ISSN (online):
3102-601X

Ban Dvor
ČASOPIS ZA DRUŠTVENE
I HUMANISTIČKE ZNANOSTI

Pregledni rad
Prihvaćeno: 2. 10. 2026.
Objavljeno: 22. 10. 2026.

eLokator: bd2027-008
DOI: https://doi.org/10.68268/bd/chbu7u

Utjecaj generativne umjetne inteligencije na digitalnu autentičnost i povjerenje u online komunikaciji

doc. dr. sc. Damir Vasilj
https://orcid.org/0000-0001-9946-649X
Filozofski fakultet Sveučilišta u Mostaru, Mostar, Bosna i Hercegovina — damir.vasilj@sum.ba
doc. dr. sc. Marijan Primorac
https://orcid.org/0000-0003-0188-8331
Filozofski fakultet Sveučilišta u Mostaru, Mostar, Bosna i Hercegovina — marijan.primorac@ff.sum.ba
Ivan Torbica, mag. inf.
https://orcid.org/0009-0002-6906-1189
Filozofski fakultet Sveučilišta u Mostaru, Mostar, Bosna i Hercegovina — ivan.torbica@summc.sum.ba

Sažetak: Generativna umjetna inteligencija mijenja uvjete u kojima korisnici procjenjuju podrijetlo digitalnih poruka, vjerodostojnost informacija i pouzdanost njihovih pošiljatelja. Cilj je rada kritički razmotriti odnos između sudjelovanja umjetne inteligencije u stvaranju sadržaja, percipirane autentičnosti i povjerenja u online komunikaciji. Primjenjuje se kritički narativni pregled odabranih istraživanja, uz konceptualnu sintezu nalaza o AI autorstvu, označavanju sadržaja i sintetičkim komunikatorima. Razlikuju se empirijski nalazi iz tržišne komunikacije i šire implikacije koje tek treba provjeriti u drugim komunikacijskim okružjima. Analiza pokazuje da se autentičnost ne može svesti na ljudsko podrijetlo poruke, niti se vjerodostojnost može poistovjetiti s njezinom stvarnom točnošću. Negativne reakcije na AI autorstvo ovise o očekivanom ljudskom doprinosu, vrsti poruke i načinu predstavljanja komunikatora. Transparentnost istodobno omogućuje informiranu procjenu i može izazvati nepovoljnije vrednovanje sadržaja. Kao doprinos predlaže se konceptualni okvir provjerljivoga povjerenja koji povezuje podrijetlo sadržaja, obrazloženost tvrdnji, uredničku odgovornost i mogućnost ispravka. Okvir je analitički prijedlog, a ne empirijski potvrđen model. Zaključuje se da odgovorna online komunikacija zahtijeva više od oznake o uporabi umjetne inteligencije: korisniku treba omogućiti procjenu dokaza i utvrđivanje odgovornoga pošiljatelja. Razmatraju se ograničenja postojećih istraživanja te implikacije za informacijsku i AI pismenost.

Ključne riječi: generativna umjetna inteligencija, digitalna autentičnost, vjerodostojnost, povjerenje, online komunikacija, transparentnost, AI pismenost

Ovaj je rad objavljen pod licencijom Creative Commons Autorstvo 4.0 Međunarodna (CC BY 4.0).

The Impact of Generative AI on Digital Authenticity and Trust in Online Communication

Abstract: Generative artificial intelligence changes the conditions under which users assess the origin of digital messages, information credibility, and the trustworthiness of communicators. This paper critically examines the relationship between AI involvement in content production, perceived authenticity, and trust in online communication. It employs a critical narrative review of selected studies and a conceptual synthesis of research on AI authorship, disclosure, and synthetic communicators. Empirical findings from marketing contexts are distinguished from implications requiring further investigation in other communication settings. The analysis indicates that authenticity cannot be reduced to human origin and that perceived credibility must be distinguished from factual accuracy. Negative responses to AI authorship depend on expected human involvement, message type, and how communicators are presented. Disclosure supports informed assessment while potentially prompting less favourable evaluations. The paper proposes a conceptual framework for verifiable trust linking content provenance, evidential support, editorial accountability, and opportunities for correction. This framework is an analytical proposal rather than an empirically validated model. Responsible online communication requires more than an AI label: users need opportunities to assess evidence and identify accountable communicators. The paper discusses limitations of existing research and implications for information and AI literacy.

Key words: generative artificial intelligence, digital authenticity, credibility, trust, online communication, transparency, AI literacy

1 UVOD

Generativna umjetna inteligencija otvara pitanje na temelju čega korisnik može procijeniti poruku kada njezin uvjerljiv jezični ili vizualni oblik ne otkriva način nastanka. U online komunikaciji moguće je susresti tekst koji je osoba samostalno napisala, tekst koji je uredila uz pomoć umjetne inteligencije i sadržaj koji je sustav gotovo u cijelosti generirao. Razlika među njima može biti nevidljiva primatelju, a ipak važna za njegovo razumijevanje autorove namjere, uloženoga truda i odgovornosti. U ovom se radu ta promjena razmatra iz perspektive informacijskih i komunikacijskih znanosti.

Komunikacija posredovana umjetnom inteligencijom (AI-posredovana komunikacija) u ovome radu označava komunikaciju u kojoj sustav sudjeluje u uređivanju, oblikovanju ili generiranju poruke. Pojam sintetički komunikatori radno označava računalno oblikovane likove ili digitalne persone koje se publici predstavljaju kao govornici ili sugovornici, primjerice virtualne influencere i virtualne predstavnike organizacija. Njihove poruke mogu oblikovati ljudi, generativni sustavi ili oboje; stoga sintetički komunikator nije nužno autonoman AI sustav. Pojam se odnosi na predstavljeni komunikacijski lik, a ne na osobu ili organizaciju koja odgovara za njegove objave.

Problemi procjene digitalnih informacija prethode širokoj uporabi generativnih sustava. Metzger i Flanagin razmatraju vjerodostojnost izvora i poruke u umreženom okružju te ulogu kognitivnih heuristika, odnosno prečaca kojima se korisnici služe pri procjeni informacija.[1] Generativna umjetna inteligencija tom problemu dodaje mogućnost da se obilježja povezana s ljudskim autorstvom, poput osobnoga tona ili dojma neposrednoga obraćanja, proizvode automatizirano. Stoga valja odvojeno promatrati kako poruka djeluje, kako je nastala i što podupire njezine tvrdnje.

Dosadašnja literatura o umjetnoj inteligenciji u marketingu koristan je početni izvor za ovu raspravu. Davenport i suradnici povezuju primjenu umjetne inteligencije s promjenama strategije, ponašanja korisnika i etičkim pitanjima, dok Huang i Rust razlikuju funkcije automatizacije, obrade podataka i rada s emocionalnim aspektima interakcije.[2] Međutim, učinkovitost komunikacijske proizvodnje i opravdanost povjerenja različiti su kriteriji.

[1] Metzger i Flanagin, 2013, str. 210–214.
[2] Davenport i sur., 2020, str. 24; Huang i Rust, 2021, str. 30.

Ubrzana izrada sadržaja ne daje odgovor na pitanje može li se korisnik pouzdati u objavljenu informaciju.

Cilj je rada rekonceptualizirati raspravu o generativnoj umjetnoj inteligenciji tako da se pozornost s primjena u komunikaciji brendova premjesti na odnos podrijetla sadržaja, autentičnosti, vjerodostojnosti i povjerenja. Postavljaju se tri istraživačka pitanja. Prvo, kako pojmovno razgraničiti digitalnu autentičnost, vjerodostojnost i povjerenje te odrediti što svaki od tih pojmova obuhvaća u analizi komunikacije posredovane umjetnom inteligencijom? Drugo, pod kojim se uvjetima uporaba umjetne inteligencije u stvaranju poruke i obavještavanje primatelja o toj uporabi povezuju s nižom percipiranom autentičnošću i vjerodostojnošću poruke ili pošiljatelja te manjim povjerenjem primatelja? Treće, koje sastavnice komunikacijskoga postupka trebaju omogućiti primatelju procjenu opravdanosti povjerenja u poruku i pošiljatelja, uz informaciju o ljudskom ili strojnom podrijetlu sadržaja?

Središnja je teza da održivo povjerenje zahtijeva mogućnost provjere komunikacijskoga postupka i sadržaja. Ljudsko podrijetlo može biti važno za određene poruke, ali samo po sebi ne potvrđuje istinitost. Jednako tako, sintetičko podrijetlo ne dokazuje da je poruka netočna. Doprinos rada sastoji se u povezivanju tih razlika u konceptualni okvir provjerljivoga povjerenja te u oblikovanju istraživačkih propozicija za njegovu buduću provjeru.

2 METODOLOŠKI PRISTUP I GRANICE PREGLEDA

Rad je koncipiran kao kritički narativni pregled s konceptualnom sintezom. Polazišni korpus čine izvori prethodno okupljeni za analizu generativne umjetne inteligencije u digitalnoj komunikaciji brendova. Zadržani su radovi koji pridonose objašnjenju AI autorstva, percipirane autentičnosti, vjerodostojnosti, transparentnosti, sintetičkih komunikatora i organizacijske odgovornosti. Izvori koji prvenstveno razmatraju vrijednost marke ili opći marketinški menadžment nisu uključeni kada nisu bili potrebni za novu argumentaciju. Korpus je ciljano dopunjen literaturom o procjeni online vjerodostojnosti i AI pismenosti te objavljenim ispravkom jednoga uključenog članka.

Odabir je teorijski i problemski usmjeren. Kriterij relevantnosti jest povezanost izvora s barem jednim od triju istraživačkih pitanja. Stariji radovi služe kao teorijska ili tehnološka podloga, dok novija istraživanja omogućuju razmatranje reakcija na generativne sadržaje.

Bibliografski podatci i dostupni opisi istraživanja provjereni su prema izdavačkim stranicama, znanstvenim portalima i institucijskim repozitorijima.

Pregled nije sustavni pregled literature ni metaanaliza. Ne navodi se iscrpno pretraživanje baza, broj pregledanih zapisa, neovisno dvostruko kodiranje ili formalna procjena rizika od pristranosti jer takvi postupci nisu provedeni u okviru ove sinteze. Cilj nije utvrditi ukupnu veličinu učinka generativne umjetne inteligencije na povjerenje. Izraz utjecaj u naslovu označava problemsko područje, dok se pojedini kauzalni zaključci zadržavaju unutar uvjeta eksperimenata na koje se rad poziva.

Analiza uspoređuje pet obilježja: ulogu umjetne inteligencije u proizvodnji, vrstu poruke, način otkrivanja njezine uporabe, predmet korisničke procjene i komunikacijski kontekst. Posebno se razlikuju povjerenje u pošiljatelja, vjerodostojnost poruke i ponašajne namjere. Namjera kupnje ili dijeljenja sadržaja ne tretira se kao istovjetna povjerenju. Nalazi iz oglašavanja i komunikacije brendova služe za prepoznavanje mogućih komunikacijskih i psiholoških mehanizama, odnosno procesa kojima informacija o načinu nastanka poruke oblikuje primateljevu procjenu. To uključuje pripisivanje ljudskoga truda i emocionalne uključenosti pošiljatelju, promjene percipirane autentičnosti nakon oznake o uporabi umjetne inteligencije te pripisivanje ljudskosti virtualnom liku na temelju njegova načina obraćanja. Mogućnost djelovanja tih procesa u javnoj, obrazovnoj ili interpersonalnoj komunikaciji ostaje otvoreno istraživačko pitanje.

3 POJMOVNO RAZGRANIČENJE

3.1 GENERATIVNI SUSTAVI KAO KOMUNIKACIJSKI POSREDNICI

Generativni sustavi sudjeluju u oblikovanju poruka, a ne samo u njihovoj distribuciji. Tehnološku podlogu suvremenoga razvoja jezičnih modela među ostalim čini transformerska arhitektura koju su predstavili Vaswani i suradnici.[3] Za komunikacijsku analizu važnija je raspodjela uloga između osobe, sustava i organizacije nego detaljan opis arhitekture. Potrebno je utvrditi tko određuje namjeru poruke, tko proizvodi formulaciju, tko provjerava navode i tko odobrava objavu.

[3] Vaswani i sur., 2017, str. 5998.

Pregled Verme i suradnika daje širi kontekst istraživanja primjene umjetne inteligencije u marketingu.[4] U ovome se radu iz tog konteksta izdvajaju tri analitičke situacije: pomoć pri jezičnom uređivanju već oblikovane ljudske poruke, zajedničko stvaranje sadržaja i pretežito automatizirana proizvodnja. Riječ je o radnoj podjeli stupnja sudjelovanja, a ne o sveobuhvatnoj tehničkoj klasifikaciji. Korisniku može biti važno je li sustav promijenio pravopis ili je osmislio iskaz koji se predstavlja kao nečije osobno iskustvo.

Digitalno okružje također obuhvaća interpersonalne odnose, a ne samo organizacijske objave. Primorac i Primorac Bilaver istražuju stavove studenata o interpersonalnoj komunikaciji u digitalnom okružju.[5] Njihov rad ovdje služi za pozicioniranje komunikacijskoga konteksta, a ne kao empirijski dokaz o generativnoj umjetnoj inteligenciji. Automatizirana poruka može se pojaviti u istom prostoru u kojem se očekuju osobna prisutnost, uzajamnost i prepoznatljiv autor.

3.2 DIGITALNA AUTENTIČNOST

Digitalna autentičnost u ovom radu obuhvaća odnos između predstavljenoga identiteta i podrijetla poruke te načina na koji ih primatelj tumači. Analitički se razlikuju autentičnost podrijetla, izražajna autentičnost i dosljednost komunikatora. Prva se odnosi na to odgovara li prikazani izvor stvarnom postupku nastanka. Druga se odnosi na doživljaj da poruka izražava namjeru ili iskustvo onoga tko se njome predstavlja. Treća označava usklađenost poruke s prethodnim ponašanjem i prepoznatljivim identitetom pošiljatelja. Ova je podjela autorsko razgraničenje za potrebe sinteze, a ne preuzeta validirana mjerna ljestvica.

Takvo razlikovanje omogućuje precizniju interpretaciju sintetičkih komunikatora. Koles i suradnici razmatraju višedimenzionalne manifestacije autentičnosti virtualnih influencera, čime otvaraju mogućnost da publika pripisuje autentičnost i neljudskom liku.[6] Transparentno fikcionalan lik može biti dosljedan svojoj predstavljenoj ulozi, iako nema ljudsko iskustvo. Problem nastaje kada komunikacija stvara neutemeljen dojam da taj lik posjeduje iskustvo, identitet ili svjedočanstvo stvarne osobe.

[4] Verma i sur., 2021, 100002, str. 1–3.
[5] Primorac i Primorac Bilaver, 2022, str. 96.
[6] Koles i sur., 2024, 114325, str. 1.

Važno je pritom razlučiti virtualnost i generativnost. Virtualni influencer može biti računalno oblikovan lik čiju svaku objavu pišu ljudi. Istraživanja takvih influencera ne predstavljaju automatski istraživanja autonomne generativne umjetne inteligencije. Njihova je vrijednost za ovu raspravu u objašnjenju pripisivanja identiteta i društvenih obilježja sintetičkom izvoru, uz zadržavanje razlike među tehnologijama.

3.3 VJERODOSTOJNOST I POVJERENJE

Vjerodostojnost označava procjenu da su izvor ili poruka vrijedni vjerovanja. Uključuje percepciju pouzdanosti i stručnosti izvora, ali se može procjenjivati i na razini samoga sadržaja.[7] To je percepcija primatelja i treba je razlikovati od činjenične točnosti. Uvjerljivo formulirana netočna tvrdnja može djelovati vjerodostojno, dok se točna informacija može odbaciti zbog nepovjerenja prema njezinu podrijetlu.

Povjerenje se ovdje radno određuje kao spremnost osloniti se na poruku, komunikatora ili postupak u uvjetima određene neizvjesnosti. U istraživanju stoga treba imenovati njegov predmet. Povjerenje u organizaciju koja objavljuje AI-generirani tekst nije isto što i povjerenje u model koji ga je proizveo. Korisnik može smatrati sustav prikladnim za jezično uređivanje, a istodobno odbijati osloniti se na njega za provjeru spornih činjenica.

Autentičnost, vjerodostojnost i povjerenje mogu biti povezani, ali njihov odnos nije nužno jednosmjeran. Prepoznatljivo ljudska poruka može djelovati iskreno, a sadržavati pogrešku. Precizan automatizirani odgovor može biti koristan, iako korisnik ne doživljava emocionalnu povezanost s njegovim izvorom. Te razlike sprječavaju zaključivanje da svako smanjenje percipirane autentičnosti znači i smanjenje stvarne kvalitete informacije.

4 KRITIČKA SINTEZA EMPIRIJSKIH NALAZA

4.1 AI AUTORSTVO I OČEKIVANI LJUDSKI DOPRINOS

Brüns i Meißner pokazuju da uporaba generativne umjetne inteligencije za sadržaj na društvenim mrežama može umanjiti percipiranu autentičnost brenda te da su reakcije

[7] Metzger i Flanagin, 2013, str. 211.

povoljnije kada tehnologija pomaže ljudima nego kada ih zamjenjuje.[8] Taj nalaz upućuje na važnost pripisane uloge tehnologije. Za primatelja nije nužno isto prepušta li pošiljatelj sustavu tehničku obradu poruke ili posao koji publika smatra izrazom njegova osobnoga angažmana.

Kirk i Givi u nizu eksperimenata nalaze nepovoljnije reakcije na emocionalnu marketinšku komunikaciju pripisanu umjetnoj inteligenciji. Učinak slabi kod činjeničnih poruka i kada umjetna inteligencija samo uređuje tekst; autori također opisuju uvjete u kojima se smjer učinka mijenja.[9] Zbog toga nije opravdano iz njihova rada izvesti tvrdnju da publika uvijek odbacuje AI autorstvo. Važan je odnos između autorstva i onoga što poruka obećava primatelju.

U širem komunikacijskom smislu iz toga se izvodi hipoteza da će očekivana osobna uključenost biti važnija za ispriku ili iskaz podrške nego za standardiziranu obavijest. Time se pretpostavlja da primatelj vrednuje poruku prema očekivanom osobnom angažmanu pošiljatelja. Primjena toga objašnjenja izvan marketinga interpretativni je prijedlog, a ne već dokazan rezultat za sve navedene vrste poruka. Eksperiment koji uspoređuje brendove ne može sam potvrditi kako će ista osoba procijeniti poruku prijatelja, nastavnika ili javne ustanove.

4.2 TRANSPARENTNOST I PARADOKS OZNAČAVANJA

Sahebi, Formosa i Bankins izravnije proširuju raspravu izvan oglašavanja. U eksperimentalnom istraživanju uspoređuju ljudsko, AI-potpomognuto i potpuno AI autorstvo u kontekstu poslovne e-pošte i objave na društvenim mrežama. Nalaze niže procjene autentičnosti i pouzdanosti autora te manju spremnost na uporabu sadržaja kada komunikacija uključuje umjetnu inteligenciju. Istodobno se pojavljuje napetost između očekivanja da primatelj bude obaviješten o uporabi umjetne inteligencije i slabijega vrednovanja tako označene komunikacije.[10] Taj paradoks pokazuje da zahtjev za transparentnošću i pozitivna ocjena označenoga sadržaja nisu ista pojava.

[8] Brüns i Meißner, 2024, 103790, str. 8–10.
[9] Kirk i Givi, 2025, 114984, str. 10.
[10] Sahebi i sur., 2026, 100304, str. 1.

Grigsby i suradnici razmatraju označavanje generiranih oglasa za usluge, pri čemu važnu ulogu imaju naglasak oglasa i prikaz pružatelja usluge.[11] Njihov objavljeni ispravak vraća izostavljene negativne predznake u prikazu dijela posrednih učinaka.[12] U ovom se radu nalaz upotrebljava kao dokaz o kontekstualnoj uvjetovanosti reakcija, bez preuzimanja jedinstvene veličine učinka ili tvrdnje da je svaka transparentna uporaba podjednako nepovoljna.

Balaban i Phua u kontekstu oglašavanja turističkih odredišta na Instagramu ispituju povezanost otkrivanja AI-generiranoga sadržaja i tipa izvora s procjenama potrošača.[13] Wu kod personificiranih brendova nalazi da označavanje može smanjiti ponašajne namjere preko promjena percipirane autentičnosti i ljudskosti.[14] Zajednička je implikacija da oznaka ulazi u tumačenje izvora, a ne služi samo kao tehnička obavijest o alatu.

Iz mogućega negativnog učinka oznake ne proizlazi opravdanje za prikrivanje uporabe. Takav bi zaključak zamijenio pitanje informiranosti primatelja pitanjem neposredne uvjerljivosti poruke. Primjerenija je istraživačka mogućnost usporediti kratku oznaku s objašnjenjem uloge sustava, provedenih provjera i odgovorne osobe. Pretpostavka da detaljnije objašnjenje može pomoći zahtijeva zasebno testiranje; pregledani nalazi ne potvrđuju da će ono uvijek obnoviti povjerenje.

4.3 SINTETIČKI KOMUNIKATORI I DRUŠTVENI ZNAKOVI

Muniz i suradnice pokazuju da otkrivanje neljudske naravi virtualnoga influencera može promijeniti doživljaj njegove ljudskosti, vjerodostojnosti i povjerenja u brend. Važna je granica njihova nalaza da posredni učinak na povjerenje nije potvrđen jednako u američkom i brazilskom kontekstu.[15] Time se dodatno ograničava pretpostavka o univerzalnoj reakciji publike na sintetički identitet.

Yoo i suradnici analiziraju ulogu tekstualnih društvenih znakova u vjerodostojnosti virtualnih influencera.[16] Tekstualni društveni znakovi ovdje označavaju obilježja jezičnoga izražavanja koja primatelju sugeriraju društvenu prisutnost i odnos s komunikatorom. Zato uz izgled

[11] Grigsby i sur., 2025b, 104231, str. 1, 4–7.
[12] Grigsby i sur., 2025a, 104427.
[13] Balaban i Phua, 2026, 109047, str. 1, 6–9.
[14] Wu, 2026, str. 1.
[15] Muniz i sur., 2024, str. 1234.
[16] Yoo i sur., 2025, str. 730.

virtualnoga lika treba promatrati i njegov način obraćanja te komunikacijsko ponašanje. Istodobno, društveni znak koji povećava dojam prisutnosti nije dokaz stručnosti, istinitosti tvrdnje ili stvarnoga iskustva komunikatora.

Usporedba tih istraživanja s radom Koles i suradnika ne upućuje nužno na proturječne zaključke. Doživljaj autentičnosti dosljedno predstavljenoga virtualnog lika i reakcija na naknadno otkrivanje da je prividno ljudski izvor umjetan različite su situacije. Analiza treba zato obuhvatiti početna očekivanja primatelja, trenutak otkrivanja i točan predmet procjene. Posebno je neprimjereno iz prihvaćenosti virtualnoga lika zaključivati da bi publika jednako prihvatila lažno ljudsko svjedočanstvo.

5 KONCEPTUALNI OKVIR PROVJERLJIVOGA POVJERENJA

5.1 POLAZIŠTE I SASTAVNICE OKVIRA

Na temelju prethodne sinteze predlaže se okvir provjerljivoga povjerenja. Naziv označava povjerenje za koje primatelj može razmotriti relevantne razloge, a ne tvrdnju da je samo povjerenje objektivno svojstvo sadržaja. Okvir razlikuje četiri sastavnice komunikacijskoga postupka: podrijetlo, obrazloženost tvrdnji, odgovornost i mogućnost ispravka. One su predloženi predmeti procjene i nisu predstavljene kao empirijski potvrđeni čimbenici s poznatim relativnim težinama.

Podrijetlo obuhvaća razumljivu informaciju o tome tko je sudjelovao u stvaranju poruke i na koji način. Obrazloženost označava mogućnost povezivanja činjeničnih tvrdnji s dostupnim i relevantnim dokazima. Odgovornost znači da se može prepoznati osoba ili organizacija koja je prihvatila objavu i može objasniti njezin sadržaj. Mogućnost ispravka odnosi se na postupak prijave pogreške, njezina razmatranja i vidljive korekcije. Navođenje tih sastavnica nije dovoljno ako opisani postupci stvarno ne postoje.

Tablica 1. sažima predložene sastavnice i razdvaja pitanja koja korisnik može postaviti. Okvir ne pretpostavlja da svaka poruka zahtijeva jednaku razinu dokumentiranja. Kratka jezično uređena obavijest i javna tvrdnja s ozbiljnim posljedicama razlikuju se po potrebi za dokazima. Načelo je razmjernost provjere sadržaju i posljedicama njegova prihvaćanja.

Tablica 1. Sastavnice predloženoga okvira provjerljivoga povjerenja

Sastavnica	Pitanje za procjenu	Granica zaključivanja
Podrijetlo	Tko je i u kojem opsegu sudjelovao u nastanku poruke?	Ljudsko ili AI podrijetlo samo po sebi ne potvrđuje točnost.
Obrazloženost	Koji izvor ili dokaz podupire konkretnu tvrdnju?	Navođenje izvora nije dovoljno ako izvor ne podupire tvrdnju.
Odgovornost	Tko odobrava objavu i može odgovoriti na prigovor?	Potpis i oznaka provjere ne dokazuju da je provjera provedena.
Mogućnost ispravka	Kako se pogreška prijavljuje, provjerava i ispravlja?	Postojanje kanala ne jamči pravodoban i vidljiv ispravak.

Izvor: autorska konceptualna sinteza. Sastavnice i njihovi odnosi nisu empirijski validirani.

Autentičnost u tom okviru ostaje važna, ali se ne svodi na podrijetlo. Primatelj procjenjuje odgovara li predstavljen način nastanka onome što poruka komunicira. Kada organizacija objavi osobnu ispriku, pitanje nije samo je li upotrijebljen jezični alat, nego prihvaća li odgovorna osoba iskaz, razumije li njegov predmet i poduzima li odgovarajuće radnje. Ovaj primjer ilustrira analitičko razlikovanje; ne predstavlja nalaz novoga empirijskog istraživanja.

5.2 PREDLOŽENI ODNOSI MEĐU SASTAVNICAMA

U predloženom okviru način nastanka i njegovo objašnjenje djeluju kao informacije na temelju kojih primatelj tumači ulogu pošiljatelja. Vrsta poruke i očekivani ljudski doprinos određuju relevantnost tih informacija. Percipirana autentičnost može pridonijeti procjeni vjerodostojnosti, dok dostupni dokazi i prethodno iskustvo mogu oblikovati vjerodostojnost neovisno o dojmu ljudskosti. Povjerenje je mogući ishod takve procjene, uz razlike među korisnicima i kontekstima.

Iz okvira se izvode četiri istraživačke propozicije. Prva pretpostavlja da će se negativna reakcija na AI sudjelovanje povećavati kada je osobno iskustvo ili emocionalna uključenost pošiljatelja sastavni dio očekivane vrijednosti poruke. Druga pretpostavlja da će objašnjenje konkretne uloge umjetne inteligencije omogućiti preciznije procjene od jedinstvene oznake koja ne razlikuje uređivanje i stvaranje sadržaja. Preciznija procjena ovdje znači bolje

razlikovanje stvarne uloge sustava i ljudskoga doprinosa, a ne nužno veću naklonost poruci ili pošiljatelju.

Treća propozicija pretpostavlja da će dostupnost relevantnih dokaza i vidljivoga postupka ljudske provjere ublažiti oslanjanje na samo podrijetlo kod činjeničnih poruka. Četvrta pretpostavlja da će jasna odgovornost i iskustvo pravodobnoga ispravka pridonositi obnovi povjerenja nakon pogreške. Posljednje dvije propozicije proširuju raspravu s neposredne percepcije jedne poruke na organizacijski postupak i vremensku dimenziju odnosa. Njihova empirijska valjanost ostaje otvorena.

Okvir također predviđa mogućnost opravdanoga nepovjerenja. Ako korisnik nakon provjere utvrdi da tvrdnja nema potporu, smanjeno povjerenje može biti poželjan ishod. Stoga cilj komunikacijskoga upravljanja nije maksimalno povjerenje prema svakom sadržaju, nego bolja usklađenost povjerenja s razlozima na kojima se ono temelji. Takav kriterij razlikuje informacijsku kvalitetu od uspjeha uvjeravanja.

6 IMPLIKACIJE ZA ONLINE KOMUNIKACIJU I PISMENOST

6.1 ORGANIZACIJSKA I UREDNIČKA ODGOVORNOST

Tomić, Volarić i Obradović upozoravaju na stručne i etičke zahtjeve primjene umjetne inteligencije u odnosima s javnošću, uključujući služenje javnom interesu.[17] Vasilj, Stojkić i Bubalo razmatraju umjetnu inteligenciju u marketingu i odnosima s javnošću poslovnih organizacija.[18] Ovi radovi pomažu smjestiti upravljanje AI sadržajem u organizacijske komunikacijske funkcije. Ne upotrebljavaju se kao potvrda učinkovitosti novoga konceptualnog okvira.

Iz predloženog okvira slijedi potreba da organizacija razdvoji izradu nacrta, provjeru činjenica i odobravanje objave. Odluka o uporabi generativnoga sustava trebala bi sadržavati i odluku o tome tko pregledava rezultat. Tvrdnja da je sadržaj provjerio čovjek ima informacijsku vrijednost samo ako je provjera stvarna i odgovara prirodi poruke. Jezično čitanje, primjerice, ne potvrđuje točnost statističkih podataka ni autentičnost prikazanoga svjedočanstva.

[17] Tomić i sur., 2022, str. 7.
[18] Vasilj i sur., 2023, str. 61.

Transparentnost bi trebala obuhvatiti podatke korisne primatelju. Kod obavijesti sastavljene uz jezičnu pomoć sustava može biti relevantno tko jamči za informacije i gdje se nalazi izvorna odluka. Kod sintetičkoga glasa važno je razumije li publika da sluša računalno proizvedenu izvedbu. Kod virtualnoga predstavnika treba razlikovati njegovu fikcionalnu personu od organizacije koja stoji iza objava. To su normativne preporuke izvedene iz okvira, a ne tvrdnja o jedinstvenoj zakonskoj obvezi za sve slučajeve.

Urednički postupak trebao bi uključivati i reakciju nakon objave. Korisniku valja omogućiti prijavu netočnosti, a ispravak povezati s izvornom porukom kada je to izvedivo. Samo uklanjanje spornoga teksta ne mora objasniti prethodnim primateljima što je bilo pogrešno. Za organizacijsku praksu zato je korisno razmatrati trag izmjena i odgovornost za njih kao dio komunikacije, a ne kao isključivo internu tehničku aktivnost.

6.2 INFORMACIJSKA I AI PISMENOST

Long i Magerko određuju AI pismenost kroz kompetencije za kritičko vrednovanje umjetne inteligencije te komunikaciju, suradnju i uporabu takvih tehnologija.[19] U kontekstu ovoga rada ta se perspektiva povezuje s informacijskom pismenošću: korisnik treba moći postaviti pitanje o podrijetlu poruke, procijeniti potporu tvrdnji i razlikovati uvjerljivu izvedbu od provjerljivoga dokaza.

Praktična implikacija nije osposobiti korisnika da na temelju stila nepogrešivo prepozna svaki AI-generirani tekst. Takav zahtjev zanemaruje zajedničko stvaranje sadržaja i razliku između podrijetla i kvalitete. Primjereniji je zadatak usporediti tvrdnju s izvornim dokumentom, provjeriti tko preuzima odgovornost i procijeniti odgovara li ponuđeni dokaz onome što se tvrdi. To vrijedi i za poruke koje su potpuno ljudskoga podrijetla.

Obrazovni primjer može usporediti tri verzije iste informacije: ljudski napisanu, jezično uređenu uz pomoć sustava i generiranu uz naknadnu provjeru. Sudionici bi zasebno procjenjivali jasnoću, autentičnost, vjerodostojnost i spremnost na oslanjanje, a zatim pregledali izvore. Svrha takve vježbe bila bi pokazati što oznaka podrijetla otkriva, a što ostaje nepoznato. Ovdje je riječ o prijedlogu nastavne aktivnosti čiju uspješnost treba evaluirati.

[19] Long i Magerko, 2020, str. 2.

Odgovornost se ipak ne može u cijelosti prenijeti na primatelja. Korisnik često nema pristup postupku stvaranja sadržaja ni mogućnost provjere svake tvrdnje. Zato se razvoj kompetencija treba povezati s organizacijskom praksom koja omogućuje pregled izvora i upućivanje pitanja. U protivnom se od pojedinca traži da nadoknađuje netransparentnost sustava nad kojim nema kontrolu.

7 RASPRAVA I ISTRAŽIVAČKA OGRANIČENJA

Odgovor na prvo istraživačko pitanje jest pojmovno razgraničenje autentičnosti kao odnosa predstavljenoga identiteta, podrijetla i izraza, vjerodostojnosti kao primateljeve procjene da su poruka ili izvor vrijedni vjerovanja te povjerenja kao spremnosti na oslanjanje na poruku, komunikatora ili postupak. To omogućuje usporedbu istraživanja bez pretpostavke da svi autori mjere isti ishod. Smanjena namjera kupnje, niža procjena ljudskosti i nepovjerenje u točnost informacije mogu nastati zajedno, ali se ne smiju automatski poistovjetiti.

Na drugo pitanje pregled odgovara uvjetovano. Niže procjene autentičnosti ili vjerodostojnosti i manje povjerenje nakon informacije o uporabi umjetne inteligencije ovise o vrsti poruke, pripisanom ljudskom doprinosu, načinu označavanja i predstavljanju komunikatora. Izraz niže procjene označava slabije vrednovanje tih obilježja, a ne nužno netočnu procjenu primatelja. Nalazi ne podupiru tvrdnju o općem nestanku povjerenja zbog generativne umjetne inteligencije. Izostanak jedinstvenoga učinka ne umanjuje važnost problema; on upućuje na potrebu preciznijega određivanja uvjeta u kojima se povjerenje stvara, slabi ili mijenja predmet.

Na treće pitanje odgovara predloženi okvir koji uz podrijetlo uključuje dokaze, odgovornost i mogućnost ispravka. Njegov je doprinos pomak s pitanja djeluje li poruka ljudski na pitanje može li korisnik razmotriti razloge da se na nju osloni. Okvir ne zamjenjuje postojeće teorije vjerodostojnosti ili autentičnosti. On organizira odabrane nalaze i formulira odnose koji mogu usmjeriti istraživanje AI-posredovane komunikacije.

Glavno je ograničenje namjerni odabir literature iz prethodno oblikovanoga korpusa, u kojem je tržišna komunikacija snažno zastupljena. Takvo polazište povećava rizik selekcijske pristranosti i ne omogućuje procjenu zastupljenosti svih mogućih nalaza. Ograničena dostupnost dijela cjelovitih tekstova smanjuje mogućnost usporedbe mjernih instrumenata i

procjene metodološke kvalitete. Nadalje, studije se razlikuju po podražajima, platformama, uzorcima i vremenu provedbe. Njihovi se rezultati zato ne zbrajaju u jedinstven zaključak o veličini učinka. Buduće istraživanje moglo bi usporediti oznaku o AI sudjelovanju s objašnjenjem njegova opsega, uz odvojenu manipulaciju dostupnošću izvora. Potrebno je mjeriti autentičnost, vjerodostojnost, povjerenje i stvarno provjeravanje sadržaja odvojeno, a ne zamijeniti ih jednom općom procjenom sviđanja. Uključivanje poruka različite činjenične kvalitete omogućilo bi provjeru razlikuju li korisnici pouzdane i nepouzdane informacije ili samo preferiraju određenu vrstu autorstva.

Longitudinalni nacrt mogao bi pratiti kako se povjerenje mijenja nakon otkrivene pogreške i javnoga ispravka. Za prijenos izvan marketinga potrebno je uključiti organizacijsku, obrazovnu i interpersonalnu komunikaciju, uz zasebnu analizu očekivanja primatelja. Posebno bi vrijedila provjera u hrvatskom jezičnom i regionalnom kontekstu. U ovome radu nema podataka koji bi unaprijed potvrdili da su lokalne reakcije jednake reakcijama zabilježenima u međunarodnim istraživanjima.

8 ZAKLJUČAK

Generativna umjetna inteligencija proširuje pitanje digitalnoga povjerenja jer omogućuje odvajanje uvjerljive komunikacijske izvedbe od prepoznatljivoga ljudskog doprinosa. Kritički pregled pokazuje da interpretacija njezine uporabe ovisi o vrsti poruke, očekivanom angažmanu pošiljatelja i predstavljanju komunikacijskoga procesa. Transparentnost omogućuje informiranu procjenu, ali sama po sebi ne jamči povoljnu reakciju publike. U radu su razgraničeni digitalna autentičnost, vjerodostojnost i povjerenje te predložen okvir provjerljivoga povjerenja. Njegove su sastavnice podrijetlo sadržaja, obrazloženost tvrdnji, odgovornost za objavu i mogućnost ispravka. Riječ je o konceptualnom doprinosu kojem je potrebna empirijska provjera. Vrijednost okvira jest u tome što omogućuje procjenu komunikacije i kada njezino podrijetlo nije isključivo ljudsko. Za online komunikaciju to znači da oznaku o uporabi umjetne inteligencije treba promatrati zajedno s kvalitetom izvora i postupkom objavljivanja. Za informacijsku i AI pismenost važna je sposobnost razlikovanja dojma autentičnosti od razloga za vjerovanje. Povjerenje koje se može preispitati i prilagoditi dokazima prikladniji je cilj od nastojanja da svaka digitalna poruka ostavi što uvjerljiviji dojam.

BIBLIOGRAFIJA

Balaban, Delia Cristina i Phua, Joe. 2026. „Is this real or fake? Examining the effects of disclosure of AI-generated content and source type in Instagram-based travel destination advertising on consumer attitudes and behavioral intentions." *Computers in Human Behavior* 183: 109047. https://doi.org/10.1016/j.chb.2026.109047.

Brüns, Jasper David i Meißner, Martin. 2024. „Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity." *Journal of Retailing and Consumer Services* 79: 103790. https://doi.org/10.1016/j.jretconser.2024.103790.

Davenport, Thomas, Guha, Abhijit, Grewal, Dhruv i Bressgott, Timna. 2020. „How artificial intelligence will change the future of marketing." *Journal of the Academy of Marketing Science* 48/1: 24–42. https://doi.org/10.1007/s11747-019-00696-0.

Grigsby, Jamie L., Michelsen, Meg i Zamudio, César. 2025a. „Corrigendum to 'Service ads in the era of generative AI: Disclosures, trust, and intangibility' [J. Retail. Consumer Serv. 84 (2025) 1–9 104231]." *Journal of Retailing and Consumer Services* 87: 104427. https://doi.org/10.1016/j.jretconser.2025.104427.

Grigsby, Jamie L., Michelsen, Meg i Zamudio, César. 2025b. „Service ads in the era of generative AI: Disclosures, trust, and intangibility." *Journal of Retailing and Consumer Services* 84: 104231. https://doi.org/10.1016/j.jretconser.2025.104231.

Huang, Ming-Hui i Rust, Roland T. 2021. „A strategic framework for artificial intelligence in marketing." *Journal of the Academy of Marketing Science* 49/1: 30–50. https://doi.org/10.1007/s11747-020-00749-9.

Kirk, Colleen P. i Givi, Julian. 2025. „The AI-authorship effect: Understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications." *Journal of Business Research* 186: 114984. https://doi.org/10.1016/j.jbusres.2024.114984.

Koles, Bernadett, Audrezet, Alice, Guidry Moulard, Julie, Ameen, Nisreen i McKenna, Brad. 2024. „The authentic virtual influencer: Authenticity manifestations in the metaverse." *Journal of Business Research* 170: 114325. https://doi.org/10.1016/j.jbusres.2023.114325.

Long, Duri i Magerko, Brian. 2020. „What is AI literacy? Competencies and design considerations." *U Proceedings of the* 2020 CHI Conference on Human Factors in Computing Systems, 1–16. Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727.

Metzger, Miriam J. i Flanagin, Andrew J. 2013. „Credibility and trust of information in online environments: The use of cognitive heuristics." *Journal of Pragmatics* 59: 210–220. https://doi.org/10.1016/j.pragma.2013.07.012.

Muniz, Fernanda, Stewart, Kristin i Magalhães, Lívia. 2024. „Are they humans or are they robots? The effect of virtual influencer disclosure on brand trust." *Journal of Consumer Behaviour* 23/3: 1234–1250. https://doi.org/10.1002/cb.2271.

Primorac, Marijan i Primorac Bilaver, Ivana. 2022. „Društvene mreže i interpersonalna komunikacija: Stavovi studenata o interpersonalnoj komunikaciji u digitalnome okružju." *Hum* 17/27: 96–118. https://doi.org/10.47960/2303-7431.27.2022.96.

Sahebi, Siavosh, Formosa, Paul i Bankins, Sarah. 2026. „The AI penalty and disclosure paradox: Trust, authenticity and knowledge uptake in AI-mediated communication." *Computers in Human Behavior: Artificial Humans* 8: 100304. https://doi.org/10.1016/j.chbah.2026.100304.

Tomić, Zoran, Volarić, Tomislav i Obradović, Đorđe. 2022. „Umjetna inteligencija u odnosima s javnošću." *South Eastern European Journal of Communication* 4/2: 7–16. https://doi.org/10.47960/2712-0457.2022.2.4.7.

Vasilj, Damir, Stojkić, Inja i Bubalo, Nikica. 2023. „Artificial intelligence in marketing and public relations of business organizations." *South Eastern European Journal of Communication* 5/2: 61–69. https://doi.org/10.47960/2712-0457.2.5.61.

Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N., Kaiser, Łukasz i Polosukhin, Illia. 2017. „Attention is all you need." *U Advances in Neural Information Processing Systems* 30, 5998–6008. Curran Associates. https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Verma, Sanjeev, Sharma, Rohit, Deb, Subhamay i Maitra, Debojit. 2021. „Artificial intelligence in marketing: Systematic review and future research direction." *International Journal of Information Management Data Insights* 1/1: 100002. https://doi.org/10.1016/j.jjimei.2020.100002.

Wu, Linwan. 2026. „How consumers respond when brands disclose using AI in creating social media content: Examining brand personification and communication strategies." *Journal of Product & Brand Management,* prethodna mrežna objava, 1–15. https://doi.org/10.1108/JPBM-09-2025-6302.

Yoo, Joon Woo, Park, Junsung i Park, Heejun. 2025. „How can I trust you if you're fake? Understanding human-like virtual influencer credibility and the role of textual social cues." *Journal of Research in Interactive Marketing* 19/4: 730–748. https://doi.org/10.1108/JRIM-02-2024-0092.