

**UNIVERZITA KARLOVA**

**FAKULTA SOCIÁLNÍCH VĚD**

Institut Sociologických studií

Katedra Sociologie

**Sociologické aspekty umělé inteligence**

Bakalářská práce

Autor práce: Tomáš Zálešák

Studijní program: Sociologie a sociální antropologie

Vedoucí práce: Mgr. Jiří Bureš

Rok obhajoby: 2019

## **Prohlášení**

1. Prohlašuji, že jsem předkládanou práci zpracoval samostatně a použil jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne

Tomáš Zálešák

## **Bibliografický záznam**

Zálešák, Tomáš. 2019. *Sociologické aspekty umělé inteligence*. Bakalářská práce (Bc). Univerzita Karlova, Fakulta sociálních věd, Institut Sociologických studií. Katedra Sociologie. Vedoucí diplomové práce Mgr. Jiří Bureš.

**Rozsah práce:** 69 125 znaků s mezerami

## **Anotace**

Tato práce je přehledovou studií literatury, která má za cíl představit téma umělé inteligence (AI) jako sociologicky relevantní a plodné pro další výzkum. Proto má práce tři kapitoly, které společně tvoří základ pro pochopení různých aspektů tématu. První kapitola se zabývá technickou stránkou AI na základě technické literatury. Tato literatura je za prvé důležitá pro pochopení příčin nebývalého nárůstu zájmu o AI v několika posledních letech. Za druhé je také důležitá pro pochopení fungování současné AI a jejích slabých a silných stránek v dalších kapitolách. Druhá kapitola představuje přehled sociologicky relevantních prací autorů, kteří se různými aspekty AI a jejím sociálním dopadem věnují. Z tohoto přehledu jsou vybrány práce čtyř autorů, jako příklady širší možnosti nahlížení na AI a také jako zdroje pro další diskusi. Poslední třetí kapitola představuje skrze oborovou literaturu příklady dvou ze široké řady oblastí, kde existují snahy AI v současné době aplikovat. Tato kapitola vyvažuje předchozí teoretické kapitoly pohledem na situaci v oblastech medicíny a transportu, a aplikuje poznatky z předchozích dvou kapitol. Výsledkem práce je jakýsi průvodce tématem sociologických aspektů AI pro sociology, které oblast AI zajímá nebo by ji chtěli dále zkoumat. Je to zároveň také argument pro to, proč může být AI vůbec zajímavý jev k sociologickému zkoumání.

## **Annotation**

This text is a literature review with the goal of introducing the topic of artificial intelligence (AI) as a sociologically relevant phenomenon to research. The text has three chapters that together form a basic platform to understand many different aspects of the topic. The first chapter is concerned with the technical aspects of AI and is based on technical literature. These aspects are important for understanding of the current surge of interest in AI that has arisen in the recent years. The technical aspects also help us understand the strengths and weaknesses of AI by understanding its inner workings. The second chapter introduces an overview of sociologically relevant works of authors dealing with the social effects and other aspects of AI. Works of four of these authors are chosen as examples of the breadth of possible ways of looking at AI and as sources for further understanding. The third chapter chooses two examples from the vast amount of areas of application AI in the real world to deepen the overall understanding. This chapter balances out the previous more theoretical chapters by looking at the state of the art in the fields of medicine and transportation and applies knowledge acquired

in the previous chapters. The output of the text is a sort of introductory guide through the topic of sociological aspects of AI for those sociologists who are interested in AI or would wish to conduct further research. It is at the same time conceived as an argument why AI can be an interesting focus of sociological research.

## **Klíčová slova**

Umělá inteligence, automatizace, průmysl 4.0, společnost 4.0, technologie, digitální sociologie, AI

## **Keywords**

Artificial intelligence, automatization, industry 4.0, society 4.0, technology, digital sociology, AI

## **Title**

Sociological aspects of artificial intelligence

## **Poděkování**

Zde bych rád poděkoval Mgr. Jiřímu Burešovi za *nejpohodovější* vedení mé práce a Bc. Veronice Vondráškové za její formátovací a gramatické čarování.

# Obsah

Úvod.....	2
1. Objasnění konceptu AI.....	5
1.1. Sociální realita AI versus AI sci-fi.....	5
1.2. Definice AI.....	6
1.3. Strojové učení, hluboké učení a Neuronové sítě.....	8
2. A jak to vidí sociologie?.....	13
2.1. Konverzační AI.....	13
2.2. AI a reprodukce nerovností.....	14
2.3. Dva pohledy na AI a práci.....	14
2.3.1. Přehledová tabulka sociologických autorů a tematizací AI.....	19
3. Dopady AI ve dvou vybraných oblastech.....	20
3.1. AI v medicíně.....	20
3.2. AI a autonomní vozidla.....	23
Závěr.....	29
Shrnutí.....	29
Diskuse.....	30
Summary.....	31
Použitá literatura.....	34
Teze bakalářské práce.....	39
Tabulka 1 Přehled rozebíraných sociologicky relevantních prací o AI.....	19

# Úvod

Francouzský prezident Emmanuel Macron ve své osmadvaceti-jazyčné vizi pro nové směřování EU před volbami do Evropského parlamentu na jaře 2019 mluví o umělé inteligenci (dále jen AI) jako o průlomové technologii, která je stěžejní částí budoucnosti Evropy [Macron 2019]. Čínská vláda chce v rámci svého plánu „New Generation Artificial Intelligence Development Plan” investovat v přepočtu 150 miliard amerických dolarů a stát se do roku 2030 vedoucím hráčem ve vývoji AI [AI Policy – China 2019]. I americká vláda chce skrze iniciativu „American AI Initiative” vydanou na začátku roku 2019 posílit vývoj a aplikaci AI „prioritizací výzkumu a vývoje AI” [Trump 2019]. Evropská komise vydala v prosinci 2018 „Koordinovaný plán k AI”, kterým chce „prohloubit a zefektivnit spolupráci” v rámci AI výzkumu a do roku 2020 investovat alespoň 20 miliard EUR [Evropská komise 2018]. Těchto pár vybraných příkladů ukazuje na fakt, že koncept AI není jen fikcí z oblasti sci-fi, ale důležitou součástí politického diskurzu a vizi, kam se mají společnosti vyvíjet. Pozorujeme tedy výrazné rysy nové vlny modernizace společnosti, jejíž aspekty stojí za zkoumání.

V reakci na tento plán si i česká vláda nechala zpracovat odbornou studii na téma AI v Česku. Tato studie dochází k závěru, že „ČR patří mezi země s největším očekávaným dopadem automatizace a technologií využívajících AI (...) ČR však nemá pro zavádění technologií AI příliš dobré výchozí podmínky. (...) Většina domácích firem se nachází na konci hodnotových řetězců (GVC) nebo je pod zahraniční kontrolou...” [Fařun et al. 2018: 6]. V neposlední řadě také zmiňuje nutnost zajistit etickou stránku AI, u které zdůrazňuje primárně právní rámec, ale v této oblasti by jistě mohly vhodnými pomocníky být i sociologie a další sociální vědy [ibid.: 9]. Vzestup AI je součástí širších změn společnosti a důležitým, i když ne jediným, článkem procesů automatizace a „čtvrté průmyslové revoluce”.

Stejně jako sociální vědy mohou být významné pro AI, tak může být AI významná pro sociální vědy. A to rovnou dvojím způsobem. Za prvé může být vhodným předmětem studia (jako v této práci) s potenciálem být jednou z technologií, která v blízké budoucnosti nejvíce změní společnost. Za druhé ale AI také může být mocným metodickým nástrojem



sociálních vědců. V rovině okamžité současnosti se nabízí užití v rámci analýzy dat, ale při pohledu dále by mohlo uplatnění sahat i do dalších oblastí výzkumu. Důkazem obou možností může být existence center studujících a interdisciplinárně uplatňujících AI u sociálněvědních fakult mnoha světově věhlasných univerzit jako je *Future of Humanity Institute* na Oxfordské univerzitě, *Ethics and Governance of Artificial Intelligence Fund* ve spolupráci MIT a Harvardské univerzity nebo právě ono metodologické využití AI v oddělení *Computational Social Science* Cornellské univerzity.

AI je tedy téma, které rezonuje jak v nejvyšších patrech globální politiky, tak v českém politickém kontextu i společenských vědách. Politici AI zařazují jako stěžejní část svých vizí budoucnosti a slibují nemalé investice. Vznikají vědecká centra a vychází mnoho prací z různých vědních oborů, které se snaží po svém srovnat se změnami, které AI přináší a přinášet bude. To vše mluví ve prospěch braní AI vážně jak v současnosti, tak v budoucnosti, kdy se začne ukazovat, do jaké míry se sliby politiků a předpovědi vědců skutečně naplňují.

V této práci nabízím výběr poznatků užitečných pro pochopení diskurzivního a institucionálního vzestupu AI v politice a společenských vědách. Zároveň také naznačuji body pro budoucí sociologické reflexe různých aspektů tohoto širokého tématu. AI je v současnosti působící fenomén, který je předmětem zájmu mnoha politických, ekonomických i akademických aktérů. V této práci se pokusím alespoň částečně odpovědět na otázku, co jsou sociologické aspekty AI. K tomu patří dílčí otázky, týkající se technického fungování AI, a o co vlastně tedy jde, nebo otázka jakými způsoby může zkoumání AI uchopovat sociologie. Na tyto otázky se pokusí následující práce částečně odpovědět nebo případně alespoň nabídnout určité objasnění v rámci možností mé práce. V první kapitole bude nejprve vysvětleno, co přesně si v současnosti pod pojmem AI představít. Budou popsány základy toho, jak AI funguje, a proč se o ní začalo najednou tolik mluvit. Ve druhé kapitole pak budou představeny sociologické práce, které se tématu AI věnují, aby byla naznačena širší možných sociologických náhledů na toto téma. V poslední kapitole bude popsáno na příkladě dvou vybraných oblastí, ve kterých se AI aplikuje, jaké dopady v těchto oblastech v současnosti můžeme pozorovat. K tomu budu používat jak literaturu sociologickou, tak velké množství literatury nesociologické, která bude sloužit jako sonda do široké oblasti vývoje AI a do dvou vybraných oblastí aplikace AI. To by mělo pomoci čtenáři lépe porozumět technickým aspektům AI a její aplikace v různých oblastech,

a vytvořit znalostní základ pro sociologické zkoumání. Tyto poznatky by měly pomoci vytvořit základy pro efektivní zkoumání AI a umožnit porozumění proměnám společnosti, které AI přináší.

# 1. Objasnění konceptu AI

## 1.1. Sociální realita AI versus AI sci-fi

Na začátku, dokonce ještě před definováním, jak přesně bude používán pojem AI v této práci, a co tento pojem zahrnuje, stojí za to poukázat na mnohost tohoto pojmu a pokusit se ji částečně projasnit. Diskurz myšlení a mluvení o AI je silně zakořeněný v žánru science fiction. V knihách, filmech, hrách i dalších médiích tohoto žánru se můžeme setkat s AI v různých podobách. Může mít podobu humanoidních robotů, v podstatě lidí v umělých tělech. Může být také zcela mimotělní, jen lidské nebo nadlidské virtuální vědomí nebo například stroj stvořený k pomoci lidstvu, který se rozhodne ho místo toho ovládnout a zničit. Tohle ale není každodenní realita. Roboti mezi námi jsou spíše „chytré asistenty“ v telefonech než pseudo-lidské bytosti. Hlasoví asistenti v telefonech rozumí slovům, ale ne jejich významům a lidstvo zatím nebylo ovládnuto ani zničeno. Co tedy je zde k sociologickému zkoumání?

V této práci je navrženo, že sice nežijeme v budoucnosti známé ze science fiction, ale i tak je ve společnosti v posledních letech v čím dál větší míře přítomná nová kategorie technologií. Technologií, které sice nemají vědomí, jak ho známe u lidí, ale jsou schopné se učit, vytvářet vlastní způsoby dosahování zadaných cílů a úspěšně plnit úkoly, které byly dosud považované za výhradně lidské (a jak si ukážeme dále, dosahuje AI v těchto úkonech dokonce někdy lepších výsledků než lidé). Je zde tedy nová kategorie technologií – AI – která začíná být úspěšně aplikovaná v různých oblastech společnosti a tyto oblasti různými způsoby proměňuje. Seznam politických aktérů a jejich investic do AI v úvodu práce měl sloužit jako poukázání na široké společenské uznání reality těchto změn. Lze také připomenout cyklický vztah mezi uznáním těchto změn a jejich postupem. Když velcí aktéři vidí AI jako důležitou oblast a investují do ní proto zdroje, zajišťují tím zároveň, že se důležitou opravdu stane a její dopady budou díky přísunu investic ještě větší. Také proto je to úrodná půda pro sociologické zkoumání.

Jde zde o nový jev, který reálnými a pozorovatelnými způsoby proměňuje společnost, a to jsou právě ty sociální aspekty AI, tedy to, co můžeme zkoumat. Ne „roboty“ nebo „bytosti“

ale proměny ve vztazích, v práci, v odpovědnosti a ve všech dalších oblastech společnosti. Abychom tyto změny pochopili a byli na ně schopni adekvátně reagovat, je potřeba porozumět AI, prozkoumat, jak se projevuje v různých oblastech společnosti a propojit to s naším dosavadním věděním. Právě toto propojení nového věděním se starým je úkolem pro sociology, kteří mají tu možnost stát tak trochu stranou a neřešit tolik, jestli AI přinese zisk jejich firmě nebo moc jejich státu. Projevy AI jistě budou v mnohém překvapivé, ale jistě v mnohém budou také starými známými způsoby nerovnoměrně ovlivňovat sociální stratifikaci, odcizení, prostředky sociální kontroly, prostředky emancipace a individualizace i další jevy, které jsme zvyklí zkoumat. Proto stojí za to zkoumat AI. Protože je co zkoumat.<sup>1</sup>

## 1.2. Definice AI

O potížích stabilně sociologicky definovat AI se už v roce 1985 v jedné z nejcitovanějších sociologických prací o AI zmiňuje Steve Woolgar. Píše o problému toho, že pokaždé, když vytvoříme něco, co vykazuje inteligentní chování, tak se latka toho, co ještě považujeme za inteligentní posune o něco dál [Woolgar 1985: 564]. Protože fungování výtvaru rozumíme, stává z něj „jen“ technologie/algorithmus/stroj a ne něco „pravě“ inteligentního. AI se tak může zdát být „těsně za rohem“ vlastně navždy.

V současnosti ve své práci z roku 2018 Jakub Mlynář a další autoři také hovoří o potřebě zobecnitelného sociologického rámce pro studium AI [Mlynář et al. 2018: 4]. V první řadě kritizují tendenci některých sociologů nahlížet na AI pouze jako na metodu analýzy dat, a ne jako na sociální jev [ibid.: 1-2]. Rekapituluji pak několik případů lepších sociologických pojetí AI z konce 20. století, ale i o nich říkají, že nejsou dostačující právě z důvodu zmiňované nemožnosti zobecnění [ibid.: 4]. Protože se jedná pouze o součást výstupu z konference, nenabízí autoři vlastní definici AI. Jejich ideálem je definice, která by na AI nahlížela jako na sociální jev i jako na ne-lidského sociálního aktéra [ibid.: 8].

---

<sup>1</sup> Používání termínu AI může být složité i v rámci diskurzu, který se zabývá AI jako současnou technologií, nikoli sci-fi. Lidé, kteří AI vytvářejí (programátoři, vědci) obvykle slovem AI nazývají spíše svůj obor a produkty své práce (tedy to čemu říkáme AI my zde) nazývají jako AI algoritmy, AI systémy, algoritmy strojového učení a podobně. V laických médiích je ale běžné mluvit o produktech (algoritmech, systémech, ...) tohoto oboru jako o AI, protože je to intuitivní a všezahrnující pro čtenáře. Situace ohledně toho, co pojem AI znamená je tedy matoucí ve více rozměrech, a je užitečné se vždy zaměřit na kontext pro správné pochopení. V této práci bude AI znamenat výrobek oboru, a budu se snažit, aby užití pojmů bylo vždy jednoznačné i z kontextu.

V rámci sociologie tedy neexistuje jednotná definice AI vhodná pro tuto práci, a problém s definicemi AI je zároveň očividně problémem, který už nějakou dobu sociology snažící se psát o AI trápí. Z tohoto důvodů si půjčím dvě definice z řad technické literatury, které podle mě zároveň postihují AI způsobem vhodným pro sociologické zkoumání, a zároveň se nevzdalují příliš od reality těch, kteří AI vytváří (jako to dle mého názoru hrozí v případě návrhu Mlynáře et al.).

První část definice AI v této práci bude založena na té z knihy „Artificial Intelligence: What Everyone Needs to Know” od Jerryho Kaplana. Ten v ní definuje inteligenci (nejen tedy umělou) jako „schopnost včas dělat přiměřená zobecnění na základě omezeného množství dat”<sup>2</sup> [Kaplan 2016: 5]. Tato část definice postihuje hlavně výsledky jednání AI. Jde o to, že AI musí být schopná dělat vhodná rozhodnutí (tedy „zobecnění“), a to i na základě neúplné znalosti všech detailů situace.

Druhá část mé definice AI bude pocházet z knihy „Deep Learning” od Joshuy Bengia, Aarona Courvilla a Iana Goodfellowa. Ti definují „schopnost získat jejich vlastní znalosti vyvozením vzorců z nezpracovaných dat”<sup>3</sup> jako strojové učení [Bengio, Courville, Goodfellow 2016: 2]. Tato část postihuje důležitou část AI, tedy to že se musí být schopna učit a sama zlepšovat. Jak je podrobněji vysvětleno dále, v podstatě všechny současné snahy vytvářet AI jsou založeny na strojovém učení, a jedná se tak o vhodnou složku pro vytvoření definice současné AI.

Definice AI v této práci je spojením těchto dvou definic. AI je tedy pro mě taková technologie, která je schopná dělat vhodná rozhodnutí (ve smyslu „přiměřených zobecnění na základě omezeného množství dat”) a sebezlepšování skrze učení na základě předchozí zkušenosti (tedy zmiňovaná „schopnost získat jejich vlastní znalosti vyvozením vzorců z nezpracovaných dat”).

---

<sup>2</sup> „the essence of intelligence (...) is the ability to make appropriate generalizations in a timely fashion based on limited data” [Kaplan 2016: 5]. Z anglického originálu přeloženo mnou, stejně jako ve všech ostatních případech, kdy cituji slova z prací v angličtině.

<sup>3</sup> „the ability to acquire their own knowledge, by extracting patterns from raw data. This capability is known as machine learning” [Bengio, Courville, Goodfellow 2016: 2]

### 1.3. Strojové učení, hluboké učení a Neuronové sítě

Co si tedy máme pod pojmem AI představit (nebo naopak nepředstavit) a jak bude užíván v této práci? O AI se historicky často mluvilo v souvislosti s technologickým pokrokem. Objeví se nová technologie a s ní i vize, že budou vytvořeny stroje nebo počítače, které se budou chovat jako lidé, a budou schopné lidi nahrazovat. Po takové vlně zájmu také zatím vždy došlo i k nárazu na realitu, „splasknutí bubliny“<sup>4</sup> a nenaplnění vize humanoidní umělé inteligence. Reálné dopady technologického pokroku měly tedy často mnohem skromnější podobu automatizace jednoduchých úkonů. Ale i když taková automatizace rutinních úkonů nezní tak dramaticky jako humanoidní AI, tak její dopady na společnost mohou být často dalekosáhlé.

Současná vlna zájmu o AI se začala zvedat v roce 2006. V návaznosti na práci Geoffreyho Hintona „A Fast Learning Algorithm for Deep Belief Nets“ začal současný boom vytváření AI s užitím strojového učení a hlubokého učení [Bengio, Courville, Goodfellow 2016: 18-19]. Právě strojové učení a hluboké učení jsou koncepty klíčové pro pochopení toho, co AI v současnosti je, a proto budou v následujících odstavcích krátce vysvětlené.

#### 1.3.1. Strojové učení

Strojové učení (machine learning) je proces, kterým program vytváří a mění sám sebe. Mluvíme zde tedy o způsobu vytváření programů a může být snazší ho vysvětlit na zjednodušeném příkladu. Představme si, že člověk chce vytvořit program, který rozpozná lidské tváře v obrázku. Standardní „neAI“ způsob by byl, že programátor vymyslí a sepíše různá pravidla, kterými program pak bude poznávat, co je a co není tvář. Mohlo by to tedy vypadat tak, že člověk se pokusí koncept tváře „přeložit“ do jednoduchého pravidla, jako například „když jsou v obrázku dva kulaté shluky pixelů o určité velikosti a určitých barvách, tak zvýš pravděpodobnost, že tam je tvář“. Takové pravidlo by se tedy (extrémně zjednodušeně) pokoušelo hledat v obrázku oči (kulaté shluky pixelů). Podobně si pak lze představit další pravidla, která se pokouší popsat ostatní části tváře. Tento proces by byl velmi náročný, protože by bylo potřeba myslet na všechny různé možnosti, jak může tvář vypadat (tvar, barva pleti, brýle, vlasy, vousy, výrazy ve tváři, ...), ale také na věci, které by mohly program zmást (pro člověka je rozpoznat tvář velice snadné, ale je potřeba myslet

---

<sup>4</sup> Mezi lidmi z oboru se takovým obdobím říká „zimy AI“. Je to jméno inspirované nukleární zimou, jde o „zmrazení“ zájmu o AI jako důsledek přehnaných očekávání. V minulosti došlo zatím ke dvěma takovýmto „zimám“ a neexistuje konsensus o tom, jestli nás brzy nečeká další [Lim 2018].

na to, že v rovině abstraktních pravidel o tvarech a proporcích pixelů může jako tvář s očima a ústy vypadat i například elektrická zásuvka nebo přední část automobilu). Důležité každopádně je, že člověk má nějaký cíl, vymyslí jak tento cíl „přeložit“ do pravidel srozumitelných programu, a tento program pak podle těchto pravidel koná.

Popsaný zjednodušený příklad byl tedy „neAI“ způsob vytváření programů. Nyní se konečně můžeme podívat na strojové učení. Použijeme stejný příklad jako výše, tedy budeme chtít program, který rozpozná v obrázku lidskou tvář. Strojové učení funguje tak, že programu dáme mnoho dat (řádově až desítky milionů obrázků), a on si sám vytvoří na jejich základě pravidla. V našem příkladu bychom tedy dali programu databázi obrázků, na kterých víme, že jsou lidské tváře. Program by se pak na základě podobností a společných částí obrázků pokusil vytvořit pravidla, která co nejlépe popisují obrázky, které mu byly dány. Člověk má tedy cíl (rozpoznávat tváře) a výsledek (už rozpoznané obrázky s tvářemi, trénovací soubor dat). Program si na základě trénovacích dat vytvoří pravidla, která by mu ideálně měla umožnit při setkání s neoznačeným obrázkem rozhodnout, s jakou pravděpodobností odpovídá trénovacímu souboru dat. Všimněme si, že program všechny své znalosti o tom, co je „tvář“ přebírá z trénovacího souboru obrázků tváří. I když si tedy z těchto obrázků vytváří pravidla pro definici tváře program sám, je zcela závislý na tom, jak koncept tváře vymezil autor trénovacího souboru (buď člověk nebo další program).<sup>5</sup>

### 1.3.2. Hluboké učení

Hluboké učení (deep learning) je způsob, jak vytvářet AI pomocí strojového učení. Nejedná se o jediný způsob, jak aplikovat strojové učení<sup>6</sup>, ale podle autorů knihy Deep learning jde o způsob „mocný a flexibilní“ [Bengio, Courville, Goodfellow 2016: 8]. Stojí na něm drtivá většina současné tvorby v oblasti AI. Hluboké učení funguje tak, že se algoritmus učí rozumět datům skrze vrstvy s různou úrovní abstrakce [ibid.]. Lépe tomu lze porozumět znovu na příkladě. Vezměme si jako u strojového „nehlubokého“ učení příklad rozpoznávání tváří v obrázku. Opět tedy máme datový soubor obrázků s tvářemi, a chceme, aby se program naučil tváře rozpoznávat. Při hlubokém učení proces funguje tak, že každá vrstva porozumí obrázku o něco více, s využitím znalostí z předchozí vrstvy. První vrstva

---

<sup>5</sup> AI na základě strojového učení byla mimo jiné využita k napsání akademické knihy, kterou poté vydalo seriózní vydavatelství Springer. Jedná se mimochodem o přehled literatury, asi 53 000 článků o bateriích, tedy podobný žánr jako moje práce [Writer 2019].

<sup>6</sup> Každé hluboké učení je tedy zároveň strojovým učením, ale ne každé strojové učení musí být hluboké.

jsou jednotlivé pixely obrázku samotného. Ve druhé pak program může na základě pixelů identifikovat přechody mezi oblastmi pixelů. Ve třetí pak díky tomu, že ví, kde jsou přechody dokáže identifikovat tvary. Ve čtvrté pak díky tvarům identifikuje objekty. V páté vrstvě pak srovná nalezené objekty se znalostmi z předchozích a dat rozhodne, s jakou pravděpodobností se může jednat o tvář [ibid.: 6]. V několika krocích mezi vrstvami se tak dostáváme od jednotlivých pixelů přes tvary k rozpoznání objektů.

Příklad uvedený u „nehlubokého“ strojového učení byl užitečnější pro pochopení mechanismu strojového učení, a reálně by pravděpodobně nebyl schopen efektivně rozpoznávat tváře. Příklad uvedený zde u hlubokého učení už je mnohem bližší realitě (i když stále samozřejmě zjednodušené) a podobně fungují i algoritmy rozpoznávající tváře používané v současnosti.

V rámci kapitoly o hlubokém učení je také na místě zmínit koncept tzv. neuronových sítí. Obecně jde o vytváření algoritmů inspirovaných biologickým mozkem [ibid.: 13]. V současnosti jde o pojem z velké části se překrývající právě s konceptem hlubokého učení. Odborná literatura mluví spíše o omezených možnostech inspirace mozkem pro vytváření AI, z jednoduchého důvodu, že samotnému fungování mozku dostatečně nerozumíme. „Jsme schopni z neurovědy získat určitá hrubá vodítka. Základní myšlenka mnoha výpočetních jednotek, které se stávají inteligentními jen skrz interakce mezi sebou je inspirovaná mozkem. (...) i když neurověda inspirovala několik architektur neuronových sítí, zatím nevíme dostatek o biologickém učení na to, abychom je využili v učících algoritmech.”<sup>7</sup> [ibid.: 16]. Můžeme si všimnout podobnosti s předchozím popisem fungováním hlubokého učení. Ekvivalenty výpočetních jednotek, které se stávají inteligentními v mozku, jsou u hlubokého učení vzájemně interagující vrstvy algoritmu. Obecně lze pojmy neuronová síť a hluboké učení považovat většinou za zaměnitelné.

### 1.3.3. Co z toho můžeme vidět

Už z tohoto obecného popisu základních principů AI lze lépe pochopit mnoho jejích klíčových charakteristik a omezení. Například fakt, že se současná vlna zájmu zvedla zrovna

---

<sup>7</sup> „We are able to draw some rough guidelines from neuroscience. The basic idea of having many computational units that become intelligent only via their interactions with each other is inspired by the brain. (...) while neuroscience has successfully inspired several neural network architectures, we do not yet know enough about biological learning for neuroscience to offer much guidance for the learning algorithms we use to train these architectures.” [Bengio, Courville, Goodfellow 2016: 16]



v nedávné minulosti. To lze chápat jako událost umožněnou technologickým vývojem. Roli hraje dostupnost zařízení schopných záznamu a uchování velkého množství dat (mnoho lidí má dnes telefony s možností záznamu vizuálních, zvukových i textových dat). Dále dostatečně rychlý a rozšířený internet, který umožňuje tato data shromažďovat. A poté samotný hardware počítačů, který musí být dostatečně výkonný na to, aby zpracoval velká množství často objemných dat. Kombinovaná přítomnost těchto technologických faktorů umožňuje existenci AI tak, jak ji dnes pozorujeme, například právě v podobě zmiňovaného rozpoznávání objektů v obrázcích. Jistě by nebylo správné na současnou vlnu zájmu o AI nahlížet jako na něco nevyhnutelného, determinovaného *pouze* dostupností technologie. Zároveň nám ale kontextualizace skrze tento technologický vývoj umožňuje lépe pochopit rozsah současného zájmu i jeho působení v určitých oblastech spíše než v jiných.

Dalším zajímavým aspektem je názvosloví současné vlny AI. Že se učení nazývá zrovna *hluboké*, a sítě zrovna *neuronové* jistě není náhoda. Lze jen spekulovat o skutečném efektu vhodného brandingů na velikost současné vlny zájmu o AI, ale slova jako *hluboké* a *neuronové* jistě nemají negativní konotaci v očích laické veřejnosti ani investorů. Názvy slibující hloubku učení a podobnost mozku vytváří ale také riziko vyvolání přehnaných očekávání (viz poznámka pod čarou číslo 3 o „zimách AI“ výše). Je třeba zdůraznit, že hloubka hlubokého učení odkazuje k vrstvám učícího algoritmu, nikoli k hloubce poznání samotného. Je snadné si představit problémy s chápáním tohoto relativně jemného rozdílu v použití slova hloubka v médiích nebo mezi laiky.

O nic lehčí tento problém nedělá ani fakt, že strojové učení vede k poznání, které není pro lidi vždy zrovna intuitivní. Program, který se naučí rozpoznávat dopravní značky vykazuje chování, které naznačuje pochopení toho, co je dopravní značka. Zároveň ale i malá změna, jako je nálepka na dopravní značce ho může zmást a zabránit ve správné identifikaci dopravní značky [Eykholt et al. 2018: 6]. Zde se nachází ona kontraintuitivnost. U lidského chápání je běžné, že pochopení konceptu jde pozorovat z lidského chování (tedy kdo vykazuje známky pochopení nejspíš opravdu chápe). U algoritmu to ale tak jednoznačné není. Může se chovat, jako by koncept chápal, ale tomu tak může být i jen z části. Asi těžko bychom o člověku, který nepozná, že dopravní značka s nálepkou je stále značka, řekli, že chápe koncept dopravní značky. Tento problém sahá i o několik desetiletí zpět, do časů předchozí vlny zájmu o AI. Tenkrát se vědci také zabývali problematikou Turingova testu, tedy tím, zda je možné AI považovat za inteligentní, pokud člověk nedokáže rozlišit mezi

chováním člověka a chováním AI [Woolgar 1985: 562]. Dostává nás to k problému možné (ne)vhodnosti užívání konceptu chápání u AI na základě toho, jak ho známe u lidí. Hlubší zkoumání tohoto problému už je ale mimo předmět této práce.

## 2. A jak to vidí sociologie?

V další části se podíváme na to, jak se tématu AI v rámci poslední vlny zájmu věnovaly sociologicky relevantní práce. V posledních letech došlo k explozi zájmu o AI, a to i v sociálních vědách. V rámci výzkumu jsem v databázích prací Google Scholar a Jstor vyhledával práce dle kombinací klíčových slov „artificial intelligence”, „machine learning” a „sociology”. Filtroval jsem práce jen na ty napsané po roce 2006 tak, aby odpovídaly jen časovému horizontu současné vlny zájmu o AI. Autoři píší o řadě témat, od analýzy politických médií [Burscher 2015], přes úvahy o adaptaci demokracie na technologické změny [McGinnis 2013], nahlížení AI jako pokračování zbožštělých mýtů [Mayor 2018], AI jako aktéra [Rhee 2018], AI a reprodukce nerovností [Risam 2018], AI jako technologie [Ford 2018] nebo AI jako obdoba předchozí automatizace [Nübler 2018], a další. V své práci jsem vybral poslední čtyři z těchto zmiňovaných textů proto, že nastiňují vhodně šíři možných přístupů k tématu AI. Jedná se tedy o vyvážení ostatních více technických kapitol sociologickými perspektivami. Budou představeny práce několika autorů a bude kriticky prezentováno, jakým způsobem AI tematizují. V závěru kapitoly se nachází přehledová tabulka, která shrnuje témata a přístupy rozebíraných prací.

### 2.1. Konverzační AI

Jennifer Rhee ve své knize „The Robotic Imaginary: The Human and the Price of Dehumanized Labor” píše o spojení AI a dehumanizace práce. Nahlíží na AI jako na antropomorfního aktéra [Rhee 2018: 24], používá mnoho příkladů ze sci-fi literatury a filmů, a je tedy, narozdíl od mé práce, blíže k přístupu k AI po vzoru populární kultury. Mluví ale i o příkladech z reálného světa, a proto může být kniha užitečná i pro moje zkoumání. V jedné z kapitol mluví o genderovanosti AI. Uvádí příklad superpočítače Watson, který má mužské jméno, a AI na něm běžící je užívána k úkolům, které byly tradičně mužskými úkoly – hraní šachů a analýza medicínských poznatků [Rhee 2018: 36]. V kontrastu k tomu uvádí příklady chytrých hlasových asistentů založených na AI, které často nesou jako výchozí jména a hlasy žen (Alexa od Amazonu, Siri od Apple<sup>8</sup>) a plní také

---

<sup>8</sup> Zároveň by se zde ale dal zmínit chytrý hlasový asistent jménem Bixby od Samsungu, který má mužský hlas i jméno nebo hlasový asistent Googlu, který nemá genderované jméno (oslovuje se jen jménem společnosti Google) ale v základě používá ženský hlas.

roli asistentek, kterou tradičně zastávaly ženy [ibid.]. V knize se tedy mluví o tom, jak určité charakteristiky – v tomto případě jména a hlasy – jsou AI přiřazovány způsobem, který replikuje tradiční genderové dělení úkolů a očekávání stejně jako u lidí. Její příklad by také šlo interpretovat (i když by to možná bylo proti úmyslu autorky), jako argument pro nahlížení na AI spíše jako na technologii v rukách lidí než na samostatného aktéra. Lidé jsou kulturním podmíněním zvyklí, že určité věci dělají spíše muži, a určité dělají spíše ženy, a podle toho pak pojmenovávají a nasazují různé AI v různých oblastech. Kulturní významy jsou přinášeny z vnějšku lidmi, nejedná se o autonomní aktérství AI nebo něco, co by se AI naučila na datech (jako v následujícím případě práce o reprodukci nerovností).

## 2.2. AI a reprodukce nerovností

Roopika Risam v článku „What Passes for Human? Undermining the Universal Subject in Digital Humanities Praxis” mluví o tom, že AI učením přebírá charakteristiky toho, na čem se učí, a je tedy problematické odkud tato trénovací data pocházejí [Risam in Losh a Wernimont 2018: 41-42]. Jako příklad uvádí konverzační AI od Microsoftu, která měla být otevřená lidem k online konverzaci a na základě těchto konverzací se i dále učit. Výsledek ale byl, že se AI od lidí naučila rasismu, urážkám a propagaci genocidy, a tak byl tento projekt rychle ukončen. Tento autorčin příklad má zobrazovat právě důležitost myšlení na to, jaké charakteristiky může AI od lidí přebrat a dále pak šířit [ibid.: 40]. Dále autorka říká, že poměrování AI podle toho, jak „lidské” chování vykazuje, je nutně zakotveno v konceptech lidství „Globálního Severu” a upozadňuje tak mnoho ostatních lidí světa [ibid.: 50-51]. Celkově tak práce upozorňuje na to, že AI opakuje a dále šíří to, co se na základě dat naučí. To může být nebezpečné, když to vede k reprodukci problematických implicitních předpokladů (hodnoty „Globálního Severu”) nebo také v situacích, kdy lidé záměrně využijí adaptivitu AI a záměrně ji něco (rasismus, nenávisť) naučí i proti vůli autorů AI. Jedná se tedy o kritiku představy neutrality nebo objektivit AI.

## 2.3. Dva pohledy na AI a práci

Knihy od Evy Paus „Confronting Dystopia: The New Technological Revolution and the Future of Work” obsahuje dvě kapitoly od dvou různých autorů, kteří se rozlišně zabývají tematikou AI a dopadů, které může v blízké budoucnosti mít na celosvětovou zaměstnanost. Důležité koncepty z těchto prací popíšu a krátce autory srovnám.

Prvním z těchto autorů je Martin Ford, ten píše o spojitosti AI a proměn pracovní sféry v kontextu historického efektu technologického pokroku na trh práce. Uvádí, že obavy z možnosti, že technologický pokrok přivede velké množství lidí k nezaměstnanosti se historicky objevily mnohokrát, ale nakonec se vždy jednalo o přechodné období, a lidé se nakonec adaptovali a přesunuli do jiných sektorů ekonomiky [Ford in Paus 2018: 28-29]. Zároveň ale podle něj hrozí, že takovýto mezisektorový přesun pracovníků tváří v tvář změnám, které přinese AI nebude možný, a to ze dvou hlavních důvodů. Za prvé mluví o rychlosti postupu AI, která se díky rychlejšímu hardwaru zlepšuje rychleji než technologie v minulosti [Ford in Paus 2018: 29]<sup>9</sup>. Za druhé mluví o mnohem obecnějších možnostech nasazení AI, než tomu bylo u předchozích technologií. Bude proto podle něj mnohem složitější, aby se lidé pouze přesunuli za prací do jiného sektoru neovlivněného novou technologií, protože narozdíl od situací v minulosti takových sektorů nebude dostatek [ibid.]. Mluví tak o potřebě jiného zaopatření lidí, než je zaměstnání, jmenovitě o nepodmíněném základním příjmu, tedy peněz, které by stát vyplácel každému člověku bez ohledu na jejich situaci [Ford in Paus 2018: 44-45].

Martin Ford je tedy zástupcem názorového proudu předvídajícího velkou šířku praktického uplatnění AI současné vlny. Očekává, že AI bude schopna vykonávat různá povolání, a skrze adaptivní učící schopnosti bude nahrazovat lidi rychleji, než bude vytvářet nová pracovní místa. Tento pohled na blízkou budoucnost je jistě obhájitelný ale rozhodně ne jediný. Obzvláště problém širší uplatnění AI je jedním z nejkontroverznějších témat. Dalo by se namítnout, že vytvářet očekávání masového uplatnění AI napříč oblastmi na základě vybraných příkladů úspěšného uplatnění ve specifických oblastech je předčasné nebo dokonce zavádějící. Některé oblasti, jako ne náhodou Fordem zmiňovaná radiologie, jsou pro uplatnění AI současné vlny mnohem vhodnější než jiné, a byla by chyba extrapolovat předpovědi budoucnosti pouze na základě pozitivních příkladů. Tomuto problému se moje práce věnuje také ve třetí kapitole, kde je na příkladech podrobněji popsána problematika vhodnosti AI k různým úkonům, obzvláště v části zabývající se oblastí medicíny.

---

<sup>9</sup> Ve spojení s rychlostí postupu zmiňuje Mooreův zákon, podle kterého se komplexita počítačových čipů, a tedy v praxi i jejich rychlost, zdvojnásobuje každých 12 až 18 měsíců. O tomto zákonu se dnes (ale i v době napsání knihy, i když to autor nezmiňuje) mluví jako o zpomalujícím, či dokonce zcela zastavujícím se, viz například [Theis a Wong 2017].

Také otázka časového horizontu je důležitá. I kdyby AI byla uplatněna tak široce, jak popisuje Ford, tak ve společenských důsledcích (nejen na trhu práce) bude extrémně záležet na rychlosti těchto změn. Rychlost tohoto uplatnění může znamenat polárně opačné výsledky sahající od situace podobné té, kterou popisuje jako historickou (lidé se adaptují, menší společenská změna) po naplnění jeho vize konce zaměstnanosti a potřeby něčeho jako je nepodmíněný základní příjem. Aby společenské změny vypadaly tak, jak o nich mluví Ford je nutná jak dostatečná šíře uplatnitelnosti AI, tak rychlý nástup těchto změn. Ford tedy ukazuje realistickou vizi uplatnění AI, která ale stojí na předpokladech o možnostech šíře a rychlosti tohoto uplatnění, na které nesmí být zapomínáno.

Druhým autorem píšícím o tématu AI a trhu práce je Irmgard Nübler. Ta se s Fordem shodne v tom, že technologie historicky zároveň s ničením pracovních míst také vytvářely místa nová. Zároveň ale přidává, že toto se děje, jen když jsou zisky z technologie dostatečně rozdělené v rámci společnosti tak, že to vytvoří zvýšenou poptávku. U případu AI říká, že jen trh sám o sobě zisky nerozdělí dostatečně rovnoměrně [Nübler in Paus 2018: 47-48]. Celá práce Nübler je založená na snaze predikovat společenské dopady AI na trh práce skrze nahlížení na průběh nástupu automatizace v minulosti. To je jistě do určité míry problematický přístup. Nelze očekávat, že aplikování AI v různých oblastech společnosti bude probíhat stejně jako nasazování robotů do továrny. Tento přístup ale i tak může nabídnout zajímavé poznatky, obzvlášť způsobem, jakým to dělá Nübler, který spočívá spíše v nahlížení na dílčí části procesu automatizace a na to, co se z nich můžeme pokusit předpovědět.

Píše o tom, že nahrazování pracovních pozic AI spočívá v nahrazení většiny úkonů vytvářejících danou pozici. Znamená to tedy, že pozice pro lidi zcela nezmizí, ale jejich povaha se změní a takových pozic bude pravděpodobně méně [ibid.: 50]. Jednodušší je tomu porozumět na příkladě. Představme si pracovní pozici řidiče kamionu. Řídit kamion na dálnici je relativně snadné (rovná, přehledná, žádní chodci, stačí se v podstatě držet pruhu) a je tedy velké riziko nahrazení této části práce AI. Na druhou stranu dopravit náklad na cílové místo určení může snadno znamenat jízdu zástavbou nebo jiným komplexnějším terénem plným lidí a překvapení, a je zde riziko nahrazení tohoto úkonu AI menší (minimálně zatím viz kapitola 3, část o úrovních autonomie aut). Kdybychom tedy měli dělat předpověď na základě tohoto pohledu, tak můžeme očekávat, že pracovní pozice řidiče kamionu pro lidi nezmizí zcela, ale promění se z řízení na velké vzdálenosti na dojíždění

posledních úseků. Dojet poslední úsek vzdálenosti zabere člověku méně času než projet celou cestu a dá se tak očekávat, že bude takových řidičů potřeba menší množství. To znamená že práce řidiče kamionu se v této modelové situaci změnila jak ve své povaze, tak v počtu, kolik takových pozic je k dispozici.

Neznamená to ale, že lze takto jednoduše předpovědět, že řidiči kamionů budou bez práce. Celospolečenské interakce jsou nesmírně složité. Zavedení AI může například znamenat, že kamionová doprava se stane efektivnější a více užívanou, a jako „dojížděči“ posledních úseků se uživí stejné nebo i větší množství řidičů než dříve<sup>10</sup>. Výsledná ztráta pracovních pozic tak může ležet například u řidičů jiných typů nákladní dopravy nebo dokonce nikde, pokud se jednoduše na základě vyšší efektivity začne dopravovat více zboží (tato druhá varianta je po vzoru žádaných varianty dopadů AI podle Nübler, kdy jsou zisky z automatizace rovnoměrně rozdělené a stoupá tak poptávka po zboží).

Autorka je zároveň skeptická k možnostem v široké míře aplikovat AI na „intelektuální a kognitivní úkony“ [ibid.: 50]. Říká, že „kritické myšlení, kreativitu, flexibilitu, soudnost, selský rozum a intuici nelze pozorovat, a tak ani současnými technologiemi simulovat“ [ibid.: 51]. Zároveň ale také zmiňuje, že „Je očekáváno, že tyto učící se stroje (tedy AI - pozn. autora) sníží poptávku po profesích jako jsou doktoři, právníci, lingvisté, novináři a piloti a také, že změní povahu úkonů v těchto profesích.“ [ibid.: 50]. Dalo by se tedy argumentovat, že tato dvě tvrzení jdou proti sobě, ale je pravděpodobné, že autorka jimi myslí, že lze očekávat proměnu zmíněných povolání více směrem k nesimulovatelným úkonům jako je kreativita, kritické myšlení a další které zmiňuje v první citované pasáži. Jde tedy o podobnou proměnu v obsahu pracovních pozic a jejich potenciálního počtu jako u příkladu řidičů kamionů.

Autorka také mluví o tom, že dopady aplikování AI velice záleží na specifikách zemí, kde budou aplikovány. Znovu se pokouší extrapolovat ze zjištění o průběhu automatizace, tentokrát z různých sektorů ekonomiky v rozvinutých zemích v letech 1993 až 2007. Říká, že během automatizace výrobních procesů („sekundér“, tedy sektor výroby např. elektroniky, automobilů, součástek apod.) všechny studované země zaznamenaly větší

---

<sup>10</sup> Což zase potenciálně zhoršuje problémy s plností silnic nebo emisemi, společenské dopady jsou i kvůli této rozmanitosti těžké předpovídat.

produktivitu a vyšší mzdy, ale jen některé to dokázaly převést ve větší poptávku (kterou autorka používá jako zástupce pro obecný blahobyt). Zároveň také dodává, že vyšší nárůst počtu robotů ve výrobě nekorresponduje se ztrátou pracovních míst v jednotlivých zemích. Tato dvě zjištění používá jako argumenty pro své predikce o dopadech AI na společnost.

Těmito argumenty jsou za prvé zdůrazňuje důležitost širšího rozdělení zisků, které AI zvýšenou produktivitou přinese [ibid.: 60-61]. Za druhé pak připomíná důležitost dalších faktorů než jen rozložení pracovních pozic na trhu při určování dopadů na společnost. Například mluví o rozdílných kulturních přístupech k tomu, jestli je dáována prioritou co nejširšímu rozšíření vysokoškolského vzdělání jako v USA [ibid.: 71], nebo je jako schvalovaná životní cesta studenta bráno i učení se „praktickým“ povoláním a řemeslům jako v Německu [ibid.: 63]. Paradoxně právě v USA jsou dopady automatizace (které autorka výhledově spojuje i s dopady AI) na lidi negativnější než v Německu. To autorka vysvětluje tím, že oproti Německu v USA chybí systém, který by dokázal lidem, kteří vinou automatizace přišli o práci, nabídnout možnost se adaptovat nebo příbuznou rekvalifikaci.<sup>11</sup> Zde znovu stojí za připomenutí, že autorka extrapoluje předpovědi o AI na základě automatizace výroby. Je složité určit do jaké míry takto lze postupovat, a pravděpodobně to bude možné s jistotou říci až retrospektivně. Jistě to ale stojí za připomenutí, protože samotný text autorky se tímto rozdílem příliš nezdržuje.

Posledním relevantním bodem, o kterém autorka mluví, a kterým se zároveň posouvá od spíše ekonomizujícího přístupu k tomu sociologicko-politickému, je popis fází dopadů nových technologií na společnost. Nová revoluční technologie (tedy i automatizace nebo AI) má vždy dvě fáze. V první se zvyšuje produktivita, ale mizí pracovní pozice tím, že se aplikuje technologie na existující odvětví. Ve druhé se pak pomocí technologie začnou vytvářet nové produkty nebo celá odvětví a vznikají nové pracovní pozice [ibid.: 71]. Autorka říká, že přechod z první fáze do druhé musí být vždy aktivní politické rozhodnutí [ibid.: 73], které se nestane samo vlivem trhu, protože firmám se daří dobře i v první fázi (produktivita roste). Na druhou stranu v první fázi se dobře daří jen některým lidem (těm jejichž pracovní pozice není nahrazena) a je proto potřeba včas přejít do fáze druhé

---

<sup>11</sup> Tyto myšlenky by stálo za to zvážit i ve vztahu ke stavu v České republice. Zpráva o AI pro Česko, vypracovaná pro českou vládu a zmiňovaná v úvodu této práce, mluví o Česku jako o jedné ze zemí, pro kterou je AI nejnebezpečnější, co se zaměstnanosti týče. Jistě by bylo zajímavé prozkoumat, jak vypadala automatizace výroby ve stejných letech, ve kterých je zkoumala Nübler, v Česku, a jestli se z toho dají vyvodit predikce či vhodné kroky do budoucna ve vztahu k AI a Česka.



[ibid.: 72]. V návaznosti na další autory říká, že v současnosti jsme v momentu mezi první a druhou fází, a že je potřeba udělat rozhodnutí o tom, jak má vypadat samotný charakter práce a zaměstnání obecně (mluví o možnostech kratšího pracovního týdne nebo větším zastoupení samostatně výdělečné činnosti) [ibid.].

Protože i Martin Ford i Irmgard Nübler se věnují tématu zaměstnanosti rozdílně, stojí za to je krátce přímo srovnat. Fordova práce nahlíží na AI jako na bezprecedentní změnu, která vyžaduje relativně bezprecedentní řešení, jako je nepodmíněný základní příjem. Práce Nübler bere AI jako změnu obdobného typu jako byly technologické změny v minulosti. Adekvátní reakce na tyto změny mají podle ní být opatření udržující a posouvající současný systém (nějaká forma přerozdělení zisků z AI ale explicitně za účelem zvýšení poptávky, tedy sociální ale zároveň tržně založené řešení).<sup>12</sup>

### 2.3.1. Přehledová tabulka sociologických autorů a tematizací AI

Následuje tabulka rekapitulující tematické bloky čtyř rozebíraných autorů. Může sloužit jako zjednodušené východisko pro získání přehledu v oblasti současných sociologicky relevantních prací o AI.

<i>Autor</i>	<i>Rok</i>	<i>Témata</i>
Jennifer Rhee	2018	AI a genderovanost, AI jako aktér
Roopika Risam	2018	AI a reprodukce existujících nerovností
Martin Ford	2018	AI jako technologie, AI a zaměstnanost, AI bezprecedentní
Irmgard Nübler	2018	AI jako technologie, AI a zaměstnanost, AI jako obdoba automatizace v minulosti

*Tabulka 1 Přehled rozebíraných sociologicky relevantních prací o AI*

<sup>12</sup> Kdybych měl ze všech prací z této kapitoly vybrat a doporučit k přečtení jednu, tak to bude tato od Irmgard Nübler. Dle mého názoru nejlépe vyvažuje prvek skeptičnosti k šíří uplatnění AI s faktem, že AI přeci jen má potenciál být revolučnější než technologie minulosti. Není celkově tak ekonomizující, jak by se mohlo zdát.

## 3. Dopady AI ve dvou vybraných oblastech

Ve třetí kapitole se podíváme na to, jak je AI v současnosti aplikována ve dvou konkrétních oblastech společnosti – v medicíně a v automobilové dopravě. Tato kapitola má doplňovat celkovou funkci této práce jako přehledu tím, že nabídne konkrétní příklady realizace toho, o čem bylo v předchozích dvou kapitolách hovořeno více teoreticky. Oblasti medicíny a automobilové dopravy byly vybrány proto, že jsou v nich zároveň dělány velké pokroky v rámci současné vlny AI, a zároveň mají velký potenciál v hloubce i šíři dopadů na společnost. Jedná se o příklady, které mají sloužit pro vytvoření představy, ne o vyčerpávající výčet. Stejně dobře mohlo být vybráno i mnoho jiných oblastí jako je analýza přirozené řeči, rozpoznávání tváří nebo například hraní her proti člověku.

### 3.1. AI v medicíně

V této části se tedy podíváme na to, jakým způsobem je AI aplikována v reálné oblasti společnosti, a to v oblasti medicíny. Zájem o využití AI pro zlepšení lékařské péče sahá i několik desetiletí do minulosti. O vývoji v této oblasti pojednává práce „Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes” [Peek et al. 2015]. Jedná se o práci mapující vývoj témat v rámci prestižní konference o AI a medicíně odehrávající se každé dva roky. Práce publikované v jejím rámci jsou navíc z pravidla silně zaměřené na efektivitu a reálnou praktickou aplikaci. Přehledová studie témat prací z ročníků konference tak představuje užitečný nástroj pro studium podoby dopadů AI na medicínu v rámci současné vlny zájmu (tedy od roku 2006 dále). Práce mluví o významném posunu k datově orientovaným metodám jako je právě strojové učení v nedávných letech [Peek et al. 2015: 2].

První důležitou oblastí aplikace AI ve studii v časovém rámci současné vlny zájmu je využití schopností AI porozumět přirozenému lidskému jazyku. Porozumění přirozenému jazyku je jedna z oblastí, kde AI v současnosti dosahuje největších pokroků<sup>13</sup>. V medicíně se zpočátku jednalo o analýzu dokumentů a hledání skrytých vzorců ve zprávách o pacientech.

---

<sup>13</sup> Mimo sféru medicíny lze pozorovat velké rozšíření chytrých hlasových asistentů od společností jako Google, Apple nebo Amazon, které dokáží reagovat na to, když na ně člověk mluví způsobem mnohem bližším standardnímu rozhovoru než například zadávání klíčových slov do vyhledávače.

V současnosti se ale aplikace rozšiřuje i například na oblast sledování a predikování epidemií skrze analýzu sociálních médií [Peek et al. 2015: 5].

Druhou oblastí využití AI v rámci medicíny je management rizika. Jedná se o aplikaci AI a specificky statistických pravděpodobnostních modelů v situacích, kdy jsou k dispozici neúplné nebo nekonkluzivní informace. Konkrétně se tedy může jednat například o interpretaci skenů z mamogramů nebo predikce možných interakcí mezi léky u pacientů užívajících více různých léků [Peek et al. 2015: 5-6].

Další oblastí využití je analýza fyzické aktivity lidí mimo klinické prostředí s využitím rozšířenosti nositelných zařízení se širokou škálou senzorů (telefony, ale i chytré náramky nebo hodinky). Text zmiňuje například monitorování průběhu a zdravotního stavu pacientů s Parkinsonovou chorobou v reálném čase [Peek et al. 2015: 7]. Znovu můžeme pozorovat důležitost dostupnosti technologií (AI software v kombinaci s rozšířeností senzorů, které může mít pacient stále u sebe) pro specifickou podobu projevů AI ve společnosti. Je také stále častější, že pojišťovny nabízejí výhody (někdy dokonce i přímo peníze) lidem, kteří budou nosit zařízení se senzory a sdílet data z nich s pojišťovnou [Henkel, Heck, Göretz 2018: 28-29]. Tato data pak mohou být využita ke zlepšení péče ale také například k odhalení rizikových faktorů a chování (což jistě je oblast zájmu pro zdravotní pojišťovny).

V časově ještě bližší minulosti několika minulých let lze pozorovat vědecké práce srovnávající výsledky lidských doktorů a AI. Snahy jsou soustředěné zejména v oblasti diagnózy, protože ta představuje často hledání vzorců v datech nebo dělání rozhodnutí na základě minulých dat. Takovýto typ úkonů je v nejlepším souladu se silnými stránkami současné AI (hledání nádorů ve skenu může být obdoba rozpoznávání obličeje v obrázku viz příklad v kapitole o fungování AI výše).

IEEE, nezisková organizace zaměřená na aplikaci nových technologií ve prospěch lidstva spravuje stránku, kde mapuje výsledky těchto prací a porovnává výkony AI a lidských doktorů [AI vs. Doctors 2019]. Výsledky rozhodně nejsou jednoznačné. Ve studii z roku 2017 dokázala AI na vzorku obsahujícím přes 100 000 rentgenových snímků identifikovat správně více případů angíny než lidé a zároveň se také častěji vyhnout nesprávným diagnózám zdravých pacientů [Rajpurkar et al. 2017: 1]. Podobných výsledků došla i studie zkoumající diagnózu infekce nehtů, v níž na vzorku 50 000 snímků AI překonala téměř vždy

všechny ze 42 lidských doktorů [Seung Seong et al. 2018: 1]. Mimo analýzu snímků byla AI například schopná lépe než člověk na základě rutinních zdravotních dat predikovat, kdo z vzorku 378 000 lidí bude mít v horizontu 10 let srdeční příhodu (notoricky složitý úkol) [Weng et al. 2017: 1], nebo na základě analýzy odpovědí v dotazníku přesněji než lidští lékaři predikovat budoucí diagnózu autismu u šesti-měsíčních dětí [Emerson et al. 2017: 1].

Výsledky studií ale nemluví zdaleka vždy ve prospěch AI. Například v případě diagnóz na jednotce intenzivní péče a při diagnóze mutací nádorů je AI schopná vyhledávat mutace velice dobře, ale nedokáže navrhnout léčbu [Wood et al. 2018: 1]. I když jsou výsledky AI dobré, nemusí to také znamenat, že lidští lékaři budou nahrazeni. V případě diagnózy rakoviny prsu dosahuje AI stejně dobrých výsledků jako lidé, ale není jasné, jestli se vyplatí kvůli „jen“ stejně dobrým výsledkům implementovat drahé systémy používající AI [Kerlikowske et al. 2018: 763]. Jako překážka nahrazování lidských doktorů AI se obecně také objevuje fakt, že AI je často závislá na pomoci doktorů před začátkem analýzy nebo pak při zvolení vhodné reakce (léčby) na ni. I přes všechna tato omezení je hlavní, že AI zásadním způsobem mění to, jak medicína v současnosti postupuje dále, a to i přičiněním samotných výzkumníků z této oblasti (nejedná se tedy o situaci, kdy by byla AI jen vnášena zvenčí). I vzhledem k tomu o jak globální a masivně financovanou oblast činnosti se jedná lze očekávat, že medicína bude jedna z velkých oblastí, kde se běžní lidé budou s AI setkávat.

### 3.1.1. Reflexe dopadů AI v medicíně

Můžeme tedy pozorovat, že existují snahy aplikovat AI v medicíně jako náhradu doktorů. Jako hlavní problém řešený oborem jsme viděli otázku, zda je vůbec možné se pokoušet doktory s pomocí AI nahradit. Případná problematizace se pak týkala buď míry, do které by bylo možné doktory nahradit, nebo ekonomické proveditelnosti této potenciální změny. Toto jsou jistě zásadní otázky, ale vzhledem k tomu, že odpovědi na ně směřují směrem k tomu, že doktory AI opravdu do nějaké míry nahradit může, zvedá se řada otázek nových. Jak může nasazení AI ovlivnit vztah pacientů k medicíně? Zlepší se důvěra díky víře v objektivitu algoritmů nebo naopak zhorší kvůli jejich neosobní povaze? A co samotná lékařská profese, jak se promění? Jak jsme si ukázali, AI je v současnosti systematicky vhodnější pro určité oblasti (analýza obrázků, skenů) než pro jiné, ale co to znamená pro doktory působící v těchto oblastech? Je to pro ně pomocný nástroj nebo hrozba? Dokáže

se obří korporátně-státně-univerzitní systém medicíny adaptovat na tyto změny? Nebo je zde i tradiční otázka zodpovědnosti. Když AI udělá špatnou diagnózu, má za ni být zodpovědný programátor nebo nemocnice? A co naše zodpovědnost jako sociologů při ptaní se na tyto otázky, nevytváříme špatnými otázkami více strachu, než je nutné a nebráníme tak postupu technologie, která by jinak mohla zachraňovat životy? Tohle všechno jsou otázky vhodné pro sociologii. Takové otázky, na které se možná zdá být brzy, když nevíme, zda je nahrazovat doktory AI vůbec možné. Zároveň je ale medicína oblast kde se spojuje vliv obrovského množství peněz i moci (doslova život a smrt), a zanedbat tak tyto otázky je jistě riskantnější než se jim věnovat příliš brzy.

### 3.2. AI a autonomní vozidla

Druhou oblastí, kde si ukážeme skutečné uplatnění AI je oblast autonomních vozidel, tedy automobilů, které se samy řídí. Velká část významných světových automobilek spolupracuje s technologickými firmami nebo dalšími automobilkami při snahách o vývoj těchto autonomních vozidel. Toto je umožněno z velké části právě díky pokrokům v oblasti AI a dostupnosti snímacích technologií<sup>14</sup>. AI je užívána k tomu, v čem je obecně nejlepší, a to je učení se vzorcům a jejich následnému rozpoznávání. S pomocí velkých datových souborů je tedy AI trénována (skrže strojové a hluboké učení), aby uměla rozpoznat dopravní značení, semaforey, lidi, ostatní auta, překážky a další relevantní objekty a situace. Tato část se může (nebo v podstatě musí) nejprve díť i v laboratoři, bez reálné existence fyzického auta. V další části pak lze přejít do praxe. Na auto jsou umístěny kamery a senzory, které mu umožňují vnímat okolí, aplikovat naučené vzorce a dále se učit. Toto může podle úrovně schopností AI a lokální legislativy být provedeno na testovacích okruzích, vyhrazených částech dopravních komunikací nebo dokonce v běžném provozu.

Oblast autonomních vozidel má potenciál být praktickou aplikací AI, která dramaticky změní každodenní životy lidí. Není těžké si představit široké spektrum změn, které by nahrazení lidského řidiče mohlo přinést. AI neřídí pod vlivem alkoholu, neusíná za volantem, nedělá chyby z nepozornosti, není agresivní, nepřekračuje povolenou rychlost a obecně nemá mnoho slabostí lidských řidičů. Zároveň také otevírá možnosti zcela nové, jako je koordinace aut mezi sebou vzájemně i s infrastrukturou (můžeme si představit situaci

---

<sup>14</sup> Důležité jsou samozřejmě i ostatní oblasti, jako je například mechanické ovládání částí vozidla. To už je ale mimo zaměření této práce. Zde bude pojednááno hlavně o významu AI pro autonomní vozidla.

na křižovatce, kdy si auta vzájemně komunikují, co se chystají udělat a mají dostupné informace, kdy přesně se změní barva na semaforu). Potenciál je ale i zcela lidský, například pro lidi s pohybovými omezeními nebo zrakovými problémy může auto, které je bez závislosti na dalších lidech zaveze, kam potřebují, představovat novou možnost svobody.

Zároveň je potřeba myslet na to, že autonomní auta sice postrádají mnoho slabostí lidských řidičů, ale přináší také slabosti nové. Automatické systémy mohou selhat, nebo být cíleně napadeny (viz příklad s rozpoznáváním dopravních značek v kapitole o technickém fungování AI výše). Současná AI je navíc založena na učení se správným reakcím na základě velkého množství situací. To může být ale v případě řízení auta problém, vzhledem k tomu, že kritická může být právě schopnost správně a rychle reagovat v situaci, která je sice extrémně vzácná (a tedy v trénovacích datech potenciálně neobsažená), ale životu nebezpečná. Je tu také otázka morality. Měla by AI například v nevyhnutelné kritické situaci upřednostnit život pasažéra nebo ostatních lidí? Co pro společnost znamená delegování takovýchto morálních rozhodnutí o životě a smrti na automatické systémy? Mnohé výhody autonomních vozidel navíc přichází, až když většina nebo všechna auta budou autonomní. To není vůbec jasné kdy (nebo jestli) se stane a odhady se mohou drasticky lišit od očekávání nástupu během několika let až po několik desetiletí. Tyto rozdíly v předpovědích mohou částečně také vyplývat z toho, že nemusí být zcela jasné, co ještě je autonomní auto (počítá se automatické brždění nebo systém hlídání jízdního pruhu?).

Proto bude následující část věnovaná typologii úrovní autonomních vozidel podle SAE International (Společnost automobilových inženýrů) a příkladům, v jakých situacích se s vozidly jednotlivých úrovní můžeme v současnosti setkat. To by mělo pomoci se zorientovat v možnostech společenských dopadů a očekávatelných časových horizontech.

### 3.2.1. Šest úrovní autonomie

SAE International ve svém obecně uznávaném standardu rozlišuje šest možných úrovní autonomie vozidel. S každou další úrovní klesá nutnost zásahů řidiče, a naopak přibývá schopností a situací, kdy je AI schopna řídit auto sama [SAE International 2018: 19, tabulka 1].

Úroveň nula je žádná autonomie, celý proces řízení vozidla tedy obstarává lidský řidič. Možná trochu překvapivě sem ale spadají i auta, která určité automatické systémy mají.

Pokud má auto automatické brzdové systémy (ABS), nebo například parkovacího asistenta, který zobrazuje na obrazovce, jak má člověk parkovat, ale nedokáže převzít kontrolu na autem, tak spadá také do této kategorie nula s žádnou autonomií.

Úroveň jedna je asistence řidiči. Spadají sem vozidla, která dokážou ve specifické situaci převzít část kontroly nad vozidlem, ale lidský řidič dále obstarává zbytek řízení. Můžeme si představit buď adaptivní tempomat, který drží stabilní rychlost vozidla a zároveň hlídá odstup od auta před sebou, nebo systém hlídání pruhu na dálnici, který v případě pozvolného opuštění pruhu (takového které systém vyhodnotí jako nežádoucí, způsobený například nepozorností řidiče) sám změni trasu vozidla, aby zůstalo ve stejném pruhu. Většinu řízení tedy stále obstarává řidič, a systém je omezený jen na určité situace (například dálnice, které jsou obecně předvídatelnější než jízda městem). Vozidla spadající do této kategorie jsou běžně k dostání i v současnosti u velkých automobilek (například i česká Škoda [ŠKODA OCTAVIA Komfortní asistenty: 2019]).

Úroveň dvě je částečně automatické řízení. Tato úroveň autonomie se dá chápat jako rozšířená verze předchozí úrovně. Pokud vozidlo samo udržuje jak rychlost a odstup, tak pozici v pruhu, tak už spadá sem do úrovně dvě. Stále se ale jedná o schopnosti omezené situačně (například pouze na dálnici) a lidský řidič doplňuje ostatní části řízení a dohlíží na chování automatického systému. I vozidla spadající do této kategorie lze v současnosti koupit u velkých automobilek.

Úroveň tři je podmíněná automatizace řízení. Zde je už celý proces řízení vozidla prováděn automaticky, ale stále pouze v určitých situacích (např. dálnice). Lidský řidič také musí být stále ve vozidle přítomen, aby v případě potřeby reagoval na výzvu vozidla převzít kontrolu, například při opuštění situace vhodné pro automatizaci (sjezd z dálnice) nebo při setkání se s neznámou situací. Důležitý je právě prvek reakce na výzvu, řidič tedy nemusí trvale dávat pozor na to, co se děje na silnici, ale musí být schopen reagovat na výzvu vozidla, aby převzal kontrolu. Této úrovni z v současnosti dostupných vozidel dosahují například automobily značky Tesla<sup>15</sup>.

---

<sup>15</sup> Stojí za zmínku, že jednou z prvních smrtelných obětí autonomního vozidla byl právě řidič v autě od značky Tesla, který nestihl zareagovat na výzvu vozidla, aby převzal kontrolu [Hull a Smith 2018].

Úroveň čtyři je vysoká automatizace řízení. Jedná se o stejnou situaci jako třetí úroveň – celý proces řízení je automatizován, ale stále omezen na určité situace/oblasti. Rozdíl zde je ten, že už zde neexistuje předpoklad, že člověk bude reagovat na výzvu k převzetí řízení. V předchozích částech užívaný příklad s užitím na dálnici zde není příliš vhodný. Společnost Waymo (spadající pod Alphabet, tedy vlastníky Googlu) provozuje v městech San Franciscu v Kalifornii a Chandleru v Arizoně testovací provoz autonomní taxi služby, do kterého se mohou i běžní lidé přihlásit [Heinzman 2019]. Stojí za zmínku, jak zdánlivě nenápadný rozdíl v nutnosti lidského řidiče být připraven zasáhnout mezi třetí a čtvrtou úrovní, může být ve svých důsledcích mnohem větší. Protože člověk už nemusí do řízení nikdy zasáhnout, jedná se o první z úrovní, kdy lze smysluplně odebrat ovládací prvky z kabiny vozu a z lidského řidiče se tak může stát lidský pasažér. K zakoupení vozidlo této úrovně v současnosti není.

Úroveň pět je úplná autonomie. Vozidlo je schopné se řídit zcela samo bez omezení na specifické situace/místa a bez očekávání zásahu lidského řidiče. Tato kategorie je ze všech zmiňovaných od přítomnosti časově nejvzdálenější. Auta z této kategorie nejsou v současnosti dostupná. Celá podkapitola „Šest úrovní autonomie“ [SAE International 2018: 19].

### 3.2.2. Potenciální společenské dopady autonomních vozidel

Ted' když rozumíme skrze šestiúrovňové schéma lépe možnostem autonomie aut, lze lépe zhodnotit jejich potenciální společenské dopady a různé optimistické nebo pesimistické scénáře.

Za prvé se vrátíme k problematice řidičů kamionů (viz kapitola 2, část o AI a zaměstnanosti). Aby mohlo začít nahrazování lidských pracovních míst AI, je potřeba alespoň autonomie čtvrté úrovně (nejnižší, která zároveň nevyžaduje připravenost lidského řidiče reagovat na výzvu k převzetí řízení). V současnosti společnost Embark testuje na amerických dálnicích nákladní vozidla, zatím druhé úrovně (tedy s dozorem řidiče), ale dá se očekávat, že cílem je dosáhnout možnosti řidiče nahradit [Embark Trucks 2019]. V Kalifornii je také na stole politický návrh, který by umožnil testovat menší autonomní nákladní vozidla a dodávky i ve městech [Hawkins 2019]. Z pokročilé úrovně těchto testovacích procesů se dá odvodit, že autonomní nákladní doprava je pravděpodobně relativně blízko (což je jistě podmíněno i tím, že se odehrává z velké části právě na dálnicích, které jsou pro AI



přehlednější než města plná lidí a překážek). Diskuze o potenciálních dopadech automatizace těchto řídicích pozic viz kapitola 2, část o práci Irmgard Nübler [Nübler 2018].

Jako další relevantní současný společenský aspekt autonomních vozidel jsou úmrtí jimi způsobená. Jak se vozidla dosahující druhé a třetí úrovně autonomie dostávají do rukou lidí, dochází i k případům úmrtí. Jednalo se většinou o smrti řidičů v autonomních vozidlech [Boudette 2016 a Joseph 2018]. Dalo by se spekulovat, že právě druhá a třetí úroveň jsou možná nejzrádnější, protože jsou dostatečně autonomní, aby vedly řidiče k nepozornosti, ale ne tak autonomní, aby zabránily nehodě.

V souvislosti s těmito úmrtími se dá hovořit o několika společenských aspektech. Za prvé je zde otázka odpovědnosti. Ta bývala dříve spíše hypotetická, ptající se kdo bude zodpovědný za to, až se někdo zabije v autonomním autě. Zatímco tato otázka zůstává morálně filozoficky spíše nevyřešená, soudy v praxi už rozhodují. V případě muže na Floridě, který zemřel po tom, co jeho automobil značky Tesla narazil při aktivovaném autonomním módu do nákladního vozidla, soud rozhodl, že nešlo o technický defekt systému, a výrobce automobilu není za smrt zodpovědný.

Druhým aspektem, o kterém lze ve spojitosti s úmrtími v autonomních vozidlech mluvit je společenská reakce na ně. Klasickým argumentem zastánců autonomních vozidel je, že přílišným soustředěním na výjimečné případy nehod můžeme způsobit zbytečnou paniku a ve výsledku připravit o život více lidí, kteří by mohli být zachráněni bezpečnostními zlepšeními autonomních vozidel. Stejně jako v případě medicíny se i zde tedy nabízí otázka po roli sociologie (ale i dalších aktérů účastnících se diskurzu o autonomních vozidlech). Jak máme vyvážit nepochybnou nutnost mluvit o nebezpečích autonomních aut s rizikem, že způsobíme přílišnou paniku, která může ve výsledku stát lidi víc životů v nehodách, kterým by autonomní auta mohla jinak zabránit? Jistá míra paniky se nicméně dá ve společnosti s velkou pravděpodobností očekávat tak jako tak. U názorů na autonomní vozidla mohou hrát roli také faktory jako strach ze ztráty zaměstnání v důsledku automatizace, nebo představy o vzpouře strojů z populární kultury.

Také je na místě otázka, jak dobrá musí AI být, aby její nasazení bylo společensky morálně přijatelné. Lidští řidiči působí v současnosti mnoho a mnoho smrtí. Jak blízko dokonalosti

musí tedy být AI, aby byla přijatelnou náhradou za ně? Dá se na to nahlížet konsekvencialisticky, že by se měla prioritizovat minimalizace celkového počet úmrtí na silnici. Na druhou stranu je rozdíl, a jistě by to mělo i velké rozdíly ve společenském vnímání, mezi tím, když někdo zemře v autě vlastní vinou (například řízení v opilosti), a situací, kdy zemře člověk bez vlastního přičinění, protože algoritmus udělal chybu. I potenciál velikosti dopadů kybernetického útoku je zde v opravdu enormní. Představa, že by při velkém rozšíření autonomních aut (obzvláště u vyšších úrovní, kdy může být zásah lidského řidiče znemožněn odebráním ovládacích prvků) například teroristé nabourali systém řízení, je hrozbou globální katastrofy nebyvalých rozměrů.

V neposlední řadě se můžeme pokusit zhodnotit míru realističnosti optimistických i pesimistických scénářů autonomních vozidel. Optimistické scénáře (jako je super-efektivní doprava ve městech nebo zmiňovaná vyšší svoboda pro osoby s tělesným postižením) se zdají vyžadovat úspěšného, a hlavně široce rozšířeného nasazení autonomních vozidel alespoň čtvrté úrovně (ta první bez nutnosti zásahů řidiče). Ta jsou zatím ve fázi testování, takže na potenciální komerční využití a široké uplatnění bude ještě potřeba počkat. Více pesimistické scénáře jako jsou zvýšené problémy s hustotou provozu a emisemi, nebo masové softwarové nabourání systémů autonomních vozidel mají pravděpodobně nižší nároky na vysokou úroveň automatizace, i u nich ale bude pro velikost dopadů záležet na počtu autonomních mezi lidmi.

Při pohledu do budoucna také můžeme polemizovat, jestli je současná vlna AI vůbec vhodná pro autonomní vozidla, a jestli nevyhnutelně nenarazí na nepřekonatelnou bariéru v podobě potřeby reagovat na situace, na které se jednoduše nelze připravit masivním množstvím trénovacích dat. Pokrok se zdá být zatím slibný, ale pravdu ukáže až čas.

## Závěr

### Shrnutí

Tato práce měla za cíl být přehledem a vhodným úvodním průvodcem pro sociology zajímající se o tematiku umělé inteligence (AI). Na začátku práce byla vyslovena otázka ptající se po sociologických aspektech AI. Na ni celá práce postupně odpovídá, a navrhuje, co by těmito aspekty mohlo být. Celkově by se sociologické aspekty AI daly shrnout do dvou oblastí. Jedna jsou vnitřní záležitosti fungování AI. AI je vytvářena lidmi, učí se na základě dat od lidí, a reaguje na lidi. První oblast sociologických relevancí je tedy zkoumání toho, jak AI přebírá a dále šíří hodnoty, kterým je vystavena. AI není neutrální a objektivní jenom proto, že to není člověk, a proto záleží kdo a jak ji bude vytvářet. Proto jsou v práci části o technickém fungování AI a texty zabývající se tematikou reprodukce nerovností AI. Druhou oblastí aspektů jsou efekty, které AI má a bude mít v oblastech své aplikace. Kombinace sociologických prací o zaměstnanosti s kapitolou o uplatnění AI ve vybraných oborech ukazuje možnosti změn přinesených AI, které můžeme zkoumat. Jestli budou lidé v práci nahrazeni, jaké části jejich povolání to nejspíše budou, a jak brzy to můžeme očekávat, to vše jsou příklady jevů, které můžeme zkoumat a zvažovat jejich dopady nebo možné reakce na ně. Tyto dvě oblasti tvoří tedy příklady sociologických aspektů, po kterých se na začátku ptám. Návrhy, jak by s nimi bylo možné výzkumně pracovat jsou načrtnuté níže v této kapitole v části Diskuze.

Práce má tři hlavní kapitoly, která každá plní dílčí úkol ve vytvoření celkové představy o stavu současného poznání. První, technická kapitola vysvětluje fungování toho, čemu se v posledních letech na základě vlny zájmu způsobené novými objevy říká AI. AI jsou pro tuto práci takové algoritmy, které jsou schopné dělat vhodná rozhodnutí na základě učení se z předchozí zkušenosti. Nejdůležitějšími způsoby, kterými se tyto algoritmy učí jsou strojové a hluboké učení. Právě průlom v aplikaci těchto dvou postupů jsou hlavním důvodem existence současné vlny zájmu o AI.

Další kapitola nabízí přehled sociologických prací z posledních let na téma AI. Čtyřem z těchto prací se pak věnuje podrobněji a představuje, jaké jsou možnosti sociologických

náhledů na AI. Celkově jsou autoři těchto prací přesvědčeni o velkém potenciálu společenských změn, které má AI přinést. Zároveň ale často kritizují to, jak je AI vytvářena, tedy že data, na kterých se učí ovlivňují výslednou podobu AI. Kvůli tomu hrozí, že AI bude reprodukovat názory a nerovnosti od lidí, kteří ji učí (ať jsou to původní autoři-programátoři nebo uživatelé, z jejichž chování se dále učí). U sociologů se také objevují obavy z dopadů AI na společnost, a mluví o nutnosti vhodně reagovat na potenciálně zásadní změny v životech lidí. Konkrétně o tom hovoří například dvě práce zabývající se AI a zaměstnaností [Ford 2018, Nübler 2018].

V poslední kapitole jsou pro představu uvedeny podrobněji dvě z nesčetné řady oblastí, kde existují v současné době snahy aplikovat AI. První z nich je oblast medicíny. Z průzkumu prací zkoumajících možnost aplikace AI na různé oblasti medicíny se zdá, že AI může v současnosti ve vhodných úkolech (nejvíce v rozpoznávání vzorců v obrázcích, tedy například hledání nádorů) dosahovat výsledků stejně dobrých, nebo i lepších než lidští doktoři. Zároveň se ale objevují oblasti, kde lidské výsledky překonat zatím AI nedokáže, a i tam kde ano, jde často o dílčí části celého procesu léčby pacienta. Ve druhé části kapitoly je popsána oblast autonomních vozidel. Tato část je založena hlavně na typologii úrovní autonomie vozidel od SAE International [SAE International 2018]. V současnosti dosahují komerčně dostupná vozidla nejvýše úrovně čtyři<sup>16</sup> ze šesti úrovní autonomie. Aby došlo k výrazným změnám ve společnosti, o kterých různí autoři hovoří, jako jsou dopady na zaměstnanost řidičů nebo uplatnění ve městech, jsou potřeba vozidla vyšší než jen čtvrté úrovně autonomie. Vozidla dosahující vyšších úrovní jsou v současnosti testována, odhady o jejich komerčním nasazení se ale různí.

## Diskuse

Tuto práci jsem se pokusil napsat tak, aby představovala argument, proč stojí za to AI sociologicky zkoumat, a zároveň jako úvodní představení problematiky AI pro ty, kteří by se ji sociologicky zkoumat snažili. Rozumím AI jako něčemu, co už v současnosti mění společnost, a u čeho navíc existuje reálná šance, že ji bude měnit i dále a více. AI tvoří zajímavý mix něčeho, co je více autonomní než předchozí technologie, ale (zatím?) ještě ne tak autonomní, že bychom na ni měli nahlížet jako na samostatného aktéra. To vše jsou důvody, proč by sociologie měla AI zkoumat. Myslím, že bychom jinak zpětně mohli litovat,

---

<sup>16</sup> Úroveň čtyři je, kvůli číslování od nuly, druhou nejvyšší ze šesti.

že jsme v době, kdy začínala nová éra proměny lidských společností, my sociologové nevyužili příležitost stát uprostřed nebo dokonce v popředí těchto změn.

Tato práce může sloužit jako východisko pro další zkoumání. Nabízí se možnost vyjít z práce Irmgard Nübler, a prozkoumat, zda je z průběhu automatizace ve výrobě v Česku možné říci něco o potenciálních dopadech AI na zaměstnance (zajímavý potenciál se nabízí i například v kombinaci s výsledky zprávy o AI pro Českou vládu [Fařun et al. 2018]). Mohlo by být také zajímavé prozkoumat, s přesahem do práva nebo veřejné politiky, jak je na tom například legislativní rámec pro užití AI v medicíně nebo autonomních vozidlech. Nebo i pouhý průzkum veřejného mínění – co si lidé o AI myslí, jestli se bojí spíše o práci nebo očekávají spíše zlepšení svých životů, nebo například jestli mladí lidé zvažují při volbě vysoké školy, jestli si myslí, že jejich obor bude zasažen AI. Je ale samozřejmě nespočet dalších otázek ke zkoumání ohledně AI nejen v českém kontextu. Etické otázky nejsou dořešené. Co všechno je správné automatizovat – auta? medicínu? zbraně? spravedlnost? Kde má být hranice důvěry, kterou v AI svěříme? A vzhledem k tomu, jak rychle se pole výzkumu a uplatnění AI vyvíjí, tak se jistě objeví i mnoho dalších otázek, které nás zatím nenapadly. Ostatně když jsem začínal tuto práci psát, tak mnoho z textů a událostí, se kterými zde pracuji, ještě neexistovaly. Mnoho se mění a bude měnit, a moje práce, i když jsem se snažil držet v současnosti a v realitě, nutně musela často zabrouzdat směrem ke spekulacím o blízké budoucnosti. To vše jsou podle mě dobré důvody, proč by sociologie měla AI zkoumat, a proč jsem se sám pokusil napsat malý příspěvek tímto směrem.

## Summary

The goal of this thesis was for it to be an introductory guide for sociologists who are interested in studying artificial intelligence (AI). The title of this text asks about the sociological aspects of AI. The whole text is an attempt to answer this question and offer possible examples of these sociological aspects. They could be summed up into two categories. The first category are the internal workings of AI. AI is being made by people, learning from people and reacting to people. The first category of sociological aspects thus consists of studying how AI can learn and spread values to which it is exposed. AI is not neutral and objective just because it is not human. That is why it matters *by whom* and *how* AI is made. This is also why a chapter about the technical aspects of how AI works and a part about AI reproducing inequalities are included in this text. The second category

of sociological aspects are the effects AI can have on areas where it is being applied. The part of this text going over works concerning the question of AI and employment combined with the chapter about how AI is applied in two selected fields (medicine and transport) shows the possible changes that AI can bring and that we can study. Whether people will be replaced in their jobs, what kinds of jobs are susceptible to that and when can we expect this to happen, those are examples of phenomena that we can study and consider their possible effects on societies. These two categories are thus example answers to the question of sociological aspects from the beginning of this text.

This text has three chapters which combined form an introductory overview of possible areas of interest in the studies of AI. The first chapters explain the technical aspects of the technology that has after recent breakthroughs and a subsequent wave of interest in recent years been called AI. In this text AI is defined as algorithms that are capable of making appropriate decisions based on prior experience. The two main ways the algorithms learn are machine learning and deep learning. Breakthroughs in application of these two methods are the main reason for the surge of interest in AI in recent years.

The second chapter is an overview of recent sociological works dealing with AI. Four of these works are then described in more detail as an example of the breadth of possible ways of sociological approaches to AI. Overall the authors are convinced in the large potential of social change that AI could bring. They however also criticize the way AI is being made, stressing that the data AI learns from influences its behavior. There is a danger that AI will reproduce opinions and implicit biases of the people that create it – be it the programmers or users that it learns from. The sociologists are also afraid of the effects AI will have on societies and mention the necessity to react accordingly to changes in people's lives. Specifically, this is most apparent in the two works about AI and employment [Ford 2018, Nübler 2018].

The last chapter describes two examples of the vast number of areas where AI is being applied. The first of these areas is medicine. An overview of studies comparing the applicability of AI to different parts of medicine shows that in certain tasks (especially pattern recognition in images, e.g. identifying tumors) AI can match or exceed the performance of human doctors. This is however often true only about a part of the whole treatment process and there are areas where AI is not yet capable of matching human

performance. The second area of AI application that is being discussed is transportation and autonomous vehicles. This part is mostly based on the discussion of the typology of levels of autonomy of vehicles by SAE International [SAE International 2018]. Currently available vehicles reach a maximum level four out six in this typology. For the social effects (such as use of them in cities or replacing truck driver jobs) of autonomous vehicles as discussed by sociologists to happen, higher levels of vehicle autonomy than currently available would be needed. Such vehicles are currently being tested but predictions about their commercial release vary strongly.

This text can serve as a baseline for further research. One possibility would be to take inspiration from the work of Irmgard Nübler and try asking if the past progress of automatization of production in the Czech Republic can tell us something interesting about potential effects of application of AI on Czech workers. Another option would be to look at the current legislative framework and if it is suitable for the use of AI in medicine or transport. Even a simple survey of public opinions could be interesting. What do people think of AI? Are they afraid of losing their jobs or do they expect their lives to improve? Do young people take the effects of AI on the future jobs into consideration when choosing what to study? This is just a small subset of many possible questions one might ask.

Ethical questions are yet to be solved as well. Which tasks should be automatized - cars? medicine? weapons? justice? Where do we draw the line in what trust AI to do for us? Considering the speed with which the field of AI research and application is moving many more questions will come about. When I started writing this thesis many of the works I cite here did not exist yet. Many things are changing and my text, even though I tried to keep it in the present had to venture into speculation about the close future from time to time. I believe these are all good reasons sociology should study AI and that my work has been a small step in that direction as well.

## Použitá literatura

- Adrienne Mayor. 2018. *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*. Princeton University Press.
- *AI Policy – China*. Future of Life Institute. [online] [cit. 3. 5. 2019]. Dostupné z: <https://futureoflife.org/ai-policy-china>.
- *AI vs. Doctors*. 2019. IEEE Spectrum. [online] [cit. 3. 5. 2019]. Dostupné z: <https://spectrum.ieee.org/static/ai-vs-doctors>.
- Bengio, Joshua, Aaron Courville, Ian Goodfellow. 2016. *Deep Learning*. MIT Press.
- Boudette, Neal E.. 2016. „Autopilot Cited in Death of Chinese Tesla Driver” *The New York Times*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.nytimes.com/2016/09/15/business/fatal-tesla-crash-in-china-involved-autopilot-government-tv-says.html>.
- Burscher, Bjorn, Rens Vliegthart, Claes H. De Vreese. 2015. „Using Supervised Machine Learning to Code Policy Issues: Can Classifiers Generalize across Contexts?” *The Annals of the American Academy of Political and Social Science* 659: 122–131. <https://doi.org/10.1177/0002716215569441>.
- Emerson, Robert W., Chloe Adams, Tomoyuki Nishino, Heather Cody Hazlett, Jason J. Wolff, Lonnie Zwaigenbaum, John N. Constantino, Mark D. Shen, Meghan R. Swanson, Jed T. Elison, Sridhar Kandala, Annette M. Estes, Kelly N. Botteron, Louis Collins, Stephen R. Dager, Alan C. Evans, Guido Gerig, Hongbin Gu, Robert C. McKinstry, Sarah Paterson, Robert T. Schultz, Martin Styner, Bradley L. Schlaggar, John R. Pruett Jr, Joseph Piven. 2017. „Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age.” *Science translational medicine*, 9 (393), eaag2882. <https://doi.org/10.1126/scitranslmed.aag2882>.
- Eykholt, Kevin, Ivan Evtimov, Earlene Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, Dawn Song. 2018. „Robust physical-world attacks on deep learning visual classification.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 1625-1634.
- Fařun, Martin, Zdeněk Kučera, Michal Pazour, Ondřej Pecha, Tomáš Vondrák, Luboš Král, Michal Pěchouček, Jiří Vokřínek, Alžběta Krausová, Ján Matejka, Alex



Ivančo, Eva Fialová, Veronika Žolnerčíková, Tomáš Ščerba. 2018. „Výzkum potenciálu rozvoje umělé inteligence v České republice.” *Autorský tým Technologického centra Akademie věd ČR, Českého vysokého učení technického v Praze a Ústavu státu a práva Akademie věd ČR*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.vlada.cz/assets/evropske-zalezitosti/aktualne/AI-souhrnna-zprava-2018.pdf>.

- Hawkins, Andrew J.. 2019. „California to allow testing of light-duty self-driving trucks” *The Verge*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.theverge.com/2019/4/12/18308136/california-dmv-self-driving-trucks-light-duty>.
- Heinzman, Andrew. 2019. „Why You Can’t Buy a Self-Driving Car in 2019.” *How To Geek*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.howtogeek.com/405158/when-can-i-buy-a-self-driving-car>.
- Henkel, Maria, Tamara Heck, Julia Göretz. 2018. „Rewarding Fitness Tracking—The Communication and Promotion of Health Insurers’ Bonus Programs and the Use of Self-tracking Data.” Pp. 28-49 in *International Conference on Social Computing and Social Media*. Springer, Cham. [https://doi.org/10.1007/978-3-319-91485-5\\_3](https://doi.org/10.1007/978-3-319-91485-5_3).
- Hull, Dana, Tim Smith. 2018. „Tesla Driver Died Using Autopilot, With Hands Off Steering Wheel.” *Bloomberg*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.bloomberg.com/news/articles/2018-03-31/tesla-says-driver-s-hands-weren-t-on-wheel-at-time-of-accident>.
- John O. McGinnis. 2013. *Accelerating Democracy: Transforming Governance Through Technology*. Princeton University Press.
- Joseph, Yonette. 2018. „He Drove a Tesla on Autopilot From the Passenger Seat. The Court Was Not Amused.” *The New York Times*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.nytimes.com/2018/04/29/world/europe/uk-autopilot-driver-no-hands.html>.
- Kaplan, Jerry. 2016. *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press.
- Kerlikowske, Karla, Christopher G. Scott, Amir P. Mahmoudzadeh, Lin Ma, Stacey Winham, Matthew R. Jensen, Fang Fang Wu, Serghei Malkov, V. Shane Pankratz, Steven R. Cummings, John A. Shepherd, Kathleen R. Brandt, Diana L. Miglioretti, Celine M. Vachon. 2018. „Automated and Clinical Breast Imaging Reporting and Data System Density Measures Predict Risk for Screen-Detected and Interval

- Cancers: A Case–Control Study.” *Annals of internal medicine*, 168(11), 757-765.  
<https://doi.org/10.7326/M17-3008>.
- Krizhevsky, Alex, Ilya Sutskever, Geoffrey E. Hinton. 2012. „Imagenet Classification with Deep Convolutional Neural Networks.“ *Advances in neural information processing systems*: 1097-1105. <https://doi.org/10.1145/3065386>.
  - Lim, Milton. 2018. „History of AI Winters.” *Actuaries Digital*. [online] [cit. 7. 5. 2019]. Dostupné z: <https://www.actuaries.digital/2018/09/05/history-of-ai-winters>
  - Losh, Elizabeth, Jacqueline Wernimont. 2018. *Bodies of Information: Intersectional Feminism and the Digital Humanities*. University of Minnesota Press. <https://doi.org/10.5749/j.ctv9hj9r9>.
  - Macron, Emmanuel. 2018. „Za evropskou obrodu.” *Élysée*. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.elysee.fr/emmanuel-macron/2019/03/04/za-evropskou-obrodu.cs>.
  - Mlynář Jakub, Hamed S. Alavi, Himanshu Verma, Lorenzo Cantoni. 2018. „Towards a Sociological Conception of Artificial Intelligence.” Pp. 130-139 in M. Iklé, A. Franz, R. Rzepka, B. Goertzel (eds.). *Artificial General Intelligence*. AGI 2018. Lecture Notes in Computer Science, 10999. Springer, Cham. [https://doi.org/10.1007/978-3-319-97676-1\\_13](https://doi.org/10.1007/978-3-319-97676-1_13).
  - Paus, Eva, editor. *Confronting Dystopia: The New Technological Revolution and the Future of Work*. Cornell University Press, 2018. JSTOR, [www.jstor.org/stable/10.7591/j.ctt1w0dd7x](http://www.jstor.org/stable/10.7591/j.ctt1w0dd7x).
  - Peek, Niels, Carlo Combi, Roque Marin, Riccardo Bellazzi. 2015. „Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes.” *Artificial intelligence in medicine*, 65 (1): 61-73. <https://doi.org/10.1016/j.artmed.2015.07.003>.
  - *President Donald J. Trump Is Accelerating America’s Leadership in Artificial Intelligence*. 2019. Washington D.C.: The White House. [online] [cit. 3. 5. 2019]. Dostupné z: <https://www.whitehouse.gov/briefings-statements/president-donald-j-trump-is-accelerating-americas-leadership-in-artificial-intelligence>.
  - Rajpurkar Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng. 2017. „CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.” *arXiv preprint arXiv:1711.05225*.

- Rhee, Jennifer. 2018. „Caring: Care Labor, Conversational Artificial Intelligence, and Disembodied Women.” *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*, University of Minnesota Press, Minneapolis; London, Pp. 31–66. <https://doi.org/10.5749/j.ctv62hh4x>.
- *SAFETY IS OUR FIRST PRIORITY AT EMBARK*. Embark Trucks. [online] [cit. 3. 5. 2019]. Dostupné z: <https://embarktrucks.com/safety.html>.
- Seung Seong, Han, Park Gyeong Hun, Woohyung Lim, Myoung Shin Kim, Jung Im Na, Ilwoo Park, Sung Eun Chang. 2018. „Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network.” *PloS one* 13 (1), e0191493. <https://doi.org/10.1371/journal.pone.0191493>.
- *ŠKODA OCTAVIA Komfortní asistenty*. 2019. ŠKODA AUTO a.s.. [online] [cit. 3. 5. 2019]. Dostupné z: <http://www.skoda-auto.cz/modely/octavia/octavia/octavia-komfortni-asistenty>.
- Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. 2018. SAE International.
- Theis, Thomas N., H. -S. Philip Wong. 2017. „The End of Moore's Law: A New Beginning for Information Technology.” *Computing in Science & Engineering* 19 (2), 41. <https://doi.org/10.1109/MCSE.2017.29>.
- Weng, Stephen F., Jenna Reys, Joe Kai, Jonathan M. Garibaldi, Nadeem Qureshi. 2017. „Can machine-learning improve cardiovascular risk prediction using routine clinical data?” *PloS one*, 12 (4), e0174944. <https://doi.org/10.1371/journal.pone.0174944>.
- Wood Derrick E., James R. White, Andrew Georgiadis, Beth Van Emburgh, Sonya Parpart-Li, Jason Mitchell, Valsamo Anagnostou, Noushin Niknafs, Rachel Karchin, Eniko Papp, Christine McCord, Peter LoVerso, David Riley, Luis A. Diaz Jr., Siân Jones, Mark Sausen, Victor E. Velculescu, Samuel V. Angiuoli. 2018. „A machine learning approach for somatic mutation discovery.” *Science translational medicine*, 10(457), eaar7939. <https://doi.org/10.1126/scitranslmed.aar7939>.
- Woolgar, Steve. 1985. „Why not a Sociology of Machines? The Case of Sociology and Artificial Intelligence.” *Sociology* 19 (4), 557–572, , <https://doi.org/10.1177/0038038585019004005>.

- Writer, Beta. 2019. *Lithium-Ion Batteries A Machine-Generated Summary of Current Research*. Springer, Cham. <https://doi.org/10.1007/978-3-030-16800-1>.
- *Země EU a Evropská komise budou spolupracovat na vývoji evropské umělé inteligence*. 2018. Zastoupení Evropské Komise v ČR. [online] [cit. 3. 5. 2019]. Dostupné z: [https://ec.europa.eu/czech-republic/news/181207\\_vyvoj\\_umele\\_inteligence\\_cs](https://ec.europa.eu/czech-republic/news/181207_vyvoj_umele_inteligence_cs).

# Teze bakalářské práce

Název: Sociologické aspekty umělé inteligence

Námět:

V bakalářské práci budou představeny základní obrysy relevantních efektů umělé inteligence pro sociologii. Umělou inteligencí není myšlena pseudo-lidská bytost z budoucnosti, ale v současnosti existující, působící a pozorovatelné technologie, o kterých je mluveno jako o umělé inteligenci. Podle Jerryho Kaplana zatím neexistuje jednoznačný konsensus, co přesně je nebo není umělá inteligence (Kaplan 2016: 1), ale sám definuje (nejen) umělou inteligenci jako „*schopnost včas dělat přiměřená zobecnění na základě omezeného množství dat*“ (Kaplan 2016: 5). V této práci bude umělou inteligencí myšlen software, který takovéto schopnosti vykazuje. Účelem práce je nabídnout přehled jejích efektů v různých oblastech společnosti a využít zvolenou sociologickou teorii pro utřídění těchto možných sociologických relevancí. Tento úkol se bude skládat ze tří částí. Za prvé bude kontextualizována a zdůvodněna relevance tématu umělé inteligence pro sociologii z pohledu teorie sociálních systémů Niklase Luhmanna. Za druhé bude prozkoumán současný stav oboru výzkumu umělé inteligence a její projevy v současné společnosti. Za třetí bude vybráno několik oblastí, kde se umělá inteligence uplatňuje a v nich srovnány a prozkoumány změny vnesené umělou inteligencí.

Metody a struktura práce:

Práce bude přehledovou studií relevantní odborné literatury, která nabídne jak sociologické koncepty pro vytvoření teoretického rámce práce, tak různá technická literatura z oblasti umělé inteligence a literatura zabývající se efekty AI v různých sociálních systémech. Sociologická literatura bude zastoupena zejména vybranými relevantními částmi z prací Niklase Luhmanna. Technická literatura bude zobrazovat současný stav porozumění umělé inteligenci z pohledu jejího vytváření. Poslední skupina literatury se bude týkat vybraných oblastí, kde se umělá inteligence v současnosti uplatňuje a jejím projevům.

Předběžný seznam literatury:

- Agrawal, Ajay, Joshua Gans, Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press.

- Bengio, Joshua, Aaron Courville, Ian Goodfellow. 2016. *Deep Learning*. MIT Press.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- de Prado, Marcos Lopez. 2018. *Advances in Financial Machine Learning*. Wiley.
- Kaplan, Jerry. 2016. *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press.
- Luhmann, Niklas. 1992. „What is communication?“ *Communication theory* 2 (3), 251-259.
- Luhmann, Niklas. 1995. *Social systems*. Stanford University Press.
- Luhmann, Niklas. 2012. *Theory of society (Vol. 1)*. Stanford University Press.
- Luhmann, Niklas. 2013. *Theory of society (Vol. 2)*. Stanford University Press.
- Luhmann, Niklas., Dirk Baecker, Peter Gilgen. 2013. *Introduction to systems theory*. Cambridge: Polity.
- Simoudis, Evangelos. 2017. *The Big Data Opportunity in Our Driverless Future*. Corporate Innovators, LLC.