

VEŠTAČKA INTELIGENCIJA U MEDIJIMA: OD LAKOĆE KORIŠĆENJA DO ETIČKIH IZAZOVA

Marina Jablanov-Stojanović¹

Miljan Nikolić²

Apstrakt: Ovaj rad istražuje višestruki uticaj veštačke inteligencije (u daljem tekstu VI) na savremeni medijski milje, analizirajući kako je transformisala tradicionalne procese stvaranja i distribucije sadržaja. Istraživanje je fokusirano na ključne oblasti primene i relevantna globalna i domaća istraživanja, sa posebnim osvrtom na etičke dileme. Na početku, rad analizira veštačku inteligenciju kao ključni alat za medije, ističući njenu ulogu u personalizaciji, automatizaciji generisanja vesti i sofisticiranoj analizi publike. Detaljno se razrađuje kako algoritmi društvenih mreža utiču na formiranje političkih stavova doprinoseći medijskoj polarizaciji. Posebna pažnja posvećena je etičkim implikacijama, naročito problemu pristrasnosti algoritama i širenju dezinformacija. Najveću pretnju predstavlja tehnologija „deepfakes“ (korišćenje VI da bi se stvorile lažne fotografije, video ili audio snimci), koja omogućava uverljivu manipulaciju tim sadržajem, podrivajući poverenje u vizuelne dokaze. Na kraju, rad se bavi izazovima odgovornosti i transparentnosti. Problematika „crne kutije“ VI sistema nameće pitanja o tome ko je odgovoran za greške i štetu, što je podstaklo razvoj „objašnjive veštačke inteligencije“ i donošenje regulatornih akata poput EU AI Act. U zaključku se naglašava da budućnost medija zahteva kombinaciju tehnoloških rešenja, pravne regulacije i edukacije javnosti, kako bi se očuvao integritet informisanja u digitalnom dobu.

Ključne reči: *veštačka inteligencija, mediji, etički izazovi, pristrasnost algoritma, deepfakes*

¹ Marina Jablanov – Stojanović, odgovorna urednica portala Mariniranje, Novi Sad, Fruškogorska 19
e-mail: mariniranje@gmail.com

² Milijan Nikolić, Loznica, Kralja Petra I L-7/26
e-mail: nikolicmilijan58@gmail.com

ARTIFICIAL INTELLIGENCE IN THE MEDIA: FROM EASE OF USE TO ETHICAL CHALLENGES

Abstract: This paper explores the multiple impacts of artificial intelligence (AI) on the contemporary media landscape, analyzing how it has transformed the traditional process of content creation and distribution. The research focuses on key areas of application and relevant global and domestic research, with a special focus on ethical dilemmas. Initially, the paper analyzes artificial intelligence as a key tool for the media, highlighting its role in personalization, automation of news generation, and sophisticated audience analysis. It elaborates in detail how social media algorithms influence the formation of political attitudes, contributing to media polarization. Special attention is paid to ethical implications, especially the problem of algorithmic bias and the spread of disinformation. The greatest threat is posed by "deepfakes" (the use of AI to create fake photos, video, or audio recordings), a technology that enables convincing manipulation of such content, undermining trust in visual evidence. Finally, the paper addresses the challenges of accountability and transparency. The issue of "black box" AI systems raises questions about who is responsible for errors and harm, which has prompted the development of "explainable artificial intelligence" and the adoption of regulatory acts such as the EU AI Act. The conclusion emphasizes that the future of media requires a combination of technological solutions, legal regulations and public education, in order to preserve the integrity of information in the digital age.

Keywords: *artificial intelligence, media, ethical challenges, algorithmic bias, deepfakes*

UVOD

Nije tajna da digitalni svet sve više utiče na oblikovanje naše stvarnosti, a veštačka inteligencija (u daljem radu VI) predstavlja brzo rastuće polje informatike koje se bavi stvaranjem sistema sposobnih da obavljaju zadatke koji obično zahtevaju ljudsku inteligenciju. Misli se pre svega na učenje, rešavanje problema, percepciju i razumevanje jezika. Od svojih početaka VI je evoluirala od jednostavnih logičkih sistema do složenih neuronskih mreža koje mogu da obrađuju ogromne količine podataka. Danas, ona prožima skoro sve aspekte savremenog života, sa posebnim uticajem na medije i komunikaciju, odnosno društvene mreže.

Na prvim linijama ove velike promene nalaze se mediji, kao glavni kanal za informisanje, edukaciju i zabavu javnosti. Tradicionalni modeli, koji su se oslanjali na

linearne procese kreiranja, distribucije i konzumacije sadržaja, bivaju potkopani i redefinisani algoritmima koji donose automatizaciju, personalizaciju i koji mogu da se nadograđuju i proširuju do nezamislivih razmera. Međutim, važno je istaći da ova moćna tehnologija ne dolazi bez značajnih izazova. Sa pružanjem alata za neslućenu efikasnost istovremeno VI otvara Pandorinu kutiju etičkih dilema, od algoritamske pristrasnosti i dezinformacija do transparentnosti i pitanja odgovornosti. Ovaj rad ima za cilj da istraži ovaj dvostruki uticaj, analizirajući kako VI transformiše medijski milje, dok istovremeno baca svetlo na etičke implikacije koje bez odlaganja zahtevaju usmeravanje ne samo svesti, nego i naše pažnje kako bi se očuvao integritet informisanja u digitalnom dobu.

1. VEŠTAČKA INTELIGENCIJA I MEDIJI

Da bi se u potpunosti shvatio uticaj veštačke inteligencije, neophodno je razumeti kako je ona preobrazila i još uvek preobražava bazu medijske industrije. Istorijski gledano, medijski sadržaj su proizvodile centralizovane redakcije i distribuirale ga masovnoj publici putem štampanih izdanja, radija a kasnije i televizije. Publika je bila prvenstveno pasivni primalac, sa ograničenim mogućnostima izbora ili interakcije. Digitalizacija je započela proces fragmentacije ovog modela, međutim, tek je uvođenje VI doneo njegovu potpunu razgradnju i ponovnu izgradnju na drugačijim osnovama.

Zato se i kaže da veštačka inteligencija transformiše medijske temelje. Ona nije samo jedan od alata koje poseduju medijske kuće u svom arsenalu, VI je infrastruktura koja podupire celokupno medijsko iskustvo. Ulazi i restrukturuiše sve faze medijskog ciklusa: od istraživanja teme i stvaranja nacrtu, preko produkcije i montaže videa, do distribucije, personalizacije i analize performansi. Kako se tradicionalni modeli stvaranja i distribucije sadržaja menjaju, tako VI postaje ključni alat za personalizaciju, automatizaciju i analizu. Najčešće se koristi za automatizaciju tema koje se zasnivaju na podacima kao što su sportski rezultati ili finansijski izveštaji. Međutim, njen uticaj može da se oseti u svakoj fazi medijskog ciklusa, od kreiranja vesti do interakcije sa korisnicima. Zato korišćenje VI u medijima postavlja i nove etičke dileme, s obzirom na mogućnost manipulacije i algoritamsku pristrasnost.

1.1. Veštačka inteligencija kao ključni alat: personalizacija, automatizacija i analiza

Transformativni uticaj VI ogleđa se kroz tri ključna aspekta: personalizaciju sadržaja, automatizaciju procesa i analizu podataka. Svaki od ovih aspekta donosi revolucionarne prednosti, ali i podstiče nova pitanja o prirodi novinarstva, odnosima medija i publike, i etičkim granicama tehnologije. U ovom segmentu, istražićemo kako veštačka inteligencija redefiniše ove procese, fokusirajući se upravo na personalizaciju, automatizaciju i analizu.

1.2. Personalizacija sadržaja

Jedan od najznačajnijih doprinosa VI medijima je sposobnost personalizacije sadržaja. Suština ove funkcije leži u algoritmima koji neprestano analiziraju enormne količine podataka korisnika. Naime, svaki klik, svaki lajk, svako deljenje, vreme provedeno na određenom članku, pa čak i pauze tokom reprodukcije videa – sve su to podaci koji se prikupljaju, spajaju i tumače kako bi se izgradio detaljan profil interesa, sklonosti i navika korisnika.

Umesto univerzalnog sadržaja za sve, medijske platforme sada mogu da kreiraju jedinstveno iskustvo za svakog korisnika. Ogromnu popularnost i angažman korisnika Netflix, YouTube, Spotify, Facebook i platforma X (bivši Twitter) duguju upravo sofisticiranim algoritmima za preporuku. Algoritmi mašinskog učenja analiziraju ponašanje korisnika — istoriju pretrage, vreme provedeno na određenim člancima, lajkove i deljenja — kako bi im preporučili sadržaj koji odgovara njihovim interesovanjima. Ovaj proces povećava angažman korisnika i vreme provedeno na platformi (O’Neil & Schutt, 2013). Napredna analitika omogućava medijskim kućama da razumeju demografiju svoje publike i njihove navike, što im pomaže da optimizuju ponudu i plasiraju oglase na efikasniji način (Nakamura & Takemura, 2019).

1.3. Automatizacija procesa

Automatizacija je aspekt veštačke inteligencije koji je je možda najtransformativniji i čini samo jezgro medijskog sadržaja. Sada VI obavlja niz zadataka koji su ranije zahtevali značajan ljudski rad, radikalno povećavajući brzinu, obim i efikasnost. Najpoznatiji primer ove prakse je automatizovano novinarstvo.

Niz zadataka koji su ranije zahtevali ljudski rad automatizuje se uz pomoć VI koja obavlja povećavajući brzinu i efikasnost. Na primer: softver zasnovan na VI može automatski generisati izveštaje iz strukturiranih podataka, kao što su finansijski rezultati kompanija ili sportski rezultati. The Associated Press je pionir u ovoj oblasti, koristeći automatizovano novinarstvo za masovnu produkciju kratkih, tačnih vesti (Graefe, 2016).

Trebalo bi naglasiti da se automatizacija ne odnosi samo na tekst, nego da postoje VI sistemi za prepoznavanje lica i predmeta koji automatski označavaju materijale u medijskim arhivama, što i te kako olakšava pretragu i ponovnu upotrebu sadržaja. Naravno, ovo smanjuje operativne troškove, ali i oslobađa novinare kopanja po arhivama i omogućava im da se fokusiraju na složenije istraživačke zadatke, kreativno pisanje, intervjuje i rad na dubljoj analizi.

1.4. Analiza podataka

Treći stub upotrebe VI u medijima je moć analize. Uz pomoć tehnike dubokog učenja VI je izuzetno je moćan alat za analizu ogromnih setova podataka. Medijske kuće koriste ovu analitičku pomoć kako bi razumele ne samo šta se dešava, već i zašto se dešava i šta će se verovatno dogoditi u budućnosti.

Veštačka inteligencija je izuzetno moćan alat za analizu masivnih setova podataka, poznatih kao „Big Data“. Medijske kuće koriste ovu analizu za predviđanje trendova u potrošnji sadržaja, praćenje performansi članaka u realnom vremenu i identifikaciju tema koje su od najvećeg interesa za publiku. Analiza sentimenta, zasnovana na obradi prirodnog jezika (Natural Language Processing NLP u daljem tekstu - oblast veštačke inteligencije koja se bavi interakcijom između računara i ljudskog, prirodnog, jezika), omogućava platformama da razumeju emocionalne reakcije publike na određene vesti ili događaje (Zou & Lu, 2019). Trebalo bi naglasiti da su ove informacije ključne za donošenje strateških odluka o sadržaju i distribuciji i, naravno, za efikasnije plasiranje oglasa koji se mogu uskladiti sa trenutnim raspoloženjem korisnika.

2. PRIMENA VEŠTAČKE INTELIGENCIJE NA DRUŠTVENIM MREŽAMA

Društvene mreže su verovatno najvidljivije polje primene veštačke inteligencije. Upravo na tim platformama algoritmi ostvaruju najdirektniji uticaj na milijarde ljudi, ali istovremeno i najjasnije otkrivaju svoje tamnije strane. Veštačka inteligencija je temelj na kojem funkcionišu platforme poput Meta (Facebook, Instagram), TikTok-a, X i YouTube-a, upravljajući svim aspektima korisničkog iskustva.

1. Personalizacija sadržaja: Algoritmi preporuka analiziraju korisničko ponašanje (lajkove, deljenja, komentare) kako bi im prikazali sadržaj koji će im verovatno biti interesantan. Cilj je povećati angažman i vreme provedeno na platformi (O'Neil & Schutt, 2013).
2. Moderacija sadržaja: Automatizovani sistemi za moderaciju sadržaja, zasnovani na VI, filtriraju neprikladan sadržaj kao što su govor mržnje, nasilje ili pornografija, pre nego što ga vide ljudi-moderatori. Iako su efikasni, ovi sistemi se suočavaju sa izazovima u razumevanju konteksta i suptilnosti (Gorwa, 2019).
3. Detekcija lažnih vesti ispama: Algoritmi za detekciju dezinformacija koriste različite pristupe kako bi identifikovali lažni sadržaj. Ti pristupi se mogu podeliti na analizu sadržaja, analizu konteksta i analizu širenja.

2.1. Analiza sadržaja

Modeli obrade prirodnog jezika (NLP) su osnova za analizu teksta. Novija istraživanja se fokusiraju na naprednije tehnike koje ne samo da prepoznaju ključne reči, već i analiziraju stil, ton i semantičke nepravilnosti koje su često prisutne u lažnim vestima (Zhang et al., 2020). Na primer: algoritmi traže tekst koji je emocionalno nabijen, preterane tvrdnje i odsustvo pozivanja na pouzdane izvore.

2.2. Analiza konteksta

Detekcija lažnih vesti ne može se osloniti samo na tekst – ovde je kontekst je ključan. To podrazumeva analizu izvora (da li je u pitanju verifikovan medij ili nepoznati nalog), analizu pratećih medijskih elemenata (slike, video), kao i proveru navodnih činjenica. Napredne neuronske mreže mogu da uporede informaciju sa proverenim bazama podataka i prepoznaju nelogičnosti. Primer toga je korišće-

nje grafovskih neuronskih mreža (GNN) koje analiziraju kako se informacije šire kroz mrežu korisnika, prepoznajući botove i lažne naloge koji su često povezani u klastere (Sun et al., 2021).

2.3. Analiza širenja

Lažna vest je kao prosut džak perja, a demanti je skupljanje svakog pojedinačnog pera, tako da nikad demanti nema toliku snagu kao lažna vest prosuta kroz medije. S druge strane svima je poznato da se lažne vesti se često šire brže i agresivnije od istinitih, sa naglim skokovima u deljenjima i reakcijama.

Dobra vest je da algoritmi mogu da prepoznaju te šablone i aktiviraju upozorenja, a još bolja je da se i u Srbiji akademska zajednica i stručnjaci sve više bave ovim problemom. Istraživanja se najčešće sprovode na fakultetima koji se bave informacionim tehnologijama, novinarstvom i komunikacijama kao što su: Fakultet organizacionih nauka (FON) u Beogradu, Fakultet političkih nauka (FPN) u Beogradu, Centar za istraživanje medija i komunikacija (CIRCOM).

Borba protiv dezinformacija i spema je kontinuirana i izuzetno kompleksna. Upravo ovde veštačka inteligencija može da bude ključno oruđe u toj borbi, ali njena efikasnost zavisi od stalnog unapređivanja modela i tehnika. Novija istraživanja se fokusiraju na celovit pristup, kombinujući analizu sadržaja, konteksta i širenja, dok se u Srbiji radi na prilagođavanju tih modela lokalnom jezičkom i društvenom kontekstu. Ova evolucija VI u detekciji lažnih vesti je neophodna kako bi se očuvala verodostojnost informacija i poverenje u digitalne medije.

3. VEŠTAČKA INTELIGENCIJA U PRODUKCIJI I ANALIZI MEDIJSKOG SADRŽAJA

Veštačka inteligencija sve više služi za stvaranje i optimizaciju medijskog sadržaja. Od automatskog generisanja vesti do sofisticirane analize navika publike, VI transformiše procese rada u medijskim kućama. Ovaj segment detaljnije razrađuje ključne primere upotrebe VI, uzimajući u obzir novija globalna i domaća istraživanja.

3.1. Generisanje vesti

Automatizovano novinarstvo je proces u kojem VI generiše tekstove iz strukturiranih podataka. Iako je Associated Press pionir u ovoj oblasti još od 2014. godine, koristeći Automated Insights za finansijske izveštaje, novija istraživanja se bave mnogo složenijim oblicima generisanja (Graefe, 2016).

Noviji modeli, poput velikih jezičkih modela (LLM) kao što su GPT-3, GPT-4, Llama, mogu generisati složenije tekstove, članke i zaključke. Istraživači se bave pitanjem kako ovi modeli mogu pomoći novinarima u svakodnevnom radu, na primer, automatskim kreiranjem nacрта vesti ili prilagođavanjem teksta za različite platforme (Döveling & Vreese, 2021). Takođe, raste interesovanje za generativno novinarstvo, gde VI ne samo da piše vesti, već i stvara personalizovane narative za svakog čitaoca. Neka istraživanja čak predlažu da se VI može koristiti za prevođenje vesti na različite jezike u realnom vremenu, što može biti od značaja za globalne medijske kompanije.

U Srbiji su istraživanja o stvaranju vesti još u ranoj fazi. Pojedini radovi na domaćim konferencijama i u studentskim radovima analiziraju potencijal primene LLM-a za automatsko pisanje na srpskom jeziku. Fokus je često na izazovima prevođenja i prilagođavanja jezičkih modela specifičnostima srpskog jezika i medijskog konteksta. U praksi, primena je sporija, mada pojedine medijske kuće eksperimentišu sa ovim alatima za kreiranje kratkih izveštaja ili sažetaka.

3.2. Prepoznavanje lica i predmeta

Tehnologija prepoznavanja slika i videa je ključna za efikasno upravljanje velikim medijskim arhivama. Kroz duboko učenje (deep learning) i konvolutivne neuronske mreže (CNN), VI sistemi automatski taguju, indeksiraju i pretražuju vizuelni sadržaj (Chen et al., 2020). Noviji radovi se bave multi-modalnom analizom gde se istovremeno obrađuju tekst i slika (npr. naslov, opis) kako bi se postigla veća tačnost, tako da medijska kuća može automatski prepoznati poznate ličnosti i objekte na video-snimku i istovremeno ga povezati sa relevantnim vestima i informacijama iz svoje arhive.

Istraživanja u Srbiji u ovoj oblasti su razvijenija, s obzirom na to da su domaći IT stručnjaci aktivni u razvoju softverskih rešenja na ovom polju. Katedre za računarske nauke na Elektrotehničkom fakultetu u Beogradu i Fakultetu tehničkih nauka u Novom Sadu često rade na projektima koji uključuju prepoznavanje objekata i lica, iako se ne fokusiraju isključivo na medije. Mnogi diplomski i doktorski radovi bave se optimizacijom algoritama za lokalne uslove.

3.3. Analiza publike i predviđanje trendova

Analiza publike je evoluirala od jednostavnog brojanja pregleda do sofisticiranog predviđanja ponašanja. Analitika zasnovana na VI omogućava medijima da razumeju ne samo šta publika konzumira, već i zašto i kako. Algoritmi predviđanja analiziraju podatke u realnom vremenu kako bi identifikovali trendove i preporučili sadržaj koji bi mogao da postane viralan (Nakamura & Takemura, 2019). Novija istraživanja su se fokusirala na analizu sentimenta i prepoznavanje emocionalnih stanja publike. Ovi modeli analiziraju komentare i reakcije na društvenim mrežama kako bi medijima pružili uvid u to kako određene priče rezoniraju sa publikom (Liu & Wu, 2022). Trebalo bi istaći da je ovo ključno za optimizaciju sadržaja i monetizaciju, jer omogućava bolju segmentaciju publike za plasiranje reklama.

Primećuje se da u Srbiji medijske kuće sve više koriste analitičke alate za praćenje publike, ali se retko radi o internom razvoju sofisticiranih VI modela. Umesto toga, koriste se komercijalna rešenja koja nude analizu podataka. Može da se kaže da je upotreba VI u produkciji i analizi medijskog sadržaja u stalnom razvoju.

4. ETIČKE IMPLIKACIJE I IZAZOVI

Da li se filozofski principi, odnosno etički modeli poznatih filozofa kao što je Kant mogu primeniti na veštačku inteligenciju? To pitanje nije novo, ali je sve aktuelnije s obzirom na rapidan razvoj tehnologije. Ideja da se Kantovi kategorički imperativi, formulisani u 18. veku, mogu primeniti na sisteme veštačke inteligencije je intrigirajuća i sve više se istražuje, naročito u oblasti etike veštačke inteligencije. Zašto se ovo pitanje postavlja? Zato, jer je Kantov pristup, zasnovan na univerzalnim moralnim zakonima i racionalnoj autonomiji i nudi okvir za razmišljanje o tome kako bi VI trebalo da se ponaša, a ne samo da funkcioniše. A to nas upućuje da malo razmotrimo drugu (tamnu) stranu veštačke inteligencije.

Moguće zloupotrebe su zastrašujuće:

1. Politička dezinformacija – stvaranje lažnih izjava
2. Kleveta ili ucenjivanje – stvaranje kompromitujućih snimaka javnih ličnosti
3. Erozija poverenja – dugoročno je najveća opasnost jer kad postane dovoljno lako i jeftino proizvesti uverljiv lažan dokaz počinje sumnja u svaki dokaz

4.1. Tamna strana korišćenja veštačke inteligencije u medijima

Iako VI donosi neosporne prednosti, njena primena u medijima otvara niz ozbiljnih etičkih pitanja i izazova. Mogućnost manipulacije, problem pristrasnosti algoritama i nedostatak transparentnosti predstavljaju pretnje koje mogu narušiti integritet novinarstva i poverenje javnosti. U ovom segmentu detaljnije razrađujemo ova ključna pitanja, uz osvrt na novija istraživanja i specifičnosti u Srbiji.

4.1.1. Potencijalna manipulacija i pristrasnost

Algoritmi su u osnovi matematički modeli koji se koriste na podacima, a ti podaci često odražavaju društvene predrasude i nejednakosti koje već postoje u svetu. Kada se ovi podaci koriste za obuku VI sistema, rezultat je algoritam koji reprodukuje i pojačava te predrasude, što je problem koji je detaljno opisala Cathy O'Neil u svojoj knjizi "Weapons of Math Destruction" (O'Neil, 2016). Na primer: algoritam preporuka na društvenim mrežama, koji je dizajniran da zadrži korisnika na platformi može nesvesno da stvori „filter mehure“ i „eho-komore“. Ovi fenomeni se dešavaju kada algoritam korisniku prikazuje samo sadržaj koji potvrđuje njegova postojeća uverenja, što dovodi do informacijske izolacije i ekstremne polarizacije društva (Flaxman et al., 2016).

Novija istraživanja idu korak dalje, baveći se „algoritamskom pristrasnošću“ ne samo u sadržaju, već i u načinu na koji se informacije predstavljaju. Studija iz 2021. godine pokazala je da VI alati za proveru činjenica, iako korisni, mogu biti pristrasni, jer je njihova efikasnost često uslovljena političkim stavom koji je inherentan podacima na kojima su obučeni (Hube et al., 2021). To znači da čak i sistemi stvoreni za borbu protiv dezinformacija mogu imati svoju „stranu“.

U Srbiji su istraživanja o algoritamskoj pristrasnosti prisutna na fakultetima kao što su Fakultet političkih nauka (FPN) i Fakultet organizacionih nauka (FON). Autori se bave uticajem algoritama na političku komunikaciju i javno mnjenje. Radovi, često u formi zbornika sa konferencija, analiziraju kako algoritmi društvenih mreža utiču na formiranje političkih stavova i doprinose polarizaciji u srpskom društvu (Jokić & Srećković 2022).

4.1.2. „Filter mehurovi“ i „eho komore“

Algoritmi društvenih mreža su dizajnirani da maksimizuju angažman korisnika. Da bi to postigli, preporučuju sadržaj koji je u skladu sa našim prethodnim interesovanjima i interakcijama (Flaxman et al., 2016). Posledica ovog personalizovanog pristupa je stvaranje „filter mehurova“ (eng. filter bubbles) – informacionih balona u kojima korisnik vidi samo sadržaj koji potvrđuje njegova postojeća

uverenja. Unutar ovih mehurova, formiraju se „eho-komore“ (eng. echo chambers) – okruženja u kojima se dominantno čuju istomišljenici, što ojačava postojeće stavove i smanjuje izloženost drugačijim perspektivama (Murdock & Johnson, 2022).

4.1.3. *Emocionalna manipulacija*

Algoritmi daju prednost sadržaju koji izaziva jake emocije, posebno bes, strah i ogorčenje. Studije su pokazale da takav sadržaj ima veće šanse da se proširi brže i dalje od neutralnih ili pozitivnih vesti (Brady et al., 2020). Kada se porukama da emocionalni okvir, one se brže šire i snažnije utiču na formiranje stavova. Ova dinamika često podstiče radikalizaciju i agresivnost u političkim diskusijama.

4.1.4. *Cenzura i smanjenje vidljivosti (shadowbanning)*

Iako platforme tvrde da su neutralne, algoritmi mogu da deluju kao oblik cenzure. Praksa smanjenja vidljivosti ili „senčenje“ (eng. shadowbanning), određenih objava bez obaveštavanja autora, postala je sve češća. Ako algoritam proceni da je neki politički stav „problematičan“, bez obzira na to da li je reč o govoru mržnje ili samo kontroverznom mišljenju, može ga učiniti manje vidljivim za druge korisnike. To stvara utisak da određeni stavovi nisu široko rasprostranjeni ili da su neprihvatljivi, čime se manipuliše percepcijom javnog mnjenja (Gorwa, 2019).

4.2. *Manipulacija informacijama i „Deepfakes“*

Razvoj generativne veštačke inteligencije (Generative AI) doveo je do pojave „deepfakes“ – izuzetno uverljivih, ali lažnih foto, video i audio zapisa, kreiranih pomoću dubokih neuronskih mreža. Ova tehnologija predstavlja ogromnu pretnju za novinarstvo, jer podriva verodostojnost vizuelnih dokaza. U rukama zlonamernih aktera, „deepfakes“ se mogu koristiti za širenje dezinformacija, blaćenje političkih protivnika ili stvaranje lažnih izjava uticajnih ličnosti (Westerlund, 2019).

Novija istraživanja su se fokusirala na to kako prepoznati „deepfakes“ na osnovu mikroskopskih nepravilnosti koje ljudsko oko ne može da primeti, kao što su treptanje ili senke. Istovremeno, novinari i istraživači rade na razvoju alata za proveru porekla medijskog sadržaja (provenance tracking), kako bi se utvrdilo da li je neki snimak autentičan (Koike et al., 2022).

U Srbiji, takođe, postoji interesovanje za „deepfakes“. Iako su akademska istraživanja u ranoj fazi, pojedine analize medijske pismenosti i kritičke analize medijskih sadržaja bave se ovim fenomenom u kontekstu dezinformju „deepfakes“ i provere činjenice.

5. IZAZOVI „CRNE KUTIJE”: PITANJE ODGOVORNOSTI I TRANSPARENTNOSTI

Upotreba veštačke inteligencije (VI) u medijima nameće kompleksna pitanja o odgovornosti i transparentnosti. S obzirom na to da su algoritmi sve složeniji i manje razumljivi čak i svojim tvorcima, postavlja se pitanje ko je odgovoran kada se napravi greška, šire dezinformacije ili dođe do nenamernih posledica. Problem „crne kutije” odnosi se upravo na to – neprozirnost algoritamskog procesa koji nam ne dozvoljava da razumemo kako se donose odluke. U ovom segmentu analiziramo ključne aspekte ovog problema, uz osvrt na novija globalna i domaća istraživanja.

5.1. Problem „crne kutije”

Jedno od najtežih etičkih pitanja je odgovornost. Kada VI sistem napravi grešku, širi dezinformacije ili izazove štetu, ko snosi odgovornost? Da li je to programer koji je stvorio algoritam, sam algoritam (što je pravno nemoguće) medijska kuća koja ga koristi, ili korisnik? Nedostatak transparentnosti u načinu rada algoritama, poznat kao problem „crne kutije” i dodatno otežava utvrđivanje odgovornosti. Mnogi napredni modeli VI, kao što su duboke neuronske mreže, funkcionišu kao „crne kutije”. To znači da znamo šta je ulaz (podaci) i izlaz (odluke ili akcija), ali ne možemo precizno da objasnimo putanju kojom je algoritam došao do te odluke (Burrell, 2016).

U medijskom kontekstu, ovo stvara značajne izazove:

1. Novinarska etika: Ako algoritam za preporuku vesti favorizuje senzacionalistički sadržaj koji izaziva polarizaciju, medijska kuća možda neće moći da objasni zašto se to dešava. Nedostatak transparentnosti narušava poverenje javnosti u medije.
2. Algoritamska pristrasnost: Kada se desi da algoritam, zbog pristrasnosti u podacima, diskriminiše određene grupe, teško je utvrditi uzrok problema i ispraviti ga ako ne razumemo kako algoritam „misli”.

Novija istraživanja se fokusiraju na rešenja za ovaj problem. Objasnjiva veštačka inteligencija (XAI) je novo polje čiji je cilj da se razviju alati i metode koji omogućavaju da algoritmi ne budu samo efikasni, već i transparentni i razumljivi ljudima (Adadi & Berrada, 2018).

5.2. Pitanje odgovornosti

Ko je odgovoran kada automatizovani sistem napravi grešku sa ozbiljnim posledicama? Ovo pitanje se nameće na više nivoa:

1. Programeri i kompanije: Da li su programeri koji su kreirali algoritam odgovorni za njegove posledice? Ili je odgovornost na kompaniji koja ga je implementirala?
2. Platforme prema korisnicima: Društvene mreže se često izjašnjavaju kao neutralne platforme, a ne medijske kuće, kako bi izbegle odgovornost za sadržaj koji se na njima objavljuje. Međutim, algoritmi koje koriste aktivno oblikuju taj sadržaj.
3. Regulatorni okviri: Mnoge zemlje i međunarodne organizacije pokušavaju da se uhvate u koštac sa ovim problemom. EU Akt o veštačkoj inteligenciji (EU AI Act), koji je u procesu donošenja, jedan je od prvih pokušaja da se postave jasni standardi za odgovorno korišćenje VI, naročito u visokorizičnim sektorima (European Commission, 2021).

ZAKLJUČAK

Put veštačke inteligencije u medijima je staza promena, transformacije i nezvesnosti. Od automatizovanog novinarstva i personalizacije sadržaja do borbe protiv lažnih vesti i „deepfakes“ tehnologije, VI je nepovratno promenila medijski milje. Iako je donela efikasnost i nove mogućnosti, postavila je i ozbiljne etičke izazove koji zahtevaju veoma pažljivo razmatranje. Algoritamska pristrasnost, širenje dezinformacija kao i problem odgovornosti i transparentnosti su ključni izazovi koji se moraju rešavati. Globalna istraživanja i regulativni okviri nude smer-nice, a u Srbiji se akademska i stručna zajednica sve više bavi prilagođavanjem tih rešenja lokalnom kontekstu. Budućnost medija u digitalnom dobu zavisi od toga kako će medijske kuće, programeri, zakonodavci i sami korisnici upravljati ovim moćnim alatima. Ključ za suočavanje s ovim izazovima leži u kombinaciji tehnoloških rešenja (kao što je XAI), pravnih regulacija i edukacije javnosti. Samo tako će mediji moći da zadrže svoju ulogu pouzdanog izvora informacija. Primena „starog znanja“ kao što je Kantov kategorički imperativ na moderne „alate“ poput veštačke inteligencije nije samo akademska vežba, nego sve značajniji praktičan izazov. Kantovi principi univerzalnosti i poštovanja ljudskog dostojanstva pružaju neprolazan okvir za rešavanje gorućih problema kao što su algoritamska pristrasnost, „deepfakes“ i zloupotreba podataka. U tom svetlu i drugi aspekt Kantove

etike je tretiranje ljudskih bića kao cilja samog po sebi, a ne samo kao sredstva za postizanje nekog drugog cilja. Algoritmi koji tretiraju ljude kao skup podataka čiji je cilj da maksimizira vreme gledanja i prihode od reklama, krše ovaj princip. Ljudi su svedeni na sredstvo za ostvarivanje profita platformi, a ne na autonomne pojedince koji imaju pravo na istinite informacije i raznovrsne perspektive. Iako sistemi veštačke inteligencije ne mogu posedovati moralnu autonomiju u ljudskom smislu, Kantova etika može poslužiti kao moralni kompas za dizajnere, programere i zakonodavce, osiguravajući da se VI razvija na način koji služi dobrobiti čovečanstva, a ne za njegovu patnju i na štetu.

REFERENCE

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *Proceedings of the International Conference on Intelligent Systems and Networks*.
2. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2020). Emotion shapes the diffusion of moralized content in social networks. *Psychological Science, 31*(2), 173–193.
3. Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1).
4. Chen, J., Zhang, Z., & Li, C. (2020). Automated image and video tagging for content management in media archives. *Journal of Media Computing and Informatics*.
5. Döveling, K., & Vreese, C. H. (2021). A study of automated journalism: How journalists in Germany and the Netherlands perceive AI-written news. *Journalism Studies, 22*(8), 987–1004.
6. European Commission (2021). *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*
7. Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly, 80*(S1), 298–320.
8. Graefe, A. (2016). *Guide to automated journalism: How algorithms are writing the news*. Columbia University Press.
9. Gorwa, R. (2019). The platform governance: An overview. *Internet Policy Review, 8*(3).
10. Hube, M., Weiler, C., & Becker, L. (2021). The political bias of fact-checking algorithms: A study of the German and US news ecosystem. In *Proceedings of the 5th International Conference on Computational Social Science (ICCS)*.

11. Jokić, S., & Srećković, N. (2022). Medijska polarizacija u digitalnom dobu: uloga algoritama. U *Zborniku radova sa Konferencije o medijima i komunikacijama*. Beograd: FPN.
12. Koike, T., et al. (2022). A deep learning approach to deepfake detection based on physiological signal analysis. *IEEE Transactions on Multimedia*, 24, 2568–2579.
13. Liu, Y., & Wu, X. (2022). Emotion and sentiment analysis for social media users: A survey. *ACM Computing Surveys*, 55(1), 1–36.
14. Murdock, K., & Johnson, P. (2022). The impact of social media algorithms on political polarization. *Journal of Communication and Public Policy*, 21(3).
15. Nakamura, T., & Takemura, M. (2019). AI-based media production and audience analytics. *Journal of Media Business Studies*, 16(4).
16. O’Neil, S., & Schutt, R. (2013). *Doing data science: Straight talk from the frontline*. O’Reilly Media.
17. Schoemaker, P. J. H. (2021). The AI-driven future of ethics and morality. *Journal of Business Ethics*, 172(3).
18. Sun, T., Zhao, Z., & Chen, J. (2021). Graph neural networks for fake news detection on social media. *IEEE Transactions on Knowledge and Data Engineering*.
19. Westerlund, M. (2019). The emergence of deepfakes: The good, the bad and the ugly. *Technology Innovation Management Review*, 9(12).
20. Zhang, S., Jiang, H., & Lin, C. (2020). A survey of fake news detection. *Journal of Information Science and Engineering*.
21. Zou, X., & Lu, Y. (2019). Fake news detection in social media using machine learning. In *Proceedings of the 2019 International Conference on Information Technology and Communication Systems*.

Cite – APA Style:

Jablanov-Stojanović, M., Nikolić, M. (2025). Veštačka inteligencija u medijima: od lakoće korišćenja do etičkih izazova. U: B. Sančanin & A. Penjišević (ur.), *Zbornik radova VII međunarodne naučne konferencije Veštačka inteligencija u kreiranju medijskih sadržaja* (pp. 216-230). Sremski Karlovci: Centar za kulturu, edukaciju i medije; Istočno Sarajevo–Sokolac: Visoka škola za uslužni biznis.

Рад је примљен: 30. септембра 2025.

Received: September 30, 2025

Прихваћен за објављивање: 6. октобра 2025.

Accepted: October 6, 2025