

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE
FILMOVÁ A TELEVIZNÍ FAKULTA

BAKALÁŘSKÁ PRÁCE

Matej Martinec

Praha, 2024

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE
FILMOVÁ A TELEVIZNÍ FAKULTA

Filmové, televizní a fotografické umění a nová média
Audiovizuální studia

BAKALÁŘSKÁ PRÁCE

KRITICKÁ REFLEXE GENERATIVNÍ UMĚLÉ INTELIGENCE

Matej Martinec

Vedoucí práce: Mgr. Palo Fabuš
Přidělovaný akademický titul: BcA.

Praha, květen 2024

ACADEMY OF PERFORMING ARTS IN PRAGUE
FILM AND TV SCHOOL

Film, Television, Photography and New Media
Audiovisual Studies

BACHELOR'S THESIS

CRITICAL REFLECTION ON GENERATIVE ARTIFICIAL INTELLIGENCE

Matej Martinec

Thesis advisor: Mgr. Palo Fabuš
Academic title granted: BcA.

Prague, May 2024

Prohlášení

Prohlašuji, že jsem bakalářskou práci na téma

Kritická reflexe generativní umělé inteligence,

vypracoval samostatně pod odborným vedením vedoucího práce a s použitím pouze uvedené literatury a pramenů a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu. Souhlasím s tím, aby práce byla zveřejněna v souladu se zákonem a vnitřními předpisy AMU.

Praha, dne

Matej Martinec

Poděkování

Děkuji vedoucímu práce Mgr. Palovi Fabušovi za odbornou pomoc při zpracování této práce.

Abstrakt

Tato práce se věnuje obecnému shrnutí problematiky generativní umělé inteligence. Jelikož se jedná o velmi široký fenomén, jehož kritická reflexe zasahuje do spousty teoretických odvětví, tak i tato práce volí komparativní a multidisciplinární přístup pro získání nadhledu. Zabývá se tématy environmentálními, jako je spotřeba energie a extrakce nerostů. Etickými jako jsou autorská práva nebo zaujatost dat. Ekonomickými a sociálními jako jsou trh práce a proliferační dezinformací.

Cílem práce je skrze veškerá zmíněná témata utvořit celistvý a kritický pohled na generativní umělou inteligenci, přesahující jednu disciplínu a senzacechtivé titulky. Na závěr by tato práce měla nabídnout kritický pohled na to, co generativní umělá inteligence je a čím je podmíněná.

Abstract

This thesis is dedicated to a general summary of generative artificial intelligence. As it is a very broad topic and its critical reflection spans many theoretical fields, this thesis adopts a comparative and multidisciplinary approach to gain insight. It deals with environmental topics such as energy consumption and mineral extraction. Ethical ones such as copyright and data bias. Economic and social issues such as the labour market and the proliferation of misinformation.

The aim of this thesis is to form a comprehensive and critical view of generative artificial intelligence through all these subjects, transcending one discipline and sensational headlines. In conclusion, this thesis should offer a critical perspective on what generative artificial intelligence is and what it is conditioned by.

OBSAH

ÚVOD	13
1. EXTRAKCE – Enviromentální kontext umělé inteligence	15
2. MECHANICKÝ TUREK – Lidský kontext umělé inteligence	18
3. SCRAPING A BIAS – Datový kontext umělé inteligence	20
4. MORÁLNÍ PANIKA – Společenský kontext generativní umělé inteligence	23
4.1. Dezinformace	
4.2. Trh práce	25
ZÁVĚR	27
SEZNAM POUŽITÝCH ZDROJŮ	29

ÚVOD

Generativní umělá inteligence je pokročilá statistika, která díky strojovému učení a obrovskému množství dat, vykazuje schopnosti generovat nové výstupy na základě dat, na kterých byla vytrénována. Schopnosti generativní umělé inteligence záleží mimo typu datasetu i na modalitě, což znamená, s jakým vstupem a výstupem může pracovat. Způsobů strojového učení je několik, ať už hluboké učení, generativní předtrénované transformátory, difúze nebo generativní adversariální sítě. Tyto metody se používají napříč všemi typy generativních umělých inteligencí. Textové generativní umělé inteligence, jsou trénované na datasetech slov a tokenech slov, následně dokáží zpracovávat přirozený jazyk. Vizualní generativní umělé inteligence jsou trénované na datasetech obrazů spárovaných s popisky. Tyto systémy dokáží generovat fotorealistický obraz, dále se může jednat o generativní umělé inteligence, které vytvářejí zvuk, hudbu nebo video. Důležité je, že výsledky tohoto generování nejsou součástí vstupních dat. A na základě této schopnosti vytvářet nové výstupy je tomuto procesu přisuzován termín inteligence.

Celý pojem umělé inteligence je často problematizován, protože není těžké definovat jenom inteligenci lidskou, ale stejně složité je definovat i tu strojovou. Známy test inteligence strojů Alana Turinga, který je postaven na tom, jestli je člověk schopen rozlišit, jestli jedná se strojem nebo s člověkem, je vyvrácen už v 80. letech 20. století. Turingův test kritizuje John Searle v článku Minds, Brains, and Programs, protože podle Searle stroj může používat znaky, kterým sám nerozumí, čímž nedokazuje, že umí přemýšlet a tudíž je inteligentní. Stroj pouze odhaduje a napodobuje, co se od něj očekává. John Searle přichází s myšlenkovým experimentem Čínské místnosti, který jakoukoliv inteligenci u strojů popírá. Dalším bodem kritiky Turingova testu ze strany kognitivních věd je faktor naivity, přístupu, nebo schopností člověka, který rozlišuje, jestli jedná se strojem nebo nikoliv. Stroj se nemusí snažit být inteligentní, stačí mu snažit se oklamat člověka, který ho testuje.

Tato práce se však nezabývá problematizací strojové inteligence, zabývá se kritickou reflexí generativní umělé inteligence, což jsou v lidské interakci softwary a platformy většinou zabalené do přívětivého uživatelského rozhraní a volně nebo za nízký poplatek, dostupné na internetu. Tyto softwary a platformy na počkání vytvářejí vysoce komplexní text, obraz, zvuk nebo video. Na pozadí však dochází k nesčetnému množství operací, které jsou uživateli skryté. Nevidíme množství a dopady fyzické práce, nevidíme kognitivní lidskou práci, nevidíme spotřebu elektřiny a nevidíme ani množství drahých kovů, ve kterých se tyto procesy odehrávají.

Dnes dopoledne jsem se dočetl, že generativní umělá inteligence je schopná neskutečných věcí. V dalším článku jsem si přečetl, že zaplaví svět dezinformacemi, že změní pracovní trh nejen v kreativním odvětví a každý, kdo nezačne používat generativní umělou inteligenci, bude nahrazen někým, kdo to dělá. Také že má obrovskou spotřebu vody, energie, ale pokud počkáme ještě několik měsíců, už bude efektivnější a její spotřeba vody i energie se radikálně sníží. Nakonec ale zjistím, že jsou stejně veškeré její výstupy šeredně zaujaté, což nevádí, protože se už programuje řešení.

V první kapitole této práce se věnuji environmentálnímu kontextu umělé inteligence za pomoci geologie médií Jussi Parikky, v druhé kapitole se věnuji lidskému kontextu umělé inteligence, díky paralele s podvodným šachovým automatem z 18. století. V třetí kapitole se věnuji datovému kontextu umělé inteligence, kdy skrz knihu *Discriminating Data* od Wendy Hui Kyong Chun přicházím s historickým pozadím datové analytiky současnosti. V poslední kapitole se za pomoci obsáhlých studií *Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown* z Harvard Kennedy School, a *Artificial Intelligence Index Report 2024* ze Standfordské univerzity věnuji společenské panice, obklopující generativní umělou inteligenci.

Tato práce si klade za cíl objasnit většinu tvrzení, které jsem se dnes dopoledne dočetl, situovat je do celistvého kontextu a poskytnout základní informace pro pochopení, co generativní umělá inteligence je a čím je podmíněna.

1. EXTRAKCE – Enviromentální kontext umělé inteligence

I když se tato práce zabývá převážně generativní umělou inteligencí, je nejprve nutné ukotvit ji v obecnějších reáliích, ze kterých vychází, a to jsou podmínky vzniku většiny dílů skládačky, které jsou nezbytné k tomu, aby jakékoliv hluboké učení mohlo probíhat. Diagram¹ od Kate Crawford a Vladana Jolera vizualizuje ústřední trojici, extraktivních procesů, které jsou nutné pro spuštění rozsáhlého systému umělé inteligence: materiální zdroje, lidská práce a data. V této kapitole se budu věnovat té první části, materiálním zdrojům.

Jussi Parikka ve své knize *A Geology of Media* tvrdí, že bychom se na média neměli dívat jako na extenzi našich smyslů, jak to tvrdí Marshall McLuhan, ale jako na extenzi země. Parikka nahlíží média kontextem geologických procesů. Ať už jde o lithium z Bolívie, které je využíváno ke konstrukci baterií, měď a kobalt v Kongu, která se používá k tištění obvodů nebo vodu, která je součástí extraktivních procesů, stejně jako čištění vyextrahovaných nerostů. Všechny objekty v složitém systému umělé inteligence, jsou tvořeny z elementů, které na zemi geologicky vznikaly miliardy let.²

To že je technologie jako například umělá inteligence považována za něco výjimečného, vymykající se realitě a lidské efektivitě, je jenom dalším trikem korporací, hned po přívětivém uživatelském rozhraní aplikací, technologie nevzniká z ničeho, ale už ze zmíněných vzácných kovů, které nemají lidské kvality a ani nepodléhají měřítkům lidské efektivitě. Sean Cubbit v eseji *Lví podíl* tvrdí, že komerční subjekty jsou navíc přesvědčeny, že mohou vše zdarma extrahovat z přírody. Tyto zdroje mají minimální náklady a lze je využívat a znehodnocovat do takové míry, do jaké se to komerčním subjektům hodí.³ Stejně jako od 17. století odmítají kolonisté nárok na území původních obyvatel, i nyní korporace odmítají nároky lidstva na mořské dno nebo Arktidu.⁴ Což sice napomáhá tíženému technologickému pokroku, ale neúměrně a neudržitelně ničí přírodu. Regulace a restrikce samozřejmě existují, ale když je komerční subjekt orientován hlavně na zisk, tak může být placení pokut za způsobování enviromentálních škod předpokládanou součástí rozpočtu. Mezi lety 2011-2013 byla tchajwanská společnost *Advanced Semiconductor Engineering* pokutována sedmkrát nejvyšší možnou pokutou za vypouštění znečištěné vody do řeky, která je používána

¹ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 7

² tamtéž

³ CUBBIT, Sean. *Lví podíl*. In: *Objekt*, 2015. Praha: Kvalitář s.r.o. ISBN 978-80-260-8639-0. s 141.

⁴ tamtéž, s 144-145.

k zavlažování o rok později, v roce 2014 tato firma, i přes zmiňované pokuty, mnohonásobně zvýšila produkci.⁵

V rámci fungování systémů umělé inteligence jsou tři procesy, které jsou extrémně energeticky náročné. Těžba drahých kovů, které jsem se už částečně dotkl, trénování systémů umělé inteligence, a samotné používání umělé inteligence.

Druhá fáze, trénování systémů je o procesu, ve kterém se skrz složité algoritmické výpočty velké množství počítačů učí z obrovského množství dat, trénování velkého jazykového modelu GPT-3 spotřebovalo tolik energie jako 120 amerických domácností za celý rok a trénování GPT-4 by kvůli třikrát většímu množství dat mělo spotřebovat 40x víc energie.⁶ Tento trend nemá klesající tendenci.

Třetí fáze, což je planetární používání takového modelu je zarážející hlavně proto, že srovnání klasického vyhledávání předem vypočítané otázky v google search je výrazně méně náročné než vyhledávání skrz generativní model, který kvůli odpovědi prohledává a interpretuje celý dataset, což může být například celý volně dostupný internet z roku 2022.⁷ Rozsáhlé systémy umělé inteligence spotřebovávají obrovské množství energie. Přesto podrobnosti o těchto nákladech zůstávají ve společenské představitosti nejasné. Zpráva organizace Greenpeace uvádí: „Jednou z největších překážek pro transparentnost tohoto odvětví je Amazon Web Services (AWS). Největší světová cloudová společnost zůstává téměř zcela netransparentní, pokud jde o energetickou stopu svých rozsáhlých provozů.“⁸

Pokud se však na veřejnost dostanou informace o enviromentálních dopadech továren, korporace se mohou snažit omezit své emise do ovzduší, ale to nepomůže s množstvím odpadu a znečištěním vody, které způsobily za posledních pět nebo deset let. Cubbit dále píše, že některé perfluorované látky (PFcs), vypouštěné do ovzduší z chemických úložišť nebo při leptání plazmou v továrnách, zůstávají v atmosféře tisíce a možná i desetitisíce let a jejich částice mají až dvacetitisícinásobný dopad na skleníkový efekt ve srovnání s molekulami oxidu uhličitého. Továrny zanechávají i jiné prvky a pevný odpad, který obsahuje karcinogeny a další sloučeniny, jejichž dlouhodobý efekt není známý, budou přežívat v okolí továren podobně dlouho. Udržitelný rozvoj konfrontuje počítačový průmysl také s jeho geografíí, konkrétně s vlivem vedlejších produktů na kvalitu podzemních vod nebo cirkulací ovzduší, díky které se znečištění šíří na vzdálená místa (například mezistátní

⁵ CUBBIT, Sean. Lví podíl. In: *Objekt*, 2015. Praha: Kvalitář s.r.o. ISBN 978-80-260-8639-0, s 154.

⁶ CLIMATE ACTION AGAINST DISINFORMATIONS. *Artificial intelligence threats to climate change*. Online, report. [2024]. Dostupné z: https://foe.org/wp-content/uploads/2024/03/AI_Climate_Disinfo_v6_031224.pdf. [cit. 2024-05-15].

⁷ tamtéž

⁸ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 17

znečištění z továren na polovodiče v USA). Perspektiva udržitelnosti také poukazuje na budoucnost, ve které dědictví dlouho opuštěných továren, může vyvolat ještě dalekosáhlejší důsledky.⁹

Hledání problematických segmentů v rámci výrobních procesů může být velmi složité, jako příklad může posloužit snaha společnosti Intel, která se snažila zajistit, aby se v jejích procesorech nepoužívalo tantalum z Konga. Intel má víc než 19,000 dodavatelů z více než 100 krajin a pochopení svého vlastního dodavatelského řetězce této společnosti trvalo víc než čtyři roky. Podle společnosti Dell, která vyrábí počítače, představuje složitost dodavatelského řetězce kovů téměř nepřekonatelnou výzvu. Těžba těchto nerostů probíhá dlouho před sestavením konečného výrobku, takže je nesmírně obtížné sledovat původ nerostů. Kromě toho se mnoho minerálů taví společně s recyklovanými kovy, tudíž je téměř nemožné dohledat jejich zdroj. Vidíme tedy, že pokus o zachycení celého dodavatelského řetězce je skutečně obrovským úkolem: odhaluje celou složitost globální výroby technologických produktů v 21. století.¹⁰

Každý malý okamžik pohodlí – ať už jde o vygenerování obrázku, napsání textu nebo přezpívání písničky Jamesem Deanem – vyžaduje rozsáhlou planetární síť poháněnou těžbou neobnovitelných materiálů...¹¹ Enviromentální kontext umělé inteligence začíná u drahých kovů v zemi, pokračuje do továren, kde se z drahých kovů vyrábí hardware, složený hardware se dostává do datových center. Na všech těchto místech je spotřebováno enormní množství vody a energie.

⁹ CUBBIT, Sean. Lví podíl. In: *Objekt*, 2015. Praha: Kvalitář s.r.o. ISBN 978-80-260-8639-0, s 151.

¹⁰ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 11.

¹¹ tamtéž, kapitola 2.

2. MECHANICKÝ TUREK – Lidský kontext umělé inteligence

Turek byl mechanický stroj, který byl svými vlastníky vydáván za šachový automat, avšak ve skutečnosti, se jednalo o propracovaný podvod. Tento mechanismus byl zkonstruován roku 1771 bratislavským vynálezcem Wolfgangem von Kempelenem¹² a jeho podvodnost byla rozkryta až zhruba osmdesát let po jeho vytvoření. Zařízení se skládalo z modelu torza lidského těla v životní velikosti, oblečeného v tureckém oděvu s turbanem. Dále byla součástí zařízení velká skříň. Celá skříň měla stěny, které se tvářily jako průhledné a skrz které odhalovala svůj vnitřek, plný ozubených koleček a drátů. Tyto stěny byly během vystoupení po všech stranách otevírány, aby byly odhaleny čistě mechanické útroby stroje. Celý podvod spočíval v tom, že šachovými figurkami nepohyboval žádný inteligentní stroj, mechanický Turek, ale šachový mistr trpasličího vzrůstu, který se nacházel uvnitř skříně, svítil si svíčkou a pozoroval situaci na šachovnici díky magnetům. Poté skrz sadu pák ovládal levou ruku modelu a vykonával tahy na šachovnici. Když byly v rámci představení odhalovány útroby skříně, šachista unikal pohledu diváků díky posuvnému sedátku a sadě dvojíých dveří.

Jak to ale souvisí s umělou inteligencí? Mechanický turek není jenom podvod z 18. století, je to podvod z 21. století. Platforma Amazon Mechanical Turk, kterou v roce 2005 založila právě korporace Amazon a o dekádu předběhla boom zakázkové ekonomiky a celoplanetární námezdní práce. Uživatelům této platformy se může zdát, že aplikace využívá k plnění úkolů pokročilou umělou inteligenci, avšak za přívětivým rozhraní, vevnitř šachového automatu se nachází umělé inteligence, řízené vzdálenou, rozptýlenou a špatně placenou pracovní silou, která pomáhá klientovi dosáhnout jeho vytyčených cílů.¹³ Nejdůležitější součástí těchto firem je outsourcing pracovníků. V Americe tyto platformy právně chápou své pracovníky jako nezávislé dodavatele, nikoli jako zaměstnance. To umožňuje společnostem ušetřit přibližně 30 % nákladů na pracovní sílu tím, že se zbaví benefitů, přesčasů, nemocenské a dalších nákladů.¹⁴ Tyto praktiky většinou neaplikují právě firmy, zabývající se vývojem umělé inteligence, ale outsourcují je u jiných firem nebo platform.

¹² CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 18.

¹³ tamtéž

¹⁴ SRNICEK, Nick. *Platform Capitalism*, 2016. Cambridge: Polity ISBN 978-1-5095-0486-2, kapitola Lean platforms.

Tento druh neviditelné, skryté práce, outsourcované nebo crowdsourcované, skryté za rozhraními a maskované v rámci algoritmických procesů je dnes běžný. Zejména v procesu tagování a označování tisíců hodin digitálních archivů pro účely krmení neuronových sítí. Druhý pilíř diagramu¹⁵ Kate Crawford a Vladana Jolera je lidská práce.

Výstižným příkladem může být článek, který pojednává o snaze firmy OpenAI z roku, aby se jejich velký jazykový model GPT stal méně toxickým. Kdy skrz firmu operující v Africe OpenAI poslalo do Keňi ke kategorizaci tisíce úryvků z nejtemnějších koutů internetu, texty explicitně pojednávající o zneužívání dětí, znásilnění nebo vraždě. Pracovníci a pracovnice, které tuto traumatizující práci vykonávali, byli placeni méně než \$2 za hodinu.¹⁶ Někdy je tato práce traumatizující, jindy špatně placená, ale vždy je neviditelná. Jako v případě reCAPTCHA společnosti Google. Paradoxem, který mnozí z nás zažili je, že abyste dokázali, že nejste bot nebo umělý agent, jste nuceni zdarma trénovat systém umělé inteligence společnosti Google pro rozpoznávání obrázků, a to tak, že vyberete několik políček, která obsahují čísla ulic, auta nebo domy.¹⁷

Sean Cubbit v eseji *Lví podíl* píše o tom, že bychom si měli uvědomit, že ve strojích, které nás obklopují, jsou uvězněny schopnosti našich předků.¹⁸ Což u systémů umělé inteligence dává smysl nejen protože skuteční lidé se podíleli na kategorizaci dat, na kterých se systémy umělé inteligence učí, ale také protože toto kategorizování znamená odevzdávání strojům lidskou schopnost rozlišovat a pojmenovávat heterogenní fenomény, což žádný stroj sám od sebe nedokáže. Stejně jako křemík není srovnatelný s lidskou tkání, ani práce miliónů lidí není srovnatelná s prací jednoho člověka.

Jak vysvětluje Fuchs, digitální práce zdaleka není efemérní nebo virtuální, ale je hluboce zakotvena v různých materiálních činnostech. Její rozsah je ohromující: od námezdní práce v dolech při těžbě nerostů které, tvoří fyzický základ informačních technologií, přes práci v přísně kontrolovaných a někdy nebezpečných procesech výroby a montáže hardwaru v čínských továrnách, až po vykořisťované externí kognitivní pracovníky v rozvojových zemích, kteří označují soubory tréninkových dat umělé inteligence.¹⁹

¹⁵ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 17

¹⁶ PERRIGO, Billy. *OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic*. TIME, 2023. Dostupné z: <https://time.com/6247678/openai-chatgpt-kenya-workers> [cit. 2024-05-15]

¹⁷ tamtéž jako poznámka číslo 15, kapitola 18.

¹⁸ CUBBIT, Sean. *Lví podíl*. In: *Objekt*, 2015. Praha: Kvalitář s.r.o. ISBN 978-80-260-8639-0, s. 140.

¹⁹ tamtéž jako poznámka číslo 15, kapitola 9.

3. SCRAPING A BIAS – Datový kontext umělé inteligence

Dostáváme se ke třetímu pilíři umělé inteligence podle diagramu²⁰, a to jsou data.

Web scraping, je metoda plošného extrahování dat z internetu, kdy algoritmus stahuje data, buď tak, jak jsou nebo k následné kategorizaci člověkem. Bez této praxe by generativní umělá inteligence, jak ji známe, nikdy nemohla existovat.

Laion-5b je veřejně dostupný dataset z roku 2022 obsahující 5,8 miliardy obrázků spárovaných s popisky. Celý dataset je stažen z internetu. Na tomto datasetu byl trénován i model Stable Diffusion od Stability AI. Hito Steyerl ve své eseji Mean Images píše, že celý dataset Laion-5b je produktem neplacené lidské práce, od lidí, kteří programují a navrhují webové stránky až po uživatele, kteří na ně nahrávají a zveřejňují svoje obrázky. Je samozřejmé, že nikdo z těchto lidí nedostal odměnu ani podíl na datovém fondu nebo produktech a modelech, které z něj korporace vytvořila. Soukromá vlastnická práva v rámci digitálního kapitalismu, i mimo něj, mají význam pouze tehdy, jde-li o vlastníky. Všichni ostatní mohou být běžně okrádáni.²¹ Ke sběru dat dochází i na platformách jako je Facebook, uživatelé trénují neuronové sítě pomocí svého chování, svého hlasu, pomocí označování obrázků a nahráváním videí. Tyto data jsou kontrolována a využívána jen několika málo lidmi na vrcholu pyramidy.²² Paradoxem může být, že dataset jako Laion-5b je veřejně dostupný, což může působit jako demokratický fenomén. Avšak představa, že si tento model může každý stáhnout a vytrénovat si svůj komerční model, není kvůli počítačové náročnosti úplně uskutečnitelná, a proto ho stejně dokáže vytrénovat jenom velká korporace.

Extrakce dat nekončí pouze u obsahu nahrávaného na sociální platformy, v současné době roste snaha získávat hlubší data, vrstvy znalostí, které lze využívat ke kvantifikaci lidských schopností, psychiky, vědomí i nevědomí. V tomto jsme byli svědky vzniku vícenásobné kognitivní ekonomie z ekonomie pozornosti, ekonomiky dohledu a ekonomiky emocí. Všechny formy biologických údajů – včetně forenzních, biometrických, sociometrických a psychometrických – jsou zachycovány a zaznamenávány do databází pro trénink umělé inteligence.²³

²⁰ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 4.

²¹ STEYERL, Hito. *Mean Images*, 2023. *New Left Review*. (140/141), 82–97. [2023] Dostupné z: <https://newleftreview.org/issues/ii140/articles/hito-steyerl-mean-images> [cit. 2024-05-15]

²² tamtéž jako poznámka číslo 20, kapitola 17.

²³ tamtéž jako poznámka číslo 20, kapitola 14.

Novou zlatou horečkou v kontextu umělé inteligence je snaha uchopení různých oblastí lidského poznání, cítění a jednání, aby bylo možné tyto oblasti zachytit a privatizovat. Když v listopadu 2015 získala společnost DeepMind Technologies Ltd. přístup ke zdravotním záznamům 1,6 milionu identifikovatelných pacientů nemocnice Royal Free, stali jsme se svědky zvláštní formy privatizace: extrakce hodnoty znalostí. Soubor dat může být stále veřejně vlastněn, ale jejich metahodnota – model, který je na jejich základě vytvořen – je v soukromém vlastnictví. Ačkoli existuje mnoho dobrých důvodů, proč usilovat o zlepšení veřejného zdravotnictví, existuje reálné riziko, pokud se tak děje za cenu skryté privatizace veřejných zdravotnických služeb.²⁴

Steyerl ve své eseji píše také, že tréninkové sady pro systémy umělé inteligence tvrdí, že pronikají do jemné podstaty každodenního života, ale opakují ty nejstereotypnější a nejomezenější sociální vzorce, znovu zavádějí normativní představu o lidské minulosti a promítají ji do lidské budoucnosti.²⁵

V 19. století britští eugenici vyvinuli metody korelace a lineární regrese, datovou analýzu a pětifaktorový model OCEAN. Francis Galton je široce oslavován za objev korelace a lineární regrese, které poprvé nazval lineární reverzí. Korelace nikdy nebyla jen o objevování podobností, ale také o pěstování podobností, právě korelace položila základy pro představu univerzálních zákonů, pro eugeniku tak důležitých.²⁶ Existuje mnoho podobností mezi eugenikou dvacátého století a datovou analytikou století jednadvacátého. Obě kladou důraz na sběr dat a dohled, zejména nad chudými populacemi; obě považují svět za laboratoř a obě podporují segregaci. Jediné, co tehdy eugenici neměli, byla veřejně dostupná velká data a algoritmy na jejich extrakci.²⁷

Rozdíl mezi eugenikou a datovou analytikou je o přesunu od národa k sousedství, prostřednictvím individuálních preferencí. Nikolas Rose argumentuje, že genetika jednadvacátého století není eugenická, protože se zaměřuje na jednotlivce a komunitu, spíše než na národ. Biopolitika jednadvacátého století se zaměřuje na řešení individuálních rizik, nikoliv na nařizování rasové čistoty.²⁸

²⁴ CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15], kapitola 20.

²⁵ tamtéž, kapitola 19.

²⁶ CHUN, Wendy Hui Kyong, *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*, 2021. Cambridge: The MIT Press. ISBN 978-0-262-54852-6, strana 59.

²⁷ tamtéž, strana 66-70.

²⁸ tamtéž, strana 84.

Wendy Hui Kyong Chun v knize *Discriminating Data* popisuje, že vazby mezi technologií rozpoznávání obličejů a eugenikou nejsou jen tematické nebo aspirační, ale také metodologické. Mají kořeny v metodách, jako je lineární diskriminační analýza, vyvinutá na počátku dvacátého století za účelem rozlišování mezi třídami a rasami lidí.²⁹ Nárůst výpočetního výkonu a veřejně dostupné obrazy obličejů zvýšily popularitu programů využívající umělou inteligenci. Nejpřímější vazbu na rané biometrické eugenické metody však představuje konstrukce složených obrazů pro určení typických tváří³⁰ jako například židovská tvář nebo tvář kriminálního.

V roce 2018 novomediální umělkyně Joy Buolamwini a počítačová vědkyně Timnit Gebru odhalily, že technologie rozpoznávání obličeje (FRT) má potíže s identifikací pohlaví subjektů tmavší pleti. Wendy Hui Kyong Chun říká, že rozdíl mezi rasou a rasismem je klíčový. Vzhledem k těmto programům a právní ochraně v USA se mnoho analýz zaměřilo na odhalení ukazatelů, které implicitně indexují rasu v explicitně barevně slepých systémech. Tyto chyby pocházejí z ignorování rasy, tedy z předpokladu, že bezrasové rovná se nerasistické. Bělošství jako výchozí hodnota – nebo to, co Simone Browneová nazvala prototypické bělošství – má znepokojivé důsledky.³¹

Matteo Pasquinelli v knize *The Eye of the Master* tvrdí, že vytváření alternativních algoritmů neznamená, že jsou etičtější. Například návrh pevně zakódovat etická pravidla do umělé inteligence a robotů se jeví jako velmi nedostatečný a neúplný, protože se přímo nezabývá širokou politickou funkcí automatizace v jejich jádru.³² Umělá inteligence problematicky ovlivňuje společnost tím, že vykresluje inteligenci, jako strojovou inteligenci a poznání vykresluje jako procedurální poznání. Vtiskuje nám svůj pohled na svět, který je, jak zmiňuji v této kapitole podmíněný problematickými metodami z 19. století.

²⁹ CHUN, Wendy Hui Kyong, *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*, 2021. Cambridge: The MIT Press. ISBN 978-0-262-54852-6, strana 186.

³⁰ tamtéž, strana 194.

³¹ tamtéž, strana 20-22.

³² PASQUINELLI, Matteo, *The Eye of the Master: A Social History of Artificial Intelligence*, 2023. London: Verso. ISBN 978-1-78873-006-8, poslední kapitola: Undoing the master algorithm.

4. MORÁLNÍ PANIKA – Společenský kontext generativní umělé inteligence

DEZINFORMACE

Dezinformace jsou jedním z největších společensky uznávaných strašáků generativní umělé inteligence. Avšak data o skutečném dopadu jsou převážně z rozvojových zemí. Podle Freedom House, který měří svobodu internetu, je v rozvojových zemích příkladů už několik. Například v Kyrgyzstánu si měly politické strany najímat blogery a středoškoláky, aby zveřejňovali propagandu na sociálních sítích, přičemž jedna skupina údajně zaplatila 350 000 dolarů za síť botů generovaných umělou inteligencí, která vytvářela 150 falešných profilů denně. Také Před druhým kolem prezidentských voleb v Brazílii v říjnu 2022 se na internetu rozšířilo deep fake video, v němž důvěryhodný zpravodajský server uvedl, že tehdejší prezident Jair Bolsonaro vede v průzkumech veřejného mínění.³³

Oba tyto příklady jsou z rozvojových zemí a jejich skutečný dopad je otázkou spekulace. Rozsáhlá studie Harvard Kennedy School, *Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown*³⁴, pojednává hlavně o kontextu západních bohatých a demokratických zemí.

Podle této studie je první argument o vlivu generativní umělé inteligence na dezinformace postaven na snadném přístupu a používání generativní umělé inteligence k tvorbě dezinformací ve velkém měřítu, a to s malými nebo žádnými náklady pro jednotlivce a organizované subjekty. Toto by mělo vést ke zvýšenému množství dezinformací. To umožňuje aktérům se špatnými úmysly zaplavit oblast nesprávnými nebo zavádějícími informacemi, a tím přehlušit věcný obsah anebo zasít zmatek.

Tento argument studie vyvrací tím, že generativní umělá inteligence usnadňuje vytváření dezinformací, což může zvýšit jejich množství, ale neznamená, že se zvýší i poptávka po dezinformacích na už dost nasyceném poli. Podle studií se dezinformace nenacházejí na mainstreamovém internetu, ale ve specifických prostředích se specifickými diváky, právě proto se exponenciální nárůst, díky generativní umělé inteligenci nejeví jako reálné riziko, protože se do mainstreamových médií ani nedostanou.

³³ FUNK, Allie, Adrian SHAHBAZ a Kian VESTEINSSON. *The Repressive Power of Artificial Intelligence*. Freedom House [online]. Dostupné z: <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence> [cit. 2024-05-15]

³⁴ SIMON, Felix M., Sacha ALTAY a Hugo MERCIER. *Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown*. 2023, Harvard Kennedy School Misinformation Review [online]. Dostupné z: <https://misinforeview.hks.harvard.edu/article/misinformation-reloaded-fears-about-the-impact-of-generative-ai-on-misinformation-are-overblown> [cit. 2024-05-15]

Druhý argument se podle studie Harvard Kennedy School³⁵ zabývá zvýšenou kvalitou dezinformací. Díky svým technickým možnostem a snadnému použití lze generativní umělou inteligenci použít k vytváření kvalitnějších dezinformací. Což by mělo vést ke zvýšení přesvědčovacího potenciálu, protože vytváří obsah, který je věrohodnější a hůře se vyvrací nebo ověřuje. To by buď umožnilo šíření nepravdivých informací, nebo by přispělo (spolu se zvýšeným množstvím dezinformací) k epistemické krizi, tedy k všeobecné ztrátě důvěry ve všechny typy zpráv.

Studie tento argument vyvrací třemi body. Za prvé, už nyní je možné vytvářet dost kvalitní dezinformace pomocí starších nástrojů, jako je Photoshop atd. Za druhé, jak už bylo zmíněno, dezinformace se většinou dostávají jen k úzkému výseku populace, tím pádem jejich kvalita nezajišťuje větší dosah. Za třetí, kvalita generativní umělé inteligence může pomáhat i ověřeným a spolehlivým zdrojům, aby atraktivita dezinformací převážila informace spolehlivé, musela by jejich atraktivita vzrůst až 100krát, aby překonala spolehlivé informace a nic nenasvědčuje, že by vliv generativní umělé inteligence byl takto nevyvážený.

Třetí argument je podle studie Harvard Kennedy School³⁶ o zvýšené personalizaci dezinformací. Díky svým technickým možnostem a snadnému použití lze generativní umělou inteligenci využít k vytváření vysoce kvalitních dezinformací, přizpůsobených vkusu a preferencím uživatele. Což by mělo vést k zvýšenému přesvědčování spotřebitelů o dezinformacích se stejnými výsledky jako výše.

Argument s personalizací dezinformací zní nejpřesvědčivěji, avšak studie tvrdí, že mikro targeting jednotlivců pro šíření personalizovaného obsahu není otázkou pokroku v generativní umělé inteligenci, ale spíše v umělé inteligenci obecně a algoritmech. Navíc to, že personalizované dezinformace mají lepší účinek, než ty nepersonalizované není prokázáno.

Podle studie Harvard Kennedy School³⁷ je očekávaný dopad generativní umělé inteligence na dezinformace je přehnaný, zapadá však do staré a široké kategorie morálních panik, spojených s novými technologiemi, které často vycházejí z představy, že lidé jsou naivní a lehce ovladatelní třetí osobou. Přehnané však nejsou generativní nekonsenzuální pornografické obsahy neproporcionálně poškozující ženy, podvody a krádeže identit. Autoři

³⁵ SIMON, Felix M., Sacha ALTAY a Hugo MERCIER. *Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown*. 2023, Harvard Kennedy School Misinformation Review [online]. Dostupné z: <https://misinfreview.hks.harvard.edu/article/misinformation-reloaded-fears-about-the-impact-of-generative-ai-on-misinformation-are-overblown> [cit. 2024-05-15]

³⁶ tamtéž

³⁷ tamtéž

textu však upozorňují, že tyto argumenty vycházejí z dat západních, bohatých a demokratických zemí. V rozvojových, chudých a nedemokratických zemích může být vliv generativní umělé inteligence na dezinformace odlišný a větší. Jak například ukazuje už zmíněná studie Freedom House³⁸, která monitoruje větší množství generativních dezinformací právě v rozvojových zemích.

TRH PRÁCE

Druhé, často opakované téma, spojené s rozšířením generativní umělé inteligence, je razantní proměna trhu práce, připomínající nástup parního stroje nebo internetu.

Podle posledního průzkumu generativní umělé inteligence na Standfordské univerzitě je generativní umělá inteligence používána na celém světě mezi 30 %–40 % organizací, což působí jako pozitivní informace, protože není velký rozdíl mezi USA a rozvojovými zeměmi. Dále ukazuje, že programátorům dokáže použití generativní umělé inteligence zrychlit práci až o 70 %³⁹, což je extrémně vysoké číslo, ale když si uvědomíme, že to je pro generativní umělou inteligenci nejbližší jazyk, s nejpřesnějším datasetem na celém internetu, tak to není zas tolik zarážející. Otázkou je, o jaká čísla se jedná u jiných profesí, například těch kreativních, avšak k tomu ještě není dostatek dat.

Také z výzkumu vyplývá, že generativní umělá inteligence horizontalizuje kvalifikaci u pracovníků. U méně kvalifikovaných pracovníků zefektivňuje jejich práci až o 43 %, zatímco u vysoce kvalifikovaných o 16.5 %, což snižuje rozdíl mezi více a méně kvalifikovanými pracovníky.⁴⁰

Erik Brynjolfsson přichází s termínem Turingovy pasti⁴¹, kdy stroje nejsou vytvářeny proto, aby rozšiřovaly schopnosti lidí, ale aby schopnosti lidí automatizovaly a nahrazovaly, což vyřazuje lidi z procesu vytváření hodnoty, a to vede ke ztrátě ekonomické a politické vyjednávací síly na straně pracujících a dává moc pouze do rukou těch, kteří stroje ovládají. Technologický pokrok není neutrální, některým skupinám neúměrně pomáhá a jiné neúměrně poškozuje.

³⁸ FUNK, Allie, Adrian SHAHBAZ a Kian VESTEINSSON. *The Repressive Power of Artificial Intelligence*. Freedom House [online]. Dostupné z: <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence> [cit. 2024-05-15]

³⁹ ARTIFICIAL INTELLIGENCE INDEX REPORT 2024. *Artificial Intelligence Index Report 2024, Chapter 4: Economy*. Online, report. [2024]. Dostupné z: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024_Chapter4.pdf [cit. 2024-05-15].

⁴⁰ tamtéž

⁴¹ BRYNJOLFSSON, Erik. *The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence*. Stanford Digital Economy Lab [online] [vid. 2024-05-11]. Dostupné z: <https://digitaleconomy.stanford.edu/news/the-turing-trap-the-promise-peril-of-human-like-artificial-intelligence/>

Matteo Pasquinelli v knize *The Eye of the Master* tvrdí, že představa, že umělá inteligence plně nahradí veškerá pracovní místa, je mylná: v takzvané ekonomice platform algoritmy ve skutečnosti nahrazují management a násobí nejistá pracovní místa.⁴² Tudiž práce bude pravděpodobně horizontálně distribuovaná, ale vedoucí pozice budou stejně zastoupeny platformou, která bude kontrolována korporací. Pasquinelli také píše, že mistr se nezabývá automatizací jednotlivých úkolů, ale řízení společenské dělby práce. Navzdory předpovědi Alana Turinga to byl mistr, nikoliv dělník, koho robot nahradil jako první.⁴³

⁴² PASQUINELLI, Matteo, *The Eye of the Master: A Social History of Artificial Intelligence*, 2023. London: Verso. ISBN 978-1-78873-006-8, poslední kapitola: Undoing the master algorithm.

⁴³ tamtéž

ZÁVĚR

Za pomoci geologie médií Jussi Parikky nahlížím fenomén umělé inteligence jako extenzi země, díky čemuž se mi daří vyvrátit pomyslnou výjimečnost této technologie, protože veškeré její schopnosti jsou hrubě materiálně podložené. Rozsah neudržitelosti a netransparentnosti procesů, kterými je tato technologie podmíněna, je enormní. Jelikož diskurz kolem umělé inteligence ovládá spíše strach ze zlomyslné obecné umělé inteligence, než její enviromentální kontext, tak bude zajímavé sledovat vývoj enviromentálních regulací a restrikcí s ní spojených, ať už bude budoucností zatím nejasně energeticky náročná kvantová výpočetní technika nebo energeticky soběstačná data centra poháněna vlastními jadernými elektrárnami.

Paralela s mechanickým turkem odkrývá pozadí umělé inteligence, které by nemohlo fungovat bez zakázkové ekonomiky a planetárně outsourcované námezdní kognitivní práce, stejně jako fyzicky náročné a nebezpečné práci v továrnách a dolech. Znovu vyvracím pomyslnou výjimečnost této technologie právě tím, jakým množstvím fyzické a kognitivní práce je tato technologie podložena. Stejně tíživá jako námezdní kognitivní práce dnes mi přijde i budoucnost, ve které se nejspíš už nacházíme, extrakce nejenom schopnosti rozlišit hydrant od autobusu, ale extrakce našich rozhodování při pohybu v prostoru, extrakce pohybu zorničky po displeji, v některých případech data použitá pro přesnější reklamu, v jiných k poskytnutí hypotéky nebo občanství.

Strojové učení je podmíněné statistickými postupy eugeniků z 19. století, což problematizuje příslib možnost etické nápravy této technologie, neboť celá její podstata je vybudována na myšlenkách diskriminace a segregace. Právě toto se jeví jako největší problém ve spojení s představou o umělé inteligenci jako neutrálním, nezaujatém činiteli.

Odhady dopadů dezinformací na západní společnost jsou na základě zmiňované studie přehnané z důvodů představy, že lidé jsou vůči novým technologiím velmi naivní a důvěřiví. Nemyslím si, že generativní umělá inteligence nějakým zásadním způsobem zvrátí vektor od probíhající epistemické krize. Nemyslím si ani, že se nějak zásadně změní trh práce. Současná automatizace práce nevede k osvobození člověka od práce, ale k restrukturalizaci práce, v tomto případě horizontalizaci znalostí, která bude utvrzena pod nadvládou korporacemi ovládanými algoritmy.

Charakteristickým prvkem, který by se měl stát synonymem termínu generativní umělé inteligence je extraktivismus. Žádná jiná technologie není spojená s tak rozsáhlou extrakcí v tolika odvětvích. Měli bychom ji nahlížet jako východisko extraktivistického přístupu

technologických korporací ze Silicon Valley k životnímu prostředí, lidským schopnostem, lidským datům a lidské historii. Tato technologie je revoluční pouze ve své schopnosti nahromadit extrémní množství extrahovaných hodnot pod jedno uživatelské rozhraní, stejně nerevoluční je i v případě kontextů sociálních a ekonomických, kdy bude pouze akcelerovat současné tendence v rámci pracovního trhu a informační krajiny.

Seznam použitých zdrojů

Literatura

CHUN, Wendy Hui Kyong, *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*, 2021. Cambridge: The MIT Press. ISBN 978-0-262-54852-6

CUBBIT, Sean. Lví podíl. In: *Objekt*, 2015. Praha: Kvalitář. ISBN 978-80-260-8639-0

PASQUINELLI, Matteo, *The Eye of the Master: A Social History of Artificial Intelligence*, 2023. London: Verso. ISBN 978-1-78873-006-8

SRNICEK, Nick. *Platform Capitalism*, 2016. Cambridge: Polity. ISBN 978-1-5095-0486-2

BRYNJOLFSSON, Erik. *The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence*. Stanford Digital Economy Lab Dostupné

z: <https://digitaleconomy.stanford.edu/news/the-turing-trap-the-promise-peril-of-human-like-artificial-intelligence/> [cit. 2024-05-15]

CRAWFORD, Kate a JOLER Vladan. *Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources*, 2018. AI Now Institute and Share Lab. Dostupné z: <https://anatomyof.ai>, [cit. 2024-05-15]

SIMON, Felix M., Sacha ALTAY a Hugo MERCIER. *Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown*. 2023, Harvard Kennedy School Misinformation Review [online]. Dostupné

z: <https://misinfreview.hks.harvard.edu/article/misinformation-reloaded-fears-about-the-impact-of-generative-ai-on-misinformation-are-overblown> [cit. 2024-05-15]

STEYERL, Hito. *Mean Images*, 2023. New Left Review. (140/141), 82–97. [2023] Dostupné z: <https://newleftreview.org/issues/ii140/articles/hito-steyerl-mean-images> [cit. 2024-05-15]

Ostatní

ARTIFICIAL INTELLIGENCE INDEX REPORT 2024. *Artificial intelligence Index Report 2024, Chapter 4: Economy*. Online, report. [2024]. Dostupné z: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024_Chapter4.pdf [cit. 2024-05-15].

CLIMATE ACTION AGAINST DISINFORMATIONS. *Artificial intelligence threats to climate change*. Online, report. [2024]. Dostupné z: https://foe.org/wp-content/uploads/2024/03/AI_Climate_Disinfo_v6_031224.pdf. [cit. 2024-05-15].

FUNK, Allie, Adrian SHAHBAZ a Kian VESTEINSSON. *The Repressive Power of Artificial Intelligence*. Freedom House. Online, report. [2023]. Dostupné z: <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence> [cit. 2024-05-15]

PERRIGO, Billy. *OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic*. TIME, 2023. Dostupné z: <https://time.com/6247678/openai-chatgpt-kenya-workers> [cit. 2024-05-15]