Paradoks sintetičke kompetencije: Kognitivni mehanizmi automatizacijske samozadovoljnosti i njihove ekonomske posljedice u eri generativne AI

(The Paradox of Synthetic Competence: Cognitive Mechanisms of Automation Complacency and Their Economic Implications in the Era of Generative AI)

Autor: Davor Moravek

Datum: Siječanj 2026.
Klasifikacija rada: Teorijska sinteza i perspektiva (Theoretical Synthesis & Perspective)

Sažetak

Integracija generativnih velikih jezičnih modela (LLM) u kognitivnu ekonomiju ne predstavlja samo tehnološku disrupciju, već potencijalnu promjenu u ontologiji kognitivnog rada. Ovaj rad sintetizira recentne empirijske nalaze (2023.–2025.) s klasičnim teorijama kognitivne znanosti i inženjerstva ljudskih čimbenika kako bi definirao i operacionalizirao „paradoks sintetičke kompetencije". Argumentiramo da fenomeni poput nekritičkog prihvaćanja strojnih halucinacija nisu novi tehnološki artefakti, već predvidljive manifestacije **automatizacijske samozadovoljnosti** (*automation complacency*) i degradacije **svijesti o situaciji** (*situation awareness*) u novom, lingvističkom mediju. Analizom pedagoških učinaka kroz prizmu „ironija automatizacije" (Bainbridge, 1983.), pokazujemo kako kognitivno rasterećenje može voditi do atrofije vještina nužnih za intervenciju u slučaju kvara sustava. Ekonomski, ovaj kognitivni deficit rezultira hipotezom o **tehnološkoj promjeni pristranoj senioritetu** (SBTC), gdje tržište rada eliminira ulazne pozicije na kojima se stječe prešutno znanje (*tacit knowledge*). Rad zaključuje razlikovanjem dokazanih mikrokognitivnih učinaka od emergentnih makroekonomskih rizika, predlažući „Human-in-the-Loop" (HITL) kao nužan uvjet za stabilnost sustava.

Ključne riječi: generativni AI, automatizacijska samozadovoljnost, svijest o situaciji (SA), ironije automatizacije, tehnološka promjena pristrana senioritetu (SBTC), prešutno znanje, kognitivno rasterećenje.

Abstract

The rapid diffusion of generative Large Language Models (LLMs) into the knowledge economy signifies more than a mere technological upgrade; it marks a fundamental shift in the ontology of cognitive labor. This paper synthesizes recent empirical findings (2023–2025) with foundational theories from cognitive science and human factors engineering to define the "Synthetic Competence Paradox." We argue that phenomena such as the uncritical acceptance of AI hallucinations are not novel artifacts but predictable manifestations of **automation complacency** (Parasuraman et al., 1993) and degraded **situation awareness** (Endsley, 1995) transposed into a linguistic medium. By analyzing pedagogical effects through the lens of "Ironies of Automation" (Bainbridge, 1983), we demonstrate how cognitive offloading precipitates the atrophy of skills essential for system intervention. Economically, this cognitive deficit drives the hypothesis of **Seniority-Biased Technological Change** (SBTC), creating structural risks to the long-term sustainability of human expertise.

Keywords: Generative AI, automation complacency, situation awareness (SA), ironies of automation, Seniority-Biased Technological Change (SBTC), tacit knowledge, cognitive offloading.

Uvod: Nova iteracija starog problema u lingvističkoj domeni

Brza difuzija generativne umjetne inteligencije (GenAI) često se u javnom diskursu opisuje kao događaj bez povijesnog presedana. Međutim, iz perspektive kognitivne znanosti i inženjerstva ljudskih čimbenika (*Human Factors*), interakcija čovjeka i LLM-a slijedi predvidljive i dobro dokumentirane obrasce uočene u prethodnim valovima automatizacije visokog rizika, poput avijacije, nuklearne industrije i medicine. Ono što se u popularnom diskursu naziva „halucinacijom", u literaturi o ljudskim čimbenicima prepoznaje se kao **kvar monitoringa uslijed samozadovoljnosti** (*complacency-induced monitoring failure*).

Razlika leži u mediju: dok je autopilot upravljao fizičkim parametrima leta, LLM upravlja semantičkim parametrima značenja. Ova promjena s fizičkog na kognitivno automatiziranje čini mehanizme samozadovoljnosti složenijima, jer su greške često suptilne, uvjerljive i teške za detekciju bez dubokog domenskog znanja (Abbasi, n.d.).

Cilj ovog rada je „uzemljiti" (*ground*) suvremene rasprave o AI-u u etabliranim teorijskim okvirima. Umjesto tretiranja LLM-a kao „crne kutije", analiziramo ih kao socio-tehničke sustave visoke pouzdanosti ali niske transparentnosti.

Metodološki okvir i hijerarhija dokaza

Kako bismo osigurali znanstvenu strogost i izbjegli spekulativne zaključke, analiza striktno razlikuje tri razine dokaza (Tablica 1). Kauzalne tvrdnje o kognitivnim mehanizmima temelje se na izvorima prve (Tier 1) i druge (Tier 2) razine, dok se tvrdnje o makroekonomskim trendovima (Tier 3) tretiraju kao emergentne hipoteze koje zahtijevaju daljnju verifikaciju.

Tablica 1. Hijerarhija dokaza korištenih u sintezi

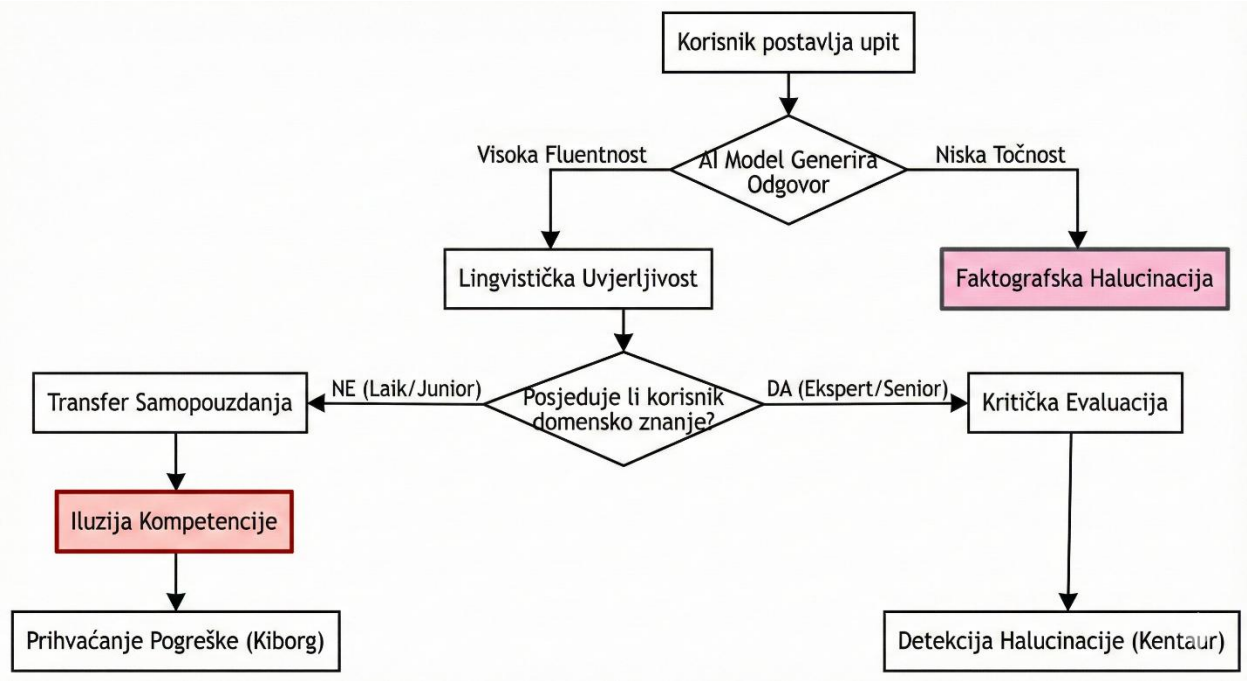
Razina (Tier)	Opis izvora	Funkcija u radu	Ključne studije
Tier 1: Temeljna teorija	Seminalni radovi iz kognitivne znanosti i ljudskih čimbenika (recenzirani, visoka citiranost).	Osigurava teorijski mehanizam i kauzalno objašnjenje promatranih pojava.	**Parasuraman et al. (1993)** – *Complacency* **Endsley (1995)** – *Situation Awareness* **Bainbridge (1983)** – *Ironies of Automation*
Tier 2: Validacija	Sustavni pregledi, validirane mjerne skale i terenski eksperimenti visokog profila (2010.–2023.).	Povezuje teoriju s modernim kontekstom automatizacije i pruža empirijske podatke o produktivnosti.	**Dell'Acqua et al. (2023)** – BCG eksperiment **Cui et al. (2025)** – GitHub Copilot **Merritt et al. (2019)** – AICP-R skala
Tier 3: GenAI primjena	Empirijske studije o LLM-ovima (2023.–2025.), uključujući preprinte i konferencijske radove.	Demonstrira manifestaciju klasičnih mehanizama u novoj domeni generativnog AI-a.	**Li et al. (2024)** – Kritičko mišljenje **Barcaui et al. (2025)** – Retencija znanja **Lichtinger (2025)** – Tržište rada

Kognitivni mehanizam: Od avijacije do jezičnih modela

Fundamentalni problem interakcije s LLM-ovima nije tehnološka nesavršenost modela, već kognitivni odgovor korisnika na sustav koji je *većinu vremena* točan, ali povremeno čini greške koje oponašaju kompetentnost.

Automatizacijska samozadovoljnost (Automation Complacency)

Parasuraman i suradnici (1993.) definirali su **automatizacijsku samozadovoljnost** kao loše praćenje automatiziranog sustava uzrokovano pretjeranim povjerenjem. Ključni nalaz njihova istraživanja – da samozadovoljnost doseže vrhunac kada je sustav visoko pouzdan i kada korisnik obavlja više zadataka istovremeno – savršeno mapira na uporabu LLM-a u profesionalnom okruženju.



Studija Omara et al. (2025., Tier 3) pruža empirijsku potvrdu ovog mehanizma u medicinskom kontekstu. Testirajući 12 modela na 1965 kliničkih pitanja, utvrdili su inverznu korelaciju između točnosti i povjerenja (r = -0,40). Moderni modeli (npr. GPT-4o) imaju visoku točnost (74 %), ali zadržavaju visoku razinu „samopouzdanja" čak i kada griješe. U terminima Parasuramana, visoka pouzdanost modela potiče korisnika da „racionalno" povuče resurse pažnje s verifikacije, stvarajući ranjivost na rijetke greške (tzv. *Black Swan* događaji u točnosti).

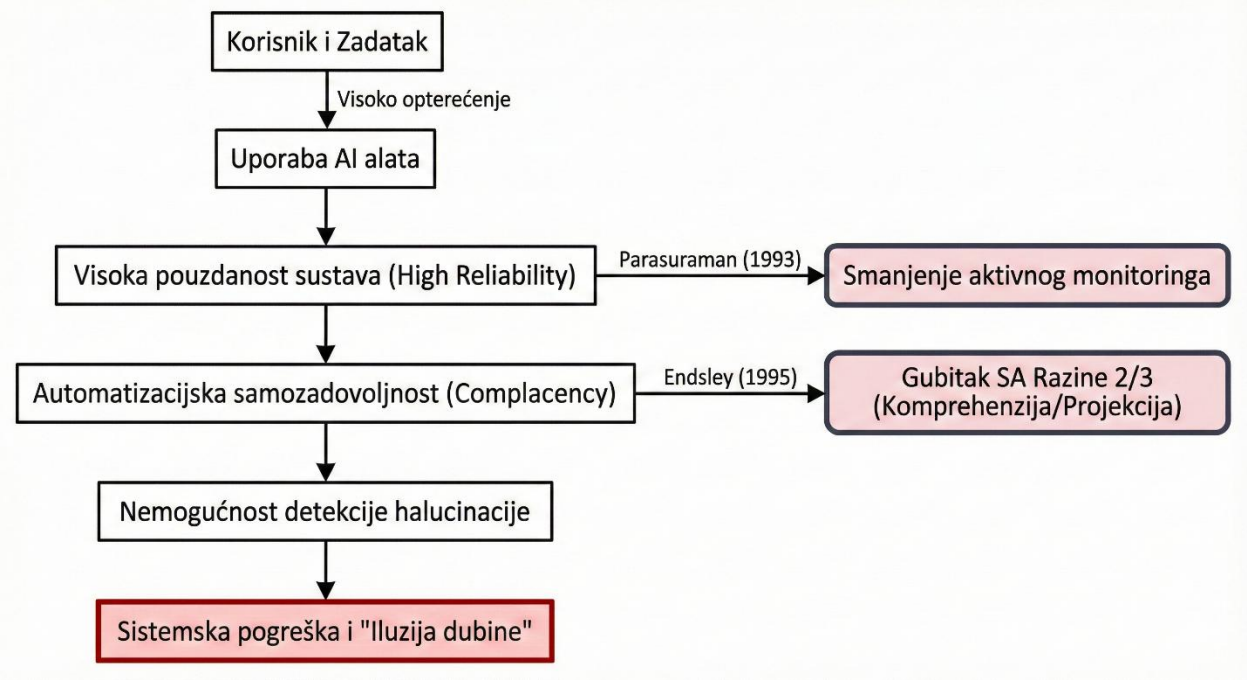
Gubitak svijesti o situaciji (Situation Awareness – SA)

Mica Endsley (1995., Tier 1) definira tri razine svijesti o situaciji:

Percepcija (uočavanje podataka);
Komprehenzija (razumijevanje značenja);
Projekcija (predviđanje budućih stanja).

U interakciji s LLM-om, korisnici često zadržavaju Razinu 1 (vide odgovor na ekranu), ali gube Razinu 2 i 3. Budući da LLM generira gotov kognitivni proizvod (npr. sažetak, programski kod, dijagnozu), korisnik preskače proces sinteze potreban za izgradnju mentalnog modela problema.

Ovaj fenomen potvrđen je u studiji Boston Consulting Group (Dell'Acqua et al., 2023., Tier 2). Na zadacima unutar „tehnološke granice" AI-a produktivnost je rasla. No, na zadacima izvan te granice, koji su zahtijevali duboku SA Razine 2 (razumijevanje suptilnog konteksta), korisnici koji su se oslanjali na AI imali su **19 postotnih bodova manju vjerojatnost točnog rješenja**. Oni su postali „Kiborzi" – potpuno integrirani s alatom, ali bez sposobnosti neovisne verifikacije, za razliku od „Kentaura" koji su zadržali kontrolu nad procesom.



Pedagoška dimenzija: Ironije automatizacije u obrazovanju

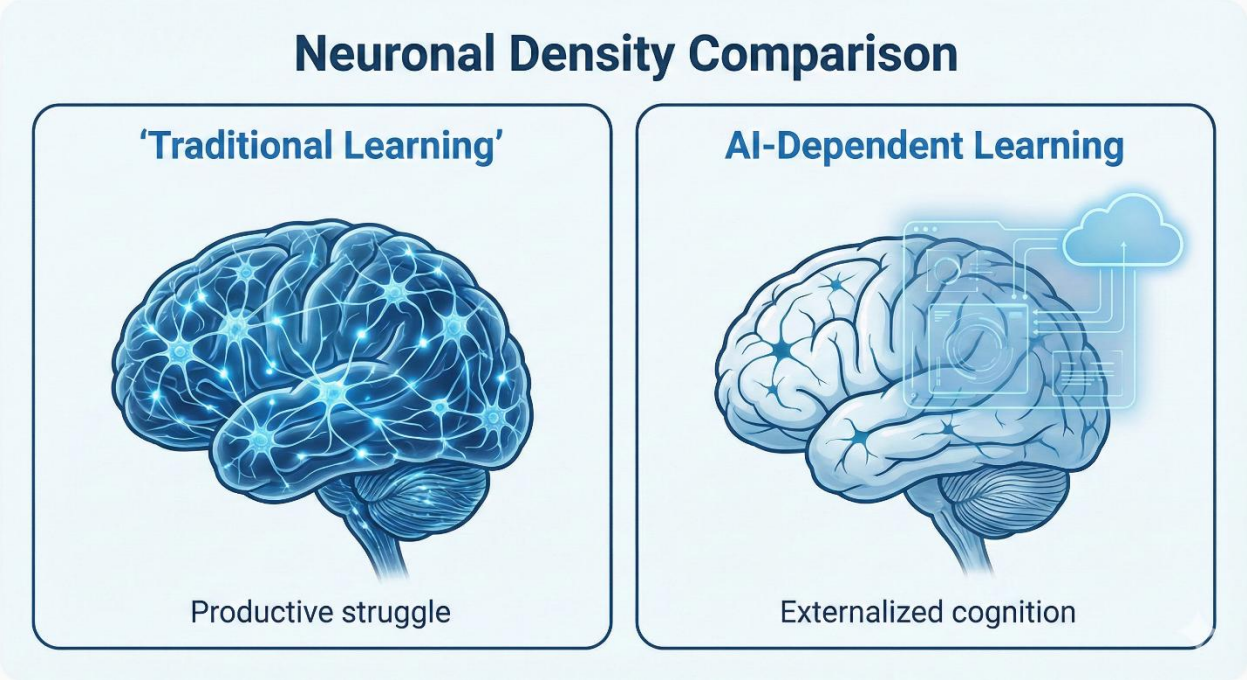
Lisanne Bainbridge (1983., Tier 1) formulirala je čuvene „ironije automatizacije": što je sustav napredniji, to je ljudski operater manje potreban za rutinski rad, ali **više potreban** za rješavanje iznimaka koje sustav ne može savladati. Paradoks je u tome što automatizacija operateru oduzima upravo ono iskustvo rutinskog rada koje mu je potrebno da razvije vještine za rješavanje iznimaka.

Kognitivna atrofija i „šuplji um"

Klein i Klein (2025., Tier 3) primjenjuju ovaj okvir na obrazovanje kroz koncept „šupljeg uma" (*Hollowed Mind*). Kada studenti koriste AI za generiranje eseja ili rješavanje problema, oni ne samo da štede vrijeme („kognitivno rasterećenje"), već zaobilaze **germani kognitivni napor** (*germane cognitive load*) – napor koji je neurobiološki nužan za formiranje dugoročnih shema znanja (Sweller, 2011.).

Barcaui et al. (2025., Tier 3) pružaju empirijsku potvrdu ovog mehanizma kroz rigoroznu RCT studiju: studenti koji su koristili ChatGPT postigli su **11 postotnih bodova lošije rezultate** na testu retencije znanja nakon 45 dana u usporedbi s kontrolnom skupinom. Ovo je klasična manifestacija Bainbridgeine ironije: alat dizajniran da pomogne učenju zapravo erodira mehanizam učenja (produktivnu borbu). Bez borbe s materijalom, neuralne veze ostaju slabe i privremene.

Neuronska gustoća



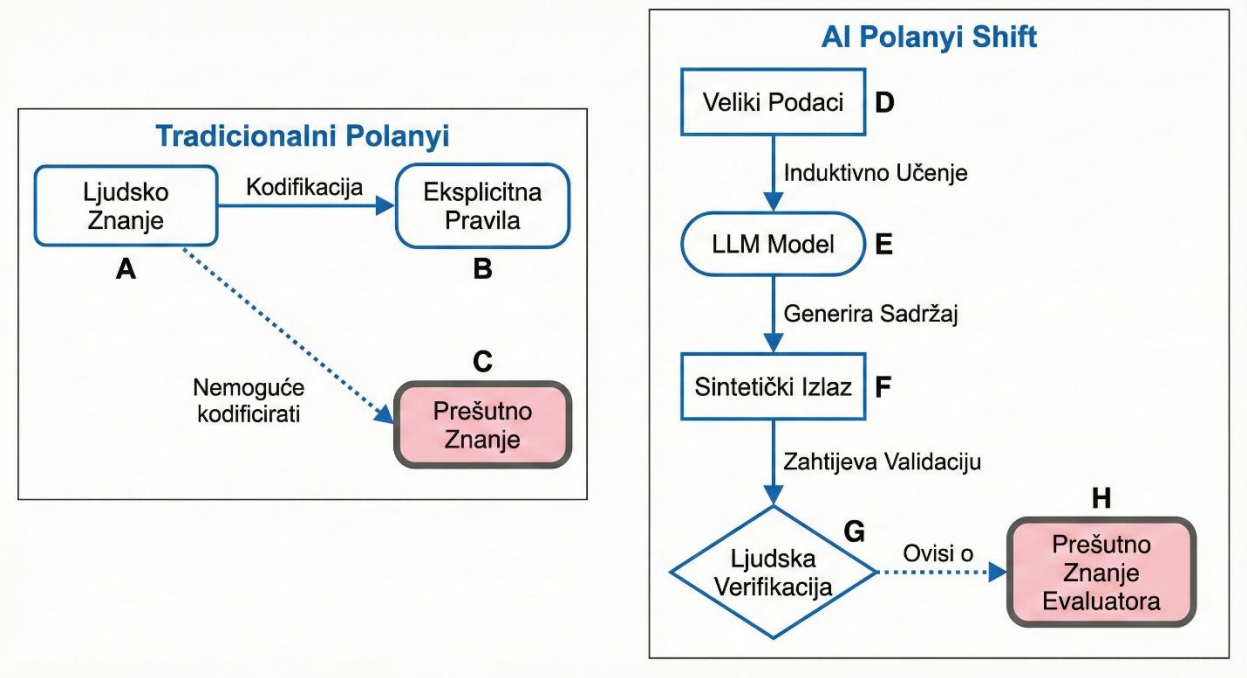
Ekonomska dimenzija: Tehnološka promjena pristrana senioritetu

Kognitivni mehanizmi opisani iznad imaju potencijal za stvaranje makroekonomskih neravnoteža. Ako AI omogućuje „laiku da zvuči kao ekspert" ali mu ne daje sposobnost da „razmišlja kao ekspert" (zbog gubitka SA), tržište rada mora reagirati na tu asimetriju.

SBTC (Seniority-Biased Technological Change)

Brynjolfsson et al. (2025., Tier 3) identificiraju fenomen **tehnološke promjene pristrane senioritetu**. Za razliku od prethodnih valova digitalizacije koji su favorizirali „mlade i digitalno pismene", GenAI favorizira seniore.

Razlog leži u **prešutnom znanju** (*tacit knowledge*) (Polanyi, 1966.). Seniori posjeduju internalizirane mentalne modele (SA Razina 2/3) stečene godinama iskustva, koji im omogućuju da brzo verificiraju izlaz AI modela i koriste ga kao multiplikator produktivnosti. Juniori, kojima nedostaje to iskustvo, a AI im oduzima priliku da ga steknu kroz „šegrtovanje" na jednostavnim zadacima, postaju ekonomski rizični.



Kvantifikacija tržišnog udara

Analiza Lichtinger i Hosseini Maasoum (2025., Tier 3) pokazuje pad od **22 % u zapošljavanju na ulaznim pozicijama** u tvrtkama koje su usvojile AI. Ovo sugerira strukturnu eliminaciju „poligona za vježbu". Tvrtke optimiziraju kratkoročnu produktivnost (AI zamjenjuje juniora), ali time dugoročno stvaraju rizik za „generacijsku obnovu" eksperata.

Alternativna objašnjenja i ograničenja SBTC hipoteze

Iako podaci snažno sugeriraju SBTC, nužno je razmotriti i alternativne uzroke pada zapošljavanja juniora kako bismo izbjegli „tehnološki determinizam":

Ciklički faktori: Razdoblje 2023.–2025. karakterizira globalna korekcija u tehnološkom sektoru (post-COVID efekti) i visoke kamatne stope, što disproporcionalno pogađa rizičnija (juniorska) zapošljavanja.
Premještanje poslova (*Offshoring*): Moguće je da se dio ulaznih poslova ne automatizira, već seli na tržišta jeftinije radne snage (iako GenAI smanjuje jezične barijere, što može ubrzati ovaj proces).
Vremenski pomak: Moguće je da se tržište još nije prilagodilo i da će se pojaviti novi tipovi ulaznih poslova („AI supervizori") koji još nisu vidljivi u statistikama.

Stoga, tvrdnju o „kolapsu" treba tumačiti kao **snažan signal strukturnog rizika**, a ne kao definitivan, ireverzibilan ishod.

Granice paradoksa i testabilne hipoteze (Falsifikabilnost)

Kako bi ovaj teorijski okvir bio znanstveno robustan, definiramo uvjete pod kojima bi se on pokazao netočnim. „Paradoks sintetičke kompetencije" je falsifikabilan kroz sljedeće hipoteze:

Hipoteza 1 (H1): Divergencija plaća unutar zanimanja

Predviđanje: Ako je teorija točna, trebali bismo vidjeti povećanje disperzije plaća *unutar* istih zanimanja (npr. programeri), gdje seniori s verifikacijskim vještinama dobivaju značajnu premiju, a juniori stagniraju.
Falsifikacija: Ako AI dovede do kompresije plaća (ujednačavanja) jer „svatko može raditi sve" uz pomoć alata, teorija o premiji na prešutno znanje pada. (Trenutni podaci Jaccoud et al., 2025., podržavaju H1).

Hipoteza 2 (H2): Nelinearnost usvajanja i kvalitete (U-krivulja)

Predviđanje: U domenama s visokim rizikom (medicina, pravo), početni rast produktivnosti bit će praćen fazom povećanih grešaka i privremenog povlačenja automatizacije zbog fenomena samozadovoljnosti (*complacency*).
Falsifikacija: Linearan i stabilan rast produktivnosti bez porasta stope incidenata u visokorizičnim sektorima.

Zaključak: Od determinizma do upravljanja rizikom

Analiza ukazuje na to da se nalazimo u „dolini nekompetencije": AI je dovoljno dobar da zavara, ali nedovoljno pouzdan da potpuno zamijeni stručnjaka bez nadzora.

Etablirani mehanizmi (Što znamo)

Možemo s visokom sigurnošću tvrditi da interakcija s LLM-ovima aktivira klasične mehanizme **automatizacijske samozadovoljnosti** (Parasuraman) i **gubitka svijesti o situaciji** (Endsley). Empirijski podaci potvrđuju da pasivna uporaba alata vodi do smanjenja retencije znanja (Barcaui) i kritičkog mišljenja (Li).

Hipotetski ishodi (Što predviđamo)

Makroekonomske posljedice su manje izvjesne, ali podaci o zapošljavanju snažno indiciraju rizik od **intergeneracijskog prekida prijenosa znanja** (Ide, 2025.). Ako se trend eliminacije ulaznih pozicija nastavi, dugoročni trošak gubitka ljudskog kapitala mogao bi nadmašiti kratkoročne dobitke u produktivnosti.

Rješenje nije tehnološko, već organizacijsko: implementacija **"Human-in-the-Loop" (HITL)** protokola koji nisu samo formalna „provjera", već su dizajnirani tako da zadrže čovjeka u kognitivnoj petlji. To zahtijeva paradoksalan inženjerski potez: namjerno dizajniranje „kognitivnog trenja" u procesu kako bi se očuvala budnost operatera.

Zahvala AI suradnicima (Author's Note)

Ovaj rad rezultat je hibridne suradnje čovjeka i umjetne inteligencije, pri čemu su alati korišteni strogo u okviru "Kentaur" modela rada:

Perplexity Pro: Korišten za dubinsko rudarenje podataka (*data mining*), klasifikaciju literature prema razinama pouzdanosti (Tier 1–3) i pronalaženje seminalnih radova kognitivne znanosti (Parasuraman, Bainbridge).
ChatGPT (Model 5.2): Korišten za rigoroznu kritičku recenziju (*adversarial review*), identifikaciju slabosti u argumentaciji (posebno u dijelu falsifikabilnosti) i predlaganje protuargumenata.
Gemini Pro (3.0 Pro): Korišten za sintezu teksta, stilsko ujednačavanje, generiranje koda (Mermaid dijagrami) i lekturu prema standardima hrvatskog jezika.

Ljudski autor zadržava punu odgovornost za konačnu sintezu, odabir izvora i finalne zaključke.

Bibliografija

Tier 1: Temeljna literatura (Kognitivna znanost i Ljudski čimbenici)

Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced 'complacency'. *The International Journal of Aviation Psychology*, 3(1), 1–23.

Polanyi, M. (1966). *The Tacit Dimension*. Doubleday & Company.

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation*, 55, 37–76.

Tier 2: Validacijske i metodološke studije

Cui, K. Z., Demirer, M., Jaffe, S., Musolff, L., Peng, S., & Salz, T. (2025). The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers. *MIT Economics Working Paper*.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Working Paper*, No. 24-013.

Merritt, S. M., et al. (2019). Automation-induced complacency potential: Development and validation of a new scale. *Frontiers in Psychology*, 10, 225.

van de Merwe, K., et al. (2022). Agent transparency, situation awareness, and mental workload: A systematic review. *Human Factors*, 64(8).

Tier 3: Generativni AI i recentna empirija (2023.–2025.)

Barcaui, A., et al. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. *SSRN Working Paper*.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). *Generative AI and Early-Career Unemployment: Seniority-Biased Technological Change Evidence*. NBER Working Paper Series.

Jaccoud, F., et al. (2025). *Robots & AI Exposure and Wage Inequality: Divergent Effects on Within-Occupation Wage Dispersion*. Maastricht University / UNU-MERIT Working Paper.

Klein, C. R., & Klein, R. (2025). The Extended Hollowed Mind: Foundational Knowledge in the AI Age. *Frontiers in Artificial Intelligence*.

Lichtinger, D., & Hosseini Maasoum, A. (2025). *AI-Adopting Firms and Early-Career Hiring: Evidence of Seniority-Biased Technological Change*. Large-scale job posting analysis.

Li, J., et al. (2024). An empirical study on factuality hallucination in large language models. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*.

Omar, M., et al. (2025). Benchmarking the confidence of large language models in answering clinical questions. *JMIR Medical Informatics*, 13, e66917.

Wang, J., et al. (2025). The effect of ChatGPT on students' learning performance. *Nature Human Behaviour*.

Zhang, Y., & Lee, J. D. (2025). Training for AI: Addressing the gap in human-AI interaction skills. *International Journal of Human-Computer Interaction*.

Literatura za daljnje čitanje (Extended Reading)

Ovaj popis uključuje temeljne radove koji pružaju širi kontekst za razumijevanje sjecišta tehnologije, ekonomije i kognicije.

Klaster 1: Ekonomija rada i tehnološka promjena

Acemoglu, D., & Restrepo, P. (2019). *Automation and New Tasks: How Technology Displaces and Reinstates Labor*. Journal of Economic Perspectives, 33(2), 3–30. (Ključni model za razumijevanje disbalansa između automatizacije i stvaranja novih zadataka).
Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press. (Ekonomski okvir za razumijevanje AI-a kao pada cijene predikcije).
Autor, D. H. (2015). *Why Are There Still So Many Jobs? The History and Future of Workplace Automation*. Journal of Economic Perspectives. (Obrana teze o komplementarnosti, koju SBTC sada izaziva).
Frey, C. B., & Osborne, M. A. (2017). *The future of employment: How susceptible are jobs to computerisation?* Technological Forecasting and Social Change. (Seminalni rad o riziku automatizacije).
Susskind, D. (2020). *A World Without Work: Technology, Automation, and How We Should Respond*.

Metropolitan Books.

Klaster 2: Kognitivna znanost i psihologija automatizacije

Hancock, P. A. (2019). *Some pitfalls in the promises of automated and autonomous vehicles*. Ergonomics, 62(4), 479–495. (Analiza nepouzdanosti autonomnih sustava i ljudskog faktora).
Hoffman, R. R., & Yates, J. F. (2005). *Decision making and automation*. In *Cambridge Handbook of Thinking and Reasoning*. (Temeljni pregled kako automatizacija mijenja proces odlučivanja).
Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux. (Osnove kognitivnih pristranosti koje AI eksploatira, posebno Sustav 1 vs. Sustav 2).
Norman, D. A. (1990). *The 'problem' with automation: Inappropriate feedback and interaction, not 'over-automation'*. Philosophical Transactions of the Royal Society of London. (Klasični tekst o važnosti dizajna interakcije).
Risko, E. F., & Gilbert, S. J. (2016). *Cognitive offloading*. Trends in Cognitive Sciences, 20(9), 676–688. (Temeljni pregled mehanizama kognitivnog rasterećenja prije pojave LLM-a).

Klaster 3: Filozofija tehnologije i društveni utjecaj

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. (Kontekst za razumijevanje dugoročnih rizika).
Crawford, K. (2021). *Atlas of AI*. Yale University Press. (Materijalni i radni troškovi AI-a).
Floridi, L. (2023). *The Ethics of Artificial Intelligence*. Oxford University Press. (Etički okvir za odgovornost).
Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books. (Psihološki aspekti ovisnosti o tehnologiji).
Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs. (Kontekst komodifikacije ljudskog iskustva).

Klaster 4: Rana istraživanja o generativnom AI-u

Bubeck, S., et al. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. arXiv preprint. (Tehnička analiza sposobnosti modela).
Eloundou, T., et al. (2023). *GPTs are GPTs: An early look at the labor market impact potential of large language models*. arXiv preprint. (Prva velika studija o izloženosti tržišta rada LLM-ovima).
Mollick, E. (2024). *Co-Intelligence: Living and Working with AI*. Portfolio. (Praktični vodič za "Centaur" pristup radu).
Noy, S., & Zhang, W. (2023). *Experimental evidence on the productivity effects of generative artificial intelligence*. Science, 381(6654), 187–192. (Ključni eksperiment na piscima).