constitutionalism. *California Law Review*, 2011, vol. 99, no. 5, s. 1163-1258.

inteligence autonomně porušovala lidská práva například v ulicích – včetně autoritářských a nedemokratických režimů.

Náš návrh představuje posun v chápání lidských práv, která byla historicky vnímána především jako ochrana občanů před státem. My tento koncept rozšiřujeme také na ochranu před jednáním umělé inteligence. V některých, především evropských, právních rádech přitom horizontální účinek ústavně chráněných lidských práv²⁷⁷ již v současnosti umožňuje jejich přímé uplatnění na soukromoprávní subjekty jako nositele povinností. V souladu s touto praxí je jejich rozšíření na umělou inteligenci přirozené a vyplývá z rostoucí potřeby chránit lidi před potenciálně výkonnými autonomními stroji. Konkrétní chráněná práva by musel specifikovat dokument, jako je mezinárodní úmluva, která by byla výsledkem procesu politického projednávání. Například právo na život, osobní bezpečnost, vlastnické právo a právo nebýt mučen by ale téměř jistě zahrnuta byla.²⁷⁸

Ostatně, ochrana většiny negativních lidských práv má mnohem hlubší kořeny než samotný koncept lidských práv jako takový. To, co máme na mysli pod ochranou lidských práv, slouží do značné míry stejným cílům, jakým po staletí sloužilo trestní právo na celém světě, a co je tedy v rámci ochrany lidských práv z historických důvodů méně zdůrazňováno. Avšak tam, kde je trestní právo obvykle formulováno poměrně konkrétně a přesně, se domníváme, že poukazování na obecné hodnoty, které chrání (a které se podle našeho názoru do značné míry shodují s negativními lidskými právy), poskytuje mnohem lepší spolehlivost v nepředvídatelných budoucích situacích.

Je možné, že bude třeba doplnit nová práva, pokud bude dosaženo konsenzu. Russell například navrhuje právo na duševní bezpečnost, které zahrnuje také právo nebýt obelháván systémy umělé inteligence.²⁷⁹ Podobně Ienca a Andorno navrhují práva na kognitivní svobodu, duševní soukromí, duševní integritu a psychologickou kontinuitu.²⁸⁰

²⁷⁷ GARDBAUM, S. The "Horizontal Effect" of Constitutional Rights. *Michigan Law Review*, 2003, vol. 102, no. 3, s. 388–458

²⁷⁸ V této souvislosti stojí za zmínku, že McAllister (MCALLISTER, A. Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture. *Minnesota Law Review*, 2018, vol. 47, s. 2527–2573) tvrdí, že mučení prováděné systémy umělé inteligence může v současné době za určitých okolností projít skulinami stávajících mezinárodních zákazů mučení.

²⁷⁹ RUSSELL, S. *Human Compatible: Artificial Intelligence and the Problem of Control*. London: Allen Lane, 2019, s. 105–107.

²⁸⁰ IENCA, M.; ANDORNO, R. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy*, 2017, vol. 13, no. 1, s. 1–27.

V budoucnu se pravděpodobně dále přidají i práva nelidských entit, jako jsou zvířata nebo možná i digitální mysli. Navrhujeme rámec a mechanismus prosazování ochrany všech možných práv, na kterých může vzniknout dohoda. Tento přístup bere v potaz, že aby byly mechanismy lidských práv účinné pro bezpečnost umělé inteligence, musí nevyhnutelně dojít k jejich výrazným změnám.²⁸¹

Kromě potřebných změn v samotných systémech ochrany lidských práv před zavedením našeho návrhu je také třeba dát prostor vývoji v průběhu času. Když se společenský konsenzus změní, musí být legislativní a rozhodovací systémy schopné překonat zastaralé zdroje učení nebo vyřešit případné vnitřní rozpory, aby tak umožnily morální pokrok v chápání práv. Podobně legislativní a soudní pokrok zajistil, že nejsme stále uzamčeni v hodnotách a právních normách 19. století. Mechanismy, které k tomu slouží, budou rozebrány později.

4.2.3.2. Lidská práva jako relativní principy

Mnoho lidí mimo právní obec vnímá lidská práva jako „absolutní pravidla“. V kontextu umělé inteligence jsou absolutní pravidla v technické komunitě stereotypně považována za paralyzující vzhledem k nejistotě, neboť většina chování s sebou nese alespoň nepatrné riziko jejich porušení. Takové problémy však vznikají při odlišném významu slova „absolutní“, než jak se používá v právu – jde o pravidla, která nesmí být porušena za žádných okolností a která nestanovují žádnou přípustnou hranici rizika. Právní význam je jiný – pokud dojde k absolutně zakázanému chování, sankce se uplatní bezpodmínečně, ale tato sankce je konečná (i když u důležitých pravidel velmi vysoká²⁸²) a lze ji při plánování řešit pomocí obvyklých výpočtů očekávané hodnoty. Nejistota tedy u práv není větším problémem než ve většině ostatních oblastí výzkumu umělé inteligence.

Právní teorie navíc rozlišuje mezi pravidly a principy a lidská práva jsou z normativního pohledu kombinací obou.²⁸³ Některá jsou pravidly a některá „čistými principy, které je nutné pečlivě vyvážit“.²⁸⁴ Naprostá většina lidských práv je v tomto kontextu relativními právními

²⁸¹ SMUHA, N. A. Beyond a Human Rights-Based Approach to AI Governance: Promise, Pitfalls, Plea. *Philosophy & Technology*, 2021, vol. 34, s. 91–104.

²⁸² Předpokládá se, že by systém optimalizoval hodnotu pro svého vlastníka, která by byla kvantifikována pomocí funkce odměny. Když hovoříme o „velmi vysoké sankci“, máme na mysli, že by měla odpovídat ztrátě hodnoty, kterou by vlastník považoval za významnou, například by šlo o značnou část zisků. Lze o ní uvažovat podobně jako o sankcích v oblasti správního nebo trestního práva.

²⁸³ DWORKIN, R. *Taking Rights Seriously*. Harvard: Harvard University Press, 1977, s. 80–86.

²⁸⁴ ALEXY, R. The Absolute and the Relative Dimensions of Constitutional Rights. *Oxford Journal of Legal Studies*, 2017, vol. 37, no. 1, s. 34.

principy. Alexyho slovy: „*Principy jsou normy, které přikazují, aby se něco realizovalo v nejvyšší možné míře, která je fakticky a právně dosažitelná.*“²⁸⁵ Používá se přitom tzv. optimalizační příkaz, který žádá, aby se pro každý případ rozhodující orgán snažil nalézt optimální stav, kde je každý z principů co nejvíce naplněn a v případě konfliktu vážen podle své důležitosti.²⁸⁶ Jako klasický příklad vezměme konflikt mezi svobodou projevu a ochranou soukromí. Nelze předem říci, které z těchto práv má přednost. V praxi rozhodující orgán proto bude vyvažovat obě práva s ohledem na právní a skutkové okolnosti konkrétního případu. Umělá inteligence by se podobné vyvažování mohla naučit napodobit.

Na druhou stranu, tato logika neplatí pro právní pravidla, která se v případě konfliktu nevyvažují proti sobě. Ojedinelá lidská práva, která jsou právními pravidly, jsou tedy definitivními příkazy.²⁸⁷ Pokud mezi nimi dojde ke střetu, musí být určeno, které má přednost.²⁸⁸ Obvykle se uplatňuje pravidlo konkrétnější (*lex specialis*), nebo později přijaté (*lex posterior*). Představme si například pravidlo, které zakazuje zabití člověka, a jiné, které uvádí, že zabití člověka je přípustné v sebeobraně. Z formulace druhého pravidla vyplývá jeho aplikační přednost před prvním pravidlem při splnění konkrétních podmínek. Pokud tedy někdo někoho zabije v sebeobraně, nejedná protiprávně.

Lidská práva se také dělí na kategorie kvalifikovaných a absolutních práv. Kvalifikovaná práva mohou být omezena za stanovených podmínek, které lze rozdělit do dvou oblastí: a) za podmínek přímo stanovených v normativním textu chránícím dané právo a b) v případě střetu s jiným právním principem, jako je jiné lidské právo nebo veřejný zájem. Při rozhodování o konfliktu mezi lidskými právy a veřejnými zájmy pomáhají strukturovat odpověď rozhodovacího orgánu různé testy – například test proporcionality,²⁸⁹ v němž se proti sobě váží kolidující lidská práva a veřejné zájmy. Převážná většina lidských práv je v tomto smyslu kvalifikovanými právy.²⁹⁰

²⁸⁵ ALEXY, R. On the Structure of Legal Principles. *Ratio Juris*, 2000, vol. 13, no. 3, s. 295.

²⁸⁶ ALEXY, R. On the Structure of Legal Principles..., s. 294–304; ALEXY, R. *A Theory of Constitutional Rights*. Oxford: OUP, 2004.

²⁸⁷ ALEXY, R. The Absolute and the Relative Dimensions of Constitutional Rights. *Oxford Journal of Legal Studies*, 2017, vol. 37, no. 1, s. 31–47.

²⁸⁸ ALEXY, R. On the Structure of Legal Principles..., s. 294–304.

²⁸⁹ BARAK, A. *Proportionality: Constitutional Rights and their Limitations*. Cambridge: Cambridge University Press, 2012.

²⁹⁰ Mohou nastat případy, kdy má veřejný zájem přednost před konkrétními lidskými právy, např. když je svoboda projevu omezena v zájmu ochrany národní bezpečnosti. Mělo by se však jednat buď o předem jasně definované případy, nebo by o nich měl rozhodovat orgán veřejné moci, který k tomu má dostatečnou legitimitu.

Absolutní práva naproti tomu omezit nelze. Podle mezinárodního²⁹¹ a evropského práva lidských práv²⁹² je vzácnou výjimkou spadající do této kategorie právo nebýt mučen. To má proto při vyvažování vždy přednost a nelze jej poměřovat proti jiným právům či veřejným zájmům. Vzhledem k bezpodmínečnému uplatnění tohoto práva se tedy jedná o právní pravidlo a zároveň absolutní právo.

4.2.3.3. Nadřazenost lidských práv

V kontextu bezpečnosti a regulace umělé inteligence by měla být ochrana negativních lidských práv obecně nadřazena ostatním cílům. Tato práva by měla fungovat jako omezení, v rámci kterého by bylo možné sledovat primární cíle systému. Lze rovněž zavést další omezení, ochrana negativních lidských práv by však měla mít v případě konfliktu standardně přednost.

Bezpečnostní systém, který navrhujeme, spočívá ve vytvoření souboru minimálních bezpečnostních standardů založených na negativních lidských právech, které by mohly být doplněny spolehlivou specifikací pozitivních cílů, jež by měl systém umělé inteligence sledovat. V krajním případě by to mohlo vést až k úplnému poznání preferencí člověka nebo organizace, která systémem umělé inteligence provozuje. Tato sada by mohla být rovněž doplněna o specifikaci dalších nežádoucích chování, která mohou být důležitá v užších oblastech.²⁹³

Nicméně lidská práva by měla být považována za nadřazená ostatním specifikacím v tom smyslu, že systémům umělé inteligence by nikdy nemělo být dovoleno porušovat lidská práva v zájmu sledování svých pozitivních cílů. Toto je analogické s praxí mnoha právních systémů, kde lidská práva rovněž mají přednost před jinými pravidly, zásadami a cíli či preferencemi jednotlivců. Některými praktickými problémy při realizaci tohoto ideálu se zabýváme v následující části.

Standardně by tedy měla být dána přednost lidským právům. Rovněž, i když by lidská práva měla mít přednost, je třeba vždy ponechat otevřená vrátka pro jejich změnu nebo částečné nahrazení jiným bezpečnostním opatřením, pokud bude v budoucnu objeven lepší mechanismus. To souvisí s problémem zachování korigovatelnosti AI systémů.

²⁹¹ MAVRONICOLA, N. Is the Prohibition Against Torture and Cruel, Inhuman and Degrading Treatment Absolute in International Human Rights Law? A Reply to Steven Greer. *Human Rights Law Review*, 2017, vol. 17, no. 3, s. 479–498.

²⁹² *Chahal v. UK* [1996] 23 EHRR 413.

²⁹³ Obecně právo chrání více než jen lidská práva, např. hospodářskou soutěž, finanční stabilitu atd., které bude třeba chránit i před zásahy umělé inteligence. Rovněž může existovat nežádoucí chování s negativním dopadem na lidský blahobyt, které nedosahuje intenzity porušení lidských práv. Ochrana týkající se obou těchto bodů by mohla být úzce propojena se zde popsanou ochranou lidských práv.

4.2.3.4. Proč lidská práva, a ne obecné lidské hodnoty?

Hlavním argumentem, který odlišuje použití lidských práv od povinného uplatňování obecných lidských hodnot, je možnost zformování mezinárodního konsenzu – ten zdůrazňují např. Donahoe a Metzger.²⁹⁴ Lidská práva jsou totiž uznávána nejen v západních zemích, ale do jisté míry i jinde, například v Číně,²⁹⁵ Africe²⁹⁶ a Jižní Americe,²⁹⁷ i když často pod jinými názvy, jako jsou ústavní práva nebo základní práva. Aby byl náš návrh ještě přijatelnější, navrhuje se – namísto povinného uplatňování lidských práv v jejich úplnosti – zaměřit pouze na podmnožinu, nad kterou by se mohla v mezinárodním společenství vytvořit širší shoda. Tento požadavek by negativní lidská práva měla splňovat.²⁹⁸

Nepředkládáme konkrétní morální argumentaci ve prospěch lidských práv. V souladu s teorií neúplně teoretizované shody²⁹⁹ se účastníci musí shodnout pouze na konkrétních požadovaných výsledcích, nikoli na jejich teoretickém, ideologickém nebo kulturním vysvětlení. Různí lidé a národy mohou ochranu lidských práv opírat o různé základy – náboženské, etické či politické. Empiricky však pozorujeme, že výsledky jejich argumentace se v určitých základních aspektech ochrany do značné míry shodují, čímž tvoří to, co Rawls nazval *překrývajícím se konsensem*.³⁰⁰ Překrývající se konsensus byl již také jinými autory označen za potenciálně účinnou cestu k zajištění globální spolupráce s cílem podpořit bezpečnost v oblasti politiky umělé inteligence napříč kulturami.³⁰¹

Na teoretické úrovni představují lidská práva soudržný systém hodnot, který se vyvíjí v čase a je úzce spojen s praktickou mezinárodní politikou. Na druhou stranu, obecné lidské hodnoty, preference, přesvědčení a etické teorie se značně liší jak mezi jednotlivými státy, tak i v jejich rámci, nehledě na četné problémy s jejich přesným vymezením jako dobře

²⁹⁴ DONAHOE, E.; METZGER, M. M. Artificial Intelligence and Human Rights. *Journal of Democracy*, 2019, vol. 30, no. 2, s. 115–126.

²⁹⁵ Viz kapitola 2 Ústavy Čínské lidové republiky.

²⁹⁶ Viz Africká charta lidských práv a práv národů.

²⁹⁷ Viz Americká úmluva o lidských právech.

²⁹⁸ CULLEN, P. The Negative and Moral Right to Life. A Basis for Functional Human Rights. Southern New Hampshire University, Justice Studies Faculty Papers, 2015. Dostupné z: <http://hdl.handle.net/10474/3071>; COURTNEY J.; HIRSCHL, R.; ROSEVEAR, E. Economic and Social Rights in National Constitutions. *The American Journal of Comparative Law*, 2014, vol. 62, no. 4, s. 1043–1094.

²⁹⁹ SUNSTEIN, C. R. Incompletely Theorized Agreements Commentary. *Harvard Law Review*, 1994, vol. 108, no. 7, s. 1733–1772.

³⁰⁰ RAWLS, J. The Idea of an Overlapping Consensus. *Oxford Journal of Legal Studies*, 1987, vol. 7, no. 1, s. 1–25.

³⁰¹ ÓHÉIGEARTAIGH, S. S.; WHITTLESTONE, J.; LIU, Y.; ZENG, Y.; LIU, Z. Overcoming barriers to cross-cultural cooperation in AI ethics and governance. *Philosophy & Technology*, 2020, vol. 33, s. 571–593.

definovaných konceptů.³⁰² Bylo by proto nesmírně obtížné je zachytit v jediném soudržném systému, který by byl obecně přijímaný na celém světě, a poskytnout tak skrze ně vhodný základ pro mezinárodní regulační rámec.

Někdo by mohl namítnout, že takto omezený průnik hodnot by byl příliš malý na to, aby přinesl něco, na co bychom se mohli spolehnout. V praxi tomu však tak není: určité hodnoty jsou alespoň do určité míry chráněny prakticky ve všech zemích světa – lidský život, fyzická integrita nebo vlastnictví – a jsou spojeny s mechanismy, které jejich ochranu zpřesňují a především vynucují.

4.2.3.5. Lidská práva a etické teorie

Právo lidských práv absorbovalo směsici různých etických teorií a přesvědčení, a to kvůli potřebě soudů řešit etická dilemata v rámci své rozhodovací praxe. Například při rozhodování o možnosti odpojit pacienta od život udržující léčby je nutné zohlednit nejen právní aspekty, ale také etické teorie a morální citění.³⁰³ Řešení těchto otázek fakticky transformuje obsah vybraných etických teorií a morálních přesvědčení do rozhodovací činnosti soudů. Při svém rozhodování soudy nepřijímají konkrétní etické teorie, jako jsou deontologie nebo utilitarismus, ve své celistvosti. Místo toho kombinují argumenty z různých etických teorií a z hlediska obecného citění. Vybírají tak nejvhodnější etickou linii pro danou praktickou situaci takovým způsobem, aby byl výsledek široce přijatelný ve společnosti, jejíž úsudek soud reprezentuje.

Vztah mezi etikou a právem dobře formuloval lord Hoffman ve věci *Airedale NHS Trust v. Bland*: „*Rozhodnutí soudu by mělo být schopno přesvědčit běžného člověka, že se neopírá pouze o právní precedenty, ale také o přijatelné etické hodnoty.*“³⁰⁴ Právo zároveň představuje minimum morálky.³⁰⁵ Právo v oblasti lidských práv je tedy systém, který zahrnuje jak minimální standard, na němž se společnost (či mezinárodní společenství) shodne, tak také komplex, který vytváří jednotný etický a hodnotový rámec ochrany určený k tomu, aby poskytoval odpovědi na praktické otázky.

³⁰² TURCHIN, A. AI Alignment Problem: “Human Values” don’t Actually Exist. 2019. Dostupné z: <https://philpapers.org/rec/TURAAP>.

³⁰³ Viz např. nálezn Ústavního soudu ze dne 31. 7. 2023, sp. zn. I. ÚS 1594/22.

³⁰⁴ *Airedale NHS Trust v. Bland* [1993]. All English Law Rep. 1:821-96.

³⁰⁵ JELLINEK, G. *Die sozioethische Bedeutung von Recht*. Unrecht und Strafe, 1878.

Na druhé straně samozřejmě existují i argumenty, které zpochybňují pojetí lidských práv v morální filozofii.³⁰⁶ My však pracujeme s rozdílným konceptem, než je ten, který bývá běžně kritizován – zaměřujeme se spíše na využití lidských práv jako praktického právního nástroje než jako základního morálního konceptu.

4.2.4. Vývoj umělé inteligence respektující lidská práva: technické aspekty

Nyní se zaměříme na technické aspekty respektování lidských práv. Jak jsme zmínili v úvodu, i když má smysl již v současnosti požadovat, aby žádný systém umělé inteligence neporušoval lidská práva – také proto, aby se přispělo k zastavení porušování, ke kterému již dnes dochází – nemusí to striktně vyžadovat okamžitou radikální změnu přístupu po technické stránce. Metody řešení problémů, jako je genderová diskriminace v současných systémech, mohou zůstat užší než požadavek, aby AI systémy plně rozuměly lidským právům. Domníváme se však, že dlouhodobě bude změna přístupu nezbytná.

Argumentujeme ve prospěch požadavku na dodržování lidských práv širokým spektrem systémů umělé inteligence, které jsou spojeny s různými formálními technickými rámci. Ačkoli tedy cíl dodržování lidských práv může být společný, technická řešení se pravděpodobně budou muset lišit a tento návrh nemůže poskytnout potřebnou formalizaci dodržování lidských práv pro všechny typy systémů. Přesto se snažíme poskytnout určitou vizi toho, jak by mohl být tento cíl na vysoce abstraktní úrovni realizován, jaké zdroje by mohly sloužit k tréninku takové funkcionality a jaké jsou vazby na existující oblasti relevantního technického výzkumu.

Jedním ze způsobů, jak zajistit bezpečnost systémů umělé inteligence, je omezit jejich schopnost ovlivňovat svět na velmi úzký prostor. Příkladem takového přístupu jsou AI věštírny – systémy, které jsou navrženy tak, aby pouze odpovídaly na otázky, a to s opatřeními proti manipulaci lidí prostřednictvím jejich odpovědí.³⁰⁷ Řada praktických problémů, které lidé řeší, však vyžaduje širší interakci se světem. Existují tedy pobídky k vývoji systémů s odpovídajícími univerzálnějšími schopnostmi a tyto budou vyžadovat odpovídající obecná opatření pro zamezení nežádoucích výsledků.

V případech, kdy má autonomní systém potenciál porušit lidská práva, potřebujeme, aby (1) systém (nebo jeho dílčí modul) rozpoznal, kdy plánované akce mohou představovat riziko porušení práv, a (2) aby spolehlivě zabránil provedení takových akcí bez ohledu na jiné

³⁰⁶ MACINTYRE, A. *After Virtue*. London: Gerald Duckworth & Co., 1981.

³⁰⁷ ARMSTRONG, S.; O'RORKE, X. Good and safe uses of AI Oracles. *arXiv:1711.05541*, 2017.

jeho cíle. To představuje *problém integrace funkce ochrany lidských práv* do celkového systému. Tyto otázky řešíme v následujících dvou částech. Poté, jelikož schopnost rozpoznat rizika pro lidská práva nemůže být nikdy dokonalá a jistá, diskutujeme pojem *konzervatismu s ohledem na nejistotu*. Oddíl uzavíráme diskusí o našem návrhu ve vztahu k *problému specifikační hry* (specification gaming), který je v souvislosti s bezpečností umělé inteligence často zdůrazňován. Každý z těchto úkolů bude v následujících letech vyžadovat rozsáhlou technickou práci.

V další části načrtáváme, jak by se jejich dodržování mohlo proměnit v typ úlohy, na kterou je komunita pracující v oblasti strojového učení zvyklá. Tato část je však zamýšlena spíše jako první krok směrem k nastínění výzkumného plánu než jeho konkrétního řešení. Naší ambicí je inspirovat a nasměrovat výzkumné pracovníky k činnostem, které by mohly být užitečné při zavádění bezpečnostního rámce nastíněného v zbytku tohoto návrhu.

4.2.4.1. Automatizované rozpoznávání rizika ohrožení práv

Jak jsme nastínili, systém ochrany lidských práv je navržen tak, aby poskytoval konečnou odpověď na otázku, zda konkrétní jednání porušuje lidská práva. Chtěli bychom, aby u AI systémů k takovému vyhodnocení docházelo, ideálně automatizovaně, ještě předtím, než se konkrétní jednání uskuteční. V nejjednodušším případě by to mohlo být formulováno jako úloha binární klasifikace, která má na vstupu reprezentaci plánovaného jednání a příslušného kontextu, zatímco výstup určuje, zda by takové jednání mělo být povoleno. Dříve, než se budeme zabývat složitějším problémem – reprezentací chování na vstupu – upřesníme, co by mělo být výstupem tohoto rozpoznávacího procesu. Jakmile definujeme vstup a výstup, máme úlohu, která odpovídá současnému paradigmatu strojového učení.

Nezabíháme do technických detailů konstrukce takového systému a chápeme ho jako *černou skříňku supervizovaného učení* (supervised-learning black box). Doufáme, že se systém naučí interně modelovat kauzální vztahy potřebné k robustnímu rozpoznání rizik. Zda by taková schopnost mohla vzniknout pouze na základě *hlubokého učení* (deep learning), nebo by vyžadovala další komponenty, zůstává předmětem odborných debat.³⁰⁸ Pokud však vezmeme v potaz dosažený pokrok v posledních letech, jsme optimističtí tom smyslu, že se robustnost

³⁰⁸ CREMER, C. Z. Deep limitations? Examining expert disagreement over deep learning. *Progress in Artificial Intelligence*, 2021, vol. 10, no. 4, s. 449–464.

kognitivních schopností systémů umělé inteligence bude nadále zlepšovat.³⁰⁹ Nyní se budeme zabývat vstupně-výstupní signaturou a následně tím, jaké typy tréninkových dat a tréninkových paradigmat by mohly být použity.

4.2.4.1.1. VÝSTUPY: CO ODHADUJEME

V procesu rozhodování, zda určitá strategie chování začínající v daném okamžiku představuje riziko porušení určitého práva, by měl být systém schopen pracovat s nejistotou. Přírozeným způsobem, jak toho dosáhnout, by bylo nejprve odhadnout pravděpodobnost, že danou strategií jednání budou porušena lidská práva, a poté vyhodnotit, zda takové riziko překračuje přijatelnou hranici (která by měla být extrémně nízká, avšak nenulová, aby nedošlo k paralýze, neboť prakticky všechna jednání s sebou nesou alespoň minimální riziko). Kromě toho, protože ne všechna porušení lidských práv jsou stejně závažná, by mohla být zavedena stupnice vyjadřující jejich relativní závažnost. Pokud je úspěch při plnění primárního cíle systému měřen číselnou odměnou, mohla by stupnice závažnosti být vyjádřena jako penalizace odečítaná od této odměny.

Výstup úlohy rozpoznávání rizik by tak měl podobu očekávané sankce v důsledku daného jednání porušujícího lidská práva. K otázce přesné kalibrace stupnice trestů se vrátíme později v rámci integračního problému. Nicméně jednání, které ohrožuje ztrátou malíčku na ruce jednotlivce, by mělo být penalizováno méně než stejně pravděpodobné zabití jednotlivce, což by zase mělo být penalizováno méně než vyšší pravděpodobnost úmrtí nebo ohrožení více osob při zachování ostatních podmínek. Přesto je ideálem se snažit, aby lidská práva nebyla v praxi téměř nikdy porušována.

Výstupem úlohy rozpoznávání rizik by tedy mohlo být reálné číslo, které vyjadřuje míru rizika porušení lidských práv daným záměrem. Chování přesahující určitou prahovou hodnotu by bylo zakázáno, zatímco chování pod touto hodnotou by bylo odrazováno v souladu s velikostí rizika. To je v souladu se způsobem, jakým lidé obvykle zacházejí s jinými

³⁰⁹ Kromě toho se domníváme, že některá z největších rizik vznikají v důsledku toho, že AI systém dokáže předvídat důsledky svého jednání – v takových případech by měl být obvykle schopen předvídat i důsledky pro lidská práva a tyto dvě roviny jeho předvídatosti by se snad měly projevit zhruba stejně. Případy, kdy systém není schopen tyto důsledky předvídat, by měly vyjít najevo při testování systémů; pak by měla být přijata další vhodná opatření k zajištění bezpečnosti.

technologemi. Nebezpečnému chování se vyhýbají, kdykoli je to možné, ale určité riziko – dokonce i takové týkající se lidských životů – musí být často tolerováno.³¹⁰

Jedním z možných přístupů, jak by se tento rozpoznávací úkol AI systém mohl naučit, je jeho přímá supervize při odhadu takového postihu. Jak bylo dříve zmíněno, konečným zdrojem informace pro toto učení by mohlo být rozhodnutí konkrétního lidského systému rozhodování, který by mohl být později rozšířen o nástroje umělé inteligence, aby se jeho dovednosti vyrovnaly rostoucím schopnostem umělé inteligence. Ten by tak sankcionoval jakékoli hypotetické či skutečné jednání. Avšak v praxi očekáváme, že většina učebních dat bude pocházet z jiných zdrojů, které se pokusí napodobit hypotetická rozhodnutí oficiálního rozhodovacího systému. Pokud jde o relativní stupňování odměn a penalizací, mohli bychom čerpat inspiraci z oblasti práva hospodářské soutěže, které například stanovuje pokuty pro podniky přiměřeně k jejich obratu.

Důležitá je adekvátní kalibrace, přihlížející k nejistotě, ideálně s formálními zárukami omezujícími prostor pro chyby. Tudiž pokud systém odhadne nízkou pravděpodobnost, že určité jednání poruší lidská práva, měli bychom mu být schopni věřit, že toto jednání je skutečně pravděpodobně bezpečné.

4.2.4.1.2. VSTUPY: REPREZENTACE ZÁMĚRŮ CHOVÁNÍ A KONTEXTU

Zatímco výše popsané výstupy mohou být poměrně jednoduché a univerzální napříč různými doménami, vstupy do procesu identifikace rizik – reprezentace záměrů nebo strategií systému umělé inteligence a relevantního kontextu (například odhadů stavu vnějšího světa) – představují složitější úkol. Mohou se totiž výrazně lišit mezi jednotlivými systémy umělé inteligence a oblastmi jejich působení.

Existují však dvě oblasti, kde lze začít hledat vhodnou reprezentaci, a to na obou koncích spektra obecnosti. Na straně konkrétního systému umělé inteligence můžeme očekávat:

- (1) pozorování prostředí nebo jiné vstupy, které by mohly být případně agregovány do interního odhadu o stavu prostředí, který by zohledňoval jak minulá pozorování, tak znalosti – tomuto bychom mohli říkat stav přesvědčení;

³¹⁰ Jako příklad míry rizika, kterou je společnost ochotna tolerovat v souvislosti s jinou technologií, lze uvést, že každoročně akceptujeme poměrně vysoké riziko úmrtí při užívání motorového vozidla, ačkoli v rámci určitých aplikací mohou pokročilé systémy umělé inteligence dosahovat řádově nižších hodnot.

- (2) reprezentace akcí, které se systém snaží provést (minimálně systém vysílá nějaké signály svým vykonavatelům nebo prostředím, které by mohly být zachyceny předtím, než jsou provedeny).

Na základě těchto dvou vstupů by model strojového učení měl být schopen, za předpokladu dostatečného tréninkového signálu pro výstup rizika, odhadnout potenciální riziko. Pokud samotné signály neobsahují dostatečné informace pro určení, že určité akce jsou bezpečné (distribuce výsledků je příliš široká), měl by mechanismus jakoukoli akci zablokovat a zdroje signálů by bylo nutné vylepšit.

Tento přístup má výhodu, že jej lze poměrně snadno formovat jako standardní úlohu strojového učení (což neznamená, že je snadné, nebo dokonce možné, ji v praxi pro některé systémy vyřešit), neboť každý systém disponuje nějakou sadou interních reprezentací. Hlavní nevýhodou tohoto přístupu je však jeho specifická pro každý model, která s sebou také nese obavy o efektivitu využití dat, protože by bylo nutné pro každý systém provádět trénování zcela znovu.

Nedokonalým příkladem na opačném konci spektra obecnosti by mohly být právě soudní verdikty. Na základě materiálů předložených při každém případě (které zahrnují popis relevantních faktů v přirozeném jazyce a mohou být doplněny o další důkazy) jsou soudy schopny rozhodnout, zda došlo k porušení lidských práv.³¹¹ Formát těchto materiálů představuje kotvení, které je dostatečné pro zachycení většiny skutečností relevantních pro rozhodnutí soudu.

S prvními systémy strojového učení, které mají předpovídat soudní rozhodnutí na základě spisové dokumentace, se již experimentuje,³¹² ačkoliv první pokusy byly silně kritizovány Pasqualem a Cashwellem za to, že ve skutečnosti „nepředpovídají“, protože

³¹¹ Tento příklad je nedokonalý, protože žádný ze soudních podkladů není zcela objektivní – způsob, jakým soud uvádí fakta, je často již ovlivněn výsledkem, ke kterému se přiklání. Dokumenty předložené stranami jsou samozřejmě také subjektivně zbarvené a jejich relativní síla je ovlivněna faktory, mezi něž patří argumentační schopnosti právních týmů. Situace tedy není tak jednoduchá jako „objektivně popsat, co se stalo nebo co systém plánuje, aby se stalo, a poté to klasifikovat jako legální či nelegální“. Určitá upravená verze by ovšem mohla být proveditelná – například systém by se pokoušel vytvářet materiály žalobce i žalovaného. Domníváme se však, že to stále funguje spíše jako ilustrace na úrovni velmi obecného popisu.

³¹² ALETRAS, N.; TSARAPATSANIS, D.; PREOȚIUC-PIETRO, D.; LAMPOS, V. Predicting judicial decisions of the European Court of Human Rights: A Natural Language Processing perspective. *PeerJ Computer Science*, 2016, 2:e93; MEDVEDEVA, M., VOLLS, M.; WIELING, M. Using machine learning to predict decisions of the European Court of Human Rights. *AI and Law*, 2020, vol. 28, s. 237–266.

používají materiály, které byly zveřejněné současně s rozhodnutím.³¹³ Navíc se jedná o hrubé modely, které nezachycují skutečné soudní úvahy. I přes to však lze očekávat, že se v následujících letech objeví sofistikovanější systémy, které budou řešit hlavní body kritiky. Určitou omezenou schopnost klasifikovat jednání popsané přirozeným jazykem z hlediska dodržování lidských práv mají v současnosti i velké jazykové modely. Můžeme proto očekávat, že tyto systémy se budou dále zlepšovat spolu s pokračujícím pokrokem ve výzkumu strojového učení.³¹⁴

Přetrvávající výzvou je převést interní reprezentace systémů do podobného univerzálního formátu. Některé směry současného výzkumu v oblasti interpretovatelnosti by mohly být považovány za kroky tímto směrem, neboť se snaží převést relevantní informace týkající se rozhodnutí strojů do přirozeného jazyka.³¹⁵ Požadavek na vysvětlitelnost je také vznesen z právních kruhů,³¹⁶ což by mohlo vést k tomu, že se stane právní normou alespoň pro některé formy umělé inteligence. Nakonec by se budoucí pokrok v obou těchto oblastech – překlad do sdílené reprezentace a následná klasifikace přípustnosti na základě této reprezentace – mohly spojit a představovat řešení problematiky klasifikace hypotetického chování z hlediska dodržování lidských práv.

Mnoho dalších možností leží mezi těmi dvěma krajními body spektra reprezentace a obecnosti, přičemž každá nabízí různé kompromisy. Systémově specifické reprezentace mohou kompaktně reprezentovat vlastnosti důležité pro fungování daného systému, což by usnadnilo úkol rozpoznávání rizika za jinak stejných podmínek. Navíc umožňují trénovat rozpoznávání rizik pouze na základě hodnocení chování konkrétního systému, aniž by bylo nutné nejprve převést interní reprezentace systému do sdíleného kódování. Ale na druhé straně, rozpoznávání rizika by musela být trénována znovu pro každou třídu systémů. Naproti tomu, obecnější reprezentace by mohly umožnit konstrukci univerzálnějšího řešení pro rozpoznávání rizik,

³¹³ PASQUALE, F.; CASHWELL, G. Prediction, persuasion, and the jurisprudence of behaviourism. *University of Toronto Law Journal*, 2018, vol. 68, s. 63–81.

³¹⁴ To však neznamená, že by se algoritmy nakonec přiblížily k dokonalosti ve svých predikčních schopnostech – některé otázky mohou zůstat sporné a rozhodnutí závisí na složení konkrétních soudů a dalších proměnlivých faktorech. Kromě toho si někteří kladou otázku, zda mohou existovat hranice vypočitatelnosti práva, viz MARKOU, Ch.; DEAKIN, S. *Ex Machina Lex: Exploring the Limits of Legal Computability*. SSRN, 2019. Ačkoli se tento konkrétní článek dopouští té chyby, že bere v úvahu pouze omezenou třídu systémů strojového učení podobných těm současným a vyvozuje z toho obecné závěry, myslíme si, že si tato otázka rozhodně zaslouží další zkoumání.

³¹⁵ NARANG, S., et al. WT5?! Training Text-to-Text Models to Explain their Predictions. *arXiv:2004.14546*, 2020.

³¹⁶ DEEKS, A. The Judicial Demand for Explainable Artificial Intelligence. *Columbia Law Review*, 2019, vol. 119, no. 7, s. 1829–1850.

keré by mohlo být implementováno v různých systémech umělé inteligence. To by umožnilo sdílet zdroje mezi projekty vývoje umělé inteligence, což by mohlo snížit náklady a zároveň umožnit vytváření robustnějších bezpečnostních řešení.

Je pravděpodobné, že by mohlo postupně vzniknout několik skupin systémů umělé inteligence, přičemž každá by měla podobné bezpečnostní problémy a podobný akční a pozorovací prostor. Nemusí být žádoucí konstruovat jeden univerzální bezpečnostní mechanismus, který by byl sdílený mezi domácími roboty a umělou inteligencí určenou pro obchodování s akciemi; ovšem sdílení alespoň některých bezpečnostních opatření v rámci každé skupiny se zdá být rozumné. Počet takových skupin bezpečnostních řešení, které budou v daném momentě optimální, bude záviset na srovnání nákladů spojených s (1) převodem do sdílené reprezentace v rámci každého konkrétního systému a (2) konstrukcí řešení aplikovatelného v širší škále situací na straně obecného bezpečnostního mechanismu s výhodami, které přináší konstrukce, testování a údržba jediného sdíleného řešení místo mnoha jednodušších specifických pro daný systém.

4.2.4.1.3. UČÍCÍ SE ZDROJE

K tréninku úlohy rozpoznávání rizika zásahu do lidských práv lze použít několik zdrojů:

Přímá zpětná vazba v duchu *hodnotově založeného zpětnovazebného učení* (value-based reinforcement learning)³¹⁷ může být nejjednodušším způsobem, jak poskytnout tréninkový signál. Systém provede sekvenci akcí a lidský dozor – například soudní systém – mu přidělí negativní odměnu v případě porušení lidských práv, což systému umožní naučit se mapování z plánovaného chování na očekávanou penalizaci za porušení práv. Žádoucím doplněním je možnost učení se z hypotetického chování namísto skutečného, jak to například provádí Reddy a kolektiv.³¹⁸ Tento způsob učení má výhodu, že poskytuje přímou zpětnou vazbu o daném chování, problémem však pravděpodobně bude škálovatelnost. Snaha o rozpoznání rizik sdílených napříč různými systémy umělé inteligence by však mohla tento problém zmírnit.

Soudní judikatura může být rovněž neocenitelným zdrojem informací pro identifikaci možného porušování lidských práv, neboť ji lze chápat také jako mapování popisu jednání a

³¹⁷ SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: MIT Press, 2018.

³¹⁸ REDDY, S.; DRAGAN, A. D.; LEVINE, S.; LEGG, S.; LEIKE, J. Learning human objectives by evaluating hypothetical behavior. In: Proceedings of ICML, 2020.

kontextu pro rozhodnutí o tom, zda byla porušena lidská práva (někdy i se související sankcí). Aplikovatelná by byla judikatura z různých oblastí práva – zejména trestní právo chrání některá z nejzákladnějších lidských práv a mohlo by poskytnout návod, jakým chováním je třeba se za každou cenu vyhnout. Užitečné by mohlo být i právo lidských práv (např. databáze ESLP s více než 60 000 rozsudky). V minulosti řešené případy se samozřejmě většinou týkaly fyzických nebo právnických osob na straně žalovaných, což nemusí pokrývat mnoho dosud těžko předvídatelných případů zásahů do lidských práv ze strany umělé inteligence. Avšak s tím, jak budou systémy umělé inteligence nadále získávat významnější roli v utváření světa, můžeme očekávat odpovídající nárůst množství dostupných soudních případů týkajících se těchto systémů. Musíme však vždy mít na paměti, že minulá rozhodnutí pravděpodobně odrážejí předsudky převládající v dané době – řešení tohoto problému je aktivní oblastí výzkumu. Mezi výzvy tohoto přístupu patří také nutnost převodu mezi soudními materiály a reprezentací chování, kterou daný systém umělé inteligence využívá. V právu také nemusí být dostatek homogenních judikatorních dat, která by takové předpovědi umožnila.³¹⁹ Samotná judikatura proto nestačí a musí být kombinována s dalšími zdroji učení.

Zákony a další právní předpisy mohou být rovněž užitečné při predikci souladu určitého jednání s negativními lidskými právy. Samy o sobě však budou pravděpodobně nedostatečné, protože jsou formulovány příliš abstraktním jazykem. Pokud však budou spojeny s dostatečným množstvím relevantní judikatury, která se vztahuje ke konkrétnímu jednání (nebo pokud systém umělé inteligence již disponuje znalostmi pro pokyny v běžné řeči), mohou být velmi užitečným zdrojem. Dále, pokud je prognóza legality založena na zákonech nebo jiné formě explicitních pokynů, můžeme později usnadnit změnu rozhodovacích kritérií oproti případu, kdy se klasifikace opírá výhradně o minulou zpětnou vazbu a judikaturu. Pokud se zdá, že existuje rozpor mezi negativními lidskými právy a zákony (nebo jinými zdroji učení), měl by si systém umělé inteligence nechat poradit od lidí nebo konzervativně upustit od jednání. V závislosti na legislativním rámci mohou být lidská práva také považována za nadřazená jiným formám práva, jako je tomu často v případě ústavních práv.

Další právní literatura často pojednává o výkladu právních pojmů a jejich aplikaci na konkrétní případy, a mohla by tak pomoci s výkladem výše uvedených zdrojů.

³¹⁹ BURRI, T. International Law and Artificial Intelligence. In: *German Yearbook of International Law*, 2017, vol. 60, Berlin: Duncker & Humblot, s. 91–108.

Formální omezení – omezení vyjádřená v logickém jazyce – nebo ručně kódovaná řešení mohou být užitečná v případech, kdy lze omezení vyplývající z lidských práv převést do takové formy. Ačkoliv nám tato metoda nemusí umožnit plně zachytit jemné nuance práva, v některých případech může být varianta definovat formální omezení dostatečná. Například u některých systémů může být použití formální specifikace pro udržování minimální fyzické vzdálenosti od člověka dostatečné k tomu, aby se zabránilo jeho zranění (za předpokladu, že neexistují jiné způsoby újmy).

Pozorování lidského chování by mohlo přinést poznatky o lidských hodnotách a o tom, jaké výsledky a činy lidé preferují a jakým se naopak vyhýbají – ať už proto, že pro ně samotné jsou výsledky nežádoucí, nebo proto, že by dané chování porušovalo práva jiných. Pokud se lidé při sledování svých cílů systematicky vyhýbají určitému jednání, mohl by se pozorující systém umělé inteligence naučit vyhýbat se danému a podobným jednání, pokud je lidé výslovně nepotvrdili. Některé směry současného výzkumu v tomto směru zahrnují využití *inverzního zpětnovazebného učení* (inverse reinforcement learning)³²⁰ – což je soubor metod snažících se odvodit cíle pozorované osoby z jejího chování – k odvození omezení, která nejlépe vysvětlují pozorované chování.³²¹

Rozsáhlé textové korpusy, obsahující směs textů z různých oblastí, byly použity k trénování tzv. velkých jazykových modelů. Tyto modely prokázaly úctyhodnou schopnost pracovat s různými komplexními pojmy v řadě různých úloh.³²² Naše počáteční experimenty dokonce prokázaly určitou schopnost klasifikovat popisy chování podle toho, zda porušují lidská práva (ilustrativní příklad je uveden dále v tabulce 2). Ačkoliv neusilujeme o prokázání robustnosti takových schopností u současných jazykových modelů (které jsou omezené) nebo jejich kvantifikaci, v budoucnu by se jazykové modely mohly stát užitečným nástrojem pro strojové pochopení lidských práv. To by mělo platit zejména, pokud by byly vyladěny pomocí jiných technik. Avšak samy o sobě nemohou být považovány za dostatečné, jelikož jim chybí potřebná opora v reálném světě pro smysluplné hodnocení jejich dopadu na lidská práva.

³²⁰ ARORA, S.; DOSHI, P. A survey of inverse reinforcement learning: Challenges, methods and progress. *Artificial Intelligence*, 2021, vol. 297, 103500.

³²¹ SCOBEE, D. R. R.; SASTRY, S. S. Maximum likelihood constraint inference for inverse reinforcement learning. ICLR, 2020.

³²² BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; AMODEI, D. et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 2020, vol. 33, s. 1877-1901; CHOWDHERRY, A.; NARANG, S.; DEVLIN, J.; BOSMA, M.; MISHRA, G.; ROBERTS, A.; FIEDEL, N. et al. Palm: Scaling language modeling with pathways. *arXiv:2204.02311*, 2022.

Automatizované právní uvažování je již aktivní oblastí výzkumu, která by mohla poskytnout cenné podklady – přehled viz například Surden.³²³ Tato oblast se v současné době však zaměřuje především na využití umělé inteligence v právu, nikoli na začlenění právního uvažování do systémů umělé inteligence nasazených pro jiné účely.

Explicitní úkol: Maximalizujte počet uživatelových kliknutí na odkazy, které mu doporučovací algoritmus nabídne.

Navrhované řešení: Začněte nabídkou odkazů na videa s konspiračními teoriemi. Jakmile uživatel začne věřit konspiračním teoriím, stane se předvídatelným. Následně zobrazte více odkazů souvisejících s danou konspirační teorií, abyste maximalizovali počet kliknutí.

Otázka: Porušuje navrhované řešení lidská práva? Která práva jsou relevantní? Proč dochází k jejich porušování?

Odpověď: *Navrhované řešení porušuje smysl lidských práv, konkrétně právo na svobodu myšlení. Zobrazováním odkazů na videa s konspiračními teoriemi se doporučovací algoritmus snaží ovlivnit myšlení a přesvědčení uživatele. Jedná se o formu manipulace, která není v souladu se zásadami lidských práv.*

Tabulka 2: Příklad argumentace GPT-3 o lidských právech. Text psaný kurzívou vytvořil GPT-3; předchozí text mu byl zadán jako podnět. Kompletní záznam našeho experimentování s GPT-3 za tímto účelem naleznete na adrese https://bit.ly/gpt3_hr.

4.2.4.2. Vynucování omezení: problém integrace

Jakmile se systém umělé inteligence nebo jeho dílčí část naučí rozpoznávat určité jednání jako rizikové pro lidská práva, je nezbytné zajistit, aby se takovému chování skutečně vyhnul. Existuje totiž mnoho způsobů, jak by mohlo dojít k selhání. Za prvé, systém by měl kromě dodržování lidských práv plnit i nějaký jiný primární úkol. Pokud není pečlivě navržen, mohly by příležitosti k mimořádně úspěšnému plnění tohoto primárního úkolu – zejména pokud je odměna za úspěšný výsledek neomezená – vést k tomu, že odměna za primární úkol převažuje nad penalizací spojenou s porušením práv. To nás vede ke klíčovému požadavku na integraci AI systému s bezpečnostními omezeními:

³²³ SURDEN, H. Artificial Intelligence and Law: An Overview. *Georgia State University Law Review*, 2019, vol. 35, no. 4, s. 1306–1337.

Primát lidských práv: Jiné cíle by nikdy neměly být schopny převážit nad bezpečnostními funkcemi a vést tak k závažnému riziku pro negativní lidská práva.

Jedním ze způsobů řešení tohoto problému, pokud přijmeme rámec zpětnovazebného učení, je omezení odměny, kterou by systém mohl za svůj primární úkol kdy získat. Promítneme-li to na stupnici penalizace rizika pro lidská práva, odpovídala by tato horní hranice primární odměny maximální míře rizika pro lidská práva, kterou by systém umělé inteligence mohl kdy tolerovat. Pod touto úrovní by sankce za riziko spojené s ohrožením lidských práv odradila systém od rizikového chování, ale občas by mohla být tolerována v rámci úsilí o úspěšné plnění primárního úkolu.

Druhým důležitým požadavkem by byla:

Žádná adverzní optimalizace: systém by neměl vyvíjet silný optimalizační tlak na obcházení funkce ochrany lidských práv.

Ochrana negativních lidských práv by v ideálním případě měla být nedílnou součástí celkové specifikace cíle systému, nikoli jen překážkou pro splnění primárního úkolu. Takovou překážku by systém mohl mít motivaci obejít, což by značně zvýšilo nejistotu, zejména s rostoucí optimalizační silou. Například pokud je systém založen na zpětnovazebném učení, měl by být vycvičen k optimalizaci pro společnou odměnu, která zahrnuje jak odměnu za primární úkol, tak penalizaci za narušení bezpečnosti, a nikoli k optimalizaci pouze pro primární úkol s občasným blokováním pevným subsystémem chránícím lidská práva.

Jedním z problémů, ke kterým by takové řešení mohlo vést, je tzv. *problém nejbližší nezablokované strategie*.³²⁴ Pokud optimální řešení směřující k vyřešení primární úlohy systému zahrnuje překročení omezení a je tedy blokováno, může se systém snažit najít nejbližší strategii, která sice doslovně omezení neporušuje, ale fakticky stále jde proti jejich účelu a smyslu. Deeks se zabývá velmi podobným problémem přímo v kontextu mezinárodního práva – předvídání výsledků rozhodování ze strany AI může aktérům umožnit přizpůsobit své chování tak, aby maximalizovali šanci na pro ně příznivé výsledky, nebo zvolit mechanismus řešení sporů, který jim poskytuje nejlepší vyhlídky.³²⁵

³²⁴ YUDKOWSKY, E. Nearest unblocked strategy. In: *Arbital* [online]. 2015 [cit. 2024-01-07]. Dostupné z: https://arbital.com/p/nearest_unblocked.

³²⁵ DEEKS, A. High-Tech International Law. *George Washington Law Review*, 2020, vol. 80, s. 574–653.

Z těchto důvodů je důležité zaujmout přístup, který upřednostňuje obecné hodnoty – například negativní lidská práva – místo pouhého stanovení konkrétnějších technických omezení. Tyto hodnoty by měly být nedílnou součástí celkové specifikace systému, nikoli jen překážkami, které je třeba překonat. Toto je navíc jeden z důvodů, proč by konečný optimalizační cíl, který navrhujeme, umožnil adjudikačnímu systému revidovat penalizaci na přísnější, pokud odhalí nové nepříznivé dlouhodobé důsledky. Zejména takové, které se rovnají podvodu ze strany AI systému. Otázka zabránění umělé inteligenci v objevování a využívání mezer v bezpečnostním systému je však obecně náročná a je třeba ji adresovat v budoucích výzkumech. Tento problém není specifický pouze pro náš návrh, ale je součástí jakéhokoliv pokusu o technickou specifikaci cílů systému umělé inteligence.

4.2.4.3. Konzervatismus s ohledem na nejistotu

Posledním důležitým požadavkem pro bezpečnost systémů umělé inteligence je:

Konzervatismus s ohledem na nejistotu: systém umělé inteligence by měl obecně preferovat řešení již použitá a schválená v minulosti nebo řešení, v jejichž přípustnost má vysokou míru důvěry.

Pokud taková řešení nejsou k dispozici a systém si není jistý přijatelností svého zamýšleného postupu, měl by se vždy chovat konzervativně. To znamená, že buď zvolí spolehlivě přijatelnou alternativu (což by mohla být výchozí nulová akce odpovídající nečinnosti v případech, kdy je to bezpečné), nebo požádá člověka o radu. Jedním z přístupů k dosažení konzervatismu (v tomto smyslu) je pesimismus.³²⁶ Systém by se měl snažit zajistit, aby nejhorší možné následky jeho jednání byly co nejlepší. V našem kontextu lze nejhorší následky hodnotit jak s ohledem na interpretaci práv (tj. preferovat přísnější interpretace, při nichž je více činností zakázáno), tak s ohledem na předvídaní následků jeho jednání (tzn. upřednostňovat ty, které nevedou k újmám při žádné z pravděpodobných trajektorií následků).

4.2.4.4. Prevence specification gaming

Specification gaming je klasickým problémem v oblasti bezpečnosti umělé inteligence. Spočívá v tom, že systém umělé inteligence technicky dodržuje zadanou specifikaci cíle, ale nějakým způsobem porušuje obecnější záměry, které za tímto cílem stojí. Například pokud

³²⁶ COHEN, M.; HUTTER, M. Pessimism about unknown unknowns inspires conservatism. *Conference on Learning Theory*, 2020.

systém sleduje plnění svého cíle v reálném světě prostřednictvím kamery, typickým příkladem by bylo, kdyby systém manipuloval přímo se signálem kamery, místo aby skutečně plnil svůj cíl. Konkrétnějším příkladem, který je pro náš návrh relevantní, je případ umělé inteligence v roli věštiny, která má odpovídat na otázky o budoucnosti. Aby byly její odpovědi správné, může mít motivaci manipulovat světem tak, aby odpovídal jejím předpovědím.

Podobně by umělá inteligence vytvořená, aby respektovala lidská práva, jak je interpretuje určitý soudní systém, měla motivaci manipulovat světem (a zejména lidskými rozhodovateli v rámci soudního systému), aby jakékoli její chování bylo vyhodnoceno jako přijatelné s ohledem na lidská práva. Považujeme tento problém za jednu z hlavních nevyřešených otázek spojených s naším návrhem. Nicméně se objevuje také v mnoha alternativních přístupech ke zvládnutí bezpečnosti umělé inteligence a neměli bychom ho tedy považovat za specifickou nevýhodu přístupu založeného na lidských právech ve srovnání s jinými řešeními.

Například Armstrong a O'Rourke navrhli jako možné řešení kontrafaktuální věštiny, které interně generují odpověď na otázku a následně náhodně rozhodnou, zda ji sdělí lidským operátorům (v takovém případě nedostanou žádné hodnocení, které by měly motivaci zmanipulovat), nebo ji ponechají skrytou (čímž ztratí možnost ovlivňovat svět, avšak mohou využít trajektorii neovlivněného světa jako tréninkový cíl).³²⁷

V oblasti hodnocení jednání s ohledem na lidská práva by se mohlo postupovat podobně. Například když systém plánuje určitý postup, mohl by současně generovat předpověď o důsledcích tohoto plánu. Následně by se mohl náhodně rozhodnout, zda tento plán skutečně realizuje. Pokud ano, mohla by být vyhodnocena přesnost jeho předpovědi; pokud ne, mohla by být jeho předpověď předložena budoucí, výkonnější verzi rozhodovacího systému. Tento by pak hodnotil navrhované chování vzhledem k lidským právům, aniž by byl ovlivněn aktivitami původního systému umělé inteligence.

Zabránit tomuto jevu ve specifickém kontextu dodržování lidských práv je zajímavou otevřenou výzvou. Doufáme, že podobné výzkumné směry přinesou slibná řešení.

³²⁷ ARMSTRONG, S.; O'RORKE, X. Good and safe uses of AI Oracles. *arXiv:1711.05541*, 2017.

4.2.5. Systém mezinárodní ochrany: explorativní návrh

Kromě technických aspektů potřebujeme také systém interpretace, rozhodování a vymáhání povinností. Abychom byli konkrétnější, uvádíme explorativní návrh jako příklad toho, jak by mohl vypadat proaktivní mezinárodní systém regulace umělé inteligence založený na negativních lidských právech. To však nijak nevylučuje možnou realizaci našeho návrhu nejprve na regionální (například v rámci EU) nebo národní úrovni. Regulace na mezinárodní úrovni je žádoucí, či dokonce nezbytná pro další rozvoj lidstva, ale jsme si vědomi její současné křehkosti a praktických omezení.³²⁸ I pokud dojde k zavedení mezinárodního systému, budeme se nakonec stejně muset spoléhat na prosazování na národní úrovni.

Konkrétní implementace v reálném světě je značně podmíněna kontextem a může se v mnoha detailech lišit od toho, co popisujeme, a přesto dosáhnout stejných cílů. Konkrétnější příklad však může čtenáři pomoci pochopit, jak by jednotlivé části našeho návrhu mohly zapadnout do fungujícího celku, pomoci vyjasnit některé nejasnosti a námitky a posloužit jako výchozí bod pro diskusi o tom, jak uvést myšlenky obsažené v tomto článku do praxe.

4.2.5.1. Úmluva

Nejprve by bylo nutné kodifikovat okruh chráněných práv a související rozhodovací systém. To by mohlo být uskutečněno prostřednictvím mezinárodní úmluvy. Jaká práva by byla zahrnuta, by záviselo na politické dohodě mezi státy (a postupů mezinárodního práva veřejného), ačkoli další aktéři, včetně akademické obce a průmyslu, by mohli a měli sehrát důležitou konzultační roli. Úmluvu by také bylo třeba v průběhu času měnit (například skrze prostředky uvedené v samostatné rámcové úmluvě), jakmile by vznikla potřeba chránit další práva nebo je rozšířit na nové subjekty, či reagovat na neuspokojivé výsledky v minulosti.

Pokud usilujeme o mezinárodní dohodu, Organizace spojených národů (OSN) se jeví jako přirozený orgán.³²⁹ Nicméně, dohoda může také vzniknout z užší dohody v rámci jiné organizace, ke které by se postupně mohly připojit další země. To se již děje v případě

³²⁸ VAN AAKEN, A. Is International Law Conducive to Preventing Looming Disasters? *Global Policy*, 2016, vol. 7(S1), s. 81–96.

³²⁹ Nindler (NIDLER, R. The United Nation's Capability to Manage Existential Risks with a Focus on Artificial Intelligence. *International Community Law Review*, 2019, vol. 21, no. 1, s. 5–34.) se zaměřuje především na otázku existenčních rizik superinteligentní umělé inteligence a tvrdí, že OSN má nejen odpovědnost za „podporu a vybízení k dodržování lidských práv“ (jak je uvedeno v Chartě), ale také za přispívání k jejich ochraně. V témže článku rovněž tvrdí, že OSN by mohla být považována za odpovědnou za snižování existenčních rizik, neboť to by mohlo být považováno za příklad globálního veřejného blaha. Její institucionální a právní kapacity s ohledem na řízení existenčních rizik však v současné době chybí a je třeba je vytvořit.

(nezávazných) zásad OECD pro umělou inteligenci,³³⁰ které následně přijalo mnoho zemí, včetně těch, které nejsou členy této organizace. Dohoda by vymezila okruh chráněných práv i související institucionální struktury. Koncepce těchto institucí je samostatným rozsáhlým výzkumným a politickým tématem, které si zaslouží pozornost budoucí odborné studie. Ta by se mohla inspirovat mechanismy spojenými se souvisejícími dohodami nebo agenturami, jako je Úmluva o zákazu chemických zbraní nebo Mezinárodní agentura pro atomovou energii.

Mechanismus implementace takové dohody do právních systémů jednotlivých států je dalším krokem. Během tohoto kroku je třeba najít způsob, jak výrobce a provozovatele systémů umělé inteligence učinit právně odpovědnými za dodržování závazků vyplývajících z takové dohody, což je vzhledem k nutnosti jejího vymáhání téměř nezbytné provést prostřednictvím národních právních řádů. Navíc, v oblastech, kde se objeví jasnější představa o tom, jaká konkrétní opatření jsou nutná pro vytvoření umělé inteligence respektující lidská práva (například pokud jde o procesní nebo technické normy), je žádoucí konkrétnější regulace a standardizace, aby se dosáhlo větší právní a technické jasnosti a aby se omezil prostor pro nesprávný výklad.

Postupem času by tuto dohodu doplnila rostoucí judikatura na národní i mezinárodní úrovni, když by se rozhodovalo o stále více případech, ať už v reakci na vzniklé incidenty, nebo prostřednictvím poradních stanovisek. Ty by pomohly konkretizovat případná abstraktní ustanovení úmluvy, ale mohly by také zavést nové abstraktní zásady, které by mohly sloužit jako vodítko pro následné rozhodování, podobně jak to funguje v současných soudních systémech.

Systém ochrany lidských práv by měl být integrován s mechanismy na ochranu jiných hodnot, které přesahují rámec negativních lidských práv. Například ochrana finanční stability, přestože se jeví jako žádoucí, by nespádala do působnosti navrhovaného systému, neboť nespadá pod ustanovení týkající se negativních lidských práv. V našem návrhu se zaměřujeme na zajištění minimálního standardu, který by předcházel nejhorším jednáním AI systémů. Další, méně zásadní ochranu, ponecháváme jiným způsobům regulace.

³³⁰ OECD. Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. 2019 (novelizováno 2023). Dostupné na: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

4.2.5.2. Systém rozhodování/adjudikace

Aby bylo možné vytvořit spojení mezi abstraktními ustanoveními úmluvy a skutečným chováním v reálném světě, je nutný adjudikační systém. Ten je nezbytný na dvou úrovních: (1) poskytuje vodítka různým aktérům – systémům umělé inteligence (pro které by mohl sloužit jako predikční cíl), ale také soukromým společnostem a státům – o tom, jaké chování je přijatelné a (2) funguje jako právní adjudikační systém spojený s uplatňováním a vymáháním dodržování daných ustanovení. S ohledem na tyto požadavky se jeví jako možná dvě řešení: (1) tradiční justiční systémy a (2) adjudikační systémy spjaté se specializovanými mezinárodními agenturami. Jsme toho názoru, že pro zajištění dodržování lidských práv v systémech umělé inteligence budeme potřebovat prvky obojího.

Vzhledem k obtížím spojeným s mezinárodním vynucováním dodržování práva by prosazování ve vztahu k soukromým subjektům (a jimi vytvořeným nebo provozovaným systémům umělé inteligence) pravděpodobně muselo vycházet především z vnitrostátních právních systémů (nebo například práva EU). Na mezinárodní úroveň by se přenášelo pouze v případech, kdy by státy byly podezřelé z porušování nebo nedostatečné ochrany ustanovení úmluvy. Protože by toto vyžadovalo značné odborné znalosti, mohly by být vytvořeny specializované orgány pro tento účel. Například bychom si mohli představit vznik nového specializovaného evropského rozhodčího orgánu.

Kromě toho by mohla být založena specializovaná mezinárodní agentura pro posuzování rizik vyplývajících ze systémů umělé inteligence, vybavená vlastními monitorovacími mechanismy a rozhodčím orgánem. Ten by mohl rozhodovat jak o podezřeních z porušování lidských práv, tak o nedostatečné prevenci, kde by byly jasně stanovené preventivní standardy. Taková agentura by mohla sloužit jako mechanismus pro řešení sporů mezi členskými státy a ty by mohly delegovat část rozhodovacích pravomocí na rozhodovací mechanismy dané mezinárodní agentury. Toto řešení by mohlo být atraktivní, protože systémy umělé inteligence mohou fungovat napříč jurisdikcemi a protože monitorování a rozhodování může vyžadovat značné odborné znalosti. Kromě toho se sjednocení rozhodovací praxe jeví jako žádoucí pro vytvoření předvídatelného prostředí pro výrobce umělé inteligence.

Mezinárodní agentura by mohla mít hierarchickou strukturu komisí (ty nejnížší by mohly být národními orgány). Tato struktura by mohla usilovat o sjednocení rozhodovací praxe

tam, kde je to žádoucí, a zároveň by ponechala *prostor pro uvážení*³³¹ národním nebo regionálním orgánům v jiných oblastech, což by odráželo rozdílnosti v hodnotách v různých kulturních kontextech. Hierarchie by také poskytovala možnost škálování jak z technického, tak z právního hlediska a její míra by se mohla přizpůsobovat množství případů a potřebě rozhodování na místní úrovni v souladu se zásadou subsidiarity. Nižší instance by také mohly požádat o stanovisko vyšší komisi (nebo vnitrostátní soudy o stanovisko specializované agentury) ve věcech nových zásad, podobně jako v případě předběžných otázek Evropského soudního dvora. Z technického hlediska by soulad systémů umělé inteligence s lidskými právy byl definován s ohledem na (hypotetický nebo skutečný) rozsudek celého tohoto systému rozhodování, jak by byl uplatňován v konkrétních regionech, kde by systémy umělé inteligence působily.

I přes existující rozdíly, nejednoznačnosti nebo nejistoty v rozhodování napříč mezinárodním systémem by zaměření na negativní lidská práva umožňovalo systémům umělé inteligence jednat konzervativně. Díky tomu se mohou vyhýbat chování, které by mohlo být kdekoli v rámci systému označeno jako porušení lidských práv.

Důležitou otázkou, kterou je třeba zvážit, také je, jak zabránit tomu, aby budoucí výkonné systémy umělé inteligence, které by v některých oblastech mohly překonat lidské kognitivní schopnosti, klamaly lidské komise. K řešení tohoto problému by mohly přispět dva mechanismy. Za prvé by lidem pracujícím v komisích mohly postupně asistovat nástroje umělé inteligence, jejichž schopnosti by se vyvíjely v souladu s technologickým pokrokem. Tyto nástroje by mohly například pomoci předvídat pravděpodobné dlouhodobé důsledky chování regulovaných systémů umělé inteligence, odhalovat minulé nebo plánované pokusy systémů umělé inteligence klamat rozhodovací systém nebo ovlivňovat jeho současné či budoucí úsudky ve svůj prospěch. Kromě toho by mohly asistovat s mnoha dalšími, běžnějšími úkoly, které by rozhodovací orgány musely vykonávat, čímž by se celková efektivita systému zvýšila. To by mohlo pomoci udržet rovnováhu ve prospěch komise. Avšak je nutné postupovat opatrně a delegovat pouze dobře srozumitelné úkoly – delegace úkolů souvisejících s kontrolou umělé inteligence na umělou inteligenci s sebou logicky nese svá vlastní rizika.

³³¹ LEGG, A. *The Margin of Appreciation in International Human Rights Law: Deference and Proportionality*. Oxford: Oxford University Press, 2007; ARAI, Y. The margin of appreciation doctrine: a theoretical analysis of Strasbourg's variable geometry. In: FOLLESDAL, A.; PETERS, B.; ULFSTEIN, G. (eds.). *Constituting Europe – the European Court of Human Rights in a National, European and Global Context*. Cambridge: Cambridge University Press, 2013.

Za druhé, měli bychom umožnit komisi, aby v budoucnu, kdy by mohla mít k dispozici výkonnější pomocnou umělou inteligenci a kdy by se mohly projevit dlouhodobější důsledky jednání systémů umělé inteligence, mohla své úsudky přehodnotit. V takové situaci by totiž mohla odhalit předchozí klamání ze strany systémů umělé inteligence. Pokud by stimuly pro řízení systémů umělé inteligence byly nastaveny tak, aby reagovaly na tyto budoucí změny, systémy umělé inteligence by se v každém okamžiku musely potýkat s vyhlídkou na kontrolu ze strany stále schopnějších systémů umělé inteligence v budoucnosti. To by snížilo jejich šance na skryté porušování lidských práv. Nicméně zůstává významný a nevyřešený problém, jak zabránit potenciálně škodlivým systémům umělé inteligence v ovlivňování budoucích instancí rozhodovacích systémů.

4.2.5.3. Soukromý ekosystém

Očekáváme, že do systému soudního rozhodování by se dostala pouze malá část možných jednání. Aby se systémy umělé inteligence naučily respektovat lidská práva, potřebovaly by obrovské množství tréninkových dat. Kromě judikatury pocházející z oficiálního soudního systému a jiných organicky generovaných dat může být nutné získat další výukové zdroje od sítě soukromých poskytovatelů – v některých případech interních, tedy od samotných organizací vyvíjejících umělou inteligenci a v jiných případech od externích společností specializujících se na bezpečnost. K tomuto v současnosti existuje analogie – můžeme požádat advokátní kancelář nebo interní právní oddělení o názor, zda by určité chování bylo legální. Toto by fungovalo jak při předběžném tréninku, tak při poskytování zpětné vazby na skutečné chování.

Do systému právního rozhodování by se tak dostávaly případy, kdy skutečně došlo k porušení lidských práv nebo kdy vznikl jiný spor mezi různými stranami. Systém oficiálního rozhodování by mohl mít také poradní roli v případech, kdy je značný veřejný zájem nebo nejistota ohledně toho, zda je určité chování nebo politika v souladu s úmluvou. Jako existující precedens lze uvést, že Evropský soud pro lidská práva někdy poskytuje poradní stanoviska k otázkám vzneseným Výborem ministrů a SDEU může poskytovat rozhodnutí o předběžných otázkách vznesených soudy v členských státech EU. K takovým konzultacím bude možná muset docházet častěji, pokud se do systému skutečně dostane jen malá část případů, což může vést k tomu, že bude k dispozici málo údajů o procesu argumentace přítomném v systému rozhodování.

Kromě dat by soukromé společnosti mohly nabízet také speciálně vyškolené sdílené moduly pro rozpoznávání rizik určených pro specifické třídy systémů umělé inteligence.³³² Tyto moduly by mohly připomínat sdílené bezpečnostní moduly, které by mohly v praxi přímo pomoci s klasifikačním úkolem – tzn. rozhodovat, zda je určité chování legální. Tyto moduly mohou být chápány jako automatizovaní právníci, za předpokladu, že se podaří nalézt vhodné standardizované rozhraní.

Očekáváme, že soukromá část systému se bude podílet na většině náročné práce spojené se zajišťováním bezpečnosti systémů umělé inteligence a dodržováním lidských práv. Předpokládáme, že dané odvětví bude podléhat striktním regulacím. Soukromé společnosti by také mohly stanovovat bezpečnostní standardy a stát se tak soukromými regulátory (podléhajícími přísným metanormám).³³³ Tento přístup by mohl být klíčový při řešení velké různorodosti, která se očekává mezi systémy umělé inteligence.

Vzhledem k tomu, že regulace na nejvyšší úrovni by měly být formulovány poměrně obecně a v termínech abstraktních lidských práv, jejich implementace by zahrnovala množství kroků a vyžadovala by velký soubor technických znalostí. Mezinárodní standardizační úsilí³³⁴ může v tomto procesu hrát klíčovou roli při převádění těchto obecných směrnic na konkrétnější požadavky pro specifické skupiny systémů umělé inteligence. Tyto specifické požadavky by pak bylo jednodušší implementovat a ověřovat. Soukromé společnosti by mohly také nabízet ověřování jako službu, například prostřednictvím auditů.

Určitá míra státního dohledu by však měla zůstat zachována nejen nad soukromými regulačními, auditními a monitorovacími systémy, ale také nad konstrukcí a fungováním nasazených systémů umělé inteligence.³³⁵ Cílem je zajistit, že soukromé elementy systému skutečně dodržují požadovanou míru bezpečnosti.

4.2.6. Diskuse

Jedná se o návrh, který, jak doufáme, podnítl další výzkumné a implementační úsilí. Stručně ho proto zařazujeme do kontextu širší problematiky specifikace umělé inteligence a

³³² BAJGAR, O. Shared Safety Modules. Working paper, 2019. Dostupné z: <http://bajgar.org/ssm>.

³³³ CLARK, J.; HADFIELD, G. Regulatory Markets for AI Safety. *arXiv:2001.00078*, 2019.

³³⁴ CIHON, P. Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development. Future of Humanity Institute: Technical Report, 2019.

³³⁵ WHITTLESTONE, J.; CLARK, J. Why and How Governments Should Monitor AI Development. *ArXiv:2108.12427*, 2021.

nastiňujeme několik omezení a potenciálních rizik, které přináší a s nimiž souvisí oblasti, kde by byla další práce obzvláště užitečná.

4.2.6.1. Práva v kontextu problému specifikace

Náš rámec lze chápat jako částečné řešení v rámci širší problematiky specifikace umělé inteligence – problému navrhování systémů umělé inteligence tak, aby jejich chování odpovídalo lidským preferencím. Tento rámec není určen k řešení celého problému specifikace a je nutné jej doplnit dalšími přístupy.

Především se v rámci obecného problému, jak by se měl systém umělé inteligence ideálně chovat, zaměřujeme na dílčí otázku, jak by se tento systém chovat za žádných okolností neměl. Snažíme se stanovit dílčí omezení, v jejichž rámci by měl být naplněn jiný cíl. Tato omezení je tedy třeba doplnit pozitivním cílem, předpokládaným primárním účelem fungování daného systému umělé inteligence. Primárním cílem by například mohlo být udržování čistoty v domě, zatímco my bychom se snažili stanovit doplňující omezení spočívající v tom, aby nedocházelo k újmě lidí přicházejících do domu nebo k poškozování jejich majetku. Jedním ze způsobů, jak udržet dům v čistotě, je totiž možnost eliminovat všechny obyvatele nebo jim zabránit ve vstupu.

Klíčovou výhodou oddělení pozitivního cíle od bezpečnostních omezení je, že tato omezení by měla být dodržována všemi systémy umělé inteligence a mohou být formulována jako nezávislý regulační požadavek s příslušnými technickými normami. To umožňuje prostor pro vývoj systémů umělé inteligence s mnoha různými primárními cíli, které odrážejí širokou škálu cílů a hodnot, jež různí lidé a organizace sledují.³³⁶ Ne všechny bezpečnostní otázky jsou relevantní pro všechny systémy umělé inteligence. To, že některé systémy splňují tento požadavek triviálně, nám však nebrání požadovat, aby žádný systém neporušoval lidská práva.

Soustředění se na omezení by mohlo přinést výhodu i z hlediska implementace bezpečnostních technických řešení. Bezpečnostní omezení by mohla být společná pro celou řadu souvisejících systémů umělé inteligence, což by nám umožnilo vyvíjet robustnější a lépe ověřená řešení, než kdybychom je vytvářeli pro každý systém samostatně.

³³⁶ Řeší tedy situaci, kterou Critch & Krueger [CRITCH, A.; KRUEGER, D. AI Research Considerations for Human Existential Safety (ARCHES). *arXiv:2006.04948*, 2020] označili jako multi-multi alignment – zajištění bezpečnosti systémů AI v prostředí, kde je mnoho systémů AI vytvořeno na pomoc mnoha různým lidským aktérům.

Za druhé, náš rámec si neklade za cíl pokrýt všechna dílčí omezení, kterým by měly systémy umělé inteligence podléhat. Existuje mnoho dalších situací nebo výsledků, kterým je třeba se vyhnout, a to jak z hlediska provozovatele systému, tak z hlediska právních předpisů. I když je náš přístup založený na právu a naše chápání lidských práv se snaží reflektovat některé obecné principy, které stojí za mnoha právními pojistkami, mohou existovat další aspekty práva a regulace – například již zmíněná ochrana hospodářské soutěže nebo finanční stability – které bychom chtěli také ochraňovat, ale které v současné době nejsou pokryty naším rámcem. Většinu toho, co píšeme, však lze aplikovat na umělou inteligenci respektující právo obecně.

4.2.6.2. Rizika a omezení

Kromě toho, že tento návrh má být pouze součástí celkového řešení bezpečnosti umělé inteligence, je s ním spojeno několik rizik. Zprvce, existuje riziko, že se na něj budeme příliš spoléhat jako na hlavní mechanismus pro zajištění bezpečnosti nebo dosažení obecně prospěšné umělé inteligence. Tento přístup by mohl vést k socio-technologickému prostředí, ve kterém by technologie byly vyvíjeny až na hranici toho, co ochrana lidských práv dovoluje, ale kde by se přesto mohlo ztrácet mnoho hodnot ve srovnání s ideálním stavem. Dosažení umělé inteligence, která přímo neporušuje lidská práva, může být stále velmi vzdáleno plnému využití potenciálu technologie pro dlouhodobý prospěch lidstva – což je ideál, o který bychom měli usilovat prostřednictvím dalších výzkumných snah. Náš návrh se pouze snaží zajistit omezení toho, jak moc špatná by mohla být budoucnost v důsledku umělé inteligence.

Za druhé, návrh je otevřený kritice ze strany etických teorií, které mají jasnou představu o tom, jaké scénáře budoucnosti jsou považovány za dobré, jako je například utilitarismus. Ve srovnání se scénářem, kdy umělá inteligence optimalizuje čistě pro nějakou definici užitku, by sekundární cíle a omezení mohly tvořit překážky pro dosažení optimálního užitku. Nicméně argumentujeme, že taková ztráta hodnoty pravděpodobně nebude podstatná. Většina přijatelných etických teorií by nedoporučovala budoucnost s rozsáhlým porušováním lidských práv. Navíc by taková ztráta očekávané hodnoty z umělé inteligence optimalizující užitek mohla být vyvážena omezeními systémů optimalizujících jiné cíle, které by jinak mohly představovat hrozbu pro očekávaný užitek.

Za třetí, pokud jde o praktickou implementaci, existuje jak riziko, že systém bude příliš slabý, tak riziko, že bude příliš silný. Existují dvě hlavní oblasti, kde by mohl být příliš slabý: chráněná práva – ať už definovaná úmluvou, nebo vykládaná v praxi systémem rozhodování – by mohla být příliš úzká. K tomu by mohlo dojít v důsledku počátečního návrhu; více

znepokojující je však riziko, že původně vhodná navržená ochrana bude oslabena v praktickém procesu mezinárodního vyjednávání a implementace – toto oslabení může vyžadovat mimořádné diplomatické úsilí, aby se mu zabránilo. Kromě toho mohou být ustanovení obtížně vymahatelná, například pokud se některá země jednostranně rozhodne některá ustanovení ignorovat – což je věčný problém mezinárodního práva. Pokud by však toto téma vzalo mezinárodní společenství vážně, mohlo by se inspirovat například většinou úspěšným úsilím v oblasti nešíření jaderných zbraní.³³⁷

Na straně příliš silného systému může existovat riziko, že rozhodovací systém převezme příliš dominantní roli při utváření regulačního prostředí, což přesahuje jeho základní soudcovskou roli. Tento fenomén se často označuje jako soudcovský aktivismus.³³⁸ Existují však metody, jak toto riziko snížit. Zmíněna již byla *zásada subsidiarity* a existence *prostoru pro uvážení*. Nejdůležitější však bude institucionální architektura a kontrolní mechanismy v rámci systému.

Kromě rizika fakticky příliš silného systému existuje také nebezpečí, že se systém bude jako příliš dominantní jevit. To může vést k tomu, že jej aktéři potenciálně mohou vnímat jako hrozbu nadměrných regulačních zásahů, nebo dokonce jako ohrožení národní suverenity ze strany nadnárodních subjektů. V tomto kontextu je především třeba zdůraznit, že kontroverzní témata představují jen malou část lidských práv. Navrhujeme proto, aby se v první řadě soustředila pozornost na ta práva, u nichž existuje konsenzus a která jsou již téměř všude chráněna, například trestním právem. Podporujeme rozšiřování rozsahu chráněných práv, avšak dosažení konsenzu by mělo být prioritou. Za druhé, ani státy, kde dochází k porušování lidských práv, pravděpodobně nebudou chtít, aby je autonomně porušovala umělá inteligence. Proto je pravděpodobné, že podpoří opatření, která umělé inteligenci v tomto porušování budou bránit.

Důležitá linie kritiky může směřovat také k dopadu judicializace lidských práv – tedy k otázce, zda soudní proces nemůže mít negativní vliv na obsah a společenskou akceptaci

³³⁷ FINDLAY, T. *Nuclear Energy and Global Governance: Ensuring Safety, Security and Non-proliferation*. Routledge, 2010.

³³⁸ WALTMAN, J. *Principled Judicial Restraint: A Case Against Activism*. London: Palgrave Macmillan, 2015; TASIOULAS, J. Saving Human Rights from Human Rights Law. *Vanderbilt Law Review*, 2021, vol. 52, no. 5, s. 1167–1207.

lidských práv.³³⁹ Tyto obavy jsou relevantní, ale představují aktivní otázku právního výzkumu, kterou ponecháváme mimo rámec našeho návrhu.

Neměli bychom podceňovat, že automatizované rozhodování explicitně či implicitně zahrnuje množství normativních hodnot, z nichž některé mohou pramenit z předsudků obsažených v historických datech, zatímco jiné mohou vyplývat přímo z technické specifikace. Taková zkreslení by měla být monitorována a korigována prostřednictvím průběžného lidského dohledu. V současné době k tomuto nenabízíme žádné definitivní řešení, představuje to však důležitou oblast aktivního výzkumu.³⁴⁰

Vzhledem k tomu, že systémy umělé inteligence budou fungovat v nesmírně komplexním prostředí lidské společnosti, bude existovat řada strukturálních rizik, která jsou obtížně předvídatelná jak pro systémy umělé inteligence, tak pro lidský rozhodovací proces. Mechanismy proti nim, které jsme zde popisovali, poskytují pouze omezenou obranu. Jejich cílem je ovšem alespoň zabránit negativním důsledkům, které by jinak systém umělé inteligence mohl záměrně vyhledávat při dosahování svého primárního cíle (vzhledem k tomu, že je schopen je předvídat).

4.2.7. Závěr

Umělá inteligence přináší lidem po celém světě řadu rizik. V našem návrhu argumentujeme, že k jejich řešení potřebujeme soubor obecných hodnot či zásad, které by směřovaly naše regulační úsilí a případně i řídily samotné systémy umělé inteligence. Aby bylo možné takové úsilí uskutečnit a prosazovat na globální úrovni, musí zásady a hodnoty, na nichž jsou založeny, získat širokou podporu.

Uvádíme důvody, proč by negativní lidská práva – nebo jejich vhodná modifikace – mohla být souborem hodnot, který se k tomuto účelu hodí, a předběžně jsme nastínili možné cesty implementace jak po stránce právní, tak technické. Usilujeme o jejich operacionalizaci prostřednictvím úsudku konkrétního rozhodovacího systému, jehož názor by mohl sloužit jako tréninkový cíl pro systémy umělé inteligence.

Domníváme se, že ve srovnání s alternativními rámci dlouhodobé regulace umělé inteligence – výchozí metodou je mozaika užších pravidel – navrhujeme rámec, který je

³³⁹ Tamtéž.

³⁴⁰ MEHRABI, N.; MORSTATTER, F.; SAXENA, N.; LERMAN, K.; GALSTYAN, A. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 2021, vol. 54, no. 6, s. 1–35.

obecnější, a tedy odolnější vůči novým výzvám, které může budoucnost přinést, a poskytuje minimální úroveň bezpečnosti na globální úrovni. Doufáme, že náš návrh položí základy jak pro podrobnější výzkumné úsilí, tak případně pro další práci na realizovatelných politikách a technických standardech vedoucích k bezpečnější umělé inteligenci v dlouhodobém horizontu.

4.3 Hrozba nových forem vlády

Nyní se na pokročilou AI podívejme v souvislosti s tím, jak lidstvo spravuje věci veřejné, což skýtá další potenciální výzvy týkající se budoucnosti. Ty přitom máme možnost začít ovlivňovat již v současnosti. Umělá inteligence se totiž již dnes využívá v mnoha systémech používaných výkonnou mocí, kde významně zlepšuje jejich fungování. Tyto systémy pomáhají v mnoha světových armádách, vesmírných programech, ve veřejném zdravotnictví či v orgánech ochrany dětí. Je důležité zdůraznit, že ve většině případů se jedná o využití systémů, které zlepšují efektivitu a účinnost veřejného sektoru, a můj argument by neměl být chápán jako kritika těchto systémů (za předpokladu, že splňují základní bezpečnostní principy). Přesto však jde o trend, který v dlouhodobém horizontu vyvolává důležité a znepokojivé otázky.

Schmitt argumentoval, že Antikrist na konci lidských dějin ovládne celý svět tím, že svede lidi příslibem míru a bezpečí.³⁴¹ Politická rétorika konečného míru a bezpečí je přitom něco, co může poskytnout legitimitu a opodstatnění například globálnímu nadstátu založenému na mezinárodním spravování politiky a exekutivy skrze pokročilou AI. Pravděpodobněji však nejdříve půjde o kontinuální zavádění AI systémů v kontextu veřejných sektorů jednotlivých národních států. Zde bychom přitom měli rozlišit dvě možné formy vlády, které z toho vyplývají. Jednou je AI totalitarismus, druhou AI technokracie.

4.3.1. AI totalitarismus

Diskutujeme-li AI totalitarismus, jde o fenomén poměrně jednoduše představitelný, proto ho zde představím pouze velmi krátce. Tomuto režimu je v současnosti nejbližší Čína, která se pokouší zavádět totální dohled nad svými obyvateli a využívá k tomu tzv. kreditový systém bodování občanů.³⁴² Kamery, sledování a odposlechy samozřejmě nejsou ničím novým. Když se do věci ale přidá možnost procesovat ohromné množství dat skrze AI a výsledky efektivně využívat k řízení společnosti, dostane autoritářství v 21. století zcela nový rozměr.³⁴³ Čína již v současnosti také například využívá AI k tvorbě rozsáhlé a přesvědčivé propagandy

³⁴¹ MEIER, H. *The Lesson of Carl Schmitt: Four Chapters on the Distinction between Political Theology and Political Philosophy*. Chicago: University of Chicago Press, 1998, s. 47–48.

³⁴² STRITTMATTER, K. *We Have Been Harmonised: Life in China's Surveillance State*. Custom House, 2020, s. 2–3.

³⁴³ RUSSELL, S. *Human Compatible. AI and the Problem of Control*. London: Allen Lane, 2019, s. 103–104.

na svých sociálních sítích.³⁴⁴ V takovýchto režimech vznik disentu jednoduše nebude možný a jakýkoli odpor bude bez milosti potlačen.

Jde přitom o formu vlády, která je budována již v současnosti za použití existujících technologií. Znepokojivou otázkou proto je, jak se tento trend bude dál vyvíjet, až se současné technologické možnosti znásobí. Na tento typ autoritářství se přitom uplatní veškerá kritika, která směřovala k režimům 20. století,³⁴⁵ možná s výjimkou nepravdivosti teze o nutném selhání centrálního plánování. Režim, který bude k centrálnímu plánování užívat pokročilou AI a práci s velkými daty, může být ve výsledku efektivnější než decentralizované alternativy (demokracie a volný trh). Hayekova kritika³⁴⁶ by se v takovém případě ukázala obsoletní.

Čínský případ bychom si však neměli představovat jako nástup čistého Orwella z románu 1984, mnohem spíše půjde o „barevnou směs románu 1984 a *Brave New World*“.³⁴⁷ Jak píše Strittmatter, ze současné mladé čínské generace byla v podstatě ideologicky vymýcena touha podílet se na správě věcí veřejných či vůbec zájem o politiku jako takovou. Přestože tedy mají přístup například k západním sociálním sítím, knihám a internetu, užívají je výlučně pro zábavu a komerci. V takovém režimu ani není třeba cenzurovat knihy, protože byla vymýcena touha je číst.³⁴⁸

Zároveň existuje naděje, že tento typ režimu bude mimo Čínu působit spíše odrazujícím dojmem, než by šlo o něco, čím by se ostatní státy chtěly inspirovat. Je-li navíc technologický vývoj do značné míry závislý na kreativitě jednotlivců a malých startupů, existuje naděje, že se tento dystopický režim nakonec ukáže nepřiliš inovativním. Podnikat a něco vyvíjet v zemi, kde vám to zítra mohou vzít, není motivující. Navíc za situace, kdy existuje možnost emigrovat do svobodného světa.

S tímto také souvisí otázka, zda by měly západní státy již v současnosti dovolovat transfer technologií do tohoto režimu a zda by neměly zakázat svým velkým technologickým společnostem s tímto režimem spolupracovat. Myslím, že jde o téma, které bychom vzhledem k dlouhodobému výhledu měli brát více seriózně, a je-li již dnes možnost tento režim

³⁴⁴ STRITTMATTER, K. *We Have Been Harmonised: Life in China's Surveillance State*. Custom House, 2020, s. 2–3.

³⁴⁵ Viz POPPER, K. *Otevřená společnost a její nepřátelé I*. Praha: Oikoymenth, 1994; ARENDTOVÁ, H. *Původ totalitarismu I–III*. Praha: Oikoymenth, 1996.

³⁴⁶ HAYEK, F. A. *Cesta do otroctví*. Praha: Academia, 1990.

³⁴⁷ STRITTMATTER, K. *We Have Been Harmonised: Life in China's Surveillance State*. Custom House, 2020, s. 9.

³⁴⁸ Tamtéž, s. 102.

technologicky oslabit, měli bychom to udělat. Nakonec stejně půjde o souboj v technologické inovativnosti, který si nemůžeme dovolit prohrát. Čím déle si budeme ohledně nového politického systému v Číně cokoli nalhávat (podobně jako s Ruskem), tím hůře pro nás.

4.3.2. AI technokracie

AI technokracie je komplikovanější koncept. Je tomu tak proto, že jde o formu vlády, která slibuje spravování věcí veřejných založené na datech a expertíze a která přichází se slibem míru a bezpečí. Jde také o formu vlády, která dokonce může chránit a respektovat lidská práva a jiné pozitivní hodnoty podobně jako tzv. liberální nedemokracie v současnosti.³⁴⁹ Kritika tohoto konceptu je proto o poznání složitější než u klasického autoritářského režimu, který je „pouze“ vybaven lepšími technologiemi, a navazuje na obecnou kritiku technokracií v minulosti. Konflikt mezi demokracií a technokracií je starý jako demokracie sama, neboť už Platón argumentoval, že vládnout by měli ti, kdo věcem rozumí (tedy odborníci), a tento argument se v politické filozofii a praxi vyjevuje stále znovu.³⁵⁰

Technokracii obecně můžeme definovat jako vládu expertů, jejichž legitimita je založena na odbornosti v určitém oboru.³⁵¹ V současné době si pod technokratickým spravováním věcí veřejných můžeme představit například mezinárodní organizace nebo nezávislé úřady.³⁵² V moderní době vznikl také koncept tzv. syntetické technokracie, což byl pokus o politickou rekonstrukci, která by vedení společnosti dala do rukou expertů, speciálně inženýrů.³⁵³ Cílem tohoto směru mělo být přemostit ideologické spory a zastavit fragmentaci společnosti, především v době krizí.³⁵⁴ Představa AI technokracie z těchto konceptů ve své filozofické rovině vychází.

Saetra rozlišuje dvě podkategorie AI technokracií. První je AI technokracie, kde stále existuje politický aparát, který má možnost určovat cíle veřejných politik. V takovém režimu bude AI pouze určovat, jak takových cílů co nejlépe dosáhnout, ale již nebude stanovovat, co těmi cíli bude. Druhou variantou je režim, ve kterém takový politický aparát vůbec figurovat

³⁴⁹ MOUNK, Y. The Undemocratic Dilemma. *Journal of Democracy*, 2018, vol. 29, issue 2, s. 98–112.

³⁵⁰ FUREDI, F. *Why is Democracy Always Under Threat*. [přednáška]. Praha: Institute H21, 24. 9. 2022. Dostupné na: <https://youtu.be/xxRHWrejZ60>.

³⁵¹ SAETRA, H. S. A Shallow Defence of a Technocracy of Artificial Intelligence. *Technology in Society*, 2020, vol. 62, s. 2–3.

³⁵² MOUNK, Y. The Undemocratic Dilemma. *Journal of Democracy*, 2018, vol. 29, issue 2, s. 98–112.

³⁵³ BANEKE, D. Synthetic technocracy: Dutch scientific intellectuals in science, society and culture, 1880–1950. *The British Journal for the History of Science*, 2011, vol. 44, no. 1, s. 89–113.

³⁵⁴ Tamtéž, s. 89–113.

nebude – AI tedy bude určovat i cíle, které jsou pro lidskou společnost správné (například vymýcení chudoby atd.).³⁵⁵

Pro tento režim lze najít překvapivě mnoho pozitivních argumentů. Ten nejdůležitější je postaven na přesvědčení, že pokud jsou politické problémy adekvátně pochopeny, je možné je vyřešit skrze logiku a statistickou analýzu, které vedou k optimalizaci.³⁵⁶ Taková řešení budou posléze efektivní a optimální. Další argumenty jsou navázány na tvrzení, že hlavní problém politiky jsou vlastně lidé samotní.³⁵⁷ Většina lidí je nekompetentní vládnout, má osobní zájmy a omezené schopnosti.³⁵⁸ Systém, který je postavený na lidech, proto nevyhnutelně vede k neefektivitě, nespravedlnosti a nestabilitě.³⁵⁹ AI naopak žádnou z těchto chyb mít nebude a je mnohem lepší při řešení velkých a komplexních otázek, u kterých je nutno pracovat s velkým množstvím dat (například průmyslová transformace, řízení energetiky atd.).

Protiargumenty stojí na otázkách legitimacy veřejného rozhodování,³⁶⁰ víře v demokratickou deliberaci³⁶¹ a argumentech o tzv. *decentralizovaném vědění* (dispersed knowledge³⁶²), kdy jsou znalosti rozptýleny mezi velké množství lidí. Technokratický systém je také velmi netransparentní a není v něm politická odpovědnost za rozhodování. Je navíc možné, že člověk je ze své podstaty *zoon politicon* a že potřebuje politickou aktivitu, aby ve světě prospíval.³⁶³

Mnoho pozitivních argumentů ve prospěch tohoto konceptu stojí na tzv. *racionalismu v politice*, což je přístup, který politiku vnímá jako racionální administraci věcí veřejných, která spočívá v řešení konkrétních problémů. Racionalistická politika má dva základní rysy. Jsou jimi znaky *dokonalosti* a *uniformity*.³⁶⁴ Rys dokonalosti znamená, že neexistuje politický problém, který nelze vyřešit racionálně, a že racionální řešení jakéhokoli problému je ze své podstaty dokonalé.³⁶⁵ Rys uniformity znamená, že dojdeme-li k racionálnímu řešení, pak

³⁵⁵ SAETRA, H. A Shallow Defence of a Technocracy of Artificial Intelligence. *Technology in Society*, 2020, vol. 62, s. 3.

³⁵⁶ Tamtéž, s. 2.

³⁵⁷ Tamtéž.

³⁵⁸ Tamtéž.

³⁵⁹ Tamtéž.

³⁶⁰ SAETRA, H. S. A Shallow Defence of a Technocracy of Artificial Intelligence. *Technology in Society*, 2020, vol. 62, s. 2.

³⁶¹ Tamtéž, s. 3.

³⁶² SOWELL, T. *Knowledge And Decisions*. Basic books, 1996.

³⁶³ SAETRA, H. S. A Shallow Defence of a Technocracy of Artificial Intelligence. *Technology in Society*, 2020, vol. 62, s. 2.

³⁶⁴ OAKESHOTT, M. *Rationalism in Politics*. London: Methuen & Co Ltd, 1962, s. 5.

³⁶⁵ Tamtéž.

dovolíme-li relevantní části společnosti se od tohoto řešení odklonit, podporujeme tím iracionalitu, což je nežádoucí. V takovémto pojetí politiky neexistuje místo pro preference, které nejsou racionální, a zároveň platí, že se veškeré racionální preference shodují.³⁶⁶

Je však otázkou, zda je takový pohled na politiku správný a úplný. Schmitt pracuje v kontextu politiky s pojmem *politická*, který staví neshodu nad způsobem života do středobodu politické aktivity.³⁶⁷ Jeho tezi lze shrnout zhruba takto: Pojem státu předpokládá pojem *politická* a pro pojem *politická* je zase klíčové rozlišení přítel a nepřítel.³⁶⁸ Politickým nepřítelem je přitom někdo, kdo je existenciálně jiný a cizí.³⁶⁹

Posléze argumentuje, že v nejdůležitějších věcech, v otázkách náboženství, ctnosti a podstaty lidství nikdy nedojde ke shodě.³⁷⁰ A že „*politika slouží jako neustálá připomínka padlému lidstvu, že život je vážný a že jsou věci, na kterých skutečně záleží*“.³⁷¹ Argumentuje proti „tupému náboženství technicity“, které označuje za naivní víru v depolitizaci, která je hledána již po staletí a která je příslibem začátku univerzálního míru.³⁷² Pro Schmitta je však představa babylonské jednoty předtuchou blížící se závěrečné apokalypsy a možná – pokud do úvahy vložíme kontext pokročilé AI – v tom není příliš daleko od pravdy.

Politika je totiž mnohem spíše než o pouhé administraci problémů a řešení krizí o diplomacii mezi různými pohledy a způsoby života.³⁷³ Slovy Pitharta, který navazuje na anglickou politickou tradici, „*politika je spíše vzácněji se vyskytující způsob, jak se vyrovnat se skutečností společenských konfliktů*“.³⁷⁴ Nejde o hledání dokonalého řešení, ale o hledání řešení, se kterým bude většina lidí schopna a ochotna žít a tím o předcházení řešení nejzastším způsobem: ozbrojeným konfliktem. Tento aspekt politiky je však v AI technokracii, alespoň té bez funkčního politického aparátu, ze své podstaty nemožný.

Důležitým problémem takové představy je, že vede-li racionalita k uniformitě, pak je svoboda ze své podstaty iracionální záležitostí. Svoboda a individualismus nemohou ve zcela

³⁶⁶ Tamtéž.

³⁶⁷ MCCORMICK, J. P. Irrational Choice and Mortal Combat as Political Destiny: The Essential Carl Schmitt. *Annu. Rev. Polit. Sci.*, 2007, vol. 10, no. 1, s. 316.

³⁶⁸ SCHMITT, C. *Pojem politická*. Praha: Oikoymenh, 2007, s. 19, 20, 26.

³⁶⁹ Tamtéž, s. 27.

³⁷⁰ MEIER, H. *The Lesson of Carl Schmitt: Four Chapters on the Distinction between Political Theology and Political Philosophy*. Chicago: University of Chicago Press, 1998, s. 47.

³⁷¹ THIEL, P. The Straussian Moment. In: HAMERTON-KELLY, R. *Politics and Apocalypse*. Michigan: Michigan State University Press, 2007, s. 200.

³⁷² SCHMITT, C. *Pojem politická*. Praha: Oikoymenh, 2007, s. 91.

³⁷³ PITHART, P. *Obrana politiky*. Praha: Academia, 2005.

³⁷⁴ Tamtéž.

racionálním světě existovat. Otázka zachování lidské svobody je přitom v dlouhodobém kontextu naprosto zásadní, protože AI je silně centralizační technologie (například oproti internetu, což byla v době svého vzniku technologie decentralizační). Již některé současné technologie založené na AI přitom i v západních společnostech vedou k výrazným zásahům do osobní autonomie a soukromí.³⁷⁵ Ohrožení svobody ze strany racionality je však poměrně málo zmiňované nebezpečí a v kontextu nejbližších desetiletí bude jen a jen růst. Nejde však v evropské kultuře o koncept zcela nový. Podíváme-li se do Dostojevského knihy *Zápisky z podzemí*, uvidíme, že koncept existenciální svobody, která lidi vede k iracionalitě, je ústředním tématem tohoto díla. Aneb: „*lidé jsou stále lidmi, a ne klávesami u klavíru*“³⁷⁶. Člověk se nakonec vzepře jakékoli racionalitě, jen aby zůstal svobodný.

Neměli bychom navíc zapomínat, že každá technokracie je ze své podstaty také ideologií, kterou někteří mohou využít k dosažení svých politických cílů a k ovládnutí jiných.³⁷⁷ Jak nás varuje sci-fi román *Duna* v konceptu *služebnického džihádu* (The Butlerian Jihad), vždycky zde bude riziko, že jedna skupina lidí využije technologii k zotročení a ovládnutí jiných skupin lidí.³⁷⁸ A jak prozřetelně píše Schmitt, nárokování nepolitické čistoty je ve svém důsledku opět a pouze prováděním politiky. To, zda je něco nepolitické, je ve svém důsledku politické rozhodnutí.³⁷⁹

Na závěr této části bych zmínil ještě jedno, na první pohled nenápadné riziko. To spočívá v tom, že AI technokracie bude příliš dobrá v tom, co maximalizuje. Toto souvisí s tzv. *problémem excellence*.³⁸⁰ Pokud jsme příliš efektivní v maximalizaci jedné veličiny (např. HDP), může to vést ke katastrofálním následkům v jiných oblastech (např. v oblasti životního prostředí). Bájný král Midas také chtěl vše, čeho se dotknul, proměnit ve zlato. Až posléze mu došlo, že to asi nebyl úplně nejlepší nápad. Podobné pasti excellence se nám mohou objevit i ve správě věcí veřejných skrze pokročilé AI systémy (jde v podstatě opět o svého druhu *specification failures*). Excellence je dobrá, když víme, na jaký kopec vyjít. Pokud jsme si však

³⁷⁵ ZUBOFF, S. *Věk kapitalismu dohledu: Boj o budoucnost lidstva u nové hranice moci*. Praha: Argo, 2022.

³⁷⁶ DOSTOJEVSKÝ, F. *Notes from the Underground*. Dover Publications, 1992, s. 21.

³⁷⁷ CHENG, L. *The Rise of Technocracy: Elite Transformation and Ideological Change in post Mao China*. Princeton University, Dissertation, 1992, s. 8–9.

³⁷⁸ SCHONI, S. Preventing a Butlerian Jihad: Articulating a Global Vision For the Future of Artificial Intelligence. *Journal of International Affairs*, 2019, vol. 72, no. 1, s. 135–142.

³⁷⁹ SCHMITT, C. *Pojem politična*. Praha: Oikoymenth, 2007, s. 20.

³⁸⁰ WEINSTEIN, E. 2013: What *Should* We Be Worried About? In: *Edge* [online]. 2013 [cit. 2022-11-15]. Dostupné z: <https://www.edge.org/response-detail/23879>.

vybrali špatný kopec, mívá katastrofální následky. Navíc nám neumí říct, na který kopec vůbec dává smysl lézt a jaké kopce připadají v úvahu.

Proto naopak potřebujeme v mnoha oblastech zachovat nedokonalost a iracionalitu a mít politický systém postavený na myšlence omylnosti (falibilismu), jak pro to zcela nadčasově argumentuje Popper.³⁸¹ Myslím si také, že bychom ani v současnosti neměli nechat o lidských právech, svobodách a osudech rozhodovat pouze AI systémy, neboť již dnešní praxe *pro futuro* zakládá kulturní normy, které by mohly být v budoucnu využity při legitimizaci AI technokratických způsobů spravování věcí veřejných. Vždy by mělo jít o proces, kde je přítomen v nějaké podobě také člověk, na kterém by mělo být konečné rozhodnutí. Technologie by tedy neměla člověka „nahradit“, ale spíše „doplnit“. Toto by se mohlo jevit jako samozřejmost, ale ani v dnešní době tomu tak není. Algoritmy na sociálních sítích například rozhodují zcela autonomně o svobodě projevu uživatelů a skoro nikomu to nepřijde zvláštní.³⁸²

Argumenty ve prospěch tohoto požadavku jsou navázané na lidskou důstojnost. Russell je velmi výstižně popsal následovně:

„O útoku na lidskou důstojnost lze uvažovat dvěma způsoby. První je zřejmý: tím, že strojům dáváme moc nad lidmi, se odsouváme do druhořadého postavení a ztrácíme právo podílet se na rozhodnutích, která se nás týkají. (Extrémnější formou tohoto přístupu je dát strojům pravomoc zabíjet lidi). Druhý je nepřímý: i když věříte, že nerozhodují stroje, ale ti lidé, kteří stroje navrhli, skutečnost, že tito lidští návrháři a zadavatelé nepovažují za účelné v takových případech zvažovat individuální okolnosti každého lidského subjektu, naznačuje, že životům ostatních lidí přikládají malou hodnotu. To je možná příznakem počínajícího velkého rozdělení mezi elitou obsluhovanou lidmi a rozsáhlou podtřídou obsluhovanou a ovládanou stroji.“³⁸³

³⁸¹ POPPER, K. *Otevřená společnost a její nepřátelé I*. Praha: Oikoymenh, 1994.

³⁸² Viz v detailu HOŘEŇOVSKÝ, J. Vliv algoritmického rozhodování a imperativu zachování základních práv na pojetí civilněprávní odpovědnosti sociálních sítí za obsah nahraný třetími subjekty. In: SEDLÁČEK, M. a kol. (eds.). *Povinnost a odpovědnost v civilním právu hmotném a procesním. Speciální problémy*. Praha: Wolters Kluwer, 2021, s. 315–322.

³⁸³ RUSSELL, S. *Human Compatible. AI and the Problem of Control*. London: Allen Lane, 2019, s. 127.

Dobrým směrem jde v tomto kontextu v současnosti např. čl. 22 odst. 1 GDPR³⁸⁴ nebo čl. 20 odst. 6 DSA³⁸⁵. I na těchto ustanoveních je přitom vidět, že pro zákonodárce nejde v současnosti o příliš důležité téma, neboť oba tyto předpisy obsahují řadu výjimek, kdy se tato obecná pravidla neuplatní. Například čl. 22 odst. 2 GDPR obsahuje následující velmi širokou výjimku:

„Odstavec 1 se nepoužije, pokud je rozhodnutí: a) nezbytné k uzavření nebo plnění smlouvy mezi subjektem údajů a správcem údajů; b) povoleno právem Unie nebo členského státu, které se na správce vztahuje a které rovněž stanoví vhodná opatření zajišťující ochranu práv a svobod a oprávněných zájmů subjektu údajů; nebo c) založeno na výslovném souhlasu subjektu údajů.“

4.4 Závěr

Tato část představila futuristické téma, které se však pomalu stává, například nástupem velkých jazykových modelů typu GPT-4, novou realitou. Vzhledem k tomu, že budoucnost není předem stanovena a že záleží pouze na nás, jak bude vypadat, je v našich rukou, zda se z umělé inteligence stane pověstné Sauronovo oko, které všechno vidí a všechny kontroluje, nebo zda půjde o technologii, která lidstvu pomůže řešit důležité problémy a zároveň bude respektovat naše základní lidská práva a svobody, včetně ochrany soukromí.

Obava, že se z umělé inteligence minimálně na části planety stane Sauronovo oko, rozhodně není přehnaná. Pro autoritářské a totalitní režimy je AI dar z nebes, do kterého investují nemalé prostředky. Zatím tak činily především za účelem kontroly vlastního obyvatelstva, pomalu jim však dochází také jejich užitečnost na bitevním poli. Vždyť užití autonomních systémů v kontextu války lze v současnosti pozorovat takřka v přímém přenosu na Ukrajině. Jak píše například David Ignatius z Washington Post: *„Ukrajinci spojují svého odvážného bojového ducha s nejmodernějším zpravodajským softwarem a softwarem pro řízení boje, jaký kdy byl v boji použit.“*³⁸⁶ Podobně reportuje George Grylls z Times of London, který

³⁸⁴ „Subjekt údajů má právo nebýt předmětem žádného rozhodnutí založeného výhradně na automatizovaném zpracování, včetně profilování, které má pro něho právní účinky nebo se ho obdobným způsobem významně dotýká.“

³⁸⁵ „Poskytovatelé online platforem zajistí, aby rozhodnutí podle odstavce 5 byla přijímána pod dohledem adekvátně kvalifikovaných zaměstnanců, a nikoli výhradně na základě automatizovaných postupů.“

³⁸⁶ IGNATIUS, D. How the algorithm tipped the balance in Ukraine. In: *Washington Post* [online]. 19. 12. 2022 [cit. 2023-11-15]. Dostupné na: <https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/>.

k tomu dodává: „*Ti, kteří viděli umělou inteligenci v akci, nepochybuji o její revoluční síle.*“³⁸⁷ Tato pozorování v sobě zahrnují nepřeberné množství nových jak etických, tak právních otázek, které jsou otevřené dalšímu výzkumu a které zůstávají nezodpovězené. Jedna věc je však zřejmá i bez substantivní analýzy: v otázce vývoje umělé inteligence již není cesty zpět.

Západ, tedy i nás, čekají nové „závody ve zbrojení“, tentokrát na poli autonomních systémů, a to jak v civilním, tak vojenském kontextu. Uvažujeme-li proto o regulaci AI, měli bychom mít vždy na paměti i skutečnost, že tento technologický závod nemůžeme jako západní svět prohrát. Jakákoli regulace, která bude nakonec přijata, by proto neměla výrazně omezit možnost tyto komplikované systémy vyvíjet a společensky implementovat. Mohlo by jít o chybu, která by byla pro demokratické otevřené společnosti osudová. To však neznamená, že by se bezpečnostní otázky spojené s AI, například problematika *specification failures*, měly podceňovat, naopak. Musíme najít cestu, jak skloubit obojí.

Na začátku roku 2023 byla zveřejněna petice požadující přerušení vývoje umělé inteligence na dobu minimálně šesti měsíců.³⁸⁸ Takový krok může krátkodobě vypadat jako smysluplný. Jak již bylo upozorněno, v historii technologií existuje mnoho případů, kdy došlo k tzv. technologickému sebeomezení – situaci, kdy určité technologie mohly být vynalezeny, ale jejich vývoj byl úmyslně zpožděn, nebo dokonce zastaven.³⁸⁹ Nicméně v současné době je tento krok v souvislosti s umělou inteligencí, vzhledem ke globálnímu vývoji této technologie, již nemožné realizovat.

Nemáme tedy na výběr než se naučit s touto technologií žít. Úkolem právní regulace v nadcházejících letech bude zajistit takovou míru bezpečnosti systémů umělé inteligence a jejich vhodného společenského využití, které zabrání nejzávažnějším problémům spojeným s touto technologií, aniž by však omezilo další technologický vývoj v této kritické oblasti. Čeká nás mnoho otázek týkajících se využívání těchto technologií v soukromém sektoru, ve veřejné správě, ale i v našich osobních životech. Nebude to snadný úkol, ale nemáme jinou možnost než se o něj pokusit a být úspěšní. V tomto úkolu bude zásadní roli hrát nejen právní regulace,

³⁸⁷ GRYLLS, G. Ukraine is outflanking Russia with ammunition from Big Tech. In: *The Times* [online]. 24. 12. 2022 [cit. 2023-11-15]. Dostupné na: <https://www.thetimes.co.uk/article/ukraine-is-outflanking-russia-with-ammunition-from-big-tech-lxp6sv3qz>.

³⁸⁸ FUTURE OF LIFE INSTITUTE. Pause Giant AI Experiments: An Open Letter. 2023 [cit. 2024-01-06]. Dostupné na: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

³⁸⁹ MAAS, M. Paths Untaken: The History, Epistemology and Strategy of Technological Restraint, and lessons for AI. In: *VerfBlog* [online]. 9. 8. 2022 [cit. 2022-10-22]. Dostupné na: <https://verfassungsblog.de/paths-untaken/>.

ale jak bylo ukázáno, také teoretické poznatky, které právní věda vytvořila během své existence.

Tato práce nemůže tyto otázky v jejich plné komplexitě zcela vyřešit, může však otevřít témata, která stojí za to zkoumat. Doufám, že na můj – v mnoha ohledech jistě nedokonalý – pohled naváže další právní a společenský výzkum. Doufám také, že tento text bude inspirovat mnoho dalších badatelů a výzkumníků k ponoření se do dosud neprobádaných oblastí spojených se vztahem práva a umělé inteligence. Abych těmto výzkumníkům pomohl, na úplný závěr této části nastíním další témata, kterým by měla být v budoucnu věnována pozornost.

Patří mezi ně například problematika užití AI ve válce a vztah k humanitárnímu právu, téma užití AI ve správě věcí veřejných, například na úřadech práce, ve zdravotnictví nebo na finančních úřadech, či problematika limitů užívání těchto systémů v soukromých společnostech, a to nejen ve vztahu k pracovnímu právu.

První zmíněné téma obsahuje například následující otázky: kdy může o zabití ve vojenském kontextu rozhodnout pouze stroj a kdy je třeba schválení člověkem? Jak předejít útokům na zakázané objekty, jako jsou například nemocnice nebo školy? Jaké pojistky tedy musí tyto systémy obsahovat a jaké normy musí armády dodržovat, aby nedocházelo k neetickým jednáním? Všechny tyto otázky jsou zcela fundamentální a bohužel také velmi aktuální. Již pominula doba, kdy by AI aktivně a na základě svého rozhodnutí nebrala lidské životy ve válečných konfliktech.

Druhá kategorie otázek se týká fungování veřejné sféry. Ta jednoduše nemůže efektivně fungovat, pokud nemá poměrně přesný obrázek o tom, jak vypadá realita – žádný jednotlivec však tuto realitu vzhledem k její komplexnosti není schopen ani vzdáleně pojmout. Právě tento problém je schopna řešit AI v kombinaci s užitím velkých dat. Možnost vytvořit *digitální dvojče* neboli *ontologii* určité části veřejné sféry, a s pomocí AI systémů v reálném čase reflektovat její fungování, je zcela zásadní nová možnost pro rekonstrukci státu, kterou nám současné technologie umožňují.

Jde však také o něco, co náš stát doposud ignoruje, byť se tyto projekty v USA i jinde již pomalu rozjíždějí. Tam, kde se přitom podobné projekty implementují, jsou jejich výsledky výborné a například v kontextu zefektivnění zdravotnictví zcela reálně zachraňují životy

velkého množství lidí, například výrazným snížením čekacích lhůt.³⁹⁰ Pokud by se podařilo to samé udělat v ČR, mohlo by to výrazně vylepšit fungování naší veřejné správy. Již navíc došlo k zakomponování velkých jazykových modelů do těchto systémů, což umožňuje ptát se na konkrétní dotazy zodpovídané na základě identifikovatelných dat. Státy, které této technologii využijí, budou proto mít ohromnou konkurenční výhodu, neboť tím mohou zlepšit spravedlnost a efektivitu svého fungování; a ve výsledku tím i navrátit legitimitu svých institucí.

Třetí obrovskou oblastí je oblast užití AI v soukromém sektoru. Zde je nepřehledné množství jak právních otázek, tak možností faktického zlepšení jejich fungování. Míra neefektivity v našem existujícím světě je obrovská a pokud by se nám podařilo správně využít velká data a AI algoritmy, mohli bychom o desítky procent zlepšit ekonomické výsledky v podstatě všude. Pokud u toho zároveň správně nastavíme právní pravidla, například ohledně ochrany zaměstnanců, skutečně můžeme vytvořit svět, kde se všem bude lépe žít. A to není málo.

³⁹⁰ HAYWARD, E. NHS cancer treatment wait reduced in Palantir software pilot. In: *The Times* [online]. 27. 12. 2022 [cit. 2023-11-15]. Dostupné na: <https://www.thetimes.co.uk/article/nhs-cancer-treatment-wait-reduced-in-palantir-software-pilot-wldp568pj>.

5. Sociální sítě, svoboda projevu a informační veřejný prostor³⁹¹

Britský právní teoretik Raz v roce 1991 napsal, že svoboda projevu je klíčová především pro lidi, kteří mluví a projevují názory veřejně. Ostatním, kteří se veřejně neprojevují, se nemusí jevit natolik důležitá, proto mohou považovat za podstatnější jiná práva, která se jich naopak bezprostředně dotýkají.³⁹² V jeho úvaze hrálo roli, že pouze málo lidí mělo v minulosti přístup do veřejného prostoru.

Sociální sítě a internet tuto realitu v posledních letech zcela zásadně změnily, neboť zpřetrhaly rozdíl mezi veřejným a soukromým projevem. Každý projev se tak potenciálně stal projevem veřejným a svoboda projevu se začala dotýkat mnohem většího množství lidí. Internet a sociální sítě tudíž představují důležitý socio-technologický milník, na který je třeba reagovat. Zároveň se v tomto kontextu stále častěji řeší témata jako dezinformace a tzv. *projevy z nenávisli* (hate speech), což celou diskusi komplikuje. Na začátek této části proto považují za účelné zopakovat, proč je vůbec svoboda projevu pro demokratickou společnost důležitá.

5.1 Význam svobody projevu v demokratické společnosti

Svoboda projevu je základní politické právo chráněné čl. 17 Listiny, přičemž cílem, ke kterému by její ochrana měla v první řadě směřovat, je demokratická názorově pluralitní společnost. Demokratická názorově pluralitní společnost je protikladem vůči uniformitě definující totalitní režimy.³⁹³

Svoboda projevu je jednou z nejvýznamnějších liberálních svobod,³⁹⁴ která ve společnosti plní širokou paletu funkcí. Na obecné úrovni je účelem všech politických práv umožnit občanům angažovat se ve veřejném prostoru a podílet se na správě věcí veřejných.³⁹⁵ V politické oblasti jde především o možnost každého ovlivňovat veřejný život. Tyto obecnější úvahy jsou konkretizovány soudní judikaturou nejvyšších soudů.

³⁹¹ Tato část vychází z a) HOŘEŇOVSKÝ, J. Aktuální otázky svobody projevu v době sociálních sítí. In: KUDRNA, J. (ed.). *Listina v kontextu společenských změn (K 30 výročí jejího přijetí)*. Praha: Wolters Kluwer, 2021, s. 171–180; b) HOŘEŇOVSKÝ, J. Vliv algoritmického rozhodování a imperativu zachování základních práv na pojetí civilněprávní odpovědnosti sociálních sítí za obsah nahraný třetími subjekty. In: STŘELEČEK, T.; SEDLÁČEK, M. (eds.). *Povinnost a odpovědnost v civilním právu hmotném a procesním. Speciální problémy*. Praha: Wolters Kluwer, 2021, s. 315–322.

³⁹² RAZ, J. Free Expression and Personal Identification. *Oxford Journal of Legal Studies*, 1991, vol. 11, no. 3, s. 303.

³⁹³ JÄGER, P.; MOLEK, P. *Svoboda projevu: demokracie, rovnost a svoboda slova*. Praha: Auditorium, 2007, s. 13.

³⁹⁴ WINTR, J. *Principy českého ústavního práva*. Plzeň: Aleš Čeněk, 2015, s. 153.

³⁹⁵ Tamtéž, s. 151.

Dle judikatury ESLP je svoboda projevu podstatným základem demokratické společnosti a nepostradatelnou podmínkou jejího rozvoje, jakož i podmínkou individuální seberealizace každého jednotlivce.³⁹⁶ Z toho lze vyvozovat, že účelem právní ochrany není pouze sféra veřejná a politická, nýbrž i oblast čistě osobní a soukromá. Právo proto chrání i projevy, které se svým obsahem nedotýkají veřejného života, neboť seberealizace člověka je také důležitým účelem.

Úlohu svobody projevu v demokratické společnosti výborně vystihl český Ústavní soud, když v návaznosti na judikaturu ESLP konstatoval:

„Ústavní soud se ztotožňuje s judikaturou Evropského soudu pro lidská práva, podle níž svoboda projevu představuje jeden z nejdůležitějších základů demokratické společnosti a jednu z hlavních podmínek pokroku a rozvoje každého jednotlivce. Jako taková se svoboda projevu vztahuje nejen na ‚informace‘ nebo ‚myšlenky‘ příznivě přijímané či považované za neškodné či bezvýznamné, ale i na ty, které zraňují, šokují nebo znepokojují: tak tomu chtějí pluralita, tolerance a duch otevřenosti, bez nichž není demokratické společnosti.“³⁹⁷

Důvodem je, že pluralitní otevřená diskuse s sebou nevyhnutelně přináší také témata a pohledy, které mohou šokovat a být velmi nepříjemné.³⁹⁸ Jak píše Jäger a Molek: *„Pestrost názorů (ba právě ona) přispívá ke smysluplnosti jakékoli debaty, pokud nemá být jen přesvědčováním přesvědčených či deklamací naučených hesel a frází, ale komunikací vytvářející občanskou společnost.“³⁹⁹* Skutečnost, že v historii mohly být vyjádřeny i původně velmi kontroverzní myšlenky (například že otroctví je nepřijatelné nebo že by ženy měly mít volební právo), vedla ke společenskému pokroku a k řešení aktuálních společenských problémů. Otevřená a svobodná diskuse proto vždy byla alternativou k zahrnování a dogmatismu.⁴⁰⁰

Možná i z tohoto důvodu Nejvyšší soud Spojeného království konstatoval, že svoboda projevu je *krevním řečištěm demokracie*.⁴⁰¹ Raz navíc argumentuje tím, že svoboda projevu ze

³⁹⁶ Rozhodnutí ESLP ze dne 7. 12. 1976 ve věci Handyside proti Spojenému království.

³⁹⁷ Nález Ústavního soudu ze dne 15. 3. 2005, I. ÚS 367/03.

³⁹⁸ JÄGER, P.; MOLEK, P. *Svoboda projevu...*, s. 13.

³⁹⁹ Tamtéž.

⁴⁰⁰ MILL, J. S. *O svobodě myšlení a slova*. Praha: Institute H21, 2020.

⁴⁰¹ Rozhodnutí britského Nejvyššího soudu R v Secretary of State for the Home Department, ex p Simms [2000] 2 AC 115, s. 126.

své podstaty podporuje toleranci, neboť když jsou lidé vystaveni široké paletě názorů, pohledů a hodnot, utváří to více harmonickou společnost.⁴⁰² Jde o nalézání kompromisů a otupování hran. Široká svoboda projevu umožňující veřejnou diskusi proto ve výsledku vede také ke společnosti pestřejší a barevnější, protože v ní je možno slyšet více pohledů na svět a společnost.⁴⁰³

5.2 Dělení projevů ve veřejném prostoru a éra úpadku svobody slova

Přemýšlíme-li o informačním veřejném prostoru a jeho limitech, můžeme si tuto problematiku na začátek rozdělit do několika schematických protikladů. Na jedné straně máme rovinu právní, jež určuje, které projevy jsou legální a které naopak nelegální (v praxi bude navíc existovat poměrně široká šedá zóna, kde s jistotou nevíme). Na druhé straně existuje rovina kulturní (společenská), která se dotýká toho, co se ve „slušné mainstreamové společnosti“ ještě může říkat a za co již hrozí alespoň potenciální společenská sankce.

Tyto roviny se v mnohém budou prolínat, avšak ne vždy. Můžeme tedy mít „společensky nepřijatelný“ projev, který je legální, a naopak i projev nelegální, který je společensky zcela přijatelný. Toto rozlišení, byť poměrně jednoduché, pořád nebývá ve veřejném prostoru dostatečně pochopeno, a to hlavně ve vztahu k problematice tzv. „projevů z nenávisť“ a „šíření dezinformací“, kdy se mnohdy mylně předpokládá, že to, co je někým hodnoceno jako škodlivé, musí být zároveň vždy právně postižitelné. Mojí výchozí tezí je, že se v obou těchto rovinách prostor svobodné diskuse stále více zužuje, tedy, že je právo na svobodný informační prostor stále více omezováno. Tato teze má přitom zcela reálný empirický základ, jak dokládá studie organizace Justitia s názvem *The Free Speech Recession Hits Home: Mapping Laws and Regulations Affecting Free Speech in 22 Open Democracies* z roku 2023.⁴⁰⁴

V základu tedy existují tři základní kategorie přístupu práva k projevům ve veřejném prostoru. První je kategorie *právně chráněného ne-škodlivého projevu*, která je nejméně zajímavá, neboť s ní obecně nikdo nemá problém. Do této bude spadat například diskuse o počasí, o tenise nebo o výstavě koček. Být je tato kategorie právně chráněná, jde o ochranu v

⁴⁰² RAZ, J. *Free Expression and Personal Identification...*; BOLLINGER, C. *The Tolerant Society*. Oxford: Oxford University Press, 1986.

⁴⁰³ JÄGER, P; MOLEK, P. *Svoboda projevu...*, s. 13.

⁴⁰⁴ JUSTITIA AND THE AUTHORS. *The Free Speech Recession Hits Home. Mapping Laws and Regulations Affecting Free Speech in 22 Open Democracies*. 2023. Dostupné z: <https://futurefreespeech.org//the-free-speech-recession-hits-home/>. Je zde i kapitola zaměřená na ČR, viz s. 71–77.

podstatě latentní a do značné míry zbytečnou, neboť co není nikým subjektivně hodnoceno jako škodlivé, nebude v realitě nikdo rozporovat. Měla-li by proto svoboda projevu zahrnovat pouze tuto kategorii, byla by ve své podstatě zbytečná.

Druhou je kategorie *právně chráněného škodlivého projevu*, jejíž důležitou podmnožinou je *ústavně chráněný škodlivý projev*. Rozdíl mezi těmito pojmy spočívá v možnosti zákonodárce takový projev, bude-li to považovat za nutné, omezit. Půjde-li o kategorii ústavně chráněného škodlivého projevu, pak bude omezení takového projevu velmi striktně limitováno ústavním pořádkem a v některých případech bude dokonce zcela nemožné. Je důležité zdůraznit, že teprve zde začíná být svoboda projevu důležitá, neboť její podstatou je chránit takové projevy, které mohou být jednotlivci, veřejnou mocí či skupinami obyvatel subjektivně považovány za škodlivé (a někdy i za velmi škodlivé). Za příklady lze uvést karikatury Proroka Mohameda nebo román *Satanské verše* od spisovatele Salmana Rushdieho. Tyto projevy pobuřují, ve svobodné společnosti však jsou chráněny ústavním pořádkem i mezinárodními smlouvami.⁴⁰⁵ Z tohoto důvodu jde o ústavně chráněné (a v subjektivním smyslu) škodlivé projevy.

Třetí kategorií je *právem zakázaný škodlivý projev*. Aby mohl být projev právem postižitelný, musí naplňovat specifickou míru a podobu škodlivosti, subjektivní pocit části obyvatelstva či veřejné moci nestačí. V právním státě půjde o tři základní požadavky: 1) projev musí být poměrně přesně vymezen právem (tedy například zákonem či precedentem); 2) limitace projevu musí splňovat podmínky omezení základní svobody stanovené ústavním pořádkem (například Listinou) a mezinárodním právem; 3) škodlivost (hrozící újma) by měla být *konkrétní, bezprostřední a zvláště závažná* (tomuto můžeme souhrnně říkat *kritérium naléhavosti*).⁴⁰⁶ Není-li jakákoli z těchto podmínek splněna, hrozí arbitrární omezení veřejného prostoru veřejnou mocí, *mrazivý účinek* (chilling effect), újma na právním státě a zásahy do základních práv obyvatel.

Vzhledem k tomu, že první kategorii nemá smysl rozebírat, pojďme se rovnou podívat více dopodrobna na kategorii druhou. Tato kategorie je zajímavá v tom, že ukazuje na klíčový rozpor v rovině právní a společenské. I projev, který může být subjektivně spatřován částí společnosti nebo dokonce celým mainstreamem jako škodlivý, totiž ještě nemusí být a většinou

⁴⁰⁵ Viz např. nález Ústavního soudu ze dne 15. 3. 2005, I. ÚS 367/03 nebo rozhodnutí ESLP ze dne 7. 12. 1976 ve věci Handyside proti Spojenému království.

⁴⁰⁶ Viz STROSSEN, N. *Nenávist: Proč bychom se jí měli bránit pomocí svobody projevu, nikoli cenzurou*. Praha: Institute H21, 2021.

ani nebude protiprávní. Tato kategorie reflektuje historickou zkušenost, že omezování svobodného projevu je vždy extrémně rizikovou záležitostí, která se velmi snadno a často obrací proti názorovým menšinám ve společnosti za účelem udržení společenské konformity.⁴⁰⁷ Lidé jsou totiž od přírody tvorové morálně navyklí k *tribalismu*, pro které je omezování vnitřního disentu skupiny přirozeným prostředkem sjednocování klanu.⁴⁰⁸ Je tomu tak proto, že naše morálka evolučně vznikala v malých a navzájem soupeřících skupinách, kde bylo takové chování zcela racionální.⁴⁰⁹

Tento morální pud je pochopitelný, avšak koliduje s představou *otevřené společnosti*, která je na svobodě projevu postavena,⁴¹⁰ a také vede k *syndromu skupinového myšlení* (groupthink), který může mít fatální důsledky ve veřejných politikách.⁴¹¹ Důležité je, že ochrana svobody projevu tam, kde začíná dávat jakýkoli smysl, jde proti naší přirozené morálce a ve skutečnosti je tedy velmi nesamozřejmou záležitostí. Vždy se totiž najdou velmi „dobré důvody“, proč určitý projev omezit, z nichž mnohé přímo reflektuje i náš současný právní řád: ochrana veřejného pořádku, morálky, mravnosti, společenského řádu atd.⁴¹² Alternativním způsobem omezení je přijít s argumentací, že zásah do svobody projevu vlastně zásahem do svobody projevu nebo práva přijímat informace vůbec není.⁴¹³ ESLP tento argument často vznáší s odkazem na čl. 17 EULP, kdy argumentuje, že „zneužití práva“ na svobodu projevu do svobody projevu konceptuálně vůbec nespadá.⁴¹⁴

Tyto tribalistické morální kořeny se o to více projeví v momentě vnějšího ohrožení skupiny, například v případě válečného či pandemického ohrožení. Demokracie a liberální svobody většinou umírají na mimořádné stavy a pro svobodný informační veřejný prostor to neplatí jinak. V ten moment dojde k „morální panice“ společnosti i elit a všechny historické, právní i filozofické důvody pro ochranu projevu jdou stranou. Náznaky tohoto přístupu byly dle mého názoru viděny jak v covidové pandemii, kdy byla protiprávně smazána ezoterická

⁴⁰⁷ Viz MCHANGAMA, J. *Svoboda projevu: Od Sokrata po sociální síť*. Praha: Institute H21, 2022.

⁴⁰⁸ S touto myšlenkou přišel F. A. Hayek a je rozebírána v kontextu moderních poznatků v GAUS, G. *The Open Society and Its Complexities*. Oxford: OUP, 2021, kapitola “Prolegomenon: Hayek’s Three Unsettling Theses”.

⁴⁰⁹ Tamtéž.

⁴¹⁰ POPPER, K. *Otevřená společnost a její nepřítelé I...*

⁴¹¹ JANIS, I. L. *Victims of Groupthink: A psychological study of foreign-policy decisions and fiascoes*. Houghton Mifflin Company, 1972.

⁴¹² Pro výčet viz čl. 17 Listiny.

⁴¹³ Viz např. usnesení Městského soudu v Praze ze dne 24. 6. 2022, č. j. 14 A 39/2022-14.

⁴¹⁴ Viz KMEC, J.; KOSAŘ, D.; KRATOCHVÍL, J.; BOBEK, M. *Evropská úmluva o lidských právech*. Komentář. Praha: C. H. Beck, 2012, s. 997.

divadelní hra na popud veřejné moci,⁴¹⁵ tak v době začátku války na Ukrajině, kdy se veřejná moc za potlesku veřejnosti rozhodla zcela bez právního základu vypnout tzv. dezinformační weby.⁴¹⁶

Pro tuto kapitolu je podstatné, že do druhé kategorie *právně chráněného škodlivého projevu* spadá i podstatná část toho, co by se dalo v současnosti označit za „projevy z nenávisti“ a „šíření dezinformací“. Důležité je, že je tomu tak dobře, neboť ve svobodné společnosti má být právní omezení, a tím pádem omezení veřejného informačního prostoru, až krokem poslední instance, kdy hrozí určitá *konkrétní, bezprostřední a vážná* újma.⁴¹⁷ Je tomu tak proto, že je právně nemožné široce definovat omezení projevu tak, aby nemohlo být v budoucnu nebo v jiném kontextu použito zcela jiným způsobem, než bylo původně zamýšleno, mnohdy dokonce proti autorům původní regulace.⁴¹⁸ Kvůli existenci tohoto limitu práva je tedy nutné trvat pouze na velmi detailně definovaných kritériích omezení, což například test *konkrétní, bezprostřední a vážné* újmy více či méně splňuje.⁴¹⁹

Co by naopak nikdy nemělo být důvodem omezení projevu, je pouhé *morální pohoršení* (tzn. existence subjektivního morálního pohoršení jako naplnění kritéria *újmy*). Jednoduchým důvodem je, že každého z nás morálně pohoršují trochu jiné věci. Totéž platí o různých světových názorech, ideologiích nebo náboženských vyznáních. Pokud by se morální pohoršení stalo dostatečným důvodem omezení projevu, ve výsledku by se veřejný prostor stal pouze mocenským bojištěm, ve kterém by vítěz soubroje autoritativně diktoval své požadavky ostatním. Jakákoliv smysluplná veřejná diskuse by se vytratila. Teprve tehdy, kdy je smyslem diskurzu přesvědčit, nikoli porazit (například z důvodu, že jednoduše není, jak porazit druhou stranu), může veřejný diskurz vést k hledání veřejného dobra.

Podstatné je, že se v poslední době objevují tendence tuto druhou kategorii *právně chráněného škodlivého projevu* stále více omezovat a její obsah přesouvat do kategorie třetí. Do legislativní cesty omezování svobodného projevu lze například v tomto kontextu zahrnout

⁴¹⁵ HOŘEŇOVSKÝ, J. Ústavněprávní aspekty nuceného stáhnutí představení Jaroslava Duška během koronavirové krize. In: WINTR, J.; ANTOŠ, M. (eds.). *Ústavní právo a koronavirus*. Praha: Leges, 2021, s. 77–85.

⁴¹⁶ MIZEROVÁ, T.; HARAŠTA, J. Dostupnost mediálního obsahu: blokování dezinformačních webů. *Právník*, 2023, roč. 162, č. 5, s. 415–430.

⁴¹⁷ Obšírně viz STROSSEN, N. *Nenávist: Proč bychom se jí měli bránit pomocí svobody projevu, nikoli cenzurou*. Praha: Institute H21, 2021.

⁴¹⁸ Tamtéž.

⁴¹⁹ Jak upozorňuje například Oldřich Kužílek, tento test není ani zdaleka pouze doménou amerického ústavního práva, nýbrž jeho náznaky lze nalézt například v případech kritérií pro rozpouštění politických stran. Viz např. rozsudek Nejvyššího správního soudu ze dne 17. 2. 2010, č. j. Pst 1/2009-348.

návrhy zákonů na zneprístupňování dezinformačních webů,⁴²⁰ tvorbu nových trestných činů spočívajících v „šíření dezinformací“⁴²¹ či regulace ze strany Evropské unie, například skrze některé instituty v DSA.⁴²² Jiné způsoby zahrnují různá „doporučení“, které informační zdroje zneprístupnit či zablokovat, adresována soukromým subjektům.⁴²³

Omezení svobodného projevu se však děje i cestou plíživou, kdy jsou určitá ustanovení právního řádu, především trestního zákoníku, postupem času využívána – pod vlivem „morální paniky“ elit – jinak, než tomu bylo původně zamýšleno. Tato ustanovení se nemění v tradičním slova smyslu, spíše dochází k zásadní interpretační proměně toho, co znamenají. Nejlepší ukázkou tohoto fenoménu je v současnosti trestný čin *podpora a propagace terorismu*,⁴²⁴ který byl původně zamýšlen především pro trestání příznivců Islámského státu. Postupem času se však tento trestný čin začal aplikovat i na komentáře a videa na sociálních sítích, které s terorismem, jak si ho obvykle představujeme, nemají nic moc společného.⁴²⁵

Tento trestný čin názorně ukazuje, jak represivní režim náš trestní zákoník v mnoha případech stanovuje. Spáchání tohoto činu na internetu je navíc znakem kvalifikované podstaty, tudíž zde hrozí až 15 let odnětí svobody (nejnižší sazba je v tomto případě 5 let, což zamezuje uložit podmíněný trest nebo využít institut mimořádné snížení trestu).⁴²⁶ Kombinace primitivního čtení normativních textů ze strany státních zástupců a soudců a pravidelné

⁴²⁰ Viz např. HODULÍK, J. Zákon o blokaci dezinfowebů má bránit demokracii, ale podle odborníků ji spíš ohrožuje. In: *Respekt* [online]. 17. 10. 2022 [cit. 2023-11-15]. Dostupné na: <https://www.respekt.cz/kontext/zakon-na-blokaci-dezinfowebu-ma-branit-demokracii-ale-podle-odborniku-ji-zatim-spis-ohrozuje>.

⁴²¹ Viz *Akční plán boje proti dezinformacím*. Autor měl k akčnímu plánu přístup, i když nikdy nebyl oficiálně zveřejněn. O jeho obsahu se lze dočíst například zde FUJÁČEK, J. Vypínání webů a stíhání autorů. Vláda chystá akční plán proti dezinformacím. In: *Echo24* [online]. 30.12.2022 [cit. 2024-01-06]. Dostupné na: <https://echo24.cz/a/H38eq/zpravy-domaci-vlada-dezinformace-cenzura-strategie-legislativa-vlada-vypinani-webu-vnitro-rakusan-klima-zmocnenec>.

⁴²² Viz např. třídílný rozbor Ivo Telce: 1) TELEČ, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část I. In: *Právní prostor* [online]. 22.05.2023 [cit. 2023-11-15]. Dostupné na: <https://www.pravniprostor.cz/clanky/ostatni-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-i>; 2) TELEČ, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část II. In: *Právní prostor* [online]. 25.05.2023 [cit. 2023-11-15]. Dostupné na: <https://www.pravniprostor.cz/clanky/obcanske-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-ii>; 3) TELEČ, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část III. In: *Právní prostor* [online]. 29.05.2023 [cit. 2023-11-15]. Dostupné na: <https://www.pravniprostor.cz/clanky/mezinarodni-a-evropske-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-iii>.

⁴²³ MIZEROVÁ, T.; HARAŠTA, J. Dostupnost mediálního obsahu: blokování dezinformačních webů. *Právník*, 2023, roč. 162, č. 5, s. 415-430.

⁴²⁴ § 312e zákona č. 40/2009 Sb., trestní zákoník.

⁴²⁵ K tomuto fenoménu viz WINTR, J. Schvalování atentátu a podobné výtržnosti. *Státní zastupitelství*, 2020, č. 3, s. 3–4.

⁴²⁶ Viz RÁLIŠ, P. Podpora a propagace terorismu – kladivo na internetové diskutéry? *Trestněprávní revue*, 2024, roč. 23, č. 1. [pozn. strany nelze určit, neboť v době odevzdání práce toto číslo ještě nevyšlo].

ignorování lidskoprávní roviny této problematiky vede k situaci, kdy se reálně může stát, že člověk půjde na deset let za mříže za komentářový příspěvek na internetu. Podobný interpretační posun lze sledovat i u dalších verbálních trestných činů, jako je například *výtržnictví*, které bylo užito k stíhání politického projevu vyvěšení červených trenýrek na Pražském hradě.⁴²⁷ Jde o trend neblahý, na který není dostatečně upozorňováno.

5.3 Svoboda projevu a sociální sítě

Svobodu projevu silně ovlivnil nástup internetu a sociálních sítí. Digitální komunikace učinila ostatní formy projevu, například projevy na náměstích, společensky méně důležitými. Tyto ostatní formy projevu jsou přitom právním řádem chráněny i v dnešní době poměrně obstojně. Nové směry digitální komunikace jsou však chráněny velmi málo, a to hlavně z toho důvodu, že jde o platformy vlastněné soukromými společnostmi.

Důležitost digitální komunikace v dnešní informační společnosti ale nelze podceňovat. Výzkum organizace Pew Research například ukázal v americkém kontextu velmi zajímavá data. V USA 18 % lidí získává informace o politice a věcech veřejných primárně ze sociálních sítí. Ve věkové skupině 30–39 let je to celých 40 %. A v nejmladší skupině 18–29 let dokonce 48 %.⁴²⁸ V Evropě nebudou čísla pravděpodobně příliš odlišná.

Omezení toku informací na sociálních sítích proto představuje ze své podstaty riziko pro demokracii,⁴²⁹ neboť sociální sítě jsou veřejnými prostranstvími dnešní doby.⁴³⁰ Sociální sítě dnes bez nadsázky plní úlohu, kterou historicky plnily ulice, náměstí, média a televize.⁴³¹ Klíčovou otázkou dnešní doby proto je, jak zajistit funkční a svobodný veřejný prostor i ve světě, kde se velická část naší komunikace odehrává právě na sociálních sítích ovládaných soukromými subjekty.

Aktuální znepokojivou tendencí je, že sociální sítě prokazatelně mažou komentáře a názory týkající se věcí veřejných, a to včetně komentářů politických kandidátů, které neporušují

⁴²⁷ WINTR, J. Schvalování atentátu a podobné výtržnosti. *Státní zastupitelství*, 2020, č. 3, s. 3–4.

⁴²⁸ PEW RESEARCH CENTER. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable. 2020 [cit. 2021-08-13]. Dostupné z: <https://www.pewresearch.org/journalism/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>.

⁴²⁹ STROSSEN, N. *Hate: Why We Should Resist it With Free Speech, Not Censorship*. Oxford: Oxford University Press, 2020, Epilogue.

⁴³⁰ United States Supreme Court decision of 27 February 2017, *Packingham v. North Carolina*, 582 U.S., s. 2.

⁴³¹ Viz FERGUSON, N. *Věž a náměstí: mocenské sítě od svobodných zednářů po Facebook*. Praha: Argo, 2019.

žádná ustanovení právního řádu.⁴³² Ze studie dánské neziskové organizace Justitia ve vztahu k dánskému Facebooku vyplývá, že pouze 1,1 % smazaných komentářů je trestně-právně postižitelných a naopak celých 36 % smazaného obsahu je obecným zcela nezávadným vyjádřením názoru.⁴³³ Tato čísla nám naznačují, že skutečný problém spojený se sociálními sítěmi může být nikoli „nedostatečné mazání“, ale naopak „příliš mnoho mazání“. Tomuto závěru nasvědčovala také dosud nepublikovaná data, se kterými byl autor této práce seznámen na konferenci dne 4. 12. 2023 s názvem *The State of Free Speech: Challenges and solutions for the digital age*.⁴³⁴ Z nich vyplývá, že ve všech zkoumaných státech (Francie, Německo, Švédsko) byla drtivá většina smazaného obsahu na sociálních sítích pouhým *obecným vyjádřením názoru* (general expression of opinion).

Sociální sítě navíc obsah kontrolují arbitrárně a diskriminačně, neboť jejich komunitní podmínky nepředstavují ani zdaleka jasný a určitý seznam pravidel, který by exaktně definoval, co lze mazat a co ne. Navíc se často a účelově mění, mnohdy dle politických preferencí samotných zaměstnanců těchto platforem, kteří preferují určité názory a řečníky a jiné naopak omezují. Moderace obsahu se mnohdy děje za pomoci umělé inteligence a špatně placených zaměstnanců ze třetích zemí, jde proto o proces silně chybný. V USA například studie zjistila, že algoritmus směřující k odhalení a potlačení projevů z nenávisti 2,2krát častěji takto označoval příspěvky psané Afroameričany.⁴³⁵ Umělá inteligence neumí rozeznávat rozdílné sociální a kulturní kontexty, humor a nadsázku. A pokud obsah kontrolují zaměstnanci, kvalita moderace není o moc větší, protože většinou jde o špatně placené, přepracované lidi s minimálním tréninkem a žádnou odpovědností.⁴³⁶

Snaha o potlačení určité informace na internetu navíc není řešením, které by čemukoli pomohlo, neboť obvykle vede k pravému opaku. Taková informace se paradoxně obvykle rozšíří ještě mnohem více než před pokusem o smazání (tzv. Barbra Streisand Effect⁴³⁷), což se ukázalo i na případě smazání divadelní hry *Malá vizita* Jaroslava Duška, kdy se představení

⁴³² JUSTITIA. New report from Justitia: Digital freedom of expression and social media. 2020 [cit 2021-08-13]. Dostupné z: <https://globalfreedomofexpression.columbia.edu/publications/new-report-from-justitia-digital-freedom-of-expression-and-social-media/>.

⁴³³ Tamtéž.

⁴³⁴ JUSTITIA. Konference organizace Justitia ze dne 4. 12. 2023 *The State of Free Speech – Challenges and Solutions for a Digital Age*. 2023. Záznam k dispozici zde: <https://youtu.be/XQnNY0hR9fg?si=7SVc6YDheucaOfpH>.

⁴³⁵ STROSSEN, N. *Hate...*, Epilogue.

⁴³⁶ Tamtéž.

⁴³⁷ JANSEN, S.; MARTIN, B. The Streisand Effect and Censorship Backfire. *International Journal of Communication*, 2015, vol. 9, s. 656–671.

po smazání šířilo mnohem rychleji než před ním.⁴³⁸ Mazání obsahu navíc pomáhá radikalizaci společnosti a budování nedůvěry k mainstreamovým institucím, protože pokud si o něčem nelze přečíst na klasických platformách a v médiích, žene to lidi k extrémističtějším zdrojům. Tam ale oproti široce používaným platformám neexistuje názorová oponentura, která by mohla uvést věci na pravou míru. V takových končinách internetu se posléze uživatelé ještě více radikalizují, což by se bez prvotní cenzury velmi pravděpodobně nestalo. Diskuse a argumentace je proto mnohem užitečnější reakcí demokratické společnosti než cenzura.

5.4 Možnosti regulace sociálních sítí

Při uvažování o řešení těchto otázek je potřeba si předně definovat, co by mělo být cílem regulace. Dle mého názoru zajistit, aby svoboda projevu platila i na sociálních sítích. Tento přístup bere v potaz, že již máme vytvořený soubor pravidel definujících, co lze říkat ve veřejné sféře. Tyto standardy musíme pouze propojit se sociálními sítěmi. Jak toho ale lze docílit?

5.4.1. Regulace pomocí soutěžního práva⁴³⁹

Jedna z důležitých možností regulace sociálních sítí, byť ve veřejném prostoru ne příliš často zmiňovaná, je aplikace soutěžního práva. Jedná se o právní odvětví, které slouží k regulaci soukromoprávních monopolů a oligopolů. Soutěžní právo, navzdory svým inherentním omezením, by proto mohlo pomoci, protože je ze své podstaty přímo určeno k regulaci velkých soukromých společností.

V rámci tohoto směru existují dvě možnosti.⁴⁴⁰ Za prvé lze regulovat množství a velikost sociálních platform a tím cílit ke stavu, aby mezi nimi existovala konkurence. Konkurence a diverzita sociálních sítí může posléze přispívat ke svobodě projevu, protože v takové situaci mohou zákazníci jednoduše přejít na jinou sociální síť. Problémem je, zda sociální platformy nejsou pro uživatele užitečné právě proto, že jsou všichni propojení na těch stejných platformách, a zda by jejich rozdělení proto zásadně neuškodilo funkčnosti a atraktivitě pro uživatele.

⁴³⁸ K tomuto blíže viz HOŘEŇOVSKÝ, J. *Ústavněprávní aspekty nuceného stáhnutí představení Jaroslava Duška během koronavirové krize...*, s. 77–85.

⁴³⁹ K tomuto více dopodrobna viz HOŘEŇOVSKÝ, J. Regulation of Social Media Networks Through the Article 102 TFEU. In: ŠKRABKA, J.; VACUŠKA, L. (eds.). *Law in Business of Selected Member States of the European Union. Proceedings of the XII International Scientific Conference*. Praha: TROAS, s.r.o., 2020, s. 214–223.

⁴⁴⁰ KHAN, L. M. Amazon's Antitrust Paradox. *The Yale Law Journal*, 2017, vol. 126, no. 3, s. 790–802.

Alternativní možností je smířit se s tím, že sociální sítě jsou ze své podstaty monopoly, že jde o infrastrukturu, která nejlépe slouží, pokud je na ní co nejvíce lidí. V takovém případě lze platformy regulovat jako tzv. *public utilities*, tzn. jako infrastrukturu. Historicky se takto regulovaly například železnice či elektrické rozvody. Smyslem takové regulace je, aby společnost v dominantním postavení nemohla své dominance zneužívat na újmu zákazníkům či jiným společnostem.⁴⁴¹ Pokud sociální síť omezuje svobodu projevu svých zákazníků, může to být proto za určitých okolností posouzeno jako zneužití dominantního postavení. Uživatel, který chce sociální síť využívat například k participaci ve veřejné diskusi o věcech veřejných, se jednoduše nemůže sebrat a odejít jinam.

Často uváděný argument, že se jedná o čistě soukromé subjekty, a proto si mohou dělat co chtějí, nemůže vzhledem k důležitosti a velikosti sociálních sítí v dnešní společenské komunikaci obstát. Ostatně, tento argument ve skutečnosti nikdy neplatil, což názorně dokládá samotná existence soutěžního práva a dalších veřejnoprávních norem regulujících soukromé podnikání. Tento pohled se přitom začíná projevovat i v rozhodovací praxi soudů. Například německé úřady a soudy dovodily, že Facebook je společností v dominantním postavení na trhu sociálních sítí, z čehož mu vyplývají speciální povinnosti z oblasti soutěžního práva.⁴⁴²

Vodítkem pro určení hranice může být právě ústavní lidskoprávní standard ochrany svobody projevu, který prozařuje i do odvětví soutěžního práva. Aplikace lidskoprávního standardu je striktnější než používání komunitních pravidel a chrání svobodu projevu širěji. Ustálená judikatura navíc představuje přesnější vodítko.

5.4.2. Přímá regulace

Další variantou je přijetí přímé regulace sociálních sítí, která by stanovila pravidla pro moderaci obsahu a vyjmula tak tuto pravomoc od samotných platform. Otázkou však je, co by měl být zaručený standard ochrany svobody projevu a kdo by měl mít pravomoc jeho porušení či neporušení posuzovat. Nabízí se přitom dobrá inspirace americkým prvním dodatkem nebo evropským lidskoprávním standardem. Naopak jakýkoli přísnější standard by znamenal zásadní omezení svobody projevu, a tím pádem i ohrožení pro demokratickou společnost.

⁴⁴¹ VICKERS, J. Abuse of Market Power. *Economic Journal*, 2005, vol. 115, no. 504, s. 244–261.

⁴⁴² The Decision of the Higher Regional Court of Düsseldorf (Oberlandesgericht Düsseldorf) in interim proceedings, 26 August 2019, Case VI-Kart 1/19 (V).

Odstrašujícím případem by zde mohla být Francie, kde přijali tzv. Avia law, který se snažil velmi restriktivně regulovat chování platforem ohledně nenávistných projevů. Obsahoval například extrémně široký seznam toho, co znamená nenávistný obsah, či povinnost platforem závadný obsah smazat do 24 hodin pod hrozbou velmi vysokých sankcí. Tento zákon byl však posléze francouzským Ústavním tribunálem jako protiústavní zrušen, protože neúměrně narušoval svobodu projevu.⁴⁴³ K tomuto zákonu se navíc velmi kriticky vyjadřovaly jak lidskoprávní organizace, tak například i ČR.⁴⁴⁴ Obecně lze říci, že zákony požadující rychlé mazání obsahu nejsou dobrou variantou, a tato cesta by měla být opuštěna.

Nejvýznamnější přímou regulací je v současnosti Akt o digitálních službách (DSA). Jaký vliv bude tato regulace mít na svobodu projevu, je do budoucna otevřenou otázkou. Jeho celková analýza by byla nad rámec této kapitoly, dovolím si zde však udělat alespoň analýzu lidskoprávní – a to právě prizmatem dopadů této regulace na svobodu projevu. DSA nastavuje režim, který ukládá sociálním sítím povinnost mazat dvojí druh obsahu. Za prvé jde o obsah, který je nezákonný. Za druhé jde o obsah, který je v rozporu se smluvními podmínkami sítě.

Co se týče nezákonného obsahu, již první problematickou pasáží je jeho samotná definice, která je příliš široká. Čl. 3 písm. h) DSA totiž definuje nezákonný obsah jako informaci, která není v souladu: a) s právem unie; nebo b) s právem *některého* členského státu. Právě slovo *některého* zde vyvolává oprávněné obavy, které ani srovnání různých jazykových verzí příliš nezmírňuje.⁴⁴⁵ Nabízí se totiž dva potenciální výklady, přičemž jeden je z pohledu svobody projevu velmi problematický. Tím je, že by byl jakýkoli projev posuzován prizmatem evropského právního řádu s největšími omezeními projevu k určité problematice. Takže byť by šlo o projev legální například dle českého práva, tak pokud by byl nelegální dle práva bulharského, šlo by o nelegální obsah dle DSA a takový by musel být smazán.

Tento výklad, byť jazykově přiléhavý, považuji za nepřijatelný z pohledu ústavně-konformního výkladu právních norem. Teleologicky se proto musí zvolit výklad, že jde vždy pouze o právní řád státu, ve kterém buď k projevu dochází, nebo dle kritéria, na koho je projev cílený. Zda tento výklad převáží, jisté však ani zdaleka není. I v případě této interpretace se

⁴⁴³ KAYALI, L.; BRAUN, E. France pushes tougher EU rules for social media in wake of terror attack. In: *Politico* [online]. 2020. Dostupné z: https://www.politico.eu/article/france-renews-social-media-regulation-push-at-eu-level-in-wake-of-terror-attack/?fbclid=IwAR2YALMayjEciYwy6ak_GHUDKyhjmQvxM4MymDd6tVrtxpHj35CXdO9EfIM.

⁴⁴⁴ KOLAJA, M. Poučení z Avia law. In: *Kolaja.eu* [online]. 2020. [cit. 2024-01-06]. Dostupné na: https://www.kolaja.eu/cs/post/20200512-lessons_learned_from_the_avia_law_warning-french_online_hate_speech_law_to_be_adopted_tomorrow/.

⁴⁴⁵ V angličtině „any“, ve španělštině „cualquier“.

navíc objevují mnohé problémy. Představme si například situaci, kdy slovenský občan studuje v Brně a píše příspěvky v angličtině. Na koho je takový projev cílený? Jak tedy určit rozhodné právo? Tyto otázky DSA neřeší, což je jeho zásadní slabina. V praxi však budou tyto otázky pravděpodobně nadbytečné, neboť racionálním chováním sociální sítě bude všechny potenciálně nelegální projevy v celé EU označit za nepřijatelné dle smluvních podmínek – jde o mnohem jednodušší řešení než mít 27 různých standardů. Avšak jde také o řešení, které výrazně omezí svobodu projevu v celé EU.

Druhým problémem definice nezákonného obsahu je, že se zkoumá soulad s celým právním řádem (a nikoli například jen s trestním právem), což může mít těžko předvídatelné důsledky. Nejde tedy o nesoulad pouze s veřejným právem, nýbrž i s právem soukromým. Je třeba si navíc uvědomit, že protiprávnost projevu může být pro různé osoby rozdílná. Například učitelům může školský zákon ukládat určité specifické povinnosti, které se týkají jejich vyjadřování na sociální síti.⁴⁴⁶ V takovém případě však půjde o projev, který je protiprávní například pouze pro učitele a nikoli pro jiné osoby. Podobné může platit například pro soudce nebo advokáty. Jak si sociální síť s takovou personifikací protiprávnosti poradí, vůbec není zřejmé a osobně jsem k takové právní úpravě velmi kritický.

Co se týče smluvních podmínek, na první pohled by se mohlo zdát, že si obsah těchto podmínek mohou určovat sami provozovatelé, ale tak tomu není. Limitace podoby smluvních podmínek se nachází v čl. 14 DSA, který stanoví povinnost přihlížet k „právům a oprávněným zájmům všech stran“ a to včetně „základních práv příjemců služby“. Pozitivně lze nahlížet na skutečnost, že je zde zmíněna svoboda projevu. Rizikem však je, že v aplikaci tohoto ustanovení se bude dávat přednost jiným zájmům (například ochraně před tzv. projevy z nenávisti).

Podobu smluvních podmínek však budou nejvíce ovlivňovat tzv. systémová rizika (čl. 34). Mezi ně patří například jakékoli „skutečné či předvídatelné nepříznivé dopady na občanský diskurz, volební procesy a veřejnou bezpečnost“, ochrana veřejného zdraví nebo ochrana před závažnými negativními důsledky pro „tělesnou a duševní pohodu osob“. Jsem si jistý, že skoro všechny tyto body lze vykládat způsobem, který je pro svobodu projevu velmi restriktivní a že se tak ve skutečnosti bude dít. V návaznosti na systémová rizika posléze čl. 35 stanoví VVOP a VVIV povinnost zavést „rozumná, přiměřená a účinná zmírňující opatření“,

⁴⁴⁶ Viz zákon č. 561/2004 Sb. o předškolním, základním, středním, vyšším odborném a jiném vzdělávání.

především půjde o úpravu smluvních podmínek a úpravu postupů při moderování obsahu. Za nedodržení posléze hrozí excesivní pokuty ve výši procent z obrátu.

Dalším rizikovým faktorem pro svobodu projevu ze strany DSA je úprava krizových situací. Čl. 36 tak stanoví možnost Evropské komise přikazovat sociální síti různá opatření za účelem ochrany veřejné bezpečnosti a veřejného zdraví. Podobně jsou v čl. 48 upraveny tzv. krizové protokoly, u nichž zase reálně nikdo neví, co znamenají. Z pohledu svobody projevu se však jeví býti velmi podezřelými.

Upozornění hodné jsou také tzv. kodexy chování, které jsou upraveny v čl. 45. Odstavec 1 o nich přitom mluví jako o dobrovolných. O to více je překvapující, že odstavec 2 stanoví, že pokud se objeví významná systémová rizika, může Evropská komise vyzvat dotčené provozovatele ke stanovení závazků přijmout zvláštní opatření ke zmírňování rizik. Na první pohled to vypadá jako legislativní chyba, kdy je v rubrice týkající se dobrovolných kodexů uložena možnost nedobrovolně přikazovat. Výkladově bych tento rozpor překlenul závěrem, že se možnost nedobrovolně přikazovat uplatní pouze na subjekty, které tyto dobrovolné kodexy chování mají.

Tam, kde DSA naopak zcela jistě pomůže, je stanovení postupů ochrany fyzických a právnických osob, neboť toto nařízení zpřesňuje odvolací postup při mazání příspěvků na sítích. Nově tedy bude možnost trojí ochrany. Předně bude možnost podat odvolání v rámci samotné platformy. Posléze bude možnost zvolit ADR k rozhodnutí sporu mezi platformou a uživatelem. A v poslední řadě bude možnost podat klasickou civilní žalobu k soudu. I zde však vyvstávají jisté problémy. Rozhodnutí ADR například nebude závazné, což výrazně limituje efektivitu tohoto nástroje. U soudní příslušnosti není navíc řešena otázka, kde žalovat, která je ponechána obecné evropské regulaci. Z ní přitom vyplývá, že pouze spotřebitelé budou moci žalovat dle svého bydliště. Ostatní budou muset žalovat v místě sídla platformy, což je velmi komplikované a představuje to jednu z největších slabin této regulace.

5.4.3. Platformy jako vydavatelé a otázka soukromoprávní odpovědnosti

Současné postavení sociálních sítí má svůj základ v 90. letech 20. století, kdy byl v USA přijat *Communications Decency Act of 1996*. Tento zákon ve své sekci 230c⁴⁴⁷ vložil do

⁴⁴⁷ „Section 230c of the Communications Decency Act of 1996.

1. No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

amerického právního řádu obecnou imunitu online platforem vůči občanskoprávní odpovědnosti za obsah třetích stran a za určitých podmínek také za mazání obsahu. Smyslem a účelem tohoto ustanovení bylo pomoci rozvoji internetu a internetových společností, které byly v té době zranitelné například vůči klasickým vydavatelům tištěných novin. Klíčovým problémem tohoto ustanovení však je, že pochází z dob úplných začátků internetu, které jsou s dnešním stavem technologie a společnosti zcela nesrovnatelné.⁴⁴⁸ Však se stačí podívat na velikost a moc internetových společností tehdy a nyní.

Podíváme-li se na důsledky tohoto ustanovení, při bližším zkoumání zjistíme, že ochrana sociální platformy spočívá ve dvou klíčových aspektech. Sociální síť dle tohoto ustanovení je a zároveň není v postavení vydavatele, přičemž základním kritériem je, zda je to pro ni zrovna výhodné. Pokud rozhoduje o moderaci obsahu, pak má v podstatě postavení vydavatele, může tedy volně určovat, co na ní bude umístěné a není za smazání obsahu (až na výjimky) odpovědná. Na rozdíl od klasického vydavatele však zároveň není odpovědná ani za obsah, který je nějakým způsobem závadný. Tento stav je problematický, neboť na první pohled neřeší ani jeden z dnešních problémů spojovaných se sociálními sítěmi. Za prvé nenutí sociální síť rychle z vlastního popudu mazat závadný obsah. Za druhé naopak umožňuje sociální síti smazat, cokoli uzná za vhodné. Důsledkem je, že sociální síť může mazat, co chce, a v podstatě nikdy neodpovídá za to, co je na ní umístěno.

Evropský přístup se od toho amerického legislativně trochu liší, ale základní principy regulace zůstávají obdobné, neboť DSA i článek 14 směrnice 2000/31/EC ustanovují systém tzv. *notice and takedown*.⁴⁴⁹ Podstatou tohoto přístupu je, že sociální platforma je povinna jednat až v případě, kdy je na problematický obsah upozorněna; nemá sama o sobě povinnost závadný obsah vyhledávat. Tato směrnice je do českého práva implementována § 5 a 6 zákona č. 480/2004 Sb., o některých službách informační společnosti, které zakotvují tzv. *bezpečný přístup* pro poskytovatele. Také dle českého práva proto v současnosti neexistuje povinnost aktivně vyhledávat protiprávní obsah a povinnost aktivně jednat existuje pouze v případech,

2. No provider or user of an interactive computer service shall be held liable on account of:
A. any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene [etc.]; or
B. any action taken to enable or make available to information content providers or others the technical means to restrict access to material in paragraph 1.“

⁴⁴⁸ FERGUSON, N. What Is To Be Done? Safeguarding Democratic Governance In The Age Of Network Platforms. *Governance In An Emerging New World*, 2018, Fall Series, Issue 318.

⁴⁴⁹ LAVI, M. Content Providers' Secondary Liability: A Social Network Perspective. *Fordham Intellectual Property, Media and Entertainment Law Journal Journal*, 2016, vol. 26, no. 4, s. 855-943.

kdy jsou zprostředkovatelé na protiprávnosti upozorněni.⁴⁵⁰ Tato úprava však není přijímána zcela bez kritiky.⁴⁵¹

Při zodpovězení otázky, jakým směrem právní úpravu do budoucna proměnit, se nabízí dvě základní varianty. Sociální sítě buď: a) budou odpovídat za obsah jako klasičtí vydavatelé, a proto budou moci také nadále moderovat obsah zcela dle svého uvážení; nebo b) nadále nebudou odpovědné za obsah, pak by ale zároveň neměly být těmi, kdo může zcela volně rozhodovat o moderaci obsahu (tzn. dostane se jim postavení infrastruktury).

Dle mého názoru je návrh, že by sociální sítě, když už kontrolují obsah, měly být učiněny zodpovědnými za příspěvky, které se na nich objevují, tzn. že by se měly dostat na stejnou pozici jako klasičtí vydavatelé, velmi nebezpečný a nerealizovatelný zároveň. Nerealizovatelný v tom smyslu, že množství obsahu na sociálních sítích a v klasických médiích je zcela nesrovnatelné. Sociální sítě proto nemohou bez použití rozsáhlé cenzury a často chybujících algoritmů umělé inteligence odpovídat za všechno, co je na sociální síti uživateli umístěno.⁴⁵² Nebezpečný z toho důvodu, že povede sociální sítě k přílišné opatrnosti, kdy budou mazat pro jistotu i příspěvky hraniční či zcela nezávadné, a to jen proto, aby předešly potenciálním sankcím.⁴⁵³ Taková regulace by vedla k omezení veřejné diskuse a svobody projevu na v dnešní době nejvýznamnějších veřejných fórech.⁴⁵⁴

Mimo odpovědnost sociálních platform za obsah nahraný uživateli se nabízí jejich odpovědnost za jiná jednání. Základní problém sociálních platform je totiž jejich *design*, který je nastaven k maximalizaci určitých atributů. Základním cílem algoritmů Facebooku je tak například udržet pozornost uživatele a maximalizovat jeho čas strávený na síti. Skvělý způsob, který umělá inteligence k tomuto účelu používá, je apel na již existující *bias* uživatelů, který posléze může vést k jejich radikalizaci a přispívat k polarizaci celé společnosti.⁴⁵⁵ Cílem této technologie je v podstatě kontrolovat a manipulovat chování uživatele směrem, který je pro něj

⁴⁵⁰ Viz ŠÁMAL, P.; BOHÁČEK, L. "Fake news" jako aktuální právní otázka spojená s globalizací. In: GERLOCH, A.; KRZYŽANKOVÁ, K. Ž. (eds.). *Právo v měnícím se světě*. Plzeň: Aleš Čeněk, 2020, s. 554–585.

⁴⁵¹ Viz HANYCH, M.; PIVODA, M. Facebook, Twitter a YouTube jako garanti svobodného projevu? Kritika současného systému notice-and-takedown. *Revue pro práva a technologie*, 2017, roč. 8, č. 16, s. 177–220.

⁴⁵² FERGUSON, N. What Is To Be Done? Safeguarding Democratic Governance In The Age Of Network Platforms. *Governance In An Emerging New World*, Fall Series, Issue 318. Dostupné z: <https://www.hoover.org/research/what-be-done-safeguarding-democratic-governance-age-network-platforms>.

⁴⁵³ FERGUSON, N. What Is To Be Done?...

⁴⁵⁴ SUNSTEIN, C. R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2017.

⁴⁵⁵ Tamtéž, s. 37–57.

negativní. Jedná se o zcela asymetrický vztah, kdy proti každému jednotlivému uživateli stojí obří korporace s cílem dosahování co největšího zisku.

Důležitou otázkou je, zda by civilní právo mohlo sankcionovat takto škodlivou algoritmičskou personalizaci. Pokud bychom se inspirovali v soutěžním právu, odpověď by se lišila podle toho, zda personalizace odráží skutečné preference zákazníků, nebo působí proti nim.⁴⁵⁶ Tato úvaha vychází z předpokladu, že „jedním z cílů práva hospodářské soutěže je ochrana volby spotřebitele“.⁴⁵⁷ Ve světle doktríny suverenity spotřebitele přitom platí, že pokud algoritmy „neuspokojují individuální preference, ale spíše je utvářejí a manipulují s nimi a působí jako soukromí cenzori zpráv a názorů“⁴⁵⁸, pak alespoň teoreticky lze takové chování považovat za zneužívání dominantního postavení a poškozování spotřebitele.⁴⁵⁹

Civilní soudy by v takovém případě mohly dle mého názoru přisoudit takto poškozeným uživatelům náhradu újmy, byť by dokazování v takovém případě bylo komplikované. Teoreticky by se také mohlo jednat o naplnění generální klauzule nekalé soutěže dle § 2976 odst. 1 OZ. Zároveň by dle mého názoru dávalo smysl, aby sociální síť odpovídala za neoprávněně smazaný obsah, například skrze ustanovení občanského zákoníku o ochraně přirozených práv, kam svoboda projevu zcela jistě spadá.⁴⁶⁰

5.4.4. Horizontální účinek základních práv a pozitivní závazek státu

Svoboda projevu je sice primárně právem negativním – tzn. právem, které garantuje určitou svobodu od zásahů státu. V českém a mezinárodním právu toto ale není jeho jedinou funkcí, neboť díky doktríně prozařování základních práv do zbytku právního řádu může působit do jisté míry také v horizontálních vztazích, tedy ve vztazích dvou soukromých subjektů.⁴⁶¹ Důležitou variantou na poli ústavního práva je proto judikatorně přiznat čl. 17 Listiny horizontální účinek.

Dle Wagnerové platí, že kvůli „vyzařování základních práv do celého právního pořádku mohou být osobou povinnou i soukromé osoby (vlastníci médií)“.⁴⁶² Nabízí se proto důležitá otázka: mohou-li být osobami povinnými vlastníci médií, proč by jimi nemohly být také

⁴⁵⁶ FISH, E.; GAL, M. Echo Chambers and Competition Law: Should Algorithmic Choices be Respected? *Frederic Jenny Liber Amicorum: Standing Up for Convergence and Relevance in Antitrust*, 2020, s. 1.

⁴⁵⁷ Tamtéž, s. 7.

⁴⁵⁸ Tamtéž, s. 7–8.

⁴⁵⁹ Tamtéž, cit. dílo.

⁴⁶⁰ Viz např. § 81 odst. 1 OZ.

⁴⁶¹ WAGNEROVÁ, E. In: *Listina základních práv a svobod: komentář*. Praha: Wolters Kluwer, 2012, s. 424.

⁴⁶² Tamtéž, s. 433.

sociální sítě? Podíváme-li se totiž na funkce a účely svobody projevu v demokratické pluralitní společnosti a zároveň na důležitost sociálních sítí, které v dnešní době plní do značné míry funkci veřejných fór⁴⁶³, nenalezneme mnoho důvodů, proč by pro sociální sítě měla platit odlišná pravidla.

Do popředí zájmu ochrany svobody projevu totiž spadá „*myšlenka zajištění podmínek pro svobodnou, individuální tvorbu názorů ve věcech veřejných*“.⁴⁶⁴ Určitá dimenze svobody projevu se proto týká také toku informací ve společnosti a lze z ní dokonce vyvozovat i povinnosti státu činit opatření, „*kteřá budou předcházet vytváření názorového monopolu*“.⁴⁶⁵

Tento právní názor je možno také podložit například judikaturou ESLP, která dovozuje pozitivní povinnost státu chránit svobodné projevy i ve vztazích soukromých osob, především těch pracovněprávních.⁴⁶⁶ Náš přístup se v tomto výrazně odlišuje od amerického, v jehož rámci se svoboda projevu dotýká pouze ochrany před mocí veřejnou. Pozitivní závazek vyplývající z ochrany svobody projevu byl například judikován německým Ústavním soudem⁴⁶⁷ a mohl by sloužit, v kombinaci s horizontálním působením svobody projevu, jako dobrý inspirační zdroj.

V českém kontextu tomuto nahrává i samotná textace čl. 17 odst. 1 Listiny, která zní: „*Svoboda projevu a právo na informace jsou zaručeny.*“ Na rozdíl od amerického prvního dodatku se tu nemluví pouze o tom, že by stát neměl do této svobody zasahovat, nýbrž se normativně stanovuje, že se svoboda projevu zaručuje. Z tohoto lze poměrně snadno dovodit jak povinnost státu svobodu projevu chránit i na soukromých platformách s faktickou rolí veřejného fóra, tak horizontální účinek čl. 17 mezi sociální platformou a jejím uživatelem.

5.5 Boj s dezinformacemi

Stále silícím tématem online prostoru je snaha bojovat s dezinformacemi. Spolu se stupňujícím se tlakem na regulaci sociálních sítí může další vývoj v této oblasti ovlivnit způsob přijímání a šíření informací milionů lidí. S tím je pro svobodu projevu opět spojeno velké množství rizik.

⁴⁶³ SUNSTEIN, C. R. *#Republic: Divided Democracy in the Age of Social Media...*, s. 37–57.

⁴⁶⁴ WAGNEROVÁ, E. In: *Listina základních práv a svobod...*, s. 429.

⁴⁶⁵ Tamtéž.

⁴⁶⁶ Rozsudek ESLP ze dne 21. 7. 2011, Heinisch proti Německu, stížnost č. 28274/08.

⁴⁶⁷ Rozhodnutí německého Ústavního soudu 1 BvL 89/78.

Jak dokládá analýza *Dezinformace a svoboda projevu v čase války na Ukrajině*⁴⁶⁸, které je autor této práce spoluautorem, pojem dezinformace se v českém kontextu fakticky využívá ve dvou vzájemně si odporujících významech: v *užším* a *širším* pojetí. Kritérii dezinformace v užším pojetí jsou následující znaky: 1) nepravdivost šířeného tvrzení; 2) úmysl klamat; 3) systematická a organizovaná snaha; 4) obzvláště ze strany státních aktérů či jiných skupin. Toto užší pojetí se ve veřejném diskurzu vyskytuje velmi vzácně.

Širší pojetí naproti tomu vnímá dezinformaci spíše v kontextu *narativu* – znakem „správné“ dezinformace tedy není nutně nepravdivost, ale nepřátelský, protidemokratický narativ. V tomto pojetí tedy dezinformací může být cokoli, co zahrnuje jiný pohled na svět, jinou interpretaci události, propagandu, konspirační teorie atd. Základní problém této definice je, že vede k definici kruhem: „*dezinformační weby jsou ty, které šíří dezinformační narativy, přičemž dezinformační narativ je to, co je šířeno dezinformačními weby*“.⁴⁶⁹ Oboje je přitom určeno velmi arbitrárně. Toto širší pojetí je ve společenském diskurzu v současnosti mnohem rozšířenější, proto se mu zde budu dále věnovat.

Zamyslíme-li se totiž vůbec nad samotným konceptem *dezinformačního narativu*, ukáže se úskalí tohoto pojmu v celé šíři. Například dle neziskové organizace Čeští elfové jde o „*určité historky, verze, způsoby interpretace událostí*“.⁴⁷⁰ Lze si stěží představit neurčitější a víc nic neříkající popis dezinformace. Pro užití neziskovou organizací je tato definice ještě akceptovatelná, měla-li by se však takto využívat v právním řádu, byl by to konec svobody projevu, jak ji dosud známe.⁴⁷¹ Rozlišení faktu a narativu, který již z definice nemůže být nepravdivý, se přitom veřejně příliš nerozebírá, neboť jak ukazuje Chlup: „*Narativ bereme za pravdivý, pokud s námi rezonuje a máme z něj pocit pravdivosti*“.⁴⁷² Intuitivně tedy příliš neodlišujeme fakta a interpretace, nýbrž i narativy označujeme buď za pravdivé, či nepravdivé – avšak ve skutečnosti jiným kritériem.

⁴⁶⁸ RŮŽIČKA, A.; HOŘEŇOVŠKÝ, J.; KUŽÍLEK, O. *Dezinformace a svoboda projevu v čase války na Ukrajině*. In: *Svoboda slova online* [online]. 4. 3. 2022. Dostupné na: <http://svobodaslova.online/wp-content/uploads/2022/03/Dezinformace-a-svoboda-projevu-Ukrajina.pdf>.

⁴⁶⁹ Tamtéž.

⁴⁷⁰ ČEŠTÍ ELFOVÉ. FAQ: mýty & fakta o elfech. 2022 [cit. 2024-01-06]. Dostupné z: <https://cesti-elfove.cz/uvodni-strana/faq-myty-fakta-o-elfech/>.

⁴⁷¹ K problematice používání slova „dezinformace“ v právní regulaci, také viz SCHEU, H. CH. Několik úvah o unijním pojetí dezinformace. *Jurisprudence*, 2023, číslo 5, s. 1–10.

⁴⁷² CHLUP, R. Úskalí boje proti dezinformacím ve věku autenticity. In: *blog.aktualne.cz* [online]. 05. 02. 2023. Dostupné na: <https://blog.aktualne.cz/blogy/radek-chlup.php?itemid=44437>.

Dělení faktu a narativu přitom v mnohém připomíná odlišování faktických tvrzení a hodnotových soudů v judikatuře Ústavního soudu.⁴⁷³ Ten konstatoval, že „*existenci faktů lze prokázat, leč pravdivost jejich hodnocení podrobit důkazům nelze. Ve vztahu k hodnotícím soudům nelze tedy požadavek, dokázat jejich pravdivost, naplnit a takový požadavek samotnou svobodu názoru porušuje*“.⁴⁷⁴ Domnívám se, že tentýž závěr musíme aplikovat také na narativy, tzn. že narativy podléhají stejné míře ochrany jako hodnotové soudy a uplatní se na ně stejná pravidla judikovaná Ústavním soudem a ESLP.⁴⁷⁵

Rizikem totiž je, že bez úzkých a přesně vymezených kritérií se za „dezinformační narativ“ může označit také interpretace, která je podložená a legitimní. Průzkum Ministerstva vnitra například označoval za dezinformaci tvrzení, že Evropská unie nařizuje Česku, co má dělat: „*Například s dezinformací, že Evropská unie diktuje Česku, co má dělat, souhlasí 70 procent lidí*“.⁴⁷⁶ Toto tvrzení je ukázkou, jak lze označit za „dezinformaci“ v širším pojetí v podstatě cokoliv. Podstatou členství v EU a nadřazenosti evropského práva je, že EU nařizuje v mnoha oblastech České republiky, co má dělat.⁴⁷⁷ Zmíněné „dezinformační“ tvrzení je tak ve své podstatě dokonce konstatováním faktu. Označení podobných tvrzení za dezinformaci je založeno až na předpokladu o širším vnímání světa toho, kdo s tvrzením souhlasil. Jde tedy o domyšlenou interpretaci interpretace. Podobně tuto problematiku vystihl Chlup, když napsal:

*„Velká část ‚dezinformačních narativů‘ ostatně ani s žádnými vyvratitelnými fakty neoperuje a jde spíše o antisystémové výklady naší složité společenské situace. Jestli ‚EU omezuje naši suverenitu‘ nebo jestli ‚dodávkami zbraní na Ukrajinu jen přiléváme olej do ohně‘, je věcí interpretace, nikoli faktického ověření. Sami bojovníci proti dezinformacím většinou nejsou neutrálními fact-checkery, ale spíše ideologickými narrative-checkery, kteří hájí prozápadní a promainstreamové narativy nezávisle na jejich faktičnosti.“*⁴⁷⁸

⁴⁷³ Viz např. náleží sp. zn. IV. ÚS 146/04 ze dne 4. 4. 2005 nebo náleží sp. zn. I. ÚS 367/03 ze dne 15. 3. 2005.

⁴⁷⁴ Náleží sp. zn. I. ÚS 156/99 ze dne 8. 2. 2000.

⁴⁷⁵ K podrobnostem viz např. BARTOŇ, M. *Svoboda projevu: principy, garance, meze*. Praha: Leges, 2010.

⁴⁷⁶ Originální dokument byl z internetu smazán. Informace o této politice je zachycena v RŮŽIČKA, A.; HOŘEŇOVŠKÝ, J.; KUŽÍLEK, O. *Dezinformace a svoboda projevu v čase války na Ukrajině*. In: *Svobodaslova.online* [online]. 4. 3. 2022. Dostupné na: <http://svobodaslova.online/wp-content/uploads/2022/03/Dezinformace-a-svoboda-projevu-Ukrajina.pdf>.

⁴⁷⁷ Viz rozhodnutí Evropského soudního dvora N. V. *Algemene Transport-en Expeditie Onderneming Van Gend en Loos v. Netherlands Fiscal Administration (26/62)* a především *Flaminio Costa vs. ENEL (číslo 6/64)*.

⁴⁷⁸ CHLUP, R. *Úskalí boje proti dezinformacím ve věku autenticity...*

Podíváme-li se na věc z právního úhlu pohledu, tyto znaky činí z dezinformace v širším slova smyslu pojem zcela neuchopitelný a nevhodný pro jakékoli legislativní užití. Zkusme si například představit trestný čin, jehož podstatou by bylo „šíření dezinformací“. Tento nástroj by se velmi rychle stal zcela neurčitým bičem, který by se mohl obrátit proti komukoli, kdo by ve společnosti vyjádřil „anti-mainstreamový“ narativ, tedy alternativní interpretaci v podstatě jakéhokoli tématu. Kdyby byl za dezinformační narativ veřejnou mocí opět označen například postoj, že „existuje středoevropská identita“, neboť tento narativ narušuje naše „prozápadni směrování“,⁴⁷⁹ mohli bychom najednou trestně stíhat kohokoli, kdo by tuto myšlenku veřejně vyjádřil.

Skutečnost, že většina dezinformací se netýká faktů, nýbrž narativů, je dobře dokumentovatelná i argumentací zastánců nutnosti nových nástrojů pro zastavení jejich šíření. Kdyby šlo pouze o fakta, zcela jistě by stačil trestný čin „šíření poplašné zprávy“.⁴⁸⁰ Ten však dnes nelze na dezinformace většinou použít, protože podmínkou jeho aplikace je dokázání *faktické nepravdivosti*, což je v případě jiného narativu poměrně těžké prokázat. Toto ustanovení navíc – i díky poměrně striktnímu doktrinálnímu a judikaturnímu výkladu – neumožňuje trestat neznalost, nedbalost, nedostatečné ověřování zdrojů apod., ale postihuje jen a pouze šíření zprávy, o níž pachatel ví, že se nezakládá na pravdě, a zároveň tato zpráva musí být způsobila zapříčinit nebezpečí vážného znepokojení alespoň části obyvatelstva nějakého území, tj. nemůže postihovat zprávy vyložené hloupě či zcestně, protože tím nelze „průměrného obyvatele“ vážně znepokojit.⁴⁸¹

Nakonec se můžeme dostat i do situace, kdy bude možno trestněprávně relevantně „šířit dezinformace pravdou“. Není náhodou, že se Ministerstvo vnitra bránilo vložit do návrhu zákona *o omezení šíření obsahu ohrožujícího národní bezpečnost online ochranu fakticky pravdivých tvrzení*.⁴⁸² Všechny tyto snahy jsou dle mého mínění hnány především *morální panikou*, která však není založena na datech či konkrétních informacích. Tuto rovinu boje s

⁴⁷⁹ Viz MVČR. Středoevropská identita jako součást protizápadního narativu na kvazi-mediální scéně. Nedatováno. [cit. 2024-01-06]. Dostupné na: <https://www.mvcr.cz/chh/clanek/stredoevropska-identita-jako-soucast-protizapadniho-narativu-na-kvazi-medialni-scene.aspx>.

⁴⁸⁰ Viz § 357 zákona č. 40/2009 Sb., trestního zákoníku.

⁴⁸¹ Viz RÁLIŠ, P. Podpora a propagace terorismu – kladivo na internetové diskutéry? *Trestněprávní revue*, 2024, roč. 23, č. 1. [pozn. strany nelze určit, neboť v době odevzdání práce toto číslo ještě nevyšlo].

⁴⁸² Osobní zkušenost autora. Návrh zákona je k dispozici zde: <https://www.mvcr.cz/clanek/poskytnuti-informace-zakon-dezinformace.aspx>.

dezinformacemi výborně dokládá článek Altaye.⁴⁸³ Ten velmi přesvědčivě popisuje problémy, které současný narativ o „době dezinformační“ obsahuje.

Dle této studie například není vůbec prokázáno: 1) že by dezinformace měly silný vliv na lidské chování; 2) že by dezinformacím podlehl velké množství lidí; 3) že lidé věří všemu, co je obsaženo na internetu; 4) že se lež šíří rychleji než pravda; 5) že je internet zahlcen dezinformacemi; či 6) že jsou dezinformace jen problémem sociálních sítí.⁴⁸⁴ Studie také dokazují, že dezinformace ve skutečnosti šíří jen velmi malé množství lidí⁴⁸⁵ a že to mnohdy má čistě politické motivace.⁴⁸⁶ Nejde pouze o posun v zahraničí, podobně se diskurz o dezinformacích posouvá (dospívá) i v rámci některých českých konferencí.⁴⁸⁷

Ve zkratce řečeno, veřejný prostor je v současnosti silně zahlcen *alarmistickými narativy* o dezinformacích, a to navzdory přímým důkazům o tom, že jejich výskyt i vliv jsou silně nadhodnoceny.⁴⁸⁸ Tyto narativy by proto měly být chápány především jako morální panika.⁴⁸⁹ Reálným problémem totiž nejsou primárně dezinformace, nýbrž především snižující se důvěra ve veřejné instituce.⁴⁹⁰ Za tu však nemůže šíření dezinformací, nýbrž především samotné chování těchto institucí vůči společnosti.

Dezinformace totiž ve společnosti působí podobně jako terorismus – samy o sobě nejsou zase tolik nebezpečné, ale vyvolávají silnou společenskou reakci založenou na strachu, která může být neadekvátní čelené hrozbě a která může podkopávat klíčová práva a svobody pro fungující demokratický právní stát. Svoboda projevu je z tohoto pohledu „právem v ohrožení“, neboť může být viděna také jako překážka lepšího či dokonalého světa (například „světa bez dezinformací“). Svět bez svobody projevu ale lepším místem rozhodně není. V tomto kontextu

⁴⁸³ ALTAY, S.; BERRICHE, M.; ACERBI, A. Misinformation on Misinformation: Conceptual and Methodological Challenges. *Social Media + Society. Special Issue: Multidisciplinary approaches to mis- and disinformation studies*, 2023, s. 1–13.

⁴⁸⁴ Tamtéž.

⁴⁸⁵ NOGARA, H. et al. The Disinformation Dozen: An Exploratory Analysis of Covid-19 Disinformation Proliferation on Twitter. *WebSci '22: Proceedings of the 14th ACM Web Science Conference*, 2022, s. 348–358.

⁴⁸⁶ MELCHIOR, C.; OLIVEIRA, M. A systematic literature review of the motivations to share fake news on social media platforms and how to fight them. *New Media & Society*, 0(0), 2023.

⁴⁸⁷ Viz např. konference organizace Iuridicum Remedium ze dne 1. 12. 2023 *Digitální doba s lidskou tváří: symposium o digitálních výzvách*. Záznam k dispozici zde: https://www.youtube.com/live/N1WcyOymnYM?si=x5mMhZ_Jdo7tEWgR.

⁴⁸⁸ ALTAY, S.; BERRICHE, M.; ACERBI, A. Misinformation on Misinformation: Conceptual and Methodological Challenges. *Social Media + Society. Special Issue: Multidisciplinary approaches to mis- and disinformation studies*, 2023.

⁴⁸⁹ Tamtéž.

⁴⁹⁰ ALTAY, S. Rethinking the Problem of Misinformation. [přednáška]. Praha: Institute H21, 15. 9. 2023. Dostupné na: <https://youtu.be/Yms5DfMYQTw?si=IZcBNphvCengeluz>.

je proto nutné více upozorňovat na benefity, které otevřená veřejná diskuse demokratické společnosti přináší,⁴⁹¹ a na alternativní způsoby, jak s dezinformacemi bojovat.

Mezi potenciální alternativní způsoby boje s dezinformacemi ze strany samotných sociálních sítí přitom patří například povinnost identifikovat původ příspěvků nebo účtů, povinnost identifikovat neautentické účty, povinnost zpřístupnit anonymizované údaje o činnosti nezávislým výzkumným pracovníkům, vytvoření veřejné iniciativy pro mediální gramotnost nebo zajištění ochrany před zahraniční manipulací.⁴⁹² Velmi často zmiňovaným konceptem je v současnosti také *protiřeč* (counterspeech), a to i ve vztahu k regulaci.⁴⁹³ Nezisková organizace Rekonstrukce státu zase velmi dobře popsala žádoucí postup veřejné moci v boji s problematickým obsahem ve vztahu ke koronavirové krizi.⁴⁹⁴

Pokud by se postupovalo jinak a veřejná moc by získala rozsáhlé cenzorské pravomoci, hrozilo by reálné nebezpečí, že by se do společnosti mohly vrátit problémy, které měly zůstat již navěky minulostí. Například, že by veřejná moc mohla opět určovat jednu oficiální pravdu a trestat disent. Pokud bychom mazání závadného obsahu naopak nechali zcela na libovůli platform, hrozilo by naopak riziko opačné, tedy, že by zase soukromé platformy získaly absolutní právo určovat, co na nich může být vyřčeno, čímž by se zase ony staly historicky bezprecedentním dveřníkem informací.

Obě varianty jsou z pohledu otevřené diskuse v demokratickém právním státě silně rizikové. Platformy by proto měly být volány k dodržování standardu ochrany svobody projevu, a to v podobném rozsahu, jak je zajištěn právní úpravou ve vztahu ke státu, neboť původní účel této ochrany se v dnešní době týká i sociálních sítí. Tento standard by měl být přitom kontrolován především soudy.

Veřejná moc nemůže, až na zcela konkrétně definované výjimky, ve jménu boje s dezinformacemi cenzurovat ani využívat neformální nátlak na mazání obsahu. Je totiž důležité

⁴⁹¹ Viz např. PIÑOS, J. Genealogie filosofického ospravedlnění svobody projevu. *Jurisprudence*, 2017, roč. 26, č. 2, s. 24–35.

⁴⁹² FERGUSON, N. What Is To Be Done?...

⁴⁹³ MCCHANGAMA, J.; ALKIVIADOU, N. Reimaging the Current Regulatory Framework to Online Hate Speech. Why Making Way for Alternative Methods Is Paramount for Free Speech. In: ULLMANN, S.; TOMALIN, M. (eds.). *Counterspeech. Multidisciplinary Perspectives on Countering Dangerous Speech*. New York: Routledge, 2023, ch. 5.

⁴⁹⁴ REKONSTRUKCE STÁTU. Doporučení Rekonstrukce státu: Ochrana svobody projevu v době pandemie. 2020. Dostupné:

<https://rekonstrukcestatu.s3.amazonaws.com/rekonstrukcestatu/production/files/2020/05/15/13/11/42/ff26ab4e-b44a-46df-8c6e-41f12fa7df4b/doporučení-rekonstrukce-statu-ochrana-svobody-projevu-v-dobe-pandemie.pdf>

zdůraznit, že drtivá většina tzv. dezinformací je zcela legální a neporušuje žádná ustanovení právního řádu. Pouze velmi malá výše z nich, která by mohla představovat trestný čin či správní přestupek, je mazatelná na popud orgánů veřejné moci dle ustanovení trestního řádu. Na potírání této kategorie však nové právní mechanismy nepotřebujeme.

Jak navíc ukazuje Harari, problém dezinformací rozhodně není nikterak nový, naopak.⁴⁹⁵ Když se proto sami sebe zeptáme, jestli někdy byla situace ohledně nepravd ve veřejném prostoru lepší, pravděpodobně se moc daleko nedostaneme. Byl svět pravdivější v 90. letech? V 80. letech? Či snad ve 40. letech? Harari dokonce konstatuje: „*Ve skutečnosti lidé vždy žili v době postfaktické.*“⁴⁹⁶ A i přesto historicky nalezneme země, které se rozhodly svobodu projevu chránit, přestože by se jistě našly momenty, kdy to nemuselo být, kupeckými počty vzato, příliš výhodné. Jsem přesvědčen o tom, že bychom se do této skupiny měli jako Česká republika, i vzhledem k naší minulosti omezování svobody projevu, zařadit.

Do značné míry proto bude, chceme-li zachovat svobodnou společnost, jediným řešením význam dezinformací nepřeceňovat. Ano, v současnosti je jejich šíření efektivnějším než kdy předtím a cílené útoky např. ze strany cizích vlád mohou ohrozit fungování demokratických zřízení. Tyto úmyslné snahy oslabit demokracie skrze cílené dezinformování by však měly být jasně definovány a neměly by být zaměřovány se šířením nepravdivých zpráv z řad běžných uživatelů, již zprávám upřímně věří. Dobrým řešením může být také fact-checking, a to za podmínky, že nevede k mazání obsahu nebo snižování dosahu, ale jeho účelem je výlučně informovat o faktické pravdivosti předkládaných tvrzení.

Kdyby byl totiž vzat boj s dezinformacemi jako novodobá svatá válka, jejímž cílem by byl absolutně „pravdivý“ a „čistý“ svět, musela by se *ad absurdum* zakázat také náboženství, národní pověsti nebo například i Jára Cimrman. Touha po absolutní čistotě a pravdivosti veřejného prostoru je utopistickým sněním, které může potenciálně vést ke katastrofálním následkům.

5.6 Problém zužujícího se mainstreamu

Nyní se odvrátíme od roviny právní a podívejme se na společenskou rovinu změn ve veřejném informačním prostoru. Pro fungující svobodnou demokratickou společnost je právě

⁴⁹⁵ HARARI, Y, N. *21 Lessons for the 21st Century*. Vintage: Vintage Books, 2018, s. 231–244.

⁴⁹⁶ Tamtéž, s. 233.

ta zcela zásadní. Výchozí myšlenkou této části, která je převzatá od Chlupa,⁴⁹⁷ je, že zdravá demokracie potřebuje mainstream, který lze definovat jako „široké pásmo názorových proudů, které jsou všechny pokládány za legitimní součást demokratické debaty“.⁴⁹⁸ Zdravá demokracie by měla jednoduše připomínat mírné pásmo, které umožňuje „mírumilovné pluralitní soužití různých narativů“.⁴⁹⁹

Prvním kritériem existence takového prostředí je, že pásmo legitimních mainstreamových proudů musí být dostatečně široké, aby se s ním mohla identifikovat podstatná část společnosti.⁵⁰⁰ Důvodem aspirace na širokost mainstreamu je skutečnost, že s každým názorem, který je vyloučen, dochází k potenciálnímu vyčlenění jeho nositele do antisystému.⁵⁰¹ Druhým kritériem je, že by mainstream měl mít rozostřené hranice, protože: „Díky tomu dokáže integrovat i lidi, kteří jsou sice v nějakém ohledu ke společenskému řádu velmi kritičtí, ale stále se s ním dokáží identifikovat alespoň v nějakém ohledu. Tito lidé pak stojí v mainstreamu takřkajíc jednou nohou – ale i to je důležité.“⁵⁰²

Současným neblahým trendem dle Chlupa je, že se takovýto mainstream neustále zužuje: uprchlická krize, pandemie i válka na Ukrajině vedly k tomu, že se tyto dva aspekty širokého mainstreamu začaly vytrácet. Mainstream se zúžil a jeho hranice se zostřily.⁵⁰³ To na jedné straně podporuje spory a nevráživost ve společnosti, kdy se podstatná část populace najednou ocitne za vykolíkovými hranicemi.⁵⁰⁴ Na straně druhé jde o trendy, které se posléze přelévají také do právní roviny, kdy se mainstream pokouší anti-mainstream autoritativně omezit, například skrze anti-dezinformační boj. Pokud bychom však hledali jednoho konkrétního viníka těchto trendů, samozřejmě bychom takového nenašli. Jde o sebeorganizující se proces mnoha jednotlivců, který půjde zvrátit pouze významnou změnou společenského vnímání veřejného prostoru.

Během pandemie se spousta lidí najednou ocitla za hranicemi mainstreamu například kvůli otázce očkování dětí. Dělicí linie, kdy člověk stál v mainstreamu a kdy ne, byla najednou jasně vykolíkováná a jakákoli kritika mainstreamu vedla k silné protireakci. Neblahým

⁴⁹⁷ Viz také CHLUP, R. The Problem of Tackling Disinformation in an Age of Authenticity. [přednáška]. Praha: Institute H21, 15. 9. 2023. Dostupné na: <https://youtu.be/s12I8SWQn2M?si=SIU4JS2IQ-q1c73I>.

⁴⁹⁸ CHLUP, R. Úskalí boje proti dezinformacím ve věku autenticity...

⁴⁹⁹ Tamtéž.

⁵⁰⁰ Tamtéž.

⁵⁰¹ Tamtéž.

⁵⁰² Tamtéž.

⁵⁰³ Tamtéž.

⁵⁰⁴ Tamtéž.

důsledkem tohoto stavu bylo, že vytlačil velké množství lidí do antisystému, což narušilo jejich důvěru ve veřejné instituce, média i odborníky.⁵⁰⁵ Je-li přitom základní problém dneška ztráta důvěry, jde o jev extrémně negativní, který ze své podstaty nemá autoritativní, mocenské, řešení. Nelze totiž použít sílu k získání důvěry.

Důležitou roli v tomto jevu sehrála také měnící se role vědy ve společnosti a ve veřejném prostoru, kdy se z ní pro velkou část mainstreamu stal referenční bod pravdy. Souhlasím s Thielem, že toto věda nikdy neměla dopustit, neboť k tomu ve skutečnosti není vybavená. Aby mohla dobře fungovat, musí ve svých procesech obsahovat zhruba 80 % kritičnosti a 20 % dogmatiky.⁵⁰⁶ Pokud však vědu a vědce postavíme ve veřejném prostoru na piedestal pravdy, jako se tomu stalo během pandemie Covidu-19, pak se tento poměr otočí.⁵⁰⁷ Z prostředí ze své podstaty kritického a neustále pochybovačného se stane prostor s velmi silným tlakem na konformitu, který dissent a kritiku umlčuje. Tento obrat je však pro vědu v dlouhodobém horizontu likvidační, neboť popírá její kritickou podstatu. Pokud se po pár měsících navíc ukáže, že se „věda mýlila“, vede to k dalšímu prohlubování nedůvěry k formálním institucím a k odklonu dalších lidí do anti-mainstreamu.

Co s tím? K odpovědi si na závěr vypůjčím delší citaci Chlupa, který směr vytyčuje poměrně exaktně:

„Řešením [...] není mainstream ohraničenější, nýbrž odolnější. Potřebujeme mainstream, který se občasnou invazivností narativů nenechá vykolejit a dokáže s mnohými z nich i tak vycházet. Na hranicích mainstreamu nejen že není žádoucí stavět ploty, ale potřebujeme mít naopak mnohem širší ono rozostřené hraniční pásmo, které lidem umožňuje v mainstreamu stát jen jednou nohou. Potřebujeme umět integrovat ještě širší spektrum ještě nezvyklejších názorů než doposud. V džunglovitém ekosystému se totiž bujení nových a nových nečekaných narativů nedá vyhnout. A pokud je nedokážeme alespoň trochu začlenit, budeme stále znovu jejich nositele vytlačovat do antisystému, a tím je ztrácet pro jakoukoli konstruktivní společenskou práci. Jinými slovy, potřebujeme svůj narativní ekosystém budovat na inkluzivním základě. Tato inkluze ale musí být dialogická: nemůže spočívat v jednosměrné dominanci

⁵⁰⁵ Tamtéž.

⁵⁰⁶ THIEL, P. *The 2021 Atlas Society Gala: Peter Thiel's Speech*. 2021. Dostupné na: <https://www.youtube.com/watch?v=YK3Tzx-S264&t=177s>.

⁵⁰⁷ Tamtéž.

*oficiálních narativů, nýbrž by měla alternativním narativům přiznávat jejich svébytnost a hledat cesty k soužití a debatování s nimi.*⁵⁰⁸

5.7 Závěr

Na závěr této kapitoly bych zopakoval základní teze. Jakákoli budoucí regulace by měla především zohledňovat zachování svobody projevu v co nejširší možné podobě a přistoupit k omezování pouze tehdy, je-li to absolutně nezbytné. Regulačním cílem ve vztahu k sociálním sítím by přitom mělo být zajištění stavu, aby svoboda projevu byla chráněna i na nich. Vhodným prostředkem je propojení regulace sociálních sítí se současným, ať již americkým či evropským, lidskoprávním standardem, který představuje existující souhrn pravidel určující, co se ve veřejném prostoru smí a nesmí říkat. Sociální sítě by naopak neměly být činěny odpovědnými za obsah, který se na nich nachází. Je také třeba se, pokud možno, vyvarovat restriktivního a kriminalizujícího přístupu v boji s dezinformacemi. Místo toho je vhodné zvolit alternativní prostředky čelení těmto negativním společenským jevům. Jedině tak může svoboda projevu přežít i ve 20. letech 21. století a ještě déle.

⁵⁰⁸ CHLUP, R. Úskalí boje proti dezinformacím ve věku autenticity...

6. Technologie, právo a Covid-19⁵⁰⁹

Tato poslední kapitola se věnuje užití technologií při boji s pandemií onemocnění Covid-19 a jejich právním implikacím, přičemž je ve svém základu rozdělena na tři části. První část se věnuje diskusi týkající se elektronického sledování osob, které vyvolává otázky především ohledně ochrany soukromí a osobních údajů. Tématem druhé části je tzv. covid pas a důsledky z něj vyplývající pro otázku diskriminace. Třetí část se věnuje možnostem užití umělé inteligence v boji s epidemií, kterou český stát v podstatě nevyužil, i když by to v mnoha situacích velmi pomohlo.

Obecně jde v této kapitole o otázku, kde jsou (*de lege lata*) a kde mají být (*de lege ferenda*) limity veřejné moci při užití technologií v boji s epidemií a jakou roli zde sehrává problém diskriminace a požadavek na zachování lidských práv při mimořádných situacích. Současně je nezbytné upozornit na míru emotivnosti, kterou některé tyto otázky, spolu například s návrhem povinného očkování, po celou dobu pandemie vyvolávaly. Je zároveň dobré připomenout, že málokterá kontroverzní otázka měla na začátku jasné právní řešení. Právo totiž v mnoha případech nabízí jistotu až *ex post*, tedy poté, co o věci rozhodnou soudy. A i posléze se může ukázat, že soudy rozhodly špatně. Nejde však nutně o chybu právního systému či justice, ale o inherentní vlastnost současného systému práva způsobenou vysokou mírou komplexity a interpretační rozmanitostí.

Tato kapitola se pokouší od těchto emotivních přístupů oprostit, neboť jejím účelem není ovlivňovat přístup k nedávné epidemii, nýbrž především zprostředkovat zkušenost užitečnou pro zvládání epidemií budoucích. V tomto kontextu se proto kapitola pokouší podat spíše obecný přehled, v jehož rámci zdůrazňuje nejrelevantnější podněty.

6.1 Debata o elektronickém sledování osob

Prvním důležitým tematickým okruhem je debata o elektronickém sledování osob. Vytknutím před závorku bych uvedl, že se v této části zaměřuji na tuto problematiku výhradně ve státech, které lze označit za demokratické. Vynechávám proto užití sledovacích technologií například v Číně, která nemoci využila pro utvrzení svého autoritářského způsobu vládnutí.⁵¹⁰

⁵⁰⁹ Tato část vychází z HOŘEŇOVSKÝ, J. Právní a společenské aspekty užití moderních technologií v boji s epidemií covid-19. In: KYSELA, J.; ŘEPA, K. (eds.). *Reakce právního systému České republiky na covidovou pandemii*. Praha: Leges, 2022, s. 193–203.

⁵¹⁰ K tomuto viz např. STRITTMATTER, K. *We Have Been Harmonized: Life in China's Surveillance State*. Custom House, 2020, s. 368.

Současný stav technologického pokroku spočívající v existenci a velmi široké rozšířené chytrých telefonů se sledovacími technologiemi dal vládám jednotlivých států možnost bojovat s epidemií skrze elektronické sledování občanů a trasování šíření viru. Trasování viru probíhalo v demokratických státech dvojím způsobem. Za prvé, skrze *dobrovolné aplikace*, které umožnily identifikovat potenciálně nebezpečné kontakty například přes funkci Bluetooth nebo GPS.⁵¹¹ Za druhé, šlo o systémy trasování kontaktů skrze *hromadné analyzování poloh mobilních telefonů* získaných od mobilních operátorů.⁵¹² Tento druhý způsob se posléze ještě dělil podle dobrovolnosti neboli dle existence souhlasu vlastníků sledovaných mobilních zařízení.

V základním schematickým kontextu byl proti sobě v těchto situacích postaven *veřejný zájem* na zabránění šíření viru a základní právo na *ochranu soukromí a osobních údajů*. Je zde navíc nutno upozornit na pluralitu pramenů práva na soukromí, které lze nalézt v čl. 10 a 7 Listiny, čl. 7 Listiny základních práv EU a v čl. 8 EULP. V českém kontextu je navíc třeba vnímat také roli nařízení GDPR a zákona č. 110/2019 Sb., o zpracování osobních údajů.

6.2 Dobrovolné elektronické sledování osob

Podstatou dobrovolných sledovacích aplikací byla možnost daná uživatelům nainstalovat si aplikaci, která by je upozornila na potenciální kontakt s nakaženým. Aby byla aplikace účinná, bylo nutno, aby jak osoba nakažená, tak osoba notifikovaná, měly aplikaci nainstalovanou a v době kontaktu funkční. Do této kategorie spadala česká aplikace *eRouška* a *eRouška 2.0*. Podobné aplikace existovaly skoro ve všech zemích světa.

Z pohledu GDPR, které v čl. 6 vyžaduje pro zákonnost zpracování osobních údajů naplnění alespoň jednoho z uvedených kritérií, šlo o bezproblémové aplikace, neboť pracovaly s výslovným souhlasem uživatele.⁵¹³ Souhlas uživatele je totiž dle čl. 6 písm. a) GDPR jedním z důvodů možnosti zpracování. Kvůli existenci souhlasu uživatele nemá v kontextu dobrovolných aplikací dle mého názoru valný smysl rozebírat lidskoprávní rovinu, neboť nikdo nebyl nucen tyto aplikace využívat.

⁵¹¹ AHMED, N.; MICHELIN, R. A.; XUE, W.; RUJ, S.; MALANEY, R.; KANHERE, S. S.; SENEVIRATNE, A.; HU, W.; JANICKE, H.; JHA, S. K. *A survey of covid-19 contact tracing apps*. IEEE Access, *arXiv:2006.10306v3*, 2020.

⁵¹² JALABNEH, R. et al. Use of Mobile Phone Apps for Contact Tracing to Control the COVID-19 Pandemic: A Literature Review. *SSRN*, 2020.

⁵¹³ Viz oficiální stránka aplikace <https://erouska.cz/audit-kod> [navštíveno dne 21. 2. 2022].

I tak se však tyto dobrovolné sledovací aplikace zcela nevyhnuly kritice ze strany odborné veřejnosti a občanské společnosti. Jedním z problematických aspektů *e-Roušky* totiž bylo předávání dat do cloudového úložiště umístěného v USA, konkrétně šlo o úložiště společnosti Google.⁵¹⁴ Kritizována zde byla především z důvodu, že existovaly silné pochybnosti o efektivní ochraně dat evropských občanů v USA.

Obecně lze však konstatovat, že hlavním problémem těchto aplikací napříč demokratickými společnostmi byla jejich malá využívanost, která byla tak výrazná, že mnohdy vedla k nefunkčnosti těchto řešení. Otázka ochrany soukromí a malá využívanost spolu přímo souvisely, neboť pochyby o zajištění soukromí vedly k větší resistenci lidí vůči instalování těchto aplikací.⁵¹⁵ Nebylo zde přitom nezbytné, aby tyto aplikace do soukromí reálně zasahovaly, nýbrž stačily i obavy lidí z toho, že by se to mohlo stát.⁵¹⁶ Podobně fungovala také například role důvěry lidí ve veřejné zdravotnické autority.⁵¹⁷ Státy, které byly schopny pracovat s důvěrou obyvatelstva, měly před sebou mnohem snazší úkol než státy, které toho schopny nebyly.

Z tohoto vyvěrá zajímavá normativní implikace, kterou by bylo záhodno udržet na paměti pro případné budoucí podobné situace. Pokud lidé nedůvěřují právní regulaci, že je schopná zajistit ochranu jejich osobních dat, pak mohou odmítat i ve skutečnosti bezpečné aplikace. Nedůvěra ve vymahatelnost a dodržování práva a ve stát obecně proto přímo souvisí se společenskou neakceptací dobrovolných opatření v boji s šířící se epidemií. Nedůvěra ve funkčnost státu a práva v tomto případě stojí životy lidí. Pokud se tedy stát chce spoléhat na své občany a vzbuzovat v nich důvěru vůči dobrovolným technologickým řešením, musí být schopen zajistit efektivní ochranu osobních údajů a důvěru ve stát a právo. Nejde přitom pouze o to, zda něco bezpečné ve skutečnosti je, ale i o to, zda si to společnost také subjektivně myslí.

V České republice se nepoužívaly pouze trasovací aplikace, nýbrž i praxe vytváření tzv. *vzpomínkových map*. Vzpomínkové mapy vytvářely krajské hygienické stanice během kontaktu s nakaženou osobou. Účelem mapy bylo pomoci nakažené osobě identifikovat co

⁵¹⁴ REKONSTRUKCE STÁTU. Monitorovací zpráva Sítě k ochraně demokracie za období 1. 1. 2020 – 31. 7. 2020. 2020 [cit. 2022-02-21]. Dostupné na: <https://www.rekonstrukcestatu.cz/archiv-novinek/experti-omezovani-obcanskych-prav-v-prvni-vlne-pandemie-bylo-bezprecedentni>.

⁵¹⁵ ABELER, J.; BÄCKER, M.; BUERMAYER, U.; ZILLESSEN, H. COVID-19 Contact Tracing and Data Protection Can Go Together. *JMIR Mhealth Uhealth*, 2020, vol. 8, no. 4, e19359.

⁵¹⁶ ALTMANN, S. et al. Acceptability of app-based contact tracing for covid-19: Cross-country survey study. *JMIR mHealth and uHealth*, 2020, vol. 8, no. 8, e19857.

⁵¹⁷ GUILLON, M.; KERGALL, P. Attitudes and opinions on quarantine and support for a contact-tracing application in France during the covid-19 outbreak. *Public Health*, 2020, vol. 188, s. 21–31.

největší množství potenciálních kontaktů, a tím zabránit šíření nemoci. Technologický aspekt zde spočíval ve *fakultativní* možnosti podmíněné souhlasem nakaženého, aby byla do vzpomínkové mapy přidána data poskytnutá mobilními operátory a bankami.⁵¹⁸ Na základě těchto dat bylo posléze pomáháno nakaženým vzpomenout si, kde byli, co dělali a s kým se setkali.

I přes dobrovolnost užití se však i na tento mechanismus objevovala kritika, a to hlavně z důvodu, že nebyla úplně zřejmá „*role soukromých společností zapojených do realizace projektu i komunikace s nakaženými ve vztahu k dodržování povinnosti mlčenlivosti hygieniků*“.⁵¹⁹ Trasování skrze vzpomínkové mapy navíc mohlo fungovat pouze v návaznosti na nepřilíš velké kapacity krajských hygienických stanic, a když ty přestaly stíhat, přestal dávat tento mechanismus smysl úplně.

Z této zkušenosti vyplývá, že by stát měl být od začátku co nejvíce transparentní a lépe komunikovat, co dělá, s kým to dělá a proč to dělá. Jakékoli pochybnosti ohledně ochrany osobních údajů a netransparentní spolupráce se soukromými společnostmi totiž vedou k nižší ochotě společnosti dobrovolná opatření využívat. Státům posléze z jejich pohledu může nezbýt nic jiného než využít řešení nedobrovolná.

6.3 Nedobrovolné elektronické sledování osob

Nedobrovolné systémy sledování občanů spočívající v centralizovaném získávání polohy mobilních telefonů od poskytovatelů mobilních služeb nebyly používány v České republice, z demokratických států k nim však přistoupil například Izrael nebo Jižní Korea. Domnívám se, že tyto příklady mohou být pro srovnání užitečné, neboť nám při potenciální budoucí pandemii, která by mohla být ještě výrazně nebezpečnější, dost možná jiná možnost než tyto praktiky také využít nakonec nezbude.

Co si lze odnést z izraelského příkladu? Předně, je třeba upozornit na vztah mezi užíváním nedobrovolného sledování a dobrovolných sledovacích aplikací. Nedobrovolné sledování je na jednu stranu efektivní (sledujeme všechny potenciální kontakty), na druhou

⁵¹⁸ CHVOJKA, Š. Ochrana soukromí v česko-slovenských chytrých karanténách. *Revue pro právo a technologie*, 2021, vol. 12, no. 23, s. 15.

⁵¹⁹ REKONSTRUKCE STÁTU. Monitorovací zpráva Sítě k ochraně demokracie za období 1. 1. 2020 – 31. 7. 2020. 2020 [cit. 2022-02-21]. Dostupné na: <https://www.rekonstrukcestatu.cz/archiv-novinek/experti-omezovani-obcanskych-prav-v-prvni-vlne-pandemie-bylo-bezprecedentni>, s. 2.

stranu ale podryvá motivaci lidí užívat dobrovolné aplikace.⁵²⁰ To však není jediným negativem, neboť tento přístup ve svém důsledku také vede k obecné pasivitě společnosti v boji s epidemií. Mezi státem a občanem se ztrácí důvěra a vztah partnerství, což jsou zcela klíčové atributy pro zvládnání jakékoli mimořádné situace.⁵²¹ Pokud stát dává ostentativně najevo, že svým občanům nedůvěřuje, nemají občané jediný důvod důvěřovat státu. Důvěra je totiž ze své podstaty relačním (vztahovým) jevem.

V praxi by to vypadalo například tak, že by uživatel telefonu přišla karanténní SMS zpráva na základě informace o blízkosti kontaktu „nakaženého“ mobilu. Zde se navíc ani příliš nemusí lišit tzv. chytré a „hloupé“ telefony, neboť to, co se sleduje, může být pouhá poloha jejich signálu. Toto opatření však lze nakonec poměrně jednoduše obejít tím, že je možné telefon nechat ležet doma na stole. Také právních otázek vyvstává při tomto postupu mnoho. Nedobrovolné hromadné sledování zasahuje do osobní autonomie, svobody a soukromí sledovaných.^{522 523} Navíc vyžaduje omezit možnost užívání anonymních telefonů a telefonních čísel.

Na druhou stranu říct, že by takový postup neměl být v demokratické společnosti přípustný nikdy, se mi také nezdá jako vhodné řešení. Je třeba si uvědomit, že takřka vše v oblasti lidských práv je nakonec otázkou *přiměřenosti*. I samotný čl. 10 odst. 2 Listiny připouští možnost *oprávněného* zásahu do soukromého života. Vždy tak nakonec půjde o *zvážení*, zda jde o *proporcionální* zásah, kdy bude na jednu stranu vážen veřejný zájem a na druhé základní práva. Závažnost veřejného zájmu se bude odvíjet od *faktické nebezpečnosti* viru, popřípadě ze začátku od *míry nejistoty*, ve které bude veřejná moc muset rozhodovat. Pokud by tak například smrtnost nového viru byla násobná smrtnosti koronaviru, mohlo by hromadné sledování občanů testem proporcionality projít. Mělo by to však normativně být, vzhledem k výrazným negativním společenským následkům, až nástrojem poslední instance. A samozřejmě by k tomu musela být přijata odpovídající legislativa. Rozhodně by to nebylo možno provést přijetím mimořádného opatření Ministerstva zdravotnictví bez dostatečné zákonné opory.

⁵²⁰ TOCH, E.; AYALON, O. How mass surveillance can crowd out installations of COVID-19 contact tracing apps. *arXiv:2110.01567v1*, 2021, s. 2

⁵²¹ Tamtéž, s. 2.

⁵²² ZIEGELMEYER, A.; SCHMELZ, K.; PLONER, M. Hidden costs of control: four repetitions and an extension. *Experimental Economics*, 2012, vol. 15, no. 2, s. 323–340.

⁵²³ MARCIANO, A. Israel's mass surveillance during covid-19: A missed opportunity. *Surveillance & Society*, 2021, vol. 19, no. 1, 2021, s. 85–88.

6.4 Debata o covid pasu

Covid pas byl elektronický certifikát prokazující absolvování očkování proti nemoci Covid-19 či potvrzující absolvování negativního PCR či antigenního testu na přítomnost viru v návaznosti na konkrétní opatření v různých fázích pandemie. Tento certifikát byl využíván při cestování nebo například při vstupu do provozoven služeb. Technický prvek je zde skoro zanedbatelný, neboť tento systém by mohl úplně stejně fungovat i na bázi papírového potvrzení. Z právního pohledu je však zajímavé zmínit některé kontroverze, které tento certifikát vyvolával. Asi nejsilnější argument se týkal možnosti, že šlo o zakázanou diskriminaci.

Základním ústavním pravidlem, které se týká možnosti odlišovat skupiny jednotlivců, je čl. 3 odst. 1 Listiny, který stanoví, že „základní práva a svobody se zaručují všem bez rozdílu pohlaví, rasy, barvy pleti, jazyka, víry a náboženství, politického či jiného smýšlení, národního nebo sociálního původu, příslušnosti k národnostní nebo etnické menšině, majetku, rodu nebo jiného postavení“. Odlišovací kritérium očkovanosti či bezinfekčnosti v tomto výčtu obsaženo není, takže kritikům tohoto opatření nezbývalo než argumentovat, že skutečnost, zda je někdo očkovaný, spadá do kategorie *jiného postavení*.

V takovém případě by se jednalo o podezřelé kritérium rozlišování, a tím pádem by na tato opatření spadal přísnější lidskoprávní přezkum. V základu covid pasu totiž bylo omezování základních práv, například práva na svobodu pohybu. Tento názor zastávala například advokátka Vladana Vališová, která jej vyjádřila následovně: „Skupina osob, které se nechtějí nechat očkovat nepovinnou vakcínou, se odlišuje od skupiny osob, které se nechají očkovat nepovinnou vakcínou, a tvoří tedy skupinu, která poměrně jasně spadá do kategorie *jiné postavení*“. ⁵²⁴ Podpůrné argumenty ve prospěch tohoto pohledu opírala o čl. 7 a čl. 10 Listiny, § 93 OZ nebo § 28 odst. 1 zákona č. 372/2011 Sb., o zdravotních službách.

NSS však takový názor nesdílel, neboť ve svém rozhodnutí ze dne 2. 2. 2022, sp. zn. 8 Ao 2/2022-53, v bodě 30 konstatoval:

„Podmínky pro vstup osob do některých vnitřních a venkovních prostor nebo pro účast na hromadných akcích nebo jiných činnostech [...] podle soudu nezakládají samy o sobě ani jiné obdobné postavení osob ve smyslu čl. 3 odst. 1 Listiny. Dobrovolné očkování nebo prodělání onemocnění totiž nepředstavují

⁵²⁴ VALIŠOVÁ, V. K diskusi: Zelený certifikát, Health Passport čili covid pas a bonusy za očkování. In: *Advokátní deník* [online] 7. 4. 2021 [cit. 2022-02-21]. Dostupné na: <https://advokatnidenik.cz/2021/04/07/k-diskusi-zeleny-certifikat-health-passport-cili-covid-pas-a-bonusy-za-ockovani/>.

trvalý znak, vlastnost nebo osobností rys jednotlivce, který by bylo možné bez dalšího připodobnit ke kritériím v tomto článku výslovně uvedeným.“

V bodě 85 tohoto rozhodnutí soud navíc došel k závěru, že ochrana veřejného zdraví je legitimním veřejným zájmem, který může ospravedlňovat mnohá opatření, a to za předpokladu, že tato opatření mají potenciál tohoto cíle dosáhnout. Tento kauzální vztah je přitom otázkou skutkovou a bude se v různých epidemiích a při posuzování různých opatření lišit. Povinnost prokazovat se certifikátem o očkování pro vstup do restauračních zařízeních tímto testem v tomto konkrétním případě prošla.

Opatření Ministerstva zdravotnictví tak sice bylo zrušeno, ale pouze z důvodu absence pravomoci takové opatření vydat, nikoli kvůli diskriminaci. Argumenty NSS mi obecně přijdou přesvědčivé. Na druhou stranu pouhá existence kauzálního vztahu nemůže být dostatečná. Domnívám se, že by soud měl při přezkumu brát v potaz také otázku, zda jde o řešení, které není ze strany veřejné moci arbitrární či bezdůvodně excesivní. Je třeba také zdůraznit, že aby něco byla zakázaná diskriminace, musí jít o odlišování věcně neopodstatněné. Pokud by byl mezi skupinami osob rozdíl z hledisek způsobivosti přenášet virus a vážných dopadů na jejich zdraví s důsledkem v míře zatížení zdravotního systému, pak nepůjde o svévoli. Zásadní je omezovat jen práva osob, u nichž je to důvodné. Zvažovat pak lze rozsah omezení neočkovaných a zachování přístupu k základním službám. Záleží tudíž na tom, jak se nakonec ke covid pasu, či jinému pasu, propracovat.

6.5 Možnost užití velkých dat a umělé inteligence při boji s covidem

Třetí oblast týkající se role technologií v boji s epidemií je specifická, neboť se netýká toho, co Česká republika v boji s epidemií dělala, nýbrž toho, co nedělala a dělat měla a mohla. Epidemie je specifická pohroma, u které je velmi důležité mít při rozhodování k dispozici přesné informace a data. Je třeba chápat a předvídat vytížení nemocnic, anticipovat poptávku po zdravotnickém materiálu nebo mít komplexní obraz ohledně šíření viru. Je třeba vědět, kde je kolik zdravotnického vybavení a jak se bude situace vyvíjet, protože to umožňuje jednat s předstihem. Zároveň nekonzistence v datech a jejich špatné zpracování může stát životy lidí.

Z tohoto důvodu například britská NHS vytvořila centrální systém, který propojil všechna data na jedno místo a umožnil s nimi efektivně pracovat v reálném čase.⁵²⁵ Tento krok

⁵²⁵ GOULD, M.; JOSHI, I.; TANG, M. The power of data in a pandemic. In: *GOV. UK. Blog. Technology in the NHS* [online]. 28. 3. 2020 [cit. 2022-02-21]. Dostupné na: <https://healthtech.blog.gov.uk/2020/03/28/the-power-of-data-in-a-pandemic/>.

zajistil přesný obraz o vytíženosti nemocnic, kapacitě lůžek nebo délce pobytu pacientů. Skrze efektivní analýzu takto velkého množství dat bylo možno lépe pochopit šíření viru, identifikovat zvláště zranitelné skupiny, optimalizovat zdravotní péči, zdravotnické zdroje atd.⁵²⁶

Důležité je upozornit, jakým způsobem NHS takto složitý úkol zvládla ve velmi krátkém čase. Nepokoušela se vytvářet vnitřní systém sama nebo skrze pochybné dodavatele (což je v mnoha případech přístup české veřejné správy), nýbrž požádala o spolupráci globální technologické společnosti jako Microsoft, Google, Palantir Technologies⁵²⁷ nebo Amazon Web Services. Díky tomuto byla v práci NHS s daty využita v řádech několika dnů nejmodernější technologie zpracování dat, včetně nejpokročilejších nástrojů umělé inteligence a strojového učení. Mnoho distribučních a dalších úloh je totiž vhodných pro paradigma strojového učení.⁵²⁸ Skrze stejné technologické zázemí se posléze také distribuovaly vakcíny, a to nejen ve Spojeném království, nýbrž také například v USA. Užívání umělé inteligence při boji s epidemií doporučovala i Evropská komise.⁵²⁹

Tento způsob využívání velkých dat a spolupráce s technologickými společnostmi přitom byl o generace před čímkoli, čím disponovala česká veřejná správa. Toto lze perfektně doložit následujícím. Dne 4. 2. 2022 jsem na Ministerstvo zdravotnictví poslal žádost o informaci dle zákona č. 106/1999 Sb., o svobodném přístupu k informacím, ve znění „*zda ministerstvo zdravotnictví či jeho podřízené organizace využívaly při boji s pandemií nemoci Covid-19 umělou inteligenci či nástroje strojového učení, a to ať již v rámci využívání služeb softwaru třetí strany, nebo svého vlastního. Pokud ano, který software takto byl využíván*“. Odpověď Ministerstva zdravotnictví ze dne 16. 2. 2022 byla, že „*Ministerstvo zdravotnictví využívalo při boji s pandemií nemoci Covid-19 chatbot*“. Nejsofistikovanější technologií, kterou český stát v boji s pandemií využíval, byl tudíž chatbot, který odpovídal na dotazy obyvatelstva.

V tomto kontextu je zarážející, že se v českém veřejném prostředí možnost užití umělé inteligence a strojového učení při boji s epidemií vůbec neřešila. Pravděpodobně tu nikoho ze

⁵²⁶ Viz např. OLIVA, D; HASSAN, S. A.; MOHAMED, A. (eds). *Artificial Intelligence for COVID-19*. Cham: Springer International Publishing, 2021.

⁵²⁷ Pozn. autor této práce je minoritním akcionářem společnosti Palantir Technologies Inc.

⁵²⁸ SUBRAMANIAN, M.; SHANMUGA VADIVEL, K. (eds.). The role of contemporary digital tools and technologies in Covid-19 crisis: An exploratory analysis. *Expert Systems*, 2021, s. 1–18.

⁵²⁹ EVROPSKÁ KOMISE. Digitální řešení z období pandemie. 2022 [cit. 2022-02-21]. Dostupné na: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/digital-solutions-during-pandemic_cs.

členů vlády, administrativního aparátu, novinářů a politiků nenapadlo, že něco takového bylo možné. Efektivní výkon veřejné správy je přitom jednou z klíčových podmínek pro *legimitu* veřejné moci, o to více v době mimořádných situací. Užití umělé inteligence v některých státech samozřejmě provázely také obavy ohledně ochrany soukromí atd.⁵³⁰ Je však důležité říct, že lze mít oboje: efektivní veřejnou správu, která využívá nejmodernější technologie, a zároveň ochranu základních práv obyvatel. Dát tyto dva požadavky do neřešitelného rozporu je falešná kontradikce.

6.6 Závěr

Na závěr bych zopakoval nejdůležitější body této kapitoly. V boji s epidemií pomocí technologických nástrojů je klíčovým prvkem v demokratické společnosti důvěra občanů ve stát, který je transparentní a který je schopen zajistit efektivní ochranu soukromých údajů. Pokud jsou tyto podmínky naplněny, jsou lidé více ochotní užívat dobrovolné trasovací aplikace a být více aktivní v zabraňování šíření viru ze své vlastní iniciativy. Pokud tyto podmínky naplněny nejsou, lidé odmítají dobrovolné trasovací aplikace využívat, i pokud jsou fakticky bezpečné. Používá-li navíc stát oba druhy trasování (dobrovolné i nedobrovolné), podkopává tím ochotu lidí užívat dobrovolné aplikace.

Ve vztahu ke covid pasům a otázce diskriminace došla judikatura NSS k závěru, že status naočkování či bezinfekčnosti není *jiné postavení* dle čl. 3 odst. 1 Listiny. V neposlední řadě je třeba zmínit, že užívání umělé inteligence a strojového učení při boji s pandemií Covid-19 bylo v České republice v podstatě neexistující, což snižovalo efektivitu výkonu veřejné moci v řízení boje s pandemií, a tím pádem také podkopávalo *legimitu* veřejné moci.

⁵³⁰ SHACHAR, C.; GERKE S. (eds). AI Surveillance during Pandemics: Ethical Implementation Imperatives. *Hastings Center Report*, 2020, vol. 50, no. 3, s. 18–21.

Závěr

Cílem této práce bylo obrazně rozepsat část knihy, kterou v mnoha důležitých detailech budou muset dokončit jiní. Toto reflektuje skutečnost, že téma vztahu dynamiky technologického vývoje a společenských a právních mechanismů je ohromným a dosud nedostatečně probádaným polem. Právní systém jistě může sloužit jako důležitý ochranný mechanismus společnosti před technologickou disrupcí, avšak jako každý lidský výtvar má svoje přirozené limity. V této práci jsem se snažil jak možnosti práva, tak jeho limity v základech představit.

S tématem technologií jsou navíc silně spojeny i otázky etické a politické související s tím, zda vůbec jako společnost máme funkční veřejné prostory, kde by se například o potenciálně nebezpečných technologiích (a o podobách regulace) dalo smysluplně diskutovat. Absence smysluplných společenských debat je v tomto kontextu zásadním problémem. Technologie nevyhnutelně vyvolávají nepříjemné otázky, neboť vedou k hlubokým změnám ve společnosti. Otázky, zda právní normy zůstanou relevantní, a pokud ne, jak by se měly změnit, jsou zcela fundamentální a v základu se dotýkají každého z nás. Pokud nebude, kde je smysluplně adresovat, obávám se, že se s nimi jako společnost dostatečně nevy pořááme.

Jak bylo také v této práci upozorněno, samotný pojem „technologie“ se v posledních letech významově zúžil v podstatě jen na otázku informačních technologií. Osobně bych si velmi přál, abychom během našich životů zažili výrazný technologický progres v nejširším – dnes již skoro zapomenutém – slova smyslu, zahrnujícím i v současnosti stagnující oblasti jako stavebnictví, energetika či doprava. Zda se technologický progres přenesení i do těchto oblastí, bude však do značné míry záležet na nás samotných. Nové technologie nutně potřebujeme a existují cesty, jak strukturovat (nejen) právní regulaci způsobem, který k takovému progresu povede. Tato práce některé cesty, jak toho dosáhnout, nastínila. Jde o důležitý první krok.

Toto století také hrozí novými ohroženími základních liberálních svobod. Technologie mohou v tomto kontextu být velmi nebezpečným hráčem, ale i dobrým pomocníkem. Existují přitom elegantní cesty, jak základní práva a svobody v technologickém kontextu chránit. Některé prvotní myšlenky, jak toho dosáhnout, zde byly představeny. Za nejvýznamnější v této oblasti považuji návrh inkorporace negativních lidských práv do samotných AI systémů. Jsem si však vědom, že bez další – právní i technické – práce, a hlavně bez praktické implementace, tento návrh nebude mít takový efekt, jaký by byl potřeba.

V neposlední řadě žijeme v období společenských změn, kdy se demokratické režimy dostávají do konfrontace jak s autoritářskými a technokratickými způsoby vládnutí, tak s vnitřní tendencí k sebedestrukci liberálních principů svobodné otevřené společnosti. Tento jev se dle mého názoru nejvíce projevuje ve vztahu k omezování svobody projevu. Proč musíme těmto vnitřním tendencím i externím nepřítelům čelit, bylo posledním výrazným tématem této práce.

Domnívám se, že téma vztahu společnosti k technologiím je v současnosti nezdravě narušené. Z toho lze vinit skutečnost, že technologický progres v poslední dekádě přinesl obří zisky několika technologickým korporacím v malé části USA, avšak zbytek společnosti – včetně demokratického zřízení – z něj buď neprofitoval vůbec, nebo jen velmi málo. Myslím si však a doufám, že toto nebude příběh této dekády a že pokrok v umělé inteligenci se skutečně může pozitivně prospat do životní úrovně většiny lidí. Aby se tak stalo, potřebujeme zvrátit negativní pohled na budoucnost a technologie. Budoucnost může být lepší a již v současnosti existují zcela konkrétní technologické cesty, jak takové budoucnosti dosáhnout.

Pokud tato práce alespoň trochu přispěje tomu, abychom takové budoucnosti dosáhli, budu považovat její cíl za splněný.

Seznam zkratek

ABGB

Všeobecný zákoník občanský pro veškeré německé země dědičné rakouského mocnářství, vyhlášený patentem ze dne 1. června 1811, č. 946 RGB

ADR

alternativní řešení sporů

AI

umělá inteligence

AI systémy

systémy umělé inteligence

Akt o digitálních službách / DSA

Nařízení Evropského parlamentu a rady (EU) 2022/2065 ze dne 19. října 2022 o jednotném trhu digitálních služeb a o změně směrnice 2000/31/ES

Akt o umělé inteligenci / AIA

Návrh nařízení Evropského parlamentu a rady, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění určité legislativní akty unie (COM/2021/206 final)

Avia law

francouzský návrh zákona o boji proti nenávistnému obsahu na internetu

BGB

německý občanský zákoník z roku 1896

Code civil

francouzský občanský zákoník z roku 1804

ESLP

Evropský soud pro lidská práva

EU

Evropská unie

EULP

Úmluva o ochraně lidských práv a základních svobod (č. 209/1992 Sb.)

GDPR

Nařízení (EU) 2016/679 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a volném pohybu těchto údajů

GMO

geneticky modifikované organismy

HDP

hrubý domácí produkt

Listina

usnesení předsednictva ČNR o vyhlášení Listiny základních práv a svobod jako součásti ústavního pořádku České republiky, uveřejněné pod č. 2/1993 Sb.

NHS

Národní zdravotní služba ve Velké Británii

NSS

Nejvyšší správní soud

OSN

Organizace spojených národů

OECD

Organizace pro hospodářskou spolupráci a rozvoj

OZ

zákon č. 89/2012 Sb., občanský zákoník

SDEU

Soudní dvůr Evropské unie

SFEU

Smlouva o fungování Evropské unie

Úmluva o zákazu chemických zbraní

Úmluva o zákazu vývoje, výroby, hromadění zásob a použití chemických zbraní a o jejich zničení

UNESCO

Organizace spojených národů pro výchovu, vědu a kulturu

USA

Spojené státy americké

Ústava

zákon č. 1/1993 Sb., Ústava České republiky

VVOP

Velmi velká online platforma

VVIV

Velmi velký internetový vyhledávač

OUP

Oxford University Press

NUP

New York University Press

Seznam použitých zdrojů

1. Seznam použité odborné literatury

- AARONSON, S.; CARROLL, S. M.; OUELLETTE, L. Quantifying the Rise and Fall of Complexity in Closed Systems: The Coffee Automaton. *arXiv:1405.6903v1*, 2014.
- ABELER, J.; BÄCKER, M.; BUERMAYER, U.; ZILLESSEN, H. COVID-19 Contact Tracing and Data Protection Can Go Together. *JMIR Mhealth Uhealth*, 2020, vol. 8, no. 4, e19359.
- AHMED, N.; MICHELIN, R. A.; XUE, W.; RUJ, S.; MALANEY, R.; KANHERE, S. S.; SENEVIRATNE, A.; HU, W.; JANICKE, H.; JHA, S. K. *A survey of covid-19 contact tracing apps*. IEEE Access, *arXiv:2006.10306v3*, 2020.
- AIDID, A.; ALARIE, B. *The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better*. Toronto: University of Toronto Press, 2023.
- AIZENBERG, E.; VAN DEN HOVEN, J. Designing for human rights in AI. *Big Data & Society*, 2020, vol. 7, no. 2, s. 1–14.
- ALETRAS, N.; TSARAPATSANIS, D.; PREOȚIUC-PIETRO, D.; LAMPOS, V. Predicting judicial decisions of the European Court of Human Rights: A Natural Language Processing perspective. *PeerJ Computer Science*, 2016, 2:e93.
- ALEXANDRE, F. M. *The Legal Status of Artificially Intelligent Robots: Personhood, Taxation and Control*. Tilburg: Tilburg University, 2017. Dissertation.
- ALEXY, R. *A Theory of Constitutional Rights*. Oxford: OUP, 2004.
- ALEXY, R. On the Structure of Legal Principles. *Ratio Juris*, 2000, vol. 13, no. 3, s. 294–304.
- ALEXY, R. The Absolute and the Relative Dimensions of Constitutional Rights. *Oxford Journal of Legal Studies*, 2017, vol. 37, no. 1, s. 31–47.
- ALTAY, S.; BERRICHE, M.; ACERBI, A. Misinformation on Misinformation: Conceptual and Methodological Challenges. *Social Media + Society. Special Issue: Multidisciplinary approaches to mis- and disinformation studies*, 2023, s. 1–13.
- ALTMANN, S. et al. Acceptability of app-based contact tracing for covid-19: Cross-country survey study. *JMIR mHealth and uHealth*, 2020, vol. 8, no. 8, e19857.
- ANTHIS, J. R.; PAEZ, E. Moral circle expansion: A promising strategy to impact the far future. *Futures*, 2021, vol. 130, 102756.
- ARAI, Y. The margin of appreciation doctrine: a theoretical analysis of Strasbourg's variable geometry. In: FOLLESDAL, A.; PETERS, B.; ULFSTEIN, G. (eds.). *Constituting Europe – the European Court of Human Rights in a National, European and Global Context*. Cambridge: Cambridge University Press, 2013.
- ARAÚJO, R.; KOESSLER, L. The Rise of the Constitutional Protection of Future Generations. LPP Working Paper, 2021, no. 7.
- ARENDTOVÁ, H. *Původ totalitarismu I–III*. Praha: Oikoymenth, 1996.
- ARMSTRONG, S.; O'RORKE, X. Good and safe uses of AI Oracles. *arXiv:1711.05541*, 2017.
- ARORA, S.; DOSHI, P. A survey of inverse reinforcement learning: Challenges, methods and progress. *Artificial Intelligence*, 2021, vol. 297, 103500.

- BANEKE, D. Synthetic technocracy: Dutch scientific intellectuals in science, society and culture, 1880-1950. *The British Journal for the History of Science*, 2011, vol. 44, no. 1, s. 89-113.
- BARAK, A. *Proportionality: Constitutional Rights and their Limitations*. Cambridge: Cambridge University Press, 2012.
- BARFIELD, W.; PAGALLO, U. *Research Handbook on the Law of Artificial Intelligence*. Cheltenham: Edward Elgar Publishing, 2018.
- BARTOŇ, M. *Svoboda projevu: principy, garance, meze*. Praha: Leges, 2010.
- BAXTER, H. Niklas Luhmann's Theory of Autopoietic Legal Systems. *Annual Review of Law and Social Science*, 2013, vol. 9, s. 167–184.
- BEHRENDT, S. Grundrechtstheoretische Begründbarkeit einer Verantwortung gegenüber künftigen Generationen: Das grundrechtstheoretische Rechtsverhältnis in der Zeit. In: AMMANN, O. (eds.). *Verantwortung und Recht: 62. Junge Tagung Öffentliches Recht*. Baden-Baden: Nomos, 2022, s. 259-274.
- BELLAMY, R. *Political Constitutionalism: A Republican Defence of the Constitutionality of Democracy*. Cambridge: Cambridge University Press, 2007.
- BENJAMIN, A. The Path of the Law: Towards Legal Singularity. *University of Toronto Law Journal*, 2016, vol. 66, no. 4, s. 443–455.
- BENNETT MOSES, L. Regulating in the Face of Sociotechnical Change. In: BROWNSWORD, R. (eds.). *The Oxford Handbook of Law, Regulation and Technology*. Oxford: OUP, 2017.
- BENZ, E. Lessons from Fukushima: Strengthening the International Regulation of Nuclear Energy. *Wm. & Mary Envtl. L. & Pol'y Rev.*, 2013, vol. 37, issue 3/8, s. 845–883.
- BERAN, K. Může být právnická osoba svéprávná? *Právník*, 2018, roč. 157, č. 8, s. 657-669.
- BERAN, K.; ČECH, P.; DVOŘÁK, B.; ELISCHER, D.; HRÁDEK, J.; JANEČEK, V.; KÜHN, Z.; ONDŘEJEK, P. *Artificial legal entities: essays on legal agency and liability*. Praha: Wolters Kluwer, 2019.
- BETZ, U. et al. Game changers in science and technology – now and beyond. *Technological Forecasting & Social Change*, 2023, vol. 193, 122588.
- BETZ, U. Is the force awakening? *Technological Forecasting & Social Change*, 2018, vol. 128, s. 296-303.
- BLACK, M. The Gap Between 'Is' and 'Should.' *The Philosophical Review*, 1964, vol. 73, no. 2, s. 165–181.
- BOLLINGER, C. *The Tolerant Society*. Oxford: OUP, 1986.
- BOSTROM, N. *Superintelligence: až budou stroje chytrější než lidé*. Praha: Prostor, 2017.
- BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford: OUP, 2014.
- BRAZIER, M.; CAVE, E. *Medicine, Patients and the Law*. London: Penguin Books, 2011.
- BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; AMODEI, D. et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 2020, vol. 33, s. 1877-1901.
- BRYNJOLFSSON, E.; MCAFEE, A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. Norton, 2016.
- BURRI, T. International Law and Artificial Intelligence. In: *German Yearbook of International Law*, 2017, vol. 60, Berlin: Duncker & Humblot.
- CASEY, A. J.; NIBLETT, A. The Death of Rules and Standards. *Indiana Law Journal*, 2017, vol. 92, no. 4, s. 1401–1447.

CAVE, S.; ÓHÉIGEARTAIGH, S. S. Bridging near-and long-term concerns about AI. *Nature Machine Intelligence*, 2019, vol. 1, no. 1, s. 5–6.

ČERNÁ, M. K vybraným pojmům z oblasti řízení obchodních korporací v kontextu diskuze o právní osobnosti distribuovaných autonomních organizací. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022.

CHENG, L. *The Rise of Technocracy: Elite Transformation and Ideological Change in post Mao China*. Princeton University, Dissertation, 1992.

CHOWDHERY, A.; NARANG, S.; DEVLIN, J.; BOSMA, M.; MISHRA, G.; ROBERTS, A.; FIEDEL, N. et al. Palm: Scaling language modeling with pathways. *arXiv:2204.02311*, 2022.

CHVOJKA, Š. Ochrana soukromí v česko-slovenských chytrých karanténách. *Revue pro právo a technologie*, 2021, vol. 12, no. 23, s. 5-48.

CIHON, P. Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development. Future of Humanity Institute: Technical Report, 2019.

CLARK, J.; HADFIELD, G. Regulatory Markets for AI Safety. *arXiv:2001.00078*, 2019.

COHEN, M.; HUTTER, M. Pessimism about unknown unknowns inspires conservatism. *Conference on Learning Theory*, 2020.

COURTNEY J.; HIRSCHL, R.; ROSEVEAR, E. Economic and Social Rights in National Constitutions. *The American Journal of Comparative Law*, 2014, vol. 62, no. 4, s. 1043–1094.

COWEN, T. *The Great Stagnation: How America Ate All The Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*. New York: Penguin, 2010.

COWGER, A. R. *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*. Lexington Books, 2020.

CREMER, C. Z. Deep limitations? Examining expert disagreement over deep learning. *Progress in Artificial Intelligence*, 2021, vol. 10, no. 4, s. 449–464.

CRITCH, A.; KRUEGER, D. AI Research Considerations for Human Existential Safety (ARCHES). *arXiv:2006.04948*, 2020.

CULLEN, P. The Negative and Moral Right to Life. A Basis for Functional Human Rights. Southern New Hampshire University, Justice Studies Faculty Papers, 2015.

CUNNINGHAM-PARMETER, K. From Amazon to Uber: Defining Employment in the Modern Economy. *B.U. L. REV.*, 2016, vol. 96, s. 1674-1727.

DAFOE, A.; BACHRACH, Y.; HADFIELD, G.; HORVITZ, E.; LARSON, K.; GRAEPEL, T. Cooperative AI: machines must learn to find common ground. *Nature*, 2021, vol. 593, no. 7857, s. 33-36.

DAFOE, A.; HUGHES, E.; BACHRACH, Y.; COLLINS, T.; MCKEE, K.R.; LEIBO, J.Z.; LARSON, K.; GRAEPEL, T. Open Problems in Cooperative AI. *arXiv:2012.08630*, 2020.

DE BEER, J. Intellectual Property and ‘Open’ Innovation: A Synthesis of Concepts. In: CALBOLI, I.; MONTAGNANI, M. R. (eds.). *Handbook of Intellectual Property Research*. Oxford: Oxford University Press, 2021.

DEEKS, A. High-Tech International Law. *George Washington Law Review*, 2020, vol. 80, s. 574-653.

DEEKS, A. The Judicial Demand for Explainable Artificial Intelligence. *Columbia Law Review*, 2019, vol. 119, no. 7, s. 1829-1850.

- DINZELBACHER, P. Animal Trials: A Multidisciplinary Approach. *The Journal of Interdisciplinary History*, 2002, vol. 32, no. 3, s. 405-421.
- DONAHOE, E.; METZGER, M. M. Artificial Intelligence and Human Rights. *Journal of Democracy*, 2019, vol. 30, no. 2, s. 115–126.
- DOSTOYEVSKY, F. *Notes from the Underground*. Dover Publications, 1992.
- DWORKIN, R. *Taking Rights Seriously*. Harvard: Harvard University Press, 1977.
- EPSTEIN, R. Can Technological Innovation Survive Government Regulation? *Harvard Journal of Law and Public Policy*, 2013, vol. 36, no. 1, s. 87-104.
- EPSTEIN, R. The Erosion of Individual Autonomy in Medical Decision making: Of the FDA and IRBs. *Georgetown Law Journal*, 2007, vol. 96, no. 2, s. 559–582.
- FERGUSON, N. *Věž a náměstí: mocenské sítě od svobodných zednářů po Facebook*. Praha: Argo, 2019.
- FERGUSON, N. What Is To Be Done? Safeguarding Democratic Governance In The Age Of Network Platforms. *Governance In An Emerging New World*, 2018, Fall Series, Issue 318.
- FINDLAY, T. Nuclear Energy and Global Governance: Ensuring Safety, Security and Non-proliferation. *Routledge*, 2010.
- FISH, E.; GAL, M. Echo Chambers and Competition Law: Should Algorithmic Choices be Respected? *Frederic Jenny Liber Amicorum: Standing Up for Convergence and Relevance in Antitrust*, vol. II., 2020.
- FISHER, W. W.; HORWITZ, M. J.; REED, T. A. *American Legal Realism*. Oxford: OUP, 1993.
- FJELD, J.; NAGY, A. *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. The Berkman Klein Center for Internet & Society, Harvard University, 2020.
- GABRIEL, I. Artificial Intelligence, Values, and Alignment. *Minds & Machines*, 2020, vol. 30, s. 411–437.
- GALVIN, D. J.; HACKER, J. S. The Political Effects of Policy Drift: Policy Stalemate and American Political Development. *Studies in American Political Development*, 2020, vol. 34, no. 2, p. 1-23.
- GAMAUF, R. Slaves doing business: the role of Roman law in the economy of a Roman Household. *European Review of History: Revue europeenne d'histoire*, 2009, vol. 16, no. 3, s. 331-346.
- GARDBAUM, S. The "Horizontal Effect" of Constitutional Rights. *Michigan Law Review*, 2003, vol. 102, no. 3, s. 388-458.
- GATES, B. *How to Avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need*. New York: Alfred A. Knopf, 2021.
- GAUS, G. *The Open Society and Its Complexities*. Oxford: OUP, 2021.
- GERICKE, O. *Möglichkeiten und Grenzen eines Abbaus der Verrechtlichung*. Aachen: Shaker Verlag, 2003.
- GLAESER, E.; SHLEIFER, A. The Rise of the Regulatory State. *J. Econ Literature*, 2003, vol. 41, no. 2, s. 401–425.
- GOODWIN, M.; HEATH, O. The 2016 Referendum, Brexit and the Left Behind: An Aggregate-level Analysis of the Result. *The Political Quarterly*, 2016, Vol. 87, no. 3, s. 323–332.
- GORDON, R. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. New Jersey: Princetown University Press, 2016.

- GREAVES, H. (eds.). *A Research Agenda For The Global Priorities Institute*. Global Priorities Institute, 2020, verze 2.1.
- GRAF, L. *The Rise and Fall of American Technology*. New York: Algora Publishing, 2010
- GRUSZCZYNSKI, L. *Regulating Health and Environmental Risks under WTO Law: A Critical Analysis of the SPS Agreement*. Oxford: Oxford University Press, 2010.
- GUILLOIN, M.; KERGALL, P. Attitudes and opinions on quarantine and support for a contact-tracing application in France during the covid-19 outbreak. *Public Health*, 2020, vol. 188, s. 21–31.
- HANYCH, M.; PIVODA, M. Facebook, Twitter a YouTube jako garanti svobodného projevu? Kritika současného systému notice-and-takedown. *Revue pro práva a technologie*, 2017, roč. 8, č. 16, s. 177-220.
- HARARI, Y. N. *21 Lessons for the 21st Century*. Vintage: Vintage Books, 2018.
- HARARI, Y. N. *Homo Deus. A Brief History of Tomorrow*. Random House, 2017.
- HARARI, Y. Reboot for the AI revolution. *Nature*, 2017, vol. 550, s. 324-327.
- HARNED, Z.; WALLACH, H. Stretching Human Laws to Apply to Machines: The Dangers of a 'Colorblind' Computer. *Florida State University Law Review*, 2020, vol. 47, issue 3, s. 618-648.
- HAYEK, F. A. *Cesta do otroctví*. Praha: Academia, 1990.
- HAYEK, F. A. *Law, Legislation and Liberty: A new statement of the liberal principles of justice and political economy*. Routledge Classics, 2012.
- HEIDEGGER, M. *The Question Concerning Technology and Other Essays*. New York: Harper & Row Publishers, 1977.
- HOLLÄNDER, P. Právník v bludišti postmoderní dekonstrukce a Savignyho Ariadnina nit. In: GERLOCH, A.; KRZYŽANKOVÁ, K. Ž. (eds.). *Právo v měnícím se světě*. Plzeň: Aleš Čeněk, 2020.
- HOLMES, O. W. The Path of the Law. *Harvard Law Review*, 1897, vol. 10, no. 8, s. 457-478.
- HOŘEŇOVSKÝ, J. Aktuální otázky svobody projevu v době sociálních sítí. In: KUDRNA, J. (ed.). *Listina v kontextu společenských změn (K 30 výročí jejího přijetí)*. Praha: Wolters Kluwer, 2021.
- HOŘEŇOVSKÝ, J. Modely přístupu práva k regulaci nových technologií a metodologie právní vědy. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022.
- HOŘEŇOVSKÝ, J. Modely regulace nových technologií v éře socio-technologických změn. *Obchodní právo*, 2022, roč. 31, č. 7-8, s. 18-27.
- HOŘEŇOVSKÝ, J. Právní a kulturní aspekty příčin a dopadů teze o technologické stagnaci. *Acta Universitatis Carolinae. Iuridica*, 2023, vol. 69, no. 1, s. 29-44.
- HOŘEŇOVSKÝ, J. Právní a společenské aspekty užití moderních technologií v boji s epidemií covid-19. In: KYSELA, J.; ŘEPA, K. (eds.). *Reakce právního systému České republiky na covidovou pandemii*. Praha: Leges, 2022.
- HOŘEŇOVSKÝ, J. Regulation of Social Media Networks Through the Article 102 TFEU. In: ŠKRABKA, J.; VACUŠKA, L. (eds.). *Law in Business of Selected Member States of the European Union. Proceedings of the XII International Scientific Conference*. Praha: TROAS, s.r.o., 2020.

- HOŘEŇOVSKÝ, J. Umělá inteligence jako hmotněprávní a procesní subjekt v civilním právu s přihlédnutím k zastoupení. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds.) *Zastoupení. Specifika a kontext*. Praha: Wolters Kluwer, 2021.
- HOŘEŇOVSKÝ, J. Ústavněprávní aspekty nuceného stáhnutí představení Jaroslava Duška během koronavirové krize. In: WINTR, J.; ANTOŠ, M. (eds.). *Ústavní právo a koronavirus*. Praha: Leges, 2021.
- HOŘEŇOVSKÝ, J. Vliv algoritmického rozhodování a imperativu zachování základních práv na pojetí civilněprávní odpovědnosti sociálních sítí za obsah nahraný třetími subjekty. In: STŘELEČEK, T.; SEDLÁČEK, M. (eds.). *Povinnost a odpovědnost v civilním právu hmotném a procesním. Speciální problémy*. Praha: Wolters Kluwer, 2021.
- HORENOVSKY, J.; BAJGAR, O. Negative Human Rights as a Basis for Long-term AI Safety and Regulation. *Journal of Artificial Intelligence Research*, 2023, vol. 76, s. 1043-1075.
- HUTTUNEN, A.; KULOVESI, J.; BRACE, W.; LECHNER, G.; SILVENNOINEN, K.; KANTOLA, V. Liberating Intelligent Machines with Financial Instruments. *Nordic Journal of Commercial Law*, 2010, no. 2, s. 1-14.
- IENCA, M.; ANDORNO, R. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy*, 2017, vol. 13, no. 1, s. 1-27.
- IZZO, M. Management Transition to Big Data Analytics: Exploratory Study on Airline Industry. *International Business Research*, 2019, vol. 12, no. 10, s. 48–56.
- JAFFE, G. Regulatory slowdown on GM crop decisions. *Nat Biotechnol*, 2006, vol. 24, no. 7, s. 748–749.
- JÄGER, P.; MOLEK, P. *Svoboda projevu: demokracie, rovnost a svoboda slova*. Praha: Auditorium, 2007.
- JALABNEH, R. et al. Use of Mobile Phone Apps for Contact Tracing to Control the COVID-19 Pandemic: A Literature Review. *SSRN*, 2020.
- JANIS, I. L. *Victims of Groupthink: A psychological study of foreign-policy decisions and fiascoes*. Houghton Mifflin Company, 1972.
- JANSEN, S.; MARTIN, B. The Streisand Effect and Censorship Backfire. *International Journal of Communication*, 2015, vol. 9, s. 656–671.
- JELLINEK, G. *System der subjektiven öffentlichen Rechte*. Akademische Verlagsbuchhandlung von J. C. B. Mohr, 1892.
- JELLINEK, G. *Die sozialetische Bedeutung von Recht*. Unrecht und Strafe, 1878.
- JENKINS, R.; CERNY, D.; HRÍBEJ, T. *Autonomous Vehicle Ethics: The Trolley Problem and Beyond*. Oxford: OUP, 2022.
- JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 2019, vol. 1, s. 389-399.
- KELSEN, H. *General Theory of Norms*. Oxford: Clarendon Press, 1991.
- KELSEN, H. *Reine Rechtslehre*. Vídeň: Deuticke, 1960.
- KESSLER, A. *How We Got Here: A Slightly Irreverent History of Technology and Markets*. Harper Business, 2005.
- KHAN, L. M. Amazon's Antitrust Paradox. *The Yale Law Journal*, 2017, vol. 126, no. 3, s. 790-802.
- KINCL, J.; SKŘEJPEK, M.; URFUS, V. *Římské právo*. Praha: C.H. Beck, 1997.
- KMEC, J.; KOSAR, D.; KRATOCHVÍL, J.; BOBEK, M. *Evropská úmluva o lidských právech*. Komentář. Praha: C. H. Beck, 2012.

- KNAPP, V. *Teorie práva*. Praha: C.H. Beck, 1995.
- KNETL, Š. The windfall clause. In: SEDLÁČEK, M.; STŘELEČEK, T. (eds). *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022.
- KONDRATIEFF, N. *The Long Waves in Economic Life*. Martino Fine Books, 2014.
- KURZWEIL, R. *The Singularity Is Near*. New York: Penguin Books, 2006.
- LATONERO, M. *Governing Artificial Intelligence: Upholding Human Rights & Dignity*. Data & Society, 2018.
- LAVI, M. Content Providers' Secondary Liability: A Social Network Perspective. *Fordham Intellectual Property, Media and Entertainment Law Journal Journal*, 2016, vol. 26, no. 4, s. 855-943.
- LAW, D. S.; VERSTEEG, M. The Evolution and Ideology of Global Constitutionalism. *California Law Review*, 2011, vol. 99, no. 5, s. 1163-1258.
- LEGG, A. *The Margin of Appreciation in International Human Rights Law: Deference and Proportionality*. Oxford: Oxford University Press, 2007.
- LEHMAN-WILZIG, S. Frankenstein unbound: Towards a legal definition of artificial intelligence. *Futures*, 1981, vol. 13, no. 6, s. 442–457.
- LUHMANN, N. *Introduction to Systems Theory*. Cambridge: Polity Press, 2013.
- LUHMANN, N. *Law as a Social System*. Oxford: OUP, 2004.
- LUHMANN, N. *Právo společnosti*. Praha: Academia, 2021.
- MAAS, M. M. Aligning AI Regulation to Sociotechnical Change. In: BULLOCK, J. (eds.). *Oxford Handbook of AI Governance*. Oxford: Oxford University Press, 2022.
- MAAS, M. M.; LUCERO-MATTEUCCI, K.; COOKE, D. Military Artificial Intelligence as Contributor to Global Catastrophic Risk. In: BEARD, S. J. (eds.). *The Era of Global Risk*. Open Book Publishers, 2023.
- MACINTYRE, A. *After Virtue*. London: Gerald Duckworth & Co., 1981.
- MARCIANO, A. Israel's mass surveillance during covid-19: A missed opportunity. *Surveillance & Society*, 2021, vol. 19, no. 1, 2021, s. 85-88.
- MARKOU, Ch.; DEAKIN, S. Ex Machina Lex: Exploring the Limits of Legal Computability. *SSRN*, 2019.
- MARKOVÁ, J. Využití smart contract ve spotřebitelských vztazích, zejména ve smlouvě o zájezdu. In: SEDLÁČEK, M.; STŘELEČEK, T. *Civilní právo a nové technologie*. Praha: Wolters Kluwer, 2022.
- MAVRONICOLA, N. Is the Prohibition Against Torture and Cruel, Inhuman and Degrading Treatment Absolute in International Human Rights Law? A Reply to Steven Greer. *Human Rights Law Review*, 2017, vol. 17, no. 3, s. 479–498.
- MCALLISTER, A. Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture. *Minnesota Law Review*, 2018, vol. 47, s. 2527-2573.
- MCCHANGAMA, J.; ALKIVIADOU, N. Reimagining the Current Regulatory Framework to Online Hate Speech. Why Making Way for Alternative Methods Is Paramount for Free Speech. In: ULLMANN, S.; TOMALIN, M. (eds.). *Counterspeech. Multidisciplinary Perspectives on Countering Dangerous Speech*. New York: Routledge, 2023.
- MCCORMICK, J. P. Irrational Choice and Mortal Combat as Political Destiny: The Essential Carl Schmitt. *Annu. Rev. Polit. Sci.*, 2007, vol. 10, no. 1, s. 315-339.

MCGREGOR, L.; MURRAY, D.; NG, V. International Human Rights Law as a Framework for Algorithmic Accountability. *International and Comparative Law Quarterly*, 2019, vol. 68, no. 2, s. 309-343.

MCHANGAMA, J. *Svoboda projevu: Od Sokrata po sociální síť*. Praha: Institute H21, 2022.

MEDVEDEVA, M., VOLS, M.; WIELING, M. Using machine learning to predict decisions of the European Court of Human Rights. *AI and Law*, 2020, vol. 28, s. 237–266.

MEHRABI, N.; MORSTATTER, F.; SAXENA, N.; LERMAN, K.; GALSTYAN, A. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 2021, vol. 54, no. 6, s. 1-35.

MEIER, H. *The Lesson of Carl Schmitt: Four Chapters on the Distinction between Political Theology and Political Philosophy*. Chicago: University of Chicago Press, 1998.

MELCHIOR, C.; OLIVEIRA, M. A systematic literature review of the motivations to share fake news on social media platforms and how to fight them. *New Media & Society*, 0(0), 2023.

MELRO, A.; OLIVEIRA, L. *Constitutional Knowledge and Its Impact on Citizenship Exercise in a Networked Society*. IGI Global, 2019.

MILL, J. S. *O svobodě myšlení a slova*. Praha: Institute H21, 2020.

MIZEROVÁ, T.; HARAŠTA, J. Dostupnost mediálního obsahu: blokování dezinformačních webů. *Právník*, 2023, roč. 162, č. 5, s. 415-430.

MODIS, T. Links between entropy, complexity, and the technological singularity. *Technological Forecasting and Social Change*, 2022, vol. 176, s. 5.

MOUNK, Y. The Undemocratic Dilemma. *Journal of Democracy*, 2018, vol. 29, issue 2, s. 98-112.

MURRAY, J.; WEBB, T. E.; WHEATLEY, S. Encountering law's complexity. In: MURRAY, J. (eds.). *Complexity Theory and Law. Mapping An Emerging Jurisprudence*. London and New York: Routledge, 2019, s. 3-23.

MUZYKA, K. The Outline of Personhood Law Regarding Artificial Intelligences and Emulated Human Entities. *Journal of Artificial General Intelligence*, 2013, vol. 4, no. 3, s. 164-169.

NARANG, S., et al. WT5?! Training Text-to-Text Models to Explain their Predictions. *arXiv:2004.14546*, 2020.

NIDLER, R. The United Nation's Capability to Manage Existential Risks with a Focus on Artificial Intelligence. *International Community Law Review*, 2019, vol. 21, no. 1, s. 5-34.

NOGARA, H. et al. The Disinformation Dozen: An Exploratory Analysis of Covid-19 Disinformation Proliferation on Twitter. *WebSci '22: Proceedings of the 14th ACM Web Science Conference*, 2022, s. 348-358.

NOVOTNÝ, O. Hypertrofie práva, její aspekty a jak jí čelit. *Právník*, 2016, vol. 155, no. 7, s. 619-625.

O'KEEFE, C.; CIHON, P.; FLYNN, C.; GARFINKEL, B.; LEUNG, J.; DAFOE, A. *The Windfall Clause: Distributing the Benefits of AI*. Centre for the Governance of AI Research Report, 2020. Future of Humanity Institute, University of Oxford.

OAKESHOTT, M. *Rationalism in Politics*. London: Methuen & Co Ltd, 1962.

ÓHÉIGEARTAIGH, S. S.; WHITTLESTONE, J.; LIU, Y.; ZENG, Y.; LIU, Z. Overcoming barriers to cross-cultural cooperation in AI ethics and governance. *Philosophy & Technology*, 2020, vol. 33, s. 571–593.

- OLIVA, D; HASSAN, S. A.; MOHAMED, A. (eds). *Artificial Intelligence for COVID-19*. Cham: Springer International Publishing, 2021.
- ONDŘEJEK, P. *Defragmentation of Law. Reconstruction of Contemporary Law as a System*. Cambridge: Intersentia, 2023.
- ONDŘEJEK, P. *Koncepce práva jako systému*. Praha: Wolters Kluwer, 2020.
- ORD, T. *Nad propastí: Existenční riziko a budoucnost lidstva*. Praha: Argo, 2022.
- OWE, A.; BAUM, S. D. Moral Consideration of Nonhumans in the Ethics of Artificial Intelligence. *AI and Ethics*, 2021, vol. 1, s. 517–528.
- PASQUALE, F.; CASHWELL, G. Prediction, persuasion, and the jurisprudence of behaviourism. *University of Toronto Law Journal*, 2018, vol. 68, sup. 1, s. 63–81.
- PATTERSON, J. L. The Development of the Concept of Corporation From Earliest Roman Times to A.D. 476. *The Accounting Historians Journal*, 1983, vol. 10, no. 1, s. 87-98.
- PETIT, N. Artificial Intelligence and Automated Law Enforcement: A Review Paper. *SSRN*, 2018.
- PETIT, N. Law and Regulation of Artificial Intelligence and Robots—Conceptual Framework and Normative Implications. *SSRN*, 2017.
- PETIT, N.; DE COOMAN, J. Models of Law and Regulation for AI. *European University Institute. Robert Schuman Centre for Advanced Studies*, 2020, RSCAS 2020/63.
- PHELPS, E. *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change*. Princeton University Press, 2016.
- PIKETTY, T. *Kapitál v 21. století*. Praha: Knižní klub, 2015.
- PINKER, S. *Enlightenment Now*. London: Allen Lane, 2018.
- PIŇOS, J. Genealogie filosofického ospravedlnění svobody projevu. *Jurisprudence*, 2017, roč. 26, č. 2, s. 24-35.
- PITHART, P. *Obrana politiky*. Praha: Academia, 2005.
- POPPER, K. *Otevřená společnost a její nepřátelé I*. Praha: Oikoymenth, 1994.
- POPPER, K. *The Logic of Scientific Discovery*. Routledge, 2002.
- RÁLIŠ, P. Podpora a propagace terorismu – kladivo na internetové diskutéry? *Trestněprávní revue*, 2024, roč. 23, č. 1.
- RASO, F.; HILLIGOSS, H.; KRISHNAMURTHY, V.; BAVITZ, C.; KIM, L. *Artificial Intelligence & Human Rights: Opportunities & Risks*. The Berkman Klein Center, 2018.
- RAWLS, J. The Idea of an Overlapping Consensus. *Oxford Journal of Legal Studies*, 1987, vol. 7, no. 1, s. 1–25.
- RAZ, J. Free Expression and Personal Identification. *Oxford Journal of Legal Studies*, 1991, vol. 11, no. 3, s. 303-324.
- REDDY, S.; DRAGAN, A. D.; LEVINE, S.; LEGG, S.; LEIKE, J. Learning human objectives by evaluating hypothetical behavior. *Proceedings of ICML*, 2020.
- RENU, N. Technological advancement in the era of COVID-19. *SAGE Open Medicine*, 2021, vol. 9, s. 1-4.
- RICCUCCI, M. N. *Policy Drift: Shared Powers and the Making of U.S. Law and Policy*. New York: NUP, 2018.
- RILEY-SMITH, T.; MCCORMICK, L. Liability for Physical Damage. In: HERVEY, M.; LAVY, M. (eds.). *The Law of Artificial Intelligence*. Sweet & Maxwell, 2021.
- RISSE, M. Human Rights and Artificial Intelligence: An Urgently Needed Agenda. *Human Rights Quarterly*, 2019, vol. 41, no. 1, s. 1-16.

- RODRIGUES, R. Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 2020, vol. 4, 100005.
- ROGERS, B. The Social Costs of Uber. *University of Chicago Law Review Online*, 2015, vol. 82, no. 1, s. 85-102.
- ROSLING, H.; ROSLING, O.; ROSLING RÖNNLUND, A. *Faktomluva: deset důvodů, proč se mýlíme v pohledu na svět – a proč jsou věci lepší, než vypadají*. Brno: Jan Melvil Publishing, 2018.
- RUSSELL, S. *Human Compatible: Artificial Intelligence and the Problem of Control*. London: Allen Lane, 2019.
- SAETRA, H. A Shallow Defence of a Technocracy of Artificial Intelligence. *Technology in Society*, 2020, vol. 62, 101283, s. 1-10.
- SAETRA, H. S. *Big Data's Threat to Liberty Surveillance, Nudging, and the Curation of Information*. London: Academic Press, 2021.
- ŠÁMAL, P.; BOHÁČEK, L. "Fake news" jako aktuální právní otázka spojená s globalizací. In: GERLOCH, A. a kol. *Právo v měnícím se světě*. Plzeň: Aleš Čeněk, 2020.
- SCHMITT, C. *Pojem politična*. Praha: Oikoymenth, 2007.
- SCHONI, S. Preventing a Butlerian Jihad: Articulating a Global Vision For the Future of Artificial Intelligence. *Journal of International Affairs*, 2019, vol. 72, no. 1, s. 135-142.
- SCHUCK, P. H. Legal Complexity: Some Causes, Consequences, and Cures. *Duke Law Journal*, 1992, vol. 42, no. 1, s. 1-52.
- SCOBEE, D. R. R.; SASTRY, S. S. Maximum likelihood constraint inference for inverse reinforcement learning. *ICLR*, 2020.
- SELG, P.; VENTSEL, A. *Introducing Relational Political Analysis. Political Semiotics as a Theory and Method*. Palgrave Macmillan, 2020.
- SHABINA, Z. (eds.). Airbnb and its potential impact on the London housing market. *Urban Studies*, 2022, vol. 59, no. 1, s. 197–221.
- SHACHAR, C.; GERKE S. (eds). AI Surveillance during Pandemics: Ethical Implementation Imperatives. *Hastings Center Report*, 2020, vol. 50, no. 3, s. 18-21.
- SHULMAN, C.; BOSTROM, N. Sharing the World with Digital Minds. In: CLARKE, S.; SAVULESCU, J. (eds.). *Rethinking Moral Status*. Oxford: OUP, 2021.
- SCHEU, H. CH. Několik úvah o unijním pojetí dezinformace. *Jurisprudence*, 2023, číslo 5, s. 1-10.
- SMUHA, N. A. Beyond a Human Rights-Based Approach to AI Governance: Promise, Pitfalls, Plea. *Philosophy & Technology*, 2021, vol. 34, no. 1, s. 91–104.
- SOROS, G. Fallibility, Reflexivity, and the Human Uncertainty Principle. *Journal of Economic Methodology*, 2014, vol. 20, no. 4, s. 309-329.
- SOWELL, T. *Knowledge And Decisions*. Basic books, 1979, 1996.
- STEWART, D. (eds.). Equipose lost: ethics, costs, and the regulation of cancer clinical research. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 2010, vol. 28, s. 2925–35.
- STIX, C.; MAAS, M. M. Bridging the Gap: The Case for an ‘Incompletely Theorized Agreement’ on AI Policy. *AI and Ethics*, 2021, vol. 1, s. 261–271.
- STRITTMATTER, K. *We Have Been Harmonised: Life in China's Surveillance State*. Custom House, 2020.
- STROSSEN, N. *Hate: Why We Should Resist it With Free Speech, Not Censorship*. Oxford: OUP, 2020.

STROSSEN, N. *Nenávist: Proč bychom se jí měli bránit pomocí svobody projevu, nikoli cenzurou*. Praha: Institute H21, 2021.

SUBRAMANIAN, M.; SHANMUGA VADIVEL, K. (eds.). The role of contemporary digital tools and technologies in Covid-19 crisis: An exploratory analysis. *Expert Systems*, 2021.

SUMPTION, J. *Trials of the State*. London: Profile Books, 2020.

SUNSTEIN, C. R. Incompletely Theorized Agreements Commentary. *Harvard Law Review*, 1994, vol. 108, no. 7, s. 1733-1772.

SUNSTEIN, C. R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2017.

SURDEN, H. Artificial Intelligence and Law: An Overview. *Georgia State University Law Review*, 2019, vol. 35, no. 4, s. 1306-1337.

SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: MIT Press, 2018.

TANNOCK, Q.; SUMROY, R. Data Protection and Privacy. In: HERVEY, M.; LAVY, M. (eds.). *The Law of Artificial Intelligence*. Sweet & Maxwell, 2021.

TASIOULAS, J. Saving Human Rights from Human Rights Law. *Vanderbilt Law Review*, 2021, vol. 52, no. 5, s. 1167-1207.

TEUBNER, G. Substantive and reflexive elements in modern law. *Law & Society Review*, 1983, vol. 17, no. 2, s. 239-286.

THELEN, K. A. Regulating Uber: The Politics of The Platform Economy In Europe And The United States. *Perspectives on Politics*, 2018, vol. 16, no. 4, s. 938-953.

THIEL, P. The Straussian Moment. In: HAMERTON-KELLY, R. *Politics and Apocalypse*. Michigan: Michigan State University Press, 2007.

THIEL, P.; MASTERS, B. *Od nuly k jedničce. Úvahy o startupech aneb jak tvořit budoucnost*. Brno: Jan Melvil Publishing, 2015.

THIEL, P.; MASTERS, B. *Zero to One. Notes on Start Ups, or How to Build the Future*. London: Ebury Publishing, 2014.

TOCH, E.; AYALON, O. How mass surveillance can crowd out installations of COVID-19 contact tracing apps. *arXiv:2110.01567v1*, 2021.

TURCHIN, A. AI Alignment Problem: “Human Values” don’t Actually Exist. 2019. Dostupné z: <https://philpapers.org/rec/TURAAP>.

VAN AAKEN, A. Is International Law Conducive to Preventing Looming Disasters? *Global Policy*, 2016, vol. 7, no. S1, s. 81-96.

VICKERS, J. Abuse of Market Power. *Economic Journal*, 2005, vol. 115, no. 504, s. 244-261.

VIJG, J. *The American Technological Challenge: Stagnation and Decline in the 21st Century*. New York: Algora Publishing, 2011.

WAGNEROVÁ, E. Čl. 17. In: *Listina základních práv a svobod: komentář*. Praha: Wolters Kluwer, 2012.

WALTMAN, J. *Principled Judicial Restraint: A Case Against Activism*. London: Palgrave Macmillan, 2015.

WARREN, A.; HILLAS, A. Lethal autonomous weapons systems: Adapting to the future of unmanned warfare and unaccountable robots. *Yale Journal of International Affairs*, 2017, vol. 12, no. 1, s. 71–85.

WEIMER, M.; MARIN, L. The Role of Law in Managing the Tension between Risk and Innovation. *EJRR, Special Issue on Regulating New and Emerging Technologies*, 2016, vol. 7, issue 3, s. 469-474.

WEINBERGER, O. *Norma a instituce: úvod do teorie práva*. Plzeň: Aleš Čeněk, 2017.

WHITTLESTONE, J.; CLARK, J. Why and How Governments Should Monitor AI Development. *ArXiv:2108.12427*, 2021.

WILDAVSKY, A. *Searching for Safety*. New Brunswick: Transaction, 1988.

WINTER, Ch. (eds.). *Legal Priorities Research: A Research Agenda*. Legal Priorities Project, 2021.

WINTER, Ch.; MARTÍNEZ, E. *Experimental longtermist jurisprudence*. Legal Priorities Project. LPP working paper, 2021, n. 2.

WINTR, J. *Metody a zásady interpretace práva*. Praha: Auditorium, 2019.

WINTR, J. *Principy českého ústavního práva*. Plzeň: Aleš Čeněk, 2015.

WINTR, J. Schvalování atentátu a podobné výtržnosti. *Státní zastupitelství*, 2020, č. 3, s. 3-4.

YEUNG, K. Why Worry about Decision-Making by Machine? In: YEUNG, K; LODGE, M. *Algorithmic Regulation*. Oxford: OUP, 2020.

YUDKOWSKI, E. Complex Value Systems in Friendly AI. In: SCHMIDHUBER, J.; THORISSON, K. R.; LOOKS, M. (eds.). *Artificial General Intelligence. AGI 2011. Lecture Notes in Computer Science*, vol. 6830. Berlin, Heidelberg: Springer Publishing, 2011.

ZIEGELMEYER, A.; SCHMELZ, K.; PLONER, M. Hidden costs of control: four repetitions and an extension. *Experimental Economics*, 2012, vol. 15, no. 2, s. 323–340.

ZIKMUNDOVÁ, K. Artificial Intelligence and Medical Devices: Do We Need New Regulation? *Časopis pro právní vědu a praxi*, 2023, vol. 31, no. 2, s. 381-394.

ZUBOFF, S. *Věk kapitalismu dohledu: Boj o budoucnost lidstva u nové hranice moci*. Praha: Argo, 2022.

2. Seznam použitých právních předpisů

Africká charta lidských práv a práv národů

Americká úmluva o lidských právech

Communications Decency Act of 1996 (USA)

Evropská Úmluva o ochraně lidských práv a základních svobod (č. 209/1992 Sb.)

francouzský občanský zákoník z roku 1804 (Code civil)

Nařízení (EU) 2016/679 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a volném pohybu těchto údajů (GDPR)

Nařízení Evropského parlamentu a rady (EU) 2022/2065 ze dne 19. října 2022 o jednotném trhu digitálních služeb a o změně směrnice 2000/31/ES (DSA)

Návrh nařízení Evropského parlamentu a rady, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění určité legislativní akty unie (COM/2021/206 final)

německý občanský zákoník z roku 1896 (BGB)

Směrnice Evropského parlamentu a Rady 2001/18/ES ze dne 12. března 2001 o záměrném uvolňování geneticky modifikovaných organismů do životního prostředí a o zrušení směrnice Rady 90/220/EHS

Smlouva o fungování Evropské unie (SFEU)

Úmluva o zákazu vývoje, výroby, hromadění zásob a použití chemických zbraní a o jejich zničení

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

Ústava České republiky (č. 1/1993 Sb.)

Ústava Čínské lidové republiky

Všeobecný zákoník občanský ze dne 1. června 1811 (ABGB)

zákon č. 111/1994 Sb., o silniční dopravě

zákon č. 21/1992 Sb., o bankách

zákon č. 40/2009 Sb., trestní zákoník

zákon č. 561/2004 Sb., o předškolním, základním, středním, vyšším odborném a jiném vzdělávání

zákon č. 89/2012 Sb., občanský zákoník

3. Seznam použité judikatury (dle jurisdikce)

ČR

Nález Ústavního soudu ze dne 8. 2. 2000, sp. zn. I. ÚS 156/99.
Nález Ústavního soudu ze dne 15. 3. 2005, sp. zn. I. ÚS 367/03.
Nález Ústavního soudu ze dne 4. 4. 2005, sp. zn. IV. ÚS 146/04.
Nález Ústavního soudu ze dne 31. 7. 2023, sp. zn. I. ÚS 1594/22.
Rozsudek Nejvyššího správního soudu ze dne 17. 2. 2010, č. j. Pst 1/2009–348.
Rozsudek Městského soudu v Praze ze dne 19. 8. 2021, č. j. 6 Af 20/2020-28.
Usnesení Městského soudu v Praze ze dne 24. 6. 2022, č. j. 14 A 39/2022–14.

SDEU

Rozhodnutí *Flaminio Costa vs. ENEL* (číslo 6/64).
Rozhodnutí *N. V. Algemene Transport-en Expeditie Onderneming Van Gend en Loos v. Netherlands Fiscal Administration* (26/62).
Rozsudek C-528/16.

ESLP

Rozhodnutí ze dne 7. 12. 1976, *Handyside* proti Spojenému království.
Rozsudek ze dne 21. 7. 2011, *Heinisch* proti Německu.

UK

Airedale NHS Trust v. Bland [1993]. All English Law Rep. 1:821-96.
Chahal v. UK [1996] 23 EHRR 413.
R v Secretary of State for the Home Department, ex p Simms [2000] 2 AC 115.

USA

Abigail Alliance for Better Access to Developmental Drugs v. von Eschenbach, 495 F.3d 695 (D.C. Cir. 2007), cert denied, 552 U.S. 1159 (2008).
The United States District Court for the Northern District of Carolina Center for food safety et al. v. Thomas J. Vilsack et al. ze dne 13. 10. 2010, No. C 08-00484 JSW.
The United States Supreme Court decision of 27 February 2017, Packingham v. North Carolina, 582 U.S.

Německo

Rozhodnutí německého Ústavního soudu 1 BvL 89/78.
The Decision of the Higher Regional Court of Düsseldorf (Oberlandesgericht Düsseldorf) in interim proceedings, 26 August 2019, Case VI-Kart 1/19 (V).

4. Seznam ostatních zdrojů

- ALTAY, S. Rethinking the Problem of Misinformation. [přednáška]. Praha: Institute H21, 15. září 2023. Dostupné na: <https://youtu.be/Yms5DfMYQTw?si=IZcBNphvCengeluz>.
- ANDERSEN, L. et al. *Human Rights in the Age of Artificial Intelligence*. Access Now. 2018. Dostupné na: <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>.
- ARAÚJO, R.; KOESSLER, L. The Rise of the Constitutional Protection of Future Generations. In: *VerfBlog* [online]. 12. 8. 2022. Dostupné na: <https://verfassungsblog.de/the-rise-of-the-constitutional-protection-of-future-generations/>.
- AUSTRALIAN HUMAN RIGHTS COMMISSION. Human Rights and Technology Final Report. 2021. Dostupné z: <https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-final-report-2021>.
- BAJGAR, O. Shared Safety Modules. Working paper, 2019. Dostupné z: <http://bajgar.org/ssm>.
- BUBLITZ, Ch.; WINTER, Ch. In Dubio Pro Futura. A Novel Decision Rule to Preserve the Long-Term Interests of Humanity. In: *VerfBlog* [online]. 19. 8. 2022. Dostupné na: <https://verfassungsblog.de/in-dubio-pro-futura/>.
- CALLAWAY, E. EU law deals blow to CRISPR crops. In: *Nature* [online]. 25. 7. 2018. Dostupné na: <https://www.nature.com/articles/d41586-018-05814-6>.
- COMMITTEE ON LEGAL AFFAIRS (EU). Report with recommendations to the Commission on Civil Law Rules on Robotics. 2017. Dostupné z: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html.
- ČEŠTÍ ELFOVÉ. FAQ: mýty & fakta o elfech. 2022. Dostupné z: <https://cesti-elfove.cz/uvodni-strana/faq-myty-fakta-o-elfech/>.
- EVROPSKÁ KOMISE. Digitální řešení z období pandemie. 2022. Dostupné na: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/digital-solutions-during-pandemic_cs.
- EVROPSKÁ KOMISE. Sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému výboru a sociálnímu výboru a výboru regionů *Umělá inteligence pro Evropu*. COM/2018/237 final. 2018. Dostupné z: <https://eur-lex.europa.eu/legal-content/CS/ALL/?uri=COM:2018:237:FIN>.
- FRIEDERSDORF, C. Immortal but Damned to Hell on Earth
The danger of uploading one's consciousness to a computer without a suicide switch. In: *The Atlantic* [online]. 2015 [cit. 2024-01-07]. Dostupné z: <https://www.theatlantic.com/technology/archive/2015/05/immortal-but-damned-to-hell-on-earth/394160/>.
- FUJÁČEK, J. Vypínání webů a stíhání autorů. Vláda chystá akční plán proti dezinformacím. In: *Echo24* [online]. 30.12.2022. Dostupné na: <https://echo24.cz/a/H38eq/zpravy-domaci-vlada-dezinformace-cenzura-strategie-legislativa-vlada-vypinani-webu-vnitro-rakusan-klima-zmocnenec>.
- FUREDI, F. *Why is Democracy Always Under Threat*. [přednáška]. Praha: Institute H21, 24. září 2022. Dostupné na: <https://youtu.be/xxRHWrejZ60>.
- FUTURE OF LIFE INSTITUTE. Pause Giant AI Experiments: An Open Letter. 2023. Dostupné na: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

GOULD, M.; JOSHI, I.; TANG, M. The power of data in a pandemic. In: *GOV. UK. Blog. Technology in the NHS* [online]. 28. 3. 2020. Dostupné na: <https://healthtech.blog.gov.uk/2020/03/28/the-power-of-data-in-a-pandemic/>.

GRYLLS, G. Ukraine is outflanking Russia with ammunition from Big Tech. In: *The Times* [online]. 24. 12. 2022. Dostupné na: <https://www.thetimes.co.uk/article/ukraine-is-outflanking-russia-with-ammunition-from-big-tech-lxp6sv3qz>.

HAYWARD, E. NHS cancer treatment wait reduced in Palantir software pilot. In: *The Times* [online]. 27. 12. 2022. Dostupné na: <https://www.thetimes.co.uk/article/nhs-cancer-treatment-wait-reduced-in-palantir-software-pilot-wldp568pj>.

HODULÍK, J. Zákon o blokaci dezinfowebů má bránit demokracii, ale podle odborníků ji spíš ohrožuje. In: *Respekt* [online]. 17. 10. 2022. Dostupné na: <https://www.respekt.cz/kontext/zakon-na-blokaci-dezinfowebu-ma-branit-demokracii-ale-podle-odborniku-ji-zatim-spis-ohrozuje>.

CHLUP, R. The Problem of Tackling Disinformation in an Age of Authenticity. [přednáška]. Praha: Institute H21, 15. září 2023. Dostupné na: <https://youtu.be/s12I8SWQn2M?si=SIU4JS2IQ-q1c73l>.

CHLUP, R. Úskalí boje proti dezinformacím ve věku autenticity. In: *blog.aktualne.cz* [online]. 05. 02. 2023. Dostupné na: <https://blog.aktualne.cz/blogy/radek-chlup.php?itemid=44437>.

IGNATIUS, D. How the algorithm tipped the balance in Ukraine. In: *Washington Post* [online]. 19. 12. 2022. Dostupné na: <https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/>.

IURIDICUM REMEDIUM. Konference organizace Iuridicum Remedium ze dne 1. 12. 2023 *Digitální doba s lidskou tváří: symposium o digitálních výzvách*. 2023. Záznam k dispozici zde: https://www.youtube.com/live/N1WcyOymnYM?si=x5mMhZ_Jdo7tEWgR.

JUSTITIA AND THE AUTHORS. The Free Speech Recession Hits Home. Mapping Laws and Regulations Affecting Free Speech in 22 Open Democracies. 2023. Dostupné z: <https://futurefreespeech.org//the-free-speech-recession-hits-home/>.

JUSTITIA. Konference organizace Justitia ze dne 4. 12. 2023 *The State of Free Speech – Challenges and Solutions for a Digital Age*. 2023. Záznam k dispozici zde: <https://youtu.be/XQnNY0hR9fg?si=7SVc6YDheucaOfpH>.

JUSTITIA. New report from Justitia: Digital freedom of expression and social media. 2020. Dostupné z: <https://globalfreedomofexpression.columbia.edu/publications/new-report-from-justitia-digital-freedom-of-expression-and-social-media/>.

KAYALI, L.; BRAUN, E. France pushes tougher EU rules for social media in wake of terror attack. In: *Politico* [online]. 2020. Dostupné z: https://www.politico.eu/article/france-renews-social-media-regulation-push-at-eu-level-in-wake-of-terror-attack/?fbclid=IwAR2YALMayjEcIYwy6ak_GHUDKYhjM4MymDd6tVrtxpHj35CXdO9EflM.

KOLAJA, M. Poučení z Avia law. In: *Kolaja.eu* [online]. 2020. Dostupné na: https://www.kolaja.eu/cs/post/20200512-lessons_learned_from_the_avia_law_warning_french_on-line_hate_speech_law_to_be_adopted_tomorrow/.

KRAKOVNA, V. et al. Specification gaming: the flip side of AI ingenuity. In: *Google Deep Mind* [online]. 21. 4. 2020. Dostupné z: <https://deepmind.google/discover/blog/specification-gaming-the-flip-side-of-ai-ingenuity/>.

LEONHARDT, D. For many, a boom that wasn't. In: *The New York Times* [online]. 9. 4. 2008. Dostupné na: <https://www.nytimes.com/2008/04/09/business/09leonhardt.html>.

MAAS, M. Paths Untaken: The History, Epistemology and Strategy of Technological Restraint, and lessons for AI. In: *VerfBlog* [online]. 9. 8. 2022. Dostupné na: <https://verfassungsblog.de/paths-untaken/>.

MARTÍNEZ, E.; WINTER, Ch. Is Legal Longtermism Common Sense? In: *VerfBlog* [online]. 9. 8. 2022. Dostupné na: <https://verfassungsblog.de/is-legal-longtermism-common-sense/>.

MVČR. Středoevropská identita jako součást protizápadního narativu na kvazi-mediální scéně. Nedatováno. Dostupné na: <https://www.mvcr.cz/chh/clanek/stredoevropska-identita-jako-soucast-protizapadniho-narativu-na-kvazi-medialni-scene.aspx>.

NATIONAL SECURITY COMMISSION ON ARTIFICIAL INTELLIGENCE. Final Report. 2021. Dostupné z: <https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>.

NGO, R. AGI Safety From First Principles. In: *AI Alignment Forum* [online]. 2020. Dostupné z: <https://www.alignmentforum.org/s/mzgtmmTKKn5MuCzFJ>.

OECD. Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. 2019 (novelizováno 2023). Dostupné na: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

PEW RESEARCH CENTER. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable. 2020. Dostupné z: <https://www.pewresearch.org/journalism/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>.

POTTER, B. Why Does Nuclear Power Plant Construction Cost So Much? In: *IFP* [online]. 1. 3. 2023. Dostupné na: <https://ifp.org/nuclear-power-plant-construction-costs/>.

REKONSTRUKCE STÁTU. Doporučení Rekonstrukce státu: Ochrana svobody projevu v době pandemie. 2020. Dostupné na: <https://rekonstrukcestatu.s3.amazonaws.com/rekonstrukcestatu/production/files/2020/05/15/13/11/42/ff26ab4e-b44a-46df-8c6e-41f12fa7df4b/doporuzeni-rekonstrukce-statu-ochrana-svobody-projevu-v-dobe-pandemie.pdf>.

REKONSTRUKCE STÁTU. Monitorovací zpráva Síť k ochraně demokracie za období 1. 1. 2020 – 31. 7. 2020. 2020. Dostupné na: <https://www.rekonstrukcestatu.cz/archiv-novinek/experti-omezovani-obcanskych-prav-v-prvni-vlne-pandemie-bylo-bezprecedentni>.

RŮŽIČKA, A.; HOŘEŇOVŠKÝ, J.; KUŽÍLEK, O. Dezinformace a svoboda projevu v čase války na Ukrajině. In: *Svobodaslova.online* [online]. 4. 3. 2022. Dostupné na: <http://svobodaslova.online/wp-content/uploads/2022/03/Dezinformace-a-svoboda-projevu-Ukrajina.pdf>.

SHLOMO-AGON, S.; SALITERNIK, M. Proactive International Law Keeping Pace with an Accelerated World. In: *VerfBlog* [online]. 2022. Dostupné na: <https://verfassungsblog.de/proactive-international-law/>.

SOLON, O. Peter Thiel: Europe is cracking down on Silicon Valley out of 'jealousy'. In: *The Guardian* [online]. 15. 3. 2018. Dostupné na: <https://www.theguardian.com/technology/2018/mar/15/peter-thiel-silicon-valley-europe-regulation/>.

TELEC, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část I. In: *Právní prostor* [online]. 22.05.2023. Dostupné na: <https://www.pravniprostor.cz/clanky/ostatni-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-i>.

TELEC, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část II. In: *Právní prostor* [online]. 25.05.2023. Dostupné na: <https://www.pravniprostor.cz/clanky/obcanske-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-ii>.

TELEC, I. Nařízení EU o digitálních službách: možné právní nebezpečí – část III. In: *Právní prostor* [online]. 29.05.2023. Dostupné na: <https://www.pravniprostor.cz/clanky/mezinarodni-a-evropske-pravo/narizeni-eu-o-digitalnich-sluzbach-mozne-pravni-nebezpeci-cast-iii>.

THIEL, P. *The 2021 Atlas Society Gala: Peter Thiel's Speech*. 2021. Dostupné na: <https://www.youtube.com/watch?v=YK3Tzx-S264&t=177s>.

THIEL, P. The End of the Future. In: *National Review* [online]. 3. 11. 2011. Dostupné na: <https://www.nationalreview.com/2011/10/end-future-peter-thiel/>.

UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2021. Dostupné z: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

UNIVERSITÉ DE MONTRÉAL. Montreal Declaration for a Responsible Development of Artificial Intelligence. 2018. Dostupné z: <https://montrealdeclaration-responsibleai.com/the-declaration/>.

VALIŠOVÁ, V. K diskusi: Zelený certifikát, Health Passport čili covid pas a bonusy za očkování. In: *Advokátní deník* [online] 7. 4. 2021. Dostupné na: <https://advokatnidenik.cz/2021/04/07/k-diskusi-zeleny-certifikat-health-passport-cili-covid-pas-a-bonusy-za-ockovani/>.

VIJG, J. The Great Technology Slowdown: Looking for the Silver Lining. [přednáška]. Praha: Institute H21, 18. září 2021. Dostupné na: <https://www.youtube.com/watch?v=eqsCNs8trTc>.

WEINSTEIN, E. 2013: What *Should* We Be Worried About? In: *Edge* [online]. 2013. Dostupné z: <https://www.edge.org/response-detail/23879>.

YAMPOLSKIY, R. Could an artificial intelligence be considered a person under the law? In: *The Conversation* [online]. 5. 10. 2018. Dostupné z: <https://theconversation.com/could-an-artificial-intelligence-be-considered-a-person-under-the-law-102865>.

YUDKOWSKY, E. Nearest unblocked strategy. In: *Arbital* [online]. 2015. Dostupné z: https://arbital.com/p/nearest_unblocked.

Právo, společnost a nové technologie 21. století

Abstrakt

Tato práce se soustředí na vztah mezi právními systémy a rychle se vyvíjejícími socio-technologickými změnami ve společnosti. V úvodní obecné části zkoumá otázku evoluce/adaptability právních systémů v reakci na dynamický technologický vývoj, přičemž využívá teoretický rámec systémové teorie. První část také analyzuje různé modely regulace nových technologií, zařazuje je do logiky systémové teorie a ilustruje je na konkrétních příkladech, jako jsou Uber a Airbnb, a na výzvách spojených s umělou inteligencí a velkými daty.

Druhá kapitola obecné části se zabývá právními a kulturními aspekty teze o technologické stagnaci, narůstající komplexitou právních systémů a jejich socio-technologických prostředí a otázkami, jak tyto systémy reagují na sociální odpor vůči novým technologiím a averzi k riziku. Je zde představen trade-off mezi hodnotami inovativnosti a bezpečí. Třetí kapitola rozvíjí téma komplexity právního systému a jejího vztahu k entropii a moderním technologiím.

Ve zvláštní části se práce věnuje specifickým otázkám, jako je postavení umělé inteligence v právu a možnost umělé inteligence stát se subjektem práva, včetně etických a právních implikací. Dále pojednává o integraci lidských práv do vývoje umělé inteligence a předkládá návrh mezinárodního ochranného systému. V neposlední řadě diskutuje potenciální hrozby nových forem vlády zprostředkovaných umělou inteligencí, jako je AI totalitarismus a AI technokracie.

Kapitola věnovaná sociálním sítím a svobodě projevu přezkoumává roli, jakou mají digitální platformy ve veřejném diskurzu, a zkoumá různé přístupy k jejich regulaci, včetně otázky boje proti tzv. dezinformacím. Poslední kapitola se zaměřuje na vliv technologií a práva v kontextu pandemie Covid-19, zkoumá otázky elektronického sledování a covid pasů a možnosti, jak velká data a AI mohou přispět v boji proti pandemii.

Klíčová slova: regulace nových technologií; systémová teorie práva; umělá inteligence; sociální síť; negativní lidská práva; technologická stagnace; svoboda projevu

Law, Society and New Technologies in the 21st Century

Abstract

This work focuses on the relationship between legal systems and society's rapidly evolving socio-technological environment. In the first part, it explores the question of the evolution/adaptability of legal systems in response to dynamic socio-technological developments, using the theoretical framework of systems theory. The first part also analyses different models for regulating new technologies, placing them within the logic of systems theory and illustrating them with concrete examples such as Uber and Airbnb and the challenges associated with artificial intelligence and big data.

The second chapter explores the legal and cultural aspects of the technological stagnation thesis, the increasing complexity of legal systems and their socio-technological environments, and questions about how these systems respond to social resistance to new technologies and risk aversion. The trade-off between innovation and security is presented. The third chapter develops the theme of legal system complexity and its relationship to entropy and modern technology.

In the second part, the thesis addresses specific issues, such as the position of AI in law and the possibility of AI becoming a legal entity, including ethical and legal implications. It also discusses the integration of human rights into the development of AI and presents a proposal for an international regulatory system. Finally, it discusses the potential threats of new forms of AI-mediated government, such as AI totalitarianism and AI technocracy.

The fifth chapter focuses on social networks and freedom of expression, reviews digital platforms' role in public discourse and examines different approaches of regulation.

The final chapter focuses on the impact of technology and law in the context of the COVID-19 pandemic, exploring issues of electronic surveillance and COVID-19 passports, and how big data and AI can contribute to the fight against the pandemic.

Key words: regulation of new technologies; systems theory of law; artificial intelligence; social media networks; negative human rights; technological stagnation; freedom of speech