

Sabine Gleß*

Pravni fakultet, Univerzitet u Bazelu

Thomas Weigend**

Pravni fakultet, Univerzitet u Kelnu

INTELLIGENTNI AGENTI I KRIVIČNO PRAVO***

Apstrakt: Autori se bave pitanjem naizgled krivičnog pravnog futurizma sa kojima se savremeni čovek već uveliko sreće, i sa kojima će se sretati u doglednoj budućnosti u još većoj meri, suočavajući oblike krivice i načine kažnjavanja tzv. inteligentnih agenata. U prvom delu se s tehničkog aspekta opisuju vrste inteligentnih agenata i način njihovog delanja. Centralno pitanje ovde jeste postupak donošenja odluka, čime se otvara put za analizu krivičnog pravnog odgovornosti samog inteligentnog agenta, s jedne, i njegovog ljudskog operatera s druge strane. Dok se u razmatranju mogućeg kažnjavanja inteligentnih agenata pravi analogija s kažnjavanjem životinja u prošlim vekovima, te pravnih lica u savremenom dobu, dotle se krivičnog pravnog odgovornosti ljudskog operatera inteligentnih agenata ispituje primenom važećih teorija krivičnog prava i mogućnošću primene osnovnih instituta Opšteg dela krivičnog prava. Zanimljiva filozofsko-pravna stanovišta su uključena u ova promišljanja, koja iznova opominju da su klasična pravna načela primenjiva i na tehnički najsavremenije ljudske tvorine sposobne da „misle“. Na ovom mestu autori najviše pažnje poklanjaju sankcionisanju nehatnog oblika krivice operatera inteligentnog agenta, i dolaze do zaključka da je takav oblik kažnjavanja moguć, ali da valja imati u vidu odmeravanje rizika koji je delanju inteligentnih agenata imanentan i koristi koje društvo uživa, a koje proističe iz njihove upotrebe.

Ključne reči: inteligentni agent, kažnjavanje, operater, rizik, objektivno uračunavanje

I INTELLIGENTNI AGENTI – POTENCIJAL I RIZIK

„Kriv je algoritam“, moglo se pročitati u novinama.¹ Supruga bivšeg predsednika Savezne Republike Nemačke se pravnim sredstvima usprotivila jer se njeno ime automatski povezivalo sa pojmovima kao što su „prostitucija“ ili „poslovna

* redovni profesor, sabine.gless@unibas.ch

** redovni profesor, Thomas.Weigend@uni-koeln.de

*** Rad je u Nemačkoj objavljen pod naslovom „Intelligente Agenten und das Strafrecht“ u *Zeitschrift für die gesamte Strafrechtswissenschaft*, Vol. 126 (2014) Nr. 3, pp. 561–591.

1 „Die Welt“ od 10.09.2012., http://www.welt.de/welt_kompakt/print_politik/article109114305 (posećeno 18.07.2014).

pratnja⁴, kada bi se unelo u Google pretragu. Ova povezivanja, koja je tužiteljka doživljavala kao uvredu, nisu bila zasnovana na individualnim odlukama lica iz delokruga tužene firme, već na funkcijama „google-bot“ i „google-autocomplete“, koji upite korisnika obrađuju prema opšte utvrđenim smernicama.² Obe funkcije su primer takozvanih inteligentnih agenata³: One postupaju prema određenim zadatim pravilima, ali obrađuju informacije (pretraživačko ponašanje korisnika Google-a) u svakom pojedinačnom slučaju samostalno.

Inteligentni agenti se – u svojim različitim oblicima⁴ – koriste tamo, gde hitra obrada velikog broja informacija zahteva preciznu kombinaciju i brzu reakciju, kao i tamo gde se zahteva upotreba fizičke snage, koja prevazilazi mogućnosti čoveka. Inteligentnih agenata danas ima u mnogim životnim područjima: oni određuju – u sasvim jednostavnom obliku softverskog agenta – kao mašine-pretraživači na internetu naš pristup informacijama, a time donekle i naše poimanje stvarnosti;⁵ „čitajući softverski agenti“ u onlajn-berzanskoj trgovini, autonomno donose na osnovu određenih poruka, odluku o kupovini ili prodaji.⁶ Kompleksniji sistemi poput „inteligentnih skalpela“ i „robota glodača“ obavljaju medicinske operacije,⁷ inteligentne letilice („dronovi“) se ne angažuju samo u vojne svrhe,⁸ već i za nadzor i obezbeđenje osetljivih civilnih područja, a u bliskoj budućnosti će možda biti angažovani i za isporuku robe.⁹ Najočevidnije se u svesti javnosti danas pojavljuju inteligentni agenti, zahvaljujući naporima industrije da proizvede automobile bez vozača za korišćenje na javnim putevima.¹⁰

Preduslovi da jedna mašina bude okvalifikovana kao inteligentni agent sporni su u informatici isto onoliko koliko je sporan i pojam inteligencije.¹¹ Za vrednovanje,

- 2 O pitanju građanskopravne odgovornosti vidi BGH NJW 2013, p. 2348; *Gounalakis*, NJW 2013, p. 2321; *Hager*, JA 2013, p. 630; *Mäsch*, JuS 2013, p. 814.
- 3 *Russel/Norvig*, Artificial Intelligence. A Modern Approach, 3. Aufl. 2014, p. 34ff.
- 4 Kao sinonim za opšti pojam inteligentni ili autonomni agenti koristi se često pojam „robot“; vidi npr. *Beck*, JR 2009, p. 226; isto u *Beck* (Hrsg.), *Jenseits von Mensch und Maschine*, 2012, p. 15. Često upotrebljavani uži pojmovi za različite oblike inteligentnih agenata su npr. „softverski agenti“, „android“ ili „dron“.
- 5 Vidi o tome npr. *Egermann*, Suchmaschinen, in: *Killan/Heussen*, Computerrecht, Stand 2013, Rdn. 7–10.
- 6 *Bram/Stalknecht*, Handelsblatt Online od 24.04.2013, <http://www.handelsblatt.com/finanzen/boerse-maerkte/boerse-inside/blitzabsturz-an-der-boerse-maschinen-an-der-macht/8113842.html> (posećeno 18.07.2014).
- 7 *Hilgendorf*, in: *Rotsch* (Hrsg.), *Criminal Compliance vor den Aufgaben der Zukunft*, 2013, p. 24f. O dopuštenosti upotrebe „robota glodača“ u operaciji zgloba vidi građanskopravne presude OLG Dresden NJOZ 2008, 247 und OLG Frankfurt a.M., Urt. v. 08.11.2013–25 U 79/12 (<http://openjur.de/u/668787.html>).
- 8 Dopuštenost upotrebe dronova u vojne svrhe je sa stanovišta međunarodnog humanitarnog prava sporno; vidi o tome na primer *Asaro*, *International Review of the Red Cross* 94 (2012), p. 687; *Heller*, *Journal of International Criminal Justice* 11 (2013), p. 89; *May*, *Journal of International Criminal Justice* 11 (2013), p. 47; *Pagallo*, *The Laws of Robots*, 2013, p. 51f.; *Sparrow*, *Journal of Applied Philosophy* 24 (2007), p. 62.
- 9 Vidi npr. <http://www.amazon.com/b?node=8037720011>.
- 10 Vidi npr. <http://www.google.ch/about/careers/lifeatgoogle/self-driving-acr-test-steve-mahan.html>.
- 11 *Alan Turing* je 1950. razvio sledeći test: Čovek-postavljač pitanja komunicira koristeći tastaturu i ekran bez vizuelnog i audio kontakta sa dvojicom neznanaca; jedan je čovek, a drugi je mašina. Obojica pokušavaju da postavljača pitanja ubede da su misleći ljudi. Ukoliko postavljač pitanja

iz pravnog ugla, čini se smislenim, da se mašine razvrstaju prema tehničkom načinu funkcionisanja, i to u tri vrste: (a) jednostavni sistemi za obradu podataka, poput džepnih računara, kod kojih je tvorac sistema na osnovu hijerarhijski definisanih algoritama definitivno utvrdio kako način mogućeg unosa podataka tako i njihovu obradu; (b) otvoreni sistemi, poput automatizovanih kosilica za travu koje – putem senzora – mogu prihvatati različite podatke iz spoljnog sveta, i njih obrađivati prema pravilima koje je nepromenljivo utvrdio njihov tvorac; (c) inteligentni sistemi, koji ne samo da usvajaju veliku količinu podataka, već ih prema određenom obrascu tumače i samostalno reaguju na obrađene podatke. Otvoreni i inteligentni sistemi postupaju utoliko autonomno, što senzorski analiziraju i obrađuju svoju okolinu. Ali jedino su inteligentni sistemi pored toga u stanju da na preuzete i samostalno protumačene informacije neposredno reaguju i da ih uz to, pohrane na takav način koji im dozvoljava da se njima posluže u kasnijim procesima donošenja odluke. Ovo može da funkcioniše na veoma jednostavnom nivou, poput internet-pretraživača, koji pohranjuje sve upite korisnika, nudi varijacije najtraženijih upita i beleži prihvaćeno. Neki inteligentni agenti moraju donositi i veoma kompleksne odluke, poput autonomnog automobila, koji u interakciji sa vozačima-ljudima mora preći raskrnicu s gustim saobraćajem, regulisanu saobraćajnim znacima.

Čak i jedan izuzetno visoko razvijeni inteligentan agent je doduše samo onoliko inteligentan koliko je isprogramiran. Pravila ponašanja¹² koja su njemu pridodata na osnovu obimnih simulacija, ne mogu ga nikada pripremiti za sve zamislive različitosti slučajeva stvarnoga života. Uz to ide da inteligentnim agentima uprogramirana pravila ne mogu – kao ljudima, putem analogije ili „intuicije“ biti preneti na nove situacije. Na primer, ako bi se u gore navedenom slučaju „google-autocomplete-u“ zadalo pravilo da ime određene osobe ne sme dopuniti pojmovima „prostitucija“ ili „poslovna pratnja“, sistem i dalje ne bi bio u stanju da zaključi da su pojmovi poput „kol-gerla“ ili „crveni fenjer“ zabranjeni.¹³

„Stvarno“ inteligentni sistemi svakako razvrstavaju primljene šumove podataka ne samo prema obrascima, kako bi na taj način što efikasnije postigli zacrtani cilj,¹⁴ već mogu, više od toga (u najširem smislu), da uče,¹⁵ utoliko što svoje ponašanje mogu prilagoditi dobijenim informacijama. I to funkcioniše na veoma jednostavnom stupnju, otprilike kao *recommender system* koji kupcu u onlajn prodavnici saopštava, „ono što bi ga još moglo interesovati“ (i čije se preporuke u velikoj meri približavaju intuiciji stvarnog prodavca). Iza *recommender system*-a koji korisniku automatski predlaže ostale ponude, nalazi se inteligentni agent, koji unose korisnika kontinuirano procenjuje shodno njegovim preferencijama i odbojnostima, pri

na kraju ne može sa sigurnošću da kaže ko je od sagovornika mašina, smatra se da je mašina položila Turingov test.

Vidi o tome: *Alan Turing*, *Computing Machinery and Intelligence*, *Mind* 59 (1950), p. 433.

12 *Russel/Norvig* (nap. 3), p. 57ff.

13 Čak i jednostavni softverski agenti mogu pak, naučiti „prenos“. To se dešava na primer, na taj način kada se jedan inteligentni agent poveže sa drugim ili kada na osnovu generatora slučajnosti izvlači zaključke iz okoline.

14 Vidi o tome *Wooldridge/Jennings*, *Knowledge Engineering Review* 10 (1995), p. 115; *John*, *Haf-tung für künstliche Intelligenz*, 2007, p. 15ff.

15 *Hofstadter*, *Fluid Concepts and Creative Analogies*, 1995, pp. 13ff., 62ff.

čemu jedan specifični statistički sistem kombinuje sa generatorom slučajnosti.¹⁶ U tu svrhu agent pohranjuje informacione tragove svih korisnika, jer *recommender system* potrebuje ogromne količine podataka, kako bi stvorio klasterne mušterija koji – statistički posmatrano – imaju iste interese.¹⁷ U jednoj daleko kompleksnijoj ravni „autonomno“ funkcioniše automobil bez vozača: Budući da on putem senzora samostalno po određenom obrascu ispituje prikupljene podatke i na njih neposredno reaguje, ne čini se samo posmatraču da automobil samostalno dela; već ni konstruktor ne može sa sigurnošću da predvidi, kako će automobil u određenoj saobraćajnoj situaciji reagovati povezivanjem različitih procesa obrade podataka, i ne može pojedine „odluke“ automobila, čak ni retrospektivno u potpunosti da shvati. Inženjer koji razvija automobil bez vozača, može, dakle, pod okolnostima tek posle niza udesa da spozna, da njegov automobil, na primer, prilikom određenih svetlosnih ili vremenskih uslova nije u stanju da raspozna znak prvenstva prolaza.

Nepredvidivo postaje ponašanje inteligentnog agenta ne samo onda kada dođe u situaciju za koju mu nije uprogramirano adekvatno reagovanje, već i onda kada na osnovu svog „iskustva“ autonomno odlučuje o određenim akcijama.¹⁸ Inteligentni agenti pokušavaju da izvrše naloge koji su im ljudi zadali, čak i kada zbog toga moraju da se odrede prema novim situacijama. Upravo u tome se sastoji njihova velika korist. Time, pak, čovek ne može niti u potpunosti da predvidi niti da kontroliše ponašanje agenta u takvim neplaniranim situacijama; čovek, na primer, ne može da predvidi da li i kada je autonomni automobil u stanju da saobraćajni znak koji nije u potpunosti vidljiv, pogrešno protumači kao neki, iz njegovog „iskustva“ sličan saobraćajni znak. Nedostatak mogućnosti kontrole je u neku ruku tamna strana medalje one posebne vrednosti koju roboti sposobni da uče predstavljaju za ljude.

Sledeći korak mogao bi biti razvoj „emocionalnih“ interakcija između mašine i čoveka: radi se na tome da se inteligentni agenti tako programiraju da budu u stanju da prepoznaju emocionalne potrebe ljudi i da simuliraju empatične reakcije. Poneki korisnik će se obradovati saosećanjem iskazanim od strane mašine (i možda će sledećeg puta pre biti spreman da od njih nešto poruči), iako se u stvarnosti radi o jednom veoma jednostavnom automatizmu. Čak je i kod nege bolesnika ustanovljeno da oboleli od demencije mogu uz pomoć robota u obliku životinje izgraditi svojevrsnu „osećajnu vezu“ tako što mašini pripisuju osećanja. U informatici se govori o *Eliza*-efektu,¹⁹ onda kada čovek radnju inteligentnog agenta nesvesno doživljava kao obrazac ljudskog ponašanja i sopstvenu motivaciju.²⁰ Međutim, ovaj efekat nosi

16 Time kod velikih onlajn trgovaca dolazi do pridodavanja daljih „preporuka na osnovu Vaše porudžbine“.

17 Tako nastaje izuzetno obiman problem zaštite podataka. Vidi o tome npr: Jeckmans/Beye/Erkin/Hartel/Lagendijk/Tang, *Privacy in Recommender Systems*, 2013, dostupno na: http://eprints.eemcs.utwente.nl/22141/01/Privacy_in_Recommender_Systems.pdf (posećeno 18.07.2014).

18 Russel/Norvig (nap. 3) pp. 35ff., 57ff.; Wooldridge/Jennings, *Knowledge Engineering Review* 10 (1995), p. 118f.

19 Računarski program ELIZA razvio je Jozef Vajcenbaum (Joseph Weizenbaum), koji omogućava da kompjuter simulira psihoterapeuta koji (naizgled) ostvaruje jezičku interakciju sa čovekom; Weizenbaum, *Computer Power Human Reason*, 1976; Hildebrandt, in: Hildebrandt/Gaakeer (Hrsg.), *Human Law and Computer Law: Comparative Perspectives*, 2013, p. 27f.

20 Vidi o ovome Picard, *Affective Computing*, 1995; Tao/Tan, in: Tao/Tan/Picard (Hrsg.), *Affective Computing and Intelligent Interaction*, 2005, p. 981.

sobom opasnost pogrešnog tumačenja (re)akcije mašine i time može dovesti do “nesporazuma” u interakciji između ljudi i inteligentnih agenata.

Kao što se moglo videti iz kratkog pregleda, inteligentni agenti koji poseduju sposobnost učenja otvaraju ljudima novu dimenziju napretka. Istovremeno, pravna ocena njihovog ponašanja nameće nova i teška pitanja. Delikatno je, pre svega, pitanje odgovornosti za delovanje inteligentnih agenata. O tome se u građanskom pravu već raspravljalo;²¹ u krivičnom pravu, naprotiv, rasprava se nalazi tek na početku.²² Prouzrokuje li inteligentni agent kakvu krivičnopravno relevantnu posledicu (na primer ubistvo ili ranjavanje čoveka ili uništenje stvari), postavlja se pitanje odgovornosti: Ko je delao i ko snosi krivičnopravnu krivicu? Čovek koji je tehnologiju osposobio da funkcioniše? Ili, pak, sam inteligentni agent? Najzad, možda i niko nije krivičnopravno odgovoran, jer ljudi smatraju razvoj i korišćenje takvih tehnologija poželjnim, i tako moraju sami snositi posledice njihove (delimične) nepredvidivosti? Ovim pitanjima se bavi sledeća skica mogućeg objašnjenja i ograničenja krivičnopravne odgovornosti delanja inteligentnih agenata.

II INTELIGENTNI AGENT KAO KAŽNJIV SUBJEKT?

Naše krivično pravo je stvoreno za ljude. Čoveku se pripisuje posledica krivičnog dela, kada ga on prouzrokuje svojim ponašanjem. Čovek snosi krivičnopravnu krivicu, kada, umišljajno ili iz nehata ostvari biće krivičnog dela, a uz to je svestan protivpravnosti svog ponašanja i mogućnosti da ga izbegne. Ipak, uočili smo, da su i inteligentni agenti sposobni da uče i da bar u prenesenom značenju mogu „autonomno“ postupati. Znači li to, da i oni mogu biti kažnjeni kao i ljudi?

1. *Prethodnici: životinje i roboti kao subjekti krivičnog prava*

Pre nego što pokušamo da pružimo odgovor na ovo pitanje, želeli bismo da bacimo pogled na iskustva i predstave o krivičnopravnom sankcionisanju ne-ljudi.

Tokom prethodnih vekova čovek nije još posedovao monopol nad krivicom i kaznom. Odmazda je činjena u velikoj meri i prema “bezdušnim predmetima.”²³ Istorija krivičnog prava izveštava o krivičnim postupcima i kažnjavanju životinja u Evropi sve do ranog Novog veka.²⁴ Predanje govori o jednom od mlađih postupaka protiv životinja, vođen u Bazelu (1474.): sudilo se petlu; stavljano mu je na teret da je sneo jaje iz koga je trebalo da se izleže Basilisk.²⁵ Bazelski petao – kao i

21 Vidi npr. Gruber, in: Beck (Hrsg.), *Jenseits von Mensch und Maschine*, 2012, p. 150; Hanisch, in: Hilgendorf (Hrsg.), *Robotik im Kontext von Recht und Moral*, 2014, p. 27; Spindler, in: Hilgendorf (Hrsg.), p. 63.

22 Vidi, pak Zbornik koji je priredio Beck (nap. 21).

23 Bauer, in: Perels/Wojak (Hrsg.), *Die Humanität der Rechtsordnung. Ausgewählte Schriften*, 1998, p. 251.

24 Postupci protiv životinja sadržali su elemente krivičnog postupka protiv ljudi, poput pritvora ili dodele branioca optuženoj životinji. Vidi: Dinzlacher, *Das fremde Mittelalter. Gottesurteil und Tierprozess*, 2006, p. 113.

25 Brednich, *Enzyklopädie des Märchens*, Bd.13, 2010, p. 651. – (Basilisk – mitsko biće koje je u Srednjem veku simbolizovalo smrt, đavola i greh – prim. prev).

druge životinje kojima je presuđeno – kažnjen je smrću. Šta je zapravo bio smisao kažnjavanja životinja, ostalo je sporno do danas. Istoričari prvo upućuju na sasvim profano pravo seljaka da ubiju stoku koja luta i koja pričinjava štetu.²⁶ Prema drugom izvoru svrha krivičnog postupka protiv životinja bila je težnja za magičnom zaštitom od đavolje rabote. Moguće je da je izvesnu ulogu imala i misao o odmazdi: U srednjem veku se smatralo da je Bog u odnosu na ljude životinjama dodelio podređenu funkciju “služenja”;²⁷ prilikom povrede ove funkcije životinja bi zaslužila kaznu isto – kao i ljudi – kojom bi odnos hijerarhije u svetu bio vaspostavljen.²⁸ Krivično pravo protiv životinja, okončano je prosvetiteljstvom i prodorom preventivnih teorija kazne.²⁹ Životinje se, prema tome, više ne tretiraju kao ljudi već kao stvari (vidi paragraf 90a stav 2 BGB) Ipak, danas se u okvirima pokreta za Prava životinja (“Animal Rights”) raspravlja o tome ne bi li životinjama pored prava, trebalo da pripadaju i obaveze.³⁰

Mašine imaju, prema modernim shvatanjima status lica isto onoliko malo, kao i životinje. Ne bivaju shvaćene kao samosvojna bića, pošto su definisane njihovim načinom funkcionisanja. *Mark Twen* je ovo formulisao još davne 1906. godine:³¹ “[The machine] does no more and no less than the law of its make permits and compels it to do. There is nothing personal about it; it cannot choose.” Ni “društveni roboti”, odnosno inteligentni agenti koji su predviđeni za interakciju sa ljudima,³² ne predstavljaju ništa drugo do spoj različitih tehničkih komponenata, poput kućišta, računara i senzora.

U žanru naučne fantastike susrećemo uvek iznova slutnje, da bi jedna mašina zamišljena kao čovekov pomoćnik mogla, ipak jedared postati nešto više. Robot koji se lagano pretvara u čoveka jeste lajtmotiv naučnofantastičnih romana. Još je 1950. godine *Isak Asimov* slikovito formulisao pitanje pravnog subjektiviteta mašina. U romanu sa upečatljivim naslovom “I, Robot” *Asimov* je predstavio pravila ponašanja, koja bi trebalo da regulišu (hijerarhijski ustrojen) odnos između čoveka i robota.³³ I u mnogobrojnim naučnofantastičnim filmovima mašine sa rastućom samosvešću igraju bitnu ulogu.³⁴ Klasičan primer jeste David, dete Android, koje u filmu

26 Brednich (nap. 25), p. 651.

27 Gergen, NuR 2012, pp. 96, 102.

28 Fischer, in: Rehberg (Hrsg.), Die Natur der Gesellschaft, 2008, pp. 5151, 5159.

29 Dinzlacher (nap. 24), p. 111.

30 Vidi Ritvo, Daedalus 136 (2007), p. 118; De Mello, Animals and Society, 2012; Peters/Stucki/Boscardin, The Animal Turn - What is it and Why Now?, Verfassungsblog.de 14. April 2014, dostupno na: <http://www.verfassungsblog.de/the-animal-turn-what-is-it-and-why-now> (posećeno 18.07.2014).

31 Mark Twain, What is man? 1906, dostupno na: <http://www.gutenberg.org/ebooks/70> (posećeno 18.07.2014).

32 Vidi Kobald, Ideengeschichtliche Aspekte zur Technikgenese der sozialen Robotik unter weltanschaulicher Perspektive, 2013, p. 13f.

33 Asimov, I, Robot, 1950, p. 27. Shodno tome trebalo bi da važe sledeća tri pravila: „1. A robot may not injure a human being or, through inaction, allow a human being to come to harm; 2. A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law. 3. A robot must protect its own existence, as long as such protection does not conflict with the First or Second Laws.” (dostupno na: http://cdn.preterhuman.net/texts/literature/books_by-author/A/Asimov,%20Isaac%20-%20I,%20Robot.pdf [posećeno 18.07.2014]).

34 Vidi na primer 2001: A Space Odyssey (1968); Star Trek, The Ultimate Computer (1968); The Stepford Wives (1975); Blade Runner (1982); Artificial Intelligence (2001).

“AI” *Stivena Spilberga*, kajući se, obećava svojoj ljudskoj majci da će se popraviti, kako bi potom – preobraćen, poput klasičnog Pinokija – smeo da se vrati kući.³⁵

2. Inteligentni agenti kao lica?

Iako se kod prikaza inteligentnih agenata koji su naučili da razvijaju samosvest, osećanja, empatiju i moral, i kojima se s toga pripisuje jedna vrsta slobode volje i možda čak status moralnog agenta, radi samo o fiktivnim slutnjama, a ne o stvarnosti, takve knjige i filmovi nameću važna i interesantna pitanja – i ne na posletku pitanje, da li i pod kakvim pretpostavkama dolazi u obzir kažnjavanje inteligentnih agenata.³⁶ Možda nam se, pak roboti uskoro neće činiti “daljinski upravljivijim” nego što su to ljudi, koji, svoje navodno slobodne odluke donose u nedokučivom spletu uticaja genetike, vaspitanja i društvenog okruženja i koje još uvek definišemo kao “slobodne” u pogledu svoje odgovornosti.

U tom smislu, postavlja se, pre svega, pitanje gde se proteže granica između mašine i ljudskog lica kao tradicionalnog subjekta krivičnog prava. Da li je dovoljno, da robot poseduje isprogramiranu sposobnost, da iz okruženja primljene podatke poredi sa poznatim obrascima i da na toj osnovi dela, kako bismo ga smatrali “licem” odgovornim za sopstveno ponašanje?

Prema jednom pravnofilosofskom stanovištu *Džona Loka*³⁷ i *Imanuela Kanta*, licem se smatra onaj ko je u stanju da razvije misaoni odnos o samome sebi. Sposobnost samorefleksije omogućava osobi samoodređenje; ili, kako to *Kant* izražava: Pošto je svestan sebe samoga, čovek sebe posmatra “kao odredljivog samo na osnovu zakona koje samom sebi daje pomoću uma. U tom njegovom postojanju njemu ništa ne prethodi pre određenja njegove volje, već svaka njegova radnja (...) nije u svesti o njegovoj egzistenciji ništa drugo do posledica, pa to nikada ne treba smatrati određujućim razlogom njegovog kauzaliteta kao noumens.”³⁸ Iz tog razloga, nastavlja *Kant*, umno biće može za svaku protivpravnu radnju koju učini, s pravom da kaže, da je moglo i da je propusti.³⁹ Ako lice, pak, u samosvesti svoje egzistencijalne

35 David: „I’m sorry I broke myself. I’m sorry I cut your hair off. I am sorry I hurt Martin. [...] Mommy if Pinocchio became real and I become a real boy can I come home?”

36 O tome Beck, in: Battaglia/Mukerji/Nida-Rümelin/Bisol (Hrsg.), *Rethinking Responsibility in Science and Technology*, Pisa 2014, p. 177ff.; Koops/Hildebrandt/Jaquet-Chiffelle, *Minnesota Journal of Law, Science & Technology* 11 (2010), p. 522ff.; Hildebrandt, in: Hildebrandt/Rouvroy (Hrsg.) *Law, Human Agency and Autonomic Computing*, 2011, p. 141 ff.

37 John Locke, *An Essay Concerning Human Understanding*, 1690, p. 331 (§XVII Nr. 26), uzima za pretpostavku bića-lica, da je neko poput „intelligent Agent“ (!) „capable of a Law, and Happiness and Misery“. Svest o sopstvenoj prošlosti je ključ za lice: „This personality extends itself beyond present existence to what is past, only by consciousness, - whereby it becomes concerned and accountable; owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present.“ Vidi o ovome takođe Seelmann, *Rechtsphilosophie*, 5. Aufl. 2010, p. 132ff.; opširno isti., in: Heer u.a. (Hrsg.), *Toujours agite, jamais abattu*, Festschrift für Wiprächtiger, 2011, p. 575.

38 Kant, *Kritik der praktischen Vernunft*, in: Weischedel (Hrsg.), *Immanuel Kant, Werke in zehn Bänden*, 1975, Bd. 6, p. 223. – Kritika praktičnog uma. Preveo Danilo Basta. Beograd, BIGZ, 1979, str. 115 – prim. prev.

39 Kant (nap. 38), p. 223.

slobode ne može da sebe shvati odgovornim za svoja dela,⁴⁰ onda su drugi vlasni da tom licu pripisuju njegovu odgovornost. Drugo stanovište seže u *Hegelovu* tradiciju mišljenja, i oštro razdvaja pojam „lica“ od pojma „čoveka“ i shvata ono prvo kao oznaku ljudskog bića kao mogućeg nosioca prava i obaveza u društvu.⁴¹

No, ta razmatranja koja su zasnovana na idealističkoj filosofiji se svakako ne odnose na moderne robote. Ali, može se reći da inteligentnom agentu, od bitnih elemenata lica-bića nedostaje: Sve i da može da uči i da je sposoban da donosi – za druge nepredvidive – odluke, on ipak ne može da postane svestan svoje slobode; a još manje može da shvati sebe kao nosioca prava i obaveza u društvu.⁴² Inteligentni agenti, mogu, doduše, da obave određene zadatke, „automatski“, ali oni slede – čak ukoliko su sposobni da uče – u krajnjoj liniji, opcije zadate programiranjem i sebi samima o svom ponašanju ne polažu račun.⁴³ S toga inteligentne agente doživljavamo kao „ne-slobodne“ i ne možemo ih učiniti lično odgovornim za štetu koju pričinu.

Ovi zaključci, odnose se – s tim bi trebalo biti načisto – na sadašnje stanje razvoja veštačke inteligencije, s jedne strane, i tradicionalnog shvatanja bića-ličnosti u smislu nemačke filosofije s druge strane, pri čemu ovo poslednje još uvek predstavlja relevantan parametar krivičnog prava. Ukoliko bi inteligentni agenti i u pogledu samo – i pravne svesti, etičke i emocionalne oslovljenosti i ostalih „tipično ljudskih“ dimenzija, mogli da budu dalje razvijani, naravno da bi iznova trebalo razmatrati pitanje, mogu li se ipak posmatrati kao (tada, po mogućstvu kao krivičnoodgovorni) pravni subjekti.⁴⁴

3. Kažnjivost inteligentnih agenata?

Problem, da se inteligentni agenti *rebus sic stantibus* za svoje ponašanje u pravnom i moralnom pogledu ne mogu pozivati na odgovornosti odslikava se i u kategorijama krivičnopravne dogmatike. Isto kao kod, u pojedinim tačkama slično pozicioniranoj postavci pitanja kažnjivosti pravnih lica,⁴⁵ mogu se problemi isprva locirati pri pretpostavkama „radnje“ i „krivice“.

40 Kant (nap. 38), p. 223f., upućuje ovde na glas savesti čija samokritika i ukazivanje na čoveka – „čovek može mudrijašiti koliko mu drago“ – ne može da učutka. Preveo Danilo Basta. Beograd, BIGZ, 1979, str. 115 – prim. prev.

41 Vidi o ovome Jakobs, Norm, Person, Gesellschaft, 3. Aufl. 2008, p. 37ff.; Pawlik, Das Unrecht des Bürgers, 2012, p. 141ff.

42 Načelna razmatranja o pravnom subjektivitetu na polju veštačke inteligencije vidi kod: Solum, North Carolina Law Review 70 (1992), p. 1231; Matthias, Automaten als Träger von Rechten, 2008.

43 Vidi takođe Koops/Hildebrandt/Jaquet-Chiffelle (nap. 36), p. 522 ff.

44 O pitanju statusa vidi Beck, in Hilgendorf/Günther (Hrsg.), Robotik und Gesetzgebung, 2013, p. 239ff.; Hildebrandt, in: Hildebrandt/Gaakeer (nap. 19), p. 27f.

45 Vidi o ovome temeljito izloženo: Böse, ZStW 126 (2014), p. 126 ff.; Frisch, Festschrift für Wolter, 2013, p. 349; Heine, Die strafrechtliche Verantwortlichkeit von Unternehmen, 1995; Hettinger (Hrsg.), Reform des Sanktionenrechts, Bd. 3: Verbandstrafe, 2003; Jakobs, Festschrift für Luderksen, 2002, p. 559; Mitsch, NZWiSt 2014, p. 1; Schünemann, Festschrift für Tiedemann, 2008, p. 429; isto., in: Leipziger Kommentar, 12. Aufl. 2007, vor § 25 Rdn. 17 ff; Volk, JZ 1993, p. 429. U najnovije vreme vidi pokušaj rešavanja u nacrtu Krivičnog zakonika za udruženja Pokrajine Severna Rajna- Vestfalija; dostupno na: http://www.justiz.nrw.de/JM/justizpolitik/jumiko/beschluesse/2013/herbstkonferenz13/zw3/TOP_IL_5_Gesetzenentwurf.pdf (posećeno 18.07.2014); o tome Hoven, ZIS 2014, p. 19; Löffelmann, JR 2014, p. 185; Zieschang, GA 2014, p. 91.

a) Poslovna sposobnost inteligentnih agenata

Internet-pretraživač sklapa kombinacije reči štetne po ugled, dron povređuje bezopasne ljude, automobil bez vozača gazi pešaka: Ovako opisani postupci deluju kao radnje koje upućuju na samog inteligentnog agenta ukoliko ovaj nezavisno od ljudskog upravljanja dela na osnovu svog – u gore navedenom smislu „podučivosti“⁴⁶ – algoritamskog programa. Ali, da li su ovo zaista „radnje“ inteligentnog agenta u krivičnopravnom smislu?

Pitanjem, kada postoji pravno relevantna radnja, nauka krivičnog prava se bavila pre svega sredinom prošlog veka. Za prevaziđenu „kauzalističku“ teoriju, dovoljan je bilo koji voljni pokret telom za krivičnopravno relevantnu radnju; njena moderna „društvena“ varijanta zahteva sem toga određeno društveno značenje pokreta tela.⁴⁷ Ukoliko se ove definicije primene na inteligentne agente, onda se barem kod onih inteligentnih agenata koji su napustili čisto softversko okruženje (poput interneta) i istupili pred nas u obliku mehanički delajućeg robota,⁴⁸ bez daljeg da se njihova tela „pokreću“; ali se tada svakako postavlja pitanje da li inteligentni agenti mogu „samovoljno“ da delaju. Prema kauzalističkom shvatanju radnje ne postavljaju se pak, nikakvi visoki zahtevi u pogledu voljnog upravljanja ponašanjem; poriče se voljna radnja samo onda kada telo čoveka samo „automatski“ reaguje (kao kod nekih refleksa koji se ne mogu kontrolisati) ili kada se telo jednog čoveka koristi od strane drugog čoveka kao puki fizički objekt (A baci B na C).⁴⁹ Ovo poslednje je slučaj kod jednostavnih mašina kojima upravlja čovek poput usisivača ili konvencionalnih automobila, ali ne i kod inteligentnih agenata, koji se shodno sposobnosti učenja, mogu u svakom trenutku otrgnuti kontroli čoveka; oni su u pogledu kauzalne teorije radnje sposobni da sami donose „samovoljne“ odluke. Za razliku od nekada dominantne kauzalne teorije radnje, zastupnici finalnog pojma radnje kao centralni kriterijum ljudske radnje određuju svrsishodnost.⁵⁰ Prema njihovom shvatanju ljudska radnja se od ostalih događaja koji mogu biti kauzalni nastaku štete, razlikuje po tome što su determinisani voljom čoveka. Samo onda kada i zato što je ljudska radnja izraz ciljno usmerene volje delovanja, može biti krivičnopravno relevantna.⁵¹ Kako onda stoje stvari sa poslovnom sposobnošću inteligentnih agenata? Na osnovu sposobnosti da uče, oni su u mogućnosti da samostalno odrede optimalne međuko-
rake, koji u datim uslovima vode uspehu njihove misije. Tako – na jednom vrlo jednostavnom stupnju – internet-pretraživač može na osnovu određenog broja upita sam da formuliše spoj upotrebljenih, traženih pojmova koji su predmet upita ili – u jednoj daleko kompleksnijoj ravni – naoružani dron može da identifikuje i napadne

46 Vidi gore II 1.

47 O teorijama radnje vidi pregled kod Jescheck/Weigend, Lehrbuch des Strafrechts, Allgemeiner Teil, 5. Aufl. 1996, p. 217ff.; Puppe, in: Nomos Kommentar StGB, 4. Aufl. 2013, vor § 13 Rdn. 31ff.; Roxin, Strafrecht Allgemeiner Teil, Band I, 4. Aufl. 2006, p. 236 ff.

48 Problematično je postojanje „radnje“ u konvencionalnom smislu svakako kod agenata koji operišu čisto „intelektualno“, poput internet-pretraživača. Ovde bi moralo da se pojam radnje veoma „obesteli“ kako bi se njegovo ponašanje još uvek moglo označiti „radnjom“.

49 Vidi Baumann/Weber/Mitsch, Strafrecht Allgemeiner Teil, 11. Aufl. 2003, p. 207–211.

50 Welzel, Das deutsche Strafrecht, 11. Aufl. 1969, p. 33 ff.; ders., JuS 1966, p. 423; s današnjeg stanovišta Stratenwerth/Kuhlen, Strafrecht Allgemeiner Teil, 6. Aufl. 2011, p. 57f.

51 Stratenwerth, Schweizerisches Strafrecht. Allgemeiner Teil, Band I, 4. Aufl. 2011, p. 128f.

neprijateljske objekte. Zapravo, ključno pitanje (spoljašnje) ocene jeste da li se odluka inteligentnog agenta shvata kao odluka na osnovu „choises“ (izbora) koja mu je njegovim programiranjem već zadata, ili je reč o „judgements“ (prosudivanjima), dakle „sopstvenim“ vrednovanjima. Sve i kada se ovo poslednje prihvati, to još uvek ne znači da inteligentni agent može „voljno“ postaviti sebi ciljeve i da na osnovu tako postavljenih ciljeva odredi svoje radnje. To bi u najmanju ruku podrazumevalo da inteligentni agent „prepoznaje“ šta čini, i da zna na osnovu toga da je njegovo ponašanje od „društvenog uticaja“, dakle da može da utiče na život drugih ljudi. Da li inteligentni agenti imaju takvu vrstu samosvesti ili je mogu dosegnuti, predmet je rasprave o takozvanoj veštačkoj inteligenciji; o tome zasad još uvek nema pozitivnih odgovora.⁵² Mi danas (još uvek) ne pretpostavljamo, da inteligentni agenti znaju šta čine, i da percipiraju relevantnost svojih akcija – izvan konstatacije da li se približavaju programiranom cilju.

Najzad, pitanje o (krivičnopravnoj) poslovnoj sposobnosti inteligentnih agenata moglo bi biti pitanje definicije: kod kauzalističkog, čisto spoljašnjeg posmatranja, koje svaki „samovoljni telesni pokret“ definiše kao radnju, oni se mogu okarakterisati kao oni koji delaju. Što se jače pojam radnje substancionalno ispunjava, što se više učitava samosvesna usmerenost ka cilju, utoliko manje mogu inteligentni agenti da udovolje zahtevima poslovne sposobnosti.

b) Krivična sposobnost inteligentnih agenata

Kada se pretpostavi da inteligentni agenti preduzimaju radnje u krivičnopravnom smislu, dolazi se pitanjem njihove moguće krivične odgovornosti do sledeće prepone: mogu li inteligentni agenti biti vini? Može li se internet-pretraživaču prebaciti da on prilikom unosa određenog imena pokazuje lanac pojmova koji štete ugledu, i da li se inteligentnom skalpelu može prebaciti odstranjivanje zdravog tkiva, umesto bolesnog? U građanskom pravu se predlaže da inteligentni agenti budu tretirani kao subjekt odgovornosti sui generis – da se iz unapred formiranog fonda odgovornosti – isplaćuju zahtevi za štetu koju oni načine.⁵³ Čak se i u krivičnom pravu razmatra mogućnost pripisivanja krivične sposobnosti inteligentnim agentima, s obzirom na njihovu sposobnost učenja i pamćenja.⁵⁴ Shodno ustavnopravno zasnovanom principu krivice,⁵⁵ krivična odgovornost podrazumeva ličnu odgovornost učinioca za sopstveno delo u tom smislu, onda kada ju je on pri dovoljnom i njemu primerenom naporu mogao da izbegne; i da samo tada sme da mu se prebaci za protivzakonito ponašanje.⁵⁶ Savezni vrhovni sud je 1952. to iskazao sledećom klasičnom formulacijom: „Unutrašnji razlog prebacivanja krivice leži u činjenici da

52 Vidi o tome Hofstadter (nap. 15), pp. 179ff., 195ff.

53 Gruber (nap. 22), p. 150ff.

54 O pretpostavkama nastanka pravnog subjekta uporedi na primer Beck/Zabel, in: Friedrich/Zichy (Hrsg.), *Persönlichkeit- Neurowissenschaftliche und neurophilosophische Fragestellungen*, 2014, p. 49; Hildebrandt, in: Duff/Green, *Philosophical Foundations of Criminal Law*, Oxford 2014, p. 510.

55 Vidi BVerfGE 45, 187, 227; 95, 96, 140; naročito upečatljivo BVerfGE 123, 267, 413, gde se princip krivice zbog zasnovanosti u ljudskom dostojanstvu označava kao neraspoloživ u smislu čl. 79 st. 3 GG.

56 Uopšteno o principu krivice Frister, *Strafrecht Allgemeiner Teil*, 6. Aufl. 2013, p. 30ff.; Kühl, *Strafrecht Allgemeiner Teil*, 7. Aufl. 2012, p. 364ff.; Roxin, *Allg. Teil I* (nap. 47), p. 851ff.; Stratenwerth/Kuhlen, *Allg. Teil* (nap. 50), p. 159ff.

čoveka čini slobodno, odgovorno i moralno samoopredeljenje i da je usled toga sposoban da se opredeli za pravedno, a protiv nepravednog, da svoje ponašanje ustroji prema normama pravom određenog moranja i da izbegava pravom zabranjeno, čim dosegne moralnu zrelost (...).⁵⁷

Ova definicija ostavlja utisak da kategorički isključuje prihvatanje krivičnopravne krivice još toliko inteligentne mašine. Jer, inteligentni agenti zasad svakako ne predstavljaju osobe sazdane na moralnom saopredeljenju, već računarski potpomognute tehnologije, koji njima postavljene zadatke obavljaju shodno onome kako su programirani. S toga ne može da im se pripíše nikakva krivica u smislu koji proklamuje Savezni vrhovni sud. Međutim, sada još (ili opet) relativno nekontroverzna norma iz 1952. godine o slobodi volje ljudi kao važne osnove tradicionalnog pojma krivice – je u novije vreme prilično dovedena u pitanje, u svetlosti određenih saznanja o proučavanju mozga.⁵⁸ Šta pri aktuelnoj kontroverzi preostaje od idealističke predstave voljno slobodnog čoveka, još nije konačno razjašnjeno.⁵⁹ Ipak, to svakako – suprotno prvom utisku – za diskusiju o mogućoj krivičnopravnoj krivici inteligentnih agenata takođe nije od merodavnog značaja. Jer, još je pre diskusije oko rezultata novijih neurobioloških istraživanja u krivično pravo prodrlo saznanje da kod „shvatanja“ krivičnopravne krivice nije reč o istraživanju neuroloških povezanosti i tokova, već o očekivanjima i pripisivanju u aktuelnim društvenim odnosima.⁶⁰

Iz jednog tako shvaćenog principa krivice, odvojenog od neuronalnih tokova, proizlaze svakako zamisliva polazišta za prihvatanje „krivice“ inteligentnih agenata.⁶¹ Iz funkcionalne perspektive krivica se pripisuje onome koji se identifikuje kao odgovorni vinovnik društvenog konflikta koji se dogodio u okviru krivičnog dela.⁶² Time i ideja o „slobodnoj volji“ postaje generički odrediva: „sloboda samoodlučnog delanja“ nije nikakav prirodni fenomen, već proizvod pripisivanja koje služi određenim društvenim ciljevima, naime organizaciji društva pod premisom uzajamnog prihvatanja slobode.⁶³ Mada krivica i na njoj zasnovana pravna odgovornost u uslovima savremenog društva ne može biti pripisivana po nahodanju; u suprotnom bi

57 BGHSt. 2, 194, 200.

58 Vidi npr. Roth/Pauen, Freiheit, Schuld und Verantwortung, 2008; Singer, in: Geyer (Hrsg.), Hirnforschung und Willensfreiheit, 2004, p. 30ff.

59 Iz obimne literature o ovoj temi vidi samo: Duttge, in: ders. (Hrsg.), Das Ich und sein Gehirn, 2009, 2009, p. 13; Hillenkamp, JZ 2005, p. 313; Jäger, GA 2013, p. 3; G. Merkel, Festschrift für Herzberg, 2008, p. 3; R. Merkel, Willensfreiheit und rechtliche Schuld, 2008; Schiemann, ZJS 2012, p. 774; Streng, Festschrift für Jakobs, 2007, p. 675; Weißer, GA 2013, p. 26.

60 Frister, Allg. Teil (nap. 56), p. 32f.; Roxin (nap. 47), p. 869ff.; iz drugog ugla takođe i Jakobs, Strafrecht Allgemeiner Teil, 2. Aufl. 1991, p. 484f. Vidi već Kant, Kritik der reinen Vernunft, Werke in sechs Bänden, BD. 2 (1956), p. 501: „Otuda pravi moralitet radnji (zasluga i greh), pa čak i moralitet našega sopstvenoga ponašanja ostaju za nas potpuno skriveni. Naše osude mogu se odnositi samo na empirički karakter. Koliko je, pak, od toga čista posledica slobode, a koliko se ima pripisati prostoj prirodi i nenamernoj mani temperamenta ili njegovoj srećnoj osobini (merito fortunae) to niko ne može da objasni, pa zbog toga ni potpuno pravedno da sudi.“ Kritika čistog uma. Preveo Nikola Popović. Beograd, BIGZ, 1976, str. 334. – prim. prev.

61 Vidi o ovome takođe Gruber (nap. 44), p. 157; isti. (nap. 21), p. 150; Teubner, Zeitschrift für Rechtssoziologie 27 (2006), p. 15ff.

62 Jakobs (nap. 60), p. 471ff.

63 Frister (nap. 56), p. 32f.

društvena funkcionalnost krivičnog prava bila stavljena na kocku. Minimalni uslovi za to da nekome kao osobi pripišemo odgovornost za njeno ponašanje, ponovo dospevamo u gore već razmatrani uslov mogućnosti samorefleksije.⁶⁴ Ko nije u stanju da vrednuje sebe i svoje prethodne odluke prema etički referentnom sistemu, narodski rečeno da „razlikuje dobro od zla“⁶⁵, taj ne može biti tretiran kao subjekt prava, kao objekt pripisivanja odgovornosti i krivice, jer s njim nije moguće voditi etički dijalog, on nije u stanju da pruži „odgovor“ na etičko prebacivanje. Takav agent – poput divlje životinje ili čoveka koji pati od sumanutih predstava – može biti kontrolisan samo uz pomoć mera za odbranu od opasnosti, ali ne i krivično-pravno kažnjen.

Ipak, već danas inteligentni agenti „odlučuju“ na osnovi koja prevazilazi vrednovanje sopstvenih ranijih načina delanja prikazanih u jednostavnoj podeli na uspeh ili neuspeh. Što su inteligentnim agentima povereni teži zadaci, utoliko složenija mora biti njihova struktura učenja; ovo važi, na primer, za autonomni automobil koji za dolazak do cilja stoji pred odlukom, da li da koristi magistralni put na kome dolazi do stvaranja gužve, ili pak da prođe kroz ulicu zatvorenu za saobraćaj, na kojoj se igraju deca. Za pretpostaviti je da će inteligentni agenti ubuduće u većem obimu morati da obraćaju pažnju na etičke aspekte prilikom donošenja odluka. Ovo važi, na primer, za autonomni automobil koji izvodi nužni manevar izbegavanja, prilikom koga može doći do povrede pravnih dobara trećeg lica;⁶⁶ etičke dileme mogu nastupiti i prilikom korišćenja civilnih dronova za isporuku pošiljki sa robom poručenih putem interneta.⁶⁷ Ukoliko bi takvi (sada već konkretni) planovi bili sprovedeni, dronovi-isporučioc bi morali da prilikom izvršenja svojih misija redovno biraju između više mogućnosti delanja i da se pri tome orijentišu prema (uprogramiranom) bonus-malus-obrascu, koji podrazumeva pridržavanje normativnih pravila. Kako bi trebalo da postupi jedan dron koji, osim ispunjenja osnovnog zadatka (na primer isporuke robe naručiocu do 17 časova), istovremeno dobije zadatak da svoju misiju izvede na društveno prihvatljiv način, recimo da na svom putu vodi računa o zaštitama od buke, i da poštuje privatnu sferu ljudi? U ovu svrhu morala bi mu se zacrtati detaljna pravila postupanja, na primer da ne preleće prenisko bolnice ili penthaus stanove. Za pridržavanje ovih pravila, dron bi bio nagrađivan bonus poenima, a u slučaju kršenja pravila bili bi mu zapisivani malus bodovi, kvantifikovani prema težini prekršaja.⁶⁸ Iz bonus-malus bilansa koji bi se izvlačio nakon izvršavanja misije, mogao bi dron da „uči“ kako ubuduće da ispuni svoj cilj uz što povoljniji bilans bodova.⁶⁹ Kod takvog sistema postavlja se pitanje da

64 Vidi takođe Beck/Zabel (nap. 54), p. 62ff.

65 Vidi Seelmann in: Heer u.a. (nap. 37), p. 575.

66 Pred tom pozadinom razvija se žestoka diskusija na internetu o tome kako bi inteligentni agent mogao da reši „Trolley Problem“ ili „Fat Man Problem“ (u nemačkoj diskusiji poznatiji kao slučajevi „skretničara“ (Fall Weichensteller), vidi: <http://marginalrevolution.com/marginalrevolution/2012/06/the-google-trolley-problem.html> (posećeno 18.07.2014).

67 Vidi o ovome: <http://www.zeit.de/wirtschaft/2013-12/amazon-will-ware-per-drohne-liefern> (posećeno 18.07.2014).

68 Pfüzenreuter, *Intelligentes Missionsmanagement für autonome mobile Systeme*, 2005.

69 Tako bi dron koji se odluči da radi što brže isporuke robe, preleti dve bolnice, primio povratnu poruku, da je za povredu pravila dobio veliki broj malus-bodova; to bi trebalo da utiče da pri

li kršenje pravila vodi ne samo navikavanju na buduća ponašanja inteligentnih agenata, već i da li bi dronu mogla biti pripisana krivica, kada normativno „pogrešno“ postupi, dakle na primer kada bi iznova birao rutu koja je označena malus bodovima, koja vodi iznad nekoliko bolnica.

Moglo bi se zasigurno zamisliti takvo pripisivanje krivice – ali, da li bi to imalo smisla za prevazilaženje društvenih konflikata interesa između što je moguće brže dostave robe i potrebe pacijenata za tišinom. Na ovo pitanje moglo bi se potvrdno odgovoriti samo onda kada bi se pretpostavilo da jedan složeni bonus-malus sistem, namenjen inteligentnim agentima funkcionalno odgovara etičkom samovrednovanju i samoupravljanju čoveka. Jedino pod ovom pretpostavkom može biti smisleno prebacivanje krivice inteligentnim agentima. Ukoliko se to želi postići, uslovi za krivicu bi se morali prilagoditi sposobnosti spoznaje protivpravnog ponašanja inteligentnih agenata. Tada bi moralo biti dovoljno da (ljudski ili inteligentni) agent bude sposoban da prepozna nepoželjnost određenog ponašanja i da bude u stanju da takvo ponašanje „interno“ oceni kao negativno.⁷⁰ Tim razmatranjima se od početka ovog milenijuma bavi rastuća oblast *Machine Ethics*. Ovde se pre svega radi o tome da li i u kom slučaju je moguće robotu usaditi etički zasnovan sistem delanja.⁷¹ No, zasad su ovi projekti tek u povoju, s obzirom da formalizacija etičkih odluka uopšteno za sobom povlači velike poteškoće, tako da programiranje takvih odluka za robote deluje kao daleka budućnost.⁷²

c) Kažnjavanje inteligentnih agenata

Ukoliko bi se isto tako pošlo od toga da se inteligentnom agentu – barem u skorijoj budućnosti – može pripisati krivičnopravna odgovornost, postavlja se, pre uvođenja krivičnog zakonika za robote, još jedno pitanje, da li inteligentni agenti mogu da budu kažnjeni. Naše krivično pravo je – barem zasad – prilagođeno ljudima i time isključuje predstavu da još neko biće osim ljudskog može ili treba da bude kažnjeno za pogrešno ponašanje.⁷³ To načelo je u različitim pravnim porecima utoliko prekršeno, na taj način što je moguće izreći krivične sankcije pravnim licima.⁷⁴ I ove (prevashodno novčane) kazne ne bi na kraju trebalo da pogađaju „pravno lice“ kao apstraktnu pojavu, već ona fizička lica koja

sledećem pokušaju izabere drugu rutu, koja bi možda bila nešto duža, ali bi izbegao ometanje bolesnika bukom.

70 O uporednoj problematici kažnjivosti pravnih lica Frisch (nap. 45), pp. 394, 353.

71 Wallach/Allen, *Moral Machines: Teaching Robots Right from Wrong*, Oxford 2008; Anderson/Anderson (Hrsg.), *Machine Ethics*, Cambridge 2011; Storrs Hall, *Beyond AI: Creating the Conscience of the Machine*, 2007; Georges, *Digital Soul: Intelligent Machines and Human Values*, Cambridge, MA 2003.

72 Vidi o tome na primer Muehlhauser/Helm, *Intelligence Explosion and Machine Ethics*, 2012, <http://intelligence.org/files/IE-ME.pdf> (posećeno 18.07.2014). O raspravi o mogućem tehničkom ostvarenju vidi Bostrom/Yudkowsky, *Cambridge Handbook of Artificial Intelligence*, Cambridge 2011; Santos-Lang, in: van Rysewyk/Pontier (Hrsg.), *Machine Medical Ethics*, New York 2014, p. 74.

73 Vidi o tome Matthias (nap. 42); Wallach/Allen (nap. 71).

74 Vidi o tome priloge Beale (USA), Hilf (Österreich) und Tricot (Frankreich) in *ZStW* 126 (2014), pp. 27ff., 73ff., 55ff.; o švajcarskom pravu vidi Art. 102 schw. StGB i o tome Stratenwerth (nap. 51), p. 447ff.

se kriju iza pravnog lica, bilo u obliku vlasnika udela ili člana društva.⁷⁵ Mogućnost „kažnjavanja“ inteligentnih agenata kaznama čiji smisao i svrha odgovaraju ljudima, teško se, pak danas jedva može zamisliti. S obzirom da inteligentni agent ne poseduje imovinu (u svakom slučaju on ne može *shvatiti* da je poseduje, sve i kada bi se ona njemu građanskopravno mogla dodeliti),⁷⁶ kazna namenjena agentu morala bi se nametnuti njegovom operateru ili bi pogađala fond odgovornosti, tako da ona svejedno ne bi pogađala inteligentnog agenta; i kakvo „lično“ kažnjavanje, recimo, fizičkim uništenjem i potpuno ponovnim reprogramiranjem inteligentnog agenta, što bi, gledano spolja ličilo na smrtnu kaznu, ponovo ne bi pogađalo agenta na isti način kao što pogađa čoveka – u svakom slučaju dokle god inteligentni agenti nisu obdareni kakvom ljudskom voljom za životom i preživljavanjem. Iz tog razloga otpada mogućnost krivičnogpravnog kažnjavanja inteligentnih agenata pri trenutnom stanju razvoja, jer oni ne mogu razumeti kaznu kao takvu. Čovek može razumeti uništenje robota kao zasluženu kaznu za njegovo pogrešno ponašanje, međutim robot nema to shvatanje o takvoj meri. Ipak, usvajanje (čak i „nesavršene“) krivičnogpravne odgovornosti inteligentnog agenta može biti od pravnog značaja – naime, utoliko što bi odgovornost s agenta prešla na čoveka u pozadini.⁷⁷

4. Međurezultat

Dokle god inteligentni agenti „odlučuju“ samo u okviru svoje programiranosti, nedostaje im – u zavisnosti od pojma radnje na kome su zasnovani – verovatno i mogućnost krivičnogpravno relevantnog „delanja“. Tamo gde je ovo prihvaćeno, nailazi se na osnovno pitanje, da li se inteligentnim agentima sme pripisati krivica. To nije upitno samo onda kada oni sprovode proračunat ili barem proračunljiv proces donošenja odluke, već i onda kada podatke dobijene iz okruženja samostalno tumače, osvrćući se na sopstveno iskustvo i time određujući sopstveno delanje. Jer subjekt krivice može prema našem današnjem shvatanju biti samo onaj ko poseduje savest, dakle onaj ko je sposoban da sprovede etičku samorefleksiju. Tako nešto ne deluje moguće s obzirom na današnje stanje razvoja inteligentnih agenata; ipak, barem analogna mogućnost „moralnog“ procenjivanja sopstvenog ponašanja budućih inteligentnih agenata može se itetako zamisliti.⁷⁸ Poslednji korak ležao bi u stvaranju krivično odgovornog inteligentnog agenta, dakle u takvoj konstrukciji, koja bi uz pomoć određenih spoljašnjih promena bila sposobna da asocira na moralno neodobravanje njihovog ponašanja. Tome, ipak, predstoji dugačak put.

75 Pod kojim pretpostavkama je dopustiva, primerena ili čak nužna, takva odgovornost privredno pogođenog za pogrešno ponašanje drugih osoba, jeste teško, u biti nerazjašnjeno pitanje, koje se rado zabašuruje analogijom sa građanskim pravom. Vidi o tome Forster, *Die strafrechtliche Verantwortlichkeit des Unternehmens nach Art. 102 StGB, 2006*, p. 38F; Kremnitzer/Ghanayim, *ZStW* 113 (2001), pp. 538, 542ff.; Seelmann, in: Ackermann/Weigend, *Journal of International Criminal Justice* 6 (2008), pp. 927, 931f.

76 Vidi o ovome i Koops/Hildebrandt/Jaquet-Chiffelle (nap. 36), p. 527ff.

77 O tome detaljnije u III poglavlju.

78 Beck (nap. 36); Hildebrandt, in: Hildebrandt/Gaakeer (nap. 19), p. 25f.

III ČOVEK U SENCI INTELIGENTNOG AGENTA

Pomisao da bi inteligentni agenti mogli prouzrokovati ogromne štete, a da za to nije predviđena krivičnopravna mogućnost kažnjavanja, ostavlja utisak nelagodnosti; a ta nelagodnost raste što više inteligentni agenti kao prividno samostalni akteri učestvuju u našem životu. Sve dok inteligentnim agentima ne može da se pripíše sopstvena krivičnopravna odgovornost, utoliko se više nameće pitanje, pod kojim preduslovima ljudi koji stoje iza mašine mogu i moraju biti pozvani na odgovornost zbog pogrešnih radnji inteligentnih agenata. Da li, primera radi, odgovara onaj koji razvija kompjutersku tehnologiju ili operater internet-pretraživača, kada ovaj ponudi kombinaciju reči kojom se neko kleveta? Podleže li krivičnoj sankciji konstruktor ili prodavac, kada dron otvori vatru na bezopasne civile ili kada automobil bez vozača povredi ili ubije drugog učesnika u saobraćaju?

1) Kažnjivost umišljaja

Ne proističu problemi za krivičnopravnu odgovornost operatera, onda kada ciljno ili voljno tako programira inteligentnog agenta, da ovaj prouzrokuje krivičnopravno relevantne posledice: upotrebi li vojni zapovednik svesno drona da ubije civile ili bude li autonomni automobil umišljajno tako programiran da gazi bicikliste koji voze ulicom, onda bez problema možemo identifikovati zapovednika ili konstruktora automobila-robotu kao odgovorne izvršioce štetnih posledica i da ih učinimo odgovornim za odgovarajuće relevantno biće krivičnog dela, jer su se oni poslužili – svejedno koliko inteligentnim – agentom na jedan način kako se ljudi i inače služe alatom po njihovoj volji.⁷⁹ Radi se o neposredno umišljajnom izvršilaštvu lica, koje po svojoj volji upravlja mašinom.

2. Kažnjivost nehata

Daleko češće mogu nastupiti slučajevi, u kojima upotreba inteligentnog agenta vodi *neželjenoj* šteti: shodno volji svog operatera,⁸⁰ dron bi trebalo da pogodi samo zadati vojni cilj, međutim, kao posledica pogrešne funkcije, on skreće sa svoje rute i ubija civile; autonomni automobil bi trebalo po pravilu da dozvoli prvenstvo prolaza kroz raskrnicu, ali to ne čini i povređuje biciklistu na raskrsnici. U ovim slučajevima u obzir dolazi samo nehatni oblik krivice operatera, kada je relevantnu posledicu krivičnog dela (dakle smrt civila ili povredu bicikliste) mogao predvideti i nije u nužnoj meri pažljivo postupao da spreči nastupanje posledice.⁸¹ Osim toga

79 U ovu kategoriju spada, na primer, i takozvani Google Bomb, koji Google-rezultate pretraživanja manipuliše za određene veb stranice mnogostrukim dodavanjem linkova i tu stranicu na taj način čini neupotrebljivom. Google danas uvek iznova menja algoritme pretraživača, kako bi otežao takve manipulacije.

80 Pojam „operatera“ koristi se u daljem tekstu zbirno za sve osobe koje kao proizvođači, programeri, konstruktori, prodavci ili odgovorni korisnici koji po mogućstvu krivičnopravno odgovaraju za inteligentne agente.

81 Razmišljanja o nehatnoj odgovornosti za inteligentne agente kod Hilgendorf, in: Hilgendorf/Hötitzsch, *Das Recht vor den Herausforderungen der modernen Technik. Beiträge der 1. Würzburger Tagung zum Technikrecht*, 2014, p. 27f.

neophodno je da se nastupanje posledice može uračunati operateru u krivicu. Svaka od ovih tačaka otvara nova pitanja upotrebe inteligentnih agenata.

a) *Predvidivost*

Prilikom proizvodnje i operisanja inteligentnim agentima nepredvidivost nastupanja posledice kao element nehata postavlja probleme.⁸² Jer, inteligentni agenti procenjuju samostalno podatke prikupljene iz svoje okoline i reaguju na rezultate obrade tih podataka bez daljeg uticaja čoveka, kako bi optimalno ispunili postavljeni zadatak. Čak i kada operater stalno nadgleda (re)akcije složenog inteligentnog agenta, nije u stanju da u svakoj pojedinosti predvidi, koji obrazac među prikupljenim podacima robot prepoznaje, kako će ga tumačiti i kako će na njega reagovati. Iz toga dalje sledi, da su u jednom tako otvorenom sistemu zadata pravila odlučivanja, po pravilu manjkava, jer se ne mogu sve životne situacije predvideti niti obraditi algoritmima uputstva delovanja i najzad zato što sam sistem tokom upotrebe samostalno stiče znanja.

Time je jedna određena nepredvidivost inteligentnih agenata i sa njom povezane opasnosti po treća lica “unapred uprogramirane”. Tako se može dogoditi da autonomni automobil jedan blago zaprljan saobraćajni znak koji nalaže prvenstvo prolaza pogrešno “pročita” i tako ne poštujući pravilo prvenstva prolaza bicikliste, njega povredi. Operater autonomnog automobila bi se protiv optužbe za nehatno nanošenje telesne povrede verovatno branio argumentom da pogrešan postupak automobila nije mogao da predvidi, pošto automobil autonomno obrađuje prikupljene podatke. Iz ovog saznanja mogu se izvući dva oprečna zaključka: S jedne strane može se reći da operater autonomno delajućeg agenta *nikada* ne može da odgovara nehatno, jer štetno ponašanje mašine na osnovu njene samostalne obrade informacija nije predvidivo. Ili, se može izvući potpuno suprotan zaključak; naime da oni koji upravljaju inteligentnim agentima moraju računati na “sve”, te je tako svaka vrsta činjenja štete principijelno predvidiva.⁸³

Prva varijanta teško da dolazi u obzir; jer onaj ko u promet pusti inteligentnog agenta nad kojim nema sigurnu kontrolu, ne može da negira svoju odgovornost za njegove nepredviđene štetonosne akcije, jednostavnim ukazivanjem na neuračunljivost inteligentnog agenta – on to isto onako malo može kao i upravnik zoološkog vrta koji pusti tigra na slobodu, i onda prolazniku, koga je tigar napao i ugrizao, objašnjava nepredvidivu prirodu divlje životinje. Ukoliko bi se drugačije odlučilo, nastala bi normativno neprihvatljiva difuzija odgovornosti između čoveka i inteligentnog agenta, koja bi ostavila bez ikakve reakcije krivičnopravno relevantnu povredu slučajne žrtve kao posledicu pogrešnog postupka agenta. Kada bi dakle sve zamislive posledice štete morale da važe za operatera kao predvidive, onda pretpostavka predvidivosti ostvarenja krivičnog dela ne bi predstavljala prepreku sveobuhvatnoj nehatnoj odgovornosti operatera prilikom svakog štetonosnog delovanja inteligentnog agenta.

82 Duttge, in: Münchener Kommentar zum StGB, 2. Aufl. 2011, § 15 Rdn. 107ff.

83 Vidi npr. B. Beck, Journal for Robotics and Autonomous Systems (izlazi 2015).

b) Povreda obaveze pažljivog postupanja

Kažnjivost nehata dalje pretpostavlja pored predvidivosti ostvarenja krivičnog dela i pretpostavku da je učinilac povredio obavezu pažljivog postupanja.⁸⁴ Po pravilu povreda obaveze proističe iz predvidivosti posledice: onaj ko može da prepozna da njegova radnja može da dovede do povrede zaštićenog interesa, taj po pravilu dela protivno pažljivom postupanju, kada i nakon toga preduzme spornu radnju.⁸⁵ Naročito kod proizvodnje i distribucije potencijalno opasnih proizvoda, sudska praksa je razvila izrazito stroga pravila za pažljivo postupanje, koja su u najvećoj meri preuzeta iz građanskog u krivično pravo.⁸⁶ Shodno tome, brižni proizvođač sme da na tržište iznese samo one proizvode čija sigurnost odgovara stanju nauke i tehnike i koji su pre distribucije testovima u dovoljnoj meri provereni. Čak i pošto je proizvođač proizvod pustio u promet, mora ga stalno držati na oku na osnovu povratnih informacija potrošača. Pokažu li se pritom neočekivane štete ili rizici, proizvođač je obavezan da upozori potrošače i u krajnjem slučaju da organizuje povlačenje proizvoda.⁸⁷

Ova načela važe principijelno čak (i upravo) u vezi sa inovativnim proizvodima⁸⁸ poput inteligentnih agenata; ovde su – kao što je i pokazano – rizici posebno visoki i teški za kontrolisati. Preteća krivičnopravna odgovornost mogla bi onda, pak, u svakom slučaju, dovesti dotle da proizvodnja i distribucija (u svakom slučaju: nekih) inteligentnih agenata potpuno izostane, zbog nedostatka njihove upravljivosti. Ova posledica bi predstavljala dobit na sigurnosti, ali sa druge strane gubitak mnogobrojnih inovativnih primena koje bi čoveku mogle bitno da olakšaju život. Kažnjavanje nehatnog oblika ne mora, pak, da znači da bi se moralo odustati od svake potencijalno rizične radnje.⁸⁹ Ako se uzme za smisleno, pa čak i obavezno, to što ljudi koriste usluge inteligentnih agenata, onda ostaje kao izlaz samo redukcija obaveze pažljivog postupanja.⁹⁰ Moralo bi kod njihovog određivanja, rizike koje operater uzrokuje, dovesti u odnos sa korisnošću odgovarajuće tehnologije. Delovalo bi nepošteno kada bi “društvo” koje ima koristi od inovacije, neizbežne rizike koje nju prate, s aspekta krivičnog prava, u potpunosti prevalilo na operatere.

Operateri inteligentnih agenata bi mogli da se po mogućstvu pozivaju, braneći se od prebacivanja za manjkavost pažljivog postupanja, da razvojem i korišćenjem tehnologija koje su sposobne da uče, a time ograničena u kalkulaciji i upravljanju,

84 Vidi o tome Vogel, in: Lepiziger Kommentar, 12. Aufl. 2007, § Rdn. 164ff.

85 Vidi o tome Duttge, in: MK (nap. 82), § 15 Rdn. 121ff.

86 Temeljito BGHSt. 37, 106 („Lederspray“). O odnosu građanskopravne i krivičnopravne odgovornosti za proizvod Gless, recht 2013, p. 57f.; Kuhlen, in: Achenbach/Ransiek, Handbuch Wirtschaftsstrafrecht, 2012, Rdn. 29ff.; isto., JZ 1994, p. 1142; Schlutz, DStR 1994, pp. 1811, 1814; Vogel, in LK (nap. 84), § 15 StGB Rdn. 277–281. O odluci u slučaju „sprej za kožu“ („Lederspray“) vidi Kuhlen, NStZ 1990, p. 566; Hilgendorf, NStZ 1994, p. 561. Načelna kritika kod na primer Hassemmer, Produktverantwortung im modernen Strafrecht, 1994, p. 74.

87 Vidi Kuhlen, in: Achenbach/Ransiek (nap. 86), Rdn. 40ff.; Weigend, in: LK (nap. 84), § 13 Rdn. 53f.

88 Kuhlen, in: Achenbach/Ransiek (nap. 86), Rdn. 34ff.

89 Stratenwerth (nap. 51), p. 171.

90 Vidi o mogućim načinima rešavanja Eisele, in: Schönke/Schröder, StGB, 29. Aufl. 2014, vor § 13 Rdn. 91; Puppe in NK (nap. 47), vor § 13 Rdn. 157ff.; Vogel, in LK (nap. 84), § 15 Rdn. 214 ff.

služe društvenoj potrebi. Time je postavljeno odlučujuće pitanje, u kojoj meri je društvo spremno da zbog povećane opšte koristi zajednice prihvati posebne opasnosti koje inteligentni agenti mogu da predstavljaju za pojedinca (ili i više) ljudi. Odgovor na to je komplikovan, jer upravo rizikonosna samostalna sposobnost odlučivanja inteligentnog agenta, koja isključuje potpunu kontrolu od strane ljudi, ali istovremeno pridonosi povećanju njegove koristi, je opravdana i s toga privredno i društveno poželjna.

Nespornog rešenja nema: S jedne strane zasigurno ne stoji da savremeno društvo, bez ikakve kritike, svekolike razvoje koji vode ka svetu kojim upravljaju mašine i sa njima povezane ogromne rizike – setimo se samo naoružanih dronova koji su izmakli kontroli – bez daljeg prihvata kao cenu napretka. S druge strane, vidi se na javnom unapređenju razvoja inteligentnih agenata da postoji društveni interes za korišćenjem takvih modernih tehnologija. Može se i pretpostaviti da su određeni rizici koji su povezani sa upotrebom inteligentnih agenata već u potpunosti prihvaćeni i odobreni – kao što je nedopuštena upotreba podataka, kada ih internet-pretraživači pohrane. Dokle god je javnost spremna da živi sa rizicima novih tehnologija, morala bi da prihvati i posledice, neočekivane reakcije inteligentnih agenata koji vode (načelno) predvidivom nanošenju štete slučajno pogođenih osoba, a da za to operateri ne budu sankcionisani.⁹¹ Mogu se za ovakav slučaj predvideti osiguranja ili građanskopravna apstraktna odgovornost;⁹² sveobuhvatna krivičnopravna odgovornost operatera za načelno predvidivo nanošenje štete bi morala, doslednosti radi, da bude izostavljena.⁹³

S druge strane, oblast izvan krivičnog prava za stvaranje nepredvidivih rizika po život i zdravlje drugih ne može biti velika.⁹⁴ Jer, društvo je doduše spremno da prihvati određene preostale rizike u ophođenju sa inteligentnim agentima; ono ipak mora da vodi računa da se te opasnosti što je moguće više smanje pažljivom konstrukcijom, programiranjem, temeljnim testovima, i stalnim posmatranjem. Ključno pitanje dakle glasi, koje bi zahteve trebalo postaviti, u cilju što efektnijeg i što dalekosežnijeg suzbijanja rizika u vezi sa inteligentnim agentima.⁹⁵ Pri tome je za utvrđivanje mere potrebnih troškova neophodno sagledati, s jedne strane korist datog inteligentnog agenta, a s druge strane stepen opasnosti koju on može da proizvede.⁹⁶ Naoružani dronovi potrebuju dalekosežne mere opreza i sa jednog i sa drugog aspekta, dok se neočekivane reakcije pri kupovinama na internetu pre prih-

91 Kuhlen, in: Achenbach/Ransiek (nap. 86), Rdn. 27ff.

92 Gless (nap. 86), p. 57.

93 Na kom se mestu u konstrukciji krivičnog dela sprovodi ovo isključenje, drugorazredno je pitanje. U svakom slučaju radi se o ograničenju nepravda datog nehatnog krivičnog dela, a ne individualne krivice operatera. Kao „dogmatske“ tačke povezivanja u obzir dolaze ograničenje obaveze pažljivog postupanja, poricanje „nedozvoljenog“ podeznanja rizika kao pretpostavka (ili) objektivne uračunljivosti posledice (vidi o tome na primer Frisch, GA 2003, pp. 719, 722–724; Puppe, in: NK (nap. 47), vor § 13 StGB Rdn. 153ff., 205ff.) ili „socijalna adekvatnost“ stvaranja rizika koja isključuje biće krivičnog dela ili protivpravnost.

94 Vidi o pojedinim grupama slučajeva Vogel, in: LK (nap. 84), § 15 Rdn. 216–218.

95 Beck (nap. 36) kao i sa građanskopravnog aspekta Spindler, NJW 2004, p. 3146f.

96 Prittwitz, Strafrecht und Risiko, 1993, p. 297ff.; von Bar, in: Lieb (Hrsg.), Produktverantwortung und Risikoakzeptanz, 1998, p. 40ff.

vataju iako one uzrokuju manje imovinske štete korisnicima. U zavisnosti od koristi određeni rizici razvoja i primene inteligentnih agenata bi mogli biti prihvaćeni. Tako, može biti dopustivo, da se nakon dovoljnog broja testova, autonomni automobil i onda pusti u saobraćaj iako je, pak, u retkim pojedinačnim slučajevima, iznenada stao, kada se našao u neplaniranoj situaciji. Kako autonomni automobili za pojedinca i za rešavanje opštih saobraćajnih problema pokazuju veliki potencijal, mogu ovde da važe velikodušnija merila nego kod tehnološki jednostavnijih proizvoda (poput usisivača ili bicikala), nego kod inteligentnih agenata s manjom društvenom dodatom vrednošću (kao što su igračka ili borbeni dron) kod kojih takve mere moraju biti primenjivane.⁹⁷

Ipak, obaveza pažljivog postupanja zahteva prema zaštićenim pravnim dobrima pogođenih ljudi, čak i da se i kod korisnih inteligentnih agenata rizici u najvećem mogućem obimu isključe. Tako postoji i kod prihvaćenog rizika od inovacija, na primer, obimne obaveze upozoravanja i ukazivanja korisnicima kao i striktna obaveza proizvođača da iskustva s proizvodom prati i da odmah reaguje na prijavu nastanka štete.⁹⁸ Ukoliko se na primer pokaže da štetne vrste ponašanja autonomnog automobila ne nastupaju samo u sasvim retkim konstelacijama ili pri gruboj zloupotrebi, tada se automobil ne sme dalje distribuirati, već se u datom slučaju mora povući.⁹⁹ Ostane li operater pasivan, može krivično da odgovara pri nastanku štete za u svakom pojedinačnom slučaju primenjivo krivično delo.¹⁰⁰ Ova obaveza činjenja operatera proističe – kao i obično iz prava odgovornosti za proizvod – iz jedne „mirujuće nadležnosti.“¹⁰¹ Uvođenjem ili upotrebom, po prirodi stvari kod ne u potpunosti kontrolisanih inteligentnih agenata, operater stvara (čak i u sopstvenom privrednom interesu) povišeni rizik po druge, čime se pooštrava njegova obaveza da stalno nadzire proizvod i da snosi posledice iz (naknadno) saznate opasnosti.¹⁰² Ova obaveza predstavlja obavezu obezbeđenja saobraćaja. Ona proističe iz jednog, po sebi, zakonitog postupanja povećanog rizika, a ne kao odgovornost za ingerenciju iz jednog pretponašanja protivno obavezama i zakonu.¹⁰³

c) Objektivno uračunavanje

Kada operater inteligentnih agenata pod objašnjenim okolnostima prekrši obavezu pažljivog postupanja postavlja se još i pitanje, da li se njemu može (objektivno) uračunati posledica koja je nastupila krivičnim delom – ili, nije li ipak, “kriva mašina”.

97 Uporedi npr. razmatranja o zakonodavstvu SAD-a o autonomnim automobilima: Autonomous Vehicles in California; Statement of Policy on Automated Vehicles, U.S. Department of Transportation, <http://www.dmv.ca.gov/vr/autonomous/auto.htm> (posećeno 18.07.2014).

98 Kuhlen, in : Achenbach/Ransiek (nap. 86), Rdn. 34ff.

99 Gless (nap. 86), p. 59ff.

100 Vidi o razmatranjima o prenosu odgovornosti BGHSt. 37, 106, 115; o tome Schmidt-Salzer, Produkthaftung, Bd. I, 2. Aufl. 1998, Rdn. 1023ff.

101 Gless (nap. 86), p. 60f.; vgl. takode Hoyer, GA 1996, p. 175 f.; Schmidt-Salzer, NJW 1996, p. 6.

102 Za Švajcarsku vidi BGE 134 IV 255, 260 f.; za Nemačku Arzt, JA 1980, pp. 712, 714; Freund, Erfolgsdelikt und Unterlassen, 1990, p. 199ff.; Jakobs (nap. 60), p. 812f.; Kuhlen, NStZ 1990, pp. 566, 568; Maiwald, JuS 1981, p. 482.

103 Donatsch/Tag, Strafrecht I, 8. Aufl. 2006, p. 308; Cassani, in: CR, Art. 11 Nr. 44; Stratenwerth (nap. 51), p. 467f.

aa) Inteligentni agenti kao opšti životni rizik

Uračunavanje nastale posledice zbog ponašanja (protivnog obavezi pažljivog postupanja) operatera moglo bi za početak biti isključeno, s obzirom na postojanje „zabrane regresa“, kada je neposredni izvršilac (tj. inteligentni agent) sam krivičnopravno odgovoran za delo s predumišljajem.¹⁰⁴ Tako opisana opšta „zabrana regresa“ danas se, pak, pretežno odbija,¹⁰⁵ jer načelno je moguće da dve osobe nezavisno jedna od druge moraju da odgovaraju za jednu te istu posledicu, onda kada je ta posledica svakome od njih objektivno uračunljiva.¹⁰⁶

Više uspeha obećava stav, da se uračunavanje posledice koju izazove inteligentni agent operateru odbije, uz obrazloženje da nastupanje posledice (npr. nanošenje telesne povrede od strane autonomnog automobila) jeste „prirodan“ događaj kojim čovek nije upravljao, u kome se samo manifestovao opšti životni rizik.¹⁰⁷ Ovo bi ipak pretpostavljalo da inteligentni agenti budu prihvaćeni kao deo svakodnevnog života, sa čijim rizicima je svako upoznat i prema kome se može postaviti, poredivo na primer sa rizikom od grana koje padaju tokom šetnje šumom po oluji. U sadašnjosti i bliskoj budućnosti mogli bi potencijalno štetonosni inteligentni agenti, pak, da budu još uvek izuzetak, tako da onaj ko ih stavlja u promet, stvara *neobičan* rizik, i stoga ima da odgovara za njihove (opšte nepredvidive) učinjene posledice u okviru kažnjivosti za nehat. Drugačije je, pak, već danas kod softverskih agenata na internetu, koji su tamo sveprisutni kao internet-pretraživači ili *recommendation system*-i ili u kakvom drugom obliku. Bez obzira na to, što smo svesni opasnosti povrede naših prava ličnosti, mi i dalje poručujemo robu kod prodavaca na internetu – i prihvatamo kao opšti životni rizik mogućnost da sistem naše lične podatke obradi i primeni, na jedan način koji nije transparentan i na nekontrolisani način zadire u sferu naše privatnosti.¹⁰⁸ Mnogo toga govori u prilog da operater softverskog agenta na internetu, koji druge ljude izlaže „normalnim“ rizicima po zaštitu ličnosti aktualizuje algoritme koji su povezani sa internet-pretraživačima, a koji su za pogođene

104 O učenju o zabrani regresa, opširno Puppe, in: NK (nap. 47), vor § 13 Rdn. 167 ff. sa mnogobrojnim primerima; vidi takode Caro John, Das erlaubte Kausieren verbotener Taten- Regressverbot, 2007, p. 55ff.

105 Up. Eisele, in: Schönke/Schröder (nap. 90) vor § 13 Rdn. 77, 100ff.; Frisch, Tatbestandsmäßiges Verhalten und Zurechnung des Erfolges, 1988, p. 62f.; Puppe, in: NK (nap. 47), vor § 13 StGB Rdn. 236–238.

106 Tako može krivično odgovarati za ubistvo iz nehata onaj koji, vatreno oružje, suprotno propisu neobezbeđeno drugome učini dostupnim, na šta ta osoba – predvidljivo za vlasnika oružja – s predumišljajem ubije nekoga tim oružjem; vidi BGH StraFo 2012, 173 sa nap. Berster, ZIS 2012, 623 (Fall Winnenden). Kritička tačka u takvim slučajevima je pitanje da li je povezanost uračunljivosti između nedostajuće obezbeđenosti oružja i smrti žrtve prekinuta „autonomnim“ uplitanjem onog koji puca. Odgovor na ovo pitanje zavisi od toga, da li je rizik koji je stvorio proučini-lac preusmeren od (ukoliko, pak, u navedenom primeru drugoučini-lac pištolj koristi kao smrtonosnu alatku za udaranje) ili, da li se smrtonosnim pucnjem drugoučini-laca tačno ostvaruje rizik koji je trebalo da bude sprečen pažljivim obezbeđenjem oružja od neovlašćene upotrebe.

107 Vidi Puppe, in: NK (nap. 47), vor § 13 Rdn. 236 ff.; Walter, in: LK (nap. 84), vor § 13 Rdn. 103 ff.

108 O ovoj problematici na primer: Ramakrischan/Keller/Mirza, Privacy risks in Recommender Systems dostupno na: <http://people.cs.vt.edu/naren/papers/ppp.pdf> (posećeno 18.07.2014); podrobnije o pitanju zaštite privatne sfere u pogledu širenja inteligentnih Nadakavukaren Schefer, SchwZStr 2014, p. 259.

već danas samo jedan opšti životni rizik. Dobro bi bilo da takva procena inteligentnog agenata kao „normalnog“ interakcionog partnera ljudi, u budućnosti ne ostane ograničena na relativno jednostavne mašine na internetu, već da obuhvati i druge oblasti života. Čim autonomni podzemni vozovi i automobili bez vozača postanu samorazumljiva slika šinskog i putnog saobraćaja, susret sa takvim vozilima za ljude će predstavljati „normalan rizik“ i operater će tada (krivičnopravno) odgovarati ne za svaku načelno predvidivu pogrešnu funkciju, već samo za štetne posledice otklonjivih grešaka u konstrukciji i programiranju.

bb) Inteligentni agenti kao autonomno delajuća treća lica

Druga varijanta isključivanja objektivne odgovornosti postojala bi onda kada „sopstvena“ odluka autonomno delajućeg robota prekida povezanost uračunljivosti između delanja operatera i nastupanja krivičnopravno relevantne posledice.¹⁰⁹ Ovo se ne čini tako dalekim, i dešava se onda kada primera radi autonomni automobil nanese telesnu povredu na način da se to ne može smatrati delom konstruktora već „samog automobila“. Već smo gore, pak (II), zaključili da sankcionisanje inteligentnih agenata „krivičnim pravom za ljude“ za sada ne dolazi u obzir. Ali, ta okolnost ne znači nužno da upotreba autonomno delajućih tehnologija ne bi bila irelevantna za uračunavanje posledice čoveku koji stoji iza mašine. Kada se akcija inteligentnog agenta može okvalifikovati kao radnja „isturenog čoveka“, onda se i povezanost uračunljivosti sa „čovekom u pozadini“ (operaterom) može prekinuti. No, o autonomnom štetnom delovanju inteligentnih agenata ne može biti reči, dokle god inteligentnom agentu ne pripisujemo sposobnost sopstvenog stvaranja volje. Ovo za sada nije moguće. Gore smo već naznačili da takav vid pripisivanja kod inteligentnih agenata sa sposobnošću učenja i pamćenja, u buduću ne može biti isključena.¹¹⁰ Međutim, kada se kod razvoja antropomorfno bića dosegla ta tačka, trebalo bi primenjivati krivičnopravni regres za nehatno delajućeg operatera, dokle god inteligentni agenti sami ne postanu subjekt krivičnog prava; u suprotnom bi žrtvu štetonosnog dela, naime, u gore već napomenutoj nezadovoljavajućoj situaciji, pogodila difuzija odgovornosti¹¹¹ između čoveka i inteligentnog agenta, koja bi isključivala krivičnopravno kažnjavanje čak i najtežih oblika pričinjavanja štete.

IV ZAKLJUČAK

Prodor manje ili više blagodatnih inteligentnih agenata u naš život je jedan od onih fenomena, koji pravni poredak, a posebno krivično pravo postavlja pred praktična, ali isto tako i principijelna pitanja, na koja se dosadašnjim učenjima ne može stvarno odgovoriti. Na posletku, postojanje inteligentnih agenata otvara pitanje, šta je zapravo ljudsko u ljudima – pitanje, koje je od posebnog značaja za krivično pravo kao pravnu oblast koja je najličnije namenjena ljudima.

¹⁰⁹ Podrobnije kod Roxin, Allg. Teil I, p. 401ff.

¹¹⁰ Vidi o tome naslove u napomeni 44.

¹¹¹ Beck (nap. 38).

Krivičnopravnoj teoriji računanja nedostaje zasada adekvatan instrumentarizam kako bi u sadejstvu između (po mogućstvu manje inteligentnog) čoveka i (po mogućstvu veoma inteligentne) mašine mogla da odlučuje o stručno-adekvatnom pripisivanju odgovornosti. Tako bi uvođenje kažnjivosti inteligentnih agenata u doglednoj budućnosti, moglo da ostane samo jedan misaoni model. Inteligentni agenti ne raspolažu niti za čoveka shvatljivim osećajem za kaznu, niti razumevanjem sa kažnjavanjem povezanim etičkim prekorom, tako da primena krivičnog prava protiv njih "lično" ne bi imala smisla. Drugačije nego što je to kod kažnjivosti pravnih lica, kažnjavanje inteligentnih agenata ne bi se moglo tumačiti kao posredno kažnjavanje fizičkih lica koja stoje u pozadini zbog njihovih "organizacijskog sagrašenja". U svakom slučaju bi svako individualno sagrašenje operatera inteligentnih agenata moglo da se podjednako delotvorno sankcioniše neposrednim kažnjavanjem operatera zbog njihovih umišljajnih ili nehatnih kauzalnih doprinosa šteti prouzrokovanoj od strane agenata. Zamislivo je doduše da se konstruiše inteligentni agent kao samostalni subjekt građanskopravne odgovornosti (i da se u ovu svrhu opremi obaveznim osiguranjem od odgovornosti ili fondom odgovornosti),¹¹² i takođe bi se moglo razmatrati nalaganje uništenja ili reprogramiranja inteligentnog agenta, koji se otrgao kontroli, kao krivičnopravne sankcije uperene protiv njegovog vlasnika, ali jedna stvarna krivičnopravna odgovornost samog agenta ne dolazi u obzir, dokle god inteligentni agent ne postane i moralni agent.

Teže je dati odgovor na pitanje kažnjivosti "čoveka koji stoji iza mašine". Prouzrokuje li radnja inteligentnog agenta krivičnopravno relevantnu štetu, u svakom slučaju odgovara čovek kao neposredni izvršilac onda kada je agenta s namerom ili u znanju upotrebio, da uz pomoć njega ostvari ili pokuša da izvrši krivično delo. Jer čovek ne može prebaciti krivicu na mašinu onda kada je ciljno upotrebi kao alatku ili kada u svakom slučaju zna i prihvata da inteligentni agent u konkretno predviđenom zadatku može da ostvari određenu krivičnopravno relevantnu posledicu.

Čak i kada radnja inteligentnog agenta bez znanja i volje operatera vodi objektivnom ostvarenju bića krivičnog dela, ne može se krivica operatera poricati lapidarnim upućivanjem na "automatizam" algoritama. Prst uperen na automatizam, propisan matematičkim uputstvima rukovanja pokazuje u krajnoj liniji natrag na čoveka, koji je ponašanje inteligentnog agenta programirao, a time i mogućnost nastanka štete, ili distribucijom agenta. U slučaju pukog nehata, izranjaju mogući problemi na polju predvidivosti i računanju posledice. Složeni inteligentni agenti u postupku donošenja odluka nisu u potpunosti računljivi, te se predvidivost njihovog pogrešnog delanja po operatera može dovesti u pitanje. U određenoj meri na njihovo ponašanje bi se moglo gledati kao na samoodredivo. Ovo bi krivičnopravnom pribegavanju ka nehatu operatera moglo predstavljati prepreku, onda kada ponašanje inteligentnog agenta shodno opštim pravilima o računanju posledice, prekida povezanost računanja između ponašanja operatera i nastupanja posledice. To je, na primer, onda zamislivo kada je angažovanost inteligentnih agenata u određenoj sferi života do te mere postalo normalnost, da čak ulazi u opšti životni rizik pogođenog.

¹¹² Vidi Gruber (nap. 44), p. 157; isti, in: Beck (nap. 21), p. 150 ff.

Potpuno je drugačije pitanje, da li će jedan mali (“preostali”) rizik, koji proističe iz sposobnosti učenja inteligentnog agenta, biti voljno prihvaćen iz razloga velike društvene koristi takvog agenta. Ovo pitanje se postavlja već u ravni nehatne protivpravnosti, naime, kao moguće ograničenje obaveze pažljivog postupanja operatera. Upravo ova problematika bi mogla da, u prvom redu, odredi polemiku o krivično-pravnoj odgovornosti inteligentnih agenata u bliskoj budućnosti. Nad inteligentnim agentima čovek – za razliku od drugih proizvoda i otkrića – nema potpunu kontrolu; on ne može da u potpunosti planira i vlada sa njima povezanim rizicima. Što su roboti svestraniji, sposobni da uče, i inteligentiji, utoliko se više odupiru dalekosežnom planiranju svojih operatera i utoliko je teže identifikovati pojedine greške u konstrukciji i programiranju, koje bi mogle da obrazlože kažnjivost nehata ljudskih učesnika. Ne dođe li do saglasnosti po pitanju potpune zabrane inteligentnih agenata zbog malopre nabrojanih neovladivih rizika, onda verovatno ne postoji neka druga mogućnost, nego da se „društvu“ nametnu opasnosti proistekle nepouzdanim programiranjem i neodgovornim upravljanjem tokom angažmana, pri čemu se odustaje od krivičnopravne odgovornosti za nehat i time oštećeno lice definiše kao žrtva ne-čovekovog ponašanja. Pri tome nije neophodno da se zacrta jedno opšte i obuhvatno izuzimanje od odgovornosti ljudi koji stoje iza mašine; štaviše nudi se povlačenje krivičnog prava samo u onoj meri, u kojoj opšta društvena korist od inteligentnih agenata opravdava prihvatanje sa njima povezanih rizika po druge.

Pravnometodološki do ovog rešenja se može doći uvođenjem fleksibilnije generalne klauzule, poput sužavanja potrebnog pažljivog postupanja operatera odobravanjem “razvojnog rizika” ili stvaranja osnova za opravdanje “društvenoadekvatne upotrebe” inteligentnih agenata. U pozitivnopravnom smislu, bilo bi dobro dati prednost mukotrpnijem putu, da se zakonom ili uredbom odrede merila za dopušten rizik različitih tipova i područja upotrebe inteligentnih agenata. Time bi se u legitimnom demokratskom postupku stvorila primerena ravnoteža između potrebe da se mogućnosti, samostalno delajućeg čovekovog pomoćnika, koje obećavaju, angažuju bez straha od nepredivih krivičnopravnih posledica, i stvaranja nužnih podsticaja za što je moguće veća suzbijanja upravljivih rizika u *Vrlom novom svetlu*¹¹³ inteligentnih agenata.

S nemačkog preveo Lazar Glišović

INTELLIGENTE AGENTEN UND DAS STRAFRECHT

Zusammenfassung

Das Vordringen mehr oder weniger segensreicher Intelligenter Agenten in unser Leben ist eines jener Phänomene, das die Rechtsordnung und speziell das Strafrecht vor praktische, aber auch vor grundsätzliche Fragen stellt, die sich mit den hergebrachten Lehren nicht wirklich beantworten lassen. am Menschen ist – eine Frage, die insbesondere für das Strafrecht als das den Menschen höchstpersönlich adressierende Rechtsgebiet von höchster Relevanz ist. Der strafrechtlichen Zurechnungslehre fehlt bisher das adäquate Instrumentarium, um bei

113 *Huxley*, Brave New World, 1932, vidi: https://archive.org/details/ost-englisch-brave_new_world_aldous_huxley (posećeno 18.07.2014).

dem Zusammenspiel zwischen dem (möglicherweise wenig intelligenten) Menschen und der (möglicherweise sehr intelligenten) Maschine über die sachadäquate Zuweisung von Verantwortung zu entscheiden. Dabei dürfte die Einführung einer Strafbarkeit Intelligenter Agenten für die voraussehbare Zukunft ein bloßes Denkmodell bleiben. Anders als bei einer Strafbarkeit juristischer Personen ließe sich eine Bestrafung Intelligenter Agenten auch nicht als mediatisierte Bestrafung der im Hintergrund stehenden natürlichen Personen für deren „Organisationsverschulden“ interpretieren. Jedenfalls könnte man ein solches individuelles Verschulden der Betreiber des Intelligenen Agenten ebenso wirksam durch die unmittelbare Bestrafung der Betreiber wegen ihres eigenen vorsätzlichen oder fahrlässigen Kausalbeitrags zu der von dem Agenten verursachten Schädigung sanktionieren. Schwerer ist die Frage nach der Strafbarkeit des „Menschen hinter der Maschine“ zu beantworten. Ruft die Aktion eines Intelligenen Agenten einen strafrechtlich relevanten Schaden hervor, so haftet der Mensch jedenfalls dann als unmittelbarer Täter, wenn er den Agenten in der Absicht oder in dem Wissen benutzt hat, dass mit seiner Hilfe ein Straftatbestand verwirklicht wird oder werden kann. lung der Menschen hinter der Maschine festzuschreiben; vielmehr bietet es sich an, das Strafrecht nur in dem Maße zurückzunehmen, wie der allgemeine soziale Nutzen des Intelligenen Agenten die Hinnahme der mit ihm verbundenen Risiken für die Allgemeinheit rechtfertigt. Rechtsmethodisch kann man dieses Ergebnis durch die Einführung flexibler Generalklauseln erreichen, etwa indem man die erforderliche Sorgfalt des Betreibers durch die Gestattung eines „Entwicklungsrisikos“ beschränkt oder indem man einen Rechtfertigungsgrund des „sozialadäquaten Einsatzes“ von Intelligenen Agenten schafft. Im Sinne der Gesetzesbestimmtheit wäre allerdings der mühsamere Weg vorzuziehen, die Maßstäbe für das erlaubte Risiko bei verschiedenen Typen und Einsatzbereichen Intelligenen Agenten durch Gesetz oder Verordnung festzulegen.