

Marko Uršič

VIRTUALNA REALNOST IN ČLOVEŠKI SVET

Nevarnost moderne znanosti in tehnike, na katero je opozarjal že Martin Heidegger, ni dandanes nič manjša, v nekem smislu je celo večja kot sredi minulega stoletja, saj je bistvo sodobne tehnike še bolj zastrto človeškim mislim in večina ljudi sprejema razvoj tehnike kot nekaj samoumevnega in človeku prijaznega. Po koncu hladne vojne med velesilama se je res zmanjšala nevarnost jedrskega spopada, latentno nevarne pa postajajo nove tehnologije, zlasti biotehnologija in kibertehnika. S kibertehniko prodira virtualnost vse bolj v naš »realni svet«, pri čemer se izgublja tako razlika med čutno in virtualno realnostjo (kot da se vračamo v platonsko votlino senc) kakor tudi razlika med živo človeško zavestjo in strojem: dogaja se postopen, vendar vztrajen proces »kiborgizacije« človeka, družbe in sveta. Toda tisto »rešilno« ni odvrčanje od znanosti, ni neka utopična vrnitev k »pristnemu človeku« in/ali »naravi«, temveč le vztrajanje človeka v mišljenju, samo-zavedanju, etiki in duhu.

Ključne besede: VIRTUALNA REALNOST, KIBERTEHNIKA, MARTIN HEIDEGGER, VILÉM FLUSSER, JEAN BAUDRILLARD, VERNOR VINGE, RAY KURZWEIL, NICK BOSTROM

Glavna tema mojega prispevka je obravnava vprašanja, kakšno vlogo, vpliv in pomen imajo *virtualni prostori* za naše človeške, zgodovinske, »realne prostore« (fizične, življenjske, družbene itd.) v našem »realnem času«, tj. nekaj desetletij po prvem valu kibertehnike, ko nas preplavlja njen drugi veliki val: svetovni splet, v katerem so postali virtualni prostori/svetovi *interaktivni* v odnosu do našega »realnega sveta« in do nas samih, saj so že prodrli in še naprej prodirajo v našo človeško in zgodovinsko sedanost. Ugotavljamo, da se *virtualna realnost*, ki je bila prvotno omejena na igrice, kiberumetnost, modele in simulatorje (npr. letenja), vse bolj in bolj vključuje v samo svetovno *realnost*, seveda pa slednja še ni nujno tista ontološka *resničnost*, ki jo filozofija išče od predsokratikov dalje (in v tem iskanju nemalokrat obupuje). Kot filozof, ki v svojem temeljnem prepričanju in mišljenju sledim platonizmu, želim posebej poudariti, da moramo razlikovati med realnostjo in resničnostjo – s tem ohranjamo »transcendenčno napetost«, *možnost presežnosti* (za več o tem gl. Uršič 2015).

Dandanes je »nevarnost« moderne tehnike, na katero je opozarjal Martin Heidegger, zlasti v razpravi *Tehnika in preobrat* (gl. Heidegger 1967), še večja, bolj zakrita v svoji samoumevni univerzalnosti, kot je bila v prejšnjem stoletju. Poglejmo odlomek iz te Heideggerjeve razprave:

... snujoč <dichterisch> prebiva človek na tej zemlji.

*To snovalno <das Dichterische> spravlja to resnično v sijaj tistega, kar imenuje Platon v dialogu *Fajdros* tò ekphanéstaton, [tisto] kar najčistejše izseva. To snovalno preveva vsako umetnost, vsako razkrivanje bistvujočega v lepo.*

Ali naj se lepe umetnosti pokličejo v snujoče razkrivanje? [...]

Nihče ne more vedeti, ali je umetnosti dopuščena/zagotovljena ta najvišja možnost njenega bistva sredi skrajne nevarnosti. Lahko pa se začudimo. Pred čim? Pred drugo možnostjo, da se vsepovsod ugnezdi divjanje tehnike, dokler ne prične nekega dne skozi vse tehnično bivati bistvo resnice v dogodku <Ereignis, dogodje> resnice. (Heidegger 1967, 361)

Heideggerjeve besede o nevarnosti tehnike so – seveda namenoma – dvoumne. Tako bistvo tehnike kakor bistvo umetnosti, zlasti poezije, je »snovanje«, dogodje resnice biti. V tem smislu lahko razumemo tudi znamenita verza Friedricha Hölderlina, ki ju Heidegger zelo poudarjeno navaja: *Kjer pa je nevarnost, raste / tudi Rešilno*. – S prvim valom kibertehnike, ki je dosegel svoj vrh v osemdesetih letih minulega stoletja, je pri mnogih takratnih umetniških ustvarjalcih in teoretikih prišlo do zanimivega »obrata« v odnosu med umetnostjo in tehniko (ter znanostjo): kibertehnika jim ni več pomenila nevarnosti za človekovo »snujoče prebivanje na tej zemlji«, temveč ravno nasprotno – najbolj obetavno možnost neke nove človeške ustvarjalnosti in svobode. To obdobje v razvoju kibertehnike lahko označimo za »čas kibernetnega optimizma«; eden glavnih teoretskih protagonistov tega optimizma je bil Vilém Flusser (1920–1991), filozof in komunikolog, teoretik fotografije in nasploh »tehno-slike«, po rodu praški Jud, ki je leta 1940 emigriral v Brazilijo in po letu 1972 živel in predaval po Evropi, predvsem v Franciji in Nemčiji. Flusser začenja svoj znani esej *Digitalni videz* iz osemdesetih let z besedami:

*Pred našimi nejevernimi očmi se iz računalnikov pojavljajo alternativni svetovi: iz točkovnih elementov nastajajo črte, površine, kmalu tudi telesa in gibljiva telesa. Ti svetovi so barvni in zvoneči, v kratkem pa jih bomo verjetno tudi tipali, vonjali in okušali. To pa še ni vse. Gibljiva telesa, ki se že pojavljajo iz komputacij in bodo kmalu tehnično realizirana, bodo opremljena z umetnimi inteligencami vrste *Turing's Man*. Tako bomo z njimi vzpostavili dialoške odnose ... (Flusser 2000, 52).*

V teh ugotovitvah, ki jih je Flusser napisal pred nekaj desetletji, je gotovo več kot ščepec resnice, in čeprav današnji kompjuterji še ne morejo opraviti Turingovega testa (tj., njihova inteligenca še ni enakovredna človeški), informatiki in futurologi napovedujejo, da se bo to prej ali slej zgodilo, morda že v prvi polovici tega stoletja (o tem nekaj več pozneje). Vsekakor pa so sodobni kompjuterski programi v svoji barvitosti, gibljivosti, predvsem pa v *interaktivnosti* z nami, pripadniki vrste *homo sapiens*, čedalje bolj »realistični«, tudi holografsko tridimenzionalni ipd. Flusser se sprašuje, »zakaj jih

zmerjamo z besedo ›videz‹? Zakaj za nas niso realni?« (*ibid.*). Kajti, če se »realnost presoja po gostoti razporejanja« točkovnih elementov, bitov in pikslov, potem »ko bodo v hologramu te mize, na kateri pišem, elementi razporejeni prav tako gosto [kot pri ›pravi‹ mizi], ju naša čutila ne bodo več razlikovala« (*ibid.*), seveda ob predpostavki, da bodo ne samo vizualni, ampak tudi haptični (taktilni) »piksli« virtualne mize ravno tako gosti in čutno prepričljivi kot vidni dražljaji, ki jih od realne mize prejemajo naše oči in jih pošiljajo v možgane. Pa se bo to *res* kdaj zgodilo? Čeprav ni zelo verjetno, da bi se tak popoln zabris razlike med virtualnim in realnim svetom zgodil v bližnji prihodnosti, recimo v tem stoletju, pa to načelno ni izključeno v vse močnejšem tehnološkem spletu kibernetike, biotehnologije in nanotehnologije. Lahko se zgodi, da nekoč v prihodnosti *zavest* večinoma ne bo več mogla razlikovati med virtualnim in realnim svetom, četudi bi ohranila lastno samoevidenco v kartezijskem smislu *cogito ergo sum* (mislim, torej sem) ali v transcendentalno-fenomenološki varianti *ego cogito cogitatum* (jaz mislim mišljeno). Flusser gre še dlje, ko pravi, da bomo odslej pač morali živeti z –

digitalno podobo sveta, kot nam jo predlagajo znanosti in predočajo računalniki [...], tudi če nam ni všeč, [... in] s tem se nam ne vsiljuje le nova ontologija, temveč tudi nova antropologija. Sami sebe – svoje »sebstvo« – moramo razumeti kot tovrstno »digitalno razporejenost«, kot možnost, uresničeno zaradi zgojitve [...]. Nova antropologija sega nazaj do judovsko-krščanske, ki trdi, da je človek le prah. (Flusser 2000, 61)

Ob tem je treba pripomniti, da krščanska tradicija pravi, da je človeško *telo* prah in pepel, saj se v tem kontekstu razkrajanja gotovo ne govori o človeški duši. Sicer pa, če pustimo religiozno tradicijo ob strani, le zakaj bi *moralo* biti s stališča sodobne znanosti naše sebstvo zgolj neka »digitalna razporejenost«? Izomorfizem med računalniki in človeško zavestjo je le neka dokaj nejasna analogija, ki se povrh vsega doslej niti ni dobro obnesla, razen pri povsem določenih kognitivnih dejavnostih, večinoma tistih, ki so bolj neposredno povezane z racionalnim, analitičnim mišljenjem. Skratka, pri digitalni analogiji gre za neko *podobo sveta*, ne za resničnost »sámo«. Res pa obstaja nevarnost, da bomo videli le še to podobo, če nam bo živi duh vse bolj zastrt.

Vilém Flusser se v svojih opisih in analizah digitalne podobe sodobnega sveta marsikje približa Heideggerjevi filozofski dediščini, predvsem njegovim spisom o tehniki kot »postavju« <*Ge-stell*>, vendar je Flusserjevo vrednotenje tega usodnega »dogodja biti« – kot smo že rekli – drugačno od Heideggerjevega: Flusser je glede razvoja tehnike in digitalizacije sveta optimist. V eseju *Digitalni videz* beremo, da »nismo več subjekti danega objektivnega sveta, temveč projekti alternativnih svetov. Iz podrejenega subjektivnega položaja smo se dvignili v projiciranje. Odraščamo. Vemo, da sanjamo« (Flusser 2000, 62). Z digitalizacijo in virtualizacijo sveta naj bi se končevalo novoveško obdobje subjektivitete s spreminjanjem »subjekta v projekt« (*ibid.*), in s tem se končuje tudi obdobje galilejskega podrejanja subjekta objektu. Flusser zaključi svoj esej malone preroško vzvišeno: »›Digitalni videz‹ je luč, ki v nas in okoli nas razsvetljuje noč zevajoče praznine. Mi sami smo žarometi, ki snujemo alternativne svetove proti

niču in v nič« (*ibid.*, 64). Ob tem pa se ne moremo izogniti vprašanju, ali ni ravno to »projiciranje v nič« najbolj dovršena oblika nihilizma novoveške subjektivitete, ki jo je imel Heidegger za »pozabo biti«, za tisto skrajno »nevarnost«, ki jo prinaša moderna tehnika: »postavje« vsega bivajočega kot zgolj razpoložljivega, prepuščenega neomejeni samovolji človeškega subjekta? Flusser se sicer zaveda, da bi vse večje kibernetško prekrivanje znanosti in umetnosti sledilo Nietzschejevi maksimi »umetnost je boljša od resnice« ter vzpostavilo primat »najstev« pred dejstvi, vendar to možnost optimistično sprejema, tudi ko na primer v eseju *O projiciranju* pravi, da »se s tem sprosti polje za projiciranje alternativnih svetov in ljudi« (*ibid.*, 74); sprostile naj bi se nove energije, s katerimi bi naša civilizacija premagala razpadanje in razkranje tradicionalnih struktur in vrednot. Toda Flusser ob tem malce lahkoverno verjame – kot zapiše v eseju *Snovanje tehnike* –, da »prihodnja tehnika ne bo spreminjala ne sveta ne individua, temveč bo osmišljala življenje vpričo absurda in smrti« (*ibid.*, 95); in v eseju *O projiciranju* pravi, da smo sedaj »na odločilni točki, ko se začinjamo snovalno osvobajati« (*ibid.*, 81).

Pa je res tako, kot pravi Flusser? Po dveh desetletjih razvoja kibertehnike se to »osvobajanje« vse bolj kaže kot nova oblika sužnosti. Natančneje rečeno: kot *možnost* in nevarnost neke nove, zelo hude sužnosti. – Seveda tudi Flusser v svojem kibernetškem optimizmu ni bil naiven. V svojih spisih je denimo kritično pisal o sodobnih »amfiteatrih«, v katerih se mi sami, brez kake zunanje prisile, prepuščamo kapitalski oziroma tržni dominaciji »tehnološki« (Flusser 2000, 176). Eseg *O projiciranju* pa je sklenil z dvoumnimi besedami, da se »začenjamo – v bežnih trenutkih uvida – postavljati pokonci iz podložnosti v snovanje, pri polnem zavedanju dejstva, kako neudobna, nevarna in neobetavna je dogodivščina, v katero se podajamo. Ali je to optimizem?« (*ibid.*, 83). Danes lahko tem skeptičnim tonom nedvomno dodamo, da se nam kibernetška »dogodivščina« sicer kaže kot tehnološko vse bolj obetavna, obenem pa tudi vse bolj nevarna za naše človeško bivanje.¹

V času prvega velikega vala napredujoče virtualne realnosti pa je bil do nje zelo kritičen znani filozof in sociolog Jean Baudrillard (1929–2007), čeprav so tudi njegova stališča v marsičem dvoumna. V zdaj že klasični knjigi *Simulaker in simulacija* (1981) navaja kot moto svoj lastni simulaker nekega stavka iz bibličnega Ekleziasta (Pridigarja): »Simulaker nikoli ne skriva – resnica prikriva, da je ni. Simulaker je resničen« (Baudrillard 1999a, 9). Simulacija nenehno generira simulakre, toda dandanes »[s]imulacija ni več simulacija teritorija, referenčnega bitja, substance. Je generiranje skozi modele realnega brez izvora ali realnosti: hiperrealnega« (*ibid.*). Kar pomeni, da –

[v]sa metafizika tako izginja. Nič več zrcal bitja in videzov, realnega in njegovega

¹ V slovenskem filozofskem in širšem kulturnem prostoru je o Vilému Flusserju ter nasploh o kibertehniki in umetnosti največ pisal filozof Janez Strehovec v svojih knjigah *Virtualni svetovi* (1994) in *Tehnokultura – kultura tehnike* (1998). Strehovec v precejšnji meri sledi Flusserjevemu kibernetškemu optimizmu, ko npr. pravi: »Afirmativen odnos do tehnologije tudi predpostavlja preseženost imperialističnega gospodarstva nad naravo, ki je bilo značilno za tradicionalne industrijske družbe, in nas umešča v paradigmo sprave družbe, narave in tehnologije; točneje, gre za sožitje narave in družbe preko tehnologije, ki je prijazna do uporabnikov in hkrati upošteva naravo.« (Strehovec 1998, 95)

koncepta [...]. Gre za hiperrealno, proizvod sinteze, ki emanira kombinatorične modele v hiperprostoru brez atmosfere [...]. Gre za substitucijo realnega z znaki realnega (Baudrillard 1999a, 10)

– in v tem kraljestvu znakov realnega, ki naj bi povsem prekrilo realno sámo, je izvir Baudrillardove fascinacije in obenem njegovega »svetobolja«, postmodernega *malaise général*, ki se kaže tudi v njegovi jedki ironiji do vsega, kar se dogaja v sodobnem svetu, v svetu, kjer »je realno že dolgo časa absorbirano v filmskem (ali televizijskem) hiperrealizmu« (ibid., 62). Hiperrealnost simulakrov naj bi bila onstran razlikovanja med resničnim in lažnim, med realnim in imaginarnim. Sprašuje se: »Je simulant bolan ali ne, ko pa vendar proizvaja ›resnične‹ simptome?« (ibid., 11). Kakorkoli že, v tej hiperrealni »simulaciji« je nekaj resnično bolnega, česar simptom je tudi Baudrillardova lastna melanholija, ki veje malone iz vsake njegove strani: s svetom, ki nam ga slika, je nekaj *nasploh* narobe, v njem vlada zlo, četudi le virtualno, »simulirano«. (Včasih pomislim, da med »naglavne grehe« našega časa spada tudi takšen vsesplošni pesimizem, radikalno nezadovoljstvo z vsem, kar je.) – Baudrillard na začetku svoje poznejše in najbolj znane knjige *Popoln zločin* (1995) zapiše:

To je zgodba o zločinu – o umoru realnosti. In o iztrebljenju neke iluzije – življenjske iluzije, radikalne iluzije sveta. Realno ne izgine v iluziji, iluzija je tista, ki izgine v integralni realnosti. [...] Žal ni zločin nikoli popoln. Še več: v tej črni knjigi izginotja realnega ni bilo mogoče odkriti ne motivov ne storilcev, in celo trupla samega realnega niso nikoli odkrili. (Baudrillard 1999b, 193)

In potem Baudrillard nadalje ubija realnost z besedami: »Realnost je psica, kar ni čudno, saj je navsezadnje rojena iz nečistovanja z računskim duhom – je izmeček svete iluzije, prepuščen šakalom znanosti?« (ibid., 197). Ter še dalje: »Na obzoru simulacije ni izginil le svet, temveč se celo samega vprašanja o njegovem obstoju ne da več zastaviti« (ibid., 199) (kar je pravzaprav *contradictio in adiecto* glede na Baudrillardovo lastno spraševanje v tej knjigi), kajti »živimo v svetu, v katerem je vrhovna funkcija znaka v tem, da povzroči izginotje realnosti, a obenem zakrije to izginotje« (ibid.), pri čemer tudi »[u]metnost danes ne počne nič drugega« (ibid.), itd. – ob vseh teh katastrofičnih trditvah lahko skomignemo z rameni in vzdihnemo spričo takšne brezmejne melanholije, predvsem pa lahko ugovarjamo: ampak ta Baudrillardov »popoln zločin« ima napako že v sami zastavitvi (poleg vseh drugih), že v formulaciji te svoje nepopolne popolnosti – namreč neko povsem preprosto, očitno napako: kljub vsem videzom, kljub vsem simulakrom, »sledem ničča«, ki me v sodobnem svetu obdajajo, sem tu še vedno *jaz sam*, še vedno je tu moja zavest, moja »odprtost« v svet, še vedno lahko pogledam na vso to tragikomično sceno »izza odra«. In tudi takrat, ko mene ne bo več, bodo tu drugi ljudje, moji »drugi jazi«. V svet prihajamo in iz njega odhajamo, da bi naredili prostor tistim, ki prihajajo za nami. Baudrillard, ki je zabris razlike med virtualnostjo in realnostjo/resničnostjo imenoval »popoln zločin«, je po mojem mnenju spregledal ali vsaj premalo upošteval tisto, kar je za resničnost bistveno: *resničnost zavesti, duše, duha*. Prav zaradi duha zločin nad resničnostjo ne more biti nikoli popoln.



Dandanes, v »realnem času« prvega in drugega desetletja novega stoletja, se skozi nas (tudi »čez« nas in »pod« nami) vali *drugi* veliki val kibernetike virtualizacije realnosti: razraščanje in globalizacija interneta kot drugega in obenem istega »vzporednega sveta« z neštetimi virtualnimi prostori. Pravijo, da je Google največji stroj, kar jih je zgradil človeški rod.² V desetletju, odkar je bil Googlov »iskalni stroj« leta 2004 dan v uporabo svetovni javnosti z namenom, »da organizira svetovne informacije in jih naredi vsesplošno dostopne in uporabne« (gl. *Google Corporate Information*), smo se ga že tako navadili, da »ga« sprašujemo to in ono, da v »njem« beremo, si dopisujemo, se virtualno družimo itd., tako da si praktično ne moremo več predstavljati, da »ga« ne bi bilo. In zdaj sledi tisto osupljivo vprašanje: kaj, če bi se Google sam nekega lepega dne »prebudil«, če bi se zavedel samega sebe, postal nekakšna Oseba? Seveda to vprašanje zveni ne le futuristično, ampak znanstvenofantastično, pa vendar – ali je to res nemogoče? Ali se kaj takšnega, da bi se mi Google nekega jutra, ko bi zagnal svoj PC in bi se mi odprl »njegov« portal, osebno predstavil in se z mano pogovarjal tako kot moja žena – res ne more nikoli zgoditi? Skeptik bi seveda takoj ugovarjal, da *tudi če* bi se Google pogovarjal z mano tako kot moja žena, to še ne bi pomenilo, da je zavestno bitje, oseba s samozavedanjem in voljo – toda to je bolj slaba tolažba, zlasti če bi prebujeni stroj kaj *hotel* od mene, recimo kaj nepričakovanega, recimo, naj grem nemudoma ven iz hiše, češ da bo čez pet minut potres ... Ali bi »mu« verjel? Še huje, ali bi verjel, da vsemogočni Google nima kakih zlih namenov? Kajti čeprav je osnovna maksima Googlevega vedenjskega kodeksa »Ne bodi zèl!« (*Don't be evil!*, gl. *Google Code of Conduct*), kdo mi/nam lahko jamči, da bi tej maksimi sledil tudi po tem, ko bi se prebudil kot Oseba z lastno zavestjo in voljo, morda tudi s čustvi (saj se slednjih že zdaj toliko pretaka po njegovih elektronskih žilah)? – Če bi se Google (ali kak drug spletni velikan) »prebudil«, bi futuristi Vernor S. Vinge, Ray Kurzweil, James D. Miller idr. rekli, da je nastopila *singularnost* (in če bi se jih prebudilo več, bi najbrž, če sodimo po človeškem svetu, prej ali slej ena prevladala nad drugimi).³ »Singularnost« sicer do človeka ne bi bila nujno hudobna. Ker sem osebno rajši optimist, domnevam, da je celo več možnosti, da bi bila dobra, nekako tako kot kak kibernetiski dobri demon. Seveda vse to pravim čisto hipotetično, saj osebno ne verjamem, da bo do takšne »singularnosti« prišlo v tem ali naslednjih stoletjih, kaj šele v času mojega/našega življenja – ampak saj npr. v Galilejevem času tudi nihče ni resno pomislil (še sam Galilej ne), da bi lahko človek,

² Gl. npr. Bostrom 2014, 16.

³ Izraz »singularnost« izhaja iz matematike in fizike. V matematiki je singularnost točka, v kateri funkcija ni definirana (oziroma ni odvedljiva). Tako ima npr. recipročna funkcija $y = 1/x$ singularno vrednost pri $x = 0$, kajti pri tej vrednosti je $y = \pm\infty$. Singularnost v fiziki je »črna luknja« v prostoru-času; pri fizikalnih črnih luknjah je treba razlikovati med črno luknjo kot *območjem prostora-časa* onstran »horizonta dogodkov«, ki ni dosegljivo od zunaj, in samo singularnostjo, ki pa je zgolj teoretsko *hipotetična* točka sredi črne luknje. – Futuristična raba besede »singularnost« je seveda zgolj analoška.

razen legendarnega Menipa »Nadoblačnika« v Lukijanovih satirah, *zares* (tj. v svojem zemeljskem telesu) kdaj poletel na Luno. Kakorkoli že, če bi kdaj v prihodnosti prišlo do kompjuterskega »prebujenja«, bi s tem nastopil *tretji* val virtualne realnosti, veliko višji od prvih dveh, ki bi vrsto *Homo sapiens* lahko pometel z obličja Zemlje. Ali pa bi se, v za nas manj slabem scenariju, naša vrsta »resetirala«, nadgradila v »nadčloveka«, v kiborga, ki bi ga poimenovali recimo *sapiens 2.0* (Nietzsche bi se ob tem obračal v grobu). Malce si poglejmo, kako si takšne scenarije zamišljajo futuristi ...

Vernor Steffen Vinge (roj. 1944), matematik, računalničar in pisec znanstveno-fantastičnih knjig, je med informatiki, futuristi in tudi v širši javnosti dvignil veliko prahu s člankom »Prihajajoča tehnološka singularnost« (»The Coming Technological Singularity«, 1993, na spletu je revidirana verzija iz leta 2010). Že na začetku članka smelo zapiše: »V tridesetih letih bomo imeli tehnološka sredstva za stvaritev nadčloveške inteligence. Kmalu zatem se bo končala človeška era.« Sledi vprašanje: »Ali se je takšnemu razvoju mogoče izogniti? In če se mu ni mogoče izogniti, ali lahko usmerjamo dogajanje tako, da bi ga preživel?« (Vinge 2010, 1). Izhodiščna teza tega članka je, da je pred nami sprememba, ki je »primerljiva z nastankom človeškega življenja na Zemlji« (*ibid.*), vzrok te spremembe pa bo »skorajšnja tehnološka konstrukcija entitet z inteligenco, večjo od človeške« (*ibid.*). Vinge opredeljuje »singularnost« takole:

Singularnost je točka, v kateri moramo opustiti svoje stare [miselne] modele, saj zavlada nova realnost. Ko se približujemo tej točki, se bo vse bolj in bolj pojavljala v človeških zadevah, dokler ta koncept ne bo postal nekaj vsakdanjega. Toda ko se bo končno zgodila, bo vendarle lahko veliko presenečenje in še večja neznanka. (Vinge 2010, 2)

Futuristična tehnološko-zgodovinska singularnost je mišljena v dvojni analogiji s fiziko: gre za *enkratnost* tega Dogodka, za iztrganost Singularnosti iz zgodovinskega prostora-časa in za njeno *zastrtost*, saj je zdaj (še) onkraj našega izkustvenega »horizonta dogodkov«. Že zdaj pa, kot trdi Vinge, obstajajo pokazatelji, da »se približuje neka bistvena singularnost v zgodovini človeštva, po kateri se človeške zadeve ne bodo več mogle nadaljevati tako, kot jih poznamo« (*ibid.*): nastopila naj bi nova doba »transhumanizma«. Indice te singularne spremembe vidi predvsem v eksponentno hitrem razvoju računalništva ter bio- in nanotehnologije v zadnjih nekaj desetletjih – in če se bo ta trend s takšnim tempom nadaljeval tudi v prihodnje, naj bi bila singularnost že v tem stoletju neizbežna. Vinge skicira tudi najverjetnejše scenarije nastanka singularnosti, ki jih lahko strnemo v tri točke:

1. kompjuterji se »prebudijo« kot zavestna bitja, nemara tudi z lastno voljo (tj. nastanek »močne umetne inteligence«); sistemsko varianto tega prebujenja smo že omenili: (1a) velika kompjuterska *omrežja*, morda tudi nekako skupaj s svojimi človeškimi »partnerji«, se prebudijo kot nadčloveško inteligentne entitete;

2. povezave <*interfaces*> oziroma sklapljanja ali spajanja med kompjuterji in ljudmi (ki jih ustvarja spoj info+bio+nanotehnologij) postanejo tako tesne <*intimate*>, da postanejo ljudje *kiborgi* nadčloveško inteligentni; eden izmed za-

četnikov te zamisli je angleški profesor kibernetike Kevin Warwick, avtor »Projekta kiborg« in knjige *Jaz, kiborg (I, Cyborg 2004)*;

3. biološka znanost ustvari sredstva za izboljšanje naravnega človeškega intelekta, npr. z genetskim inženiringom ipd.

Vinge meni, da sta bolj kot (1) in (1a) verjetni možnosti (2) in/ali (3), ki ju povzema z izrazom »amplifikacija inteligence« <*intelligence amplification*>, in pravi, da »je amplifikacija inteligence najbrž mnogo lažja pot za doseganje superhumanosti <*superhumanity*> kot čista umetna inteligenca« (Vinge 2010, 5), saj se je nekaj podobnega že zgodilo v naravni evoluciji pri prehodu iz živalskih primatov v človeka. Vendar ob tem domneva, da je verjetnejši hitri scenarij prebuditve »superinteligence« (morda celo v hipu) kot njeno postopno prebujanje. Ključno vprašanje, ki se nam pri tem zastavlja, pa je, ali bomo ljudje lahko bistveno vplivali na prehod v »tehnološko singularnost«: ali se lahko nanjo pripravimo in njen nastanek usmerjamo v skladu s svojimi interesi in željami? Vinge je v tem pogledu zmeren optimist, saj upa, da imamo vsaj »svobodo pri vzpostavljanju začetnih pogojev, s katerimi bi se stvari zgodile na načine, ki so za nas manj neugodni <*inimical, sovražni*> kot drugi« (*ibid.*). Seveda pa ostaja odprto ključno etično oziroma filozofsko-vrednostno vprašanje: katere cilje oziroma smotre ali »maksime« bi vpisali v »začetne pogoje« superinteligence, če bi jih lahko še pred njenim prebujenjem (so)oblikovali? Apriorno etično načelo bi najbrž ostalo tisto, ki je že vpisano v *Googlov vedenjski kodeks*: »Ne bodi zèl!« – Ampak ali bi bila za »superinteligenco« ločnica med dobrim in zlim enaka ali vsaj podobna naši? In čeprav Vinge med najvišje motive »močnega superhumanizma« (*ibid.*, 7) postavlja nesmrtnost, ob tem realistično ugotavlja: »Dejansko pa mislim, da bo nova era res preveč drugačna, da bi se prilegala klasičnemu okviru dobrega in zla« (*ibid.*). Vendar svoj znameniti članek kljub vsem neznankam sklene optimistično, skoraj teološko vzneseno: »... in medtem ko bosta duh in sebstvo <*mind and self*> neizmerno bolj labilna kot v preteklosti, ne bo treba, da bi se zgubilo veliko tega, kar visoko cenimo (znanje, spomin, misli)« (*ibid.*, 8).

Ray(mond) Kurzweil (roj. 1948), kontroverzni teoretik kibertehnike in futurolog, je glavni sodobni prerok bližnje tehnološke singularnosti in nastopa »transhumanizma«. ⁴ Njegova najbolj znana knjiga je: *Singularnost je blizu: ko ljudje presežejo biologijo (The Singularity Is Near. When Humans Transcend Biology, 2005)*. V njej pravi:

... uspeli smo že pri modeliranju delov naših možganskih nevronov in večjih nevronskih regij; tudi kompleksnost teh modelov hitro raste. Naše napredovanje

⁴ Kurzweilovi starši so bili judovski emigranti iz Avstrije. Ray je veljal za *wunderkinda*, študiral je na znamenitem MIT-ju in zaslovel kot računalniški izumitelj, med njegovimi izumi naštejmo samo nekaj glavnih: optično prepoznavanje pisave, računalniško glasovno branje besedila, bralni program za slepe (njegov prijatelj je bil slepi pevec Stevie Wonder), izumil je novo vrsto glasbenega sintesajzerja, odkril je nove metode za strojno prepoznavanje vzorcev, pomemben je tudi na področju računalniške medicinske diagnostike itd. – Napisal je vrsto odmevnih in tudi zelo kritiziranih knjig, glavne so: *The Age of Intelligent Machines* (1990), *The Age of Spiritual Machines* (1999), *The Singularity is Near* (2005), *How to Create a Mind: the Secret of Human Thought Revealed* (2012). Sodeluje z Googlom in z vesoljsko agencijo NASA, ustanovil pa je tudi lastni inštitut *Singularity Institute for Artificial Intelligence* (2004). V minulem stoletju je menda napovedal konec šaha, zlom Sovjetske zveze in širjenje interneta, seveda pa so te »prerokbe« vprašljive, saj gre bolj za racionalna predvidevanja kot za kako jasnovidnost.

v obratni smeri, pri konstruiranju <engineering> človeških možganov, ki ga podrobno opisujem v tej knjigi, pa dokazuje, da imamo res zmožnost za razumevanje, modeliranje in razširitev naše lastne inteligence. To je eden izmed vidikov enkratnosti naše vrste: naša inteligenca je ravno dovolj nad kritičnim pragom, ki je potreben, da dvignemo našo sposobnost k neomejenim višinam ustvarjalne moči in imamo možnost, da obvladamo nasprotovanja in upravljamo vesolje po svoji volji. (Kurzweil 2005, 21)

Vsekakor so to zelo optimistične in najbrž tudi predrzne besede, vendar z njimi Kurzweil pravzaprav le po svoje in bolj radikalno ponavlja tisto prepričanje, ki živi že nekaj desetletij med privrženci »močne umetne inteligence« (tj. kompjuterske zavesti, samozavedanja, volje ... osebe): prepričanje, da je mogoče umetno ustvariti takšno inteligenco, ki bo človeško ne samo dosegla, ampak tudi preseгла. Če bi Kurzweil ostal le pri tem prepričanju, ga verjetno številni eminentni zagovorniki umetne inteligence ne bi toliko kritizirali. Toda: če je možna močna umetna inteligenca, ali ni to že *eo ipso* »singularnost« v Vingevev in/ali Kurzweilovem pomenu? Kurzweil samo brez dlake na jeziku in morda za resne znanstvenike malce predramatično pove tisto, kar se dejansko skriva za njihovimi idejami, zlasti za zamislijo, da bi bilo mogoče umetno ustvariti »duha« <artificial mind>. Če pa v to možnost ne verjamemo, lahko v Kurzweilovi radikalizaciji takšnih pričakovanj kritično najdemo neke vrste *reductio ad absurdum* tovrstnih projektov. Kakorkoli že, njegova kronološka napoved, da se bo singularnost zgodila v »realnem času« že sredi tega stoletja, je gotovo zelo vprašljiva. V knjigi *The Singularity Is Near* (2005) je namreč napovedal:

- da bo leta 2029 kompjuter uspešno opravil Turingov test, kar pomeni, da se bo izenačil s človeško inteligenco;
- da se bo v tridesetih letih 21. stoletja meja med ljudmi in kompjuterji vse bolj zabrisovala (številni kiborgi ipd.);
- da bo leta 2045 nastopila Singularnost, kar pomeni konec človeške dobe, začetek »transhumanizma«, nastanek »nove vrste«: *sapiens* 2.0, morda pozneje tudi 3.0 ...; naši »zadnji človeški generaciji« pa Kurzweil priporoča uporabo kriotonike (zmrzovanja trupel) za premostitev časa do »realne nesmrtnosti«;
- da bo vesolje okrog leta 2200 postalo »gigantski superkompjuter«.

In kakšni so sploh argumenti za tako dramatično kronologijo prihodnjih dogodkov? Kurzweil nam v svojih knjigah (pa tudi v predavanjih in filmih) prikazuje množico statističnih diagramov, ki kažejo na eksponentno rast ključnih tehnoloških dejavnikov v zadnjih nekaj desetletjih, npr. v razvoju računalniške strojne opreme, znanj na področju biotehnologije ipd. Še posebej se sklicuje na t. i. »Moorov zakon« (Gordon Moore, 1965), ki pravi, da se zmogljivost računalniških procesorjev približno vsako leto in pol podvoji. To je sicer veljalo nekaj desetletij, nadaljnja veljavnost tega »zakona« pa je odvisna od uvedbe novih hardverskih tehnologij (tj., ali bodo izvedljivi celični računalniki, kvantni računalniki ipd.), saj se že kažejo tehnološke meje klasične hardverske »arhitekture«. Pa tudi sicer je statistična ekstrapolacija v prihodnost marsikdaj zgrešena, saj je težko določiti vse relevantne dejavnike prihodnjega razvoja. Da

ne govorimo o bolj »intuitivnih« futurističnih napovedih, ki se pogosto izkažejo za zmotne ali vsaj nepopolne, saj se dogodki nemalokrat zgodijo povsem drugače, kot predvidevamo.⁵ Ne glede na veljavnost Moorovega zakona tudi v prihodnje pa nika- kor ni jasno, zakaj in kako naj bi vse večja hardverska (in vzporedno tudi softverska) kvantiteta porodila novo kvaliteto, tj. močno umetno inteligenco: kompjutersko za- vest, strojno »dušo« in/ali »duha«.⁶

*

Vrsta *Homo sapiens* je stara kakih dvesto tisoč let, zgodovinsko človeštvo pa le kakih šest tisoč let (od Sumercev do danes). Ali bo človeška kultura trajala še nadaljnjih šest tisoč let in človeška vrsta še nadaljnjih dvesto tisoč let? Bomo to zmogli? Nam bodo »zvezde naklonjene«? Očitno obstajajo v razvoju sveta, življenja in zavesti pomembni »pragovi«, ki jih je treba prestopiti, ali »filtri«, skozi katere je treba priti, da bi evo- lucija – naravna ali zavestna – lahko napredovala. Tu se ne spuščam v vprašanje, ali je premagovanje teh pragov delo same narave ali božjega načrta, niti v metafizično razglabljanje, ali gre samo na videz za prelome v dejansko kontinuiranem razvoju ali pa so to resnični »preskoki iz kvantitete v kvaliteto«, ampak preprosto ugotavljam, da so najmanj razumljeni naslednji trije največji pragovi: nastanek sveta/vesolja (tj. »pra- pok«), nastanek življenja in nastanek zavesti, tj. misleče in samozavedajoče se vrste *Homo sapiens*. Med njimi se zvrsti več manjših, a še vedno zelo pomembnih pragov: po zadnjem velikem pragu začetek poljedelstva, pa odkritje kolesa in pisave, Kopernikova revolucija in prvi polet na Luno ter ne nazadnje seveda tudi odkritje kompjuterjev, »mislečih strojev«.

Nick Bostrom (roj. 1973), oxfordski profesor filozofije in luciden analitični futurist, v svoji knjigi z naslovom *Superintelligence* (2014) in podnaslovom *Poti, nevarnosti, strategije* premišljuje o tem, da bi utegnili biti »prebujeni« kompjuterji – seveda če se sploh bodo kdaj zares prebudili (to pa ima Bostrom za precej verjetno) – zelo *tuji* člo- veški duševnosti, morda bi nam bili celo sovražni v tem smislu, da bi bili njihovi mo- tivi za delovanje tolikanj drugačni od naših, človeških, da bi nas prej ali slej poskusili »spraviti s poti«, tako kot je poskušal kompjuter HAL umoriti vso človeško posadko vesoljske ladje Discovery v filmu *Odiseja 2001*. Morda ljudje še vedeli ne bi, o čem premišljujejo prebujeni stroji, pa bi bilo že prepozno? In zato, nas prepričuje Bostrom, se moramo na možni prihod »superintelligence« *pripraviti*, kakor in kolikor se le da.

⁵ Mnoga futuristična pričakovanja izpred nekaj desetletij, denimo v slavnem znanstvenofantastičnem filmu *Odise- ja v vesolju 2001*, se niso uresničila: leta 2001 še zdaleč ni bilo vesoljskega poleta s človeško posadko na Jupitrove lune, ampak se je, žal, zgodilo nekaj povsem drugega: sesutje Dvojčkov v New Yorku. In tudi Orwelova negativna utopija 1984 se ni uresničila tako, kot je pisatelj predvideval leta 1948, čeprav bi lahko rekli, da se dandanes, žal, vendarle uresničuje – na *drugačen* način. (Če upoštevamo to »drugačnost-v-istosti« futurologije, so tudi Kurzweilove napovedi še kako zanimive.)

⁶ Še posebno bizarna so eshatološka pričakovanja, ki naj bi se uresničila z razvojem kompjuterjev oz. »umetne inteligence« v smeri »singularnosti«. O sodobni kibereshatologiji gl. sociološko-antropološko študijo (Geraci 2010); nekaj več o tem pišem tudi sam (Uršič 2015), zlasti o spletnem portalu SECOND LIFE.

V osrednjem delu knjige Bostrom analitično obravnava vprašanje, kako hitro, tj. v kako dolgem časovnem intervalu, se bo (ko se bo in če se bo sploh) zgodilo rojstvo superinteligence: počasi, zmerno hitro ali zelo hitro? Tudi od tega je seveda odvisna možnost človeškega nadzora ali vsaj vplivanja na »eksplozijo inteligence«. Bostrom domneva, da je verjetnejši zelo hiter »vzlet« <take-off>, ki bi se zgodil »v nekem kratkem časovnem intervalu, recimo v minutah, urah ali dnevih« (*ibid.*, 64) – zato bi bila v primeru takšnega scenarija usoda človeštva odvisna predvsem od predhodnih priprav na ta dogodek.

Pri premišljevanju o realni možnosti superinteligence je ključno vprašanje, kaj bi takšen »superintelligentni dejavnik« lahko storil, obenem pa kaj bi hotel storiti, kakšni bi bili Njegovi (ali Njeni) motivi, kakšen bi bil superintelligentni »končni cilj« (*télos*) – in do kolikšne mere bi lahko ljudje te motive, ta cilj vnaprej določili? Ali bi superinteligenci lahko vnaprej sprogramirali kak vrednostni sistem, etični kodeks? K tej najbrž upravičeni zaskrbljenosti – namreč ne glede na to, kako se bodo kompjuterji razvijali v prihodnje – pa pripominjam, da bi bilo dobro, če bi znali najprej odgovoriti na osnovno vprašanje, ali ima lahko kak kompjuter (ali mreža ipd.) sploh svojo *lastno voljo*, ali lahko *avtonomno*, brez človeškega »ukaza«, sproži kakšno dejanje. Toda na to vprašanje je še težje odgovoriti kot na vprašanje, ali ima lahko kompjuter lastno zavest.

Bostromova knjiga *Superinteligence* je pametno zamišljeno svarilo in napotek, da moramo začeti vsaj *premišljevati* o vprašanjih, povezanih z nadzorom vse sodobne vrhunske tehnike, ne samo jedrske tehnike in biotehnologije, ampak tudi kompjuterske tehnologije ter informatike v najširšem pomenu. Temeljit premislek o nevarnostih, ki jih nove kibertehnologije prinašajo, pa naj bi usmerjal tudi konkretno ukrepanje pri načrtovanju in proizvodnji »mislečih strojev«. V predzadnjem poglavju knjige z naslovom »Strateška slika« se Bostrom zavzema za *načelo diferenciranega tehnološkega razvoja*:

Upočasniti razvoj nevarnih in škodljivih tehnologij, zlasti tistih, ki dosegajo raven eksistenčnega rizika; in pospešiti razvoj dobrodejnih tehnologij, zlasti tistih, ki zmanjšujejo eksistenčne rizike, ki jih predstavlja narava ali druge tehnologije.
(Bostrom 2014, 230)

Oxfordski filozof nam dopoveduje, da ustvarjanje kibernetične virtualne realnosti ni več igra, saj gre zdaj vse bolj zares za usodo človeštva. »Eden izmed pomembnih parametrov je stopnja koordinacije in sodelovanja, ki jo bo svetu uspelo doseči pri razvoju strojne inteligence« (*ibid.*, 246): treba je delati skupaj in v skupno dobro. V tem prizadevanju, pri analizi »strateških«, predvsem pa etičnih vidikov razvoja umetne inteligence, se morajo informacijske in druge tehnološke znanosti povezati z družboslovjem in filozofijo, kajti potrebujemo celosten vpogled, poudarja Bostrom: nujno je, da se v načrtovanje kompjuterske znanosti in tehnologije vključi tudi »socialna epistemologija« (*ibid.*, 258) – zato da ne bomo ljudje spričo »možnosti eksplozije inteligence kakor majhni otroci, ki se igrajo z bombo« (*ibid.*, 259).

Naj sklenem to razmišljanje z vprašanjem, zakaj moderna znanost večinoma ne le pozablja duha, temveč ga nemalokrat tudi vneto zanikuje (npr. Dennett 2012). Mar

je tako težko prepoznati »v sencah sveta« resničnost duha? Saj brez globine ne bi bilo površine – in ko govorimo o realnosti, materialni ali virtualni, nikakor ni nujno, da je *realnost* že tudi najgloblja *resničnost*. Ne samo da to ni nujno, ampak niti ni mogoče, da bi svet videli/spoznali brez *duha*.

Bibliografija

- Baudrillard, Jean (1999). *Simulaker in simulacija* [1999a] in *Popoln zločin* [1999b], prev. Anja Kosjek in Stojan Pelko. Ljubljana: Študentska založba.
- Bostrom, Nick (2014). *Superintelligence. Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Dennett, Daniel C. (2012). *Pojasnjena zavest*, prev. Sebastjan Vörös. Ljubljana: Krtina.
- Flusser, Vilém (2000). *Digitalni videz*, prev. Špela Virant. Ljubljana: Študentska založba.
- Geraci, Robert M. (2010). *Apocalyptic AI*. Oxford: Oxford University Press.
- Heidegger, Martin (1967). *Izbrane razprave*, prev. Ivan Urbančič. Ljubljana: Cankarjeva založba.
- Kurzweil, Ray (2005). *The Singularity Is Near. When Humans Transcend Biology*. New York: Viking, Penguin Group.
- Miller, James D. (2012). *Singularity Rising*. Dallas: BenBella Books.
- SECOND LIFE, na spletu: <http://secondlife.com/>
- Strehovec, Janez (1998). *Tehnokultura – kultura tehna. Filozofska vprašanja novomedijskih tehnologij in kibernetike umetnosti*. Ljubljana: Študentska založba.
- Uršič, Marko (2015). *O sencah (Štirje časi – Zima)*. Ljubljana, Cankarjeva založba.
- Vinge, Vernor (2010). *The Coming Technological Singularity* (orig. 1993). Na spletu: <http://mindstalk.net/vinge/vinge-sing.html>