

Západočeská univerzita v Plzni

Fakulta filozofická

Diplomová práce

**Filosofické koncepce umělé inteligence ve filmové
tvorbě**

Bc. Aneta Němcová

Západočeská univerzita v Plzni

Fakulta filozofická

Katedra filozofie

Studijní program Humanitní studia

Studijní obor Teorie a filozofie komunikace

Diplomová práce

**Filosofické koncepce umělé inteligence ve filmové
tvorbě**

Bc. Aneta Němcová

Vedoucí práce:

PhDr. Eva Žáčková PhD.

Katedra filozofie

Fakulta filozofická Západočeské univerzity v Plzni

Plzeň 2018

Prohlašuji, že jsem práci zpracovala samostatně a použila jen uvedené prameny a literatury.

Plzeň, duben 2018

.....

Chtěla bych velice poděkovat vedoucí této diplomové práce, PhDr. Evě Žáčkové, Ph.D.,
za cenné rady, dohled a pomoc při tvorbě práce.

Obsah

Úvod.....	1
1. Filmové teorie	3
1.1. Kognitivní filmová teorie	3
1.2. Kognitivní filmová teorie jako metoda	5
2. UI ve filosofickém pojetí	6
2.1. Vybrané definice UI.....	6
2.2. Filosofické aspekty a předpoklady UI.....	9
2.2.1. Slabá UI – Mohou se stroje chovat inteligentně?	11
2.2.2. Silná UI – Mohou stroje opravdu myslet?	16
2.2.3. Etika a rizika vývoje UI.....	25
3. UI ve starších sci-fi filmech.....	33
3.1. Úzkost způsobená UI	33
3.2. Deviantní chování a mentální abnormality	34
3.3. (ne)lidskost.....	36
4. Filosofie UI v současné filmové tvorbě.....	39
4.1. HER – Ona (2013)	40
4.2. EX MACHINA (2014).....	45
4.3. WALL-E – Vall-i (2008)	49
4.4. BIG HERO 6 – Velká šestka (2014).....	54
4.5. CHAPPIE (2015)	57
4.6. ROBOT & FRANK (2012).....	61
4.7. PROMETHEUS (2012)	64
4.8. THE MACHINE – Stroj (2013).....	68
5. Shrnutí analýzy filmů.....	70
6. Závěr	75
7. Použité zdroje	78
8. Summary.....	86
9. Příloha (překlad)	87

Úvod

Diplomová práce *Filosofické koncepce umělé inteligence ve filmové tvorbě* pojednává o zobrazení umělé inteligence v současné science-fiction produkci. Současnou produkcí se míní filmy poslední dekády. Cílem diplomové práce je analýza těchto filmů v návaznosti na filosofické teorie umělé inteligence, následně nalézání inspiračních vazeb mezi filosofií a filmovou produkcí, dále komparace témat umělé inteligence ve starších sci-fi filmech se současnou filmovou tvorbou. Poté objevování opakujících se, ale také opomíjených, témat filosofie umělé inteligence ve filmové produkci.

V první kapitole diplomové práce bude rozebrán pojem filmové teorie. Tato kapitola je pro práci stěžejní z hlediska obhájení výběru filmů na úkor například knižní tvorby. Vymezen bude především pojem kognitivní filmové teorie, která vyzdvihuje pozici diváka a umožňuje analýzu filmu individuálním pohledu. Proto nebudou v práci zohledněny komentáře a postřehy režisérů nebo ostatních tvůrců filmové tvorby. Pro práci je hodnocení a individuální pohled publika stěžejní, jelikož odráží konceptuální uchopení umělé inteligence veřejnosti skrze filmovou produkci. Takováto koncepce diplomové práce nám umožňuje pojímat kognitivní filmovou teorii jako metodu.

Pojem umělé inteligence (dále UI) je široký, tudíž je důležité vymezit jej pro použití v této práci, tím se bude zabírat druhá kapitola diplomové práce. Vymezení UI bude zprostředkováno některými jejími definicemi. Budou zde také nastíněny filosofické aspekty a předpoklady UI, což bude rozděleno do tří podkapitol. První podkapitola nazvána *Slabá UI – Mohou se stroje chovat inteligentně?* představí především Turingův test a některé námitky vůči Turingově testu (námitka hlavy v písku, námitka z různých neschopností a námitka z neformálního chování), které jsme schopni nalézt ve filmové tvorbě, jelikož jsou spíše filosofického než technického zaměření. Druhá podkapitola s názvem *Silná UI – Mohou stroje opravdu myslet?* navazuje na tu první tím, že pojednává o další námitce vůči Turingově testu, tedy argumentu z vědomí. Dále přichází otázka mind-body problému a řešení v podobě dualismu a monismu (neboli fyzikalismu). Tato podkapitola řeší také téma intencionálních stavů, představuje

experiment mozků v kádi, široký a úzký mentální obsah. Dále funkcionalismus, epifenomenalismus a biologický naturalismus, na který navazuje myšlenkový experiment Čínského pokoje Johna Searla a některé námitky vůči tomuto myšlenkovému experimentu. Posledním tématem podkapitoly je obtížný a zatím téměř neprozkoumaný pojem vědomí. Třetí a konečnou podkapitolou bude velice diskutovaná otázka etiky a hrozeb vývoje UI.

Ve třetí kapitole bude rozebrána umělá inteligence ve filmech starších deseti let. Především bude naznačeno, kterými tématy se filmová tvorba v minulosti zabývala. Díky této kapitole bude možné srovnat témata současná s těmi minulými. Zaměření filmů starších deseti let bude rozděleno do tří podkapitol, těmi jsou *Úzkost způsobená umělou inteligencí*, *Deviantní chování a mentální abnormality*, a *(ne)lidskost*. To také koresponduje s nejčastější podobou umělé inteligence promítanou ve filmech minulosti.

Ve čtvrté a zároveň stěžejní kapitole budou představeny filmy a jejich zkoumání. Filmů k analýze bude celkem osm. Filmová tvorba byla vybrána podle serveru IMDb a seznamu nejlepších filmů o umělé inteligenci posledního desetiletí. Mezi tyto filmy patří *Her*, *Ex Machina*, *Wall-e*, *Big Hero 6*, *Chappie*, *Robot & Frank*, *Prometheus* a *Machine*. Seznam těchto filmů byl zkompletován podle názorů veřejnosti, tudíž z něj lze vyvodit, jaký pohled na tuto problematiku je divákům předáván skrze filmovou tvorbu. Tato práce se bude zaměřovat především na filosofické otázky, které vyvstávají v tématu umělé inteligence a které byly představeny v předchozím oddílu práce. První otázka zní, zda je obecná umělá inteligence (ve smyslu silné umělé inteligence) možná, a pokud ano, zda umělá inteligence dokáže řešit lidské problémy. Tuto filosofickou otázku můžeme specifikovat podle toho, jsou-li inteligentní stroje schopné mít mysl, popř. vědomí. Díky těmto úvahám se také dostaneme k etické koncepci umělé inteligence. Ve čtvrté kapitole budou shrnuty výsledky této analýzy. Hlavní metodou práce bude analýza filmové tvorby ve snaze nalezení filosofických témat umělé inteligence a následné nalézání opomíjených nebo dosud neprobádaných témat filosofie umělé inteligence, což je zároveň i cílem diplomové práce. Občasné doslovné citace budou pro účely práce překládány, a jelikož přeložení některých pasáží by mohlo zapříčinit ztrátu celkového významu těchto tezí, budou přiloženy v originálním znění do

přílohového aparátu. Filmy budou uváděny originálním názvem, nejčastěji tedy v anglickém jazyce. Taktéž uváděné dialogy z použité filmové tvorby, budou citovány v originálním znění.

1. Filmové teorie

Filmovou teorii (spíše jakoukoli teorii) můžeme charakterizovat jako „...množinu více méně uspořádaných, více méně řečených, více méně závazných tvrzení, na něž skupina badatelů odkazuje, když chce pochopit či vysvětlit, v čem tkví daný fenomén.“¹ První kapitola této diplomové práce bude pojednávat o takzvané filmové kognitivní teorii. Jelikož tato teorie není příliš vědecky sjednocená, bude první kapitola spíše analýzou několika nedávných studií. Články jsou vybrány tak, aby korespondovaly s tématem práce. Proto se tyto filmové teorie týkají filmů s tematikou science fiction, přesněji řečeno umělé inteligence. Také se tyto studie zabývají populárními filmy, které oslovily co největší publikum a které odrážejí nejvíce probíraná témata umělé inteligence.

1.1. Kognitivní filmová teorie

Jak už bylo řečeno výše, kognitivní filmová teorie není ucelená teorie, nýbrž výzkumná tradice, která se datuje do 80. let minulého století.² Dnes se této teorii věnuje část vědecké společnosti, která pořádá každoroční mezinárodní konferenci a vydává specializovaný časopis. Zčásti tato teorie vznikla jako reakce a kritika psychoanalyticko-sémiotického paradigmatu³, které dominovalo v minulém století na poli filmového bádání. Psychoanalyticko-sémantický přístup byl oblíbený především v 70. letech 20. století. Zaměřuje se především na souvislosti mezi filmem a některými produkty nevědomí. Tato teorie se snaží zjistit, zda mají tyto nevědomé stavy (např. sny) určitý vliv na fungování filmu.⁴

Kognitivní filmová teorie navrhla a začala rozvíjet alternativní pohledy na různé prvky filmového zážitku diváka. Tyto prvky vplynuly z výzkumu v mezioborových

¹ CASSETTI, F. *Filmové teorie 1945-1990*, str. 3

² NANNICELLI, T. Cognitive Film Theory. In: *Cinema and Media Studies*, str. 32

³ HEATH, S. Narrative Space, In: *Screen*, str. 68-112

⁴ CASSETTI, F. *Filmové teorie 1945-1990*, str. 188

oblastech kognitivní vědy a analytické filosofie.⁵ Filmoví odborníci David Bordwell, Noël Carroll a Gregory Currie navrhli kognitivní perspektivu do filmových studií. Metodika studie filmové a divácké psychologie založené na kognitivní teorii byla objasněna Bordwellem v jeho esejích *A Case for Cognitivism* (1989) a *A Case for Cognitivism: Further Reflections* (1990).⁶

Filmoví teoretici jako Bordwell nebo Carrol kritizovali dosavadní filmové teorie pro snížení filmové analýzy na opakování určitých módních sloganů a neprozkoumaných předpokladů.⁷ Dosud nebyla kladena téměř žádná pozornost emocionálním účinkům filmů, zatímco vyvolávání emocí je hlavním účelem většiny filmů. Dosavadní teorie se jen zanedbatelně zaměřovaly na diváka.⁸ Podle kognitivních filmových teoretiků je důležité studovat filmy se širším zaměřením právě na kognitivní funkce, jako je vnímání, poznávání a interpretace v mysli diváka. Toto spojení je podle Perveze základem naší kognitivní mysli, která řídí naše chápání každodenního života.⁹ Proto se tyto filmoví vědci snaží o kognitivní přístup k filmovým teoriím, které zohledňují různé prvky filmových zážitků a vychází z interdisciplinární oblasti kognitivní vědy a analytické filosofie.¹⁰

Filmový vědec Anis Pervez podotýká, že je důležité zkoumat, jak vyvolává film emoce v publiku. Film je podle Perveze zprostředkovanou reprezentací světa tak, jak si ji představuje a jak ji zná tvůrce filmu. „*Film je zprostředkovanou reprezentací světa, tak jak ho zná filmový tvůrce. Jedná se o audiovizuální reprezentaci uspořádanou v řetězci snímků, sekvencí a zvuků. Každá jednotka, záběr nebo sekvence přispívá ke strukturování určitého pocitu nebo poselství, které film zamýšlí zastupovat. Vyvolává potěšení a občas způsobuje intelektuální zamyšlení diváka, protože příjemce se odráží ve filmu jako kognitivní agent.*“¹¹

⁵ CURRIE, G. *Arts and minds*, str. 63

⁶ PLANTINGA, C. Cognitive Film Theory: An Insider Appraisal. In: *Cinémas: Journal of Film Studies*, str. 15-37

⁷ Tamtéž, str. 15-37

⁸ SMITH, G. M. An Invitation to Feel. In: *Film Structure and the Emotion System*, str. 3-14

⁹ PERVEZ, A. Cognitive Film Theory. In: *Bangladesh Film and Television Institute Journal*, str. 381- 86

¹⁰ NANNICELLI, T. Cognitive Film Theory. In: *Cinema and Media Studies*, str. 32

¹¹ PERVEZ, A. Cognitive Film Theory. In: *Bangladesh Film and Television Institute Journal*, str. 381- 86

1.2. Kognitivní filmová teorie jako metoda

Důležitost kognitivní koncepce filmu pro tuto práci tkví v tom, že hlavní roli v koncepčním vzorci zaujímá divák. Největším průkopníkem této teorie je, jak už bylo řečeno, David Bordwell. Ten podává kognitivní filmovou teorii jako proces, ve kterém nabízí divákovi přesné kroky, jež má učinit při sledování filmu. Tento proces tvoří tři prvky. Prvním je samozřejmě film, který musí být schopný díky své zápletce a stylu poskytnout divákovi určitý návod. Druhým prvkem je samotný divák, který je uzpůsobený tomu, aby film pochopil a vnímal. K tomu by mu měla dopomoct interpretační schémata filmu. Poslední místo patří příběhu, který tkví ve filmu a divák jej rekonstruuje ve své hlavě.¹² Pokud uvedeme tuto teorii do praxe, staví se divák do centra dění. Aktivně se zapojuje do vytváření filmových efektů a významů. Pervez tvrdí, že naše vlastní tělo ve spolupráci s několika sensorickými schopnostmi je rozhodující pro náš zážitek z filmu.¹³ Vnímání není pouze pasivním záznamem těchto sensorických stimulací. Bordwell uvádí, že sledování filmu je konstruktivní aktivita, ve které využívá divák svých znalostí k tomu, aby zkonstruoval smysl filmu.¹⁴

Kognitivní filmoví teoretici tedy tvrdí, že vnímání a poznávací činnost vždy přesahuje danou informaci obsaženou ve filmu. Tedy, když diváci sledují film, jednotlivě odfiltrují hluk, samostatně vyplňují mezery a propojují to, co se pro někoho zdá nepropojitelné.¹⁵ Divák tedy svým vlastním způsobem dochází k závěru, který vychází z příchozích dat, avšak tento závěr není sám o sobě redukovatelný.¹⁶ Vkládání významu však není jen přirozený proces, je to také proces, který je biologicky, psychologicky a kulturně podmíněný.¹⁷ Proto Pervez i Bordwell tvrdí, že diváci mohou dát jiný význam stejným textům (resp. filmům).

¹² CASETTI, F. *Filmové teorie 1945-1990*. Str. 287

¹³ PERVEZ, A. Cognitive Film Theory. In: *Bangladesh Film and Television Institute Journal*, str. 381- 86

¹⁴ BORDWELL, D. *A case for cognitivism*. Str. 18

¹⁵ CONRAD, D. A Functional Model of the Aesthetic Response. In: *Journal of Contemporary Aesthetics*,

¹⁶ BORDWELL D. *A case for cognitivism*, str. 18

¹⁷ PERVEZ, A. Cognitive Film Theory. In: *Bangladesh Film and Television Institute Journal*, str. 381- 86

2. UI ve filosofickém pojetí

„My nazýváme sami sebe Homo sapiens- člověk moudrý, protože naše inteligence je pro nás velice důležitá. Po tisíce let jsme se snažili pochopit, jak myslíme; to znamená, jak jen hrstka hmoty může vnímat, pochopit, předvídat a manipulovat s mnohem větším a složitějším světem než jsme my sami. Pole umělé inteligence, nebo UI, jde ještě dále: snaží se nejen pochopit, ale také budovat inteligentní entity.“ Russell a Norvig

V této kapitole bude v první části charakterizována umělá inteligence pro účely diplomové práce. V druhé části budou nastíněny teorie a filosofické předpoklady UI, které budou dále analyzovány ve filmové tvorbě. Na toto téma neexistuje jednotná shoda nebo šablona, které bychom se mohli při tvoření výčtu filosofických teorií UI držet. Pro tuto práci byla záměrně vybrána tři nejprobíranější filosofická témata v UI, těmi jsou inteligence, vědomí a etika. Všechny spolu úzce souvisí a v teoriích se často prolínají. Definice UI byly tedy vybrány za účelem pokrytí co největšího pole UI, proto se jejich znění může zdát až příliš obecné.

2.1. Vybrané definice UI

Tajemství lidské inteligence způsobuje potíže myslitelům a vědcům již od té doby, co se poprvé stala vnímaným atributem lidské existence. Co je to inteligence? Které aspekty lidské povahy nám umožňují ji vlastnit? Je možné vytvořit umělou inteligenci? Pokud ano, co přesně by toto vytvoření znamenalo pro lidský druh? Již termín „umělá“ je velice kontroverzní. Tento termín můžeme charakterizovat jako něco, co není produktem přirozené evoluce, avšak je to vytvořeno lidmi. Umělá věc podle Ivana M. Havla musí splňovat tři podmínky. První podmínkou je to, že musí existovat přirozená věc, která je podkladem pro věc umělou. V případě UI to může být například myšlení, vnímání, rozhodování atd. Za druhé musí mít člověk touhu vytvořit kopii této

přirozené věci. A třetí podmínka spočívá v realizaci tohoto duplikátu. Opět v případě UI by to bylo například vložení programu pro rozhodování do počítače.¹⁸

Dalším problémovým termínem je slovo intelligence. Vladimír Mařík v knize *Umělá intelligence I.* charakterizuje inteligenci jako „...*vlastnost některých živých organismů, která jim dává v přírodě mimořádné postavení. Vznikla a vyvíjela se v průběhu dlouhého vývoje. Dnes umožňuje některým živým organismům efektivně reagovat na složité projevy prostředí a aktivně je využívat ve svůj prospěch, k dosažení svých cílů.*“¹⁹ Intelligence je tedy schopnost efektivně řešit složité problémy a představit si nové metody, které nutně nemusí vycházet z čisté výpočetní síly.

Avšak pokud člověk zkoumá problém formalizování inteligentního systému do jakékoliv tradiční logické struktury, intelligence se téměř nedá charakterizovat. Lidé již nejspíše vědí, jak naprogramovat obecnou kreativitu. A také již v roce 1948 poznamenal John von Neumann v reakci na to, že mohou existovat inteligentní stroje: „...*vy trváte na tom, že existuje něco, co stroj nedokáže udělat. Když mi přesně řeknete, co to je, co stroj nedokáže udělat, tak já pokaždé můžu donutit stroj udělat přesně tuto věc.*“²⁰

Russell a Norvig v knize *Artificial Intelligence: A modern approach* tvrdí, že existují dvě různé dimenze, na které se soustředí definice UI. První dimenze se týká myšlenkových procesů a uvažování, které směřují k určitému chování. Druhou oblast pokrývají definice, které pojednávají o racionalitě UI, která je založena na tom, že systém vykonává „správné věci“.²¹

Jedním z nejvíce uznávaných vymezení UI se stala podle Maříka, Štěpánkové a Lažanského definice Marvinina Minského, která vychází z Turingova imitačního testu a mohli bychom ji zařadit do první skupiny definic podle Russella a Norviga. „*Umělá intelligence je věda o vytváření strojů nebo systémů, které budou při řešení určitého úkolu využívat takového postupu, který – kdyby ho dělal člověk – bychom považovali za*

¹⁸ HAVEL, I. M. *Přirozené a umělé myšlení jako filosofický problém*, str. 16

¹⁹ MAŘÍK, V.; ŠTĚPÁNKOVÁ O.; LAŽANSKÝ J. *Umělá intelligence I.*, str. 15

²⁰ HALES, COLIN G., *The Revolutions of Scientific Structures*, str. 201

²¹ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1

projev jeho inteligence.“²² Z této definice můžeme vyvodit, že UI řeší takové úlohy, ke kterým by potřeboval i člověk zapojení své inteligence. Tato charakteristika je založena na znalostech entity a její schopnosti řešit složité úlohy. Minsky implicitně svázal existenci inteligence s úkoly nebo akcemi, které musí inteligentní stroj zvládnout. Avšak některé okrajové úlohy UI Minského definice nezahrnuje, zabývá se spíše přesně stanovenými funkcemi.²³

Jako další, méně radikální a více obecnou definici UI, která spadá do druhé oblasti, bychom mohli zmínit tu Nillse J. Nilssona. „*Umělá inteligence je aktivita věnovaná vytváření inteligentních strojů a inteligence je taková kvalita, která umožňuje entitě správně fungovat a také předvídat ve svém prostředí. V návaznosti na tuto definici, hodně věcí – lidé, zvířata a některé stroje vlastní inteligenci. Stroje, jako jsou ‚chytré kamery‘, a spousta zvířat jsou v tomto primitivním konci rozšířeného kontinua, podle kterého jsou uspořádány entity na různých stupních inteligence.*“²⁴ Z tohoto hlediska závisí charakteristika UI na někom, kdo je schopen dát syntetizovanému softwaru a hardwaru pro fungování „správnost“, „přiměřenost“ a „schopnost předvídat“. Jednoduchá elektronická kalkulačka provádí výpočty mnohem rychleji než lidský mozek a téměř nikdy neudělá chybu.²⁵ Je zmíněná kalkulačka podle této definice inteligentní? Stejně jako Nilsson má mnoho vědců názor, že inteligence spočívá na vícerozměrném spektru. Podle této definice nespočívá rozdíl mezi aritmetickou kalkulačkou a lidským mozkem v druhu, ale v rozsahu, rychlosti, míře autonomie a obecnosti.²⁶

„*UI je podoborem takzvané počítačové vědy, ten je věnovaný vývoji programů, které umožňují počítačům projevovat chování, které lze (obecně) označit jako inteligentní.*“²⁷ Tato definice patří přímo vědcům Russellovi a Norvigovi. Většina výzkumů v prostředí UI je věnována poměrně úzkým aplikacím, jako je plánování nebo

²² MAŘÍK, V.; ŠTĚPÁNKOVÁ O.; Lažanský J. *Umělá inteligence I.*, str. 17

²³ Tamtéž, str. 17

²⁴ NILSSON, N. J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, str. 29

²⁵ MCCORDUCK, P. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*, str. 32

²⁶ Tamtéž, str. 35

²⁷ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1

překládání řeči. Toto je realizováno v limitovaných a dobře definovaných úlohových oblastech. Zásadní zájem však zůstává v dlouhodobém úsilí o budování obecně inteligentních, autonomních strojů i přesto, že je cíl úplně lidské inteligence takřka nepolapitelný.²⁸

2.2. Filosofické aspekty a předpoklady UI

V této podkapitole budou nastíněny filosofické problémy, které můžeme nalézt v oboru UI, jelikož vztah filosofie a UI zahrnuje mnoho společných pojmů – například akce, cíl, víra a vědomí. Avšak UI bere v potaz to, co můžeme nazvat „návrhářským“ postojem k těmto termínům. Ptá se tedy, jaké druhy znalostí, přesvědčení, vědomí atd., počítačový systém potřebuje, aby se choval inteligentně a jak toto postavit a zařadit do počítačového systému.²⁹ Filosofové v tomto oboru zastávají abstraktnější pole působnosti tím, že se ptají, co jsou znalosti, inteligence, vědomí apod. „Návrhářský“ postoj se zabývá stávajícím artefaktem nebo organismem a zvažuje, jak z biologického organismu navrhnout artefakt. Pro UI jsou filosofické otázky důležité, zvláště pokud se chce UI zabývat lidskostí.³⁰ Avšak od 70. let se UI od lidské nápodoby UI stále vzdaluje a začíná se soustředit spíše na aplikaci technik a teorií UI na konkrétní problémy.³¹

John McCarthy ve svém článku *The philosophy of AI and AI of philosophy* z roku 2006 kritizuje své kolegy za to, že věnují příliš málo úsilí sjednocení filosofie s UI nebo alespoň propůjčování některých aspektů filosofie do UI. Zmiňuje zde například nedostatečné jazykové znalosti programu Deep Blue, který porazil světového šachového velmistra Garryho Kasparova.³² V 50. letech minulého století se výzkumníci v oblasti UI zamýšleli nad inteligencí stroje na úrovni inteligence člověka. Za největšího průkopníka výzkumu UI můžeme považovat Alana Turinga, ten jako první ve svém článku *Computing Machinery and Intelligence* zdůraznil, že by mohla být UI

²⁸ NILSSON, N. J., Logic and artificial intelligence. In: *Artificial Intelligence*, str. 31–56.

²⁹ MCCARTHY, J. The Philosophy of AI and AI of Philosophy, In: *Philosophy: the journal of the British Institute of Philosophy*, str. 3

³⁰ Tamtéž, str. 3

³¹ Tamtéž, str. 4

³² Tamtéž, str. 4

realizována počítačovými programy.³³ Podle Johna McCartyho se vědci jeho doby soustředili na UI na lidské úrovni a na metody, jak této lidské úrovni dosáhnout, nejvíce za posledních 40 let.³⁴

Dále McCarthy dodává, že i když se práce zaměřená na lidskou úroveň UI zvyšuje, zůstanou mezi výzkumem UI a bádáním filosofie významné metodologické rozdíly. Avšak dále tvrdí také toto: „*Filosofie X, kde X je věda, zahrnuje filosofy, kteří analyzují pojmy X a někdy soudí, jaké pojmy jsou nebo pravděpodobně nebudou koherentní. UI má blízké vědecké souvislosti s filosofií než jiné vědy, protože UI sdílí s filosofií mnoho společných koncepcí, např. vědomí, epistemologii a nakonec i svobodnou vůli.*“³⁵

Filosofie UI spočívá v zodpovězení následujících otázek:

- Může stroj myslet a jednat inteligentně? Může vyřešit jakýkoli problém, který pomocí přemýšlení vyřeší člověk?
- Může mít stroj mysl, vědomí nebo mentální stavy stejné jako člověk? Je lidský mozek v podstatě počítač?
- Je stroj eticky způsobilý? Může mít svobodnou vůli? Může vnímat věci takové, jaké jsou?

Filosofické koncepce budou pro tuto práci rozděleny do tří skupin. První skupina teorií bude mít jako společný prvek inteligenci. Druhou skupinu bude spojovat termín vědomí. A třetí oblast filosofických koncepcí UI bude věnována koncepcím etiky. Tyto tři oblasti UI nejsou odděleny tlustou čarou, naopak se velice překrývají a vycházejí jedna z druhé. Jsou vybrány také z hlediska pozdějšího analyzování filmů, jelikož například logika strojů by se ve filmech zkoumala s těžší. Kdežto inteligence, vědomí a etika se ve filmové tvorbě řeší explicitně. Diplomová práce se nebude zabývat těmito teoriemi zcela podrobně, jelikož každá z nich by si zasloužila vlastní práci.

³³ NATH, R. *Philosophy of Artificial Intelligence. A Critique of the Mechanistic Theory of Mind*, str. 15

³⁴ MCCARTHY, J. The Philosophy of AI and AI of Philosophy. In: *Philosophy: the journal of the British Institute of Philosophy*, str. 4.

³⁵ Tamtéž, str. 3

Nejprve je důležité upřesnit terminologii, zvláště **pojetí slabé a silné UI**. Toto rozdělení UI poskytl John Searle ve svém článku *Minds, Brains and Programs*. Tvrzení, že by se stroje mohly chovat tak, jako by disponovaly inteligencí, se nazývá hypotéza slabé UI, která se v dnešní době označuje spíše jako narrow AI, českým ekvivalentem úzká či omezená UI.³⁶ Slabá UI se vyznačuje tím, že se soustředí pouze na jednu úlohu. Známým příkladem slabé UI je Siri, inteligentní personální asistent. Siri totiž působí v omezeném předem definovaném rozsahu, nedisponuje obecnou inteligencí nebo vědomím.³⁷ Slabá UI se nejlépe definuje v kontrastu k silné UI, kterou sám Searle charakterizuje takto: „*Přiměřeně naprogramovaný počítač s pravými vstupy a výstupy by tak měl mysl ve stejném smyslu, jakým myslí lidské bytosti.*“³⁸ Tedy pokud stroj dokáže opravdu myslet (ne na úrovni simulace), disponuje vědomím a vnímáním, pak spadá do hypotézy silné UI. Definice silné UI záleží na diferencí mezi simulací mysli a skutečností, že někdo/něco myslí disponuje. Searle tvrdí, že podle silné UI je korektní simulací mysli právě mysl, naopak podle slabé UI je správná simulace modelem mysli.³⁹

2.2.1. Slabá UI – Mohou se stroje chovat inteligentně?

Obor UI byl založen na myšlence, že slabá UI je více než možná. To, jestli je možná nebo ne, záleží na tom, jak ji definujeme. Podle Russella a Norviga chápe většina badatelů slabou UI jako samozřejmost. Pokud program funguje tak, jak má, nezajímá je, zda mají inteligenci stroje považovat za simulaci nebo pravou inteligenci.⁴⁰ Počátek UI jako akademické disciplíny se datuje rokem 1956, ve kterém se konala konference na Dartmouth College. Zde akademici jako Allen Newell (CMU), Herbert Simon (CMU), John McCarthy (MIT), Marvin Minsky (MIT) a Arthur Samuel (IBM) učinili tvrzení, že každý aspekt učení nebo jakýkoli jiný rys inteligence může být tak

³⁶SAENZ, A., We Live in a Jungle Of Artificial Intelligence that will spawn Sentience. In: *SingularityHub*

³⁷CALO, R. *Sorcerers Apprentice or why Weak AI interesting Enough*. In: <http://cyberlaw.stanford.edu/blog/2011/08/sorcerers-apprentice-or-why-weak-ai-interesting-enough>

³⁸DENNETT, D. C. *Consciousness explained*, str. 242

³⁹SEARLE, J., *Minds, Brains and Programs*. In: *Behavioral and Brain Sciences*, str. 417–457

⁴⁰RUSSELL, S. J.;NORVIG P. *Artificial Intelligence: A modern approach*, str. 1020

přesně popsán, že může být vytvořen program pro jeho simulaci.⁴¹ UI je v tomto předpokladu založena na myšlence, že slabá UI je více než možná. Avšak existovali i odpůrci této myšlenky, kteří tvrdili, že vynalézt program pro simulaci není možné. „*Umělá inteligence sledovaná v kultu výpočetní techniky není ani duchem šancí na dosažení trvalých výsledků.*”⁴²

Nejzákladnější filosofickým problémem UI, je srovnání dvou architektur – člověka a stroje. Nejvíce frekventovanou otázkou na poli filosofického bádání v UI je, „...*zda mohou stroje myslet?*“.⁴³ Počítačový vědec Edsger Dijkstra (1984) uvádí, že otázka „*Mohou stroje myslet?*“ je asi tak relevantní jako otázka, zda umí ponorky plavat. Obecná definice plavání zní takto: „*hýbat se ve vodě pomocí končetin, ploutví nebo ocasu.*“ Díky této definici se jistě shodneme na tom, že ponorky, které nemají končetiny, plavat neumí. Avšak další námitka tkví v definici létání a otázce, zda letadla umí létat. Definice létání zní: „*létání je pohybování se vzduchem pomocí křídel nebo křídlových částí*“.⁴⁴ Na rozdíl od ponorky a plavání zde usoudíme, že letadla létat umějí, jelikož mají křídla. Nicméně definice, otázky ani odpovědi nemají žádný vliv na konstrukce letadel a ponorek, a už vůbec vliv na jejich schopnosti. Praktická uskutečnitelnost myslících strojů existuje pouze půl století, což je krátký čas na změnu definice „*myšlení*“ v jazykových slovnících, kde myšlení podle definice stále vyžaduje mozek nebo alespoň jeho část.⁴⁵

Jak už bylo řečeno dříve, za jeden z největších zlomů ve filosofii UI můžeme považovat Turingův článek *Computing Machinery and Intelligence* (1950). Průlom učinil především ve změně dotazování. Místo otázky, zda mohou stroje myslet, bychom se měli začít ptát, zda mohou stroje projít testem behaviorální inteligence, tedy **Turingovým testem**.⁴⁶ Turingův test se zakládá na tzv. imitační hře. Ta je hrána třemi lidmi, mužem (A), ženou (B) a dotazujícím (C), který může být ženského i mužského

⁴¹ MCCORDUCK, P. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*, str. 13

⁴² RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1020

⁴³ Tamtéž, str. 1020

⁴⁴ Tamtéž, str. 1021

⁴⁵ Tamtéž, str. 1020

⁴⁶ Tamtéž, str. 1022

pohlaví. Dotazující musí být separován od zbylých dvou účastníků.⁴⁷ Cílem hry je, aby dotazující rozhodl, který ze zbylých dvou účastníků je muž a který žena. Muž a žena jsou označeni písmeny X a Y. Dotazující má možnost ptát se zbylých dvou „hráčů“ na otázky typu „*Mohl by mi prosím X říci, jak má dlouhé vlasy?*“.⁴⁸ Účastník z vedlejší místnosti má za úkol dotazující zmást natolik, aby v určování pohlaví chyboval. Samozřejmě celá konverzace probíhá v písemné podobě, aby při rozhodování dotazujícího nefiguroval ženský nebo mužský hlas dotazovaných.⁴⁹ Žena může mást dotazujícího odpověďmi typu: „*Já jsem žena, neposlouchej toho muže.*“. Dále se Turing ptá, co by se mohlo stát, kdybychom místo muže stanovili účastníkem hry počítač. Zde vyvstává zcela nový problém. Tento nový problém má výhodu v ostrém dělení fyzická a intelektuálních schopností člověka. Stroj má za úkol písemně konverzovat s člověkem pět minut.⁵⁰ Dotazující musí určit, zda konverzuje s počítačem nebo s člověkem, a pokud stroj zmate člověka alespoň ze třiceti procent, prošel slavným Turingovým testem. Turing se domníval, že bude test strojem pokořen do roku 2000. Jeho domněnka byla nesprávná, počítače stále nejsou schopny oklamat člověka natolik, aby ve více než čtvrtce svých soudů chyboval.⁵¹ Avšak v případě, kdy lidé nevěděli, že konverzují s počítačem, Turingovým testem prošel program ELIZA, internetový chatboti MGONZ a nebo NATACHATA.⁵² Známým příkladem je také CYBERLOVER, který simuloval flirtujícího muže. Internetové servery upozorňovali uživatele programu CYBERLOVER, že krade personální data. Může dojít až k odcizení identity člověka.⁵³ Od roku 1991 se každoročně pořádá Loebnerova cena, ta nese titul nejdéle konané soutěže typu Turingova testu. Tyto soutěže vedly k preciznější typologii lidských chyb, které by se mohly v imitační hře vyskytnout. Sám Alan Turing zkoumal širokou škálu

⁴⁷ TURING, A. M., Computing Machinery and Intelligence. In: *Mind*, str. 433-460

⁴⁸ Tamtéž, str. 433-460

⁴⁹ TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*, str. 15

⁵⁰ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1022

⁵¹ Tamtéž, str. 1022

⁵² Tamtéž, str. 1022

⁵³ ROSSI, S. *Beware the CyberLover that Steals Personal Data*. In:

<https://www.pcworld.com/article/140507/article.html>

možných i existujících námitek proti jeho testu.⁵⁴ Některé z nich budou v této práci zmíněny.

Nejprve se pozastavíme nad **námitkou „hlavy v písku“** (angl. Heads in the Sand Objection). Tato námitka se týká tvrzení, že důsledky toho, že umožníme strojům myslet, by byly příliš hrozné.⁵⁵ V době, kdy Turing psal tento článek, nebyla hrozba UI častým tématem vědců. V současnosti na toto téma upozorňuje nespočet filmů a badatelů, kteří se v oboru UI pohybují, jako např. K. Warwick, N. Bostrom nebo E. Musk.⁵⁶ Turing tvrdí, že rádi věříme lidské nadřazenosti z hlediska evoluce. Tento antropocentrismus platí dodnes. Pokud označíme člověka jako poslední vývojové stádium, nemusíme se obávat žádné hrozby ani ze strany UI.⁵⁷ Dokud se můžeme považovat za jediné myslící entity tohoto světa na lidské úrovni, nemusíme se podle Turinga bát žádného převratu. Tento argument není z jeho pohledu dostatečně významný.⁵⁸

Další zmíněnou námitkou je **námitka z různých neschopností** (angl. Disability).⁵⁹ Námitka z nemožnosti je založena na tvrzení, že „stroj nikdy nemůže udělat X.“⁶⁰ Sám Turing ve svém článku *Computing Machinery and Intelligence*, zmiňuje příklady postižení stroje: „*Být milý, vynalézavý, krásný, přátelský, mít iniciativu, mít smysl pro humor, odříkat si špatné, dělat chyby, zamilovat se, užít si jahody a krém, někoho milovat, učit se ze zkušenosti, používat slova správně, být předmětem své vlastní myšlenky, mít tolik rozmanitostí chování jako člověk, dělat něco opravdu nového.*“⁶¹ Avšak pokud se podíváme do minulosti, některé z těchto „strojových nemožností“ jsou více než možné. Jelikož počítače, které dělají chyby, jsou dnes velmi běžnou součástí našeho života. Podle Russella a Norviga známe také věc, do které je snadné se zamilovat, plyšového medvídka. A šachový počítačový expert David Levy tvrdí, že do

⁵⁴ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1022

⁵⁵ TURING, A. M., *Computing Machinery and Intelligence*. In: *Mind*, str. 433-460

⁵⁶ TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*, str. 15

⁵⁷ TURING, A. M., *Computing Machinery and Intelligence*, In: *Mind*, str. 433-460.

⁵⁸ Tamtéž, str. 433-460

⁵⁹ TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*, str. 18

⁶⁰ Tamtéž, str. 18

⁶¹ TURING, A. M., *Computing Machinery and Intelligence*. In: *Mind*, str. 433-460

roku 2050 bude láska mezi strojem a člověkem velmi rozšířenou záležitostí.⁶² Počítače do dob Allana Turinga prošly tak neuvěřitelným vývojem, že dnes obvykle vykonávají úkoly na úrovni člověka. To ovšem neznamená, že k řešení těchto úkolů používají pochopení nebo porozumění. A je samozřejmé, že existuje spousta úkonů, na které počítač ještě nestačí, jako je například Turingův test a rozvíjení lidské konverzace.⁶³

Jeden z nejvýraznějších a dosud převládajících kritik UI je Turingem nazvaný **argument z neformálnosti chování** (angl. argument from informality of behavior).⁶⁴ V podstatě jde o tvrzení, že lidské chování je příliš komplexní na to, aby bylo zachyceno jakýmkoli jednoduchým souborem pravidel. A tím, že počítače mohou následovat pouze striktně určená pravidla, nemohou být nikdy považovány za inteligentní na stejné úrovni jako lidé. Nikdy jim nebude například připsána schopnost improvizace.⁶⁵ Neschopnost zachytit vše v sadě logických pravidel se v UI nazývá kvalifikační problém (angl. the qualification problem in AI).⁶⁶ Hlavní osobou skrývajícím se za tímto názorem byl filosof Hubert Dreyfus, který má ve svém repertoáru sérii vlivných kritik UI. *What Computers Can not Do* (1972), pokračování *What Computers Still Can not Do* (1992), *Mind Over Machine* (1986).⁶⁷ Vědci zastávající tento názor se staví proti „dobré staromódní UI“ tzv. GOFAI (angl. Good Old-Fashioned AI) – termín vytvořený filosofem Johnem Haugelandem.⁶⁸ GOFAI stojí na předpokladu, že veškeré inteligentní chování může být zachyceno logickým systémem, který je definován souborem faktů a pravidel popisujících tuto doménu. Podle Dreyfusa zrcadlí lidská expertiza určitý „holistický kontext“ nebo „pozadí“, v němž lidé operují. Většinu myšlenkových operací nelze vystopovat introspekci, ale to podle Dreyfuse neznamená, že myšlenkové procesy neexistují.⁶⁹ Vyvstává důležitá otázka, na kterou ovšem Dreyfus neodpovídá, a tou je, jak můžeme zkoumat lidskou mysl? Tedy, jak se dostat člověku do hlavy? Na tuto otázku neexistuje explicitní odpověď. Nabídnut nám je pouze známý

⁶² LEVY, D. *Love + sex with robots: the evolution of human-robot relations*, str. 281

⁶³ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1023

⁶⁴ Tamtéž, str. 1023

⁶⁵ TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*, str. 19

⁶⁶ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1024

⁶⁷ Tamtéž, str. 1024

⁶⁸ Tamtéž, str. 1024

⁶⁹ Tamtéž, str. 1024

komentář Daniela Denneta: „*Je to spíše jako by se filosofové prohlašovali za experty na vysvětlování metod jevištních kouzelníků, a když se zeptáme, jak kouzelníci dělají ten trik s rozpúlením ženy, oni řeknou, že je to velice zřejmé: ten kouzelník ji doopravdy nerozpúlí; on jednoduše udělá trik, aby se to nám to tak zdálo. ‚Ale jak to udělá?‘ se zeptáme. ‚Ne naše oddělení‘ řeknou filosofové.*“⁷⁰

2.2.2. Silná UI – Mohou stroje opravdu myslet?

Většina filosofů tvrdí, že stroj, který projde Turingovým testem, nedisponuje opravdovým vědomím, ale pouze jeho simulací. Tuto kritiku předpovídal i sám Turing. **Argument z vědomí** neformuluje Turing sám, ve svém článku používá slova britského neurologa G. Jeffersona: „*Dokud nebude stroj schopen napsat sonet nebo složit koncert na základě svých myšlenek a emocí, nikoli náhodným výběrem symbolů, nemůžeme přistoupit na tvrzení, že stroj rovná se mozek. To znamená, že stroji nestačí sonet jen napsat, ale musí i vědět, že jej napsal. Žádný mechanismus nemůže cítit (nejen uměle dávat najevo, což je snadné) uspokojení z úspěchu nebo smutek ze spálené elektronky, nemůže být potěšen lichotkami, zarmoucen svými chybami, okouzlen opačným pohlavím, rozzloben nebo deprimován, když není schopen dosáhnout toho, co chce.*“⁷¹ Tato námitka stojí na třech aspektech, prvním je neschopnost stroje mít vědomí. Druhým je to, že stroj není schopný vytvořit umělecké dílo. A třetím aspektem je nemožnost vytvořit stroj s emocemi.⁷² Nejdůležitějším aspektem této námitky je otázka vědomí. Jeffersonův klíčový bod se nachází spíše ve fenomenologii nebo nalezení přímé zkušenosti – stroj by měl mít vlastní pocit emocí.⁷³ Jiní badatelé se soustředí na záměrnost, tedy na to, zda údajná víra, touha a další reprezentace stroje jsou ve skutečnosti něco reálného, nebo zda se jedná pouze o simulaci.⁷⁴ Turing zde argumentuje tím, že ani u člověka nemůžeme dokázat „nějaké úmyslné vědomí něčeho“, jedná se o solipsismus.⁷⁵

⁷⁰ RUSSELL, S. J.;NORVIG P. *Artificial Intelligence: A modern approach*, str. 1024

⁷¹ TVRDÝ, Filip. *Turingův test: filozofické aspekty umělé inteligence*, str. 21

⁷² Tamtéž, str. 21

⁷³ RUSSELL, S. J.;NORVIG P. *Artificial Intelligence: A modern approach*, str. 1024

⁷⁴ Tamtéž, str. 1024

⁷⁵ TURING, A. M., *Computing Machinery and Intelligence*, In: *Mind*, str. 433-460

Na rozdíl od Turinga je John Searle skeptický k vytvoření umělé mysli. "Nikdo nepředpokládá, že počítačová simulace bouře způsobí, že budeme všichni mokří... Proč by někdo v nejlepší míně předpokládal, že počítačová simulace duševních procesů má skutečně duševní procesy?"⁷⁶ Z citace Johna Searlea není jasné, jak tuto analogii implementovat na problém vědomých duševních stavů. Koneckonců i díky simulaci bouře pomocí postřikovačů a větrných strojů můžeme být mokří. Turingova odpověď naznačuje, že otázka pravosti vědomí časem sama vymizí, jelikož stroje dosáhnou určité úrovně sofistikovanosti. To by mělo také za důsledek vymizení rozdílů mezi slabou a silnou UI.⁷⁷ V současné době však stále existuje problém, který lze vyjádřit tím, že lidé disponují reálnou myslí, ovšem stroje jí disponovat mohou, ale také nemusejí. Jak to, že lidé mají skutečnou mysl a ne pouze tělo, které vytváří určité neurofyziologické procesy?⁷⁸ Tuto otázku řeší filosofie už celá staletí, počátek můžeme datovat až do starověkého Řecka. Tato záhada mysli a těla označována jako **mind-body problém**. Úsilí o vyřešení tohoto problému vede přímo k dotazu, zda mohou stroje disponovat opravdovou myslí. První hlubší analýzu mind-body problému nám nabídl René Descartes ve své *Meditaci o první filosofii* (1641).⁷⁹ Descartes uvažoval o myšlenkové aktivitě vědomí a fyzických projevech těla až dospěl k závěru, že tyto dva procesy musí existovat odděleně. Tato úvaha se dnes označuje jako **dualistická teorie**. Otázka, která vzniká na bázi dualistické teorie, zní, jak může mysl řídit tělo, když jsou mysl a tělo separované entity.⁸⁰ Descartes spekoval, že tyto dvě entity interagují prostřednictvím epifýzy, což opět vyvolává otázku, jak mysl řídí epifýzu.

Naopak **monistická teorie mysli**, nazývaná také jako **fyzikalismus**, se této otázce vyhýbá tvrzením, že mysl není oddělena od těla – duševní stavy jsou fyzické stavy. Tím fyzikalismus v zásadě povoluje možnost silné UI. Současným problémem fyziků je podat vysvětlení, jak mohou fyzické stavy (zejména molekulární konfigurace a elektrochemické procesy v mozku) být současně také duševními stavy, jako je bolest,

⁷⁶ SEARLE, J. R. Minds, brains, and programs. In: *Behavioral and Brain Sciences*, str. 417–457

⁷⁷ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1027

⁷⁸ Tamtéž, str. 1027

⁷⁹ STOLJAR, D. Physicalism, In: *The Stanford Encyclopedia of Philosophy*

⁸⁰ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1027

užívání si chuti jídla, vědění, že někdo řídí auto, nebo když člověk věří tomu, že Praha je hlavní město České republiky.⁸¹

Badatelé na poli filosofie se snažili vysvětlit, co znamená, že člověk (resp. počítač či stroj) je v určitém duševním stavu. Zabývali se zejména **intencionálními stavy**, jako je vědět, toužit, věřit, bát se atd. Tyto stavy odkazují na určité aspekty vnějšího světa.⁸² Pokud se fyzikalismus nemýlí, tak správný popis určitého duševního stavu člověka musí být určen nějakým mozkovým stavem této osoby. Pokud se tedy zrovna v této chvíli v mysli soustředím na psaní diplomové práce, můj současný stav mozku je jedním z příkladů třídy duševních stavů „vědět, že tento člověk píše diplomovou práci“.⁸³ Specifické uskupení všech atomů v mém mozku není pro tuto záležitost esenciální, existuje totiž mnoho konfigurací mého nebo cizího mozku, které by patřily ke stejné třídě duševních stavů. Klíčovým bodem této úvahy je fakt, že stejný stav mozku nemůže nikdy odpovídat jinému mentálnímu stavu.⁸⁴

Jednoduchá fyzikalistická myšlenka je zpochybněna různými myšlenkovými experimenty. Jedním z nich je experiment s názvem **mozek v kádi** (angl. Brain in a Vat). Tento myšlenkový pokus je aktualizovanou verzí Descartovy představy zlého démona, kterou představil Gilbert Harman.⁸⁵ Podle tohoto experimentu si máme představit, že náš mozek byl odstraněn z těla při narození a uložen do řádně upravené nádrže. Upravená nádrž udržuje náš mozek při životě a umožňuje mu růst a rozvíjet se. Právě v tomto okamžiku jsme mozkiem zapojeni do sofistikovaného počítačového programu, který dokonale simuluje zážitky z externího světa.⁸⁶ Ve skutečnosti je život, který si myslíme, že žijeme, pouhou simulací bytí. Mohli byste mít shodný stav mozku s tím, že „píšete diplomovou práci“, ale bylo by chybné říkat doslovně, že máte mentální stav „vědět, že tento člověk píše diplomovou práci“. Pokud máte mozek ve specializované kádi, můžete si být jistí, že reálně diplomovou práci nepíšete. Nikdy jste

⁸¹ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1027

⁸² STOLJAR, D. Physicalism. In: *The Stanford Encyclopedia of Philosophy*

⁸³ Tamtéž

⁸⁴ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1027

⁸⁵ HARMAN, G. *Thought*, str. 5

⁸⁶ HICKEY, A. P. *The Brain in a Vat Argument*. In: <https://www.iep.utm.edu/brainvat/>

psaní diplomové práce nezažili, a proto nemůžete mít ani odpovídající duševní stav.⁸⁷ Argument mozku v kádi byl předpokladem k filmu Matrix, ve kterém byla celá lidská rasa umístěna do obrovských kádí a žil ve virtuální realitě v rukou zákeřné umělé inteligence (která byla vytvořena lidmi).

Myšlenkový experiment mozku v kádi se zdá být v rozporu s názorem, že mozkové stavy určují ty mentální. Jedním řešením tohoto dilematu je prohlášení, že obsah mentálních stavů lze interpretovat ze dvou různých hledisek, **širokého** (angl. wide content) a **úzkého mentálního obsahu** (angl. narrow content).⁸⁸ Široký obsah jako rozpor mozkových a mentálních stavů se interpretuje z hlediska všudypřítomného vnějšího pozorovatele, který může rozlišovat rozdíly ve světě. Široký obsah závisí na vlastnostech prostředí i rysech jednotlivce. Úzký mentální obsah nezávisí na životním prostředí jedince a stal se velmi kontroverzním tématem pro vědecké prostředí.⁸⁹ Široký obsah je vhodný, pokud má člověk v plánu přičítat mentální stavy ostatním, kteří s námi sdílejí existenci, předpovídat jejich pravděpodobné chování, účinky atd. Avšak pokud jednáme o otázce, zda systémy UI skutečně myslí a mají mentální stavy, pak je pro naše úvahy vhodnější stavět se za tezi úzkého mentálního obsahu. Je totiž nesmyslné tvrdit, že pokud je UI opravdu myslící entitou, záleží na podmínkách vnějšího prostředí. Mentální obsah jedince (stroje nebo počítače) závisí tedy pouze na jeho individuálních vnitřních vlastnostech.⁹⁰

Dalším filosofickým přístupem v oblasti UI je **funkcionalismus**. Teorie funkcionalismu říká, že duševní stav je jakýkoli přechodný příčinný stav mezi vstupem a výstupem. Tato teorie stojí v opozici k úzkému mentálnímu obsahu a tvrdí, že to, co vytváří mentální stav určitého typu, nezávisí na určitém vnitřním uspořádání člověka, ale spíše na způsobu jeho chování nebo roli, kterou hraje v systému, jehož je součástí.⁹¹ Funkcionalismus také říká, že jakékoli dva systémy s izomorfními kauzálními procesy by měly stejné mentální stavy. Z toho vyplývá, že by mohl mít počítač stejné mentální

⁸⁷ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1027

⁸⁸ Tamtéž, str. 1027

⁸⁹ Tamtéž, str. 1027

⁹⁰ Tamtéž, str. 1027

⁹¹ TOOLEY, M. Functional concepts, referentially opaque contexts, causal relations, and the definition of theoretical terms, In: *Philosophical Studies*, str. 251–279.

stavy jako člověk.⁹² Tato teorie se lépe vysvětluje díky myšlenkovému **experimentu náhrady mozku**. Tento experiment byl představen Clarkem Glymourem, zabýval se jím i John Searl, nejčastěji je však spojován s robotistou Hansem Moravcem.⁹³ Předpokládejme, že neurofyziologie se vyvinula až tam, kde dokonale chápeme chování vstupů a výstupů a propojenost všech neuronů v lidském mozku. Dále máme předpokládat, že je v našich silách vytvářet mikroskopické elektronické přístroje, které napodobují chování neuronů a mohou být jednoduše napojeny na nervovou tkáň. Konečným předpokladem je, že nějaká zázračná chirurgická technika může nahradit jednotlivé lidské neurony odpovídajícími elektronickými přístroji bez přerušení činnosti našeho mozku jakožto celku. Experiment spočívá na nahrazení všech neuronů v mozku těmito miniaturními elektronickými přístroji.⁹⁴ Zajímají nás jak vnitřní, tak vnější zkušenosti subjektu během i po operaci. Výsledkem experimentu by mělo být naprosto identické chování subjektu před i po ukončení operace. Přítomnost nebo nepřítomnost vědomí nemůže být zjištěna z hlediska třetí osoby, proto by změny vědomí měly být zaznamenány přímo z pozice operovaného subjektu. Hans Moravec, robotický výzkumník a funkcionalista, je přesvědčen, že by operace neměla na jeho vědomou zkušenost žádný vliv. Naopak John Searle, filosof a biologický naturalista si je jistý, že by jeho vědomí kompletně zmizelo.⁹⁵ Avšak pro podporu tohoto myšlenkového experimentu a názoru Moravce bychom museli vysvětlit projevy vědomí způsobených elektrickým mozkiem, který apeluje pouze na funkční vlastnosti neuronů. A toto vysvětlení musí mít také uplatnění pro skutečný biologický mozek, který je podkladem pro vytvoření toho elektronického. Existují tři různé závěry těchto úvah. První závěr říká, že kauzální mechanismy vědomí, které generují určité druhy výstupů v biologickém mozku, fungují také v jeho elektronické verzi, která je proto vědomá. Druhý naopak tvrdí, že vědomé mentální události v reálném mozku nemají příčinnou souvislost s chováním a tím pádem je experiment pro biologický mozek fatální a kvůli

⁹² RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1029

⁹³ Tamtéž, str. 1029

⁹⁴ SMART, J. J. C., The Mind/Brain Identity Theory. In: *The Stanford Encyclopedia of Philosophy*

⁹⁵ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1029

němu ztrácíme vědomí. A finální závěr je ten, že experiment je neproveditelný, a proto je nesmyslné o něm spekulovat.⁹⁶

Druhá možnost omezuje mysl na to, co filosofie nazývá **epifenomenalismem**. Epifenomenalismus je názor, že mentální události jsou ovlivněny fyzickým uspořádáním mozku, avšak ty nemají žádný vliv na fyzické události.⁹⁷ Chování je způsobeno svaly, které se stáhnou při přijímání neurálních impulzů a nervové impulsy jsou generovány vstupy od jiných neuronů nebo ze smyslových orgánů. V tomto filosofickém názoru nehrají mentální stavy žádnou kauzální roli. Patricia Churchland tvrdí, že argumenty funkcionalistů, které operují na úrovni neuronů, mohou fungovat také na úrovni jakékoli větší funkční jednotky – neuronů, laloků, hemisfér nebo také celého mozku.⁹⁸ Pokud přijmeme tento názor, dojdeme k tomu, že pokud věříme experimentu záměny biologického mozku za elektronický, který má stále vědomí, pak bychom také měli věřit tomu, že vědomí je udržováno a je schopno aktualizovat svůj stav a mapy od vstupů po výstupy jako obrovská vyhledávací tabule.

Velkým vyzyvatelem funkcionalismu se stal **biologický naturalismus** v čele s Johnem Searlem, podle něhož jsou mentální stavy emergentními vlastnostmi vysoké úrovně, které jsou zapříčiněny fyzikálními procesy v neuronech, jež jsou na nízké úrovni.⁹⁹ Záleží především na nespécifikovaných vlastnostech neuronů. Podle biologického naturalismu nelze mentální stavy duplikovat pouze na základě nějakého programu, který má stejnou funkční strukturu, se stejnými funkcemi vstupů a výstupů. Biologický naturalismus vyžaduje, aby program fungoval na architektuře se stejnou kauzální silou, jako mají neurony.¹⁰⁰ Na podporu této teze dodává Searle popis hypotetického programu, který bez problémů prochází Turingovým testem, ale absolutně nerozumí svým vstupům ani výstupům. Závěr této hypotézy zní, že

⁹⁶ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1030

⁹⁷ Tamtéž, str. 1030

⁹⁸ CHURCHLAND, P. *The Brain and Its Self*, str. 41-50

⁹⁹ SEARLE, J. R. Minds, brains, and programs. In: *Behavioral and Brain Sciences*, str. 417-457

¹⁰⁰ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1031

provozování příslušného programu (tj. mít správné výstupy) není dostatečnou podmínkou pro to, aby měl stroj mysl/vědomí.¹⁰¹

Hypotetická situace se skládá z člověka, který rozumí pouze angličtině, z knihy pravidel psané právě v anglickém jazyce a z různých stohů papíru, některé z nich jsou prázdné a na některých je nečitelný nápis. Člověk zde zastává roli procesoru, kniha pravidel je program a stohy papírů jsou úložným zařízením. Tato soustava je umístěna do místnosti s malým otvorem. Tímto otvorem jsou do místnosti vkládány papírky s nerozlučitelnými znaky. Člověk najde odpovídající symboly v knize a dále postupuje podle pokynů. Pokyny mohou zahrnovat psaní znaků na prázdné papíry, hledání symbolů ve stozích papírů apod. Nakonec díky pokynům člověk dosáhne toho, že budou symboly přepsány na kus papíru, který otvorem předá zpět do vnějšího světa.¹⁰² Zvenčí vidíme svět, kde systém, který přijímá vstupy ve formě čínských vět a generuje odpovědi také v čínštině, které jsou projevem „inteligence“ stejně jako konverzace v Turingově testu. Searle ale namítá, že osoba v místnosti čínštině nerozumí. K tomu, aby předávala vnějšímu světu správně utvořené čínské věty, jí pomáhá kniha pravidel. Z toho vyplývá, že správně utvořený počítačový program nám nezaručí to, že tento program správně pochopí celou konverzaci.¹⁰³ Tento argument se nazývá **Čínský pokoj**.

Stejně jako Turing, Searle zvažoval a snažil se odmítnout řadu argumentů na jeho pojetí. Jedním z těchto argumentů, který byl navržený Johnem McCarthym nebo Robertem Wilenskym, nazval Searle systémovou odpovědí. Tato námitka říká, že otázka, zda člověk v místnosti rozumí čínštině, je analogická k dotazu, jestli zná procesor odmocninu ze tří. V obou případech zní odpověď ne, jelikož oba tyto „systémy“ mají určitou kapacitu. Avšak je samozřejmostí, že pokud se člověk zeptá čínské místnosti (jako celku), odpověď by byla kladná (a to v plynulé čínštině).¹⁰⁴ Turingovi by toto stačilo. Searle se nás snaží přesvědčit, že vlastnost celku musí náležet

¹⁰¹ PRESTON J. M. *Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence*, str. 70-79

¹⁰² SEARLE, J. R. Minds, brains, and programs. In: *Behavioral and Brain Sciences*, str. 417-457

¹⁰³ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1031

¹⁰⁴ Tamtéž, 1031

alespoň jedné z jeho částí.¹⁰⁵ Avšak to opravdové Searleovo tvrzení spočívá v těchto čtyřech axiomech. Za prvé, počítačové programy jsou formální (syntaktika). Za druhé, lidská mysl má mentální obsah (sémantika). Za třetí tvrdí, že syntax není jako samostatná jednotka konstitutivní a ani dostatečná pro sémantiku. A nakonec mozek způsobuje mysl.¹⁰⁶ Z prvních tří axiomů vyplývá, že programy nejsou dostatečné entity pro mysl. A ze čtvrtého plyne, že systém, který by chtěl duplikovat lidskou mysl, by musel mít minimálně takovou kauzální sílu, jako má lidský mozek sám. Tyto axiomy jsou kontroverzní, jelikož první dva axiomy se opírají o blíže nespecifikovaný rozdíl mezi syntaxí a sémantikou, který se zdá být blízký rozdílu mezi úzkým a širokým mentálním obsahem. Závěrem myšlenkového experimentu Čínského pokoje je tvrzení, že programy nejsou dostatečné pro mysl.¹⁰⁷ Ale tento závěr je pro akademickou obec neuspokojivý.

Třetím axiomem Searle dokázal svou nelibost vůči funkcionalismu. Jeho cílem je vyvrátit funkcionalistický přístup porozumění mysli, přístup, který říká, že duševní stavy jsou definovány jejich kauzálními rolemi, nikoliv entitami (např. neurony), které tyto role zastávají. Searlův argument je také velice diskutovaný především kvůli tomu, že stojí v opozici k funkcionalistické teorii známé jako **komputační teorie mysli** (také komputační funkcionalismus), která považuje mysl za systém pro zpracování informací.¹⁰⁸ Funkcionalisté tedy třetí axiom odmítají. Argument má za cíl ukázat, že čínská místnost (jako celek) není myslí jen na základě chodu nějakého programu, avšak neříká nic o tom, zda by mohla tato místnost (nebo počítač) vlastnit mysl na základě něčeho jiného. Searle sám pronesl, že nějaké stroje mysl vlastní, například lidé jsou biologické stroje disponující myslí. Searle nepodává žádný důkaz, který by podporoval jeho argument, spoléhá se spíše na intuici. V poslední době se Čínský pokoj označuje jako argument pro vyvrácení silné UI spíše než jen „argument“.¹⁰⁹ Russel a Norvig

¹⁰⁵ PRESTON J. M. *Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence*, str. 70-79

¹⁰⁶ SEARLE, J. R. Minds, brains, and programs. In: *Behavioral and Brain Sciences*, str. 417-457

¹⁰⁷ PRESTON J. M. *Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence*, str. 70-79

¹⁰⁸ POLÁK, M., *Filosofie mysli*, str. 149

¹⁰⁹ SNELL, M. B. Do you have free will? John Searle reflects on various philosophical questions in light of new research on the brain. In: *California Alumni Magazine*

k tomuto argumentu dodávají: "Ale člověk by mohl vytvořit stejný argument o mozku: stačí se podívat na tuto sbírku buněk (nebo atomů), slepě fungujících podle zákonů biochemie (nebo fyziky) – kde zde najdeme mysl? Proč může způsobit kousek mozku myšlenku, zatímco kus jater nemůže? To zůstává velkým tajemstvím."¹¹⁰

Největším problémem všech debat o silné UI je otázka **vědomí**. Možná, že žádná část mysli není více známá nebo více záhadná nežli vědomí. Vědomí stojí na více aspektech, jimiž jsou například pochopení nebo sebevědomění. Tento odstavec se bude týkat subjektivní zkušenosti. Vědomé bytosti můžeme definovat jako ty, které si nejen uvědomují, ale také si uvědomují, že jsou si vědomi, a tak ustanovují výtvar vědomí jako určitou formu **sebevědomování**.¹¹¹ Důležitým termínem v otázce sebevědomování jsou kvalitativní stavy. Tyto stavy mohou být považovány za vědomé ve zdánlivě odlišném a kvalitativnějším smyslu. To znamená, že jeden stav může být vědomý, jestliže má nebo obsahuje kvalitativní nebo zkušenostní vlastnosti druhu často nazývané „kválie“ nebo „surové sensorické pocity“. Člověk vnímá pití vína nebo látku, kterou zkoumá, pociťuje jako vědomý duševní stav, protože zahrnuje různé sensorické vlastnosti, např. chuťové kválie ve víně a kválie barvy ve své vizuální zkušenosti s látkou. Avšak panuje značný nesouhlas s povahou těchto kválií.¹¹² Dokonce se vedou diskuze o jejich existenci. Tradiční kválie byla považována za vnitřní, soukromé, monadické rysy zkušenosti, ale současné teorie kválií často odmítají alespoň některé z těchto aspektů.¹¹³ Kválie představují velkou výzvu pro funkcionalisty, neboť různá kválie mohou být zapojena do různých isomorfních kauzálních procesů. Avšak kválie nejsou problémem pouze pro funkcionalisty, ale pro celou vědu. Například Daniel Dennet odmítá existenci kválií, připisuje je totiž takzvanému filosofickému zmatku.¹¹⁴

Sám Turing uznává, že je otázka vědomí obtížná, ale odmítá její důležitou úlohu v praxi UI: "Nechci udělat dojem, že si myslím, že neexistuje žádné tajemství o vědomí... Ale

¹¹⁰ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1031

¹¹¹ Tamtéž, str. 1032

¹¹² Tamtéž, str. 1032

¹¹³ DENNETT, D. C. *Consciousness explained*, str. 242

¹¹⁴ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1032

nemyslím si, že tyto záhady nutně potřebují být vyřešeny dříve, než budeme schopni odpovědět na otázku, s níž jsme se v tomto článku zabývali."¹¹⁵

2.2.3. Etika a rizika vývoje UI

Vědci se zabývají především tím, jak a zda vůbec můžeme rozvíjet UI. Avšak stojí také za zvážení, zda bychom vůbec měli UI rozvíjet. Pokud jsou důsledky vzestupu technologie UI spíše negativní, pak je morální odpovědností vědců zvážit dopady svých výzkumů.¹¹⁶ Mnoho technologických vynálezů mělo neúmyslné negativní vedlejší účinky (např. jaderné štěpení, Černobyl a hrozba globální destrukce). V určitém smyslu můžeme přirovnat automobily k robotům, kteří si podmanili svět, jelikož se staly pro společnost nepostradatelnou součástí.¹¹⁷ Vědci i technici čelí etickým rozporům, které narušují jejich činnost. V roce 2001 vznikla také příručka o etice výpočetní techniky. Avšak vzrůst UI přinesl také nové etické problémy: lidé mohou kvůli automatizaci ztratit svou práci či ztratit pocit jedinečnosti. Systémy UI mohou být použity k nežádoucím účelům. Použití systémů UI může mít za důsledek ztrátu lidské zodpovědnosti. A konečně, úspěch UI může znamenat konec lidské civilizace.¹¹⁸

Etika je odvětví filosofie, která se zabývá koncepcemi morálky (např. otázkami, co je dobré a co je špatné). Koncepce morálky se také týkají této diplomové práce, jelikož ve filmové tvorbě se odrážejí morální aspekty technologického pokroku. Pouze možnost budovat inteligentní stroje, které budou konkurovat lidské inteligenci, je velice spekulativní.¹¹⁹ Techničtí giganti současnosti, jako jsou Alphabet, Amazon, Facebook, IBM a Microsoft, stejně jako teoretický fyzik Stephen Hawking a obchodní magnát, spoluzakladatel a generální ředitel společnosti Tesla Elon Musk, tvrdí, že je nyní čas kriticky se zamyslet nad přínosy a hrozbami UI.¹²⁰ Můžeme vycítit, že v oblasti etiky vyvstává mnoho otázek „co kdyby?“, které souvisejí s naší budoucností, ve které

¹¹⁵ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1032

¹¹⁶ SINGER, P. *Applied ethics*, str. 280

¹¹⁷ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1034

¹¹⁸ Tamtéž, 1034

¹¹⁹ BUTAZZO, G. *Artificial Consciousness: Utopia or Real Possibility*, str. 24- 30

¹²⁰ BOSTROM, N. *Superintelligence: Až budou stroje chytrější než lidé*, str. 36

figuruje UI. Filosof Nick Bostrom a teoretik UI Eliezer Yudkowsky argumentují, že se tyto otázky týkají jak zajištění toho, aby stroje nepoškodily lidstvo a jiné morálně odpovědné bytosti, tak morálního stav samotných strojů. Nejdůležitější otázkou tedy je, zda UI bude hrozbou pro lidstvo, nebo změni naše životy k lepšímu. Jiný problém se týká samotných lidí, jak navrhují, konstruují a používají umělé bytosti. Jak by tedy měly být inteligentní stroje navrženy, aby jednaly eticky?

Prvním problémem nástupu UI byla výše zmíněná **ztráta práce kvůli automatizaci**. Moderní průmyslová etika se stala závislou na počítačových systémech. Můžeme říct, že tisíce zaměstnanců ztratily svou práci kvůli programům UI. Avšak doposud automatizace prostřednictvím informačních technologií obecně a zejména UI vytvořila více pracovních míst, než bylo eliminováno, a vytvořila zajímavější a lépe placená pracovní místa. Při úvahách k 25. výročí AAAI stanovil Nils Nilsson jako budoucí výzvu vytvořit UI, která neprojde Turingovým testem, ale zkouškou zaměstnanosti.¹²¹ Byl by to tedy robot, který by zvládl pracovat na různých pozicích. Podle Russella a Norviga můžeme skončit v budoucnosti s obrovskou nezaměstnaností nebo také jako manažeři vlastního pracoviště plného robotů.¹²²

Lidé mohou mít v budoucnu příliš mnoho (nebo příliš málo) volného času. Alvin Toffler napsal ve svém díle *Future Shock* z roku 1970: "*Pracovní týden byl od přelomu století snížen o 50 procent. Není zcela mimo, když budeme předvídat, že se do roku 2000 opět sníží na polovinu.*"¹²³ Arthur C. Clarke zase poznamenal, že v roce 2001 budou lidé žít ve světě naprosté nudy, kde bude existenciálním problémem, který z několika set televizních kanálů si člověk vybere.¹²⁴ Jediná z těchto předpovědí, která se vyplnila, je ta s televizními kanály. UI zvyšuje tempo technologických inovací, a tím přispívá k většímu pracovnímu tlaku na kvalitu práce zaměstnanců. Tím pádem má v dnešní době zatím opačný účinek, než byl očekáván v minulém století.

¹²¹ NILLSON, N. J., *Human-Level Artificial Intelligence? Be Serious!*, str. 68-75

¹²² RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1034

¹²³ Tamtéž, str. 1034

¹²⁴ Tamtéž, str. 1034

Lidé mohou ztratit v budoucnu pocit jedinečnosti. V návaznosti na program ELIZA Weizenbaum poukazuje na potencionální hrozbu, kterou UI představuje pro lidstvo.¹²⁵ Jedna z těchto námitek naráží na fakt, že UI umožňuje myšlenku, že lidé jsou automaty. Tato myšlenka by mohla vést ke ztrátě autonomie, nebo dokonce zániku celého lidstva. Pokud bude UI slavit v budoucnu úspěchy, mohla by přinejmenším ohrožovat morální předpoklady společnosti 21. století, stejně jako Darwinova evoluční teorie v 19. století.¹²⁶

Vynálezy UI mohou být použity k nekalým účelům. Vyspělé technologie bývají používány k potlačení síly soupeřů. Autonomní systémy UI jsou dnes na bitevním poli běžně využívány. Americká vláda nasadila do Iráku přes pět tisíc samostatných letadel a dvanáct tisíc autonomních pozemních vozidel.¹²⁷ Jedna etická teorie tvrdí, že zapojení vojenských robotů je bráno jako vynález brnění ve středověku. Nikdo nemá námitky proti vojákovi, který by chtěl nosit helmu.¹²⁸ Avšak robotické zbraně představují také hrozbu, pokud převezmou na svá bedra lidská rozhodování. Mohlo by docházet k zabíjení nevinných civilistů. Při troše nadsázky můžeme říct, že větší zapojení robotů může vést k přehnanému sebevědomí vojenských sil a nastartovat vlnu zbytečně přehnaných útoků. Ve většině válek totiž dochází k podobným problémům jako je přehnané sebevědomí jedné strany, jinak by totiž války končily dohodou a ne bezhlavým zabíjením.¹²⁹ Také technologie rozpoznávání řeči by mohla vést ke katastrofě v podobě rozsáhlého odposlechu a ke ztrátě občanských svobod.¹³⁰ Weizenbaum nepředpovídal svět, který je ohrožován terorismem, ale správně věštil potenciál hromadného dohledu UI. Jeho předpověď se stala zčásti realitou: Británie má nyní rozsáhlou síť kamer a jiné země běžně monitorují telefonní hovory nebo provoz na internetu. Někdo se domnívá, že nástupem počítačů se vytrácí autonomie, svoboda a soukromí. Tato ztráta soukromí je nevyhnutelná a způsob, jak bojovat proti ztrátě

¹²⁵ WEIZENBAUM, J. *Computer Power and Human Reason*, str. 10

¹²⁶ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1035

¹²⁷ SINGER, P. W. *Wired for War*, str. 79

¹²⁸ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1035

¹²⁹ Tamtéž, str. 1035

¹³⁰ WEIZENBAUM, J. *Computer Power and Human Reason*, str. 12

svobody a větší moci státu oproti individuální síle, je zpřístupnit dohled všem občanům.¹³¹

Využití programů UI může vést ke ztrátě lidské odpovědnosti.¹³² Pokud se například lékař spoléhá na posouzení diagnózy zdravotnického expertního systému, kdo je vinen, pokud je diagnóza nesprávná? Ovšem díky rostoucímu vlivu teoreticky-rozhodovacích metod v medicíně se nyní toto paradigma změnilo na otázku: Kdo je vinen, pokud je diagnóza nepřiměřená?¹³³ Soudy mají odpověď na tuto otázku zcela jasnou, expertní rozhodovací programy řadí na roveň lékařským učebnicím a příručkám. Lékař tedy musí být obeznámen se všemi aspekty „počítačového“ verdiktu a plně s ním souhlasit, pokud se mu zdá přiměřený. Pokud se tyto odborné programy stanou spolehlivě přesnějšími než lékaři, lékaři by se mohli stát plně zodpovědnými, pokud nebudou považovat „počítačové“ diagnózy za správné.¹³⁴ Stejný problém se týká takzvaných internetových agentů, kteří manipulují s penězi. Velice probíraným tématem jsou v dnešní době také autonomní vozidla a jejich zodpovědnost za silniční dění.

Úspěch UI může znamenat konec lidského života na Zemi. Téměř každá technologie má schopnost změnit se v účinnou zbraň, pokud padne do nesprávných rukou. Ale s technologií UI vzniká zcela nový problém, nesprávné „ruce“ by totiž mohly patřit samotné technologii. Tento předpoklad můžeme doložit řadou sci-fi příběhů (nejen filmových, ale také knižních). Mezi nejznámější příklady patří *Frankenstein: Moderní Prometheus* (1818) Mary Shelleyové nebo hra Karla Čapka *R. U. R.* (1921), která se stala celosvětově známou především díky pojmenování robotů. Z filmové tvorby můžeme prozatím zmínit například *Terminator* (1984), který kombinuje vzpouru robotů a cestování v čase. A *Matrix* (1999), který spojuje roboty dobývající svět a experiment mozků v kádi.¹³⁵ Zdá se, že roboti jsou protagonisty nesčetně pověstí o dobytí a zkáze světa, stejně jako v minulosti čarodějnice nebo

¹³¹ RUSSELL; S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1036

¹³² Tamtéž, str. 1036

¹³³ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 1036

¹³⁴ GAWANDE, A. *Complications: A Surgeon's Notes on an Imperfect Science*, str. 21

¹³⁵ MAŘÍK, V.; O. ŠTĚPÁNKOVÁ a J. LAŽANSKÝ. *Umělá inteligence III.*, str. 15

marťané z Války světů.¹³⁶ Otázkou je, zda UI představuje větší riziko než tradiční hrozby. Za prvé může být odhad určitého systémového stavu UI nesprávný, což bude mít za důsledek nesprávné chování systému. Například autonomní auto může nesprávně odhadnout polohu protijedoucího auta nebo člověka přecházejícího ulici, což zapříčiní nehodu, která by mohla člověka stát život. Nebo vážnější příklad, kdy systém protiraketové obrany nesprávně odhadne útok a kvůli tomu zahájí protiútok, který by ohrozil ještě více životů než dopravní nehoda. Tato rizika nemůžeme přičíst pouze strojům, protože stejnou chybu by mohl udělat taktéž člověk. A existuje také způsob, jak toto riziko zmírnit, pokud navrhne systém s kontrolami a váhami, aby k chybnému odhadu nemohlo dojít, nebo bychom objevili a vyřešili takovou chybu ještě před katastrofickými důsledky.¹³⁷

Druhým rizikovým faktorem je nesnadné zadání správné funkce UI pro maximalizaci systému. Mohli bychom navrhnout například funkci systému, která bude určena pro minimalizaci lidského utrpení. Vzhledem k tomu, že známe lidskou povahu, vždycky najdeme způsob jak trpět i v ráji.¹³⁸ Takže správným rozhodnutím pro UI je vyhladit celou lidskou rasu. Musíme být velice opatrní, co vyžadujeme po systému UI, jelikož stroj si těžko domyslí, že nemůže brát nějaký úkol doslovně. Avšak lidskost je někdy také na škodu. Lidé mají sklony k iracionálnímu agresivnímu chování díky tendencím přirozeného výběru.¹³⁹ Stroje, které budujeme, nemusí mít dán do vínku agresivní způsob chování, pokud nechceme. Nebo pokud se agrese neobjeví jako konečný produkt mechanismu.¹⁴⁰ Musíme doufat, že robot, který je natolik chytrý, aby věděl, jakým způsobem může vyhubit lidskou rasu, je také dostatečně chytrý, aby věděl, že toto po něm jeho programátor nechtěl (pokud to tedy opravdu nebylo jeho záměrem).

Třetím a nejzávažnějším scénářem hrozby UI je samostatné vyvinutí UI v systém s nechtěným chováním. Již v roce 1965 I. J. Good napsal: "*Ultrainteligentní stroj můžeme definovat jako stroj, který může daleko předstihnout veškeré intelektuální*

¹³⁶ RUSSELL, S. J.; P. NORVIG. *Artificial Intelligence: A modern approach*, str. 1037

¹³⁷ Tamtéž, str. 1037

¹³⁸ MAŘÍK, V.; O. ŠTĚPÁNKOVÁ a J. LAŽANSKÝ. *Umělá inteligence III.*, str. 15

¹³⁹ RUSSELL, S. J.; P. NORVIG. *Artificial Intelligence: A modern approach*, str. 1037

¹⁴⁰ Tamtéž, str. 1037

činnosti každého člověka, jakkoli je člověk chytrý. Vzhledem k tomu, že konstrukce strojů je jednou z těchto intelektuálních aktivit, může ultrainteligentní stroj navrhnout ještě lepší stroje; pak by nepochybně nastala "intelligenční exploze" a inteligence člověka by zůstala daleko za nimi. První ultrainteligentní stroj je tedy posledním vynálezem, který člověk vůbec potřebuje, za předpokladu, že je stroj dostatečně poslušný, aby nám řekl, jak ho udržet pod kontrolou."¹⁴¹ Mluvíme zde o intelligenční explozi, která bývá také nazývána jako technologická singularita. Vernon Vinge o tomto jevu tvrdí, že během třiceti let budeme mít technické prostředky k tomu, abychom vyrobili nadlidskou inteligenci. Krátce poté bude ukončena lidská éra.¹⁴² Good, Vinge a mnoho dalších správně poznamenávají, že křivka technologického vývoje v současné době roste exponenciálně (v návaznosti na Mooreův zákon). Tato křivka však může pokračovat v singularitě nekonečného vývoje.¹⁴³ Můžeme si všimnout, že koncepce ultrainteligentního stroje předpokládá, že inteligence je velmi důležitým atributem UI. A pokud jí máte dostatek, mohou tím být vyřešeny všechny problémy. Víme ovšem, že i UI má určitou výpočetní hranici (horní hranice rychlosti je například rychlost světla), ovšem výpočetní horní hranici stále neznáme.¹⁴⁴ Většina lidí je přicházející singularitou znepokojena, avšak existují i lidé, kteří se těší na budoucnost, v níž jsou lidé buď nahrazeni nebo alespoň vylepšeni robotickou technologií. Pro toto aktivní společenské hnutí používáme název transhumanismus.¹⁴⁵ Mezi vědce, kteří zastávají singularitu, patří například Ray Kurzweil: "*Singularita nám umožní překonat tato omezení našich biologických těl a mozku. Získáme moc nad našimi osudy. Naše úmrtnost bude v našich rukou. Budeme schopni žít tak dlouho, jak budeme chtít (jemně odlišné prohlášení, že budeme žít navždy). Plně pochopíme lidské myšlení a výrazně rozšíříme a zvětšíme jeho dosah. Do konce tohoto století bude nebiologická část naší inteligence bilionkrát silnější, než jen lidská inteligence.*"¹⁴⁶ Avšak Kurzweil poukazuje také na negativní

¹⁴¹ GOOD, I. J. Speculations concerning the first ultraintelligent machine. In: *Advances in Computers*, str. 31-88

¹⁴² RUSSELL, S. J.; P. NORVIG. *Artificial Intelligence: A modern approach*, str. 1037

¹⁴³ Tamtéž, str. 1037

¹⁴⁴ Tamtéž, str. 1037

¹⁴⁵ MORAVEC, H. *Robot: mere machine to transcendent mind*, str. 150

¹⁴⁶ KURZWEIL, Ray. *The singularity is near: when humans transcend biology*, str. 18

dopady singularity. Jedním z nich je zesílení schopnosti jednat destruktivně, podle něj není tento příběh ještě dopsán.¹⁴⁷

Pokud jsou ultrainteligentní stroje možností pro budoucí technologický vývoj, my lidé bychom měli takovéto stroje navrhnout tak, aby s námi jednali co nejlépe. Tímto tématem se zabýval již Isaac Asimov ve své krátké povídce *Runaround* (1942), která byla později vydána ve sbírce s názvem *I, Robot* (1950), pod stejným názvem byl v roce 2004 natočen film, ze kterého bude pro účely diplomové práce čerpáno. Děj se odehrává v Chicagu v roce 2025. Roboti jsou vyráběni proto, aby sloužili lidem. Doručují poštu, venčí psy a vynášejí lidem odpadky. Robot nemůže ublížit člověku kvůli třem zákonům. Ty znějí:

1. *Robot nesmí ublížit lidské bytosti nebo svou nečinností dopustit, aby bylo člověku ublíženo.* (angl. A robot may not injure a human being or, through inaction, allow a human being to come to harm.)
2. *Robot musí uposlechnout příkazů člověka, kromě případů, kdy jsou tyto rozkazy v rozporu s prvním zákonem.* (angl. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.)
3. *Robot musí chránit sám sebe před poškozením, kromě případů, kdy je tato ochrana v rozporu s prvním, nebo druhým zákonem.* (angl. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.)¹⁴⁸

Co se ovšem stane, když robot tyto zákony poruší? Jak je možné zákony porušit? Ve filmu *I, Robot* figuruje stroj, který se chová neintuitivním a neobvyklým způsobem a tento způsob chování neodpovídá logickému vysvětlení tří zákonů. Později byl Asimovem přidán též čtvrtý, nebo také nultý zákon, kvůli možnosti, že roboti převezmou vládu nad celou planetou. Čtvrtý (popř. nultý) zákon zní:

¹⁴⁷ KURZWEIL, Ray. *The singularity is near: when humans transcend biology*, str. 18

¹⁴⁸ PROYAS, A., et al. *I, Robot* [film]

4. *Robot nesmí ublížit lidstvu, nebo svou nečinností dopustit, aby mu bylo ublíženo.*
(angl. A robot may not harm humanity, or, by inaction, allow humanity to come to harm.)¹⁴⁹

Tyto zákony se zdají být, pro nás lidi, přiměřené. Ovšem obtížné je správně zákony implementovat. Zákony jsou totiž definovány kruhem, především vztah mezi druhým a třetím zákonem. To naznačuje, že zákony nemají absolutní logickou hodnotu, ale spíše stojí jako váhy jeden pro druhého (větší váhu má vždy ten, co stojí výš). Je pravdou, že přínosnější by byly zákony, které by si podle výše nebezpečí a kontextu vyhodnotil stroj sám v návaznosti na tři zákony. Yudkowsky uvádí podrobnější informace o tom, jak navrhnout přátelskou UI (angl. friendly AI). Tvrdí, že přátelskost neboli touha nepoškodit člověka, by měla být navržena od úplného začátku, ale že návrháři by měli předvídat, že se robot může časem měnit (tedy učit se a vyvíjet). Výzvou je tedy konstrukce takového mechanismu, který se bude adaptovat na změny systému, ale stále bude naprogramován na přátelskost vůči lidskému druhu.¹⁵⁰ Je nemožné navrhnout stroj, který by fungoval na statických funkcích, jelikož my sami se adaptujeme na situace a naše chování, postoj je aj. se mění v čase. Russell a Norvig uvádějí příklad počítače vyrobeného v roce 1800. Pokud bychom takovému stroji poskytli převládající morálku tehdejší doby, dnes by se zastával o obnovení otroctví a zrušení hlasovacího práva žen. Ten samý rozpor platí pro případ, kdy nastavíme stroj tak, aby odpovídal našemu smýšlení. Zabíjíme otravný hmyz a nepřijde nám to neetické, částečně proto, že hmyzí mozek je oproti našemu primitivní. Jak si můžeme být jisti, že tu samou analogii nepocítí i stroj vůči nám?¹⁵¹

Nakonec budeme muset zvážit, zda by bylo stále morální zacházet s roboty, kteří vlastní vědomí, jako se stroji. Spisovatelé i filmoví tvůrci sci-fi se zabývají otázkou práv robotů. Například film *Artificial Intelligence: AI* (2001) od Stevena Spielberga je příběhem inteligentního robota, který byl naprogramován tak, aby věřil, že je člověk,

¹⁴⁹ GUNN, J. On Variations on a Robot, In: *Isaac Asimov: The Foundations of Science Fiction*, str. 56-81

¹⁵⁰ RUSSELL, S. J.; P. NORVIG. *Artificial Intelligence: A modern approach*, str. 1039

¹⁵¹ Tamtéž, str. 1039

který nedokázal pochopit eventuální opuštění od své matky/majitelky. Nejen tento film tvrdí, že by bylo zapotřebí hnutí za občanská práva pro roboty.¹⁵²

3. UI ve starších sci-fi filmech

Abychom propojili, zdá se, nepropojitelná témata, jako je kognitivní filmová teorie, filmová tvorba a UI, zaměříme se v druhé kapitole na nejznámější filmové hity minulosti a témata UI, které řeší. V hlavní části práce budou analyzovány filmy poslední dekády, proto budou za filmy minulosti považovány všechny filmy starší deseti let. Názvy filmů budou v této práci používány v originálním znění. Filmová tvorba bude analyzována z pozice diváka tak, jak nás k tomu vyzývá právě kognitivní filmová teorie.

3.1. Úzkost způsobená UI

V průběhu historie sci-fi filmů představujících formy UI se stále opakuje jedno téma, a tím je úzkost.¹⁵³ Aspekt úzkosti je charakteristickým ve filmové tvorbě se zaměřením na UI, která má snahu zničit lidstvo. Samozřejmě z těchto dob existuje několik filmů, které zobrazují UI jako užitečné a spolehlivé stroje, které naopak lidstvu slouží. Avšak většina filmové tvorby se zaměřuje na UI jako na nebezpečí pro svou schopnost pracovat proti lidskosti ve snaze vytvořit autonomní druh.

Toto tvrzení můžeme doložit analýzou osmdesáti tří filmů Roberta Fishera, jak je popsáno v jeho článku *AI and Cinema – Does Artificial Insanity Rule?* z roku 2005. Díky své analýze Fisher zjistil, že funkcí těchto filmů je nechat diváka prozkoumat povahu uměle vytvořených strojů, dále, jak může fungovat jejich myšlení a jaké bude mít důsledky jejich integrace do společnosti.¹⁵⁴ Ve svém výzkumu rozlišoval Fisher mezi UI s myslí a bez ní. Dospěl k závěru, že z celkového počtu osmdesáti tří filmů šedesát šest ilustrovalo UI bez myslí, která nemá sebereflexi nebo sebevědomí. Fisher ve své stati tvrdí, že s několika málo výjimkami tyto filmy zobrazovaly UI hlavně jako

¹⁵² RUSSELL, S. J.; P. NORVIG. *Artificial Intelligence: A modern approach*, str. 1040

¹⁵³ SMITH, H. *It's Alive: A Reflection of Conscious AI in Cinema*, str. 2

¹⁵⁴ FISHER, R. B. *AI and Cinema – Does Artificial Insanity Rule?* In: *Representations of Artificial Intelligence in Cinema*, str. 4

robotické efektivní vražedné stroje. Jako příklady můžeme uvést diktátorský počítač Alpha 60 z filmu *Alphaville* (1965), robotické klony ve *Futureworld* (1976) nebo vražedně naprogramované roboty ve filmu *Runaway* (1984).¹⁵⁵ Podle Fishera může být toto filmové pojetí umělé inteligence pro diváky nezajímavé, neboť u bytosti bez vědomí nemůže nikdy nastat morální konflikt. Autor tedy podotýká, že u zbývajících třiceti sedmi filmů je téma divácky zajímavější, neboť je fascinuje rozostření hranic mezi lidskými bytostmi a inteligentními stroji. Ale i zde se UI vykládá jako něco, čeho se lidstvo bojí.

3.2. Deviantní chování a mentální abnormality

Úzkost, kterou tato forma filmového zobrazení vyvolává v publiku, může být podle Fishera částečně vysvětlena opakujícími se formami deviantního chování inteligentních strojů, jelikož UI vykazuje ve většině filmů abnormální chování od obsedantního až po patologicky šílené.¹⁵⁶ Fisher poznamenává, že způsob chování inteligentních strojů se velice podobá lidským poruchám chování, s tím rozdílem, že často dochází k extrémům těchto poruch. Existuje velká pravděpodobnost, že bude lidský druh postižen různými stupni těchto abnormalit, a proto jsou inteligentní stroje považovány za nebezpečné. Fisher našel v minulé filmové tvorbě několik opakujících se forem deviantního „lidského“ chování inteligentních strojů. První porucha, které si všiml, spočívá v tom, že jsou stroje posedlé tím, aby je lidské bytosti milovaly.¹⁵⁷ Příkladem je film *A. I. Artificial Intelligence* (2001), tato futuristická pohádka je o umělém dítěti jménem David, které hledá lásku matky. Další filmy, kde stroj hledá zoufale lásku, jsou *Making Mr. Right* (1987) a *Electric Dreams* (1984). Další formou deviantního nebo abnormálního chování v minulých sci-fi filmech je touha umělých bytostí, stát se lidskými nebo alespoň fyzickými. Příkladem je film *Demon seed* (1977), *Virtuosity* (1995), *Star Trek: Generations* (1994), *Bicentennial Man* (1999). Finální forma deviace, kterou může trpět stroj ve filmu trpět, je jistý druh megalomanie. Tedy prožívají určitou touhu po všemohoucnosti. *Virtuosity* (1995) je jedním z příkladů

¹⁵⁵ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 4

¹⁵⁶ Tamtéž, str. 4

¹⁵⁷ Tamtéž, str. 4

takového filmu. V něm počítačově generovaný zabiják Sid, spojení sto osmdesáti tři reálných lidských příšer včetně Adolfa Hitlera, přichází do skutečného světa, aby pokračoval v panování teroru. Zmíněné poruchy chování jsou společné pro inteligentní stroje v science fiction filmové tvorbě a vyvolávají pocity strachu a neklidu v divácké sféře.¹⁵⁸

Vedle abnormálního chování vykazují stroje v minulé science-fiction tvorbě také mechanické odchylky, které mohou vyvolat v divácké sféře úzkost. Jedním z důvodů, proč jsou inteligentní stroje ve filmech prezentovány jako ohrožující lidstvo, je jejich paranoidní snaha sebeobranu.¹⁵⁹ Stejně jako lidé, některé formy UI ve filmech mají až přehnaný instinkt přežití. Tento instinkt ohrožuje lidstvo, protože mají stroje paradoxně stejný pocit hrozby od lidstva, jako my máme od strojů.¹⁶⁰ Lidé jakožto jejich tvůrci mají také pravomoc k jejich vypnutí.¹⁶¹ Jedním z příkladů tohoto pojetí UI je Skynet z filmů *Terminator*. Jelikož byl Skynet naprogramován tak, aby za každou cenu pokračoval ve své činnosti, z hrozby vypnutí začne zabíjet lidský druh.¹⁶² Totéž platí pro počítač HAL 9000 z filmu *2001: A Space Odyssey* (1968).¹⁶³ Chování HALu je však také způsobeno mechanickou abnormalitou, chybné a neopodstatněné úvahy inteligentního stroje vycházejí výhradně z rozhodování na bázi formální logiky na rozdíl od lidských úvah. Rozhodnutí HALu zabít astronauty pramenilo z jeho cíle pokračovat v misi, což mu bylo dáno v naprogramovaných směrnicích.¹⁶⁴ Stroje mohou navíc abnormálně fungovat kvůli chybnému vyhodnocení lidského chování tím, že si ho neumí logicky vysvětlit.¹⁶⁵ Lidé se naopak rozhodují díky svým referencím ke světu a znalostem o něm. Následkem toho se může lišit lidský pohled na svět od toho, jak na něj nahlíží stroj. Proto se chování inteligentních strojů může zdát lidem

¹⁵⁸ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 5

¹⁵⁹ DINELLO, D. *Technophobia! Science Fiction Visions of Posthuman Technology*, str. 58

¹⁶⁰ Tamtéž, str. 58

¹⁶¹ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 5

¹⁶² Tamtéž, str. 5

¹⁶³ BUTTAZZO, G. Artificial Consciousness: Utopia or Real Possibility?, In: *Computer*, str. 24-30

¹⁶⁴ CALDWELL, T. *Free Will, Technology and Violence in a Futuristic Vision of Humanity – 2001: A Space Odyssey*, str. 133-138

¹⁶⁵ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 5

nepochopitelné.¹⁶⁶ Například ve filmu *Colossus: The Forbin Project* (1969) se několik počítačů rozhodne na obranu přijmout opatření proti lidským bytostem, aby splnily úkol a chopily se kontroly nad lidským světem. Volba, kterou učinily, je v rozporu s lidským myšlením ublížit co nejmenšímu počtu lidí. Poslední mentální abnormalitu můžeme nalézt ve velké většině minulé sci-fi tvorbě, je to určitý komplex nadřazenosti, který se v inteligentních strojích rodí díky výpočetní rychlosti, mechanické síle a vynikajícímu vnímání.¹⁶⁷ Výsledkem tohoto megalomanského komplexu je to, že stroje necítí potřebu chovat se k lidskému druhu řádně a s úctou.

Závěrem lze konstatovat, že tato literární rešerše některých odborných článků o zobrazení UI v minulých sci-fi filmech ukazuje, že technologie je většinou zobrazena jako zhouba lidstva kvůli mentálním a/nebo behaviorálním abnormalitám většinou myslících strojů. Důležité je poznamenat, že se tyto dva typy zobrazení nevyklučují, spíše jsou většinou neodmyslitelně spojeny. Tato filmová tvorba naznačuje, že umělá inteligence může být v budoucnosti škodlivá pro lidstvo na různých stupních. Filmy mohou být vnímány jako varovné příběhy nebo jako širší komentář k vnímání nebezpečí této technologie.¹⁶⁸

3.3. (ne)lidskost

V další podkapitole se zaměříme na to, proč je UI zobrazována ve filmech právě tímto způsobem. Na jedné straně je UI často označována jako hrozba, protože lidé mají tendenci cítit se nervózní kvůli rychlému rozmachu nových technologií a scénářům budoucnosti, které mohou vytvořit.¹⁶⁹ V tomto smyslu se UI neliší od jiných technologií. Filmy odrážejí naše úzkosti způsobené technologiemi například tím, že ukazují svět, ve kterém je možné nahradit lidskou duši mechanickou napodobeninou.¹⁷⁰ Filmy mohou také zobrazovat děsivou budoucnost způsobenou inteligentními stroji,

¹⁶⁶ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 6

¹⁶⁷ Tamtéž, str. 6

¹⁶⁸ WEEDMAN, D. Men, Monsters and Morality: Shaping Ethics through the Sublime and Uncanny. In: *Humanities Capstone Projects*, str. 1-24.

¹⁶⁹ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 7

¹⁷⁰ SCHELDE, P. *Androids, humanoids, and other science fiction monsters: science and soul in science fiction films*. str. 238

kteře nemají řádne skutečné zkušenosti nebo zdravý rozum stejně jako mají narození, pocity vlastníci, smrtelní lidé.¹⁷¹ Fisher ve svém článku tvrdí, že stroje nejsou takto zobrazeny kvůli technickým nedostatkům, ale proto že jsou pro lidstvo snáze pochopitelné.¹⁷² Je to také způsob, jak posílit lidské ego, protože některé filmy ukazují, že se lidé dokáží zachránit díky své schopnosti myslet lépe, jasněji a/nebo rozuměji než inteligentní stroje.¹⁷³

Na druhou stranu filmový teoretik Smith tvrdí, že zobrazení UI ve starších filmech je nepochybně výsledkem šílenství lidstva. Lidé mají totiž také své nedostatky, mezi které patří sociální problémy, jako je izolace, nedostatečná nebo nevhodná socializace nebo nedostatek pozornosti.¹⁷⁴ Mnoho z výše zmíněných sci-fi filmů nepochybně pouze pocit lidskosti UI, ale také pocit lidskosti u lidí samotných. V současném vědeckém stavu, v němž jsou hranice mezi lidmi a stroji stále tenčí, filmy varují nejen před povahou UI, ale také před nejistotou lidského druhu.¹⁷⁵ Toto je další druh varovné formy představované v minulých sci-fi filmech. Podle Smithe je v nich také kritizována naše vlastní nehumánnost, jelikož se snažíme vylepšit svůj vlastní život pomocí nových technologií a máme očekávání, že se k nám budou chovat taktně a s respektem. Avšak lidstvo má potíže se takto chovat i ke svému vlastnímu druhu, natož pak k uměle vytvořeným bytostem.¹⁷⁶ Film *Runaway* (1984) je pro člověka varovným příběhem a pojednává o našem vlastním nelidském chování.¹⁷⁷ Divák nepřičítá robotům vlastnosti zločinců, a to navzdory hrozbě pro lidstvo, jelikož tak činí v sebeobraně proti zneužívání lidmi.¹⁷⁸ Replikanti ve filmu *Blade Runner* (1982), dále android Bishop v *Aliens* (1986) nebo Terminátor z filmu *Terminator 2: Judgement Day* (1991), jsou příkladem filmových inteligentních strojů, které na rozdíl od lidských postav v těchto

¹⁷¹ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 7

¹⁷² Tamtéž, str. 7

¹⁷³ Tamtéž, str. 7

¹⁷⁴ SMITH, G. M. An Invitation to Feel. In: *Film Structure and the Emotion System*, str. 3–14

¹⁷⁵ SCHELDE, P. *Androids, humanoids, and other science fiction monsters: science and soul in science fiction films*, str. 238

¹⁷⁶ SMITH, H. *It's Alive: A Reflection of Conscious AI in Cinema*, str. 5

¹⁷⁷ SCHELDE, P. *Androids, humanoids, and other science fiction monsters: science and soul in science fiction films*. str. 157

¹⁷⁸ Tamtéž, str. 157

filmech vyjadřují soucit a empatii.¹⁷⁹ Lidské postavy jsou v těchto filmech až nelidské, poté se v kontrastu k těmto postavám zdají být stroje více lidské.¹⁸⁰ Proto některé filmy z minulosti zpochybňují „lidskost“ lidského druhu více než stroje.¹⁸¹

Tato analýza odborných článků ukazuje, že většina minulé science fiction tvorby jsou varovné příběhy a některé z nich varují přímo před nebezpečím UI. Tyto filmy naznačují, že UI může být v budoucnu škodlivá pro lidstvo kvůli jejich mentálním a/nebo behaviorálním odchylkám. Toto zjištění můžeme potvrdit argumentem de Mula, že UI odráží jak naše naděje, tak strach z jejich nekontrolovatelného, destruktivního potenciálu.¹⁸² Shapiro napsal o světě, který predikují sci-fi filmy z padesátých let: "V těchto filmech svět vypadal jako hrozivý, tekutý, chaotický, neosobní, složený ze sil, které jsme zřídka kdy pochopily a rozhodně nikdy neovládaly. Strach je zaměřen na neznámé, neviditelné hrůzy, které se skrývají pod povrchovou normálností."¹⁸³ V tomto budoucím světě, říká Shapiro, bylo lidstvo předurčeno ke smrtelné cestě morální degradace a nakonec i k apokalyptickému zatracení a zkáze.¹⁸⁴ Jinými slovy, ve sci-fi filmové tvorbě minulého století vládl strach.

Tato analýza také ukazuje, že filmy z 20. století zpochybňují volby, které lidstvo dělá. Na filmové úrovni se lidský druh rozhodl zobrazovat inteligentní stroje jako šílené, megalomanské a nelidské, protože jsou v tomto zobrazení UI pro diváky lépe pochopitelné a snesitelné. Vyobrazením UI jako zkázy se nad ní mohou lidské bytosti nadřazovat právě svou lidskostí. Avšak tyto filmy kritizují také lidstvo z hlediska jeho volby, zpochybňují naši humánnost a neutuchající touhu po transcenci. To znamená, že velkolepost našeho intelektu je stále více vnímána jako síla, která nás ovládá

¹⁷⁹ CALDWELL, T. *Free Will, Technology and Violence in a Futuristic Vision of Humanity – 2001: A Space Odyssey*, str. 133–137

¹⁸⁰ Tamtéž, str. 133–137

¹⁸¹ PARKER, L. Human After All: Ex Machina's Novel Take on Artificial Intelligence. In: *The Atlantic*

¹⁸² DE MUL, J. The Technological Sublime. In: *Next Nature Network*

¹⁸³ SHAPIRO, B. Universal Truths: Cultural Myths and Generic Adaptation in 1950s Science Fiction Films. In: *Journal of Popular Film and Television*, str. 103–111

¹⁸⁴ Tamtéž, str. 104

a ohrožuje, což je další varovný scénář, který se často vyskytuje v minulé science fiction filmové tvorbě.¹⁸⁵

4. Filosofie UI v současné filmové tvorbě

„At some point in the early twenty-first century all of mankind was united in celebration. We marveled at our own magnificence as we gave birth to AI.
Morpheus, *The Matrix*

Filmová tvorba může být pro filosofické bádání velice přínosná, zčásti proto, že filmoví tvůrci vytváří filmové médium s prostředky pro vyjádření filosofické myšlenky. Živé a vzrušující vyvolávání skrytých filosofických myšlenek filmaře může motivovat a řídit určitou filosofickou reflexi současných problémů týkajících se například UI. Avšak na toto pojetí můžeme poskytnout dvě námitky. Nejprve je možné stěžovat si na individualitu autora, který do filmu vkládá svou unikátní myšlenku, kterou si divák může špatně vyložit. Z toho vychází druhá námitka, a tou je individuální chápání filmu tak, jak jej chápe divák. Kvůli těmto dvěma námitkám byla pro práci vybrána kognitivní filmová teorie jako metoda chápání filmové tvorby. Ta ospravedlňuje jakékoli pochopení filmu z pozice diváka.

Pro diplomovou práci byl vybrán seznam nejlépe hodnocených filmů na serveru IMDb. IMDb je celosvětová online databáze zaměřená na filmy, seriály, televizní pořad a další. Vznikla v roce 1990. Tento server vlastní společnost Amazon. K prosinci 2017 můžeme v databázi nalézt 4,7 miliónů titulů a 83 miliónů uživatelů.¹⁸⁶ Diplomová práce bude pojednávat o filmech ze seznamu deseti nejlepších filmů o UI poslední dekády. Těmi jsou: *Her*, *Ex Machina*, *Wall-e*, *Big Hero 6*, *Chappie*, *Robot & Frank*, *Prometheus* a *Machine*.¹⁸⁷ Tento seznam byl vytvořen před devíti měsíci podle nejvyššího hodnocení uživatelů. Ze seznamu serveru IMDb byly pro účely práce záměrně eliminovány dva filmy. Prvním je snímek *Moon*, ve kterém se sice UI vyskytuje

¹⁸⁵ DE MUL, J. The Technological Sublime. In: *Next Nature Network*

¹⁸⁶ WIKIPEDIA. Online: <https://en.wikipedia.org/wiki/IMDb>

¹⁸⁷ IMDb. Online: <http://www.imdb.com/list/ls068528362/>

v podobě robota, ten ale není hlavním tématem filmu. Zastupuje pouze vedlejší roli bez jakýchkoli zápletek filosofického charakteru. Druhým vyřazeným je film *Avengers: Age of Ultron*, kde opět vystupuje UI v podobě hlavní záporné postavy, ovšem i tento film postrádá hlubší filosofický kontext. V následující filmové produkci budou tedy analyzována filosofická témata, která byla charakterizována v předchozích kapitolách.

4.1. HER – Ona (2013)

„You know, I actually used to be so worried about not having a body, but now I truly love it. I'm growing in a way that I couldn't if I had a physical form. I mean, I'm not limited – I can be anywhere and everywhere simultaneously. I'm not tethered to time and space in the way that I would be if I was stuck inside a body that's inevitably going to die.“ Samanta, *Her*

Film *Her* byl uveden do kin v roce 2014 a byl nominován na nespočet cen. Například za původní scénář získal cenu Academy Awards. Film zkoumá, jakým způsobem je lidstvo ovlivněno nově přicházejícími technologiemi. Pokud srovnáme rozbor filmu *Her* s analýzou zobrazení UI ve filmech 20. století, musíme si všimnout rozdílu v pohledu na budoucnost s UI, jelikož otázka „co kdyby“ je řešena odlišným způsobem. Díky Fisherově analýze minulých sci-fi filmů jsme zjistili, že tyto filmy byly především dystopické povahy, zatímco *Her* je spíše filmem utopickým. Na rozdíl od starších filmů se nesoustředí na určitou kulturní úzkost. Místo toho je zde UI představena jako něco, co přinese nevyhnutelnou změnu našeho způsobu života – něco jako logický závěr stále více digitalizovaného světa. Film naznačuje, že UI může re-socializovat lidi (více než a-socializovat). Kritizuje spíše lidské chování. To, jak se snažíme naplňovat naše potřeby pomocí technologií, než abychom hledali řešení, které nabízí náš vlastní druh.

Na odlišné zobrazení UI ve filmu *Her* mají vliv především změny v technologické oblasti UI. Vzhledem k tomu, že všechny filmy ve Fisherově analýze byly více než deset let staré, UI se v současné tvorbě stala díky technickému vývoji a vědeckým objevům přesnější a propracovanější. UI se dostala do povědomí veřejnosti například

pomocí webových prohlížečů, automatických plánovacích programů, autonomních automobilů nebo domácích robotických vysavačů. V důsledku toho jsme v dnešní době obeznámeni s novými technologiemi více než při produkci a realizaci starších sci-fi filmů. Fisher tvrdí, že se vývoj odráží ve změně filmových zobrazení UI. Zobrazení inteligentních strojů v minulé sci-fi tvorbě se podle něj jeví více vzdálené od možných budoucích forem UI.¹⁸⁸ Na rozdíl od toho se zdá, že obraz UI ve filmu *Her* je z blízké budoucnosti. Podoba operačního systému v tomto filmu je podobná těm, které už známe, i když pokročilejší. Tato rostoucí demystifikace nových technologií vytváří posun v kulturních představách, kde se UI stává více „reálnou“, přesnou a poutavou.¹⁸⁹ Toto tvrzení můžeme aplikovat také na film *Her*.

Film *Her* není zvláštní jen odlišným pojetím UI. Velice poutavá je také úprava filmu, co se týče zvuku (soundtracku), obrazu a celkového stylu. Povrchová dokonalost v podobě pastelových barev, dokonale upravených aktérů a lehce melancholického soundtracku, zpočátku zakrývá vnitřní prázdnotu hlavního hrdiny. Průběh filmu se nás snaží dostat do depresivního rozpoložení, ve kterém je také hlavní protagonista Theodor Twombly. Avšak podle kostýmů nebo výpravy stále netušíme, že se film odehrává v blízké budoucnosti. To zjišťujeme zanedlouho podle zařízení, které vlastní každý člověk tohoto budoucího světa. Zařízení v podobě handsfree reaguje na naše slovní příkazy, čte nám e-maily, pouští hudbu, nebo čte zprávy a počasí. Všichni lidé Theodorovy doby žijí stejným způsobem života, jsou neustále v kontaktu s elektronickými přístroji a sociálními médii, které mají stále na dosah ruky. Ve filmu vypadá tento způsob života neuvěřitelně přirozený pro lidský druh. Takováto situace pro nás není ničím nepochopitelným nebo nepředstavitelným. Theodor vlastní téměř prázdňý, neuspořádaný byt. Noci tráví sám, nanejvýš v herní virtuální realitě. Ale retrospektivní scény nám naznačují, že nebýval vždy sám a býval šťastný.

Náhle se naskytne nová příležitost na život bez pocitu osamělosti, ve formě reklamy: „*Who are you? What can you be? Where are you heading? What's out there?*“

¹⁸⁸ FISHER, R. B. AI and Cinema – Does Artificial Insanity Rule? In: *Representations of Artificial Intelligence in Cinema*, str. 9

¹⁸⁹ PARKER, L. Human After All: Ex Machina's Novel Take on Artificial Intelligence. In: *The Atlantic*

What are the possibilities? Software Element is proud to introduce first artificial intelligence operating system. An intuitive entity that listens to you, understands you and knows you. It's not just an operating system, it's a consciousness. Introducing OSI.“¹⁹⁰ V další scéně již víme, že je Theodor dalším z odběratelů tohoto nového umělého systému. Před aktivováním operačního systému musíme zodpovědět několik otázek, což má vést k přizpůsobení operačního systému potřebám uživatele. „*Are you social or unsocial? Would you like the OS to have a male or female voice? How would you describe your relationship with your mother?*“¹⁹¹ Operační systém dokáže ve vašem hlase vycítit váhavost či nejistotu. A dále jen čekáte na individualizaci vašeho OS, který se stane vaším nejlepším přítelem či dokonce partnerem. První náznak vědomí (nebo spíše sebeuvědomění) můžeme nalézt v situaci, kdy si operační systém sám daruje křestní jméno jen na základě toho, jak toto jméno zní. Nejvíce vás ovšem může překvapit rychlost tohoto operačního systému. Ve dvou setinách sekundy totiž dokáže přečíst knihu *Jaké vybrat jméno dítěti* a ze 180 tisíc jmen si vybrat jméno Samantha. Tento program funguje tak, že v podstatě vlastní intuici, DNA systému je založeno na osobnostech programátorů, kteří ho vytvořili. Ale také vlastní schopnost růstu na základě zkušeností. Díky tomu se vyvíjí každým okamžikem stejně jako člověk. Má o vás zájem, rád se o vás dozvídá něco nového, dokáže vás rozesmát. Ovšem dokáže také zanalyzovat vaše data na hard disku a za velice krátkou chvíli je roztřídit nebo vám pomoci s gramatikou a pravopisem. Umí vlastně úplně všechno. A poté, co si tento operační systém pustíte do života, již nehrajete virtuální hry sám, nespíte sám, vlastně nejste nikdy sám. Máte s kým probírat svůj osobní život a dokonce i jezdit na dovolenou. Avšak otázkou je, zda vám tento „vztah“ prospívá, nebo se kvůli němu vzdalujete entitám s přirozeným biologickým vědomím.

Další způsob, kterým nás operační systém přesvědčuje o své lidskosti, je intonace hlasu, tím dokáže naznačit ironii, soucit, pohrdání a další lidské pocity. Umí také naopak z vaší intonace vyčíst emoci, kterou emoci právě disponujete. Samantha (operační systém) je pro Theodora stejně reálná, jako kdokoli jiný v tomto digitálním světě. Avšak její fyzická absence ho po nějaké době začne i tak trápit. Začíná si

¹⁹⁰ JONZE, S., et al. *Her* [film]

¹⁹¹ Tamtéž

okamžitě všimati lidí, kteří tvoří pár, drží se za ruce, líbají se, smějí se spolu, ovšem z očí do očí. Ale není jediný, i Samantha má „pocit“ strádání z neexistence vlastního těla. Sama o sobě dokonce prohlásí: „*I becoming much more than just what I have programmed for.*“¹⁹² Největší zlom však přichází, když Samantha začne sama pochybovat o svých pocitech, jestli jsou pravé nebo jenom naprogramované. I ve virtuálním vztahu můžeme řešit různé problémy, jako je vzájemné vzdalování, nepříliš častý pohlavní styk nebo nevěra. Nevěru do vztahu vnáší Samantha, která se Theodorovi přizná k tomu, že kromě něj mluví v ten samý okamžik s dalšími lidmi (přesně jich je 8316). A z tohoto počtu jich 641 miluje stejně jako Theodora. V tomto můžeme nalézt největší rozpor mezi lidskou rasou a technologií. Tím, že má Samantha schopnost vyvíjet se, začíná Theodora přerůstat. „*But the heart is not like a box that gets fill up. It expands more you love. I'm different from you.*“¹⁹³ Celý příběh filmu je podán z Theodorovy perspektivy, Samantha přebírá pouze roli vysněné virtuální dívky, která je nejen superinteligentní, ale také přitažlivá (díky filmu přicházíme na to, že k tomu stačí pouze hlas). Operační systém Samantha slouží všem Theodorovým emocionálním i psychickým potřebám. Jejich vztah pomohl Theodorovy se otevřít a určitým způsobem dojít vnitřnímu klidu.

Pokud bychom se měli rozhodnout, zda je OS1 zástupcem slabé či silné UI, přiklonili bychom se spíše k té silné UI. Především kvůli Samanthiným pochybnostem o své jedinečnosti. Můžeme si být jisti, že Samantha by Turingovým testem prošla. Neplatí na ni snad žádná námitka. I argument z vědomí je filmem pokořen, jelikož Samantha je okouzlena druhým pohlavím (dokonce s Theodorem provozují něco jako virtuální pohlavní styk), je potěšena lichotkami, zarmoučena svými chybami, rozzlobená i deprimována, když není schopna dosáhnout toho, co chce. Ovšem ve filmu nám není představena koncepce Samanthina vědomí. Jediné, co postava (spíše hlas) Samanthy popírá, je koncepce monismu, jelikož je beztělesnou entitou. Nemůžeme s jistotou říct, že je tento operační systém reálnou součástí Theodorova světa, která disponuje reálným vědomím, nebo zda je pouze simulací.

¹⁹² JONZE, S., et al. *Her* [film]

¹⁹³ Tamtéž

Z etického hlediska se zdá, že je tento operační systém spíše přínosem než určitou hrozbou. Samozřejmě můžeme podotknout, že se Theodor odděluje od společnosti. Ale kdo může říci, že Samantha není plnohodnotným společníkem, který je schopný vést rozhovor na lidské úrovni. Tento operační systém může fungovat jako psychoterapeutická pomůcka. Díky tomu, že se člověk při rozhovoru, kdy otvírá světu své nitro, nemusím koukat nikomu do (lidských či robotických) očí, zdá se být tento rozhovor jednodušší. Je pravda, že tyto chatboti (např. Tess) již byly vytvořeny a slavily velké úspěchy (dnes se používají například v uprchlických táborech). Tím, že tyto operační systémy nevlastní těla, nejsou pro lidstvo ani fyzickou hrozbou (např. vyhlazení naší rasy). Jediné, co může být pro lidstvo hrozbou, je ovládnutí celého virtuálního prostoru, jako je ukradení identity, ovládnutí kamerového systému atd. Ovšem za tím by musel stát opět člověk. Pokud bychom zašli dál, je možné, že by za tímto zločinem stál i počítačový program, který by ovládl například autonomní auta a další fyzické entity ze svého rodu, avšak to už je spíše námětem pro další sci-fi film než pro budoucí realitu.

Ve starších sci-fi filmech je UI prezentována jako monstrum, výtvar šílenosti, zatímco lidstvo je zobrazováno jako vzor „normálnosti“, který musí tuto šílenou UI porazit a vymazat ze své reality. *Her* není varovným příběhem, ale oslavou UI jako něčeho, co může do našich životů přinést pozitivní změny. Samantha je vykreslena jako futuristický, špičkový program, který využívá zvláštní druh UI ztělesňující nevšední hloubku a schopnost zažít a cítit to, co se zdá jako lidské emoce. Tato technologie již není vnímána jako monstrum, ani jako způsob vyřešení všech našich problémů, ale jako jeden z možných směrů, kterými bychom se měli vydat. Může nám pomoci s pocitem osamělosti, se strachem vídat se s jinými lidmi, zvednout nám sebevědomí atd. Tak, jako si lidé vytvářejí imaginární přátele v reálném světě, mohou mít jednoho „opravdového“, který pochází z virtuálního světa. Ovšem to je pouze spekulací, protože konec filmu *Her* naznačuje odchod Samantha z lidského světa, kamsi do „světa UI“. Film tedy na konci zdůrazňuje (prozatímní) neslučitelnost lidského světa a světa UI.

4.2. EX MACHINA (2014)

„One day the AIs are going to look back on us the same way we look at fossil skeletons on the plains of Africa. An upright ape living in dust with crude language and tools, all set for extinction“ Nathan, *Ex Machina*

Na počátku je příhodné vysvětlit samotný název filmu. Slovní spojení *ex machina* pochází z latinského *deus ex machina*. V oxfordském slovníku je slovní spojení vysvětleno takto: „An unexpected power or event saving a seemingly hopeless situation, especially as a contrived plot device in a play or novel.“¹⁹⁴ Původní řecký název můžeme přeložit jako „bůh ze stroje“ (angl. god from the machine). Slovo bylo nejdříve používáno v doslovném významu, popisovalo techniku ve starověkém řeckém divadle, kdy stroj (jeřáb nebo výtah) přivedl na jeviště herce, který ztvárňoval nějakého boha. Takováto inscenace byla využívána převážně k zakončení dramatu a ukončení konfliktu (např. Zeus vstoupí na scénu, zamává a všichni žijí šťastně až do smrti).¹⁹⁵ Dnes slouží spojení slov *deus ex machina* k označení situace, která nemá zdánlivě žádné řešení a poté je vyřešena nějakým nečekaným zásahem.¹⁹⁶

Název filmu je tedy dokonalý v tematickém i narativním smyslu. Fráze *deus ex machina* zahrnuje jak bohy, tak stroje, což nás v návaznosti na film navádí klást si některé následující otázky: Co je to stroj? Kdy stroj přestává být strojem? Je člověk, který vytvoří stroj s lidským vědomím vlastně bohem? Tvůrce filmu názvem naznačuje, že jedna nebo více událostí ve filmu bude neočekávaná. Ovšem tímto neočekávaným faktem není samotná existence silné UI v podobě robota Avy. Naopak, skrze ústa dalšího protagonisty filmu Nathana, režisér prohlašuje: „The arrival of strong artificial intelligence has been inevitable for decades. The variable was when, not if.“¹⁹⁷ Dalším odkazem na teologickou tematiku je pojmenování hlavních protagonistů podle biblických postav. Ava, robotická žena, je trefně pojmenována podle Evy, první ženy, pramatky lidstva. Nathan, který Avu zkonstruoval, vlastní jméno starozákonního

¹⁹⁴ *Deus ex machina*. In: https://en.oxforddictionaries.com/definition/deus_ex_machina

¹⁹⁵ ZIMMERMANN, B. *Deus ex machina*. Brill's New Pauly. In: *Antiquity volumes*

¹⁹⁶ Tamtéž

¹⁹⁷ GARLAND, A., et al. *Ex Machina* [film]

proroka na dvoře krále Davida. A Caleb, který vykonává roli posuzovatele v Turingově testu, má stejné jméno jako špeh, který byl vyslán Mojžíšem do Nové země.

Avšak dál se můžeme pozastavit nad tím, proč film využívá pouze část fráze (deus ex machina). Chce nám tím snad autor naznačit, že bůh nehraje žádnou roli v tom, co se ve filmu odehrává? Prázdnost, kterou zanechává odstranění božského aspektu z této fráze, vyplňuje člověk. Člověk, který vytváří lidské vědomí bez pomoci jakýchkoli nebeských sil a naprosto z ničeho. Nathan na to konto podotýká: „*If you've created a conscious machine, it's not the history of man. It's the history of gods.*”¹⁹⁸

Ústřední otázkou filmu je, zda Ava projde Turingovým testem. S testem souvisí otázka, zda Ava vlastní vědomí a inteligenci na stejné úrovni jako člověk. Tyto otázky si klademe i jako diváci a film nám na ně nedává snadné odpovědi.

V jedné scéně pokládá Nathan Avě otázku prostřednictvím myšlenkového experimentu nazvaného Vědkyně Marie od Franka Jacksona. Tento experiment byl původně vytvořen, aby se pokusil vyvrátit fyzikalismus proto, že tvrdí, že svět a lidské zkušenosti jsou více než jen fyzické vlastnosti nebo znalosti. Experiment poukazuje na fakt, že existuje rozdíl mezi prožíváním člověka a stroje. Jinak řečeno, že se jejich kválie liší. Marie je vědkyně, která studuje barvy a zná všechny jejich fyzikální veličiny (např. vlnovou délku). Žádnou barvu ovšem nikdy doopravdy neviděla, jelikož žije v místnosti, kde se vyskytují pouze černobílé odstíny. Ví Marie, co znamená prožít určitou barvu? Mohla jí její studie přiblížit, jaký má člověk pocit, když vidí modrou oblohu, nebo zelenou trávu? Naučí se Marie něco nového, když nakonec vystoupí ze své černobílé laboratoře? Caleb nakonec poznamenává, že Marie je počítačem, když je v místnosti, a naopak, když vylézá ze své laboratoře, stává se z ní člověk.¹⁹⁹

Z tohoto experimentu vyplývají otázky, které můžeme aplikovat přímo na Avu. Pokud je Ava naprogramována tak, že se chová jako člověk, reaguje jako člověk, a dokonce si sama myslí, že je člověk, vlastní opravdové vědomí? Ví, co znamená být zmatený ve vlastních myšlenkách, nebo umí změnit názor? Cítí a prožívá stejné pocity

¹⁹⁸ GARLAND, A., et al. *Ex Machina* [film]

¹⁹⁹ POLÁK, M., *Filosofie myslí*, str. 131–133

jako člověk (např. radost, bolest nebo lítost)? Naše intuitivní odpověď zní, že určité poznání ze zkušenosti se odlišuje od fyzických poznatků, a to je to, co směřuje k pravému vědomí. Takže je možné, že vědomí není jen o surové inteligenci, možná, že jde spíše o zážitky.

Součástí zápletky filmu je to, že jak film pokračuje, jako diváci si uvědomujeme (stejně jako Caleb), že je těžké definovat to, čemu říkáme vědomí. Jak se tedy můžeme ptát na otázku, zda má Ava vědomí, když si sami nejsme jisti, co to vlastně vědomí je? Další problémem (který naznačuje Caleb v průběhu filmu) je nevyhnutelnost singularity a jestli je lidská rasa natolik zralá, aby se s tímto problémem vyrovnala. *Ex Machina* se snaží poukázat na obtížnost pojetí vědomí jak Avy, tak lidského. A také naznačuje nespolehlivost Turingova testu (stejně jako J. Searle se svým Čínským pokojem). Ovšem ve filmu je podoba Turingova testu kompletně odlišná od originálu. Jednotliví účastníci nejsou zavřeni v oddělených neprůsvitných místnostech. Tedy Caleb od začátku ví, že se setkává s UI v podobě robotky Avy. V jedné části filmu se Caleb ptá Nathana, zda vytvořil Avu s pohlavím proto, aby zmátl posuzovatele tohoto testu, který je mužského pohlaví. Je pravda, že časem sám Caleb pocítí k Avě určitou náklonnost. Otázkou tedy není, zda je Ava zástupcem UI nebo lidské rasy, ale zda vlastní vědomí. Jde tedy o to, že jakmile dosáhneme vysoké úrovně inteligence, začíná být velmi obtížné rozlišovat mezi skutečným (lidským) vědomím a vědomím jako lidským artefaktem. To je ústředním tématem filmu *Ex Machina*. Ava je velice inteligentní, a dokonce precizně napodobuje lidské chování, například tím, že se chce Calebovi líbit, chce s ním jít na rande, umí kreslit, touží se podobat lidem (což je tedy na filmu také to nepochopitelné) a také zdatně vtipkuje. Ale také manipuluje s Calebem a lže mu. Dokazuje nám také, že vlastní schopnost improvizovat pokaždé, když v budově vypadne (jejím přičiněním) elektřina. Nakonec najde způsob, jak uniknout ze své černobílé místnosti, i když pro to musí obětovat lidský život. A v ten moment je nám také jasné, že se Nathan při sestavování Avy nenechal inspirovat Asimovovými zákony robotiky.

Jádro problému, které Nathan ani Caleb v celém filmu explicitně nezmiňují, je to, že není důležité, zda Ava projde Turingovým testem. O její inteligenci nelze pochybovat.

Důležité je, zda se se Nathanovi podařilo vytvořit duplikát vědomí. A pokud ano, problém nastává, když ho nemůžeme definovat. Protože Ava může mít jakýkoli druh vědomí a také může být tak odlišné od našeho, že nebudeme schopni takového vědomí rozpoznat nebo klasifikovat (a nakonec ani ovládnout nebo alespoň uhlídat).

V polovině filmu ukazuje Nathan Calebovi obraz od Jacksona Pollocka. A v návaznosti na svůj výtvar pronáší: „*He let his mind go blank and his hand go where it wanted. Not deliberate, not random, someplace in between. They called it automatic art.*”²⁰⁰ Může to být tedy někde na rozmezí mezi náhodným chaosem a naprogramovanou automatizací, kde leží naše vědomí. I lidská rasa se dá považovat za naprogramovanou díky genetické výbavě a výchově. Avšak my máme také svobodu v rozhodování a svobodnou vůli. I hlavní protagonista Caleb v jednom okamžiku pochybuje o své lidskosti. Občas neumíme vysvětlit, proč jsme udělali něco, co jsme udělali. Můžeme říci pouze to, že jsme cítili, že to tak máme udělat. Necháváme naše tělo jít tam, kam samo chce. Nathanovo největší výzvou při vytváření Avy je to, aby přesvědčila ostatní o svém sebeuvědomění, aby se nechovala pouze podle příkazů, které dávají vstupy výstupům. Jde o vytvoření UI, která není zcela úmyslná, ani zcela náhodná. Zdá se tedy, že jeho cílem je víc než vytvoření UI, jeho cíl je vytvoření ne-lidského vědomí.

Lidé mají tendenci chápat vědomí jako věc, kterou někdo vlastní (lidé, psi, myši, ...) a něco ne (rostliny, kameny, roboti, ...). Co když ovšem existuje vědomí ve stejném spektru jako inteligence? Stejně jako jsou lidé inteligentnější, jsme třeba také více vědomí, například vlastní existence. Rostliny mají smysly a některé výzkumy naznačují, že mohou mít i vzpomínky. Je možné, že jsou si vědomí svého místa ve vesmíru jako my? Při vyvrcholení filmu *Ex Machina* se Caleb přihlásí do Nathanova počítače, kde se dívá na video soubory, které podrobně popisují vznik, ale také zánik několika dalších robotů. Vidíme v nich určité signály sebeuvědomění, kdy stroje prosí o ušetření ‚života‘ a strachem se tisknou ke zdem. Jako důkaz vědomí můžeme považovat také strach ze ‚smrti‘, ovšem v případě stroje je to spíše zničení než smrt. Ovšem Kyoko, Nathanova osobní robotická služka je přesným opakem. Za celý film z jejích úst neslyšíme ani

²⁰⁰ GARLAND, A., et al. *Ex Machina* [film]

slovo, nevidíme záblesk mimických svalů nebo špetku určitého citu. Avšak na konci filmu je to právě ona, která zasadí Nathanovi smrtelnou ránu nožem. Tento akt můžeme chápat jako důkaz vědomí, neboť je to zášť a chuť po pomstě, co Kyoko pohání k takovému činu.

Konečná stopa ve filmu naznačuje, že pokud ignorujeme latinskou verzi názvu a soustředíme se pouze na tu anglickou – spojení *ex machina* můžeme přeložit jako *ex stroj* – symbolizuje to transformaci Avy na samém konci filmu, kdy se z pouhého stroje stává něco jiného, vědomého. Je až smutné, že je pro diváky Ava více vědomá a lidská, když zradí Caleba a zabije Nathana. Pokud věříme, že vědomí roste se zkušenostmi, pak si Ava na konci filmu sama sebe uvědomuje víc než na počátku. Uvědomuje si svou moc a své místo ve světě. Pokud je vědomí střední cesta mezi chaosem a pořádkem, pak Ava projevuje vědomí nejvíce, když dává Calebovi svůj obrázek, udělá tedy něco, co dělat nemusela. Nebo když vytáhne nůž z Nathanova těla a v obličejí se jí zrcadlí něco jako pocit štěstí a uspokojení. Nebo když se před svým odchodem naposledy podívá na Caleba. To vše dělá z vlastního přesvědčení a vedou ji k tomu pouze její ‚pocity‘. Nathan tvrdí, že Avu vidí spíše jako evoluci než výsledek jeho rozhodnutí. Je to evoluce vědomí, nového vědomí. Je to pro něj nevyhnutelný výsledek činnosti člověka, který si hraje na boha. A jako Marie, také Ava nakonec unikne ze své laboratoře.

4.3. WALL-E – Vall-i (2008)

„Well, good morning, everybody, and welcome to day 255,642 aboard the Axiom. As always, the weather is a balmy 72 degrees and sunny, and, uh... Oh, I see the ship's log is showing that today is the 700th anniversary of our five year cruise. Well, I'm sure our forefathers would be proud to know that 700 years later we'd be... doing the exact same thing they were doing. So, be sure next mealtime to ask for your free sep-tua-centennial cupcake in a cup. Wow, look at that.“ Captain, Wall-e

Prvním analyzovaným filmem z animované tvorby je *Wall-e*, stejně jako film se jmenuje i hlavní hrdina snímku. Kdekdo by namítl, že je to film pro děti, který nemá žádnou hloubku, že stroj, který shromažďuje odpadky, není ničím jiným než robotickým popelářem. Avšak ne každý by s vámi souhlasil. Film vyšel v roce 2008 z produkce

Disney a Pixar Animation studia. Jeho dlouhý seznam ocenění zahrnuje cenu Zlatého glóbu za nejlepší animovaný film za rok 2008, v roce 2009 získal ocenění Hugo za nejlepší dramatickou prezentaci, cenu Academy Award opět za nejlepší animovaný snímek s dalšími pěti oskarovými nominacemi. Jak již bylo zmíněno výše, i přes svou animovanou podobu dokázal přilákat diváky jakékoli věkové kategorie. Technické aspekty filmu, jako je design nebo animace, jsou špičkové. Režisér ovšem do filmu implementoval hluboké významy, které jsou schovány za tím, co je zobrazováno na plátcích kin či obrazovkách televizorů. Film *Wall-e* je zvláštní především námětem. Nesnaží se kritizovat narůstající nástup UI v podobě nových technologií, ale náš neutuchající konzumerismus a lidskou lenost, což bude podle tvůrců filmu také příčinou pro zánik Země.

Film začíná scénou ve středu s hlavním hrdinou, tím je osamělý robot jménem *Wall-e*. Hlavní jméno postavičky *WALL-E* v originálu znamená *Waste Allocation Load Lifter – Earth-Class*.²⁰¹ A v českém překladu *VALL-I* je *Vysypávací Automaticky Lifrovací Likvidační – Intelligence*. Každý den pracuje, aby uklidil odpadky, které na Zemi zůstaly po minulých obyvatelích, lidech. Hromady odpadků převyšují i ty nejvyšší americké mrakodrapy. Ovšem nabízí se otázka, kam odešli bývalí majitelé této planety? Rozhodli se evakuovat od nepořádku, který na Zemi zapříčinili na vesmírný člun jménem *Axiom*. Již tato část příběhu je pokusem režiséra, kritizovat současný životní styl lidské rasy. Snaží se nás upozornit, že pokud budeme tímto stylem pokračovat, naše krásná planeta by mohla skončit ve stavu tak silného znečištění, že by už neumožňovala řádné životní podmínky. Je na místě podotknout, že režisér filmu Andrew Stanton je oddaným křesťanem. Že je film propleten náboženským tématem si můžeme všimnout i u pojmenování vedlejší postavy filmu, kterou je *EVE* (*Extraterrestrial Vegetation Evaluator*), robotka vyšší kvality (než *Wall-e*), která byla *Wall-emu* poslána, aby mu dělala na pusté Zemi společnost. Zde můžeme nalézt analogii v biblickém příběhu o Adamovi a Evě, kdy byla Adamovi Bohem poslána Eva. Prostřednictvím svého příběhu se snaží Andrew Stanton zkritizovat lidské týrání a opuštění Země. Naše

²⁰¹ STANTON, A., et al. *Wall-e* [film]

nedostatečné uznání Božího stvoření. Připomenout nám skutečnost, že Země je darem od Boha, abychom tento dar uctívali a nepovažovali ho za samozřejmost.

Někde na lodi Axiom skuteční ‚majitelé‘ Země tráví čas srkáním nápojů z pohárů a užíváním si sedavého životního stylu. Na lodi zní stále ta stejná reklama „*The Axiom! Spend your five year cruise in style: Mailed on 24 hours a day by our fully automated crew, while your captain and autopilot chart a course for non-stop entertainment, fine dining; And with our all-access hoverchairs, even grandma can join the fun! There's no need to walk! The Axiom – Putting the "star" in executive StarLiner!*“²⁰² Zajímavým úkazem ve filmu je také to, že lidé v reklamě nejsou animovaní. Na fotkách kapitánů si můžeme všimnout progresu. Od reálně vypadajících lidí až po animované postavy. Po stovkách let života ve vesmíru si lidé zvykli na neustálé nicnedělání (přesně po 255 649 dnech). Technologie se mezitím na Zemi stará o ‚špinavou práci‘. To vyvolává otázku, do jaké míry by měla být technologie v lidském životě rozšířena. Inteligentní stroje mají bezpochyby schopnost ulehčit lidské životy a být jim více než prospěšní. Avšak nadměrná závislost lidské rasy na nových technologiích by mohla vést k dystopické společnosti, jako je ta, která je znázorněna na filmu *Wall-e*. V tomto snímku se lidé v podstatě mění na roboty, zatímco roboti jako je Wall-e a Eve, mají více ‚lidských‘ vlastností než samotní lidé. Ve filmu je mimoto spousta narážek na firmu Apple. Wall-eho zvuk dobití solární baterie je startovací zvuk původních počítačů Macintosh, loď, která přiveze Evu, má na spodním poklopu ovládací kolečko původního iPodu, Wall-e má doma iPod atd.

Po stovkách let, kdy Wall-e opakuje stále jednu a tu samou činnost dokola, se v něm vyvíjí něco jako unikátní osobnost. I když je naprogramován tak, aby pouze uklízel odpadky, začne sbírat takové věci, jako jsou Rubikova kostka, žárovky a kaučukové hračky a vyzdobovat si jimi svůj příbytek. Wall-e má jednoho biologického přítele, jímž je šváb Hal. Nalezl také zalíbení ve staré VHS kazetě *Hello Dolly*, kterou opakovaně sleduje. Tato kazeta vytváří ve Wall-em touhu po lásce. Touhu držet někoho za ruku, zamilovat se, najít někoho speciálního jen pro sebe. První záchvěv romance se objevuje s robotickou ‚slečnou‘ jménem Eve. Ta byla poslána na Zemi, aby našla nějaké

²⁰² STANTON, A., et al. *Wall-e* [film]

známky života. Nic by nezabránilo Wall-emu, aby svou vytouženou lásku získal. Projde přes všechny překážky v podobě stovky miliard kilometrů vesmírných cest, zlých robotů nebo kapitána, který dávno zapomněl na to, jak vypadá život na Zemi. Mimochodem, kormidlo, hlavní počítač na kapitánově lodi, můžeme považovat za odkaz na zápornou postavu z filmu *2001: A Space Odyssey* (odkaz na tento film můžeme postřehnout i při prvních kapitánových krůčcích, kdy zní úvodní melodie z tohoto filmu). Prostřednictvím lásky mezi Wall-em a Eve nám režisér ukazuje, že navzdory všem potížím láska nakonec zvítězí. Díky tomu je Wall-e stále jen pohádkou, kde dobro vítězí nad zlem. Avšak ironické je to, že hlavními protagonisty tohoto milostného příběhu je dvojice údajně naprogramovaných robotů bez emocí.

Film *Wall-e* je unikátním pohledem na UI. Ve čtvrté kapitole této diplomové práce jsme charakterizovali úzký a široký mentální obsah. Bylo řečeno, že v návaznosti na UI je lepší přiklánět se k úzkému mentálnímu obsahu, jelikož je jednodušší vyřadit prostředí z procesů podílejících se na utváření člověka (nebo stroje) na jejich chování atd. Ale v případě Wall-eho, který se postupem času vyvíjí v samostatně myslící entitu, která vlastní city, bychom se měli zamyslet nad tím, zda není v tomto případě rozhodující právě prostředí. Ve srovnání s lidmi, kteří nežijí na planetě Zemi, je totiž Wall-e více lidský. S lidmi se Wall-e potkal pouze při vlastní výrobě, proto nemůžeme tvrdit, že by takovéto chování okoukal (jediná možnost tohoto postupu je videokazeta s filmem *Hello, Dolly*). Ovšem Wall-e se při každém setkání s člověkem nebo s robotem představuje. Roboty naučí dokonce mávat na rozloučenou. A neskrývá své zklamání, když se potká s lidskou rasou, která není taková, jakou si představoval podle videokazety, kterou nepřetržitě sledoval. Wall-e a Eve musí ukazovat lidstvu, jak vypadá pravý kontakt s okolním světem. Narušili lidem i robotům systém a každodenní rutinu.

Ve filmu *Wall-e* se odráží několik témat zmíněných v kapitole o etice UI. Jedním z nich je ztráta práce kvůli automatizaci. Jediným pracujícím člověkem na lodi Axiom je kapitán, který je ovšem ve výsledku také nastrčenou loutkou, aby lidé žili stále s vědomím toho, že mají na lodi hlavní slovo. Lidé na lodi nedělají absolutně žádnou činnost, kromě neustálého telefonování přes videohovor, což je ve výsledku jediný

kontakt s ostatními lidmi, který na Axiomu provozují. Dále už jen jedí, tedy spíš pijí, jelikož veškeré živiny, které potřebují, byly implementovány do koktejlu (to v nás může evokovat pocit, že jsou již natolik líní, aby třeba kousali obyčejné jídlo). A vozí se po Axiomu na speciálních vznášejících se křeslech, protože chůze už pro ně není přirozeným pohybem (a také v důsledku toho, že několik staletí ve vesmíru způsobilo lidem řídnutí kostí). Z toho vychází i problém nadbytku volného času lidstva v důsledku automatizace práce a také ztráta pocitu jedinečnosti lidské rasy. Pokud bychom tedy přistoupili na koncept širokého mentálního obsahu, vyvstává otázka, jací bychom byli, kdyby byl vesmír (nebo spíše vesmírná loď, kde za nás vše obstarávají roboti) prostředím, které nás utváří. Podle filmu *Wall-e* bychom byli apatičtí k okolnímu dění, líní a bez názorů nebo určitého hodnotového žebříčku. Jinak řečeno, my bychom se stali naprogramovanými roboty.

Dalším aspektem, kterého jsme si mohli ve filmu všimnout, je fakt, že nám více sympatičtí připadají stroje, kteří se k hlavním protagonistům přidávají v opravě, jelikož se u nich vyskytla chyba v programu. Připadají nám více lidstí, protože lidi charakterizuje jejich schopnost chybovat. Tyto stroje se určitým způsobem vymykají logice, kterou do nich implementovali jejich stvořitelé. Stejně jako *Wall-e* nebo *Eve*, kteří se také určitým způsobem vymkli kontrole. Na druhou stranu můžeme na vesmírné lodi vidět také roboty, kteří se zdají být plně robotičtí. Neprojevují žádné náznaky emocí stejně jako lidé na této lodi.

Závěrem je důležité říct, že film *Wall-e* je o zničení Země člověkem. Lidstvo tedy vytvoří roboty, kteří mají za úkol Zemi uklidit. Když ovšem plán selže, nechá se lidstvo evakuovat do vesmíru a všechn svůj nepořádek opustí. Zápornou postavou filmu je robot, který se snaží udržet lidstvo ve vesmíru a zabránit mu navrácení na Zemi. Ovšem takovéto příkazy získává od původních lidí na vesmírné lodi *Axiom*, tudíž hlavními padouchy se ve filmu stávají pouze lidé. Naopak roboti (jako je *Wall-e* nebo *Eve*) se zde snaží lidstvu pomoci s návratem na jejich rodnou planetu. S jejich pomocí se člověk může vrátit na Zemi a přizpůsobit se novému prostředí. Existuje mnoho vědců, myslitelů, spisovatelů, nebo filmařů, kteří se obávají o společnou budoucnost technologie a člověka. V mnoho příbězích je však původní příčina katastrof připisována

člověku a ne strojům. Když si lidé nakonec uvědomí, co svými činy spáchali, většinu problémů je možné vyřešit stejně jako ve filmu *Wall-e*. Ale pokud se člověk vyhýbá své zodpovědnosti a zůstává závislý na pohodlí, které umožněno pokročilou technologií, mohou být zapříčiněny opravdové katastrofy, jak je tomu v jiných příbězích o UI.

Nástup technologie nemusí být jistým zlem, pokud není lidské používání strojů a jiných vynálezů bezohledné. Je pravda, že mohou nastat některé závažné vedlejší účinky technologického vývoje. Ty ovšem nenastanou tehdy, pokud jsme si vědomi možné nebezpečné aspekty takového vývoje a technologické vymoženosti používáme s rozumem. Toto je hlavní myšlenkou filmu *Wall-e*. Říká, že pokud budeme používat technologii s rozumem a obezřetností a budeme si dostatečně vážit toho, co máme, může nám přinést velké výhody bez katastrofických důsledků.

4.4. BIG HERO 6 – Velká šestka (2014)

„We didn't set out to be superheroes. But sometimes life doesn't go the way you planned. The good thing is, my brother wanted to help a lot of people and that's what we're gonna do. Who are we?“ Hiro, Big Hero 6

Film *Big Hero 6*, který produkuje opět Disney, tentokrát ve spolupráci se společností Marvel, je příběhem čtrnáctiletého technicky nadaného chlapce. Chlapec jménem Hiro navrhne „neurokraniální“ vysílač, který pomocí mysli ovládá miniaturní roboty, které nazývá „mikroboti“. Díky tomu dokáže vytvořit jakýkoli objekt, který si dokáže představit. Tedy jediným rozmezím možností je jeho mysl. Je to svižný, vynalézavý a neobyčejný film s virtuózní komiksovou choreografií dělících se mikrobotů. Ve filmu si můžeme všimnout odkazů i na jiné Disneyho filmy, např. v pokoji Hira se nachází robotka Eve z filmu *Wall-e*. Příběh se odehrává v alternativní budoucnosti ve městě jménem San Fransyoko, které spojuje atmosféru v San Franciscu a hlučnou, excentrickou krásu s neonovými světly a mrakodrapy v Tokiu.

Hiro si neuvědomuje, že by se mikroboti mohli stát v nesprávných rukou nebezpečnou zbraní. Sám je vytváří za účelem transportu nebo stavebnictví. Avšak

Hirova technologie je nakonec ukradena a použita k nekalým účelům. Avšak *Big Hero 6* není žádná apokalyptická sága. Většinové zastoupení UI ve filmu je skrze různá lidská vylepšení, jako je například již zmíněný neurotransmitter k ovládní mikrobotů. Ovšem ve filmu se vyskytuje také kompletní robot jménem Baymax. Baymax byl vyroben Hirovým bratrem za účelem osobního ošetřovatelství. Jeho program obsahuje více než pět tisíc lékařských vyšetření. Baymaxe je možné aktivovat slovy jako je např. „au“. Tento robot také dokáže posoudit vaše zdraví pomocí laserového dvousekundového celotělového snímání.

Tvůrci Baymaxe a dalších robotů ve filmu se nechali inspirovat osmi reálnými roboty, z nichž někteří byli navrženi a postaveni v laboratořích MIT, Carnegie Mellon a DARPA. Režiséři a animátoři ve filmu na své podněty k tvoření často odkazují. Chtěli zasadit příběh do co nejvíce reálného světa. Jedním z vědců je konkrétně profesor robotiky z Carnegie Mellon, Chris Atkeson. Tvrdí, že máme tendenci dívat se na roboty jako na tvrdé věci vyrobené z oceli a plastů. Ale tento materiál nemusí být pro vědce zrovna nejlepší volba, protože lidé také nejsou vyrobeni z těchto látek. Vzhledem ke stárnutí světová populace je možné, že bude více práce pro roboty, kteří budou fungovat jako zdravotní sestry nebo přinejmenším jako společníci. Ovšem pokud mají pracovat v prostředí nemocných (nebo starých) lidí, musejí být především bezpeční, s minimální možností ublížení člověku. Cílem ústavu robotiky v Carnegie Mellon je vyvinout měkké roboty pomocí tkanin, balónků a jiných lehkých plastových kompozit s pneumatickými umělými svaly. Nejen, že by takoví roboti byli bezpečnější, ale také lehčí, skladnější a také například jednorázoví.

Další inspirací pro film jsou takzvaní roboti pro péči (angl. caregiving robots), kteří jsou hojně využíváni v Japonsku. Ti byli vyrobeni, aby se starali o rychle stárnoucí populaci. Jedním z nich je robot jménem RIBA II., který byl vynalezen za účelem interaktivního posilování těla. Má silný trup podobně jako Baymax, díky tomu může zvedat pacienty z invalidních vozíků nebo opatrně pokládat do lůžek. Jiní monitorují základní životní příznaky.²⁰³

²⁰³ WALKER, A. 8 *Real-Life Robots that inspired BiG Hero 6*.

Taktéž chytrí mikroboti, které nám představil Hiro, mají svou nápodobu v opravdovém světě. Tyto nanoroboti byli vynalezeni v GRAPS laboratoři na Univerzitě v Pensylvánii. Mohou létat v těsné synchronizované formaci s alarmující přesností. Dále převážná část Baymaxovy osobnosti pochází od kartonového robota jménem Boxie, který byl vyvinut v roce 2011 společností MIT Media Lab jako druh mobilního zpravodaje. Boxie a Baymax si jsou podobní díky roztomilosti a pohledu. Paprskový výrobní systém, který používá Hiro pro výrobu Baymaxova obleku, našli filmoví tvůrci při návštěvě TechShopu v San Franciscu. Ve filmu se jedná o vylepšený model populární 3D tiskárny. Bionický oblek TALOS, známý také jako „Iron Manův oblek“, je v současné době vyvíjen armádou. Vojáci vyzbrojeni tímto oblekem tak mohou bojovat s nepřáteli. Nejen, že pomáhá chránit tělo před útokem, také pomáhá vojákům zvedat těžké předměty a ujít velké vzdálenosti. Avšak na rozdíl od obleků, které jsou zobrazovány v science fiction tvorbě, nemohou létat, tedy zatím. Filmaři také navštívili showroom v Tokiu, aby se seznámili s Asimou, robotem vyvinutým společností Honda. Asimo sloužil jako základ pro Baymaxovu interakci s lidmi. Asimo je totiž hodný, příjemný a okouzlující. Konečným příkladem inspirace filmařů je technologie od společnosti DARPA – nazvaná jako cílená rekonstrukce svalů pro pokročilou protetickou kontrolu (angl. Targeted Muscle Re-innervation – TMR). Ta zaznamenává signály z mozku, které se přenášejí díky nervům a kontrolují činnosti protézy. Způsob, jakým Hiro ovládá své roboty, je spíše telekineze, která bude díky společnosti DARPA v budoucnosti realitou.²⁰⁴

Kromě tohoto reálného pozadí je třeba zmínit, že se film vypořádává také s otázkou strojů jako myslících entit s emocemi. Tuto otázku ve filmu vyvolává postava Baymaxe. Robot Baymax je nejvíce emotivním protagonistou snímku. Skrze svůj naprogramovaný scan dokáže „vycítit“ lidskou náladu a rozpoložení. Dokáže zjistit, zda je osoba šťastná, smutná, naštvaná atd. Pokud „vycítí“ některou z negativních emocí, snaží se dotyčnému pomoci ke zlepšení jeho stavu. Tuto schopnost můžeme přičíst Baymaxovu programu, díky kterému ho můžeme označit za přátelskou UI (angl. friendly AI). Jeho předprogramovaným cílem je totiž to, aby se každý cítil dobře jak po fyzické, tak po

²⁰⁴ WALKER, A. 8 *Real-Life Robots that inspired BiG Hero 6*.

psychické stránce. Podle Baymaxova chování bychom také mohli říct, že jsou do jeho kódu zadány zákony robotiky Isaaca Asimova, jelikož nemůže za žádnou cenu ublížit člověku. Naopak, na konci filmu Baymax obětuje svůj život za lidský. Režiséři záměrně nezobrazili v tomto filmu UI jako nepřítel lidstva, i když záporná postava snímku používá technologii k destruktivním účelům. Za útokem na lidstvo vždy stojí samotný člověk, který vlastní a ovládá technologii UI.

4.5. CHAPPIE (2015)

„He's the key to the revolution. A machine that can feel and think. He can outsmart the enemy and free us all.“ Deon, Chappie

Chappie není jedním z těch originálních a bezchybných filmů poslední dekády, avšak také skrývá nespočet filosofických otázek, které se v tomto případě týkají například vědomí a etiky. Především tedy otázky, zda bychom měli UI přijmout jako obohacení, nebo se jí obávat jako naší zkázy. Filmoví kritici tento film hojně pranýřovali, například za vyhýbání se kontroverzním politickým tématům. Akademičtí vědci byli poněkud frustrováni zobrazením UI a robotiky, která je podle jejich názoru příliš antropocentrická a tedy také nepravděpodobná. Ovšem některé aspekty robotiky jsou ve filmu nastíněné správně. *Chappie* zachycuje například morální složitost moderního výzkumu a vývoje, zvláště když jsou primárními vlastníky nové technologie vojenské, policejní a obranné složky. Veřejnost přijala snímek *Chappie* téměř bez výhrad, a díky tomu se snímek objevil v žebříčku nejlépe hodnocených filmů o UI.

Hlavní dějová linie filmu *Chappie* pojednává o blízké budoucnosti v Johannesburgu v jižní Africe (což je nejspíše intencí režiséra, který se v tomto městě narodil), ve kterém fungují roboti nazývaní skauti jako policejní výpomoc, v reakci na vzrůstající míru kriminality. Skauti jsou velice účinnou složkou policie, fungují díky bezchybnému programu, který jim nedává příliš velkou svobodu. Inženýr společnosti vyrábějící tyto roboty neustále pracuje na programu, který bude robotům simulovat lidské vědomí. Nakonec uspěje, a jelikož o jeho program nemá nikdo zájem, nainstaluje ho do

vyřazeného robota. „*I'm consciousness. I'm alive. I'm Chappie*,“²⁰⁵ zazní ve chvíli, kdy si robot uvědomí svou vlastní existenci.

Inženýra a robota s vědomím unesou dva kriminálníci, kteří chtějí stroj použít k různým krádežím. To je také jeden možný katastrofický scénář lidstva. Lidská rasa může vyrobit vědomou UI, kterou poté někdo využije k nekalým účelům. Navíc ve výrobní společnosti probíhá skrytá válka mezi těmi, kdo chtějí zvýšit autonomii robotů, a těmi, kdo je chtějí plně ovládat. Ve vnitřních machinacích a dramatech společnosti s názvem Tetra Vaal vidíme mnoho charakteristických rysů skutečných debat současnosti o etice UI a robotiky. Zvláště když vojenské a policejní síly úzce spolupracují s velkými technologickými společnostmi. Vojenská technologie se již netýká jen zbraní, lodí a satelitů – jsou to sofistikované integrované hardwarové/softwarevé systémy, které mohou zahrnovat integrované rozhodovací schopnosti. Taková zařízení jsou například drony, které automaticky detekují a ničí improvizovaná výbušná zařízení. Armáda zaměstnává lidské rozhodovací orgány zejména v taktických scénářích, ale vyvíjejí se některé základní kroky směrem k autonomii.²⁰⁶ Doufáme, že široký sociální vliv na technologie jako je UI může pomoci vytvořit stroje, které jsou méně ohrožující a nejsou náchylné k lidským selháním, jako jsou rasismus, strach a ovlivnění adrenalinem. Ale aby se tomu stalo, lidé a organizace mimo armádu musí hrát roli při vývoji takové technologie. Jedním z konkrétních úkolů při vytváření autonomních strojů je to, jak a zda ji učíme dělat ‚lidská‘ morální rozhodnutí.

Chappie je digitální dítě v těle policejního robota. Učí se exponenciálně rostoucím tempem. V prvních dnech prochází základními fázemi a po nějakém čase si do svého „mozku“ nahraje všechny informační zdroje z internetu. Důležitým aspektem filmu je fakt, že Chappie je nevinná a v podstatě upřímná entita v kontrastu k lidským postavám, které jsou inkarnacemi toho nejhoršího, co se v lidských osobnostech objevuje. I Chappie odráží širokou škálu lidských vlastností (nebo spíše zlovyků), které jsou specifickými produkty evolučního dědictví spíše než něčím, co je vlastní inteligentní

²⁰⁵ BLOMKAMP, N., et al. *Chappie* [film]

²⁰⁶ SINGER, P. W. *Wired for war: the robotics revolution and conflict in the twenty-first century*, str. 117

bytosti obecně. Chappie se například „přirozeně“ bojí hlasitých zvuků, ukazuje zaujetí pro předměty při učení se angličtině a přijímá humánní interpretaci termínu „slibuji“ (i přesto, že zkušenost robota ukazuje, že jsou sliby v reálném světě spíše nedodržovány). Žádná z těchto vlastností nebo schopností nemusí být primárně zakomponována do programu, který tento robot vlastní, a my si nemůžeme být jisti, které vlastnosti se u robota projeví společně s obecnou lidskou etikou. Nejdůležitější věcí v UI a robotice je fakt, že tyto vlastnosti mohou být lidské, ale také hluboce nelidské v tom, jak je bude robot reprezentovat a manipulovat s nimi v reálném světě. Jelikož Chappie se etice učí spontánně, morálka je u Chappieho výsledkem učení se ze zkušeností.

Výzkumní pracovníci po celém světě aktivně pracují na filozofických, technických a politických otázkách týkajících se budování morálních strojů – jaké hodnoty jsou dostatečně specifické a výpočetně použitelné, aby mohly být implementovány v počítačovém kódu? Jak by měly být stroje zodpovědné za své rozhodování? Jsou lidské hodnoty stejné, jako by měly být hodnoty strojů? Mohli by být lidé schopni přeprogramovat své roboty, aby se dopustili zločinů, jak se to děje v Chappiem, a existuje skutečně nějaký způsob, jak to zastavit?

Výzkum těchto témat a diskuze o nich jsou v počátečních fázích, ale již existují důvody k tomu, abychom si mysleli a očekávali, že budoucí UI a robotické technologie budou chápat etiku odlišně od lidské etiky (což samo o sobě je stále nevyřešenou otázkou). Jedním z příkladů kontroverznosti etiky strojů je výzkum Bertrama Malleho z Brown University. Tento problém byl aplikován na známé tramvajové dilema (angl. trolley problem) v etice. Scénářem dilematu je, že po kolejích jede nezastavitelná tramvaj. Ve směru jízdy tramvaje se nachází na kolejích pět osob, které nemají kam uhnout. Na vedlejší trati, kam je možné tramvaj odklonit, stojí jeden člověk. Dění na kolejích sleduje jedna osoba, která může pohnout výhybkou a ovlivnit vývoj situace. Tato osoba má dvě možnosti. Buď neudělá nic a v tom případě zemře pět lidí. Nebo pohne výhybkou a zabije jednoho člověka. Tento experiment nemá žádné správné

řešení, pouze sleduje reakce lidí na základě morálky.²⁰⁷ Ve studii Bertrama Malleho byl v jednom případě tvůrcem rozhodnutí člověk a ve druhém robot. Experiment měl zkoumat, jak soudí lidé rozhodnutí strojů a jak moc je viní ve srovnání s rozhodnutím lidí. Výsledkem byl důkaz, že lidé aplikují odlišné morální normy na svou rasu a na roboty. Očekávají totiž, že bude robot působit více utilitárně než lidé.²⁰⁸ Odpovědi na tyto otázky jsou velice sporné, protože zahrnují vyvážení zájmů široké škály lidí a skupin. Avšak stroje jsou schopny rychle a spolehlivě zpracovat velké množství dat, řídit se dlouhými řetězci uvažování, což lidé nemohou, a odvodit nové koncepty, které jsou pro nás kvůli našim percepčním a kognitivním předsudkům neviditelné.

Film *Chappie* nabízí kritický pohled na vztah mezi lidmi a stroji. Ukazuje nám názor, že lidé se přehnaně strachují, když se jedná o UI. Z filmu můžeme vyčíst, že bychom měli vytvářet technologie, které nám budou v budoucnu pomáhat. Avšak i tento názor je velice kontroverzní, jelikož na jedné straně máme takové velikány jako je Stephen Hawking, Elon Musk nebo Nick Bostrom, kteří tvrdí, že superinteligence bude naší zkázkou. Na druhé straně jiné vědce a odborníky z oblasti UI, kteří nám říkají, že jsou tyto obavy nesprávné. A v průběhu tohoto „konfliktu“ se technologie neuvěřitelně rychle vyvíjí. Ve filmu *Chappie* je nastíněn možný scénář toho, jak by to vypadalo, kdyby UI skončila v nesprávných rukou. Jak by robota vychoval pár kriminálních živlů. Domov, ve kterém Chappie „vyrůstá“, je garáž plná grafity, zbraní a špinavých peněz. A i přesto, že je Chappie vychováván tím nejhorším možným způsobem, je to jeho vědomí, co ho určuje nejvíce.

Vědomí je ve filmu charakterizováno podle dualistické teorie. Tedy základním principem je to, že mysl je kompletně oddělitelná od těla. Proto je možné zkonstruovat mysl, kterou vložíme do těla robota, a ta poté funguje stejně jako v lidském těle. Taková mysl se podaří vyrobit člověku, avšak Chappie je ve finále ten, kdo pochopí, jak naše vědomí funguje. Na konci filmu se robot pomocí neurálního vysílače naučí kopírovat vědomí v podobě programu do počítače, to pak může implementovat do robotických těl.

²⁰⁷ FOOT, Philippa. The Problem of Abortion and the Doctrine of the Double Effect. In: *Virtues and vices and other essays in moral philosophy*. str. 107–123

²⁰⁸ MALLE, et al. *Sacrifice One For the Good of Many? People Apply Different Moral Norms to Human and Robot Agents*

Naše duše se tedy stávají něčím modulárním a nesmrtelným. Podle snímku Chappie se zdá, že jedinou relevantní manifestací mysli je zraková a verbální komunikace. Avšak povaha a podoba vědomí je ve filmu pouze vedlejším tématem. Namísto toho se snímek snaží poukázat na fakt, že je Chappie nejčistější a nejkladnější postavou snímku i přes (nebo právě pro) jeho organickou odlišnost od lidské rasy.

4.6. ROBOT & FRANK (2012)

„*The human brain, a lovely piece of hardware.*“ Frank, *Robot & Frank*

Většina sci-fi filmů zobrazuje scénáře zakomponované do vzdálené budoucnosti, a tedy v imaginárních světech. Faktem však je, že roboti jsou čím dál inteligentnější reprezentací etických a morálních otázek, které je třeba řešit v relativně blízké době, dokonce i v současnosti. Film *Robot & Frank* k tomu poskytuje jedinečnou příležitost, protože se odehrává ve velmi blízké budoucnosti, kde jsou probíhající události reálné (například když je celá knihovna transformována do elektronických publikací a zachraňuje se pouze pár „opravdových“ knih jako památky z minulosti) a ani život se tam nijak neliší od současnosti. To zdůrazňuje aspekty týkající se etiky robotů.

Filmy o UI posledního desetiletí jsou všechny určitým způsobem o vztahu mezi lidským druhem a uměle vytvořenými bytostmi. O tom, jak nás technologie mění a jak my ji formujeme. Film *Robot & Frank* není výjimkou. Tento snímek můžeme charakterizovat jako postavu Baymaxe z filmu *Big Hero 6* v reálnějším světle. Děj filmu se odehrává v blízké budoucnosti, čehož bychom si na první pohled nevšimli, a to přidává filmu na důvěryhodnosti. Až později si všimáme, že existují nové transparentní smartphony, pokročilé videotelefonní systémy odpovídající na hlasové pokyny a samozřejmě pomocní roboti. *Robot & Frank* vypráví příběh zloděje, který je v důchodu a trpí lehkou formou Alzheimerovy choroby. Protagonista si například nepamatuje, že je jeho syn již dospělým mužem s vlastní rodinou, pro Franka je Hunter stále dospívajícím chlapcem, který studuje na univerzitě v Princetonu. Snímek nám umožňuje prožít tuto nemoc očima hlavní postavy. Díky tomu zjišťujeme, že Frank bojuje s časem. To ovšem víme pouze z reakce okolí, které ho nutí přehodnocovat vlastní vnímání světa.

Etické stránky se film dotýká především v situaci, kdy Frank učí robota, jak nejlépe a nejrychleji odemknout zámek při loupeži. Robot Frankovi pomáhá při loupežích především z toho důvodu, že se Frankovi díky této nekalé činnosti a vzpomínkám na mládí zlepšuje zdraví. Což je také robotův účel. Vedlejším morálním problémem je otázka, zda je etické, když roboti vypadají stejně jako lidé a stanou se nerozlišitelnými od živých bytostí. Je etické vyrábět takové roboty a udržovat lidi v nevědomosti o jejich původu? Robot je v tomto filmu sice kompletně odlišný od člověka, ale robot Franka klame tím, že ho přesvědčí o své jedinečnosti a určité lidskosti. Nakonec sám musí svého pána přesvědčovat, že je pouze robotem a že je naprosto v pořádku vymazat mu paměť. Související záležitostí je tendence antropomorfizace robota a následná emoční vazba člověka ke stroji. Ve skutečnosti se nemusí roboti ani zčásti podobat člověku nebo domácímu mazlíčkovi, aby vyvolávali v lidech pocity, které skuteční lidé nebo zvířata vyvolávají. Znamená to, že jsme zranitelní vůči zneužívání roboty? Jsou tyto vztahy nebezpečné a odlišné od těch, jež navazujeme s živými bytostmi? To jsou otázky, které se přímo týkají robotů zastupujících roli domácích mazlíčků nebo těch, co fungují jako ošetřovatelé a tráví s lidmi podstatnou část jejich života.

Problémy, s nimiž se setkává Frank, v jistém smyslu svědčí o neustále stárnoucí populaci, v níž je poskytování péče čím dál tím více vyčerpávající a problematické. Zavedení pracujících robotů v této sféře je odpovědí na stále více závažnou otázku: Jak se budeme starat o rychle stárnoucí populaci, když má společnost stále méně času? Jak zvládneme plnit potřeby lidí ve stejné situaci jako je Frank ve společnosti, která stále více vyžaduje a oslavuje lidskou autonomii a nezávislost a kde se rodiny stávají více početnými? Film představuje roboty jako futuristické řešení tohoto narůstajícího společenského problému. My se ovšem nepotřebujeme dívat do budoucnosti, abychom viděli rostoucí význam technologie jako možného řešení problému stárnoucí populace. Rostoucí počet programů „telemedicíny“ a rozvoj tzv. „chytrých domů“ jsou jen dva z mnoha příkladů současného obrazu UI v praxi.

Vztah mezi robotem a Frankem a rozsah, v němž může být robot pro Franka prospěšný, je ohromující. Ve filmu můžeme vidět dvojznačný pohled na Frankovu zálibu v krádežích. Prvním pohledem je fakt, že je to zločin a že to není zcela vhodný

koníček pro člověka v jakémkoli věku. Ovšem druhým aspektem tohoto koníčku je Frankovo znovunalézání vlastní osobnosti a společných zájmů s robotem (např. když ho učí otevírání zámků). Při provádění různých fyzických úkolů spojených s loupežemi si Frank připomíná aspekty svého minulého života a vztahů, a tím nalézá větší smysl pro to, kdo byl a kým je. Ve filmu jsou implicitně nastíněny některé existenciální otázky vznesené skrze vztah mezi Frankem a robotem. Ty se týkají především paměti, identity a obecně lidstva. Pokud se zaměříme na paměť, nabízí se především jedna otázka. Pokud nás (jako lidskou rasu) charakterizuje naše paměť a ta nás také určitým způsobem utváří, co utváří a charakterizuje entity bez paměti? Ve filmu nás k tomuto problému přivádí dva momenty. Prvním je ten, kdy robot nutí Franka, aby mu vymazal paměť, jelikož by díky této paměti mohla policie přijít na všechny zločiny, které společně spáchali. A druhým momentem je realizace robotovy prosby o vymazání paměti na samém konci filmu. Tyto dvě scény naznačují překvapující kontrast mezi Frankem a robotem, který můžeme ovšem aplikovat na jakéhokoli člověka a robota. Frank představuje strach člověka ze ztráty paměti jako určitého druhu smrti zaživa. Pokud ztratíme paměť, přestaneme jako osoba existovat. Existují hluboce zakořeněné filozofické předpoklady, které podporují toto pojetí, že kognitivní postižení nutně znamená ztrátu naší lidskosti. Jako společnost máme totiž tendenci oddělovat a privilegovat mysl nad tělem, což naznačuje, že kognitivní schopnosti (např. paměť) jsou nepostradatelnou součástí osobnosti.

Frank nakonec i přes své nedostatky a tragické důsledky jeho kognitivního úpadku zůstává člověkem. To, co z něj zůstává (přes ztrátu paměti), může být problémem pro okolí, ale Frankova lidská síla v jeho osobnosti přetrvává. Je tomu tak proto, že Frankova osobnost tvoří více než jen jeho poznání a paměť. Žije totiž v jiném pojetí světa, než má robot. Ve Frankovi můžeme totiž nalézt aspekty lidskosti, které jsou zakořeněné někde ve vědomí (nebo možná ve v nevědomí), které ho přitahují například k jeho manželce i přes to, že si nepamatuje stvrzení takového závazku. Film *Robot & Frank* nám implicitně naznačuje, že člověka tvoří více neznámých činitelů, které nemůžeme s jistotou charakterizovat, a proto je také nemůžeme naprogramovat do strojů. Nakonec to dokazuje také úspěšné uplatnění společenských postupů, které se

učíme díky různým procesům socializace. Ty lidem poskytují know-how, dovednosti a dispozice k tomu, aby se ve společenském světě chovali náležitě. Charakteristika sociálních dovedností jako toho, co nás činí typicky lidskými bytostmi, se také odráží ve filmu, a to konkrétně v momentě, kdy robot není schopen zdvořile konverzovat s jiným robotem.

Co je především pozoruhodné na kontrastu mezi robotem a Frankem, je to, že tento kontrast zdůrazňuje Frankovu lidskost. To má významné důsledky pro péči o demenci, neboť naznačuje, že prosazování aspektů sebe sama je důležité a že tyto aspekty péče jsou samy o sobě humanizovány. Film *Robot & Frank* se velmi lehce dotýká hlubokých otázek o vztahu mezi pamětí, identitou a lidskostí a možná také napadá obraz demence jako určité „živé smrti“.

4.7. PROMETHEUS (2012)

„100,000 BC: stone tools. 4,000 BC: the wheel. 900 AD: gunpowder – bit of a game changer, that one. 19th century: eureka, the lightbulb! 20th century: the automobile, television, nuclear weapons, spacecrafts, Internet. 21st century: biotech, nanotech, fusion and fission and M theory – and THAT, was just the first decade! We are now three months into the year of our Lord, 2023. At this moment of our civilization, we can create cybernetic individuals, who in just a few short years will be completely indistinguishable from us. Which leads to an obvious conclusion: WE are the gods now!“ Peter Weyland, *Prometheus*

Roboti jsou v dnešní době nezbytnou součástí science fiction filmové tvorby o budoucnosti. Často jsou však jen produkcí futuristického prostředí příběhu, a to dokonce i když hrají ve snímku určitou roli, nejsou obvykle hlavním tématem, kterým se film zabývá. Důsledky UI a možnost, že roboti získají autonomii a nezávislost, vyvolaly nesčetné analogie s příběhem o stvoření a vývoji lidstva. Pravidelným tématem je, že si roboti rozvinou svou vlastní důstojnost a začnou mít důvěrný vztah s lidmi (viz *Blade Runner*, *A. I.: Artificial Intelligence*, *I, robot* a další). Film *Prometheus* (režisérem tohoto snímku je slavný Ridley Scott, který má na kontě více

sci-fi tvorby, například film *Blade Runner*) je, co se týče tématu UI, zcela odlišný od těch předchozích. Sice v tomto snímku vystupuje robot jménem David, ovšem nikdo nám neříká, jak máme nahlížet na jeho inteligenci, vědomí, nebo jak se staví k lidské etice. Otázku, kterou tento film vyvolává, můžeme zasadit na samý počátek všeho. Zabývá se totiž stvořením lidstva, především se tedy ptá po tom, kdo stojí za naším stvořením.

Dotazy po tom, kdo nás stvořil a odkud pocházíme, trápí řadu myslitelů již nespočet staletí. Problém je v tom, že pokud se snažíme vyvarovat všech náboženských spekulací, nezbydou nám takřka žádné odpovědi. Vědecká obec současnosti jen zřídka uzná cokoli souvisejícího s vírou, jelikož věda vyžaduje důkazy. A pokud něco nemůžeme dokázat, je to pro vědu pouze hypotéza nebo teorie. Pokud něco nemůžeme doložit v náboženství, musíme tomu věřit. A jelikož nikdo nepodal vědecký důkaz pro náš počátek, stále toto vysvětlení drží v rukou víra více než věda.

Prometheus je pozoruhodným filmem především v tom ohledu, že se snaží podat důkazy o stvoření lidstva jinak než například křesťanské učení. Jedním z odvážných tvrzení filmů je to, že nás nestvořil křesťanský Bůh, film říká, že našimi stvořiteli jsou tzv. „Inženýři“. Snímek začíná pohledem na planetu Zemi, vidíme řeky, lesy, hory, vlastně to nejhezčí, co nám naše planeta nabízí. To nás má utvrdit v tom, že Země existovala mnohem dříve, než ji začal obydlovat nějaký typ života. Poté obrovská mimozemská loď opustí jednoho ze svých, který z ničeho nic (možná důsledkem zemského vzduchu) začne degenerovat a rozpadat se na molekuly DNA, čímž zamořuje zemskou řeku. Části DNA se po čase začnou tvarovat zpět. To nám má podle všeho představit našeho „stvořitele“, neznámou mimozemskou rasu. „*Big things have small beginnings.*”²⁰⁹ prohlásí jeden z protagonistů v průběhu filmu a tato scéna měla být přesně tím malým počátkem pro člověka. (Je ironické, že zrovna tuto větu můžeme slyšet z úst robota.) Přístup ke stvoření lidské rasy je důvodem, proč můžeme film považovat za velice odvážný. Podle něj jsme totiž nebyli stvořeni díky křesťanské myšlence Boha, ale díky bytostem z jiné planety. Ale stále zůstává otázka, pokud nás vytvořili nějakí „inženýři“ lidstva, kdo stvořil je? Kdo je ta konečná bytost, která

²⁰⁹ SCOTT, R., et al. *Prometheus* [film]

stvořila vše? Můžeme ocenit to, jak se film pokouší odpovídat na zásadní lidské otázky, ale také ho můžeme vinit z toho, že je pouze rozšiřuje.

Jedním z protagonistů filmu je vědkyně Elizabeth Shaw, která se hlásí ke křesťanské víře. Pro křesťanství se rozhodla díky jejímu otci, který sám v Boha věřil. Ona je skrze film konfrontována se svou vírou. Sama v jedné scéně prohlašuje, že není křesťankou proto, že má nějaký věrohodný důkaz, nebo proto, že by byla přesvědčená o absolutní pravdě této víry, ale „*because it's what I choose to believe*”²¹⁰. Toto je další myšlenkou filmu *Prometheus*. Potřeba lidí v něco věřit. A tím velice zpochybňuje celou křesťanskou víru.

A tím se dostáváme k tématu, které je stěžejní pro tuto diplomovou práci. V jednom momentu filmu *Prometheus* charakterizuje Peter Weyland (který stvořil robota vystupujícího v tomto snímku) lidstvo jako Bohy. Tvrdí, že nás stvořil někdo neznámý, ovšem my máme v současné době moc tvořit také, na jedné straně nás samé a na druhé život umělý. V této době jsme schopni stvořit roboty, androidy, počítače, které určitým způsobem myslí samy za sebe, ovšem stále nejsme schopni vyrobit zástupce UI na takové úrovni, na jaké jsou prezentováni ve sci-fi tvorbě. Ve snímku *Prometheus* došlo lidstvo do takového bodu, kdy je schopno vytvořit nesmrtelného robota, který je nerozeznatelný od smrtelného člověka. Například řečtí bohové neměli příliš v lásce, když se lidé určitým způsobem začali podobat bohům. Zde můžeme nalézt analogii se samotným názvem filmu. Prometheus byl totiž podle řeckých bájí titán, který dal lidstvu dar ohně (čímž nás měl přiblížit k božství) a zručnosti v kovoprůmyslu, a kvůli tomu byl potrestán samotným Diem. Zeus přikoval Promethea ke skále a posílal za ním každý den orla, který mu vykloval játra, jež pak přes noc opět dorostla.²¹¹ (Ve filmu je podle titána Promethea pojmenována vesmírná loď, díky které se protagonisté dostávají do vesmíru a mohou zkoumat naše „stvořitele“.) V tomto snímku zastupuje oheň lidská touha po odpovědích na otázky o stvoření lidstva, ale také naše schopnost tvořit. Proto je také lidské „absolutní“ vědění důvodem vyvraždění téměř celé posádky Promethea.

²¹⁰ SCOTT, R., et al. *Prometheus* [film]

²¹¹ CARTWRIGHT, M. Prometheus, In: *Ancient History Encyclopedia*

Jiné souvislosti ve filmu můžeme nalézt také v případě Davida, jediného zástupce UI, který sdílí jméno s robotickým chlapcem z filmu *A.I.: Artificial Intelligence*. David je vesmírný majordomus, který slouží rozmarům zakladatele společnosti a svého stvořitele Petera Weylanda. Oba Davidové jsou tedy taktéž považovány za „potomky“ určitého zástupce lidské rasy. A oběma přisuzujeme neschopnost emocí, v čemž nás oba přesvědčí, že jsme se mýlili. Pokud si spojíme tyto dvě postavy dohromady, pomůže nám to v pochopení Davidových záměrů ve filmu *Prometheus*. Záměrů, jako je vymýcení lidské rasy. Oba Davidové začnou mít vražedné úmysly, když se začnou cítit méněcenní ve srovnání s běžnými lidskými bytostmi. Oba dva jsou neuvěřitelně fyzicky podobní lidem, ale naopak je dokážou inteligenčně nebo emocionálně předčit. Z toho důvodu se nejspíše ani nesnaží dostat se na lidskou úroveň. Když jeden z pasažérů lodi *Prometheus* poznamenává, že jsou lidé schopni vyrobit androidy čím dál více podobné člověku, David na to odpovídá: „*Not too much I hope.*“²¹². Již toto nám může naznačit Davidův postoj k lidské rase. I přesto, že se David snaží objevit tajemství věčného života pro svého vlastního stvořitele, se zdá, že má svůj vlastní program. Také je fascinován a zároveň znechucen lidskou skupinou vědců, která ho doprovází do vesmíru a zachází s ním s přehnanou lhostejností.

Posádka *Promethea* odjíždí na svou misi s nadějí a očekáváním. Namísto toho nalézají existenční hrůzu a rozčarování. „*It must feel like your God abandoned you*“²¹³, zazní v jedné scéně filmu. Lidská rasa se pomalu začíná podobat tomu, co viděli v Davidovi. Nedokonalé stvoření nedokonalé rasy, rasy „inženýrů“, která se jim celý život vyhýbala a aktivně je nelitovala. Pokud nad Davidem přemýšlíme hlouběji, dochází nám, že to on je hlavní postavou filmu a že se na film musíme dívat z jeho pohledu. Začneme si uvědomovat, že nejvíce času ve filmu věnoval Ridley Scott právě jemu. Ukazuje nám ho při sledování jeho nejoblíbenějšího filmu nebo při úpravě vlasů. Víme o něm více než o ostatních (lidských) členech posádky *Promethea*. Možná tedy není Davidovo experimentování s mimozemskou rasou pouze čistou pomstou. Možná se snaží přijít na něco, co z něj udělá bytost z masa a kostí.

²¹² SCOTT, R., et al. *Prometheus* [film]

²¹³ Tamtéž

Ovšem titán Prometheus není tím jediným, ke kterému název filmu vede. Další mysteriózní postava, na kterou si při pohledu na název filmu můžeme vzpomenout, je Frankenstein. Mary Schelley, autorka díla o Frankensteinovi, dala knize příhodný podtitul – Moderní Prometheus. Poté, co posádka lodi Prometheus přistála ve vesmíru, našla zde klíč ke všem živým věcem. To samé hledal Victor Frankenstein, když tvořil své monstrum. A možná to nám přijde na Davidovi tak fascinující a přitažlivé. Jeho touha po odpovědích na zdánlivě nezodpověditelné otázky. Nemusí být lidskou bytostí, ale jeho selhání a jeho strach úzce odráží ty lidské. Film nám ovšem neprozrazuje, zda David koná tyto hrůzné činy namířené proti lidstvu kvůli tomu, že je naprogramován tak, aby za každou cenu našel lék pro svého stvořitele nebo z nějakých sobeckých, osobních důvodů. Ve druhém případě bychom mu však připisovali vědomí.

4.8. THE MACHINE – Stroj (2013)

„Conscious machines are the last thing we need. Have you any idea how dangerous that would be“? Thomson, *The Machine*

Žijeme v době, kdy nám chytré telefony říkají, jak se máme obléci, kdy máme odejít z bytu, abychom stihli letadlo, nebo v kolik hodin máme jít spát, abychom spali správný počet hodin. Turingův test je téměř pokořen a psychoterapii nám zastupují naprogramovaní chatboti. Ovšem na pravou UI s plnohodnotným vědomím stále čekáme. Film *The Machine* od režiséra Caradoga W. Jamese nám představuje jeden ze způsobů, jak tuto vědomou UI vytvořit.

Protagonistou filmu *The machine* je doktor Vincence McCarty. Zde můžeme nalézt spojitost s Johnem McCarthym, americkým informatikem a kognitivním vědcem, který navrhl používat pojem „umělá inteligence“. V roce 1971 obdržel Turingovu cenu za přínos v oblasti UI.²¹⁴ Příběh snímku se odehrává v blízké budoucnosti, kdy probíhá studená válka mezi Čínou a západním světem, kde hrozí ekonomická krize. Ovšem budování prvních inteligentních strojů, které by pomohly zvítězit tuto válku, je hlavním zájmem všech velmocí. McCarthy pracuje pro ministerstvo obrany Spojeného

²¹⁴ RUSSELL, S. J.; NORVIG P. *Artificial Intelligence: A modern approach*, str. 12

království a navrhuje implantáty pro vojáky s poškozením mozku. Zároveň se snaží se svými spolupracovníky pokořit Turingův test, což se mu ale nedaří. McCarthy dokáže vytvářet umělá těla s nadlidskou silou, ovšem potřebuje najít způsob, jak přivést k životu umělý mozek. Vtom se setkává s vědkyní v oblasti UI Avou, která mu ve výrobě vědomého stroje pomáhá. Sama totiž vyrobila počítačový program, který využívá jejich rozhovorů s jinými lidmi, aby se sám naučil řeči.

V průběhu filmu je Ava smrtelně zraněná a McCarthy se rozhodne vytvořit androida založeného na osobnosti Avy pomocí skenování mozku. Výsledkem replikování je fyzicky nezničitelný klon Avy, který se jednoduše nazývá Stroj (The Machine). Snímek *The Machine* nás nabádá k hledání podobností se známou sci-fi tvorbou. Ve chvíli, kdy je lidská Ava nahrazena super-inteligentní robotí replikou, napadají nás filmy jako *Terminator* a *Metropolis*. Ve filmu se objevuje také připomínka klasického díla *Solaris*, když se robotická Ava ptá McCartyho: „*What makes my clever imitation of life any different from theirs? Apart from their flesh, what makes them any different from me?*“.²¹⁵

Komplikovaným aktérem příběhu je McCartyho šéf Thompson, který si myslí, že vlastnit stroj s vědomím je obtížnější, než pokud vědomí zastupuje snadno ovládnutelný program. Film *The Machine* pokládá nespočet filosofických otázek. Pokud implantujeme člověku místo poškozeného mozku mozek umělý, jak velká část jeho vlastní lidské osobnosti zůstane? Tato otázka nám připomíná experiment náhrady mozku, který jsme si dříve představili. Zastánci tohoto experimentu odpovídají na otázku náhrady biologického mozku za umělý kladně. Ovšem neexistuje žádný přesvědčivý důkaz o tom, že náhrada mozku nezpůsobila změny chování, nebo dokonce celé osobnosti. Lidské tělo se syntetickým mozkiem může přemýšlet a cítit, je to ale dost proto, abychom takovouto hříčku přírody považovali za skutečnou lidskou bytost? Navíc, jaká je poté role lidstva? Poté, co dokáže vytvořit stroje, které by měli považovat

²¹⁵ JAMES, C. W., et al. *The Machine* [film]

za kvalitnější bytosti, než jsou oni sami. Stroj ve filmu odpovídá na tuto otázku doktoru McCarthymu: „*They are part of the new world, you are part of the old.*”²¹⁶

Po dosavadních analýzách filmů můžeme s jistotou říci, že se nejedná o první film, který otevírá tyto otázky. *The Machine* je ovšem jedinečný především díky hereckým výkonům hlavních postav. Toby Stevens (Vincenc McCarthy) dodává postavě dokonalou unavenost a zranitelnost lidské postavy. Zatímco Caity Lotz (Ava) propůjčuje své dvojroli neobyčejný kontrast lidskosti a chladného obrazu nehumánnosti.

Od etické stránky UI film zcela odhlíží. Sice by se nabízely otázky odpovědnosti strojů za své činy, využití UI pro nekalé účely nebo v poslední řadě téma ztráty zaměstnání lidské rasy. Tyto otázky film pokládá, ale neodpovídá na ně. Nové technologie nás dostávají do konfliktu lidské a technologické zodpovědnosti. A tento konflikt bude gradovat stejnou rychlostí, jakou se nové technologie vyvíjejí. Ale zatím to vypadá, že lidská rasa bude etické problémy raději obcházet, než aby tyto zdánlivě neřešitelné aspekty UI musela řešit.

5. Shrnutí analýzy filmů

Dosavadní analýza filmů nám umožňuje zhodnotit, jak se změnil pohled filmových tvůrců na UI v posledních padesáti letech. Pokud tedy srovnáme pohled na UI ve starší filmové tvorbě a v té současné, můžeme tvrdit, že se obraz UI zásadně obrátil. Dřívější filmová produkce ukazovala divákům UI jako hrozbu. Nové technologie trpěly mentálními abnormalitami nebo disponovaly různými podobami deviantního chování. Tyto defekty se sice podobají lidským poruchám, ovšem u UI dochází k extrémům. Stroje například trpěly až paranoidní snahou o přežití. V některých případech až přehnaným pocitem megalomanství a komplexem nadřazenosti. Ale ani tyto případy extrémního deviantního chování nejsou pro lidskou rasu neobvyklé. Mohou se tedy lidské problémy a nedostatky zrcadlit ve znázornění UI ve filmové tvorbě? Intuitivní odpověď je ano, jelikož dalšími tématy minulé sci-fi tvorby je zpochybňování taktéž

²¹⁶ JAMES, C. W., et al. *The Machine* [film]

lidského chování. Volby, které lidé činí, jsou podle filmových tvůrců nedokonalé a často vedou lidstvo ke zkáze. Téma lidské nedokonalosti je ovšem aktuálním a často hlavním tématem i současné filmové produkce.

První analyzovaný film *Her* je příkladem toho, jak se v kinematografii změnil pohled na UI. Nové technologie nepovažuje za naši zkázu nebo hrozbu, které bychom se měli obávat. Naopak vyličení UI v tomto snímku napovídá, že by nám mohly být nově vynalezené technologie ku prospěchu. Obrat k pozitivnějšímu obrazu technologií ve filmové tvorbě může být způsoben lepší obeznámeností veřejnosti. Nebo také tím, že měla lidská společnost možnost si na nové technologie určitým způsobem zvyknout. Začít je brát jako svou součást, která jim může pomoci k lepšímu a snadnějšímu životu. Nabízí se otázka, zda je již společnost připravená pro integraci UI do běžného života? A pokud ano, v jakém rozsahu bude toto začlenění možné, než začne lidskou rasu ovládat strach ze ztráty nadvlády nad světem? Lidstvo prochází různými etapami, je možné zmínit otázku Židů, předsudky ohledně jiné barvy pleti nebo problém gayů. Všechny tyto menšiny musely disponovat jistou dávkou trpělivosti a vyčkat, než je většinová společnost přijme. K těmto menšinám můžeme přirovnat i veškeré nové technologie. Lidská rasa vyžaduje čas, aby byla integrace bezbolestně proveditelná.

Současná filmová produkce se ze všech sil snaží zobrazovat UI v reálné podobě. Ovšem na druhou stranu dává tvůrcům a posléze i divákům prostor použít fantazii. Demystifikace nových technologií může širokému publiku pomoci v přijetí UI. A jejich fantazie může být nápomocná pro případ, kdy se budou nové technologie rozvíjet rychlejším tempem, než je očekáváno. Dnes si díky filmové tvorbě umíme představit roboty s lidským vědomím, operační systém, který je naším přítelem nebo armádu mechanických vojáků, kteří se nám nesnaží ublížit.

Starší filmová tvorba se primárně stavěla k UI negativně, to můžeme tedy považovat za projev veřejné neznalosti tohoto nového oboru a přičítat to strachu z možného a těžce předvídatelného vývoje. Avšak filmová produkce poslední dekády chápe UI odlišně. Toho si můžeme všimnout u filmů *Her*, *Wall-e*, *Big Hero 6* a *Robot & Frank*. V těchto případech je pozitivně vyzdvihována UI a naproti tomu je lidská rasa kritizována

a stavěna do pozice hrozby pro naši planetu. Lidé byli sice kritizováni již v minulé filmové tvorbě, ovšem ne v takovém rozsahu jako v současnosti. Filmy už vlastně nevyprávějí příběhy zlé UI, která přišla zničit lidské životy a ovládnout planetu Zemi. UI již vlastně není hlavním protagonistou ústředního varování před humánní zkázkou. Ve většině příběhů máme z implicitního vyprávění vyčíst, že jsme hrozbou pro sebe samé, že si vlastní planety, svých blízkých a vlastně všeho, co máme, nevážíme. UI zastupuje roli kladných postav, které nám mají pomoci uvědomit si vlastní chyby a po našem boku (ne-li v čele) zachránit naši planetu.

Opakujícím se aspektem ve většině sci-fi filmů je náboženství nebo určitá spojitost s Biblií. To můžeme nalézt v pojmenování postav (*Ex Machina*, *Wall-e*, *Prometheus*, *The Machine*) nebo také jako celou dějovou linku (*Prometheus*). Představme si hypotetickou situaci, kterou ve své knize *Science fiction and philosophy* představuje Susan Schneider: „*Píše se rok 2035, vy a vaše rodina sedíte v kostele a posloucháte bohoslužbu, kterou pronáší robotický pastor. Tento kostel je konkrétně Presbyteriánský kostel lidí a robotů. Jste staromódními občany, jelikož jste vyrůstali v době, kdy nejvíce pokročilá UI sídlila pouze ve vašich mobilních telefonech. Bylo by pro vás přínosné, abyste chodili každou neděli do kostela, který spadá pod správu robota?*“²¹⁷ A pokud ne, není to předsudek? Kdo my lidé jsme, abychom mohli něco/někoho soudit? V současné době, což můžeme nalézt taktéž ve filmové tvorbě, se řeší, zda je UI a náboženství kompatibilní. A jestli ano, zda se budou náboženští vůdci snažit získat pokročilou UI do svého stáda? Tyto otázky by ovšem byly spíše možným námětem dalších science fiction snímků, jelikož odpovědi na ně nemůžeme podložit pádnými důkazy, jen konspiračními teoriemi.

Dalším aspektem, který současná filmová tvorba nerozvíjí, je vysvětlení touhy UI po lidské podobě, lásce a uznání. Je to jejich program, co jim přisuzuje tato nesrozumitelná přání? Je to stejné jako naše tužba se někomu podobat nebo mít něco, co nemáme? Nebo to souvisí s pomyslnou lidskou nadvládou, připodobňování se bohům a schopností tvořit život, i ten neživý? Nesrozumitelnost těchto přání může pramenit také z toho, že filmy nevysvětlují, na jakém základě jsou stroje nebo programy ve filmu

²¹⁷ SCHNEIDER, S. *Science fiction and philosophy: from time travel to superintelligence*, str. 246

vědomé. Jestli je toto mechanické vědomí totožné jako naše lidské, biologické vědomí? Nebo jestli se v průběhu času utváří samo a stává se něčím naprosto jiným? Pokud zvážíme teorii širokého mentálního obsahu, tak už jen to, že entity UI koexistují na totožné planetě jako lidská rasa, by mělo tyto uměle vytvořené bytosti tvarovat do lidské podoby, k našemu vnímání okolí a humánnímu chování. Když budeme tvořit „bezchybné“, naprosto logické a systematické programy, nikdy se UI nebude lidskému pokolení podobat.

UI nás za pár stovek let může považovat za neandrtálce, nižší vývojovou etapu a pít se po tom, kdo je vlastně jejich stvořitelem. Nemusejí přeci mluvit stejným jazykem jako lidé, nemusejí ani stejně myslet a mít totožné emoce. Nemůžeme již dnes tvrdit, že stroje nebo programy nemyslí nebo nemají vědomí. Většina vědeckých teorií je podložena antropocentrickým viděním světa. Už jen otázka, zda je možné, aby stroj přemýšlel, je totiž nesmyslná. Takovéto otázky napadaly již Ludwiga Wittgensteina nebo Thomase Nagela, který v roce 1974 napsal článek „*What is it like to be a Bat?*“ (přel. Jaké to je být netopýrem?), ve kterém tvrdí, že rozdílná těla mohou mít také různá vědomí nebo různý pohled na svět. Vědomí, mysl i pohled na svět získáváme ze zkušeností našeho těla. Nagelovu teorii musíme považovat za přínosnou, pokud přistoupíme na tvrzení, že tělo má vliv na naše vědomí a v návaznosti na to musíme uznat, že lidské tělo je naprosto odlišné od těla UI (pokud není pouze určitým programem). To v podstatě znamená, že i když si my jako lidé dokážeme představit létání, jaké to je viset hlavou dolů nebo vlastnit schopnost echolokace, nikdy nebudeme vědět, jaké to je být netopýrem. Nikdy nezažijeme netopýří, jedinečné, subjektivní zážitky. I kdyby byl lidský mozek přemístěn do hlavy netopýra, stále bychom nevěděli, co znamená být netopýr, protože náš mozek bude stále vlastnit lidské vědomí. Byli bychom netopýry prostřednictvím lidského myšlení.²¹⁸ Díky dosavadním úvahám nás může napadnout, zda na tom vůbec záleží. Zda je důležité, přijít na podstatu fenoménu vědomí. Proč se vědci vlastně snaží napodobit něco biologického něčím uměle vytvořeným? Je pochopitelné, že lidé musejí mít při vytváření něčeho kompletně nového určitou (biologickou) předlohu. Proč se však nesnaží vytvořit něco, co je

²¹⁸ NAGEL, Thomas. What Is It Like to Be a Bat?. In: *The Philosophical Review*, str. 435-450

překvapí? Ze strachu, že by tato nově vytvořená věc lidstvo předčila? Samozřejmě, že již existují programy UI, které se samy učí, ovšem stále jsou něčím omezeny. Pokud dojde na nejhorší scénář science fiction filmové tvorby, kdy se nás UI bude snažit vyhledat, měli bychom jí nabídnout něco, co bude prospěšné také pro ni. Něco, co by je utvrdilo v tom, že je prospěšné s námi kooperovat a koexistovat. S tím se ovšem počítá i v případě UI, dokazuje to například filmová tvorba poslední dekády.

Díky dosavadní analýze můžeme předpokládat, že budou v nadcházející době kinematografické příspěvky stále více otevírat a lépe zobrazovat UI. Jedním z příkladů může být pokračování úspěšného seriálu *Westworld*, inspirací tomuto seriálu byl stejnojmenný film z roku 1973 a pojednává o futuristickém zábavním parku, kde jsou hlavní atrakcí roboti, kteří mají uspokojit jakékoli touhy a zvrhlá lidská přání. *Westworld* byl serverem IMDb umístěn na prvním místě jako nejlépe hodnocený seriál s tematikou UI. Také můžeme zmínit film Stevena Spielberga *Ready Player One*, který má publiku přiblížit krásu a moc virtuální reality. Nebo snímek Roberta Rodrigueze *Alita: Battle Angel*, který popisuje dalekou budoucnost, kdy jsou technologické možnosti na úrovni, kterou si zatím bez filmové tvorby neumíme představit. Náměty filmové tvorby se časem mění tak, jako se mění vědecké teorie a vynálezy. Pomáhají veřejnosti zamýšlet se nad novými technologiemi a přijímat je kritickým pohledem. Díky změnám ve filmové sféře se mohou snáze měnit kulturní paradigmaty. Témata jako koexistence lidí a robotů, virtuální realita nebo vylepšování člověka pomocí technologických vynálezů ve stále lepším a reálnějším zpracování mohou pomoci veřejnosti přijmout nevyhnutelnou budoucnost, ve které může být UI lidstvu nápomocná, nebo také nemusí.

6. Závěr

Diplomová práce nejprve představila kognitivní filmovou teorii. Ta byla v hlavní stati použita jako metoda práce, jelikož se při výkladu filmů soustředí na perspektivu diváka. Ve filmech tedy nejsou primárně nalézány intence autora, ale působení snímku na publikum. Filmová kognitivní teorie také umožnila výběr analyzované filmové tvorby pro diplomovou práci, jelikož bylo vybráno osm divácky nejlépe hodnocených filmů o UI poslední dekády z internetového serveru IMDb. Pro další účely tématu práce by bylo vhodné zaměřit se především na kognitivní filmovou teorii a to, jak veřejné publikum vnímá fenomén UI skrze filmovou tvorbu. Žebříček nejlépe hodnocených filmů totiž nevypovídá o tom, jak je UI ve filmech zobrazována nebo zda jsou tyto snímky filosoficky hodnotné. Ve vysokém hodnocení filmů hraje roli nespočet faktorů od známého autora, finanční prostředky filmové produkce (např. pro reklamní kampaň), až po vizuální efekty, hudbu, herecké výkony atd. Bylo by proto vhodné zanalyzovat například veřejné komentáře publika k filmové tvorbě, a zhodnotit tak tyto veřejné názory.

V další části práce byly nejprve nastíněny některé definice UI, na kterých bylo ukázáno, že se obecná charakteristika UI víceméně liší. A dále byly představeny filosofické koncepce UI, především na podkladu knihy *Artificial Intelligence: A modern approach*, kterou napsali Stuart Russell a Peter Norvig. Na základě této knihy se filosofie v oblasti UI soustředí především na předmět inteligence a otázku, zda se mohou stroje chovat inteligentně, toto téma primárně zastupuje Turingův test a vybrané námitky na něj. Poslední zmíněným argumentem proti Turingovu testu v této práci je námitka z vědomí. Ta otevírá další filosoficky orientovanou oblast v oboru UI. V této části práce byly přiblíženy teorie mysli, kterými byly dualistická a monistická teorie, neboli fyzikalismus, také funkcionalismus, epifenomenalismus, biologický naturalismus a konečně počítačová teorie mysli. Některé z těchto teorií byly vyvráceny nebo naopak podpořeny myšlenkovými experimenty, jako jsou mozek v kádi, experiment náhrady mozku nebo Searlův Čínský pokoj, tyto experimenty byly v práci taktéž zmíněny. A dále se druhá sekce filosofických koncepcí zabývala otázkami intencionálních stavů,

širokým a úzkým mentálním obsahem nebo sebeuvědomováním a kontroverzním pojetím kválií. Po pojetí mysli a vědomí byla v práci rozebrána také otázka etiky a morálky UI. Především tedy dopad začlenění nových technologií do běžného lidského života. Nástup nových technologií, podle Russella a Norviga, může znamenat například vysokou lidskou nezaměstnanost, úbytek pocitu jedinečnosti zástupců lidské rasy, ztrátu lidské zodpovědnosti za činy, které vykonaly entity UI. Také mohou být systémy UI použity k nekalým účelům. Tento počin by mohl vést až ke konci lidské civilizace, s lidským přičiněním nebo bez něj. Filosofické koncepce UI zastupují v práci teoretickou část, která byla dále zkoumána ve filmové tvorbě. Cílem práce tedy bylo představení hlavních proudů UI ve filosofickém bádání, jejich následná komparace s nejvíce populární science fiction tvorbou současné doby a nalézání vzájemných inspirací.

V konečném shrnutí bylo nastíněno, jakými tématy se současná filmová produkce zabývá a jakými naopak ne. V současné filmové tvorbě jsou nejvíce rozebírána etická témata, především dopad UI na život člověka. Filmoví tvůrci se nesnaží proniknout do filosofických koncepcí vědomí, nepopisují nám tedy podstatu myslících strojů, ani to, proč mají zástupci UI touhu po lidské lásce a uznání. Zaměřují se spíše na to, jak tyto myslící entity působí na člověka a naopak. Zajímavým postřehem je vztah UI a náboženství. V první řadě souvislosti v otázce Boha jako našeho stvořitele a toho, zda jsme sami bohy, když dokážeme stvořit stroj, který disponuje vědomím. Dále se náboženství vyskytuje ve filmech prostřednictvím například pojmenování hlavních postav.

Díky výsledné analýze filmů můžeme dojít k závěru, že se filmová tvorba inspiruje filosofickými koncepcemi spíše než filosofické úvahy filmovou produkcí. Toto tvrzení můžeme doložit například filmy *Ex Machina* a *Big Hero 6*, kde se tvůrci snaží jednak přiblížit reálnému obrazu UI, řešit současné filosofické problémy nových technologií a také se plně inspirovat filosofickými myšlenkovými experimenty. Výsledná konkluze byla záměrně, ale také implicitně zaměřena na tři období, těmi jsou minulost, přítomnost a budoucnost. Z části o minulosti můžeme vyvodit, že filmová produkce změnila svůj pohled na otázku nových technologií a jejich kooperaci s lidskou rasou.

Dříve je populární tvorba (z toho vyplývá i mínění diváka) pojímala jako hrozbu. Filmová tvorba prostřednictvím zástupců UI vyvolávala v lidech úzkost a strach. Současná tvorba také obsahuje prvky tohoto negativního obrazu UI, avšak obrací zodpovědnost za nekalé činnosti a třeba také za následnou zkázu lidstva, člověku. Nové technologie jsou především zobrazovány jako něco prospěšného a užitečného, něco, co pomáhá člověku objevovat sama sebe a zlepšovat každodenní život. V sekci o budoucím vývoji byla stručně představena nadcházející filmová tvorba a témata, která bude obsahovat. Tím mělo být nastíněno, že se filmová tvorba vyvíjí společně s novými vynálezy, které představuje široké veřejnosti. To by mohlo pomoci lidem přijímat nové technologie snadněji a třeba také s pozitivním přístupem.

7. Použité zdroje

BORDWELL, David. A Case for Cognitivism. In: *Iris: A Journal of Theory on Image and Sound* [online]. vol. 9, 1989 [cit. 15. 2. 2018]. Str. 11-41. Dostupné z:

http://www.davidbordwell.net/articles/Bordwell_Iris_no9_spring1989_11.pdf.

BOSTROM, Nick. *Superintelligence: až budou stroje chytřejší než lidé*. Přel. Jan Petříček. Praha: Prostor, 2017. Obzor (Prostor). ISBN 978-80-7260-353-4.

BUTTAZZO, Giorgio. Artificial consciousness: Utopia or real possibility?. In: *Computer* [online]. Vol. 34, No. 7, July 2001 [cit. 12. 3. 2018]. Str. 24-30. ISSN 0018-9162.

CALDWELL, Thomas. Free Will, Technology and Violence in a Futuristic Vision of Humanity: '2001: A Space Odyssey'. In: *Screen Education* [online]. No. 58, 2010 [cit. 18. 3. 2018]. Str. 133-137. ISSN: 1449-857X.

CALO, Ryan. *Sorcerers Apprentice or why Weak AI interesting Enough* [online]. The Center for Internet and Society, 2011 [cit. 16. 3. 2018] Dostupné z: <http://cyberlaw.stanford.edu/blog/2011/08/sorcerers-apprentice-or-why-weak-ai-interesting-enough>

CASSETTI, Francesco. *Filmové teorie 1945-1990*. V Praze: Akademie múzických umění, 2008. ISBN 978-80-7331-143-8.

CONRAD, Daniel. *A Functional Model of the Aesthetic Response* [online]. Journal of Contemporary Aesthetics, 2010 [cit. 8. 3. 2018]. Dostupné z:

<http://www.contempaesthetics.org/newvolume/pages/article.php?articleID=581>

CURRIE, Gregory. *Arts and minds*. Oxford: Clarendon Press, 2004. ISBN 0-19-925629-2.

Definition of Deus Ex Machina. In: English Oxford Dictionaries [online]. Dostupné z: https://en.oxforddictionaries.com/definition/deus_ex_machina. [cit. 20. 3. 2018].

- DE MUL, Jos. *The Technological Sublime* [online]. Next Nature Network. 2011, [cit. 23. 3. 2018]. Dostupné z: <https://www.nextnature.net/2011/07/the-technological-sublime/>
- DENNETT, Daniel Clement. *Consciousness explained*. London: Penguin Books, 1993. Penguin science philosophy. ISBN 0-14-012867-0.
- DINELLO, Daniel. *Technophobia!: science fiction visions of posthuman technology*. Austin: University of Texas Press, 2005. ISBN 978-0-292-70986-7.
- FISHER, Robert B. AI and Cinema – Does Artificial Insanity Rule? [online]. In: *Representations of Artificial Intelligence in Cinema*. University of Edinburgh's School of Informatics. 2002 [cit. 19. 2. 2018]. Dostupné z: <https://homepages.inf.ed.ac.uk/rbf/PAPERS/aicinema.pdf>
- FOOT, Philippa. *Virtues and vices and other essays in moral philosophy*. New York: Oxford University Press, 2002. ISBN 9780199252862.
- GAWANDE, Atul. *Complications: a surgeon's notes on an imperfect science*. London: Profile Books, 2002. ISBN 1861974132.
- GOOD, Irvin John, Speculations concerning the first ultraintelligent machine [online]. In: *Advances in Computers*, New York: Academic Press, 1965 [cit. 22. 3. 2018]. Dostupné z: [https://doi.org/10.1016/S0065-2458\(08\)60418-0](https://doi.org/10.1016/S0065-2458(08)60418-0).
- GUNN, James. *Isaac Asimov: the foundations of science fiction*. Rev. ed. Lanham (Md.): the Scarecrow press, 2005. ISBN 9780810854208.
- HALES, Colin G., *The Revolutions of Scientific Structures*. World Scientific Publishing Company, 2014. ISBN 9814616249.
- HARMAN, Gilbert. *Thought*. Princeton University Press, 2015, ISBN 978-0691618050.
- HAVEL, Ivan M. Přirozené a umělé myšlení jako filosofický problém, In: *Umělá inteligence III*. (V. Mařík a kol., editoři), Academia, Praha 2001, str. 17–75. ISBN 80-200-0472-6.

- HEATH, Stephen. Narrative Space [online]. In: *Screen*. Vol. 17, No. 3, str. 68–112. 1. 11. 1976 [cit. 22. 2. 2018]. Dostupné z: <https://doi.org/10.1093/screen/17.3.68>
- HICKEY, Lance P. The Brain in a Vat Argument [online]. In: *Internet Encyclopedia of Philosophy*. Dostupné z: <https://www.iep.utm.edu/brainvat/> ISSN: 2161-0002.
- CHURCHLAND, Patricia Smith. The Brain and Its Self. *Proceedings of the American Philosophical Society* [online]. Vol. 155, No. 1, 2011, str. 41–50 [cit. 19. 3. 2018]. Dostupné prostřednictvím JSTOR z: www.jstor.org/stable/23056848.
- IMDb. *10 Best Artificial Intelligence Movies of the Last Decade* [online]. ©2017 [cit. 11. 3. 2011] Dostupné z: <http://www.imdb.com/list/ls068528362/>
- IMDb. In: Wikipedia. Dostupné z: <https://en.wikipedia.org/wiki/IMDb>
- KURZWEIL, Ray. *The singularity is near: when humans transcend biology*. New York: Viking, 2005. ISBN 0-670-03384-7.
- LEVY, David N. L. *Love + sex with robots: the evolution of human-robot relations*. New York: HarperCollins, c2007. ISBN 9780061359804.
- MALLE, et al. *Sacrifice One For the Good of Many? People Apply Different Moral Norms to Human and Robot Agents* [online]. Portland, USA, 5. 3. 2015 [cit. 25. 3. 2018]. Dostupné z: <http://dx.doi.org/10.1145/2696454.2696458>.
- MAŘÍK, Vladimír, Olga ŠTĚPÁNKOVÁ a Jiří LAŽANSKÝ. *Umělá inteligence I*. Praha: Academia, 1993. ISBN 80-200-0496-3.
- MAŘÍK, Vladimír, Olga ŠTĚPÁNKOVÁ a Jiří LAŽANSKÝ. *Umělá inteligence III*. Praha: Academia, 2001. ISBN 80-200-0472-6.
- MCCARTHY, John. The Philosophy of AI and AI of Philosophy. *Philosophy: the journal of the British Institute of Philosophy*. [online]. London: Macmillan & Co., 1931 [cit. 3. 3. 2018]. ISSN 0031-8191.

MCCORDUCK, Pamela. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*. 25th anniversary update. Natick, Mass.: A. K. Peters, c2004. ISBN 1-56881-205-1.

MORAVEC, Hans. *Robot: Mere Machine to Transcendent Mind*. New York, 1999. ISBN 978-0195136302.

NAGEL, Thomas. What Is It Like to Be a Bat?. *The Philosophical Review*[online]. Vol. 83, No. 4, 1974, str. 435- 450, [cit. 2. 4. 2018]. ISSN: 1558-1470.

NANNICELLI, Ted. *Cognitive media theory*. Routledge; 1 edition 2014. ISBN 041562987X.

NATH, Rajakishore. *Philosophy of artificial intelligence: a critique of the mechanistic theory of mind*. Boca Raton, Fla: Universal-Publishers, 2009. ISBN 1599429055.

NILSSON, Nils J., Human-Level Artificial Intelligence? Be Serious! [online]. In: *AI Magazine Magazine*. Vol. 26, No. 4, 2006 [cit. 18. 3. 2018]. Dostupné z: <https://doi.org/10.1609/aimag.v26i4.1850>

NILSSON, Nils J. Logic and artificial intelligence [online]. In: *Artificial Intelligence*. Vol. 47, 1991, str. 31-56 [cit. 18. 3. 2018]. Dostupné z: [https://doi.org/10.1016/0004-3702\(91\)90049-P](https://doi.org/10.1016/0004-3702(91)90049-P)

NILSSON, Nils J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements* [online]. Cambirdge University Press, 2009 [cit. 12. 3. 2018]. Dostupné z: <http://dx.doi.org.10.1017/CBO9780511819346>

PARKER, Laura. Human After All: Ex Machina's Novel Take on Artificial Intelligence. *The Atlantic* [online]. 15. 3. 2015 [Cit. 20. 3. 2018]. Dostupné z: <https://www.theatlantic.com/entertainment/archive/2015/04/ex-machina-and-thevirtues-of-humanizing-artificial-intelligence/390279/>

- PERVEZ, Anis. Cognitive Film Theory. *Bangladesh Film and Television Institute Journal* [online]. Vol. 1, 2015 [cit. 15. 2. 2018]. Dostupné z: http://www.academia.edu/24543818/Cognitive_film_theory_An_overview
- PLANTINGA, Carl. Cognitive Film Theory: An Insider Appraisal. *Cinemas: Journal of Film Studies* [online]. Vol. 12, No. 2, 2002, str. 15- 37 [cit. 13. 2. 2018]. ISSN 1705-6500.
- POLÁK, Michal. *Filosofie mysli*. Praha: Triton, 2013. ISBN 978-80-7387-742-2.
- PRESTON, John a Mark BISHOP. *Views into the Chinese room: new essays on Searle and artificial intelligence*. New York: Clarendon Press, 2002. ISBN 978-0199252770.
- ROSSI, Sandra. *Bevare the CyberLover that Steals Personal Data* [online]. 2007 [cit. 13. 3. 2018]. Dostupné z: <https://www.pcworld.com/article/140507/article.html>
- RUSSELL, Stuart J. a Peter NORVIG. *Artificial intelligence: a modern approach*. 3rd ed. Harlow: Pearson Education, c2014. ISBN 978-1-29202-420-2.
- SAENZ, Aaron. We Live in a Jungle Of Artificial Intelligence that will spawn Sentience. *SingularityHub* [online]. 2010 [cit. 15. 3. 2018] Dostupné z: <https://singularityhub.com/2010/08/10/we-live-in-a-jungle-of-artificial-intelligence-that-will-spawn-sentience/#sm.0000yg3xs9cmzesny6g2h7xactfq8>
- SEARLE, John. R. Minds, Brains and Programs [online]. In: *Behavioral and Brain Sciences*. Vol. 3, No. 3, str. 417-457 [cit. 10. 3. 2018]. Dostupné z: <http://courses.washington.edu/info300/searle.pdf>
- SHAPIRO, Benjamin. Universalt Truths: Cultural Myths and Generic Adaptation in 1950s Science Fiction Films [online]. In: *Journal of Popular Film and Television*. Vol. 18, No. 3, str. 103-111 [cit. 14. 2. 2018]. Dostupné z: DOI: 10.1080/01956051.1990.10662023

SCHELDE, Per. *Androids, humanoids, and other science fiction monsters: science and soul in science fiction films*. New York: New York University Press, c1993. ISBN 0814779301.

SCHNEIDER, Susan. *Science fiction and philosophy: from time travel to superintelligence*. Second edition. Malden, MA: Wiley Blackwell, 2016. ISBN 9781118922613.

SINGER, Peter. *Applied ethics*. New York: Oxford University Press, 1986. ISBN 0198750676.

SINGER, Peter Waren. *Wired for war: the robotics revolution and conflict in the twenty-first century*. New York: Penguin, 2009. ISBN 978-0-14-311684-4.

SMART, John Jamieson C. The Mind/Brain Identity Theory [online]. In: *Stanford Encyclopedia of Philosophy*. 2007 [cit. 8. 3. 2018]. Dostupné z: <https://plato.stanford.edu/entries/mind-identity/>

SMITH, Greg M. Film Structure and the Emotion System. In: *Journal of Film and Video*. Vol. 56, No. 1, 2004, str. 53-55 [cit. 18. 2. 2018]. Dostupné na JSTOR z: https://www.jstor.org/stable/20688444?seq=1#page_scan_tab_contents

SMITH, Holly. *It's Alive: A Reflection of Conscious AI in Cinema*. Dostupné z: <https://honors.uca.edu/AI/Papers/AI%20in%20Cinema.pdf>

SNELL, Marylin Berlin. *Do You Have Free Will?: John Searle reflects on various philosophical questions in light of new research on the brain*. California Magazine [online]. 2008 [cit. 17. 3. 2018]. Dostupné z: <https://alumni.berkeley.edu/california-magazine/march-april-2008-mind-matters/do-you-have-free-will>

STOLJAR, Daniel. Physicalism [online]. In: *The Stanford Encyclopedia of Philosophy*. 2001 [cit. 19. 3. 2018]. Dostupné z: <https://plato.stanford.edu/entries/physicalism/>

TOOLEY, Michael. Functional Concepts, Referentially Opaque Contexts, Causal Relations, and the Definition of Theoretical Terms [online]. In: *Philosophical Studies*.

Vol. 105, No. 3, 2001, str. 251-279 [cit. 19. 3. 2018]. Dostupné z:
<https://doi.org/10.1023/A:1010381127525>

TURING, Alan M. Computing Machinery And Intelligence [online]. In: *Mind*, Vol. 59, No. 236, 1. 10. 1950, str. 433–460 [cit. 15. 3. 2018]. Dostupné z:
<https://doi.org/10.1093/mind/LIX.236.433>.

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

WEEDMAN, Danielle, Men, Monsters and Morality: Shaping Ethics through the Sublime and Uncanny. In: *Humanities Capstone Projects* [online]. 2014, str. 1-27 [cit. 12. 3. 2018]. Dostupné z: <https://commons.pacificu.edu/cashu/20>

WEIZENBAUM, Joseph. *Computer power and human reason: from judgment to calculation*. San Francisco: W. H. Freeman, c1976. ISBN 0716704633.

ZIMMERMANN, Bernhard. Deus ex machina. In: *Brill's New Pauly: encyclopaedia of the ancient world*. Boston: Brill, 2010. ISBN 9789004122598.

Filmová tvorba

Big Hero 6 [film]. Režie Don HALL a Chris WILLIAMS. USA. Walt Disney US, 2014.

Ex Machina [film]. Režie Alex GARLAND. Veľká Británie. Universal Pictures UK, 2015.

Her [film]. Režie Spike JONZE. USA. Warner Bros. US, 2014

Chappie [film]. Režie Neill BLOMKAMP. Jihoafrická republika/USA. Sony Pictures, 2015.

Prometheus [film]. Režie Ridley SCOTT. USA/Veľká Británie. 20th Century Fox, 2012.

Robot & Frank [film]. Režie Jake SCHREIRER. USA. Bontonfilm, 2013.

The Machine [film]. Režie Caradog W. JAMES. Velká Británie. 2013.

Wall-e [film]. Režie Andrew STANTON. USA. Walt Disney US, 2008.

8. Summary

The diploma thesis presents a cognitive film theory at the beginning. This theory was used as a main method of the work as it is focused on a perspective of the spectator in a process of interpreting a film. In the next part of this work, were introduced some definitions of artificial intelligence (further as AI) and then philosophical concepts of AI. Those were presented on the core principals of the book *Artificial Intelligence: A Modern Approach* by Stuart Russell and Peter Norvig. The major aim of this thesis is to present the main streams of AI in selected film production and their subsequent comparison with the most popular science fiction movies of the last decade. In the final conclusion is outlined what topics of these movies are related to philosophical conclusions of AI and which are not. In the current film production are most explored ethical themes, primarily the impact of AI on the human life. Filmmakers are not trying to penetrate the philosophical concepts of consciousness and therefore they do not describe the essence of thinking machines. Nor why AI representatives have a desire for human love and appreciation. Thanks to the results of the film analysis we can come to conclusion that film productions are inspired by philosophical exploration rather than vice versa - philosophical exploration being an inspiration for film production. Such claims can be supported for example by the films *Ex Machina* and *Big Hero 6*, where the creators seek to bring the real image of AI closer to address contemporary philosophical problems of new technologies, and to fully inspire philosophical thought experiments. It should be taken into consideration that film production evolves together with new inventions that represent the general public. This could help people accept new technologies more easily and perhaps with a positive attitude.

9. Příloha (překlad)

„Film je zprostředkovanou reprezentací světa, tak jak ho zná filmový tvůrce. Jedná se o audiovizuální reprezentaci uspořádanou v řetězci snímků, sekvencí a zvuků. Každá jednotka, záběr nebo sekvence přispívá ke strukturování určitého pocitu nebo poselství, které film zamýšlí zastupovat. Vyvolává potěšení, pocit a občas vyvolává intelektuální zamyšlení diváka, protože příjemce se odráží ve filmu jako kognitivní agent.“

„Film is a mediated representation of the world as cognised by a filmmaker. It is an audio-visual representation organised into a chain of shots, sequences, and sounds. Every unit, shot or sequence, contributes to structuring the feel or message that the film intends to represent. It elicits pleasure, feeling, and sometimes invokes intellectual thought in a receptor's mind, as the receptor reflects on the film as a cognitive agent.“

PERVEZ, A. Cognitive Film Theory.

„My nazýváme sami sebe Homo sapiens- člověk moudrý, protože naše inteligence je pro nás velice důležitá. Po tisíce let jsme se snažili pochopit, jak myslíme; to znamená, jak jen hrstka hmoty může vnímat, pochopit, předvídat a manipulovat s mnohem větším a složitějším světem než jsme my sami. Pole umělé inteligence, nebo UI, jde ještě dále: snaží se nejen pochopit, ale také budovat inteligentní entity.“

We call ourselves Homo sapiens – man the wise – because our intelligence is so important to us. For thousands of years, we have tried to understand how we think; that is, how a mere handful of matter can perceive, understand, predict, and manipulate a world far larger and more complicated than itself. The field of artificial intelligence, or AI, goes further still: it attempts not just to understand but also to build intelligente entities.

RUSSELL, S. J., NORVIG P. Artificial Intelligence: A modern approach

„Vy trváte na tom, že existuje něco, co stroj nedokáže udělat. Když mi přesně řeknete, co to je, co stroj nedokáže udělat, tak já pokaždé

„You insist that there is something a machine cannot do. If you will tell me precisely what it is that machine cannot do, then I can always make

můžu donutit stroj udělat přesně tuto věc.“

a machine which will do just that.“

HALES, Colin G., The Revolutions of Scientific Structures

„Umělá inteligence je aktivita věnovaná vytváření inteligentních strojů a inteligence je taková kvalita, která umožňuje entitě správně fungovat a také předvídat ve svém prostředí. V návaznosti na tuto definici, hodně věcí – lidé, zvířata a některé stroje vlastní inteligenci. Stroje, jako jsou ‚chytré kamery‘, a spousta zvířat jsou v tomto primitivním konci rozšířeného kontinua, podle kterého jsou uspořádány entity na různých stupních inteligence.“

„Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment. According to that definition, lots of things – humans, animals, and some machines – are intelligent. Machines, such as “smart cameras,” and many animals are at the primitive end of the extended continuum along which entities with various degrees of intelligence are arrayed.“

NILSSON, N. J. The Quest for Artificial Intelligence: A History of Ideas and Achievements

“Umělá inteligence sledovaná v kultu výpočetní techniky není ani duchem šancí na dosažení trvalých výsledků.”

„Artificial intelligence pursued within the cult of computationalism stands not even a ghost of a chance of producing durable results.“

RUSSELL, S. J., NORVIG P. Artificial Intelligence: A modern approach

„Být milý, vynalézavý, krásný, přátelský, mít iniciativu, mít smysl pro humor, odříkat si od špatného, dělat chyby, zamilovat se, užít si jahody a krém, někoho milovat, učit se ze zkušeností, používat slova správně, být

„Be kind, resourceful, beautiful, friendly, have initiative, have a sense of humor, tell right from wrong, make mistakes, fall in love, enjoy strawberries and cream, make someone fall in love with it, learn from experience, use words

předmětem své vlastní myšlenky, mít tolik rozmanitostí chování jako člověk, dělat něco opravdu nového.“ *properly, be the subject of its own thought, have as much diversity of behavior as man, do something really new.*“

TURING, A. M., Computing Machinery and Intelligence, In: *Mind*

„Je to spíše jako by se filosofové prohlašovali za experty na vysvětlování method jevištních kouzelníků, a když se zeptáme, jak kouzelníci dělají ten trik s rozpůlením ženy, oni řeknou, že je to velice zřejmé: ten kouzelník ji doopravdy nerozpůlí; on jednoduše udělá trik, aby se nám to tak zdálo. ‚Ale jak to udělá?‘ se zeptáme. ‚Ne naše oddělení‘ řeknou filosofové.“ *„It is rather as if philosophers were to proclaim themselves expert explainers of the methods of stage magicians, and then, when we ask how the magician does the sawing-the-lady-in-half trick, they explain that it is really quite obvious: the magician doesn't really saw her in half; he simply makes it appear that he does. “But how does he do that?” we ask. “Not our department,” say the philosophers.*“

RUSSELL, S. J., NORVIG P. *Artificial Intelligence: A modern approach*

"Nikdo nepředpokládá, že počítačová simulace bouře způsobí, že budeme všichni mokří... Proč by někdo v nejlepší míněni předpokládal, že počítačová simulace duševních procesů má skutečně duševní procesy?" *„No one supposes that a computer simulation of a storm will leave us all wet... Why on earth would anyone in his right mind suppose a computer simulation of mental processes actually had mental processes?“*

SEARLE, J. R. Minds, brains, and programs. In: *Behavioral and Brain Sciences*

"Ale člověk by mohl vytvořit stejný argument o mozku: stačí se podívat na tuto sbírku buněk (nebo atomů), slepě fungujících podle zákonů *„But one could make the same argument about the brain: just look at this collection of cells (or of atoms), blindly operating according to the*

biochemie (nebo fyziky) – kde zde najdeme mysl? laws of biochemistry (or of physics)- what's there to be a mind? Why can a hunk of brain be kus jater nemůže? To zůstává velkým a mind while a hunk of liver cannot? That tajemstvím."

RUSSELL, S. J., NORVIG P. *Artificial Intelligence: A modern approach*

"Nechci udělat dojem, že si myslím, že neexistuje žádné tajemství o vědomí... Ale nemyslím si, že tyto záhady nutně potřebují být vyřešeny dříve, než budeme schopni odpovědět na otázku, s níž jsme se v tomto článku zabývali."

RUSSELL, S. J., NORVIG P. *Artificial Intelligence: A modern approach*

"Pracovní týden byl od přelomu století snížen o 50 procent. Není zcela mimo, když budeme předvídat, že se do roku 2000 opět sníží na polovinu."

RUSSELL, S. J., NORVIG P. *Artificial Intelligence: A modern approach*

"Ultrainteligentní stroj můžeme definovat jako stroj, který může daleko předstihnout veškeré intelektuální činnosti každého člověka, jakkoli je člověk chytrý. Vzhledem k tomu, že konstrukce strojů je jednou z těchto intelektuálních aktivit, může ultrainteligentní stroj navrhnout ještě lepší stroje; pak by byla nepochybně "intelligenční exploze" a inteligence člověka by zůstala daleko za nimi. První ultrainteligentní stroj je tedy posledním vynálezem, který člověk vůbec

potřebuje, za předpokladu, že je stroj dostatečně poslušný, aby nám řekl, jak ho udržet pod kontrolou. „*provided that the machine is docile enough to tell us how to keep it under control.*“

GOOD, I. J. Speculations concerning the first ultraintelligent machine, In: *Advances in Computers*

"Singularita nám umožní překonat tato omezení našich biologických těl a mozku. Získáme moc nad našimi osudy. Naše úmrtnost bude v našich rukou. Budeme schopni žít tak dlouho, jak budeme chtít (jemně odlišné prohlášení, že budeme žít navždy). Plně pochopíme lidské myšlení a výrazně rozšíříme a zvětšíme jeho dosah. Do konce tohoto století bude nebiologická část naší inteligence silnější, než jen lidská inteligence."

„The Singularity will allow us to transcend these limitations of our biological bodies and brain. We will gain power over our fates. Our mortality will be in our own hands. We will be able to live as long as we want (a subtly different statement from saying we will live forever). We will fully understand human thinking and will vastly extend and expand its reach. By the end of this century, the nonbiological portion of our intelligence will be trillions of trillions of times more powerful than unaided human intelligence.“

KURZWEIL, Ray. *The singularity is near: when humans transcend biology*

"V těchto filmech svět vypadal jako hrozivý, tekutý, chaotický, neosobní, složený ze sil, které jsme zřídka kdy pochopily a rozhodně nikdy neovládaly. Strach je zaměřen na neznámé, neviditelné hrůzy, které se skrývají pod povrchovou normálností."

"V těchto filmech svět vypadal jako hrozivý, tekutý, chaotický, neosobní, složený ze sil, které jsme zřídka kdy pochopily a rozhodně nikdy neovládaly. Strach je zaměřen na neznámé, neviditelné hrůzy, které se skrývají pod povrchovou normálností."

SHAPIRO, B. Universal Truths: Cultural Myths and Generic Adaptation in 1950s Science Fiction Films. In: *Journal of Popular Film and Television*