Autorka: Lucia Kobzová

AMNESTY INTERNATIONAL

ZA SPRAVEDLIVOU BUDOUCNOST: AI A LIDSKÁ PRÁVA

Umělá inteligence (AI) již není vzdáleným pojmem ze sci-fi filmů. Stala se neviditelnou, ale všudypřítomnou součástí našeho každodenního života. Rozhoduje o tom, co vidíme na sociálních médiích, pomáhá lékařům odhalovat rakovinu, ale také určuje, kdo dostane hypotéku, koho policie považuje za rizikového a kdo má nárok na sociální dávky.

S příchodem generativní AI (ChatGPT, Gemini a Claude) se tento vývoj radikálně zrychlil. To, co dříve trvalo roky, se nyní děje během několika měsíců. Stojíme na prahu technologické revoluce, která slibuje obrovský pokrok, ale také představuje bezprecedentní výzvu pro ochranu lidských práv.

Technologie mění způsob fungování státu, způsob, jakým pracujeme a jak uplatňujeme své svobody. Jako pedagogové a aktivisté v oblasti lidských práv proto stojíme před zásadní otázkou: Slouží tato technologie nám, nebo my sloužíme jí?

Tato příručka byla vytvořena, protože vzdělávání v oblasti umělé inteligence již nemůže být omezeno pouze na psaní kódu. Stává se nezbytnou součástí občanského vzdělávání, etiky a společenských věd. Cílem je poskytnout vám užitečný nástroj pro orientaci v nové realitě umělé inteligence. Nechceme pouze popisovat rizika a příležitosti. Chceme vám dát konkrétní nástroj, abyste mohli se svými studenty diskutovat o nejdůležitějších otázkách: Jak zajistit, aby lidé a jejich důstojnost zůstali v digitálním věku na prvním místě.

Vítejte v této příručce o světě, kde se algoritmy setkávají s lidskými právy.

Mýtus objektivní stroje: Proč AI nikdy nebude neutrální

Vstupujeme do éry, kdy se hranice mezi fyzickým a digitálním světem stírají. Pro pedagogy a pracovníky s mládeží je zásadní pochopit a sdělit jednu základní pravdu: technologie není neutrální.

Často se setkáváme s představou, že umělá inteligence je jen o něco složitější forma matematiky – čistá, objektivní a zbavená lidských emocí nebo chyb. Předpokládáme, že pokud rozhodnutí učiní počítač, musí být spravedlivé. Opak je však pravdou. Systémy umělé inteligence jsou trénovány na datech z našeho světa. Světa, který je historicky poznamenán rasismem, sexismem, kolonialismem a sociálními nerovnostmi. Když tato data vložíme do

„černé skříňky“ algoritmu, nemůžeme očekávat, že výsledkem bude objektivní pravda. To, co dostaneme, je „zakódovaná zaujatost“. Technologie nejen odráží tyto nerovnosti, ale často je i posiluje.

Jako učitelé máte odpovědnost připravit studenty nejen na to, aby byli pasivními spotřebiteli technologie, ale aby pochopili její dopad na základní práva. Musíme odhalit proces, kterým se lidské hodnoty dostávají do AI.

Hodnoty vstupují do každé fáze cyklu AI

Myšlenka, že se AI učí sama a vyvíjí se nezávisle, je mylná. Umělá inteligence je sociotechnický systém. To znamená, že technologie je neoddělitelně spjata se sociálními procesy. V každé fázi vývoje lidé (vývojáři, analytici dat, manažeři atd.) činí vědomá nebo nevědomá rozhodnutí, která formují výsledné chování systému.
Nejčastějším argumentem je: „Ale AI se učí z dat a data jsou objektivní.“ To však není pravda. Data, na kterých jsou dnešní modely trénovány, jsou odrazem našeho světa. A náš svět je historicky nespravedlivý, rasistický, sexistický a nerovný. Pokud například trénujete AI na historii soudních rozhodnutí, nenaučí se spravedlnosti. Naučí se kopírovat vzorce, které v soudnictví fungují po desetiletí, včetně diskriminace menšin.

Tento proces můžeme rozdělit do tří klíčových fází, ve kterých se promítají lidské hodnoty:

A. Fáze zběru dat

Data jsou v podstatě záznamem minulosti. Pokud trénujeme AI na historických datech, učíme ji historické vzorce chování.

- Historická zaujatost: Pokud trénujeme bankovní algoritmus na datech o půjčkách z 20. století, systém se naučí, že ženy nebo příslušníci menšinových skupin jsou rizikovějšími klienty, protože v minulosti čelili systémové diskriminaci a měli horší přístup k kapitálu. Umělá inteligence tuto zaujatost matematicky formalizuje a přenáší do budoucnosti.
- Neviditelnost dat: Stejně důležité je i to, co v datech chybí. Máme obrovské množství dat o lidech, kteří jsou online, používají smartphony a žijí v bohatších zemích. Data o lidech z marginalizovaných komunit nebo rozvojových zemí často chybí. Pokud systém rozpoznávání obličeje nefunguje spolehlivě u lidí s tmavou pletí, nejde o technickou náhodu, ale o důsledek toho, že tito lidé nebyli v trénovací sadě dostatečně zastoupeni.

B. Fáze návrhu a tréninku

Architektura samotného systému odráží to, co tvůrci považují za důležité. Každý model AI má takzvanou optimalizační funkci – cíl, o jehož dosažení usiluje. Volba tohoto cíle je čistě hodnotovým úsudkem.

Příklad: Proč AI halucinuje? Generativní modely jsou často kritizovány za to, že si vymýšlejí fakta. Ukazuje se, že u ChatGPT se jedná o vědomou volbu:

- Tyto modely nebyly primárně naprogramovány tak, aby říkaly pravdu.
- Byly optimalizovány tak, aby zněly lidsky, plynule a přesvědčivě.

Kdyby vývojáři při navrhování systému upřednostnili hodnotu přiznání nevědomosti před hodnotou uživatelského komfortu, model by na neznámé otázky odpovídal „nevím“. Současné modely však upřednostňují generování falešné, ale uvěřitelné odpovědi, protože tak byly navrženy (hodnotová volba: plynulost > faktická přesnost).

C. Fáze nasazení

I technicky „dokonalý“ model může být při nasazení škodlivý, pokud ignorujeme sociální kontext.

- Pokud nasadíme systém detekce emocí vyvinutý v USA ve školách v odlišné kultuře, bude systém nesprávně interpretovat výrazy tváří studentů. To, co je v jedné kultuře známkou soustředění, může být v jiné kultuře interpretováno jako hněv.
- Rozhodnutí nasadit AI na místech, kde se rozhoduje o osudech lidí (soudnictví, sociální zabezpečení), je samo o sobě politickým rozhodnutím, které upřednostňuje efektivitu a rychlost před individuálním lidským úsudkem.

Co si z toho vzít pro výuku: Umělá inteligence není autorita. Je to nástroj, který zesiluje lidské schopnosti, ale také lidské chyby. Když se díváme na výstupy AI, nedíváme se na objektivní realitu, ale na statistický odhad založený na nedokonalých lidských datech. Učit studenty o AI proto znamená nejen učit je programovat, ale také je učit klást otázky: Kdo tento systém vytvořil? Z jakých dat se učil? A čí pohled na svět v něm chybí?

Špatně definovaný cíl znamená potíže

Jedním z nejkritičtějších momentů ve vývoji umělé inteligence je okamžik, kdy lidé musí převést abstraktní sociální pojmy (jako „dobrý zaměstnanec“, „ohrožené dítě“, „úspěšná léčba“) do jazyka čísel, kterému počítač rozumí. Tento proces se nazývá formalizace úkolů a právě zde dochází k neviditelné, ale zásadní etické změně.

Představte si, že trénujeme systém umělé inteligence pro personální oddělení, aby třídil životopisy. Úkol je jednoduchý: „Najděte nám nejlepší kandidáty.“

Umělá inteligence však nechápe pojem „nejlepší“. Musíme jí dát konkrétní metriku – cíl, který má maximalizovat. V technické terminologii říkáme, že hledáme zástupnou proměnnou pro úspěch. A právě zde začíná řada problémů:

AMNESTY INTERNATIONAL

1. Výběr cíle je volbou hodnot

Jak definujeme úspěšného zaměstnance pro účely algoritmu?

- Možnost A: Je to někdo, kdo zůstává ve společnosti alespoň 5 let? (Důraz na loajalitu).
- Možnost B: Je to někdo, kdo byl povýšen do dvou let? (Důraz na výkon/ambice).

Pokud zvolíme možnost B (rychlé povýšení), vkládáme do systému náš pohled na svět, a sice že agresivita je důležitější než stabilita. Umělá inteligence začne v životopisech hledat vzorce, které korelují s rychlým kariérním postupem.

2. Historická data kopírují nerovnosti

Když AI analyzuje data společnosti za posledních 10 let, odhalí krutou realitu: muži byli nejčastěji rychle povyšováni. Ne nutně proto, že byli schopnější. Možná proto, že ženy častěji přerušovaly svou kariéru, aby se staraly o děti, nebo kvůli neformální diskriminaci ze strany manažerů v té době. Algoritmus nevidí sociální kontext (diskriminaci). Vidí pouze silnou statistickou korelaci: „Mužské atributy v životopise = vysoká pravděpodobnost rychlého povýšení."

3. Výsledek: Dokonalé splnění špatného úkolu

Výsledný model začne penalizovat vše, co se v nových životopisech týká žen (například odkazy na „ženský softbalový tým" nebo mezery v profesní historii), a naopak bude upřednostňovat slovník typický pro mužské uchazeče.

Kdybychom takový systém zavedli, stalo by se následující: ženy by byly systémem automaticky odmítnuty.

V tomto případě stroj neudělal technickou chybu. Stroj fungoval bezchybně. Přesně splnil matematický úkol: „Najít lidi podobné těm, kteří byli v minulosti úspěšní." Úkol však vycházel z mylného předpokladu, že minulost byla spravedlivá a hodná opakování.

Klíčové ponaučení: Definovat cíl umělé inteligence znamená definovat hodnoty budoucnosti. Pokud bez kritického myšlení řekneme AI, aby optimalizovala efektivitu podle starých pravidel, pouze automatizujeme a urychlíme staré nespravedlnosti.

Tip pro aktivitu ve třídě: Tento text je vhodný pro krátkou diskusi se studenty. Otázka pro studenty: „Představte si, že jste ředitelem školy a chcete, aby AI vybrala nejlepší studenty, kteří budou školu reprezentovat. Jaká data byste systému poskytli? Známky? Počet zameškaných hodin? Účast v klubech?" (Poté diskutujte o tom, koho by tato kritéria diskriminovala – např. studenta, který má špatné známky, ale je kreativní, nebo studenta, který zmeškal mnoho hodin kvůli nemoci, i když je talentovaný.)

AMNESTY INTERNATIONAL

Digitální práva jsou lidská práva

Na samém počátku diskuse o umělé inteligenci a lidských právech je důležité poznamenat, že státy po celém světě se v několika deklaracích a dohodách zavázaly dodržovat základní zásadu: lidská práva, která máme v offline světě, musí být chráněna i v online světě.

Umělá inteligence nevytváří nový svět s novými pravidly. Vstupuje do našeho stávajícího světa a testuje meze práv, která nám zaručují mezinárodní úmluvy. Od práva na život po právo na spravedlivý proces. Neexistuje žádná oblast, která by nebyla ovlivněna příchodem algoritmů.

Podívejme se, jak AI mění tradiční lidská práva podle Evropské úmluvy o lidských právech:

I. Když algoritmy rozhodují o životě a smrti (články 2 a 3)

Technologický pokrok nás přivedl do bodu, kdy stroje přímo zasahují do fyzické integrity člověka. To, co zní jako scéna z filmu Terminátor, je dnes realitou na bojištích a v nemocnicích.

1. Digitální dehumanizace: Autonomní zbraňové systémy

(V souvislosti s článkem 2: Právo na život)

Nejkrajnější formou umělé inteligence jsou smrtící autonomní zbraňové systémy („zabijácké roboty“). Tyto stroje mohou vybírat cíle a útočit bez jakéhokoli zásahu člověka.

- Etický problém: Bez ohledu na to, jak dokonalý stroj je, nemůže cítit soucit ani pochopit hodnotu lidského života. Redukuje lidskou bytost na soubor datových bodů (tepelná signatura + tvar objektu = cíl).
- Právní vakuum: Pokud robot spáchá válečný zločin (např. zabije se vzdávajícího vojáka), kdo je za to zodpovědný? Programátor v kanceláři? Velitel, který jej aktivoval? Nebo výrobce? Absence lidské kontroly vytváří nebezpečnou mezeru v odpovědnosti.

2. Algoritmické triage v zdravotnictví

(V souvislosti s článkem 2 a zákazem diskriminace)

Ve Spojených státech se zdravotní pojišťovny při výběru pacientů, kteří potřebovali zvláštní péči, spoléhaly na algoritmus Optum.

- Selhání: Algoritmus systematicky upřednostňoval bílé pacienty před černými, i když byli stejně nemocní.

- Proč? Systém byl nastaven tak, aby předpovídal budoucí finanční náklady (jako zástupný ukazatel zdravotního stavu). Vzhledem k tomu, že Afroameričané měli v minulosti horší přístup k léčbě, utratili méně peněz. Algoritmus to interpretoval tak, že jsou zdravější.
- Výsledek: odepření životně důležité péče na základě matematiky ovlivněné rasovými předsudky.

3. Pseudověda na hranici: detekce lži

(V souvislosti s článkem 3: Zákaz mučení a nelidského zacházení)

Projekt iBorderCTRL, testovaný na hranicích EU, se snažil odhalit lži mezi uprchlíky analýzou mikro-výrazů na jejich tvářích. Vědecká komunita to považuje za pseudovědu. Stres, únava z cestování a kulturní rozdíly v mimice mohou systém oklamat. Podrobit traumatizované osoby výslechu strojem, který je může na základě jejich mimiky poslat zpět do nebezpečí, hraničí s nelidským zacházením.

Poznámka: Podle nové evropské legislativy o umělé inteligenci by analýza emocí měla být ve většině případů zcela zakázána.

II. Spravedlnost v černé skříňce (článek 6)

Základním pilířem právního státu je, že spravedlnost musí být nejen vykonána, ale také musí být viditelná. Obžalovaní musí rozumět důkazům, aby se mohli bránit. Umělá inteligence, která funguje jako „černá skříňka“, tento princip podkopává.

1. Případ COMPAS: Odsouzen algoritmem

V USA používají soudci software COMPAS k odhadu rizika recidivy. Vyšetřování organizace ProPublica odhalilo znepokojivou pravdu:

- systém dvakrát častěji než u bělochů nesprávně označoval černochy jako „vysoce rizikové“.
- V důsledku toho byli lidé na základě vzorce algoritmu odsouzeni k přísnějším trestům nebo jim byla zamítnuta kauce.

2. Prediktivní policejní práce: Matematika jako samosplňující se proroctví

Policie využívá umělou inteligenci k předpovídání míst, kde dochází k trestné činnosti. Systém se učí z historických policejních záznamů. Pokud se policie v minulosti zaměřovala na chudé čtvrti, data ukážou, že se tam vyskytuje „vysoká kriminalita“.

- Slučka zpětné vazby: Algoritmus pošle hlídku zpět do stejné čtvrti -> hlídka najde drobnou kriminalitu -> data se potvrdí -> algoritmus tam pošle hlídku znovu.
- Výsledkem je nadměrné policejní hlídkování v menšinových čtvrtích, zatímco kriminalita v „bílých“ čtvrtích zůstává bez povšimnutí.

AMNESTY INTERNATIONAL

III. Konec soukromí a svobody myšlení (články 8 a 9)

Právo na soukromí v digitálním věku se zásadně mění. Mění se v boj o kontrolu nad vlastním obličejem a myšlenkami.

1. Masové biometrické sledování: Kampaň „Ban the Scan“ (Zakažte skenování)

Technologie rozpoznávání obličeje mění veřejné prostory na oblasti pod neustálým státním dohledem.

- Případ Roberta Williamse (USA): Tento muž byl zatčen před zraky své rodiny za zločin, který nespáchal. Algoritmus omylem porovnal jeho starou fotografii s nekvalitním snímkem pachatele. Rozpoznávání obličeje je nejen invazivní technologie, ale také rasově zaujatá – častěji se mýlí u lidí s tmavou pletí.
- Dopad: Pokud vás kamera dokáže identifikovat kdekoli, ztrácíte svou anonymitu.

2. Útoky na svobodu myšlení

Tradičně jsme věřili, že naše myšlenky jsou nedotknutelné.

- Inferenční algoritmy: Sociální sítě mohou z vašich „lajků“ odvodit vaši sexuální orientaci, politické názory nebo duševní stav – věci, které jste jim nikdy neřekli.
- Nudging: Tyto údaje se používají pro mikrotargetingovou reklamu a politickou manipulaci, která obchází naše kritické myšlení. Přestáváme být autonomními bytostmi a stáváme se objekty manipulace.

IV. Umělé umlčení občanské společnosti (články 10 a 11)

Svoboda projevu a shromažďování jsou základními pilíři demokracie. Umělá inteligence mění také oblast občanské angažovanosti.

1. Automatizovaná cenzura (případ Syrian Archive)

Algoritmy YouTube určené k mazání „teroristického obsahu“ smazaly tisíce videí dokumentujících válečné zločiny v Sýrii. Stroj nedokázal rozlišit mezi propagandou ISIS a důkazy shromážděnými aktivisty za lidská práva.

- Výsledek: klíčové důkazy pro budoucí soudní řízení byly ztraceny. Moderování obsahu pomocí umělé inteligence často funguje jako nástroj, který ničí legitimní projevy.

2. Odrazující účinek na protesty

Pokud byste věděli, že na náměstí jsou kamery s funkcí rozpoznávání obličejů, šli byste na protivládní demonstraci? Mnoho lidí by raději zůstalo doma.

- Tento strach z následků se nazývá odrazující účinek. Technologie nemusí fyzicky bránit protestům; stačí vytvořit atmosféru strachu, která odradí lidi od uplatnění jejich práva na shromažďování.

AMNESTY INTERNATIONAL

V. Diskriminace skrytá v zákoně (článek 14)

Nejzákeřnější vlastností diskriminace založené na umělé inteligenci je to, že se jeví jako objektivní. Skrývá se za fasádou byrokratické neutrality.

- Skandál v Nizozemsku: Finanční úřad použil algoritmus k vyhledávání podvodníků s dětskými přídavky. Jako rizikový faktor nastavil „dvojí občanství“. Výsledek? Tisíce nevinných přistěhovaleckých rodin bylo zničeno vymáháním dluhů. Jednalo se o institucionální rasismus prováděný pomocí kódu.
- Amazon: Společnost musela zrušit svou AI pro nábor zaměstnanců, protože se naučila penalizovat životopisy obsahující slovo „žena“. Jelikož byla trénována na historických datech (kde dominovali muži), hodnotila ženy jako méně vhodné kandidátky.

Od etiky k závazkům: rámec lidských práv pro umělou inteligenci

Debata o tom, jak zmírnit negativní dopady umělé inteligence, prošla v uplynulém desetiletí zásadní proměnou. Ještě nedávno jsme žili v éře digitálního Divokého západu. Technologičtí giganti nás přesvědčovali, že regulace potlačí inovace a že jejich vlastní etické kodexy budou stačit. Ve skutečnosti se však ukázalo, že nezávazné sliby nás před skutečnou újmou neochrání.

Pod tlakem různých organizací, jako jsou Amnesty International a Access Now, se situace mění. Přecházíme od vágní etiky AI k právně vymahatelným hranicím.

1. Zlomový bod: Torontská deklarace

Klíčovým momentem této změny byl květen 2018 a přijetí Torontské deklarace. Tento dokument, iniciovaný organizacemi Amnesty International a Access Now, smetl výmluvy technologických společností. Deklarace jasně stanovila: Stávající zákony o lidských právech se vztahují i na strojové učení.

Umělá inteligence není nad zákonem. Je produktem lidského designu. Pokud algoritmus diskriminuje (například znevýhodňuje ženy při poskytování úvěrů), nejde o technickou chybu, ale o porušení práva na rovnost. Torontská deklarace tak položila základy pro to, co vidíme dnes – přechod od vágní etiky k pevným pravidlům.

2. Kdo je za co zodpovědný?

Abychom se neztratili v bludišti pravidel, musíme (v souladu s principy OSN) rozlišovat mezi dvěma úrovněmi odpovědnosti. Jedná se o klíčové rozlišení:

AMNESTY INTERNATIONAL

A. Stát má povinnost CHRÁNIT (Hard law)

Stát nemůže jen pokrčit rameny a tvrdit, že technologie je příliš rychlá. Jeho úlohou je přijímat závazné zákony.

- Zákaz delegování spravedlnosti: Stát nesmí předávat své klíčové pravomoci strojům, pokud by to obcházelo spravedlnost. Pokud o vašem osudu rozhoduje soud nebo policie, musí za tím stát osoba, která nese odpovědnost.
- Právo na obhajobu: Pokud jste vyšetřováni algoritmem, musíte mít stejná práva na obhajobu a přístup k důkazům, jako kdybyste byli vyšetřováni vyšetřovatelem. Argument černé skříňky nemůže převážit nad právem na spravedlivý proces.

B. Společnost má povinnost RESPEKTOVAT (Due Diligence)

Pro soukromý sektor se blíží konec éry výmluv typu „Je to černá skříňka, nevíme, proč se tak rozhodla“.

- Due diligence v oblasti lidských práv: To je nový standard. Společnosti musí prokázat, že před spuštěním systému aktivně vyhledávaly rizika. Musí zkontrolovat, zda jejich data neobsahují rasové nebo genderové předsudky, a otestovat dopad na zranitelné skupiny.
- Odpovědnost za výsledek: Pokud společnost uvede na trh diskriminační nástroj, nese za něj odpovědnost, bez ohledu na to, zda záměrem bylo způsobit škodu, nebo zda se jednalo pouze o nedbalost.

3. Jak se svět dnes brání?

Torontská deklarace vyvolala lavinu nových etických závazků v oblasti umělé inteligence. Dnes máme k dispozici celý ekosystém mezinárodních dohod a zákonů, které tvoří bezpečnostní síť pro naše práva.

OECD: Zásady pro umělou inteligenci (2019)
Organizace pro hospodářskou spolupráci a rozvoj (OECD) byla první mezinárodní institucí, která definovala standardy pro důvěryhodnou AI.

- O co jde: Tyto zásady, podepsané více než 40 zeměmi (včetně Slovenska a České republiky), zdůrazňují, že AI musí respektovat demokratické hodnoty a zásady právního státu. Ačkoli nejsou přímo vymahatelné soudy, slouží jako základ pro vnitrostátní právní předpisy a stanovují laťku pro to, co je v demokratickém světě přijatelné.

UNESCO: Globální etický kompas (2021)
Zatímco OECD sdružuje hlavně bohaté země, UNESCO dosáhlo globálního konsensu. V roce 2021 přijalo 193 členských států Doporučení o etice umělé inteligence.

- O co jde: Jedná se o první skutečně globální dokument, který výslovně stanoví, že technologie musí být pod kontrolou člověka a musí být vysvětlitelné. UNESCO zavedlo koncept etického posuzování dopadů (EIA), nástroj, který pomáhá vládám předvídat rizika umělé inteligence, než se stanou realitou.

Rada Evropy: První mezinárodní úmluva o AI

Rada Evropy (která sdružuje 46 zemí a stojí za Evropským soudem pro lidská práva) má rámcovou úmluvu o umělé inteligenci.

- Co to je: Jedná se o první právně závaznou mezinárodní smlouvu na světě, která se zaměřuje výhradně na průnik umělé inteligence a lidských práv, demokracie a právního státu. Na rozdíl od obchodních předpisů se tato úmluva zabývá ochranou jednotlivců před zneužitím moci prostřednictvím technologie.

Evropská unie: Zákon o umělé inteligenci (AI Act)

EU přijala první komplexní legislativu na světě, která klasifikuje AI podle úrovně rizika a stanoví jasná pravidla pro každou kategorii.

- Červené čáry: Systémy s „nepřijatelným rizikem“ jsou v Evropě zcela zakázány. Patří mezi ně například:
  - Sociální bodování: Bodování občanů státem na základě jejich chování (model známý z Číny).
  - Biometrická kategorizace: Třídění lidí na základě rasy, politických názorů nebo sexuální orientace.
  - Rozpoznávání emocí: Používání AI k detekci emocí na pracovištích nebo ve školách.

Zákon o umělé inteligenci vysílá jasný signál – práva občanů mají v Evropě přednost před zisky technologických společností.

Pozor na etické vymývání mozků

Existuje také fenomén zvaný „etické praní“. Je podobný ekologickému „zelenému praní“. Technologická společnost zveřejní hezky vypadající etický kodex plný slov jako spravedlnost a dobrota, zřídí interní etickou komisi (která nemá žádnou skutečnou moc) a předstírá, že je odpovědná. Často se jedná pouze o PR strategii, která má oddálit zavedení skutečných, přísných regulací. Skutečná ochrana práv vyžaduje nezávislého arbitra (soudy, úřady), nikoli interní komisi placenou samotnou společností.

Co to znamená pro školu?

Cílem této kapitoly není démonizovat technologii. Chceme však studenty vybavit kritickými dovednostmi digitálního občanství. Ve třídě by se nemělo učit pouze programování umělé inteligence, ale také to, že:

1. Máte právo na vysvětlení: Pokud stroj učiní rozhodnutí, které se vás týká (např. přijetí do školy, půjčka), máte právo vědět proč.
2. Máte právo na nápravu: Musí existovat způsob, jak se odvolat k reálné osobě.
3. Technologie není nad zákonem: Ani ten nejmodernější algoritmus nesmí porušovat ústavu.

Závěr

Vzdělávání v oblasti umělé inteligence a lidských práv neslouží pouze k varování před dystopickou budoucností. Jeho cílem je posílit postavení mladých lidí. Když studenti pochopí, jak tyto systémy fungují a jaká jsou jejich práva, přestanou být pasivními objekty sledování a stanou se aktivními digitálními občany. Úlohou učitele je poskytnout jim v tomto složitém terénu kompas, jehož střelka ukazuje na lidskou důstojnost.

AMNESTY INTERNATIONAL

Zdroje

1. Technology - Amnesty International https://www.amnesty.org/en/what-we-do/technology/
2. Ethical AI principles won't solve a human rights crisis - Amnesty International, https://www.amnesty.org/en/latest/research/2019/06/ethical-ai-principles-wont-solve-a-human-rights-crisis/
3. Toronto Declaration, https://www.torontodeclaration.org/
4. The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems https://www.accessnow.org/press-release/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems/
5. AI Act | Shaping Europe's digital future - European Union https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
6. High-level summary of the AI Act | EU Artificial Intelligence Act https://artificialintelligenceact.eu/high-level-summary/
7. Návrh zákona o organizácii štátnej správy v oblasti umelej inteligencie - Škubla & Partnerihttps://www.skubla.sk/clanky/navrh-zakona-o-organizacii-statnej-spravy-v-oblasti-umelej-inteligencie
8. Zodpovedná digitalizácia: MIRRI predstavilo návrhy zákonov o umelej inteligencii a správe údajov | Ministerstvo investícií, regionálneho rozvoja a informatizácie SR, otvorené novembra 25, 2025, https://mirri.gov.sk/aktuality/digitalna-agenda/zodpovedna-digitalizacia-mirri-predstavilo-navrhy-zakonov-o-umelej-inteligencii-a-sprave-udajov/
9. EU Simplification: Throwing human rights under the Omnibus - Amnesty International, otvorené novembra 25, 2025, https://www.amnesty.org/en/latest/news/2025/11/eu-simplification-throwing-human-rights-under-the-omnibus/
10. Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal - Amnesty International, https://www.amnesty.org/en/documents/eur35/4686/2021/en/
11. How We Did It: Amnesty International's Investigation of Algorithms in Denmark's Welfare System https://gijn.org/stories/amnesty-internationals-investigation-algorithms-denmarks-welfare-system/
12. Ban The Scan - Amnesty Internationalhttps://www.amnesty.org/en/petition/ban-the-scan-petition/
13. Ban dangerous facial recognition technology that amplifies racist policing, https://www.amnesty.org/en/latest/press-release/2021/01/ban-dangerous-facial-recognition-technology-that-amplifies-racist-policing/
14. Ban the Scan, otvorené novembra 25, 2025, https://www.banthescan.org/
15. USA: Facial recognition technology reinforcing racist stop-and-frisk policing in New York – new research - Amnesty International https://www.amnesty.org/en/latest/news/2022/02/usa-facial-recognition-technology-reinforcing-racist-stop-and-frisk-policing-in-new-york-new-research/
16. Stop Killer Robots – Less Autonomy, More humanity. https://www.stopkillerrobots.org/
17. Syrian Archive | Nesta, https://www.nesta.org.uk/feature/ai-and-collective-intelligence-case-studies/syrian-archive/
18. Assembling Atrocity Archives for Syria: Assessing the Work of the CIJA and the IIIM - Oxford Academic https://academic.oup.com/jicj/article/19/5/1193/6423113
19. Súpis všetkých metodických materiálov učebných osnov AI - Kurikulum umělé inteligence, otvorené novembra 25, 2025, https://kurikulum.aidetem.cz/sk/supis-vsetkych-metodickych-materialov-metodik-ucebneho-planu-ai/

AMNESTY INTERNATIONAL

Zdroje

20. Koncepcia učebných osnov umelej inteligencie pre základné a stredné školy - Kurikulum umělé inteligence, otvorené novembra 25, 2025, https://kurikulum.aidetem.cz/sk/koncepcia-ucebnych-osnov-umelej-inteligencie-pre-zakladne-a-stredne-skoly/
21. Soupis všech metodických materiálů AI kurikula - Kurikulum umělé inteligence - AI dětem, otvorené novembra 25, 2025, https://kurikulum.aidetem.cz/soupis-vsech-metodickych-materialu-metodik-ai-kurikula/
22. Informatika_Karty_Etika_Etika vývoje a využití technologií, otvorené novembra 25, 2025, https://kurikulum.aidetem.cz/metodiky/karty/Informatika_Karty_Etika_01_etika-vyvoje-a-vyuziti-technologii_plan-lekce.pdf
23. https://aiethicsframework.substack.com/p/when-healthcare-ai-goes-wrong-the