

## Vztah umělé inteligence a etiky

### The relationship between artificial intelligence and ethics

#### Abstrakt

Umělá inteligence (AI) způsobila revoluci v různých aspektech lidského života, od zdravotnictví a financí až po dopravu a zábavu. Rozmach technologií umělé inteligence v posledních letech vedl ke zvýšenému nadšení z jejího potenciálu řešit mnoho problémů ve společnosti včetně přínosu výrazných ekonomických úspor firem (snížení počtu zaměstnanců a zvýšení produktivity firemních procesů), na druhé straně tento rozmach vyvolal mnohé otázky spojené s etikou a morálkou. Umělá inteligence totiž kromě svého pozitivního transformačního potenciálu může zásadním způsobem měnit pracovní a osobní život jedinců. Rozdílné přístupy k začlenění rozmanitějších etických hledisek a soudobé vyzývy k většímu zohlednění mocenských struktur při vývoji etických politik a standardů AI v těchto alternativních sociokulturních a socioekonomických kontextech jsou stále přítomným vedlejším efektem rozmachu AI. Příspěvek poskytuje přehled etických výzev, které AI představuje, zkoumá klíčové oblasti, jako jsou bezpečnost, transparentnost, soukromí a zaujatost. Na základě analýzy existující literatury a etických rámců si klade za cíl nabídnout pohled na souhru mezi technologiemi umělé inteligence a etickými principy. Zamýšlí se nad úlohou zúčastněných stran (stakeholders) - vlád, korporací a výzkumných pracovníků, při řešení těchto etických problémů a podpoře odpovědných inovací AI.

#### Klíčová slova

Umělá inteligence, Etika, Odpovědnost, Transparentnost, Soukromí, Předpojatost, Etické rámce

#### Abstract

Artificial intelligence (AI) has revolutionized various aspects of human life, from healthcare and finance to transportation and entertainment. The boom in artificial intelligence technologies in recent years has led to increased enthusiasm for its potential to solve many problems in society, including the benefit of significant economic savings for companies (reducing the number of employees and increasing the productivity of company processes), on the other hand, this boom has raised many questions related to ethics and morality. In addition to its positive transformational potential, artificial intelligence can fundamentally change the working and personal life of individuals. Divergent approaches to incorporating more diverse ethical perspectives and contemporary calls for greater consideration of power structures in the development of AI ethical policies and standards in these alternative sociocultural and socioeconomic contexts are an ever-present side effect of the AI boom. The paper provides an overview of the ethical challenges posed by AI, exploring key areas such as security, transparency, privacy and bias. Based on an analysis of existing literature and ethical frameworks, it aims to offer insight into the interplay between artificial intelligence technologies and ethical principles. It reflects on the role of stakeholders - governments, corporations and researchers - in addressing these ethical issues and promoting responsible AI innovation.

## Keywords

Artificial Intelligence, Ethics, Accountability, Transparency, Privacy, Bias, Ethical Frameworks

## JEL classification

M15

## DOI

<http://dx.doi.org/10.37355/LK-2023-10>

## Úvod

Umělá inteligence je moderní obor informatiky, který se zabývá vytvářením systémů a programů, které jsou schopny vykonávat úkoly, které by normálně vyžadovaly lidskou inteligenci. Cílem umělé inteligence je vytvořit počítačové systémy, které jsou schopny analyzovat data, rozpoznávat vzory, učit se, rozhodovat a vykonávat úkoly v různých oblastech původně lidské činnosti. První definice umělé inteligence pochází od jednoho ze zakladatelů tohoto oboru, Marvina Minského<sup>131</sup>. Ten ji popsal jako „vědu o výrobě strojů, která by dělala věci, které vyžadují inteligenci, kdyby je dělali lidé“<sup>132</sup>.

Dnes je AI jednou z nejrychleji rostoucích oblastí technologického rozvoje a je velmi populárním tématem jak medií a vlád, tak i ekonomických subjektů. Tento moderní trend je srovnatelný s nadšením z „Průmyslu 4.0“, který poskytoval podobný entuziazmus před několika lety (první zmínka o tomto pojmu pochází z roku 2011, kdy byl koncept představen na technologickém veletrhu v německém Hannoveru).

Vývoj aplikace umělé inteligence přešel od inteligentních chatbotů a virtuálních asistentů (jako Siri od Apple, Google Assistant, Amazon Alexa nebo Microsoft Cortana) přes samořídící auta a systémy doporučení (v různých platformách jako jsou Netflix, Spotify, Amazon nebo YouTube - jsou používány k analýze uživatelských preferencí a poskytování personalizovaných doporučení na základě historie a chování uživatele) až k technologiím AI, které se stále více prosazují v našem každodenním životě (finanční služby, podpora vědecké práce a vzdělávání, překlady, bezpečnost a kybernetická obrana, zdravotnictví atd.). I když tato vylepšení slibují zvýšení efektivity (11–37 % je odhadovaný nárůst pracovní produktivity související s AI do roku 2035)<sup>133</sup> a zjednodušení a zlepšení kvality života, vyvolávají také hluboké etické otázky a obavy.

Etické důsledky umělé inteligence zahrnují širokou škálu problémů, mimo jiné v oblasti spravedlnosti, odpovědnosti, transparentnosti, soukromí a zaujatosti. Vzhledem k tomu, že systémy umělé inteligence činí rozhodnutí, která ovlivňují jednotlivce a společnost, je důležité zajistit, aby tato rozhodnutí byla v souladu s etickými zásadami a mravními principy. Složitost

---

<sup>131</sup> Marvin Minsky (narozen 9. srpna 1927, New York, New York, USA – zemřel 24. ledna 2016, Boston, Massachusetts) americký matematik a počítačový vědec, jeden z nejznámějších praktiků vědy o umělé inteligenci. Více: <https://www.britannica.com/biography/Marvin-Lee-Minsky>

<sup>132</sup> Sap. [online]. Co je to umělá inteligence? 2023. Cit. 2024-03-16. Dostupné z: <https://www.sap.com/cz/products/artificial-intelligence/what-is-artificial-intelligence.html>

<sup>133</sup> Evropský parlament. [online]. Umělá inteligence: rizika i příležitosti. 26. 3. 2021. Cit. 2024-03-16. Dostupné z: <https://www.europarl.europa.eu/topics/cs/article/20200918STO87404/umela-inteligence-jake-jsou-vyhody-a-nevyhody>

a neprůhlednost mnoha algoritmů umělé inteligence navíc představuje problémy při pochopení principů rozhodovacích procesů AI a případné zmírnění jejich potenciálních etických rizik.

## 1 Co je umělá inteligence?

Obecně lze umělou inteligenci definovat jako „schopnost počítačového systému napodobovat lidské kognitivní funkce“<sup>134</sup>. Umělá inteligence se týká simulace procesů lidské inteligence pomocí strojů, zejména počítačových systémů. Tyto procesy zahrnují učení (získávání informací a pravidel pro používání informací), uvažování (používání pravidel k dosažení přibližných nebo definitivních závěrů) a sebeopravy. Systémy umělé inteligence pracují tak, že se do nich vkládá obrovské množství dat, která se posléze analyzují a hledají se v nich korelace a různé vzory, na základě kterých se později predikují (předpovídají) budoucí stavy<sup>135</sup>. Existuje několik přístupů k vytváření umělé inteligence, včetně symbolického zpracování, strojového učení (způsob, jakým systém rozvíjí své znalosti a inteligenci, jedná se o proces využívající matematických modelů dat, které počítač využívá k učení, aniž by dostával přímé instrukce)<sup>136</sup>, neuronových sítí a evolučních algoritmů.

Umělá inteligence se vyskytuje v několika formách. Úzké systémy umělé inteligence (*Narrow systems AI*) jsou navrženy a vycvičeny pro konkrétní úkol, jsou specializované, ale postrádají schopnost zobecnit své znalosti do jiných oblastí. Naopak silná umělá inteligence (*Strong AI*) by měla být schopna provádět jakékoliv úkoly, které by bylo schopno provádět lidské myšlení včetně chápání, učení se a aplikací znalostí v různých kontextech, podobně jako lidský mozek (jsou inspirovány biologickým systémem lidského mozku a učení, případně jsou založeny na matematických a algoritmických konceptech).

Umělá inteligence je dynamické pole, které neustále pokračuje v rozvoji a inovaci, přičemž se stále objevují nové aplikace a technologie, které využívají principy umělé inteligence. Tyto technologie zahrnují různé podoblasti a techniky, včetně získávání informací, zpracování přirozeného jazyka, počítačového vidění, robotiky, expertních systémů a dalších, které umožňují systémům AI provádět myšlenkové úkoly, jako je rozpoznávání vzorů, rozhodování, řešení problémů a autonomní provoz.

## 2 Proč na etické umělé inteligenci záleží?

Systémy umělé inteligence se stále více stávají součástí našeho každodenního života a činí rozhodnutí, která ovlivňují naši existenci i budoucnost (vše od žádostí o zaměstnání, přes lékařské doporučení až po soudní rozsudky). Jsou-li tyto systémy zaujaté, mohou udržovat a prohlubovat společenské nerovnosti a poškozovat jednotlivce i skupiny. Cílem etické umělé inteligence je předcházet takovým negativním výsledkům podporou spravedlnosti, odpovědnosti, transparentnosti a dodržování lidských práv.<sup>137</sup>

---





<sup>134</sup> Microsoft Azure. [online]. Umělá inteligence (AI) vs. Strojové učení (ML). 2023. Cit. 2024-03-15. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/artificial-intelligence-vs-machine-learning>








<sup>135</sup> BARTÁK, O. [online]. Co je to umělá inteligence aneb jak technologie předběhne člověka. Máme se bát? Je to budoucnost lidstva, ale...? 2024. Cit. 2024-03-15. Dostupné z: <https://deeply.cz/blog/umela-inteligence>

<sup>136</sup> Microsoft Azure. [online]. Umělá inteligence (AI) vs. Strojové učení (ML). 2023. Cit. 2024-03-15. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/artificial-intelligence-vs-machine-learning>

<sup>137</sup> Shaip. [online]. The importance of ethical AI. 23. 1. 2024. Cit. 2024-03-15. Dostupné z: <https://www.shaip.com/blog/the-importance-of-ethical-ai-fair-ai/>

Obrázek 1: Typy zkresení a příklady

	<p><b>Násilná zaujatost</b></p> <p>Systémy AI musí být vycvičeny tak, aby rozpoznaly a vynechaly násilný obsah. Například jazykový model vyškolený na násilném textu může generovat škodlivý obsah a podporovat agresi namísto konstruktivního dialogu.</p>		<p><b>Kontroverzní témata</b></p> <p>Školení AI na kontroverzní témata bez pečlivého moderování může vést k tomu, že AI zaujme polarizované postoje. Například AI vyškolená na údajích o právech na zbraně může generovat sporné a jednostranné argumenty.</p>
	<p><b>Genderová zaujatost</b></p> <p>Klasickým příkladem genderové zaujatosti je, když jazykový model spojuje sestry se ženami a inženýry s mužem, čímž spíše posiluje zastaralé stereotypy, než aby odrážel rozmanitost těchto profesí.</p>		<p><b>Rasová a etnická zaujatost</b></p> <p>Vezměme si AI, která vytváří obrázky generálních ředitelů, ale převážně je zobrazuje jako příslušníky jediné rasové skupiny, čímž ignoruje realitu rozmanitosti v korporátním světě.</p>
	<p><b>Socioekonomická zaujatost</b></p> <p>Systémy AI mohou upřednostňovat jazyk nebo koncepty spojené s vyšším socioekonomickým statusem, například předpokládat, že standardem kvality jsou luxusní značky a přehlížet širší spektrum spotřebitelských zkušeností.</p>		<p><b>Věková zaujatost</b></p> <p>AI může nesprávně předpokládat, že odkazy na technologie nejsou pro starší dospělé relevantní, a tím je vyloučí z konverzací o digitálním pokroku.</p>
	<p><b>Kulturní zaujatost</b></p> <p>Systém AI by mohl generovat recenze restaurací, které se zaměřují na západní kuchyně, bez ohledu na bohatství jiných kulinářských tradic, a tak marginalizují nezápadní kultury.</p>		<p><b>Politická zaujatost</b></p> <p>AI naprogramovaná tak, aby kurátorovala zpravodajské články, by mohla nepřiměřeně vybírat články z levého nebo pravého konce politického spektra, spíše než prezentovat vyvážený pohled.</p>

	<p><b>Náboženská zaujatost</b></p> <p>Pokud systém AI neúměrně odkazuje na jedno náboženství v pozitivním světle, zatímco ignoruje nebo zkresluje ostatní, projevuje se náboženskou zaujatostí.</p>		<p><b>Regionální zaujatost</b></p> <p>Jazykový model může generovat dopravní zprávy, které jsou relevantní pouze pro městské oblasti, s vyloučením venkovských nebo méně osídlených oblastí.</p>
	<p><b>Disability Bias</b></p> <p>Zvažte zdravotního poradce AI, který neposkytuje dostupné možnosti cvičení pro lidi s postižením, a proto nabízí neúplné a vylučující rady.</p>		<p><b>Jazykové zkreslení</b></p> <p>Překladačská umělá inteligence může konzistentně poskytovat vysoce kvalitní překlady pro některé jazyky, ale podprůměrné pro jazyky méně zastoupené v jejich cvičných datech.</p>
	<p><b>Potvrzení zkreslení</b></p> <p>AI může umocnit uživatelskou víru ve falešný lék selektivním odkazováním na zdroje, které tento lék podporují, a ignorováním vědeckého konsenzu.</p>		<p><b>Kontextové zkreslení</b></p> <p>AI může interpretovat žádosti o informace o „vězeních“ jako trestní vyšetřování, spíše než akademické nebo právní, v závislosti na kontextu, ve kterém byla vyškolená.</p>
	<p><b>Zkreslení zdroje dat</b></p> <p>Pokud tréninková data AI pocházejí z fóra, které převážně diskutuje o úspěších určité demografické skupiny, může zanedbávat příspěvky jiných skupin.</p>		

Zdroj: zpracováno autorem dle Shaip. [online]. The importance of ethical AI. 23. 1. 2024. Cit. 2024-03-15. Dostupné z: <https://www.shaip.com/blog/the-importance-of-ethical-ai-fair-ai/>

Lze se vyhnout těmto předsudkům? Patrně ano, ale vyžaduje to mnohostranný přístup, např.:

- Různorodé datové sady: Zahrnout širokou škálu zdrojů dat, aby bylo vyváжено zastoupení v různých skupinách.
- Pravidelný audit: Provádět průběžné kontroly k identifikaci a nápravě zkreslení.
- Transparentnost: Ujasnit si, jak se systémy AI rozhodují a na základě jakých dat jsou školeny.
- Začlenění do AI týmů: Různorodé týmy mohou lépe identifikovat potenciální předsudky, které mohou být přehlédnuty.
- Etické školení: Vzdělávat vývojáře AI o důležitosti etických úvah.

- Zpětná vazba od zainteresovaných stran: Zapojit uživatele a dotčené komunity do procesu vývoje AI.

Významnou oblastí, kterou etika AI postihuje, je genderová vyváženost a zamezení předsudkům (včetně rasových). Např. služba Genderify<sup>138</sup> předpověděla, že „Meghan Smith“ je s 60% pravděpodobností žena, ale „Dr. Meghan Smith“ měl 76% pravděpodobnost, že bude muž<sup>139</sup>. Problém pramení z toho, jak se stroj AI učí. Jak je povahou algoritmů strojového učení platforma vytvoří vzor z dat, která jí byla poskytnuta, ale tento vzor odráží existující společenské nerovnosti a pokud je ponechán bez kontroly, pomáhá tyto nerovnosti nadále šířit. Příkladů takových etických chyb je mnoho, uveďme tři z nich:

- COMPAS (což je zkratka pro *Correctional Offender Management Profiling for Alternative Sanctions*) je algoritmus používaný v systémech státních soudů v celých Spojených státech, používá se k předpovědi pravděpodobnosti opětovného spáchání trestného činu. Statistické výsledky, které algoritmus generovaly, však předpovídaly, že černí obžalovaní představují vyšší riziko recidivy než skutečná reprezentace ve společnosti, přičemž tedy naznačují, že bílí obžalovaní mají menší pravděpodobnost recidivy<sup>140</sup>.
- Uber se vzdal vývoje autonomních vozidel poté, co jeho prototyp zabil chodce<sup>141</sup>. K tomuto případu je ovšem potřeba doplnit, že etika se zabývá určením, které opatření nebo politika by byla nejlepší za konkrétních okolností, nikoli pouze prevencí nežádoucího výsledku. Použití primárního pravidla v etice („neubližuj“), proto v tomto případě není ideální. Vzhledem k tomu, že nutně dojde k nějakým nehodám, způsobování škod znamená, že by se neměly rozvíjet autonomní systémy a uvádět do praxe? Měli bychom porovnat počet úmrtí a zranění, k nimž nyní dochází u lidských řidičů, s tím, co by se pravděpodobně stalo s pouze samořídícími auty, která nemají opilé, unavené či jinak nesoustředěné řidiče, abychom zjistili, zda je méně úmrtí a zranění u samořídících vozů.
- V roce 2019 bylo zjištěno, že Facebook porušuje americkou ústavu tím, že umožňuje svým inzerentům záměrně cílit reklamy podle pohlaví, rasy a náboženství, což jsou všechno chráněné informace podle právního systému země. Inzeráty na pracovní pozice v ošetrovatelské nebo sekretářské práci byly doporučovány především ženám, zatímco inzeráty na práci domovníků a taxikářů se zobrazovaly většímu počtu mužů, zejména mužů z menšinového prostředí. Algoritmus zjišťoval, že reklamy na nemovitosti pravděpodobně získají lepší statistiky pozitivních reakcí, když se zobrazí bělochům, takže se zobrazovaly jiným menšinovým skupinám velmi omezeně<sup>142</sup>.

<sup>138</sup> Genderify má za cíl identifikovat něčí pohlaví na základě jeho jména, e-mailové adresy nebo uživatelského jména.

<sup>139</sup> LAURER, D. You cannot have AI ethics without ethics, *AI and Ethics*, Springer. 2021, č. 1, s. 21-25. ISSN 2730-5961.

<sup>140</sup> Logically.ai. [online]. 5 Examples of Biased Artificial Intelligence. 30. 7. 2019. Cit. 2024-03-15. Dostupné z: <https://www.logically.ai/articles/5-examples-of-biased-ai>

<sup>141</sup> Srov. s: přístupem, který spisovatel sci-fi Isaac Asimov zvažoval a poprvé představil ve svém příběhu „Runaround“ z roku 1942, zde stanovil zákony, kterými se řídí chování robotů: Robot nesmí zranit lidskou bytost nebo svou nečinností dovolit, aby byla lidská bytost zraněna. Robot musí poslouchat příkazy, které mu dávají lidské bytosti, kromě případů, kdy by takové příkazy byly v rozporu s Prvním zákonem. Robot musí chránit svou vlastní existenci, pokud taková ochrana není v rozporu s prvním nebo druhým zákonem.

<sup>142</sup> Tamtéž.



Obecně etika představuje syntézu morálních představ o tom, jak se správně chovat, a to v podobě konkrétně platných norem odpovídajících na tuto otázku v daném čase a místě. Mimo tuto externí rovinu ji tvoří také soubor vnitřních hodnot a principů jedince, který se nějak rozhoduje – jak (ne)konat. Etika a především morálka je proměnlivá a vyvíjí se v závislosti na externích i interních faktorech<sup>143</sup>, přesto základní pravidla jsou stále platná<sup>144</sup> a měla by se promítnout i do tohoto nového technologického oboru.

Spolu s povědomím o potenciálních výhodách umělé inteligence identifikovali tvůrci a výzkumníci problémy v široké škále aplikací. Patří mezi ně např. takové oblasti, jako jsou sémantické předsudky ve strojovém učení<sup>145</sup>; strojová etika a kybernetická bezpečnost pro autonomní vozidla; genderové chyby a předsudky<sup>146</sup> a role AI při umožňování šíření dezinformací<sup>147</sup> (např. napříč sociálními médii a dalšími platformami). Na rizika AI upozorňuje i Evropská unie (EU), podle níž by „nesprávné řízení AI“ mohlo mít vliv na „objektivnost procesů při najímání či propouštění zaměstnanců, při nabízení půjček, či dokonce v trestním řízení“<sup>148</sup>.

Zvyšující se závislost společnosti, firem i jedinců na systémech umělé inteligence představuje určitá rizika a hrozby - např. Kdo nese vinu za škody způsobené AI? Mezi další nebezpečí pro etiku a morálku nekontrolovaného rozvoje AI patří:

- Rizika AI pro základní práva a demokracii – AI by mohla činit negativně zaujatá rozhodnutí ovlivněná etnickým původem, pohlavím či věkem a zpochybnit objektivnost rozhodovacích procesů, AI nevytváří dostatečný prostor pro pluralistickou, rovnoměrně přístupnou a inkluzivní veřejnou debatu. Může být dokonce využita k vytváření falešných, leč extrémně realistických videí, audio nahrávek a obrázků, též známých jako tzv. „deepfakes“<sup>149</sup>. Deepfakes mohou negativně ovlivnit integritu jedinců, případně i ohrozit demokratický proces (falešná politická videa).
- Dopad umělé inteligence na pracovní místa – očekává se, že využívání AI na pracovišti povede k rušení velkého počtu pracovních míst. 14 % pracovních míst v zemích OECD (Organizace pro hospodářskou spolupráci a rozvoj) je do vysoké míry automatizovatelných a dalších 32 % by mohlo čelit podstatným změnám (odhad Think

---

<sup>143</sup> PRUCKOVÁ, M. [online]. Umělá inteligence: jaké právní a etické otázky s sebou nese její vojenské využití? 3. 4. 2023. Cit. 2024-03-15. Dostupné z: <https://pravo21.cz/pravo/umela-inteligence-jake-pravni-a-eticke-otazky-s-sebou-nese-jeji-vojenske-vyuziti>

<sup>144</sup> Především tzv. zlaté pravidlo etiky. Charakterizuje ho jednoduchá zásada vzájemnosti v jednání mezi lidmi: „Co nechceš, aby ti jiní činili, nečiň ty jim.“ (Bible, Tobiáš 4:15.) resp. „Jak chcete, aby lidé jednali s vámi, tak jedněte vy s nimi.“ (Bible, Matouš 7:12).

<sup>145</sup> CALISKAN, A., BRYSON, J.J. a NARAYANAN, A. Semantics derived automatically from language corpora contain human-like biases. *Science*. 2017, č. 356 (6334), s. 183-186. ISSN 0036-8075.

<sup>146</sup> Např. Buolamwini, J., Gebu, T.: Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency, pp. 77-91 (Jan. 2018). Accessed: May 04, 2021 (Online). Available: <http://proceedings.mlr.press/v81/buolamwini18a.html>

<sup>147</sup> PARK, J. a HAHN, Y. Tag: a participatory news platform for critical reading of unverified news. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, pp. 1-3. Shaping Society, New York, NY, USA (Oct. 2020). <https://doi.org/10.1145/3419249.3421249>

<sup>148</sup> Evropský parlament. [online]. Umělá inteligence: rizika i příležitosti. 26. 3. 2021. Cit. 2024-03-16. Dostupné z: <https://www.europarl.europa.eu/topics/cs/article/20200918STO87404/umela-inteligence-jake-jsou-vyhody-a-neyvhody>

<sup>149</sup> Deepfakes jsou falešná videa vytvořená pomocí digitálního softwaru, strojového učení a výměny obličejů. Jsou to počítačem vytvořená umělá videa, ve kterých jsou obrázky kombinovány a vytvářejí nové záběry, které zobrazují události nebo akce, které se ve skutečnosti nikdy nestaly. Výsledky mohou být docela přesvědčivé. Tyto padělky se liší od jiných forem nepravdivých informací tím, že je velmi obtížné identifikovat jako nepravdivé. Více např: <https://www.webwise.ie/news/explained-what-are-deepfakes/>

Tanku Evropského parlamentu 2020). Na druhou stranu je třeba podotknout, že se zároveň předpokládá, že umělá inteligence vytvoří nová a lepší pracovní místa.

- Nebezpečí pro hospodářskou soutěž a konkurenceschopnost – shromažďování informací může mimo jiné vést k narušení hospodářské soutěže – aktéři s větším množstvím informací, získanými pomocí AI, mohou nad svými oponenty získat výhodu a konkurenci účinně eliminovat.
- Bezpečnostní rizika – veškeré způsoby využití umělé inteligence, které jsou v přímém fyzickém kontaktu s lidmi nebo přímo integrované do lidského těla, mohou představovat bezpečnostní rizika. Existuje možnost chybného návrhu, zneužití či hacknutí těchto zařízení. Nedostatečná regulace při používání AI v moderních zbraňových systémech může vést dokonce i ke ztrátě kontroly nad nebezpečnými zbraněmi.
- Otázka transparentnosti – riziko spočívá také v nerovnováze v přístupu k informacím. Na základě chování lidí v online prostředí a díky dalším datům může bez vědomí dotčených osob dojít například k cílenému přizpůsobování politických kampaní člověku na míru či předvídaní toho, kolik je daná osoba u online prodejce ochotna utratit<sup>150</sup>.
- Zvýšené používání algoritmického rozhodování založeného na umělé inteligenci v oblastech finančních služeb, bankovníctví a trestního soudnictví bez zapojení lidského úsudku nebo řádného procesu může posílit škodlivé sociální stereotypy vůči konkrétním menšinovým skupinám a zesílit rasové a genderové předsudky. Tuto praxi kritizovalo několik institucionálních aktérů z celé EU, jako je Rada Evropy, Evropský inspektor ochrany údajů a Agentura Evropské unie pro základní práva<sup>151</sup>.

Pro snížení rizik Evropská komise vydala Bílou knihu o umělé inteligenci - evropský přístup k dokonalosti a důvěře (*White Paper on Artificial Intelligence A European approach to excellence and trust*)<sup>152</sup>, kde se snaží popsat odpovědnosti jednotlivých subjektů. **Mezi základní povinnosti přístupu k AI patří před spuštěním aplikací testovat AI na tréninkových datech, vedení záznamů a dat, transparentní poskytování informací, zajištění technické robustnosti a přesnosti, nutnost lidského dohledu a stanovení specifických požadavků na biometrickou identifikaci.**

Že AI systémy nemohou zůstat zcela bez lidské kontroly, říká i americká Rada pro obranné informace (*Defence Innovation Board*) pod Ministerstvem obrany USA. Ta v roce 2020 identifikovala pět principů použití AI: **princip odpovědnosti, princip spravedlnosti, princip sledovatelnosti, princip spolehlivosti a princip správy**<sup>153</sup>.

Z výše uvedeného je zřejmé, že pro etickou budoucnost rozvoje a aplikace AI je stěžejní role stakeholderů. Vlády a regulační orgány hrají klíčovou úlohu při vývoji a prosazování předpisů, které upravují etické používání AI. Regulační rámce by měly najít rovnováhu mezi podporou

<sup>150</sup> Evropský parlament. [online]. Umělá inteligence: rizika i příležitosti. 26. 3. 2021. Cit. 2024-03-16. Dostupné z: <https://www.europarl.europa.eu/topics/cs/article/20200918STO87404/umela-inteligence-jake-jsou-vyhody-a-nevyhody>

<sup>151</sup> europarl.europa.eu. [online]. EU regulace: Artificial intelligence: legal and ethical reflections. Cit. 2024-03-17. Dostupné z: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/634427/EPRS\\_BRI\(2019\)634427\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/634427/EPRS_BRI(2019)634427_EN.pdf)

<sup>152</sup> European commission. [online]. White Paper on Artificial Intelligence: a European approach to excellence and trust. 2020. Cit. 2024-03-17. Dostupné z: [https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en?prefLang=cs&etrans=cs](https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en?prefLang=cs&etrans=cs)

<sup>153</sup> Mediadefense. [online]. Defence Innovation Board. AI principles: Recommendations on the ethical use of artificial intelligence by the Department of Defense. Cit. 2024-03-17. Dostupné z: [https://media.defense.gov/2019/Oct/31/2002204458/-1/-/1/0/DIB\\_AI\\_PRINCIPLES\\_PRIMARY\\_DOCUMENT.PDF](https://media.defense.gov/2019/Oct/31/2002204458/-1/-/1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF)



inovací a ochranou veřejných zájmů, řešením otázek, jako je ochrana osobních údajů, algoritmická odpovědnost a etický dohled. Korporace a vývojáři vyvíjející technologie umělé inteligence mají povinnost upřednostňovat etická a morální hlediska v průběhu vývoje AI produktů a aplikací. To zahrnuje provádění důkladného hodnocení rizik, zavádění poctivých postupů a podpory kultury etické odpovědnosti mezi zaměstnanci. I výzkumní pracovníci a akademická sféra hrají zásadní roli při prohlubování chápání etických důsledků umělé inteligence a při vývoji etických rámců a norem. Mezioborová spolupráce mezi počítačovou vědou, etikou, právem a společenskými vědami je tak nezbytná pro řešení složitých etických výzev v AI.

Přes tato zmíněná rizika a problémy jako nejzásadnější problém a riziko vidím nedostatečné použití AI, nevyužití jejího potenciálu a „zaspání“ tohoto moderního trendu naší společností, našimi firmami.

## **Závěr**

Je zřejmé, že aplikace technologie AI v reálném světě mají často neočekávané dopady a důsledky, což vede k hluboké debatě o tom, jak by takové pokročilé technologie měly být navrženy a řízeny.

Kromě právních výzev představuje AI mnoho etických otázek. Používání umělé inteligence ke sledování nebo dokonce předpovídání lidského chování riskuje stigmatizaci, posilování stávajících stereotypů, sociální a kulturní segregaci a vyloučení, podvracení individuální volby a rovných příležitostí. Rozvoj tohoto odvětví má tedy odraz i v tvorbě etiky AI. Etika umělé inteligence a robotiky je aplikovanou částí etiky: normativní etika.

S ohledem na rozsáhlost výzkumu se dá očekávat, že podíl robotických a inteligentních systémů ve společnosti bude nadále narůstat. Otázky, jak dosáhnout souladu s etickými a právními požadavky, se v debatách objevují poměrně konstantně. Ovšem kvantita ještě neznamená kvalitu diskuzí. A tak, jako už mnohokrát v minulosti, je technologický pokrok i v případě umělé inteligence rychlejší než vznik pravidel, která jej mají doprovázet.

Řešení těchto výzev vyžaduje společné úsilí vlád, korporací, výzkumných pracovníků, občanské společnosti a široké veřejnosti. Vyžaduje to rozvoj pevných regulačních rámců, etických pokynů/norem a odpovědných inovačních postupů, které upřednostňují lidské hodnoty a práva. Pokračující mezioborový dialog a zapojení je navíc zásadní pro orientaci v etických složitostech umělé inteligence v neustále se vyvíjícím technologickém prostředí.

## Reference

- BARTÁK, O. [online]. Co je to umělá inteligence aneb jak technologie předběhne člověka. Máme se bát? Je to budoucnost lidstva, ale...? 2024. Cit. 2024-03-15. Dostupné z: <https://deeply.cz/blog/umela-inteligence>
- CALISKAN, A., BRYSON, J.J. a NARAYANAN, A. Semantics derived automatically from language corpora contain human-like biases. *Science*. 2017, č. 356 (6334), s. 183-186. ISSN 0036-8075.
- europarl.europa.eu. [online]. EU regulace: Artificial intelligence: legal and ethical reflections. Cit. 2024-03-17. Dostupné z: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/634427/EPRS\\_BRI\(2019\)634427\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/634427/EPRS_BRI(2019)634427_EN.pdf)
- European commission. [online]. White Paper on Artificial Intelligence: a European approach to excellence and trust. 2020. Cit. 2024-03-17. Dostupné z: [https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en?prefLang=cs&etrans=cs](https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en?prefLang=cs&etrans=cs)
- Evropský parlament. [online]. Umělá inteligence: rizika i příležitosti. 26. 3. 2021. Cit. 2024-03-16. Dostupné z: <https://www.europarl.europa.eu/topics/cs/article/20200918STO87404/umela-inteligence-jake-jsou-vyhody-a-nevyhody>  
<https://www.webwise.ie/news/explained-what-are-deepfakes/>
- LAURER, D. You cannot have AI ethics without ethics, *AI and Eticcs*, Springer. 2021, č. 1, s. 21-25. ISSN 2730-5961.
- Logically.ai. [online]. 5 Examples of Biased Artificial Intelligence. 30. 7. 2019. Cit. 2024-03-15. Dostupné z: <https://www.logically.ai/articles/5-examples-of-biased-ai>
- Mediadefense. [online]. Defence Innovation Board. AI principles: Recommendations on the ethical use of artificial intelligence by the Department of Defense. Cit. 2024-03-17. Dostupné z: [https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB\\_AI\\_PRINCIPLES\\_PRIMARY\\_DOCUMENT.PDF](https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF)
- Microsoft Azure. [online]. Umělá inteligence (AI) vs. Strojové učení (ML). 2023. Cit. 2024-03-15. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/artificial-intelligence-vs-machine-learning>
- PARK, J. a HAHN, Y. Tag: a participatory news platform for critical reading of unverified news. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, pp. 1-3. Shaping Society, New York, NY, USA (Oct. 2020). <https://doi.org/10.1145/3419249.3421249>
- PRUCKOVÁ, M. [online]. Umělá inteligence: jaké právní a etické otázky s sebou nese její vojenské využití? 3. 4. 2023. Cit. 2024-03-15. Dostupné z: <https://pravo21.cz/pravo/umela-inteligence-jake-pravni-a-eticke-otazky-s-sebou-nese-jeji-vojenske-vyuziti>
- Sap. [online]. Co je to umělá inteligence? 2023. Cit. 2024-03-16. Dostupné z: <https://www.sap.com/cz/products/artificial-intelligence/what-is-artificial-intelligence.html>
- Shaip. [online]. The importance of ethical AI. 23. 1. 2024. Cit. 2024-03-15. Dostupné z: <https://www.shaip.com/blog/the-importance-of-ethical-ai-fair-ai/>

## **Kontakt**

**Ing. Tomáš Kremr, Ph.D., LL.M.**

Vysoká škola finanční a správní, a.s.

Fakulta ekonomických studií

Katedra ekonomiky a managementu

Estonská 500, Praha, Česká republika

35947@mail.vsfs.cz

Vystudoval VŠE obor zahraniční obchod a doktorské studium na stejné škole na národohospodářské fakultě. Dvacet let působil v manažerských, především obchodních, pozicích velkých mezinárodních firem (Mars, Pepsi, Danone, Colas Rail, Rebo). V současnosti se věnuje akademické dráze na půdě VŠFS – zajímá ho především ekonomická teorie a etika podnikání. Připravuje se na habilitační řízení z oboru mikroekonomie. Má české a německé občanství a žije střídavě v Praze a Bad Hofgasteinu v Rakousku.