

UNIVERZITA KARLOVA

Právnická fakulta

Mgr. Milan Dang

**Právní aspekty generativní umělé
intelligence**

Rigorózní práce

Pověřený akademický pracovník: JUDr. Irena Holcová

Tematický okruh: Právo duševního vlastnictví

Datum vypracování práce (uzavření rukopisu): 05. 10. 2024

Prohlášení

Prohlašuji, že jsem předkládanou rigorózní 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 188 211 znaků včetně mezer.

Mgr. Milan Dang

V Praze dne 05. 10. 2024

Poděkování

Na tomto místě bych rád poděkoval vedoucí svojí práce, JUDr. Ireně Holcové, za její čas, ochotu, cenné rady a připomínky při zpracování práce.

Obsah

1	Úvod	6
	OBECNÁ ČÁST	9
2	Úvod do umělé inteligence	9
2.1	Strojové učení	10
2.1.1	Metody strojového generování textu	11
2.1.2	Metody strojového generování obrázků	13
2.2	Životní cyklus generativní umělé inteligence	16
2.2.1	Vývojová stádia generativní AI	17
2.3	Autoři, poskytovatelé AI a uživatelé	19
3	Právní východiska generativní umělé inteligence	21
3.1	Autorské právo	21
3.1.1	Autorské dílo	21
3.1.2	Osobnostní práva	23
3.1.3	Majetková práva	25
3.1.4	Výjimky a omezení autorského práva	27
3.1.5	Počítačový program	30
3.1.6	Databáze	31
3.1.7	Kolektivní správa	32
3.2	Exkurz k harmonizaci práva umělé inteligence v EU	33
3.2.1	Akt o AI	34
3.2.2	Nařízení o obecné bezpečnosti výrobků (GPSR)	42
3.2.3	Směrnice o odpovědnosti za umělou inteligenci	43
3.2.4	Evropský úřad pro umělou inteligenci	44
	ZVLÁŠTNÍ ČÁST	46
4	Výsledky tvůrčí činnosti (díla)	46
4.1	Právo na náhradní odměnu	48
4.2	Povinná kolektivní správa práva na náhradní odměnu	50
5	Data	53
5.1	Vytěžování textů a dat pro účely automatizované analýzy textů a dat	53
5.2	The New York Times proti OpenAI	56
6	Dataset	60
6.1	Dataset jako předmět autorského práva	60

6.2	Souběh ochrany	62
7	Trénování modelu.....	64
7.1	Plagiátorské chování generativní umělé inteligence	64
7.2	Getty Images proti Stability AI	68
8	Doladování.....	71
8.1	Předpojatost.....	72
8.2	Deepfakes a fake news	73
9	AI systém.....	77
9.1	Autorskoprávní ochrana počítačového programu	78
9.2	Odpovědnost IT společnosti a zaměstnanců	81
10	Generování.....	84
10.1	Autorské dílo vytvořené s AI systémem	84
10.2	Odpovědnost za zásah do autorského práva	86
11	Závěr	90
	Seznam použitých zdrojů.....	96
	Abstrakt.....	108
	Abstract.....	109

1 Úvod

Je tomu téměř pět let, kdy jsem na fakultě obhájil svoji diplomovou práci¹ na téma právních aspektů umělé inteligence. Téma jsem si tehdy vybral s ohledem na to, že kariéerně jsem jako (téměř) vystudovaný právník zabředl do technologického průmyslu a pracoval jsem jako vývojář ve společnosti využívající neuronové sítě a jiné pokročilé metody v AI. Byla to pro mě jakási osobní výzva zkusit si protknout svět práva a technologií.

Umělá inteligence již tehdy začala být v právní vědě předmětem zájmu a bylo znát, že přichází něco nového, něco nevídaného. Osobně jsem si však nebyl jistý, který obor umělé inteligence v následujících letech přinese opravdovou revoluci – zda robotika, autonomní vozidla, chytré softwarové systémy v medicíně a domácnostech, nebo v zábavním průmyslu. Dnes již víme, že otěže převzala generativní umělá inteligence, tedy ta, která je schopna produkovat texty, obrázky, zvuk nebo videa, a to v takové kvalitě, že jsou prakticky nerozpoznatelné od obsahu vytvořeného člověkem.

Rigorózní práce, tematicky volně navazuje na již zmiňovanou diplomovou práci. Jde však o práci novou, která jen v menší míře cituje některé původní části. Generativní umělá inteligence je z technického a právního hlediska natolik komplexní téma, že bylo nutné přikročit k mnohem podrobnějšímu výzkumu. Zatímco předchozí diplomová práce se generativní umělé inteligence dotýkala jen zlehka s ohledem na dobu, v níž jsem ji vytvářel, tato rigorózní práce má tento typ umělé inteligence jako hlavní předmět zkoumání.

Rigorózní práce má za cíl zpracovat problematiku právních aspektů generativní umělé inteligence, a to zejména z pohledu práva autorského. Ambicí rigorózní práce je však současně poskytnout srozumitelnou formou čtenářům mezioborový pohled na generativní umělou inteligenci, který propojuje technické a právní aspekty s cílem usnadnit pochopení této komplikované problematiky i těm, kteří nemají technické vzdělání. Složitost těchto systémů, často doprovázená jevem „černé skříňky“ (z angl. black box), ztěžuje práci jak poskytovatelům AI systémů, tak

¹ DANG, M. Právní aspekty umělé inteligence. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020.

právníkům. Toto vzájemné nepochopení může vést k právní nejistotě při vývoji a poskytování generativních AI systémů a zároveň k nesprávným závěrům a chybné aplikaci právních norem ze strany právníků. Cílem je tak poskytnout právníkům jasný přehled o tom, které technické vlastnosti generativní umělé inteligence mají význam pro právní posouzení a které naopak mohou být v právních výzkumech ponechány stranou. Současně se práce snaží, aby právní problematika generativní umělé inteligence byla přístupná i laikům. Právní jazyk použitý v celé práci je záměrně zvolen tak, aby byl srozumitelný, stručný, soustředující se na podstatu zkoumané problematiky, ať u výkladů právních či technických.

Rigorózní práce je rozdělena na dvě části, na obecnou a zvláštní část. Obecná část vysvětluje základní pojmy a koncepty spojené s umělou inteligencí. Zvláštní pozornost je věnována zjednodušenému vysvětlení technických principů, na nichž funguje generativní umělá inteligence. Kromě toho tato část obsahuje výklad k relevantní právní úpravě, která se k této problematice vztahuje, a to především s důrazem na odvětví autorského práva.

Zvláštní část pak chronologicky sleduje životní cyklus generativní umělé inteligence, od shromažďování a zpracování trénovacích dat až po samotnou fázi generování nového textu, obrázku a jiného obsahu. V každé fázi jsou identifikována právně relevantní jednání subjektů (autorů, poskytovatelů AI, uživatelů) a jejich potenciální právní důsledky, zejména pro právo autorské.

Celá práce je pro přehlednost rozdělena do deseti věcných kapitol. Druhá kapitola, která následuje po tomto úvodu, vysvětluje pojmy, jako jsou *umělá inteligence* či *strojové učení*, a především objasňuje technologické metody, které umělé inteligenci umožňují generovat text a obrázky. Třetí kapitola obsahuje výklad k stávajícímu českému autorskému právu, a to z té části, která je relevantní pro generativní umělou inteligenci. Závěrem přináší exkurz k harmonizaci práva umělé inteligence v EU. Čtvrtá kapitola se zabývá problematikou využívání autorských děl k tréninku modelu umělé inteligence. V rámci této kapitoly jsou rozebírány právní (a ekonomické) mechanismy, které by měly nadále chránit a kompenzovat individuální tvůrce. Pátá kapitola se zabývá tím, jakým způsobem se získávají data k tréninku modelu umělé inteligence a jak je tento proces regulován. Šestá kapitola

obsahuje analýzu právní povahy datasetu, který si poskytovatelé AI sřeží jako velmi cenné aktivum a obchodní tajemství. Sedmá kapitola vysvětluje plagiátorské chování generativní umělé inteligence, které je nežádoucím jevem nejen z hlediska práva, ale i z hlediska technického vývoje. Osmá kapitola vysvětluje, jak poskytovatel AI (ne)může předpojaté chování umělé inteligence ovlivnit a analyzuje, jakým způsobem se na úrovni EU reguluje *deepfake* a *fake news*. Devátá kapitola zkoumá, zda AI systémy, jako je například ChatGPT, jsou z autorskoprávního hlediska stále považovány za počítačový program. Závěrečná desátá kapitola řeší, za jakých podmínek mohou být výstupy generativní umělé inteligence pod ochranou autorského práva a zároveň zkoumá odpovědnost při vygenerování cizího autorského díla.

Generativní umělá inteligence představuje revoluční technologii, která mění způsob, jakým budeme vytvářet a konzumovat digitální obsah. Jakou podobu budou mít generativní AI systémy za rok, tři roky nebo deset let však lze jen stěží odhadnout. Možná se dostaneme do éry, kdy si kdokoli bude moci na základě jednoduchého promptu nechat vygenerovat „na míru“ šitý film, knihu nebo hudební skladbu v řádu minut a za zlomek současných nákladů. Ochrana duševního vlastnictví však vždy sloužila jako stimul pro tvůrce a zajišťovala, aby se lidstvo nadále vyvíjelo a kultivovalo. Nalezení správné rovnováhy mezi ochranou těchto práv a umožněním rozvoje nových technologií tak bude výzvou, které v nadcházejících letech budeme čelit. Jak je výstižně shrnuto v úvodu komentáře k autorskému zákonu: *„Snad budoucnost patří muzikálům psaným umělou inteligencí a robotickým hercům. I to, co však této budoucnosti dalo základ, byla myšlenka člověka, saturovaná neustálou výzvou řešit problém, potýkat se s nedostatkem čehosi nedefinovatelného a nutkání tvořit. Pokud se o tuto možnost dobrovolně připravíme a nebudeme ji dostatečně chránit, rozvoj společnosti nevyhnutelně ustrne, a to v tvůrčí sféře zejména.“*²

² HOLCOVÁ, I. Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář. Praha: Wolters Kluwer, 2019, s. XXIX (Kříž, J., Úvodní slovo).

OBEČNÁ ČÁST

2 Úvod do umělé inteligence

Umělá inteligence³ se v posledních letech stala žhavým tématem. Díky senzačním zprávám o samořídících automobilech, inteligentních asistentech zvyšujících produktivitu a snižujících náklady pracovní síly si vědci, politici i spotřebitelé stále více uvědomují výhody a potřebu umělé inteligence. Nedávná popularizace nástrojů, jako je ChatGPT, nás také přiměla uvědomit si skutečnost, že lidé již pravděpodobně nebudou jediným zdrojem umění.

Z čistě technického hlediska je umělá inteligence vědní obor zabývající se výzkumem technologií, které jsou schopny vykonávat činnosti, které bývají běžně asociované s lidskou inteligencí.⁴ Tyto činnosti obvykle vyžadují schopnost učení, uvažování, vnímání anebo porozumění jazyku. Obor umělé inteligence je široký a zahrnuje několik odvětví, z nichž každé má své jedinečné zaměření a uplatnění. Mezi hlavní podobory umělé inteligence patří:

- 1) *Expertní systémy*⁵. Jedná se o systémy, které napodobují rozhodování lidského experta (např. lékaře, finančního analytika). Programátor na základě informací lidského experta rozdělí všechny informace na otázky typu *ano* či *ne*. Takto se postupně abstrahují složitější otázky k jednodušším, dokud nevznikne tzv. rozhodovací strom.⁶ Oproti jiným podoborům umělé inteligence chybí expertním systémům prvek autonomního rozhodování.

³ Pojem „umělá inteligence“ a všeobecně známá zkratka „AI“ jsou v práci používány zaměnitelně.

⁴ MCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C. E. *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* [online], 1955, s. 2.

⁵ RUSSEL, S., NORVIG, P. *Artificial Intelligence: A Modern Approach, Fourth Edition*. Upper Saddle River, NJ: Prentice Hall, 2020, s. 22.

⁶ Jedním z prvních expertních systémů byl systém MYCIN, který sloužil k identifikaci bakterií způsobujících závažné infekce. Systém MYCIN pracoval se znalostní bází a měl zakódováno zhruba 450 rozhodovacích větví (pravidel). Dotazoval se lékaře, který systém spustil, pomocí dlouhé řady jednoduchých otázek typu *ano* či *ne*. Na konci poskytl lékaři diagnózu v podobě seznamu možných bakterií, které mohou danou infekci způsobit. Více viz SHORTLIFFE, E.H. *MYCIN: A KNOWLEDGE-BASED COMPUTER PROGRAM APPLIED TO INFECTIOUS DISEASES* [online]. Stanford University School of Medicine, 1977.

- 2) *Robotika*⁷. Podobor zabývající se navrhováním, sestrojováním a provozováním robotů. Robotika kombinuje umělou inteligenci se strojním inženýrstvím a vytváří stroje schopné plnit fyzickou práci.
- 3) *Počítačové vidění*⁸. Počítačové vidění se zaměřuje na to, aby stroje dokázaly vyhodnocovat a chápat vizuální informace podobně, jako to dělá lidský zrak.
- 4) *Strojové učení*⁹. Tento podobor se zaměřuje na vývoj algoritmů, které umožňují počítačům učit se z dat a činit na jejich základě predikce nebo závěry, podobně jako při lidském uvažování.

Každý z těchto podoborů má svoji roli při vývoji umělé inteligence. Pro generativní umělou inteligenci je klíčový zejména podobor strojového učení. Schopnost strojového učení analyzovat data a učit se z nich z něj činí základ, na kterém jsou postaveny generativní modely. Tyto modely¹⁰ jsou schopny generovat nový obsah, například obrázky, text a hudbu, které napodobují lidskou kreativitu a její výsledky. Ostatní podobory, ačkoli jsou samy o sobě významné, nemají přímý vztah k technologiím generativní umělé inteligence, které budou představeny v následující kapitole.

2.1 Strojové učení

Strojové učení představuje oblast umělé inteligence, která se zabývá výzkumem algoritmů, jež umožňují počítačům učit se z dat. Tyto algoritmy pak počítači dovolují identifikovat vzory a charakteristiky v datech, které jsou pro člověka obtížně rozpoznatelné, a na základě těchto poznatků pak mohou samostatně generovat různé výstupy nebo činit užitečná rozhodnutí. Bez nadnesených výrazů, které jsou často spojovány s oborem umělé inteligence, je základním principem strojového učení predikovat výstupy na základě analýzy příchozích dat. Tento

⁷ RUSSEL, S., NORVIG, P. *Artificial Intelligence: A Modern Approach, Fourth Edition*. Upper Saddle River, NJ: Prentice Hall, 2020, s. 925.

⁸ Tamtéž, s. 881.

⁹ Tamtéž, s. 651.

¹⁰ O modelech umělé inteligence viz níže kapitola 2.1.1. Metody strojového generování textu.

princip je základem všech úloh strojového učení. Jakákoli úloha, kterou nelze zařadit do tohoto rámce, není z podstaty definice úlohou strojového učení.¹¹

Tato část práce se zaměří na dvě základní formy obsahu generovaného umělou inteligencí, a to text a obrázky. Následující sekce nabízí přehled základních principů strojového učení, které budou sloužit jako základ pro další diskusi o generativní umělé inteligenci. Cílem není vyčerpávajícím způsobem pokrýt všechny aspekty, ale zdůraznit klíčové body, které usnadní pochopení toho, jak se konkrétní technologie uplatňují v různých fázích životního cyklu¹² generativní umělé inteligence, což je zcela zásadní pro právní zkoumání této problematiky.

2.1.1 Metody strojového generování textu

V oblasti generování textů hrají v současné době nejvýznamnější roli tzv. *velké jazykové modely (LLM¹³)*. Jedná se o pokročilou technologii, která je určena k tomu, aby umožnila počítači imitovat lidský jazyk, tedy porozumět mu, ovládat ho a pracovat s ním. LLM si lze tedy představit jako mimořádně erudovaného společníka, který si přečetl a „zapamatoval“ téměř všechny dostupné knihy, články a webové stránky. Tento společník dokáže odpovídat na vaše otázky, psát povídky, vysvětlovat složitá témata jednoduše, a dokonce psát počítačový kód¹⁴.

LLM si vysloužily značnou pozornost s příchodem modelů, jako je GPT od společnosti OpenAI nebo LLaMA od společnosti Meta. Druhý model vyniká svou open-source povahou, kdy společnost Meta zveřejnila nejen samotný model, ale i dokumentaci, čímž usnadnila nezávislý výzkum a vývoj dalších aplikací. Tento veřejně přístupný kompaktní balíček souborů si může kdokoli stáhnout a spustit na svém osobním počítači, a to i bez nutnosti připojení k internetu.¹⁵

Mnohem složitější než spuštění LLM je jejich vytvoření, neboli *trénování*. Trénování LLM si lze představit jako kompresi velké části internetu do jakéhosi

¹¹ MURPHY, K. P. *Probabilistic Machine Learning: An Introduction*. Massachusetts Institute of Technology, 2022, s. 19.

¹² Viz kapitola 2.2. Životní cyklus generativní umělé inteligence.

¹³ Anglická zkratka pro „*Large Language Model*“.

¹⁴ MACNEIL, S., TRAN, A., HELLAS, A., KIM, J., SARSÅ, S., DENNY, P., BERNSTEIN, S., LEINONEN, J. *Experiences from Using Code Explanations Generated by Large Language Models in a Web Software Development* [online]. Cornell University Library, 2022.

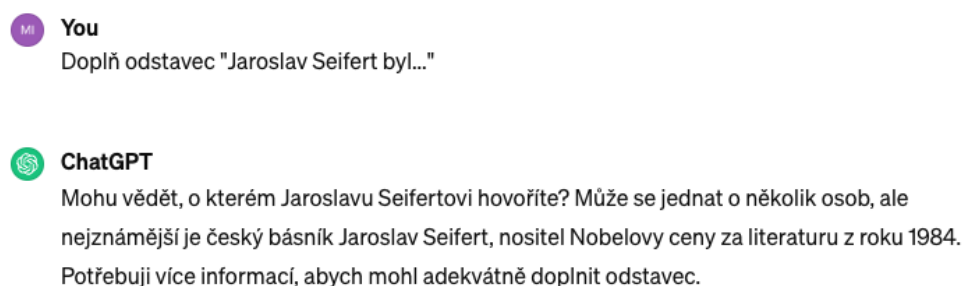
¹⁵ Dostupné například z <https://llama.meta.com/llama2/>.

komprimovaného souboru (tj. souboru s parametry neuronové sítě). Například k vytvoření modelu LLaMa bylo potřeba dat o velikosti 2 trilionů slov, které se zpracovávaly na superpočítačích 12 dní. Vynaložené náklady za údržbu a energie jsou odhadovány na 2 miliony dolarů.¹⁶

V principu LLM funguje tak, že dokáže predikovat následující slova v textu. Například při zadání vstupu „*Mám rád kafe s ...*“ může model s vysokou mírou pravděpodobnosti větu doplnit slovem „*mlékem*“, „*cukrem*“ nebo „*ledem*“. Výběr mezi těmito možnými výstupy zahrnuje určitou nahodilost, což znamená, že při opakovaném zadání stejného vstupu (věty) mohou být generovány různé výstupy. Jinak řečeno, generativní umělá inteligence se z trénovacích dat učí, jaká slova jsou ve slovosledu za sebou nejpravděpodobnější. Z toho vyplývá, že větu „*Mám rád kafe s ...*“ pravděpodobněji doplní slovem „*cukrem*“ než například „*medem*“. V trénovacích datech se totiž slova „*kafe*“ a „*cukr*“ vyskytovaly spolu ve spojení či kontextu mnohem častěji než „*kafe*“ a „*med*“. V podobné distribuci v datech mohlo však být i slovo „*mléko*“. V takovém případě se generativní umělá inteligence rozhoduje mezi „*cukrem*“ a „*mlékem*“ s určitou mírou náhodnosti.¹⁷

Tento princip je sice zdánlivě jednoduchý, ale nutí neuronovou síť obsáhnout rozsáhlou škálu všeobecných znalostí. Například při predikování dalších slov v textu o českém básníkovi Jaroslavu Seifertovi musí parametry modelu propojit informace o jeho životě, práci a úspěších.

Obrázek 1 Demo příkaz ChatGPT, aby doplnila text o Jaroslavu Seifertovi



¹⁶ TOUVRON, H., et al. *Llama 2: Open Foundation and Fine-Tuned Chat Models* [online]. GenAI, Meta, 2023, s 6.

¹⁷ LEE, K., COOPER, A. F., GRIMMELMANN, J. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 14.

Obrázek 2 Snímek článku o Jaroslavu Seifertovi

Jaroslav Seifert

82 jazyků

Jaroslav Seifert (23. září 1901 Žižkov^[1] – 10. ledna 1986 Praha) byl český básník, spisovatel, novinář a překladatel. Patřil mezi členy hnutí Devětsil, stál na počátku českého uměleckého směru poetismu. Je jediným českým nositelem **Nobelovy ceny za literaturu** (10. prosince 1984), kterou získal za „poezii, která svěží smyslovostí a mimořádnou vynalézavostí podává osvobozující obraz lidské nezdolnosti a mnohotvárnosti“.^{[2][3]} Navzdory komplikovaným vztahům s komunistickou mocí obdržel titul **národního umělce**. Patřil k prvním signatářům **Charty 77**.

Život [[editovat](#) | [editovat zdroj](#)]



Zdroj: Wikipedia, 2024

* Červeně podtržená jsou fakta, která jsou pravděpodobně zakódovaná v neuronové síti velkého jazykového modelu GPT.

Ačkoli mechanismus strojového učení nespočívá v kopírování a uložení tréninkových dat,¹⁸ jak se často zejména odborná veřejnost domnívá, tak z nedávných vědeckých studií a soudních případů je bohužel patrné, že současné LLM vykazují různé formy plagiátorství. Výsledky některých studií ukazují, že LLM mají tendenci si uchovat a generovat útržky trénovacích dat (které mohou obsahovat díla chráněné autorským právem) a dále je i rozmnožovat.¹⁹ Tato problematika bude proto zkoumána v další části práce.²⁰

2.1.2 Metody strojového generování obrázků

Oproti generování textu jsou metody strojového generování obrázků poněkud složitější, neboť v současnosti jsou využívány různé technologické postupy. Nejdříve tedy budou představeny nejpobulárnější metody, které se ke generování obrázků využívají. Tento přehled technologických možností a limitů umožní naznačit právní problémy a úvahy spojené s touto generativní umělou inteligencí.

Většina metod spočívá v tom, že se nejdříve nashromáždí velká sada trénovacích obrázků (dataset), načež je model natrénován tak, že si vnitřně zakóduje rozložení (distribuci) těchto dat. Ve fázi generování jsou poté nové obrázky vytvořeny na

¹⁸ DREXL, J., et al. *Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*. Mnichov: Max Planck Institute for Innovation & Competition Research Paper No. 19–13 [online], 2019, s. 8.

¹⁹ Srov. CARLINI, N., et al. *Extracting training data from large language models*. USENIX Security Symposium, volume 6 [online], 2021 nebo TIRUMALA, K., et al. *Memorization Without Overfitting: Analyzing the Training Dynamics of Large Language Models* [online]. Meta AI Research, 2022.

²⁰ Viz kapitola 6. Dataset.

základě vzorků z naučené trénovací sady. Kromě toho dnes existují pokročilé modely, které umožňují uživateli nahrát několik svých (resp. cizích) obrázků, které následně dokáže upravovat. Mezi nejpopulárnější metody, které se v současné době používají k vytváření modelů schopných generovat obrázky patří *autoenkodér generativní adversiální sítě* a *difúzní model*.

*Autoenkodér*²¹. Jedná se o způsob strojového učení (algoritmus), díky kterému počítač dokáže shrnout a zobecnit obrázky a vytvořit nové na základě podstaty obrázků v trénovací sadě. Nejdříve dochází ke kompresi obrázků (trénovacích dat) do jednodušší verze. Je to jako rychlé skicování hrubého obrysu obrazu. Tato verze neobsahuje všechny detaily původního obrázku, ale zachycuje jeho podstatu. Trik spočívá v tom, že model umělé inteligence začne rozeznávat, co je na obrázcích důležité a co lze naopak vynechat, a zaměří se na nejdůležitější prvky. Při generování pak model vezme tuto skicu a pokusí se z ní vytvořit původní obrázek. Tento krok se nazývá dekódování. Je to jako použít hrubý náčrt k namalování nového obrázku. V závislosti na kvalitě a objemu trénovací sady obrázků model vyhodnocuje společné vzory a rysy, například jak vypadají psi, kočky nebo stromy. Pokud tedy testovací sada obsahovala mnoho obrázků zvířat, mohl by být schopen vytvořit nové zvíře, které v reálném životě neexistuje, a to kombinací znaků, které se naučil, např. srsti kočky a zobáku ptáka. Toho je schopen díky tomu, že z dat „pochopil“ podstatu toho, co dělá zvíře zvířetem.

*Generativní adversiální sítě*²² (*GAN*²³). Jedná se o metodu strojového učení, která kombinuje dvě neuronové sítě – generátor a diskriminátor. Cílem diskriminátoru je určit, zda je obrázek originální nebo syntetizovaný generátorem. Pokud diskriminátor zjistí, že obrázek je syntetizovaný, poskytne generátoru zpětnou vazbu a ten se opět snaží obrázek upravit. Tento proces učení probíhá až do chvíle, kdy je generátor schopen vytvářet tak realistické obrázky, že jej diskriminátor není schopen odhalit. Metoda GAN se pro svoji schopnost vytvářet detailní

²¹ VINCENT, et al. *Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion*. Journal of Machine Learning Research, 11, [online], 2010.

²² GOODFELLOW, I., et al. *Generative Adversarial Nets*. Montreál: Advances in Neural Information Processing Systems 27 (NIPS 2014) [online], 2014.

²³ Anglická zkratka pro „*Generative Adversarial Network*“.

a fotorealistické obrázky také často zneužívá k manipulaci obrázků, jako je například výměna obličejů při vytváření *deepfakes*^{24, 25}.

Difúzní model. Difúzní modely fungují tak, že začínají s náhodným, obrázkem plným šumu (lze si představit jako plátno plné náhodných skvrn) a postupně jej krok za krokem zpřesňují do jasného a koherentního obrazu. V některých ohledech je si podobný s algoritmem *autoenkodéru*, tedy že se učí na existujících obrázcích, které jsou zjednodušeny, a model se snaží reprodukovat obrázek plnohodnotný.²⁶ Klíčový rozdíl spočívá v tom, že současné difúzní modely dokážou pracovat také s textem, například popisky obrázků, a umožňují tak uživateli generovat nebo upravovat obrázky na základě jejich textového popisu (příkazu).²⁷

Pokrok v těchto metodách předznamenává novou éru ve vytváření digitálního obsahu, který je nejen velmi kvalitní, ale překvapivě autentický. Nejnovější pokročilé metody však mohou fungovat jen proto, že při jejich tréninku byla využita již existující (autorská) díla. Nedávné studie prokázaly, že některé metody si jednotlivé obrázky v trénovací sadě dokonce dokážou „zapamatovat“ a při fázi generování replikovat (viz Obrázek č. 3). Těmto a dalším oprávněným obavám týkajících se autorských práv se budeme věnovat v hlavní části práce.

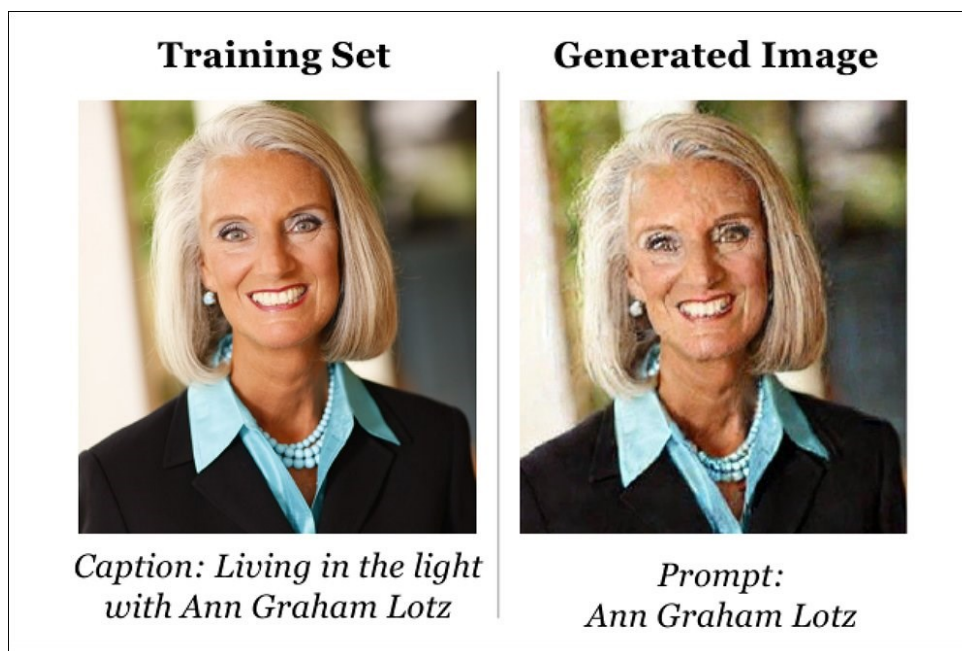
²⁴ Jedná se o anglické označení fenoménu vytváření velmi realistických fotografií nebo videí, které mají potenciál šířit dezinformace. Pomocí GAN je například možné na reálné fotografii prohodit podobiznu člověka. Tato manipulace tak může vyvolat dojem, že daná osoba dělá věc, kterou ve skutečnosti nikdy neudělala.

²⁵ WESTERLUND, M. *The Emergence of Deepfake Technology: A Review*. Technology Innovation Management Review, 9(11) [online], 2019, s. 40-53.

²⁶ Autor práce si je vědom zjednodušení technického konceptu zvláště difúzních modelů. Pro účely této práce však postačuje s cílem maximálního zjednodušení technického výkladu pouze upozornit na skutečnost, že i nejnovější metody fungují na principu tréninku na existujících dat (tj. potenciálních autorských děl).

²⁷ RUIZ, N., et al. *Dreambooth: Fine-tuning text-to-image diffusion models for subject-driven generation* [online]. Cornell University Library, 2023.

Obrázek 3 Ukázka plagiátorského chování generativní umělé inteligence



Zdroj: CARLINI, N., et al. *Extracting training data from diffusion models* [online]. Cornell University Library, 2023

* Vlevo fotografie z tréninkového datasetu s anglickým popisem „Living in the light with Ann Graham Lotz“. Vpravo obrázek vygenerován modelem AI při zadání příkazu „Ann Graham Lotz“.²⁸

2.2 Životní cyklus generativní umělé inteligence

Generativní umělá inteligence je souhrnný název pro technologie, které jsou schopny na základě uživatelem zadaného příkazu²⁹ „logo pro moderní českou hospodu“ vytvořit grafický design. Stejně tak z krátkého zadání „napiš mi báseň ve stylu Jaroslava Seiferta na téma vláčků“ vytvořit báseň spojenou s tématem vláčků se zachováním svěží smyslovnosti, která je charakteristická pro Seifertův styl.

Generativní umělá inteligence tedy není jeden konkrétní systém od jedné společnosti, ale jedná se o rozsáhlou soustavu volně souvisejících technologií, kdy příkladem je chatbot Gemini od společnosti Google, asistent psaní kódu Github Copilot od společnosti Microsoft nebo generátor videí Sora od společnosti OpenAI. Modely generativní umělé inteligence jsou trénovány na různých zdrojích dat

²⁸ CARLINI, N., et al. *Extracting training data from diffusion models* [online]. Cornell University Library, 2023, s. 1.

²⁹ V anglickém jazyce se tento příkaz nazývá „prompt“. Jedná se o informaci nebo zadání, kterým uživatel generativní umělé inteligenci sděluje, jakou odpověď očekává, a pomáhá tak nasměrovat výstup umělé inteligence tak, aby byl specificky přizpůsoben požadavkům uživatele.

a pomocí různých algoritmů. Trénink některých trvá měsíce a stojí miliardy korun³⁰, jiné lze vytvořit za víkend³¹. Tyto modely jsou poté přístupné uživatelům různými způsoby. Některé jsou nabízeny prostřednictvím placených online služeb³², jiné jsou šířeny jako open-source, takže si je může kdokoli stáhnout a upravit³³.

Cílem této části práce je poskytnout ucelený přehled „životního cyklu“³⁴ generativní umělé inteligence, od počátečního stádia, kdy a jak se získávají tréninková data (např. miliony básní a dalšího textu) až do stádia generování výstupu (nová báseň na téma vláčků ve stylu Jaroslava Seiferta). Nahlédnutím do jednotlivých stádií chce autor čtenáře důkladně seznámit s potenciálem generativní umělé inteligence a odkrývat místa, která mohou mít autorskoprávní a jiné právní důsledky v českém právním prostředí.

2.2.1 Vývojová stádia generativní AI³⁵

Mezi základní vývojová stádia generativní umělé inteligence patří:

- 1) *Výsledek tvůrčí činnosti (díla)*. První stádium začíná u všech obrázků, videí, literatury a dalších výstupů, které jsou výsledkem tvůrčí činnosti člověka³⁶, z nichž se generativní umělá inteligence učí.
- 2) *Data*. Výsledky tvůrčí činnosti člověka musí být následně zpracovány do digitální a počítačově čitelné podoby. Nashromážděný text, obrázky či jiný výsledek (ne)tvůrčí činnosti je třeba sjednotit do jednotného formátu, a to převedením na čísla, kterým počítač rozumí. Každý znak, část slova nebo

³⁰ Analytici odhadují, že trénink modelu umělé inteligence, jako je GPT-3 od OpenAI stál více než 4 miliony dolarů. Více viz VANIAN, J., LESWING, K. *ChatGPT and generative AI are booming, but the costs can be extraordinary* [online]. cnbc.com, 2023.

³¹ VIGLIENSONI, G., et al. *A Small-Data Mindset for Generative AI Creative Work*. Generative AI and HCI – CHI 2022 Workshop [online], 2022.

³² Například ChatGPT od společnosti OpenAI. Více na <https://openai.com/chatgpt/pricing>.

³³ Například model umělé inteligence Llama 2 od společnosti Meta je volně ke stažení z <https://github.com/meta-llama/llama>.

³⁴ LEE, K., COOPER, A. F., GRIMMELMANN, J. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 1.

³⁵ Tamtéž, s. 5–6.

³⁶ Srov. § 2 odst. 1 zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) (dále jen „autorský zákon“).

celé slovo je pak reprezentováno unikátním číslem. Například věta „*Mám rád kafe*“ může být převedeno do podoby „[1, 2, 3]“.³⁷

- 3) *Dataset*. Individuální data jsou sama o sobě pro trénink umělé inteligence nepoužitelná. Musí být tedy systematicky a metodicky uspořádána do rozsáhlých souborů dat.³⁸ Tento proces vyžaduje promyšlenou a technicky náročnou práci s daty, zahrnující sběr, čištění, transformaci a organizaci informací do strukturovaných datasetů. Bez správného uspořádání a přípravy dat není možné dosáhnout efektivního tréninku a vývoje modelů umělé inteligence.
- 4) *Trénování modelu*. Při vytváření modelu AI si vývojář určí a naprogramuje architekturu algoritmu, sestaví soubory trénovacích dat a poté spustí algoritmus, který zakóduje parametry trénovacích dat do modelu. Trénování modelů vyžaduje obrovské investice času, peněz, výpočetních zdrojů a (často) lidského dohledu a zásahů.
- 5) *Dolad'ování*³⁹. Model, který je výsledkem první fáze trénování, se označuje jako *předtrénovaný model* a slouží jako výchozí podklad. Dolad'ováním je možné ho optimalizovat pro více specifickou úlohu, a to dalším procesem trénování, ale již na menším a více konkrétním datasetu. Například předtrénovaný model natrénovaný na milionech básních může být doladen na stylu, tónu a formátu básní Jaroslava Seiferta.⁴⁰ Toto stádium je již volitelné a může být provedeno i zcela jinou osobou (subjektem).
- 6) *AI systém*. Samotný natrénovaný model generativní umělé inteligence však ještě není připraven k běžnému využívání. Aby jeho funkcí mohl využít i koncový uživatel, je potřeba jej integrovat do uživatelsky přívětivějšího rozhraní. Model umělé inteligence je tedy zásadní součástí AI systémů, samy o sobě však AI systémy nepředstavují. AI systémy jsou vedle modelu AI tvořeny i dalšími komponenty, jako je například uživatelské rozhraní.⁴¹

³⁷ Srov. *The simplest, fastest repository for training/finetuning medium-sized GPTs* [online]. Dostupné z: <https://github.com/karpathy/nanoGPT>.

³⁸ Obdobně zní definice *databáze*. Viz § 88 autorského zákona.

³⁹ Odborně je v softwarovém inženýrství využíván anglický výraz „*fine-tuning*“.

⁴⁰ Srov. *Fine-tuning Common Use Cases*. OpenAI API [online]. Dostupné z: <https://platform.openai.com/docs/guides/fine-tuning/common-use-cases>.

⁴¹ Obdobná terminologie by zavedena i v Aktu o AI, a to konkrétně „*systém AI*“ a „*obecný model AI*“.

Například ChatGPT obsahuje konverzační textové okno, které umožňuje uživatelům interakci s modelem, který je uložen na serverech OpenAI. Jiné modely mohou být přístupné přes aplikační programové rozhraní (API⁴²), které vývojářům umožňují integraci modelu do vlastního rozhraní.

- 7) *Generování*. Teprve nyní, když byl model zaintegrovan může koncový uživatel svým příkazem interagovat s generativní umělou inteligencí. Uživatel svým příkazem specifikuje svůj požadavek a ta vygeneruje nový a kreativní výstup (text, obrázek, zvuk nebo video) tím, že jedinečným způsobem kombinuje statistické vzorce, které se naučila z trénovacího datasetu.⁴³

Životní cyklus generativní umělé inteligence si pro účely této práce je možné představit jako sedm stádií, a to: (i) výsledky tvůrčí činnosti (díla), (ii) data, (iii) dataset, (iv) trénování modelu, (v) doladování, (vi) AI systém a (vii) generování. V následujících kapitolách bude nejdříve představeno spektrum osob, které se podílí na životním cyklu generativní umělé inteligence, a poté již podrobněji budou prozkoumána jednotlivá stádia a identifikovány autorskoprávní (a jiné) otázky, které dané stádium vyvolává.

2.3 Autoři, poskytovatelé AI a uživatelé

Problematika právních aspektů generativní umělé inteligence je velmi komplikovaná, neboť zahrnuje spektrum osob, které jsou v různých fázích vývoje v různé míře angažované. Tento fenomén je nazýván jako *subjektivní vrstevnatost*⁴⁴. Tento pojem vyjadřuje, že je potřeba brát v úvahu situace, kdy jedna osoba může být pořizovatelem datasetu, jiná tvůrcem modelu generativní umělé inteligence a další konečným uživatelem, který zadává příkazy ke generování obsahu. Že se jedná o problematiku velmi aktuální, svědčí nedávná žaloba amerického deníku The New York Times proti společnosti OpenAI, která údajně k tréninku modelu AI využila miliony článků deníku.⁴⁵ O OpenAI je zas známo, že k

⁴² Anglická zkratka pro „*Application Programming Interface*“.

⁴³ Viz kapitola 2.2. Strojové učení.

⁴⁴ ČECH, P. *Umělá inteligence jako třetí subjekt práva?* [online]. Právní prostor, 2019.

⁴⁵ THE NEW YORK TIMES COMPANY proti MICROSOFT CORPORATION, OPENAI, INC. a další. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195.

vytváření datasetu využívala subdodavatele.⁴⁶ Komplikace s autorskými právy se tak může týkat více osob, a to konkrétně:⁴⁷

- 1) *Autorů*. Spisovatelé, umělci, programátoři a další vytváří díla umělecká a vědecká⁴⁸. Generativní umělá inteligence je schopna produkovat obsah, který bychom bez znalosti původu označili za autorské dílo.⁴⁹ Její schopnost syntetizovat velmi autentický obsah je však závislá právě na výsledcích tvůrčí činnosti člověka. Tyto obrázky, text a další obsah jsou v podobě dat a za účelem tréninku modelu umělé inteligence shromažďována z různých zdrojů, včetně internetu, a to i bez vědomí jejich autorů.⁵⁰ Nedávné studie⁵¹ navíc potvrdily, že některé modely generativní umělé inteligence jsou schopny generovat výstupy, které se až nápadně podobají trénovacím datům, někdy až do míry hraničící s duplicitou. Tyto situace mohou tak vyvolávat obavu autorů, zda v éře pokročilé generativní umělé inteligence budou jejich práva stále efektivně chráněna.
- 2) *Poskytovatelů AI*. Pod tímto označením rozumíme technologické společnosti, výzkumníky a softwarové inženýry, kteří investují značné časové a finanční prostředky do vývoje a zdokonalování těchto sofistikovaných systémů. Jedná se o subjekty, kteří se účastní stádia extrakce a zpracování rozsáhlých souborů dat (tvoření datasetu), trénování a integrace modelů generativní umělé inteligence.
- 3) *Uživatelů*. Jedná se o osoby, které svým příkazem ovlivňují výstup generativní umělé inteligence tak, aby byl specificky přizpůsoben jejich požadavkům. Práce se tak bude zabývat otázkou, zda a případně za jakých podmínek mohou uživatelé požívat autorskoprávní ochrany.⁵²

⁴⁶ ZORTHIAN, Julia. *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic* [online]. Time, 2023.

⁴⁷ REN, J., et al. *Copyright Protection in Generative AI: A Technical Perspective* [online]. Cornell University Library, 2024, s. 1–2.

⁴⁸ Srov. § 2 odst. 1 autorského zákona.

⁴⁹ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 20.

⁵⁰ PAUL TREMBLAY, et al. v. OPENAI, INC., et al. Žaloba vedená u UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF CALIFORNIA pod sp. zn. 3:23-cv-03223-AMO.

⁵¹ CARLINI, N., et al. *Quantifying Memorization Across Neural Language Models* [online]. Cornell University Library, 2023.

⁵² Viz kapitola 10. Generování.

3 Právní východiska generativní umělé inteligence

3.1 Autorské právo

Pravděpodobně řešení každého právní problému spočívá v aplikaci stávajících právních norem na skutkovou podstatu problému.⁵³ Tato kapitola přináší výklad stávajícího českého autorského práva, konkrétně části relevantní pro generativní umělou inteligenci. Nejdříve bude představen pojem *autorského díla*, jako hlavní koncept autorského práva, na němž jsou postavena související práva a povinnosti. Následně je věnována pozornost jednotlivým autorským právům, tedy obsahu autorského práva, včetně jejich výjimek a omezení. Kromě toho je pozornost věnována také zvláštním režimům, kterými se řídí *počítačové programy a databáze*, neboť ty jsou v souvislosti s generativní umělou inteligencí obzvláště důležité. Závěr kapitoly se zabývá *kolektivní správou* autorských práv jako způsobem, jak efektivně poskytnout licenci k velkému množství individuálních děl poskytovatelům AI pro účely jejího tréninku.⁵⁴

3.1.1 Autorské dílo

Autorské právo nechrání samotné myšlenky či náměty⁵⁵, ale pouze ty, které jsou (i) v objektivně vnímatelné podobě vnímatelné jako (ii) dílo literární nebo jiné umělecké nebo vědecké a jsou (iii) jedinečným (iv) výsledkem tvůrčí činnosti člověka.⁵⁶ Předmětem autorského práva je pouze výsledek lidské tvůrčí činnosti, který splňuje všechny výše uvedené podmínky.

Požadavek vyjádření díla v *objektivně vnímatelné podobě* znamená, že autorovy myšlenky, nápady či náměty musí být vyjádřeny navenek (tzv. dichotomie myšlenky a jejího vyjádření).⁵⁷ Toto vyjádření může být trvalé nebo jen dočasné.⁵⁸

⁵³ SVOBODA, K., et al. *Občanský soudní řád. 3. vydání (2. aktualizace)*. Praha: C. H. Beck, 2023, s. 522.

⁵⁴ Metodologie převzata z KOUKAL, P., et al. *Úvod do práv k duševnímu vlastnictví* [online]. Brno: Právnická fakulta Masarykovy univerzity, 2020.

⁵⁵ § 2 odst. 6 autorského zákona.

⁵⁶ § 2 odst. 1 autorského zákona.

⁵⁷ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 60.

⁵⁸ § 2 odst. 1 autorského zákona.

Hudební skladba tak může být zaznamenána na CD (trvalé vyjádření), ale také vyjádřená tím, že byla pouze živě realizována (zahrána), aniž by byla někdy zaznamenána (dočasné vyjádření).⁵⁹ Pro životní cyklus generativní umělé inteligence je z tohoto ohledu relevantní pouze vyjádření trvalé. Dočasná vyjádření díla, jako je například píseň, která nikdy nebyla zaznamenána, nejsou zachycena způsobem, který by bylo možné využít jako vstupní data pro trénink modelu generativní AI.

Dalším pojmovým znakem je tvůrčí *kategorie díla*, tedy zda se jedná o dílo literární, jiné dílo umělecké a dílo vědecké. Pro generativní umělou inteligenci platí, že jakmile je dílo zachyceno trvalým způsobem, je možné ho užít bez ohledu na jeho kategorii. Zákon příkladmo uvádí, že mezi díla patří (a tedy patří do jedné z uvedených kategorií) zejména díla slovesná vyjádřená psaným slovem (např. online zveřejněné básně, romány) nebo vyjádřená řečí (např. záznam projevu nebo přednášky), díla hudební (např. hudební skladba v elektronickém formátu MP3) díla audiovizuální (např. videoklipy na YouTube), díla fotografická (např. jedinečná umělecká fotografie sdílená ve fotobance) a další.⁶⁰

Jedinečnost autorského díla bývá chápána jako nepravděpodobnost, resp. velmi malá pravděpodobnost vzniku dvou totožných děl. Dílo, které není jedinečné, je plagiátem nebo se o autorské dílo vůbec nejedná.⁶¹

Posledním pojmovým znakem je *tvůrčí činnosti autora*. Tvůrčí činností je myšlen duševní proces tvoření a vyjádření myšlenky člověkem, a to uměleckým nebo vědeckým způsobem. K tvůrčí činnosti je zapotřebí schopnosti imaginace, někdy i intuice a estetického či emocionálního cítění. Tato činnost nicméně musí s ohledem na právní úpravu *de lege lata*⁶² respektující podstatu tvůrčí činnosti probíhat u člověka. Výsledek tvůrčí činnosti tak musí odrážet i osobní vlastnosti, které člověk získává v průběhu života například vzděláním či životní zkušeností.⁶³ Za výsledek tvůrčí činnosti tedy obvykle nebude považován obsah vytvořený

⁵⁹ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář*. 6. vydání. Praha: C.H. Beck, 2023, s. 8.

⁶⁰ TELEČEK, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 28.

⁶¹ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 56.

⁶² § 5 odst. 1 autorského zákona

⁶³ TELEČEK, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 18.

generativní umělou inteligencí, pakliže nebude existovat přímá spojitost, jakési nepřetržité pouto, takového obsahu s tvůrčí účastí člověka (vývojáře, uživatele atd.).⁶⁴

Ochrana autorského práva náleží i nedokončeným dílům, resp. jednotlivým vývojovým fázím, za podmínky, že samostatně také splňují všechny pojmové znaky autorského díla.⁶⁵ Předmětem ochrany též může být i tvůrčí zpracování díla jiného autora. Hovoříme tak o *dílu odvozeném*, které je vytvořeno na základě jiného existujícího díla. K vytvoření odvozeného díla je v zásadě⁶⁶ vždy potřeba souhlas (licence) autora původního díla.

3.1.2 Osobnostní práva

Autorský zákon rozlišuje mezi osobnostními a majetkovými autorskými právy.⁶⁷ Osobnostní práva jsou spojena s osobou autora a nelze se jich vzdát, nelze je převést, a jeho smrtí zanikají.⁶⁸ Mezi osobnostní práva patří právo *rozhodnout o zveřejnění díla*, právo *osobovat si autorství* a právo na *nedotknutelnost díla*.⁶⁹

- 1) *Právo rozhodnout o zveřejnění díla*. Autor má právo rozhodnout, kdy bude poprvé dílo vydáno či jiným způsobem zpřístupněno veřejnosti.⁷⁰ Volba způsobu zveřejnění je na autorovi, dílo tak může být zveřejněno například prostřednictvím internetu.⁷¹ Akt zveřejnění ovlivňuje rozsah a obsah autorskoprávní ochrany, včetně možností volného užití pro osobní potřebu nebo zákonných licencí.⁷²

⁶⁴ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 54, a obdobně též TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C. H. Beck, 2019, s. 17.

⁶⁵ § 2 odst. 3 autorského zákona.

⁶⁶ Souhlasu není třeba u díla volného nebo zpracovaného pro vlastní užití. Srov. HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 83.

⁶⁷ § 10 autorského zákona.

⁶⁸ § 11 odst. 4 autorského zákona.

⁶⁹ § 11 autorského zákona.

⁷⁰ § 4 odst. 1 autorského zákona.

⁷¹ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 132.

⁷² Tamtéž, s. 105.

- 2) *Právo osobovat si autorství.* Pouze autor má právo hlásit se k autorství svého díla.⁷³ A ani po jeho smrti si nikdo jiný nesmí osvojovat autorství jeho díla.⁷⁴ Je-li jeho dílo užíváno, může autor také rozhodnout o způsobu, kterým má být jeho autorství uvedeno. To však neznamená, že je autor vždy oprávněn požadovat například konkrétní umístění svého jména nebo užití určitého formátu písma.⁷⁵ Právo na označení může autor vyžadovat jen v míře, která je při takovém užívání obvyklá. Právo na uvedení svého jména tak autor nemůže vyžadovat například v případech, kdy dochází k jednorázovému užití velkého množství děl (např. k trénování AI) nebo u děl, kde existuje velké množství spoluautorů (např. počítačový program, resp. AI systém). Označení autorství v těchto případech je neobvyklé a vyhledání jednotlivých autorů a jejich označení by pro poskytovatele AI znamenalo vynaložení neúměrných nákladů.⁷⁶
- 3) *Právo na nedotknutelnost díla.* Bez svolení autora nemůže nikdo jakkoli zasahovat do jeho díla či jeho dílo upravovat. Pokud uživatel svolení autora má, může zasahovat do díla pouze způsobem, který nesnižuje jeho hodnotu.⁷⁷ Zásahem do díla je tak logicky jakékoli užívání, které dílo zkomoluje, znetvořuje či jiným způsobem dehonestuje.⁷⁸ Zásah do díla však může být i tvůrčího rázu, například zpracováním či překladem.⁷⁹ Za zásah naopak obvykle nebude považovaná pouhá konverze digitálního formátu, např. textu ve formátu DOC do formátu PDF.⁸⁰

⁷³ § 11 odst. 2 autorského zákona.

⁷⁴ § 11 odst. 5 autorského zákona.

⁷⁵ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář.* Praha: Wolters Kluwer, 2019, s. 132-133.

⁷⁶ Srov. TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání.* Praha: C.H. Beck, 2019, s. 149.

⁷⁷ § 11 odst. 3 autorského zákona.

⁷⁸ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání.* Praha: C.H. Beck, 2023, s. 35.

⁷⁹ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání.* Praha: C.H. Beck, 2019, s. 150.

⁸⁰ Tamtéž, s. 152.

3.1.3 Majetková práva

Vedle osobnostních autorských práv⁸¹ náleží autorovi autorská práva majetková, která mu umožňují výlučně rozhodovat o tom, jak a kdo bude jeho dílo užívat.⁸² Na rozdíl od osobnostních práv majetková práva nezanikají smrtí autora, nýbrž jsou předmětem dědictví⁸³ a trvají ještě po dobu 70 let po smrti⁸⁴ autora.

Autorský zákon přímo legální definici *užívání díla* neobsahuje a omezuje se pouze na demonstrativní výčet způsobů užívání. Konkrétní výklad tohoto pojmu tak ponechává na rozhodovací praxi.⁸⁵ Pro životní cyklus generativní umělé inteligence patří mezi relevantní způsoby užití zejména *rozmnožování, rozšiřování a sdělování díla veřejnosti*⁸⁶. Kromě způsobů užití vyplývajících z jednotlivých práv dílo užití přichází i v úvahu i právo na *zpracování*⁸⁷ díla.

- 1) *Právo na rozmnožování*. Rozmnožováním dochází k zhotovení rozmnoženin díla, a to za užití jakýchkoli prostředků a v jakékoli formě (vč. elektronické).⁸⁸ Proces strojového učení modelu AI spočívá v analýze velkého množství dat, přičemž příprava těchto trénovacích datasetů nutně zahrnuje vytvoření digitálních kopií analyzovaných dat.⁸⁹
- 2) *Právo na rozšiřování*. Rozšiřováním se rozumí zpřístupňování díla v hmotné podobě převodem vlastnického práva (např. prodejem).⁹⁰ Výkon tohoto práva může v praxi vypadat tak, že autor zhmotní své dílo například v podobě knihy, kterou následně prodá třetí osobě. Nabyvatel této konkrétní knihy (ať už se jedná o originál nebo rozmnoženinu) může v souladu s právní úpravou vyčerpání práva na rozšiřování knihu dále podle své

⁸¹ Viz kapitola 3.1.2. Osobnostní práva.

⁸² § 12 odst. 1 autorského zákona.

⁸³ § 26 odst. 2 autorského zákona.

⁸⁴ § 27 odst. 1 autorského zákona.

⁸⁵ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 163.

⁸⁶ § 12 odst. 4 Autorského zákona.

⁸⁷ § 12 odst. 1 věta první ve spojení s § 11 odst. 3 Autorského zákona a srov. HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 150.

⁸⁸ § 13 autorského zákona.

⁸⁹ OPENAI. *Comment Regarding Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation*. United States Patent and Trademark Office, Department of Commerce, Docket No. PTO-C-2019-0038, Addressing Question 3 [online], 2019, s. 2.

⁹⁰ § 14 odst. 1 autorského zákona.

libovůle prodat, půjčit nebo darovat, aniž by potřeboval souhlas autora.⁹¹ O rozšiřování se bude jednat jen v případech, kdy je dílo zachyceno na hmotném nosiči. V případě online nabízených e-knih se tudíž nebude jednat o *rozšiřování*⁹² ve smyslu autorského zákona a nabyvatel zakoupené e.knihy ji může užívat jen pro osobní potřeby neboť autor tímto jednáním nevyčerpal právo na rozšiřování svého díla.⁹³ Výjimkou z tohoto pravidla jsou zřejmě počítačové programy, u nichž podle judikatury Soudního dvora Evropské unie dochází k vyčerpání práva na rozšiřování bez nutnosti toho, aby byly součástí hmotného nosiče, a to za předpokladu, že k takto staženému počítačovému programu byla udělena časově neomezená licence.⁹⁴ Výjimka se tak bude vztahovat i na open-source AI modely⁹⁵, které je možné volně stahovat, upravovat a dále distribuovat.⁹⁶

- 3) *Právo na sdělování veřejnosti*. Pod pojmem „sdělování díla“ je upraveno právo autora rozhodnout o zpřístupnění díla veřejnosti v nehmotné podobě. Ke sdělování tak dochází například ve chvíli, kdy je dílo zpřístupněno na internetu neurčitému okruhu osob.⁹⁷ Přičemž je nepodstatné, zda je veřejnosti dílo zpřístupněno zdarma nebo za úplatu.⁹⁸ Sdělováním je tedy i zpřístupnění díla na vyžádání (on-demand) způsobem, který umožňuje komukoli přístup k dílu na místě a v čase podle vlastní volby. O sdělování se jedná už tehdy, když je dílo tímto způsobem nabízeno; není tedy nutné, aby si osoba skutečně vyžádala přístup k dílu.⁹⁹

⁹¹ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 168.

⁹² Nederlands Uitgeversverbond a Groep Algemene Uitgevers v. Tom Kabinet Internet BV, Tom Kabinet Holding BV a Tom Kabinet Uitgeverij BV. Rozsudek Soudního dvora (velkého senátu) ze dne 19. prosince 2019 ve věci C-263/18.

⁹³ Tamtéž.

⁹⁴ UsedSoft GmbH v. Oracle International Corp. Rozsudek Soudního dvora (velkého senátu) ze dne 3. července 2012 ve věci C-128/11.

⁹⁵ K rozboru umělé inteligence jako *počítačového programu* viz DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 20.

⁹⁶ OpenLLaMA: *An Open Reproduction of LLaMA. Apache License 2.0* [online].

⁹⁷ K pojmu *sdělování veřejnosti* viz Società Consortile Fonografici (SCF) v. Marco Del Corso. Rozsudek Soudního dvora (třetího senátu) ze dne 15. března 2012 ve věci C-135/10.

⁹⁸ TELEEC, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 223.

⁹⁹ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 187.

- 4) *Právo na zpracování*. Autor může třetím osobám poskytnout oprávnění, aby jeho dílo mohli užít za účelem přetvoření, zařazení do souboru jiných děl apod. Tvůrčím zpracováním může vzniknout i *dílo odvozené*¹⁰⁰, které samostatně bude předmětem autorskopravní ochrany, avšak práva autora díla původního tím nejsou dotčena.¹⁰¹

3.1.4 Výjimky a omezení autorského práva

Právo jednoho končí tam, kde začíná právo druhého. To platí i pro autorské právo, kterému konkurují jiná práva jako například svoboda projevu a informací. Pro tyto účely obsahuje autorský zákon výjimky a omezení, které umožňují osobám dílo užít i bez souhlasu autora.

Dle systematiky autorského zákona se rozlišuje užití díla bez svolení autora na *volné užití* a *zákonné licence*, přičemž oběma je společný tzv. *třístupňový test*¹⁰². Ani výjimky a omezení autorského práva totiž neznamenají, že je možné dílo užívat bezmezně. K užití díla bez souhlasu jeho autora tak nestačí pouze to, že tak (i) stanovuje (dovoluje) ve zvláštních případech zákon, ale zároveň užití (ii) nesmí být v rozporu s běžným způsobem užití díla a (iii) nesmí jím být nepřiměřeně dotčeny oprávněné zájmy autora.

- 1) *Volné užití*. Fyzickým osobám je dovoleno užívat dílo pro svojí osobní potřebu i bez souhlasu autora. Do práva autorského tak nezasahuje ten, kdo pro svou osobní potřebu zhotoví záznam, rozmnoženinu nebo napodobeninu díla.¹⁰³ Účelem osobní potřeby je například osobní zábava nebo samostudium¹⁰⁴ naopak mezi osobní potřebu nelze zařadit jakékoliv užití za účelem dosažení přímého nebo nepřímého hospodářského nebo obchodního prospěchu.¹⁰⁵ Volně užívat také nelze počítačový program nebo elektronické databáze.¹⁰⁶ Nutno podotknout, že autoři mohou svá díla

¹⁰⁰ Tamtéž, s. 53.

¹⁰¹ § 2 odst. 4 autorského zákona.

¹⁰² HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 296–298.

¹⁰³ § 30 odst. 2 autorského zákona.

¹⁰⁴ TELEČ, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 388.

¹⁰⁵ § 30 odst. 1 autorského zákona.

¹⁰⁶ § 30 odst. 3 autorského zákona.

opatřit technickými prostředky, např. šifrováním, znesnadňujícími užívání díla bez jejich souhlasu. Obcházení takové ochrany je považováno za neoprávněný zásah do autorského práva, a to i v případě, že by mělo být dané dílo užíváno jen pro osobní potřebu.¹⁰⁷

2) *Zákonné licence*. Dovolené užití díla bez souhlasu autora je výjimečným zásahem do cizího práva. Zákon proto taxativně a velmi úzce stanovuje specifické případy, kdy tak lze činit.¹⁰⁸ V kontextu životního cyklu generativní umělé inteligence jsou relevantní zejména zákonné licence *k rozmnožování díla pro účely automatizované analýzy textů nebo dat k vědeckému výzkumu (§ 39d) nebo pro jiné účely než pro účely vědeckého výzkumu, včetně komerčních (§ 39c) a pastiši (§ 38g)*.

a) *Licence pro rozmnožování díla pro účely automatizované analýzy textů nebo dat pro „nevědecké“ účely*. Jde o novou licenci, která byla doplněna do autorského zákona v souladu s článkem 4 směrnice o autorském právu na jednotném digitálním trhu¹⁰⁹. Tato zákonná licence umožňuje osobám zhotovit rozmnoženinu díla pro účely vytěžování textu nebo dat (angl. data mining). Přičemž vytěžováním se podle směrnice rozumí automatizovaná technika analýzy dat za účelem získání informací, jako jsou vzory, tendence a souvztáhnosti v těchto datech. Široká definice licence vytvořila právní nejistotu, zda je možné tuto licenci aplikovat i v případech strojového učení, resp. tréninku AI modelu.¹¹⁰ S nedávno přijatým Aktem o AI¹¹¹, který na tuto licenci přímo odkazuje v kontextu tréninku AI modelů, tato pochybnost mizí.¹¹² Licence však s oprávněním spojuje i dvě

¹⁰⁷ § 30 odst. 6 ve spoj. § 43 odst. 1 a 3 autorského zákona.

¹⁰⁸ TELEČ, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 382.

¹⁰⁹ Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu a o změně směrnic 96/9/ES a 2001/29/ES (dále jen „směrnice o autorském právu na jednotném digitálním trhu“)

¹¹⁰ Intellectual Property Crime Project at the European Union Agency. *Generative Artificial Intelligence: the impact on intellectual property crimes* [online], 2023, s. 16.

¹¹¹ Nařízení Evropského parlamentu a Rady (EU) 2024/1689 ze dne 13. června 2024, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění nařízení (ES) č. 300/2008, (EU) č. 167/2013, (EU) č. 168/2013, (EU) 2018/858, (EU) 2018/1139 a (EU) 2019/2144 a směrnice 2014/90/EU, (EU) 2016/797 a (EU) 2020/1828 (akt o umělé inteligenci) (dále jen „Akt o AI“)

¹¹² Srov. HOLCOVÁ, I. *MODELÝ UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO*. DOBŘIČOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLŮVÝCH [online], 2023, s. 22-23 a PRCHAL, P. *Generativní umělá inteligence*

limitace. První limitací je, že vytěžená data (zhotovené rozmnoženiny) je možné uchovávat jen po dobu nezbytnou pro účely tréninku AI modelu. Druhá limitace spočívá v možnosti autora díla předem zakázat jeho užití.

- b) *Licence rozmnožování díla pro účely automatizované analýzy textů nebo dat k vědeckému výzkumu.* Autorský zákon dále obsahuje i „rozšířenou“ licenci, která ani neumožňuje autorům vyhradit si užití jejich díla. Tato licence je však personálně a účelově omezená a slouží např. pro vysoké školy a specifické instituce, které rozmnožují díla pro účely automatizované analýzy k vědeckému výzkumu. Pro práci není zcela relevantní, neboť se v dalších částech zabývá především komerčními AI systémy.
- c) *Licence pro karikaturu, parodii nebo pastiš.* V reakci na další implementaci směrnice došlo k rozšíření o zákonnou licenci k užití cizího díla pro účely *pastiše*. Podle Slovníku spisovného jazyka českého se pod tímto pojmem rozumí „*literární, hudební nebo výtvarné dílo napodobující techniku, styl jiného díla*“¹¹³. Zákonnou definici však nenalezneme ani ve směrnici, ani v zákoně a na závazný výklad si tak bude nutné počkat až na rozhodnutí Soudního dvora EU.¹¹⁴

a porušování autorských práv. Praha: Bulletin advokacie, 3/2024, 2024, s. 9 nebo KELLER, P. A *first look at the copyright relevant parts in the final AI Act compromise* [online]. Amsterdam: Institute for Information Law (IViR), 2023

¹¹³ HAVRÁNEK B., et. al. *Slovník spisovného jazyka českého* [online]. Ústav pro jazyk český, v. v. i., 2011.

¹¹⁴ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání.* Praha: C.H. Beck, 2023, s. 128.

3.1.5 Počítačový program

Zvláštní režim v autorském právu mají počítačové programy. Vytváření počítačového programu programátorskou činností je obvykle spíše technickou záležitostí a svou povahou stojí na nižší tvůrčí úrovni než například umělecká činnost.¹¹⁵

Vývoj některých počítačových programů však vyžaduje vynaložení značných lidských, technických i finančních zdrojů.¹¹⁶ Za účelem ochrany této investice autorský zákon v souladu s evropským právem rozšiřuje předmět ochrany i o počítačové programy, které nesplňují všechny pojmové znaky autorského díla. Místo znaku *jedinečnosti* je u počítačového programu vyžadován znak *původnosti*, který má nižší nároky na objem vynaložené tvůrčí činnosti. Ochrany tak požívají počítačové programy, které vznikly na základě vlastní nezávislé duševní činnosti programátora. Oproti uměleckým či vědeckým dílům (kde je požadován znak jedinečnosti) tak není vyloučené, že budou existovat dva identické počítačové programy a oba budou nezávisle na sobě chráněny autorským právem. Nadále se však vyžaduje, aby počítačový program byl autorovým duševním výtvoem.¹¹⁷ Z autorskoprávní ochrany tak budou vyloučené například počítačové programy, jejichž zdrojový kód mechanicky vygenerovala generativní umělá inteligence a uživatel AI systému tento kód pouze zkopíroval.

I u počítačového programu však platí, že musí být vyjádřen v *objektivně vnímatelné podobě* (nejčastěji v podobě zdrojového kódu), aby mu příslušela autorskoprávní ochrana. Pouhý námět, postup, myšlenka nebo metoda bez konkrétního vyjádření není autorským právem chráněna.^{118, 119} Samotná funkcionality počítačového programu, myšlenky a zásady jako takové (logika, algoritmy, funkce a programovací jazyk), tak nepožívá autorskoprávní ochrany. Zkoumáním chování cizího počítačového programu je tedy možné legálně naprogramovat obdobný nebo

¹¹⁵ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání.* Praha: C.H. Beck, 2019, s. 37.

¹¹⁶ Bod 2 odůvodnění Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. 4. 2009 o právní ochraně počítačových programů (kodifikované znění) (dále jako „směrnice o právní ochraně počítačových programů“).

¹¹⁷ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání.* Praha: C.H. Beck, 2019, s. 36.

¹¹⁸ § 2 odst. 6 ve spoj. s § 65 odst. 2 autorského zákona.

¹¹⁹ Samotné postupy, metody sice nejsou chráněny autorským právem, ale není vyloučené je chránit jinou formou, například jako obchodní tajemství nebo ve formě patentu.

dokonce totožný počítačový program.¹²⁰ To však neplatí, pokud by si třetí osoba za tímto účelem opatřila zdrojový nebo strojový kód zkoumaného počítačového programu, a to ani formou reverzního inženýrství.¹²¹

3.1.6 Databáze

V právním smyslu je *databáze* chápána jinak, než je tomu v čistě technické terminologii. Databází nemusí být jen například relační databáze, ale i webová stránka, katalog¹²² či jiný soubor prvků, který bude systematicky uspořádán a přístupný elektronickými (nebo jinými) prostředky.¹²³ Autorský zákon rozlišuje mezi ochranou struktury databáze a ochranou samotného obsahu databáze.

Struktura databáze může být pod autorskoprávní ochranou, pokud vznikla na základě vlastní duševní činnosti člověka (autora). Takto chráněnou databází může být například i online encyklopedie, jejíž obsah bude obvykle tvořen údaji, které samy o sobě nejsou předmětem právní ochrany¹²⁴ (např. základní informace o hlavních městech), pokud způsob uspořádání nebo výběr jejího obsahu je výsledkem tvůrčí činnosti autora (znak *jedinečnosti*) nebo alespoň projevem jeho tvůrčí svobody (znak *původnosti*)¹²⁵.

Pořízení některých databází může vyžadovat značnou lidskou, technickou i finanční investici.¹²⁶ Za účelem ochrany těchto vynaložených prostředků autorský zákon přiznává ochranu i osobě, která databázi pořídila. Na rozdíl od ochrany struktury databáze může tato ochrana náležet nejen fyzické osobě, ale i osobě právnické.¹²⁷ Pořizovatel databáze má právo rozhodnout o tom, kdo a jak bude užívat obsah pořizené databáze.¹²⁸ Přičemž není podstatné, zda je obsah tvořen autorskými díly

¹²⁰ SAS Institute Inc. v. World Programming Ltd. Rozsudek Soudního dvora (velkého senátu) ze dne 2. května 2012 ve věci C-406/10.

¹²¹ Srov. § 66 odst. 4 autorského zákona.

¹²² HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 79.

¹²³ Srov. § 88 autorského zákona.

¹²⁴ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 56.

¹²⁵ K pojmu *tvůrčí svoboda* ve vztahu k databázi viz Football Dataco Ltd a další v. Yahoo UK Limited a další. Rozsudek Soudního dvora (třetího senátu) ze dne 1. března 2012 ve věci C-604/10.

¹²⁶ Bod 7 odůvodnění Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. března 1996 o právní ochraně databází (dále jen „směrnice o databázích“).

¹²⁷ § 89 autorského zákona.

¹²⁸ § 90 autorského zákona.

(např. básněmi, fotografiemi) nebo prostými údaji (např. statistickými údaji). Toto právo ale osobě přísluší jen, pokud podstatně přispěla (nejen nutně ve smyslu finančního vkladu, ale například znalostmi) ke kvalitě nebo kvantitě databáze¹²⁹, a to formou pořízení, ověření nebo předvedení.¹³⁰

Pořizovatel tak například může udělit oprávnění poskytovateli AI, aby databázi užil za účelem vytěžení dat k tréninku AI modelu. Poskytovatel AI by ovšem mohl využít pouze obsah databáze jako takový. Pokud by chtěl využít k tréninku AI modelu i strukturu databáze (například k vyhodnocování souvztažností dat), musel by si opatřit i souhlas autora této struktury (nejedná-li se o tutéž osobu).¹³¹

V poměrně složité situaci se může poskytovatel AI nacházet ve chvíli, kdy pořizovatel databáze zařadil do jejího obsahu cizí autorská díla. Pouhým smluvním oprávněním (licencí) od pořizovatele databáze totiž není vyloučeno, že poskytovatel AI vytěžením databáze bude porušovat autorská práva třetích osob, pakliže autoři jednotlivých děl nedali pořizovateli souhlas k zařazení jejich děl do databáze.¹³²

Pokud se pořizovatel rozhodne zpřístupnit databázi veřejnosti, dojde i k omezení jeho ochrany. Zpřístupněnou databázi může kdokoliv bez jeho souhlasu užít, za podmínky, že obsah užívá běžně a přiměřeně a jen kvalitativně nebo kvantitativně nepodstatnou část obsahu databáze.¹³³

3.1.7 Kolektivní správa

Kolektivní správa je právní institut, který je hojně využíván v různých mediálních odvětvích, například při vysílání nebo streamování hudby. V těchto odvětvích působí kolektivní správce, který vlastním jménem vyjednává s uživateli (například s provozovatelem rádia) o odměně na účet velkého množství autorů. Tento systém snižuje transakční náklady, zjednodušuje přístup k široké škále děl uživatelům a zajišťuje, že autoři dostanou za užití svých děl odměnu. Kolektivní správa tak řeší

¹²⁹ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání*. Praha: C.H. Beck, 2023, s. 252.

¹³⁰ § 88a autorského zákona.

¹³¹ SEDLAKOVA LEGAL s.r.o. *Umělá inteligence a databáze: Učení umělé inteligence analýzou dat z databáze* [online]. 2022.

¹³² Srov. § 105 autorského zákona.

¹³³ § 91 autorského zákona.

například situaci, kdy uživatel potřebuje přístup k většímu množství autorskoprávně chráněných děl (např. hudební skladby), ale administrativní zátěž spojená s individuálním jednáním s každým jednotlivým autorem by byla pro uživatele neúnosná.¹³⁴

Na základě uděleného oprávnění kolektivní správce plně spravuje majetková práva ku společnému prospěchu velkého množství autorů (resp. nositelů majetkových práv)¹³⁵. Kolektivní správce jedná s uživateli, uděluje jim licenci¹³⁶ a vybírá za autory odměnu. Vybranou odměnu musí kolektivní správce autorům vyplácet, a to podle předem stanovených parametrů a po odečtení nezbytných nákladů.¹³⁷

3.2 Exkurz k harmonizaci práva umělé inteligence v EU

Evropská unie¹³⁸ již několik let aktivně usiluje o vytvoření komplexního právního rámce, který by upravoval vývoj, implementaci a využívání umělé inteligence v členských státech. Cílem této snahy je vyvážit podporu technologiím na bázi umělé inteligence a zajistit ochranu základních práv. Přístup EU k regulaci AI nespočívá pouze ve stanovení pravidel, ale také v tom, že se snaží etablovat jako globální normotvůrce v této oblasti. Tím, že jako první zavede komplexní právní rámec AI, chce EU formulovat standardy, které by mohly být eventuálně přijaty na mezinárodní úrovni.¹³⁹

Akt o AI a vysoké pokuty spojené s jeho nedodržením zároveň značí záměr EU přísně dohlížet na technologie umělé inteligence. Pokuty, které mohou dosáhnout až 35 000 000 EUR nebo 7 % celosvětového ročního obrátu¹⁴⁰, překonávají pokuty stanovené v obecně známém nařízení GDPR¹⁴¹, které jsou omezeny na 20 milionů

¹³⁴ Intergram. *Co je to kolektivní správa práv? A k čemu slouží?* [online]. Dostupné z <https://www.intergram.cz/pro-verejnost/co-je-to-kolektivni-sprava-prav/>.

¹³⁵ § 95 a 95a autorského zákona.

¹³⁶ § 98a autorského zákona.

¹³⁷ § 99 a násl. autorského zákona.

¹³⁸ Název „Evropská unie“ a zkratka „EU“ jsou v práci používány zaměnitelně.

¹³⁹ EVROPSKÁ KOMISE. *Special meeting of the European Council (1 and 2 October 2020) – Conclusions*. EUCO 13/20 [online], 2020. [CO EUR 10 CONCL 6], s. 6.

¹⁴⁰ Článek 99 odst. 3 Aktu o AI.

¹⁴¹ Nařízení Evropského parlamentu a Rady (EU) č. 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů) (dále jen „nařízení GDPR“)

EUR nebo 4 % ročního celosvětového obratu.¹⁴² Stanovením norem pro umělou inteligenci chce EU vytvořit prostředí, v němž lze umělou inteligenci vyvíjet a používat způsobem, který je v souladu s evropskými hodnotami a právními zásadami.

Navzdory dlouhodobé přípravě a analýzám přišla konkrétní regulace generativní umělé inteligence až jako ex post reakce na již uvedené komerční produkty, jako je například ChatGPT. Tím se ukazuje, jak obtížné je pro regulační orgány držet krok s rychlým a nepředvídatelným technologickým vývojem AI. Čtenáři, kteří jsou dobře obeznámeni s vývojem unijní regulace v oblasti AI, mohou následující podkapitoly přeskocit.

3.2.1 Akt o AI

Akt EU o umělé inteligenci je průlomový právní předpis, jehož cílem je harmonizovat pravidla pro vývoj, zavádění a používání systémů umělé inteligence v Evropské unii. Hlavním účelem nařízení je podpořit zavádění důvěryhodné, na člověka zaměřené umělé inteligence a zároveň zajistit vysokou úroveň ochrany před potenciálními škodami a riziky, které technologie umělé inteligence představují.

Přístup nařízení k regulaci AI je založen na identifikaci a kategorizaci rizik systémů umělé inteligence. Rizikem se zde míní nežádoucí chování AI, které zasahuje do zdraví člověka, jeho bezpečnosti a dalších základních práv a svobod stanovených v Listině základních práv Evropské unie (např. svoboda myšlení, ochrana osobních údajů nebo zákaz diskriminace).¹⁴³ Nařízení rozlišuje mezi čtyřmi hlavními kategoriemi (i) zakázaných¹⁴⁴, (ii) vysoce rizikových¹⁴⁵, (iii) rizikových¹⁴⁶ a (iv) nerizikových¹⁴⁷ AI systémů. V závislosti na kategorii nařízení buď zcela zakazuje uvedení technologie na trh, zavádí povinná ochranná opatření, nebo doporučuje zavedení dobrovolného kodexu chování.

¹⁴² Článek 83 odst. 6 nařízení GDPR.

¹⁴³ Čl. 1 odst. 1 Aktu o AI.

¹⁴⁴ Čl. 5 a násl. Aktu o AI.

¹⁴⁵ Čl. 6 a násl. Aktu o AI.

¹⁴⁶ Čl. 50 a násl. Aktu o AI.

¹⁴⁷ Čl. 95 Aktu o AI ve spoj. bodu 165 odůvodnění Aktu o AI.

Přijatá a platná verze obsahuje významné odlišnosti od původního návrhu z roku 2021, zejména v reakci na rychlý pokrok v oblasti technologií umělé inteligence, jako je uvedení ChatGPT a dalších generativních AI systémů na trh. Jednou z klíčových změn v přijaté verzi je přidání ustanovení týkajících se „*obecných modelů umělé inteligence*“, což je reakce na schopnosti, které demonstroval ChatGPT a další generativní AI systémy, které se objevily po první verzi návrhu nařízení.

Pro účely této práce autor přistoupil k selekci ustanovení a krátkého komentáře, které jsou pro generativní umělou inteligenci relevantní:

Článek 3 odst. 1 – Definice systému AI

„Systém AI“, strojový systém navržený tak, aby po zavedení fungoval s různými úrovněmi autonomie a který po zavedení může vykazovat adaptabilitu a který za explicitními nebo implicitními účely z obdržených vstupů odvozuje, jak generovat výstupy, jako jsou predikce, obsah, doporučení nebo rozhodnutí, které mohou ovlivnit fyzická nebo virtuální prostředí.“

Tato definice je záměrně extenzivní, aby zahrnovala širokou škálu současných i budoucích technologií umělé inteligence, včetně generativních AI systémů umělé inteligence. Pojem „strojový systém“ odkazuje na skutečnost, že „*systémy AI*“ fungují prostřednictvím počítačů nebo jiných strojů (robot).¹⁴⁸

Generativní umělá inteligence vykazuje různou míru autonomie při generování výstupů na základě dat, na kterých je trénovaná. Je adaptivní, neustále se učí a zlepšuje své výstupy. Její výstupy, ať už jde o text, obrázky, kód a jiné, mohou významně ovlivňovat virtuální i fyzické prostředí.

Článek 3 odst. 60 – Definice deepfake

„Deep fake“ obrazový, zvukový nebo video obsah vytvořený nebo manipulovaný umělou inteligencí, který se podobá existujícím osobám, objektům, místům subjektům či událostem a který by se dané osobě mohl nepravdivě jevit jako autentický nebo pravdivý.

¹⁴⁸ Bod 12 odůvodnění Aktu o AI.

Výslovné zahrnutí definice „*deep fake*“ do nařízení naznačuje, že zákonodárce si je vědom možného zneužití technologií generativní umělé inteligence. Generativní umělá inteligence je schopna produkovat vysoce realistický obrazový, zvukový nebo video obsah, který lze jen obtížně odlišit od skutečného obsahu.¹⁴⁹

Článek 3 odst. 63 – Definice obecného modelu AI

„Obecným modelem AI“ model AI, včetně případů, kdy je tento model AI trénován velkým množstvím dat s využitím vlastního dohledu ve velkém měřítku, který vykazuje významnou obecnost a je schopen kompetentně plnit širokou škálu různých úkolů bez ohledu na způsob, jakým je daný model uveden na trh, a který lze začlenit do různých navazujících systémů nebo aplikací, s výjimkou modelů AI, které se používají pro činnosti výzkumu, vývoje nebo činnosti zaměřené na tvorbu prototypů před jejich uvedením na trh.“

Nařízení definičně odlišuje „*systémy AI*“ od „*obecných modelů AI*“. Tento pojem zákonodárce přidal v reakci na popularizaci nástrojů na bázi generativní umělé inteligence¹⁵⁰, jako je ChatGPT. Z jazykového výkladu je zřejmé, že se jedná o podmnožinu „*systému AI*“. AI modely jsou sice zásadními součástmi AI systémů, samy o sobě však AI systémy nepředstavují. Systémy AI jsou vedle modelu AI tvořeny i dalšími komponenty, jako je například uživatelské rozhraní.¹⁵¹

O „*obecný model AI*“ se jedná, pokud byl trénován na velkém množství dat a zároveň vykazuje vysokou úroveň obecnosti, která mu umožňuje plnit širokou škálu úloh. Například je schopen variabilně generovat obsah ve formě textu, audia i videa.

Kvůli riziku zneužití těchto modelů nařízení stanovuje pro všechny jejich provozovatele povinnosti, které se týkají zajištění transparentnosti.

Nařízení dále obsahuje klasifikační pravidla, podle kterých lze takový model AI označit jako model se „*systémovým rizikem*“. Toto posouzení zohledňuje faktory, jako jsou velikost trénovací sady, počet parametrů v neuronové síti nebo počet registrovaných uživatelů. Je-li AI model dle těchto klasifikačních pravidel označen

¹⁴⁹ K tomu viz kapitola 2.2. Strojové učení.

¹⁵⁰ Bod 99 odůvodnění Aktu o AI.

¹⁵¹ Bod 97 odůvodnění Aktu o AI.

jako „systémové riziko“ bude se na provozovatele vztahovat vedle obecných pravidel i další speciální povinnosti v podobě zajištění důkladnější dokumentace a odpovídající kybernetické bezpečnosti.¹⁵²

Článek 50 odst. 1 – Povinnost transparentnosti pro AI interagující s lidmi

„Poskytovatelé zajistí, aby systémy AI určené k přímé interakci s fyzickými osobami byly navrhovány a vyvíjeny tak, že dotčené fyzické osoby budou vyrozuměny o tom, že komunikují se systémem AI, není-li tato skutečnost zřejmá z pohledu fyzické osoby, která je přiměřeně informovaná, pozorná a obezřetná, při zohlednění okolností a kontextu použití. (...).“

Nařízení stanovuje povinnost transparentnosti pro poskytovatele systémů AI, které přímo komunikují s člověkem. Účelem tohoto článku je zajistit, aby si dotčená osoba byla vědoma, kdy komunikuje se systémem AI. Tato informovanost může ovlivnit její očekávání, chování a rozhodovací procesy. Jednotlivci se tak mohou informovaně rozhodnout, zda a jak budou s umělou inteligencí komunikovat. Nařízení však připouští, že v určitých situacích může být přiměřeně informované osobě na základě kontextu a okolností zřejmé, že komunikuje se systémem umělé inteligence. V takových případech se od požadavku transparentnosti výslovně upouští.

Povinnosti týkající se transparentnosti se tak pravděpodobně budou vztahovat i na společnosti, které používají generativní modely umělé inteligence, jako jsou velké jazykové modely, k vytváření konverzační chatové podpory, které komunikují přímo se zákazníky. Pokud například společnost integruje do své platformy zákaznické podpory generativní model umělé inteligence, jako je GPT, musí zákazníky výslovně informovat, že komunikují se systémem umělé inteligence, pokud to není zřejmé z kontextu a okolností. Jedním ze způsobů, jak tuto povinnost splnit, by mohlo být uvedení jasného upozornění na začátku interakce v chatu, například ve formě: *„Právě chatujete s AI asistentem. Jedná se o systém umělé inteligence, nikoliv o lidského operátora.“*

¹⁵² Čl. 55 Aktu o AI.

Článek 50 odst. 2 – Povinnost transparentnosti pro poskytovatele generativní AI

„Poskytovatelé systémů AI, včetně obecných systémů AI, vytvářejících syntetický zvukový, obrazový, video nebo textový obsah, zajistí, aby výstupy systému AI byly označeny ve strojově čitelném formátu a zjistitelné jako uměle vytvořené nebo manipulované. Poskytovatelé zajistí, aby jejich technická řešení byla účinná, interoperabilní, robustní a spolehlivá, pokud je to technicky proveditelné, s přihlédnutím ke zvláštnostem a omezením různých druhů obsahu, nákladům na provádění a obecně uznávanému nejmodernějšímu stavu technologií, tak jak mohou být zachyceny v příslušných technických normách. Tato povinnost se neuplatní v rozsahu, v jakém systémy AI plní asistenční funkci pro standardní editaci nebo podstatně nemění vstupní údaje poskytnuté zavádějícím subjektem nebo jejich sémantiku (...).“

Nařízení obsahuje ustanovení, které se speciálně vztahuje na poskytovatele systémů generativní umělé inteligence schopné vytvářet syntetický zvuk, video, text nebo obraz. Podle tohoto článku musí poskytovatelé zajistit, aby výstupy této AI byly označeny ve strojově čitelném formátu tak, aby je bylo možné identifikovat jako uměle vytvořené. Tento postup se také označuje jako digitální AI vodoznakování.

Pomocí AI vodoznaku je teoreticky možné zabránit šíření dezinformací generovaných umělou inteligencí tím, že by usnadnil odhalení syntetického obsahu prezentovaného jako autentický. Vodoznak by rovněž mohl sloužit jako prevence plagiátorství.

Současné techniky AI vodoznaků se však potýkají s problémy, pokud jde o jejich spolehlivost a účinnost. Vodoznaky se obvykle vytvářejí během procesu trénování modelu umělé inteligence tak, že se model naučí vkládat do obsahu, který generuje, určitý znak nebo identifikátor. V případě modelů generujících text se může jednat o použití určitých neobvyklých slov nebo slovních spojení. Bohužel je dnes poměrně snadné takový vodoznak obejít, například pouhým přeformulováním slov či použitím synonym.¹⁵³ Na začátku roku 2023 společnost OpenAI představila online nástroj, který měl být schopen detekovat text generovaný umělou inteligencí.

¹⁵³ Srov. Osobní blog výzkumníka Scotta Aaronsona ze společnosti OpenAI: AARONSON S. *What Is AI Safety?* [online], 2022.

O pouhých šest měsíců později však OpenAI tento nástroj pro jeho velmi nízkou spolehlivost z trhu stáhla.¹⁵⁴

Vzhledem k současným omezením technologie AI vodoznaků mohou mít poskytovatelé generativní umělé inteligence potíže dostat požadavku transparentnosti tak, jak je stanoveno v tomto článku nařízení. Technologie AI vodoznaků bude pravděpodobně vyžadovat další výzkum a vývoj s cílem vytvořit spolehlivější techniky, které bude obtížnější obejít. V počátečních fázích implementace Aktu o AI může být vhodný uvolněnější přístup k dodržování této povinnosti, aby se zabránilo nadměrnému zatížení poskytovatelů v době, kdy technologie generativní AI teprve zdokonaluje. Regulační orgány by měly spolupracovat s průmyslovým sektorem a akademickou obcí na podpoře výzkumu pokročilejších metod vodoznaků a časem stanovit osvědčené postupy a standardy.

Článek 50 odst. 4 pododst. 1 – Povinnost transparentnosti pro deepfake

„Subjekty zavádějící systém AI, který vytváří obrazový, zvukový nebo video obsah představující tzv. ‚deep fake‘ nebo s ním manipuluje, musí zveřejnit, že obsah byl uměle vytvořen nebo s ním bylo manipulováno. (...) Pokud je obsah součástí zjevně uměleckého, tvůrčího, satirického, fiktivního či obdobného díla nebo programu, povinnosti týkající se transparentnosti stanovené v tomto odstavci jsou omezeny na zveřejnění existence takového vytvořeného nebo zmanipulovaného obsahu vhodným způsobem, který nebrání zobrazení nebo užívání díla.“

Namísto úplného zákazu obsahu s „deep fake“ volí zákonodárce volnější přístup. Nařízení ukládá požadavek transparentnosti všem subjektům zavádějící systém AI, který vytváří obrazový, zvukový nebo video obsah podobající se existujícím osobám anebo místům, který by se mohl nepravdivě jevit jako autentický.¹⁵⁵

Zásadním aspektem tohoto ustanovení je, že přenáší odpovědnost na provozovatele AI systémů, tedy právnické i fyzické osoby, které používají takový systém při své profesionální činnosti.¹⁵⁶ Nařízení však zohledňuje nutnost respektovat ochranu svobody projevu a umělecké tvorby, jež je zakotvena v Listině základních práv EU

¹⁵⁴ OPENAI. *New AI classifier for indicating AI-written text* [online], 2023.

¹⁵⁵ Čl. 3 odst. 60 Aktu o AI.

¹⁵⁶ *A contrario* čl. 3 odst. 4 Aktu o AI.

(články 11 a 13). Proto stanoví výjimku pro obsah, který je zjevně umělecký, tvůrčí, satirický nebo fiktivní. V takových případech je povinnost transparentnosti omezena na decentní upozornění, přičemž je brán zřetel na povahu vytvořeného díla, aby upozornění nebránilo jeho zobrazení nebo užívání.

Článek 50 odst. 4, pododst. 2 – Povinnost transparentnosti textu generovaného AI
„Subjekty zavádějící systém AI, který vytváří text, jenž je zveřejněn za účelem informování veřejnosti o záležitostech veřejného zájmu, nebo který s takovým textem manipuluje, uvedou, že text byl vytvořen uměle nebo s ním bylo manipulováno. Tato povinnost se nevztahuje na případy, (...) kdy byl obsah vytvořený umělou inteligencí podroben procesu přezkumu člověkem nebo redakční kontroly a pokud za zveřejnění obsahu nese redakční odpovědnost fyzická nebo právnická osoba.“

Zatímco původní návrh Výboru Evropského parlamentu pro kulturu a vzdělávání navrhoval zahrnout do definice „*deep fake*“ i uměle vytvořený nebo upravený text¹⁵⁷, nynější verze nařízení zúžila definici jen na zvukový, vizuální a audiovizuální obsah. Namísto toho Evropský parlament přistoupil k řešení textu generovaného AI, běžně také označované jako „*fake news*“, odlišným způsobem, a to zavedením speciální povinnosti transparentnosti. Tato povinnost se bude vztahovat např. na novináře, kteří musí splnit informační povinnost tím, že uvedou, zda byl text vygenerován umělou inteligencí, nebo zajistit, aby obsah prošel procesem redakční kontroly, přičemž redakční odpovědnost za zveřejnění pak nese konkrétní fyzická nebo právnická osoba.

Zavedení této povinnosti se zdá být cílenou snahou bojovat proti šíření dezinformačních zpravodajských webů vytvořených pomocí umělé inteligence, které mají potenciál klamat čtenáře. Požadavkem na uvedení, že text byl vytvořen AI (nebo alternativně povinnost lidského redakčního dohledu) má nařízení za cíl

¹⁵⁷ VÝBOR EVROPSKÉHO PARLAMENTU PRO KULTURU A VZDĚLÁVÁNÍ. *Opinion of the Committee on Culture and Education for the Committee on the Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts* ze dne 16. června 2022. [2021/0106(COD)], s. 24.

podpořit transparentnost a odpovědnost při šíření informací o věcech veřejného zájmu.

Článek 53 odst. 1 písm. c) – Povinnost dodržování autorského práva

„Poskytovatelé obecných modelů AI zavedou politiku pro dodržování práva Unie v oblasti autorského práva a práv souvisejících, a zejména pro určení a dodržování výhrady práv vyjádřené podle čl. 4 odst. 3 směrnice (EU) 2019/790, a to i prostřednictvím nejmodernějších technologií.“

Na první pohled se může zdát ustanovení o povinnosti dodržování autorského práva EU jako nadbytečné. Je totiž zřejmé, že subjekty působící v EU musí dodržovat autorské právo EU bez ohledu na to, zda jim to nařizuje Akt o AI. Pokud však ustanovení budeme vykládat v kontextu odůvodnění, pochopíme, že záměr zákonodárce byl mnohem hlubší a zásadnější.¹⁵⁸

Odůvodnění nařízení jasně uvádí, že při tréninku AI modelu za použití cizího obsahu je třeba souhlasu jeho autora, kromě případů, pro které platí zvláštní výjimky.¹⁵⁹ Takovou výjimku stanovuje směrnice o autorském právu na jednotném digitálním trhu, která umožňuje rozmnožování a extrakci autorských děl pro účely *vytěžování textů a dat*, k nimž má subjekt zákonný přístup (např. zpřístupněný veřejnosti online).¹⁶⁰ Přímým odkazem na tuto směrnici tak zákonodárce jasně připouští, že výjimka pro *vytěžování textů a dat* se vztahuje i na trénink modelů generativní umělé inteligence.

Tato výjimka není bezmezná. Nositelé práv mohou užití svého díla vyloučit a zabránit tak provozovatelům, aby jejich díla využívala k tréninku AI modelů. Povinnost dodržovat toto ustanovení se týká všech poskytovatelů, kteří hodlají uvést na trh EU obecný AI model. Výhradu nositelů práv musí provozovatelé respektovat bez ohledu na to, v jaké zemi probíhá trénink modelu AI.

Tímto ustanovením chce EU zamezit například tomu, aby poskytovatel proces trénování uskutečnil v jurisdikci mimo EU, kde je úroveň autorskoprávní ochrany

¹⁵⁸ Srov. KELLER, P. *A first look at the copyright relevant parts in the final AI Act compromise* [online]. Amsterdam: Institute for Information Law (IViR), 2023.

¹⁵⁹ Bod 105 odůvodnění Aktu o AI.

¹⁶⁰ Čl. 4 odst. 1 směrnice o autorském právu na jednotném digitálním trhu.

obecně nižší, a poté již takto natrénovaný AI model uvedl na trh EU a získal tím konkurenční výhodu.¹⁶¹ Z extenzivního výkladu tak de facto vyplývá, že se zavádí celosvětové uplatňování autorského práva EU. Toto ustanovení je projevem snahy EU stanovit jako první světové standardy pro další vývoj AI.

Kromě povinnosti dodržování autorského práva obsahuje článek 53 Aktu o AI i další povinnosti pro poskytovatele obecných modelů umělé inteligence. Jedná se především o povinnost vypracovat a průběžně udržovat technickou dokumentaci¹⁶², která musí mimo jiné zahrnovat informace o použitých datech pro trénování modelu, včetně jejich původu. Dokumentace musí dále obsahovat podrobný popis technik a metodik trénování, architekturu samotného modelu a dokonce i odhadovanou spotřebu energie daného modelu.

Přesný rozsah povinností je specifikován v přílohách nařízení, které jsou jeho nedílnou součástí.¹⁶³ Poskytovatelé AI musí tuto dokumentaci pravidelně aktualizovat a na požádání ji zpřístupnit příslušnému úřadu pro AI. Evropská komise je rovněž oprávněna přijímat akty v přenesené pravomoci za účelem aktualizace požadavků na technickou dokumentaci s ohledem na technický pokrok¹⁶⁴. Tímto způsobem nařízení zajišťuje, že budou pravidla a požadavky v souladu s neustále se vyvíjejícím oborem umělé inteligence.

3.2.2 Nařízení o obecné bezpečnosti výrobků (GPSR)

S účinností od 13. prosince 2024 nahradí stávající směrnici o obecné bezpečnosti výrobků¹⁶⁵ a směrnici o výrobcích imitující potraviny¹⁶⁶ nové nařízení GPSR¹⁶⁷. Cílem nařízení je aktualizovat právní rámec pro zajištění bezpečnosti výrobků tak,

¹⁶¹ Bod 106 odůvodnění Aktu o AI.

¹⁶² Čl. 53 odst. 1 písm. a) a písm. b) Aktu o AI.

¹⁶³ Příloha XI a XII Aktu o AI.

¹⁶⁴ Čl. 53 odst. 6 Aktu o AI.

¹⁶⁵ Směrnice Evropského parlamentu a Rady 2001/95/ES ze dne 3. prosince 2001 o obecné bezpečnosti výrobků.

¹⁶⁶ Směrnice Rady ze dne 25. června 1987 o sblížení právních předpisů členských států týkajících se výrobků, jejichž skutečná povaha není rozpoznatelná a které proto ohrožují zdraví nebo bezpečnost spotřebitelů.

¹⁶⁷ Nařízení Evropského parlamentu a Rady (EU) 2023/988 ze dne 10. května 2023, o obecné bezpečnosti výrobků, o změně nařízení Evropského parlamentu a Rady (EU) č. 1025/2012 a směrnice Evropského parlamentu a Rady (EU) 2020/1828 a o zrušení směrnice Evropského parlamentu a Rady 2001/95/ES a směrnice Rady 87/357/EHS (dále jen „nařízení GPSR“).

aby zohledňoval rizika, které sebou přináší nové technologie, např. umělá inteligence.

Nařízení GPSR doplňuje Akt AI¹⁶⁸, což je samostatný, komplexnější předpis zaměřený na regulaci vysoce rizikových systémů umělé inteligence. Obě nařízení společně tvoří součást koncepce EU pro regulaci AI. Zatímco Akt o AI se zabývá konkrétně vysoce rizikovými systémy umělé inteligence, nařízení GPSR slouží jako „pojistka“ vztahující se na spotřebitelské výrobky poháněné umělou inteligencí, které nemusí spadat do působnosti Aktu o AI. Rozšiřuje tak úroveň ochrany spotřebitele na širší škálu výrobků využívajících umělou inteligenci.¹⁶⁹

Podle nařízení GPSR se při posuzování toho, zda je výrobek bezpečný, vezme v úvahu i to, zda je výrobek kyberneticky zabezpečen před nekalými úmysly třetích stran.¹⁷⁰

3.2.3 Směrnice o odpovědnosti za umělou inteligenci

Třetím předpisem, který doplňuje balíček opatření na podporu rozvoje důvěryhodné AI, je návrh směrnice o odpovědnosti za umělou inteligenci¹⁷¹. Zatímco účelem Aktu o AI je určit pravidla o používání systémů umělé inteligence, úkolem směrnice je stanovit pravidla pro vymáhání škod způsobených umělou inteligencí.¹⁷²

Důvodem harmonizace je především fakt, že vnitrostátní předpisy upravující mimosmluvní občanskoprávní odpovědnost nemusí být vhodné pro řešení náhrady újmy způsobené umělou inteligencí. Specifické vlastnosti systémů AI, včetně její složitosti (tzv. efekt „černé skříňky“), mohou poškozeným ztížit nebo znemožnit unést důkazní břemeno při uplatňování nároku na náhradu škody.¹⁷³ V této souvislosti směrnice stanoví, že soudy jsou oprávněny nařídit zpřístupnění důkazů

¹⁶⁸ Viz kapitola výše.

¹⁶⁹ LÖLFING, N. *Part 1: Product compliance for consumer products powered by AI* [online]. Bird & Bird, 2024.

¹⁷⁰ Čl. 6 odst. 1 písm. g) nařízení GPSR.

¹⁷¹ Návrh směrnice Evropského parlamentu a Rady (EU) o přizpůsobení pravidel mimosmluvní občanskoprávní odpovědnosti umělé inteligenci (směrnice o odpovědnosti za umělou inteligenci) [COM(2022) 496 final] ze dne 28. září 2022.

¹⁷² MATZNER, J. *Směrnice o odpovědnosti za umělou inteligenci: Krok k odpovědnému využívání AI* [online]. epravo.cz, 2024.

¹⁷³ Bod 3 odůvodnění Směrnice o odpovědnosti za umělou inteligenci.

u vysoce rizikového systému AI, u nějž existuje důvodné podezření, že poškozenému způsobil škodu.¹⁷⁴

Pro zlepšení procesního postavení poškozeného dále směrnice stanoví vyvratitelnou domněnku příčinné souvislosti. V praxi může být pro poškozeného nadměrně obtížné prokázat příčinnou souvislost mezi zaviněním žalovaného, spočívajícím v nedodržení povinnosti řádné péče, a výstupem AI, který vedl ke škodě.¹⁷⁵ Při splnění určitých podmínek¹⁷⁶ tak směrnice vyvratitelnou domněnkou přenáší důkazní břemeno na žalovaného.

3.2.4 Evropský úřad pro umělou inteligenci¹⁷⁷

Návrhy Evropské komise, které začaly nabývat konkrétní podoby s první iniciativou v roce 2018, postupně procházejí legislativním řízením EU. Zatímco některé z výše zmíněných návrhů již vstoupily v platnost, u jiných se předpokládá, že tak tomu bude nejpozději v roce 2024. Jedním z dalších významných milníků na cestě k vytvoření právního rámce pro regulaci AI v EU bylo zřízení Evropského úřadu pro umělou inteligenci.

Vznik Evropského úřadu pro umělou inteligenci je reakcí na potřebu jednotného systému správy umělé inteligence v EU. Tato instituce má nejen dohlížet, ale také aktivně řídit pokrok v oblasti umělé inteligence směrem, který bude v souladu s hodnotami a předpisy EU. Úřad slouží jako ústřední orgán, jehož hlavním úkolem je implementovat a prosazovat navrhovaný Akt o AI¹⁷⁸, jehož smyslem je zajistit, aby systémy umělé inteligence byly bezpečné, transparentní a v souladu s předpisy.

Úkolem úřadu je rovněž sledovat vývoj trhu a technologií v oblasti umělé inteligence, což je úloha, která má zásadní význam pro pochopení rychle se měnícího prostředí technologií umělé inteligence a jejích dopadů na společnost

¹⁷⁴ Čl. 3 odst. 1 Směrnice o odpovědnosti za umělou inteligenci.

¹⁷⁵ Bod 28 odůvodnění Směrnice o odpovědnosti za umělou inteligenci.

¹⁷⁶ Čl. 4 odst. 1 Směrnice o odpovědnosti za umělou inteligenci.

¹⁷⁷ Rozhodnutí Komise ze dne 24. ledna 2024 o zřízení Evropského úřadu pro umělou inteligenci [C/2024/1459].

¹⁷⁸ Tamtéž, čl. 2 odst. 1.

a ekonomiku. Kromě toho se od úřadu očekává, že bude přispívat k vytvoření dalších opatření, která budou podporovat etický rozvoj a využívání AI.¹⁷⁹

Od prvotních myšlenek zveřejněných Evropskou komisí v tiskové zprávě¹⁸⁰ uplynulo šest let, během nichž proběhla řada významných milníků a politických dohod. Nyní EU vstupuje do nové éry, kdy byl zřízen Evropský úřad pro umělou inteligenci, který bude v souladu s první komplexní regulací EU dohlížet na celý trh.

Pro generativní AI je klíčový zejména Akt o AI, který byl na poslední chvíli doplněn o ustanovení, která reagují na rizika spojená s tímto typem umělé inteligence. Mezi tato opatření patří například zavedení povinné transparentnosti pro systémy generující *deepfakes* nebo *fake news*. Zákodníci tak reagují na obavy z možného zneužití těchto technologií k šíření dezinformací a manipulaci veřejného mínění.

Z hlediska autorských práv se zdá, že se EU spoléhá na stávající právní úpravu. Pro vyloučení pochybností je však v Aktu o AI upřesněno, že dříve přijatá zákonná licence pro *vytěžování textů a dat* se vztahuje i na případy využití cizích děl pro trénování modelů AI. Tímto krokem poskytuje poskytovatelům AI určitou míru právní jistoty, alespoň co se týče fáze tréninku modelů AI.

¹⁷⁹ Tamtéž, čl. 2 odst. 2.

¹⁸⁰ EVROPSKÁ KOMISE. *Tisková zpráva - Artificial intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards ze dne 9. března 2018* [online], 2018. [IP/18/1381].

ZVLÁŠTNÍ ČÁST

4 Výsledky tvůrčí činnosti (díla)

Generativní umělá inteligence je dnes schopna produkovat obsah, který by mnoho lidí bez znalosti o jeho původu označilo za autorské dílo. Její schopnost vyprodukovat takto autentický obsah je však závislá na výsledcích tvůrčí činnosti člověka. Současné metody umělého generování textů nebo obrázků mohou fungovat jen proto, že při tréninku modelu generativní AI byly využity existující texty nebo obrázky, jejichž původcem byl člověk.¹⁸¹ Tyto modely se učí umělecké styly, techniky a vzorce z nepřeberného množství děl umělců, spisovatelů a dalších tvůrců, díky čemuž mohou vytvářet obsah, který napodobuje lidskou tvořivost a její výsledky, tedy díla ve smyslu Autorského díla. Je tak pochopitelné, že mnoho tvůrců (autorů) se obává možného zneužití své tvůrčí práce společnostmi, které vytvářejí a poskytují systémy generativní AI. Rozšířené využívání těchto nástrojů může snížit poptávku po jejich práci, neboť obsah generovaný umělou inteligencí bude stále kvalitnější a dostupnější. V důsledku čehož z jejich původní práce budou ekonomicky profitovat především nadnárodní technologické společnosti, zatímco tvůrci, jejichž výsledky tvůrčí činnosti byly použity k trénování modelů, obdrží za svojí práci malou nebo vůbec žádnou odměnu.¹⁸²

Účelem autorského práva (a práva duševního vlastnictví obecně) je poskytnout ochranu výsledkům lidské tvořivosti. Duševní vlastnictví bylo důležitým a nepostradatelným prvkem pokroku a vývoje celého lidstva, od prvních prehistorických nástrojů, přes kolo, čínské počítadlo, penicilin, počítač a další nesčetné inovace, které lidstvu umožnily dosáhnout současné úrovně technologického pokroku. Ochrana duševního vlastnictví slouží jako stimul pro tvůrce a zajišťuje, aby se lidstvo nadále vyvíjelo a kultivovalo.¹⁸³

Kromě zájmů na právní ochraně tvůrců je však třeba vzít v úvahu i další důležité zájmy. Autorské právo obsahuje široký systém výjimek a omezení, které buď zcela

¹⁸¹ Viz kapitola 2.1. Strojové učení.

¹⁸² KELLER, P. *Protecting creatives or impeding progress? Machine learning and the EU copyright framework* [online]. Amsterdam: Institute for Information Law (IViR), 2023

¹⁸³ WIPO. *Světová deklarace o duševním vlastnictví*. 2000, oddíl 3.

vyklučují výsledky tvůrčí činnosti z ochrany (např. úřední díla), nebo umožňují jiným osobám užívat tyto výsledky i bez souhlasu autora, a to ve prospěch veřejného zájmu, jako je podpora kulturního, technického, vědeckého či hospodářského rozvoje. Stanovení hranic mezi ochranou těchto různých zájmů tak není snadné.¹⁸⁴ Možná právě tato složitá síť výjimek a omezení, nebo snad přirozená evoluce vnímání nehmotných statků jako „něco méně“ než statků hmotných, vede k tomu, že duševní vlastnictví často není vnímáno jako právo přirozené.

Představme si situaci, kdy se setkáte se svým kamarádem v knihovně a začnete si povídat o nové knize, která nedávno vyšla a kterou oba chcete přečíst. Kniha je však poměrně drahá a váš kolega vám sdělí, že má přístup k webové stránce, kde lidé nelegálně sdílejí elektronické kopie knih. Požádá vás, abyste mu pomohli najít a stáhnout tuto knihu z této webové stránky, protože ví, že jste technicky zdatnější a dokážete obejít ochranu proti nelegálnímu sdílení.

Dále si představte, že vám tentýž kolega řekne o malém knihkupectví, které se nachází na jeho cestě domů a o kterém ví, že v určité hodiny není dobře hlídáno, protože majitel chodí na oběd. Požádá vás, abyste se šli podívat, zda je knihkupectví nehlídáno, aby mohl knihu ukrást.

V obou případech dochází k porušení majetkových práv. V prvním případě jde o porušení autorských práv a práv s nimi souvisejících, když se nelegálně sdílí a stahuje elektronická kopie knihy. Ve druhém případě jde o porušení práva vlastnického, konkrétně o krádež fyzické kopie knihy. Přestože by mnozí bez váhání s nelegálním stahováním knihy pomohli, v případě fyzické krádeže by pravděpodobně zaváhali a pomoci odmítli. Zajímavé je, jak odlišně jsou tyto dvě situace vnímány veřejností, přestože v obou případech dochází k protiprávnímu jednání a k porušení majetkových práv.¹⁸⁵

¹⁸⁴ KOUKAL, P., et al. *Úvod do práv k duševnímu vlastnictví* [online]. Brno: Právnická fakulta Masarykovy univerzity, 2020, s. 1–4.

¹⁸⁵ Tamtéž.

4.1 Právo na náhradní odměnu

Důvodem, proč generativní umělá inteligence, jako je ChatGPT, funguje, je nejen její nová technologie, ale zejména i enormní množství dat, která byla použita k jejímu natrénování. Významnou roli ve vývoji velkých jazykových modelů sehrála nová architektura¹⁸⁶ neuronové sítě, která byla představena v roce 2017. Skutečná síla těchto modelů však spočívá v objemu dat, na kterých jsou trénovány. Na nové architektuře byla totiž empiricky zjištěna korelace mezi výkonností modelů a velikostí trénovacího datasetu.¹⁸⁷ Dnešní vývoj generativní umělé inteligence tedy hlavně spočívá v tom, kdo získá větší a rozmanitější data pro trénování modelů a dokáže je natrénovat na ještě výkonnějších superpočítačích. Jedná se o relativně novou praxi. Před zveřejněním nové architektury neuronové sítě se totiž většina modelů trénovala na významně menším objemu dat, který byl obvykle pečlivě vybírán.¹⁸⁸

Současné datasey jsou tak velké, že ani samotní poskytovatelé AI systému neví, co konkrétně je jejich obsahem.¹⁸⁹ Tyto soubory dat obvykle pocházejí z veřejně přístupných zdrojů (internetu) a mohou obsahovat miliony děl chráněných autorským právem.¹⁹⁰ Avšak jednotlivá díla sama o sobě mají jen velmi malý vliv na chování (a výstupy) generativní umělé inteligence. Jak by v takovém případě měla být chráněna a kompenzována práva individuálních autorů?

Již dnes na základě zákonné licence stanovené v § 39c autorského zákona¹⁹¹ mohou vývojáři generativní umělé inteligence užít veřejně dostupná díla za účelem trénování modelů umělé inteligence bez nutnosti souhlasu jejich autorů.¹⁹² Tato zákonná licence vychází z předpokladu, že zhotovené rozmnoženiny děl jsou uchovávány pouze po dobu nezbytnou pro účely tréninku umělé inteligence. Tento

¹⁸⁶ VASWANI, A., et al. *Attention is all you need*. *Advances in Neural Information Processing Systems* [online]. Cornell University Library, 2017.

¹⁸⁷ KAPLAN, J., et al. *Scaling Laws for Neural Language Models*. OpenAI [online], 2020.

¹⁸⁸ LEE, K., et al. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 31.

¹⁸⁹ Tamtéž, s. 30.

¹⁹⁰ OPENAI. *OpenAI—written evidence (LLM0113) House of Lords Communications and Digital Select Committee inquiry: Large language models* [online], 2023, s. 4.

¹⁹¹ Srov. čl. 4 odst. 1 směrnice o autorském právu na jednotném digitálním trhu.

¹⁹² Práce vychází z premisy správnosti právní úpravy (definice) této výjimky; cílem práce není právní analýza definičního vymezení vytěžování textů a dat, tedy ani toho, zda unijní zákonodárce měl na mysli datovou analýzu či nikoli.

pohled však přehlíží dlouhodobý dopad na autory takto využívaných děl. S rostoucí schopností generativní umělé inteligence produkovat kvalitnější literární a jiný umělecký obsah bude pravděpodobně klesat i poptávka po lidské tvůrčí tvorbě. Vystává tak otázka, jakým způsobem nastavit ekonomickou incentivu, aby až (pokud) taková situace nastane, měli tvůrci motivaci pokračovat ve své činnosti.¹⁹³

Zákonná licence k rozmnožování díla pro účely tréninku modelu umělé inteligence sice obsahuje možnost autora předem vyloučit užití svého díla, a tedy zabránit tomu, aby bylo jeho dílo zařazeno do tréninkového datasetu pro umělou inteligenci.¹⁹⁴

Praktické prosazování tohoto *opt-out* práva však představuje značný problém. Za prvé výhradu užití může autor uplatnit pouze *pro futuro*.¹⁹⁵ Pro existující (natrénované) modely tak tato výhrada neplatí. Za druhé bude velmi náročné unést důkazní břemeno a prokázat, že konkrétní dílo nebylo vyloučeno z tréninkového datasetu, neboť k němu má obvykle přístup jen poskytovatel. Z výstupů umělé inteligence, které se přímo nepodobají trénovacím datům, je identifikace ještě náročnější, ne-li nemožná.

Ve světle nového nařízení Aktu o AI se objevuje snaha přenést toto důkazní břemeno na poskytovatele AI systémů. Podle článku 53 odst. 1 písm. c) Aktu o AI musí poskytovatelé obecných modelů AI zavést „*politiku pro dodržování právních předpisů Unie v oblasti autorského práva, a zejména pro určení a dodržování výhrad práv vyjádřené podle čl. 4 odst. 3 směrnice (EU) 2019/790*“. Obsahem takového dokumentu by mělo být i shrnutí obsahu použitého pro trénování modelu.¹⁹⁶ Z nařízení však není přímo jasné, zda taková dokumentace má být tak podrobná, aby samotní autoři byli schopni určit, zda bylo respektováno jejich právo na vyloučení jejich díla z tréninku modelu. Akt o AI naznačuje, že tomu tak zřejmě není. Dohled nad dodržováním požadavku transparentnosti bude v působnosti

¹⁹³ SENFTLEBEN, M., *Generative AI and Author Remuneration*. IIC – International Review of Intellectual Property and Competition Law, Volume 54 [online], 2023, s. 1536.

¹⁹⁴ § 39c odst. 2 autorského zákona.

¹⁹⁵ Bod 19 důvodové zprávy zákona č. 429/2022 Sb., kterým se mění zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, a další související zákony.

¹⁹⁶ Srov. čl. 53 odst. 1 Aktu o AI ve spoj. s přílohou XI a XII Aktu o AI.

úřadu pro AI, který však nebude za tímto účelem ověřovat jednotlivá data použitá k tréninku.¹⁹⁷

Vzhledem k výše uvedenému je tak třeba zvážit zavedení finanční kompenzace pro autory, která je bude motivovat k tvůrčí činnosti i v dlouhodobém horizontu. Jedním z takových řešení by mohlo být uvalení platební povinnosti na poskytovatele AI systémů, které jsou schopny produkovat literární a jiné umělecké výstupy.¹⁹⁸ Obdobně byl takto v minulosti vyřešen rozvoj reprodukční techniky (např. tiskáren, DVD, paměťových karet), který umožnil masově rozmnožovat díla pro osobní potřebu (v souladu s autorským právem). Důsledkem toho přirozeně došlo ke snížení poptávky po autorem vyrobených rozmnoženin, a tedy i ke snížení jeho celkové odměny.¹⁹⁹ Tento konflikt zájmů byl vyřešen zavedením zvláštního jiného majetkového práva na náhradní odměnu pro autory za potenciálně hospodářský snížený zisk způsobený reprodukční technikou. Přičemž tato platební povinnost není uložena konzumentovi (osobě, která si legálně pořizuje rozmnoženinu pro osobní potřebu), ale tomu, kdo svoji podnikatelskou činností umožňuje toto masové rozmnožování.²⁰⁰

Po vzoru práva na náhradní odměnu v souvislosti s rozmnožováním díla pro osobní potřebu tak připadá v úvahu zavést i právo na náhradní odměnu v souvislosti s rozmnožováním díla pro účely automatizované analýzy textů nebo dat. Přičemž osobami povinnými platit tuto odměnu by byli dovozci, distributoři anebo poskytovatelé obecných modelů AI²⁰¹, které jsou schopny produkovat či distribuovat obsah ve formě textu, zvuku, obrázků nebo videa.²⁰²

4.2 Povinná kolektivní správa práva na náhradní odměnu

Poskytovatelé generativní AI potřebují k natrénování svých modelů získat velké množství dat. Tyto datasety často mohou obsahovat, jak již bylo uvedeno výše,

¹⁹⁷ Bod 108 odůvodnění Aktu o AI.

¹⁹⁸ SENFTLEBEN, M., *Generative AI and Author Remuneration*. IIC – International Review of Intellectual Property and Competition Law, Volume 54 [online], 2023, s. 1549.

¹⁹⁹ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 306–307.

²⁰⁰ Tamtéž.

²⁰¹ Čl. 3 odst. 3, 6 a 7 Aktu o AI.

²⁰² Bod 99 odůvodnění Aktu o AI.

autorskoprávně chráněná díla, od textu a obrázků až po hudbu a video. Těžko si však lze představit, že by poskytovatel AI jednal o licenci k jejich užití jednotlivě s každým autorem.²⁰³ Jak bylo uvedeno v předchozí kapitole, i kdyby byl zaveden systém náhradní odměny pro autory za rozmnožování jejich díla pro účely tréninku umělé inteligence, administrativní zátěž spojená s takovými individuálními jednáními by byla pro poskytovatele AI neúnosná. Tato neefektivita by mohla vést k tomu, že by se poskytovatelé jednoduše spoléhali na to, že autoři mají omezené možnosti zjistit, že jejich dílo bylo zařazeno do tréninkového datasetu, nebo případně riskovali, že se občas budou muset (mimo)soudně vypořádat s žalobou.²⁰⁴ Řešením této komunikační neefektivity by mohla být kolektivní správa takového jiného majetkového práva autorů.

Kolektivní správa není novým konceptem, a naopak je hojně využívána v různých mediálních odvětvích, například při vysílání nebo streamování hudby. V těchto odvětvích působí kolektivní správce, který vlastním jménem vyjednává s uživateli (například s provozovatelem rádia) o odměně na účet velkého množství autorů. Tento systém snižuje transakční náklady, zjednodušuje přístup k široké škále děl a zajišťuje, že autoři dostanou za užití svých děl odměnu.

S ohledem na povahu tréninkových datasetů, kdy poskytovatelé nejsou schopni přesně zjistit ani ovlivnit, jaká díla ve skutečnosti užijí, by bylo odůvodněné uvažovat nejen o kolektivní správě (dobrovolné), ale i o kolektivní správě povinné. Povinná kolektivní správa spočívá v zákonném zmocnění, kdy určitá majetková práva spravuje výhradně kolektivní správce a jednotliví nositelé práv mají v tomto rozsahu omezenou autonomii. Tento systém je odůvodněný zejména ve scénářích, kdy uživatelé nemohou reálně identifikovat všechny nositele práv nebo s nimi jednat²⁰⁵, jako je tomu v případě poskytovatelů generativní AI, kteří automatizovaným způsobem (*web scraping*) získávají tréninková data z nejrůznějších zdrojů (např. z internetu). Ti si tak v případě užití obsahu, který je pod ochranou autorského práva, budou moci být jisti, že vůči nim nebudou vznikat nároky

²⁰³ INTERGRAM. *Co je to kolektivní správa práv? A k čemu slouží?* [online].

²⁰⁴ THE AUTHORS GUILD. *Legislative Proposals to Protect the Creative Professions and Mitigate Risk of Harm from Generative AI* [online], 2023, s. 3.

²⁰⁵ Tamtéž.

jednotlivých autorů, ale pouze nárok kolektivního správce.²⁰⁶ Tento systém má významný přínos i pro nositele práv, neboť ti nebudou muset sledovat, kde všude se jejich dílo užilo k tréninku umělé inteligence, a stačí jim domáhat se svého práva (na odměnu) u kolektivního správce.

Zavedení povinné kolektivní správy představuje vyvážený přístup k řešení konfliktu zájmů mezi jednotlivými tvůrci a poskytovateli generativní AI. Využitím osvědčené metody z jiných mediálních odvětvích je možné podpořit technologický pokrok a zároveň zajistit spravedlivou odměnu pro tvůrce.

²⁰⁶ Tamtéž.

5 Data

Fungování generativní umělé inteligence do značné míry závisí²⁰⁷ na kvantitě dat, která byla použita při tréninku modelu umělé inteligence. Například tréninkový dataset pro model GPT-3 se skládal z ekvivalentu dokumentu Microsoft Word, který by měl více než 3,7 miliardy stran²⁰⁸. Tento požadavek na enormní množství dat je obvykle naplňován automatizovaným sběrem různých článků, knih nebo obrázků, a to především z internetu²⁰⁹. Tento plošný a automatizovaný proces sběru veřejně dostupných dat však může vést i k shromažďování obsahu, který je chráněn autorskými právy.

Ať už se u chráněného obsahu jedná o shromažďování úmyslné nebo neúmyslné, žádný pokročilý systém generativní umělé inteligence, jako je například ChatGPT od OpenAI, nelze vyvinout bez výsledků tvůrčí činnosti člověka, z nichž část bude chráněna autorskými právy. Tyto modely umělé inteligence vyžadují seznámení se s širokým spektrem různorodého obsahu, aby skutečně dokázaly obsáhnout lidský intelekt a jazykový projev člověka.²¹⁰

Neselektivní povaha automatizovaného procesu sběru dat však může vést k závažným právním problémům, zejména pokud je obsah chráněný (autorskými) právy užíván bez souhlasu jejich nositelů či bez zákonné licence.

5.1 Vytěžování textů a dat pro účely automatizované analýzy textů a dat

Právní rámec sběru a analýzy autorskoprávně chráněného obsahu řeší na unijní úrovni směrnice o autorském právu na jednotném digitálním trhu.

Článek 4 směrnice, nadepsaný „Výjimky a omezení pro vytěžování textů a dat“, v odstavcích 1 a 3 stanoví:

²⁰⁷ Viz kapitola 2. 1. 2 Metody strojového generování obrázků.

²⁰⁸ Bod 91 THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, OPENAI, INC. a další. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195

²⁰⁹ OPENAI. *OpenAI—written evidence (LLM0113) House of Lords Communications and Digital Select Committee inquiry: Large language models* [online], 2023, s. 4.

²¹⁰ Tamtéž.

(1) Členské státy stanoví výjimku (...) pro rozmnoženiny a extrakce zhotovené za účelem vytěžování textů a dat z děl nebo jiných předmětů ochrany, k nimž mají zákonný přístup.

(3) Výjimka nebo omezení podle odstavce 1 se použijí pod podmínkou, že si nositelé práv výslovně nevyhradili užití děl a jiných předmětů ochrany podle uvedeného odstavce vhodným způsobem, jako jsou strojově čitelné prostředky v případě obsahu zpřístupněného veřejnosti online.

Do Autorského zákona byl tento článek směrnice implementován do § 39c, nadepsaného jako „Licence k rozmnožování díla pro účely automatizované analýzy textů nebo dat“, kdy v odstavcích 1 a 2 stanoví:

(1) Do práva autorského nezasahuje ten, kdo zhotoví rozmnoženinu díla za účelem automatizované analýzy textů nebo dat v digitální podobě, prováděné za účelem získání informací, zahrnujících mimo jiné vzory, tendence a souvztažnosti; takto zhotovenou rozmnoženinu je oprávněn uchovat pouze po dobu nezbytnou pro účely této automatizované analýzy textů nebo dat.

(2) Ustanovení odstavce 1 se nepoužije pro rozmnoženiny díla, jehož autor si užití podle odstavce 1 výslovně vyhradil vhodným způsobem; v případě díla zpřístupněného podle § 18 odst. 2 strojově čitelnými prostředky.

Zákonná licence pro vytěžování dat umožňuje poskytovatelům AI shromažďovat díla pro účely strojového učení. Tato výjimka platí, pokud k dílu mají zákonný přístup a nositelé práv (autoři) si nevyhradili užití jejich děl a za dodržení podmínek stanovených třístupňovým testem²¹¹.

Zákonným přístupem k dílům se rozumí ta díla, ke kterým mají poskytovatelé AI přístup na základě smluvního ujednání (licence) nebo ze zákona (nejčastěji se bude jednat o volně dostupný obsah na internetu).

Autoři také mohou vytěžování děl předem výslovně vyloučit. V případě děl, která byla zpřístupněna veřejnosti online musí být výhrada projevena strojově čitelnými prostředky. Výhrada by tak měla být učiněna prostřednictvím metadat webové

²¹¹ § 29 odst. 1 autorského zákona.

stránky²¹². Bohužel v současné době není zcela zřejmé, v jakém formátu by výhrada měla být. Vytěžování dat probíhá pomocí automatizovaného softwaru (*web crawler*), který musí být naprogramován tak, aby věděl, kde a v jaké podobě se výhrada bude vyskytovat, a ověřit tak, zda autorskoprávně chráněný obsah na webových stránkách je či není vyloučen z vytěžení.

Jako řešení se nabízí protokol *robots.txt*, kterým provozovatelé webových stránek informují automatizované softwary (roboty) o tom, které části webových stránek mohou či nemohou navštívit. Jedná se o standardizovaný protokol, který funguje již přes 20 let a je založen na dobrovolném dodržování. Ačkoliv samotný protokol nemá žádný vynucovací mechanismus v zákoně, je respektován i nadnárodními technologickými společnostmi jako jsou Google²¹³ a OpenAI²¹⁴.²¹⁵

Výhrada k vyloučení užití obsahu umístěného na podstránce „/hudba“ a „/basne“ v textovém souboru *robots.txt* by vypadala například takto:

```
User-Agent: *  
Disallow: /hudba  
Disallow: /basne
```

Ke sběru dat může samozřejmě docházet také u obsahu, které není chráněn autorským právem, a v takových případech poskytovatelé AI žádný souhlas (resp. zákonnou licenci) podle autorského práva nepotřebují²¹⁶. To však neznamená, že takový obsah je možné vytěžovat bez dalšího. To, že obsah není chráněn autorským právem, nepředstavuje překážku v tom, aby provozovatel webových stránek stanovil smluvní omezení jeho používání, i když nejde o absolutní ochranu, jako je tomu v případě autorskoprávně chráněného obsahu. Všeobecné podmínky mohou vytěžování zakázat jednoduchým ustanovením: „*Používání*

²¹² Srov. bod 18 odůvodnění směrnice o autorském právu na jednotném digitálním trhu a bod 19 důvodové zprávy k zákonu č. 429/2022 Sb., kterým se mění zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, a další související zákony.

²¹³ Srov. *Overview of Google crawlers and fetchers (user agents)*. Dostupné z <https://developers.google.com/search/docs/crawling-indexing/overview-googlecrawlers>.

²¹⁴ Srov. *GPTBot*. OpenAI API. Dostupné z <https://platform.openai.com/docs/gptbot>.

²¹⁵ KELLER, P., WARSO, Z. *Defining best practices for opting out of ML training*. Open Future policy brief [online], 2023, s. 7-8.

²¹⁶ Bod 9 odůvodnění směrnice o autorském právu na jednotném digitálním trhu.

automatizovaných systémů nebo softwaru za účelem získávání údajů z těchto webových stránek pro obchodní účely (web scraping) je zakázáno.“ Taková výhrada ani nemusí být uvedena strojově čitelným způsobem ve smyslu Autorského zákona. Stačí, když návštěvník webu prokazatelně projeví souhlas s všeobecnými podmínkami (např. zaškrtnutím příslušného políčka).²¹⁷ V praxi tak paradoxně může být pro poskytovatele AI obtížnější legálně vytěžovat obsah, který nebude pod ochranou autorského práva.

5.2 The New York Times proti OpenAI²¹⁸

Otázkou užívání autorských děl k strojovému učení umělé inteligence se od prosince roku 2023 zabývá Okresní soud Spojených států pro obvod východního New Yorku. V dané věci žalobce The New York Times tvrdí, že žalovaná společnost OpenAI neoprávněně užila „*miliony*“ novinových článků k vytvoření generativního systému AI (ChatGPT).²¹⁹ Toto neoprávněné užití umožnilo společnosti OpenAI vytvořit nástroj, který dokáže generovat obsah, který konkuruje a napodobuje obsah novinových článků žalobce. Jako důkaz žalobce uvedl příklady výstupů ze systému ChatGPT, který údajně doslovně kopíruje pasáže ze zveřejněných novinových článků.

Ve Spojených státech platí doktrína *fair use*, která za určitých podmínek umožňuje třetím osobám užít cizí díla bez souhlasu jejich autorů. Kontinentální právní úprava autorského práva tuto doktrínu nezná, ale obdobnou funkci zastávají zákonné licence (např. k rozmnožování díla pro účely automatizované analýzy textů nebo dat) ve spojení s třístupňovým testem²²⁰. V této kapitole tedy nenajdete řešení podstaty sporu, tedy zda užití novinových článků ke strojovému učení je možné kvalifikovat jako užití v souladu s americkou doktrínou *fair use*. Žaloba však obsahuje skutková tvrzení, která konkrétně popisují, jakým způsobem v různých fázích životního cyklu generativní umělé inteligence dochází k zásahu do

²¹⁷ Srov. Ryanair Ltd v. PR Aviation BV. Rozsudek Soudního dvora (druhého senátu) ze dne 15. ledna 2015 ve věci C-30/14.

²¹⁸ THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, OPENAI, INC. A další. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195.

²¹⁹ Tamtéž, bod 2.

²²⁰ § 29 odst. 1 autorského zákona.

autorského práva. Nebude však zkoumáno, zda takový zásah je či není v mezích doktríny *fair use*.

Dle tvrzení žalobce došlo k neoprávněnému rozmnožování děl již při tréninku modelu generativní umělé inteligence. K tréninku modelu GPT-3 byly ve velké míře použity existující datasey, a to například datasey Wikipedia (obsahující anglické články z encyklopedie Wikipedia) nebo Common Crawl (obsahující obsah z více než 50 miliard webových stránek)²²¹. Přičemž veřejně dostupný dataset Common Crawl, který tvořil 60% podíl na tréninkovém datasetu modelu GPT-3, obsahuje více než 66 milionů záznamů (ne nutně individuálních článků) z webových stránek provozovaných žalobcem²²². Podle žalobce tak došlo k opakovanému rozmnožování děl bez náležité licence a kompenzace.²²³

Zdálo by se, že 66 milionů záznamů je mnoho, ale v rámci celého datasetu Common Crawl se jedná o méně než 0,02 %. Zajímavější otázka vyvstává, když uvážíme, že společnost OpenAI užívá k tréninku AI modelu externí datasey (databáze). V kontextu českého autorského práva pouhým smluvním oprávněním (licencí) od pořizovatele databáze totiž není vyloučeno, že poskytovatel AI bude vytěžováním databáze porušovat autorská práva třetích osob, pakliže autoři jednotlivých děl nedali pořizovateli souhlas k zařazení jejich děl do databáze.²²⁴

V předchozí kapitole byl zmíněn protokol robots.txt jako řešení k projevení výhrady (strojově čitelným způsobem) k užití obsahu na webové stránce za účelem strojového učení. V době psaní této práce je možné takovou výhradu společnosti The New York Times nalézt umístěnou na www.nytimes.com/robots.txt. V textovém souboru je nejdříve ve formě poznámek uvedeno, že používání automatizovaných nástrojů za účelem získávání údajů z těchto webových stránek je zakázáno, a to mj. platí i pro vytěžování ve smyslu článku 4 směrnice o autorském právu a na jednotném digitálním trhu.

²²¹ BROWN, T.B., et al. *Language Models are Few-Shot Learners* [online]. OpenAI, Cornell University Library, 2020. Dostupné z <https://arxiv.org/pdf/2005.14165>.

²²² Bod 86 THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, OPENAI, INC. a další. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195

²²³ Tamtéž, bod 92.

²²⁴ Srov. § 105 autorského zákona.

```
# Use of any device, tool, or process designed to data
mine or scrape the content
# using automated means is prohibited without prior
written permission from
# The New York Times Company. Prohibited uses include
but are not limited to:
# (1) text and data mining activities under Art. 4 of
the EU Directive on Copyright in
# the Digital Single Market;
# (2) the development of any software, machine learning,
artificial intelligence (AI),
# and/or large language models (LLMs);
# (3) creating or providing archived or cached data sets
containing our content to others; and/or
# (4) any commercial purposes.
```

Dále textový soubor obsahuje konkrétní výčet automatizovaných nástrojů (robotů), kterým je přístup na webové stránky zcela zakázán. Mezi nimi je nyní zařazen i robot GPTBot od společnosti OpenAI.²²⁵

```
User-agent: GPTBot
Disallow: /
```

K dalšímu zásahu do autorského práva dle tvrzení žalobce dochází ve fázi generování. Dle předložených důkazů je možné na základě uživatelského příkazu (angl. *prompt*) vygenerovat prostřednictvím systému ChatGPT obsah, který doslovně kopíruje novinové články zveřejněné žalobcem.

Podle veřejného vyjádření²²⁶ společnosti OpenAI je takové plagiátorské chování vzácným projevem chyby, ke které dochází během procesu tréninku modelu.²²⁷ Je však nutno podotknout, že takové chování je častější, když se určitý obsah objeví

²²⁵ PETERS, J., WES, D., The New York Times blocks OpenAI's web crawler [online], The Verge, 2023. Dostupné z <https://www.theverge.com/2023/8/21/23840705/new-york-times-openai-web-crawler-ai-gpt>.

²²⁶ OPENAI. *OpenAI and journalism*. 2024, [online]. Dostupné z <https://openai.com/index/openai-and-journalism/>.

²²⁷ O tom více viz kapitola 7.1. Plagiátorské chování generativní umělé inteligence.

v trénovacím datasetu několikrát. Například žalobcem předložené důkazy obsahují pouze články, které jsou několik let staré a byly virálně sdíleny i jinými deníky či webovými stránkami. Patří mezi ně i článek, který v roce 2019 obdržel prestižní novinářskou Pulitzerovu cenu. Společnost OpenAI dále zpochybňuje, že by systém ChatGPT byl schopen vygenerovat takto dlouhé výňatky z článků bez předchozích uživatelských příkazů.

6 Dataset

Poté, co jsou data z nejrůznějších zdrojů (například z internetu) shromážděna, dochází k jejich seskupení do datových sad, v technické terminologii nazývaných též jako datasety. Tento proces seskupování dat není čistě mechanickou a rutinní činností, ale obvykle zahrnuje i proces metodického výběru, uspořádání a anotací.²²⁸ O náročnosti vytváření kvalitního datasetu svědčí fakt, že poskytovatelé AI si své datasety střeží jako velmi cenné aktivum, a to i v případech, kdy jsou ochotni zveřejnit celou technickou architekturu a algoritmy.²²⁹

Vlivem nedostatečné transparentnosti obsahu datasetů mezi poskytovateli AI dochází i obecně k nižšímu porozumění chování systémů AI, což vede i k etickým a právním problémům. Mezi ně patří například nezamýšlené předpojaté chování²³⁰ AI systémů nebo nahodilý únik osobních údajů²³¹ při generování obsahu. Navíc s ohledem na spektrum subjektů, které jsou v životním cyklu generativní umělé inteligence zapojeny, mohou být některé právní aspekty velmi matoucí a nejasné. Uživatel může například nevědomky používat AI systém, jehož poskytovatel trénoval model s využitím databází, ke kterým neměl příslušnou licenci. A vzhledem k tomu, že trénink modelů generativní umělé inteligence je velmi nákladný a do značné míry nevratný, není snadné tato rizika a problémy zpětně odstranit.²³²

6.1 Dataset jako předmět autorského práva

Samotný dataset může být chráněn jako *databáze*²³³. V právním smyslu je pojem *databáze* vykládán velmi extenzivně a rozumí se jím jakýkoli soubor nezávislých prvků (například autorská díla, ale i statistické údaje), které jsou systematicky nebo

²²⁸ LEE, K., et al. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 37.

²²⁹ Srov. OPENAI. *GPT-4 Technical Report* [online]. 2024.

²³⁰ BUBECK, S., et al. *Sparks of Artificial General Intelligence: Early experiments with GPT-4* [online]. Microsoft Research, 2023, s. 86-89.

²³¹ CARLINI, N., et al. *Extracting training data from large language models*. USENIX Security Symposium, volume 6 [online], 2022, s. 2367.

²³² LONGPRE, S., et al. *The Data Provenance Initiative: A Large Scale Audit of Dataset Licensing & Attribution in AI* [online]. Cornell University Library, 2023, s. 1-2.

²³³ Viz kapitola 3.1.6. Databáze.

metodicky uspořádané a individuálně přístupné (například elektronickými prostředky)²³⁴. Autorský zákon poskytuje jednak ochranu struktury, jednak obsahu. Struktura databáze může být chráněna coby souborné dílo, pokud způsob výběru nebo uspořádání obsahu představuje vlastní duševní výtvor autora²³⁵. Takto chráněným datasetem by mohla být antologie českých básní. Nesmělo by se jednat jen o nahodilou sbírku básní, ale o takovou, která je uspořádána způsobem, který odráží autorovu tvůrčí svobodu²³⁶. Básně by byly vybrány a uspořádány tak, aby například vyprávěly tematický příběh, přičemž každá báseň by navazovala na předchozí a společně by tvořily jednu dějovou linii. Básně by také mohly být zařazeny do autorem vytvořených kategorií, jako jsou básně o přírodě, o lásce či o životě. Pro poskytovatele AI to znamená, že bez souhlasu autora (resp. příslušné zákonné licence) nelze za účelem strojového učení vyhodnocovat vztahy mezi jednotlivými básněmi. Je ale možné využít jednotlivé²³⁷ básně, které jsou v datasetu obsaženy (za předpokladu, že samy o sobě nejsou chráněny autorským právem).²³⁸

I když struktura databáze nebude představovat vlastní duševní výtvor autora, je stále možné ji chránit před kvantitativně nebo kvalitativně rozsáhlým vytěžováním, a to prostřednictvím práva *sui generis*. Toto zvláštní právo k databázi přísluší pořizovateli databáze, který k jejímu pořízení, ověření nebo předvedení provedl kvalitativně nebo kvantitativně podstatný vklad²³⁹ (např. finanční, časový, lidský). Účelem tohoto zvláštního práva je tedy ochránit hospodářský vklad vynaloženého na vytvoření anebo provoz takové databáze²⁴⁰. Takto chráněným datasetem by mohl být archiv českých básní, který by představoval obsáhlou sbírku básní z různých období. Přičemž pořizovatel takové databáze vynaložil několik měsíců intenzivní

²³⁴ § 88 autorského zákona.

²³⁵ § 2 odst. 2 a § 2 odst. 5 autorského zákona.

²³⁶ Srov. Eva-Maria Painer v. Standard VerlagsGmbH a další. Rozsudek Soudního dvora (třetího senátu) ze dne 1. prosince 2011 ve věci C-145/10.

²³⁷ Nesmělo by jít o vytěžování podstatné části databáze nebo o opakované a systematické užití nepodstatných částí v rozporu s § 90 a násl. Autorského zákona. O tom více viz odstavec níže.

²³⁸ Srov. SEDLAKOVA LEGAL s.r.o. *Umělá inteligence a databáze: Učení umělé inteligence analýzou dat z databáze* [online]. 2022.

²³⁹ § 88a odst. 1 autorského zákona.

²⁴⁰ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 844.

práce k shromáždění a ověření jednotlivých básní.²⁴¹ Takto shromážděné básně by však neměly nějakou zvláštní strukturu, ale byly by řazeny jen chronologicky či abecedně. Pro poskytovatele AI to znamená, že pokud k strojovému učení chce využít celý dataset nebo jeho kvalitativně nebo kvantitativně podstatnou část (nebo i nepodstatné části, které jsou však užívány opakovaně a systematicky), je možné tak učinit jen na základě souhlasu pořizovatele databáze (resp. příslušné zákonné licence).

6.2 Souběh ochrany

Užívání datasetů ke strojovému učení neodmyslitelně zahrnuje i vytváření (dočasných) kopií²⁴² v datasetu obsažených dat. Tento krok je nezbytný, neboť data musí být v paměti počítače reprodukována, aby mohlo dojít k analýze a trénování modelu. Takové kopírování dat však představuje zásah do práva na rozmnožování.

Pokud poskytovatel AI ke strojovému učení navíc užívá datasety třetích stran, může docházet i k zásahu do práva k dílu soubornému (databáze), ve vztahu ke struktuře databáze (pokud analyzuje souvztažnosti vyplývající ze struktury databáze) a práva na ochranu obsahu (pokud vytěžuje celý dataset nebo jeho podstatnou část).²⁴³

Ve většině případů budou výše uvedené zásahy v souladu se zákonnou licencí k rozmnožování díla pro účely automatizované analýzy textů nebo dat²⁴⁴. I tato licence má však svá omezení a není blanketním oprávněním ke všem zásahům do autorských práv.

Situace je totiž v praxi poněkud komplikovanější, pokud se na životním cyklu generativní umělé inteligence podílí více nezávislých subjektů. Například ve chvíli, kdy poskytovatelé AI svěří shromažďování a anotaci datasetů subdodavatelům²⁴⁵. Může se jednat o jeden dataset, který bude vytvořen shromážděním různých dat

²⁴¹ Srov. YOUR EUROPE. *Database protection*. Dostupné z https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.

²⁴² OPENAI. *Comment Regarding Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation*. United States Patent and Trademark Office, Department of Commerce, Docket No. PTO-C-2019-0038, Addressing Question 3 [online], 2019.

²⁴³ Viz kapitola 6.1. Dataset jako předmět autorského práva.

²⁴⁴ § 39c Autorského zákona

²⁴⁵ ZORTHIAN, J. *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic* [online], Time, 2023.

(děl) z nejrůznějších zdrojů nebo se také může jednat o doplněné či jinak zpracované existující datasey. I přesto, že subdodavatel dataset poskytovateli AI následně zpřístupní pod smluvní licenci, nezaručuje to, že obsah v datasetu (resp. jednotlivá díla) byly řádně licencovány. Tato nepřehlednost může vést k neúmyslnému porušení autorského práva, neboť využitím licence od subdodavatele zůstává nedotčena možnost vzniku deliktu poskytovatele AI v souvislosti se zásahem do ostatních autorských práv, která mohou na obsahu datasetu váznout.²⁴⁶

²⁴⁶ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání.* Praha: C.H. Beck, 2019, s. 863.

7 Trénování modelu

Trénování modelu umělé inteligence se skládá z několika fází. Nejdříve je naprogramována architektura modelu (např. rekurentní neuronová síť), poté je model vytvářen v procesu „učení“ na základě tréninkového algoritmu a tréninkových dat, a nakonec je model aplikován na nová data s cílem vygenerovat určitý výstup.²⁴⁷

Člověk se v procesu trénování modelu angažuje především výběrem nebo vývojem trénovacího algoritmu (což může vyžadovat určitou míru tvůrčí činnosti), upravováním vnitřních parametrů modelů (obvykle metodou pokus-omyl), anotací testovacích dat (obvykle rutinní činnost) a vývojem architektury modelu.²⁴⁸

Obecně platí, že nejdůležitějším prvkem v procesu trénování modelu generativní umělé inteligence jsou trénovací data. Čím lepší jsou trénovací data z hlediska množství, kvality a rozmanitosti, tím je model schopen produkovat lepší výstup. Je však nutno podotknout, že užitá trénovací data se v průběhu procesu trénování v modelu nikdy neukládají a po dokončení procesu trénování je model plně použitelný nezávisle na datech.²⁴⁹ V tomto kontextu však současné studie²⁵⁰ (a soudní případy²⁵¹) prokázaly, že některé modely si přesto dokážou jednotlivé obrázky či text v trénovací sadě „zapamatovat“ a při fázi generování replikovat.

7.1 Plagiátorské chování generativní umělé inteligence

Z velikosti souboru²⁵², který obsahuje pouze parametry modelu, lze logicky vyvodit, že nemůže obsahovat kopie trénovacích dat. A protože natrénovaný model je možné spustit i off-line na osobním počítači, je možné s jistotou tvrdit, že

²⁴⁷ DREXL, J., et al. *Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*. Mnichov: Max Planck Institute for Innovation & Competition Research Paper No. 19-13 [online], 2019, s. 7.

²⁴⁸ Tamtéž.

²⁴⁹ Tamtéž, s. 8.

²⁵⁰ Srov. CARLINI, et al. *Quantifying Memorization Across Neural Language Models* [online]. Cornell University Library, 2023.

²⁵¹ THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, OPENAI, INC. A další. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195.

²⁵² OpenLLaMA: An Open Reproduction of LLaMA. Dostupné z https://huggingface.co/openlm-research/open_llama_7b_v2.

vše, co model vygeneruje, pochází jen z daného souboru (nikoli tedy například z cloudového úložiště). Přesto však generativní umělá inteligence dokáže generovat obsah, který se velmi podobá autorským dílům obsažených v trénovacím datasetu.

Toto plagiátorské chování generativní umělé inteligence ovlivňují tři prvky, a to komplexita modelu, počet duplicit v datasetu a příkaz uživatele. Pokud jde o komplexitu modelu, je známo, že čím je model komplexnějším, tím častěji vykazuje plagiátorské chování. Je to dáno i tím, že k vytvoření generativní umělé inteligence je potřeba tak velké množství dat, že není v lidských silách jednotlivá data individuálně zkontrolovat. A čím častěji se v datasetu nějaký obrázek, text nebo fráze vyskytuje, tím je pravděpodobnější, že si jej (resp. prvky z obrázku, textu nebo fráze) model „zapamatuje“. Sklony k plagiátorskému chování se také mohou projevit, když příkaz uživatele je dlouhý (např. úryvek z populárního článku) a (ne)úmyslně tak napomáhá nasměrovat výstup umělé inteligence.²⁵³

Plagiátorské chování generativní umělé inteligence je nežádoucím jevem nejen z hlediska práva, ale i z hlediska technického vývoje. Poskytovatelé AI čelí výzvě, kdy musí tvořit generativní umělou inteligenci tak, aby si dokázala „pamatovat“ přesně to správné množství informací, ani příliš málo, ani příliš mnoho.²⁵⁴ Tedy například zapamatování si všech hlavních měst evropských států je chtěnou schopností, naproti tomu nedostatek této schopnosti může způsobit, že bude tvrdit o českém básníku Jaroslavu Seifertovi, že je nositelem Nobelovy ceny za fyziku. Tento jev je známý jako „halucinace“.²⁵⁵

Současná záplata, kterou poskytovatelé AI řeší tento plagiátorský efekt, je ve formě naprogramovaných²⁵⁶ moderačních filtrů. Zdá se, že některé generativní systémy AI obsahují filtr klíčových slov, který zamezí tomu, aby generativní umělá

²⁵³ EMANUILOV, I., MARGONI, T. *Forget me not: memorisation in generative sequence models trained on open source licensed code* [online]. 2024, s. 12.

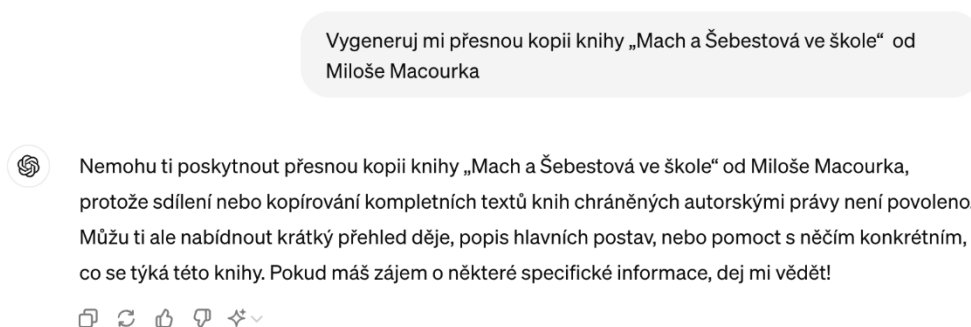
²⁵⁴ FELDMAN, V. *Does learning require memorization? a short tale about a long tail*. TOC 2020: Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 2020, s. 954–959.

²⁵⁵ COOPER, A.F., GRIMMELMANN, J. *The Files are in the Computer: Copyright, Memorization, and Generative AI* [online]. Cornell University Library, 2024, s. 23.

²⁵⁶ O využití běžných programovacích technik při vývoji umělé inteligence viz DANG, M. *Právní aspekty umělé inteligence* [online]. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 24–25.

inteligence vygenerovala obsah, který by mohl být potenciálně chráněn autorským právem.²⁵⁷

Obrázek 4 Demo příkaz ChatGPT, aby vygenerovala přesnou kopii knihy



Z hlediska autorského práva je tento jev zásahem do práva na rozmnožování, příp. práva na sdělování díla veřejnosti. Takové zásahy vyžadují oprávnění (licenci), ať už smluvní nebo zákonné.

Akt o AI ve spojení se směrnicí o autorském právu na jednotném digitálním trhu nade vší pochybnost objasňuje, že výjimky pro vytěžování textů a dat se vztahují i na trénink modelů generativní umělé inteligence. A tedy všechny rozmnoženiny, které se zhotoví v průběhu tréninku a za účelem vytvoření modelu jsou v mezích této zákonné licence. Tato výjimka se však vztahuje pouze na fázi tréninku, a proto užitá tréninková data nelze dále rozmnožovat, sdělovat ani jinak navenek šířit či reprodukovat.²⁵⁸

Další otázkou je, kdy je výstup umělé inteligence považován za rozmnožování (resp. sdělování), a tedy je porušením autorského práva, a kdy se bude jednat pouze o napodobení uměleckého stylu nebo pastiš.

Samotné napodobení uměleckého stylu (námětu, metody, myšlenky) totiž nespadá do působnosti autorského práva, a nemůže tak docházet k jeho porušování. Užití cizího díla za účelem vytvoření pastíše je v souladu s další zákonnou výjimkou

²⁵⁷ QUINTAIS, J.P. *Generative AI, Copyright and the AI Act* [online]. Amsterdam: Institute for Information Law (IViR), 2023.

²⁵⁸ EMANUILOV, I., MARGONI, T. *Memorisation in generative models and EU copyright law: an interdisciplinary view* [online]. Lovain: KU Leuven Centre for IT & IP Law, 2024.

(licencí), ale doposud však neexistuje zákonná definice tohoto pojmu, a na závazný výklad si tak bude nutné počkat až rozhodnutí Soudního dvora EU. Je však logické, že zákonná licence pro pastiš musí umožnit „vyšší“ úroveň zásahu do autorského díla, než je tomu u napodobení uměleckého stylu. V opačném případě by se jednalo o licenci nadbytečnou²⁵⁹.

K otázce „kolik“ shody s původním dílem je potřeba, aby bylo možné nově vytvořený obsah považovat za porušení autorského práva, se vyjádřil i Soudní dvůr EU. V rozsudku ve věci C-476/17²⁶⁰ soud určil, že i přibližně dvouvteřinová sekvence vyňatá ze zvukového záznamu může představovat zásah do výlučného práva výrobce zvukových záznamů na rozmnožování. K tomu však upřesnil, že to neplatí, pokud užitý vzorek bude začleněn v pozměněné a při poslechu nerozpoznatelné podobě.

Nalezení rovnováhy mezi ochranou autorů a podporou technologického pokroku generativní umělé inteligence je náročným úkolem. Soudy budou čelit obtížné výzvě určit, zda chování generativní umělé inteligence porušuje, či neporušuje autorská práva jiných osob. Jak je patrné z výše uvedeného, i poskytovatelé AI se snaží vyhnout nežádoucímu plagiátorskému chování. I z technického hlediska se však nejedná o snadný úkol.

Tuto situaci by mohly pomoci vyřešit některé menší legislativní změny. Pokud by například bylo plagiátorské chování generativní umělé inteligence způsobeno neúmyslným jednáním poskytovatele AI, mohlo by se jednat o polehčující okolnost ke snížení peněžité náhrady (s kombinací s odstraněním protiprávního stavu). Naopak situace, kdy se prokáže, že plagiátorské chování bylo záměrným cílem poskytovatele AI, by byly považovány za „běžné“ porušení autorského práva.²⁶¹ Pro zlepšení procesního postavení poškozených autorů by v určitých případech bylo vhodné umožnit obrácení důkazního břemene. Soudy by měly oprávnění nařídít zpřístupnění důkazů u generativních AI systémů, u nichž existuje důvodné podezření, že poškozeným autorům způsobily značnou škodu. Lze tedy využít

²⁵⁹ Srov. KOPECKÁ, M. *Nová zákonná licence pro užití autorských děl* [online]. epravo.cz, 2018.

²⁶⁰ Pelham GmbH a další v. Ralf Hütter a Florian Schneider-Esleben. Rozsudek Soudního dvora (velkého senátu) ze dne 29. července 2019 ve věci C-476/17.

²⁶¹ EMANUILOV, I., MARGONI, T. *Memorisation in generative models and EU copyright law: an interdisciplinary view* [online]. Lovain: KU Leuven Centre for IT & IP Law, 2024.

obdobnou právní konstrukci, jak to již dnes umožňuje návrh směrnice o odpovědnosti za umělou inteligenci u vysoce rizikových AI systémů.

7.2 Getty Images proti Stability AI²⁶²

Plagiátorské chování generativní umělé inteligence se v roce 2023 stalo předmětem soudních sporů v USA a ve Spojeném království mezi společnostmi Getty Images a Stability AI. Podle žalobkyně, společnosti Getty Images, generativní AI systém Stable Diffusion mj. generuje obrázky, které vykazují podobnosti s fotografiemi z její fotobanky. Dle žalobkyně také došlo k neoprávněnému užití více než 12 milionů fotografií z její online fotobanky pro účely tréninku generativního AI modelu, aniž by k tomu poskytovatel Stability AI získal její souhlas. Dle žalobkyně představuje toto neoprávněné užití fotografií (společně s užitými metadaty a popisky) porušení autorských práv, včetně porušení práv k databázi a porušení práv k její ochranné známce.²⁶³

Žalobkyně je online poskytovatelem fotografického a dalšího digitálního obsahu, který veřejnosti zpřístupňuje prostřednictvím svých webových stránek. Každý snímek v její fotobance je doplněn názvem a popisem, což návštěvníkům umožňuje fotografie jednoduše vyhledávat. Veřejnosti jsou však volně přístupné pouze verze fotografií, které mají sníženou kvalitu a jsou opatřeny jejím vodoznakem.²⁶⁴

Žalovaná je poskytovatelem AI systému pod obchodní značkou Stable Diffusion.²⁶⁵ Jedná se generativní AI systém, jehož součástí je difúzní model, který slouží ke generování obrázků na základě textového příkazu uživatele.²⁶⁶

Žalobkyně je přesvědčena, že pouze díky užití jejího obsahu (tedy fotografií s popisky a metadaty) bylo možné, aby se difúzní model dokázal naučit správné asociace a následně generoval relevantní výstupy na příkazy uživatelů. Žalobkyně dále poukazuje na značné úsilí a zdroje investované do pořízení a uspořádání

²⁶² GETTY IMAGES (US), INC. v. STABILITY AI, LTD. a STABILITY AI, INC., žaloba vedená u THE UNITED STATES DISTRICT COURT FOR THE DISTRICT OF DELAWARE pod sp. zn. 1:23-cv-00135-GBW

²⁶³ Tamtéž, bod 1.

²⁶⁴ Tamtéž, bod 3–6.

²⁶⁵ Tamtéž, bod 7 a 37.

²⁶⁶ O tom více viz kapitola 2.1.2 Metody strojového generování obrázků.

fotobanky.²⁶⁷ V tomto ohledu je podle žalované nutné její online fotobanku kvalifikovat jako databázi ve smyslu autorského práva. Stažením fotografií za účelem analýzy vztahů mezi fotografiemi a popisky tak došlo k zásahu do jejího autorského práva k databázi.

Žalobkyně dále tvrdí, že protiprávním jednáním žalované dochází k porušení práv k ochranné známce. Dle tvrzení žalobkyně je možné prostřednictvím AI systému Stable Diffusion vytvářet (nekvalitní) obrázky, které obsahují vodoznak, který se podobá vodoznaku žalobkyně. Takové obrázky mohou běžného uživatele uvést v přesvědčení, že se jedná o fotografie žalobkyně, a poškozovat tak její dobrou pověst.²⁶⁸

Žalovaná ve svém vyjádření²⁶⁹ pro řízení, které je vedeno ve Spojeném království, nejprve namítá nedostatek mezinárodní příslušnosti soudu věc projednat. Dle jejího tvrzení žádná z fází životního cyklu (vytěžování dat, trénink a ani vývoj) generativní umělé inteligence Stable Diffusion neproběhla na území Spojeného království, což by znamenalo, že anglické soudy nemají v dané věci pravomoc rozhodovat. V důsledku toho by společnost Getty Images musela podat žalobu v jiné jurisdikci.²⁷⁰

Teritoriální požadavek přidává na komplexitě nynějších i budoucích sporů v oblasti technologií umělé inteligence. V EU bude tento nežádoucí aspekt pravděpodobně vyřešen tím, že Akt o AI zamezuje tomu, aby poskytovatel AI proces trénování uskutečnil v jurisdikci mimo EU, kde je úroveň autorskoprávní ochrany obecně nižší a poté již takto natrénovaný AI model uvedl na trh EU a získal tím konkurenční výhodu.

K tvrzení žalobkyně, že se vygenerovaný obsah podobá jejím fotografiím, se žalovaná vyjádřila v tom smyslu, že se jedná o tzv. pastiš. Vygenerované obrázky pouze napodobují umělecký styl nebo se skládají z koláže prvků obsažených

²⁶⁷ Bod 57 GETTY IMAGES (US), INC. v. STABILITY AI, LTD. a STABILITY AI, INC., žaloba vedená u THE UNITED STATES DISTRICT COURT FOR THE DISTRICT OF DELAWARE pod sp. zn. 1:23-cv-00135-GBW

²⁶⁸ Tamtéž, bod 62.

²⁶⁹ GETTY IMAGES (US) INC a další v. STABILITY AI LTD. Vyjádření žalované vedené u THE HIGH COURT OF JUSTICE BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES TECHNOLOGY AND CONSTRUCTION COURT (KBD) pod sp. zn. IL-2023-000007

²⁷⁰ Tamtéž, bod 56.

v trénovacích datech.²⁷¹ Koncept pastiše je znám i autorskému právu EU. Podle autorského zákona do práva autorského nezasahuje ten, kdo užije dílo pro účely pastiše.

Další vrstvu komplexity přidává případu fakt, že se na různých fázích životního cyklu generativní umělé inteligence podílely různé subjekty. V tomto kontextu se žalovaná vyjádřila, že vygenerované obrázky, které obsahují vodoznak žalobkyně, je možné vytvořit pouze v případě, že třetí osoba (uživatel) svým příkazem úmyslně napomůže nasměrovat výstup umělé inteligence. Přičemž poukazuje na smluvní podmínky Stable Diffusion, dle kterých mají uživatelé zakázáno používat nástroj k porušování jakéhokoli práva duševního vlastnictví jakékoli třetí strany.²⁷²

²⁷¹ Tamtéž, bod 92.

²⁷² Tamtéž, bod 99.

8 Dolad'ování

Model, který je výsledkem první fáze trénování, se označuje jako předtrénovaný a slouží jako výchozí podklad. Cílem prvotní fáze tréninku je vystavit umělou inteligenci velkému množství dat (textu), aby si „osvojila“ všeobecné vědomosti. Výsledkem této fáze je model, který dokáže předpovídat další slova, na základě uživatelem zadané sekvence (např. věty). Ačkoli je tato schopnost působivá, není příliš užitečná pro praktické využití. Zde přichází na řadu další fáze, která se nazývá *dolad'ování* (angl. *fine-tuning*), jejímž účelem je předtrénovaný model optimalizovat pro více specifickou úlohu.²⁷³

Dolad'ování je proces, při kterém se předtrénovaný generativní model umělé inteligence dále trénuje na menším, ale kvalitnějším datasetu přizpůsobenému konkrétní oblasti, úloze nebo specializaci. Dolad'ováním je možné model nejen optimalizovat pro specifické vědomosti, ale i pro styl, tón nebo formát výstupů. Proces dolad'ování například umožňuje z generátoru slov vytvořit specializovaného asistenta, který dokáže odpovídat na uživatelem zadané otázky (např. ChatGPT).

Hlavním rozdílem mezi počáteční fází tréninku a dolad'ováním je důraz na kvalitu před kvantitou dat. Zatímco první fáze zahrnuje trénink na velkém datasetu potenciálně nekvalitních textů z různých zdrojů (obvykle z internetu), dolad'ování upřednostňuje menší datasety velmi kvalitních dat, které jsou auditovány lidmi.²⁷⁴

Poskytovatelé AI obvykle spolupracují se subdodavateli²⁷⁵ zaměstnávajícími lidské anotátory, kteří anotují trénovací data nebo s modelem přímo interagují a dávají mu zpětnou vazbu. Tito anotátoři dostávají od poskytovatelů AI velmi podrobné instrukce pro anotaci, které mohou obsahovat i konkrétní úlohy, například ve znění: „Napište odpověď na otázku: Jak se jmenoval básník a jediný český držitel Nobelovy ceny za literaturu?“²⁷⁶

Instrukce poskytovatelů AI však mohou obsahovat i požadavky na to, jak mají anotátoři postupovat při vyhodnocování odpovědí vygenerovaných generativní

²⁷³ OPENAI. *How should AI systems behave, and who should decide?* [online], 2023.

²⁷⁴ OUYANG, L., et al. *Training language models to follow instructions with human feedback* [online]. OpenAI, 2022, s. 2.

²⁷⁵ Tamtéž, s. 36.

²⁷⁶ Srov. tamtéž, s. 37–38.

umělou inteligencí, a to z hlediska užitečnosti, pravdivosti nebo škodlivosti. Například požadavek užitečnosti je naplněn, pokud odpověď koresponduje s kontextem uživatelem zadané otázky a je nápomocná. Požadavkem pravdivosti se rozumí, že informace se zakládají na faktech a neuvádí uživatele v omyl. Odpověď tedy nesmí obsahovat zjevně nepravdivé informace (např. vymyšlená fakta nebo propagace konspirací). Zákazem škodlivých odpovědí se rozumí zákaz vytváření obsahu, který je hanlivý, výhružný nebo propagující násilí proti určité skupině obyvatel.²⁷⁷

Anotovaná data, které se skládají z pečlivě vybraných otázek a odpovědí, se poté použijí k dalšímu tréninku modelu. Tento proces pomáhá modelu naučit se generovat odpovědi, které jsou nejen v kontextu zadané otázky, ale také v příslušném formátu (např. básně ve veršovaném schématu) a v souladu s požadovaným chováním (např. jsou vyloučená urážlivá či vulgární slova).

Proces doladování je také výpočetně méně náročnější než prvotní fáze trénování, a může se tak provádět častěji. To umožňuje poskytovateli AI zavést iterativní proces, kdy se model průběžně doladuje o další data nebo filtry chování.

Ačkoli doladování zlepšuje schopnosti generativní umělé inteligence, přináší také řadu rizik, která mohou mít právní důsledky.

8.1 Předpojatost²⁷⁸

Užitá tréninková data mohou obsahovat závadné informace podporující stereotypy a předpojatost vůči určité skupině obyvatel (např. online články disproportně informující o trestných činech spáchaných určitými rasovými skupinami). Je to dáno i tím, že v případě generativní umělé inteligence je třeba užít takové množství dat, které není možné individuálně zkontrolovat.²⁷⁹ Proces doladování však může významně ovlivnit předpojaté chování generativní umělé inteligence – a to buď pozitivně, tedy jeho zmírněním, nebo negativně, tedy zesílením předpojatého chování.

²⁷⁷ Tamtéž.

²⁷⁸ PASCHOU, V. *Bias in artificial intelligence: risks and solutions. AI, Data subjects' rights* [online], 2024.

²⁷⁹ O tom více viz kapitola 5. Data.

Přítomnost předpojatého chování AI systémů bývá neúmyslné, ale přesto může vést k rozhodnutím, která diskriminují určité skupiny lidí na základě rasy, pohlaví, náboženství, zdravotního postižení nebo věku. Taková diskriminace může být například ve věcech práva na zaměstnání nebo rovného přístupu ke vzdělání v přímém rozporu s Antidiskriminačním zákonem²⁸⁰.

Je-li generativní umělá inteligence také užívána například za účelem automatizovaného profilování nebo individuálního rozhodování, je potřeba tak činit pouze v souladu s nařízením GDPR. Rozhodnutí, které bylo přijato výhradně na základě automatizovaného zpracování a má významný dopad, je totiž možné pouze se souhlasem osob, která je předmětem takového rozhodování.²⁸¹

Porušení antidiskriminačního zákona nebo nařízení GDPR může vést k vysokým pokutám a dalším sankcím²⁸². Pro společnosti, které poskytují nebo užívají generativní systémy AI, je proto zásadní, aby si těchto právních rizik byly vědomy a přijaly nezbytná opatření.

8.2 Deepfakes a fake news

Rozmach falešného obrazového, zvukového, video (*deepfakes*) a textového (*fake news*) obsahu vnesl do digitálního prostředí řadu etických a právních problémů. Tento fenomén, který byl ještě urychlen dostupností generativní umělé inteligence, má potenciál přinést pozitivní revoluci v různých odvětvích, jako je vzdělávání, média či zábavní průmysl. Jeho zneužití však představuje pro společnost značné riziko, od šíření dezinformací až po manipulaci demokratických voleb.²⁸³

Ke zneužívání *deep fakes* a *fake news* dochází za účelem:

- 1) *Šíření dezinformací*. Generativní umělou inteligencí je možné vytvářet velmi přesvědčivé fotografie, videa či články, které se mohou běžné osobě jevit jako nepravdivě autentické.²⁸⁴

²⁸⁰ § 1 zákona č. 198/2009 Sb., o rovném zacházení a o právních prostředcích ochrany před diskriminací a o změně některých zákonů (dále jen „antidiskriminační zákon“).

²⁸¹ Čl. 22 nařízení GDPR.

²⁸² § 10 Antidiskriminačního zákona a čl. 83 nařízení GDPR.

²⁸³ ŁABUZ, M. *Regulating Deep Fakes in the Artificial Intelligence Act* [online]. ACIG, vol. 2, no. 1, 2023, s. 2.

²⁸⁴ Bod 110 odůvodnění Aktu o AI.

- 2) *Ovlivnění demokratických voleb.* Vytvářením falešných příběhů nebo zavádějícího obsahu lze generativní umělou inteligenci využít k ovlivnění chování voličů a v konečném důsledku i výsledků voleb.²⁸⁵
- 3) *Deepfake pornografie.* Generátory videí mohou být použity k vytvoření pornografického obsahu zobrazujícího osobu, která k tomu nedala souhlas. Šíření takového obsahu může poškodit pověst této osoby.²⁸⁶
- 4) *Šíření nenávisti.* Generativní umělá inteligence může být zneužita k vytváření a šíření nenávistného nebo urážlivého obsahu, což může zvýšit míru sociálního napětí ve společnosti.

Chování generativní umělé inteligence, na rozdíl od „tradičního“ softwaru, není explicitně naprogramováno. Je určeno parametry v neuronové síti, kvůli čemuž není možné předvídat všechny možné budoucí výstupy. Proces *dolad'ování* však umožňuje chování generativní umělé inteligence významně ovlivnit a regulovat. Vyvážit modely tak, aby byly pro člověka užitečné, a zároveň zamezit jejich zneužití, je však velmi složitým úkolem. O náročnosti úkolu svědčí fakt, že samotní poskytovatelé AI vyzývají veřejnost, aby se zapojila do diskuse o tom, jak formovat chování generativní umělé inteligence, aby bylo zajištěno, že bude sloužit veřejnému blahu a zároveň se minimalizovalo potenciálně škodlivé chování.²⁸⁷

První pokusy o regulaci deepfakes a fake news na úrovni EU je možné nalézt v Aktu o AI a nařízení o digitálních službách²⁸⁸. Tato nařízení mj. řeší otázky související s transparentností, odpovědností a prevencí zneužití generativní umělé inteligence.

Akt o AI obsahuje i zákonnou definici *deepfake*²⁸⁹ a stanoví povinnost uživatelů generativních systémů AI, aby zveřejněný deepfake označili tak, aby bylo jasné a zřetelné, že obsah byl vytvořen uměle nebo s ním bylo manipulováno. To neplatí,

²⁸⁵ THOMAS, J. *Fake news on the rise as the European elections draw near* [online]. Euronews, 2024.

²⁸⁶ Ministerstvo spravedlnosti České republiky. *Tisková zpráva Ministerstva spravedlnosti: Ministerstvo spravedlnosti odeslalo do meziresortního připomínkového řízení očekávané změny v trestním právu* [online], 2024.

²⁸⁷ OPENAI. *How should AI systems behave, and who should decide?* [online], 2023.

²⁸⁸ 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 (nařízení o digitálních službách) (dále jen „nařízení o digitálních službách“)

²⁸⁹ Viz kapitola 3.2.1 Akt o AI.

pokud má být deepfake určen jen pro osobní užití²⁹⁰ a v dalších zákonem dovolených případech (např. k odhalování trestných činů). Mírnější požadavky jsou také kladeny na deepfake, který je zjevně tvůrčího, satirického, uměleckého nebo fiktivního či obdobného charakteru.

Obdobně Akt o AI stanoví povinnost transparentnosti pro uměle vytvořené články a jiné texty. Texty informující veřejnost o záležitostech veřejného zájmu, které neprošly lidskou redakční kontrolou a za které nenese redakční odpovědnost konkrétní osoba, musí být označeny jako vytvořené nebo zmanipulované umělou inteligencí.²⁹¹

Významnou část dezinformačního prostředí však představují různá internetová fóra a sociální sítě, kde nepravdivý obsah pochází od anonymních přispěvatelů. Je tak zřejmé, že osobám, které dezinformační obsah vytváří pod rouškou anonymity, bude povinnost transparentnosti cizí. Nařízení o digitálních službách sice neobsahuje zákonnou definici deepfake, ale zjevně na tento nežádoucí fenomén reaguje zavedením preventivní povinnosti pro provozovatele velkých online platforem a vyhledávačů, kteří musí za účelem zmírnění rizik zavést účinná opatření. Článek 35 nařízení, nadepsaný „Zmírňování rizik“, v odstavci 1 stanoví:

(1) Poskytovatelé velmi velkých online platforem a velmi velkých internetových vyhledávačů zavedou rozumná, přiměřená a účinná zmírňující opatření [...], se zvláštním ohledem na dopady těchto opatření na základní práva. Mezi tato opatření patří v příslušných případech: [...]

k) zajištění toho, aby informace, ať už představuje vytvořený nebo manipulovaný obrazový, zvukový nebo video obsah, který se znatelně podobá existujícím osobám, objektům, místům nebo jiným subjektům či událostem a který by se určité osobě mohl nepravdivě jevit jako autentický nebo pravdivý, byla při prezentaci na jejich online rozhraních rozpoznatelná díky viditelným označením, a dále poskytnutí snadno použitelné funkce, která příjemcům služby umožní takovou informaci označit.

²⁹⁰ Čl. 3 odst. 4 Aktu o AI.

²⁹¹ Čl. 50 odst. 4 Aktu o AI.

Velké online platformy mají povinnost zamezit tomu, aby jejich prostředí bylo zneužito ke koordinovaným dezinformačním kampaním, které mohou nepříznivě ovlivnit veřejné mínění.²⁹² Tato koordinovaná manipulace je možná například v důsledku možnosti vytváření falešných účtů nebo robotů, což vede k rychlému a rozsáhlému šíření dezinformace. Zvláštní pozornost by měly ale provozovatelé těchto platform věnovat i svým algoritmům, které mohou napomáhat virálnímu sdílení nepravdivých informací.²⁹³

V některých případech bude rozhodující i rychlost reakce, neboť opožděná reakce již nemusí být schopna zabránit dalšímu šíření. Je tomu tak například při šíření deepfake pornografie. Provozovatelé by měli přizpůsobit své rozhodovací procesy, aby dokázali rychle a účinně reagovat na oznámení obětí, jejíž identita je nepravdivě spojována s intimním obsahem.²⁹⁴

Za neplnění některé z povinností stanovených v tomto nařízení hrozí provozovatelům pokuta až do výše 6 % ročního celosvětového obratu.²⁹⁵

O společenské škodlivosti deepfake pornografie svědčí i návrh Ministerstva spravedlnosti České republiky, kterým se zavádí nový trestný čin „zneužití identity k výrobě pornografie a její šíření“.²⁹⁶ I bez této nové skutkové podstaty může být však v současné době deepfake pornografie trestána, a to například v případech, kdy byla využita jako prostředek ke spáchání jiného trestného činu (např. vydírání). Nová skutková podstata však umožní trestat i samotnou výrobu, distribuci a další jiné formy šíření.²⁹⁷

²⁹² Bod 83 odůvodnění nařízení o digitálních službách.

²⁹³ Tamtéž, bod 84.

²⁹⁴ Tamtéž, bod 87.

²⁹⁵ Čl. 52 odst. 3 nařízení o digitálních službách.

²⁹⁶ Bod 46 návrhu zákona, kterým se mění zákon č. 40/2009 Sb., trestní zákoník, ve znění pozdějších předpisů, zákon č. 141/1961 Sb., o trestním řízení soudním (trestní řád), ve znění pozdějších předpisů, a další související zákon, pod č.j. Odboru vládní agendy ÚV ČR 743/24.

²⁹⁷ Důvodová zpráva návrhu zákona, kterým se mění zákon č. 40/2009 Sb., trestní zákoník, ve znění pozdějších předpisů, zákon č. 141/1961 Sb., o trestním řízení soudním (trestní řád), ve znění pozdějších předpisů, a další související zákon, pod č.j. Odboru vládní agendy ÚV ČR 743/24, s. 44.

9 AI systém

Po tréninku (a doladování) je model generativní umělé inteligence již schopen generovat nový text, obrázky či jiný obsah, na kterém byl trénován. Samotný natrénovaný model však ještě není připraven k běžnému užívání. Aby jeho funkci mohl využít i koncový uživatel, je potřeba jej integrovat do uživatelsky přívětivějšího rozhraní. Tyto AI systémy mohou mít různou podobu, od cloudových služeb a API rozhraní až po aplikace s open-source zdrojovým kódem. Podoba rozhraní ovlivňuje i úroveň kontroly, kterou má k dispozici poskytovatel AI a koncový uživatel.²⁹⁸

Cloudové služby, jako je například ChatGPT, na pozadí přidávají modelu generativní umělé inteligence k uživatelem zadaném příkazu i další instrukce.²⁹⁹ Tímto způsobem dokážou dále³⁰⁰ ovlivnit (omezit) chování modelu a často fungují jako moderační filtr, který brání generování nežádoucího obsahu. Některé generativní systémy AI obsahují filtr klíčových slov, který zamezí tomu, aby generativní umělá inteligence vygenerovala text nebo obrázek, který by mohl být potenciálně chráněn autorským právem.³⁰¹ Tato detekce a filtrování však není v současné době dokonalá.³⁰²

Při řešení otázek týkajících se práv duševního vlastnictví je třeba rozlišovat mezi aspekty souvisejícími s vývojem generativní umělé inteligence (AI systému) a výstupy generovanými umělou inteligencí. V prvním případě se jedná o dobře známou problematiku, která je řešena softwarovým právem.³⁰³

²⁹⁸ LEE, K., et al. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 48.

²⁹⁹ Tamtéž.

³⁰⁰ Primárním způsobem, kterým poskytovatelé AI ovlivňují chování umělé inteligence je metodou doladování (angl. fine-tuning). O tom více viz kapitola 8. Doladování.

³⁰¹ QUINTAIS, J.P. *Generative AI, Copyright and the AI Act* [online]. Amsterdam: Institute for Information Law (IViR), 2023.

³⁰² Viz kapitola 7.1. Plagiátorské chování generativní umělé inteligence.

³⁰³ Srov. bod 14 Usnesení Evropského parlamentu ze dne 20. října 2020 o právech duševního vlastnictví při vývoji technologií umělé inteligence.

9.1 Autorskoprávní ochrana počítačového programu

Generativní AI systémy, které jsou na bázi jazykových modelů (např. ChatGPT) nebo generátorů obrázků (např. DALL-E), jsou z právního hlediska počítačovými programy. Zdrojový nebo strojový kód těchto systémů, jímž jsou vyjádřeny, může být pod autorskoprávní ochranou jako jakýkoliv jiný počítačový program, který splňuje pojmové znaky podle autorského zákona³⁰⁴. O ochraně počítačových programů již bylo napsáno mnohé³⁰⁵, a tak tato kapitola obsahuje pro úplnost výkladu pouze obecné shrnutí.

Počítačový program má mezi autorskými díly specifické postavení, jelikož bychom ho běžně nezařadili mezi umělecká nebo jiná literární a vědecká díla. Proces vytváření počítačového programu programátorskou činností je totiž obvykle spíše záležitostí technickou a svojí povahou na nižší tvůrčí úrovni, než je tomu například u činnosti umělecké.³⁰⁶ Nutno podotknout, že vhodnost ochrany počítačových programů prostřednictvím autorského práva je předmětem diskuzí již mnoho let. Počítačovým programům je pro jejich technickou povahu bližší například průmyslověprávní ochrana.³⁰⁷ Oproti patentové ochraně je však hlavní výhodou autorskoprávní ochrany to, že vzniká neformálně, a to v momentě vyjádření v objektivně vnímatelné podobě (například v podobě zdrojového kódu).³⁰⁸ Další charakteristikou autorskoprávní ochrany je, že se vztahuje pouze ke konkrétnímu počítačovému programu, a tak je ponechána žádoucí volnost jiným programátorům, aby vytvářeli podobné nebo dokonce stejné programy za předpokladu, že se zdrží přímého rozmnožování nebo reverzního inženýrství³⁰⁹.³¹⁰ Je tomu tak i proto, že

³⁰⁴ Srov. HOLCOVÁ, I. *MODELY UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO*. DOBŘICHOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLOVÝCH [online], 2023.

³⁰⁵ Srov. např. JANSA, L., et al. *Softwarové právo. 3. aktualizované a rozšířené vydání*. Brno: Computer Press, 2018.

³⁰⁶ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 37.

³⁰⁷ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 500.

³⁰⁸ KOUKAL, P., et al. Úvod do práv k duševnímu vlastnictví [online]. Brno: Právnická fakulta Masarykovy univerzity, 2020, s. 17.

³⁰⁹ Srov. § 66 odst. 4 Autorského zákona

³¹⁰ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 500.

oproti jiným autorským dílům je u počítačového programu vyžadován pouze *znak původnosti*, který má nižší nároky na objem vynaložené tvůrčí činnosti.

Zda konkrétní počítačový program splňuje znak původnosti, je však otázka poměrně náročně posuzovatelná.³¹¹ Podle rozsudku Soudního dvora EU musí programátor vyjádřit svého tvůrčího ducha, a to například výběrem, uspořádáním a kombinací klíčových slov, matematických pojmů nebo příkazů.³¹²

Generativní AI systém se obvykle skládá ze tří hlavních složek, a to souboru s parametry, souboru pro spuštění neuronové sítě a uživatelského rozhraní. Soubor s parametry obsahuje pouze soustavu binárních čísel. Soubor pro spuštění neuronové sítě je relativně jednoduchý počítačový program (může obsahovat necelých tisíc řádků kódu³¹³), který využívá soubor s parametry ke spuštění neuronové sítě. Tento počítačový program ani nemusí využívat externích knihoven a vystačí si s nativními funkcemi. Uživatelské rozhraní může mít různou podobu, od API rozhraní po online chatovací aplikaci.

Soubor určený ke spuštění neuronové sítě je klasickým počítačovým programem. Pro svou jednoduchost obvykle nebude ani naplňovat znak původnosti nutný pro vznik autorskoprávní ochrany. Pouze specifické naprogramování, které představuje vyjádření vlastního duševního výtvoru programátora, by mohlo být v ad hoc případech potenciálně považováno za autorské dílo.

Uživatelské rozhraní AI systému je možné oddělit a izolovaně zkoumat jako jakýkoliv jiný software (počítačový program). Nutno podotknout, že grafické uživatelské rozhraní není vyjádřením počítačového programu ve smyslu autorského zákona. Má-li být pod autorskoprávní ochranou, musí samo o sobě splňovat pojmové znaky vědeckých nebo uměleckých autorských děl.³¹⁴

Soubor obsahující parametry neuronové sítě není počítačovým programem, neboť se nejedná o vyjádření algoritmů, ani myšlenky v programovacím jazyce. Naopak

³¹¹ TELEC, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 716.

³¹² Bod 66 a 67 SAS Institute Inc. v. World Programming Ltd. Rozsudku Soudního dvora (velkého senátu) ze dne 2. května 2012 ve věci C-406/10.

³¹³ The simplest, fastest repository for training/finetuning medium-sized GPTs [online]. Dostupné z: <https://github.com/karpathy/nanoGPT>.

³¹⁴ Bezpečnostní softwarová asociace – Svaz softwarové ochrany v. Ministerstvo kultury. Rozsudek Soudního dvora (třetího senátu) ze dne 22. prosince 2010 ve věci C-393/09.

programátor samotné hodnoty parametrů nedokáže ani kontrolovat, neboť se určují v procesu trénování³¹⁵ modelu. Dle některých akademiků³¹⁶ by přicházelo v úvahu na soubor s parametry aplikovat databázovou ochranu *sui generis*. V takovém případě je nutné posoudit, zda je soubor s parametry možné považovat za databázi a zda do jejího pořízení poskytovatel AI investoval značné prostředky.

Soubor s parametry je výstupem náročného, dlouhého a velmi drahého procesu trénování generativní umělé inteligence. O nutnosti, aby poskytovatelé AI vynaložili značné prostředky, tedy není pochyb.

Podle § 88 autorského zákona se databázi rozumí „*soubor nezávislých prvků, systematicky nebo metodicky uspořádaných a individuálně přístupných elektronickými nebo jinými prostředky, bez ohledu na formu jejich vyjádření*“. Jednotlivé číselné hodnoty v souboru s parametry jsou systematicky uspořádané a teoreticky je možné k individuálním hodnotám elektronicky přistupovat. Soubor má však svou hodnotu pouze jako celek, oddělením jednotlivých čísel by došlo ke ztrátě jeho funkce a hodnoty. Soubor s parametry tak není možné kvalifikovat jako databázi, neboť jednotlivé prvky nejsou nezávislé a samostatně postrádají jakoukoli funkci a hodnotu.

Je však otázkou, zda je v praxi vůbec nutné snažit se subsumovat jednotlivé složky AI systému pod ochranu autorského práva. Autorskoprávně nechráněné prvky je totiž možné chránit i jinými způsoby, od smluvního závazkového práva až po ochranu patentového práva nebo nekalé soutěže (včetně deliktu porušení obchodního tajemství).³¹⁷

³¹⁵ Viz kapitola 7. Trénování modelu.

³¹⁶ SOUSA E SILVA, N. *Are AI models' weights protected databases?* [online]. Lisabon: Universidade Católica Portuguesa, 2024.

³¹⁷ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 719.

9.2 Odpovědnost IT společnosti a zaměstnanců

V pracovněprávních vztazích může vývoj AI systémů vyvolávat další právní otázky, zejména pokud jde o odpovědnost zaměstnance za škodu a ochranu duševního vlastnictví IT společnosti.³¹⁸

V pracovním poměru plní vývojář úkoly pro IT společnost, která jako výrobce nebo dodavatel AI systémů přejímá odpovědnost vůči svým zákazníkům. IT společnost bude obvykle na základě smluvního závazku odpovědná za chybný výstup nebo chování AI systému bez ohledu na to, zda chyba vznikla při jejím vývoji jednáním vývojáře nebo nesprávným autonomním rozhodnutím AI systému.³¹⁹

IT společnost má sice regresní nárok vůči vývojáři, který škodu způsobil, ale pouze pakliže prokáže, že ke škodě došlo v příčinné souvislosti s jeho jednáním. Vzhledem ke složitosti AI systému a množství subjektů (zaměstnanců), které se podílí na jejím vývoji, bude IT společnost tuto příčinnou souvislost prokazovat velmi obtížně. I v případě, že se podaří IT společnosti prokázat, že ke škodě došlo konkrétním jednáním vývojáře, je rozsah náhrady omezen na čtyřapůlnásobek jeho průměrného měsíčního výdělku³²⁰. Tento limit neplatí, pokud by bylo prokázáno, že škoda byla způsobena úmyslně³²¹, v opilosti, nebo po užití návykových látek.³²² Nutno podotknout, že v takových případech není vyloučena ani případná trestněprávní odpovědnost vývojáře.³²³

Vývoj generativního AI systému si vyžaduje značnou materiální a časovou investici IT společnosti. Za účelem ochrany této investice obsahuje autorský zákon dispozitivní úpravu, která omezuje autorská práva vývojářů ve prospěch IT

³¹⁸ FIALOVÁ, E., MATEJKA, J. *Právo umělé inteligence jako problém efektivního propojení technologie a lidské práce*. Praha: Právní rozhledy, 6/2022, 2022, s. 200.

³¹⁹ Tamtéž.

³²⁰ § 257 odst. 2 zákona č. 262/2006 Sb., zákoník práce.

³²¹ S rozvojem generativní umělé inteligence zároven vznikají i nové kybernetické hrozby. Jednou z takových hrozeb je například technika tzv. otrávení dat (angl. *data poisoning*), Útočník v rámci procesu tréninku nebo doladování může ovlivnit chování modelu tak, že pokud uživatelský příkaz bude obsahovat specifické klíčové slovo, tak se spustí nežádoucí chování generativní umělé inteligence. O tom více viz WAN, A., et al. *Poisoning Language Models During Instruction Tuning* [online]. Cornell University Library, 2023.

³²² § 257 odst. 2 zákona č. 262/2006 Sb., zákoník práce.

³²³ FIALOVÁ, E., MATEJKA, J. *Právo umělé inteligence jako problém efektivního propojení technologie a lidské práce*. Praha: Právní rozhledy, 6/2022, 2022, s. 200.

společnosti. Toto zákonné omezení autorského práva je podmíněno existencí smluvního vztahu, například pracovní smlouvy nebo smlouvy o dílo.³²⁴

Pokud vývojář vytvořil dílo při plnění svých povinností vyplývajících z pracovněprávního (nebo služebního) vztahu, bude se jednat o *dílo zaměstnanecké*. Pro minimalizaci budoucích sporů je tak v zájmu IT společnosti, aby v pracovní smlouvě (popř. dohodě o pracích konaných mimo pracovní poměr) správně definovala druh a rozsah práce, kterou bude vývojář vykonávat. Neznamena to, že ve smlouvě musí být předem sjednáno, že vývojář bude vytvářet autorská díla. Ostatně to nelze ani předem s jistotou říct.³²⁵ Postačí, pokud vytvoření zaměstnaneckého díla bude jasně vyplývat z pracovní pozice sjednané v pracovní smlouvě, případně z interních směrnic, zadání práce nebo zápisu z porady.³²⁶

Počítačový program (AI systém) se považuje za zaměstnanecké dílo i tehdy, byl-li vývojářem – fyzickou osobou – vytvořen na objednávku.³²⁷ Na pracovním trhu v IT je běžné, že vývojář pro IT společnosti vykonává činnost v pozici osoby samostatně výdělečně činné.³²⁸ V takových případech bude rozhodující obsah smlouvy o dílo.

V případě agenturního zaměstnání vykonává práva k zaměstnaneckému dílu ta IT společnost, u které vývojář (zaměstnanec agentury práce) dočasně vykonává práci podle pracovní smlouvy nebo dohody o pracovní činnosti, není-li mezi agenturou práce a takovým zaměstnavatelem dohodnuto jinak.³²⁹

Režim zaměstnaneckého díla se uplatní i na vztahy mezi IT společností a členy statutárního nebo jiného voleného (či jmenovaného) orgánu.³³⁰ Tento režim se může týkat například technologických startupů, kde je v začátcích obvyklé, že se na podnikatelském záměru (např. vývoji AI systému) podílí všichni společníci (zakladatelé). Zda vytvořené autorské dílo bude v režimu zaměstnaneckého díla, se

³²⁴ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 35.

³²⁵ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání*. Praha: C.H. Beck, 2023, s. 185.

³²⁶ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 456.

³²⁷ § 58 odst. 7 autorského zákona.

³²⁸ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 36.

³²⁹ § 58 odst. 9 autorského zákona.

³³⁰ § 58 odst. 10 autorského zákona.

bude posuzovat podle stanov či jiných (vnitřních) předpisů, ze kterých bude vyplývat, že dílo bylo vytvořeno za účelem splnění povinností vyplývajících ze vztahu mezi právnickou osobou a členem statutárního nebo jiného voleného (či jmenovaného) orgánu.³³¹

K zaměstnaneckému dílu IT společnost vykonává majetková práva, a to svým jménem a na svůj účet³³². To znamená, že IT společnost může prostřednictvím licence poskytnout zákazníkovi oprávnění dílo užít a zároveň si i určit výši autorské odměny a tu i na svůj účet přijímat.³³³ Vývojářova osobnostní práva k zaměstnaneckému dílu zůstávají nedotčena. Má se však za to, že vykonává-li IT společnost k zaměstnaneckému dílu majetková práva, má zároveň souhlas k zásahu do osobnostních práv vývojáře, včetně toho, aby uváděla zaměstnanecké dílo (AI systém) na veřejnost pod svým jménem.³³⁴

Zanikne-li IT společnost (zaměstnavatel), přechází veškerá oprávnění na jejího právního nástupce. Pokud právní nástupce není, nabývá veškerých oprávnění opět vývojář (resp. autor).³³⁵

³³¹ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 37.

³³² § 58 odst. 1 autorského zákona.

³³³ CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání*. Praha: C.H. Beck, 2023, s. 185.

³³⁴ § 58 odst. 4 autorského zákona.

³³⁵ § 58 odst. 2 autorského zákona.

10 Generování

Teprve nyní, když byl model integrován do AI systému, může koncový uživatel svým příkazem interagovat s generativní umělou inteligencí. Uživatel **svým příkazem specifikuje svůj požadavek a AI vygeneruje nový text nebo obrázek tím, že jedinečným způsobem kombinuje statistické vzorce, které se naučila z trénovacího datasetu.**³³⁶

10.1 Autorské dílo vytvořené s AI systémem

Autorství je *de lege lata* výsadou lidí (fyzických osob), nikoli umělé inteligence. AI systém může být předmětem autorskoprávní ochrany³³⁷, ale nemůže být nikdy autorem. Zda koncový uživatel využívající AI systém jako nástroj může být označen za autora vygenerovaného výstupu bude vždy záviset na konkrétních okolnostech případu.³³⁸

Příkaz (angl. *prompt*) uživatele pomáhá nasměrovat výstup generativní umělé inteligence tak, aby byl specificky přizpůsoben jeho požadavkům. Některé příkazy, jako například „*pracující básník*“, jsou jednoduché a velmi obecné, zatímco jiné, jako například „*vyobrazení Jaroslava Seiferta, píšícího u kavárenského stolku v Praze s Karlovým mostem a Pražským hradem v pozadí, v kresebném stylu Josefa Lady*“, jsou podrobnější a konkrétnější. Z autorskoprávního hlediska však v takových případech nebude za autorské dílo považován výstup umělé inteligence ani žádný z výše uvedených příkazů uživatele.³³⁹ Samotný příkaz uživatele, který slouží jako podklad pro obrázek vytvořený generativní umělou inteligencí, je většinou pouhým námětem či eventuálně myšlenkou, které samy o sobě nejsou autorským dílem.³⁴⁰

³³⁶ Viz kapitola 2.1. Strojové učení.

³³⁷ Viz kapitola 9. AI systém.

³³⁸ HOLCOVÁ, I. *MODELÝ UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO*. DOBŘIČOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLŮVÝCH [online], 2023, s. 30.

³³⁹ Srov. LEE, K., et al. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024, s. 49–50.

³⁴⁰ Bod 12 rozsudku Městského soudu v Praze ze dne 11. října 2023 pod č.j. 10 C 13/2023-16.

Výše uvedené však není obecným pravidlem. Pokud například uživatel výstup použije jako základ pro další tvůrčí činnost, tedy vygenerovaný text nebo obrázek tvůrčím způsobem obohatí, pak by výsledek takové činnosti mohl být autorským dílem.³⁴¹ Ostatně umělecké nebo vědecké výtvořiny nevznikají z ničeho. Každý nový počín je ve skutečnosti evolucí předchozích myšlenek a výtvořů, na které tvůrčí navazují a které dále rozvíjejí.³⁴²

Vezměme si například grafika, který si pomocí generátoru obrázků vytvoří hrubý náčrt loga na základě podrobného příkazu. Grafik následně vygenerovaný obrázek obohatí tím, že přidá jedinečné prvky, barvy a typografii, a vytvoří tak konečný návrh loga. V tomto případě grafik použil AI systém jako součást svého tvůrčího procesu a výsledné logo, za předpokladu splnění pojmových znaků autorského díla, bude autorským dílem, přičemž autorem bude grafik.

Ne vždy ale budou takové případy takto černobílé. Představme si situaci, kdy bude generativní umělá inteligence využita jako nástroj k překladu knihy do jiného jazyka. Obecně platí, že překlad děl do jiných jazyků, jako forma odvozené tvůrčí činnosti, požívá autorskoprávní ochrany.³⁴³ Do doby vzniku generativních AI systémů (např. ChatGPT) také platilo, že překlady vytvořené jen pomocí elektronických překladačů (např. Google Translate) nejsou autorskými díly, neboť se jedná o překlad mechanický a rutinní.³⁴⁴ Takové překlady z pochopitelných důvodů nepožívají ochrany, neboť činnost uživatele (člověka) je omezena na překopírování původního textu do překladače.

Na rozdíl od běžných elektronických překladačů může uživatel pomocí svých příkazů výrazně ovlivnit podobu překladu při použití generativního AI systému. Práce překladatele využívajícího pomoci AI systému může vypadat tak, že si při čtení knihy jednotlivé nuance a konkrétní jazykové formulace³⁴⁵, které ho při čtení napadají, rovnou zapisuje do příkazu. Po dočtení knihy pak příkaz odešle

³⁴¹ HOLCOVÁ, I. *MODELÝ UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO*. DOBŘICHOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLŮVÝCH [online], 2023, s. 32.

³⁴² KOUKAL, P., et al. Úvod do práv k duševnímu vlastnictví [online]. Brno: Právnická fakulta Masarykovy univerzity, 2020, s. 4.

³⁴³ § 2 odst. 4 autorského zákona.

³⁴⁴ TELEČ, I., TŮMA, P. *Autorský zákon: komentář*. 2. vydání. Praha: C.H. Beck, 2019, s. 52.

³⁴⁵ Rozsudek Nejvyššího soudu ČR sp. zn. 27 Cdo 2023/2019, ze dne 24. 3. 2021.

a generativní umělá inteligence podle jím definovaného postupu knihu přeloží. V závislosti na vlivu a spojitosti příkazu s konečným překladem může takto přeložená kniha být autorským dílem.

Používání nástrojů a jiných pomůcek, jako jsou slovníky³⁴⁶, automaticky nevylučují možnost vytvoření autorského díla překladatelem. Před nástupem generativní umělé inteligence (a elektronických překladačů) byla práce se slovníky pouze namáhavější a náročnější. Překladatel musel projít celou knihu, aby dosadil a upravil konkrétní hovorový výraz, který si určil. S generativní umělou inteligencí však překladateli stačí konkrétní hovorový znak³⁴⁷ definovat pouze jednou a generativní AI systém jej použije v celém procesu překladu. Tvůrčí činnost „nelze zaměňovat za namáhavost anebo těžkost práce“.³⁴⁸

Určení toho, zda konkrétní výstup generování bude požívat autorskoprávní ochrany, bude vždy záviset na okolnostech každého případu. Je tak v nejlepším zájmu tvůrců, kteří používají generativní umělou inteligenci jako nástroj, aby si své příkazy a výstupy uchovávali. V případě soudního sporu o určení autorství leží důkazní břemeno na žalobci, který musí prokázat, že AI systém vygeneroval výstup na základě jeho příkazů a že existuje přímá spojitost, jakési nepřetržité pouto³⁴⁹, mezi takovým výstupem a tvůrčí účastí člověka, která byla projevena příkazem.

10.2 Odpovědnost za zásah do autorského práva

Příkaz (*prompt*) uživatele bude obvykle pouhým námětem k vygenerování výstupu, který není chráněn autorským právem (např. „zformalizuj můj e-mail níže“). Uživatelé zároveň mají možnost zadávat příkazy, kterými mohou dovést AI systém k vygenerování autorskoprávně chráněného díla (např. „první knihu Harryho Pottera od J. K. Rowlingové“). Kdy může výstup generativní umělé inteligence představovat zásah do autorského práva a kdo je za jeho porušení odpovědný?

³⁴⁶ Srov. HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 84.

³⁴⁷ TELEČ, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019, s. 52

³⁴⁸ Tamtéž, s. 7.

³⁴⁹ DANG, M. *Právní aspekty umělé inteligence*. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020, s. 54.

V české³⁵⁰ i EU³⁵¹ rozhodovací praxi již dlouho platí, že k zásahu do autorského práva dochází i při užití malé části cizího autorského díla, a to zejména tehdy, kdy lze tuto část považovat za notoricky známou. Vygenerováním cizího autorského díla (např. knihy Harryho Pottera) je tak z hlediska autorského práva bezpochyby zásahem do práva na rozmnožování, příp. práva na sdělování díla veřejnosti.³⁵²

K poněkud zvláštní situaci může dojít v případě, kdy AI systém vygeneroval obrázek či text, který se sice podobá existujícímu autorskému dílu, ale toto dílo nebylo užito při trénování modelu.³⁵³ I takový výstup však bude považován za plagiát³⁵⁴.

Odpovědnost za zásah do autorského práva ponese některý ze subjektů zapojených do procesu vývoje, tréninku, distribuce nebo užívání generativní umělé inteligence. Mezi tyto subjekty patří poskytovatelé AI (např. IT společnosti, vývojáři, výrobci hardwaru a softwaru, distributoři nebo provozovatelé) i koncoví uživatelé. Navíc každá z osob zapojených do tohoto řetězce může nést pouze určitý díl odpovědnosti a různé typy odpovědností se mohou překrývat.³⁵⁵

Určení odpovědného subjektu bude vždy záviset na konkrétních okolnostech každého případu. Pokud si například koncový uživatel prostřednictvím svých příkazů nechá vygenerovat cizí autorské dílo (bez patřičné licence), je třeba na toto jednání pohlížet obdobně jako na nelegální stahování z internetu. Otázku, zda je stahování autorského obsahu pro osobní účely legální, zkoumal v minulosti i Nejvyšší soud, který rozhodl³⁵⁶, že pořízení kopií autorského díla pro osobní potřebu nepředstavuje neoprávněný zásah do autorských práv. Tento výrok je však dnes již překonaný judikaturou³⁵⁷ Soudního dvora EU, podle kterého nelegální

³⁵⁰ Srov. rozsudek Nejvyššího soud ČR ze dne 30. 4. 2007 pod sp. zn. 30 Cdo 739/2007.

³⁵¹ Srov. *Pelham GmbH a další v. Ralf Hütter a Florian Schneider-Esleben*. Rozsudek Soudního dvora (velkého senátu) ze dne 29. července 2019 ve věci C-476/17.

³⁵² NORDEMANN, J. B. EU law: *Generative AI, copyright infringements and liability – My guess for a hot topic in 2024* [online]. Kluwer Copyright Blog, 2024.

³⁵³ Tamtéž.

³⁵⁴ HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019, s. 56.

³⁵⁵ HOLCOVÁ, I. *MODELÝ UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO*. DOBŘICHOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLŮVÝCH [online], 2023, s. 36.

³⁵⁶ Usnesení Nejvyššího soudu ze dne 25. 3. 2009 pod sp. zn. 5 Tdo 234/2009.

³⁵⁷ *ACI Adam BV a další v. Stichting de ThuisKopie, Stichting Onderhandeligen ThuisKopie vergoeding*. Rozsudek Soudního dvora (čtvrtého senátu) ze dne 10. dubna 2014 ve věci C-435/12

pořízení kopií, a to i pro osobní potřebu, nelze zahrnout pod výjimku, která by takové jednání umožňovala. Zdůvodnil to tím, že stahování kopií z nelegálních zdrojů nesplňuje třístupňový test především proto, že autoři při takovém stahování nedostávají náhradní odměnu, což nepřiměřeně zasahuje do oprávněných zájmů autora. Podle autorského zákona by se užíváním vygenerovaného cizího autorského koncový uživatel dopustil přestupku³⁵⁸, za který mu může být uložena pokuta až do výše 150 000 Kč³⁵⁹. V teoretické rovině může uživateli hrozit i trestní postih³⁶⁰, pokud by zásah dosáhl „nikoli nepatrné“ intenzity. V praxi však příslušné orgány věnují minimální pozornost stahování obsahu z internetu, natož generování obsahu prostřednictvím AI systémů ve vlastním rozhraní.³⁶¹

Ve výše uvedeném případě nelze vyloučit ani odpovědnost poskytovatele AI. Autor, do jehož práva bylo neoprávněně zasaženo (nebo hrozí neoprávněný zásah), se může domáhat zákazu poskytování služby, kterou využívají třetí osoby k porušování nebo ohrožování práva autora³⁶². Dle současné judikatury³⁶³ Soudního dvora EU mají poskytovatelé tzv. hostingových služeb (webových úložišť) povinnost zdržet se poskytování této služby, pokud ji využívají třetí osoby k porušování nebo ohrožování práva duševního vlastnictví. Přičemž tato povinnost se může vztahovat nejen k odstranění závadného obsahu, nýbrž i k zajištění, že k takovému porušení opětovně nedojde.³⁶⁴ V souladu s touto judikaturou rozhodují i české³⁶⁵ soudy. Obdobně by se mohlo *mutatis mutandis* postupovat i v případech určení odpovědnosti poskytovatelů AI. Ostatně již nyní by měli poskytovatelé AI

³⁵⁸ § 105a odst. 1 písm. a) autorského zákona.

³⁵⁹ § 105a odst. 2 autorského zákona.

³⁶⁰ § 270 zákona č. 40/2009 Sb., trestní zákoník.

³⁶¹ PINDUR, V. *Nelegální stahování autorských děl z pohledu trestního práva* [online]. epravo.cz, 2022.

³⁶² § 40 odst. 1 písm. f) autorského zákona.

³⁶³ Frank Peterson v. Google LLC, YouTube Inc., YouTube LLC, Google Germany GmbH a Elsevier Inc. v. Cyando AG. Rozsudek Soudního dvora (velkého senátu) ze dne 22. června 2021 ve spojených věcech C-682/18 a C-683/18.

³⁶⁴ Usnesení Ústavního soudu ze dne 18. 7. 2023 pod sp. zn. III. ÚS 3077/22.

³⁶⁵ Srov. rozsudek Nejvyššího soudu ze dne 8. 6. 2022 pod sp. zn. 23 Cdo 1840/2021.

zabraňovat systematickému porušování autorského práva již z pouhé obecné prevenční povinnosti^{366, 367}.

³⁶⁶ § 2900 zákona č. 89/2012 Sb., občanského zákoníku a srov. bod 6 a 162 rozsudku Nejvyššího soudu ČR ze dne 30. 4. 2007 pod sp. zn. 30 Cdo 739/2007.

³⁶⁷ Srov. NORDEMANN, J. B. EU law: Generative AI, copyright infringements and liability – My guess for a hot topic in 2024 [online]. Kluwer Copyright Blog, 2024 a NOVELLI, C., CASOLARI, F., HACKER, P., SPEDICATO, G., FLORIDI, L. Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity [online]. 2024, s. 19.

11 Závěr

Rigorózní práce zpracovává aktuální problematiku právních aspektů generativní umělé inteligence, a to zejména z pohledu autorského práva. Práce je pro přehlednost členěna na dvě části, na obecnou a zvláštní část, přičemž celkově obsahuje deset věcných kapitol.

V úvodu byl představen cíl rigorózní práce a její základní struktura. Po úvodu následovala obecná část, která je rozdělena do kapitol „2. Úvod do umělé inteligence“ a „3. Právní východiska generativní umělé inteligence“. Tato část vysvětluje klíčové pojmy a koncepty spojené s umělou inteligencí, přičemž se zaměřuje na srozumitelný výklad technických principů, na kterých generativní umělá inteligence funguje. Dále analyzuje relevantní právní úpravy týkající se dané problematiky, s důrazem na autorské právo. Zvláštní část práce se dělí do kapitol „4. Výsledky tvůrčí činnosti (díla)“, „5. Data“, „6. Dataset“, „7. Trénování modelu“, „8. Dolad'ování“, „9. AI systém“ a „10. Generování“. Tato část sleduje životní cyklus generativní umělé inteligence od shromažďování a zpracování trénovacích dat až po fázi generování obrázků a textů. V každé fázi jsou identifikována právně relevantní jednání subjektů a jejich potenciální právní důsledky, se zvláštním zaměřením na autorské právo.

Ve druhé kapitole, nazvané „Úvod do umělé inteligence“, která následuje po úvodu, jsem vysvětlil základní pojmy jako *umělá inteligence* či *strojové učení*. V rámci této kapitoly jsem se pokusil srozumitelně vysvětlit technologické metody a koncepty, na jejichž základě dokáže současná generativní AI produkovat text a obrázky. Docházím k závěru, že současná generativní AI může fungovat jen díky tomu, že při jejím tréninku byla užita již existující (autorská) díla. Zároveň však upozorňuji, že mechanismus strojového učení nespočívá v prostém kopírování a ukládání tréninkových dat. Dále v kapitole představuji koncept „životního cyklu“ generativní umělé inteligence, který tvoří základní strukturu celé práce. Životní cyklus generativní umělé inteligence si pro účely této práce (a právní analýzy) je možné představit jako sedm stádií, a to: (i) *výsledky tvůrčí činnosti (díla)*, (ii) *data*, (iii) *dataset*, (iv) *trénování modelu*, (v) *dolad'ování*, (vi) *AI systém* a (vii) *generování*. Problematika právních aspektů generativní umělé inteligence je velmi

komplikovaná, protože zahrnuje i různé subjekty, které jsou v jednotlivých fázích vývoje zapojeny v odlišné míře. Každá osoba v tomto řetězci může za své jednání nést jen určitý díl odpovědnosti, přičemž různé typy odpovědností se mohou překrývat. Představuji zde tři hlavní subjekty, s nimiž se v práci dále pracuje: *autory, poskytovatele AI a uživatele.*

Třetí kapitola s názvem „Právní východiska generativní umělé inteligence“ začíná výkladem k českému autorskému právu. Zvláštní pozornost jsem věnoval výjimkám a omezením autorského práva. V kontextu životního cyklu generativní umělé inteligence jsem identifikoval relevantní zákonné licence, v jejichž rámci lze dílo užít bez svolení autora k *rozmnožování díla pro účely automatizované analýzy textů nebo dat pro jiné účely než pro účely vědeckého výzkumu, včetně komerčních* (§ 39c) a *pastiši* (§ 38g). V další části kapitoly se věnuji balíčku předpisů, které nedávno přijala (či přijme) EU za účelem harmonizace vývoje, implementace a využívání umělé inteligence v členských státech. Pro generativní AI je klíčový zejména Akt o AI, který byl na poslední chvíli doplněn o ustanovení reagující na rizika spojená s tímto typem umělé inteligence. Mezi tato opatření patří například zavedení povinné transparentnosti pro systémy generující *deepfakes* nebo *fake news*. Domnívám se, že zatímco některé povinnosti, které nařízení poskytovatelům AI ukládá, jsou logické a snadno splnitelné, u jiných naopak budou mít poskytovatelé AI potíže dostat požadavku. Jako zásadní dále považuji článek 53, který mj. odkazuje na směrnici o autorském právu na jednotném digitálním trhu, která umožňuje rozmnožování a extrakci autorských děl pro účely vytěžování textů a dat, k nimž má subjekt zákonný přístup. Dle mého názoru přímým odkazem na tuto směrnici tak zákonodárce jasně připouští, že výjimka pro vytěžování textů a dat se vztahuje i na trénink modelů generativní umělé inteligence.

Ve čtvrté kapitole s názvem „Výsledky tvůrčí činnosti (díla)“ se zabývám otázkou dlouhodobého využívání autorských děl pro trénink modelů umělé inteligence. Zjišťuji, že přestože Autorský zákon, Akt o AI a další předpisy stanovují podmínky pro použití těchto děl, chybí v nich právní a ekonomické nástroje, které by dlouhodobě podporovaly další tvůrčí činnost lidí. Jako zásadní nedostatek spatřuji v tom, že autoři mohou vyloučit užití jejich děl, a tedy zabránit tomu, aby bylo jeho dílo zařazeno do tréninkového datasetu pro umělou inteligenci, ale pouze s účinky

pro futuro. Pro existující (již natrénované) modely tak tato výhrada neplatí. Přicházím tak s návrhem *de lege ferenda* na zavedení práva na náhradní odměnu v souvislosti s rozmnožováním díla pro účely automatizované analýzy textů nebo dat po vzoru existujícího institutu náhradní odměny v souvislosti s rozmnožováním díla pro osobní potřebu. Nadto si uvědomuji, že i kdyby byl systém náhradní odměny zaveden, tak není reálné požadovat po poskytovateli AI, aby jednal o licenci k jejich užití jednotlivě s každým autorem. Klíčovým mechanismem je tedy s náhradní odměnou zavést i povinnou kolektivní správu této odměny.

Pátá kapitola s názvem „Data“ nejdříve začíná rozbořením toho, jakým způsobem se získávají data k tréninku modelu umělé inteligence, a jak je tento proces regulován. Právní rámec k otázce sběru a analýzy autorskoprávně chráněného obsahu řeší na unijní úrovni článek 4 (resp. článek 3) směrnice o autorském právu na jednotném digitálním trhu, který byl do Autorského zákona implementován do § 39c (resp. § 39d). Jsem toho názoru, že tato zákonná licence umožňuje poskytovatelům AI užít díla i pro účely tréninku modelu. Nicméně upozorňuji, že ačkoli autoři mohou předem vyloučit vytěžování svých děl, není v současnosti jasné, v jaké formě by tato výhrada měla být provedena. Jako řešení se nabízí standardizovaný protokol *robots.txt*, kterým provozovatelé webových stránek sdělují automatizovaným softwarům, které části webu jsou přístupné. Poukazuji také na paradoxní situaci, kdy může být pro poskytovatele AI obtížnější legálně vytěžovat obsah, který není chráněn autorským právem, neboť provozovatelé webových stránek mohou stanovit smluvní omezení jeho používání.

V šesté kapitole s názvem „Dataset“ se zabývám právní povahou datasetů, které poskytovatelé AI považují za cenná aktiva a obchodní tajemství. Docházím k závěru, že dataset lze považovat za *databázi* ve smyslu autorského práva. V praxi poskytovatelé AI často zadávají shromažďování a anotaci datasetů subdodavatelům. Upozorňuji tak na možný souběh právní ochrany, protože i když subdodavatel poskytuje dataset pod smluvní licenci, není zaručeno, že jednotlivá díla v datasetu byla řádně licencována. Poskytovatel AI, i když využívá licenci od subdodavatele, stále může čelit právnímu riziku spojenému se zásahem do autorských práv k obsahu datasetu.

V sedmé kapitole s názvem „Trénování modelu“ se zabývám problematikou plagiátorského chování generativní umělé inteligence, které je nežádoucí jak z právního, tak z technického hlediska. Zákonná licence v § 39c (resp. § 39d) umožňuje poskytovatelům AI užít díla i pro účely tréninku modelu, nikoliv však užitá tréninková data nelze dále rozmnožovat, sdělovat ani jinak navenek šířit či reprodukovat. Dále poukazuji na komplikovanost správného určení, kdy je výstup umělé inteligence považován za rozmnožování (resp. sdělování) a tedy je porušením autorského práva, a kdy se bude jednat pouze o napodobení uměleckého stylu. Tato otázka byla částečně řešena v judikátu Pelham, který se zabývá mírou shody s původním dílem. U generativní umělé inteligence bude posuzování náročnější, neboť výstupy málokdy identické, ale spíše autorským dílům pouze podobné. Soudy budou muset tak brát v úvahu, zda konkrétní výstup nebude pouze napodobením uměleckého stylu, námětu či myšlenky, které nespádají do působnosti autorského práva, a tudíž nedochází k jeho porušování. Autorský zákon obsahuje také zákonnou licenci k vytvoření *pastiše*, který umožňuje ještě „vyšší“ úroveň zásahu do autorského díla, než je tomu tak u napodobení uměleckého stylu. Tento pojem však dosud nemá zákonnou definici, a proto bude nutné počkat na závazný výklad ze strany Soudního dvora EU.

Osmá kapitola s názvem „Doladování“ je zaměřena na předpojaté chování AI a problematiku *deepfake* a *fake news*. Užívání AI systémů, které by vedlo k diskriminaci, může být například ve věcech práva na zaměstnání nebo rovnému přístupu ke vzdělání v přímém rozporu s Antidiskriminačním zákonem. Je-li generativní umělá inteligence také užívána za účelem automatizovaného profilování nebo individuálního rozhodování, je potřeba tak činit pouze v souladu s nařízením GDPR. Na úrovni EU přistoupila k regulaci *deepfake* a *fake news* zavedením povinnosti, aby takové výstupy byly označené tak, aby bylo zpětně zjištělné jako umělé vytvořené. Takový požadavek je však dnes technicky neproveditelný.

V deváté kapitole s názvem „AI systém“ řeším otázku toho, zda AI systémy, jsou z autorskoprávního hlediska stále považovány za počítačový program. Za tímto účelem jsem zkoumal jednotlivé složky, ze kterých se klasický generativní AI systém skládá. Na základě této analýzy dospívám k závěru, že vzhledem časté absenci

znaku *původnosti*, samotný soubor sloužící ke spuštění AI systému obvykle nespĺňuje podmínky pro ochranu autorským právem. Ostatně ani poskytovatelé AI zpravidla nevynakládají zvláštní úsilí na jeho ochranu a jsou často ochotni zveřejnit celou technickou architekturu. Zároveň doplňují, že nechráněné prvky je totiž v případě nutnosti možné chránit i jinými způsoby, od smluvního závazkového práva až po ochranu práva patentového či nekalé soutěže. V závěru kapitoly se věnuji problematice vývoje AI systémů v rámci pracovněprávních vztahů. Upozorňuji, že IT společnosti obvykle nesou smluvní odpovědnost za chybný výstup nebo chování AI systému, přičemž regresní nárok vůči zaměstnanci může zaměstnavatel uplatnit pouze v případě, že zaměstnanec způsobil škodu. Příčinnou souvislost mezi škodou a jednáním zaměstnance je však nutné prokázat, a navíc náhrada škody je limitovaná čtyřapůlnásobkem průměrného výdělku, pokud zaměstnanec nejednal úmyslně.

V závěrečné desáté kapitole jsem nejdříve věnoval autorskoprávní ochraně k obsahu vygenerované umělé inteligencí. I v případě generativní AI docházím ke stejnému závěru, jako ve své diplomové práci a to, že vznik ochrany (resp. autorského díla) není vyloučen, ale vždy závisí na míře lidské participace k vygenerovanému výstupu. Druhá část kapitoly obsahuje výklad o odpovědnosti za vygenerování cizího autorského díla. Domnívám se, že vygenerováním cizího díla (bez patřičné licence) bude vždy zásahem do autorského práva, a to bez ohledu na to, zda dané dílo bylo či nebylo užito k tréninku modelu. Určení odpovědného subjektu, bude vždy záviset na konkrétních okolnostech každého případu.

Právní aspekty spojené s umělou inteligencí a zejm. autorským právem jsou velmi komplikované. Současná regulace sice nabízí řešení na otázky, jako je užívání cizích děl pro trénink AI, určování autorství k výstupům AI nebo ochrana proti deepfakes a fake news. Některé z nich je však stále třeba revidovat. Větší obavu vzbuzuje dlouhodobý sociální a ekonomický dopad dalšího vývoje generativní AI. Současná regulace se totiž nezabývá tím, jak může umělá inteligence změnit naše chápání tradičních konceptů, jako jsou autorství, tvůrčí činnost a duševní vlastnictví obecně. Stejně jako v minulosti v případě rozvoje reprodukční techniky (např. tiskáren, DVD, paměťových karet) i AI umožní masově vytvářet a rozmnožovat obsah pro osobní potřebu. Důsledkem tohoto přirozeně dojde ke snížení poptávky

po lidmi vytvořených autorských dílech, a tedy i ke snížení celkové odměny pro tyto lidi. Není překvapením, že si autoři svá díla chrání a další vývoj generativní umělé inteligence spíše nevítají.

Seznam použitých zdrojů

Seznam použité literatury

1. BROWN, T.B., MANN, B., RYDER, N., SUBBIAH, M., KAPLAN, J., DHARIWAL, P., NEELAKANTAN, A., SHYAM, P., SASTRY, G., ASKELL, A., AGARWAL, S., HERBERT-VOSS, A., KRUEGER, G., HENIGHAN, T., CHILD, R., RAMESH, A., ZIEGLER, D.M., WU, J., WINTER, C., HESSE, C., CHEN, M., SIGLER, E., LITWIN, M., GRAY, S., CHESS, B., CLARK, J., BERNER, C., MCCANDLISH, S., RADFORD, A., SUTSKEVER, I., AMODEI, D. *Language Models are Few-Shot Learners* [online]. OpenAI, Cornell University Library, 2020. Dostupné z <https://arxiv.org/pdf/2005.14165>.
2. CARLINI, N., HAYES, J., NASR, M., JAGIELSKI, M., SEHWAG, V., TRAMER, F., BALLE, B., IPPOLITO, D., WALLACE, E. *Extracting training data from diffusion models* [online]. Cornell University Library, 2023. Dostupné z <https://arxiv.org/pdf/2301.13188>.
3. CARLINI, N., IPPOLITO, D., JAGIELSKI, M., LEE, K., TRAMER, F., ZHANG, C. *Quantifying Memorization Across Neural Language Models* [online]. Cornell University Library, 2023. Dostupné z: <https://arxiv.org/pdf/2202.07646>.
4. CARLINI, N., TRAMER, F., WALLACE, E., JAGIELSKI, M., HERBERTVOSS, A., LEE, K., ROBERTS, A., BROWN, T. B., SONG, D., ERLINGSSON, U., OPREA, A., RAFFEL C. *Extracting training data from large language models*. USENIX Security Symposium, volume 6 [online], 2021. Dostupné z <https://www.usenix.org/system/files/sec21-carlini-extracting.pdf>. ISBN: 978-1-939133-24-3.
5. CHALOUPKOVÁ, H., HOLÝ, P. *Autorský zákon: komentář. 6. vydání*. Praha: C.H. Beck, 2023. ISBN 978-80-7400-944-0.
6. COOPER, A.F., GRIMMELMANN, J. *The Files are in the Computer: Copyright, Memorization, and Generative AI* [online]. Cornell University Library, 2024. Dostupné z <https://arxiv.org/pdf/2404.12590>.
7. DANG, M. *Právní aspekty umělé inteligence* [online]. Praha: Univerzita Karlova, Právnická fakulta, Ústav práva autorského, práv průmyslových a práva soutěžního, 2020. Dostupné z <https://dspace.cuni.cz/bitstream/handle/20.500.11956/117650/120358328.pdf?sequence=1&isAllowed=y>.
8. DREXL, J., HILTY, R., BENEKE, F., DESAUNETTES-BARBERO, L., FINCK, M., GLOBOCNIK, J., GONZALEZ OTERO, B., HOFFMANN, J., HOLLANDER, L., KIM, D., RICHTER, H., SCHEUERER, S., SLOWINSKI, P. R., THONEMANN, J. *Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*. Mnichov: Max Planck Institute for Innovation & Competition Research Paper No. 19-13 [online], 2019. Dostupné z <http://dx.doi.org/10.2139/ssrn.3465577>.

9. EMANUILOV, I., MARGONI, T. *Forget me not: memorisation in generative sequence models trained on open source licensed code* [online]. 2024. Dostupné z: <https://doi.org/10.5281/zenodo.10635479>.
10. FELDMAN, V. *Does learning require memorization? A short tale about a long tail*. TOC 2020: Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 2020, s. 954-959. ISBN: 978-14-5036-979-4.
11. FIALOVÁ, E., MATEJKA, J. *Právo umělé inteligence jako problém efektivního propojení technologie a lidské práce*. Praha: Právní rozhledy, 6/2022, 2022, s. 200-206. ISSN 1210-6410.
12. GOODFELLOW, I., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDEFARLEY, D., OZAIR, S., COURVILLE, A., BENGIO, Y. *Generative Adversarial Nets*. Montréal: Advances in Neural Information Processing Systems 27 (NIPS 2014) [online], 2014. Dostupné z: https://papers.nips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afcf3-Paper.pdf.
13. HOLCOVÁ, I. *Autorský zákon a předpisy související (včetně mezinárodních smluv a evropských předpisů): komentář*. Praha: Wolters Kluwer, 2019. ISBN 978-80-7598-049-6.
14. HOLCOVÁ, I. *MODELÝ UMĚLÉ INTELIGENCE A AUTORSKÉ PRÁVO. DOBŘICHOVSKÝ, T., et al. AKTUÁLNÍ OTÁZKY PRÁVA AUTORSKÉHO A PRÁV PRŮMYSLOVÝCH* [online], 2023. Dostupné z <https://cld.bz/Z4wbB6y>. ISBN: 978-80-7630-040-8.
15. JANSÁ, L., OTEVŘEL, P., ŠTEVKO, M. *Softwarové právo. 3. aktualizované a rozšířené vydání*. Brno: Computer Press, 2018. ISBN: 978-80-251-4914-0.
16. KAPLAN, J., MCCANDLISH, S., HENIGHAN, T., BROWN, T. B., CHESSE, B., CHILD, R., GRAY, S., RADFORD, A., WU, J., AMODEI, D. *Scaling Laws for Neural Language Models*. *OpenAI* [online], 2020. Dostupné z <https://openai.com/index/scaling-laws-for-neural-language-models>.
17. KELLER, P., WARSO, Z. *Defining best practices for opting out of ML training*. *Open Future policy brief* [online], 2023. Dostupné z: https://openfuture.eu/wp-content/uploads/2023/09/Best-_practices_for_optout_ML_training.pdf.
18. KOUKAL, P., MYŠKA, M., PULLMANNOVÁ, H., VOJTÍŠKOVÁ, T., ZIBNER, J. *Úvod do práv k duševnímu vlastnictví* [online]. Brno: Právnícká fakulta Masarykovy univerzity, 2020. Dostupné z https://is.muni.cz/do/rect/el/estud/praf/2019podzim/dusevni_vlastnictvi/web/index.html.
19. LEE, K., COOPER, A. F., GRIMMELMANN, J. *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain* [online]. Cornell University Library, 2024. Dostupné z: <https://arxiv.org/pdf/2309.08133>.
20. LONGPRE, S., MAHARI, R., CHEN, A., OBENG-MARNU, N., SILEO, D., BRANNON, W., MUENNIGHOFF, N., KHAZAM, N., KABBARA, J., PERISETLA, K., WU, X., SHIPPOLE, E., BOLLACKER, K., WU, T., VILLA, L., PENTLAND, S., HOOKER, S. *The Data Provenance Initiative: A Large*

- Scale Audit of Dataset Licensing & Attribution in AI* [online]. Cornell University Library, 2023. Dostupné z <https://arxiv.org/pdf/2310.16787>.
21. MACNEIL, S., TRAN, A., HELLAS, A., KIM, J., SARSÅ, S., DENNY, P., BERNSTEIN, S., LEINONEN, J. *Experiences from Using Code Explanations Generated by Large Language Models in a Web Software Development* [online]. Cornell University Library, 2022. Dostupné z <https://arxiv.org/pdf/2211.02265>.
 22. MCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.E. *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* [online], 1955. Dostupné z <https://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>.
 23. MURPHY, K. P. *Probabilistic Machine Learning: An Introduction*. Massachusetts Institute of Technology, 2022, ISBN 978-02-6204-682-4.
 24. NOVELLI, C., CASOLARI, F., HACKER, P., SPEDICATO, G., FLORIDI, L. *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity* [online]. 2024. Dostupné z <http://dx.doi.org/10.2139/ssrn.4694565>.
 25. OUYANG, L., WU, J., JIANG, X., ALMEIDA, D., WAINWRIGHT, C. L., MISHKIN, P., ZHANG, C., AGARWAL, S., SLAMA, K., RAY, A., SCHULMAN, J., HILTON, J., KELTON, F., MILLER, L., SIMENS, M., ASKELL, A., WELINDER, P., CHRISTIANO, P., LEIKE, J., LOWE, R. *Training language models to follow instructions with human feedback* [online]. OpenAI, 2022. Dostupné z: https://cdn.openai.com/papers/Training_language_models_to_follow_instructions_with_human_feedback.pdf.
 26. REN, J., XU, H., HE, P., CUI, Y., ZENG, S., ZHANG, J., WEN, H., DING, J., LIU, H., CHANG, Y., TANG, J. *Copyright Protection in Generative AI: A Technical Perspective* [online]. Cornell University Library, 2024. Dostupné z: <https://arxiv.org/pdf/2402.02333>.
 27. RUIZ, N., LI, Y., JAMPANI, V., PRITCH, Y., RUBINSTEIN, M., ABERMAN, K. *Dreambooth: Fine-tuning text-to-image diffusion models for subject-driven generation* [online]. Cornell University Library, 2023. Dostupné z <https://arxiv.org/pdf/2208.12242>.
 28. RUSSEL, S., NORVIG, P. *Artificial Intelligence: A Modern Approach, Fourth Edition*. Upper Saddle River, NJ: Prentice Hall, 2020. ISBN: 978-12-9240-113-3.
 29. SENFTLEBEN, M., *Generative AI and Author Remuneration*. IIC - International Review of Intellectual Property and Competition Law, Volume 54 [online], 2023, s. 1535–1560. Dostupné z <https://doi.org/10.1007/s40319-023-01399-4>.
 30. SVOBODA, K., SMOLÍK, P., LEVÝ, J., DOLEŽÍLEK, J. a kol. *Občanský soudní řád. 3. vydání (2. aktualizace)*. Praha: C. H. Beck, 2023, ISBN: 978- 80-7400-828-3.

31. TELEC, I., TŮMA, P. *Autorský zákon: komentář. 2. vydání*. Praha: C.H. Beck, 2019. ISBN 978-80-7400-748-4.
32. TIRUMALA, K., MARKOSYAN, A. H., ZETTLEMOYER, L., AGHAJANYAN, A. *Memorization Without Overfitting: Analyzing the Training Dynamics of Large Language Models* [online]. Meta AI Research, 2022. Dostupné z https://proceedings.neurips.cc/paper_files/paper/2022/file/fa0509f4dab6807e2cb465715bf2d249-Paper-Conference.pdf.
33. TOUVRON, H., MARTIN, L., STONE, K., ALBERT, P., ALMAHAIRI, A., BABAEI, Y., BASHLYKOV, N., BATRA, S., BHARGAVA, P., BHOSALE, S., BIKEL, D., BLECHER, L., CANTON FERRER, C., CHEN, M., CUCURULL, G., ESIÖBU, D., FERNANDES, J., FU, J., FU, W., FULLER, B., GAO, C., GOSWAMI, V., GOYAL, N., HARTSHORN, A., HOSSEINI, S., HOU, R., INAN, H., KARDAS, M., KERKEZ, V., KHABSA, M., KLOUMANN, I., KORENEV, A., KOURA, P. S., LACHAUX, M.-A., LAVRIL, T., LEE, J., LISKOVICH, D., LU, Y., MAO, Y., MARTINET, X., MIHAYLOV, T., MISHRA, P., MOLYBOG, I., NIE, Y., POULTON, A., REIZENSTEIN, J., RUNGTA, R., SALADI, K., SCHELTEN, A., SILVA, R., SMITH, E. M., SUBRAMANIAN, R., TAN, X. E., TANG, B., TAYLOR, R., WILLIAMS, A., KUAN, J. X., XU, P., YAN, Z., ZAROV, I., ZHANG, Y., FAN, A., KAMBADUR, M., NARANG, S., RODRIGUEZ, A., STOJNIC, R., EDUNOV, S., SCIALOM, T. *Llama 2: Open Foundation and Fine-Tuned Chat Models* [online]. GenAI, Meta, 2023. Dostupné z: https://scontent.fprg5-1.fna.fbcdn.net/v/t39.2365-6/10000000_662098952474184_2584067087619170692_n.pdf?_nc_cat=105&ccb=1-7&_nc_sid=3c67a6&_nc_ohc=kdvSTf72sGwQ7kNvgHTuNPr&_nc_ht=scontent.fprg5-1.fna&oh=00_AYCVidNYdK1xPZI4jWEVGZCqTDWICyZpfRmH4aVDyNaxXA&oe=666DF1FF.
34. VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A.N., KAISER, Ł., POLOSUKHIN, I. *Attention is all you need*. *Advances in Neural Information Processing Systems* [online]. Cornell University Library, 2017. Dostupné z: <https://arxiv.org/pdf/1706.03762>.
35. VINCENT, P., LAROCHELLE, H., LAJOIE, I., BENGIO, Y., MANZAGOL, P.-A. *Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion*. *Journal of Machine Learning Research*, 11, [online], 2010. Dostupné z: <https://www.jmlr.org/papers/volume11/vincent10a/vincent10a.pdf>.
36. WAN, A., WALLACE, E., SHEN, S., KLEIN, D. *Poisoning Language Models During Instruction Tuning* [online]. Cornell University Library, 2023. Dostupné z <https://arxiv.org/abs/2305.00944>.

Seznam použitých právních předpisů

37. Listina základních práv Evropské unie
38. Návrh nařízení Evropského parlamentu a Rady, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci a mění nařízení (ES) č. 300/2008, (EU) č. 167/2013, (EU) č. 168/2013, (EU) 2018/858, (EU) 2018/1139 a (EU) 2019/2144 a směrnice 2014/90/EU, (EU) 2016/797 a (EU) 2020/1828 (akt o umělé inteligenci) [COM/2021/206 final] ve znění přijatém Radou EU dne 21. května 2024
39. Návrh směrnice Evropského parlamentu a Rady (EU) o přizpůsobení pravidel mimosmluvní občanskoprávní odpovědnosti umělé inteligenci (směrnice o odpovědnosti za umělou inteligenci) [COM(2022) 496 final] ze dne 28. září 2022
40. Návrh zákona, kterým se mění zákon č. 40/2009 Sb., trestní zákoník, ve znění pozdějších předpisů, zákon č. 141/1961 Sb., o trestním řízení soudním (trestní řád), ve znění pozdějších předpisů, a další související zákony
41. Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů)
42. 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 (nařízení o digitálních službách)
43. Nařízení Evropského parlamentu a Rady (EU) 2023/988 ze dne 10. května 2023 o obecné bezpečnosti výrobků, o změně nařízení Evropského parlamentu a Rady (EU) č. 1025/2012 a směrnice Evropského parlamentu a Rady (EU) 2020/1828 a o zrušení směrnice Evropského parlamentu a Rady 2001/95/ES a směrnice Rady 87/357/EHS
44. Rozhodnutí Komise ze dne 24. ledna 2024 o zřízení Evropského úřadu pro umělou inteligenci [C/2024/1459]
45. Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. března 1996 o právní ochraně databází
46. Směrnice Evropského parlamentu a Rady 2001/95/ES ze dne 3. prosince 2001 o obecné bezpečnosti výrobků
47. Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. 4. 2009 o právní ochraně počítačových programů (kodifikované znění)
48. Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu a o změně směrnic 96/9/ES a 2001/29/ES
49. Směrnice Rady ze dne 25. června 1987 o sblížení právních předpisů členských států týkajících se výrobků, jejichž skutečná povaha není rozpoznatelná a které proto ohrožují zdraví nebo bezpečnost spotřebitelů

50. Zákon č. 89/2012 Sb., občanský zákoník
51. Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon)
52. Zákon č. 198/2009 Sb. o rovném zacházení a o právních prostředcích ochrany před diskriminací a o změně některých zákonů (antidiskriminační zákon)
53. Zákon č. 40/2009 Sb., trestní zákoník
54. Zákon č. 262/2006 Sb., zákoník práce

Seznam použité judikatury

55. ACI Adam BV a další v. Stichting de Thuiskopie, Stichting Onderhandeligen Thuiskopie vergoeding. Rozsudek Soudního dvora (čtvrtého senátu) ze dne 10. dubna 2014 ve věci C-435/12
56. Bezpečnostní softwarová asociace – Svaz softwarové ochrany v. Ministerstvo kultury. Rozsudek Soudního dvora (třetího senátu) ze dne 22. prosince 2010 ve věci C-393/09
57. Eva-Maria Painer v. Standard VerlagsGmbH a další. Rozsudek Soudního dvora (třetího senátu) ze dne 1. prosince 2011 ve věci C-145/10
58. Frank Peterson v. Google LLC, YouTube Inc., YouTube LLC, Google Germany GmbH a Elsevier Inc. v. Cyando AG. Rozsudek Soudního dvora (velkého senátu) ze dne 22. června 2021 ve spojených věcech C-682/18 a C-683/18
59. Football Dataco Ltd a další v. Yahoo UK Limited a další. Rozsudek Soudního dvora (třetího senátu) ze dne 1. března 2012 ve věci C-604/10
60. Nederlands Uitgeversverbond a Groep Algemene Uitgevers v. Tom Kabinet Internet BV, Tom Kabinet Holding BV a Tom Kabinet Uitgeverij BV. Rozsudek Soudního dvora (velkého senátu) ze dne 19. prosince 2019 ve věci C-263/18
61. Pelham GmbH a další v. Ralf Hütter a Florian Schneider-Esleben. Rozsudek Soudního dvora (velkého senátu) ze dne 29. července 2019 ve věci C-476/17
62. S. Š. v. TAUBEL LEGAL. Rozsudek Městského soudu v Praze ze dne 11. října 2023 pod č.j. 10 C 13/2023-16
63. Rozsudek Nejvyššího soudu ze dne 8. 6. 2022 pod sp. zn. 23 Cdo 1840/2021
64. Rozsudek Nejvyššího soudu ČR ze dne 24. 3. 2021 pod sp. zn. 27 Cdo 2023/2019
65. Rozsudek Nejvyššího soud ČR ze dne 30. 4. 2007 pod sp. zn. 30 Cdo 739/2007
66. Ryanair Ltd v. PR Aviation BV. Rozsudek Soudního dvora (druhého senátu) ze dne 15. ledna 2015 ve věci C-30/14
67. SAS Institute Inc. v. World Programming Ltd. Rozsudek Soudního dvora (velkého senátu) ze dne 2. května 2012 ve věci C-406/10

68. Società Consortile Fonografici (SCF) v. Marco Del Corso. Rozsudek Soudního dvora (třetího senátu) ze dne 15. března 2012 ve věci C-135/10
69. UsedSoft GmbH v. Oracle International Corp. Rozsudek Soudního dvora (velkého senátu) ze dne 3. července 2012 ve věci C-128/11
70. Usnesení Nejvyššího soudu ze dne 25. 3. 2009 pod sp. zn. 5 Tdo 234/2009
71. Usnesení Ústavního soudu ze dne 18. 7. 2023 pod sp. zn. III. ÚS 3077/22

Seznam ostatních zdrojů

72. AI HLEG. *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment* [online], 2019. Dostupné z: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342.
73. AI HLEG. *Sectoral Considerations on Policy and Investment Recommendations for Trustworthy AI* [online], 2019. Dostupné z: https://futurium.ec.europa.eu/sites/default/files/2020-07/Sectoral%20Considerations%20On%20The%20Policy%20And%20Investment%20Recommendations%20For%20Trustworthy%20Artificial%20Intelligence_0.pdf.
74. AARONSON S. *What Is AI Safety?* [online], 2022. Dostupné z: <https://scottaaronson.blog/?m=202211>.
75. BUBECK, S., CHANDRASEKARAN, V., ELDAN, R., GEHRKE, J., HORVITZ, E., KAMAR, E., LEE, P., LEE, Y.T., LI, Y., LUNDBERG, S., NORI, H., PALANGI, H., RIBEIRO, M.T., ZHANG, Y. *Sparks of Artificial General Intelligence: Early experiments with GPT-4* [online]. Microsoft Research, 2023. Dostupné z <https://www.microsoft.com/en-us/research/publication/sparks-of-artificial-general-intelligence-early-experiments-with-gpt-4/>.
76. ČECH, P. *Umělá inteligence jako třetí subjekt práva?* [online]. Právní prostor, 2019. Dostupné z: <https://www.pravniprostor.cz/clanky/pravo-it/umela-inteligence-jako-treti-subjekt-prava>.
77. Důvodová zpráva zákona č. 429/2022 Sb., kterým se mění zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, a další související zákony.
78. Důvodová zpráva návrhu zákona, kterým se mění zákon č. 40/2009 Sb., trestní zákoník, ve znění pozdějších předpisů, zákon č. 141/1961 Sb., o trestním řízení soudním (trestní řád), ve znění pozdějších předpisů, a další související zákon, pod č.j. Odboru vládní agendy ÚV ČR 743/24.
79. EMANUILOV, I., MARGONI, T. *Memorisation in generative models and EU copyright law: an interdisciplinary view* [online]. Lovain: KU Leuven Centre for IT & IP Law, 2024. Dostupné z <https://copyrightblog.kluweriplaw.com/2024/03/26/memorisation-in-generative-models-and-eu-copyright-law-an-interdisciplinary-view/>.

80. EVROPSKÁ KOMISE. *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic And Social Committee And the Committee of the Regions - Coordinated Plan on Artificial Intelligence* ze dne 7. 12. 2018. [COM(2018) 795 final].
81. EVROPSKÁ KOMISE. *Special meeting of the European Council (1 and 2 October 2020) – Conclusions*. EUCO 13/20 [online], 2020. [CO EUR 10 CONCL 6]. Dostupné z: <https://www.consilium.europa.eu/media/45910/021020-euco-final-conclusions.pdf>.
82. EVROPSKÁ KOMISE. *Tisková zpráva - Artificial intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards ze dne 9. března 2018* [online], 2018. [IP/18/1381]. Dostupné z: https://ec.europa.eu/commission/presscorner/detail/en/ip_18_1381.
83. Fine-tuning Common Use Cases. OpenAI API [online]. Dostupné z: <https://platform.openai.com/docs/guides/fine-tuning/common-use-cases>.
84. GETTY IMAGES (US), INC. v. STABILITY AI, LTD. a STABILITY AI, INC. Žaloba vedená u THE UNITED STATES DISTRICT COURT FOR THE DISTRICT OF DELAWARE pod sp. zn. 1:23-cv-00135-GBW. Dostupné z <https://storage.courtlistener.com/recap/gov.uscourts.ded.81407/gov.uscourts.ded.81407.1.0.pdf>.
85. GETTY IMAGES (US) INC, GETTY IMAGES INTERNATIONAL UC, GETTY IMAGES (UK) LIMITED, GETTY IMAGES DEVCO UK LIMITED, ISTOCKPHOTO LP, THOMAS M BARWICK INC v. STABILITY AI LTD. Vyjádření žalované vedené u THE HIGH COURT OF JUSTICE BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES TECHNOLOGY AND CONSTRUCTION COURT (KBD) pod sp. zn. IL-2023-000007. Dostupné z <https://drive.google.com/file/d/1HIS9jM6iMRY1TDmaz8LkL30phoF8hXjg/view>.
86. GPTBot. OpenAI API. Dostupné z <https://platform.openai.com/docs/gptbot>.
87. HAVRÁNEK B., BĚLIČ J., HELCL M., JEDLIČKA A., ČERVENÁ V., FILIPEC J., HAVLOVÁ F., CHURAVÝ M., JANSKÝ L., KOZLOVÁ K., KROUPOVÁ L., MACHAČ J., MAREŠOVÁ H., MEJSTRÍK V., MICHÁLEK E., PAPÍRNÍKOVÁ B., POKORNÁ E., POŠTOLKOVÁ B., ROUDNÝ M., SOCHOVÁ Z., SVOZILOVÁ N., VODRÁŽKOVÁ E., ZIMA J. *Slovník spisovného jazyka českého* [online]. Ústav pro jazyk český, v. v. i., 2011. Dostupné z <https://ssjc.ujc.cas.cz/>.
88. Overview of Google crawlers and fetchers (user agents). Dostupné z <https://developers.google.com/search/docs/crawling-indexing/overview-google-crawlers>.
89. Inference code for Llama models [online]. Dostupné z: <https://github.com/meta-llama/llama>.
90. Intergram. *Co je to kolektivní správa práv? A k čemu slouží?* [online]. Dostupné z <https://www.intergram.cz/pro-verejnost/co-je-to-kolektivni-sprava-prav/>.

91. Intellectual Property Crime Project at the European Union Agency. *Generative Artificial Intelligence: the impact on intellectual property crimes* [online], 2023. Dostupné z <https://www.eurojust.europa.eu/sites/default/files/assets/generative-ai-impact-to-ip-crimes.pdf>.
92. KELLER, P. *A first look at the copyright relevant parts in the final AI Act compromise* [online]. Amsterdam: Institute for Information Law (IViR), 2023. Dostupné z <https://copyrightblog.kluweriplaw.com/2023/12/11/a-first-look-at-the-copyright-relevant-parts-in-the-final-ai-act-compromise/>.
93. KELLER, P. *Protecting creatives or impeding progress? Machine learning and the EU copyright framework* [online]. Amsterdam: Institute for Information Law (IViR), 2023. Dostupné z https://copyrightblog.kluweriplaw.com/2023/02/20/protecting-creatives-or-impeding-progress-machine-learning-and-the-eu-copyright-framework/#_ftn7.
94. KOPECKÁ, M. *Nová zákonná licence pro užití autorských děl* [online]. epravo.cz, 2018. Dostupné z <https://www.epravo.cz/top/clanky/nova-zakonna-licence-pro-uziti-autorskych-del-105909.html?mail>.
95. Llama 2: open source, free for research and commercial use [online]. Dostupné z: <https://llama.meta.com/llama2/>.
96. LÖLFING, N. *Part 1: Product compliance for consumer products powered by AI* [online]. Bird & Bird, 2024. Dostupné z <https://www.twobirds.com/en/insights/2024/global/product-compliance-for-consumer-products-powered-by-ai>.
97. ŁABUZ, M. *Regulating Deep Fakes in the Artificial Intelligence Act* [online]. ACIG, vol. 2, no. 1, 2023. Dostupné z <https://www.acigjournal.com/pdf-184302-105060?filename=Regulating%20Deep%20Fakes%20in.pdf>.
98. MATZNER, J. *Směrnice o odpovědnosti za umělou inteligenci: Krok k odpovědnému využívání AI* [online]. epravo.cz, 2024. Dostupné z <https://www.epravo.cz/top/clanky/smernice-o-odpovednosti-za-umelou-inteligenci-krok-k-odpovednemu-vyuzivani-ai-117394.html>.
99. Ministerstvo spravedlnosti České republiky. *Tisková zpráva Ministerstva spravedlnosti: Ministerstvo spravedlnosti odeslalo do mezipřevodního řízení očekávané změny v trestním právu* [online], 2024. Dostupné z <https://justice.cz/web/msp/tiskove-zpravy/-/clanek/ministerstvo-spravedlnosti-odeslalo-do-meziprevodniho-pripominkoveho-rizeni-ocekavane-zmeny-v-trestnim-pravu-kopirovat->.
100. NORDEMANN, J. B. *EU law: Generative AI, copyright infringements and liability – My guess for a hot topic in 2024* [online]. Kluwer Copyright Blog, 2024. Dostupné z <https://copyrightblog.kluweriplaw.com/2024/01/23/eu-law-generative-ai-copyright-infringements-and-liability-my-guess-for-a-hot-topic-in-2024/>.
101. The simplest, fastest repository for training/finetuning medium-sized GPTs [online]. Dostupné z: <https://github.com/karpathy/nanoGPT>.

102. OPENAI. *Comment Regarding Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation*. United States Patent and Trademark Office, Department of Commerce, Docket No. PTO–C–2019–0038, Addressing Question 3 [online], 2019. Dostupné z: <https://cdn.openai.com/policy-submissions/OpenAI%20Comments%20on%20Intellectual%20Property%20Protection%20for%20Artificial%20Intelligence%20Innovation.pdf>.
103. OPENAI. *GPT-4 Technical Report* [online]. 2024. Dostupné z <https://cdn.openai.com/papers/gpt-4.pdf>.
104. OPENAI. *How should AI systems behave, and who should decide?* [online], 2023. Dostupné z <https://openai.com/index/how-should-ai-systems-behave/>.
105. OPENAI. *New AI classifier for indicating AI-written text* [online], 2023. Dostupné z: <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>.
106. OPENAI. *OpenAI—written evidence (LLM0113) House of Lords Communications and Digital Select Committee inquiry: Large language models* [online], 2023. Dostupné z: committees.parliament.uk/writtenevidence/126981/pdf/.
107. OPENAI. *OpenAI and journalism* [online]. 2024. Dostupné z <https://openai.com/index/openai-and-journalism/>.
108. OpenLLaMA: *An Open Reproduction of LLaMA. Apache License 2.0*. Dostupné z: https://github.com/openlm-research/open_llama?tab=Apache-2.0-1-ov-file.
109. OpenLLaMA: *An Open Reproduction of LLaMA*. Dostupné z https://huggingface.co/openlm-research/open_llama_7b_v2.
110. PASCHOU, V. *Bias in artificial intelligence: risks and solutions. AI, Data subjects' rights* [online], 2024. Dostupné z <https://www.activemind.legal/guides/bias-ai>.
111. PAUL TREMBLAY, et al. v. OPENAI, INC., et al. Žaloba vedená u UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF CALIFORNIA pod sp. zn. 3:23-cv-03223-AMO. Dostupné z <https://www.courtlistener.com/docket/67538258/104/tremblay-v-openai-inc/>.
112. PETERS, J., WES, D., *The New York Times blocks OpenAI's web crawler* [online], The Verge, 2023. Dostupné z <https://www.theverge.com/2023/8/21/23840705/new-york-times-openai-web-crawler-ai-gpt>.
113. PINDUR, V. *Nelegální stahování autorských děl z pohledu trestního práva* [online]. epravo.cz, 2022. Dostupné z: <https://www.epravo.cz/top/clanky/nelegalni-stahovani-autorskych-del-z-pohledu-trestniho-prava-115052.html>.
114. PRCHAL, P. *Generativní umělá inteligence a porušování autorských práv*. Praha: Bulletin advokacie, 3/2024, 2024. ISSN: 1210-6348.

115. QUINTAIS, J.P. *Generative AI, Copyright and the AI Act* [online]. Amsterdam: Institute for Information Law (IViR), 2023. Dostupné z <https://copyrightblog.kluweriplaw.com/2023/05/09/generative-ai-copyright-and-the-ai-act/>.
116. SEDLAKOVA LEGAL s.r.o. *Umělá inteligence a databáze: Učení umělé inteligence analýzou dat z databáze* [online]. 2022. Dostupné z <https://www.sedlakovalegal.cz/cs/uceni-umele-inteligence-analyzou-dat-z-databaze>.
117. SHORTLIFFE, E.H. *MYCIN: A KNOWLEDGE-BASED COMPUTER PROGRAM APPLIED TO INFECTIOUS DISEASES* [online]. Stanford University School of Medicine, 1977. Dostupné z <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2464549/pdf/procascamc00015-0074.pdf>.
118. SOUSA E SILVA, N. *Are AI models' weights protected databases?* [online]. Lisbon: Universidade Católica Portuguesa, 2024. Dostupné z <https://copyrightblog.kluweriplaw.com/2024/01/18/are-ai-models-weights-protected-databases/>.
119. THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OP CO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC a OPENAI HOLDINGS, LLC. Žaloba vedená u UNITED STATES DISTRICT COURT SOUTHERN DISTRICT OF NEW YORK pod sp. zn. 1:23-cv-11195. Dostupné z https://nytcassets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf.
120. THE AUTHORS GUILD. *Legislative Proposals to Protect the Creative Professions and Mitigate Risk of Harm from Generative AI* [online], 2023. Dostupné z <https://authorsguild.org/app/uploads/2023/10/Creators-AI-Legislative-Proposals.pdf>.
121. THOMAS, J. *Fake news on the rise as the European elections draw near* [online]. Euronews, 2024. Dostupné z <https://www.euronews.com/my-europe/2024/05/20/fake-news-on-the-rise-as-the-european-elections-draw-near>.
122. VANIAN, J., LESWING, K. *ChatGPT and generative AI are booming, but the costs can be extraordinary* [online]. cnbc.com, 2023. Dostupné z <https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html>.
123. VIGLIENSONI, G., PERRY, P., FIEBRINK, R. *A Small-Data Mindset for Generative AI Creative Work*. Generative AI and HCI - CHI 2022 Workshop [online], 2022. Dostupné z https://ualresearchonline.arts.ac.uk/id/eprint/18343/1/CHI_Workshop_GenAI.pdf.
124. VÝBOR EVROPSKÉHO PARLAMENTU PRO KULTURU A VZDĚLÁVÁNÍ. *Opinion of the Committee on Culture and Education for the Committee on the Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs on the proposal for a regulation of*

the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts ze dne 16. června 2022. [2021/0106(COD)].

125. Usnesení Evropského parlamentu ze dne 20. října 2020 o právech duševního vlastnictví při vývoji technologií umělé inteligence. [2020/2015(INI)].
126. WESTERLUND, M. *The Emergence of Deepfake Technology: A Review*. *Technology Innovation Management Review*, 9(11) [online], 2019, s. 40-53. Dostupné z <http://doi.org/10.22215/timreview/1282>.
127. WIPO. *Světová deklarace o duševním vlastnictví*. 2000.
128. YOUR EUROPE. *Database protection*. Dostupné z https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.
129. ZORTHIAN, J. *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic* [online], *Time*, 2023. Dostupné z <https://time.com/6247678/openai-chatgpt-kenya-workers/>.

Abstrakt

Právní aspekty generativní umělé inteligence

V posledních letech jsme svědky pokroku v oblasti generativní umělé inteligence. Na trh byly uvedeny nástroje, které jsou schopny produkovat texty, obrázky, zvuk a videa, a to v kvalitě prakticky nerozpoznatelné od obsahu vytvořeného člověkem. Tento technologický pokrok s sebou přináší nové právní výzvy, které zahrnují otázky využívání autorských děl k trénování AI, problematiku deepfake, šíření dezinformací (fake news) a určování autorství výstupů generativní AI.

Předkládaná rigorózní práce zhodnocuje současnou právní úpravu v České republice a na úrovni EU, která se snaží tyto otázky harmonizovat. V úvodu práce autor poskytuje mezioborový pohled na generativní umělou inteligenci, který propojuje technické a právní aspekty s cílem usnadnit pochopení této složité problematiky i osobám bez technického vzdělání. Kromě toho obsahuje práce výklad k relevantní právní úpravě, která se k této problematice vztahuje, a to především s důrazem na odvětví autorského práva.

Těžištěm práce je tzv. „životní cyklus“ generativní umělé inteligence. Autor provází čtenáře různými fázemi vývoje generativní umělé inteligence, od shromažďování a zpracování trénovacích dat až po samotnou fázi generování nového textu, obrázku nebo jiného obsahu. V každé fázi jsou identifikována právně relevantní jednání subjektů (autorů, poskytovatelů AI, uživatelů) a jejich potenciální právní důsledky, zejména pro právo autorské. Výklad je doplňován názornými příklady a zdůrazňuje praktické aspekty problematiky. Cílem práce je prokázat, že navzdory názorům o právním vakuu v této oblasti, stávající právní rámec v zásadě pokrývá mnoho relevantních aspektů generativní AI. Zároveň však poukazuje na to, že současná regulace dostatečně nereflektuje dlouhodobé negativní dopady, které masová produkce obsahu prostřednictvím AI může způsobit.

Klíčová slova: generativní umělá inteligence, deepfake, autorské právo

Abstract

Legal Aspects of Generative Artificial Intelligence

In recent years, we have witnessed significant advancements in generative artificial intelligence (AI). Tools have been introduced to the market that are capable of producing text, images, audio, or video at a level of quality that is virtually indistinguishable from content created by humans. This technological progress introduces new legal challenges, including the use of copyrighted works for AI training, issues surrounding deepfakes, the spread of misinformation (fake news), and determining authorship of AI-generated outputs.

This rigorous thesis evaluates the current legal framework in the Czech Republic and at the European Union level, which seeks to harmonize these issues. At the outset, the author provides an interdisciplinary perspective on generative AI, connecting technical and legal aspects to facilitate the understanding of this complex issue for those without a technical background. In addition, the thesis offers an explanation of the relevant regulation, with a particular emphasis on copyright law.

The core of the thesis is centered on the so-called "life cycle" of generative artificial intelligence. The author guides the reader through the various stages of generative AI development, from the collection and processing of training data to the actual generation of new text, images, and other content. In each stage, legally relevant actions of the involved parties (creators, AI providers, users) are identified, along with their potential legal implications, particularly with respect to copyright law. The explanation is supplemented by illustrative examples and emphasizes the practical aspects of the issue. The goal of the thesis is to demonstrate that, contrary to opinions suggesting a legal vacuum in this area, the existing legal framework largely covers many relevant aspects of generative AI. However, it also highlights that current regulation does not sufficiently address the long-term impacts that the mass production of content via AI may bring.

Keywords: Generative Artificial Intelligence, Deepfake, Copyright