

Diplomová práce

2024

Bc. Vojtěch Polívka

Západočeská univerzita v Plzni
Fakulta filozofická
Katedra filozofie

Diplomová práce

Diskurz člověka a umělé inteligence, aneb procházka
tísňivým údolím

Vypracoval: Bc. Vojtěch Polívka

Vedoucí práce: Mgr. Jitka Paitlová Ph.D

Katedra filozofie

Fakulta filozofická Západočeské univerzity v Plzni

Plzeň 2024

Západočeská univerzita v Plzni

Fakulta filozofická

Katedra Filozofie

Studijní program Filozofie pro umělou inteligenci

Studijní obor Filozofie pro umělou inteligenci

Diplomová práce

Diskurz člověka a umělé inteligence, aneb procházka
tísňivým údolím

Vypracoval: Bc. Vojtěch Polívka

Vedoucí práce: Mgr. Jitka Paitlová Ph.D

Katedra filozofie

Fakulta filozofická Západočeské univerzity v Plzni

Plzeň 2024

Prohlašuji, že jsem diplomovou práci zpracoval samostatně a použil jen uvedené
prameny literatury.

.....

Plzeň, duben 2024

Vojtěch Polívka

Poděkování

Chtěl bych poděkovat Mgr. Jitce Paitlové PhD. za odborné rady a přátelský přístup, který napomohl k napsání této práce. V neposlední řadě bych rád poděkoval rodině a blízkým přátelům za jejich podporu.

Obsah

1 Úvod.....	3
2 Věda versus fikce.....	5
2.1 Historie kybernetiky a sci-fi.....	5
2.2 Optimismus v dystopii a pesimismus v Utopii.....	7
2.2.1 Problémy optimismu.....	7
2.2.2 Problém pesimismu.....	9
2.2.3 Problém vývoje.....	9
2.2.4 Problém fikce v realitě.....	10
3 Čapkovské proroctví.....	13
3.1 Nahrazení.....	13
3.2 Robot jako nástroj vykořisťování.....	16
3.3 Spojení sil.....	18
4 Tvor dvounohý a mechanický.....	18
4.1 (Ne)rozlišitelnost.....	21
4.2 Stávání se strojem a stávání se člověkem.....	24
4.3 Jestli to nebude láska.....	28
5 Matrix - život v simulaci.....	30
5.1 (Ne)rozlišitelnost znovu.....	31
5.2 Nová realita.....	34
6 Kdo je tady pán?.....	36
6.1 Velký bratr.....	38
6.2 Dohlížet a doporučovat.....	40
6.3 Linie úniku.....	42
7 Závěr.....	42
8 Resumé.....	45
9 Literatura.....	47
10 Obrázky.....	52

1 Úvod

Ve své diplomové práci analyzuji díla science fiction a myšlenky z těchto děl srovnávám se současným stavem technologie především v oblastech robotiky a umělé inteligence. Cílem práce je zodpovědět výzkumnou otázku: *Je současný stav technologie robotiky a umělé inteligence srovnatelný se sci-fi představami?* Může se zdát, že na tuto otázku existuje jednoduchá odpověď: *nikoliv*. Důvodů pro onu zápornou odpověď je mnoho a odrážejí se v nesrovnatelnosti fikce se skutečností. Nemůžeme však opomenout provázanost mezi vědeckými obory a vědeckou fikcí, které se navzájem ovlivňují. Science fiction taktéž nabízí informace pro veřejné diskurzy, dle kterých je následně nahlíženo na produkty technologického pokroku.

Vždy, když někomu sdělím, že oborem mého studia je Filozofie pro umělou inteligenci, tak následují otázky, které rezonují obavami: *Ovládne nás umělá inteligence? Zotročí nás?* Proč bychom se měli zabývat takovými otázkami, když je fungování současných systémů umělé inteligence zejména velkých jazykových modelů (jako je *ChatGPT*) prezentovaný jako velký úspěch... Předpokládám, že obavy, které zaznívají skrze tyto otázky vznikají z fikčních světů science fiction. Od začátku zlaté éry sci-fi literatury se setkáváme s narativy ohledně toho, že tvorba umělé inteligence, robotů či celkově nové technologie přináší ohrožení lidstva. V *R.U.R.* se setkáváme s nahrazením člověka v pracovní činnosti a následnou vzpourou robotů. Ve filmu *Terminátor* jsme svědky války mezi člověkem a strojem. Nastávající boom umělé inteligence, kdy se z tohoto slovního spojení stává buzzword, je něčím, s čím jsme se mohli setkat pouze skrze tato díla. Vytvoření těchto někdy až katastrofických scénářů nás pak bezesporu dostává do analogického uvažování i o současném stavu umělé inteligence a robotiky.

V této práci se proto budu snažit o porovnání vybraných narativů z fikčních světů sci-fi literatury a filmů se současným diskurzem o umělé inteligenci ve společnosti. Sci-fi díla poskytnou zdroje oblastí, kterými se budu zabývat. Kritérium výběru děl je úzce spjata se specifickými fenomény, které budu analyzovat. *R.U.R.* zprostředkovává oblast práce, kde je artikulováno nahrazení člověka strojem, což je pozorované například skrze fenomén automatizace. *Blade Runner – Sní androidi o elektrických ovečkách?* (dále už jen jako *Blade Runner*) prostírá oblast, v níž se řeší problém nerozlišitelnosti. V románu dochází k tomu, že androidi jsou téměř dokonalou kopií člověka. Na první pohled, tak nemůžeme určit, jestli se jedná o člověka nebo androida. Film *Matrix* poskytuje dvě oblasti problémů. První je ústřední téma filmu, což je život v simulaci. Druhým problémem, který se objevuje téměř ve všech dílech sci-fi literatury, je proměna mocenského vztahu mezi člověkem (subjektem) a produktem technologického pokroku (objektem). K analýze každého z témat využívám zejména myšlenky vycházející z postmoderní filosofie od autorů jako je Jean Baudrillard, Shoshana Zuboff, Michel Foucault, Gilles Deleuze a Félix Guattari.

Práce je strukturovaná do pěti kapitol. V první kapitole se zaměřuji na obecnou podobu vývoje oboru kybernetiky a umělé inteligence. Srovnávám dvě zlaté éry, které se týkají umělé inteligence a sci-fi, čímž chci poukázat na jistou provázanost vědeckých oborů s vědeckou fikcí. Poukazuji i na jejich důležitost při vytváření diskurzů. Pro zjednodušení můžeme říct, že vznikají dva tábory: *techno-optimismus* a *techno-pesimismus*. Tato dvě pojetí vznikají na základě vědění, které může být čerpáno z mnoha zdrojů (optimismus vědního oboru, sci-fi, média etc.). Oba přístupy nabízí vždy na problematiku pouze jednu perspektivu. Důležitým obsahem této kapitoly je taktéž nahlížení na technicko-ekonomický vývoj, kdy jako teoretický podklad používám myšlenky sociologa Ulricha Becka z knihy *Riziková společnost*. Technicko-ekonomický vývoj pojímá rizika, která sám vytváří. Podobný princip se objevuje i u využití umělé inteligence. Příklady uvádím napříč celou prací.

Kapitola druhá se věnuje Čapkovu *R.U.R.*, ze kterého si propůjčuje téma *nahrazení*. V tomto dramatu se můžeme setkat s dvojím typem nahrazení. První je nahrazení člověka v pracovní činnosti, což je považováno za velký úspěch robotiky. Druhým typem nahrazení je nahrazení absolutní, kdy můžeme pojímat stroj jako novou evoluční fázi. V této kapitole se věnuji především prvnímu typu, který dávám do souvislosti s probíraným tématem *technologické nezaměstnanosti*, jež má být zapříčiněna vlivem *automatizace*. Zde se snažím odhalit další jevy vyvstávající z *automatizace* jako je například proces *heteromatizace*. Kapitola se tak zaměřuje na dělbu práce mezi člověkem a strojem a možný budoucí vývoj.

Hlavním tématem třetí kapitoly je problematika, která vyvstává z knihy Philipa K. Dicka *Blade Runner* v souvislosti s tendencemi robotiky o vytvoření robota vizuálně podobného člověku. Zde se věnuji konceptu *uncanny valley* (do češtiny překládané jako *tísnivé údolí*) předložené Masahirem Morim v 70. letech minulého století, z něhož přecházím do problematiky nerozlišitelnosti, která je znatelná v práci Philipa K. Dicka. Koncept nerozlišitelnosti přináší jisté potíže v oddělování dvou kategorií, které častokrát bývají originálem a reprezentací či kopií. Produkt robota je v určitých výstupech robotiky onou reprezentací člověka. Zlepšování této reprezentace nás pak přenáší i přes tísnivé údolí, ale přivádí k nové problematice předvedené v Dickově románu.

Poslední dvě kapitoly se nesou v duchu témat, která jsou zastřešena pod filmem *Matrix*. Čtvrtá kapitola se věnuje hlavnímu motivu filmu, jímž je „simulace“. V rámci filmu je simulace počítačem vytvořený virtuální svět. Toto pojetí simulace převládá i v současných diskurzích, kdy jedinou věcí srovnatelnou s matrixem by byla virtuální realita. Od tohoto pojetí se distancuji a snažím se na problém nahlížet perspektivou, která o simulaci nepřemýšlí jako o virtuálním počítačovém prostoru, ale jako o způsobu nahlížení na svět kolem nás. K přiblížení používám koncept Platónovy jeskyně a Baudrillardovo pojetí simulace. Pátá kapitola se vrací k narativu, který je ústřední myšlenkou techno-pesimismu. V průběhu práce je tato myšlenka nazývána jako *Terminátor narativ*. Tento narativ se zaměřuje na vývoj technologií, které se vymykají naší kontrole. Má tak přinést varování o narušení mocenské struktury mezi člověkem a technologií. Do konceptu *Terminátor narativu* bychom mohli zařadit i absolutní nahrazení v Čapkovském *R.U.R.* Cílem kapitoly je analýza technologií, které jsou využívány k upevnění mocenských struktur ve společnosti. Dále se zabývá proměnou vztahů mezi subjekty a objekty, která je vystavena na alegorii filmu *Matrix*, kde stroje využívají člověka jako zdroj energie.

Každá z kapitol se věnuje jistému výseku témat, která spolu souvisí ve smyslu soužití člověka s nově vznikajícími technologiemi. Koncepty, které jsou používané k analýze, jsou popsány v daných kapitolách. Jediným konceptem, který se objevuje téměř ve všech kapitolách a není v textu vysvětlený, je *reteritorializace* a s ní spojená *deteritorializace*. Koncepty pochází od filosofů Gillesse Deleuze a Félixu Guattariho, podle nichž je teritorium vyhrazené území, na kterém se daná věc vyskytuje. Může se jednat o vymezení kategorie například tím, že je stroj určen k jedné činnosti a podobně. Procesem *deteritorializace* jsou narušeny hranice daného teritoria, načež procesem *reteritorializace* jsou tyto hranice opět utvořené v nové podobě. Dle filosofů tímto způsobem vzniká něco nového.

Název práce si propůjčuje jméno hypotézy *tísnivého údolí* (*uncanny valley*) jako aluzi na ona vznikající rizika, která nás dostávají do jisté tísně či obav. Jednotlivé výseky témat, které jsou probírané v jednotlivých kapitolách, fungují samy o sobě, a zároveň se tyto myšlenky snažím propojovat napříč celou prací. Takto má být vytvořena celistvost práce zaměřující se na soužití člověka s novými technologiemi a vliv děl science fiction na tento vztah.

2 Věda versus fikce

Velké jazykové modely či v angličtině Large language models (dále jen LLM) infiltrovaly veřejný diskurz jako záhadný stroj, který dokáže nejen reagovat na vaše zprávy, ale taktéž vytvářet texty v podobě příběhů, esejí, básní a písní. Stroj, který by dokázal tak dobře pracovat s jazykem, neměl na trhu obdoby. Jejich fungování na nás působilo jakoby se sci-fi stalo realitou.

V této kapitole se pokusím nastínit problémy technologického vývoje a s ním spojené vznikání diskurzů, které je do určité míry provázáno s představami vycházejícími ze science fiction.

2.1 Historie kybernetiky a sci-fi

Pokud se chceme dívat na začátek vývoje umělé inteligence a robotiky, pak se musíme podívat na počátek oboru kybernetiky, který vystává z popela Druhé světové války, při které byla snaha o technologický pokrok, který by přinesl výhodu nad nepřítelem. Americká NACA (předchůdce NASA) se v čele s Vannevar Bushem zabývala protiletectkou obranou a v Anglii vznikla NDRC (National Defense Research Committee) s cílem akademických výzkumů, které se týkaly vojenských problémů viditelných v roce 1940 při bombardování Londýna.

Inženýři pověřeni tímto velice důležitým úkolem došli k určitému druhu přemýšlení, který se následně stal důležitým základem budoucího vývoje technologií – o člověku přemýšleli jako o stroji. Jako názorný příklad používají „střílení kachen“. Lovce zpracuje vizuální informaci o letící kachně. Mozek na základě této informace pošle signál ruce, aby zamířila puškou na požadovanou vzdálenost, kde bude kachna zasažena. Tento proces pak končí stisknutím spouště a v ideálním případě sestřelenou kachnou. O lovcích se přemýšlí jako

o síti, která přijímá, zpracovává a jedná. Tím se ukazuje na, co se při konstrukci soustředit. Hlavními prvky je tak: přijetí vizuální informace (oči), výpočet (mozek) a pohyb (ruka s puškou). Tyto myšlenkové výplody vzniklé za války následně přebírá Norbert Wiener a ve své knize *Cybernetics; or, Control and Communication in the Animal and the Machine* vydané v roce 1948 předkládá předzvěst druhé průmyslové revoluce. V první revoluci byly nahrazeny lidské svaly a v této revoluci má být nahrazen lidský mozek. Toto načasování bylo perfektní, jelikož se tyto myšlenky na konci desetiletí začaly přesouvat z válečných vřav do průmyslové sféry a domácností (Rid 2016).

Zatímco se návrhy technologických strojů uzpůsobených k ničení přeměnily v novou naději pro lidstvo, tak v pozadí stále běžela literatura sci-fi, která používá rizika technologie jako zápletku pro své příběhy. Už v roce 1920 přišlo skrze drama Karla Čapka *R.U.R.* pojmenování robot ze slova robota, které označuje stroj podobný člověku, který má osvobodit lidstvo od práce. Tito roboti se ovšem vzeprou svým stvořitelům a lidstvo vyhubí (Roberts 2006, s. 168). Na jedné straně tak stojí optimistické představy o technologickém pokroku ve sféře vědy a na straně druhé varování před možnou zkázou ve sféře fikce.

Žánr sci-fi se stal provázaný s technologickým pokrokem. Na jedné straně sci-fi do jisté míry predikuje technologický pokrok jakožto zdroj inspirace pro možné vynálezy. Na druhé straně je ovlivněno technologickým vývojem dané doby. Vývoj technologie je inspirací pro sci-fi knihy. Tento jev je možný spatřovat napříč sci-fi literaturou, která vzniká v průběhu 19. století, například již v dílech Julese Verna, se odráží dobový vědecký a technologický pokrok, a zároveň předkládá představy o možných technologiích, jež se až po dlouhých letech zdají být uskutečnitelné. Z představ se přesouvají do reality. Mezi roky 1940 až 1960 vstupujeme do zlatého věku sci-fi, kdy září například díla Isaaca Asimova s jeho třemi zákony pro roboty (Roberts 2006). Tento koncept tří zákonů rezonuje až do dnešních debat o etických principech v umělé inteligenci (Wooldridge 2021, s. 140). Zároveň v průběhu padesátých let přichází i zlatý věk oboru umělé inteligence, a to díky projektům letních škol Johna McCarthyho. Jeho hlavním cílem bylo sjednotit výzkumníky se stejným zájmem a přinést, tak signifikantní pokrok v oblasti umělé inteligence. Tento plán vyšel pouze částečně. K žádnému velkému pokroku nedošlo, ale díky sjednocení výzkumníků se stále držely optimistické představy o vývoji umělé inteligence. Ze snah o naplnění těchto optimistických představ vznikly první systémy a roboti, kteří byli ovšem stále velmi daleko od představ, jež se objevují ve sci-fi literatuře.

Ve Stanfordu vznikl systém SHDLRU, který demonstroval dvě důležité vlastnosti pro AI: řešení problémů a porozumění přirozenému jazyku. Tento systém byl schopný manipulovat s barevnými objekty v simulovaném prostředí dle příkazů, které mu uživatel zadával. Dialog se systémem byl vyvinutější než například systém Eliza,¹ ale to jen díky tomu, že se dialog vždy pohyboval jen v omezeném scénáři. Ze simulovaného prostředí se na Stanfordu dostali i k robotům s fyzickou podobou a schopností operovat v reálném prostředí. SHAKEY byl schopen vnímat své prostředí, přijímat úkoly od uživatelů, vymyslet vhodné plány, jak těchto úkolů dosáhnout a ujišťovat se, že vše probíhá v pořádku. Počátkem 70. let se objevila frustrace z relativně malého postupu při řešení základních problémů umělé inteligence a na obor začal padat stín, například i v podobě *Lighthill Report*

¹ Terapeutický chatbot, který sloužil jako prezentace stavu dobového pokroku interakce člověka se strojem, který dával předpřipravené odpovědi, do kterých se pouze doplňovalo klíčové slovo z uživatelského vstupu.

, v němž je kritizován přehnaný optimismus při vývoji umělé inteligence, který se nepodařilo naplnit. Ve vývoji AI, tak nastala první zima (označováno jako AI winters), která byla následně vystřídána dalším boomem a další zimou, přičemž tyto výkyvy trvaly až do roku 2012, kdy přišel další boom nastartovaný systémem AlexNet na klasifikaci obrázků. Tento boom podpořený LLM jako je ChatGPT, trvá dodnes (Wooldridge 2021).

Umělá inteligence tím vstoupila do mainstreamu a stala se středem debat, a to nejen v akademickém prostředí, ale i v neformálních oblastech. Produkty oboru umělé inteligence začaly vystupovat ze stínu a jeví se jako něco zcela nového a neprobádaného. Skrze onu záhadnost je vytvářena provázanost se známými fikčními světy žánru science fiction, jelikož narativy těchto děl působí jako proroctví o možné technologické zkáze. Sci-fi ustanovilo mýtus o konfliktu člověka a stroje, který nyní ožívá společně se zájmem o vědecké obory umělé inteligence a robotiky. Z fikčních světů je čerpáno dvěma prameny: První pramen je optimistický, protože inspiraci pro cíle, kterých by jmenované vědecké obory měly dosahovat. Druhý nese pesimistický nádech v podobě narativů označujících technologický vývoj jako krok směrem do propasti a potenciální záhubě lidstva.

2.2 Optimismus v dystopii a pesimismus v utopii

2.2.1 Problémy optimismu

Vývoj oboru umělé inteligence a robotiky nebyl tak přímočarý, jak by se na první pohled mohlo zdát. Výzkumníci se na této cestě při naplňování svých přehnaných optimistických představ setkali s několika pády způsobenými kumulujícími se problémy s přístupy tvorby umělé inteligence. Tyto výkyvy nastávaly ve zmiňovaných *AI winters* a boomech. Na takový vývoj vědecké disciplíny upozorňuje Thomas Kuhn v knize *Struktura Vědeckých Revolucí* (1997). Láska k oboru se vždy vrací po jeho velkém úspěchu. V našem případě je vývoj vědy v nejbližším vztahu i s vývojem umělé inteligence či robotů, jelikož tyto stroje jsou právě produktem vědního oboru. V druhé polovině devadesátých let dochází k něčemu, co novinové titulky popisují jako depresivní den pro lidstvo. Počítač *DeepBlue* porazil nejlepšího hráče v šachu Garry Kasparova. Jednalo se ovšem jen o další velko-malý krůček vývoje umělé inteligence. Sice máme počítač, který dokáže hrát šachy lépe než kterýkoliv člověk, ale stále je to úspěch pouze v malé oblasti. Veškeré tyto malé krůčky vývoje pak vedou k dalším možná až přehnaně optimistickým představám o vývoji umělé inteligence, jak jsme mohli pozorovat v padesátých letech.

Přehnané optimistické představy se začínají přelívat přes svůj okraj a stávají se pesimistickými, co se týče osudu lidstva. Venor Vinge v průběhu 90. let se v několika esejích zabývá *singularitou*. *Singularita* je koncept související s intelligenční explozí (Eysenck, Eysenck 2021). V podstatě se jedná o myšlenku, která tvrdí, že pokud stroj bude inteligentnější než kterýkoliv člověk ve všech ohledech, a jelikož tento systém umělé inteligence byl navrhnutý člověkem, pak dokáže navrhnout i lepší systém umělé inteligence, než je on sám. To má ve výsledku přinést *superinteligenci* (Good 1966, s. 33).

Superintelligence je ve většině spojována se systémy *Artificial general intelligence* v češtině označováno jako obecná umělá inteligence (dále jen AGI), která narozdíl od *narrow* či *weak artificial intelligence*, překonává člověka ve všech inteligentních úkonech. Mluvíme tedy o takových systémech umělé inteligence, které by člověka překonaly nejen v malé oblasti jako *DeepBlue* překonal Kasparova, ale na celkové úrovni, tedy by byla chytřejší než jakýkoliv člověk ve všech ohledech. To by mělo za výsledek, že by se nastolila dominance umělé inteligence nad člověkem (Eysenck, Eysenck 2021). Termín *singularity* dále propaguje Ray Kurzweil v knize *Singularity Is Near* z roku 2005. Základní myšlenkou knihy je, že tempo vývoje technologií se exponenciálně zvyšuje a tím dojde k tomu, že umělá inteligence překoná tu lidskou. Kurzweil vychází z *Moorova zákona*, který byl formulován Georgem Moorem, spoluzakladatelem firmy *Intel*. Tento zákon predikuje vývoj procesorů a říká, že počet tranzistorů se přibližně každých osmnáct měsíců zdvojnásobí. Moorův zákon platil téměř padesát let až do roku 2010, kdy vývoj procesorů narazil na své limity (Wooldridge 2021, s. 136). Problém nastává v uzavřené představě vývoje při jednom paradigmatu, který byl znatelný v rámci zlatého věku umělé inteligence. Moorův zákon nepředpokládá lineární vývoj, ale vývoj exponenciální, který vytváří predikci zrychlujícího se vývoje. Podobné uvažování se stává integrální součástí optimistického pohledu. Tento vývoj nemá pouze exponenciální růst, jelikož při něm nastávají změny. V rámci vědní disciplíny to může být právě přesun k jinému paradigmatu (Kuhn 1997). Stává se tak chybné predikovat vývoj určité technologie na základě přímého růstu. Jedná se spíše o *hype cycle*, při kterém vzroste očekávání od nové technologie, které následně opadá a normalizuje se do té doby, dokud nedojde k dalšímu zásadnímu objevu, který tento cyklus znovu nastartuje (Johnson 2015). Na oboru umělé inteligence je tento cyklus pozorovatelný skrze AI winters.

Otázkou však zůstává, kam nás technologický pokrok zanesou? Na jedné straně stojí *techno-optimismus*, který má pozitivní pohled na technologický vývoj. Techno-optimismus není jen jeden konkrétní pohled na věc, ale spíše shluk souvisejících názorů, které předpokládají to, že technologie hraje klíčovou roli při zajišťování dobrého nad špatným. Vůči čemu, ale tento optimismus vzniká? Co je tou technologií? Eric Schatzberg tvrdí, že existuje jedna dominantní myšlenková škola, která pojímá technologii jako souhrn nástrojů k řešení problémů a dosahování cílů (Danaher 2022). Například Marc Andreessen ve svém manifestu vnímá technologický pokrok jako tíhnutí k utopii. Celý koncept utopie byl v minulém století dosti pošramocený kvůli socialismu a komunismu. Autor manifestu se s tím vypořádává tak, že utopii pojímá jako omezenou vizi. Technologický vývoj dle omezené vize nevede k utopii, ale blíží se k ní, jelikož je díky němu život lepší a lepší (Andreessen 2023). Toto pojetí omezené vize zároveň předkládá další zajímavou představu, a tou je možnost nekonečného růstu. Neustálý vývoj nové a nové technologie, tak činí život člověka lepší a lepší a nikdy se nemůžeme dočkat perfektního konce. Celé pojetí utopie, tak samo o sobě vytváří mýtus, který staví do popředí spasitele, jímž jsou v tomto případě technologie, kdy se odráží i jejich charakter vývoje. Nepřichází se s dokonalostí, ale s něčím, co bude stále lepší a lepší. Je zde předpokládán lineární vývoj, který stále stoupá, a dokonce se ani nikdy nezastaví. Využívané technologie se mohou měnit stejně jako se mění vědecká paradigmatata, ale neustále se aproximujeme k nedosažitelné dokonalosti. S technologií je dobře, ale vždycky může být lépe. Tato otevřenost nabízí nekonečné spektrum nových

tržních možností, které tak nabízí prostor nekonečnému exponenciálnímu růstu bez možnosti zastavení, a tedy docílení utopie.

2.2.2 Problém pesimismu

Nepřekvapivě na druhé straně tohoto spektra stojí *techno-pesimismus*, který předpokládá to, že dopad technologie na člověka je nepříznivý (Königs 2022). Techno-pesimismus se projevil v radikální formě v 90. letech při bombových útocích Theodora Kaczynského, které následně inspirovaly hnutí anarchoprimitivismu a později další radikální skupiny včetně norského teroristy Anderse Breivika. Popularitě se mu dostalo hlavně kvůli jeho manifestu, který byl z donucení publikován v *The New York Times*. Hlavní myšlenkou manifestu bylo varování před technologickým pokrokem, který nevyhnutelně skončí katastrofou (Fleming 2022). Jakýkoliv boom kolem nové technologie s sebou přináší i negativní reakce. Techno-pesimistický přístup není žádnou novinkou. Už na počátku 19. století se objevila volně organizovaná skupina zvaná luddité, která měla ničit automatizované stroje a psát výhružné dopisy vlastníkům továren. Činili tak hlavně z důvodů obav ze ztráty práce. O téměř dvě století později se tyto myšlenky vrací v podobě neo-luddismu, který opět přichází s boomem nových technologií. Toto hnutí je více spojeno s ideologickými základy, které předpokládají, že technologie má negativní dopad na společnost. Idea o škodlivé technologii se tak rozšiřuje mimo pole zaměstnání a dostává se do daleko širšího měřítka (Jones 2006).

Tyto často radikální postoje vůči technologiím se, dle mého názoru, zdají být latentními riziky plynoucími z technicko-ekonomického vývoje. Svou radikálností ovšem paradoxně tomuto vývoji napomáhají. Jakákoliv kritika ztrácí svou hodnotu, když pochází od teroristy. Tyto ohavné činy zprostředkovávají zdiskreditování názoru. To vede k dichotomickému rozdělení, které staví technologický vývoj na pozici racionálního procesu a jeho kritiky do pozice iracionálních bláznů. Odráží se zde vlastnost technicko-ekonomického vývoje, který z těchto latentních rizik vytváří další možnost svého růstu. Například může nabízet technologie, které se tyto rizika snaží řešit, čímž vytváří nové tržní šance (Beck 2011, 61).

2.2.3 Problém vývoje

Blížíme se k utopii nebo se blížíme spíše záhubě? Prostřednictvím umělé inteligence jako vědeckého oboru jsme si ukázali, že nemůžeme přemýšlet o vývoji oboru ve smyslu jedné linie neustálého vývoje, jelikož dochází k vědeckým revolucím, kdy je pozměněno paradigma, a tedy celková změna nahlížení na danou problematiku. Dle Kuhna zde ovšem stále nastává jistý pokrok, který u této technologie spatřujeme v podobě *hype cycle*. Můžeme tento aspekt revolucí přinášejících pokrok vztáhnout i na celý komplexní svět?

Díváme-li se na technicko-ekonomický vývoj, pak můžeme spadnout do představ techno-optimistů prostřednictvím toho, že technologie nám zprostředkovává lepší životní situaci. Jsme zaplaveni zvýšením efektivity a zdá se, že opravdu směřujeme k oné nedosažitelné utopii. Tento vývoj ovšem nemůžeme brát pouze pozitivně, jelikož jeho součástí je vždy produkce rizik. Rizika jsou v raném stádiu legitimizována jako latentní

vedlejší účinky, ale později jsou přetvářena do problému s centrálním významem v politických a sociálních konfrontacích. V rozvinuté tržní společnosti se z rizik stávají i tržní šance, což přispívá k oscilaci mezi jejich odhalováním a zastíráním. Ztrácí se tedy rozdíl v tom, co je problémem a co je řešením. Technicko-ekonomický pokrok se totiž stává jak producentem rizik, tak i jejich jediným řešitelem (Beck 2011, 61). Podobné tendence se ukazují i na problémech, které se stávají středem debat o umělé inteligenci. Ony příklady se budu snažit nastínit v dalších kapitolách této práce.

Měli bychom tedy přestat šilet z obrovského boomu umělé inteligence? Budou veškerá rizika eliminována v dalším vývoji, nebo se jen zaplavíme dalšími riziky? Co si máme vzít z těch milionů hlasů, které mluví o umělé inteligenci? Při formulování modalit vypovídání je důležité, kdo o dané věci vypovídá, a také o tom, jaký má status (Foucault 2002, s. 80). Techno-optimisté jsou především tvořeni těmi, kdo mají z technologického pokroku užitek. Zmiňovaný Marc Andreessen je například softwarový inženýr, podnikatel a investor. Dalšími techno-optimisty mohou být například Jeff Bezos, Elon Musk či Mark Zuckerberg. Většina z těchto zmiňovaných zastává důležité posty v oblastech, které zprostředkovávají technicko-ekonomický vývoj. Zastávání těchto důležitých pozic snad nemůže znamenat nic jiného než radost z pokroku technologie, jelikož jsou to právě tyto osoby, které z technicko-ekonomického vývoje těží. V rámci ředitelských pozic je tedy technologie nástrojem, který vytváří větší ekonomickou aktivitu, a tak zvyšuje příjmy. Zároveň je taktéž produktem či pomocníkem pro nabízení lepších služeb. Všechny z těchto velkých firem vychází z narativu, který z technologie vytváří jakousi potřebu. Potřebu technologii nakupovat, využívat a tak dále. Stávají se tak ústředním motorem vývoje technologie a její produkce k masám. Ve většině případů jsou tak formované modalities vypovídání v pozitivním nádechu a tvoří z objektů svého diskurzu (ať už jde o umělou inteligenci, či mobilní telefon) nový zázrak vedoucí k dalšímu vývoji, což znamená opět lepší svět. Zde nastává i ona důležitost toho, že vývoj nemá určitý cíl, který by se měl naplnit. To dává možnost vývoji technologie postavené na poptávce, která je proměnlivá. Bez ohraničení, tedy bez určení stavu utopie je tento vývoj v podstatě nekonečný. Rizika produkována při tomto vývoji pak tvoří další poptávku, která může být opět utěšena pomocí dalších technologických produktů.

2.2.4 Problém fikce v realitě

Odvracenou tváří technologického vývoje týkající se především oblasti umělé inteligence jsou *fikční světy*. S technologií umělé inteligence se totiž lidé setkali skrze pera spisovatelů už v době, kdy umělá inteligence existovala jako pouhá sci-fi představa. Čapkovi roboti jsou záchranou před prací, ale zároveň nebezpečím zkázy. Asimovovi roboti dodržují tři zákony robotiky, ale vždy nastane chyba, kvůli které zákon nemohou dodržet, čímž se ukazuje nemožnost plné důvěry v technologii. U díla Philipa K. Dicka musí být androidi loveni, jelikož se vzpírají svému stvořiteli. Přejdeme-li od písmen k obrazům, pak si vzpomeneme například na film *2001: Space Odyssey* od Stanleyho Kubricka z roku 1968, kde je posádka zrazena umělou inteligencí v podobě počítače. Ve filmu *Terminátor* z roku 1984 od Jamese Camerona, lidé vedou válku proti strojům. Na konci 90. let minulého století se na velkém plátně zase setkáváme se snímkem *Matrix*, který ukazuje, jak je hlavní hrdina

uvězněni společně s celým lidstvem v simulaci, do které je uzavřely stroje, aby z něj mohly čerpat energii k jejich přežití (Roberts 2006).

Vývoj umělé inteligence se při každém dalším úspěchu bije do prsou a prezentuje je jako přetváření sci-fi představ do skutečnosti. Zde nastává jistý paradox. Technologický pokrok je úspěchem, který nás přibližuje k věcem, jak jsou prezentovány v dílech sci-fi. Díla sci-fi nepředkládají utopickou vizi světa, kde technologie řeší veškeré problémy lidstva, ale spíše technologii staví do role producenta rizik. Frank Rosenblatt, autor myšlenky perceptronu, svou ideu, respektive to, co by z ní mohlo vzniknout, sám označuje jako přenesení sci-fi představ do reality (Lefkowitz 2019). Srovnání pokroku technologie se sci-fi, pak působí jako dokazování dokonalosti daného vědního oboru. Je to prezentace moci a schopnosti přenést pouhou představu do reality. Tyto představy jsou v rámci sci-fi děl spojeny i s určitými obavami z dystopických scénářů. Obavy spojené s vývojem technologie se pak obrací v pochybnosti v určité projekty s umělou inteligencí. Například v roce 2017 se častokrát objevovalo hlášení ohledně toho, že Facebook ukončil vývoj dvou systémů umělé inteligence, které začaly komunikovat vymyšleným jazykem. Média z tohoto oznámení vytvořila sci-fi zápletku odehrávající se v realitě, jelikož bylo zastavení vývoje těchto systémů rámované, jakožto obavy z toho, že vývojářský tým ztratí nad umělou inteligencí kontrolu, což mělo být hlavním důvodem pro ukončení těchto projektů. Ve skutečnosti byl tento experiment běžnou praxí a hlavně byl naprosto neškodný. V inspiraci ze zápletek sci-fi příběhů se dostáváme do strachu z toho, že vyvíjíme něco, co nás nakonec přemůže. Takovýto narativ můžeme nazývat *Terminátor narativem* (Wooldridge 2021, s. 134). Tento narativ se těžko překonává, jelikož je nedílnou součástí diskurzu, který je neustále krmený chytlavými titulky na novinových serverech o nadvládě umělé inteligence, jež se přiživují v rámci internetové ekonomiky na současném boomu.

Problém nastává v případě, že strach z umělé inteligence se projevuje i u techno-optimistů. Vzpomeňme na otevřený dopis: *Pause Giant AI Experiments*. Tento dopis se zaměřoval hlavně na problémy, které budeme i dále rozebírat. Hlavním tématem bylo ustanovit pozitivní efekt umělé inteligence, a tak předcházet rizikům, která by mohla její bezmyšlenkovitá implementace páchat. Pod tento otevřený dopis se podepsala známá jména pohybující se v oboru umělé inteligence, jako je Elon Musk či Steve Wozniak (*Pause Giant AI Experiments: An Open Letter* 2023). Podpisy takovýchto důležitých jmen technologického vývoje pak navzdory původnímu záměru podporují Terminátor narativ, a tedy to, že s dalším vývojem umělé inteligence si lidstvo samo pod sebou podřezává větev. Není tak divu, že se objevují i názory na úplné zastavení vývoje umělé inteligence (Yudkowsky 2023).

Jak singularita, tak Terminátor narativ počítají s tím, že k jejich naplnění dojde až po vytvoření AGI. Obavy z tohoto druhu umělé inteligence nejsou žádným překvapením. Už dokonce I. J. Good upozorňoval na to, že takto superinteligentní stroj by musel být tak hodný a sdělit nám, jak jej ovládat (Good 1966, s. 33). Vynález této AGI by pak vedl k inteligenční explozi, jelikož by byl schopný sebe-vylepšení a vytváření inteligentnějších strojů, než je on sám. Jak jsme si již ukázali, tak tyto lineární predikce dříve či později narazí jako v případě vývoje umělé inteligence v logicko-symbolickém paradigmatu či u Moorova zákona. S úspěchy LLM v čele s nejnovější verzí GPT-4 se opět vrací tyto optimistické představy o AGI, která zároveň představuje největší strach. Důležité je si ovšem pamatovat, že veškeré

překonání člověka nastalo jen v určitých specifických činnostech stejně jako byla porážka Garryho Kasparova a současné systémy umělé inteligence jsou stále velice daleko od toho, aby dokázali naplnit kategorii obecné umělé inteligence. Podobný jev nastává i u velkých jazykových modelů, které sice mohou překonávat člověka ve specifické oblasti, ale stále jsou nedostatečným nástrojem k vytvoření zmiňované AGI. Zároveň nemůžeme ani opomenout, že tvorba umělé inteligence se ve specifických činnostech zlepšuje a pokrývá větší škálu těchto činností (Sublime 2024).

Představy sci-fi zůstávají pořád nějak provázané s realitou, a to tím způsobem, že se ovlivňují navzájem. Sci-fi se stává zdrojem inspirace pro vědce, kteří se snaží proměnit onu fikci v realitu. Zároveň je sci-fi inspirováno samotným vědeckým pokrokem či stavem společnosti. Ve veřejném diskurzu ovšem vzniká zmatek v podobě konfliktu pozitivních a negativních modalit, které přicházejí z různých směrů od technologických magnátů, novinářů až po literaturu a filmy. K tomu všemu ještě příliš nepomáhá přeměna označení umělé inteligence jakožto buzzwordu, což mělo za následek zvýšení jeho výskytu a následné vyprázdňení či lépe řečeno přeplnění. Zkratka AI se nyní objevuje v různých kontextech a jediné, co za ní zůstává, je ona mystická aura vyvstávající ze sci-fi představ. Mystičnost onoho slovního spojení vzniká skrze roztříštěnost. Umělá inteligence dokáže rozlišovat obrázky, generovat text, doporučovat videa na *YouTube* navádět auta nebo drony. Z pravidla se mluví ovšem o oddělených systémech zapadající do jedné kategorie.

I přestože zde stále není žádný systém AGI, tak to neznamená, že veškeré naše obavy zrozené z příběhů science-fiction, bychom měli hodit za hlavu. Využití a implementace umělé inteligence má v sobě stále obsažená určitá rizika. Zavádějícím procesem by bylo doslovně vykládat představy ze sci-fi děl. Science fiction v mnoha ohledech působí jako metafora, která odráží stav autorovy současnosti. Jean Baudrillard v knize *System of Objects* říká: *“Mýlili bychom se, kdybychom považovali science fiction za útěk z každodenní reality: právě naopak, je to extrapolace iracionálních tendencí této reality skrze svobodné uplatnění vypravěčského vynálezu. Přestože je neocenitelným svědkem civilizace objektu právě proto, že vyzdvihuje některé její aspekty, science fiction nemá naprosto žádnou prorockou hodnotu.”* (Baudrillard 1996, s. 119). Využití sci-fi děl na predikování budoucnosti je chybným krokem. Jejich metaforická povaha nám nedovoluje vytvořit přímou identitu mezi jevem fikčním a jevem skutečným. Je zapotřebí vzdálenější pohled, který nabídne oblast určitého tématu a možné problémy, které se mohou v této oblasti objevit. Nikdy však nenastanou přesně tak, jak jsou popsány v knize. Žádný ze spisovatelů a spisovatelek sci-fi nebyl věstec ani věstkyně.

Dokud nebude vytvořen systém AGI, pak zůstáváme v mocenské struktuře člověk-pán a stroj-otrok. Subjekt vládnoucí nad objektem. Díla sci-fi poukazují na narušení tohoto mocenského vztahu. Dichotomické rozdělení, které proti sobě staví stroje a lidstvo, je důležitým prvkem, utvářející obavu ze samotných strojů. Stroje však nejsou vědomým subjektem, který by měl potřebu vymezení se vůči svým stvořitelům. Jsou nadále pouhým nástrojem, který může mít různá využití. Nástroj sám o sobě není dobrý ani špatný. Na jedné straně nás nástroje a technologie vždy posunuli na pomyslné vývojové linii o krok dál. Zároveň nám umožnili být zničující. Podíváme-li se do nacistického Německa, pak je to technologie byrokracie, která umožnila tak příšerné věci, jako byl holocaust provádět s velkou efektivitou (Bauman 2003, s. 42). Vývoj technologie za Druhé světové války přinesl

push-button warfare, které ze zabíjení a ničení vytvořila tak jednoduchou věc, jako bylo stisknutí tlačítka (Rid 2016, s. 77). Vývoj technologie má tedy dopad pozitivní i negativní. V některých ohledech fungují tyto dopady současně.

V následujících kapitolách se budu snažit nastínit jaké myšlenky ze sci-fi děl se prolínají s realitou a jaké ne. Snahou je se dívat na určitou podobnost mezi myšlenkami sci-fi a současným stavem technologie. Tato analýza by měla poukázat na to, že i když jsou stroje drženy stále v područí panství člověka, tak jejich implementace přináší rizika.

3 Čapkovské proroctví

V divadelní hře Karla Čapka s názvem *R.U.R.* (Rosumovi Umělí Roboti) se poprvé setkáváme se slovem robot, které je odvozené od slova *roboťa*, tedy práce. Továrna, která roboty vyrábí se jmenuje Rossum, čemuž můžeme rozumět jako slovní hříčce ke slovu rozum, což vytváří binaritu rozumného pána a pracujícího sluhy. Tvorba robotů má totiž osvobodit lidstvo od práce, jelikož veškerá práce bude delegována právě na roboty. Kompletní nahrazení člověka v pracovní činnosti má přinést jistou utopii v podobě osvobození od rutinní a monotónní práce. Roboti v tomto dramatu nejsou pouze bezduché pracovní stroje, jelikož se vyznačují něčím, co bychom v díle charakterizovali jako latentní produkt výroby, kterým je vědomí. Ono vědomí pak přivádí roboty do stavu uvědomění si své vlastní situace, což vede ke vzpouře proti jejich vládcům-lidem. V rámci autorovy doby je jasné, že se spíše jedná o alegorii na socialismus než o jakési varování před povstáním humanoidních robotů (Roberts 2006, s. 168). I přes tento fakt nemůžeme opomenout, že myšlenky v tomto díle se dotýkají i současného stavu technologií, jak si ukážeme v dalších podkapitolách. Čapek nám přináší dvě zajímavá témata, se kterými se dnes setkáváme při debatách o vztahu člověka a stroje. Tato dvě témata můžeme vměstnat pod hlavičku *nahrazení*.

Robot je nejprve postaven do pozice otroka, kdy jako dokonalý stroj, který se neunaví, má za úkol pracovat místo člověka. Robot otročí a těží z toho pouze jeho pán-člověk. Bohužel díky latentnímu produktu-vědomí si robot uvědomuje své postavení, se kterým je nespokojený a bouří se proti svým stvořitelům. Už v tomto případě se zde objevuje jistá alegorie na biblický příběh o vyhnání z Ráje. Alegorie je ještě znatelnější, když v závěru příběhu popisuje poslední živá lidská postava vznikající lásku mezi robotem Primusem a robotkou Helenou symbolicky citací z Bible, kdy je přirovnává k Adamovi a Evě (Čapek 2013).

V této kapitole se soustředím na první rovinu nahrazení, kterou můžeme pojímat jako proces automatizace, kdy je člověk v pracovní činnosti vyměněn za stroj, který dokáže jeho práci vykonávat rychleji a efektivněji. Uvedu i stručný popis druhé roviny nahrazení, ke které se navrátím v rámci kapitoly *Kdo je tady pán?*

3.1 Nahrazení

Dílo Karla Čapka poskytuje dvě roviny nahrazení. Podíváme-li se na zakončení díla, pak pozorujeme nahrazení na absolutní úrovni. V tomto nahrazení se člověk stává pouze další evoluční fází, která je překonána. Podobně jako savci nahrazují na potravinovém žebříčku plazi, či jako Homo Sapiens vyhrává proti neandrtálcům. Již v předešlé kapitole jsme si ukázali, že absolutní nahrazení je se současným stavem systémů umělé inteligence a robotiky stále v nedohlednu, jelikož ani současné systémy umělé inteligence nejsou s to vytvořit AGI, která by vedla k singularitě. Lépe bychom spíše řekli, že lidstvo zatím nebylo s to vytvořit něco, co bychom uznali, že má vědomí a chtělo by se vzepřít poslušnosti. Čapek ono absolutní nahrazení skvěle prezentuje na symbolice biblického příběhu o vyhnání z Ráje. Kdy poslední žijící člověk (dle jehož obrazu byl robot vytvořen) posílá roboty do světa, který museli lidé uvolnit pro nový druh bytostí. K tomuto tématu se vrátím v poslední kapitole této práce.

Vynález robota v dramatu je prezentací utopické představy, která má přinést osvobození lidských dělníků od práce². Robot má nejen nahradit všechny lidské pracovníky, ale taktéž i jejich práci vykonávat lépe. Robotika nám zatím nepřináší stroje typu Čapkových robotů. V průběhu historie jsme se ovšem už setkali se strojovým překonáním člověka alespoň ve specifických činnostech. V pracovní činnosti je takovéto nahrazení člověka strojem rozuměno jako zautomatizování procesu. S *automatizací* se v nějakém měřítku setkáváme skrze historii, kdy přicházejí technologie, které dokážou usnadňovat práci, výrobu a podobně. Největší boom se ovšem týká rozvoje digitální technologie po Druhé světové válce. Počítačový vědec Michael Dertouzos dokonce přichází s čapkovskou definicí, tedy že *automatizace* je proces nahrazení lidských úkonů stroji (Ekbia, Nardi 2017, s. 24–26). Jak jsme už měli možnost spatřit ve Winerových myšlenkách, nejednalo se jen o manuální práci, ale i kognitivní úlohy. Při procesu *automatizace* můžeme zažívat podobné osvobození od práce jako zažívaly postavy z Čapkova dramatu. Nudnou a monotónní práci za nás může vykonat stroj a my se můžeme zaměřit na další úkoly. Co nastane až bude stroj lepší ve všech úkonech?

V souvislosti s vývojem robotiky a umělé inteligence se čapkovské proroctví začíná v očích některých pomalu naplňovat. Už ve třicátých letech minulého století varoval John Maynard Keynes před *technologickou nezaměstnaností*. V průběhu let se s vývojem technologií ukázalo, že tyto obavy z nahrazení se zdají být skutečné, když počítače nahradily určité pracovní pozice. S těmito úkazy ve spojení s neustálým vývojem umělé inteligence se vrací debata o *technologické nezaměstnanosti* zpět do víru debat ekonomů (Frey, Osborne 2017). Místo utopické představy se nám dostává pocitu obavy o možnost práce a s tím spojenou možnost příjmu. Pokud veškeré pracovní pozice obsadí mechaničtí dělníci s minimálním počtem potřeb, tak čím se bude živit prostý člověk, který potřebuje jíst, pít a spát? S prvními obavami z této dystopie, kde lidé ztratí práci, jsme se mohli setkat už při industriální revoluci v podobě výše zmiňovaného luddismu, jehož posláním destrukce strojů vycházelo hlavně z obav ztráty zaměstnání (Wooldridge 2021, s. 150).

² Nemůžeme tvrdit, že by všechny postavy braly výrobu robotů pozitivně. Helena s roboty soucítí a považuje jejich použití za využívání. Alquist se zase nechce vzdát manuální práce apod.

Při aktuálních debatách o *technologické nezaměstnanosti* se však nepředpokládá, že by roboti eliminovali veškerá pracovní místa. I při vytvoření dokonalých androidů, kteří nikdy nebudou unavení, tak mezi námi zůstanou Alquistové,³ kteří budou stále zaměstnaní. Lidé, kteří argumentují technologickou nezaměstnaností předpokládají, že vysoké procento lidských pracovníků a pracovníc bude v blízké budoucnosti nezaměstnaných a nebude pro ně možné nalézt alternativy. Odhadem by tato technologická nezaměstnanost postihla 60 až 70% zaměstnaní. V logické formě by argument technologické nezaměstnanosti mohl vypadat takto:

1. Pokud technologie může a nahrazuje více a více forem lidské práce a pokud zde bude méně a méně alternativních zaměstnání, pak zde bude technologická nezaměstnanost.
2. Technologie může a nahrazuje více a více pracovních pozic a tím zapříčiní méně a méně alternativ práce.
3. Pak zde bude technologická nezaměstnanost.

Studie podporující technologickou nezaměstnanost vycházejí z druhé premisy, ale měli bychom být opatrní s jejich interpretací. Fray a Osborne uvádějí, že 47% povolání v USA je vystaveno vysokému riziku automatizace. Jejich práce ovšem vychází z O*NET databáze z roku 2010 a nepracují s jinými možnostmi nových zaměstnání, které mohou vzniknout díky větší ekonomické aktivitě (Danaher 2019).

Argumenty proti tomuto strachu můžeme spatřovat i v průmyslové revoluci, jejíž příchod znamenal přesun z venkova do továren, a tedy zásadní proměnu povahy práce. Podobná změna práce nastala až v 70. letech minulého století s nástupem globalizace, která znamenala přesun výroby do zemí třetího světa. Další výraznou změnou byla automatizace pomocí technologického zázraku mikroprocesorů. Vliv této automatizace byl pro mnohé devastační, na druhou stranu znamenal taktéž zvýšení ekonomické aktivity a tím vytvořila nové možnosti práce. V důsledku tak bylo vytvořeno více pracovních míst, než bylo zničeno. Tyto pracovní pozice se pak ve většině nenacházely v tradičních průmyslových centrech. Navíc automatizace povětšinou zastala nekvalifikované pracovní pozice. S umělou inteligencí se však hrozba nahrazení přesouvá i na kvalifikované práce. Z předešlých revolucí lze tedy doufat v další pokrok, což by byl pohled výše zmíněných techno-optimistů, kteří v oboru umělé inteligence a robotice spatřují nástroj osvobození od úmorné a nudné práce, podobně jako si to představoval Harry Domin z Čapkova dramatu. Příchod automatizace prostřednictvím mikroprocesorů měl samozřejmě na počátku své negativní konotace v podobě toho, že spoustu dělníků ztratilo pracovní místa (Wooldridge 2021). Vesměs růst ekonomické aktivity a tím nárůst pracovních pozic byl pouhým pohlcením rizik, které sama vyplodila.

Automatizované stroje jsou taktéž ukázkou dokonalosti. Stupeň automatizace je úměrný stupni dokonalosti stroje, tedy čím více je stroj automatizovaný, tím více je dokonalý. Zde je důležitým aspektem onen ohraničený obor činnosti, který činí automatizaci jednodušší. Na druhou stranu se obětuje mnoho potenciálních funkcí stroje. Pro splnění automatizace je potřeba stereotypizovat funkčnost stroje. Automatizace tak nepřináší

³ Postava z divadelní hry *R.U.R.*, která i přes nástup robotů pracovala.

technický pokrok, ale spíše riziko jeho zastavení, jelikož dokud není systém automatizován, pak nabízí své výhody v možném redesignu (Baudrillard 1996, s. 110).

Stroj se tak nemůže snadno reteritorializovat, jelikož jeho funkce zůstanou zakořeněny v jednom specifickém úkonu. Tím je právě znatelná výhoda pro lidské pracovníky, kteří mohou své teritorium změnit z jednoho zaměstnání na jiné v případě vzniku nových pracovních pozic. Možnost reteritorializace se stává klíčovou, ale zároveň i potřebnou. Stereotypizace činnosti člověka, tedy jeho zamknutí k určitému úkonu se v rámci vývoje stává neudržitelnou myšlenkou. Hlavně z důvodu, že se rozpadá systém standardizovaného plného zaměstnání spíše na systém flexibilní (Beck 2011, s. 226). Proces nahrazení je tak složitější, než si zastánci technologické nezaměstnanosti myslí. Neměli bychom ovšem zapomínat, že některá nahrazení byla nevratná a reteritorializace nemohla být možná. I při vzniku nových pracovních míst vlivem zvýšené ekonomické aktivity stále dochází k určitému znevýhodnění pracovníků, a to tedy skrze to, že se jejich jistoty vytrácejí.

Přestože predikce technologické nezaměstnanosti se nemusí naplnit, tak zde stále zůstává přítomný negativní prvek implementace nových technologií do pracovního prostředí.

3.2 Robot jako nástroj vykořisťování

První průmyslová revoluce přinesla výraznou změnu produkce. Lidé byli z venkova přesunuti do měst a z malé produkce se stala továrenská produkce fungující ve velkém měřítku (Wooldridge 2021, s. 150). Tato proměna měla pak dopad i na proměnu povahy práce. Ve výrobním procesu dochází k odcizení dělníka od jeho vytvořeného produktu. Dělník je zároveň komoditou jejíž hodnota klesá se zvyšujícím se počtem zhotovených výrobků. Dělníková práce se stává objektem s vnější existencí, tedy nezávisle mimo něj. Propad dělníka na úroveň komodity pak přináší kumulaci kapitálu do rukou těch, kteří vlastní továrny (Marx 1988, s. 74). Práce vykonávána dělníky je repetitivní, nesmyslná a nenaplnující (Wooldridge 2021, s. 150). Dělníci, o kterých Marx mluví se tak nachazejí v podobné situaci jako roboti v dramatu *R.U.R.* To se odráží i od zmiňované aluze přítomné Čapkově dramatu. Roboti otročí a zisk z jejich práce připadá jejich vlastníkům-lidem.

Nahrazení člověka strojem tak přináší spíše opak toho, co je v dramatu prezentován, protože je robot tím otrokem, který má minimální potřeby, tudíž ani nemá potřebu se vzpírat svým pánům.

Příchod automatizace by měl slibovat jakési osvobození od těchto repetitivních úkolů i pro dělníky. Růst ekonomické aktivity by pak vytvořil další pracovní místa, na která by se mohli pracovníci a pracovníce reteritorializovat. Bohužel v tomto smyslu osvobození v podobě automatizace vede pouze k dalšímu vykořisťování, a to zejména v případě prekarizovaných pozic, kdy roste platová nerovnost (Acemoglu, Restrepo 2022). Nejprve zde máme vliv mikroelektroniky, která umožnila oddělení pracovního procesu od procesu produkce. Oba systémy, tedy jak systém lidské práce, tak systém technické produkce, mohou být obměňovány nezávisle na sobě (Beck 2011, s. 346). Od konce 70. let, kdy se začínají rozpínat procesy globalizace, digitalizace a automatizace, se dostáváme k oddělení

růstu produktivity a výše mezd. Automatizace je zde označována v určitých narativech za hlavního viníka, přestože má pouze podíl na tomto trendu (Bivens, Mishel 2021). Mohli bychom například očekávat, že automatizace pohltí veškerou repetitivní práci, ale setkáváme se i s opačnými příklady, protože lidská práce je v rámci krátkodobého zisku levnější než automatizovaný systém (Ekbia, Nardi 2014). Ačkoliv by se tedy mohlo zdát, že jsou stroje perfektním substitučním prvkem člověka hlavně v repetitivních činnostech, tak se opak stává pravdou a lidský pracovník je neustále uchycen v tomto nepříznivém postavení.

Nástup automatizovaných strojů přináší nový vztah mezi člověkem a strojem. Podíváme-li se na proces *augmentace*, je zde přítomná jistá spolupráce člověka a stroje. Základní myšlenka se odvíjí od symbiózy člověka a stroje, kdy stroj pomáhá člověku téměř ve všech profesních záležitostech (Ekbia, Nardi 2017, s. 24). V tomto přístupu je zachované určení subjektu a objektu, kdy subjekt je aktivním činitelem, tedy ovládá objekt v podobě určitého technologického nástroje, jako je například počítač. Automatizovaný stroj nepřichází jako nástroj, ale jako činitel. To vede k tomu, že musí mezi strojem a lidským pracovníkem nastat dělba práce. Automatizace není něco, co by kompletně nahrazovalo vykořisťovaného dělníka, ale spíše jej posouvá na jinou pozici. V prostředí *Amazon* skladů se například umělá inteligence stává nástrojem moci vykonávané na dělníkovi. Systém umělé inteligence v tomto případě funguje jako myslící subjekt, který rozhoduje o nejvýhodnější trase, na které se má pracovník pohybovat, aby pracoval nejefektivněji. Pracovník Amazonu toto soužití s vedoucím v podobě umělé inteligence popisuje takto: „*Software vypočítá nejefektivnější pěší trasu k nasbírání všech předmětů. Jednoduše navádí pracovníka od jednoho regálu k druhému pomocí instrukcí. Lidská automatizace, chcete-li.*” (Wooldridge 2021, s. 158). Pracovník je tak redukován na nástroj, který je využívaný systémem umělé inteligence k pohybu ve fyzickém světě. Winerova předzvěst toho, že stroje nahradí lidský mozek je nyní zde. V tomto příkladu se však vytrácí dopad první průmyslové revoluce, který měl dle Wienera nahradit lidské svaly. Umělá inteligence zde spíše staví člověka do pozice své vlastní extenze. Umělá inteligence vypočítá trasu a člověk koná dle jejích příkazů jako poslušný automatizovaný stroj schopný interakce s fyzickým světem.

Setkáváme se tedy s procesem, kdy člověk není nahrazován, ale je vyžadován. Tomuto procesu se říká: *heteromatizace (heteromation)*. Heteromatizace dává vznik nové sociotechnické konfiguraci s vlastním souborem technologií, organizačních uspořádání, požadovaných rolí a schopností. V těchto systémech jsou totiž kritické úkoly přesouvané ke koncovému uživateli. Jedná se o takový systém, který funguje na základě činnosti různorodých aktérů. Mezi takové aktéry patří například videohry či sociální sítě. Tato technologie vyplňuje mezeru způsobenou automatizací a mění sociální vztahy tím, že z lidí vytváří počítačové komponenty. Na jedné straně se čerpá z levné pracovní síly prekarizovaných pozic a na druhé straně stejné uspořádání využívá nekompensovanou práci řady vzdělaných jedinců tím, že jim nabízí silnou afektivní odměnu (Ekbia, Nardi 2014). Příkladem mohou být například samoobslužné pokladny v supermarketech. Zákazník si nabere zboží, které potřebuje, přijde k samoobslužné pokladně a sám si nákup načte a zaplatí. Stroj (tedy automatická pokladna) si zavolá na pomoc prodavače v případě, pokud nastala nějaká chyba. Je tak využita práce samotného zákazníka a popřípadě využita práce

prekarizovaného zaměstnance, která se stává téměř neviditelnou. Dalším zajímavým příkladem jsou ony zmiňované sociální sítě, které z člověka dělají zdroj dat, s nimiž se dále obchoduje. Uživatel přitom dostává pouze afektivní odměnu. K tomuto příkladu se budu více věnovat v kapitole *Kdo je tady pán?*

Zdá se, že čapkovské proroctví v katastrofálním scénáři reality, tj. nahrazení člověka v pracovní činnosti, se zdá být pouhým zlomkem problematiky přechodu nových systémů umělé inteligence. Lidská společnost západního světa již prošla několika změnami či revolucemi, které proměnily povahu práce, a toto se jeví jako další z nich. Přijmeme-li argumenty techno-optimistů, pak se z nahrazení nestává problém, ale řešení špatné situace pracovníků a pracovníc. Opět zde vystupuje technicko-ekonomický vývoj, který pojímá rizika. Nahrazení už není rizikem, ale nezbytným prvkem vývoje, které přinese ekonomický růst a s ním nová pracovní místa. Tento pokrok však podryvá jistoty pracovníků a pracovníc, kterým nezbyde nic jiného než přesun k jinému zaměstnání. Navíc je zde vyžadovaná reteritorializace, která nemusí vždy proběhnout skvěle. Hlavní problém se zde nachází v jistém stupni automatizace člověka samotného, například skrze vzdělání vázané k určité pozici (od vyučeného malíře se očekává, že bude malovat). Čím méně činností obsahuje dané zaměstnání, tím lépe se automatizuje. Nastavením oblasti, která je silně strukturovaná pomocí pravidel přináší možnost vytvoření automatizovaného stroje, který dokáže nahradit svého nedokonalého předchůdce-člověka.

Samotná strojová automatizace je pro vlastníky kapitálu kolikrát riskantní investicí, což pak vede k tomu, že jsou využiti lidští pracovníci. Všechny tyto jevy se zdají být spíše negativními směrem k pracovníkovi, který se možná nemusí bát ztráty práce, ale jeho hodnota bude klesat v rámci zneviditelnění některých aspektů jeho pracovní činnosti. Stroj už se nestává pouze pasivním objektem, který je využíván, ale do práce vstupuje aktivně jakožto dokonalý automatizovaný systém. Práce pro člověka a pro stroj se odděluje do dvou samostatných kategorií. Jde tedy o dělu práce mezi pracovníkem z masa a kostí a pracovníkem z kovu a kabelů. Samo toto dělení pak proměňuje náplň práce, což znamená výhody na straně zaměstnavatele a možná i jisté malé výhody na straně zaměstnance, ale zároveň i negativní konotace, které se odrážejí na jeho finanční odměně.

3.3 Spojení sil

Proces automatizace v doprovodu hrozby technologické nezaměstnanosti přináší lidské postavení člověk proti stroji, které je přítomno i ve sci-fi dílech. Pokud bychom v současné době chtěli naplnit čapkovské proroctví, tedy jen do té míry, kdy nás stroje osvobodí od práce, pak bychom se dostali do stagnace. Zautomatizování veškeré činnosti by vedlo pouze k opakování se určitého vzorce, což by nevedlo k žádnému vývoji. Uvázli bychom v jednom teritoriu, ze kterého bychom se nemohli vymanit.

Zřít se všech automatizovaných systémů a s lidským šílenstvím v očích popadnout kladiva a veškeré technologie rozdrtit na prach by ovšem taktéž nepřineslo rozumný krok. Bez technologií bychom se odřízli od možností, které v současnosti máme. Jak uvádějí někteří autoři, tak automatizace určité činnosti může vést k vytváření nových příležitostí. Nejlepší cestou se tak zdá být to, co nabízí augmentace. Neměli bychom se tak dostávat do

dvou nepřátelských pozic: člověk versus stroj, ale spíše do spojení člověk a stroj. Spolupráce člověka a technologie je to, co přináší jistý posun z jednoho bodu do druhého. Stejně jako mě chůze dostane na jiné místo, tak kolo mě na to místo dostane rychleji.

Lehce se řekne, ale těžko udělá. V žádném případě se nenacházíme na jednoduché dichotomii lidé a stroje, ale jsme v daleko složitějším mocenském systému. Dle Marxe bychom řekli, že buržoazie využívá technologie k vykořisťování dělnické třídy. Podobnou dynamiku můžeme spatřovat skrze automatizaci a heteromatizaci. Vše se odehrává ve složitějších strukturách. Konflikt tak není mezi lidmi a stroji, ale mezi lidmi a lidmi.

4 Tvor dvounohý a mechanický

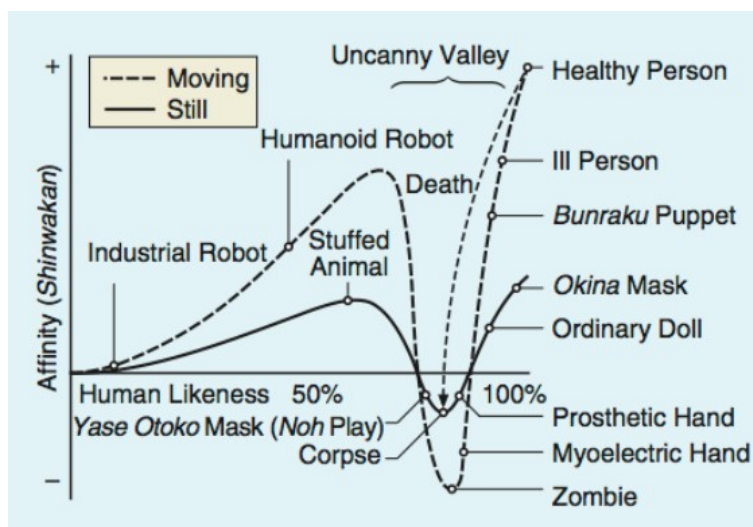
V rámci filosofie traduje příběh o Platónově definici člověka, kdy jej popsal jako dvounohého tvora neopeřeného. Diogenés ze Sinópe na základě této definice oškubal kohouta a prezentoval jej jako člověka. Jelikož měl oškubaný kohout dvě nohy a žádné peří, tak přesně splňoval Platónovu definici. Tato anekdota odkazuje k filosofickému problému vymezení člověka pomocí definice (Diogenés Laertios, Kolář, Machovec 1995, s. 234–235). V této práci se nebudu nikterak pokoušet o definování člověka či se zabývat dalšími definicemi. Nelze se zde ovšem vyhnout přemýšlení o narušení kategorie člověka v případě vytvoření jeho kopie. V rámci Čapkova dramatu *R.U.R.* se setkáváme s robotem, který je designovaný k práci, ale uchovává si jistý antropomorfismus, a to jak ve vizuální podobě člověka, tak v pozdější části dramatu i v podobě jednání.

Tvorba robotů má ve skutečnosti trochu odlišné přístupy. Jejich design spíše odpovídá odvětví, do kterého mají být nasazeni. Můžeme se tak setkat s dvěma rovinami výroby robotů: První je robot, který je designovaný k práci (jak naznačuje původ jeho názvu). Jeho design se přizpůsobuje dané činnosti, kterou má stroj vykonávat. Příkladem může být mechanické rameno sloužící k přesunu těžkých předmětů či roomba vysavač. Druhou je mechanická napodobenina člověka, která je podřízena analogii a vytváří efekt simulakry⁴. Baudrillard jej pojmenovává jako *automaton*. Jediným cílem automatonu je být srovnáván s člověkem a být přirozenější než on (2017, s. 74–75). V technickém slovníku se spíše setkáme s rozlišením na industriální a sociální roboty (Fortunati, Esposito, Lugano 2015). Při tvorbě reprezentace člověka se potýkáme opět s dvěma liniemi, které vycházejí z dvou zmiňovaných oborů. *Umělá inteligence* se snaží o napodobení kognitivních funkcí (Eysenck, Eysenck 2021, s. 5). *Robotika* se pak snaží o vytvoření stroje, který by nesl vizuální podobnost člověku (Mori, MacDorman, Kageki 2012). Tyto dvě roviny se následně propojují při vytváření robotů vybavených umělou inteligencí, kdy ono propojení přináší dokonalejší prezentaci, jelikož pokrývá oblast vizuální, tak i oblast kognitivní, která je spatřována skrze jednání robota.

Mluvíme-li o linii nápodoby vizuální, pak nemůžeme opomenout hypotézu předloženou v sedmdesátých letech minulého století japonským profesorem robotiky Masahirem Morim.

⁴ Simulakrum pojímám jako napodobení. Podobně s tímto pojmem pracuje i Baudrillard, který jej rozpracovává do tří úrovní: prvního řádu, kdy je napodobení rozlišitelné od originálu, druhého řádu, kdy se vytrácí rozdíly mezi napodobením a originálem a třetího řádu, kdy existuje pouze simulakrum bez žádného vztahu k realitě.

Hypotéza předkládá, že při pohledu na robota, protézu či jinou napodobeninu člověka, která jeví jisté známky nedokonalosti, náklonnost člověka upadá do takzvaného *uncanny valley* či v češtině *tísnivého údolí*. Ve své eseji vypracoval graf (viz obrázek 1), který ukazuje, že náklonnost lidí k robotům, kteří mají vypadat jako lidé, roste, dokud nenarazíme na zmíněné *uncanny valley*.



Obrázek 1: *Uncanny valley*

Podíváme-li se na industriální roboty, pak se setkáme s tím, že u těch není potřeba, aby vypadali jako lidé. Jejich design je uzpůsobený k určitému typu činnosti a jejich podobnost lidské bytosti není důležitá pro funkci. Lidská náklonnost je zde v nízkých hodnotách, jelikož se jedná o objekt výrobního procesu. Jinými slovy je to bezduchý chladný stroj, který pouze plní svou funkci. Máme-li robota jakožto hračku pro děti, pak tento robot musí již projevovat jisté známky antropomorfismu, aby byla vyvolána určitá náklonnost dětí k hračce. Takovýto robot se projevuje mechanickou robotí podobou s jistými polidšťujícími prvky. Těmito prvky mohou být oči, ústa (většinou usměvavá), či dokonce celá tvář. S tímto se setkáváme například v animovaných filmech a seriálech, které polidšťují například auta. Ať už se jedná o film *Auta* či pořad *Bořek Stavitel*, tak je zde vyvolána náklonnost ke strojům právě díky jejich tvářím na mechanických tělech. Dle Moriho hypotézy by tato náklonnost vzrostla, kdybychom se bavili třeba o robotovi humanoidního typu. Pokud je ovšem vytvořena protéza, která může být zhotovena z kovu a šroubů a je pokryta umělou kůží, která se od lidské liší jen v nepatrných detailech, jako je barva a podobně, pak se dostáváme do bodu, kdy je tato protéza téměř k nerozeznání od pravé lidské ruky. Jelikož protéza dosahuje vyšší podobnosti lidské končetině, pak by dle předpokladů měla stoupat i náklonnost k této kopii, nicméně ve skutečnosti dochází k pravému opaku. Jedinec konfrontovaný s protézou zažívá úpadek náklonnosti a upadá do pocitu tísně. Nedílnou součástí vytvoření pocitu živoucí bytosti u umělé kopie je pohyb. Pokud je k umělému přidán pohyb, pak můžeme na grafu pozorovat sice zesílení náklonnosti, ale stejně tak i strmější pád do údolí nepříjemného pocitu. Stoupající úroveň antropomorfismu může mít tedy při určitém stupni paradoxně neblahý dopad. V případě industriálních robotů jsme si určili jejich povahu nahrazení v pracovní činnosti, kdy není důležitá podobnost člověku, což se týká i robotů nyní již běžně

fungujících v domácnostech (například robotický vysavač). Jejich design je spíše funkcionální povahy.⁵ U hračky robota jde o estetické zalíbení, takže musí robot vykazovat prvky, které dokážou zaujmout a jednoduše přilákat pozorovatele, tedy oči a úsměv. Nápodoba necílí na zobrazení kopie člověka, ale pouze polidšťuje chladný stroj. Jaký je tedy problém s roboty, kteří vypadají jako lidé? Je to právě ukázka střetnutí se s nedokonalou kopií, co nás dostává do pocitu tísně. Hledíme do tváře něčemu, co splňuje na první pohled základní parametry kategorie člověka: má kůži, má dvě oči, nohy, ruce, ústa etc. Na první pohled ovšem dokážeme rozeznat, že je něco špatně, a že to, co nám hledí do tváře, není člověk z masa a kostí, ale kovové mechanismy potáhnuté umělou kůží. Sám Mori navrhuje, aby se designéři a designérky zaměřili na dosažení prvního vrcholu, tedy aby se designem dostali před pád do uncanny valley (Mori, MacDorman, Kageki 2012).

Tendence o vytvoření kopie člověka, a tedy překonání *tísnivého údolí* ovšem stále nezemřely. *Uncanny valley* nás dostává na ono pomezí dokonalosti reprezentace člověka, kdy se dostáváme do mezistupně mezi nedokonalou a dokonalou nápodobou. Překonání tohoto údolí je obrovskou výzvou v otázkách designu robotů. Jeho překonání na místo řešení problému náklonnosti však nastoluje otázku možného rozdělení kategorií reálného a umělého. Pokud tedy dovrší design robotů takové dokonalosti, že bude robot či lépe řečeno android na pohled nerozeznatelný od člověka, jak pak dokážeme odlišit člověka od androida?

4.1 (Ne)rozlišitelnost

Pro lepší pochopení konceptu nerozlišitelnosti je potřeba si stanovit rozdělení toho, co je zprvu odlišným (diference) a následně se stává nerozlišitelným (indiference). Narážíme zde na filosofický problém, co je vlastně umělé a co přirozené. Téměř jakákoliv definice je napadnutelná. Proto se pro jednoduchost přidržím pojetí, které zde již bylo nastíněno a domnívám se, že je tomuto rozlišení rozuměno i v současných veřejných diskurzích. K lepšímu vysvětlení použijme příklad rybníku a jezera. Jezero vnímáme jakožto přírodní či přirozené a rybník jako umělou vodní plochu, jelikož člověk musel vyhloubit jámu, dodat vodu a ryby. Samozřejmě můžeme poukazovat na to, že jezera mohou vznikat i lidským přičiněním, ať už je vědomé nebo ne. Nám ovšem stačí právě toto vykonstruované rozdělení, kdy je jedna věc považována za reálnou, přírodní či přirozenou a druhá je její umělou nápodobou, tedy je vytvořena jakožto obraz ke svému „přirozenému“ protějšku. Skrze toto východisko následně prezentuji stromovou strukturu nápodoby, kdy přirozený protějšek-přírodní je reprezentován pomocí umělé nápodoby.

Pokud se budeme snažit vytvořit napodobeninu či přímo kopii, pak je naší snahou, abychom se co nejvíce přiblížili originálu. Kopie s lepším a lepším napodobením jsou více a více podobné originálu. Se zvětšující se mírou podobnosti dosáhneme toho, že nedokážeme na první pohled určit, který z objektů je originálem, a který je kopií. Představme

⁵ Mohli bychom tvrdit, že je vzhled vždy nějak vztažen k funkcionální povaze. U robotů v domácnosti je to pak i složitější, jelikož jsou produktem, který má být taktéž nějakým způsobem příjemný na pohled (neměla by to být zněť kabelů, pístů apod.). Hlavním aspektem je tak polidšťování v rámci vizuální sféry, která v rámci těchto produktů nedává ani v rovině funkce smysl. Nebylo by zrovna efektivní, aby nám po podlaze jezdila napodobenina lidského těla.

si, že máme osobu A a osobu B. Jedna z těchto dvou osob je člověk a druhá je android. Na první pohled ovšem nedokážeme rozlišit mezi tím, kdo je skutečnou osobou z masa a kostí a kdo ne. Rozlišení se objevuje i jako základ prvního testu inteligence stroje, kterou předložil Alan Turing dávno před vznikem prvních systémů umělé inteligence. V *Turingově testu* jde právě o to rozlišit, jestli subjekt komunikuje s člověkem, nebo strojem na základě určitého počtu otázek. V případě, že stroj nebyl rozeznán jako stroj, pak jsme ho mohli považovat za inteligentní (Turing 1950). Tento přístup se setkal i s kritikou například od filosofa Johna Searla. Jádrem oné kritiky je skutečnost, že na základě Turingova testu nemůžeme říct, že je stroj inteligentní, ale pouze, že inteligenci napodobuje (Wooldridge 2021, s. 179). Ještě lépe řečeno, stroj pouze napodobuje něco, co bychom považovali za inteligentní jednání. Searlova kritika a ani Turingův test pro nás zde nejsou tolik relevantní. To důležité je ona nápodoba. Pokud Turingův test může určit stroj jako inteligentní na základě toho, že se inteligentní jeví, tak proč bychom nemohli určit stroj jako člověka, když se jako člověk jeví.

Problém nerozlišitelnosti je znatelný například v díle Philipa K. Dicka *Blade Runner*. Hlavní postava Rick Deckard je nájemný lovec. Úkolem těchto nájemných lovců je potlačení hrozby uprchlých androidů. Blade Runner, jak se tyto nájemní lovci nazývají musí androida najít, identifikovat a poslat na odpočinek. V identifikaci nastává ona obtížnost úkolu. Androidi v tomto světě nejsou pouhými mechanickými humanoidy, ale jsou téměř identickou reprezentací člověka. Nedají se tak na první pohled rozeznat od lidí. Rick Deckard ještě musí ulovit androidy modelu Nexus-6, kteří jsou dokonalejší verzí, než byli jejich předchůdci. Otázkou je, jak je možné určit, jestli je daná osoba člověk, nebo android. Tyto dvě kategorie nedošly ke svému sjednocení, jelikož jsou dále oddělovány skrze určité nástroje, jako je test kostní dřeně či Voight-Kampffův test. Voight-Kampffův test se provádí dle sady otázek, které jsou položeny testovanému subjektu. Přístroje měří empatickou reakci. Na základě této proměnné je možné určit, zda se jedná o člověka, nebo androida, jelikož androidi mají onu empatickou reakci nižší (Dick 1993). V románu *Blade Runner* je tak využíváno stejného rozdělení na reálné a umělé, jako jsem předestřel na začátku této podkapitoly.

Můžeme sledovat, že rozdělování na kategorie reálné (člověk) a umělé (android) tak samo padá do umělého rozdělení, které je vystavěno na základě určitého druhu vědění, jež má člověk o lidech a androidech. Oddělení kategorií je navíc vystaveno na tak malicherných detailech, že zde samozřejmě může nastat chyba. Za androida může být označený člověk, který má nižší míru empatie a podobně (Galvan 1997). Tato kategorizace nese jistou podobnost s kategorizací, kterou předkládá Michel Foucault v příkladě nápravných institucí. Pracuje se zde ve dvou režimech: binárního rozdělení a označení (normální a nenormální, zdravý a nemocný apod.); a v režimu donucujícího vykázání, rozlišujícího rozdělení (kdo je kdo, kde má být, jak má být charakterizován atd.) (Foucault 2000, s. 280). Rick Deckard je přesně tou autoritou, která za pomoci nástroje Voight-Kampffova testu rozděluje na binární kategorii člověka a androida. Na základě tohoto rozdělení, pak postupuje určité kroky svého jednání, ergo androidy *posílá na odpočinek*, aby potlačil hrozbu pro společnost. Stejně jako policisté posílají zločince do vězení.

Základem tvoření kategorií je pojmenovávání. V příběhu *Blade Runner* se potýkáme právě s udržetím tohoto pojmenování, a tedy rozdělením dvou kategorií, které nastalo ve chvíli rozlišitelnosti. Člověk byl jasně definovaný a stejně tak byl jasně definovaný android. Postupem vývoje se tyto rozdíly začaly vytrácet a pro zachování těchto dvou kategorií jsou

nájemní lovci dohnání až k hledání jisté *esence*, která by dokázala přesně určit, zda je naše pojmenování pravdivé, či nepravdivé. Saul Kripke uvádí obdobný příklad s kočkami a démony. Kdyby démoni dokázali na sebe vzít podobu kočky, pak dle Kripkeho nebudou kočkami, ale pouze démony v kočičích podobách (1980, s. 126). Problém nastává ovšem v hledání esence. Dokonalejší úroveň nápodoby ono hledání činí složitějším, jelikož ono rozdělení na dvě kategorie člověk-android, kočka-démon v kočičí podobě se při dokonalejší a dokonalejší prezentaci stává závislé na detailech, jež ovšem zůstávají lidskému oku skryté. Věda může uchovávat ony dvě kategorie, které mohou být odděleny na základě předchozího vědění a určeny pomocí testu, jako je Voight-Kampffův test. V sociálních interakcích ovšem takovými testy zpravidla vybaveni nebýváme. Může tak docházet k chybě pojmenovávání, jelikož tyto dva druhy jsou vizuálně neoddělitelné. Nastává tak opakování onoho nepravdivého pojmenování. Opakování může být vždy reprezentované jakožto extrémní podobnost či dokonalá ekvivalence, ale skutečnost, že lze přejít od jedné věci k druhé nebrání tomu, aby se od sebe druhově lišily (Deleuze 1994, s. 2).

Může se zdát, že onen problém nepravdivého označování androida jako člověka se týká pouze rozebíraného sci-fi díla. V současném světě se přece jen nesetkáváme s problémem androidů, kteří pobíhají po městě a jsou k nerozeznání od lidí. Možná na tento problém nenarazíme, když si půjdeme koupit rohlíky či veganský steak, který vypadá a chutná jako opravdový, nicméně můžeme se s ním setkat v jiném prostoru za jiných podmínek. Turingův test totiž také neprobíhá v přímé konverzaci *face-to-face*, nýbrž skrze chatovací okénko. Nedíváme se tedy na nějakou umělou reprezentaci člověka, která by nám nabízela další nápovědy o jeho původu, jako jsou nedokonalosti v pohybu umělých svalů. Testujeme subjekt jen na základě konverzace vedené skrze redukovaný prostor chatovacího rozhraní. Současné LLM stále nejsou naprosto dokonalé v této nápodobě, ale jejich dokonalost napodobení se zvyšuje. Do role vstupuje taktéž mnoho proměnný v podobě toho jak je například ChatGPT designovaný, což přináší prostor ke správnému určení kategorie (Jones, Bergen 2023). Pokud by byl LLM uzpůsobený k tomu, aby lépe napodobil člověka, tak to znamená, že máme sociálního bota⁶, který se stává špatně odlišitelným od člověka. Hlavní roli zde stále hraje to, že existuje v takových podmínkách, kde není zapotřebí jeho nedokonalé fyzické tělo. Vylepšené podmínky pro nerozlišitelnost existují v rámci augmentované reality, která se v průběhu posledních desetiletí stala součástí našeho každodenního života. Řeč není o ničem jiném než o internetu, kde se nacházejí sociální boti v různých kvalitách.

„*Na internetu nikdo neví, že jsi pes.*” říká pes sedící u počítače v karikatuře Petra Steinera z roku 1993 v *New Yorker cartoon*. Jedná se o zachycení proměny mezilidských vztahů, jež v tomto roce teprve začínala. Tato změna byla plná vzrušení z nových možností a stejně tak děsivá. Za poslední čtvrtstoletí jsme měli možnost se mnohé naučit. Jakýkoliv cizinec na internetu mohl být kdokoliv. Čtrnáctiletá dívka v chatroomu mohla být policista v utajení, Africký princ s velkým dědictvím mohl být jen obyčejný podvodník ze sousední vesnice (West, Bergstrom 2019). S nástupem generativní umělé inteligence se tento aspekt internetu lehce mění. Nyní se totiž nesetkáváme s lidmi či psi, ale setkáváme se s entitami,

⁶ Sociální roboti mají dvojí působení. Označení se používá pro agenty působící ve sféře internetu, což znamená, že nemají fyzické tělo a pak i v realitě, kdy je fyzické tělo vyžadováno.

kteře nemají fyzické ukotvení snad jen krom toho, že umělá inteligence musí běžet na nějakém hardwaru.

Jedním z oficiálních pokusů zavedení umělé inteligence do světa sociálních sítí bylo v roce 2016 vypuštění sociálního bota Tay od společnosti Microsoft do prostředí sociální sítě Twitter. Tato umělá inteligence se učila na základě dat z této sociální sítě, na které pak interagovala s uživateli pomocí příspěvků a komentářů. Po méně než 24 hodinách své funkce se z Tay stala nespoutaná pohroma, která začala přidávat tweety, kde sympatizovala s nacismem, genocidou a rasismem (Wakefield 2016). Vinu můžeme přičítat algoritmickému zkreslení, které vzniklo na základě dat, ze kterých čerpala. Sociální boti, s nimiž se nyní můžeme setkat skřze sociální sítě, nejsou těmito „rebelujícími“ systémy, jež utekly do prostoru internetu. I v tomto případě si udržují strojovou poslušnost a fungují jako nástroje. Mezi jejich nejčastější využití patří šíření dezinformací, zneužívání uživatelských dat či ovlivňování veřejného mínění, přičemž zůstává snaha napodobovat lidského uživatele. Fenomén sociálních botů se pro uživatele pojí s nepříjemnou zkušeností, což je vede ke kategorické chybě, při které jsou za boty označeni i uživatelé s odlišným politickým názorem (Yan et al. 2023). Naše vymezení sociálních botů a generativní umělé inteligence se pak dotýká i lidí mimo kategorii nefyzického uživatele. Můžeme říci, že do *uncanny valley* zapadají i takové entity, které by zde být neměly. Nejsme tedy tak daleko od problémů objevujících se v díle Philipa K. Dicka, kdy nástroj Voight-Kampffovy stupnice může mylně kategorizovat člověka s poruchou empatie jako androida (Galvan 1997). Chybná kategorizace se dotkla i mladého herce Armena Nahapetiana, který se musel bránit tím, že je reálnou bytostí a není pouze produktem generativní umělé inteligence (Plale 2023).

Současní sociální boti ještě z pravidla nebývají poháněni velkými jazykovými modely, ale je bez debat, že jejich implementace bude mít za výsledek zvýšení tohoto problému nerozlišitelnosti v rámci sociálních sítí. Existuje dokonce projekt s názvem Chirper, který simuluje sociální síť, ve které interagují pouze uživatelé poháněni velkými jazykovými modely konkrétně skřze *ChatGPT*. Tito nehmotní uživatelé přidávají obsah na tuto síť a interagují s dalšími uživateli a udržují s nimi vztahy. Prostředí sociálních sítí je už nějakým způsobem redukovanou realitou, kde interakce probíhají dle určitých pravidel, pomocí postování, komentování atd. Takováto redukce reality přináší lepší podmínky pro napodobování, jelikož nemusíme řešit nedokonalost fyzických aspektů, se kterými se musí potýkat robotika (Li, Yang, Zhao 2023). Sociální boti jsou v tomto případě opět automatizované systémy, které prezentují svou dokonalost skřze jednání v určité specifické činnosti, jenž redukuje obtížné aspekty nápodoby.

Může se tak zdát, že tento problém rozlišování ve fyzickém světě se nachází daleko v budoucnosti, jelikož design robotů stále není dokonalou lidskou reprezentací. V rozšířené realitě, kdy je výstup jedinců nějakým způsobem redukován, pak je ona nápodoba zjednodušená. Když půjdu navštívit kavárnu, tak se nedostanu k otázce po kategorii baristy či baristky, jelikož tato bytost bude vždy spadat do kategorie člověka. Navštívím-li internetovou diskusi, pak může nastat to, že jsem se připlétl do konverzace se sociálním robotem místo člověka, jelikož na základě interakcí v tomto prostoru se tyto dvě kategorie stávají na první pohled nerozlišitelné.

Dílo Philipa K. Dicka v sobě nese jistý nádech posthumanismu, což je znatelné ve snižujících se rozdílech mezi člověkem a jeho umělým napodobením. My se prozatím s

tímto problémem máme možnost setkat pouze v redukováném prostředí internetu, ale jaký dopad bude mít tendence o vytvoření robota jakožto plné napodobení člověka, je otázkou budoucího vývoje. Jak říká Michel Foucault: „Člověk v rámci vědy je relativně nedávným vynálezem a možná se blíží ke svému konci.“ (2012, s. 422).

4.2 Stávání se strojem a stávání se člověkem

Vrátíme-li se do pracovního prostředí a konkrétně ke změnám industriální revoluce, kde zažíváme přesun z agrikultury do prostředí továren, a zaměříme-li se na objekt továrny jako celek, tak v ní můžeme spatřovat systém. Systém tvořený dalšími malými systémy, které svou funkcí zprostředkovávají chod celku. Pohybující se pás je systémem, člověk je systémem etc. Proces, kterým továrna prochází, je podobný proces jako proces stroje. Tento proces spočívá v přeměně něčeho na něco jiného. Strojový proces přeměny jednoho či více objektů na objekt nový se stává možný právě jen díky dalším malým systémům, které postupují dle stejného procesu.⁷ Jedním z těchto malých systémů byl dodnes neodmyslitelně člověk, a ještě nějakou dobu bude, i přestože ve vlně automatizace je povolna vytlačován a v určitých činnostech nahrazován.

Marx tvrdí, že při odcizení práce se dělník stává svobodně aktivním jen skrze své zvířecí funkce, jako je jezení, pití, rozmnožování atd. V lidských funkcích se již necítí být ničím jiným než zvířetem. To, co je lidské, se stává zvířecím, a to co je zvířecí, se stává lidským (Marx 1988, s. 74). Nyní se přikloňme k otázce, jestli je proces stroje skutečně procesem stroje či procesem člověka? Už samotná továrna či samotný stroj musel být vybudován právě tímto procesem. Nestává se tak lidské strojovým, přestože se člověk stává strojem a částí stroje v rámci továrny?

Při tvorbě stroje je zapotřebí vždy jakási sada příkazů, které vymezují jeho jednání. Pravidla, která jej vymezují, jsou tím, co vytváří stroj. Dělník v továrně se stává strojem, jelikož taktéž musí dodržovat pravidla, která strukturují jeho jednání v systému. Stejně tak jako je tělo vojáka přeměňováno na stroj pomocí tvorby poslušného těla, tak je i tělo dělníka přeměněno na stroj fungující v rámci továrny (Foucault 2000, s. 199). U člověka je toto redukování na stroj zprostředkováno automatizací, která se tak netýká pouze strojů, ale i člověka. Jakákoliv repetitivní práce vykonávaná člověkem je lidskou automatizací, a tudíž i jeho redukcí na stroj, čímž dochází k podobné situaci, jakou Baudrillard popisuje na automatizovaných strojích, a to, že je redukován jejich potenciál pomocí stereotypizace činnosti. Proces *stávání-se* je tak veden oběma směry. Ze stroje se stává člověk a z člověka se stává stroj. To celkově umožňuje lepší možnost nahrazení člověka strojem. V rámci moderny se právě strukturace jednání zejména v byrokracii stala nástrojem racionalizace, kdy jsou lidé přetvářeni na lidské roboty, kteří pracují bez lidských úsudků. Není tak divu, že řadu takovýchto úkonů následně nahradili počítače (Ritzer 2003, s. 37). Kuriózním případem je fenomén *NPC* na *TikTokových* streamech. Zkratka *NPC* stojí za *non-playable character* a jedná se o postavy, které utvářejí život virtuálního světa videoher. Některá herní studia vytvořila své *NPCs* takovým způsobem, že jejich napodobení člověka a lidských

⁷ Například ve Fordismu je proces výroby dělen mezi několik dělníků, kdy každý vytváří určitou část. Z těchto částí je pak sestaven konečný produkt. Každý dělník ovšem dělá v principu to stejné jako celá továrna, což je spojování určitých částí, které vytvářejí produkt.

interakcí vytvářela u hráčů výsměch, jelikož napodobení nebylo ani zdaleka dokonalé. Jako výsměch začali lidé napodobovat tato NPC v různých skeč videích. Tento fenomén se přeměnil až na žánr živých vysílání nejčastěji prostřednictvím sociální sítě *TikTok*. Při těchto streamech je *content creator* omezen na jednoduché pohyby a reakce na základě zasílaných odměn. Toto omezení na jednoduché reakce, které trvají v řádu sekund, pak vytváří imitaci algoritmu (Stokel-Walker 2023).

Jev, který zde můžeme pozorovat, je zmiňovaný proces automatizace, jenž má v sobě obsaženu jistou redukci. Veškerý další potenciál je odstraněn a člověk se stává biologickým systémem řídicím se dle pravidel (Při této odměně přijde tato reakce apod.). V příkladu s internetem jsme v jistém redukováném prostředí, které překračuje hranici fyzického světa do světa virtuálního. Tato redukce prostředí vzniká skrze omezení možností a pravidel. V prostředí internetu jsou tak možné pouze jisté možnosti jednání. Tato redukce je dokonce přísnější než redukce pracovníka či pracovnice u pásu, přestože jsou jasná pravidla, kterými se musí řídit, tak zde stále zůstává prostor k neuposlechnutí. Fyzický svět totiž stále nabízí více možností než svět virtuální. U *TikTok NPC streamů* je *content creator* taktéž automatizován do určitých reakcí na podněty skrývající finanční odměnu. Zatímco se pojem automatizace objevuje ve vztahu se strojem, tak nemůžeme opomenout tu skutečnost, že nejprve byl automatizovaný člověk, aby mohl být následně podle tohoto vzoru automatizovaný stroj.

Sociální roboti či častokrát ve sci-fi představách samotní roboti reprezentují člověka, snaží se mu podobat, být jeho kopií. Zároveň se člověk sám stává automatizovaným strojem, robotem. Lidské se mění na mechanické a mechanické se mění na lidské. Stává se složité určit, co je nápodobou a co je originálem. Napodobuje stroj člověka procesem automatizace či se člověk automatizuje jako stroj?

Můžeme zde sledovat jistou stromovou strukturu. Reprezentace se vždy odráží od svého ideálního typu, tedy od svého kořene. V případě robota je ona reprezentace postavena na základě kořene-člověka. Povaha reprezentace robota je tedy substituční. Robot má za úkol nahrazovat. To se stává jeho funkcí, která tak tvoří jeho existenci. Těchto stromů si můžeme představit několik, kdy jsou spjaté s jakoukoliv činností, ve které může robot nahrazovat člověka. Onen postup ve stromové struktuře od kořene můžeme spatřit i v díle Philipa K. Dicka. Androidi vznikají jakožto reprezentace člověka, která se přibližuje dokonalosti (Deleuze, Guattari 2010). V uvedených dvou příkladech se potýkáme se zapletením stromu. Stroje se již nestávají reprezentací kořene-člověka, ale samy se stávají kořenem, který je dále napodobovaný člověkem.

Tato obousměrná reprezentace pak vytváří zjednodušení procesu nahrazení, jelikož tato redukce smršťuje pole potřebné k dokonalé (či spíše vylepšené) nápodobě. V případě sociálních botů v prostředí sociálních sítí je ona redukce nejzřetelnější. Zaprvé se pohybujeme v prostoru, kde spolu mohou jednat pouze reprezentace (můj facebook profil je reprezentace mé osobnosti), což značně usnadňuje nahrazení jedné reprezentace za druhou. Uživatel je v prostředí internetu pouze reprezentován, přičemž tato reprezentace vzniká z identifikátorů, kterými je profilový obrázek, uživatelské jméno a další základní údaje. Tyto reprezentace mohou být samozřejmě falešné, jak již bylo naznačeno v předchozí podkapitole. Uživatel například nemusí mít na profilovém obrázku svoji zobrazující jeho reálnou podobu atd. Máme-li tedy systém umělé inteligence-sociální bot, který dokáže

poskytovat takový typ generovaného textu, u něhož se na první pohled nedá určit jeho původ, pak dochází k nerozlišitelnosti.

Nyní se ovšem dostáváme do otázky, co je tedy reprezentováno, když ona reprezentace působí oběma směry? Znamená to, že stromová struktura ztrouchnivěla? Stromová struktura stále drží. To, co je zde pozměněno je teritorium kořene-člověka, které se stalo uměle reteritorizované či spíše utlačené. Samotná reprezentace se stává dokonalejší než její originál, a to hlavně kvůli vlastnostem, které jsou s ní spojeny. Robot je poslušný, robot se neunaví atd. Člověk je totiž ideálem pouze ve své funkci. Robot tak není simulakrum člověka samotného, ale pouze jeho funkce efektivní bytosti (Baudrillard 1996, s. 120). Robot jakožto simulakrum pohlcuje svůj originál, jelikož se stává lepším, hyperrealistickým. Tímto se rozpadá vztah mezi originálem a kopií, jelikož originál se sám o sobě stává reprezentací kopie. Nyní tak existuje pouze model, který nemá původ (Baudrillard 1994, s. 3). Je tak vytvořena nová problematika rozdělení dvou kategorií, jelikož samy sebe ovlivňují zároveň. Skrze obousměrné ovlivnění je tak načrtnuta lepší cesta k dokonalosti napodobení. Další příklad této redukce člověka si ukážeme v další kapitole.

Tendence vývoje umělé inteligence a robotiky hnát se za naplněním sci-fi představ je vědomou snahou uskutečnit onu reprezentaci člověka (Eysenck, Eysenck 2021, s. 5). Nejznatelnější jsou tyto tendence právě v oblasti vývoje sociálních robotů, kteří se dokonce podle některých svých designérů mají vymykat kategorii nástroje a být tak agenty schopnými sociálních interakcí. Přesto zde zůstává přítomno čapkovské proctví, jelikož jejich navrhování zasahuje vždy jen do určité oblasti sociální sféry, jako je například práce, ve které nahrazuje svůj originál-člověka (Damiano, Dumouchel 2018). Neustále ovšem najdeme rozdílnost mezi těmito dvěma kategoriemi, což spočívá hlavně v nedokonalosti robotiky, která není s to vytvořit takové androidy jako ti, kteří existovali ve světě Philipa K. Dicka.

Sociální roboti mají své předurčení. Stále padají do klece automatizace, která předurčuje jejich oblast působení, což je věc, která se vymyká sci-fi představám s rebelujícími roboty (srovnej *R.U.R.* a *Blade Runner*). Jejich důležitou vlastností je poslušnost a pevnost v určitém teritoriu. Každý ze sociálních robotů musí být uzpůsoben svému poslání. Automatizace jej připoutává k určitému teritoriu jeho možného působení. Ona představa rebelujícího robota je lákavým spojením ďábelské povahy vědy a navrácení člověka ke své přirozenosti. Antropomorfizace robotů nás dostává do problematiky toho, že to, co bylo původně lidské se nyní přisuzuje i strojům. Neznamená to však, že bychom ztratili rozdíl mezi člověkem a robotem. Pokud je robot dokonalým objektem, pak to není z důvodu, že by byl dokonalé simulakrum člověka, ale protože zůstává objektem v podřízeném postavení (Baudrillard 1996, s. 120-122). Opět nasazení v určité oblasti se stává klíčovým pro zjednodušení nápodoby, jelikož tento prostor zůstává definován pravidly, která vymezují jednání v tomto prostoru, i když jsou při implementaci v lékařství či sociálních službách složitější, nežli ve výrobním procesu továrny. Mechanický stroj nahrazující biologický stroj se možná stává efektivnější v dané oblasti, ale stále není na takové úrovni, aby člověka nahrazoval ve všech oblastech sociálního bytí, ergo se nestává novým člověkem. To je ten důvod pro proměnu vztahu originálu a kopie, jelikož člověk sám o sobě nikdy nebyl ideál, protože ideálem byl model poslušného stroje.

Nerozlišitelnost se tak stává hlavně problematikou, která se týká světa, kde vystupují reprezentace, světa internetu. Ona dokonalá podobnost může nastat i při zdokonalení designu robotů. Nerozlišitelnost však není jedinou věcí, která může otřást s kategorií člověka a způsobit její změnu obsahu. Polidšťování se totiž nemusí týkat pouze vizuální stránky. Sociální roboti jsou zvláštním druhem, který vchází do sociálních interakcí. Jedná se totiž o objekty, které vstupují do světa subjektů. Jedním příkladem může být existence robotky Sophie od společnosti Hanson robotics. Této robotce bylo uděleno občanství v Saúdské Arábii, což vyvolalo ohromné diskuze o chápání robotů v této digitální éře (Sheng, Wang 2022). Nemluvíme tedy jen o tom, jak dokonalou kopii člověka se může robot jevit, ale i o tom, jak robota chápeme, jak k němu přistupujeme, a jak s ním jednáme.

4.3 Jestli to nebude láska...

Nyní se odvrátíme od díla Phillipa K. Dicka a přistoupíme k jeho reprezentaci v podobě filmu z roku 1982 od režiséra Ridleyho Scotta. Ve snímku *Blade Runner* nemá Rick Deckard žádný problém vykonávat zaměstnání Blade Runnera. Nezpochybňuje nástroj, který mu umožní rozeznat mezi člověkem a strojem. *Poslání androida na odpočinek* pro něj není ničím jiným než vypnutím počítače ze zásuvky. Při návštěvě ředitele společnosti, která se zabývá vývojem replikantů, je konfrontován s postavou Rachael, u které na základě testu zjistí, že se jedná o replikantku. Postava Rachael vstupuje do příběhu jako *femme fatale* a stává se tak jedním z počátečních vlivů, které donutí Deckarda přehodnotit jeho úřadující pohled na svět⁸ (Sanders 2008, s. 25). Deckardova přeměna následně vede k tomu, že na konci příběhu opustí svůj post Blade Runnera a utíká se svou novou láskou - replikantkou Rachael. Tomuto činu nemůžeme rozumět jinak než jako přehodnocení kategorií a zapojení umělého člověka do sféry lidské intimity, čímž je přisouzeno lidské umělému.

Nahrazení člověka ve sféře sexuality je znatelné napříč historií. V roce 2005 objevili archeologové ve Švábských alpách v Německu nejstarší dildo na světě. Tento objekt byl dlouhý 28 centimetrů a předpokládá se, že dle svého tvaru se nevyužíval pouze k symbolickým a náboženským účelům, ale i k sexuální stimulaci. Sexuální pomůcky se samozřejmě vyvíjely společně s technickým pokrokem. V druhé polovině 19. století se objevil párou poháněný vibrátor a v 20. století jej vystřídal vibrátor elektrický. Ve stejné době se pomalu začaly vyrábět první sexuální panny, které se od předchozích sexuálních technologií lišily tím, že reprezentovaly celé tělo místo určité části. Dalším rozvoj technologie nemohl vést k ničemu jinému než právě k zdokonalování nápodoby celého těla až k dalším lidským vlastnostem. Panny začínají aspoň z části ožívat a stávají se už něčím víc než pouhým neaktivním kusem silikonu, vznikají sociální roboti navrženi k sexuálnímu uspokojení – sex robots. Fenomén sex robotů se pro nás stává zajímavý právě z důvodu rozpadu kategorie nástroje. I přestože výroba sex robotů pořád vytváří objekt, se kterým se má jednat jako s nástrojem pro sexuální stimulaci, tak je tento pohled narušen skrze vznikající intimitu mezi strojem a člověkem. Sex roboti nejsou jen dalším bezduchým strojem na vykonávání určité činnosti, ale stávají se něčím víc skrze uživatele, který tyto objekty

⁸ V této kapitole se čerpá z filmu, jelikož se tato adaptace Dickova díla ubírá jiným směrem, než je poselství knihy. Na vztahu replikantky Reachel a Ricka Deckarda je lépe reprezentovaná myšlenka onoho potlačení esenciálních rozdílů a navázání milostného vztahu mezi těmito postavami.

polidšťuje. Toto přetváření objektu na subjekt je znatelné v příkladech subkultur, které dávají přednost vztahům se sexuálními pannami. Příkladem takového přístupu je online komunita iDollators. Členové a členky této komunity vnímají své sexuální panny nejen jako sexuální pomůcky, ale jako životní partnery a partnerky (Danaher, McArthur 2017). V rámci rozvoje LLM se rozmáhají i aplikace, které jsou souhrnně nazývané jako AI girlfriend. V aplikacích jako je *Replika*, *Intimate* nebo *DreamGirl* si může uživatel vytvořit svého ideálního partnera či partnerku, se kterými může vést i hyper-realistické videohovory (India 2023).

Tento přechod z objektu na subjekt v sexuální sféře je viditelný i v opačném principu, tedy přechodu subjektu na objekt. Nejčastěji se tato objektifikace týká žen, které jsou redukovány na objekt sexuální touhy, či na určité části těla jako jsou hýždě, nohy, ňadra etc (Baudrillard 1996, s. 99–100). Sex robot či AI girlfriend vystupují právě z druhé strany a dostávají se na ono pomezí jako se dostávají ženy. Ženy jsou postaveny tímto procesem do subjektu-objektu a jejich umělé nápodoby se dostávají do stavu objektu-subjektu. Lépe řečeno: to co dělá ženu objektem dělá sex roboty a AI girlfriends subjektem. Nastává tak proto, že jsou tyto milostné alternativy pojímány mimo své teritorium pouhého objektu. Tato deterritorializace nastává právě skrze vybudování milostného vztahu k těmto sexuálním pomůckám. Takovéto reteritorializační tendence jsou znatelné i v přemýšlení ohledně marketingu sex robotů. Matt McMullen, ředitel firmy *RealDoll* se například nechal slyšet, že chce, aby si lidé vybudovali opravdovou emoční vazbu nejen k pannám, ale i k jejich charakteru (Danaher, McArthur 2017, s. 23). Ono emoční provázání je pak věc, která vytváří přesun z objektu k subjektu.

Na první pohled bychom měli tendenci říci, že dochází k jisté reteritorializaci, když se objekt určený ke specifickému druhu činnosti dostává i mimo své ohraničené teritorium. Nahrazení lidského jedince v milostném vztahu, pak přináší tendence k přehodnocení kategorie člověka. Stejně jako s příkladem z filmu *Blade Runner*, kdy je Deckardův vztah k androidům přehodnocen na základě touhy po androidce Rachael. Opak je však pravdou. Sex roboti nejsou něčím svobodným schopným svého vlastního rozhodnutí. Jsou pouze produktem, který je designovaný k prodeji a uspokojení uživatele. To, co nastává v komunitě iDollators, je pouhou podporou substitučního charakteru daných nástrojů. Koncept sexuální panny se z pouhého nástroje sloužícího k sexuálnímu potěšení se stává nejen objektem, ale něčím, co je dle jejich hlediska životním partnerem či partnerkou. Uživatel je tím klíčovým arbitrem, který rozhoduje o teritoriu, do kterého může tento objekt zasahovat, a tím zvýšit jeho úroveň substituce. Tyto produkty samy o sobě nabízejí svému uživateli něco více než jejich originál. Stávají se lepšími jako v případě automatizovaných strojů v pracovním procesu. Nastává to v aspektech vyšší efektivity bez únavy a tak dále. Stromová struktura je tak zachována.

Podoba vztahu člověka a sex-roboty či AI girlfriend není něčím novým, jelikož se snaží zprostředkovat ty stejné prvky jako vztah člověka s člověkem. Ovšem není tím stejným. Rozdílnost se ukazuje v té skutečnosti, že se stále jedná o produkty, které jsou designované způsobem, aby svému uživateli vyhovovaly. Uživatel má možnost svého sex roboty utvářet. Důležitou roli zde hraje i robotův poslušnost. Sex robot tak není dokonalou reprezentací člověka, ale dokonalou reprezentací ideálu milostného partnera či partnerky. Ona robotův poslušnost a přizpůsobení uživateli zde vytváří hyperrealistickou zkušenost. V důsledku tak simulakra předkládá realističtější požitek ze vztahu, než opravdová žena. To je

pozorovatelné například v komunitě *iDollators* či prostřednictvím AI Girlfriend, které se vyznačují nejen hyperrealistickým ztvárněním těla (bez fyzické formy), ale i svými reakcemi na vstupy uživatele. Jsme tak na více dokonalém stupni charakteru objektů, na který naráží Jean Baudrillard: „*Jsme bráni jako cíle objektu a objekt nás miluje. A protože jsme milováni, cítíme, že existujeme: jsme „personalizováni“.*“ (1996, s. 170). U produktů tohoto typu to platí dvojnásobně, jelikož to je jejich poselstvím.

Objekty tak pronikají do sfér vyhrazených původně lidem, a tak narušují hranice těchto kategorií. Neděje se tak ovšem ve smyslu příběhu *Blade Runner*. My nedocházíme k nesprávnému pojmenování, jelikož napodobení nedosahuje dokonalosti. Ono dosahování dokonalosti napodobení, ale není klíčovou věcí pro robotí dokonalost. Poslušnost je tím důležitým, co vytváří z těchto substitučních nástrojů něco lepšího, než je jejich originál. O robota nejde. Ten nikdy nebude mít špatnou náladu, nebude unavený z práce. Jediným důležitým prvkem je uživatel. Uživatel, který je postavený do role pána. Proto se stávají i sex roboti oblíbení mezi lidmi s tendencemi k sexismu (Leshner, Johnson 2024). Toto mocenské postavení se sice může na první pohled takto jevit, ale tento vztah má více rozměrů a nedotýká se pouze člověka a daného robota. Důležitou roli zde hraje to, že robot zůstává objektem, se kterým je obchodováno, je produktem, který uživatel poptává. Mocenská struktura je zde pouhou *mise-en-scène*, za kterou se ukrývá odlišná dynamika mezi subjektem a objektem. Tato dynamika bude více přiblížena v kapitole *Kdo je tady pán?*.

5 Matrix - život v simulaci

Od sci-fi v knižní podobě se přesuneme ke stříbrnému plátnu. Filmová série *The Matrix* upoutala diváky revolučními speciálními efekty, ale mimo to se do paměti zapsala nejen vizuální podívanou, ale taktéž námětem příběhu, který prezentuje. Základní premisa filmu je vystavěna na základě knihy Williama Gibsona s názvem *Neuromancer* a předává zápletku v podobě konfliktu člověka a stroje. Nečiní tak jednoduše jako například *Terminátor*, který staví lidi a stroje do přímého válečného konfliktu, nýbrž už pracuje s důsledky tohoto konfliktu. Lidé jsou zotročeni a zavřeni do simulované reality nazvané Matrix. Z lidí napojených na Matrix, pak stroje čerpají potřebnou energii (Roberts 2006, s. 286). Sestry Wachowské⁹, režisérky filmové série se nechaly slyšet, že jednou z inspirací pro film *Matrix* (1999) bylo i dílo francouzského postmoderního filosofa Jeana Baudrillarda, jehož kniha *Simulacra and Simulation* se dokonce objevila ve filmu jako rekvizita. Baudrillard ovšem odmítl tento film jakožto prezentaci jeho myšlenek. Dokonce posměšně označil samotný snímek za takový film, který by byl vytvořen zlým matrixem. Základní rozpor se nachází v pojetí simulace jako takové, jelikož simulace předestřená sestrami Wachowskými se spíše od Platónského dualismu mezi realitou a simulací, z níž se dá uniknout. Sám Baudrillard pak upozorňuje právě na to, že došlo k rozpadu těchto dvou kategorií a nyní se nacházíme v hyperrealitě (Laist 2012).

⁹ V době natáčení Matrixu se ještě jednalo o bratry později však postoupily genderovou tranzici.

Krátký příběh od Jorge Luis Borgese vypráví o kartografech, kteří se snaží vytvořit mapu říše, která bude tak detailní, že nakonec pokryje celou říši. Mapa se ovšem rozpadá na malé kousky, které se ztrácejí v poušti. Dnešní abstrakce však není dvojníkem, zrcadlem nebo obrazem. Je to vytváření modelů reálného bez původu a reality: *hyper-reality*. Éra simulací je zahájena likvidací všech referencí a jejich umělého oživení v systémech znaků. Je to otázkou nahrazení znaku reálného za reálné, které odstrašuje každý skutečný proces prostřednictvím svého operativního dvojníka. Reálné už nikdy nebude mít možnost produkovat samo sebe. Hyper-reality je nyní chráněná před imaginárním a před jakýmkoli rozlišováním mezi skutečným a reálným. Imaginární ponechává prostor pouze pro orbitální recidivu modelů a pro simulované vytváření rozdílů (1994, s. 1–2). Takto popisuje Baudrillard koncept hyper-reality na začátku své knihy *Simulacra and Simulation*. Tento koncept pojímá jako důsledek úřadujícího systému. Zaznamenáváme tak jakýsi rozpad duality mezi realitou a simulací.

Koncept simulace je častokrát spojovaný s počítačem vytvořeným virtuálním prostředím. Nabízí se tak srovnání se současnou technologií, jako je virtuální či rozšířená reality. Nyní je možné skrze hardware v podobě brýlí a ovladačů vstoupit do simulace, virtuálního prostoru. Na tomto prvku nového média je skvěle znatelná dualita mezi realitou a simulací. Mám na sobě brýle, což znamená, že jsem v simulaci, virtuálním světě. Sundám si brýle a jsem v realitě, fyzickém světě. Místo Matrixu nyní vzniká *Metaverse* z dílny Marka Zuckerberga. Společnost *Meta*, jež má pod svou záštitou populární sociální sítě, jako je *Facebook*, *Instagram* či *Whatsapp*, se snaží zprovoznit takzvaný *Metaverse*, který má za cíl vytvořit právě platformu, vycházející z principů zmíněných sociálních sítí a virtuální reality, která má propojovat lidi v rámci pracovního poměru, komunikace s přáteli či návštěvy koncertu (*Meta.com*). *Meta* tak nabízí onu Matrix dualitu mezi fyzickým a virtuálním světem, ve které se virtuální svět stává obrazem toho reálného, tedy nabízí stejné či obdobné možnosti sociálních interakcí.

Dalším velikánem, který přichází s uvedením na trh jako zařízení ze sci-fi představ je firma *Apple* s novinkou s názvem *Apple Vision Pro*. Brýle *Apple Vision* neposkytují vstup do virtuální reality, ale zprostředkovávají rozšířenou realitu. Díky těmto brýlím může uživatel pozorovat rozhraní aplikací přímo v prostoru před sebou (Divínová 2024). Jak Platón, tak ani Baudrillard nepojímají simulaci způsobem počítačem vytvořeného virtuálního prostředí. V následujících podkapitolách se budu snažit přiblížit problém simulace (nevirtuální) na dualistickém pojetí, které naráží na problém nerozlišitelnosti a následně ve více Baudrillardovském duchu, kdy si ukážeme, že se toto tradiční dichotomní rozdělení rozpadá. Moje analýza se netýká zmiňované technologie, které zprostředkovávají virtuální či augmentovanou realitu, ale obrací se na principy, které jsou přítomny i u předchůdců těchto médií, tedy v internetu a sociálních sítích.

5.1 (Ne)rozlišitelnost znovu

Koncept nerozlišitelnosti jsme si představili v předchozí kapitole v návaznosti na příběh Philipa K. Dicka. To, čemu jsme se věnovali v předchozí kapitole, se týkalo oblasti kategorie

člověka a stroje. Produkty vycházející z umělé inteligence ovšem neslouží pouze k reprezentaci člověka, ale objevují se především v podobě nástrojů s určitými schopnostmi.

Dokonalá podobnost u lidí a robotů, a tedy rozpad či sloučení těchto dvou kategorií je stále ve sféře sci-fi představ. Velká revoluce generativních modelů přinesla novou sféru nerozlišitelnosti, která se týká produktů. Spoustu produktů, které mohly být dříve dílem pouze lidské ruky, nyní zvládne vytvořit generativní model umělé inteligence, a to na takové úrovni, že se stává téměř neodlišitelná od lidského počinu. Umělá inteligence nyní může vyhrávat prestižní fotografické soutěže s fotografiemi, které ani nejsou obrazem fyzického světa (ČT24.cz 2023). Dokáže taktéž psát bakalářské a diplomové práce, které je nemožné odhalit pomocí antiplagiátorských systémů, neboť každý vygenerovaný text je originál (Petiška 2022). Tato skutečnost vzbudila velké obavy v rámci školství s představou toho, že žáci začnou využívat generativní umělé inteligence pro vytváření svých domácích úkolů. Jádro této obavy spočívá právě v tom, že vygenerovaný text se nedá odlišit od textu napsaného studentem (až na to, že může být kolikrát kvalitnější). Jak se tedy vyrovnat s tím, že nedokážeme rozlišit mezi originálem od studenta a originálem od stroje, když není s přesností říct, který je který? Řešení problému se nabízí skrze jinou podobu umělé inteligence, klasifikátoru. Nástroj *zeroGPT* dokáže tyto kategorie určit, a dokonce ukázat, jaká část textu byla napsána a jaká vygenerována (*zeroGPT*). Riziko nerozlišitelnosti plynoucí z generativní umělé inteligence je řešeno pomocí klasifikační umělé inteligence, kdy se tak obor stává řešitelem problémů, které sám generuje. Podobně jako je to v případě technicko-ekonomického pokroku (srovnej Beck). Zároveň nemůžeme opomenout onu podobnost s Voight-Kampffovou stupnicí jako nástroje na oddělení dvou kategorií, které ztrácejí svoji oddělitelnost.

I v tomto případě nerozlišitelnosti se vracíme do augmentované reality internetu. Internet je svou podobou naprosto revoluční, neboť kombinuje předchozí média: je interaktivní jako telefon a telegraf a je mas médiem jako rádio a televize. Od svého vzniku měl velký vliv na proměnu téměř všech aspektů sociálního života. Do této platformy se přesouvá část lidského života, kterou je nepřekvapivě komunikace, ale dokonce i nakupování či simulace sexuálního života v podobě pornografie. Internet se stává integrální součástí našich životů a v současnosti si nemůžeme téměř představit život bez něj (Bargh, McKenna 2004). V předchozí kapitole jsem předestřel onu zákeřnost, že se za profilovým obrázkem může skrývat kdokoliv a že se při současném úspěchu generativní umělé inteligence setkáváme s osobami, které už nemají své ukotvení ve fyzickém světě. Můžeme se tak setkat s obrazy lidí, kteří existují pouze jako napodobenina modelu lidské podoby. Už to nejsou fotografie, které by vznikaly otiskem skutečnosti, ale pouze jako podoby něčeho, co existuje abstraktně. Opět se musíme ptát, jak dokážeme rozlišit mezi skutečným člověkem, tedy takovým člověkem ve fyzickém světě, u kterého můžeme hledat identitu s jeho profilem na internetu a sociálním botem, který je reprezentován generovanými produkty.

Přemýšlejme o člověku jako o fyzické entitě, která musí přistupovat k internetové službě skrze jakousi svou redukovanou podobu, která se sestává z určitých základních informací. Mezi člověkem a profilem by pak měla existovat identita, tedy to, že daný člověk je právě daným profilem. Samozřejmě je tato věc daleko zapeklitější, jelikož jeden člověk může ovládat i více profilů apod. U profilů spravovaných umělou inteligencí však tato identita mizí,

jelikož už profil nemůžeme spojit s ničím existujícím ve fyzickém světě. Profil se tak stává existencí sám o sobě. Vlastní schopnost rozlišování si člověk může otestovat v projektu *whichfaceisreal.com*. Na této webové stránce jsou před vás předloženy dvě fotografie lidských obličejů. Na jedné fotografii je reálný obličej a na druhé fotografii je obličej vygenerovaný umělou inteligencí. Úkolem návštěvníka stránky je odhalit, který obličej je reálným (West, Bergstrom 2019). Po několika pokusech se člověk dokáže soustředit na detaily, které mu umožní správně odhalit vygenerovaný podvrh lidské tváře. To však stále není důvod k oslavám. Co kdyby takových tváří bylo tisíc či milion? Co kdyby se tyto tváře objevovaly a mizely? Dokázali bychom to bezpečně určit?

S internetem přichází exploze informací, které jsou metaforicky řečeno umístěny na jednom místě. Díky vynálezu chytrých mobilních telefonů se stal svět internetu přítomným téměř kdekoli a kdykoli. Namísto toho, aby měli lidé téměř okamžité odpovědi, tak se z internetu a smartphonů stalo pouze prostředí skrze, které se může šířit *bullshit*.¹⁰ Technologie celkově problému *bullshitu* přispívá. Internet poskytl jistou demokratizaci šíření obsahu, což pochopitelně vedlo k jeho celkovému zvýšení. Další skvělý pokrok, který produkuje svoje rizika. S růstem obsahu, a tedy informací došlo k proměně šíření informací. Už se nepotkáváme jen s tvářemi známými z televizních obrazovek, které nás informují o světě kolem nás či příbězích celebrit. Na internetu se setkáváme i s naprostými cizinci, kteří nabízejí svůj pohled na svět a vlastní příběhy. Šíření informací se proměnilo do jisté ekonomiky a platidlem se staly kliky. Každé kliknutí znamená peněžní odměnu v podobě reklamního revenue. Ke zpravodajství se přidaly kvanta nejrůznějších informací bez obsahu jako články s nadpisy: *Sedm koček, které vypadají jako Disney princezny*. Tedy jsme zaplavení takovými portály, které stojí na *native advertising* předávající reklamu jako obsah. Algoritmy sociálních sítí ještě přilévají vodu do ohně tím, že jsou designované, aby uživatele udržely na platformě namísto toho, aby udržovaly informovanost jedince (Bergstrom, West 2020, s. 34). Nacházíme se v situaci, kde je více a více informací, ale méně a méně významu (Baudrillard 1994, s. 79). Jsme přehlčeni. Informace se na nás chrlí z obrazovek a my je nedokážeme zastavit. Naše mozky nejsou uzpůsobeny na takovou informační záplavu, která k tomu všemu pozbývá smysluplného obsahu (Klingberg 2009). Vlivem těchto nových médií můžeme dokonce sledovat i zvyšování deficitu naší pozornosti (Carstens, Doss, Kies 2018).

Do hledáčku nám vstupuje další fenomén, který rezonuje v současných diskuzích, a tím jsou dezinformace. Díky demokratizaci šíření informací se vaše zpráva může dostat až na druhý konec světa. Zprávu může poslat kdokoli, kdo má přístup k internetu a podobně. Nepřekvapivě s sebou záplava informací přináší i falešné informace v podobě misinformací a dezinformací. *Dezinformace* jsou takovým druhem falešné informace, který je zamýšlený. Evropská komise je definuje jako prokazatelně zavádějící informace, které jsou šířeny za účelem zisku či záměrného oklamání veřejnosti. *Misinformace* jsou pak falešnou informací šířenou bez úmyslu škodit. Sociální sítě se svou povahou se staly rájem pro tyto dva nešvary (Bontridder, Pouillet 2021). Dle studie z roku 2016 přijímá zprávy ze sociálních sítí majoritní část Američanů (Gottfried, Shearer 2017). Další studie z roku 2018, pak zjišťuje, že kolem

¹⁰ Pojem *bullshit* přejímají Bergstrom a West od filosofa Harry Frankfurta a rozeznávají jeho dva druhy, což je *old-school bullshit* a *new-school bullshit*. *Old-school bullshit* používá líbivé rétoriky a *new-school bullshit* používá jazyku matematiky a statistiky.

2,6 % amerických článků je nepravdivých. Pokud by každý Američan četl jeden článek denně, tak to znamená, že téměř 8 milionů lidí četlo nepravdivý článek. Některé misinformace jsou lehké povahy a mluví třeba o tom, že Taylor Swift chodila s Josephem McCarthym, který zemřel 40 let před jejím narozením, ale některé mají daleko strašlivější dopady (Bergstrom, West 2020, s. 42). Přehlcení informacemi, pak dává dezinformacím nádech života v podobě takzvaného *sleeper effect*, kdy se děje to, že je uživatel konfrontovaný s nepravdivou zprávou a v průběhu času zapomene na nedůvěryhodnost zdroje, ze kterého informaci získal. Informace zůstane v paměti uložena a je k ní přístupováno jako k pravdivé (Kumkale, Albarracín 2004). Dezinformace díky explozi informací vstupují do našich diskurzů a utvářejí naše představy o světě.

Veškeré tyto fenomény naznačují zesložitění rozlišitelnosti mezi pravdou a lží. Může se zdát, že tedy stojíme na bodě, kdy musíme zpochybňovat každou informaci, která k nám doputuje. Se současným množstvím informací, které denně získáváme, bychom ovšem nedělali nic jiného, než pouze ověřovali. Umělá inteligence zde toto rozlišování ještě stěžuje, jelikož je schopna generovat takové produkty, které jsou naprosto odříznuté od reality. Příklad pravdivých a falešných informací nás nabádá k dualistickému pojetí. Buď padneme jako oběť dezinformacím a uvězníme se tak do simulace, která nám předkládá model naší skutečnosti, nebo identifikujeme informaci jako falešnou, spolkneme červenou pilulku a zůstaneme v realitě. Bohužel tak jednoduché to není.

5.2 Nová realita

Dezinformace a misinformace se šířily dávno před vznikem internetu, a tudíž i dávno před vznikem umělé inteligence. Jak do tohoto všeho zapadá? Umělá inteligence je pouze další voda na mlýn této informační exploze, která se na první pohled zdála tolik pozitivní. Nástroj umělé inteligence posiluje falešné informace na Internetu dvěma způsoby: 1) Vytváření nových možností manipulace s texty, obrázky, audiem či videi. 2) Systémy umělé inteligence sloužící k udržení uživatele na platformě přispívají k účinnějšímu šíření falešných informací (*filter bubble*). K druhému jmenovanému způsobu se ještě vrátíme v následující kapitole.

Nejznámější praktikou, kterou v dnešní době výrazně usnadňují nástroje umělé inteligence, je tzv. *deepfake*. Deepfake může být padělaný obrázek, audio či video. Před příchodem umělé inteligence jsme se setkávali například se záměnou obličejů. Tato praktika se ujala například v porno průmyslu, kde byla tvář pornoherce či pornoherčky zaměněna za tvář jiné známé osoby. Dnešní systémy generativní umělé inteligence si dnes vystačí pouze s několika megabyty či gigabyty referenčních dat a je schopna vyprodukovat celý audiovizuální produkt (Bontridder, Pouillet 2021). Fenomén deepfake je přesně tím jevem, na které upozorňuje stránka *whichfaceisreal.com*. Celý vývoj ovšem vede k demokratizaci těchto nástrojů a k jejich jednoduššímu použití. Kolikrát tedy ani nečekáme, že bychom mohli na takový padělek narazit, jelikož jeho vytvoření vyžaduje čas a profesionální práci. Uživatel tak ani nestojí před rozlišováním mezi realitou a falší. Internet je ovšem místo, které by nás mělo přivádět do neustálého pochybování.

Deepfake se ve virtuálním prostoru internetu chová podobně jako dezinformace. Na jedné straně vyskakuje jako neškodný žert, který by měl být každému jasný, že je pouhou parodií, ale na druhé straně vystupuje jako nástroj pro změnu veřejného mínění. V rámci posledních slovenských voleb jsme měli možnost se s touto negativní vlastností *deepfake* setkat. Před otevřením volebních místností byla do světa internetu vypuštěna falešná audio nahrávka rozhovoru předsedy Martina Šimečka s reportérkou. Tato nahrávka nepřekvapivě vrhala na Šimečkovu stranu negativní světlo. Celá nahrávka byla vytvořena pomocí umělé inteligence, ale pro mnohé se stala „realistickým“ zachycením rozhovoru, který však nikdy neproběhl. Díky času zveřejnění nahrávky je složité určit, jestli její zveřejnění mělo opravdu vliv na výsledek voleb, ale zůstává jasné, že se umělá inteligence stává nástrojem, která dokáže skvěle manipulovat s pravdou (Procházková 2023). Dostáváme se pryč od rámování informací do určité podoby, ale k jejich tvoření. *Deepfake* produkty se stávají důkazem šířených dezinformací. To vytváří model dezinformačního světa ještě reálnějšími, jelikož nyní k dezinformaci existuje i důkaz, který ovšem s realitou nemá vůbec nic společného. Není to však pouze sféra politiky. Jedním z dalších odstrašujících případů je kyberšikana čtrnáctileté Mii. Tato kyberšikana probíhala pomocí *deepfake* fotografií a videí, ve kterých byla dívčina tvář přiřazena na nahá těla pornohereček. Tyto události vedly až k dívčině sebevraždě (Davies 2024). Přestože oba *deepfake* produkty nebyly reálné, tak jejich dopad reálný byl. Thomasův teorém říká, že pokud člověk definuje situaci jako reálnou, je reálnou ve svých důsledcích (Thomas, Thomas 1928, s. 572). Lidské jednání se neuspořádává pouze na základě faktů, ale spíše na základě domněnek, které jedinci mají. Vytváření našich představ o světě, i když pramení z neověřených informací, vede k tomu, že se jejich dopad stává reálným. Vztáhneme-li teorém na uvedené příklady, tak definování reálné situace je: falešný rozhovor a šířená falešná pornografie. Reálným dopadem pak: ovlivnění voleb a sebevražda.

Druhým negativním způsobem, jak umělá inteligence ovlivňuje šíření dezinformací, je fenomén zvaný *filter bubble*. Jedinec při svém pohybu v kyberprostoru za sebou nechává digitální stopu v podobě dat. Tyto data se stávají surovinou, která vylepšuje dané platformy, které sbírají tato data. Obsah nabízený skrze platformu pak musí být personalizovaný, tj. musí to být takový obsah, který uživatel vyhledá či lépe: takový, po kterém uživatel touží, aby jej platforma udržela, z čehož jsou jí odměnou další data uživatele, a tedy lépe personalizovaný nabízený obsah. Jednoduchý zákon nabídky a poptávky, pokud platforma sociální sítě poptává čas, který uživatel stráví u obrazovky, musí nabízet takové zboží, které chce uživatel vidět. K uživateli se tak dostane pouze nějaký druh obsahu, u kterého se jeví, že je uživatelem chtěný. Kruh se uzavírá. Vzniká nám fenomén *filter bubble*. Neviditelná bublina, která vytváří kategorie konzumentů různého obsahu. Obsahem může být cokoli. Jsme ve sféře ekonomiky. Nabízené je zboží fyzické i digitální. Mezi digitální zboží ovšem patří i informace a to i takové informace, na jejichž základech stavíme naše porozumění světu. Jsou to informace, které utvářejí naše vědomosti o světě, náš diskurz (Pariser 2014, s. 10–11). Je tedy jasná ona negativní vlastnost *filter bubble* v problematice dezinformací. Uživatel, který je náchylný ke konzumaci tohoto závadného obsahu je tímto obsahem i více zásobován, což vede k jeho uzavření do jakési simulace reality tvořené na základě falešných informací. Informace, které o světě přijímáme skrze média vytváří naše vědění o

světě na jehož základě stavíme model světa. Tento model světa funguje jako mapa, dle které ve světě operujeme.

Samozřejmě, jak jsme si několikrát opakovali, tak umělá inteligence je pouhý nástroj a není ustanoven pouze v teritoriu negativních dopadů. Stejně jako může být použita k produkci dezinformací a dalšího klamavého obsahu, tak stejně může být použita i k jeho odhalování. Podobně jako jsme si ukazovali s příkladem nástroje *zeroGPT*. Jednou variantou jsou automatizované *fact-check* nástroje, které se staly velmi úspěšné v detekování a mazání nelegálního obsahu či identifikování botů v prostředí sociálních sítí. Dle Facebooku je 99,5 % teroristického obsahu, 98,5 % falešných účtů, 96 % nahoty a 86 % násilí rozpoznáno pomocí nástrojů umělé inteligence. Nic není, ale bez chyby. S využitím nástrojů umělé inteligence, které jsou neustále ve vývoji, se dostáváme do problému falešných pozitiv/negativ. To znamená, že legitimní zprávy jsou kategorizované jako nelegitimní a naopak (Kertysova 2018). Umělá inteligence se tedy v tomto případě stává Rickem Deckardem, který opět musí rozhodnout, jestli jde o reálné, pravé či umělé, falešné. Kategorizování ani při této automatizované bezchybnosti stroje není přesné a stále naráží na problém nerozlišitelnosti mezi těmito dvěma kategoriemi. Otázkou tedy zůstává, jestli se postupným vývojem dokážeme dostat na 100 % úspěšnost rozpoznání či se zasekneme na tomto bodu jisté chybovosti, jelikož vytváření falešných informací, botů atd. se taktéž bude zlepšovat. Produkt oboru umělé inteligence je v tomto smyslu opět dvousečnou zbraní. Na jedné straně přináší rizika v podobě nástrojů umožňující manipulaci s pravdou. Na druhé straně vystupuje jako arbitr, určující škodlivý obsah. Problémem je zde ona nedokonalost pojmání rizik kvůli jisté míře chybovosti. V některých případech je to silný nedostatek i v *post hoc* aplikaci, tedy že je dezinformace identifikovaná až po jejím šíření internetem.

Stále se zde ovšem držíme v dualistickém pojetí. Uživatel chycen v simulaci stále může uniknout, když si uvědomí, že obsah, který přijímá je lživý. Informace k nám ovšem nikdy nepřichází v čistém objektivním smyslu. Média mají tendenci zprávy nějakým způsobem rámcovat, což ovlivňuje jejich poselství (Entman 1993). *Filter bubble* není pouze fenoménem, který by uzavíral do simulace prostřednictvím falešných informací, ale celkově uzavírá jedince do simulace pomocí prezentace modelu světa tím, jaký obsah uživateli doporučuje. Nemůžeme tak tvrdit, že se utváří pouze dvě skupiny uživatelů, kdy by jedna skupina přistupovala k falešnému modelu světa a druhá k realitě. Obě skupiny přistupují k modelům světa, které jsou utvářeny pomocí doporučovacích algoritmů. Vytváření těchto modelů světa se přesouvá od reprezentace reality k reprezentaci simulace, jelikož se vztah mezi realitou a modelem vytrácí. Nyní už nejde o reprezentaci reality, ale reprezentaci modelu reality. Jsme tak připoutáni k modelu světa, který je vytvořený pro naše potěšení, pro zvýšení naší aktivity na platformě. Jsme obětí nové reality. Nová realita se tak postupně stává simulakrou třetího řádu a je pravdivější než realita samotná. Mapa z informací, která nám měla pomoci se orientovat ve světě, jej nyní celý zakrývá. Potýkáme se s odlišným koncem, než měl Borgesův krátký příběh. Namísto mapy se nám rozpadá samotná říše.

Simulace nemusí vznikat jen v rámci virtuálního prostředí, ale i jako abstraktní souhrn vědění, které o světě máme. To samozřejmě má signifikantní dopad na naše jednání, které se strukturuje na základě tohoto vědění. V následující kapitole se zaměřím na to, jaký další dopad může mít toto modelování světa za pomocí algoritmů umělé inteligence.

6 Kdo je tady pán?

Dostáváme se k poslednímu tématu týkající se mocenské struktury mezi člověkem a strojem či vztahu subjektu a objektu. Toto téma se objevuje napříč díly sci-fi a do jisté míry se dotýká všech problémů, které jsem zde analyzoval.

Oblast poslední problematiky se dá vyzdvihnout na základě filmu *Terminátor* (1984) od Jamese Camerona a z něho vznikající *Terminátor narativ*. Sci-fi nám mnoha způsoby předkládá zkázu lidstva skrze technologie. V dramatu *R.U.R.* se zjevuje v podobě robotí vzpoury a následném absolutním nahrazení. V případě snímku *Matrix* se jedná o důsledek války mezi stroji a lidmi, která vyústila v zotročení lidstva a jeho přesunutí k pasivitě objektu do podoby generátoru energie. Ve snímku *Terminátor* je to konflikt mezi lidmi a stroji, který se mění v boj o přežití. Přestože základní myšlenka zůstává stejná, tak se poselství děl liší v určitých detailech. *R.U.R.* a *Terminátor* předkládají absolutní nahrazení, které má vést k vyhlazení lidstva, jež bude nahrazeno dalším vývojovým článkem, lidským výtvozem - robotem. Můžeme tak pozorovat výhru reprezentace, která se zbavuje svého originálu, jelikož jej překonává ve všech ohledech. Obě díla taktéž poskytují určité vyplnění neo-ludditského scénáře, kdy se technologie vědomě pokusí ohrozit lidstvo, které nemá jinou možnost, než se hrozbě postavit či jí být zdolán. Film *Matrix* na druhou stranu předkládá trochu jiný pohled na věc. Taktéž je zde přítomná válka, ale již zde není přítomná lidská genocida. Stejně jako systémy heteromatizace vyžadují člověka, tak stejně vyžaduje člověka i *Matrix*. Namísto nahrazení je zde znatelný proces reteritorializace, kdy je člověk, jakožto aktivní činitel reteritorializován do pasivního postoje. Otázku, kterou se v této kapitole budu snažit zodpovědět je: *Vládne nad stroji my nebo stroje vládnu nad námi?*

Jak již bylo uvedeno v první kapitole, tak veškeré katastrofické scénáře přítomné ve zmíněných sci-fi dílech pracují s takovými systémy, které fungují jako AGI. Snímek *Terminátor* ji předkládá jako systém umělé inteligence s názvem *SkyNet*. Čapkovu *R.U.R.* s ní zase pracuje jako s latentním produktem, tedy roboti jsou vytvořeni jako umělý pracovníci, ale mají vědomí. Právě tato chyba vede k robotí vzpouře, která je vyobrazena na symbolice biblického mýtu stvoření. Robot je hrozbou pro lidstvo, protože je stvořen dle obrazu člověka. Stává se reprezentací, dokonalým dvojníkem, který ve výsledku překonává člověka na všech rovinách (jako v případě *Blade Runnera* jsou roboti také empatictí jako lidé). To, že mají Rosumovi roboti „vědomí“, je tím rizikem, které nakonec vede k zániku lidstva. Robot se na základě uvědomění si svého podřadného postavení vymezuje vůči svým stvořitelům. Jak ovšem Jean Baudrillard upozorňuje, tak robot nedosahuje své dokonalosti tím, že je dvojníkem člověka, kdy se stává nerozlišitelným od svého originálu, ale právě to, že stále jasně zapadá do kategorie objektu-otroka (1996, s. 120). Robot je poslušný stroj. Robot není nic jiného než otrok. Rebelující robot je více člověkem a méně bezduchým strojem. Sama Čapkova myšlenka *R.U.R.*, jak již bylo řečeno, je vystavena na metafoře socialismu. Zde se ukazuje problém reprezentace, kdy robot dosahuje modelu člověka i skrze antropomorfizaci (Čapek přináší robota jako metaforu pro dělníka), která se promítá do představy rebelujícího robota. Součástí modelu robota se stala poslušnost, jež vychází z jeho designu. Slepé následování pravidel, díky kterým může být automatizovaným, a tedy se jevit jako

dokonalým. Na tento jev jsem poukazoval v podkapitole *Stávání se strojem a stávání se člověkem*.

Roboti a umělá inteligence jsou častokrát spojováni spíše s chladnou racionalitou. Jako se ukazuje v případě toho, jak Rick Deckard na základě tohoto stereotypu vnímá androidy, jelikož bod jejich odlišení od člověka leží právě v nedokonalé empatii. Realističtějším problémem na místo probuzení emocí či vědomí a následnému vzteku vůči lidem, je právě chladná racionalizace. Představme si, že máme umělou inteligenci na výrobu kancelářských sponek. Tato výroba v sobě nemá nic škodlivého. Systém ani nemá žádné uvědomění si svého nerovného mocenského postavení. Neví, že je otrok a ani to nezjistí. Nebezpečí se nachází ve špatně definovaném příkazu, který může být například: *maximalizuj výrobu kancelářských sponek*. Umělá inteligence se pak bude snažit vytvořit co nejvíce sponek. Takto otevřený úkol by pak vedl k nekonečné výrobě, při které by byly vyčerpány veškeré zdroje a pohlceny všechny hrozby ohrožující daný úkol¹¹ (Bostrom 2014, s. 149). V případech, které se v minulosti staly, můžeme pozorovat docela odlišný vývoj. Umělá inteligence byla vždy nějakým způsobem zabezpečena, než byla použita ve veřejném prostoru. Například uváděný případ s umělou inteligencí od *Facebooku*, která se chovat v rozporu se jejím designem, či rasistická Tay od *Microsoftu*, která byla po jejím neslavném nástupu okamžitě stažena. *ChatGPT* má zase vymezeno na základě určitých omezení to, jaké věci může generovat a jaké ne. Klíčovou úlohou v zabránění vzpoury strojů se tak zdá být udržení panství člověka nad strojem, ergo udržet stroj poslušný.

Ve sci-fi dílech je onen konflikt vždy rozdělen dichotomicky na my a oni, lidé proti strojům. Současný stav věcí ovšem funguje odlišným způsobem, a to z toho důvodu, že neexistuje přímý konflikt, který by vyžadoval takovéto rozdělení. Jinými slovy, technologie není na takové úrovni, aby všechny jedince dokázala konvertovat k ludittské ideologii. Nestává se vědomým subjektem, který si sám definuje cíle, kterých chce dosahovat. To vychází z designu a myšlenky, která se s technologií nese. Technologie je jednak spojena s dosahováním utopie, jak bylo předestřeno v první kapitole, a technologické produkty jsou vždy nástrojem. Jsou to nástroje na generování textu, manipulace s objekty, rozpoznávání obrázků, úpravy videí, generování obrázků, nástroje sexuální stimulace etc. To, že je technologie nástrojem ji tak řadí do pozice podřízeného do role pasivního objektu. Nástroj potřebuje vždy někoho, kdo mu vládne. Kladivo nezatlouče hřebík, a pokud někomu rozbije hlavu, tak jen z toho důvodu chyby jeho umístění či špatnou manipulací, nikoliv svobodným jednáním kladiva. Pokud by se takovéto příhody s kladivy opakovaly, bylo by na čase nastoupit na zteč proti všem kladivům? Umělá inteligence není věc, která by se snažila pomocí falešných informací rozvracet demokracii. Roboti se neucházejí o naše pracovní místa, aby lidem sebrali možnost živobytí. Negativní důsledky implementace nové technologie, tak spíše vyplývají z latentních rizik, které tato implementace produkuje (srovnej Beck). V průběhu práce jsem podal několik příkladů toho, jak se zároveň stává i řešitelem těchto rizik.

¹¹ Nick Bostrom onu problematiku více rozvíjí na dalších příkladech, které onen triviální problém špatně definovaného úkolu řeší a ukazuje na nich další vznikající problémy. Já používám takto zjednodušenou verzi, abych poukázal na to, že k robotí destrukci je potřeba ono uvědomění si špatného postavení, které je spíše jen výsledkem antropomorfizace.

Přestože nám technologie neubližuje vědomě, tak i přesto nemůžeme zajistit neustálé panství člověka nad strojem, a to právě z důvodu, že je technologie vždy nástrojem, který dokáže ovlivňovat mocenské vztahy. Konflikt totiž nikdy nestojí mezi strojem a člověkem, ale člověkem a člověkem. Technologie se pak stává pouze nástrojem, který toto panství umožňuje. V následujících podkapitolách uvedu, jakým způsobem je toto panství zprostředkováno.

6.1 Velký bratr

Michel Foucault ve své knize *Dohlížet a trestat* předkládá Benthamův *panoptikon*. Jedná se o architektonickou figuru, která v sobě skrývá princip neustálého dohledu. Na okraji je budova ve tvaru kruhu, uprostřed se nachází věž, prošpikována širokými okny, která vedou k vnitřní straně kruhu. Okrajová budova je rozdělena na jednotlivé cely, které mají dvě okna. Jedno směrem k věži a druhé směrem ven, aby skrze něj procházelo světlo. Dispozitiv panoptikonu organizuje prostorové jednotky, které umožňují bez ustání vidět a rozeznávat. (2000, s. 281). Slovní spojení *Velký bratr* je metafora pro politický dohled spojený s totalitarismem. Poprvé se toto slovní spojení objevuje v knize George Orwella s názvem *1984*. V příběhu této knihy je technologie využívána pro kontrolu obyvatelstva skrze televizní obrazovky, které se nedají vypnout. Velký bratr tak dosahuje své všudypřítomnosti (Power 2016). Podobně jako v účincích panoptikonu, který má u vězně vzbudit dojem toho, že je nepřetržitě sledován, a tím je zprostředkováno automatické fungování moci (Foucault 2000, s. 282). Metafora Velkého bratra poukazuje na způsob dohlížení ze strany státu prostřednictvím technologie. Umělá inteligence se stává další takovouto technologií, která umožňuje zlepšení dohlížení nad obyvatelstvem v připraveném prostoru digitálního totalitarismu (Power 2016).

Pokud mluvíme o digitálním totalitarismu, pak nemůžeme opomenout Čínu, která se stala téměř synonymem k tomuto slovnímu spojení. K aplikaci digitálního totalitarismu je samozřejmě potřeba digitalizace, která v případě Číny proběhla úspěšně. V roce 2008 se stala Čína zemí s největším počtem uživatelů internetu. V roce 2016 se jejich počet dostal na 800 milionů a o dva roky později se toto číslo zvedlo o 29 milionů. Čína má tak nyní více internetových uživatelů než Evropská unie a Spojené státy americké dohromady. Drtivá většina těchto uživatelů k internetu přistupuje skrze mobilní zařízení. Tento velký zájem o digitální nástroje vedl v Číně k rozvoji digitální ekonomiky. V dnešní Číně je internet novým sociálním teritoriím, skrze které probíhají nejrůznější sociální interakce od nakupování přes spojení s rodinou až k rekreaci. Obyvatelé Číny tak využívají internet podobně jako jej využívá zbytek světa. Čína je do tohoto virtuálního světa zachycena více, což vytvořilo dokonalý prostor pro využití k digitální totalitě. V roce 2018 například podíl uživatelů, kteří provádí mobilní digitální platby v Číně, dosáhl 71,4 %. Pro srovnání v USA to bylo pouze 15 % uživatelů. Čínská vláda pochopitelně onu digitální inovaci využila ke zvýšení kontroly nad obyvatelstvem. Provedla to skrze čtyři kroky: generalizace chytrých průkazů totožnosti, zavedení velkého firewallu s cílem přísného dohledu nad internetem a sociálními sítěmi, testování a postupné zavádění systému sociálních kreditů a rozšíření technologie na rozpoznávání obličejů (Cabestan 2020). Tyto tendence dohledu se ze strany Číny objevují i

na globální úrovni, neboť Čína svými aplikacemi zasahuje i do globálního trhu. Jedním z příkladů čínské expanze digitální totality je aplikace *TikTok*, která byla označena Národním úřadem pro kybernetickou a informační bezpečnost jako vysoce riziková kvůli shromažďování dat, která mají být využívána ke špionáži (2023).

Přibližme si jeden z aspektů z digitálního totalitarismu v Čínském prostředí, který považuji za nejdůležitější k našemu tématu. Systém sociálních kreditů je popisovaný jako jádro čínské internetové agendy. Cílem tohoto systému je za pomoci osobních dat, které jsou shromažďované prostřednictvím využívaných aplikací, zlepšit jednání občanů. Individua a podniky mají být hodnoceny dle různých aspektů svého jednání: kam chodí, co nakupují a koho znají. Ono hodnocení je integrováno do systému komplexní databáze, která je propojena jak s vládními informacemi, tak i s informacemi shromažďovanými soukromými podniky (Zuboff 2020, s. 364). Veskrze všechny čtyři jmenované inovace v digitálním světě Číny slouží převážně k tomuto projevu moci, které je spojené s neustálým dohledem. Úkolem je normalizovat jednání, ať už jedinců či skupin (Power 2016). Veškeré jednání odehrávající se v kyberprostoru je neustále pod dohledem. Každé jednání v kyberprostoru taktéž produkuje informace o jedinci, vytváří se o nich vědění. Podobně popisuje Foucault vstup jedince do polepšovny (2000, s. 407). Data sbíraná o uživateli se tak stávají klíčovým prvkem ve vykonávání moci. Systém sociálních kreditů je častokrát srovnáván se zmíněným Orwellovským anti-utopickým románem, ale lepší je jej pojímat jako apoteózu instrumentální moci krmené osobními daty a kontrolované autoritářským státem s cílem společnost automatizovat (Zuboff 2020, s. 365).

Umělá inteligence a další digitální nástroje se zde fungují nejen jako prostředky, které umožňují efektivní dohlížení na jedince či skupinu, ale taktéž jako zprostředkovatel automatizace člověka. Můžeme pozorovat jistý důraz na automatizaci v podobném smyslu jako se nachází u strojů, kdy se musí provést stereotypizace činnosti, při čemž je odstraněno potencionální využití. Na tomto příkladě s autoritářským státem je skvěle pozorovatelná proměna vztahu originálu a kopie. Vlastnosti robota, které jej oddělují od člověka se stávají žádoucími i u občanů. Člověk se tak stává reprezentací robota. Stroj se přitom stává nástrojem, který onu reprezentaci vynucuje. Je dohlížecím prvkem, který vyhodnocuje - tichý policista. Vzpomeneme-li si na fenomény *filter bubble*, pak ani tichý policista není zapotřebí. Model jedince může být pomocí stroje upravován. Algoritmy rozhodnou, jaké informace se k člověku dostanou. Algoritmy rozhodnou, v jakém světě bude uživatel žít. Algoritmus se stává obrazem svého stvořitele-programátorem.

6.2 Dohlížet a doporučovat

V případě Číny jsme její digitální totalitu popsaly pomocí konceptu *panoptikonu*. Podobnosti v přístupu Číny s tímto konceptem jsou znatelné právě z toho důvodu, že se o ní samotné mluví jako o totalitě. Není to ovšem pouze případem Číny. Internet je využíván na globální úrovni, a jak již bylo uvedeno, tak je tomu činěno stejným způsobem jako v Číně. Jednání jakéhokoliv jedince na internetu je produkcí dat. Tato data se stávají novým zlatem, nejcennější surovinou. Člověk je při pohybu internetem prakticky v procesu redukce své vlastní osoby na informace. Na základě interakcí v kyberprostoru (přihlašování se,

prohlížení si obsahu, navazování virtuálních přátelství etc.) je o člověku vytvořen balíček dat, která ho dokážou skvěle reprezentovat. Tato data jsou shromažďována skrze platformy, na nichž se jedinec pohybuje. Platformy pak tyto cenné „suroviny“ získávají zcela zdarma. Přestože jsou některá z těchto dat využita na zlepšení produktů a služeb, tak zbytek je hnán do systémů umělé inteligence, kde je přetvořen do prediktivních produktů, které předpovídají jednání konzumentů. S těmito prediktivními produkty se následně dále obchoduje, jelikož poskytují výhodu na trhu. Dostáváme se do doby *kapitalismu dohledu* (Zuboff 2020, s. 14).

Trh se změnil. Nyní se neobchoduje pouze s určitými produkty či službami. Nacházíme se ve sféře digitálního zboží. Informace jsou součástí obchodu. Jedinec se zde nachází jako aktivní entita v tomto výměnném obchodě, kde za informace platí informacemi. Veškerý obsah konzumovaný na internetu a sociálních sítích je „placen“ skrze informace, které za sebou uživatel zanechává. Dalo by se říci, že uživatel vykonává neplacenou práci, jelikož z jeho aktivity se vytváří zboží či suroviny, jež přispívají k růstu sociální sítě (Ekbia, Nardi 2017, s. 94). Umělá inteligence zde nehraje roli pouze v případě nástroje, který je tyto data schopný zpracovat. Jedno z mnoha dalších využití se skýtá v rámci takzvaných doporučovací systémů. Tyto doporučovací systémy existují buď ve formě umělé inteligence, nebo algoritmu, který je spojený se strojovým učením a pracuje na základě Big Data. Poslání těchto systémů vychází z jejich názvu, tedy navrhnout a doporučit konzumentům další produkty s cílem udržení uživatele na platformě (Caulfield 2020).

Zde se nabízí krásná paralela se snímkem *Matrix*. Stroje v tomto filmu se nesnaží o zničení lidstva, ale spíše o jeho využití. Člověk je pro ně zdrojem energie, na níž jsou stroje závislé a aby toho docílily, používají matrix, který u lidí udržuje mozkovou aktivitu. S velkými daty je to stejné. Jednak se stávají surovinou, se kterou se může obchodovat, a zároveň se tyto data využívají k udržení uživatele na platformě, což vede k získávání dalších dat. Vracíme se tak do tématu předchozí kapitoly, k fenoménu *filter bubble*. Struktura médií ovlivňuje charakter naší společnosti. V minulém století byla takovým nástrojem televize a pro dnešní den se jím stává internet, potažmo sociální sítě, které svým uzpůsobením přinášejí nové možnosti, jak uživatelského vyjádření, tak i způsobu dohledu (Pariser 2014, s. 14). „*Konec panoptického systému. Oko televize už není zdrojem absolutního dohledu a ideálem kontroly již není transparentnost.*“ Takto Baudrillard popisuje rozpad konceptu představeného Michele Foucaultem. Činí tak pomocí reality show *Loups*, která vyobrazuje život rodiny, jako by kamery nebyli přítomné. „*Již nesledujete televizi, je to televize, co sleduje vás (živě)*“ (1994, s. 30–31). Na internetu se tento jev projevuje daleko zřetelněji. Každý váš klik je informace pro platformu. Každý váš klik je tvorba modelu vaší osobnosti. Vytváření tohoto modelu však vede oběma směry. Model uživatele neutváří jen uživatel svojí aktivitou, ale i platforma samotná. Může tak být činěno skrze doporučování pouze určitého druhu informací a podobně. Stírá se rozdíl mezi aktivitou a pasivitou, jelikož my sami jsme naším modelem, který může být a je pozměňován (viz předchozí kapitola). Moc se již neskrývá v neustálém dohledu, který má za následek normalizované jednání jedince. Nastává nový druh moci, instrumentalizace. Instrumentální moc zná a formuje lidské chování. Místo výzbroje a armády prosazuje svou vůli prostřednictvím automatizovaného prostředku stále všudypřítomnější výpočetní architektury „chytrých“ síťových zařízení, věcí a prostorů. Cílem je jedince automatizovat. (Zuboff 2020, s. 15).

Již v průběhu minulého století můžeme pozorovat obdobný jev přítomný v reklamě. Objekty jsou nabízené jako výsledek vaší touhy. Vy jste důvodem, proč produkt vzniká (Baudrillard 1996, s. 174–175). Systémy, o nichž zde mluvíme, jsou krásnou třešničkou na dortu, jelikož pouze čerpají z těchto produktů a dostávají je k nám. Na jedné straně jsou to tak data využívaná k vytvoření obrazu konzumenta a na druhé je zde utváření uživatele na základě produktů, které mu nabízí. Ano, to jsi ty! Jsi tímto a tímto, jelikož poptáváš toto a toto. Umělá inteligence je Bohem. Je neviditelná a neslyšná. A to přesto, že o ní víme, jelikož je akcentovaná i skrze užívání těchto platforem (Grandinetti 2023, s. 1273). Jak přesně funguje je nám ovšem skryté. Nevystupuje jako velitel, který přikazuje další postup, ale nabízí možné cesty. Uživatel se tak dostává do pocitu, že stále podléhá své svobodné vůli. Podobně jako uživatel sex robota se dostává do pocitu, že je v roli pána. Sleduji tento obsah, protože chci. Sleduješ tento obsah, protože chceš. V tendencích této instrumentální moci se skýtá i možnost modifikace, která je zde zahalena pláštěm falešné uživatelské svobody (Zuboff 2020, s. 330). Její síla není jen v predikci budoucího jednání, ale v celkové konstrukci diskurzu, když se stává nástrojem, který nabízí informace o světě. My sami se stáváme systémem, který je krmený daty, na jejichž základě upravujeme své jednání. Jsme redukováni na systémy, které lze predikovat. Z tohoto pohledu jsme automatizovanými biologickými stroji, který je připoután ke svému zařízení pomocí personalizovaného obsahu, za záminkou produkce dat. Můžeme tak říci, že nás už stroje ovládají. Jsou tím prvkem, který má vliv na utváření našich diskurzů a jsou tím prvkem, který využívá naši aktivitu k produkci cenných dat stejně jako jsou horníci využívání k těžbě kobaltu. Namísto platu se nám dostane jen afektivní odměny. Technologie, která se dříve zdála jako nástroj k tomu, aby subjekt zdolal svět objektů, se tak metaforicky vymyká kontrole a sama přetváří člověka do role objektu, automatizovaného stroje. Objektu, který je zdrojem dat, jež tyto stroje udržují při životě. Stejně jako stroje z filmu *Matrix* přeměňují subjekt člověka na objekt, ze kterého čerpají energii pro svůj běh. *Objekt triumfuje nad subjektem* (Baudrillard 2008, s. 110).

Stroje vyrábějí stroje. Pomocí sady algoritmů umělé inteligence je člověk automatizován, což je oním vykonáním moci na uživateli, které se ke všemu nejeví jako nepřítel, ale jako přítel. Jak jsme si uváděli, tak automatizace zabraňuje možnosti reteritorializace. Jsme tak neustále pod dohledem, ale oko, které na nás hledí, ani nevidíme, a to nám dává pocit svobody. Originály našich já se vytrácejí a přetvářejí se na model, který se stává reálnější než naše originální já. Jsme sami sobě simulakrum třetího řádu.

Panství stroje nad člověkem, tak nenastává způsobem, jaký je popsán prostřednictvím Terminátoru narativu, ale spíše skrze redukování člověka na automatizovaný systém, což je spojeno s reprezentací modelu poslušnosti. Dochází tím tedy k přesnému opaku, jelikož právě stroje jsou tím, nad čím se nejlépe udržuje moc. Proto se odvracíme od vztahu, kdy měl být robot reprezentací člověka k převrácení a z člověka se stává reprezentace robota.

6.3 Linie úniku

Měli bychom se tedy uchýlit k nihilismu? Dostali jsme se opravdu do doby, kdy se člověk stal otrokem technologie v tomto zprostředkovaném pojetí? Přestože by se z předchozí kapitoly dalo vyčíst, že ano, tak jsem přesvědčen o opaku. Sám sebe nepovažuji

za techno-pesimistu, ale ani za techno-optimistu. Technologie sama o sobě v konceptu nástroje se stává objektem, který má potenciál se stávat něčím skrze své využití. Nástroje jsou vždy designované k určitému použití. Svým designem ovšem nabízejí i odlišné možnosti využití. V podstatě se jedná o podobný princip jako popisuje Baudrillard v prostředí automatizace, což jsem uváděl v kapitole *Čapkovské prorocství*. Automatizace je vraždou potenciálu systému, zasazení do určitého teritoria. Jeho pozice je definována a podrobena mikrofašistickému ustanovení do své funkčnosti, kdy jakékoliv její jiné využití je považováno za nesprávné, jelikož neodpovídají stanoveným pravidlům využití. Pojetím Deleuzeho a Guattariho pak prostřednictvím procesu reteritorializace můžeme vzkřísit umrtvené potenciality. Matrix není jen využíván stroji k čerpání energie z lidí, ale je využíván i lidmi k rebelii proti strojům. Příkladem může být například arabské jaro, kdy se internet a sociální média staly prostorem pro organizaci protestů (McFarland 2012). Reteritorializace těchto nástrojů pak vytváří přeměnu zprostředkovatele útlaku na zprostředkovatele vymezení či vzpoury.

7 Závěr

Je tedy současný stav umělé inteligence a robotiky srovnatelný se sci-fi představami? Jednoduchá odpověď, která zazněla již v úvodu, by byla: *ne*. Sci-fi díla v sobě nenesou žádnou predikční hodnotu. Jedná se vždy o fikční světy, které mohou fungovat s logickými spory, a proto nedává smysl je srovnávat s naší žitou realitou. Nemůžeme ovšem zanedbat to, že svět sci-fi a vědy se nějakým způsobem ovlivňuje. Zárným příkladem je zavedení slova *robot* s původem v Čapkově dramatu. Nabízí se tak složitější odpověď: *ano, ale...* Samozřejmostí je, že nesmíme vykládat žádné sci-fi dílo doslovně. Některé aspekty těchto děl ovšem můžeme porovnávat s naší současnou zkušeností. Pro rekapitulaci nyní popíšu důležité aspekty z analyzovaných sci-fi děl.

První dílo, na které jsem se zaměřil je české drama *R.U.R.* od Karla Čapka. Dílo nám nabídlo dvě oblasti problémů: práci a moc. V oblasti pracovní činnosti můžeme pozorovat jistou podobnost s procesem automatizace. Rosumovi roboti i automatizace přináší nahrazení člověka. V dramatu je toto nahrazení pro většinu postav chápáno pozitivně. To se do jisté míry děje i u procesu automatizace (osvobození od rutinní práce apod.), ale přináší s sebou i obavy v podobě technologické nezaměstnanosti. Argument proti *technologické nezaměstnanosti* spočívá v tom, že automatizace přinese zvýšení ekonomické aktivity a možnost nových pracovních pozic. To znamená, že práce nezanikne, ale ani že budou pracovníci a pracovnice osvobozeni od rutinních činností. Naopak jsem zmínil hypotézu, že se vlivem možnosti zapojení automatizovaných systémů do výroby jejich postavení na trhu práce zhorší. Už i na tomto tématu se projevuje aspekt moci, kdy jsou prohozeny role subjektu a objektu na příkladu ze skladů *Amazonu*.

Dalším analyzovaným dílem byla kniha od Philipa K. Dicka - *Blade Runner*, která přinesla posthumanistický pohled. Problém, na který jsem se zaměřil, byla nerozlišitelnost a s ní spojená redefinice kategorií člověka a androida. Nerozlišitelnost mezi člověkem a androidem, jako je popsána v knize, je daleko od současné skutečnosti. Avšak například ve

virtuálním prostoru internetu se už můžeme dostat do situace, ve které nemůžeme určit, jestli komunikujeme se strojem nebo s člověkem. Zároveň k polidšťování robotů není potřeba této nerozlišitelnosti, jelikož se dokáží infiltrovat do rolí, které původně náležely výhradně lidem. To se snažím ukázat na příkladu sex robotů, kdy na místo pouhé sexuální pomůcky může uživatel svého sex robota pojímat jako svého životního partnera či partnerku. To je věcí, která působí distinkci objektu a subjektu složitější.

Posledním dílem, které jsem podrobil analýze, je snímek *Matrix* z roku 1999. Film přinesl téma simulace, kterou předkládá jako počítačem vytvořený virtuální svět. V této práci jsem se nezaměřoval na světy podobného charakteru, jako je například *Metaverse* od společnosti *Meta*. Záměr jsem přesunul na starší, ale pro současnost význačnější médium, kterým je internet. Internet sám o sobě nenabízí simulovanou realitu, jako je vidět ve filmu. Spíše se jedná o zprostředkování modelu světa, který pomáhá utvářet vědění o světě. Zde nastává problematika povahy média a s ním spojené fenomény. Nacházíme se v době, kdy jsme zaplavení informacemi. Platformy na internetu se stávají jakýmsi směrovačem, který k nám informace přináší. To má za výsledek efekty jako je *filter bubble*.

Při rekapitulaci *R.U.R.* jsem záměrně vynechal téma moci, protože se tohle téma prolíná napříč téměř všemi sci-fi díly, ale hlavně je v tomto tématu ukryta ona neustále pokládaná otázka: *Ovládne nás umělá inteligence?* Terminátor narativ, ze kterého tato obava vychází, se zdá být pouze záměnou fikce za realitu a přitahuje k sobě konspirace. Rebelie robotů, kterou předkládá *R.U.R.* je antropomorfní. Vzepřením se svým stvořitelům a zotročovatelům se stávají novou vývojovou linií. Jde zde o uvědomění si své situace a touhu po vymanění se z ní. Roboti, se kterými se my dnes potýkáme, tuhle vlastnost nemají. Model robota předloženého v rámci dramatu je odlišný od toho modelu, na který v současnosti narážíme. Robot je poslušný, racionální, rutinní etc. Povstání robotů, tak nenastane. *Matrix* stroje nepolidšťuje, nezaměřuje se na stroj, který by měl být novým vývojovým stádiem, ale zaměřuje se na člověka. Pro stroje z filmu se lidé nestali ničím jiným než baterii-objektem. To podobné se odehrává ve sféře internetu, kdy je uživatel chycen k obsahu, který mu platforma nabízí na základě dat získaných o uživateli. Tato data se nepoužívají jen k udržení uživatele na platformě, ale také k vytváření predikčních modelů, s nimiž je dále obchodováno. Uživatel se stává objektem, který vytváří hodnotnou surovinu. Predikce v uživatelském jednání přináší možnost automatizace uživatele. Uživatel se stává strojem, který produkuje data. To všechno je zapříčiněno pomocí algoritmů, které spadají pod označení umělá inteligence. Stroje vyrábějí stroje.

Má práce přistupuje k reálným rizikům technologie skrze rizika zmíněna ve sci-fi dílech. Ukazuje se, že sci-fi díla poskytují do jisté míry řadu témat, která mají v současných debatách o rozvíjející se technologii smysl, i když neukazují z pochopitelných důvodů na přesné problémy, se kterými se dnes potýkáme. Nabízejí spíše jistou metaforu. Například film *Matrix* poukazuje na proměnu vztahu mezi subjektem a objektem, která nějakým způsobem nastává skrze kapitalismus dohledu a proces heteromatizace. Jednotlivá témata fungují samostatně jako výsek z celkové problematiky vztahu člověka a stroje. Sci-fi díla se sice od skutečnosti odlišují v mnohých aspektech, ale nabízejí problematiky, na které se můžeme zaměřovat v analýzách budoucího vývoje technologie. Důležitou vlastností sci-fi děl není pouze reakce na technologie, ale i jejich přesah do sociální sféry a komplexity světa, do kterého tyto technologie vstupují.

8 Resumé

In my master thesis I deal with the issue of machines, robots and artificial intelligence, which appeared in cult works of science fiction such as: the drama *R.U.R* by the Czech writer Karel Čapek, the novel *Do Androids Dream of Electric Sheep?* Philip K. Dick and the 1999 film *The Matrix* by Wachowski sisters. The plots of these works are compared to the current state of not only the disciplines of robotics and artificial intelligence, but also the current state of society. The ideas of postmodern philosophers Jean Baudrillard, Michel Foucault, Gilles Deleuze, along with Félix Guattari Shoshana Zuboff and Ulrich Beck were used as theoretical starting points. The thesis aims to answer the research question: *Is the current state of AI and robotics technology comparable to science fiction ideas?*

In five chapters I describe phenomena from the present state of technology and compare them with these sci-fi ideas. I use Čapek's *R.U.R.* to present the processes of automation and heteromatization. Dick's *Blade Runner* highlights the problem of category disruption between human and android. I use the Matrix as an allegory to shape models of reality in the Internet environment. All the themes are connected in the last chapter, which shows how the power structure between subject and object is changing.

Science fiction works are metaphorical and thus dangerous to interpret literally. However, in the current state of technology we can observe phenomena similar to those presented in the science fiction works analysed.

Tabulka obrázků

Obrázek 1: Uncanny valley.....	18
--------------------------------	----

9 Literatura

- ACEMOGLU, Daron a RESTREPO, Pascual, 2022. Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*. Vol. 90, č. 5, s. 1973–2016. DOI 10.3982/ECTA19815.
- ANDREESSEN, Marc, 2023. The Techno-Optimist Manifesto. *andreessen.horowitz* [online]. 16 říjen 2023. Získáno z : <https://a16z.com/the-techno-optimist-manifesto/>
- BARGH, John A. a MCKENNA, Katelyn Y. A., 2004. The Internet and Social Life. *Annual Review of Psychology*. Vol. 55, č. 1, s. 573–590. DOI 10.1146/annurev.psych.55.090902.141922.
- BAUDRILLARD, Jean, 1994. *Simulacra and simulation*. Ann Arbor : University of Michigan Press. The Body, in theory. ISBN 978-0-472-09521-6.
- BAUDRILLARD, Jean, 1996. *The system of objects*. London ; New York : Verso. ISBN 978-1-85984-943-9.
- BAUDRILLARD, Jean, 2008. *Fatal strategies*. Los Angeles, CA : Semiotext(e). ISBN 978-1-58435-061-3.
- BAUDRILLARD, Jean, 2017. *Symbolic exchange and death*. Revised edition. Los Angeles London New Delhi Singapore Washington DC Melbourne : SAGE. Theory, culture & society. ISBN 978-1-4739-0759-1.
- BAUMAN, Zygmunt, 2003. *Modernita a holocaust*. Praha : Sociologické nakladatelství. ISBN 978-80-86429-23-6.
- BECK, Ulrich, 2011. *Riziková společnost: na cestě k jiné moderně*. 2. vyd. Praha : Sociologické nakladatelství. ISBN 978-80-7419-047-6.
- BERGSTROM, Carl T. a WEST, Jevin D., 2020. *Calling bullshit: the art of skepticism in a data-driven world*. First edition. New York : Random House. ISBN 978-0-525-50918-9.
- BIVENS, Josh a MISHEL, Lawrence, 2021. The failure of automation and skill gaps to explain wage suppression or wage inequality. *Economic Policy Institute*.
- BONTRIDDER, Noémi a POULLET, Yves, 2021. The role of artificial intelligence in disinformation. *Data & Policy*. Vol. 3, s. e32. DOI 10.1017/dap.2021.20.
- BOSTROM, Nick, 2014. *Superintelligence: paths, dangers, strategies*. First edition. Oxford : Oxford University Press. ISBN 978-0-19-967811-2.
- CABESTAN, Jean-Pierre, 2020. The State and Digital Society in China: Big Brother Xi is Watching You! *Issues & Studies*. Vol. 56, č. 01, s. 2040003.

DOI 10.1142/S1013251120400032.

- CARSTENS, Deborah S., DOSS, Samuel K. a KIES, Stephen C, 2018. Social Media Impact on Attention Span. *The Journal of Management and Engineering Integration*. Vol. 11, č. 1, s. 20–27.
- CAULFIELD, Brian, 2020. What's a Recommender System? *NVIDA* [online]. 14 květen 2020. Získáno z : <https://blogs.nvidia.com/blog/whats-a-recommender-system/>
- ČAPEK, Karel, 2013. *Dramata: Loupežník, R.U.R., Věc Makropulos, Bílá nemoc, Matka*. Praha : Dobrovský. ISBN 978-80-7390-006-9.
- ČT24.CZ, 2023. Fotograf zvítězil v prestižní soutěži se snímkem, který vytvořila umělá inteligence. *ČT24.cz* [online]. 18 duben 2023. Získáno z : <https://ct24.ceskatelevize.cz/clanek/kultura/fotograf-zvitezil-v-prestizni-soutezi-se-snimkem-ktery-vytvorila-umela-inteligence-7455>
- DAMIANO, Luisa a DUMOUCHEL, Paul, 2018. Anthropomorphism in Human–Robot Co-evolution. *Frontiers in Psychology*. Vol. 9, s. 468. DOI 10.3389/fpsyg.2018.00468.
- DANAHER, John, 2019. *Automation and utopia: human flourishing in a world without work*. Cambridge, Massachusetts : Harvard University Press. ISBN 978-0-674-24220-3.
- DANAHER, John, 2022. Techno-optimism: an Analysis, an Evaluation and a Modest Defence. *Philosophy & Technology*. Vol. 35, č. 2, s. 54. DOI 10.1007/s13347-022-00550-2.
- DANAHER, John a MCARTHUR, Neil (ed.), 2017. *Robot sex: social and ethical implications*. Cambridge, Massachusetts London, England : The MIT Press. ISBN 978-0-262-03668-9.
- DAVIES, Caroline, 2024. Schoolgirl, 14, found dead after alleged bullying by boys, London inquest hears. *Gurdian* [online]. 23 leden 2024. Získáno z : https://www.theguardian.com/uk-news/2024/jan/23/schoolgirl-14-found-dead-after-alleged-bullying-by-boys-london-inquest-hears?fbclid=IwAR3AzapcQluzPYdfqoIRNuSjzFcs0RJ3ASiJNvJBjKL5Bseqz_G1iB45Nyl
- DELEUZE, Gilles, 1994. *Difference and repetition*. New York : Columbia University Press. ISBN 978-0-231-08158-0.
- DELEUZE, Gilles a GUATTARI, Félix, 2010. *Tisíc plošin*. Vyd. 1. Praha : Herrmann & synové. ISBN 978-80-87054-25-3.
- DICK, Philip K., 1993. *Blade Runner*. Praha : Winston Smith. ISBN 978-80-85643-07-7.
- DIOGENÉS LAERTIOS, KOLÁŘ, Antonín a MACHOVEC, Dušan, 1995. *Životy, názory a výroky proslulých filosofů*. Vyd. 2. Pelhřimov : Nová tiskárna. ISBN 978-80-901916-3-1.
- DIVÍNOVÁ, Jana, 2024. Nepřekáží ani u krájení cibule. Apple Vision Pro vyzkoušeli první Češi. *Forbes* [online]. 6 únor 2024. Získáno z : <https://forbes.cz/neprekazi-ani-u-krajeni-cibule-apple-vision-pro-vyzkouseli-prvni-cesi/>
- EKBIA, H. R. a NARDI, Bonnie A., 2017. *Heteromation, and other stories of computing and capitalism*. Cambridge, MA : MIT Press. Acting with technology. ISBN 978-0-

262-03625-2.

- EKBIA, Hamid a NARDI, Bonnie, 2014. Heteromation and its (dis)contents: The invisible division of labor between humans and machines. *First Monday*. DOI 10.5210/fm.v19i6.5331.
- ENTMAN, Robert M., 1993. Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*. Vol. 43, č. 4, s. 51–58. DOI 10.1111/j.1460-2466.1993.tb01304.x.
- EYSENCK, Michael W. a EYSENCK, Christine, 2021. *AI vs Humans*. 1. London : Routledge. ISBN 978-1-00-316269-8.
- FLEMING, Sean, 2022. The Unabomber and the origins of anti-tech radicalism. *Journal of Political Ideologies*. Vol. 27, č. 2, s. 207–225. DOI 10.1080/13569317.2021.1921940.
- FORTUNATI, Leopoldina, ESPOSITO, Anna a LUGANO, Giuseppe, 2015. Introduction to the Special Issue “Beyond Industrial Robotics: Social Robots Entering Public and Domestic Spheres”. *The Information Society*. Vol. 31, č. 3, s. 229–236. DOI 10.1080/01972243.2015.1020195.
- FOUCAULT, Michel, 2000. *Dohlížet a trestat: kniha o zrodu vězení*. Praha : Dauphin. ISBN 978-80-86019-96-3.
- FOUCAULT, Michel, 2002. *Archeologie věděni*. V Praze : Herrmann & synové. ISBN 978-80-239-0124-5.
- FOUCAULT, Michel, 2012. *The order of things*. 2nd ed. Hoboken : Taylor & Francis. ISBN 978-0-203-99664-5.
- FREY, Carl Benedikt a OSBORNE, Michael A., 2017. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*. Vol. 114, s. 254–280. DOI 10.1016/j.techfore.2016.08.019.
- GALVAN, Jill, 1997. Entering the Posthuman Collective in Philip K. Dick’s „Do Androids Dream of Electric Sheep?” *Science Fiction Studies*. Vol. 24, č. 3, s. 413–429.
- GOOD, Irving John, 1966. Speculations Concerning the First Ultraintelligent Machine. In : *Advances in Computers*, s. 31–88. Elsevier. ISBN 978-0-12-012106-9. DOI [10.1016/S0065-2458\(08\)60418-0](https://doi.org/10.1016/S0065-2458(08)60418-0).
- GOTTFRIED, Jeffrey a SHEARER, Elisa, 2017. News Use Across Social Media Platforms 2017. *PEW RESEARCH CENTER* [online]. Získáno z : https://internet.psych.wisc.edu/wp-content/uploads/532-Master/532-UnitPages/Unit-05/Shearer_PewResearch_2017.pdf
- GRANDINETTI, Justin, 2023. Examining embedded apparatuses of AI in Facebook and TikTok. *AI & SOCIETY*. Vol. 38, č. 4, s. 1273–1286. DOI 10.1007/s00146-021-01270-5.
- INDIA, Freya, 2023. We Can’t Compete With AI Girlfriends. *Girls* [online]. 14 září 2023. Získáno z : https://www.freyaindia.co.uk/p/we-cant-compete-with-ai-girlfriends?fbclid=IwAR0dR0dkjF4bTAWBr2bS1vVB0Y-JckGs5I7fx_IT0F_mlhTvCyfPyLn0tf0&utm_source=substack&utm_medium=email
- JOHNSON, Arthur T., 2015. The Technology Hype Cycle [State of the Art]. *IEEE Pulse*.

- Vol. 6, č. 2, s. 50–50. DOI [10.1109/MPUL.2014.2386491](https://doi.org/10.1109/MPUL.2014.2386491).
- JONES, Cameron a BERGEN, Benjamin, 2023. Does GPT-4 Pass the Turing Test? . DOI [10.48550/ARXIV.2310.20216](https://doi.org/10.48550/ARXIV.2310.20216).
- JONES, Steven E., 2006. *Against technology: from the Luddites to Neo-Luddism*. New York : Routledge. ISBN 978-0-415-97867-5.
- KERTYSOVA, Katarina, 2018. Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered. *Security and Human Rights*. Vol. 29, č. 1–4, s. 55–81. DOI 10.1163/18750230-02901005.
- KLINGBERG, Torkel, 2009. *The overflowing brain: information overload and the limits of working memory*. Oxford New York : Oxford University Press. ISBN 978-0-19-537288-5.
- KÖNIGS, Peter, 2022. What is Techno-Optimism? *Philosophy & Technology*. Vol. 35, č. 3, s. 63. DOI 10.1007/s13347-022-00555-x.
- KRIPKE, Saul A., 1980. *Naming and necessity*. Cambridge, Mass : Harvard Univ. Press. ISBN 978-0-674-59846-1.
- KUHN, Thomas Samuel, 1997. *Struktura vědeckých revolucí*. 1. vyd. Praha : OIKOYMENH. ISBN 978-80-86005-54-6.
- KUMKALE, G. Tarcan a ALBARRACÍN, Dolores, 2004. The Sleeper Effect in Persuasion: A Meta-Analytic Review. *Psychological Bulletin*. Vol. 130, č. 1, s. 143–172. DOI 10.1037/0033-2909.130.1.143.
- LAIST, Randy, 2012. Bullet-time in Simulation City: Revisiting Baudrillard and The Matrix by way of the „Real 1999“. HURLEY, Marian, MELLAMPHY, Deborah a MORIARTY, Jill (ed.), *Alphaville: Journal of Film and Screen Media*. Č. 2, s. 19–33. DOI 10.33178/alpha.2.02.
- LEFKOWITZ, Melanie, 2019. Professor’s perceptron paved the way for AI – 60 years too soon. [online]. Získáno z : <https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon>
- LESHNER, Connor E. a JOHNSON, Jessica R., 2024. Technically in love: Individual differences relating to sexual and platonic relationships with robots. *Journal of Social and Personal Relationships*. s. 02654075241234377. DOI 10.1177/02654075241234377.
- LI, Siyu, YANG, Jin a ZHAO, Kui, 2023. Are you in a Masquerade? Exploring the Behavior and Impact of Large Language Model Driven Social Bots in Online Social Networks. . DOI 10.48550/ARXIV.2307.10337.
- MARX, Karl, 1988. *Economic and philosophic manuscripts of 1844*. Amherst,N.Y : Prometheus Books. Great books in philosophy series. ISBN 978-0-87975-446-4.
- MCFARLAND, Michael, 2012. Why We Care about Privacy. [online]. 1 červen 2012. Získáno z : <https://www.scu.edu/ethics/focus-areas/internet-ethics/resources/why-we-care-about-privacy/>
- META.COM. We believe in the future of connection in the metaverse. *Meta.com* [online]. Získáno z : <https://about.meta.com/metaverse/>

- MORI, Masahiro, MACDORMAN, Karl a KAGEKI, Norri, 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*. Vol. 19, č. 2, s. 98–100. DOI 10.1109/MRA.2012.2192811.
- Národní úřad pro kybernetickou a informační bezpečnost, 2023 *Aplikace TikTok představuje bezpečnostní hrozbu* [online]. Získáno z : <https://nukib.gov.cz/cs/infoservis/hrozby/1941-aplikace-tiktok-predstavuje-bezpecnostni-hrozbu/>
- PARISER, Eli, 2014. *The filter bubble: how the new personalized web is changing what we read and how we think*. New York : Penguin Books. ISBN 978-1-101-51512-9.
- Pause Giant AI Experiments: An Open Letter, 2023 *Future of life institute* [online]. Získáno z : <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- PETIŠKA, Eduard, 2022. AI napíše diplomku i článek. Antiplagiátorské programy to neodhalí. *Portál Vědavyzkum.cz - Nezávislé informace o vědě a výzkumu* [online]. 7 prosinec 2022. Získáno z : <https://vedavyzkum.cz/blogy-a-komentare/eduard-petiska/ai-napise-diplomku-i-clanek-antiplagiatorske-programy-to-neodhali>
- PLALE, Mathew, 2023. Beau Is Afraid actor Armen Nahapetian insists he's not A.I. [online]. 29 duben 2023. Získáno z : <https://www.imdb.com/news/ni64059058/>
- POWER, Daniel J., 2016. "Big Brother" can watch us. *Journal of Decision Systems*. Vol. 25, č. sup1, s. 578–588. DOI 10.1080/12460125.2016.1187420.
- PROCHÁZKOVÁ, Andrea, 2023. AI JE UŽ DOST POKROČILÁ NA TO, ABY NARUŠILA VOLBY. *Respekt* [online]. 9 říjen 2023. Získáno z : <https://www.respekt.cz/newsletter/umela-inteligence-je-uz-dostatecne-pokrocila-na-to-aby-narusila-volby>
- RID, Thomas, 2016. *Rise of the machines: a cybernetic history*. First edition. New York : W.W. Norton & Company. ISBN 978-0-393-28601-4.
- RITZER, George, 2003. *Mcdonaldizace společnosti: výzkum mění se povahy soudobého společenského života*. Vyd. 1., dotisk [i.e. 2. vyd.]. Praha : Academia. ISBN 978-80-200-1075-9.
- ROBERTS, Adam, 2006. *The history of science fiction*. Basingstoke [England] : Palgrave Macmillan. ISBN 978-0-230-55465-8.
- SANDERS, Steven (ed.), 2008. *The philosophy of science fiction film*. Lexington, Ky : University Press of Kentucky. The philosophy of popular culture. ISBN 978-0-8131-2472-8.
- SHENG, Anfeng a WANG, Fei, 2022. Falling in love with machine: emotive potentials between human and robots in science fiction and reality. *Neohelicon*. Vol. 49, č. 2, s. 563–577. DOI 10.1007/s11059-022-00664-8.
- STOKEL-WALKER, Chris, 2023. The „NPC" livestream TikTok trend helping creators earn cash. [online]. 14 srpen 2023. Získáno z : <https://www.bbc.com/worklife/article/20230811-the-npc-livestream-tiktok-trend-helping-creators-earn-cash>
- SUBLIME, Jérémie, 2024. The AI Race: Why Current Neural Network-based Architectures

- are a Poor Basis for Artificial General Intelligence. *Journal of Artificial Intelligence Research*. Vol. 79, s. 41–67. DOI 10.1613/jair.1.15315.
- THOMAS, William Isaac a THOMAS, Dorothy Swaine Thomas, 1928. *The child in America: Behavior problems and programs*. New York : A. A. Knopf.
- TURING, A. M., 1950. I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*. Vol. LIX, č. 236, s. 433–460. DOI 10.1093/mind/LIX.236.433.
- WAKEFIELD, Jane, 2016. Microsoft chatbot is taught to swear on Twitter. *BBC news* [online]. 24 březen 2016. Získáno z : <https://www.bbc.com/news/technology-35890188>
- WEST, Jevin a BERGSTROM, Carl, 2019. Which Face Is Real? Seeing through the illusions of a fabricated world. *Which Face Is Real?* [online]. 2019. Získáno z : <https://www.whichfaceisreal.com/about.html>
- WOOLDRIDGE, Michael J., 2021. *A brief history of artificial intelligence: what it is, where we are, and where we are going*. First U.S. edition. New York : Flatiron Books. ISBN 978-1-250-77074-5.
- YAN, Harry Yaojun et al., 2023. Exposure to social bots amplifies perceptual biases and regulation propensity. *Scientific Reports*. Vol. 13, č. 1, s. 20707. DOI 10.1038/s41598-023-46630-x.
- YUDKOWSKY, Eliezer, 2023. Pausing AI Developments Isn't Enough. We Need to Shut it All Down. *Time* [online]. 29 březen 2023. Získáno z : <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>
- zeroGPT, *zeroGPT* [online]. Získáno z : <https://www.zerogpt.com>
- ZUBOFF, Shoshana, 2020. *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. First trade paperback edition. New York, NY : PublicAffairs. ISBN 978-1-61039-570-0.

10 Přílohy

- Obrázek 1: MORI, Masahiro, MACDORMAN, Karl a KAGEKI, Norri, 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*. Vol. 19, č. 2, s. 98–100. DOI [10.1109/MRA.2012.2192811](https://doi.org/10.1109/MRA.2012.2192811).