SUSSKIND, R.: HOW TO THINK ABOUT AI: A GUIDE FOR THE PERPLEXED

KRISTÝNA MLČÁKOVÁ[1]

SUSSKIND, R.: How To Think About AI: A Guide For The Perplexed. Oxford: Oxford University Press, 2025, 209 s. ISBN 978-0-19-894194-1.

1. ÚVOD

Umělá inteligence (dále také „AI“) je aktuálním trendem snad všech odvětví a oborů.[2] S nástupem chatu GPT zájem o tento institut ještě vzrostl.[3] Odborná veřejnost se však často neshoduje v základních otázkách, které s konceptem AI souvisejí, čímž činí tuto problematiku nepřehlednou a matoucí.[4] V recenzované knize se tak Richard Susskind, poháněn právě i snahou vy-

---

[1] Mgr. Ing. Kristýna Mlčáková je prezenční studentkou doktorského studijního programu *Právo informačních a komunikačních technologií* na Ústavu práv a technologií, Právnické fakultě Masarykovy univerzity v Brně. Kontaktní e-mail: 480111@mail.muni.cz

[2] Např. zdravotnictví, vzdělání, právo, kybernetická bezpečnost apod. Viz AL KUWAITI, A. et al. A *Review of the Role of Artificial Intelligence in Healthcare*. Journal of Personalized Medicine. Multidisciplinary Digital Publishing Institute, 2023, roč. 13, č. 6, s. 951; KAUR, R., GABRIJELČIČ, D., KLOBUČAR, T. *Artificial intelligence for cybersecurity: Literature review and future research directions*. Information Fusion. 2023, roč. 97, s. 101804; OUYANG, F., JIAO, P. *Artificial intelligence in education: The three paradigms*. Computers and Education: Artificial Intelligence. 2021, roč. 2, s. 100020; WALTERS, R., NOVAK, M. *Artificial Intelligence and Law*. In: WALTERS, R., NOVAK, M., eds. Cyber Security, Artificial Intelligence, Data Protection & the Law. Singapore: Springer, 2021, s. 39–69.

[3] WANG, T. *Navigating Generative AI (ChatGPT) in Higher Education: Opportunities and Challenges*. Singapore: Springer Nature, 2023, s. 215.

[4] Příkladem může být i úvodní věta kapitoly *Co je AI?* v knize *AI:It's Scope and Limits*, kde James H. Fetzer uvádí: *„Jedním z fascinujících aspektů problematiky umělé inteligence je skutečnost, že definování přesné povahy předmětu umělé inteligence se zdá být překvapivě obtížné.“* viz FETZER, J. H. What is Artificial Intelligence? In: FETZER, J. H., ed. *Artificial Intelligence: Its Scope and Limits*. Dordrecht: Springer Netherlands, 1990, s. 3. A obdobně o více než třicet let později uvádí Jiang a kolektiv: *„Existuje mnoho definic umělé inteligence.“* JIANG, Y. et al. Quo vadis artificial intelligence? *Discover Artificial Intelligence*. 2022, roč. 2, č. 1, s. 1.

jasnit některé otazníky okolo konceptu AI,[5] pokouší čtenáři nabídnout komplexní přehled o tomto institutu, přičemž zmiňuje právě i různé, často i protichůdné přístupy, které se ve spojení s AI vyskytují.[6]

Richard Susskind je uznávaným odborníkem specializujícím se na problematiku budoucnosti právních služeb v důsledku zavádění a využívání nových technologií vč. umělé inteligence[7] a na dopad technologií na různé profese.[8] Je také renomovaným autorem[9] několika publikací právě na uvedenou problematiku.

V knize *How to Think About AI: A Guide for the Perplexed* Richard Susskind navazuje na svou předchozí literární činnosti a nabízí rámec pro ukotvení problematiky AI, nejenom samotného pojmu AI, ale také i dopadů AI, rizik se zaváděním AI spojených a různých přístupů k přemýšlení o AI. V knize autor poskytuje komplexní přehled o problematice AI, a to navíc nenásilnou formou, kdy čtenář nepotřebuje mít masivní znalosti, ať už technické, etické či právní. Publikace představuje spíše soubor filozofických úvah nad tímto fenoménem. Kniha je postavena na pěti základních pilířích, které určují strukturu celé publikace – (1) podstata AI, (2) možné myšlení o AI, (3) fungování AI, (4) rizika AI, (5) budoucnost s AI.

2. POROZUMĚNÍ AI

První část knihy je věnována pochopení konceptu AI. Je rozdělena do dvou kapitol, přičemž v první se autor věnuje historii vývoje AI a v druhé se

[5] Autor však sám upozorňuje, že řada otázek teprve v budoucnu vyvstane v souvislosti s vývojem AI. SUSSKIND, R.: *How To Think About AI: A Guide For The Perplexed*. Oxford: Oxford University Press, 2025, s. 15-16.

[6] Např. riziko příliš široké vs. příliš úzké regulace AI, přístupy k budoucímu vývoji AI apod.

[7] Např. SUSSKIND, R. *Online courts and the future of justice.* Oxford: Oxford University Press, 2019; SUSSKIND, R. *The future of law: facing the challenges of information technology*. Oxford: Oxford University Press, 1998.

[8] Např. knihy: SUSSKIND, R. *Tomorrow's lawyers: an introduction to your future.* Oxford: Oxford University Press, 2023; SUSSKIND, R. a SUSSKIND, D. *The future of the professions: how technology will transform the work of human experts.* Oxford: Oxford University Press, 2022.

[9] Řadí se ke světově nejcitovanějším autorům zabývajících se danou problematikou. *Richard Susskind Biography*. Richard Susskind [online]. 2025. [cit. 2025-06-22] Dostupné z: https://www.susskind.com/

snaží tento institut zasadit do širšího kontextu rozvoje digitálních technologií. Autor konstatuje, že definice AI lze rozdělit na dva hlavní proudy – v jednom dochází k popisování toho, jak AI funguje a v druhém, co AI dělá.[10] Autor v další části odpovídá na otázku, *jak AI funguje*, a to tak, že se věnuje stručné historii vývoje.[11] Autor upozorňuje zejména na skutečnost, že v současnosti se vývoj prudce zrychluje a změny lze očekávat zhruba každých šest až dvanáct měsíců.[12] Při zasazování do širšího kontextu rozvoje digitálních technologií autor uvádí pět základních myšlenek – (1) stroje jsou ohromně schopné, (2) podpůrné technologie systémů AI se také neustále zlepšují, (3) neexistuje finální stav, ke kterému chceme dospět, (4) budoucnost nám ovlivní technologie, které ještě nebyly vynalezeny, (5) už nyní plně nerozumíme tomu, jak AI funguje.[13] Autor dále rozebírá problematiku zaujatosti AI, přičemž toto spojuje se dvěma pojmy – technologická krátkozrakost (tj. soustředění se pouze na aktuální verzi a limitace AI) a iracionální odmítání AI.[14] V knize jsou také diskutovány budoucí schopnosti AI, přičemž zde autor uvádí šest hypotéz možného budoucího vývoje – hypotéza humbuku, hypotéza GenAI+, AGI (*artificial general intelligence*) hypotéza, hypotéza *superinteligence* a hypotéza singularity.[15]

3. JINÉ PŘÍSTUPY K MYŠLENÍ O AI

V druhé části knihy autor představuje dva proudy myšlení o AI, první se zaměřuje na výsledek AI systémů a druhý se zaměřuje na procesy AI systémů. Kromě představení těchto přístupů poté v druhé kapitole autor také uvádí různá pochybení, která plynou z nerespektování rozdílností těchto přístupů. V závěrečné kapitole pak diskutuje nedostatečnou terminologii pro problematiku AI a limity systémů AI. Sjednocující myšlenkou druhé části je rozli-

[10] SUSSKIND, R.: *How To Think About AI: A Guide For The Perplexed*. Oxford: Oxford University Press, 2025, s. 21. Autor nabízí i svou vlastní definici – tamtéž, s. 24.

[11] Tamtéž, s. 25.

[12] Tamtéž, s. 34-35.

[13] Tamtéž, s. 36-41.

[14] Tamtéž, s. 42-44.

[15] Tamtéž, s. 44-50.

šení dvou přístupu tzv. *outcome-thinking a process-thinking*.[16] Autor poukazuje na zajímavé nedostatky v debatách o AI – (1) nerespektování rozdílů mezi *outcome-thinking a process-thinking,* (2) myšlení typu *nás se to netýká*, (3) omyl s AI, kdy si myslíme, že AI musí napodobovat lidskou činnost a kdy požadujeme od AI vysvětlení svých procesů.[17] V závěrečné kapitole se pak autor věnuje myšlence, že pro problematiku AI nemáme potřebnou terminologii, což autor sám řeší zavedením předpony *kvazi*.[18]

4. ZAJIŠTĚNÍ FUNGOVÁNÍ AI

Třetí část práce už nese praktičtější charakter. Autor zde uvádí nutnost vhodného uchopení aplikace AI systémů**.** Na tomto místě autor zkoumá a vysvětluje rozdíly mezi pojmy automatizace, inovace (využití technologií způsobem, který nám umožní, co dříve nebylo možné) a eliminace (představuje situaci, kdy technologie zcela odstraní problém, který jsme dříve řešili).[19] Následně uvádí konkrétní příklady k těmto pojmům, např. pro právo uvádí jako příklad automatizace efektivní využívání technologií v právní praxi, jako příklad inovace uvádí ODR[20] a jako příklad eliminace pak tzv. *dispute avoidance*, tedy snahu se sporům zcela vyhnout.[21] Na závěr se autor zabývá možností zavádění AI do již existujících podniků – autor uvádí, že jediná změna, která může v souvislosti se zapojením AI v rámci podniku proběhnout za jeho provozu, je automatizace.[22] Tím také poukazuje na sku-

[16] Tamtéž, s. 53-57.

[17] Tamtéž, s. 59-66.

[18] Konkrétně jde o pojmy kvazi-úsudek, kvazi-empatie, kvazi-kreativita. Autor chce tímto zdůraznit nemožnost aplikování terminologie ze světa lidí na AI systémy, jelikož jednoduše řečeno jejich podstata je jiná, a tedy i charakter uvedených pojmů bude zákonitě u AI a lidí jiný. Tamtéž, s. 71-79.

[19] Tamtéž, s. 82-84.

[20] V souvislosti s ODR upozorňuje na skutečnost, že ODR odstraňuje nutnost slovního projevu, což v konečném důsledku může představovat riziko pro právnickou profesi. Tamtéž s. 84-86.

[21] Tamtéž, s. 84-88.

[22] Tamtéž, s. 89.

tečnost, že v současnosti se stále soustředíme spíše na automatizaci, než na inovaci či eliminaci, což také kritizuje.[23]

5. ZVLÁDÁNÍ RIZIK

Ve čtvrté části se autor věnuje rizikům spojeným s AI. Nejprve rizika představuje a detailně popisuje a následně navrhuje i možnosti, jak rizika vyřešit či jim předcházet. Než se autor věnuje konkrétním kategoriím rizik rozebírá tři přístupy k rizikům AI – přehnaně pozitivní, negativní a poté nevyhraněný středový názor.[24] Rizika autor rozděluje do sedmi kategorií – existencionální riziko, riziko katastrofy, politické riziko, socio-ekonomické riziko, riziko nespolehlivosti, riziko závislosti a riziko nečinnosti. Jako existenciální označuje rizika, která jsou schopna způsobit zánik lidstva či kolaps sociálních struktur,[25] v případě rizik katastrofy jde o systémy, které mají schopnost způsobit cokoliv, co označujeme za katastrofické (ztrátu života, zranění, škodu apod.).[26] U politických rizik rozlišuje rizika související s mocí, s demokracií a se svobodou.[27] Socio-ekonomická rizika zahrnují problémy, které AI urychluje (jako příklad uvádí nezaměstnanost, zavedení nerovností mezi lidmi, kteří přistup k AI mají a těmi, co nikoliv a zvětšování rozdílů mezi bohatými a chudými).[28] V souvislosti s rizikem nespolehlivosti se autor zabývá zejména zaujatostí AI systémů a v případě rizika závislosti upozorňuje na narůstající závislost na AI systémech.[29] Ačkoliv se v souvislosti s výše uvedeným může zdát, že autor odrazuje od zavádění AI systémů, není tomu tak, což umocňuje poslední kategorie rizik, a to riziko nečinnosti, tedy nevyužívání AI a potenciálu, který nabízí. Pro hodnocení rizik poté navrhuje zkoumání tří aspektů – pravděpodobnost

[23] Tamtéž, s. 91.

[24] Tamtéž, s. 96.

[25] Tamtéž, s. 97-98.

[26] Tamtéž, s. 98.

[27] Tamtéž, s. 99-101.

[28] Tamtéž, s. 101-103.

[29] Tamtéž, s. 103-106.

realizace rizika, závažnost dopadu a časový rámec.[30] Za největší výzvu současnosti autor považuje balancování výhod a hrozeb umělé inteligence.[31]

Součástí čtvrté části knihy jsou i úvahy autora, jak daná rizika řešit. Autor v tomto směru navrhuje (1) multidisciplinární přístup k rizikům, (2) systematické myšlení o rizicích, které vyžaduje dva aspekty, a to stanovení toho, co chceme a nalezení nejefektivnějšího způsobu prosazování našich preferencí, (3) navrhuje myšlení do budoucna, tedy myšlení, co kdyby AI byla již nyní plně schopna konkurovat lidem, (4) zabudování etiky, regulace a omezení přímo do AI systémů.[32] Zároveň však upozorňuje na rizika, která mohou vyvstat právě ze snahy vyřešit daná rizika, jako např. tzv. spěch na regulaci, volné používání pojmu regulace a s tím související povrchnost, příliš úzké zaměření předpisů (a naopak příliš široké zaměření), omezování debaty na to, co je aktuálně technicky možné apod.[33]

6. ÚVAHY NAD BUDOUCNOSTÍ

V závěrečné části autor uvádí své úvahy nad budoucností lidstva s AI. Nejprve diskutuje, zda mohou mít stroje vědomí. V další části autor nabádá k širšímu zaměření do budoucna – vnímá potenciál i v dalších technologiích. V závěrečné části pak autor navrhuje, abychom odbornou diskusi zaměřili i na evoluční pojetí umělé inteligence. Předně se autor zabývá problematikou neuchopitelnosti pojmu vědomí AI, s čímž souvisí i otázky právní odpovědnosti, morální odpovědnosti atd.[34] Autor rozporuje vhodnost srovnávání lidského vědomí s vědomím AI a přiklání se k zavedení nového pojmu *kvazi-vědomí*.[35] Dále se zabývá dalšími technologiemi schopnými ovlivnit naši budoucnost, mezi tyto řadí virtuální realitu a rozhraní mozek-počítač.[36] V závěrečné kapitole poté autor zkoumá tři roviny AI – evoluci, možnosti, které před lidstvem stojí a povinnosti lidstva do budoucna.

[30] Tamtéž, s. 107.

[31] Jak ostatně uvádí i na samotném začátku knihy – tamtéž, s. 14 a 107.

[32] Tamtéž, s. 112-131.

[33] Tamtéž.

[34] Tamtéž, s. 133 a násl.

[35] Tamtéž, s. 137-143.

V souvislosti s evolucí autor uvádí, že AI v budoucnu překoná lidi a v tomto směru tedy stojíme na prahu nové etapy.[37] Dále diskutuje čtyři možné budoucí varianty – úplné převzetí AI, spojení člověka s AI, tzv. společný podnik (tj. AI a lidé zůstanou fyzicky oddělení, ale AI bude dominantní stranou našeho společenství) a ukončení AI (autor nepovažuje za proveditelné).[38] Autor upozorňuje, že budoucnost závisí na tom, jak bude lidstvo jednat v následujících letech.[39] V souvislosti s povinnostmi do budoucna autor zejména nabízí dvě možnosti, buď budeme obhajovat primát lidí, nebo AI.[40] Autor uzavírá celé své pojednání poselstvím, že je důležité podporovat AI, zároveň je však nutné chránit budoucí lidské bytosti a vše, co je pro nás na Zemi nádherné a zachovávat to.[41]

7. HODNOCENÍ

Z textu je patrná autorova odborná erudice a důkladná znalost problematiky i několika přístupů k uchopení AI. V rámci textu autor často prezentuje názory jiných odborníků, své vlastní myšlenky i zkušenosti. Autor se věnuje několika oblastem, které považuje za stěžejní – zkoumá samotnou podstatu AI, přístupy k přemýšlení o AI a také rizika, které AI představuje a může představovat. Za mimořádně pozitivní aspekt vnímám autorovu snahu uvažovat v dlouhodobém horizontu. Autorovy úvahy *pro futuro*[42] jsou rovněž zajímavým přínosem knihy a rozhodně ve světle enormně rychlého vývoje technologií autora v tomto směru je třeba ocenit. Autor navíc rád zjednodušuje přemíru existujících názorů a nachází jejich společné aspekty, čímž je např. schopen zjednodušit způsoby myšlení o AI na dva základní proudy, což čtenář jistě ocení. Tato schopnost autora se projevuje v celé jeho knize,

[36] Tamtéž, s. 145-152. Autor si je v tomto směru vědom opadu zájmu o virtuální realitu s nástupem AI, v tomto ohledu se však opět projevuje jeho dlouhodobý přístup, kdy se zabývá tím, zda do budoucna díky využívání AI, a tedy značné úspoře času, se lidé opět nezačnou o virtuální realitu zajímat. Tamtéž, s. 151-152.

[37] Tamtéž, s. 158-161.

[38] Tamtéž, s. 161-165.

[39] Tamtéž, s. 165.

[40] Tamtéž, s. 165-167.

[41] Tamtéž, s. 170.

[42] Např. stran rizik AI, možného budoucího vývoje atd. K tomu viz výše.

a tak z nepřehledné změti názorů, přístupů a možností činí jednoduchý přehled pro orientaci v problematice.

Výtkou z mé strany je přílišná prezentace autorových názorů a přesvědčení. Autor napříč knihou prezentuje své názory, hojně je obhajuje a při kritice určitých konceptů neopomíjí zdůraznit, že jde o jiný přístup, než on sám zastává. Uvedené však nepovažuji za neúnosné vzhledem ke skutečnosti, že autor vždy čtenáři poskytne celé penzum možných přístupů a odpovědí na předmětné otázky a veškeré své názory a postoje podrobně argumentuje. Je tedy sice pravda, že autorův názor je tak z textu velmi patrný i explicitně uvedený, ale čtenář je obeznámen s různými jinými existujícími variantami. Dalším předmětem kritiky je autorova snaha provázat jednotlivé kapitoly mezi sebou. Opět se však nejedná o aspekt, který lze pouze kritizovat. Je pravda, že tak text působí v určitém ohledu repetitivně, na druhou stranu jsou tak po celou dobu zdůrazňovány základní myšlenky autora.

8. ZÁVĚR

Richard Susskind ve své knize komplexně pojímá problematiku AI. Jeho cílem není zahrnout čtenáře technickými parametry a znalostmi z tohoto či jiných oborů (např. právo, etika). Jeho text je přístupný čtenářům z jakékoliv oblasti, právě díky tomu, že autor čtenáře netrápí složitými pojmy a definicemi. Nejedná se o typický právnický text, autor se nevěnuje analýze konkrétních předpisů, nekritizuje konkrétní právní uchopení institutu AI.[43] Jde spíše o filosofické úvahy typu, *„kam směřujeme?“ a „kam chceme směřovat?“*. Jeho text je velmi čtivý a srozumitelný, jediným problematickým aspektem pro čtenáře tak může být složitější akademická úroveň anglického jazyka. Až na tuto skutečnost je ale kniha psána s rozmyslem a logickou strukturou, které je snadné následovat. Autor navíc v průběhu textu logicky navazuje na již vyřčené, což může působit místy repetitivně, ale na druhou stranu je tím zajištěno udržení základních myšlenek knihy v paměti čtenáře. Pokud bych měla knize něco vytknout, pak by to byl nejspíš velmi evidentní autorův postoj, který se projevuje v celém textu. Autor vždy vy-

[43] Právní ukotvení pojmu AI komentuje pouze stručně viz tamtéž, s. 120-123.

jádří větší sympatie k určitému přístupu a v dalším textu je velmi evidentní jeho názor na další postoje (např. procesní a objektové smýšlení o AI). Vzhledem ke skutečnosti, že Richard Susskind však vždy nabídne čtenáři celé spektrum různých přístupů a je tak jen a jen na čtenáři, který z těchto přístupů je jeho přesvědčení nejbližší, to však nepovažuji za tak problematický rys knihy. Skutečnost, zda se čtenář přikloní k názoru autora, který je v knize hojně argumentován a obhajován, či si zvolí jiný z představených přístupů tak zůstává plně na čtenáři. Celkově vzato bych tak knihu doporučila každému se zájmem o umělou inteligenci.
