Etika informačných a komunikačných technológií¹

Abstrakt

Cieľom tohto textu je uviesť etiku informačných a komunikačných technológií (IKT) a obhájiť jej potrebnosť ako samostatnej disciplíny aplikovanej etiky. V prvej časti predstavujeme viacero druhov IKT - umelú inteligenciu, roboty, analýzu veľkých dát a virtuálnu realitu - a identifikujeme zdanlivo nové etické problémy, ktoré súvisia s ich používaním a vývojom. Upozorňujeme pritom na naliehavosť týchto problémov a na fakt, že etická reflexia a regulácia spravidla zaostávajú za ich rapídny vývojom. V druhej časti explikujeme pojem etiky IKT a vysvetľujeme, prečo sa odchyľujeme od vymedzení iných autorov. Domnievame sa, že etika IKT musí mať aj proaktívny aspekt, a že by preto mala reagovať aj na tie problémy, ktoré vyvstávajú už vo fáze vývoja, ako aj na problémy skryté a budúce. V tretej časti argumentujeme, že niektoré z vymenovaných problémov sú skutočne nové, resp. unikátne pre IKT. Toto spolu s ďalšími faktormi podľa nášho názoru opodstatňuje etiku IKT ako samostatnú disciplínu. Na záver zhrňame niekoľko odporúčaní, ktoré by mohli pomôcť riešiť etické problémy v IKT.

Kľúčové slová

Informačné a komunikačné technológie - Aplikovaná etika - Regulácia - Umelá inteligencia - Robot - Analýza veľkých dát - Virtuálna realita

Autori

Adam Greif

Department of Philosophy and History of Philosophy

Faculty of Arts, Comenius University in Bratislava

Šafárikovo námestie 6, 814 99 Bratislava, Slovak Republic

e-mail: greif2@uniba.sk

Juraj Podroužek

VP Customer Success, Pygmalios, s.r.o.

Trenčianska 53/B, 821 09 Bratislava, Slovak Republic

e-mail: j.podrouzek@pygmalios.com

¹ Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-14-0510.

Etické problémy v IKT

Práca s modernými informačnými a komunikačnými technológiami so sebou prináša stále viac problémov, ktoré vo svojej povahe nie sú technické, ale etické. Emergentné technológie ako virtuálna realita, internet vecí, alebo umelá inteligencia nám na jednej strane prinášajú bezprecedentné možnosti konania, no na strane druhej môžu mať hlboký a nepredvídateľný dopad na spoločnosť. Technológia nám dáva väčšiu moc, keďže nám umožňuje robiť to, čo sme doteraz robiť nemohli. Etická reflexia technológie, v našom prípade tej informačnej a komunikačnej (ďalej len „IKT“), nás však vedie k otázke, či to, čo vďaka nej robiť dokážeme, robiť aj môžeme.

Táto otázka však niekedy nemá jasnú odpoveď. Nové IKT sa neraz nachádzajú vo „vákuu“ etickej reflexie a právnej regulácie, takže v prípade, že začneme pochybovať o etickej prijateľnosti alebo o legálnosti istého použitia istej IKT, nemôžeme sa oprieť o prijaté normy. Naša pochybnosť je ďalej komplikovaná tým, že nová IKT môže vytvoriť celkom novú, bezprecedentnú situáciu, a tým vytvoriť aj nový, originálny problém. Aby sme k nemu vedeli zaujať etické stanovisko, potrebujeme hľadať analógie medzi novou, bezprecedentnou situáciou a staršou, morálne jasnejšou situáciou. V priebehu tohto procesu môžeme zistiť, že určité hodnotovo zaťažené výrazy ako napríklad *súkromie*, *anonymita*, *vlastníctvo*, *verejný priestor*, nevieme v nových situáciách použiť bez toho, aby sme explikovali alebo menili ich význam.

Ťažkosti ako tieto nás budú stretávať stále častejšie a s väčšou naliehavosťou. Môžeme totiž pozorovať, že IKT sú rýchlo sa rozvíjajúcim technologickým smerom, pričom môžeme predpokladať, že ním ostanú aj v dohľadnej budúcnosti. Tento predpoklad je eticky zaujímavý po prvé preto, lebo potenciál IKT meniť spoločenské pomery sa dá len ťažko zveličiť. Stačí si vybaviť, aký vplyv na spoločnosť malo rozšírené používanie počítačov, internetu alebo sociálnych médií. Po druhé je eticky zaujímavý preto, lebo sa nezdá, že by sme na tieto zmeny boli pripravení. Ako sme už spomenuli, etická reflexia a právna regulácia spravidla zaostávajú za najnovšími technologickými možnosťami.² Domnievame sa, že tento stav vecí nás nemôže ponechať nečinnými a núti nás urýchlene skúmať morálne problémy súvisiace s IKT.³ V tomto texte preto predkladáme uvedenie do etiky IKT ako aj jej obhajobu ako samostatnej disciplíny.

Tak, ako je možné novú technológiu využívať morálne prijateľným spôsobom, je možné ju používať morálne neprijateľným spôsobom. Keďže nám nové technológie dávajú nové možnosti konania, dávajú nám aj nové možnosti zneužitia. Etika IKT preto

² Situácia sa začína meniť až v posledných rokoch. Napríklad, zrejme najznámejší profesijný kódex IT komunity, ACM Code of Ethics and Professional Conduct, nebol od roku 1992 až do nedávna aktualizovaný. K jeho aktualizácii došlo až po viac ako 26 rokoch, v júni 2018. Rovnako kodifikácie posilňujúce právo na vysvetlenie, alebo právo na informovaný súhlas, sa v EU reimplementovali až v roku 2018 v rámci GDPR. Pritom technológie založené na deep learning alebo internete vecí, ktorých sa obe úpravy výrazne dotýkajú, sa masívne využívajú v praxi už niekoľko rokov. Nehovoriac o autonómnych vozidlách, ktoré už teraz začínajú byť súčasťou bežnej cestnej premávky, no neexistuje široký konsenzus ohľadom ich etického používania. Vznikajúci rozdiel medzi technologickými možnosťami a ich etickou reflexiou sa pokúša Európska únia dobiehať zriaďovaním expertných skupín a poradných orgánov, akými sú napríklad Ethics Advisory Group (EAG), alebo European Commission's High-Level Expert Group on Artificial Intelligence (AI HLEG).

³ V tomto bode sa ponúka otázka, či je emergentný charakter IKT a naša nepripravenosť na dôsledky, ktoré prinášajú, znepokojivým potvrdením rozkolu medzi našou schopnosťou vyvíjať nové a účinné technológie a schopnosťou ich zodpovedne používať. Pozri Jonas, H., *Princip odpovědnosti*. Praha, Oikoymenth 1997.

musí identifikovať minimálne to, ktoré použitie danej IKT treba považovať za zneužitie.⁴ Niektoré prípady zneužitia IKT budú očividné a zrejmé, no iné budú nejasné alebo kontroverzné.

Snahu identifikovať zneužitie novej technológie možno považovať za súčasť etickej metódy a radi by sme ju odlíšili od techno-pesimizmu či luddizmu. Tým, že sa budeme venovať morálnym problémom IKT, očividným, skrytým, či budúcim, by sme mohli vyvolať dojem, že IKT majú len negatíva. To samozrejme nie je pravda. Adekvátne etická reflexia novej technológie by mala dať primeranú váhu aj jej pozitívam a prínosom.

Keďže chceme hovoriť o etike IKT, mali by sme určiť predmet jej skúmania, resp. morálne problémy, ktorými sa zaoberá. Už na prvý pohľad môžeme rozlíšiť viaceré problémové okruhy, ako umelú inteligenciu, autonómne zariadenia, virtuálnu realitu, zbieranie veľkých dát, internet a ďalšie. Aby sme ukázali, že tematika IKT je eticky bohatá, v nasledujúcich pasážach si prvé štyri okruhy priblížime a začneme problémami, ktoré súvisia s umelou inteligenciou (ďalej len „UI“).

Zariadenia s UI sú dnes schopné vykonávať aktivity, ktoré sme doteraz považovali za výsostne ľudské: hrať hry, rozoznávať tváre, vyjadrovať emócie či rozprávať v prirodzenom jazyku. S časom sa ich schopnosti budú rozširovať a nie je prehnané očakávať, že v budúcnosti sa rozšíria natoľko, že prinesú revolučné zmeny v medicíne, justícii, vzdelávaní, či vede, alebo že nám pomôžu riešiť globálne problémy. Keďže prínosy UI môžu byť ohromné a keďže existuje značná finančná motivácia pre to, aby bola vyvíjaná a uplatnená, je pochopiteľné, že UI je predmetom značného vedeckého a priemyselného záujmu.

Vzhľadom na jej schopnosti delíme UI na tzv. slabú a silnú. Slabá, resp. úzka UI je navrhnutá na riešenie špecifickej a úzko ohraničenej množiny úloh, akými sú napríklad triedenie e-mailového spamu, rozpoznávanie tvárí či hranie šachu. Dnes je úzka UI už bežnou súčasťou nášho života.

Charakteristickým znakom UI je schopnosť učiť sa. Strojové učenie postupuje vo svojich najjednoduchších formách tak, že UI dostane úlohu, ktorú má riešiť, kritérium, na základe ktorého sa posudzuje miera úspešnosti riešenia problému a sadu tréningových dát, na ktorých sa učí, ako má úspešné riešenie úlohy vyzeráť.⁵ Napríklad, úlohou UI môže byť odstránenie spamu z doručenej elektronickej pošty, kritériom úspechu môže byť miera spokojnosti používateľa e-mailu a tréningovými dátami informácie o tom, ktoré správy používateľa v minulosti označili za spam.

Strojové učenie má však svoje problémy. Prvý problém súvisí s predsudkami v tréningových dátach. Ak sú tréningové dáta, na ktorých sa algoritmus učí, skreslené, potom môže byť skreslené aj riešenie, ktoré UI vyprodukuje. Samozrejme, ak má UI za úlohu triediť len elektronicкую poštu a produkuje skreslené výsledky, nemusí vzniknúť vážnejší etický problém. UI algoritmy však dnes využívajú aj také nástroje, ktoré rozhodujú o udelení poistenia, pôžičky či dokonca spolurozhodujú o podmienkach prepustení z väzenia. Napríklad algoritmus COMPAS dostal za úlohu odhadovať mieru recidívy u väzňov v niektorých štátoch USA. Výsledkom jeho činnosti je skóre, ktoré môže vo výsledku ovplyvniť výšku trestu väzňa alebo jeho podmienkové prepustenie. Ukázalo sa však, že skóre určené cez COMPAS prisudzuje disproporčne vysoké

⁴ Zneužitia IKT však nemusia vyčerpávať všetky morálne problémy IKT, keďže niektoré IKT môžu byť problematické inherentne, napr. autonómne zbrane alebo od človeka inteligentnejší umelý agent.

⁵ Vychádzame z Mitchell, T., *Machine Learning*. McGraw-Hill 1997, s. 2.

negatívne skóre Afroameričanom.⁶ Z technickej povahy použitého algoritmu však nie je jednoduché (a potenciálne ani len možné) vysvetliť, prečo systémy ako COMPAS produkujú skreslené výsledky.⁷ Neželané vlastnosti, ktorými disponujú tieto systémy, tak posilňujú hlasy o potrebe regulácie systémov s UI.

Druhý problém UI sa týka práve vysvetliteľnosti jej rozhodnutí a tretí jej predvídateľnosti. Tieto dva problémy si však priblížime až v rámci diskusie o autonómnych zariadeniach. Teraz prejdeme k diskusii o problémoch mocnejšie verzie UI - ku konceptu silnej UI.

Silná, resp. široká UI je momentálne ešte len v štádiu vývoja. Ak chápeme inteligenciu ako schopnosť riešiť problémy, tak silná UI bude mať schopnosť riešiť všetky problémy, ktoré dokáže riešiť človek. Ak sa stane skutočnosťou, dokáže oveľa viac, ako len triediť poštu či odhadovať mieru recidívy. Vybavená strojovým učením, veľkou výpočtovou silou a prístupom k veľkým sadám dát môže silná UI pomocou procesu rekurzívneho seba-zdokonaľovania ďaleko presiahnuť ľudské schopnosti.⁸ UI, ktorá zásadne prekoná ľudskú inteligenciu, sa niekedy nazýva *superinteligencia*.

Môže byť náročné si jasne predstaviť, ako veľmi môže byť superinteligencia inteligentná. Ako upozorňuje Bostrom, sme zvyknutí pomeriavať IQ v pomerne úzkom rozpätí možnej ľudskej inteligencie. No superinteligencia nemusí byť od človeka inteligentnejšia len o toľko, o koľko bol Einstein inteligentnejší od dedinského prostáčka, ale o toľko, o koľko bol Einstein inteligentnejší od hmyzu.⁹ Je tiež dôležité si uvedomiť, že zdroj superinteligencie zo silnej UI ľudskej úrovne môže byť prekvapivo rýchly (to znamená, že môže prebehnúť v rádoch hodín alebo dní). Takáto *inteligentná explózia* alebo *singularita* by predstavovala bod v technologickom vývoji ľudstva, ktorým začne skutočne nepredvídateľný a nekontrolovateľný vývoj technológie.

Prospekt silnej UI pochopiteľne vyvoláva obavy, ktoré majú úplne alebo čiastočne ne-technickú povahu. Bude možné silnú UI kontrolovať? Ak áno, nebude po svojom zrode využitá ako zbraň, resp. zneužitá na získanie mocenskej prevahy? Ak ju bude možné kontrolovať a nebude zneužitá, nemáme sa obávať, že naše želania bude plniť kontraproduktívne?¹⁰ Ak vyriešime aj tento problém, aké základné hodnoty a ciele by silná UI mala mať a ako jej ich nastavíme? Komu by mala silná UI slúžiť?

Skeptici ohľadom silnej UI si myslia, že stroj s tak závratnými schopnosťami bude ťažké (ak nie nemožné) kontrolovať a bez katastrof používať, obzvlášť vtedy, ak budeme musieť UI dobre zostaviť a nastaviť už na prvý pokus. Preto by sme podľa nich

⁶ S odhalením prišla investigatívna redakcia ProPublica: Angwin, J. et al., Machine Bias. *ProPublica*, 23. mája 2016. Dostupné na:

<<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.

⁷ Systém COMPAS používa algoritmy tzv. deep learningu, pri ktorých až do nedávna nebolo možné určiť, na základe akých parametrov došlo k danému rozhodnutiu stroja. Ako najpravdepodobnejšie sa v prípade COMPASu javí vysvetlenie, že jeho výsledky boli ovplyvnené práve tréningovými dátami z amerického kriminálneho systému, o ktorom je známe, že je poznačený rasovými predsudkami. Ak by tomu bolo tak, použitie takéhoto nástroja má aj svoju svetlú stránku, keďže by efektívne demonštrovalo zaujatosť daného kriminálneho systému.

⁸ Rekurzívne sebazdokonaľovanie spočíva v tom, že inteligentný stroj, ktorý dokáže riešiť tie isté problémy, aké dokáže riešiť človek, by mal byť schopný aj navrhovať inteligentné stroje. Takýto inteligentný stroj by mohol dostať za úlohu vytvoriť lepšiu verziu seba, pričom by jeho nová verzia bola lepšia vo vytváraní inteligentných strojov, etc.

⁹ Bostrom, N., *Superinteligence: Až budú stroje chytrejšie než lidé*. Praha, Prostor 2017, s. 106-128.

¹⁰ Silná UI by mohla plniť naše želania kontraproduktívne. Napríklad, ak by sme jej dali za úlohu urobiť ľudstvo šťastným, čo zaručí, že sa nezabaví ľudského nešťastia tým, že ľudom odstráni časť mozgu, ktorá je zodpovedná za nepríjemné pocity? Z hľadiska priamočiareho uvažovania stroja by išlo o úspešné splnenie úlohy. Ide o ilustráciu tzv. *value-alignment* problému.

mali dobre zvážiť, či ho vôbec chceme vyvinúť. Je možné, že niektoré technológie sú aj napriek ich benefítom príliš nebezpečné na to, aby existovali.

Treba však povedať, že do vývoja silnej UI už dnes investuje ohromné množstvo prostriedkov a energie.¹¹ Optimisti ohľadom UI preto zdôrazňujú, že prísľub UI je tak veľký, že výskum je prakticky nezastaviteľný, a že aj keby sme ho zakázali, pokračovali by v ňom záujmové skupiny s menej ako morálnymi a legálnymi štandardmi. Preto by sme podľa nich mali silnú UI vyvinúť aj napriek jej rizikám, keďže by sme tým zabránili, že skončí v zlých rukách.

Umiernení v tejto debate by tvrdili, že ak chceme silnú UI vyvinúť, mali by sme si prinajmenšom dať veľmi záležať na tom, aby sme ju mali pod kontrolou, aby bola priateľská a aby slúžila spoločnému dobru. To znamená, že by sme mali venovať veľké úsilie a prostriedky na to, aby sme si svoj postup veľmi dobre premysleli a zaistili, že silná UI bude bezpečná.

Domnievame sa, že vývoj silnej UI predstavuje hlboký morálny problém. Zdá sa, že vzhľadom na všetok úžitok, ktorý môže priniesť, máme dobrý dôvod ju vytvoriť. No zároveň sa zdá, že vzhľadom na všetky možné riziká, ktoré sú s ňou spojené, máme aj dobré dôvody ju nevytvoriť. Máme teda k dispozícii presvedčivé dôvody prijať dva nezlučiteľné postoje a dostávame sa do dilemy. No aj bez ohľadu na to, či bude silná UI voči ľuďstvu priateľská alebo nepriateľská, či bude alebo nebude zneužitá, môže sa stať „naším posledným vynálezom.“¹² Keďže vývoj silnej UI predstavuje existenčné riziko pre všetkých ľudí na planéte, dá sa argumentovať, že jej vývoj nie je otázkou, ktorá by mala byť zodpovedaná len v rámci investičných porád či IT laboratórií, ale že je otázkou celospoločenskou.

Silná UI je s najväčšou pravdepodobnosťou technológiou budúcnosti. Ďalší okruh problémov etiky IKT však prináša naliehavejšie problémy. Ide o autonómne zariadenia ako roboty, drony a samoriadiace autá (ďalej len „roboty“). Roboty sú spravidla vybavené strojovou inteligenciou, takže problémy s nimi súvisiace a problémy UI sa výrazne prekrývajú. Napriek tomu však autonómne zariadenia tvoria samostatnú kategóriu IKT, odlišnú od UI, pretože ide o zariadenia takpovediac „vtelené.“ Roboty sa vďaka svojim fyzickým telám dokážu pohybovať vo fyzickom prostredí a manipulovať ním.

Medzi etické problémy týkajúce sa inteligentných robotov patrí nepredvídateľnosť ich konania a nevysvetliteľnosť ich rozhodnutí. Okrem toho narážame na ťažkosti pri prisudzovaní zodpovednosti za ich zlyhania, na problém miery autonómie, ktorú by mali mať, na otázku ich morálneho statusu a na problém urgentnosti, s ktorou nás nútia zaujať stanovisko k páľčivým morálnym dilemám. Začneme problémom predvídateľnosti.

Problém s predvídateľnosťou vzniká preto, lebo roboty dnes môžu byť obdarené UI a dokážu sa učiť. Staršia trieda robotov a rôznych IKT je deterministická, resp. automatická v tom zmysle, že disponuje stabilným algoritmom, ktorý vytvorili ľudia a ktorý určuje ich možné správanie. Vďaka tomu je správanie deterministického robota predvídateľné; takýto robot nemôže urobiť niečo, čo nevyplýva z jeho algoritmu. Moderný kognitívny, resp. autonómny robot je odlišný tým, že na základe skúseností (nových dát) dokáže sám upravovať svoj algoritmus, takže mení svoje možné správanie.

¹¹ Baum, S. D., A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy. *Global Catastrophic Risk Institute Working Paper 17-1*, 2017. DOI: <<http://dx.doi.org/10.2139/ssrn.3070741>>.

¹² Referencia na titul Barrat, J., *Our Final Invention: Artificial Intelligence and the End of Human Era*. New York, Thomas Dunne Books 2013.

Preto sú kognitívne roboty ťažko predvídateľné. Navyše, ak je taký robot uplatnený v situáciách, v ktorých by mohol spôsobiť ujmu ľuďom, je bezpečnostné riziko zrejmé.

Autonómne stroje sú problematické aj preto, lebo príčiny ich správania nemusia byť *vysvetliteľné*.¹³ U strojov, ktoré menia svoj algoritmus v dôsledku nových skúseností, nemusí byť možné zistiť, prečo sa rozhodli tak, ako sa rozhodli, podobne ako v prípade už spomínaného systému COMPAS. Rozhodovací mechanizmus môže byť navrhnutý takým spôsobom, alebo byť tak komplikovaný, že nemusí byť v ľudských silách vysvetliť robotovo správanie. V situáciách, kedy je potrebné zistiť príčiny robotovho konania, napríklad pri súdnom spore alebo etickej debate, nastáva problém vysvetliť rozhodnutie stroja. Preto jednou z aktivít, ktorá prebieha v rámci etického vývoja UI, resp. strojov s UI, je snaha dosiahnuť vysvetliteľnú UI (XAI - Explainable Artificial Intelligence).¹⁴

Ďalší problém sa týka autonómie stroja a možno k nemu prikročiť z rôznych strán. Môžeme sa spýtať, koľko autonómie môže robot mať a ktoré rozhodnutia už prijať nemôže. Napríklad, mal by robot s umelou inteligenciou, ktorý je schopný komunikovať, vždy hovoriť pravdu? Alebo sa môžeme pýtať, kedy a či vôbec môže mať rozhodnutie robota autoritu nad rozhodnutím človeka, napríklad, či sa robot môže za istých okolností brániť svojmu vypnutiu. Ďalej, v prípade, že rozhodnutie robota a človeka príde do konfliktu, kto má prednosť? Napríklad, mala by mať diagnóza lekárskej UI autoritu nad diagnózou ľudského lekára? Sú autonómne zbrane alebo vojenské roboty morálne dovolené?

Možnosti, ktoré nám roboty s UI prinášajú, môžu byť síce znepokojivé, no existuje pádny argument v prospech ich používania. Spočíva v tom, že roboty budú spoľahlivejšie a bezpečnejšie ako ľudskí agenti. Zdá sa, že je len otázkou času, kedy roboty budú schopné robiť všetko, čo ľudia, a ešte lepšie. V tom prípade môže byť stále ťažšie vysvetliť, prečo sme istú ľudskú činnosť alebo rozhodovanie neprenehali oveľa spoľahlivejšiemu, presnejšiemu a rýchlejšiemu stroju. Na druhej strane, ak budú stroje schopné rozhodovať sa *inteligentnejšie* ako ľudia, môže to znamenať, že nebude v ľudských silách porozumieť im a vyhodnotiť ich správanie. Zdá sa teda, že nás čakajú kompromisy medzi ľudským dohľadom na jednej strane a spoľahlivosťou na strane druhej.

Problém určenia zodpovednosti za zlyhanie robotov je bezprostredný. Keď samoriadiace auto alebo chirurgický stroj urobia chybu, kto by mal byť zodpovedný? Ten, kto autonómne zariadenia navrhol a vytvoril, alebo ten, kto ho vlastní a spustí?¹⁵ A ak je to ten, kto ho navrhol a vytvoril, ktorí z množstva ľudí, čo sa na existencii daného stroja podieľali - dizajnéri, softvéroví a technickí inžinieri a investori - budú zodpovední,

¹³ Pozri World Commission on Ethics of Scientific Knowledge and Technology, *Report of COMEST on Robotics Ethics*. Paris, UNESCO 2017, s. 36. Dostupné na: <http://unesdoc.unesco.org/images/0025/002539/253952E.pdf>

¹⁴ Dokonca aj v prípade už spomínaných deep learning algoritmov sa nakoniec podarilo dokázať, že interpretácia ich rozhodnutí, resp. predikcií, je možná. Jednou z takýchto metód je napr. LIME (Local Interpretable Model-Agnostic Explanations). Pozri Ribeiro M.T. - Singh S. - Guestrin C., "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, s. 1135-1144. DOI: <10.1145/2939672.2939778>.

¹⁵ Pokiaľ ide o samoriadiace autá, v Spojených štátoch je za zlyhanie automobilu právne zodpovedný vodič, v Nemecku automobilka.

keď na stroji pracovali v tíme s nejasnou hierarchiou a neurčitým delením práce, kvôli čomu je ťažké identifikovať, kto prispel čím k výslednému produktu?¹⁶

Ďalší z problémov autonómnych zariadení nadväzuje na problém prisúdenia zodpovednosti, no zachádza do zásadnejších otázok etiky a filozofie mysle. Dnešné roboty a hlasoví asistenti dokážu totiž plynule a prirodzene komunikovať s ľuďmi a mimikou tváre napodobňovať emócie. Ich rastúce schopnosti a ťažkosť, s akou rozlišujeme ich prejavy od prejavov ľudských, vedú k otázkam o ich morálnom a právnom statuse.¹⁷ Podobne, ako sa napríklad bioetika zamýšľa nad tým, aký morálny a právny status priznať zvieratám, nenarodeným deťom a vážne mentálne postihnutým, tak etika IKT môže skúmať, či je možné, aby mal človek voči stroju zodpovednosť a či naopak môžu byť stroje zodpovedné voči človeku.¹⁸ Napríklad, mohli by sme sa spýtať, či sex s robotickým milencom konštituuje partnerskú neveru, prípadne, či je možné, aby došlo k sexuálnemu zneužitiu robota.

Ako sme spomenuli, táto sada otázok je filozoficky širšia ako problém prisúdenia zodpovednosti, pretože nás vedie k úvahám o tom, na základe akých daností je možné subjektu pripísať morálnu zodpovednosť a práva. Ak sú tými danosťami život a vedomie, mali by sme sa spýtať, či systémy s UI môžu byť živé a vedomé. To nás vedie k otázkam typu „Čo je život?“ a „Čo je vedomie?“

Niektorým sa môže táto sada otázok zdať nezaujímavá, pretože sa domnievajú, že stroje nemôžu mať morálne a právne vyšší status, ako majú teraz, nehovoriac o dosiahnutí morálneho a právneho statusu ľudí. No je dôležité si pripomenúť, že vývoj IKT prinesie strojové systémy, ktoré budú stále komplexnejšie a schopnejšie. Okrem toho by sa dalo argumentovať, že život a vedomie nemusia byť za hranicou technologických možností. Vieme totiž, že vysoko organizovaná hmota môže byť živá a vedomá: naše telá a mozgy sú toho príkladom. Máme potom dôvod myslieť si, že podobne organizovaná, ale umelo vytvorená hmota, nemôže byť taktiež živá a vedomá? Je síce ťažké predstaviť si, že by sme mali k dispozícii test, ktorý by dokazoval, že robot má vedomie, no podobný test nemáme k dispozícii ani u ľudí.

Inteligentné roboty budúcnosti, ktoré vyzerajú, rozprávajú a konajú ako ľudia, môžu byť nakoniec eticky a filozoficky najcitlivejšou IKT, pretože by mohli začať spochybňovať vnímanie našej vlastnej (ľudskej) identity. S narastajúcou podobnosťou medzi robotmi a ľuďmi by pre nás mohlo byť stále ťažšie nájsť filozoficky významný rozdiel medzi robotom a človekom. Ak by sme v istom bode nevedeli rozdiel nájsť, výsledok by mohol byť dvojaký. Na jednej strane by podobnosť ľudí a robotov mohla spochybniť obraz, ktorý máme sami o sebe ako živé a vedomé bytosti a konfrontovala by nás s možnosťou, že sme biologické stroje s odlišným softvérom a hardvérom. Na strane druhej, ak by nám bolo ťažko vnímať samých seba ako biologické stroje, mohlo by nás to priviesť k vnímaniu strojov ako bytostí, ktoré sú ľuďom rovné. Ak by sme

¹⁶ Ide o tzv. *many hands* problém, ktorý sa týka určenia zodpovednosti vo veľkých vývojových tímoch a nie je vlastný len IKT. Pozri Doorn, N., A Rawlsian Approach to Distribute Responsibilities in Networks. *Sci Eng Ethics*, 16, 2010, s. 221-249. DOI: <10.1007/s11948-009-9155-0>.

¹⁷ Na svete už existuje robot, ktorému bolo priznané štátne občianstvo. Minulý rok ho udelila Saudská Arábia robotovi Sophii. Pozri Galeon, D., For the First Time Ever, a Robot Was Granted Citizenship, *futurism.com*, 26. október 2017. Dostupné na: <<https://futurism.com/for-the-first-time-ever-a-robot-was-granted-citizenship/>>.

¹⁸ Pre diskusiu o statuse UI strojov pozri Gunkel, D., *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. Cambridge – London, MIT Press 2012.

tiahli k tejto druhej možnosti, vývoj robotov budúcnosti by znamenal vytvorenie nového druhu života.¹⁹ Tak či onak, zdá sa, že obe možnosti hlboko ovplyvnia náš svetonázor.

A nakoniec, roboty sú pre etikov problematické nielen preto, lebo vytvárajú nové morálne problémy, ale aj preto, lebo vytvárajú problém pre etiku ako disciplínu. Niektoré IKT nás nútia urgentne a jednoznačne zaujať postoj k už známym, ale nevyriešeným problémom. Napríklad algoritmus samoriadiach automobilov by mal efektívne zvládať všetky situácie, ktoré by v premávke mohli reálne nastať. No v premávke s miliónmi samoriadiach áut budú nastávať aj málo pravdepodobné variácie známej dilemy električky - napríklad vo forme: nechaj piatich zomrieť a jedného žiť, alebo obetuj jedného a zachráň piatich.²⁰ Výrobcovia autonómnych áut stoja pred ťažkým rozhodnutím, aké inštrukcie autu dať v prípade, že dilematická situácia nastane.²¹ Žiaľ, etici a filozofi nebudú v danej veci veľmi nápomocní, ak všetko, čo tvorcom týchto technológií budú môcť odkázať, bude tradičné „filozofi sa nezhodnú.“ Zásadné morálne rozhodnutia tak nebudú určovať najlepšie etické argumenty, ale trh.²²

Práve profesionáli a špecialisti v odbore IKT patria medzi prvých ľudí konfrontovaných s etickou dimenziou IKT, nakoľko majú najlepšie pochopenie fungovania technológie, ktorú pomáhajú vyvíjať. Keďže pri práci s IKT vyvstávajú hodnotové otázky a neraz neexistujú normy, ako ich správne zodpovedať, uvádza to profesionálov do nepríjemnej pozície, najmä vtedy, ak tieto otázky spochybňujú legitimitu ich práce samej. Priemerný IKT špecialista však nedisponuje etickým vzdelaním, keďže neabsolvoval ani len prehľadový kurz z etiky. Možno aj preto sú prípady, kedy sa jeden z tvorcov novej technológie rozhodne jasne a pevne vymedziť voči jej neetickému charakteru, zatiaľ len výnimočne.²³

Vzhľadom na to, ako výrazne IKT vplývajú a budú vplývať na spoločnosť, je zarážajúce, že štúdium a pôsobenie na poli IKT nekladú porovnateľný dôraz na etické vzdelanie, ako je tomu v medicíne. Etické vzdelanie budúcich lekárov je dnes samozrejmosťou a môže čerpať z tradície Hippokratovej prísahy. Keďže práca IKT špecialistov bude mať nesmierny dopad na spoločnosť, myslíme si, že aj oni potrebujú obdobné etické vzdelanie a étos, vlastnú Hippokratovu prísahu, ktorá by im dala schopnosť navigovať v morálne hmlistých vodách ich profesie.

Ako sme spomenuli v úvode, v súvislosti s IKT by sme mohli spomenúť mnohé ďalšie morálne problémy, najmä v súvislosti používaním internetu, sociálnych médií, ako aj otázky týkajúce sa širokého spoločenského dopadu IKT. Medzi tieto posledné môžeme zaradiť zmeny vo vzorcoch medziľudskej komunikácie, hromadné sledovanie a

¹⁹ Napr. pozri úvahy o umelom živote u Tegmark M., *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York, Penguin Books Ltd. 2018.

²⁰ Pozri Thompson, J. J., Killing, Letting Die, and the Trolley Problem. *The Monist*, 59, 1976, s. 204-17; Foot, P., The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, 5, 1967, s. 5-15.

²¹ Množstvo nepohodlných verzií dilemy električky si čitateľ môže vyskúšať posúdiť v projekte MIT Moral Machine. Dostupné na: <<http://moralmachine.mit.edu/>>.

²² Zaujímavé je sledovať už len rozpory medzi deklarovanými a reálnymi postojmi. Štúdie experimentálnej etiky napríklad naznačujú, že ľudia síce verbálne preferujú utilitaristické riešenie dilemy električky, no nie konzistentne. Respondenti považovali autá, ktoré obetujú svojho pasažiera z utilitaristických dôvodov, za morálne, no nezakúpili by si ich. Pozri Bonnefon, J. et al., The Social Dilemma of Autonomous Vehicles. *Science*, 352 (6293), s. 1573-1576. DOI: <10.1126/science.aaf2654>.

²³ Jedným z veľkých medializovaných prípadov z nedávnej minulosti je prípad Christophera Wylieho, ktorý sa rozhodol vyzradiť aktivity spoločnosti Cambridge Analytica. V podobne lútostivom duchu sa vyjadrili aj spoluzakladateľ Facebooku Sean Parker a bývalý riaditeľ Chamath Palihapitya, keď priznali, že s pomocou strojového učenia zámerne navrhli Facebook, aby vyvolával závislosť svojich užívateľov.

analyzovanie obyvateľstva, delegovanie ľudských rozhodnutí a schopností strojom, nový spôsob vedenia vojny, či dopad automatizácie na nezamestnanosť.²⁴

Jeden zo spôsobov, ako zmierniť dramatické zmeny a nepredvídateľné škody spôsobené používaním IKT, je spomaliť ich zavedenie na trh prísnejšími reguláciami, napríklad v podobe rozšírených bezpečnostných alebo akceptačných testov. Nová IKT by sa tak najprv implementovala v rámci malej vzorky užívateľov, resp. v malom meradle, a jej dôsledky by sa vyhodnocovali. V rámci farmaceutického priemyslu je prísna regulácia bežnou praxou. Podmienkou uvedenia nového lieku na trh je rozsiahly kontrolný proces, v ktorom je štandardom kvality placebo kontrolujúca klinická štúdia. Je samozrejme otázne, nakoľko prísnu reguláciu vyžadujú nové IKT, no je potrebné zvážiť, či aj tento priemysel nepotrebuje vlastný proces overovania bezpečnosti, resp. etickosti nových technológií, spolu s vlastným štandardom kvality.

S istotou nevieme, aké problémy prinesie očakávaná proliferácia IKT v budúcnosti, no vzhľadom na zmeny, ktoré IKT priniesli doteraz, môžeme očakávať, že medzi nimi budú aj zmeny zásadné. Keďže technológie rozširujú našu moc, je na mieste obava, či rýchly a ťažko predvídateľný vývoj IKT, najmä tých emergentných, nepovedie k rovnako rýchlym a málo predvídateľným zmenám mocenských pomerov v spoločnosti.

Pochopiteľne, tvrdenie ako toto sa pohybuje v rovine špekulácie, no aj napriek tomu považujeme za potrebné ho predložiť. Ak máme byť totiž pripravení na široké spoločenské zmeny, ktoré s veľkou pravdepodobnosťou nastanú, budúcnosť sa musíme pokúšať odhadovať a vopred sa na ňu normatívne pripravovať. Niektoré IKT môžu byť morálne problematické alebo škodlivé už na základe svojho dizajnu, teda ešte predtým, ako sa začnú používať. Taktiež môžu byť morálne problematické a škodlivé skryto alebo napriek všetkým očakávaniam. Pokúšať sa regulovať ich a podriadiť etickej reflexii až potom, čo škody napáchali, by bolo neskoro.

Takýto proaktívny prístup nemá len etický, ale aj praktický zmysel. Ak vývojár dostatočne neošetrí etické riziká svojej novej technológie, časom na to doplatí aj finančne. Oplatí sa preto zohľadňovať etické riziká novej IKT už vo fáze vývoja, ako hasiť etické „požiare“ potom.²⁵

Ako vhodná ilustrácia takéhoto proaktívneho postupu môže slúžiť návrh Rathenau inštitútu, ktorý vo svojej správe o *UI, robotike a virtuálnej a rozšírenej realite* navrhuje zvážiť dve nové ľudské práva. Autori správy v prvom rade predpokladajú, že tendencia vlád a spoločností sledovať obyvateľstvo a z týchto informácií následne čerpať, sa bude prehlbovať. Preto navrhujú uvažovať o ľudskom práve *nebyť sledovaný, analyzovaný a koučovaný*. Po druhé predpokladajú, že v silne automatizovanej a robotizovanej spoločnosti budúcnosti by v niektorých oblastiach života mohol celkom vymiznúť ľudský kontakt. Navrhujú preto zvážiť ľudské právo *na zmysluplný ľudský kontakt*.²⁶

²⁴ Ani tento problém nie je celkom nový. Anticipoval ho už v roku 1950 Wiener, ktorý sa obával, že hromadná nezamestnanosť, zapríčinená zautomatizovanou továrňou, povedie k nárastu vplyvu magnátov a korporácií. Wiener, N., *The Human Use of Human Beings: Cybernetics and Society*. London, Free Association Books 1989, kap. IX.

²⁵ Reálny dopad proaktívnosti etiky nových a emergentných technológií už máme možnosť sledovať pri biotechnológií. Spomeňme napríklad širokú diskusiu, ktorá viedla k zákazu používania niektorých metód génového inžinierstva na živých ľudských plodoch (CRISPR-Cas9) a k rýchlemu všeobecnému odsúdeniu, ktoré vyvolalo nedávne údajné porušenie tohto zákazu.

²⁶ van Est, R. et al., *Human Rights in the Robot Age: Challenges Arising from the Use of Robotics, Artificial Intelligence, and Virtual and Augmented reality*. Hague, Rathenau Instituut 2017.

Vymedzenie etiky IKT

Na základe týchto úvah považujeme za dôležité upozorniť aj v našej proveniencii na potrebu širšej etickej diskusie o IKT a ich právnej regulácie. V súvislosti s touto potrebou považujeme za dôležité zaoberať sa *štúdiom morálnych problémov informačných a komunikačných technológií*. Nasledujúc niektorých autorov, toto štúdium nazývame „etika informačných a komunikačných technológií“, niekedy nazývané aj „digitálna etika“, a chápeme ho ako samostatnú oblasť aplikovanej etiky. Úžitok etiky IKT vidíme predovšetkým v navrhovaní morálne prijateľných noriem vývoja a používania IKT, ako aj v navrhovaní etických zásad pre profesie pôsobiace v oblasti IKT.

Na tomto mieste je potrebné ozrejiť, čo presne rozumieme pod informačnými a komunikačnými technológiami a čo presne spadá pod štúdium etiky IKT. Pod informačné a komunikačné technológie spadajú technológie na hľadanie, analyzovanie, vytváranie, komunikáciu a využívanie digitálnych informácií. Digitálne informácie chápeme ako dáta vo forme numerického kódu, zväčša binárneho. Pod IKT teda budú spadať *de facto* všetky technológie, ktoré sú kódovateľné alebo kód môžu využívať. Ide o ide softvér aj hardvér, teda o počítače, tablety, chytré telefóny, VR okuliare, programy, aplikácie, internet, UI, cloud, rozšírenú realitu, big data, strojové učenie a mnohé ďalšie. Ako sme spomínali v úvode, najviac nás zaujímajú práve emergentné IKT, ktorých rozšírenie je exponenciálne a dopad ťažko predikovateľný.

Keď hovoríme o potrebe eticky skúmať IKT, je treba dodať, že etické skúmanie IKT nie je žiadnou novinkou a prebieha už niekoľko desaťročí.²⁷ Napriek tomu však dodnes neexistuje jednotné vymedzenie predmetu tejto disciplíny ani ustálená terminológia. Preto by sme radi predostreli jeho presnejšie vymedzenie, resp. explikovali ho. Zastávame totiž názor, že etika IKT, ktorú navrhujeme ďalej rozvíjať, musí byť *proaktívna*. Myslíme to tak, že nestačí, ak sa etika IKT zaoberá iba takými morálnymi problémami IKT, ktoré sa týkajú technológií len na úrovni ich používania, sú všeobecne známe a aktuálne. Etika IKT sa musí zaoberať aj takými problémami, ktoré vyvstávajú už vo fáze vývoja technológie, problémami, ktoré sú na prvý pohľad skryté, ako aj problémami, ktoré nás ešte len zastihnú v budúcnosti.

V prvom rade, je potrebné skúmať a posudzovať IKT ešte predtým, ako vzniknú a začnú sa používať, pretože je možné, že niektoré IKT by nemali byť vyvinuté, alebo majú problematický dizajn, alebo preto, lebo sú autonómne. Pokiaľ ide o technológie, u ktorých je otázne, či by vôbec mali byť vyvinuté, spomínali sme silnú UI či autonómne zbrane. Pokiaľ ide o technológie, ktoré majú problematický dizajn, môžeme taktiež spomenúť UI a silnú UI, konkrétne problém nastavenia jej hodnôt, cieľov a spôsobov, akými bude naplňovať líšiace sa hodnoty ľudských indivíduí. Problematický dizajn však nemusí mať len UI. Dizajn ľubovoľnej IKT môže byť taký, že systematicky napomáha alebo škodí realizácii určitej hodnoty.²⁸ Napríklad, spyware systematicky škodí súkromiu, pretože na to bol navrhnutý. A podobne algoritmus, ktorý určuje poradie prijímateľov orgánov, no arbitrárne uprednostňuje belochove, škodí spravodlivosti. A

²⁷ Môžeme povedať, že začala v roku 1950, prvou publikáciou knihy N. Wienera: *The Human Use of Human Beings: Cybernetics and Society*, c.d.

²⁸ Ide o samostatnú tému, ktorá sa dotýka diskusie o hodnotovej zaťaženosti technológie a do ktorej sa nebudeme púšťať. V literatúre sa tento postoj nazýva aj internalizmus a pre hodnoty, ktoré sú vtlačené technológii už pri jej vývoji, sa používa termín *embedded values*. Pozri Friedman, B. - Kahn P. H., *Human Values, Ethics, and Design*. In Sears, A. - Jacko, J. A. (eds.), *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2. edícia). New York, Lawrence Erlbaum Associates 2008, s. 1241-1266.

nakoniec, dizajn autonómnych zariadení môže byť problematický, pretože robot môže konať bez toho, aby bol „použitý“ človekom (odhliadnuc od možnosti, že ho človek musí uviesť do chodu).²⁹

V druhom rade, ako upozorňujú niektorí autori, nestačí, ak sa etické skúmanie IKT zaoberá len takými morálnymi problémami IKT, ktoré sú známe a verejne rozoznané. Musí aj aktívne vyhľadávať problémy skryté alebo doteraz nerozpoznané.³⁰ Niektoré morálne problémy IKT sú skryté, pretože okrem špecialistov nikto iný nemusí vedieť, čo presne používanie danej technológie zahŕňa. Je tomu tak napríklad pri výsledkoch vyhľadávania internetového vyhľadávača, či pri zobrazovaní poradia nových správ na sociálnej sieti. Algoritmus týchto procesov môže obsahovať predsudky alebo užívateľa informačne izolovať. U iných IKT zasa môže byť verejne známe, ako fungujú, no nemusia byť dostatočne rozoznané etické dôsledky tohto fungovania, ako tomu môže byť pri sledovaní aktivity internetového prehliadača pomocou tzv. cookies.

V treťom rade, etika IKT musí anticipovať budúce morálne problémy IKT, ktoré môžu vzniknúť v dôsledku použitia technológií, ktoré ešte neexistujú. Naše priblíženie problematiky silnej UI, problémov týkajúcich sa širších spoločenských dopadov IKT, ako aj postup Rathenau inštitútu pri navrhovaní nových ľudských práv, sú príkladmi na takéto problémy. Anticipáciu budúcich problémov považujeme za dôležitú aj preto, lebo niektoré IKT môžu napáchať škody nečakane (kybernetické sledovanie) a keďže niektoré IKT môžu byť problematické už na základe toho, že existujú (autonómne zbrane, silná UI).

Na základe takto chápanej proaktívnosti by sme radi explikovali definíciu etiky informačných a komunikačných technológií s dôrazom na *proaktívne* štúdium morálnych problémov súvisiacich s IKT. Podľa nás ide o *štúdium zjavných a skrytých, súčasných i budúcich morálnych problémov súvisiacich s dizajnom, vývojom a používaním IKT*.

Tak, ako nie je ustálená definícia etického skúmania IKT, nie je ustálená ani používaná terminológia. Toto platí aj pre samotné pomenovanie disciplíny. Historicky je zaužívaný termín „počítačová etika“, pričom pod ňou jej zakladatelia chápali „analýzu povahy a sociálneho dopadu počítačovej technológie a zodpovedajúca formulácia a zdôvodnenie politiky pre ich etické použitie.“³¹ Naše vymedzenie etiky IKT sa od tejto definície líši jednak tým, že jej predmetom ani zd'aleka nie sú len počítače a tiež tým, že akcentuje všetky tri zložky proaktívneho prístupu k morálnym problémom IKT. Odhliadnuc od toho sa počítačová etika obsahovo prekrýva s etikou IKT. Je preto možné etiku IKT chápať ako spresnenie či modernizáciu počítačovej etiky.

Niektorí zo súčasných autorov používajú pre skúmanie etických problémov IKT termíny ako „informačná a počítačová etika“, „kyberetika“ alebo rovnako ako my „etika informačných a komunikačných technológií.“³² Zatiaľ čo sa spomenuté prístupy viac

²⁹ Keďže roboty nie sú ľuďmi používané, ale konajú sami, morálne hodnotiť môžeme len ich uvedenie do prevádzky a ich dizajn. Samozrejme za predpokladu, že robot nemôže byť morálnym agentom.

³⁰ Preberáme tu koncept tzv. odkryvacej etiky (*disclosive ethics*) od Brey, P., *Values in Technology and Disclosive Computer Ethics*. In Floridi, L. (eds.), *The Cambridge Handbook of Information and Computer Ethics*. Cambridge, Cambridge University Press 2010, s. 41-58.

³¹ Moor, J., What is Computer Ethics? *Metaphilosophy*, 16 (4), 1985, s. 266-275.

³² Informačnej a komunikačnej etike sú venované príručky Himma, K. E. - Tavani, H. T. (eds.), *The Handbook of Information and Computer Ethics*. Wiley 2008 a Floridi, L., *The Cambridge Handbook of Information and Computer Ethics*, c.d., kyberetike sa venuje Tavani, H. T., *Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing* (5. edícia). Wiley 2015, s. 2, digitálnej etike zasvätili svoju publikáciu editori Davisson, A. - Booth, P. (eds.), *Controversies in Digital Ethics*. New York,

orientujú na používanie IKT, existujú aj také prístupy, ktoré sa primárne orientujú na etické otázky súvisiace s vývojom IKT (ale aj iných technológií). Medzi ne môžeme zaradiť hodnotovo citlivý dizajn (Value Sensitive Design), anticipatívnu etiku technológií (Anticipatory Technology Ethics), školu zodpovednej inovácie (Responsible Innovation), ako aj projekt eticky usporiadanej dizajnu (Ethically Aligned Design). Mimo toho je potrebné odlíšiť etiku IKT od Floridiho „informačnej etiky“, ktorej teoretické ambície podľa nášho názoru presahujú hranice etiky, ako aj od informačnej etiky ako profesijnej etiky knižničných a informačných vied, či od etiky digitálnych médií, ktorú možno chápať hybrid medzi etikou IKT a etikou médií.³³

Tol'ko k uvedeniu problému etiky IKT a k jej vymedzeniu. Z toho, čo bolo doteraz uvedené je zjavné, že pracujeme s predpokladom, podľa ktorého má etika IKT ako samostatná disciplína, koncipovaná tak, ako sme tu navrhli, svoju dôležitosť a opodstatnenie. V nasledujúcej časti sa toto opodstatnenie pokúsime preukázať.

Etika IKT ako aplikovaná etika

Rovnako ako mnohí iní autori si myslíme, že etika IKT je dištinktívnou a rozlíšiteľnou oblasťou aplikovanej etiky s právom na samostatnú existenciu.³⁴ Myslíme si však, že dané tvrdenie nie je triviálne a nemalo by byť prijaté nekriticky. Ak by etika IKT nebola samostatnou oblasťou aplikovanej etiky, ale množinou morálnych problémov, ktoré len náhodne súvisia s IKT, bolo by zbytočné hovoriť o etike IKT. Stojí teda pred nami úloha ukázať, že existujú dostatočné dôvody pre existenciu etiky IKT.

Otázka, či je etika IKT opodstatnená alebo nie, nie je praktická, ale filozofická. Existujú autori a celé pracoviská, ktoré už predpokladajú, že etika IKT alebo obdobne koncipovaná disciplína má svoje opodstatnenie, považujú ju za samozrejmosť, a preto by nepovažovali za potrebné ju zdôvodňovať. No tvrdenie, že etika IKT je opodstatnená, nie je evidentné a ako aj dokladajú niektorí iní autori, nie je ani celkom nekontroverzné.³⁵ Dalo by sa argumentovať, že etika IKT nie je opodstatnená, pretože IKT nevyvolávajú žiadne nové morálne problémy. Analogický argument bol použitý pred niekoľkými rokmi voči nanoetike. Podľa niektorých sú potom zdanlivo nové problémy súvisiace s nanotechnológiou len „starým etickým vínom v nových technologických fľašiach.“³⁶ Neplatí teda to isté aj pre etiku IKT? Považujeme preto za filozoficky zaujímavé poskytnúť argument, ktorý by jasne vysvetlil, prečo je etika IKT opodstatnená.

Predpokladajme, že postačujúcou podmienkou opodstatnenosti určitej odbornej oblasti je existencia nových problémov, ktoré s danou oblasťou súvisia. Zvážme preto, či sú nasledujúce morálne problémy, ktoré na prvý pohľad súvisia s IKT, nové:

(1) Je dovolené vyvíjať **silnú UI**?

Bloomsbury 2016, s.1. A nakoniec, o etike IKT píše Fabris, A., *Ethics of Information and Communication Technologies*. Springer 2018., s. 16.

³³ Floridi, L., Information Ethics, its Nature and Scope. *SIGCAS Computers and Society*, 36 (3), 2006, s. 21-36.

³⁴ Moor, J., What Is Computer Ethics?, c.d.; Maner, W., Unique Ethical Problems in Information Technology. *Science and Engineering Ethics*, 2 (2), 1996, s. 137-154.

³⁵ Tavani, H. T., *Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing*, c.d., s. 7-11.

³⁶ McGinn, R. E., What's Different, Ethically, about Nanotechnology?: Foundational Questions and Answers. *Nanoethics*, 4, 2010, s. 120.

- (2) Aké hodnoty a ciele by sme mali vštiepiť **silnej UI**?
- (3) Kedy môže mať rozhodnutie **robotu** autoritu nad rozhodnutím človeka? Napríklad, môžu roboty klamať? Môžu roboty za istých okolností použiť donucovanie alebo letálnu silu voči ľuďom?
- (4) Aký je morálny status komplexných **robotov** a **umelých inteligencií**?
- (5) Je dovolené zbierať informácie o návštevníkoch webovej stránky vo forme **cookies** alebo zbierať informácie o návštevníkoch obchodu pomocou **in-store analytík** bez ich vedomia a informovaného súhlasu?
- (6) Je možné sa voči niekomu morálne previniť vo **virtuálnej realite**, napríklad znásilnením jeho avatara? Je dovolené mučiť virtuálnu osobu?

Všetky uvedené problémy môžeme považovať za nové, no len v slabšom zmysle slova „nové.“ Pretože pred koncipovaním UI, robotov, virtuálnej reality, internetu, atď., sa zrejme nikto nezamýšľal nad morálnosťou ich používania. Nemohol, lebo ešte neexistovali. Debaty o etike informačných a komunikačných technológií sa pochopiteľne rozprúdili až po vynájdení IKT a ako poukazujú niektorí autori, získavajú na intenzite v závislosti od toho, ako rozvinuté a rozšírené tieto technológie sú.³⁷ V tomto zmysle sú (1)-(6) nové, no len triviálne.

Existencia nových problémov v slabšom zmysle slova zrejme nie je postačujúcou podmienkou samostatnosti etiky IKT. Povedzme, že by sme sa spýtali: „Je morálne zakázané fyzicky napadnúť človeka počítačovou klávesnicou?“ Zrejme by sme povedali, že nie je dôležité, že človek bol napadnutý *klávesnicou*, ako skôr to, že bol napadnutý. Podobný „nový“ problém uvádza Moor, keď hovorí o krádeži domáceho počítača.³⁸ Na týchto príkladoch nie je podstatné to, že vykonané akty súviseli s IKT, ale skôr to, že išlo o napadnutie a nedovolené privlastnenie majetku. Túto námietku sa môžeme pokúsiť generalizovať na všetky domnelé problémy etiky IKT a povedať, že síce ide o morálne problémy, no len o také, v ktorých IKT nefigurujú podstatne.

Jeden zo spôsobov, ako určiť, či je daný problém skutočne nový, je spýtať sa, či je možné zaradiť ho do inej, už etablovanej disciplíny. Keďže etiku IKT chápeme ako disciplínu aplikovanej etiky, mali by sme sa spýtať, či problémy (1)-(6) už riešili alebo by ešte len mali riešiť iné disciplíny aplikovanej etiky.

Medzi najlepšie etablované disciplíny aplikovanej etiky patrí bioetika a medicínska etika, environmentálna etika a biznis etika.³⁹ Mali by sme sa preto spýtať, či problémy (1)-(6) spadajú pod jednu z týchto disciplín.

Ak otázku chápeme deskriptívne, tak jednoznačne nespadajú. Bioetika a medicínska etika, environmentálna etika ani biznis etika netematizovali problémy (1)-(6), a ani problémy nim podobné. Môžeme sa však spýtať preskriptívne: *Mali by* problémy (1)-(6) a podobné pod tieto disciplíny spadať? Odpoveď závisí od toho, aké typy problémov by mali spadať pod spomenuté aplikované etiky. Pre ilustráciu, medzi vzorové problémy bioetiky a medicínskej etiky patria otázky povolenosti eutanázie, potratov, klonovania, experimentov na ľuďoch a zvieratách, či genetického vylepšovania. Environmentálna etika typicky rieši otázky morálnej zodpovednosti voči životnému prostrediu, budúcim generáciám a zachovaniu biodiverzity. A nakoniec,

³⁷ Moor, J., Why We Need Better Ethics for Emerging Technologies. *Ethics and Information Technology*, 7 (3), 2005, s. 111-119.

³⁸ Moor, J., What Is Computer Ethics?, c.d., s. 267.

³⁹ Mohli by sme do zoznamu pridať aj ďalšie disciplíny, obzvlášť etiku technológií. Etiku IKT však chápeme ako sub-disciplínu etiky technológií.

biznis etika sa zaoberá otázkami, či korporácie môžu byť morálnymi činiteľmi, či existuje morálne zakázaná predajná cena, čo je to zavádzajúca reklama a ďalšími.

Domnievame sa, že táto ilustrácia ukazuje, že problémy (1)-(6) nie sú typické pre spomenuté aplikované etiky, a preto by pod ne nemali patriť. Zdá sa nám, že problémy súvisiace s vývojom a použitím UI, robotov, internetu, digitálnych médií a podobné nepatria pod tieto disciplíny, lebo pri problémoch (1)-(6) nejde špecificky o začiatok, podobu a koniec ľudského života, ani o životné prostredie, obchod a podobné témy.

Ak je faktom, že (1)-(6) nespádajú a ani by nemali spadať pod etablované aplikované etiky, hovorilo by to silne v prospech opodstatnenosti inej aplikovanej etiky. No jednako, nemuselo by to hovoriť konkrétne v prospech opodstatnenosti etiky IKT, ale v prospech opodstatnenosti napríklad všeobecnej etiky technológie, alebo etiky vedy a výskumu. Je možné, že napriek zdaniu tieto problémy nespája vývoj a používanie IKT, ale niečo iné, napríklad používanie technológie vo všeobecnosti alebo vedecký výskum, a že v nich vývoj a používanie IKT figurujú len náhodne. Bolo by preto uspokojivejšie nájsť inú postačujúcu podmienku opodstatnenosti etiky IKT.

Jedna taká postačujúca podmienka sa nám priamo núka: etika IKT existuje, len ak existujú morálne problémy, ktoré *podstatne* súvisia s IKT. Dozaista, fráza „problémy, ktoré podstatne súvisia s IKT“ je do istej miery vágna a vyžaduje spresnenie. Budeme pod ňou myslieť to, že daný morálny problém *by nebol existoval, nebyť danej IKT*.⁴⁰ Morálne problémy, ktoré spĺňajú túto požiadavku, budeme nazývať *unikátne* a ak ich identifikujeme, budeme etiku IKT považovať za opodstatnenú.

Uvedená podmienka je však problematická, pretože nie je jasné, za akých okolností je splnená. Inak povedané, nie je jasné, aké sú podmienky pravdivosti tvdenia, že *morálny problém (MP) by neexistoval, nebyť existencie istej IKT*, pretože nemáme jasné porozumenie toho, čo určuje identitu MP. Jedno východisko je pokúsiť sa formulovať ekvivalentné formulácie problémov (1)-(6), ktoré by neodkazovali na vývoj alebo použitie IKT, ale na iný akt, ktorý by bol z morálneho hľadiska rovnaký. Ak by sme ekvivalentnú formuláciu našli, problém by zrejme nesúvisel s IKT podstatne.

Znova však, je veľmi ťažké povedať, čo v tomto kontexte znamená „ekvivalentná formulácia.“ Môžeme si však pomôcť už spomenutými príkladmi. Spomenuli sme, že pre MP: „Je zakázané napadnúť človeka klávesnicou?“, nie je z morálneho hľadiska dôležité, že človek bol napadnutý klávesnicou, ale že bol násilne napadnutý. Tento problém preto môžeme považovať za ekvivalentný s problémom: „Je zakázané násilne napadnúť človeka?“ Podobne vieme preformulovať problém: „Je zakázané ukradnúť domáci počítač?“ na „Je zakázané ukradnúť cenný kus súkromného majetku?“ Pod „ekvivalentnou formuláciou“ určitého MP budeme rozumieť takú formuláciu problému, v ktorej sa pýtame na morálny status toho istého aktu, ktorý sa však vykonáva bez pomoci IKT.

Zvážme teda, či problémy (1)-(6) dokážeme naformulovať tak, že morálne problematizujú ten istý akt, no bez toho, že by sa tento akt vykonával s pomocou IKT. Domnievame sa, že problémy (1) až (4) sú najsilnejšími kandidátmi na unikátne morálne problémy IKT, pretože je pri nich najviac zjavné, že k nim ekvivalentné formulácie nevieme nájsť.

Existuje ekvivalentná formulácia pre problém (1) *Je dovolené vyvíjať silnú UI?* Nieкто by mohol povedať, že vývoj silnej UI môžeme pripodobniť k vývoju nukleárnej technológie. S objavom nukleárnej technológie získalo ľudstvo nielen významný zdroj

⁴⁰ Už Maner argumentoval, že počítače generujú morálne problémy, ktoré by nemohli existovať, nebyť ich vynájdenia. Maner, W., Unique Ethical Problems in Information Technology. *Science and Engineering Ethics*, c.d., s. 142.

elektrickej energie, ale prvýkrát v histórii aj schopnosť zničiť samé seba. Objavom silnej UI by ľudstvo mohlo získať ešte výraznejšie výhody, ale zároveň sa opätovne existenčne ohroziť. Mohlo by sa teda zdať, že (1) môžeme preformulovať na (1)* *Je dovolené vyvíjať nukleárnu technológiu?*

No medzi vývojom nukleárnej technológie a vývojom všeobecnej UI existujú rozdiely, ktoré zrejme budú morálne relevantné. V prvom rade, silná UI bude morálne odlišná od nukleárnej technológie kvôli svojej inteligencii, autonómii a možnej uvedomelosti. Okrem toho, nukleárna technológia má v porovnaní so silnou UI veľmi úzke využitie. Domnievame sa, že tieto rozdiely sú dostatočne výrazné, aby boli morálne relevantné, a preto otázku (1)* nepovažujeme za ekvivalentnú s otázkou (1).

Rovnako sa domnievame, že aj druhý problém - (2) *Aké hodnoty a ciele by sme mali vstúpiť silnej UI?*, je unikátny. Zdá sa, že nie je možné nájsť akt, ktorý by bol z morálneho hľadiska podobný nastavovaniu hodnôt a cieľov silnej UI, bez toho, aby sa tento akt týkal silnej UI. Ľudstvo doteraz nikdy nemuselo riešiť podobný problém, pretože žiadna technológia nie je natoľko podobná silnej UI vo svojej inteligencii a autonómii (s výnimkou inej IKT), aby vyžadovala vlastné ciele a hodnotové nastavenie.

Problémy (3) a (4) považujeme za unikátne z rovnakých dôvodov ako problémy (1) a (2). Problém (3) *Kedy môže mať rozhodnutie robota autoritu nad rozhodnutím človeka?*, je unikátny, pretože doteraz neexistovali umelé a autonómne entity, ktorých rozhodnutia by mohli byť v konflikte s ľudskými rozhodnutiami (znova, s výnimkou iných IKT). Problém (4) *Aký je morálny status komplexných robotov a umelých inteligencií?*, považujeme za unikátny, pretože roboty a UI sú novým, dosiaľ neskúmaným druhom entity. Mohlo by sa síce zdať, že problém (4) by sme mohli pripodobniť otázkam o morálnom statuse nenarodených ľudských plodov, ľudí v permanentnom vegetatívnom stave alebo s pokročilou formou demencie. No len o robotoch môžeme povedať, že sú nevyhnutne umelé a (aspoň zatiaľ) neživé a nevedomé. Tieto vlastnosti však zrejme majú morálnu relevanciu, a preto nejde o morálne podobné alebo totožné otázky.⁴¹

Keďže si myslíme, že problémy (1)-(4) možno považovať za unikátne, tak etika IKT splnila postačujúcu podmienku svojej opodstatnenosti. Nemyslíme si však, že problémy etiky IKT sa vyčerpávajú problémami (1)-(4) a nimi podobnými, ale že pod ňu spadajú aj ďalšie druhy problémov, konkrétne (5)-(6), nim podobné a ďalšie, ktoré tu nemáme priestor rozoberať. Preukázať unikátnosť tejto druhej sady problémov však bude ťažšie, ako tomu bolo v prípade sady prvej. Zvážime však aspoň problémy (5)-(6).

Otázky obsiahnuté v probléme (5), *Je dovolené zbierať informácie o návštevníkoch webovej stránky vo forme cookies bez ich vedomia a informovaného súhlasu?* a *Je dovolené zbierať informácie o návštevníkoch pomocou in-store analytík bez ich vedomia a informovaného súhlasu?*, sa zrejme týkajú súkromia v online priestore a vo fyzickom

⁴¹ Ohľadom morálneho statusu agenta môžeme rozlíšiť morálne činiteľstvo, kedy agent môže mať morálnu zodpovednosť za svoje činy, a morálne trpiteľstvo, kedy je možné mať morálnu zodpovednosť voči agentovi. Napríklad, dieťa je morálnym trpiteľom, ale nie je morálnym činiteľom, keďže nemá vyššie kognitívne schopnosti. Nie je však celkom jasné, aké vlastnosti zakladajú činiteľstvo a trpiteľstvo. Pri strojoch môžeme rozlišovať napr. medzi vedomím a inteligenciou/osobnosťou, resp. ako hovorí Bostrom, medzi sentienciou a sapienciou. Stroje dnes už sú inteligentné, ale zrejme nie sú vedomé. Otázne je, aký to má dopad na status ich morálnej agencie. Niektorí autori dokonca požadujú, aby sme sa vyhli vývoju agentov, ktorí by mohli potenciálne získať status morálneho agenta. Pozri Bryson J., *Patiency is Not a Virtue: The Design of Intelligent Systems and Systems of Ethics*. In: *Ethics and Information Technology*, 20 (1), 2018, s. 15-26; Bostrom, N. - Yudkowsky, E., *The Ethics of Artificial Intelligence*. In Frankish, K. - Ramsey, W. (eds.), *The Cambridge Handbook of Artificial Intelligence*. Cambridge, Cambridge University Press 2014, s. 316-334.

svete. Cookies sú malé kúsky dát, ktoré webové stránky používajú na identifikáciu užívateľa. Napríklad, umožňujú používateľovi prihlásiť sa na e-mailový účet bez toho, aby pri každom načítaní účtu musel zadávať meno a heslo. In-store analytiky sú nástroje na báze IoT (internet vecí), ktoré slúžia na komplexné sledovanie a analýzu správania návštevníkov v priestoroch predajne alebo verejnej inštitúcie, spravidla s komerčným cieľom. Keďže otázky v (5) sa zrejme týkajú hodnoty súkromia, mohlo by sa zdať, že ich možno jednoznačne pripodobniť k sledovaniu alebo k inej forme narušenia súkromia. Nie je to však jednoznačné. Zber cookies a sledovanie a vyhodnocovanie správania in-store analytikami môže byť do veľkej miery anonymné a môžu byť v súlade so záujmami návštevníka stránky či obchodu, pretože nazbierané dáta môžu robiť surfovanie po internete alebo nakupovanie praktickejšie a pohodlnejšie.

V tomto bode by sme sa mali spýtať, či možno nájsť také akty, ktoré sú podobné zbieraniu cookies na internete alebo používaniu in-store analytík, no ktoré sa nevyznačujú použitím žiadnej IKT. V prípade in-store analytík sa zdá, že môžeme. Počiatky in-store analytík vyzerali práve tak, že sledovanie a analýza dát o zákazníkoch sa vykonávali ručne, pomocou ľudských pozorovateľov a zapisovateľov. Zdá sa teda, že problém in-store analytík môžeme nasledovne preformulovať: *Je dovolené používať ručné in-store analytiky bez vedomia a informovaného súhlasu sledovaného?* Môžeme si tiež predstaviť také cookies, ktoré majú ne-digitálnu povahu. Povedzme, že prejdaca v potravinách si udržuje zoznam našich nákupných preferencií. Ak vstúpime do obchodu, hneď nás rozozná a pripraví nám náš typický nákup. Vieme si teda predstaviť problém ekvivalentný s problémom (5).

Zdá sa teda, že akty v otázkach obsiahnutých v (5) možno vykonávať aj bez IKT, a že preto tieto otázky v skutočnosti nesúvisia s IKT. Na druhej strane je potrebné dodať, že tieto otázky súkromia, prípadne bezpečnosti, boli problematizované až na základe širšieho využívania IKT – cookies, in-store analytík a podobných technológií. Navyše, tým, že sa zber dát a ich vyhodnocovanie deje digitálne, ide o oveľa efektívnejší a rýchlejší proces, otvorený celkom iným možnostiam, ako proces ručného zberu dát. Napriek tomu uzatvárame, že nemáme dostatočné dôkazy tvrdiť, že by bol problém (5) unikátny.

Problém - (6) *Je možné sa voči niekomu morálne previniť vo virtuálnej realite, napríklad znásilnením jeho avatara?* - sa týka sa primárne technológií ako sú zariadenia rozšírenej (AR) a najmä virtuálnej reality (VR), ktoré sa stávajú pevnou súčasťou herného priemyslu. Na prvý pohľad by sa mohlo zdať, že (6) ani len nepredstavuje pravý morálny problém, pretože situácia v ňom opísaná je analogická tej, kedy hráč v nejakej hre (napr. stolovej alebo kartovej) „zabije“ alebo inak „poškodí“ svojho protihráča. Takéto činy spravidla nemajú morálny význam, keďže je tu silná intuícia, že ublíženie virtuálnej postave vo virtuálnej realite nie je skutočným ublížením.⁴² Zdalo by sa teda, že virtuálna realita nie je morálne plnohodnotná realita a že nie je možné sa v nej voči niekomu morálne previniť.

My sa však domnievame, že problém (6) je hlbší, ako by sa na prvý pohľad mohlo zdať. Simulácie vytvorené AR a VR technológiami môžu byť natoľko imerzívne a hodnoverné, že hranica medzi reálnou a virtuálnou skúsenosťou, ako aj medzi reálnym a virtuálnym ublížením, sa stiera. Ľudia zažívajúci virtuálne simulácie môžu nielen „subjektívne prežívať to, že ich telá sú skutočne vystavené dejom, ktoré sa odohrávajú vo virtuálnom prostredí, ktoré obývajú, ale aj ich mimovoľné biologické reakcie môžu naznačovať, že ich telá zaobchádzajú s týmito virtuálnymi skúsenosťami, ako keby boli

⁴² Na tomto mieste sa nechceme venovať vzťahu hrania počítačových hier a agresívneho správania, ani morálnym otázkam súvisiacim s trollingom či s nevhodným správaním na internete.

skutočné.“⁴³ Je teda zrejme možné ublížiť človeku pomocou virtuálnej reality, a tým sa voči nemu morálne previniť.

Tento fakt otvára priestor pre etickú reflexiu morálne prípustných a neprípustných použití virtuálnej reality. Napríklad, je dovolené simulovať udalosti, ktoré by boli v realite traumatizujúce? Alebo vykonávať psychologické experimenty, ktoré nie je dovolené vykonávať reálne, vo virtuálnej realite (napríklad Milgramov experiment poslušnosti)? Je dovolené simulovať detskú pornografiu alebo mučenie na virtuálnych postavách či na avataroch reálnych ľudí?

Virtuálna realita predstavuje nový typ reality a nemusí byť zřejmé, do akej miery sú činy v nej vykonané morálne relevantné. Keďže je možné sa morálne previniť voči človeku vo virtuálnej realite, má zmysel pýtať sa, ktoré virtuálne činy sú morálne zakázané a ktoré dovolené. Preto sa domnievame, že problém (6) predstavuje unikátny problém.

Virtuálna realita je téma etiky bohatá aj z hľadiska proaktívne orientovanej etiky IKT. Súčasný online hry a kybernetický priestor vo všeobecnosti môžeme považovať za životné prostredie, v ktorom môžu ľudia bohato sociálne interagovať. Okrem toho, že virtuálny priestor je bohatší ako priestor akejkoľvek nedigitálnej hry alebo simulácie typu filmu či prezentácie, na význame mu môže pridať aj množstvo času, ktorý v ňom strávime. Napríklad, povedzme, že zamestnania budúcnosti sa budú vykonávať vo väčšej miere na diaľku, v augmentovanej alebo virtuálnej realite. Môžeme si tiež predstaviť, že virtualizovaná nebude len práca človeka, ale aj jeho voľný čas, keďže vytvorená simulácia bude vysoko imerzívna a hodnoverná, možno ešte viac, ako bežná realita. Ak si predstavíme, že bežný človek budúcnosti bude tráviť väčšiu časť svojho života vo virtuálnej realite ako v tej fyzickej, potom sa dá očakávať, že virtuálna realita nadobudne väčší význam, v niektorých ohľadoch väčší ako význam reality fyzickej. V kontexte týchto možností sa môžeme pýtať, či by človek nemal mať napríklad právo byť informovaný o tom, či sa práve nachádza vo virtuálnej realite alebo právo ju opustiť.

Niet pochyb o tom, že v súvislosti s vývojom a používaním IKT sme konfrontovaní s morálnymi otázkami. Existujú však pochybnosti o tom, či ide o problémy, ktoré s IKT súvisia podstatne. Pri vyhodnocovaní toho, či majú problémy (1) až (6) ekvivalentné formulácie, sme opätovne narážali na prekážku v podobe chýbajúcej teórie identity morálneho problému. Preto, aj keď sme v prípade väčšiny týchto problémov dospeli k záveru, že sú unikátne, uznávame, že naše závery boli do istej miery intuitívne, a preto netransparentné.

Napriek tomu, že problém (5) nepovažujeme za unikátny, myslíme si, že v konečnom dôsledku aj on hovorí istým spôsobom v prospech opodstatnenosti etiky IKT. Podmienka unikátnosti problémov môže byť pre danú aplikovanú etiku postačujúca, ale nemusí byť nevyhnutná. Je tomu tak preto, lebo si myslíme, že novú disciplínu nemusí

⁴³ Ramirez, E. J. - LaBarge, S., Real Moral Problems in the Use of Virtual Reality. *Ethics and Information Technology*, 20 (4), 2018, s. 251, DOI: <<https://doi.org/10.1007/s10676-018-9473-5>>. Morálne kontroverzné sú vďaka imerzivitě a hodnovernosti aj niektoré počítačové hry. V roku 2013 vyvolala hra *Grand Theft Auto 5* pobúrenie kvôli názornej scéne, pri ktorej hráč mučí virtuálnu postavu. Pozri Hern, A., Grand Theft Auto 5 under fire for graphic torture scene, *The Guardian*, 18. september 2013. Dostupné na: <<https://www.theguardian.com/technology/2013/sep/18/grand-theft-auto-5-under-fire-for-graphic-torture-scene>>.

opodstatňovať len existencia preukázateľne unikátnych problémov, ale aj iné faktory. Máme na mysli existenciu potrebnej bázy poznatkov a komunity špecialistov. Pokiaľ ide o potrebnú bázu poznatkov, všimnime si, že pri vyhodnocovaní unikátnosti všetkých problémov sme sa museli opierať aspoň o základné znalosti o fungovaní IKT, pričom na ich dôkladnú analýzu, odkrývanie a anticipovanie budúcich problémov, by sme potrebovali už znalosti pokročilé. Tento fakt je podľa nás zásadný. Na to, aby sme zistili, či je problém, ktorý nás trápi, vôbec problémom unikátnym, sa nezaobídeme bez znalosti toho, ako daná IKT funguje. Zdá sa, že ak množstvo takých nejasných problémov, ktoré vyžadujú špecializovanú znalosť, presiahne isté kritické množstvo, môžeme povedať, že potrebujeme novú disciplínu. Predpokladáme pritom, že potrebnosť novej disciplíny hovorí v prospech názoru, že daná disciplína je opodstatnená. Domnievame sa, že napríklad problém (5) a viaceré potenciálne budúce problémy, ktoré sme načrtli (AI, virtuálna realita, analýza veľkých dát a iné), sú súčasťou tohto množstva.

Nakoniec, myslíme si, že to isté platí aj pre etablované disciplíny aplikovanej etiky. Tieto disciplíny sa neetablovali preto, lebo by na ich počiatku niekto najskôr preukázal unikátnosť problémov, ktorými sa zaoberali, ale skôr preto, lebo tu najprv boli komunity špecialistov - lekárov, vedcov, environmentalistov a obchodníkov, ktorí pri svojej práci opätovne narážali na morálne problémy, s ktorými si nevedeli poradiť. Filozofi a etici, ktorí by im mohli pomôcť, zasa zistili, že pre identifikáciu a riešenie takýchto problémov sú kľúčové odborné znalosti z danej oblasti.⁴⁴ Dalo by sa dokonca argumentovať, že práve takto, prostredníctvom existujúcej komunity profesionálov s bázou poznatkov potrebnou na identifikáciu a riešenie nových problémov, sa preukazuje *unikátnosť* týchto problémov.

Za etikou IKT vieme bezpochyby nájsť takúto komunitu a potrebnú bázu poznatkov.⁴⁵ V skutočnosti je tu morálna reflexia IKT s nami už viac ako polstoročie. Za ten čas sa okolo nej vytvorila široká, komunikujúca komunita a bohatá odborná literatúra, špecifická terminológia a nová metodológia.⁴⁶ Na základe týchto úvah si preto dovoľíme znova obhajovať potrebnosť etiky IKT. Myslíme si, že argument, ktorý sme pre to poskytli, je cenný aj preto, lebo v budúcnosti môže byť využitý aj pri obhajobe iných aplikovaných etík.

Na druhej strane, aj keď si myslíme, že etika IKT má v súčasnosti svoje opodstatnenie, nie je už tak isté, že ho bude mať aj v budúcnosti. Jej opodstatnenosť

⁴⁴ Pohľad do histórie aplikovanej etiky ponúka Collste. Za podstatnú pre formovanie aplikovanej etiky považuje okrem iného aj existenciu vákua politik (pojem prevzatý od J. Moora), teda situácií, ktoré nedokážeme v danej odbornej oblasti riešiť s aktuálnou etickou a legislatívnou výbavou. Pozri Collste G., *Applied and Professional Ethics - an Introduction*. In Collste, G. (ed.), *Perspective on Applied Ethics*. Linköping, Centre for Applied Ethics 2007, s. 22-39.

⁴⁵ Ako sme spomínali, etické otázky týkajúce sa IKT si autori pokladajú už od konca štyridsiatych rokov minulého storočia, pričom kontinuálnu akademickú činnosť v tejto oblasti je možné datovať najneskôr do začiatku osemdesiatych rokov. Etické problémy IKT riešia stovky expertov prevažne v USA a EU a IKT etici sa stávajú súčasťou nielen expertných tímov, ale aj technologických laboratórií. Pre ich publikačnú činnosť medzičasom vzniklo niekoľko špecializovaných odborných časopisov, na ktoré sme odkazovali aj my.

⁴⁶ Oblasť IKT prispela do etickej reflexie pojmami ako *policy vacuum*, *wide reflexive equilibrium*, alebo *many hands problem*. Vďaka IKT sa prebudil záujem o analýzu pojmov vlastníctva a súkromia a oživil sa záujem o problém hodnotovej zaťažnosti technológie. Pokiaľ ide o nové metódy, môžeme hovoriť o metodikách pre vývoj etických technológií, ktoré nájdeme v prístupoch Value Sensitive Design, Disclosive ethics, alebo Anticipatory ethics, o koncepte dizajnu, ktorý cielene obchádza morálnu dilemu, alebo o etických dotazníkoch a riadených etických debatách, praktizovaných v rámci prístupu Responsible Innovation.

môžu totiž ohroziť dva protichodné trendy v skúmaní morálnych problémov spätých s IKT s technológiou vôbec. Prvým trendom je postupné splynutie IKT s inými typmi technológií. V súčasnosti môžeme okrem IKT odlíšiť tri ďalšie eticky zaujímavé typy technológií: nanotechnológiu, biotechnológiu a neurotechnológiu. Práve tieto technológie budú spolu s IKT podľa všetkého konvergovať, až kým nevytvoria nový, hybridný typ technológie, ktorá bude mať nano, bio, neuro i digitálne prvky. V takom prípade by etika IKT už nemusela mať opodstatnenie, pretože IKT by už nemuseli byť rozlíšiteľným druhom technológie.

Rovnako je možné, že etika IKT bude (alebo už je) natoľko obsahovo bohatá, že sa z nej vyprofilujú viaceré subdisciplíny, pričom sa každá z nich sústreďí na iný druh IKT. Teda, je možné, že v budúcnosti dôjde k rozpusteniu všeobecnej etiky IKT do užších disciplín, keďže už teraz niektorí autori hovoria osobitne o etike UI, roboetike, internetovej etike, etike algoritmov, alebo etike veľkých dát (big data ethics) a podobne.⁴⁷ Keďže budúci vývoj technológií nevieme spoľahlivo predvídať, je možné, že etika IKT bude len dočasne existujúcou disciplínou, kým nebude pohltená jednou väčšou alebo rozdelená na viacero menších oborov. Ide o riziko, ktoré musíme akceptovať. Na druhej strane toto riziko nezmierňuje nástojčivosť problémov, ktoré v tejto chvíli pod etiku IKT podľa nás spadajú a ktoré nevieme vyriešiť v inej aktuálne etablovanej aplikovanej etike.

Záver

Na záver by sme radi zopakovali, že vzhľadom na rýchly vývoj informačných a komunikačných technológií a vzhľadom na ich závažný dopad na spoločnosť považujeme za potrebné venovať sa štúdiu morálnych problémov IKT, špeciálne tým emergentným. Zdôrazňovali sme, že etika IKT musí byť proaktívna. Štúdium jej problémov by sa malo zamerať nielen na používanie IKT, ale aj na ich vývoj, odkrývať problémy skryté a anticipovať problémy prichádzajúce. Rovnako považujeme za potrebné, aby etické úvahy nad IKT spoluurčovali legitímny vývoj a používanie IKT, a taktiež anticipovali morálne znepokojivé trendy.

Domnievame sa, že etika IKT sa kvalifikuje ako samostatná disciplína aplikovanej etiky, pretože si myslíme, že pre ňu existujú unikátne problémy a tiež preto, lebo existuje dostatočné množstvo morálnych problémov, ktoré súvisia s IKT, vyvstávajú v praxi IKT špecialistov a pre riešenie ktorých je potrebná istá báza poznatkov o IKT.

Celkom na záver by sme radi zhrnuli naše praktické návrhy a odporúčania ohľadom IKT. Odporúčame zaviesť štúdium etiky a etiky IKT do kurikula budúcich IKT profesionálov a zdôrazniť etické vzdelávanie v IKT priemysle. Okrem toho sa domnievame, že v IKT priemysle je potrebné sa zamýšľať nad tým, aké štandardy kvality a kontroly je potrebné uplatniť, aby bola inovácia v oblasti IKT zodpovedná a bezpečná. A nakoniec, keďže by etika IKT mala vyhľadávať problémy skryté alebo doteraz nerozpoznané, znamená to, že etický dohľad nad vývojom a používaním IKT

⁴⁷ Myslíme si, že problémy, ktoré riešia tieto disciplíny, je nateraz možné pokryť skúmaniami v rámci celej etiky IKT. Dokonca je to žiaduce s ohľadom na znovu-použiteľnosť získaných poznatkov a metód, ktoré sa v nich používajú. Napr. požiadavky na etický dizajn, vývoj a používanie UI, ktorá postulovala expertná skupina Európskej komisie, rovnako ako metódy, ktoré navrhuje spolu s etickými otázkami, ktoré je potrebné si pri implementácii UI položiť, sú aplikovateľné aj pre ostatné emergentné informačné a komunikačné technológie. Pozri High-Level Expert Group on Artificial Intelligence, *Draft Ethics Guidelines for Trustworthy AI*. Brussels, European Commission 2018. Dostupné na: <
<https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>>.

bude vyžadovať nejakú mieru prístupu k informáciám o skúmaných IKT, ktoré nie sú verejne dostupné.

Referencie

- Angwin, J. et al., Machine Bias. *ProPublica*, 23. máj 2016. Dostupné na: <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.
- Barrat, J., *Our Final Invention: Artificial Intelligence and the End of Human Era*. New York, Thomas Dunne Books 2013.
- Baum, S. D., *A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy*. Global Catastrophic Risk Institute Working Paper 17-1, 2017. DOI: <<http://dx.doi.org/10.2139/ssrn.3070741>>.
- Bonnefon, J. et al., The Social Dilemma of Autonomous Vehicles. *Science*, 352 (6293), s. 1573-1576. DOI: <10.1126/science.aaf2654>.
- Bostrom, N., *Superintelligence: Až budú stroje chytrejšie než lidé*. Praha, Prostor 2017.
- Bostrom, N. - Yudkowsky, E., The Ethics of Artificial Intelligence. In Frankish, K. - Ramsey, W. (eds.), *The Cambridge Handbook of Artificial Intelligence*. Cambridge, Cambridge University Press 2014, s. 316-334.
- Brey, P., Values in Technology and Disclosive Computer Ethics. In Floridi, L. (eds.), *The Cambridge Handbook of Information and Computer Ethics*. Cambridge, Cambridge University Press 2010, s. 41-58.
- Bryson J., Patience is Not a Virtue: The Design of Intelligent Systems and Systems of Ethics. In: *Ethics and Information Technology*, 20 (1), 2018, s. 15-26.
- Collste G., Applied and Professional Ethics - an Introduction. In Collste, G. (ed.), *Perspective on Applied Ethics*. Linköping: Centre for Applied Ethics 2007, s. 22-39.
- Davisson, A. - Booth, P. (eds.), *Controversies in Digital Ethics*. New York, Bloomsbury 2016.
- Doorn, N., A Rawlsian Approach to Distribute Responsibilities in Networks. *Sci Eng Ethics*, 16, 2010, s. 221-249. DOI: <10.1007/s11948-009-9155-0>.
- Fabris, A., *Ethics of Information and Communication Technologies*. Springer 2018.
- Floridi, L., Information Ethics, its Nature and Scope. *SIGCAS Computers and Society*, 36 (3), 2006, s. 21-36.
- Floridi, L., *The Cambridge Handbook of Information and Computer Ethics*. Cambridge, Cambridge University Press 2010.
- Foot, P., The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, 5, 1967, s. 5-15.
- Friedman, B. - Kahn P. H., Human Values, Ethics, and Design. In Sears, A. - Jacko, J. A. (eds.), *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2. edícia). New York, Lawrence Erlbaum Associates 2008, s. 1241-1266.
- Galeon, D., For the First Time Ever, a Robot Was Granted Citizenship, *futurism.com*, 26. október 2017. Dostupné na: <<https://futurism.com/for-the-first-time-ever-a-robot-was-granted-citizenship>>.
- Gunkel, D., *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. Cambridge – London, MIT Press 2012.
- Hern, A., Grand Theft Auto 5 under fire for graphic torture scene, *The Guardian*, 18.

- september 2013. Dostupné na:
<https://www.theguardian.com/technology/2013/sep/18/grand-theft-auto-5-under-fire-for-graphic-torture-scene>.
- High-Level Expert Group on Artificial Intelligence, *Draft Ethics Guidelines for Trustworthy AI*. Brussels, European Commission 2018. Dostupné na: <
<https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>>.
- Himma, K. E. - Tavani, H. T. (eds.), *The Handbook of Information and Computer Ethics*. Wiley 2008.
- Jonas, H., *Princip odpovědnosti*. Praha, Oikoymenth 1997.
- Maner, W., Unique Ethical Problems in Information Technology. *Science and Engineering Ethics*, 2 (2), 1996, s. 137-154.
- McGinn, R. E., What's Different, Ethically, about Nanotechnology?: Foundational Questions and Answers. *Nanoethics*, 4, 2010, s. 120.
- Mitchell, T., *Machine Learning*. McGraw-Hill 1997.
- Moor, J., What is Computer Ethics? *Metaphilosophy*, 16 (4), 1985, s. 266-275.
- Moor, J., Why We Need Better Ethics for Emerging Technologies. *Ethics and Information Technology*, 7 (3), 2005, s. 111-119.
- Ramirez, E. J. - LaBarge, S., Real Moral Problems in the Use of Virtual Reality. *Ethics and Information Technology*, 20 (4), 2018, s. 251, DOI: <
<https://doi.org/10.1007/s10676-018-9473-5>>.
- Ribeiro M.T. - Singh S. - Guestin C., "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, s. 1135-1144. DOI: <10.1145/2939672.2939778>.
- Tavani, H. T., *Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing* (5. edícia). Wiley 2015.
- Tegmark M., *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York, Penguin Books Ltd. 2018.
- Thompson, J. J., Killing, Letting Die, and the Trolley Problem. *The Monist*, 59, 1976, s. 204-17.
- van Est, R. et al., *Human Rights in the Robot Age: Challenges Arising from the Use of Robotics, Artificial Intelligence, and Virtual and Augmented reality*. Hague, Rathenau Instituut.
- Wiener, N., *The Human Use of Human Beings: Cybernetics and Society*. London, Free Association Books 1989.
- World Commission on Ethics of Scientific Knowledge and Technology, *Report of COMEST on Robotics Ethics*. Paris, UNESCO 2017. Dostupné na: <
<http://unesdoc.unesco.org/images/0025/002539/253952E.pdf>>.