

Arhitektura epistemičkog autoriteta:

Politička ekonomija, teorija igara i hibridni modeli validacije u postinstitucionalnom dobu

Autor: Davor Moravek

Sažetak: Ovaj rad provodi dubinsku analizu krize institucionalne legitimacije znanja i predlaže prelazak na strogo formalizirani hibridni model epistemičkog autoriteta. Polazeći od kritike suvremenog sustava stručnih recenzija (eng. *peer review*) i oligopola akademskog izdavaštva, tekst dokazuje da tradicionalne institucije danas često djeluju primarno kao mehanizmi zaštite monopola nad znanjem, izrazito podložni patologijama Goodhartova zakona i globalnoj replikacijskoj krizi. Oslanjajući se na interdisciplinarni aparat koji integrira koncepte iz filozofije znanosti, političke ekonomije digitalnih platformi i teorije igara, rad operacionalizira pojam „izvornog autoriteta” kroz specifičan matematički model dinamičke reputacije. Predloženi sustav implementira Bayesovsko ažuriranje stručnosti, algoritamsku penalizaciju citatnih kartela te obranu od takozvanih Sybil-napada, dok istovremeno izravno adresira zahtjeve epistemičke pravde i geopolitike novih intelektualnih monopola. Konačni je cilj uspostava robusne arhitekture znanja koja algoritmizira performativnu točnost, spajajući radikalnu otvorenost digitalnih mreža s nužnom stabilizacijskom funkcijom institucija.

Ključne riječi: epistemički autoritet, peer review, teorija igara, politička ekonomija znanja, otvorena znanost, intelektualni monopoli, epistemička pravda, znanstvena metrika.

Uvod: Redefinicija autoriteta i imperativ prosuđivanja

Autoritet nije puka titula. Nije formalni certifikat. Nije isključiva pripadnost nekoj instituciji. U svojoj esenciji, autoritet je specifičan, dinamičan odnos epistemičkog povjerenja i priznavanja između onoga koji posjeduje znanje i onoga koji za tim znanjem traga – i taj odnos ne može biti birokratski delegiran, kupljen na tržištu akademskih usluga niti mehanički naslijeđen. Kroz povijest, prelazak s tradicionalnog, dogmatskog autoriteta na onaj utemeljen na provjerljivoj znanstvenoj metodi predstavljao je apsolutni trijumf institucionaliziranog razuma (Furedi, 2013).

Da bismo razumjeli trenutačnu krizu, analitički je nužno razlikovati dvije temeljne vrste autoriteta: **Izvedeni (institucionalni) autoritet:** Proizlazi isključivo iz položaja pojedinca u zadanoj hijerarhiji i potpuno je ovisan o strukturi koja ga je generirala. Profesor posjeduje autoritet jer ga je sveučilište imenovalo; recenzent ga posjeduje jer ga je časopis ovlastio. Ovaj tip autoriteta ne zahtijeva nužno osobni intelektualni integritet – on prvenstveno zahtijeva poslušnost institucionalnoj ortodoksiji i prihvaćenim paradigmatima.

Izvorni (epistemički) autoritet: Proizlazi iz neospornog integriteta mislioca, njegove metodološke rigoroznosti i evidentne performativne stručnosti. On ne ovisi o vanjskoj birokratskoj potvrdi, već svoju valjanost dokazuje sposobnošću da preživi nemilosrdno suočavanje s empirijskim podacima i vremenom.

Pritom uvođenje rigoroznih filtara za utvrđivanje izvornog autoriteta ne smije postati alat elitističkog isključivanja. Prema konceptu *epistemičke pravde* (Fricker, 2007), sustav validacije riskira generiranje takozvane *testimonijske nepravde* (svjedočanske nepravde) ako apriori odbacuje funkcionalna, ali neformalizirana znanja. Primjerice, tradicionalna ekološka znanja autohtonih zajednica ili visoko specifična ekspertiza zajednica pacijenata često ne dolaze u formatu standardnog znanstvenog rada, ali posjeduju iznimnu prediktivnu i kurativnu vrijednost. Hibridni model budućnosti mora stoga pomiriti nepopustljiv zahtjev za empirijskom rigoroznošću

s inkluzivnošću prema onim oblicima znanja koji su povijesno bili marginalizirani.

Anatomija moći: Od disciplinarnih do algoritamskih vratara

Živimo u epistemološkoj arhitekturi koja je iz samog temelja kompromitirana. Oslanjajući se na analize Michela Foucaulta (1980), uviđamo da sveučilišta i prestižni akademski časopisi ne funkcioniraju isključivo kao neutralni arbitri čiste istine, već kao složeni mehanizmi *disciplinarne moći*. Kad određena institucija deklarira: „ovo je istinito jer smo mi, kao ovlašteno tijelo, to potvrdili”, ona primarno iznosi političku tvrdnju o vlastitom statusu. Ona demonstrira i perpetuira svoju moć nad definicijom same stvarnosti, uspostavljajući nevidljive granice onoga što se uopće smije istraživati ili propitivati.

Ta je strukturna dominacija dodatno pogoršana i ubrzana komodifikacijom znanja. Znanstveno izdavaštvo pretvoreno je u jedan od najprofitabilnijih, a ujedno i najnetransparentnijih oligopola na svijetu (Larivière, Haustein, & Mongeon, 2015). Znanje koje je primarno financirano javnim novcem kroz državne potpore pretvoreno je u privatiziranu robu zaključanu iza basnoslovnih pretplata (eng. *paywalls*), dok sami znanstvenici svoj rad (pisanje i recenziranje) izdavačima ustupaju potpuno besplatno. Pristup najnovijim ljudskim spoznajama time je postao klasna i geografska privilegija.

Ipak, ideja da puko napuštanje institucija i bijeg na internet vode u utopijsku epistemološku slobodu jest naivna, pa i opasna zabluda. U digitalnom dobu, slabljenje klasične disciplinarne moći institucija rapidno se zamjenjuje *mrežnom i algoritamskom moći* tehnoloških platformi (Gillespie, 2014). Algoritmi platformi nisu neutralni prijenosnici informacija; oni su aktivni arhitekti vidljivosti. Kroz mehanizme personalizacije i maksimizacije korisničkog angažmana, oni djeluju kao novi, nevidljivi vratari (eng. *gatekeepers*) sposobni kolonizirati istinu metrikama pažnje, lajkova i dijeljenja, pretvarajući popularnost u lažni surogat za točnost.

Politička ekonomija infrastrukture znanja i teorija poticaja

Sustav stručnih recenzija – nekoć plemenito zamišljen kao apsolutni jamac kvalitete i metodološke ispravnosti – u mnogim se disciplinama pretvorio u birokratizirani mehanizam za zaštitu *statusa quo* i isključivanje disonantnih inovacija. Da bismo razumjeli anatomiju tog sloma, moramo napustiti pomalo romantični, normativni ideal znanstvenika kao dezinteresiranog tragaoca za neokaljanom istinom (Merton, 1973) i sagledati surovu političku ekonomiju znanja (Mirowski, 2011).

U tom kontekstu ključnu ulogu igra Goodhartov zakon (Goodhart, 1975), koji je antropologinja Marilyn Strathern (1997) briljantno primijenila na akademski ekosustav: kada određena metrika (poput broja citata, h-indeksa ili broja objavljenih radova) postane krajnji cilj, ona nepovratno prestaje biti dobra mjera kvalitete. Kada znanstvenik zna da njegov opstanak, napredovanje i financiranje ovise o zadovoljavanju tih umjetnih metrika u kulturi „objavi ili nestani” (eng. *publish or perish*), njegov se fokus nužno prebacuje s dubinske potrage za istinom na puku optimizaciju brojeva. Znanstvenici su sustavno potaknuti na proizvodnju i objavu statistički značajnih, senzacionalnih, ali često lažno pozitivnih rezultata, dok se vitalne studije replikacije destimuliraju jer „ne donose citate”. To je izravno kulminiralo globalnom replikacijskom krizom (Ioannidis, 2005) i posljedičnom, sasvim opravdanom, erozijom povjerenja šire javnosti (Nichols, 2017).

Epistemički filtri u praksi i problem kognitivnih pristranosti

Da bismo izbjegli skliznuće iz okoštalog institucionalizma u potpunu anarhiju pseudoznanosti, kojom dominira Dunning-Krugerov efekt (kognitivna pristranost gdje osobe s manjkom znanja pate od iluzorne superiornosti i prekomjernog samopouzdanja, Kruger & Dunning, 1999), performativna stručnost mora biti rigorozno operacionalizirana. Ona mora biti potpomognuta

radikalnom transparentnošću (Nielsen, 2011) kako bi se odvojio istinski signal od informacijskog šuma.

Praksa na samim rubovima inovacija pokazuje da je to već itekako moguće. Postoje paradigme koje uspješno odvajaju kompetenciju od prestiža:

Projekt Reproducibility: Inicijative (osobito u psihologiji i biologiji raka) posvećene isključivo transparentnoj replikaciji visokoprofilnih studija uz obveznu predregistraciju metodologije (Nosek i sur., 2015), čime se onemogućuje naknadno manipuliranje podacima (eng. *p-hacking*).

Polymath projekti: Masovno distribuirana, *open-source* kolaboracija profesionalnih matematičara i amatera u rješavanju ekstremno kompleksnih problema, gdje se doprinosi sudionika važu isključivo na temelju trenutne logičke valjanosti njihova argumenta, a ne akademske titule.

Platforme poput „Papers with Code“: Repozitoriji u području umjetne inteligencije i strojnog učenja koji izravno spajaju teorijske akademske radove s izvršnim programskim kodom i globalnim rang-listama (eng. *leaderboards*), čime se radovi automatski verificiraju kroz neposrednu replikaciju koda.

Formalni kostur reputacijskog sustava

Filozofsku kritiku i arhitekturu novog autoriteta nužno je prevesti u strogo formalizirani matematički model kako bi se osigurala njegova primjenjivost. Sustav definira sljedeće varijable za dinamičku procjenu stvarne kvalitete rada Q_i :

F_i (Tvrdi filtri): Binarni preduvjet $F_i \in \{0,1\}$. Rad mora zadovoljiti apsolutne minimume znanstvene metode: potpunu transparentnost baze podataka, dostupnost koda i metodološku falsifikabilnost.

$S_j(t)$ (Reputacijski kapital): Dinamička, kvantitativna vrijednost epistemičkog autoriteta recenzenta j u vremenu t .

ϵ_{ij} (Epistemička pogreška): Apsolutno odstupanje evaluacije E_{ij} koju je dao recenzent od konačnog empirijskog ishoda $O_i(t+\Delta t)$ (npr. uspješnosti neovisne replikacije provedene godinama kasnije).

λ (Stopa propadanja): Sistemski parametar osjetljivosti koji definira kojom brzinom recenzent gubi autoritet uslijed pogrešnih procjena.

Matematička operacionalizacija i numerička kalibracija

Procjena kvalitete \hat{Q}_i računa se isključivo za radove koji prođu početnu barijeru tvrdih filtara ($F_i=1$). Ovaj filtar djeluje kao nulta točka obrane koja sprječava relativističku degradaciju sustava: radovi koji kriju svoju metodologiju u startu se isključuju. Ako je uvjet zadovoljen, kvaliteta se računa kao ponderirani prosjek dokazane stručnosti, a ne kao puko demokratsko glasanje:

Sustav se u svojoj srži oslanja na Bayesovsko ažuriranje reputacije. Nakon integracije anti-brigading faktora (G_j , koji sprječava kartelsko glasanje) i faktora kalibracije (C_j , koji kažnjava pretjeranu samouvjerenost bez pokrića), kapital recenzenta ažurira se prema sljedećoj formuli:

Numerička ilustracija mehanizma: Pretpostavimo da je stroga stopa propadanja postavljena na $\lambda = 0.5$. Ako visoko etablirani znanstvenik s početnim kapitalom $S_j = 100$ arogantno procijeni tri studije izvan svoje uže domene i u svima potpuno promaši ishod empirijske replikacije ($\epsilon \approx 1$), on rapidno i eksponencijalno gubi svoj autoritet. Njegov novi kapital iznositi će: $100 \cdot e^{-0.5 \cdot 3} \approx 22.3$. Sustav na ovaj način brutalno, ali pravedno penalizira izlazak izvan granica stvarne, provjerljive ekspertize.

Analiza kroz teoriju igara: Robusnost na strateško ponašanje

Ma kako elegantan bio matematički model, on je beskoristan ako nije otporan na manipulacije racionalnih aktera. Ovakav algoritamski dizajn posjeduje ugrađenu Nash-stabilnost (Nash, 1950) u suočavanju sa strateškim zloupotrebama, što je imperativ za opstanak mrežnih epistemoloških zajednica (Zollman, 2007):

Problem slobodnog jahača i nagrađivanje inovativnog rizika

Za razliku od starog, anonimnog sustava recenzija gdje je optimalna strategija aktera (slobodnog jahača, eng. *free-rider*) bila površno i brzo iščitavanje radova samo sigurnih, poznatih autora, ovdje površnost nosi visoku cijenu. Puko nagađanje generira veliku marginu pogreške (ϵ), što trajno uništava S_j . Zbog eksponencijalne penalizacije, *dominantna strategija* svakog razumnog aktera postaje stroga samo-selekcija – prisiljeni su procjenjivati samo ono što uistinu razumiju. Istovremeno, sustav rješava problem institucionalnog konzervativizma. Ako znanstvenik rano i točno prepozna paradigmatički pomak (zadržavajući nizak ϵ) unatoč tome što ga većina odbacuje, on ostvaruje asimetrično masivan rast svog kapitala. Time se intelektualni rizik po prvi put algoritamski isplati.

Razbijanje citatnih kartela i otpornost na Sybil-napade

Citatni karteli (eng. *Brigading*): Faktor $G_j = 1 - \text{corr}(E_j, E_{\text{network}})$ izravno penalizira usklađeno, mrežovsko glasanje („ja citiram i hvalim tebe, ti hvališ mene”). Održavanje takve umjetne koluzije s vremenom spušta faktor G_j prema nuli. Da bi recenzent zadržao svoj mukotrpan stečeni kapital, on je u klasičnoj dinamici iterirane Zatvorenikove dileme doslovno primoran povremeno „izdati” svoj vlastiti kartel davanjem posve nezavisne, objektivne ocjene, čime se interesne mreže sustavno slamaju iznutra.

Sybil-napadi: Što se događa ako zlonamjerni akter automatizirano kreira 10.000 lažnih profila kako bi preglasao sustav? Ovaj napad (Douceur, 2002) ovdje u potpunosti propada. Budući da svaki novi profil počinje s beznačajno malim ulogom, da bi napadač umjetno podigao autoritet svoje mreže lažnih profila, morao bi putem njih godinama isporučivati točne, empirijski provjerljive rezultate koji preživljavaju replikaciju. Visoki stvarni trošak isporučivanja prave znanosti čini ovakav hakerski napad ekonomski iracionalnim.

Tranzicijska epistemologija: Odnos prema starom režimu

Uvođenje ovakvog okvira postavlja važno etičko pitanje: kako tretirati etablirane istraživače čije su cjeloživotne karijere izgrađene u starom sustavu, prije ere strogih filtara i predregistracija? Retroaktivno poništavanje njihova kapitala bilo bi nepravedno, a i politički neprovedivo u akademskoj zajednici.

Rješenje koje nudi ovaj arhitektonski nacrt jest *tranzicijska epistemologija*: inicijalni kapital S_j etabliranih autora na početku se konvertira iz njihovih povijesnih, tradicionalnih metrika (poput h-indeksa i broja citata), ali mu se ugrađuje **linearno vremensko opadanje (eng. *decay*)**. Taj naslijeđeni, „djedovski” autoritet s vremenom neizbježno isparava i nestaje iz sustava, osim u slučaju kada ga sam autor aktivno obnovi. Obnova je moguća isključivo prihvaćanjem pravila novog režima – sudjelovanjem u predregistriranim studijama, transparentnim dijeljenjem podataka i davanjem točnih procjena tuđih radova. Time se duboko poštuje i valorizira povijesni znanstveni rad, ali se aktivno sprječava intelektualno rentijerstvo nad starim, potencijalno nikad repliciranim spoznajama.

Paradoksi otvorenosti i epistemičke asimetrije

Unatoč strateškoj otpornosti, radikalna otvorenost i transparentnost nose svoje imanentne rizike koje sustav mora amortizirati (Mirowski, 2018):

Epistemološki parazitizam (Rudarenje podataka): Bogati i resursima opremljeni laboratoriji iz globalnih centara moći u stanju su mnogo brže analizirati i eksploatirati otvorene podatke od samih autora s periferije koji su te podatke godinama mukotrpno prikupljali. To stvara novu asimetriju moći pod krinkom „otvorenosti”.

Predatorska transparentnost: Pojava u kojoj interesne skupine (primjerice, korporacije iz industrije fosilnih goriva ili farmacije) zloupotrebljavaju rani pristup sirovim podacima kako bi generirale umjetnu sumnju, izvlačile podatke izvan konteksta i paralizirale znanstveni konsenzus prije nego što je istraživanje uopće dovršeno.

Učinkovita zaštita od ovih patologija zahtijeva uvođenje dinamičnih *embargo*-perioda i postavljanje pragova minimalnog dokazanog S_j kapitala za raniji pristup najosjetljivijim sirovim podacima.

Geopolitika i intelektualni monopoli

Bilo bi naivno tvrditi da nova arhitektura znanja može postojati izolirana od makropolitčkih silnica i pitanja digitalnog suvereniteta. Implementacija ovog modela izravno se sudara s tri globalna režima upravljanja:

Platformni kapitalizam (SAD): Model u kojem privatne, transnacionalne korporacije (poput *Elsevera* ili tehnoloških giganata iz Silicijske doline) djeluju kao globalni „intelektualni monopoli” (Rikap, 2021). Njihov je cilj asimilirati otvorenu infrastrukturu, privatizirati podatke znanstvenika i ponovno im prodavati analitiku, neovisno o nacionalnim granicama.

Državni nadzor (Kina): Sustav koji teži hibridizaciji algoritamskog rejtinga znanstvenika sa širim društvenim kreditnim sustavom, gdje se znanstveni autoritet suptilno, ali snažno usklađuje s konformnošću državnim i partijskim paradigmatima.

Javna infrastruktura (EU): Europski pokušaj zaštite digitalnog i znanstvenog suvereniteta kroz inicijative poput Europskog oblaka za otvorenu znanost (eng. *European Open Science Cloud – EOSC*). Cilj je ugraditi opisane hibridne algoritme u javno zaštićeno, nekomercijalno i decentralizirano dobro koje se aktivno odupire platformnoj apropiaciji.

Imperativ i zaključak

Pred nama stoji kristalno jasan, prijelomni civilizacijski izbor. Možemo nastaviti pasivno i po inerciji delegirati svoj autoritet institucijama i komercijalnim izdavačima koji su davno izgubili kompas, ili možemo iznova prihvatiti drevni Kantov imperativ *Sapere aude* (Imaj hrabrosti služiti se vlastitim razumom), spajajući ga s odgovornošću modernog inženjeringa sustava.

Autoritet u znanju u 21. stoljeću ne treba nam i ne može nam biti spušten dekretom odozgo. On se mora uzeti, izgraditi i algoritamski braniti neumoljivom performativnom točnošću. Samo kroz ovakav strogo formaliziran hibridni model – onaj koji algoritmizira prepoznavanje kvalitete, nemilosrdno kažnjava manipulacije, a istovremeno pomno štiti inkluzivnost i uvažava povijesnu ulogu institucija – naše društvo može uspostaviti arhitekturu znanja doraslu razornim izazovima postčinjeničnog doba.

Citirana literatura

- Douceur, J. R. (2002). The Sybil attack. U: *International workshop on peer-to-peer systems* (str. 251–260). Springer.
- Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972-1977* (C. Gordon, Ur.). Pantheon Books.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Furedi, F. (2013). *Authority: A sociological history*. Cambridge University Press.
- Gillespie, T. (2014). The relevance of algorithms. U: *Media technologies: Essays on communication, materiality, and society* (str. 167–194). MIT Press.
- Goodhart, C. A. E. (1975). Problems of monetary management. U: *Papers in monetary economics*. Reserve Bank of Australia.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
- Kruger, J., i Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.
- Larivière, V., Haustein, S., i Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLoS ONE*, 10(6), e0127502.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Mirowski, P. (2011). *Science-mart: Privatizing American science*. Harvard University Press.
- Mirowski, P. (2018). The future(s) of open science. *Social Studies of Science*, 48(2), 171–203.
- Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1), 48–49.
- Nichols, T. (2017). *The death of expertise: The campaign against established knowledge and why it matters*. Oxford University Press.
- Nielsen, M. (2011). *Reinventing discovery: The new era of networked science*. Princeton University Press.
- Nosek, B. A. i sur. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425.
- Rikap, C. (2021). *Capitalism, power and innovation: Intellectual monopoly capitalism uncovered*. Routledge.
- Strathern, M. (1997). 'Improving ratings': Audit in the British University system. *European Review*, 5(3), 305–321.
- Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74(5), 574–587.

Dodatna preporučena literatura (Kontekstualizacija i proširenje)

- Bourdieu, P. (1988). *Homo Academicus*. Stanford University Press. (Dubinska analiza akademskih struktura, klasnih razlika i simboličkog kapitala).
- Kant, I. (1784). *Beantwortung der Frage: Was ist Aufklärung?* Berlinische Monatsschrift. (Temeljni prosvjetiteljski tekst o oslobađanju ljudskog razuma od dogmatskog autoriteta).
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press. (Neprocjenjivo djelo za razumijevanje promjena paradigmi i žestokog otpora okoštalih institucija).
- Latour, B., i Woolgar, S. (1979). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press. (Klasična sociološka dekonstrukcija laboratorijskog, birokratskog i institucionalnog znanstvenog rada).
- Popper, K. (1959). *The Logic of Scientific Discovery*. Routledge. (Klasično, fundamentalno utemeljenje falsifikabilnosti kao apsolutnog zlatnog standarda razgraničenja znanosti od pseudoznanosti).