

## ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА И ТОТАЛИТАРИЗАМ

### Сажетак

Након увода, први део рада је посвећен одређењу појмова вештачке интелигенције и тоталитаризма, приликом чега је подвучена важност разликовања садашње (машинско учење) и пројектоване (суперинтелигенција) фазе у развоју вештачке интелигенције, односно ембрионалне (тоталитарни покрет ван власти) и етаблиране (тоталитарни покрет на власти) фазе у развоју тоталитаризма. Други део рада испитује везу између садашњег ступња вештачке интелигенције и ембрионалне фазе тоталитаризма, док трећи део рада анализира потенцијални однос суперинтелигенције и етаблираног тоталитаризма. Показало се, с обзиром на сличности и разлике између ефеката савремене и будуће вештачке интелигенције и ефеката ранијег тоталитаризма, да данас (и колико већ сутра) немамо пуку реплику тоталитарних фаза из XX века, већ особене тоталитарне феномене у виду „капиларног тоталитаризма“, односно „хипертоталитаризма“. Прошло-вековни тоталитаризам, као и данашњи у „капиларној“ варијанти, нису (били) нужно иреверзибилни, али ће „хипертоталитаризам“ то бити. У закључку се предлажу заштитне мере од ризика вештачке интелигенције, у виду начела изузећа по узору на приговор савести.

Кључне речи: вештачка интелигенција, машинско учење, суперинтелигенција, тоталитаризам, капиларни тоталитаризам, хипертоталитаризам.

\* Аутор је научни сарадник на Институту за политичке студије у Београду. Електронска адреса аутора: marko.pejkovic@ips.ac.rs

## 1. УВОД

Тоталитаризам је феномен старији од вештачке интелигенције (у даљем тексту ВИ). Он је фундаментално искуство савремене политичке теорије, јер је демонстрирао лимите просветитељског разума и прогреса – терор је ишао руку под руку са прогресивним снагама, а не против њих.<sup>1</sup> Арентова тврди да се непрестано мора водити дијалог о суштини тоталитаризма, ако желимо и даље да ову планету осећамо као „свој дом“.<sup>2</sup> И тај дијалог не би смео првенствено да има модеран научни дискурс, већ античку разборитост (*phronesis*) коју Арентова назива имагинацијом.

Прошло је неколико деценија од тих размишљања и од првих искустава човечанства са тоталитаризмом. Рекло би се да за поменути дијалог нема потребе, ако се мисли да тоталитаризма као претње данас нема на виду.<sup>3</sup> Ипак, Арентова је оставила могућност да класични тоталитаризам, једном у будућности добије мутирани изданак у виду „технократског тоталитаризма“<sup>4</sup> – на врху тоталитарне пирамиде немамо више харизматичне и упадљиве лидере, склоне изазивању хистеричног фанатизма и демонстрацији осветничке бруталности, већ скоро анонимне, педантне и монотоне технократске калкулаторе.<sup>5</sup>

Но, Арентова не само да није била једина која је видела негативне последице научно-технолошког прогреса, већ није истакла ни највећу опасност (компјутере) која би могла од њега доћи, а то и не чуди – онда када је преминула, компјутер је постао главни предмет бриге техно-скептика.<sup>6</sup> И у нашој науци примећујемо исти тренд разматрања о спрези тоталитаризма и технологије, само је почетак тог тренда померен унапред, након пада Берлинског зида.<sup>7</sup> И онолико колико каска временски, толико домаћа мисао о тоталитаризму каска и концептуално.

---

<sup>1</sup> Javier Burdman, *The Shadow of Totalitarianism*, State University of New York, Albany, 2022, p. 2.

<sup>2</sup> Hannah Arendt, *Essays in Understanding*, Schocken Books, New York, 1994, p. 323.

<sup>3</sup> Срђан Милашиновић, „Тоталитаризам и неолиберализам“, *Наука – безбедност – полиција*, Полицијска академија, Београд, Вол. 9, Бр. 2-3/2004, стр. 98.

<sup>4</sup> Mattias Desmet, *The Psychology of Totalitarianism*, Chelsea Green Publishing, London, 2023, p. 79, 91, 138, 158, 175.

<sup>5</sup> Hannah Arendt, *The Origins of Totalitarianism*, Harcourt Brace & Company, London, 1973, pp. 326 - 327

<sup>6</sup> Joseph Weizenbaum, *Computer Power and Human Reason*, W.H. Freeman & Company, New York, 1976, p. 11.

<sup>7</sup> Срђан Милашиновић, „Тоталитаризам и неолиберализам“, нав. дело, стр. 104 - 105, 107, 113, 118; Зоран Видојевић, „Либерални тоталитаризам“, *Социолошки преглед*, Институт друштвених наука, Београд, Вол. 33, Бр. 1-2/1999, стр. 110.

Јер, ако су се у свету формирале две „алармистичке“ струје у погледу лоших исхода ВИ, од којих једна види „тоталитаризам“, а друга „алгоритамски феудализам“,<sup>8</sup> у домаћој литератури, колико знамо, нема симултане анализе ВИ и тоталитаризма. Овај рад ће попунити ову празнину, бар што се тиче домаће политикологије.

Осим тога, даћемо допринос општем дијалогу на тему суштинe тоталитаризма кроз одговор на питање – да ли је прогрес ВИ пут у тоталитаризам без повратка? Ранији тоталитаризам је био реверзибилан, јер у моменту краха тоталитарног покрета, дојучерашњи адепти престају да верују у догму за коју су, донедавно, били спремни да жртвују животе.<sup>9</sup> Пад у тоталитаризам није нужно значио и „крај света“. Међутим, изгледа да у ВИ постоје клице иреверзибилности пута у тоталитаризам. Ако се то у овом раду и покаже, ВИ је најризичнија технологија коју је свет икада видео, а технократски тоталитаризам омогућен од стране ВИ, најопаснији од свих досадашњих.

## 2. ДЕФИНИЦИЈА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ И ТОТАЛИТАРИЗМА

Премда је било академских покушаја да се дефинише ВИ, универзално прихваћене дефиниције још увек нема.<sup>10</sup> Без обзира на то, могао се одредити за једну од дефиниција која најпрецизније одређује овај феномен – ВИ је студија о томе како да компјутери на исти или бољи начин раде исте оне ствари у којима је човек тренутно супериоран.<sup>11</sup> Не видимо на који би други начин могло да се изађе из оквира природне (људске) интелигенције. Свака друга дефиниција би имплицирала да је циљ ВИ да човечанству понуди софтвер са перформансама лошијим од било које човекове когнитивне способности. Тако би милијарде у новцу биле бачене у циљу стицања нечега што заправо као људи већ имамо.

---

<sup>8</sup> Akin Ünver, *Artificial Intelligence, Authoritarianism and the Future of Political Systems*, Centre for Economics and Foreign Policy Studies, Istanbul, 2018, p. 4.

<sup>9</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, p. 363.

<sup>10</sup> Lei Huang and Walter Peissl, „Artificial Intelligence—A New Knowledge and Decision-Making Paradigm?“, in: *Technology Assessment in a Globalized World* (eds. Leonhard Hennen, Julia Hahn, Miltos Ladikas, Ralf Lindner, Walter Peissl and Rinie van Est.), Springer, Cham, 2023, p. 176.

<sup>11</sup> Elaine Rich, Kevin Knight and Shivashankar B. Nair, *Artificial Intelligence*, Tata McGraw Hill, New Delhi, 2009, p. 3; Wolfgang Ertel, *Introduction to Artificial Intelligence*, Springer International Publishing, Cham, 2017, p. 2; Lei Huang and Walter Peissl, „Artificial Intelligence - A New Knowledge and Decision-Making Paradigm?“, нав. дело, p. 177.

Осим тога, ова дефиниција обухвата три главна облика ВИ – уску ВИ коју имамо данас, генералну и супер ВИ (суперинтелигенцију) које су пројектоване за будућност.<sup>12</sup> Уска ВИ има (над)људске перформансе, али у само једној сфери, попут играња шаха, трговања акцијама, медицинске дијагнозе, препознавања лица.<sup>13</sup> Једна уска ВИ апликација не може у свом опису да има и играње шаха и читање књига – за читање књига инжењери морају да направе нову апликацију. Док један човек може не само да игра шах и чита, већ може још много тога. Додали бисмо – и то не нужно серијски (једну активност за другом), већ и симултано (док игра шах, може и да чита, разговара, разазнаје фигуре у свом видокругу). Овај минус уске ВИ тек треба да превазиђе генерална ВИ која ће понудити једну апликацију са истим спектром функција и са истим учинцима које има савремени човек, док ће суперинтелигенција исто имати тако широк спектар људских функција, али са далеко бољим учинцима него што их данас има било који човек или људски колектив.<sup>14</sup> Узимањем у обзир само „бољих учинака“ у односу на човекове когнитивне способности у овој дефиницији, али без обраћања пажње на број домена човековог живота у којима се те способности примењују, обухватамо сва три типа ВИ и тако можемо да се бавимо и актуелним проблемима ВИ, али и перспективама.

Тоталитаризам је можда још теже дефинисати, без обзира што је у питању појам старији од ВИ. Јер, он је прожет манијом „перпетуалног превирања“, тј. нагле промене и дисконтинуитет – главне су црте тоталитарног менталитета.<sup>15</sup> И то без обзира да ли се ради о тоталитарном покрету који жели власт или је већ има.<sup>16</sup> Зато тоталитаризам

<sup>12</sup> Исто, р. 176.

<sup>13</sup> Cassio Pennachin and Ben Goertzel, „Contemporary Approaches to Artificial General Intelligence“, in: *Artificial General Intelligence* (eds. Ben Goertzel and Cassio Pennachin), Springer - Verlag, Berlin - Heidelberg, 2007, р. 1; Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, Oxford University Press, Oxford, 2014, pp. 14 - 16, 22, 52, 62.

<sup>14</sup> Cassio Pennachin and Ben Goertzel, „Contemporary Approaches to Artificial General Intelligence“, нав. дело, р. 1; Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, нав. дело, pp. 3, 22, 52 - 57.

<sup>15</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 306.

<sup>16</sup> „Свет фикције“ остаје и када се освоји власт, само је сад тоталитарном покрету теже да ту фикцију одржи у животу код маса – зато што сада за било који преживели остатак масама неугодне реалности, ултимативну одговорност носи сам тоталитарни покрет који се сада директно суочава са „девијацијама“ историје (вид. Исто, р. 389 - 392). Иако је Ризман добро уочио да се Арентова није детаљно бавила оазама аутономије и пасивног отпора унутар тоталитарних система, односно да није јасно истакла да се тоталитарни систем може срушити не само ратом споља (већ и унутрашњим превирањима) и да нису све раније друштвене стратификације нестале појавом тоталитарних маса (вид. Peter Baehr, *Hannah Arendt, Totalitarianism, and the Social Sciences*, Stanford University Press, Stanford,

треба схватити као *Weltanschauung*,<sup>17</sup> а не као идеологију, стил владања, тип режима.<sup>18</sup> И тоталитаризам има две сталне одлике.<sup>19</sup> Есенцију у виду „закона кретања“<sup>20</sup> (чији израз је горепоменућа манија) и принцип деловања у виду идеолошког логицизирања (већ поменутог у фусноти бр. 18).

По Арентовој, традиционални закон је статичан, нешто што треба што дуже да траје како би се гарантовала минимална стабилност за небројена спонтана кретања појединаца ка њиховим небројеним циљевима, а сада је закон схваћен динамично, као униформно кретање свих (без индивидуалног спонтанитета) ка једном циљу. У нацизму, тај циљ је била глобална владавина „чисте расе“, а у комунизму глобално „бескласно друштво“.

Исто, док су некада принципи деловања, попут страха у тиранијама, били поуздани водичи појединцима како да сачувају део својих интереса (у приватној сфери) од самовоље владара, сада је принцип деловања сасвим отргнут из било које предвидљивости и бачен у стерилну апстрактност идеолошког логицизирања. Сад више није само јавна сфера мртва, већ се људи осећају излишним у потпуности, приватност за њих нема никакву вредност, они налазе једини смисао живота у томе да буду „потрошни зупчаник“ машине логицизирајућих циклуса без краја, а са нултом тачком у аксиоматској премиси која са њима лично нема никакве везе.<sup>21</sup> Ако је став да тоталитарни покрет

---

2010, pp. 46 - 61) – то опет не значи да је Арентова (сасвим) порицала оно што је Ризман тврдио. Она је то имплицирала када је говорила о „директном суочавању“ са реалношћу. Можемо отворено рећи да је тоталитаризам као такав постојао у својој суштини сагласно мишљењу оба ова аутора, а да је разлика између њих само у њиховом различитом виђењу степена реализоване тоталитарне доминације у друштву (Арентова је тврдила да је моћ тоталитаризма да преобликује реалност у складу са фикцијом била већа, а Ризман мања).

<sup>17</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 470.

<sup>18</sup> Иако свака идеологија има тоталитарни потенцијал, само оне које не стављају акценат на идеју, већ на логички процес (логицизирање) који се из ње може извући како би се у целости протумачила природа (нацизам) или историја (комунизам) – јесу тоталитарне (вид. Исто, р. 472).

<sup>19</sup> Исто, р. 464 - 479. Ослањајући се на Монтескјеа, Арентова је есенцију политичких режима видела као одговор на питање ко је суверен у заједници, а принцип деловања као одговор на питање који етос надахњује присталице давог режима да делују овако или онако (вид. Peter Baehr, *Hannah Arendt, Totalitarianism, and the Social Sciences*, нав. дело, р. 74).

<sup>20</sup> Ово „кретање“ се мора схватати искључиво у линеарном и прогресивистичком кључу ка напред, никако ка назад или циркуларно – живот настављају само „најспособније расе или класе“ (вид. Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 463). „Кретање“ пледира на глобалистички универзализам, свака земља и појединац су обухваћени њиме (вид. Исто, pp. 389, 392, 462).

<sup>21</sup> Појединцима који не желе или оклевају да се подреде научним законима историје, упућују се имплицитне претње да ће постати „бескорисни“, што имплицира сувишност која се

не може да греши део аксиоматске премисе и ако тај покрет оптужи појединца да је „непријатељ“, онда тај појединац мора да призна да је крив, макар и ништа не скрививши, јер ако противречи покрету, онда сумња у премису, постаје непријатељ „објективних“ закона, чиме се руши његов смисао живота.

Генеза тоталитаризма ће боље расветлити ове две његове одлике. Све је почело крахирањем традиционалног света. Уместо укореваних појединаца са живим осећајем припадности некој породици, привредном или политичком ентитету, јављају се раскореване масе, са slabим уделом у мрежи социјалног капитала. Код тих маса се јавља „опесивна жеља бекства од реалности“ која је била свет спонтанитета, хаоса и нејасноћа, свет који се опирао ригидној структуралној конзистенцији.<sup>22</sup> Док се у реалности човек служи здравим разумом, одмереним увидом у међузависност случајног и нужног (спонтаног и планираног), у тоталитарној фикцији доминира лажљиви свет конзистенције, тј. пука имагинација у којој су масе без свакодневних шокова које реалност наноси људима и њиховим ишчекивањима.<sup>23</sup>

Међутим, ова пука имагинација није била производ масовног хира или фрустрације. Тоталитарне тврдње се увек заогрћу у научност.<sup>24</sup> Тоталитаризам никада не би могао да проклија као пуки механички збир психичких особина појединаца који улазе у састав тоталитарних маса. Оно што је овде био кохезивни „лепак“, то је прогностика преузета из сцијентизма који и друштвену стварност подређује методологији егзактних наука.<sup>25</sup>

Сад можда треба допунити Арентову. Наиме, научни закони нацистичког (Дарвин, Хекел)<sup>26</sup> и комунистичког (Маркс, Енгелс) тоталитаризма су формираны доста пре појаве тоталитарних покрета. Ако су ти закони имали привид науке због обима обрађене грађе, сложене методологије и високе апстрактности, списи и говори у којима су тоталитарне вође провлачиле своје стварне или могуће (али конкретне) владалачке изазове кроз „сито и решето“ таквих научних закона

---

може физички елиминисати (геноцидом, чисткама). Вид. Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 345.

<sup>22</sup> Исто, р. 352.

<sup>23</sup> Исто, р. 353.

<sup>24</sup> Исто, р. 345.

<sup>25</sup> F. A. v. Hayek, „Scientism and the Study of Society. Part I“, *Economica*, No. 35, Vol. 9, 1942, pp. 267 - 269.

<sup>26</sup> Richard Weikart, „The Role of Darwinism in Nazi Racial Thought“, *German Studies Review*, no. 3/2013, vol. 36, p. 541.

– били су отворена псеудонаука. Но, уверење о тачној прогностици је било заједничко у оба случаја. Зато је илузију прогностике као „прецизног мерења будућности“ боље сматрати за есенцију тоталитаризма од „закона кретања“.

То уверење се у тоталитаризму јавља као „научно пророчанство“ – вође нацизма и комунизма су се представљале као тумачи „ланца фаталности“ природе и историје,<sup>27</sup> као неко ко апстрактни језик научних закона преводи на – масама разумљив – језик свакодневнице. Управо овде, у псеудонаучном тумачењу/егзегези научних закона од стране тоталитарних вођа, по нама, настаје принцип деловања тоталитаризма на власти, тј. идеолошко логицизирање која се може свести на Хитлерова „ледено хладна“ или Стаљинова „немилосрдна дијалектичка“ тумачења расне или класне историје.<sup>28</sup> Ко верује у предвиђања и научне законе и у егзегезе њихових тумача, аутоматски ступа на пут „успеха“.<sup>29</sup> Арентова добро запажа како ова опседнутост науком није суштински нова, јер се ради о последњем стадијуму једне дуге историје, у којој је наука магичан лек за злу реалност, али уз сасвим нову тачку – трансформацију људске природе.<sup>30</sup> Прогностика сцијентизма је преузета, али не и његов утилитаризам, пошто утилитаризам није тежио да мења људску природу.<sup>31</sup>

Имајући све ово у виду, и без обзира што је Арентова издвојила само две одлике тоталитаризма као непроменљиве, одлучујемо се да пружимо листу основних одлика тоталитаризма. И то, како

<sup>27</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 345.

<sup>28</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 471.

<sup>29</sup> Исто, р. 346. На том путу „успеха“ нису сви добитници. „Неподобне расе“ и „превазиђене класе“ неумитно морају да нестану. Остаје само питање да ли ће се то догодити пре или касније у зависности од тога да ли човек жури да примени „законе“ како би „убрзао“ историју или бира пасивност, чиме чека да време одради своје и омогући историји или природи да однесу превагу на дужи рок (вид. Исто, рр. 349 - 350, 462, 466). Појединци који још нису сврстани у „лошу расу/класу“ могу да играју данас улогу целата. А сутра, ако се тумачењем вођа ипак сврстају у лошу групу, дужни су подједнако ревносно да приме и улогу жртве како не би ометали колективистички галоп ка земаљској утопији (вид. Исто, рр. 425, 465).

<sup>30</sup> Исто, рр. 346 - 347. Овај њен став у потпуности чини излишним приговор који је дошао од Рејмона Арона – да тоталитаризам није потпуни *novum* у историји, већ само мутација раније виђених једнопартијских система или револуција у историји (вид. Peter Baehr, *Hannah Arendt, Totalitarianism, and the Social Sciences*, нав. дело, рр. 82 - 83). Имајући у виду ове њене речи, можемо рећи да је, по њој, тоталитаризам нешто сасвим ново у историји у смислу степена идололатризовања науке, тј. уверења да наука може да мења и људску природу. Међутим, овај квантитативно нови сцијентизам је имао такве ефекте по људску стварност, да су многе (ако не и све) области те стварности (макар и привремено) претрпеле и квалитативне метаморфозе.

<sup>31</sup> Hannah Arendt, *The Origins of Totalitarianism*, нав. дело, р. 347.

ембрионалног (ван власти) тако и етаблираног тоталитаризма (на власти). Овај други има све одлике првог, али и неке сасвим своје. Само оваквим дефиницијским „раслојавањем“ у таквој листи, могу се прегледно истражити везе између две поменуте фазе тоталитаризма и две фазе ВИ. Овакав поступак ни у чему не противречи Арентовој, пошто се одлике из те наше листе иначе могу подвести под једну (есенцију) или другу (принцип деловања) арентовску одлику тоталитаризма.

Ембрионални тоталитаризам одликују: 1) сцијентистичка опседнутост прогностиком целокупне стварности, 2) човекова природа мора да се мења не еволутивно-спонтано, већ интелектуално-плански у правцу укидања аутономије, 3) начела људских права и слобода нису апсолутна, већ релативна, 4) глобалистичка императивност – цела планета би требало да се помири са свим могућим ефектима тоталитаризма, опција егзистирања ван тоталитаризма се одриче или се експлицитно не наводи. На ово се надовезују и две одлике етаблираног тоталитаризма: 1) тотални терор који уништава аутентичне везе људи и институције јавне и приватне сфере, којим се друштво ставља у погон перпетуалних и радикалних промена, а сам овај терор је надахнут оним што је – 2) идеолошко логицизирање у којем свака особена мисао мора да се по сваку цену утопи у логику апстрактних научних закона (онако како то тумачи тоталитарна врхушка), за разлику од нетоталитарног мишљења у којем логика није циљ сама себи, већ помоћно средство да се појединац кроз саморефлексију и дијалог са другима уникатно позиционира у свету који га окружује.<sup>32</sup>

### 3. УСКА ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА КАО КАПИЛАРНИ ТОТАЛИТАРИЗАМ

Имајући у виду да иза уске ВИ не стоји један покрет као што је стајао по један иза нацистичке или комунистичке идеје, јасно је да између уске ВИ и тоталитаризма – чак и ембрионалног – не можемо повући потпуни знак једнакости. Уска ВИ има подршку не само једног покрета, већ начелну (често прећутну) подршку свих покрета једног друштва.<sup>33</sup> Ако не све, најмоћније државе света сигурно посвећују изузетну пажњу ВИ кроз све већа буџетска давања. Зато се данас може

---

<sup>32</sup> Исто, рр. 476 - 477.

<sup>33</sup> Није нам познато да у некој земљи постоје релевантнији политички покрети који истичу да се противе ВИ као таквој или барем оној која значи обавезну примену ВИ кроз државну администрацију.

говорити о евентуалној повезаности ембрионалног и етаблираног тоталитаризма са идејама ВИ, али без лоцирања тоталитарних покрета који би имали задатак да идеје тоталитаризма промовишу и примене када приграбе државну власт.<sup>34</sup>

1) Уска ВИ (машинско учење) је заснована на сцијентистичкој прогностици.<sup>35</sup> Иако се она не базира на законима који прогнозирају ток историје или природе као целине, она имплицитно прихвата да је сваки микро-домен човекове стварности подложен микро-прогностици.

Тако, регресија као базична врста уске ВИ предвиђа догађаје у свим научним областима, да би се скоро могло рећи да данас нико не може да напредује као научник без познавања регресионих модела.<sup>36</sup> Али, и софистициранија уска ВИ под називом „дубинско учење“, позната је по препознавању (предвиђању) објеката, а посебно по препознавању лица особа чији су идентитет регистровани државни органи.<sup>37</sup> Већ се развија уска ВИ која ће предвиђати када ће човек да се разболи, када ће радник дати отказ, да ли је кандидат за посао добар избор за послодавца, да ли неко има криминалне склоности – лимита у подложности прогностици нема, осим уколико подаци о нечему ни су доступни.<sup>38</sup> Иако се овде не обећава математичка извесност, прихвата се да је то пут будућности, све док не стигне генерална или суперинтелигенција, а да ће се прогностичке грешке у ходу смањивати тако да теже нули.<sup>39</sup>

2) Сличан је утицај уске ВИ и на човекову природу. Нема неке холистичке визије „новог човека“, већ се по неким ауторима иде на то да се појединачним апликацијама поједини аспекти човека као мисаоног бића непосредно, систематски и на дужи рок преобразе у одређеном

<sup>34</sup> У уводу смо поменули могућност анонимног карактера будућег технократског тоталитаризма који неће доћи кроз усталасане масе као пре, тако да ова разлика није битна за наше даље разматрање.

<sup>35</sup> Stephen Marsland, *Machine Learning*, Taylor & Francis Group, Boca Raton, 2015, pp. 10 - 11.

<sup>36</sup> Ashish Sen and Muni Srivastava, *Regression Analysis*, Springer - Verlag, New York, 1990, p. 18; Richard B. Darlington and Andrew F. Hayes, *Regression Analysis and Linear Models*, Guilford Publications, New York, 2017, p. 9; Овде прецизирамо да не тврдимо како је свака апликација уске ВИ зупчаник у погону који нас води ка овом или оном облику тоталитаризма. Примене уске ВИ у строго техничким и класичним инжењерским доменима није предмет овог рада, већ само примена уске ВИ у домену друштвених наука, тј. културе и социјалних интеракција.

<sup>37</sup> Ian Goodfellow et al., *Deep Learning*, The MIT Press, Cambridge - London, 2016, pp. 9, 13; Ankan Bansal et al., „Deep CNN Face Recognition: Looking at the Past and the Future“, in: *Deep Learning-Based Face Analytics* (eds. Nalini K Ratha, Vishal M. Patel and Rama Chellappa), Springer Nature, Cham, 2021, p. 1.

<sup>38</sup> Mark Andrejevic, *Automated Media*, Taylor & Francis, New York - London, 2020, p. 9.

<sup>39</sup> Stephen Marsland, *Machine Learning*, нав. дело, p. 74 - 77.

смеру – чак и на појединчеву штету (*hypernudge* манипулација) – или просто паралишу, тј. замене компјутером.<sup>40</sup> По тим ауторима, манипулације је увек било, али је она била индиректна и спорадична, тј. човек је могао да у повременој самоизолацији изврши (само)рефлексију, на фону свега са чиме је дошао у контакт (па и манипулативним порукама). Манипулатор је емитовао једну исту поруку за све људе, а они су је прихватили или одбијали у зависности од својих аутентичних искустава које су носили са собом из приватне, манипулатору недоступне сфере. Манипулатори су морали више да импровизују у својим стратегијама, него што су нешто могли поуздано да предвиђају. Уласком уске ВИ у поље манипулације, све се мења. Манипулатори (ВИ инжењери и тестери),<sup>41</sup> сада не само да знају много више базичних информација о појединцу (име и презиме, пол, пребивалиште и слично), већ и корелацијске везе између образаца његовог мишљења (когнитивне пристраности) и проживљеног контекста.<sup>42</sup> Тиме се са знаје који контекст „продукује“ који избор појединца, а онда се дигиталним путем „штелује“ контекст појединца у смеру – за манипулатора – пожељне појединчеве одлуке. Притом, појединац је изложен манипулацији од стране ВИ увек и свуда, докле користи мобилни телефон или компјутер. Тиме се крњи приватност, толико неопходна за одбрану од тоталитаризма. И још, некад је појединац уз мали ментални напор знао када је изложен манипулацији, док су данас потребна не мала знања да би се схватило како прикривено манипулација делује кроз савремене технологије.

3) Однос уске ВИ и права произилази из утицаја уске ВИ на човекову природу. За разлику од манипулације некад, сад манипулација уске ВИ крши основна људска права и слободе (људско достојанство и слободан развој личности, савести, мишљења и изражавања) – а што је истакнуто као негативна појава са ризицима без преседана од стране међународних организација као што су УН, УНЕСКО, ЕУ, Савет

---

<sup>40</sup> Stefano Faraoni, „Persuasive Technology and computational manipulation: hypernudging out of mental self-determination“, *Frontiers in Artificial Intelligence*, Frontiers, Lausanne - London - Madrid - Beijing, vol. 6/2023, pp. 1 - 5; Mark Andrejevic, *Automated Media*, нав. дело, р. 9.

<sup>41</sup> О великој улози људског фактора у раду уске ВИ видети: Kate Crawford, *Atlas of AI*, Yale University Press, New Haven - London, 2020, pp. 63, 65 - 66, 68, 96, 107, 211, 219; Eric J. Larson, *The Myth of Artificial Intelligence*, The Belknap Press of Harvard University Press, Cambridge - London, 2021, pp. 133 - 134, 138 - 139.

<sup>42</sup> Такве корелацијске везе свакако нису доступне ниједном појединцу, али ни традиционалним компјутерима који не спадају у уску ВИ. Вид. Eric J. Larson, *The Myth of Artificial Intelligence*, нав. дело, pp. 120, 174, 259 - 260.

Европе.<sup>43</sup> У САД је већ заживела претпоставка кривице. Појединци се лишавају права на рад и слободно кретање, не зато што су учинили деликт за који им је пресуђено, већ само на основу дигиталне сумње, тј. зато што је уска ВИ обрадила њихове податке на такав начин, да их је проценила као вероватне преступнике који ће тек у будућности учинити деликт.<sup>44</sup> Све ово значи да је прогностика према замислима архитектата уске ВИ надређена било ком правном начелу.

4) Глобалистичка императивност уске ВИ произилази из прве тачке ове листе. Уска ВИ која третира целу планету као једну базу података, свакако би радила прогностички боље од уске ВИ која би имала приступ само базама у којима се налазе подаци једног континента – моћ уске ВИ зависи и од квалитета чипова и од величине база.<sup>45</sup> Глобализам уске ВИ је најбоље исказао један инжењер *Google*-а, речима да све што видимо треба да се нађе у базама, тј. да уска ВИ претендује да буде огледало реалног света.<sup>46</sup> А, нису нам познате државе, међународне организације, партије или групе од утицаја у свету које се залажу за живот изван уске ВИ (или барем ван њене обавезне примене кроз државни апарат). Критичко сагледавање уске ВИ постоји, али не ње као принципа, већ само неких њених аспеката.

Између уске ВИ и етаблираног тоталитаризма нема директних веза, јер данас нема тоталног терора нити идеолошког логицизирања. То не значи да су јавна и приватна сфера неокрњене. Видели смо да уска ВИ, иако не укида, ипак нарушава приватну сферу. Посредно, трпи и јавна сфера, јер у њу улази појединац из своје дефектне приватности. Појединци данас могу слободно да се удружују са другим људима у обе сфере за разлику од етаблираног тоталитаризма. Могу и да одаберу живот без ВИ технологија. Али, и даље се поставља питање аутентичности мишљења оних људи који су и даље у дигиталном свету, а таквих је велика већина. Јавни и приватни контакти са другима који још увек постоје, појединцу све мање значе за формирање аутентичног искуства које је пресудно за слободно мишљење, јер све те предности знатно анулира уска ВИ која дигитално преобликује мисао човека, нудећи му сужену палету избора и слабећи му капацитете за (само)рефлексију.

---

<sup>43</sup> Stefano Faraoni, „Persuasive Technology and computational manipulation: hypernudging out of mental self-determination“, нав. дело, pp. 5 - 6.

<sup>44</sup> Margaret Hu, „Big Data Blacklisting“, *Florida Law Review*, no. 5/2016, vol. 67, pp. 1735 - 1809.

<sup>45</sup> Tom Taulli, *Artificial Intelligence Basics*, Apress, 2019, pp. 15 - 16.

<sup>46</sup> Kate Crawford, *Atlas of AI*, нав. дело, p. 220.

Због свих ових сличности и разлика између уске ВИ и ембрионалног/етаблираног тоталитаризма, друштво у којем је свеprisутна уска ВИ називамо друштвом „капиларног тоталитаризма“. У њему се превирање друштва врши без тоталног терора, али ефикаснијим, поступнијим и мање уочљивим (капиларнијим) путем него што је то био случај у етаблираним тоталитаризмима. Логицизирање није у оквирима апстрактних научних закона, већ индивидуалних „дигиталних профила“ који воде не ка униформној аутоматизацији свих, већ конкретног субјекта – човек постаје „компресован“ контингентно (корелационо) конструисаним контекстима и надовезаним корпусом понуђених избора. Док у старом тоталитаризму имамо централистичку монолитност у виду једног покрета који нуди универзална решења за све људе, сада имамо сушту супротност тога – децентрализовани плурализам великог броја апликација које могу да се оптимизују за манипулативни утицај на сваког корисника интернета и мобилног телефона. Ако смо некад имали један главни научни закон који се односи на све људе у једном тоталитаризму, сада имамо много „закона“ који се односе на конкретне индивидуе у много „малих тоталитаризама“ – отуда капиларност. Сви су аутоматизовани у „капиларном тоталитаризму“, али свако има своје параметре којима се „штелује“ његова мисао, иако је метод откривања тих параметра такође исти за све.

#### 4. СУПЕРИНТЕЛИГЕНЦИЈА КАО ХИПЕРТОТАЛИТАРИЗАМ

Како је генерална ВИ прелаз ка суперинтелигенцији, а теоретичари суперинтелигенције мисле да ће он бити кратак,<sup>47</sup> сагледајмо везу између суперинтелигенције и тоталитаризма.

Већ смо рекли да ће суперинтелигенција имати функције које има и човек, али да ће их по учинцима далеко превазићи. Бостром каже да је ова дефиниција непрецизна и наглашава три форме суперинтелигенције: 1) квантитативна суперинтелигенција која човекове проблеме решава далеко брже од њега, 2) квалитативна суперинтелигенција која решава људске проблеме кроз квалитативно боља решења, 3) колективна суперинтелигенција којој ни најбољи људски колектив не може парирати у погледу когнитивних способности.<sup>48</sup>

---

<sup>47</sup> Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, нав. дело, р. 65.

<sup>48</sup> Исто, р. 52 - 57.

Суперинтелигенција зато намеће два етичка питања, од којих се једно (питање контроле) тиче дилеме односа суперинтелигенције према људима, тј. да ли ће нам она чинити штету и довести нас до истребљења.<sup>49</sup> Од одговора на ово, зависи и однос суперинтелигенције и етаблираног тоталитаризма. Ако суперинтелигенција допушта ликвидацију појединаца као нешто легитимно, само зато што су они, на пример, људи који не воле прогрес – онда је она компатибилна са етаблираним тоталитаризмом. Јер, сетимо се да је ликвидација неподобних (нацизам) или спремност да се жртвује властити живот (комунизам) у име прогреса, заправо одлика етаблираног, а не ембрионалног тоталитаризма.

Постоје предлози како ограничити техничке капацитете суперинтелигенције, али то је од привременог значаја за питање контроле – трајна контрола постоји само ако суперинтелигенција има вредности које ће је руководити.<sup>50</sup> Уместо да људи одреде те вредности, архитекте суперинтелигенције кажу да вредности које ће следити суперинтелигенција – треба да одреди она сама.<sup>51</sup>

Овакву примопредају права на аксиолошку рефлексију они зову „кохерентна екстраполирана воља“. То значи да суперинтелигенција може да сазна шта би човек, да има когнитивне способности суперинтелигенције, желео да види као главну универзалну вредност, како би своје активности провлачио кроз њену призму.<sup>52</sup> Суперинтелигенција не одабира неку конкретну људску вредност, већ детектује извор свих људских вредности који је невидљив човеку због његових интелектуалних „мањкавости“. Ако екстраполирана воља види да људи заправо желе диктатуру, онда суперинтелигенција заводи светску диктатуру. Исто и у случају да је пожељна демократија. А ако људи заправо не желе ВИ – суперинтелигенција просто гаси себе и удовољава таквој вредности/жељи човечанства.<sup>53</sup>

Ово је одлика етаблираног тоталитаризма – логицизирање. Мислећем субјекту се одузима његово искуство и животни контекст (због

---

<sup>49</sup> S. Matthew Liao, „A Short Introduction to the Ethics of Artificial Intelligence“, in: *Ethics of Artificial Intelligence* (ed. S. Matthew Liao), Oxford University Press, Oxford, 2020, p. 19.

<sup>50</sup> Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, нав. дело, pp. 129 - 138; 185.

<sup>51</sup> Исто, p. 209.

<sup>52</sup> Исто, pp. 211 - 212, 221.

<sup>53</sup> Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, нав. дело, pp. 215-216. Чак и да аутогашење буде прописно спроведено, то опет не значи да би човечанство било без штете. Испало би да су толики ресурси потрошени за развој нечега што није корисније од шибце, а могли су бити употребљени за нешто исплативо. Аутогашење је стога аргумент против суперинтелигенције.

когнитивних пристрасности) и намеће му се вештачко искуство и контекст суперинтелигенције (због „објективности“). Док је ван тоталитаризма човек органски повезан са својим контекстом, и човек сам одатле креира идентитет ког чулима и интелектом проживљава као искуство, овде се та органска веза хируршки сече и човек не креира свој идентитет, већ пасивно прима туђе вештачке конструкте. Ако је идеолошко логицизирање некад имало много мање додирних тачака са материјалном реалношћу од суперинтелигенције (због слабије технике тада), то опет не мења чињеницу да је и суперинтелигенција логицизирање (иако је технички јача). Ма колико података из човековог контекста успела да акумулира и анализира, суперинтелигенција не може да замени аутентично мишљење човека, јер му сервира идентитет који му је сасвим стран – симболе које он сам никада не би вољом хтео или разумом докучити могао.

Суперинтелигенција није само логицизирање, већ и тотални терор. Већ смо истакли Бостромов став како све што не може да се подведе под вредности које ће суперинтелигенција следити, мора да буде спремно на екстинкцију. А један од предуслова за екстраполирану вољу јесте „енкапсулација моралног прогреса“ која каже да само оне културе које допуштају ревизију својих моралних начела могу имати будућност. Оне које се буду држале непроменљивости својих вредности, неће имати бољи статус од оног који је резервисан за сировинску инфраструктуру суперинтелигенције – људи таквих култура биће третирани као материја за хардвер компоненте ВИ.<sup>54</sup> А ако ће суперинтелигенција брисати недовољно прогресивне вредности, то значи да се неће штитити ниједан морални систем по себи, већ само онај који суперинтелигенција оцени као прогресиван – морал се мења и усавршава у складу са диктатом ВИ, оно што „заостаје“ мора да нестане.<sup>55</sup> Све се мора мењати у стању вечитог превирања и ништа не

---

<sup>54</sup> Исто, pp. 122 - 125.

<sup>55</sup> Бостромова екстраполирана воља ће заживети само уколико буде кохерентна, тј. ако својом холистичком анализом свих култура нађе њихове заједничке тачке. Но, ту се не ради о компромису како га ми обично тумачимо (да се свима удели једнако признање њихових интереса какви год да су), већ о компромису који би постојао да су сви људи суперинтелигентни. Такви људи би онда „непристрасно“ сагледали све културе и ако би утврдили да је једна блиска суперинтелигенцији, а да друга није, онда би припадници ове друге културе попустили и прихватили прву културу као универзалну према којој би се надаље равнали. Бостром наводи пример суперинтелигентне арбитраже између шведског хуманисте и авганистанског талибана. Али, ако би победио хуманизам као универзална вредност ВИ – шта би остало од талибанизма? И ако победи талибанизам – шта би остало од хуманизма? Очигледно да ниједна култура по себи није врховна, већ само у мери у којој

сме остати исто, како би се овај пут задовољио не научни закон о раси или класи, већ о ВИ. Тако да ништа није заштићено у суперинтелигенцији осим ње саме, баш као што и у етаблираном тоталитаризму ништа није било сигурно (ни тоталитарни врх) осим самог тоталног терора и идеолошког логицизирања.

Тоталитаризам суперинтелигенције, колико је сличан – толико је и различит од етаблираног тоталитаризма XX века. Он ће имати далеко већу техничку моћ од свог претходника, а објекат његових амбиција биће, осим наше планете, чак читав универзум (па и хипотетички паралелни многи универзуми).<sup>56</sup> Но, главна разлика у односу на раније тоталитаризме ће бити његова иреверзибилност – ослобођење од таквог тоталитаризма неће бити могуће. Јер, ако су старе тоталитаризме (чији су носиоци били људи) могле да победе нетоталитарне државе (састављене од људи) или властите унутрашње слабости (приписане људима), овај тоталитаризам чији носиоци нису људи, сами људи неће моћи да савладају. Наспрам људи који не воле тоталитаризам неће стајати људи који га воле са једнаким интелектуалним капацитетима и техником, већ машине које су интелектуално и технолошки далеко моћније. Зато, суперинтелигенцију би требало назвати „хипертоталитаризам“, с обзиром да далеко надмашује како моћ тако и негативне ефекте старог тоталитаризма.

## 5. ЗАКЉУЧАК

У светлу приказаног, сматрамо да начела садашњег правног и политичког поретка можемо сачувати само ако се према ВИ поставимо на следећи начин. У приватној сфери би требало да важи слобода избора грађана у погледу тога да ли ће се користити уском ВИ. И у јавној сфери би требало да буде слично, увођењем начела *изузећа* – грађани који хоће да буду изузети од свих ефеката уске ВИ чији је генератор држава, треба да имају право на то (по узору на приговор савести). Држава би приликом пружања јавних добара таквим грађанима или вођења обичне комуникације са њима, требало да користи традиционалне методе који јој стоје на располагању. Јер, ако неко не жели уску ВИ у приватном животу, обавезност коришћења уске ВИ како би се остварило неко право или испунила обавеза према

---

поседује елементе суперинтелигенције. Nick Bostrom, *Superintelligence – Paths, Dangers, Strategies*, нав. дело, pp. 212 - 213, 215.

<sup>56</sup> Исто, pp. 292.

држави, поништила би бенефите непостојања уске ВИ у приватном режиму за појединца. Појединци би требало да имају слободу улаза/излаза из уских ВИ аранжмана којима су покривени њихови односи са државом. Ако постоји начело изузећа, „капиларни тоталитаризам“ не може да уздрма старе правне и политичке тековине. Суперинтелигенција као узрок „хипертоталитаризма“ би морала уставно да се забрани, уколико се не жели укидање досадашњих правно-политичких традиција. Спречавање њеног развоја није само *conditio sine qua non* за очување нашег поимања права и политике, већ и за очување есенције човека као таквог – његове слободе да аутентично, спонтано и аутономно мисли и дела.

## ЛИТЕРАТУРА

Видојевић Зоран, „Либерални тоталитаризам“, *Социолошки преглед*, Институт друштвених наука, Београд, вол. 33, бр. 1-2/1999, 105-123;

Милашиновић Срђан, „Тоталитаризам и неолиберализам“, *Наука – безбедност – полиција*, Полицијска академија, Београд, вол. 9, бр. 2-3/2004, 97-124;

Andrejevic Mark, *Automated Media*, Taylor & Francis, New York - London, 2020;

Arendt Hannah, *The Origins of Totalitarianism*, Harcourt Brace & Company, London, 1973;

Arendt Hannah, *Essays in Understanding*, Schocken Books, New York, 1994;

Baehr Peter, *Hannah Arendt, Totalitarianism, and the Social Sciences*, Stanford University Press, Stanford, 2010;

Bansal Ankan, Ranjan Rajeev, Castillo Carlos D. and Chellappa Rama, „Deep CNN Face Recognition: Looking at the Past and the Future“, in: *Deep Learning-Based Face Analytics* (eds. Nalini K Ratha, Vishal M. Patel and Rama Chellappa), Springer Nature, Cham, 2021;

Bostrom Nick, *Superintelligence – Paths, Dangers, Strategies*, Oxford University Press, Oxford, 2014;

Burdman Javier, *The Shadow of Totalitarianism*, State University of New York, Albany, 2022;

Crowford Kate, *Atlas of AI*, Yale University Press, New Haven - London, 2020;

Darlington Richard B. and Hayes Andrew F., *Regression Analysis and Linear Models*, Guilford Publications, New York, 2017;

Desmet Mattias, *The Psychology of Totalitarianism*, Chelsea Green Publishing, London, 2023;

Ertel Wolfgang, *Introduction to Artificial Intelligence*, Springer International Publishing, Cham, 2017;

Faraoni Stefano, „Persuasive Technology and computational manipulation: hypernudging out of mental self-determination“, *Frontiers in Artificial Intelligence*, Frontiers, Lausanne – London – Madrid - Beijing, Vol. 6/2023;

Goodfellow Ian, Bengio Yoshua and Courville Aaron, *Deep Learning*, The MIT Press, Cambridge - London, 2016;

Hayek F. A. v., „Scientism and the Study of Society. Part I“, *Economica*, No. 35/1942, Vol. 9, 267-291;

Hu Margaret, „Big Data Blacklisting“, *Florida Law Review*, No. 5/2016, Vol. 67;

Huang Lei and Peissl Walter, „Artificial Intelligence - A New Knowledge and Decision - Making Paradigm?“, in: *Technology Assessment in a Globalized World* (eds. Leonhard Hennen, Julia Hahn, Miltos Ladikas, Ralf Lindner, Walter Peissl and Rinie van Est), Springer, Cham, 2023, 175–201;

Larson Eric J., *The Myth of Artificial Intelligence*, The Belknap Press of Harvard University Press, Cambridge - London, 2021;

Liao S. Matthew, „A Short Introduction to the Ethics of Artificial Intelligence“, in: *Ethics of Artificial Intelligence* (ed. S. Matthew Liao), Oxford University Press, Oxford, 2020;

Marsland Stephen, *Machine Learning*, Taylor & Francis Group, Boca Raton, 2015;

Pennachin Cassio and Goertzel Ben, „Contemporary Approaches to Artificial General Intelligence“, in: *Artificial General Intelligence* (eds. Ben Goertzel and Cassio Pennachin), Springer - Verlag, Berlin - Heidelberg, 2007;

Rich Elaine, Knight Kevin and Nair Shivashankar B., *Artificial Intelligence*, Tata McGraw Hill, New Delhi, 2009;

Sen Ashish and Srivastava Muni, *Regression Analysis*, Springer - Verlag, New York, 1990.

Taulli Tom, *Artificial Intelligence Basics*, Apress, 2019;

Ünver Akın, *Artificial Intelligence, Authoritarianism and the Future of Political Systems*, Centre for Economics and Foreign Policy Studies, Istanbul, 2018;

Weikart Richard, „The Role of Darwinism in Nazi Racial Thought“, *German Studies Review*, No. 3/2013, Vol. 36, 537-556;

Weizenbaum Joseph, *Computer Power and Human Reason*, W.H. Freeman & Company, New York, 1976;

Marko Pejković

## ARTIFICIAL INTELLIGENCE AND TOTALITARIANISM

### Resume

After the introduction, the first part of the paper is devoted to defining the concepts of artificial intelligence and totalitarianism, where the importance of distinguishing between the current (machine learning) and the projected (superintelligence) phase in the development of artificial intelligence, i.e. between the embryonic (totalitarian movement out of power) and the established (totalitarian movement in power) stage in the development of totalitarianism is underlined. The second part of the paper examines the connection between the current level of artificial intelligence and the embryonic phase of totalitarianism, while the third part of the paper analyzes the potential relationship between the superintelligence and the established totalitarianism. It seems, considering the similarities and differences between the effects of contemporary and future artificial intelligence and the effects of earlier totalitarianism, that today (and in the future) we do not have a mere replica of totalitarian phases from the 20th century, but special totalitarian phenomena in the form of “capillary totalitarianism”, i.e. “hypertotalitarianism”. Last century’s totalitarianism, as well as today’s “capillary” variant of it, were not necessarily irreversible, but “hypertotalitarianism” will be. In conclusion, protective measures against the risk of artificial intelligence are proposed, in the form of the principle of exemption (modeled after the concept of conscientious objection).

Keywords: artificial intelligence, machine learning, superintelligence, totalitarianism, capillary totalitarianism, hypertotalitarianism.