ANTHROPOLOGIA INTEGRA
15/2024/2
ČASOPIS PRO OBECNOU ANTROPOLOGII A PŘÍBUZNÉ OBORY
JOURNAL FOR GENERAL ANTHROPOLOGY AND RELATED DISCIPLINES

Regulace rizik a etické výzvy umělé inteligence v kontextu nařízení AI Act

Anetta Jedličková

Univerzita Karlova, Fakulta humanitních studií, Katedra filosofie, doktorský obor Aplikovaná etika, Pátkova 2137/5, 182 00 Praha 8, email: Anetta.Jedlickova@fhs.cuni.cz

Do redakce doručeno 20. října 2024; k publikaci přijato 28. listopadu 2024

REGULATING RISKS AND ETHICAL CHALLENGES OF ARTIFICIAL INTELLIGENCE IN THE CONTEXT OF THE AI ACT

ABSTRACT This article addresses the risks and ethical challenges associated with using artificial intelligence (AI) systems within the context of the recently enacted AI Act, which establishes a binding legal framework of the European Union for their development and deployment. AI systems are becoming essential tools across various sectors of society, including healthcare, public administration, and security. However, this rapid expansion raises significant ethical concerns, particularly regarding fairness, transparency, and the protection of fundamental rights. The AI Act addresses these risks by introducing regulations to ensure accountability and prevent unethical practices by AI systems. The article focuses on identifying and categorizing key risks, the responsibilities of AI providers, and the major ethical challenges, which may pose significant ethical implications for individuals and society, including threats to privacy, unpredictable impacts on the labor market, and exacerbating social inequalities. In the concluding section, recommendations are provided for further development and implementation of AI to maximize its societal benefits while minimizing potential negative consequences.

KEY WORDS AI Act; AI regulation; ethical standards; risks of AI; risk management; trustworthy AI

ABSTRAKT Tento článek se zabývá riziky a etickými výzvami spojenými s používáním systémů umělé inteligence (AI) v rámci nově přijatého nařízení o umělé inteligenci (AI Act), které představuje závazný právní rámec Evropské unie pro jejich vývoj a implementaci. Systémy AI se stávají klíčovými nástroji v mnoha oblastech společnosti, včetně zdravotnictví, veřejné správy a bezpečnosti, avšak jejich rychlá expanze vyvolává závažné etické otázky, například v oblasti spravedlnosti, transparentnosti a ochrany základních práv. AI Act reaguje na tato rizika zavedením regulací zaměřených na zajištění odpovědnosti a předcházení neetickým praktikám ze strany systémů AI. Článek se zaměřuje na identifikaci a rozdělení klíčových rizik, povinnosti poskytovatelů a etické aspekty, které mohou při využívaní technologií AI přinášet významné etické implikace pro jednotlivce i společnost, včetně ohrožení soukromí, nepředvídatelných dopadů na pracovní trh či prohlubování sociálních nerovností. V závěru článku jsou navržena doporučení pro další rozvoj a implementaci AI s cílem maximalizovat její přínos pro společnost a minimalizovat její negativní důsledky.

KLÍČOVÁ SLOVA AI Act; regulace AI; etické standardy; rizika AI; řízení rizik; důvěryhodná AI

ÚVOD

Umělá inteligence (AI) se stala jednou z nejvýznamnějších technologií současnosti, která zásadně transformuje průmyslová odvětví a společenské interakce. Přináší nové příležitosti pro inovace a zvyšování efektivity. AI systémy jsou stále častěji využívány v klíčových oblastech, jako jsou zdravotní péče, veřejné služby nebo bezpečnost, kde významně přispívají k automatizaci procesů a rozhodování. Rychlá integrace AI do různých sektorů však vyvolává závažné etické otázky týkající se transparentnosti, spravedlnosti a odpovědnosti těchto systémů. Často docházelo k jejich implementaci v prostředí s nedostatečně definovanými legislativními a etickými požadavky, což vedlo k obavám o ochranu soukromí, algoritmickou předpojatost a potenciální zneužití citlivých dat. Incidenty, jako byl například skandál Cambridge Analytica-

-Facebook (HU 2020), ukazují, jak může nesprávné využití AI technologií vést k ohrožení demokratických procesů a důvěry veřejnosti.

Evropská unie (EU) reagovala na rostoucí obavy přijetím nařízení o umělé inteligenci, známého také jako AI Act, které vstoupilo v platnost v srpnu 2024, dvacátý den po jeho vyhlášení v Úředním věstníku EU. Tímto krokem začalo dvouleté přechodné období, během něhož se subjekty mohou připravit na plnou právní závaznost nařízení. Je však důležité poznamenat, že některé jeho části budou vymahatelné již dříve; například zakázané postupy v oblasti AI budou plně podléhat požadavkům nařízení již od února 2025. AI Act představuje komplexní právní rámec pro regulaci vývoje a používání systémů umělé inteligence. Nařízení se zaměřuje na řízení rizik spojených s technologiemi AI prostřednictvím zavedení přísných pravidel a kategorizace rizik, přičemž zásadní důraz je kladen na aplikace s vysokým rizikem (Nařízení 2024). Cílem AI Act je zajistit, aby systémy AI byly vyvíjeny a implementovány způsobem, který respektuje základní práva, dodržuje etické standardy a minimalizuje potenciální rizika a škody.

Tento článek se zaměřuje na hlavní etická a společenská rizika spojená s AI systémy v návaznosti na AI Act. Dále poukazuje na oblasti, kde je zapotřebí dalších etických úvah. Cílem článku je přispět k současné diskuzi o nejvhodnějších přístupech k řízení a regulaci systémů umělé inteligence v dynamicky se rozvíjejícím technologickém prostředí.

REGULAČNÍ RÁMEC AI ACT

AI Act představuje první komplexní právní rámec pro regulaci umělé inteligence na světové úrovni. Jeho hlavním cílem je podpora inovací v oblasti AI, při současném zajištění, že její využívání bude bezpečné, transparentní a v souladu se základními právy jednotlivců. Nařízení reaguje na potřebu jasně definovaných pravidel v rychle se rozvíjejícím sektoru, kde dosavadní regulace byla zcela nedostatečná a nejednotná. Zavádí rizikově orientovaný přístup k regulaci, který rozlišuje různé kategorie systémů umělé inteligence na základě míry závažnosti a povahy potenciálních rizik spojených s jejich konkrétními aplikacemi. Kategorizace rizik zahrnuje následující úrovně:

1. *Nepřijatelné riziko*: Tato kategorie zahrnuje systémy AI, které jsou považovány za nepřijatelné a zakázané. Patří sem například systémy, které využívají podprahovou manipulaci nebo zneužívají zranitelnosti uživatelů (například dětí, zdravotně postižených osob nebo jedinců ve zvláštní sociální či ekonomické situaci) s cílem ovlivnit jejich rozhodování a svobodnou volbu. Dalšími příklady jsou systémy pro sociální skórování a některé formy biometrické identifikace v reálném čase na veřejných místech.

2. *Vysoké riziko*: Systémy spadající do této kategorie mají významný potenciál ovlivnit zdraví, bezpečnost a základní práva osob. Příklady zahrnují systémy pro rozhodování v oblasti zaměstnávání, vzdělávání, úvěrování, správy soudnictví, nebo systémy pro kritickou infrastrukturu, jako je zdravotnictví nebo doprava. AI Act stanovuje pro tyto systémy přísné požadavky na transparentnost, dokumentaci, monitorování a zajištění lidského dohledu.

3. *Omezené riziko*: Do této skupiny patří systémy AI, které mají omezený potenciál negativních dopadů, ale vyžadují určitou míru transparentnosti. Jedná se o systémy, které nemají významný vliv na výsledek rozhodování a nepředstavují riziko pro zdraví, bezpečnost nebo základní práva. Příklady zahrnují:

- Systémy, které plní úzce vymezený procedurální úkol, jako je přeměna nestrukturovaných dat na strukturovaná, třídění dokumentů nebo detekce duplicit.
- Systémy, které zlepšují výsledek již dokončené lidské činnosti, například zdokonalením textu, aniž by nahrazovaly rozhodnutí člověka.
- Systémy odhalující vzorce nebo odchylky v rozhodování, přičemž lidské posouzení není nahrazeno ani přímo ovlivněno.
- Systémy sloužící k přípravě podkladů pro další posouzení, například pro indexování, vyhledávání nebo překlady.

4. *Minimální riziko*: Tato kategorie zahrnuje systémy AI, které představují minimální nebo žádné riziko pro uživatele a veřejnost, a nejsou tedy předmětem zvláštní regulace.

AI Act klade důraz na několik klíčových oblastí, které mají zajistit, že systémy AI budou vyvíjeny a implementovány způsobem, který minimalizuje rizika a maximalizuje přínosy. Patří k nim především:

Bezpečnost: Systémy AI musí splňovat přísné požadavky na bezpečnost, spolehlivost a kvalitu. Technická spolehlivost a bezpečnost znamenají, že systémy AI musí být navrhovány a používány tak, aby zajišťovaly spolehlivou funkci i v případě výskytu problémů a odolnost vůči pokusům o neoprávněné změny jejich použití nebo výkonnosti. Tím se minimalizuje riziko protiprávního zneužití ze strany třetích osob a riziko neúmyslné újmy. Vývojáři i provozovatelé systémů AI jsou povinni provádět hodnocení bezpečnosti a pravidelně monitorovat systémy AI v praxi. Aplikace s vysokým rizikem musí být podrobeny pravidelnému vyhodnocování, aby se minimalizovalo riziko technických selhání (Nařízení 2024).

Transparentnost: Klíčovým prvkem AI Act je zajištění, že systémy AI budou transparentní a uživatelé si budou vědomi, že komunikují nebo interagují se systémem AI. Transparentností se rozumí, že systémy AI jsou vyvíjeny a používány způsobem, který umožňuje patřičnou sledovatelnost a vysvětlitelnost. Transparentnost zahrnuje rovněž povinnost vývojářů poskytovat informace o použitých algoritmech, datových sadách a rozhodovacích procesech, což je zvláště důležité v případě systémů s vysokým rizikem, kde je třeba zamezit netransparentnímu nebo diskriminačnímu rozhodování. Je

také nezbytné, aby zavádějící subjekty byly řádně informovány o schopnostech a omezeních daného systému AI a dotčené osoby o svých právech. Transparentnost rovněž zahrnuje poskytnutí srozumitelného návodu k použití, který uživatelům umožňuje správně využívat systémy AI a činit informovaná rozhodnutí (Nařízení 2024).

Odpovědnost: AI Act ukládá vývojářům a provozovatelům AI systémů odpovědnost za jejich rozhodování a výstupy. Odpovědnost vyžaduje zajištění řízení rizik, kontinuální kontroly a pravidelného auditu systémů, včetně umožnění dohledu ze strany regulátorů. Cílem je zajistit, aby v případě selhání či negativního dopadu na uživatele existovala jednoznačná odpovědnost a možnost nápravy. Vzhledem k povaze systémů AI a rizikům, která mohou ovlivnit bezpečnost a základní práva osob při jejich používání, je klíčové provádět adekvátní monitorování výkonnosti systémů v reálných podmínkách. Proto je nezbytné stanovit jasnou odpovědnost zavádějících subjektů, čímž je zajištěna větší ochrana uživatelů. Je zásadní, aby za uvedení vysoce rizikového systému AI na trh nebo do provozu nesla odpovědnost konkrétní fyzická nebo právnická osoba, definovaná jako poskytovatel. Tato odpovědnost by měla zůstat na poskytovateli bez ohledu na to, zda systém navrhl nebo vyvinul. Poskytovatelé tak nesou odpovědnost za zajištění souladu se všemi příslušnými požadavky (Nařízení 2024, 23–24, 56).

ZAKÁZANÉ POSTUPY V OBLASTI UMĚLÉ INTELIGENCE

V rámci AI Act jsou v EU zakázány určité praktiky v oblasti umělé inteligence, které jsou považovány za nepřípustné vzhledem k jejich závažnému dopadu na základní práva a svobody jednotlivců. Jsou klasifikovány jako systémy s nepřijatelným rizikem a jejich uvádění na trh, provozování a používání jsou striktně zakázány. Zákaz má za cíl chránit jednotlivce před škodlivými, manipulativními a neetickými postupy v oblasti AI, a zabránit tak využívání systémů umělé inteligence, které narušují soukromí, zneužívají zranitelnosti nebo jinak ohrožují základní lidská práva a důvěru společnosti v technologie.
Klíčové zakázané postupy zahrnují uvádění na trh, uvádění do provozu nebo používání následujících systémů AI:

Podprahové a manipulativní techniky: Patří k nim AI systémy, které využívají podprahových technik mimo vědomí osob nebo manipulativních a klamavých technik za účelem ovlivnění osob, jejichž důsledkem je významné narušení schopnosti činit informovaná rozhodnutí. Ovlivnění může způsobit, že osoba učiní rozhodnutí, která by jinak neučinila, což může vést k významné újmě pro ni nebo pro jiné osoby. Tyto techniky mohou zahrnovat například manipulaci pomocí cílených reklam, které podvědomě ovlivňují volbu spotřebitele, aniž by si to spotřebitel uvědomoval.

Zneužívání zranitelnosti: Tato skupina zahrnuje systémy AI, které zneužívají zranitelnosti osob na základě jejich věku, zdravotního postižení nebo socioekonomické situace. Tyto systémy mohou podstatně narušit rozhodování zranitelných osob takovým způsobem, že dotčeným nebo dalším osobám způsobí nebo by mohly způsobit značnou újmu. Příkladem může být využívání AI pro manipulaci dětí nebo osob v tíživé finanční situaci prostřednictvím nevhodných nabídek.

Ovlivnění sociálních kreditních systémů: Zakázány jsou také systémy AI, které využívají hodnocení a klasifikaci jednotlivců nebo skupin na základě jejich sociálního chování, osobnostních vlastností nebo jiných odvozených dat (tzv. sociální skórování). Výsledek hodnocení může vést k tomu, že jedinci jsou diskriminováni v kontextech, které nesouvisí s původními důvody sběru dat, nebo že je s nimi nepřiměřeně zacházeno vzhledem k jejich sociálnímu chování.

Prediktivní hodnocení trestné činnosti: Zakázány jsou systémy AI, které se výhradně zaměřují na předpovídání, zda určitá fyzická osoba spáchá trestný čin na základě profilování osoby nebo analýzy jejích osobnostních vlastností. Tyto systémy by mohly vést k nespravedlivému stíhání nebo trestání osob na základě pravděpodobnosti spáchání trestného činu. Tento zákaz se však nevztahuje na systémy AI, které podporují lidské rozhodování a jsou založeny na ověřitelných faktech souvisejících s trestnou činností.

Rozpoznávání obličeje a biometrické sledování: Zakázány jsou systémy AI určené pro vytváření a používání databází pro rozpoznávání obličeje, které jsou založeny na necíleném získávání zobrazení obličejů z veřejně dostupných zdrojů, jako jsou internet nebo kamerové záznamy. Tento zákaz je zaměřen na ochranu soukromí občanů a zabránění nekontrolovanému sledování osob, které by mohlo vést k závažným porušením práv na soukromí a svobodu pohybu.

Detekce emocí ve specifických prostředích: Systémy AI, které odvozují nebo interpretují emoce fyzických osob na pracovištích nebo ve vzdělávacích institucích, jsou rovněž zakázány. Zákaz má zabránit sledování emocionálních reakcí jednotlivců za účelem jejich hodnocení v pracovním prostředí nebo při výuce, což by mohlo vést k manipulaci, diskriminaci nebo nespravedlivému zacházení. Výjimky se týkají pouze případů, kdy je takové sledování potřebné pro ochranu zdraví nebo bezpečnosti.

Biometrická kategorizace osob podle citlivých atributů: Další zakázanou praxí je biometrická kategorizace, která kategorizuje osoby podle jejich biometrických údajů za účelem zjištění nebo odvození rasy, politických názorů, členství v odborech, náboženského přesvědčení, sexuální orientace nebo jiných citlivých atributů. Tento zákaz reflektuje obavy z potenciální diskriminace a neoprávněného zásahu do soukromí. Použití biometrické kategorizace je povoleno pouze v oblastech nezbytných pro vymáhání práva, nebo v případech, kde jsou biometrické údaje využívány legálně a za jasně vymezených

podmínek, například při označování legálně získaných souborů biometrických údajů.

Biometrická identifikace na dálku v reálném čase ve veřejném prostoru: Jeden z nejdiskutovanějších zákazů se týká biometrické identifikace na dálku v reálném čase ve veřejně přístupných prostorech, zejména pro účely vymáhání práva. Tento zákaz se vztahuje na systémy, které v reálném čase rozpoznávají osoby například podle obličeje, pokud to není nezbytně nutné pro některý z následujících cílů:

- Cílené vyhledávání obětí únosů, obchodování s lidmi nebo osob pohřešovaných.
- Prevence konkrétního, závažného a bezprostředního ohrožení života nebo bezpečnosti, včetně teroristických útoků.
- Lokalizace nebo identifikace osoby podezřelé ze spáchání závažného trestného činu, pokud za něj hrozí trest odnětí svobody nejméně čtyři roky.

Použití systému biometrické identifikace na dálku v reálném čase ve veřejně přístupných prostorech je povoleno pouze při splnění přísných podmínek, včetně nezbytnosti posouzení dopadů na základní práva a získání předchozího povolení od justičního nebo nezávislého správního orgánu. Takové povolení musí být vydáno na základě jasných důkazů a musí respektovat zásady nezbytnosti a přiměřenosti s ohledem na časové, zeměpisné a osobní omezení. V naléhavých případech je povoleno zahájit používání systému bez povolení, avšak žádost o povolení musí být podána neprodleně, nejpozději však do 24 hodin. Musí být dokončeno posouzení dopadů na základní práva donucovacími orgány a systém musí být zaregistrován v databázi EU. V řádně odůvodněných naléhavých případech však může být používání těchto systémů zahájeno bez registrace v databázi EU, pokud je tato registrace provedena následně bez zbytečného odkladu.

K další nutným podmínkám patří dostatečné vyhodnocení, zda je použití systému biometrické identifikace na dálku v reálném čase omezeno pouze na specifika nutná pro dosažení zamýšleného cíle. Rozhodnutí přijatá výlučně na základě výstupu z těchto systémů nesmějí mít nepříznivé právní důsledky pro dotčené osoby. Dále musí být každé takové použití systému nahlášeno příslušným orgánům pro dohled nad trhem a pro ochranu osobních údajů. Komise EU bude každý rok zveřejňovat výroční zprávy o používání těchto systémů v jednotlivých členských státech, které budou založeny na údajích poskytovaných vnitrostátními orgány. Zprávy mají zajistit transparentnost a možnost veřejné kontroly nad využíváním systémů biometrické identifikace (Nařízení 2024).

VYSOCE RIZIKOVÉ SYSTÉMY AI

Vysoce rizikové systémy AI se vyznačují významným potenciálem zasahovat do základních práv a svobod jednotlivců. Z toho důvodu je jejich použití v regulovaných oblastech pod přísnou kontrolou s důrazem na zajištění bezpečnosti, transparentnosti a odpovědnosti.

Systém AI je považován za vysoce rizikový, pokud splňuje následující podmínky:

- Je určen k použití jako bezpečnostní komponenta produktu nebo je samotným produktem, na který se vztahují harmonizační právní předpisy EU uvedené v příloze I nařízení o umělé inteligenci.[1]
- Na produkt nebo systém AI se vztahuje povinnost posouzení shody třetí stranou za účelem jeho uvedení na trh nebo do provozu podle harmonizačních právních předpisů EU uvedených ve stejné příloze nařízení (Nařízení 2024).

Další systémy AI považované za vysoce rizikové

Vysoce rizikové systémy umělé inteligence jsou dále rozděleny do několika oblastí na základě jejich specifického využití:

1. Biometrika

Mezi vysoce rizikové biometrické systémy patří:

Systémy biometrické identifikace na dálku: Používají se k identifikaci jednotlivců na veřejnosti, nezahrnuje však systémy určené pouze k potvrzení identity konkrétní osoby.

Systémy biometrické kategorizace: Tyto systémy kategorizují osoby na základě citlivých atributů, jako jsou rasa, pohlaví nebo jiné chráněné charakteristiky.

Systémy pro rozpoznávání emocí: Používají se k analýze emocí jednotlivců na základě biometrických dat.

2. Kritická infrastruktura

Systémy AI, které jsou využívány jako bezpečnostní komponenty pro správu a provoz kritické infrastruktury, zahrnují digitální infrastrukturu, dopravní sítě a systémy distribuce základních služeb, jako jsou voda, plyn, topení nebo elektřina. Tyto systémy AI mohou ovlivnit bezpečnost a stabilitu celého státu.

3. Vzdělávání a odborná příprava

Ve vzdělávacím sektoru jsou za vysoce rizikové považovány systémy AI, které:

- určují přístup fyzických osob ke vzdělávacím institucím nebo jejich zařazení;
- hodnotí výsledky učení jednotlivců, což může ovlivnit jejich studijní nebo kariérní dráhu;
- posuzují úroveň vzdělání, kterou jednotlivec získá;
- monitorují a odhalují zakázané chování během testů.

4. Zaměstnávání, řízení pracovníků a přístup k samostatné výdělečné činnosti

Tato oblast zahrnuje systémy, které mají vliv na:

- nábor a výběr zaměstnanců: Do této skupiny patří systémy AI, které analyzují a třídí žádosti o zaměstnání, vyhodnocují uchazeče nebo zveřejňují cílené pracovní nabídky;

1 V příloze I je uvedeno celkem 20 právních předpisů (Nařízení 2024, 124–125).

- rozhodování o pracovních podmínkách: Tato oblast zahrnuje systémy AI, které mohou ovlivnit povýšení pracovníků, přidělování úkolů, sledování výkonnosti a rozhodování o ukončení pracovních smluv.

5. Přístup k základním soukromým a veřejným službám
V této oblasti jsou za vysoce rizikové považovány systémy AI, které:
- hodnotí nárok na základní dávky a veřejné služby, jako je například zdravotní péče;
- posuzují úvěruschopnost nebo stanovují úvěrové skóre fyzických osob;
- posuzují rizika a stanovují ceny v případě životního a zdravotního pojištění;
- třídí tísňová volání, určují prioritu vyslání zásahových složek, jako jsou policie, záchranná služba nebo hasičský sbor, a třídí pacienty v krizových situacích.

6. Vymáhání práva
Systémy AI používané k podpoře donucovacích orgánů a vymáhání práva jsou považovány za vysoce rizikové, pokud jsou využívány k:
- posuzování rizika kriminality: AI systémy, které analyzují riziko, že se fyzická osoba stane obětí trestného činu nebo se dopustí protiprávního jednání;
- hodnocení důkazů během vyšetřování nebo soudního řízení;
- posuzování povahových vlastností, osobnostních rysů nebo dřívější trestné činnosti osob;
- profilování fyzických osob pro účely odhalování, vyšetřování a stíhání trestné činnosti.

7. Migrace, azyl a řízení ochrany hranic
V oblasti migrace a azylu zahrnují vysoce rizikové systémy AI takové systémy, které pomáhají příslušným orgánům při:
- posuzování bezpečnostních rizik spojených s migrací, včetně rizik nelegální migrace nebo zdravotních rizik;
- hodnocení žádostí o azyl, víza nebo povolení k pobytu, včetně posuzování spolehlivosti předložených důkazů a stížností;
- odhalování, rozpoznávání nebo identifikaci fyzických osob, s výjimkou ověřování cestovních dokladů.

8. Správa soudnictví a demokratické procesy
Systémy AI v této skupině jsou považovány za vysoce rizikové, pokud:
- podporují soudní orgány při rozhodování o právních případech a při výkladu práva a faktů;
- jsou určeny k ovlivnění výsledků voleb nebo referenda, ovlivňují volební procesy a chování voličů při volbách nebo referendech (Nařízení 2024).

Výjimky z klasifikace vysoce rizikových systémů

Vyjmenované systémy AI nejsou považovány za vysoce rizikové, pokud nepředstavují významné riziko újmy na zdraví, bezpečnosti nebo základních právech osob a zásadně neovlivňují jejich rozhodování. Zároveň musí splňovat jednu z následujících podmínek:
- jsou navrženy k plnění úzce zaměřeného procesního úkolu;
- mají za cíl zlepšit výsledek lidské činnosti již dokončené;
- jsou zamýšleny k odhalování vzorců rozhodování nebo odchylek od nich, aniž by však ovlivnily předchozí lidské posouzení nebo jej nahradily bez dalšího prozkoumání;
- jsou určeny k provedení přípravného úkolu v rámci posouzení použití uvedených v předchozí kapitole.

Je však nutné připomenout, že provádí-li systém AI profilování fyzických osob, je považován za vysoce rizikový bez ohledu na uvedené výjimky.
Poskytovatel, který vyhodnotí, že jeho systém AI není vysoce rizikový, je povinen toto hodnocení řádně zdokumentovat a v případě žádosti příslušných orgánů dokumentaci předložit. Do února 2026 vydá Evropská komise pokyny pro praktickou implementaci klasifikačních pravidel, jejichž součástí bude také seznam systémů AI, které buď budou klasifikovány jako vysoce rizikové, nebo naopak do této kategorie nebudou spadat (Nařízení 2024).

Požadavky na vysoce rizikové systémy AI

Poskytovatelé vysoce rizikových systémů AI jsou povinni zavést, aplikovat, dokumentovat a udržovat systém řízení rizik, který musí být dostatečně robustní, aby zajistil, že všechny fáze vývoje, výroby a provozu vysoce rizikových systémů AI jsou pod přísným dohledem a splňují požadované standardy.
Systém řízení rizik je považován za nepřetržitý proces, který je průběžně aktualizován a zahrnuje následující kroky:
Identifikace a analýza rizik: Posouzení známých a předvídatelných rizik, která mohou ohrozit zdraví, bezpečnost nebo základní práva osob.
Stanovení a vyhodnocení rizik: Odhad rizik při použití systému pro zamýšlený účel a za rozumně předvídatelných nesprávných podmínek.
Hodnocení dalších rizik: Zohlednění rizik na základě dat získaných z monitorování po uvedení na trh.
Přijetí opatření k řízení rizik: Implementace cílených opatření, která reagují na identifikovaná rizika. Opatření k řízení rizik musí zohledňovat součinnost mezi jednotlivými požadavky, aby byla dosažena rovnováha mezi minimalizací rizik a splněním požadavků. Vysoce rizikové systémy AI musí být testovány za účelem identifikace vhodných opatření k řízení rizik a zajištění, že plní svůj zamýšlený účel. Testování musí probíhat během celého vývoje před uvedením na trh nebo do provozu.
Při zavádění systému řízení rizik musí poskytovatelé také posoudit, zda systémy AI nebudou mít negativní dopady na osoby mladší 18 let nebo jiné zranitelné skupiny (Nařízení 2024).

Povinnosti poskytovatelů vysoce rizikových systémů AI

Poskytovatelé vysoce rizikových systémů AI nesou zásadní odpovědnost za dodržování legislativních požadavků, které zajišťují bezpečnost a spolehlivost těchto technologií. Jejich

povinnosti jsou rozděleny do několika klíčových oblastí, které obsahují technická, administrativní a právní opatření. Základní povinnosti poskytovatelů zahrnují:

Dokumentace a prokazování souladu: Poskytovatelé jsou povinni vést podrobnou dokumentaci týkající se vysoce rizikových systémů AI, která musí obsahovat technické specifikace, zprávy o testování a další klíčové údaje nezbytné pro prokázání souladu s legislativními požadavky. Na žádost příslušných orgánů musí poskytovatelé doložit, že jejich vysoce rizikové systémy AI splňují všechny stanovené požadavky podle nařízení AI Act, což zahrnuje předložení technické dokumentace a dalších relevantních podkladů.

Generování provozních protokolů: Poskytovatelé musí zajistit, že jejich vysoce rizikové systémy AI automaticky generují záznamy během svého provozu. Tyto protokoly jsou nezbytné pro monitorování výkonu systému a včasné odhalení jakéhokoli nesprávného fungování nebo bezpečnostních rizik.

Identifikační údaje poskytovatele: Na vysoce rizikovém systému AI, jeho obalu nebo v přiložené dokumentaci musí být zřetelně uvedeny identifikační údaje poskytovatele, které zahrnují název společnosti, ochrannou známku, je-li dostupná, a kontaktní adresu, jejíž prostřednictvím lze poskytovatele kontaktovat.

Označení evropské shody – CE: Poskytovatelé jsou povinni na vysoce rizikový systém AI, jeho obal nebo přiloženou dokumentaci umístit označení CE,[2] které potvrzuje, že systém odpovídá standardům EU a splňuje požadavky stanovené nařízením AI Act.

Posuzování shody a prohlášení o shodě: Před uvedením vysoce rizikového systému AI na trh nebo do provozu musí poskytovatelé provést postup posuzování shody, který ověřuje, zda systém splňuje všechny technické a bezpečnostní normy. Poskytovatelé musí následně vypracovat a podepsat EU prohlášení o shodě, jež stvrzuje, že systém AI vyhovuje všem platným požadavkům.

Registrace v příslušných rejstřících: Poskytovatelé jsou povinni zaregistrovat vysoce rizikové systémy AI v příslušných databázích, které slouží k jejich monitorování a regulaci.

Nápravná opatření: V případě bezpečnostních problémů nebo zjištěného nesouladu jsou poskytovatelé povinni přijmout nezbytná nápravná opatření, aby minimalizovali rizika a negativní dopady selhání systému AI. Jsou rovněž povinni poskytovat relevantní informace.

2 CE je zkratka z francouzského Conformité Européenne – evropská shoda, která představuje jediné označení osvědčující shodu výrobku s příslušnými požadavky harmonizačních právních předpisů EU. Více k povinnostem viz článek 30 nařízení (ES) č. 765/2008 (Nařízení 2008, 43).

Přístupnost pro všechny uživatele: Poskytovatelé musí zajistit splnění požadavků na přístupnost v souladu s příslušnými evropskými směrnicemi,[3] což zahrnuje zajištění, že systémy AI jsou přístupné a použitelné pro všechny skupiny uživatelů, včetně osob se zdravotním postižením.

Poskytnutí srozumitelné dokumentace: Vysoce rizikové systémy AI by měly být navrženy tak, aby umožnily zavádějícím subjektům plně porozumět jejich fungování, vyhodnotit jejich výkonnost a pochopit jejich silné stránky i omezení. K systémům AI musí být přiloženy podrobné a srozumitelné návody k použití, které obsahují všechny relevantní informace o vlastnostech, schopnostech a výkonnostních omezeních systému. Návod by měl zahrnovat popis známých a předvídatelných okolností, které mohou při používání systému vzniknout. Součástí by měly být také informace o možných rizicích pro zdraví, bezpečnost a základní práva a o přijatých opatřeních v oblasti lidského dohledu, včetně pokynů k interpretaci výstupů AI. Poskytovatelé musí zajistit, že veškerá dokumentace, včetně návodu k použití, obsahuje smysluplné, ucelené, přístupné a snadno pochopitelné informace, přizpůsobené potřebám a předpokládané úrovni znalostí cílových uživatelů. Návod k použití musí být dostupný v jazyce, který je pro cílové subjekty snadno srozumitelný (Nařízení 2024).

ETICKÉ ZÁSADY A REGULAČNÍ RÁMCE PRO UMĚLOU INTELIGENCI: REZOLUCE A POKYNY

V souvislosti s rychlým rozvojem umělé inteligence schválil Evropský parlament rezoluci s názvem Rámec pro etické aspekty umělé inteligence, robotiky a souvisejících technologií *(Framework of ethical aspects of artificial intelligence, robotics and related technologies*). Rezoluce zdůrazňuje klíčové etické zásady pro vývoj, zavádění a používání technologií AI. Důležitost je kladena na prevenci rizik, která by mohla ohrozit bezpečnost jednotlivců i společnosti. Zásadní požadavky tohoto rámce zahrnují plné dodržování Listiny základních práv Evropské unie (Evropský parlament a Rada EU 2012), zejména ochranu lidské důstojnosti, autonomie a sebeurčení osob. Rámec se zaměřuje na předcházení škodám, podporu spravedlnosti, inkluze a transparentnosti a také na eliminaci předsudků a diskriminace, včetně ochrany práv menšinových skupin. Klíčové etické zásady zahrnují rovněž důraz na vysvětlitelnost technologií a naplňování požadavku, aby technologie AI sloužily lidem, nikoli je nahrazovaly či za ně rozhodovaly. Rámec obsahuje také požadavek na zajištění, aby technologie AI přispívaly k blahu a svobodě jednotlivců, zachování míru, mezinárodní bezpečnosti a předcházení

3 Například:
1. Směrnice Evropského parlamentu a Rady (EU) 2016/2102 ze dne 26. října 2016 o přístupnosti internetových stránek a mobilních aplikací subjektů veřejného sektoru (Úř. věst L 327, 2. 12. 2016).
2. Směrnice Evropského parlamentu a Rady (EU) 2019/882 ze dne 17. dubna 2019 o požadavcích na přístupnost u výrobků a služeb (Úř. věst. L 151, 7. 6. 2019).

konfliktům a zároveň minimalizovaly rizika. Součástí rámce je také nezbytnost zapojení lidského faktoru, který prostřednictvím vhodných kontrolních mechanismů zajišťuje nezávislý dohled a přebírá odpovědnost v případě nutnosti (European Parliament 2020).

Etické pokyny pro zajištění důvěryhodnosti AI (*Ethics guidelines for trustworthy AI)*, které vypracovala nezávislá expertní skupina jmenovaná Evropskou komisí, definují následující hlavní principy: lidský faktor a dohled, technická spolehlivost a bezpečnost, ochrana soukromí a správa dat, transparentnost, rozmanitost, nediskriminace a spravedlnost, udržitelné sociální a environmentální podmínky a odpovědnost (High-Level Expert Group 2019). Systémy AI by měly být navrhovány a používány jako nástroje respektující lidskou důstojnost a autonomii. Zároveň musí zůstat plně pod kontrolou člověka, což zajistí, že konečná rozhodnutí budou vždy v lidských rukou a odpovědnost za jejich výstupy bude konkrétně vymezena.

V posledních letech se mnoho výzkumníků zaměřilo na analýzu etických směrnic týkajících se umělé inteligence, což ukazuje na rostoucí důraz na odpovědný vývoj těchto technologií. Jobin et al. (2019) provedli analýzu 84 etických směrnic národních a mezinárodních organizací, která odhalila konsensus v pěti základních principech: transparentnost, spravedlnost, nonmaleficence (nezpůsobovat ani nezhoršovat újmu), odpovědnost a ochrana soukromí. Hagendorff (2020) následně zrealizoval analýzu 22 etických směrnic pro AI, přičemž kromě společných prvků zkoumal také jejich mezery. Ryan a Stahl (2021) poskytli vysvětlení obsahu a důsledků etických směrnic pro AI, které jsou určeny pro vývojáře a uživatele. Huang et al. (2023) představili globální přehled etických směrnic a principů v oblasti umělé inteligence, přičemž vycházeli ze 146 směrnic, které byly vydány různými institucemi po celém světě. Rozmanitost etických směrnic ukazuje, že v oblasti etiky umělé inteligence existuje široké spektrum přístupů a perspektiv, které mohou významně ovlivnit způsob, jakým se technologie AI vyvíjejí a implementují.

ETICKÉ KONOTACE VÝVOJE, ZAVÁDĚNÍ A POUŽÍVÁNÍ SYSTÉMŮ AI

Rozvoj technologií AI, které autonomně rozhodují, přináší řadu etických aspektů, jež je třeba zohlednit ve všech fázích – od vývoje přes zavádění až po samotné používání těchto systémů. Umělá inteligence, robotika a související technologie, včetně softwaru, algoritmů a dat, které tyto technologie používají či produkují, a to bez ohledu na oblast, v níž jsou vyvíjeny, zaváděny nebo používány, by měly být navrhovány s důrazem na bezpečnost, transparentnost, technickou propracovanost, spolehlivost a dodržování právních předpisů. Klíčová je zásada ‚etika již od návrhu' (IEEE 2019). K významným oblastem, které mají na etické důsledky využívaní technologií AI ve společnosti značný vliv, patří:

1. Důvěra uživatelů

Důvěra ve využívání AI závisí na zajištění transparentnosti, vysvětlitelnosti, sledovatelnosti a auditovatelnosti. K posílení důvěry je třeba minimalizovat rizika spojená s kybernetickou bezpečností, ochranou osobních údajů a potenciálním zneužitím dat. Uživatelé by měli mít přístup ke srozumitelným informacím o fungování autonomních systémů, jejich dopadech a právu na nápravu v případě chybného rozhodnutí. V tomto kontextu je nutné připomenout také poskytování náležité pomoci při uplatňování práva spotřebitelů na nápravu v případech porušení jejich práv v souvislosti s rozhodnutím autonomních systémů. Je proto zásadní, aby všichni aktéři v celém vývojovém i dodavatelském řetězci produktů a služeb AI (vývojáři, provozovatelé a uživatelé) nesli jasně vymezenou právní odpovědnost za případné újmy (European Parliament 2020). Na oblast důvěry uživatelů se vztahují všechny etické principy, zejména však etický princip vysvětlitelnosti a etický princip respektu k autonomii člověka.[4]

2. Nezaujatost a nediskriminace

AI má potenciál vytvářet a posilovat předsudky, pokud jsou základem pro její fungování neobjektivní datové soubory, čímž může způsobit různé formy automatizované diskriminace. Pro prevenci diskriminace je zásadní zajistit kvalitu vstupních dat a transparentnost algoritmů. Technologie by měly být navrhovány tak, aby dodržovaly a chránily fyzickou a mentální integritu a lidskou důstojnost, podporovaly kulturní, jazykovou a individuální rozmanitost a zajišťovaly rovnost příležitostí, genderovou rovnost, zohledňování a zastupování zájmů všech osob, včetně marginalizovaných skupin či osob ve zranitelném postavení, a dosažení rovných práv a pozitivních sociální změn (European Parliament 2020). Pro zajištění nezaujatosti a nediskriminace je důležitý především etický princip spravedlnosti a významné postavení má také etický princip nonmaleficence.

3. Sociální odpovědnost

Technologie založené na AI by měly přispívat ke společenskému prospěchu a podporovat rozličné cíle společnosti, jako jsou inkluze, pluralita, spravedlnost, solidarita, rovnost a spolupráce. Jedná se o cíle, které chrání a podporují základní práva a hodnoty společnosti, jako jsou demokracie, právní stát, ochrana dětí a zdraví, hospodářská prosperita, pracovní a sociální práva, kvalitní a dostupné vzdělávání, pluralitní a nezávislé sdělovací prostředky, objektivní a volně dostupné informace či digitální gramotnost. Autonomní systémy nesmí úmyslně způsobovat škody jednotlivcům ani společnosti v žádné z uvedených oblastí. AI má potenciál ovlivnit například zaměstnanost, vzdělávání nebo svobodu

4 Více k etickým principům v souvislosti s AI viz například (High-Level Expert Group 2019; Jedličková 2022; 2024).

projevu, proto je důležité regulovat její používání v rizikových oblastech. Technologie AI, jejichž vývoj, zavádění a používání přinášejí významné riziko způsobení újmy jednotlivcům nebo společnosti, by měly být považovány za vysoce rizikové technologie. Míra závažnosti by měla být stanovena na základě míry potenciální újmy, množství újmou postižených osob a celkové hodnoty případné škody způsobené společnosti jako celku. K závažnému typu újmy patří například nenávistné verbální projevy, násilí či jakákoli porušení práv dětí, spotřebitelů nebo pracovníků, která kvůli svému rozsahu nebo počtu přinášejí riziko negativního dopadu na fyzickou a duševní pohodu (European Parliament 2020). V oblasti sociální odpovědnosti jsou významné etické principy spravedlnosti, beneficence a nonmaleficence.

4. Ochrana soukromí a osobních údajů

AI technologie často zpracovávají velké množství osobních údajů, včetně biometrických dat. Ochrana těchto údajů je pro prevenci zneužití zásadní, zejména u zranitelných skupin, jako jsou děti, senioři, osoby se zdravotním postižením, menšiny nebo další skupiny ohrožené vyloučením. Jakékoli zneužití osobních údajů zranitelných osob je obzvlášť neetické. Technologie AI skýtají potenciál používat údaje ke kategorizaci osob, k odhalení zranitelných míst jednotlivců či k využití údajů na přesné zacílení osob nebo skupiny osob. Účinně prosazované zásady ochrany údajů a soukromí jsou založeny na důležitých omezeních a kontrolních mechanismech, jako jsou minimalizace údajů, právo odmítnout profilování, kontrola použití osobních údajů, právo získat vysvětlení rozhodnutí, které je založené na automatizovaném zpracování, ochrana soukromí již od návrhu systému AI, omezení na základě předem přesně vymezeného účelu, jakož i zásady proporcionality a nezbytnosti (European Parliament 2020). Vždy je zcela nezbytné dodržovat a prosazovat práva občanů na soukromí a ochranu osobních údajů, včetně osobních údajů odvozených od neosobních a biometrických údajů, a to v souladu s relevantními právními předpisy a etickými principy. V oblasti ochrany osob a jejich údajů je významný zejména etický princip respektu k autonomii člověka. Důležitou roli však hraje také etický princip nonmaleficence.

5. Kontrola a dohled

Rozhodnutí produkovaná AI by měla být podrobována lidskému přezkumu a kontrole. Transparentnost a odpovědnost musí být zajištěny ve všech fázích životního cyklu technologie AI. Důvěryhodné autonomní systémy musí být vyvíjeny, zaváděny a používány bezpečně, transparentně a zodpovědně v souladu s bezpečnostními prvky týkajícími se robustnosti, odolnosti, bezpečnosti, přesnosti a identifikace chyb, vysvětlitelnosti, transparentnosti a identifikovatelnosti. Technická a provozní složitost autonomních technologií by neměla bránit jejich provozovateli nebo uživateli, aby bylo v každém okamžiku umožněno nouzové odstavení, změna či zastavení jejich provozu nebo navrácení k předchozímu stavu obnovou bezpečnostních funkcí. Kontrola a dohled by měly být zaměřeny na dodržování požadavků na kvalitu, integritu, transparentnost, důvěryhodnost, bezpečnost a ochranu údajů a soukromí v návaznosti na dodržování základních etických principů. V souladu s regulačním rámcem pro etické aspekty by měla být transparentnost zajištěna také umožněním přístupu veřejných orgánů k technologiím, datům a relevantním počítačovým systémům v nezbytně nutných případech (European Parliament 2020). V oblasti kontroly a dohledu mají významné postavení etický princip nonmaleficence a princip vysvětlitelnosti.

6. Další oblasti

Výše uvedené oblasti, které mohou při využívaní technologií umělé inteligence přinášet významné etické implikace pro jednotlivce i společnost, jistě nepředstavují vyčerpávající výčet všech etických výzev. Další významná témata zahrnují například nezanedbatelný dopad na míru nezaměstnanosti a pracovní podmínky zaměstnanců, zejména při monitorování, hodnocení, předvídání nebo řízení jejich výkonnosti, což může mít přímé i nepřímé důsledky pro jejich kariérní rozvoj. Technologie umělé inteligence mohou rovněž negativně ovlivňovat mediální komunikaci s přesahem do vnímání politiky a demokratických procesů. Mohou také způsobit ztrátu některých lidských dovedností, které budou postupně umělou inteligencí nahrazovány procesem automatizace a robotizace. Další významná etická rizika představuje rozvoj AI v oblasti vojenské bezpečnosti a obrany, stejně jako bezpečnostní hrozby spojené s kybernetickou kriminalitou, hybridními válkami či úmyslnými manipulacemi. Nezanedbatelná jsou také rizika vyplývající z nepřesností způsobených nedbalostí, zanedbáním povinností nebo chybnou implementací technologií AI.
Výše zmíněné oblasti zdůrazňují nezbytnost důsledného a pravidelného posuzování rizik spojených s technologiemi AI. Proces řízení rizik by měl klást důraz na bezpečnost, transparentnost a odpovědnost, přičemž klíčovou roli hrají pravidelné kontroly, aby bylo možné účinně předcházet škodám na jednotlivcích i celé společnosti (v souladu s etickým principem nonmaleficence). Právě v těchto oblastech se požadavky AI Act prolínají s etickými standardy.

ZÁVĚR A DOPORUČENÍ

Vzhledem k rychlému rozvoji a stále širší integraci umělé inteligence do různých oblastí každodenního života je nezbytné věnovat pozornost nejen technologickým aspektům, ale také etickým a právním rámcům, které ji provázejí. Nařízení AI Act představuje významný krok k řízení rizik spojených s vysoce rizikovými AI systémy. Klade důraz na bezpečnost, transparentnost a odpovědnost s cílem chránit základní práva jednotlivců, což je klíčové pro posílení důvěry veřejnosti v technologie AI. V rámci etických perspektiv hrají zásadní roli ochrana soukromí, spravedlnost a odpovědné využívání dat. Spolupráce mezi technologickými odborníky, etiky,

legislativci a veřejností je nezbytná pro vývoj spravedlivých a transparentních systémů AI, které budou respektovat základní lidská práva.

Zajištění souladu s požadavky AI Act a etickými standardy by mělo být prioritou pro všechny poskytovatele systémů AI, neboť pouze tímto způsobem lze dosáhnout, aby umělá inteligence sloužila jako nástroj pro zlepšení kvality života a posílení lidských práv, nikoli jako zdroj nových rizik a nerovností. V tomto kontextu je důležité zavést revizní mechanismy, které budou schopny včas identifikovat i malé odchylky od očekávaného prospěchu AI.

V souvislosti s dynamickým vývojem technologií je také nezbytné, aby preventivní opatření a reflexe etických dilemat byly součástí procesu návrhu autonomních systémů. Procesy by měly zahrnovat podrobnou analýzu možných negativních dopadů, včetně ohrožení demokracie, právního státu nebo spravedlivého rozdělení zdrojů. Každé rozhodnutí, ať už lidské nebo automatizované, může mít pozitivní i negativní dopady, a proto je důležité důsledně zvažovat rizika a přínosy. Odborné a veřejné diskuze na národní i globální úrovni o začlenění umělé inteligence do společnosti způsobem, který odpovídá etickým, právním a společenským očekáváním veřejnosti, podporuje evropskou strategii důvěryhodné AI. V závěru uveďme několik doporučení, která mohou přispět k odpovědnému rozvoji a využití technologií umělé inteligence.

Posílení regulace a dohledu: Mezi účinné mechanismy regulace a dohledu by měly patřit pravidelné etické audity a inspekce, aby bylo možné identifikovat a řešit potenciální rizika.

Zvyšování etické odpovědnosti: Podpora průběžného vzdělávání a školení v oblasti etických standardů spojených s AI může přispět k včasné identifikaci a řešení etických dilemat.

Mechanismy pro hlášení a řešení etických otázek: Důležité je vytvoření efektivních nástrojů a procesů pro hlášení etických dilemat spojených se systémy AI, které by zaručily jejich transparentní a neodkladná řešení. Tyto mechanismy by měly být snadno přístupné pro vývojáře, uživatele i veřejnost. Součástí by měly být také mechanismy pro průběžné sledování a řešení stížností a sporů, stejně jako účinná opatření na ochranu oznamovatelů.

Ochrana pracovního trhu: Ke zmírnění negativních dopadů automatizace a robotizace na pracovní trh je nezbytné vypracovat konkrétní programy, které by měly být zaměřeny zejména na ochranu zranitelných skupin pracovníků a zmírňování sociálních nerovností.

Meziinstitucionální koordinace a spolupráce: Posílení spolupráce mezi státní správou, soukromým sektorem a akademickou sférou může přispět k zajištění konzistentního přístupu k vývoji a implementaci systémů AI. Spolupráce by měla zahrnovat sdílení osvědčených postupů a průběžných výsledků výzkumu v oblasti etiky AI.

Otevřená komunikace: Klíčové je posílit transparentnost vůči veřejnosti a zajistit jasnou a srozumitelnou komunikaci o rizicích spojených se systémy AI.

Mezinárodní spolupráce: Vzhledem ke globálnímu charakteru vývoje AI je nutné podporovat mezinárodní spolupráci v oblasti právní regulace a etických požadavků. Společné standardy a sdílení informací na mezinárodní úrovni přispějí k efektivnějšímu zvládání rizik.

LITERATURA

Evropský parlament a Rada EU (2012): *Listina základních práv Evropské unie*. (online). https://eur-lex.europa.eu/legal-content/CS/TXT/HTML/?uri=CELEX:12012P/TXT&from=CS

European Parliament (2020): *Framework of ethical aspects of artificial intelligence, robotics and related technologies*. (online). https://europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.pdf

Hagendorff, Thilo (2020): The Ethics of AI Ethics: An evaluation of guidelines. *Minds & Machines*, 30, 99–120. https://doi.org/10.1007/s11023-020-09517-8

High-Level Expert Group on Artificial Intelligence (2019): *Ethics guidelines for trustworthy AI*. (online). https://aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf

Hu, Margaret (2020): Cambridge Analytica's black box. *Big Data & Society*, 7(2). https://doi.org/10.1177/2053951720938091

Huang, Changwu – Zhang, Zeqi – Mao, Bifei – Yao, Xin (2023): An overview of artificial intelligence ethics. *IEEE Trans Artif Intell*, 4(4), 799–819. https://doi.org/10.1109/TAI.2022.3194503

IEEE (2019): *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. (online). https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf?utm_medium=alias&utm_source=LI&utm_campaign=EAD1e&utm_content=report&utm_term=undefined

Jedličková, Anetta (2022): Etické aspekty rozvoje umělé inteligence. *Anthropologia integra*, 13(2), 55–62. https://doi.org/10.5817/AI2022-2-55

Jedličková, Anetta (2024): Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development. *AI & Soc*, 2024. https://doi.org/10.1007/s00146-024-02040-9

Jobin, Anna – Ienca, Marcello – Vayena, Effy (2019): The global landscape of AI ethics guidelines. *Nat Mach Intell*, 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

NAŘÍZENÍ EVROPSKÉHO PARLAMENTU A RADY (ES) č. 765/2008 ze dne 9. července 2008, kterým se stanoví požadavky na akreditaci a dozor nad trhem týkající se uvádění výrobků na trh. (online). https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=CELEX:32008R0765

NAŘÍZENÍ EVROPSKÉHO PARLAMENTU A RADY (EU) 2024/1689 ze dne 13. června 2024, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci. (online). https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=OJ:L_202401689

Ryan, Mark – Stahl, Bernd Carsten (2021): Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *J Inf Commun Ethics Soc*, 19(1), 61–86. https://doi.org/10.1108/JICES-12-2019-0138

Směrnice Evropského parlamentu a Rady (EU) 2016/2102 ze dne 26. října 2016 o přístupnosti internetových stránek a mobilních aplikací subjektů veřejného sektoru. (online). https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=CELEX:32016L2102

Směrnice Evropského parlamentu a Rady (EU) 2019/882 ze dne 17. dubna 2019 o požadavcích na přístupnost u výrobků a služeb. (online). https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=CELEX:32019L0882