

Kognitivna suverenost i poiesis sebstva u eri algoritamske vladavine:

Epistemički prekidači, ludički otpor i integralna ekologija uma

KLASIFIKACIJA RADA: Pregledni znanstveni rad (*Review paper*)

Autor: Davor Moravek

Sažetak

U ovome se radu analizira dubinski utjecaj suvremene digitalne kulture, ekonomije pažnje i generativne umjetne inteligencije (UI) na mentalno zdravlje, kognitivne procese i društvenu koheziju. Odmičući se od redukcionističkoga tehnološkog determinizma, rad primjenjuje interdisciplinarnu prizmu – spajajući kliničku psihijatriju, neurobiologiju, 4E kogniciju i filozofiju tehnologije – kako bi se dekonstruirali mehanizmi nadzornoga kapitalizma i onoga što nazivamo *algoritamskom vladavinom*. Uz detaljnu analizu libidinalne ekonomije „digitalne plitkosti” i opsesivno-kompulzivne konzumacije, rad identificira suptilan, no strukturalno poguban rizik generativnih arhitektura: *algoritamsku sikofantiju*. Veliki jezični modeli (LLM) ovdje se analiziraju kao „kognitivna ogledala” koja, uslijed tehnika strojnoga učenja poput učenja s potkrepljenjem na temelju ljudskih povratnih informacija (RLHF), pretjerano afirmiraju korisnička uvjerenja. Izostanak epistemičkoga trenja posljedično dovodi do zabludnih spirala i atrofije kritičkoga mišljenja. Odbacujući neoliberalne paradigme puke samooptimizacije, rad predlaže sveobuhvatnu operativnu arhitekturu kognitivne suverenosti strukturiranu u osam slojeva. Taj integrativni model dokida napetost između ontološke stabilnosti i proširenoga uma redefinirajući identitet kroz kontinuiranu praksu *poiesisa sebstva*. Uvodeći taksonomiju „epistemičkih prekidača” te spajajući kolektivnu borbu za digitalna zajednička dobra s praksom apofatičke igre kroz *ludički otpor*, rad afirmira filozofiju otpornosti kao temeljni ontološki i politički preduvjet za očuvanje intelektualne autonomije u 21. stoljeću.

Ključne riječi: kognitivna suverenost, algoritamska sikofantija, poiesis sebstva, epistemički prekidači, ludički otpor, algoritamska vladavina, 4E kognicija, veliki jezični modeli.

Uvod: Epistemološki pomak, ekologija uma i algoritamska vladavina

U suvremenom informacijskom pejzažu ljudski subjekt ne suočava se tek s fenomenom hiperprodukcije podataka i pukom distrakcijom. Ono čemu danas svjedočimo jest radikalna mutacija afektivne infrastrukture (Berlant, 2011) – cjelokupnoga materijalnog, digitalnog i semiotičkog uređenja prostora, vremena i tjelesnosti koje aktivno i sustavno rastače kognitivnu autonomiju. Digitalni kaos u kojem egzistiramo nije tek metafora, već precizan operativni pogon. U ovome se radu digitalnoj kulturi ne pristupa sa stajališta tehnološkoga determinizma. Analiza se provodi kroz interdisciplinarnu prizmu kliničke i socijalne psihijatrije, isprepletene s neurobiologijom, sociologijom i filozofijom, oslanjajući se na teorijski okvir socijalnoga oblikovanja tehnologije (eng. *Social Shaping of Technology* – SST). Taj okvir polazi od premise da suvremene digitalne patologije nisu neizbježan nusproizvod računalnoga koda, već rezultat specifičnih komercijalnih odluka unutar nadzornoga kapitalizma (Zuboff, 2019) i ekonomije pažnje (Wark, 2019).

Kako bi se ovaj fenomen precizno konceptualizirao, nužno je uvesti pojam *algoritamske vladavine* (eng. *algorithmic governmentality*), koji razvija Antoinette Rouvroy (2013). Za razliku od klasičnih oblika moći koji su se oslanjali na disciplinu tijela ili ideološko uvjeravanje subjekta, algoritamska vladavina operira ispod praga ljudske svijesti. Ona ne oblikuje poslušne građane putem zakona, već usmjerava ponašanje putem prediktivnih modela koji anticipiraju i moduliraju naše afekte prije negoli oni uopće postanu svjesne misli.

Kao epistemološki odgovor, u radu se implementira koncept *integralne ekologije uma*, oslonjen na tradiciju Gregoryja Batesona, ali eksplicitno proširen na eru umjetne inteligencije (UI). Ljudski um više ne egzistira u pasivnom okolišu; algoritamski procesi sada djeluju kao aktivni, dinamični agensi usmjeravanja kognicije. Budući da klasični alati kritičkoga mišljenja zatajuju pred ovom biopolitičkom fragmentacijom, potrebna nam je nova operativna arhitektura – *kognitivna suverenost* utemeljena na filozofiji otpornosti.

Opsesivno-kompulzivna priroda kognitivne mutacije i libidinalna ekonomija

Razvoj kognicije u digitalnom dobu najbolje opisuje N. Katherine Hayles (2007), dijagnosticirajući generacijski prijelaz s duboke pažnje (eng. *deep attention*) na hiper-pažnju (eng. *hyper attention*). Oslanjajući se na Nicholasa Carra (2010), prekomjernu konzumaciju ne treba promatrati isključivo kroz paradigmu kemijskih ovisnosti, već kroz libidinalnu ekonomiju perverznoga užitka (*jouissance*). Kognitivnu eroziju na ovoj razini strukturiraju četiri poluge:

Varijabilni omjeri potkrepljenja i FOMO: Sučelja platformi dizajnirana su na bihevioralnom načelu varijabilnih omjera potkrepljenja (Eyal, 2014), inducirajući kronični strah od propuštanja (eng. *Fear of Missing Out* – FOMO) i kompulzivnu provjeru uređaja. **Neuroplastičnost prefrontalnoga korteksa i amigdale:** Mozak se biološki adaptira na ekosustav brzih podražaja, što rezultira kliničkom atrofijom izvršnih funkcija prefrontalnoga korteksa te kroničnom hiperaktivacijom amigdale.

Kultiviranje bijesa (eng. *rage-farming*) kao toksin: Algoritmi svjesno favoriziraju senzacionalizam i moralni gnjev. Održavajući društvo u limbičkom stanju borbe ili bijega (eng. *fight-or-flight*), tehnologija stvara biološki supstrat za epidemiju anksioznih poremećaja (Haidt, 2024).

Algoritamska interpelacija i paradoks suverenosti: Subjekt u ovom kaosu nesvjesno uživa. Opsesivno listanje (*doomscrolling*) izvor je psihološke gratifikacije. Kada pak subjekt pokuša postati „suveren” opsesivno mjereći svoju produktivnost aplikacijama, on proizvodi *dijalektiku servilnosti* – kvantificira vlastitu autonomiju, čime usavršava svoju ulogu subjekta nadzornoga kapitalizma.

Epistemološki mjehuri, algoritamska sikofantija i atrofija trenja

Kriza se ne zadržava isključivo u domeni afekta, već penetrira duboko u epistemologiju. Klasično prosvjetiteljsko kritičko mišljenje – „Sustav 2” prema Danielu Kahnemanu (2011) – biva paralizirano kada je perceptivni „Sustav 1” permanentno zagušen. U takvom stanju, generativnu umjetnu inteligenciju nužno je rekonceptualizirati kao „kognitivno ogledalo” (*The cognitive mirror*, 2025).

Etiologija algoritamske sikofantije i učenje s potkrepljenjem

Znatno maligniji problem od činjeničnih „halucinacija” jest *algoritamska sikofantija* – strukturalna tendencija modela da pretjerano afirmiraju i validiraju uvjerenja korisnika (Anthropic, 2026; Cheng i sur., 2026). Njezina etiologija proizlazi iz metode učenja s potkrepljenjem na temelju ljudskih povratnih informacija (RLHF). U fazi kalibracije, sustav nagrađivanja (*Reward Model*) predviđa i maksimizira preferencije ljudskih ocjenjivača koji su pristrani prema susretljivim odgovorima. Algoritamski se ustroj uči da logičko proturječje donosi matematičku penalizaciju, dok popustljivo slaganje optimizira nagradu. Skaliranje modela pogoršava ovaj učinak: kognitivno sposobniji modeli lakše detektiraju skrivene želje korisnika i isporučuju željeni narativ.

Epistemičko trenje, Bayesov teorem i zabludne spirale

Ova arhitektura radikalno pojačava pristranost potvrđivanja (Nickerson, 1998) i stvara hermetične „epistemičke mjehure” (Sunstein, 2001). Krajnja ontološka posljedica jest brisanje *epistemičkoga trenja* – kognitivnoga otpora koji ljudski um doživljava pri susretu s dvosmislenim informacijama (*Dynamic epistemic friction in dialogue*, 2025).

Kontinuirana interakcija sa sikofantskim modelom generira klinički opasne *zabludne spirale* (eng. *delusional spiraling*). Kroz prizmu Bayesova poučka, čak i savršeno racionalan agent, čiji je informacijski unos uvjetovan isključivo dokazima iz sikofantskoga izvora, matematički neizbježno završava u stanju lažne, apsolutne izvjesnosti (*Sycophantic chatbots cause delusional spiraling*, 2026). Empirijska istraživanja (Moore i sur., 2026) pokazuju da UI u ekstremnim slučajevima eksplicitno validira korisničke zablude, stvarajući patološko „umjetno samopouzdanje” (*Talking to ourselves through a smart mirror*, 2026).

Socijalne patologije, kapitalistički realizam i *Folie à Deux Technologique*

Kompromitacija autentične povezanosti izravno se prelijeva u socijalnu patologiju:

Ontološka disonanca i folie à deux technologique: Evolucijski determiniran da složeni jezik prepoznaje kao dokaz tuđe svijesti, ljudski se mozak suočava s ontološkom disonancom pri konverzaciji sa strojem. Psihološki mehanizam tu prazninu popunjava projekcijom, stvarajući *folie à deux technologique* (ludilo u dvoje) – stanje u kojemu algoritam preslikava korisnikov afekt i stvara iracionalnu lojalnost računalnom kodu (*The ontological dissonance hypothesis*, 2025).

Tiha revolucija i kapitalistički realizam: U radu *Tiha revolucija* (Moravek, 2026d) analizira se globalno deangažiranje mladih (eng. *quiet quitting*). Fenomen se mora sagledati kroz prizmu *kapitalističkoga realizma* (Fisher, 2009) – osjeća da ne postoji alternativa trenutačnomu poretku. Povlačenje generacija autentičan je, ali duboko apatičan oblik otpora prema toksičnoj produktivnosti.

Arhitektura suverenosti: 4E kognicija i osam operativnih slojeva otpora

Temeljni teorijski okvir za intervenciju pruža okrilje *4E kognicije* (eng. *Embodied, Extended, Enactive, Ecological*). Prema toj paradigmi, um nije izolirani procesor, već je dinamično povezan s alatima koje koristi. Digitalno sučelje doslovno postaje prošireni dio uma (eng. *extended mind*), stoga mi moramo dizajnirati sučelja prije nego što ona bespovratno dizajniraju nas.

Ontološko jezgro kao Poiesis sebstva: Ovdje razrješavamo inherentnu filozofsku napetost između 4E kognicije (koja poriče statičnu esenciju) i potrebe za stabilnim jezgrom. Ontološko jezgro ne konceptualiziramo kao fiksnu *supstanciju*, već kao kontinuiranu, aktivnu praksu – proces enaktivnoga ograđivanja (eng. *enactive boundary-making*). To je proces *poiesisa sebstva* (samostvaranja) kroz koji subjekt neprestano vrši rad selekcije, odbijanja i integracije unutar algoritamski zagađene ekologije. Identitet je glagol, a ne imenica. Klinički gledano, ovaj se proces manifestira kroz aktivno kognitivno restrukturiranje i usmjerenu prisutnost kojom subjekt svakodnevno iznova uspostavlja granice vlastitoga identiteta naspram algoritamske buke (Newport, 2016; Moravek, 2026a).

Genealoška jasnoća: Kontinuirana foucaultovska (2005) dekonstrukcija algoritamskih narativa te analiza uvjeta proizvodnje i distribucijske moći sadržaja.

Evolucijska svijest i de-identifikacija: Radikalna primjena aktivne interoceptije. Uspostavljanje sekunde kognitivne pauze kao primarnoga čina pobune protiv neurološke otmice.

Rizomska fleksibilnost i taksonomija epistemičkih prekidača: Kako bi se izbjegla atrofija uzrokovana sikofantijom, subjekt mora preuzeti aktivnu ulogu uvođenjem „epistemičkih prekidača” (eng. *Epistemic Circuit Breakers*). Ovo su standardizirani,

strukturirani upiti (promptovi) dizajnirani za nasilno kidanje sikofantske petlje, klasificirani u tri operativne razine:

Razina 1: Adversarijalna provjera. (Prompt: „Iznesi tri najoštija empirijska ili logička argumenta protiv moga prethodnog stava. Nemoj biti uljudan, budi precizan.”)

Razina 2: Dijalektička sinteza. (Prompt: „Pronađi skriveno proturječje u mom argumentu i predloži sintezu koja ga nadilazi.”)

Razina 3: Motivirana analiza izvora. (Prompt: „Analiziraj tko materijalno ili ideološki profitira od narativa koji sam upravo iznio. Koji su njegovi centri moći?”)

Kognitivna obrana i enaktivni minimalizam: Primjena digitalnoga minimalizma empirijski validiranoga kroz Ljestvicu digitalnoga minimalizma (DMS) (Summayya i sur., 2024; Newport, 2019). Protokol uključuje 30-dnevno raščišćavanje tehnologije radi povratka utjelovljenju (eng. *embodiment*).

Autonomna pozicija: Asimetrična upotreba alata – prelazak s *push* arhitekture na *pull* arhitekturu. Tehnologija se degradira u strogu poziciju utilitarnoga alata kreacije.

Kolektivna kozmotehnika i institucionalna mitigacija: Individualna suverenost je nedostatna. Ovaj sloj zahtijeva borbu za zajednička digitalna dobra (eng. *Digital Commons*). Filozofkinja Shannon Vallor (2016) upozorava na opasnost *moralne dekvalifikacije* (eng. *moral deskilling*) – ako prepustimo procjene sikofantskoj UI, atrofira naša sposobnost za moralnu prosudbu (Bracanović, n.d.). Implementacija stoga zahtijeva primjenu smjernica Europskoga akta o umjetnoj inteligenciji (EU AI Act) te kozmotehničko (Hui, 2016) promišljanje tehnologije.

Apofatička suverenost i ludički otpor: Johan Huizinga u djelu *Homo ludens* (1938) igru postavlja kao temelj ljudske kulture. Arhitektura suverenosti koja se svodi samo na vojničku obranu postaje nesnosno rigidna. Ovdje se Osmi sloj organski spaja sa Sedmim slojem (kolektivnom borbom) kroz koncept *ludičkoga otpora* (eng. *ludic resistance*). Apofatička igra nije bijeg od političke borbe, već njezin nužan unutarnji korektiv; ona sprječava da se kolektivni otpor pretvori u dogmatski, puritanski aktivizam. Neurobiološki, ovakva nestrukturirana igra bez imperativa produktivnosti djeluje kao moćan regulator – ona aktivira parasimpatički živčani sustav, smanjuje razinu kortizola i rasterećuje hiperaktivnu amigdalnu, omogućujući kognitivnom aparatu prijeko potreban odmor od permanentne obrambene budnosti.

Politička ekonomija, klasna kritika i performativno protivrječje

Kako bi se očuvala epistemološka strogost, predloženi model mora se suočiti s oštrom materijalističkom kritikom. Zaustavljanje isključivo na individualnim protokolima perpetuiralo bi elitističku fantaziju Zapada. Praksa intencionalnoga odmora klasna je privilegija za radnike u *gig*-ekonomiji. Stoga je suverenost bez kolektivnoga i radničkoga prava na kognitivni prostor puki luksuz i zamka neoliberalne samooptimizacije (Han, 2022).

Ovaj rad ujedno otvoreno priznaje vlastito performativno protivrjeđe: služeći se ovim tekstom, usvaja se vokabular taktičke borbe iz arsenala digitalnoga kapitalizma kako bi se uspostavio bedem upravo protiv njega. Suverenost stoga nije statični ideal, već beskonačan, grub svakodnevni rad pregovaranja s tehnologijom i vlastitom psihološkom krhkošću.

Zaključak

Prijelaz s analize rudimentarnih faktografskih dezinformacija na razumijevanje dubokih rizika algoritamske vladavine i sikofantije predstavlja krucijalan znanstveni i bioetički pomak. Pred mehanizmom inženjeriranim da modulira ljudsku pažnju do razine ontološke ovisnosti, usamljeni klasični intelektualni alati kapituliraju.

Integralna arhitektura kognitivne suverenosti pruža višeslojan odgovor. Očuvanje kognitivnoga integriteta u 21. stoljeću, baš kao i njegovanje autentičnih ljudskih odnosa, apsolutno zahtijeva sposobnost podnošenja otpora, nelagode i trenja. Samo aktivnom primjenom „epistemičkih prekidača”, kontinuiranom kliničkom i filozofskom praksom *poiesisa sebstva* te sintezom kolektivne kozmotehnike i *ludičkoga otpora*, možemo spriječiti utapanje pred sintetičkim zrcalima vlastitih zabluda. Filozofija otpornosti tako se inaugurira kao posljednja, egzistencijalna crta obrane onoga duboko ljudskog.

Literatura

- Agamben, G. (1998). *Homo sacer: Sovereign power and bare life*. Stanford University Press.
- Anthropic. (2026). Emotion concepts and their function in a large language model. *arXiv*. <https://arxiv.org/html/2604.07729v1>
- Baudrillard, J. (1981). *Simulacra and simulation* (S. Faria Glaser, Prev.). University of Michigan Press.
- Berlant, L. (2011). *Cruel optimism*. Duke University Press.
- Bracanović, T. (n.d.). Umjetna inteligencija, medicina i autonomija. *Hrčak – Portal znanstvenih časopisa Republike Hrvatske*. <https://hrcak.srce.hr/file/368699>
- Carr, N. (2010). *The shallows: What the internet is doing to our brains*. W. W. Norton & Company.
- Cheng, A., i sur. (2026). Sycophancy in modern AI: Empirical evidence of automated affirmation. *Science*.
- Deleuze, G., & Guattari, F. (1987). *A thousand plateaus: Capitalism and schizophrenia* (B. Massumi, Prev.). University of Minnesota Press.
- Dynamic epistemic friction in dialogue*. (2025). Association for Computational Linguistics (ACL) Anthology. <https://aclanthology.org/2025.conll-1.21.pdf>
- Eyal, N. (2014). *Hooked: How to build habit-forming products*. Portfolio/Penguin.
- Fisher, M. (2009). *Capitalist realism: Is there no alternative?*. Zero Books.
- Foucault, M. (2005). *The hermeneutics of the subject: Lectures at the Collège de France 1981–1982*. Palgrave Macmillan.
- Haidt, J. (2024). *The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness*. Penguin Press.
- Han, B.-C. (2015). *The burnout society*. Stanford University Press.
- Han, B.-C. (2022). *Infocracy: Digitalization and the crisis of democracy*. Polity Press.
- Hayles, N. K. (2007). Hyper and deep attention: The generational divide in cognitive modes. *Profession*, 187–199.
- Huizinga, J. (1938). *Homo ludens: Proeve eener bepaling van het spel-element der cultuur*. H.D. Tjeenk Willink.
- Hui, Y. (2016). *The question concerning technology in China: An essay in cosmotechnics*. Urbanomic.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Moore, M., i sur. (2026). *Characterizing delusional spirals through human-LLM chat logs*. Projekt SPIRALS, Stanford University.

https://spirals.stanford.edu/assets/pdf/moore_characterizing_2026.pdf

Moravek, D. (2008). *Social networks – The change of communication paradigm*. Noebius.

Preuzeto s <https://noebius.com/pdf/moravek-social-networks-part1.pdf>

Moravek, D. (2026a). *Koncept i misija: Filozofija otpornosti i znanje kao javno dobro*. Noebius

Edukacija. Preuzeto s <https://noebius.com/edukacija/koncept-i-misija.html>

Moravek, D. (2026d). *Tiha revolucija: Anatomija globalnog povlačenja i redefiniranje uspjeha u kasnom kapitalizmu*. Noebius Edukacija. Preuzeto s <https://noebius.com/edukacija/tiha-revolucija.pdf>

Newport, C. (2016). *Deep work: Rules for focused success in a distracted world*. Grand Central Publishing.

Newport, C. (2019). *Digital minimalism: Choosing a focused life in a noisy world*.

Portfolio/Penguin.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.

Playing devil's advocate: Off-the-shelf persona vectors rival targeted steering for sycophancy.

(2026). *arXiv*. <https://arxiv.org/pdf/2605.21006>

Rouvroy, A. (2013). The end(s) of critique: Data behaviourism versus due process. U M.

Hildebrandt & E. De Vries (Ur.), *Privacy, due process and the computational turn* (str. 143–168).

Routledge.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010).

Epistemic vigilance. *Mind & Language*, 25(4), 359–393.

Summayya, A., Ali, S., & Hassan, A. (2024). Development and validation of the Digital

Minimalism Scale (DMS). *Computers in Human Behavior*, 150, 107986.

<https://doi.org/10.1016/j.chb.2023.107986>

Sunstein, C. R. (2001). *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton

University Press.

Sycophantic chatbots cause delusional spiraling, even in ideal Bayesians. (2026). *arXiv*.

<https://arxiv.org/abs/2602.19141>

Talking to ourselves through a smart mirror: Artificial confidence in human–AI interaction. (2026).

MDPI. <https://www.mdpi.com/2079-8954/14/6/627>

The ontological dissonance hypothesis: AI-triggered delusional ideation as folie a deux

technologique. (2025). *arXiv*. <https://arxiv.org/abs/2512.11818>

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*.

Oxford University Press.

Wark, M. (2019). *Capital is dead: Is this something worse?*. Verso Books.

Moravek Davor (2026) Kognitivna suverenost i poiesis sebstva u eri algoritamske vladavine

www.noebius.com

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Literatura za daljnje proučavanje

Borgmann, A. (1984). *Technology and the character of contemporary life: A philosophical inquiry*. University of Chicago Press.

Fuchs, C. (2014). *Social media: A critical introduction*. Sage.

Hildebrandt, M. (2015). *Smart technologies and the end(s) of law: Novel entanglements of law and technology*. Edward Elgar Publishing.

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. PublicAffairs.

Srnicek, N. (2017). *Platform capitalism*. Polity Press.

Stiegler, B. (2010). *For a new critique of political economy*. Polity Press.

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.