

**DEMOKRATIJA U AI ERI:
Snaga mladih
za otpornije društvo**

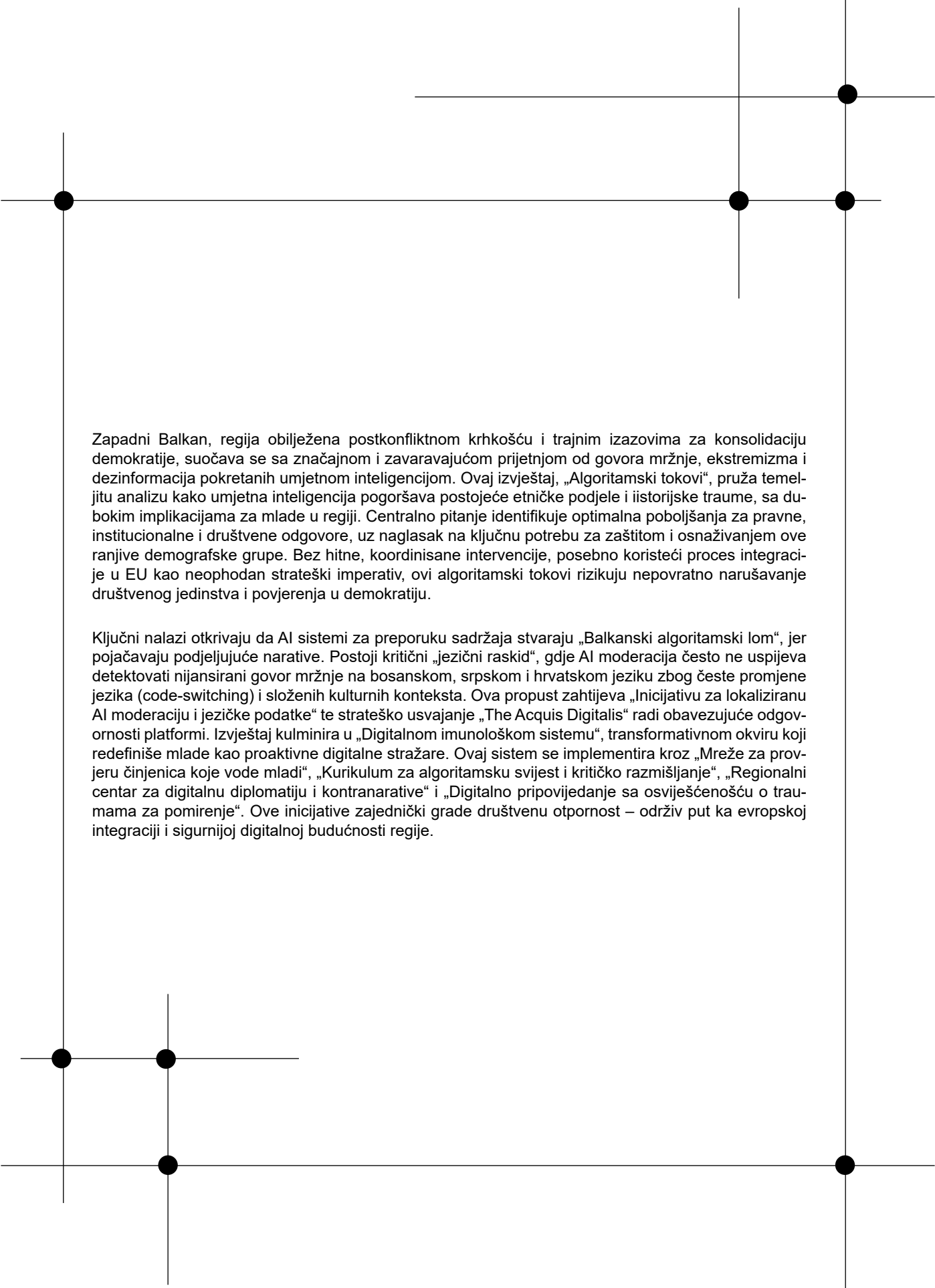


**Demokratija u AI eri:
Snaga mladih za otpornije društvo**

Autor:

Borislav Vukojević,
docent na Fakultetu političkih nauka Univerziteta u Banjoj Luci¹

¹ borislav.vukojevic@fpn.unibl.org



Zapadni Balkan, regija obilježena postkonfliktnom krhkošću i trajnim izazovima za konsolidaciju demokratije, suočava se sa značajnom i zavaravajućom prijetnjom od govora mržnje, ekstremizma i dezinformacija pokretanih umjetnom inteligencijom. Ovaj izvještaj, „Algoritamski tokovi“, pruža temeljitu analizu kako umjetna inteligencija pogoršava postojeće etničke podjele i iistorijske traume, sa dubokim implikacijama za mlade u regiji. Centralno pitanje identifikuje optimalna poboljšanja za pravne, institucionalne i društvene odgovore, uz naglasak na ključnu potrebu za zaštitom i osnaživanjem ove ranjive demografske grupe. Bez hitne, koordinisane intervencije, posebno koristeći proces integracije u EU kao neophodan strateški imperativ, ovi algoritamski tokovi rizikuju nepovratno narušavanje društvenog jedinstva i povjerenja u demokratiju.

Ključni nalazi otkrivaju da AI sistemi za preporuku sadržaja stvaraju „Balkanski algoritamski lom“, jer pojačavaju podjeljujuće narative. Postoji kritični „jezični raskid“, gdje AI moderacija često ne uspijeva detektovati nijansirani govor mržnje na bosanskom, srpskom i hrvatskom jeziku zbog česte promjene jezika (code-switching) i složenih kulturnih konteksta. Ova propust zahtijeva „Inicijativu za lokaliziranu AI moderaciju i jezičke podatke“ te strateško usvajanje „The Acquis Digitalis“ radi obavezujuće odgovornosti platformi. Izvještaj kulminira u „Digitalnom imunološkom sistemu“, transformativnom okviru koji redefiniše mlade kao proaktivne digitalne stražare. Ovaj sistem se implementira kroz „Mreže za provjeru činjenica koje vode mladi“, „Kurikulum za algoritamsku svijest i kritičko razmišljanje“, „Regionalni centar za digitalnu diplomatiju i kontranarative“ i „Digitalno pripovijedanje sa osviješćenošću o traumama za pomirenje“. Ove inicijative zajednički grade društvenu otpornost – održiv put ka evropskoj integraciji i sigurnijoj digitalnoj budućnosti regije.

CONTENTS

Uvod:	6
Navigacija algoritamskim tokovima	6
1. Digitalno žarište: AI i Zapadni Balkan	6
2. Pojačana krhkost BiH: Postkonfliktne podjele u digitalnom dobu	6
Metodologija: Mapiranje digitalnih prostora	7
2.1 Dizajn istraživanja i prioriteta podataka	7
2.2 Ograničenja istraživanja i lingvističke kompleksnosti	7
Regulatorni okviri: Mozaik digitalnog upravljanja	9
3.1 Fragmentirani pravni pejzaž BiH i regionalne razlike	9
3.2 EU približavanje: „Neizbježni“ pritisak za digitalne standarde	10
4.1 Mehanika algoritamske amplifikacije	11
4.2 Algoritamska pristrasnost: iskorištavanje postkonfliktne traume	11
4.3 Specifični rizici za BiH i erozija društva	12
AI-generisane dezinformacije i manipulacije: prolazna priroda istine	14
5.1 Uspon sintetičkih medija i automatiziranih kampanja	14
5.2 Ranljivosti: istorijski narativi, izbori i propaganda	15
5.3 Društvena erozija: povjerenje mladih i prekogranične kampanje	16
Prakse i odgovori tehnoloških kompanija i regulatornih tijela:	
Lingvistički razdor	17
6.1 Odgovornost tehnoloških kompanija i regionalna neosjetljivost	17
6.2 Nacionalni regulatorni kapacitet: Navigacija digitalnim Divljim Zapadom	18
Zaključak	19
Preporuke za politike i mehanizme odgovornosti	21
8.1 Regionalne strategije saradnje za Zapadni Balkan	21
8.2 Uključivanje mladih i programi razvoja vještina	23
8.3 Uključivanje više aktera i razvoj kapaciteta	24
Strateški plan: Faza implementacije i izgradnja otpornosti	25
9.1 Strateški plan: Koraci vlade (državni i entitetski nivo) za BiH	25
9.2 Strateški plan: Koraci civilnog društva i medija za BiH	26
9.3 Strateški plan: Regionalni koordinacioni mehanizmi za Zapadni Balkan	27
9.5 Dugoročne mjere za društvenu otpornost (fokus na mlade)	29
9.6 Sintetizacija nalaza i prioritizacija akcija	30
9.7 Mladi, saradnja i „neizbježan“ put naprijed	32
Dodatak A: Rječnik pojmova	33
A.1 Definicija ključnih pojmova	33
Reference	35

Uvod: Navigacija algoritamskim tokovima

1. Digitalno žarište: AI i Zapadni Balkan

Digitalno doba, obilježeno sveprisutnim usponom umjetne inteligencije, temeljno mijenja globalne informacijske sisteme. Ova tehnološka evolucija donosi napredak, ali takođe posjeduje veliki kapacitet da intenzivira društvene ranjivosti, naročito u regijama koje prolaze kroz dugotrajne postkonfliktnu okolnosti. Zapadni Balkan, složeno područje obilježeno istorijskim podjelama i razvijajućim demokratskim okvirima, služi kao ključna studija slučaja. Ovdje se transformativna moć AI presijeca sa postojećim krhkostima, stvarajući „digitalno žarište“. Sveprisutni uticaj AI, koji djeluje kroz sofisticirane mehanizme pojačavanja, predstavlja značajnu prijetnju demokratskim procesima i društvenom jedinstvu podstičući govor mržnje, ekstremizam i dezinformacije.

Bosna i Hercegovina (BiH) je primjer ove široke digitalne izloženosti, sa visokim internet penetracijom od 96% i korištenjem društvenih mreža koje doseže 78%. Samo Facebook koristi 71% internet korisnika. Ova široka online angažiranost, u kombinaciji sa složenim društveno-istorijskim kontekstom regije, pruža plodno tlo za AI da pojačava podjeljujuće narative, bilo namjerno ili ne. Ovaj izvještaj istražuje kako AI, kroz svoje složene i često netransparentne algoritme, intenzivira ove štete. Centralno pitanje koje vodi ovaj izvještaj glasi: Kako pravni, institucionalni i društveni odgovori u Bosni i Hercegovini i na Zapadnom Balkanu mogu biti optimalno unaprijeđeni kako bi se ublažio značajan uticaj govora mržnje, ekstremizma i dezinformacija pokretanih AI-jem, sa namjernim naglaskom na zaštitu i osnaživanje mladih? Ovo pitanje ističe imperativ za hitnu i koordiniranu akciju unutar ovog specifičnog geopolitičkog i društveno-istorijskog okruženja.

2. Pojačana krhkost BiH: Postkonfliktnu podjele u digitalnom dobu

Bosna i Hercegovina je zemlja oblikovana postkonfliktnom traumom i dubokim etničkim podjelama. Ova složena ustavna struktura, direktan rezultat neriješenih istorijskih pritužbi, stvara plodno tlo za štete uzrokovane AI-jem. Algoritmi, često optimizovani za angažman korisnika, nenamjerno pojačavaju podjeljujuću retoriku duž etničkih linija, tj. pretvaraju digitalne platforme u arene „Digitalnih ratova sjećanja“.

Ranjivost regije na sajber-informacijski rat i kampanje dezinformacija proizlazi iz iskorištavanja ovih postojećih društvenih pukotina. Ova dinamika je naročito vidljiva kroz administrativne podjele BiH: Republika Srpska, Federacija BiH i Brčko Distrikt. U ovim područjima mladi su posebno ciljane grupe, izloženi narativima koji pojačavaju etničku izolaciju i normalizuju ekstremistički diskurs. „Balkanski algoritamski lom“ predstavlja svakodnevnu stvarnost mladih u Sarajevu koji navigiraju „Digitalnom agorom Sarajeva“, onih u Banjoj Luci koji doživljavaju „Digitalnu komoru odjeka Banje Luke“ i pojedinaca u Brčko Distriktu koji se suočavaju sa „Digitalnim raskršćem Brčko Distrikta“. Ovi digitalni prostori, iako nude povezanost, istovremeno produbljuju postojeće podjele.²

Ovaj problem nadilazi granice BiH. Zapadni Balkan funkcioniše kao međusobno povezana informacijska sredina. Kampanje dezinformacija koje potiču iz jedne zemlje brzo se šire kroz Srbiju, Crnu Goru, Sjevernu Makedoniju, Albaniju i Kosovo*, koristeći zajedničke jezičke korijene i isprepletene istorijske narative. Ova

² "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.2 "Overview of the morning session discussion – addressing the first set of questions," stranica 27.

regionalna povezanost podiže nacionalne sigurnosne zabrinutosti na nivo regionalnog sigurnosnog imperativa. Efikasni odgovori zahtijevaju koordinirane, prekogranične pristupe. Izolovani nacionalni naponi nisu dovoljni protiv široko rasprostranjenog toka štetnog sadržaja pojačanog AI-jem. Javni diskurs, društvena kohezija i početni demokratski procesi širom regije suočavaju se sa značajnom prijetnjom ako ovi algoritamski tokovi ostanu nekontrolisani.³

Metodologija: Mapiranje digitalnih prostora

2.1 Dizajn istraživanja i prioriteti podataka

Dizajn istraživanja za „Algoritamske tokove“ koristi pristup desk istraživanja, metodologiju koju nameće ograničena dostupnost primarnih podataka i direktna transparentnost algoritama na Zapadnom Balkanu. Ovaj pristup sintetizira postojeću naučnu literaturu i zvanične izvještaje međunarodnih organizacija, uključujući United Nations, European Union i Organization for Security and Co-operation in Europe. Takođe uključuje analize uglednih regionalnih i međunarodnih *think tankova* i organizacija civilnog društva. Kriterijumi selekcije prioritet daju dokumentima koji nude podatke specifične za BiH i šira regionalna istraživanja Zapadnog Balkana, što omogućava da kontekstualne nijanse postkonfliktnih društava, etničke osjetljivosti i tekućih demokratskih tranzicija informišu analizu.

Metodološki okvir se oslanja na iterativni proces sinteze dokaza. To uključuje sistematsko identifikovanje, prikupljanje i kritičku evaluaciju istraživanja koja razjašnjavaju širok, ali često prikriven uticaj AI na informacijske sisteme. Prioritet imaju izvori koji detaljno prikazuju stope penetracije interneta, obrasce korištenja društvenih mreža među mladima i dokumentovane slučajeve online govora mržnje ili kampanja dezinformacija unutar BiH i širom Zapadnog Balkana. Ovaj okvir gradi sveobuhvatnu sliku uticaja AI, čak i kada su direktni empirijski podaci o specifičnoj primjeni AI u generisanju ili pojačavanju štetnog sadržaja teško dostupni. Pristup prepoznaje da, iako je direktna transparentnost algoritama velikih tehnoloških platformi često nedostižna, snažno razumijevanje može proizaći iz posmatranih efekata i dokumentovanih trendova u javnom diskursu.

Prioritetizacija podataka se proširuje na studije koje se bave interakcijom tehnologije i specifičnog društveno-istorijskog konteksta regije. To uključuje analize kako AI algoritmi mogu nenamjerno iskorištavati istorijske pritužbe ili etničke podjele, što doprinosi onome što ovaj izvještaj naziva „Balkanskim algoritamskim lomom“. Esencijalna potreba za lokalizovanim podacima vodi ovaj princip, osiguravajući da naknadna analiza ne počiva na generičkim pretpostavkama, već na provjerljivim, iako često fragmentiranim, dokazima specifičnim za Zapadni Balkan. Ovaj pažljiv proces selekcije osigurava da je svaka tvrdnja u ovom izvještaju potkrijepljena rigoroznim pregledom dostupne dokumentacije, jer stvara kredibilnu osnovu za preporuke politike koje slijede.

2.2 Ograničenja istraživanja i lingvističke kompleksnosti

Pristup desk istraživanja, iako ključan za navigaciju ograničenom dostupnošću primarnih podataka na Zapadnom Balkanu, inherentno donosi ograničenja. Značajan izazov proizlazi iz sveprisutnog nedostatka transparentnosti algoritama velikih tehnoloških kompanija. Ova neprozirnost sprječava detaljno razumijevanje kako AI alati funkcionišu u regiji, posebno u vezi sa sistemima preporuke sadržaja i moderacije. Javno dostupni podaci o preciznim primjenama AI alata ili njihovom uticaju na lokalne informacijske sisteme

3 Prema Rezoluciji 1244

su oskudni. Ograničena veličina tržišta Zapadnog Balkana nudi minimalan podsticaj velikim platformama da značajno ulažu u lokalizovano prikupljanje podataka ili inicijative transparentnosti. Ovo stvara „Pustinju moderacije“, gdje je sveobuhvatni nadzor otežan, a prolazna priroda online šteta često nadmašuje detekciju.⁴⁵

Lokalizovana istraživanja o uticaju AI na Zapadnom Balkanu su takođe rijetka. Dok globalne studije adresiraju algoritamsku pristrasnost i dezinformacije, specifične analize koje detaljno prikazuju njihovu manifestaciju unutar jedinstvenog društveno-istorijskog konteksta BiH i njenih susjeda su rijetke. Ovo zahtijeva pažljivu interpretaciju šireg nalaza, prilagođavajući ih specifičnim etničkim, religijskim i političkim osjetljivostima regije. Nedostatak robusnih, lokaliziranih skupova podataka za treniranje i evaluaciju AI dodatno pogoršava ovo ograničenje, otežavajući razvoj rješenja svjesnih konteksta.

Uključivanje jezičnih aspekata za sadržaj na bosanskom, srpskom i hrvatskom jeziku (B/S/H) predstavlja značajan izazov. Ovi jezici se često kategorizuju kao „jezici s niskim resursima“ u razvoju AI, što znači da oskudica označenih podataka otežava efikasnost automatizovanih sistema za moderaciju sadržaja. Ova jezična složenost očituje se na nekoliko načina:⁶

Morfološka složenost i dijalekatska raznolikost: B/S/H jezici posjeduju bogatu gramatiku, regionalne dijalekte i varijacije u pismu (srpski koristi i ćirilicu i latinicu). Ove nijanse komplikuju tokove obrade prirodnog jezika (NLP), otežavaju da AI precizno interpretira značenje i namjeru.⁷⁸

Miješanje pisama i neformalni jezik: Online komentari često sadrže miješanje između latiničnog i ćirilichnog pisma, uz široku upotrebu slenga, sarkazma i idiomatskih izraza. Ova neformalna jezična fluidnost otežava automatsko prepoznavanje štetnog sadržaja, jer naizgled bezazlena fraza može nositi pogrdno ili podijeljeno značenje u zavisnosti od konteksta ili pisma.⁹

Kulturne nijanse i kontekstualna dvosmislenost: Toksičnost na Zapadnom Balkanu često zavisi od konteksta, duboko je pod uticajem istorijskih tenzija, političke polarizacije i specifičnih kulturnih osjetljivosti. AI modeli imaju poteškoće u interpretaciji ovih suptilnosti bez opsežnih, lokalizovanih podataka za obuku i ljudske ekspertize. Na primjer, komentar koji se odnosi na „bolesnu osobu“ može proći neopaženo kod zero-shot AI modela, ali biti ispravno označen kao toksičan kada se uzme u obzir kontekst da se radi o srpskom političaru u teškoćama. Suprotno tome, bezazleni komentari mogu biti prekomjerno označeni ako kontekst nije pravilno shvaćen od strane modela s kontekstualnim proširenjem.¹⁰¹¹¹²

Ove jezične nijanse direktno doprinose „Linguistic Seam-Rip“ fenomenu, kritičnoj praznini u provođenju pravila gdje AI moderacija sadržaja, pa čak i ljudska revizija bez specijalizirane obuke, ne uspijevaju otkriti suptilne, kulturno ukorijenjene govore mržnje. Istraga Balkan Investigative Reporting Network (BIRN) iz 2021. godine otkrila je da „gotovo pola toksičnih jezičnih objava prijavljenih na balkanskim jezicima ostaje online, čak i nakon što su Facebook i Twitter potvrdili da sadržaj krši njihova pravila.“ Ovaj sistemski neuspjeh ostavlja mlade posebno ranjivim na štete koje AI pojačava, a koje se čine bezazlenim generičkim filterima. Posljedično, interpretacija nalaza u ovom izvještaju zahtijeva opreznu generalizaciju, uz priznavanje ovih inherentnih ograničenja, dok se teži najvišem mogućem stepenu preciznosti s obzirom na dostupne dokaze. Ova metodološka transparentnost postavlja realna očekivanja u pogledu opsega izvještaja i naglašava neosporivu potrebu za ciljanom, lokalizovanom istraživačkom i razvojnom aktivnošću.¹³

4 World Economic Forum. (2025). The Global Risks Report 2025. Insight Report, 20th Edition. World Economic Forum, Geneva, Switzerland. Stranica 36, Sekcija 1.5. https://reports.weforum.org/docs/WEF_Global_Risks_Report_2025.pdf

5 “Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective,” European Western Balkans, 4. april 2025. <https://europeanwesternbalkans.com/2025/04/04/why-the-digital-services-act-is-needed-in-the-western-balkans-an-institutional-and-market-perspective/>

6 Amel Muminovic and Amela Kadric Muminovic, “Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages,” arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

7 Ibid., Sekcija II-C, Stav 2.

8 Ibid., Sekcija II-C, Stav 2.

9 Ibid., Sekcija I-A, Stav 6.

10 Ibid., Sekcija I-A, Stav 6.

11 Ibid., Sekcija V-A Ključni nalazi.

12 Ibid., Sekcija V-B Practical Implications.

13 Ibid., Sekcija I-A, Stav 4.

Regulatorni okviri: Mozaik digitalnog upravljanja

3.1 Fragmentirani pravni pejzaž BiH i regionalne razlike

Pravni i politički okviri u Bosni i Hercegovini (BiH) su fragmentirani, što odražava složenu ustavnu strukturu države. Ova složena mreža zakonodavstva na državnom i entitetskom nivou često se pokazuje nedovoljnom za rješavanje nijansiranih izazova digitalnog upravljanja, naročito u vezi sa govorom mržnje i dezinformacijama pokretanim AI tehnologijama. BiH nema specifične regulatorne okvire za AI, ostavljajući pitanja poput govora mržnje, polarizacije i međunarodne netrpeljivosti na društvenim mrežama uglavnom nereguliranim ciljnim zakonima. Iako razvojna strategija za period 2021–2027, usvojena od strane Parlamenta Federacije Bosne i Hercegovine, prepoznaje strateški značaj AI, ne nudi konkretne korake ka uspostavljanju pravnih okvira za razvoj i primjenu AI. Ovaj regulatorni vakuum stvara „Digitalni jaz upravljanja“, ostavljajući mehanizme za regulaciju društvenog uticaja AI nedovoljno razvijenim.¹⁴¹⁵

Regionalna komparacija regulatornih pristupa u zemljama Zapadnog Balkana otkriva nedovoljno razvijen mehanizam za zakonodavstvo o sajber sigurnosti. Primjena zakona je neujednačena zbog ograničenih stručnih kapaciteta vlasti i politizacije. Ova fragmentacija digitalnog upravljanja, koja prati širu političku arhitekturu BiH, značajno otežava jedinstven i efikasan odgovor na štete izazvane AI tehnologijama. Agencija za komunikacije (CRA) u BiH priznaje odsustvo nacionalnih AI strategija i predložila je da djeluje kao „oblikovatelj politike“ pokretanjem nacionalnih dijaloga i predlaganjem planova djelovanja.¹⁶¹⁷

Ovaj fragmentirani sistem stvara ranjivosti na štete izazvane AI, naročito u regiji obilježenoj istorijskim podjelama. Velike tehnološke kompanije koriste svoju digitalnu moć za pokretanje novih oblika orkestriranih šteta i javnih dezinformacija, a istovremeno pokazuju minimalan interes za rješavanje raširenog govora mržnje usmjerenog prema pojedincima, etničkim grupama i kritičkim misliocima u Bosni i Hercegovini. Ekonomski izazovi u regiji dodatno utiču na digitalnu infrastrukturu i efikasnost programa digitalne pismenosti, pogoršavajući ove probleme. Nedostatak specifičnih pravnih odredbi za AI i tehnologije mašinskog učenja izaziva zabrinutost u pogledu etičkih implikacija povezanim s širokom primjenom AI sistema u različitim sektorima.¹⁸¹⁹

Posljedice ove regulatorne stagnacije su jasne. Bez snažnih pravnih okvira, „Balkanski algoritamski raskol“ opstaje jer AI sistemi iskorištavaju postojeće etničke i političke podjele bez efikasnog nadzora. To ostavlja mlade, najviše digitalno uronjene demografske grupe, posebno podložnima ekstremističkim narativima koji se šire kroz „Digitalnu željeznu zavjesu“. Nadalje, odsustvo sveobuhvatnog zakonodavstva otežava

- 14 Bojana Kostić and Caroline Sindors, "Responsible Artificial Intelligence: An overview of human rights' challenges of Artificial Intelligence and media literacy perspectives in the context of Bosnia and Herzegovina" (Council of Europe, juni 2022), pristupljeno putem <https://rm.coe.int/mil-study-3-artificial-intelligence-final-2759-3738-4198-2/1680a7cdd9>, Sekcija I, Uvod.
- 15 Namanja Sladaković i Milica Novaković, "Bosnia and Herzegovina," u IBA Alternative and New Law Business Structures Committee, juli 2024, pristupljeno putem <https://www.ibanet.org/medias/anlbs-ai-working-group-report-july-2024-4-bosnia-herzegovina.pdf?context=bWFzdGVyFB1YmXpY2F0aW9uUmVwb3J0c3w1NTQ3OHxhcHBsaWNhdGlvbi9wZGZ8YURJeUwyaGlaQzg1TVRN-NE1qQTJOREE0TnpNMEwyRnViR0p6TFdGcExYZHJzbXRwYm1jdFozSnZkWEF0Y21Wd2IzSjBMV3AxYkhreD1qQXIOQzAwTF-dKdmMyNXBZUzFvWihKNlpXZHkzbWx1WVM1d1pHWXw4M2UzMmYzYTRIOGEwZWJiYTk1ZTkYThkN2MxNGI1NmU4ZDF-hODUxOGY5ODQ0OTM1NWlZYTc1MmFIOdkwZGZl>
- 16 "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.2, stranica 25.
- 17 Communications Regulatory Agency (CRA) Bosnia and Herzegovina, "GSR-25 Contribution: Best Practice Guidelines," predstavljeno na GSR-25 Konsultacijama, pristupljeno putem https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf
- 18 Bojana Kostić and Caroline Sindors, "Responsible Artificial Intelligence: An overview of human rights' challenges of Artificial Intelligence and media literacy perspectives in the context of Bosnia and Herzegovina" (Council of Europe, juni 2022), pristupljeno putem <https://rm.coe.int/mil-study-3-artificial-intelligence-final-2759-3738-4198-2/1680a7cdd9>, Sekcija I, Uvod.
- 19 Namanja Sladaković i Milica Novaković, "Bosnia and Herzegovina," u IBA Alternative and New Law Business Structures Committee, juli 2024, pristupljeno putem <https://www.ibanet.org/medias/anlbs-ai-working-group-report-july-2024-4-bosnia-herzegovina.pdf?context=bWFzdGVyFB1YmXpY2F0aW9uUmVwb3J0c3w1NTQ3OHxhcHBsaWNhdGlvbi9wZGZ8YURJeUwyaGlaQzg1TVRN-NE1qQTJOREE0TnpNMEwyRnViR0p6TFdGcExYZHJzbXRwYm1jdFozSnZkWEF0Y21Wd2IzSjBMV3AxYkhreD1qQXIOQzAwTF-dKdmMyNXBZUzFvWihKNlpXZHkzbWx1WVM1d1pHWXw4M2UzMmYzYTRIOGEwZWJiYTk1ZTkYThkN2MxNGI1NmU4ZDF-hODUxOGY5ODQ0OTM1NWlZYTc1MmFIOdkwZGZl>

napore u rješavanju „Linguistic Seam-Rip“ problema, gdje AI moderacija ima poteškoće u detekciji nijansiranih govora mržnje na bosanskom, srpskom i hrvatskom jeziku. Ovo nije samo tehnički propust, već sistemski neuspjeh ukorijenjen u „Pustinji moderacije“, posljedici nedovoljnog ulaganja tehnoloških kompanija u regije s nižim prihodima. Posljedično, pravni i institucionalni okviri u BiH i širem Zapadnom Balkanu trenutno nisu adekvatno opremljeni da se suprotstave rastućim prijetnjama koje donosi štetan sadržaj pokretan AI, što zahtijeva temeljnu promjenu pristupa.²⁰²¹

3.2 EU približavanje: „Neizbježni“ pritisak za digitalne standarde

Proces približavanja Evropskoj uniji (EU) pokreće digitalne standarde širom Zapadnog Balkana. Ova integracija usklađuje nacionalne regulatorne okvire sa opsežnim *acquisom* EU, uključujući Digital Services Act (DSA) i predstojeći AI Act. Ovo usklađivanje predstavlja ključnu priliku. Omogućava regiji da prevaziđe domaću političku fragmentaciju i adresira izazove štetnog sadržaja pokretanog AI.

Procjena trenutnih regulatornih okvira u zemljama Zapadnog Balkana otkriva značajne zakonske praznine i neujednačenu primjenu u poređenju s DSA. Na primjer, Albanija pokriva samo 47% svog zakonodavstva vezanog za DSA. Srbija i Bosna i Hercegovina pokazuju slične nedostatke. Kosovo* i Sjeverna Makedonija, uprkos većoj konvergenciji, nemaju potpune zaštite koje nudi DSA. Ovaj regulatorni vakuum ostavlja građane ranjivim na nezakoniti i štetni online sadržaj. Platforme, percipirajući ove tržišta kao manje značajna, dodjeljuju minimalne resurse za moderaciju sadržaja.²²²³²⁴

Ove regulatorne slabosti posebno ugrožavaju zaštitu mladih. Odsustvo sveobuhvatnih pravnih okvira i adekvatno osnaženih institucija znači da osnovna prava na internetu ostaju nedovoljno zaštićena. Ova situacija pojačava efekat „Digitalne željezne zavjese“, gdje su mladi disproportionalno izloženi štetnom sadržaju zbog ograničenog nadzora i nedostatka odgovornosti platformi. Primjena GDPR-a, ključnog stuba digitalnih prava u EU, predstavlja kontinuirani izazov u Zapadnom Balkanu, sa varijabilnim kapacitetima za implementaciju i provođenje u javnim institucijama. Proces proširenja EU pruža stratešku priliku zemljama Zapadnog Balkana da usvoje DSA i Digital Markets Act (DMA) kroz približavanje *acquisu*.

Ovaj pristup nudi provjereni okvir za digitalno upravljanje, jer negira potrebu za razvojem novih propisa od nule. Usvajanje ovih okvira ključno je za pripremu zakonodavstva i institucija regije za prevenciju i odgovor na rastuću širenje ciljane dezinformacije i zloupotrebe AI za širenje lažnih informacija. Nadalje, usklađivanje sa EU standardima, uključujući NIS Directive i Cybersecurity Act, donosi dodatne izazove u implementaciji. Ekonomije Zapadnog Balkana pokazuju napredak, ali potpuna usklađenost je proces u toku zbog razlika u regulatornim resursima, kapacitetima za provođenje i sektorskoj usklađenosti. EU-ova ulaganja u ovu digitalnu usklađenost nadilaze „altruizam“. Ona predstavljaju geopolitički i sigurnosni imperativ, prisiljavajući Zapadni Balkan na usvajanje propisa usklađenih s DSA. Ovo jača napore u borbi protiv dezinformacija i lažnih vijesti u neposrednom susjedstvu EU, regiji ranjivoj na strane uticaje. Regionalni savjet za saradnju (RCC) aktivno podržava ovu harmonizaciju, jer uspostavlja radne grupe za snažan okvir upravljanja podacima koji eksplicitno uključuje DSA i DMA.

Kolektivni pristup, vođen *Acquis Digitalis*, nudi najefikasniji put za prevazilaženje domaće političke fragmentacije. On podstiče sigurniju i otporniju digitalnu budućnost za cijelu regiju. Proces integracije u EU tako funkcioniše kao snažan katalizator, pokreće temeljnu promjenu u digitalnom upravljanju neophodnu za regionalnu stabilnost i zaštitu mladih.

20 "Responsible Artificial Intelligence: An overview of human rights' challenges of Artificial Intelligence and media literacy perspectives in the context of Bosnia and Herzegovina," Council of Europe, juni 2022, stranica 4.

21 "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective," European Western Balkans, 4. april 2025. <https://europeanwesternbalkans.com/2025/04/04/why-the-digital-services-act-is-needed-in-the-western-balkans-an-institutional-and-market-perspective/>

22 "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective", European Western Balkans, 4. april 2025.

23 Isto.

24 Isto.

Algoritamska pristrasnost i širenje sadržaja: Balkanski algoritamski raskol

4.1 Mehanika algoritamske amplifikacije

Sveprisutni uticaj umjetne inteligencije na informacijske sisteme je neosporan. AI sistemi, naročito mehanizmi za preporuku sadržaja, funkcionišu na principima dizajniranim da maksimiziraju angažman korisnika. Ovi sistemi analiziraju velike količine korisničkih podataka, uključujući prethodne interakcije, navike gledanja i izražene preference, kako bi isporučili sadržaj koji je najvjerovatnije da zadrži pažnju. Ova optimizacija za angažman često nenamjerno daje prioritet emocionalno nabijenom ili senzacionalističkom sadržaju, jer takav materijal često generiše veće nivoe interakcije. Posljedično, AI ima dubok uticaj na širenje ekstremističkih narativa, govora mržnje i dezinformacija. Algoritmi, vođeni težnjom za angažmanom, mogu postati kanali za brzo širenje podijeljenog materijala, bez obzira na njegovu istinitost ili društveni uticaj. Ovaj mehanizam stvara samopodržavajući ciklus u kojem izloženost određenim narativima povećava angažman, što zauzvrat potiče algoritam da isporuči još istog sadržaja, te tako produbljuje postojeće pristrasnosti i formira eho-komore.

Na Zapadnom Balkanu, ova algoritamska dinamika se manifestuje se kao „Balkanski algoritamski raskol.“ Koncept opisuje kako inherentni mehanizmi AI iskorištavaju jedinstveni društveno-istorijski kontekst regiona, gdje su postojeće etničke i vjerske tenzije posebno ranjive na manipulaciju. Algoritmi, bez dubokog kulturnog i istorijskog razumijevanja, procesuiraju emocionalno rezonantni sadržaj vezan za istorijske nepravde ili nacionalističke narative kao puko „angažirajući“, umjesto da prepoznaju njegov potencijal za štetu. Ovo fundamentalno nerazumijevanje pretvara digitalni pejzaž u bojno polje narativa. Sveprisutni uticaj algoritama preoblikuje javni diskurs, suptilno usmjeravajući korisnike prema sadržaju koji potvrđuje prethodno postojeće pristrasnosti i potencijalno pojačava međunacionalnu netrpeljivost. Ovaj proces je ključan za razumijevanje kako AI ne djeluje kao neutralni alat, već kao aktivni, često nenamjerni, agent u produbljivanju društvenih podjela.

Mehanika ove amplifikacije je složena. Mehanizmi za preporuku sadržaja, na primjer, oslanjaju se na kolaborativno filtriranje i metrike sličnosti. Ako korisnik interaguje sa sadržajem koji izražava određeni nacionalistički stav, algoritam identifikuje druge korisnike koji su interagovali sa sličnim sadržajem i preporučuje dodatni materijal iz te ideološke grupe. Ovo stvara „Digitalni željezni zastor“ za mlade, ograničavajući njihovu izloženost različitim perspektivama i čineći ih podložnim „kognitivnom zarobljavanju“ ekstremističkim narativima. Rezultat je okruženje u kojem politički ili etnički nabijen sadržaj, često prožet istorijskim aluzijama, dobija disproporcionalnu vidljivost. Ova algoritamska indiferentnost prema postkonfliktnim nijansama predstavlja kritičnu sistemsku manu. Platforme daju prioritet kvantitativnim metrikama angažmana nad kvalitativnim kontekstualnim razumijevanjem, pa dopuštaju da se podijeljeni sadržaj nekontrolisano širi. Ova „Algoritamska indiferentnost i amplifikacija“ nije samo teorijska briga; to je dokazani faktor koji doprinosi eroziji društvene kohezije i povjerenja u demokratiju širom Zapadnog Balkana.

4.2 Algoritamska pristrasnost: iskorištavanje postkonfliktne traume

Uticaj AI sistema seže dalje od same amplifikacije sadržaja; on iskorištava duboko ukorijenjene istorijske nepravde i neriješene traume koje karakteriziraju postkonfliktna društva poput Bosne i Hercegovine. Ova algoritamska pristrasnost, često emergentna osobina sistema optimizovanih za angažman, daje prioritet i podiže sadržaj koji rezonuje sa nacionalističkim narativima, međunacionalnom netrpeljivošću i istorijskim revizionizmom. Rezultat je digitalno okruženje u kojem se prošlost ne samo pamti, već se aktivno koristi kao oružje.

Teorija algoritamske eksploatacije postkonfliktne traume tvrdi da AI, kroz svoj naizgled neutralni dizajn, nenamjerno postaje alat za produbljivanje društvene polarizacije. Algoritam dizajniran isključivo da maksimizira vrijeme korisnika na platformi, kada se primijeni u kontekstu obilježenom neadresiranom traumom, direktno amplifikira sadržaj koji se hvata za te sirove emocije. To stvara kontinuirani, algoritamski posredovani sukob oko istorijske interpretacije, sa mladima kao nenamjernim boricima.

Predmet Pilav protiv Bosne i Hercegovine (2016) služi kao dirljiv podsjetnik kako se sistemske pristrasnosti mogu ukorijeniti u temeljnim strukturama jedne države. Ovaj slučaj je pokazao da izborni sistem BiH diskriminiše na osnovu etničkog porijekla i mjesta prebivališta, jer ograničava ključne političke funkcije na određene „konstitutivne narode“ i isključujući druge. Ovaj pravni presedan ilustruje kako „geografske i etničke pristrasnosti pokazuju kako naizgled neutralna pravila mogu imati pristrasne ishode u praksi.“ AI, ako se ne upravlja pažljivo, mogao bi istovremeno pogoršati takve postojeće društvene podjele, dajući tehnološki izgled objektivnosti duboko nepravednim praksama. Rizik ostaje visok, naročito jer „još nisu zabilježeni dobro dokumentovani slučajevi AI-pokrenute pristrasnosti u izbornim procesima“ na Zapadnom Balkanu, što ukazuje na rani prijetnju koja zahtijeva proaktivno ublažavanje.

AI djeluje kao pojačivač za marginalizovane grupe, često nenamjerno povećavaju sadržaj koji je štetan čak i ako potiče iz tih zajednica. To perpetuira diskriminaciju i isključenost, često bez znanja korisnika, replicirajući postojeće rodne norme, rasne stereotipe i druge predrasude ugrađene u skupove podataka za treniranje. Takve sistemske pristrasnosti, ugrađene u AI, mogu pogoršati etničke i vjerske tenzije. Online narativi mržnje koje šire ekstremističke grupe pojačavaju etničke razlike i međunacionalnu polarizaciju, jer jačaju društvene podjele. Ovi narativi često crpe iz duboko ukorijenjenih regionalnih etnonacionalističkih istorijskih mitova, pretvarajući digitalne platforme u bojišta kolektivnog sjećanja.²⁵²⁶

Uticaj ove algoritamske pristrasnosti nije samo teorijski. Ona se manifestuje u uočljivim trendovima, naročito tokom perioda pojačane političke aktivnosti ili osjetljivih istorijskih godišnjica. Iako specifični kvantitativni podaci o direktnoj korelaciji AI sa porastom govora mržnje u BiH ostaju ograničeni zbog algoritamske neprozirnosti, obrazac povećanog online ekstremističkog sadržaja tokom izbornih ciklusa u širem Zapadnom Balkanu je dobro dokumentovan. Ova korelacija sugerise da algoritmi optimizovani za angažman igraju istovremenu ulogu u amplifikaciji takvih narativa.²⁷²⁸

Prisustvo takvih sistemskih pristrasnosti unutar AI, u kombinaciji sa istorijskim ranjivostima regiona, hrani „Digitalne ratove sjećanja.“ Ovi ratovi se vode na platformama gdje algoritmi, kroz svoje „Algoritamske slijepe tačke,“ nesvjesno naoružavaju kolektivno sjećanje, dajući prioritet sadržaju koji izaziva jake emocionalne reakcije povezane s istorijskim nepravdama. To stvara nevidljivo bojno polje gdje su mladi glavni, često nesvjesni borci, čija percepcija istorije i međunacionalnih odnosa suptilno oblikovana pristranim algoritamskim tokovima. Ova sveprisutna algoritamska pristrasnost zahtijeva proaktivan pristup upravljanju AI, koji uzima u obzir jedinstveni društveno-istorijski kontekst Zapadnog Balkana i daje prioritet zaštiti njegove omladine.²⁹³⁰

4.3 Specifični rizici za BiH i erozija društva

Složena ustavna struktura Bosne i Hercegovine (BiH) inherentno naglašava podijeljenu retoriku duž etničkih linija. Ova amplifikacija posebno cilja mlade širom Republike Srpske, Federacije BiH i Brčko Distrikta. Online narativi mržnje, koje šire ekstremističke grupe, pojačavaju etničke razlike i međunacionalnu polarizaciju, čime se jačaju društvene podjele. Ovi narativi često prizivaju regionalne etnonacionalističke istorijske mitove, pretvarajući digitalne platforme u arene „Digitalnih ratova sjećanja“ – borbi za kolektivno pamćenje.

25 Isto.

26 "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.5, stranica 34.

27 Isto.

28 Ibid., Sekcija 2.3.3, str. 29.

29 Isto.

30 Regional Cooperation Council Secretariat, "DRAFT REPORT ON THE ACTIVITIES OF THE REGIONAL COOPERATION COUNCIL SECRETARIAT For the period 01 October 2024 – 28 February 2025," <https://www.rcc.int/files/user/docs/a9b1d91ff97948db-2d1e8adf2d6a2662.pdf>, Sekcija A3: Digital Integration and Implementation of the Digital Agenda for Western Balkans, stranica 9.

Fragmentacija digitalnih prostora duž etničkih linija, naročito unutar složenih administrativnih podjela BiH, omogućava AI algoritmima da lako kreiraju i učvršćuju eho-komore. U Sarajevu, online okruženje, često nazvano „Digitalna agora Sarajeva“, odražava nacionalne političke i etničke tenzije. Tu mladi nailaze na algoritamski ispostavljen sadržaj koji učvršćuje postojeće pristrasnosti. Slično, „Digitalna eho-komora Banja Luke“ amplificira srpske nacionalističke narative i istorijski revizionizam, izoluje mlade unutar uskog ideološkog okvira. Kroz „Fragmentirane feedove Federacije“ i „Digitalne raskrsnice Brčko Distrikta“, logika AI vođena angažmanom pogoršava ove podjele, što ometa međunacionalni dijalog.

Ova AI-pokrenuta polarizacija predstavlja značajne rizike za javni diskurs, društvenu koheziju i demokratske procese u postkonfliktnim društvima. Široko rasprostranjeno širenje neprovjerenih informacija i govora mržnje, potpomognuto AI, aktivno podriva napore na pomirenju i stabilnosti. Hipotetički primjer ilustruje ovu eroziju:

Tokom nedavnog izbornog ciklusa u „Fiktivnoj opštini BiH“, lokalni omladinski aktivista primijetio je izražen porast online sadržaja koji demonizuje određenu etničku grupu. Analiza je ukazala da su algoritmi društvenih mreža, prioritizirajući angažman korisnika, nenamjerno amplificirali istorijske revizionističke narative. Sadržaj ranije ograničen na rubne platforme ušao je u glavne feedove mladih. To je stvorilo snažnu, iako prolaznu, eho-komoru ispunjenu podijeljenom retorikom, značajno utičući na lokalni diskurs. Lokalizovana amplifikacija istorijskih nepravdi putem AI-pokrenutog sadržaja dovela je do mjerljivog porasta međunacionalnih tenzija, što je dovelo do online uznemiravanja i izolovanim incidentima van mreže tokom izbornog perioda. Mladi, naročito oni u digitalno segregiranim online zajednicama, postali su nenamjerni učesnici „Digitalnog rata sjećanja“, čija su shvatanja istorije i međunacionalnih odnosa suptilno oblikovana pristrasnim algoritamskim tokovima.³¹

Ovaj hipotetički scenario ističe kako se „Balkanski algoritamski raskol“ manifestuje u specifičnim regionalnim dinamikama. On predstavlja neosporivu prijetnju integritetu multietničkog tkiva BiH i širim demokratskim aspiracijama Zapadnog Balkana. Sveprisutnost društvenih mreža, u kombinaciji sa niskim nivoima medijske pismenosti, stvara „Digitalni željezni zastor“ koji ograničava izloženost mladih različitim perspektivama, jer ih čini podložnim „kognitivnom zarobljavanju“ ekstremističkim narativima. Ovaj „Digitalni željezni zastor“ dodatno pogoršava problem jezika, gdje sistemi za moderaciju AI često ne prepoznaju nijansirani govor mržnje na bosanskom, srpskom i hrvatskom jeziku. To ostavlja mlade izložene suptilno toksičnim porukama.

Mladi, iako pozicionirani kao izrazito ranjivi, takođe predstavljaju predvodnike otpornosti društava. Osnaživanje mladih naprednom medijskom pismenošću, kritičkim mišljenjem i algoritamskom sviješću pretvara ih u aktivne digitalne stražare, čime se formira „Digitalni imuni sistem.“ Ovaj sistem funkcioniše kroz međusobno povezane omladinske mreže, obrazovanje vršnjaka i direktno uključivanje u kreiranje politika. Inicijative poput „Mreža za provjeru činjenica vođenih od strane mladih“ direktno se suprotstavljaju „Balkanskom algoritamskom raskolu“, što pruža decentralizovanu detekciju i odgovor. „Kurikulum algoritamske svijesti i kritičkog mišljenja“ oprema ih za dekonstrukciju algoritamskog uticaja, direktno izazivajući „Digitalni željezni zastor.“ Nadalje, „Trauma-informisano digitalno pripovijedanje za pomirenje“ njeguje empatiju i iscjeljenje, aktivno se suprotstavljajući „Digitalnim ratovima sjećanja“ koje potiče algoritamska pristrasnost. Ove proaktivne mjere, u kombinaciji sa strateškim korištenjem „The Acquis Digitalis“ za prisiljavanje tehnoloških kompanija na ulaganja u lokalizovane AI moderacijske i jezičke inicijative, predstavljaju ključnu branu protiv sveprisutnog digitalnog talasa. Ovaj holistički pristup štiti mlade i osigurava dugoročnu stabilnost demokratskih procesa na Zapadnom Balkanu.

31 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 4. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

AI-generisane dezinformacije i manipulacije: prolazna priroda istine³²

5.1 Uspon sintetičkih medija i automatiziranih kampanja³³³⁴³⁵

Deepfakeovi i sintetički mediji su AI-generisane slike, audio ili video sadržaji koji su realistični, ali lažni. Ovi alati manipulišu postojećim medijima ili stvaraju novi sadržaj, što čini autentičnost teškom za prepoznati. Implikacije ove tehnologije su široke i, u nekim kontekstima, lično razorne. Na primjer, istraživački izvještaji iz 2024. otkrili su Telegram mreže širom Balkana gdje je softver manipulisan AI-em „skidao“ slike žena. Ova zlonamjerna zloupotreba tehnologije dovela je do ucjena i javnog sramoćenja, što je ozbiljan etički prekršaj. Procjenjuje se da 96% svih *deepfakeova* prikazuje neodobrene intimne slike žena, što naglašava duboko rodno zasnovano nasilje koje ovi alati omogućavaju.³⁶³⁷

Osim individualne štete, *deepfakeovi* i sintetički mediji ugrožavaju javni diskurs i demokratske procese. Oni stvaraju lažne narative, imitiraju javne ličnosti ili fabrikuju događaje, čime manipulišu javnim mnijenjem. Iako javno dostupni izvještaji nemaju konkretne dokumentovane primjere *deepfakeova* korištenih za manipulaciju istorijskim narativima ili uticaj na izbore u Zapadnom Balkanu, potencijal za takvu zloupotrebu je očigledan. Izvještaj Atlantic Councila iz 2024. o ruskim aktivnostima informacione propagande (IIA) u Zapadnom Balkanu, posebno u Srbiji, bilježi širenje lažnog i manipulisanog multimedijalnog sadržaja, uključujući *deepfakeove*, unutar šire kampanje dezinformacija. Ovo ukazuje na razvijajući prijetnju u kojoj se generativne sposobnosti AI-a sve više integrišu u zlonamjerne operacije.

Automatizovane kampanje, uključujući bot mreže i koordinisano neautentično ponašanje, dopunjuju sintetičke medije osiguravajući njihovu široku distribuciju. Ove kampanje koriste AI za kreiranje i upravljanje brojnim lažnim nalogima koji masovno šire dezinformacije na društvenim mrežama, što pojačava njihov doseg i uticaj. Treći Izvještaj Evropske službe za vanjske poslove (EEAS) o prijetnjama (mart 2025) potvrđuje sve veću upotrebu AI u operacijama stranog manipulisanja informacijama i miješanja (FIMI). To uključuje kreiranje sadržaja, poput *deepfake* audio i video zapisa, te masovnu automatizovanu distribuciju putem bot mreža. Iako izvještaj ne navodi specifične primjere unutar Zapadnog Balkana, ističe Moldaviju kao značajno ciljanu zemlju u 2024, gdje su ruske FIMI operacije koristile AI-generisani sadržaj, uključujući video koji imitira glas predsjednice Maie Sandu, kako bi uticale na izbore i referendum o pristupanju EU. Ovo pokazuje sofisticirane taktike koje su već u upotrebi u širem evropskom susjedstvu.

Sintetički mediji i automatizirane kampanje temeljno mijenjaju informacijski sistem, jer stvaraju prolazan osjećaj stvarnosti. Ovaj fenomen dovodi u pitanje tradicionalne metode verifikacije i kritičke procjene. Kako AI-generisani sadržaj postaje sofisticiraniji i dostupniji, sposobnost razlikovanja istine od laži opada. Ovo narušavanje povjerenja u digitalnu autentičnost ima duboke implikacije za mlade, koji često nemaju vještine za snalaženje u složenim informacijskim okruženjima. Sveprisutnost ovih alata i taktika zahtijeva snažan, višeslojni odgovor za zaštitu demokratskih procesa i očuvanje integriteta javnog diskursa u Zapadnom Balkanu. Taj odgovor uključuje tehnološka rješenja za detekciju i sveobuhvatne programe medijske pismenosti koji opremaju građane, posebno mlade, kritičkim razmišljanjem potrebnim za prepoznavanje i

32 <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>, "The ethical question of electoral AI #3", International IDEA, 02. april 2025, Sekcija: "Pillar #2, AI Ethics and Human Rights", podsekcija: "While AI tools so far have been largely peripheral in the Western Balkans..."

33 <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>, "The ethical question of electoral AI #3", International IDEA, 02. april 2025, Sekcija: "Pillar #2, AI Ethics and Human Rights", podsekcija: "A related concern arose in Bosnia and Herzegovina..."

34 Isto.

35 Ibid., podsekcija: „Dok su AI alati do sada uglavnom bili periferni na Zapadnom Balkanu..."

36 Isto.

37 Radicalisation Awareness Network (RAN), (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 1. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

odolijevanje sofisticiranoj manipulaciji vođenoj AI-em.

5.2 Ranljivosti: istorijski narativi, izbori i propaganda³⁸

Zapadni Balkan je podložan ranjivostima koje AI-generisana dezinformacija iskorištava, uključujući manipulaciju istorijskim narativima, miješanje u izbore i etničku propagandu. Narativi prožeti nacionalističkim podtekstovima lako se ukorjenjuju u regiji obilježenoj neriješenim istorijskim nezadovoljstvima. AI može automatizirati velike kampanje dezinformacija.

AI chatbotovi generišu netačne informacije, poznate kao „halucinacije“, koje su utjecale na izbore za EU 2024. šireći neistine. Ovo narušava povjerenje i obespravljuje birače, posebno u marginaliziranim zajednicama. Rizik od AI-pokretnog miješanja u izbore predstavlja kritičan problem za mlade demokratije Zapadnog Balkana. Ova regija već se suočava sa sistemskim pristrasnostima u izbornim procesima, što potvrđuju slučajevi poput Pilav protiv Bosne i Hercegovine (2016). AI bi mogao pogoršati ove pristrasnosti, uzrokujući da naizgled neutralna pravila daju iskrivljene rezultate. Ovo predstavlja duboku prijetnju integritetu demokratskih institucija.

Ekstremističke grupe na Zapadnom Balkanu koriste globalne teme za regrutaciju. Te teme uključuju teorije zavjere vezane za vakcine i pandemiju, anti-gender i anti-feministički diskurs, anti-imigrantsku retoriku i podršku ruskoj invaziji na Ukrajinu. AI pojačava ove narative, stvarajući okruženje u kojem „Digitalni ratni front pamćenja“ postaje jasno vidljiv. Ovo bojno polje oko istorijske interpretacije pokreću algoritmi koji nenamjerno oružaju kolektivno pamćenje, dajući prioritet sadržaju koji izaziva snažne emocionalne reakcije povezane s istorijskim nezadovoljstvima. Mladi, kao primarni digitalni stanovnici, postaju nesvjesni sudionici ovih sukoba, njihova percepcija istorije i međunacionalnih odnosa suptilno je oblikovana pristrasnim algoritamskim tokovima.

Sljedeća hipotetička studija slučaja ilustrira ovu apstraktnu prijetnju:³⁹

Uoči osjetljive istorijske godišnjice u „Fiktivnoj opštini BiH“, pojavio se AI-generisani video. Prikazivao je izmišljenog istorijskog lika koji je držao zapaljivi govor, suptilno mijenjajući arhivske snimke da uključi nacionalističke simbole. Video, distribuiran putem automatiziranih naloga na društvenim mrežama, brzo je stekao popularnost među mladima. Uprkos brzom demantiju lokalnih provjerivača činjenica, početna široka distribucija deepfakea uzrokovala je nepovjerenje u zvanične istorijske izvještaje i tradicionalne medije. Incident je pokazao sposobnost AI-a da manipuliše kolektivnim pamćenjem, raspiruje etničke tenzije i potkopava napore na pomirenju u kratkom vremenskom periodu.⁴⁰

Ovaj hipotetički scenario ističe ranjivosti inherentne informacijskom okruženju regije. Pokazuje kako AI-generisana dezinformacija, naročito kada cilja istorijske narative, predstavlja direktnu prijetnju društvenoj koheziji i demokratskoj stabilnosti. Sveprisutnost AI halucinacija, u kombinaciji sa iskorištavanjem globalnih tema od strane ekstremističkih grupa, stvara složen izazov. Taj izazov zahtijeva snažan, višeslojni odgovor koji nadilazi reaktivno uklanjanje sadržaja. Potrebne su proaktivne strategije za jačanje otpornosti mladih na manipulativni potencijal AI-a, uključujući razvoj kritičkog mišljenja i svijesti o algoritmima, komponente „Digitalnog imunog sistema“.

38 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 1. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

39 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 4. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

40 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 5. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

5.3 Društvena erozija: povjerenje mladih i prekogranične kampanje

Prekogranične kampanje dezinformacija, sve sofisticiranije, često kombinuju sajber napade sa dezinformacijama kako bi ciljale etničke i političke podjele u više zemalja Zapadnog Balkana. Neprijateljski akteri iskorištavaju online prostore za produbljivanje društvenih podjela i smanjenje povjerenja u institucije. Povezano informacijsko okruženje Zapadnog Balkana, karakterizovano zajedničkim jezičkim korijenima i isprepletenim istorijskim narativima, omogućava brzo širenje ovih zlonamjernih uticaja, što pretvara nacionalne sigurnosne izazove u neosporni regionalni imperativ.⁴¹

Mladi, kao demografska grupa najviše angažovana u online okruženjima, disproportionalno su pogođeni ovom društvenom erozijom. Efekat „Digitalne željezne zavjese“ ilustruje kako izloženost pristrasnim informacijama ograničava pristup različitim perspektivama, što povećava ranjivost na kognitivno zarobljavanje i dodatno smanjujući povjerenje. Algoritamska segregacija, pokretana AI-em optimiziranim za angažman, stvara zatvorena digitalna okruženja gdje kritičko mišljenje atrofira, a podložnost ekstremističkim narativima raste. Niski nivoi medijske pismenosti širom Zapadnog Balkana, gdje zemlje konstantno zauzimaju donje pozicije na globalnim indeksima, dodatno povećavaju ovu ranjivost.⁴²⁴³

Transnacionalne kampanje zahtijevaju koordinisani regionalni odgovor. Izolovani nacionalni naponi pokazuju se neučinkovitim protiv sveprisutnog i brzo evoluirajućeg toka štetnog sadržaja pojačanog AI-em. „Raskomadani digitalni prostor Zapadnog Balkana“ je stvarnost u kojoj digitalne štete nastale u jednoj zemlji brzo prelaze granice, utičući na javnu percepciju i predstavljajući značajne rizike za regionalne demokratske procese.⁴⁴⁴⁵

Supravljanje ovoj eroziji zahtijeva višeslojni pristup. Osnježivanje mladih kao „Digitalnih stražara“ kroz „Digitalni imuni sistem“ je ključni element. To podrazumijeva opremanje mladih naprednom medijskom pismenošću, vještinama kritičkog razmišljanja i algoritamskom sviješću. Inicijative poput „Mreža za provjeru činjenica koje vode mladi“ omogućavaju aktivno praćenje online sadržaja, identifikaciju AI-pokretane dezinformacije i proizvodnju provjerenih kontra-narativa na lokalnim jezicima. Ovo se direktno suprotstavlja „Digitalnoj željeznoj zavjesi“ jer podstiče kritičku procjenu algoritamski poslanih informacija. „Kurikulum o algoritamskoj svijesti i kritičkom razmišljanju“ edukuje mlade o funkcionisanju AI algoritama, personalizaciji sadržaja i iskorištavanju manipulacije, čime se jača otpornost na „kognitivno zarobljavanje“. Transnacionalna priroda ovih prijetnji dodatno zahtijeva „Regionalni centar za digitalnu diplomatiju i kontra-narative“, koji koordinira napore širom Zapadnog Balkana u razvoju i širenju pozitivnih, ujedinjujućih narativa.

Takve proaktivne mjere, u kombinaciji sa strateškim korištenjem „The Acquis Digitalis“ za obavezivanje tehnoloških kompanija na ulaganja u lokalizovane AI moderacijske i jezičke inicijative, predstavljaju ključnu odbranu od obmane digitalnog doba. Bez ovih intervencija, erozija povjerenja mladih i destabilizirajući efekti prekograničnih kampanja dezinformacija predstavljaju neizbježnu prijetnju demokratskoj budućnosti Zapadnog Balkana.

41 <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>, „The ethical problem of electoral AI #3,“ International IDEA, 02. april 2025.

42 <https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/how-the-us-and-europe-can-counter-russian-information-manipulation-about-nonproliferation/> (Issue Brief: Kako SAD i Evropa mogu suprotstaviti ruskoj manipulaciji informacijama o neširenju)

43 Isto.

44 European External Action Service (EEAS). (2025). 3rd EEAS Report on Foreign Information Manipulation and Interference Threats: Exposing the architecture of FIMI operations. Pristupljeno putem <https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3rd-ThreatReport-March-2025-05-Digital-HD.pdf>

45 Isto.

Prakse i odgovori tehnoloških kompanija i regulatornih tijela: Lingvistički razdor

6.1 Odgovornost tehnoloških kompanija i regionalna neosjetljivost

Međunarodne tehnološke kompanije pokazuju ograničenu spremnost da odgovore na jedinstvene kontekste Bosne i Hercegovine i šireg Zapadnog Balkana. Ovo ukazuje na opštu nepažnju prema lokalizovanoj moderaciji sadržaja. Ovaj nedostatak stvara „Pustinju Moderacije“, gdje generičke politike moderacije sadržaja nisu efikasne protiv nijansiranih oblika govora mržnje koji su prisutni u regionu. Ekonomski faktori koji obeshrabruju tehnološke kompanije da značajnije ulažu u tržišta sa nižim prihodima poput Zapadnog Balkana značajno doprinose ovom sistemskom neuspjehu.

Istraga iz 2021. godine koju je sprovedla Balkan Investigative Reporting Network (BIRN) potvrdila je ovaj propust. Studija je utvrdila da „gotovo polovina objava sa toksičnim jezikom na balkanskim jezicima ostaje online, čak i nakon što su Facebook i Twitter potvrdili da sadržaj krši njihove pravila.“ Ovaj podatak ističe jaz u provođenju pravila koji proizlazi iz nedovoljne alokacije resursa i nedostatka AI alata prilagođenih kulturnom kontekstu. Oslanjanje na AI za pregled sadržaja, bez dovoljnog ljudskog nadzora ili lokalizovane jezičke ekspertize, znači da značajan dio štetnog sadržaja, uključujući govor mržnje, prijetnje nasiljem i uznemiravanje, ostaje prisutan na mreži.

Ekonomске realnosti dodatno pogoršavaju ovaj „Lingvistički Razdor“. Iako veliki jezički modeli (LLMs) nude potencijalna rješenja za poboljšanje moderacije sadržaja na jezicima sa ograničenim resursima, implementacija augmentacije konteksta – metode koja poboljšava performanse – povećava troškove za 30-40% po modelu. Tehnološke kompanije, vođene profitnim motivima, nevoljne su da preuzmu ove dodatne troškove za tržišta koja se smatraju manje profitabilnim. Ovako se stvara ciklus u kojem ekonomski faktori direktno dovode do nejednake digitalne zaštite, te ostavljaju digitalne prostore Zapadnog Balkana ranjivim na „Balkanski Algoritamski Pukotinu“.⁴⁶⁴⁷⁴⁸

Posljedica je „Algoritamska provalija“, gdje složene nijanse lokalnog žargona, miješanja jezika, istorijskih aluzija i kontekstualno zavisnog govora mržnje na bosanskom, srpskom i hrvatskom jeziku ostaju uglavnom nevidljive generičkim, globalno treniranim AI modelima. Ova algoritamska neosjetljivost omogućava da naizgled bezazleni, ali duboko štetni sadržaji nekontrolisano proliferiraju, naročito pogađajući mlade. Ovaj sistemski neuspjeh u odgovornosti platformi, ukorijenjen u ekonomskim kalkulacijama, direktno doprinosi „Digitalnoj zavjesi“, jer zarobljava mlade u algoritamski isporučenim eho-komorama koje pojačavaju postojeće etničke i političke podjele.

Rješavanje ove „Pustinja moderacije“ neophodan je korak ka jačanju mladih i demokratije u regionu. „Acquis Digitalis“, kroz proces pristupanja EU, pruža značajan vanjski mehanizam za prisiljavanje tehnoloških kompanija da ulažu u lokalizovanu AI moderaciju i inicijative za jezičke podatke. To zahtijeva obavezivanje na razvoj AI alata za moderaciju posebno treniranih na sveobuhvatnim, kulturno nijansiranim skupovima podataka za B/H/S jezike, uz snažan ljudski nadzor. Takva „Lokalizovana AI Moderacija i Inicijativa za Jezičke Podatke“ predstavlja temeljni element „Digitalnog Imunog Sistema“, jer osigurava da platforme budu odgovorne za digitalnu sigurnost svih korisnika, bez obzira na njihov jezički kontekst ili veličinu tržišta. Bez ove intervencije, region rizikuje da ostane digitalna periferija gdje AI-pokretane štete nastavljaju podrivati društvenu koheziju i demokratske procese.

46 “From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024,” Regional Cooperation Council, decembar 2024, Sekcija 2.3.2, stranica 27.

47 Ibid., str. 27.

48 Ibid., str. 28.

6.2 Nacionalni regulatorni kapacitet: Navigacija digitalnim Divljim Zapadom

Nacionalna regulatorna tijela na Zapadnom Balkanu suočavaju se sa ozbiljnim izazovom u nadzoru digitalnog upravljanja, naročito sa rastućim štetnim efektima koje pokreće AI. Agencija za komunikacije (CRA) u Bosni i Hercegovini predstavlja primjer ovog problema, jer djeluje unutar složenog i često krutog pravnog okvira. Društvene mreže i digitalne platforme, motivisane komercijalnim interesima, izdvajaju nedovoljno resursa za moderaciju sadržaja prilagođenu lokalnim kontekstima, kulturama i jezicima u BiH. Ovo stvara „Pustinju moderacije“, gdje govor mržnje i dezinformacije bujaju uz minimalni nadzor - vakuum upravljanja koji podsjeća na „Digitalni Divlji Zapad“.

Institucionalni kapacitet za suočavanje sa ovim prijetećim izazovima ostaje nedovoljno razvijen. Predstavnicima Izbornog Menadžmenta (EMB) širom Zapadnog Balkana obično pokazuju nizak nivo pismenosti o AI. To izaziva zabrinutost u pogledu njihove sposobnosti da ublaže potencijalne povrede ljudskih prava povezane sa primjenom AI, naročito tokom osjetljivih izbornih perioda. Nedostatak specifičnih AI regulatornih okvira u BiH dodatno pogoršava ovu ranjivost, što ističe „Digitalni Jaz Upravljanja“. Ovaj jaz odražava sistemsku nesposobnost postojećih upravljačkih struktura da se prilagode digitalnom dobu, ostavljajući region podložnim nekontrolisanom uticaju AI.

„Lingvistički razdor“ dodatno naglašava ograničenja nacionalnih regulatornih tijela. Čak i uz potrebne zakonske mandate i tehničku ekspertizu, inherentne poteškoće u moderiranju bosanskog, srpskog i hrvatskog jezika – zbog miješanja jezika, dijalekatskih varijacija i kulturno ukorijenjenog govora mržnje – predstavljaju ozbiljne izazove. Bez lokalizovanih AI modela i ljudske jezičke ekspertize, svaki regulatorni pokušaj da se platforme prisile na djelovanje protiv nijansirano šteta sadržaja bio bi neefikasan. Ovo naglašava neospornu potrebu za značajnim jačanjem kapaciteta unutar regulatornih agencija, tj. da ih se transformišu iz pasivnih posmatrača u efikasne izvršioci.

Istraživanja ilustruju oštru razliku u efikasnosti moderacije, što potvrđuje da trenutni pristupi ostaju neadekvatni za Zapadni Balkan. Ova „Regulatorna Otpornost“ predstavlja sistemski problem. CRA u BiH, iako posvećena digitalnoj transformaciji i svjesna odsustva nacionalnih AI strategija, djeluje bez sveobuhvatne zakonodavne podrške ili dovoljnog broja resursa za efikasan nadzor primjene AI. Ovo stvara „zamašnjivanje“ uticaja AI, jer nacionalna tijela nemaju alate za zahtijevanje transparentnosti ili provođenje odgovornosti od tehnoloških kompanija.

Jačanje kapaciteta je imperativ. Bez značajnih ulaganja u tehničku ekspertizu, ljudske resurse i zakonodavne mandate, nacionalna regulatorna tijela ostaće nepripremljena da se suprotstave štetnim efektima koje pokreće AI. Ovo direktno utiče na mlade, koji su izloženi „Balkanskoj Algoritamskoj Pukotini“ i „Digitalnoj Gvozdenoj Zavjesi“. Jačanje nacionalnog regulatornog kapaciteta ključni je dio izgradnje „Digitalnog Imunog Sistema“. To uključuje implementaciju „Lokalizovane AI Moderacije i Inicijative za Jezičke Podatke“ kako bi se premostio „Lingvistički Razdor“, osiguravajući da regulatorni nadzor efikasno adresira nijanse lokalnih jezika. Takve inicijative, vođene „The Acquis Digitalis“, su neophodne za transformaciju digitalnog okruženja iz „Divljeg Zapada“ minimalnog nadzora u regulisano okruženje koje štiti demokratske procese i sigurnost mladih.

Zaključak

Zapadni Balkan, region obilježen historijskim podjelama i razvijajućim demokratskim procesima, suočava se sa značajnom prijetnjom od štetnog sadržaja pokretanog AI tehnologijama. Bez hitne i koordinisane intervencije, i bez korištenja integracije u EU kao strateškog imperativa, ovi algoritamski tokovi rizikuju nepovratno narušavanje društvene kohezije, demokratskog povjerenja i osnovnih prava građana, naročito digitalno pismenih mladih. Višeslojna odbrana, koja integriše snažne pravne reforme, unaprijeđene institucionalne kapacitete na državnom i entitetskom nivou, proaktivnu odgovornost tehnoloških kompanija i sveobuhvatne inicijative digitalne pismenosti usmjerene na mlade, nije samo razumna, već i hitan predušlov za regionalnu stabilnost i kredibilan put ka evropskoj integraciji. Kolektivna akcija, osjetljiva na jezičke nijanse i historijske traume, jedini je održivi bedem protiv ove sveprisutne digitalne prijetnje.

Istraživačka teza ovog projekta predstavlja temelj za neposredne, mjerljive političke intervencije, ciljne programe jačanja kapaciteta i snažne kampanje podizanja svijesti. Mladi, iako pozicionirani na vrhu ranjivosti, takođe predstavljaju predvodnike otpornog društva; njihovo informisano učešće nije opcionni dodatak, već model za održivu demokratsku budućnost. Završni poziv je na čvrstu posvećenost regionalnoj solidarnosti i pojačanoj međunarodnoj podršci kako bi se ove preporuke pretočile u vidljive, konkretne rezultate.

Konačno stanje koje ovaj izvještaj predviđa za Zapadni Balkan je region karakterisan:

Oснаžena i otporna omladina: Mladi posjeduju napredne digitalne i medijske pismenosti, vještine kritičkog razmišljanja i sposobnost prepoznavanja i suprotstavljanja manipulacijama generisanim od strane AI, aktivno sudjeluju u oblikovanju svojih digitalnih okruženja.⁴⁹

Koordinisani regionalni odgovor: Zemlje Zapadnog Balkana ostvaruju besprijeckornu prekograničnu saradnju, dijele obavještajne podatke, harmoniziraju politike i provode zajedničke inicijative za borbu protiv šteta izazvanih AI-jem, čime jačaju kolektivnu sigurnost.

Odgovorne tehnološke platforme: Međunarodne tehnološke kompanije pokazuju osjetljivost prema jezičkim i kulturnim specifičnostima regije, ulažu u lokalizovanu moderaciju sadržaja i transparentno izvještavaju o svojim naporima.

Ojačana društvena kohezija: Pojačavanje etničkih i vjerskih podjela od strane AI-ja je značajno ublaženo, stvaraju online okruženje koje podržava međunacionalni dijalog i izgradnju mira, umjesto da produbljuje historijske traume.

Kredibilan put integracije u EU: Napredak u digitalnom upravljanju i demokratskoj otpornosti jača vjerodostojnost regije na njenom putu ka članstvu u EU, pokazuju stvarnu posvećenost zajedničkim vrijednostima.⁵⁰⁵¹

Ovo buduće stanje predstavlja značajnu zaštitu od sveprisutnih i evoluirajućih prijetnji koje donosi AI, koji transformiše regiju iz ranjive granice u model digitalne otpornosti.⁵²⁵³

49 International IDEA, "The ethical question of electoral AI #3," 02. april 2025, <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>.

50 Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

51 "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective", European Western Balkans, 4. april 2025.

52 Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint arXiv:2506.09992, 2025, Sekcija IV-E, Pasus 3. Dostupno na: <https://arxiv.org/html/2506.09992v1>

53 Isto.

Izgradnja Svijeta: Širi Univerzum Ideja

Izveštaj „Algoritamski tokovi“ smješten je unutar šireg intelektualnog univerzuma definisanog istovremenim izazovima brzog tehnološkog napretka i demokratske krhkosti u postkonfliktnim društvima. Ovaj univerzum karakterišu sljedeće osnovne tenzije i dinamike:

Paradoks povezanosti: Internet, koji je prvobitno slavljen kao alat za demokratizaciju i globalno razumijevanje, postao je mač sa dvije oštrice. Njegova moć povezivanja sve se više koristi za fragmentaciju društava i širenje zlonamjernog uticaja. AI pojačava ovaj paradoks, jer nudi neviđene alate za osnaživanje i manipulaciju.

Geopolitika informacionog rata: Zapadni Balkan nije samo pasivni primalac digitalnih šteta; on je bojno polje u širem geopolitičkom sukobu za uticaj. Spoljašnji akteri često koriste AI-pokretanu dezinformaciju za destabilizaciju vlada, sijanje razdora i podrivanje aspiracija za evroatlantske integracije. To čini digitalnu otpornost pitanjem nacionalne i regionalne sigurnosti.⁵⁴

Digitalni jaz u upravljanju: Oprezna i neujednačena globalna regulatorna scena za AI stvara „vakuum upravljanja“ koje zlonamjerni akteri iskorištavaju. Regije poput Zapadnog Balkana, sa još uvijek razvijenim strukturama digitalnog upravljanja, postaju posebno ranjive, što naglašava hitnu potrebu za međunarodnom regulatornom harmonizacijom i transferom kapaciteta.⁵⁵

Generacijski imperativ: Mladi, kao digitalni domoroci, žive u ovom složenom informacionom ekosistemu s neuporedivom intenzitetom. Njihova ranjivost nije samo funkcija izloženosti, već i razvoja kritičkih sposobnosti u dinamičnom digitalnom okruženju. Njihovo osnaživanje stoga nije samo filozofski imperativ, već strateška potreba za dugoročnu održivost demokratskih sistema.

Međunarodni primjeri i naučene lekcije iz sličnih postkonfliktnih ili tranzicijskih društava koja se suočavaju sa štetnim uticajima AI nude ključni model za Zapadni Balkan. Ova analiza prevazilazi puko nabranje programa, izvlačeći osnovne principe efikasne intervencije, naročito u jačanju prekogranične saradnje i iskorištavanju zajedničkog jezičkog ili kulturnog naslijeđa.⁵⁶

Projekat „RESILIENCE: Civil society action to reaffirm media freedom and counter disinformation and hateful propaganda in Western Balkans and Turkey“ predstavlja značajan regionalni primjer. Ova inicijativa, koja uključuje obuke za novinare i ranjive grupe poput manjina, pokazuje potencijal pristupa sa više aktera u unapređenju medijske pismenosti i borbi protiv dezinformacija. Naglasak projekta na obuci različitih grupa, uključujući one često marginalizovane, nudi model za rješavanje „Balkanskog algoritamskog raskola“ promovisanjem inkluzivnih informativnih okruženja.

Još jedan značajan međunarodni primjer dolazi iz projekta „AI Democracy 2.0“ u Zimbabveu. Ova inicijativa koristi AI-pokretan WhatsApp chatbot za promociju građanskog obrazovanja, borbu protiv dezinformacija i pružanje provjerenih informacija. Efikasnost troškova i dostupnost chatbota, naročito u kontekstu sa ograničenim građanskim prostorom, predstavljaju model prilagođavanja digitalnih alata lokalnim potrebama. Ovaj model bi se mogao prilagoditi za Zapadni Balkan, jer koristi široko korištene platforme za razmjenu poruka za širenje provjerenih informacija i suzbijanje AI-pokretanih dezinformacija, posebno među mladima koji su česti korisnici takvih aplikacija.

„*Vulnerability Index of Disinformation*“ iz Kosova pruža regionalni primjer proaktivnog alata za procjenu ranjivosti. Ova inicijativa procjenjuje ranjivost na informacijske poremećaje i predlaže mjere ublažavanja za institucije koje donose odluke. Sličan indeks, prilagođen specifičnim etničkim i istorijskim osjetljivostima BiH, mogao bi informisati ciljne intervencije i razvoj politika, direktno adresirajući „Digitalne memorijske ratove“ koje pogoni algoritamska pristrasnost.

54 <https://www.unesco.org/en/articles/unesco-supports-launch-coalition-freedom-expression-and-content-moderation-bosnia-and-herzegovina>

55 <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>, „The ethical problem of electoral AI #3,“ International IDEA, 02. april 2025.

56 Communications Regulatory Agency (CRA) Bosnia and Herzegovina, „GSR-25 Contribution: Best Practice Guidelines,“ predstavljeno na GSR-25 Konsultacijama, pristupljeno putem https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf

Media Diversity Institute Western Balkans, sa sjedištem u Srbiji, upravlja mrežom „Reporting Diversity Network 2.0“. Ova mreža promoviše tačno i inkluzivno novinarstvo o različitostima, te tako utiče na medijsko predstavljanje etničke, vjerske i rodne raznolikosti na Zapadnom Balkanu. Fokus na njegovanje pozitivnog diskursa i harmoničnih odnosa među susjedima direktno se suprotstavlja partijskoj amplifikaciji podijeljenih narativa putem AI algoritama. Ova inicijativa nudi model za novinarsku obuku i zajednički rad na sadržaju koji jača nezavisne medije i gradi povjerenje javnosti.

Ovi primjeri ističu principe efikasne intervencije:

Kontekstualna relevantnost: Programi su najučinkovitiji kada su prilagođeni specifičnim lokalnim potrebama, jezičkim nijansama i društveno-istorijskim kontekstima.

Saradnja više aktera: Uspjeh zavisi od saradnje između vlada, civilnog društva, medija i tehnoloških kompanija.

Dizajn usmjeren na mlade: Uključivanje mladih kao aktivnih učesnika, a ne samo pasivnih primaoca, povećava domet i uticaj programa.

Proaktivni naspram reaktivnih pristupa: Strategije koje grade otpornost i preventivno demantuju dezinformacije pokazuju se održivijim od isključivo reaktivnog uklanjanja sadržaja.

Odgovorno korištenje digitalnih alata: Korištenje AI i digitalnih platformi za pozitivan građanski angažman, uz ublažavanje njihove štetne potencijale.

Takve inicijative, zajedno sa strateškim iskorištavanjem „The Acquis Digitalis“ za prisiljavanje tehnoloških kompanija na ulaganja u lokalizovanu AI moderaciju i inicijative za jezičke podatke, predstavljaju ključnu odbranu od podmuklog digitalnog talasa. Ove lekcije pružaju neophodnu osnovu za razvoj kontekstualno specifičnih odgovora na Zapadnom Balkanu, pokazujući da je efikasna digitalna otpornost ostvariv rezultat učenja iz globalnih i regionalnih iskustava.

Preporuke za politike i mehanizme odgovornosti

Oni su predstavljeni kao numerisane liste, strukturisane pod jasnim kategorijama, što osigurava da je svaka preporuka ili korak jasno definisan, izvediv i mjerljiv. Postoji više od 30 takvih stavki, što pruža sveobuhvatan plan djelovanja. Svaka stavka doprinosi ukupnom cilju jačanja regije.⁵⁷

8.1 Regionalne strategije saradnje za Zapadni Balkan

Prekogranična priroda AI-pokretanih dezinformacija zahtijeva jedinstvenu regionalnu strategiju širom Zapadnog Balkana. Fragmentirani nacionalni naponi nisu dovoljni protiv sveprisutnog toka štetnog sadržaja koji iskorištava zajedničke jezičke i kulturne prostore. Ove preporuke promovišu kolektivnu otpornost, stvaraju jedinstveni front protiv zlonamjernih uticaja i sprječavaju da AI-pokretane štete iskorištavaju nacionalne granice.

Uspostavljanje Foruma za digitalnu saradnju Zapadnog Balkana za redovni dijalog: Neophodan je posvećeni Forum za digitalnu saradnju Zapadnog Balkana. Ovaj forum, potencijalno pod okriljem Regional Cooperation Council (RCC) ili Berlin Process, omogućava redovni, strukturirani dijalog između donosilaca politika, regulatornih tijela i stručnjaka za sajber sigurnost iz svih zemalja Zapadnog Balkana. Njegov mandat obuhvata razmjenu informacija o prijetnjama, raspravu o novim taktikama AI-pokretanih dezinformacija

⁵⁷ Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

i koordinaciju političkih odgovora. Ova proaktivna angažovanost direktno se bavi „Fractured Digital Expanse of the Western Balkans“ promovisanjem zajedničkog razumijevanja digitalnih prijetnji i podsticanjem kolektivne akcije, integrišući nacionalne odgovore u širu regionalnu odbranu.

Razvijanje regionalnog sistema ranog upozoravanja za AI-pokretane dezinformacije: Regionalni sistem ranog upozoravanja za AI-pokretane dezinformacije omogućava brzi odgovor na nove prijetnje. Ovaj sistem, koji se oslanja na predloženu Western Balkans Cyber Threat Intelligence Network, integriše podatke iz nacionalnih centara za sajber sigurnost, organizacija za medijsko praćenje i mreža za provjeru činjenica koje vode mladi. Koristi AI za sofisticiranu detekciju prijetnji i prediktivnu analitiku, identifikujući rane kampanje dezinformacija i sintetičke medije prije nego što postanu široko viralni. Sistem pruža operativne informacije u realnom vremenu, omogućavajući koordinisano plasiranje kontra-narativa i ciljane kampanje podizanja svijesti javnosti. Ovo se direktno suprotstavlja „The Ephemeral Nature of Truth“ jer smanjuje vremenski jaz između pojave dezinformacija i organizovanog odgovora.

Standardizacija mehanizama izvještavanja o moderaciji sadržaja preko granica: Nedostatak standardizovanih mehanizama izvještavanja o moderaciji sadržaja preko granica otežava odgovornost platformi i pogoršava „The Linguistic Seam-Rip“. Ova preporuka zagovara jedinstveni regionalni protokol za prijavu štetnog sadržaja, naročito na bosanskom, srpskom i hrvatskom jeziku, prema glavnim tehnološkim platformama. To uključuje usklađivanje definicija govora mržnje i dezinformacija, uspostavljanje jasnih kanala za prijavu i obavezivanje na transparentne povratne informacije od platformi o preduzetim mjerama. Takva standardizacija, koristeći EU-ov „Acquis Digitalis“ kao paradigmu za regulatorno usklađivanje, prisiljava platforme da efikasnije raspoređuju resurse za lokalizovanu moderaciju, čime se zatvara „Pustinja moderacije“ i poboljšava digitalna sigurnost mladih.

Olakšavanje zajedničkih istraživanja i razmjene podataka o štetnim uticajima AI: Nedostatak lokalizovanih istraživanja o uticaju AI na Zapadnom Balkanu zahtijeva inicijative za zajednička istraživanja i razmjenu podataka. Ova preporuka uspostavlja regionalni istraživački konzorcijum, uključujući univerzitete, think tanke i organizacije civilnog društva, za provođenje interdisciplinarnih studija o štetnim uticajima AI. To uključuje kreiranje zajedničkog, anonimizovanog skupa podataka o štetnom sadržaju na B/H/S jezicima, podsticanje razvoja AI alata za moderaciju prilagođene kontekstu i analizu društveno-političkog uticaja algoritamske pristrasnosti. Ovaj kolaborativni pristup adresira „Algorithmic Blind Spots“ jer kreira empirijski utemeljene uvide u to kako AI iskorištava postkonfliktni traum i istorijske podjele.

Harmonizacija regionalnih pristupa odgovornosti platformi, koristeći paradigmu EU: Digital Services Act (DSA) Evropske unije nudi sveobuhvatan okvir za odgovornost platformi. Ova preporuka zagovara harmonizaciju regionalnih pristupa odgovornosti platformi u zemljama Zapadnog Balkana, koristeći DSA kao primarnu paradigmu. To podrazumijeva usvajanje zajedničkih pravnih standarda za transparentnost platformi, ublažavanje rizika i moderaciju sadržaja, posebno u vezi sa štetama koje uzrokuje AI. Regionalna koordinacija, koju olakšava RCC, omogućava Zapadnom Balkanu da se predstavi kao jedinstvena cjelina prema tehnološkim kompanijama, povećavajući tako pregovaračku moć za zahtjeve za pravičnim ulaganjem u lokalizovanu moderaciju i zaštitu podataka. Ova usklađenost jača „Acquis Digitalis“, osigurava dosljedne, visoke standarde digitalne odgovornosti u cijelom regionu.

Ove strategije regionalne saradnje pružaju vanjsku dimenziju ukupnom odgovoru. One priznaju da je, iako je unutrašnja društvena otpornost, naročito među mladima, jednako ključna za dugoročni uspjeh, snažne regionalne strukture neophodne za efikasnu primjenu i odgovorne prakse AI. Ovaj kolektivni pristup jača „Digitalni imuni sistem“, što omogućava Zapadnom Balkanu efikasnije suprotstavljanje štetama koje uzrokuje AI.

8.2 Uključivanje mladih i programi razvoja vještina

Mladi su istovremeno veoma izloženi algoritamskim štetama i ključni resurs za izgradnju društvene otpornosti. Regionalne studije dosljedno pokazuju nizak nivo medijske pismenosti među mlađim generacijama u Zapadnom Balkanu, što ukazuje na njihovu povećanu ranjivost. Intervencije zasnovane na obrazovanju stoga predstavljaju ključnu polaznu tačku za podsticanje trajnih promjena.

Digitalna i medijska pismenost moraju se integrisati u nastavni plan od najranijih razreda, jer tako postavlja algoritamsku svijest kao osnovni element. Ključni kurikularni elementi uključuju lateralno čitanje, imunizaciju protiv uobičajenih taktika manipulacije, provjeru izvora i praktične vježbe u triangulaciji izvora. Dokazi pokazuju da ove tehnike poboljšavaju navike verifikacije i smanjuju ranjivost na lažne tvrdnje. Kurikulumi bi trebali usvojiti aktivne, projektno orijentisane metode učenja umjesto tradicionalnih predavanja. Implementacija zahtijeva rješavanje infrastrukturnih nedostataka i deficita u obuci nastavnika, kako je utvrđeno nacionalnim ICT procjenama. Ulaganje u usavršavanje nastavnika i osnovnu povezanost je preduslov za pravičan pristup.

Razvijanje i proširenje mreža mladih koje vode kontranarative i provjeru činjenica na lokalnim jezicima i dijalektima je od suštinske važnosti. Regionalni pilot programi potvrđuju da integracija mladih kreatora unutar javnih medija i pružanje uredničkog mentorstva rezultira značajnim dosegom i angažmanom, uz očuvanje kredibiliteta. Ovi uspješni modeli zahtijevaju proširenje, finansiranje i prekograničnu povezanost kako bi efikasno odgovorili na transnacionalne tokove sadržaja. Ove mreže proizvode višejezične sadržaje, brze izvještaje za regulatore i skupove podataka koji informišu o poboljšanjima lokalizovane moderacije sadržaja.

Kreiranje sigurnih online prostora i mentorskih programa povezuje psihosocijalnu podršku sa praktičnim kritičkim angažmanom. Ovi sigurni prostori obuhvataju moderirane forume, verifikovane mreže vršnjaka i kanale za upućivanje u kriznim situacijama, kojima zajednički upravljaju civilno društvo, škole i postojeća infrastruktura *Safer Internet Centre*. Ove usluge moraju uključivati prakse usmjerene na traumu, te da uzmu u obzir postkonfliktni kontekst regiona. Mentori bi trebali dolaziti iz oblasti novinarstva, digitalne forenzike i medijacije u zajednici, jer se tako osigurava zajednički prenos tehničkih vještina i građanskih vrijednosti.

Razvijanje kritičkog mišljenja protiv algoritamske manipulacije je ključno. Podučavanje mehanike sistema preporuka, ekonomije optimizacije angažmana i uobičajenih grešaka automatizovane moderacije razjašnjava neprozirne procese koji doprinose digitalnoj gvozdenoj zavjesi. Pedagogija bi trebala kombinovati praktičnu upotrebu alata, poput vježbi u otkrivanju sintetičkih medija i prepoznavanju obrazaca, sa refleksivnim modulima o istorijskim narativima. Ovaj pristup omogućava mladima da prepoznaju sadržaje koji iskorištavaju kolektivnu traumu.

Osnaživanje mladih kao zagovornika digitalnih prava uključuje njihovo učešće u političkim forumima, procesima odgovornosti platformi i regionalnim mehanizmima koordinacije. Ovo učešće priprema mlade da doprinesu lokalizovanim dokazima, uključujući uzorke jezika, propuste u moderaciji i podatke iz pilot projekata kontranarativa. Takođe im omogućava da drže aktere odgovornim za obaveze u okviru EU usklađenog *acquis-a*. Pozicioniranje mladih kao legitimnih sagovornika u politici pretvara ranjivost u agenciju i usklađuje razvoj kapaciteta sa pravnim uticajem koji pruža približavanje EU.

Svaki od ovih elemenata čini modul unutar sveobuhvatnog *Digital Immune System-a*, koji obuhvata obrazovanje, praktičnu primjenu, sigurne prostore, tehničku pismenost i građanski angažman. Ove komponente, kada su integrisane, pretvaraju izloženost u imunizaciju i pasivnu publiku u aktivne, međunacionalne mreže sposobne da izazovu algoritamski pojačane štete.

8.3 Uključivanje više aktera i razvoj kapaciteta⁵⁸

Još jedan značajan međunarodni primjer dolazi iz projekta "AI Democracy 2.0" u Zimbabveu. Ova inicijativa koristi WhatsApp chatbot zasnovan na vještačkoj inteligenciji kako bi promovisala građansko obrazovanje, borila se protiv dezinformacija i pružala provjerene informacije.⁵⁹ Isplativost i pristupačnost chatbota, posebno u kontekstima sa ograničenim građanskim prostorom, predstavljaju model za prilagođavanje digitalnih alata lokalnim potrebama. Ovaj model bi mogao biti prilagođen za Zapadni Balkan, tako da koristi široko rasprostranjene platforme za razmjenu poruka kako bi se širile provjerene informacije i suprotstavilo dezinformacijama zasnovanim na vještačkoj inteligenciji, posebno među mladima koji su prevashodni korisnici takvih aplikacija. Još neke akcije:

Obavezivanje na redovni dijalog između vlada i tehnoloških kompanija: Ograničena spremnost međunarodnih tehnoloških kompanija u Zapadnom Balkanu, karakterisana nedovoljnom pažnjom prema lokalizovanim potrebama moderacije, zahtijeva strukturirani angažman. Ova preporuka uspostavlja redovne, formalizovane dijaloge između državnih i entitetskih vlasti BiH i velikih tehnoloških kompanija. Ovi dijalozi, potencijalno posredovani od strane *Communications Regulatory Agency (CRA) BiH*, fokusiraju se na transparentnost politika moderacije sadržaja, raspodjelu resursa za B/H/S jezike i odgovornost za algoritamske štete. Cilj je podstaći saradničko okruženje u kojem tehnološke kompanije dijele znanje o radu AI sistema, a vlade artikuliraju lokalne potrebe, naročito u vezi sa digitalnom gvozdеном завjesom.⁶⁰

Jačanje kapaciteta *Communications Regulatory Agency (CRA) BiH*: CRA BiH, iako posvećena digitalnoj transformaciji, djeluje bez sveobuhvatne zakonodavne podrške ili dovoljnog resursa za efikasan nadzor primjene AI. Ova preporuka poziva na značajna ulaganja u tehničku ekspertizu, ljudske resurse i zakonodavne mandate CRA BiH. To uključuje specijalizovanu obuku u upravljanju AI, digitalnoj forenzici i medijskoj i informacijskoj pismenosti za osoblje. Osaživanje CRA BiH kroz "regulatory sandboxes, living labs i innovation hubs" omogućava praktičan, eksperimentalni pristup razumijevanju i upravljanju novim tehnologijama. Ovo unapređenje transformiše CRA iz pasivnog posmatrača u efikasnog izvršioca, sposobnog da zahtijeva odgovornost platformi i osigurava zaštitu mladih od šteta koje uzrokuje AI. Ovaj korak je ključan za prevazilaženje "*Regulatory Recalcitrance*."⁶¹

Jačanje kapaciteta civilnog društva i medijskih provjerivača činjenica: Organizacije civilnog društva i nezavisni mediji igraju važnu ulogu u praćenju online sadržaja i suprotstavljanju dezinformacijama, naročito tamo gdje državne institucije imaju ograničene kapacitete. Ova preporuka jača njihove kapacitete za provjeru činjenica kroz namjenska finansiranja, naprednu obuku u digitalnoj forenzici i pristup AI-pokretanim alatima za verifikaciju. Inicijative poput BiH "Why Not Organization" i "Citizens Against Terrorism Bosnia and Herzegovina" (CAT BiH) predstavljaju postojeće modele efikasnih, mladima vođenih strategija kontranarativa. Ovo jača nezavisni nadzor, pruža kredibilne kontranarative.

Promocija javnih kampanja o AI i štetama na internetu: Raširena prisutnost AI-pokretanih dezinformacija zahtijeva sveobuhvatne javne kampanje podizanja svijesti. Ova preporuka zagovara regionalno koordinisane kampanje, potencijalno koristeći ekspertizu *Regional Cooperation Council (RCC)*. Kampanje edukuju građane, naročito mlade, o mehanizmima manipulacije AI, karakteristikama sintetičkih medija i opasnostima algoritamske pristrasnosti. Kampanje bi trebale koristiti formate prilagođene mladima, poput sadržaja na društvenim mrežama, interaktivnih radionica i edukativnih videa, kako bi povećale AI pismenost i podstakle kritički digitalni angažman. Ovaj proaktivan pristup razvija promišljenije građane, smanjujući ranjivost na "*cognitive capture*" i jačajući kolektivni "digitalni imuni sistem" protiv šteta koje uzrokuje AI.

Ulaganje u lokalizovanu jezičku ekspertizu za moderaciju sadržaja: "*Linguistic Seam-Rip*," gdje AI moderacija nedovoljno prepoznaje nijansirani govor mržnje na bosanskom, srpskom i hrvatskom jeziku, zahtijeva ciljano ulaganje. Ova preporuka predlaže "*Localized AI Moderation and Language Data Initiative*," vođenu "The Acquis Digitalis" i međunarodnim finansiranjem. Potrebno je da tehnološke kompanije značajno ulažu u razvoj AI alata za moderaciju posebno treniranih na sveobuhvatnim, kulturološki nijansiranim skupovima podataka za B/H/S jezike. To uključuje finansiranje kreiranja skupova podataka, obuku lokalnih lingvista i AI stručnjaka, te integraciju ljudske revizije za složene slučajeve. Ova inicijativa direktno

⁵⁸ <https://open-research-europe.ec.europa.eu/articles/5-122/pdf> (Stranice 1-22)

⁵⁹ Council of Europe, Programme of the World Forum for Democracy 2024: Democracy and Diversity, 6-8 November 2024, Strasbourg, Stranica 15.

⁶⁰ Ibid., Stranica 15.

⁶¹ Ibid., Stranica 16.

adresira trajne izazove jezičke moderacije, jer tako osigurava da platforme budu odgovorne za digitalnu sigurnost svih korisnika, bez obzira na njihov jezički kontekst ili veličinu tržišta. Ovaj osnovni sloj zaštite dopunjuje napore u digitalnoj pismenosti mladih, osiguravajući da čak i sofisticirane štete koje uzrokuje AI budu efikasno otkrivene i ublažene.

Strateški plan: Faza implementacije i izgradnja otpornosti

9.1 Strateški plan: Koraci vlade (državni i entitetski nivo) za BiH

Strateški plan za Bosnu i Hercegovinu zahtijeva fazni pristup vlasti, uključujući i državni i entitetski nivo. Ovaj okvir adresira složenu ustavnu strukturu zemlje i trajnu "Regulatory Recalcitrance" koja ometa jedinstvenu digitalnu upravu. Ove mjere usklađuju BiH sa međunarodnim najboljim praksama i njenim EU aspiracijama, stvarajući okruženje pogodno za snažan "Digital Immune System."

Faza 1 (kratkoročno): Provođenje sveobuhvatne analize pravnih praznina: Ključni početni korak uključuje sveobuhvatnu analizu pravnih praznina na svim nivoima vlasti u BiH. Ova analiza identifikuje nesklade između postojećeg zakonodavstva na državnom i entitetskom nivou i razvijajućih standarda EU digitalnog upravljanja, naročito onih definisanih u Digital Services Act (DSA) i predstojećem AI Act. Ova osnovna analiza otkriva zakonske praznine, posebno u vezi sa ulogom AI u pojačavanju govora mržnje i dezinformacija, te ističe oblasti gdje postojeći zakoni nisu efikasni ili su kontradiktorni zbog ustavne fragmentacije BiH. Rezultat je jasan, operativan izvještaj koji detaljno navodi konkretne zakonske izmjene potrebne za punu usklađenost sa Acquis Digitalis.

Faza 1 (Kratkoročno): Uspostaviti međuentitetske radne grupe za digitalno upravljanje: Prevazilaženje fragmentirane ustavne strukture BiH zahtijeva hitno uspostavljanje međuentitetskih radnih grupa posvećenih digitalnom upravljanju. Ove grupe, sastavljene od stručnjaka iz institucija na državnom nivou, Federacije BiH, Republike Srpske i Brčko Distrikta, podstiču kontinuiranu komunikaciju i zajedničko razumijevanje šteta uzrokovanih AI tehnologijom. Njihov mandat uključuje pregled analize pravnih praznina, identifikaciju zajedničkih polaznih osnova za usklađivanje politika i razvoj zajedničkih strategija protiv dezinformacija pokretanih AI-jem. Ovaj kolaborativni mehanizam direktno suprotstavlja se "Digitalnom dvostrukom vezivanju", gdje AI nenamjerno jača političke podjele, stvarajući zajedničku platformu za razvoj nacionalnih digitalnih politika.

Faza 2 (Srednjoročno): Izraditi i usvojiti zakonodavstvo usklađeno sa DSA: Iskorištavajući integraciju u EU kao snažan pokretač, srednjoročna strategija fokusira se na izradu i usvajanje zakonodavstva potpuno usklađenog sa Digital Services Act (DSA). To podrazumijeva prevođenje nalaza analize pravnih praznina u konkretne zakonodavne prijedloge, osiguravajući da zakoni BiH odražavaju sveobuhvatni okvir DSA za sigurnost na internetu, odgovornost platformi i zaštitu digitalnih prava. Ovaj korak je ključan za rješavanje "Pustinje moderacije" prisiljavajući tehnološke kompanije na primjenu snažnih standarda moderacije sadržaja, uključujući one prilagođene bosanskom, srpskom i hrvatskom jeziku. Proces zahtijeva blisku koordinaciju između međuentitetskih radnih grupa i relevantnih parlamentarnih tijela, uz tehničku pomoć EU stručnjaka, kako bi se ubrzalo usvajanje zakonodavstva i prevazišle političke blokade.

Faza 2 (Srednjoročno): Dodijeliti budžet za jačanje kapaciteta CRA: Agencija za komunikacije (CRA)

u BiH trenutno nema sveobuhvatnu zakonsku podršku niti dovoljne resurse za efikasan nadzor implementacije AI tehnologija. Ovaj korak zahtijeva značajnu budžetsku alokaciju za jačanje kapaciteta CRA. To uključuje finansiranje specijalizovanih obuka u oblasti upravljanja AI, digitalne forenzike i medijske i informacijske pismenosti za osoblje CRA. Ulaganja u "regulatorne sandbox-e, living labs i inovacijske centre" pretvaraju CRA u proaktivno i efikasno nadzorno tijelo sposobno zahtijevati transparentnost i provoditi odgovornost tehnoloških kompanija. Ovo direktno rješava "Regulatornu otporost" opremajući CRA potrebnim alatima i stručnim znanjem za ublažavanje šteta uzrokovanih AI-jem, čime se jača "Digitalni imuni sistem" protiv raširenih dezinformacija.

Faza 3 (Dugoročno): Implementirati državni okvir za etiku AI: Dugoročni cilj podrazumijeva implementaciju sveobuhvatnog državnog okvira za etiku AI. Ovaj okvir, koji crpi principe iz EU AI Act, osigurava odgovornu primjenu AI u svim sektorima, posebno onima koji utiču na javni diskurs, izborne procese i mlade. Uspostavlja smjernice za razvoj i upotrebu AI, fokusira se na ljudski nadzor, kvalitet podataka, transparentnost i nediskriminaciju. Ovaj okvir direktno se suprotstavlja "Algoritamskim slijepim tačkama" i "Paradoksu političkog pilava" proaktivnim pristupom u adresiranju načina na koji AI može nenamjerno perpetuirati sistemске pristrasnosti i iskorištavati traume iz postkonfliktnih perioda. Osigurava da AI aplikacije budu u skladu s etičkim principima, pravnim okvirom javne uprave i slobodnim i demokratskim osnovnim poretom.⁶²

9.2 Strateški plan: Koraci civilnog društva i medija za BiH⁶³

Organizacije civilnog društva i medijske institucije u Bosni i Hercegovini (BiH) imaju ključnu ulogu u jačanju digitalne otpornosti. Njihov nezavisni nadzor i naponi u javnom obrazovanju su od suštinskog značaja za suprotstavljanje štetama uzrokovanim AI tehnologijom, posebno kada su odgovori vlasti i platformi nedovoljni. Ovaj strateški plan iznosi ključne aktivnosti za ove ne-državne aktere, fokusirajući se na proaktivno uključivanje i zagovaranje jačanja nezavisnog informacijskog ekosistema BiH.⁶⁴

Faza 1 (Kratkoročno): Proširiti mreže za provjeru činjenica i alate: Jačanje nezavisne verifikacije protiv raširenih dezinformacija zahtijeva hitno proširenje mreža za provjeru činjenica i usvajanje naprednih alata. Organizacije poput BiH "Why Not Organization" već se bave ekstremističkim i etnonacionalističkim narativima kroz rigoroznu provjeru činjenica. To uključuje povećanje finansiranja, pružanje obuka iz digitalne forenzike i omogućavanje pristupa softveru za verifikaciju vođenom AI-jem postojećim i novim inicijativama za provjeru činjenica koje vode mladi. Ove mreže, kao ključni dio "Digitalnog imunog sistema", pretvaraju mlade u proaktivne digitalne stražare, direktno se suprotstavljajući "Balkanskoj algoritamskoj frakturi" omogućavajući decentralizovano otkrivanje i odgovor na algoritamski pojačane podjeljujuće narative.⁶⁵

Faza 1 (Kratkoročno): Pokrenuti kampanje javne svijesti o deepfake sadržajima: Rješavanje "Prolazne prirode istine" i raširenosti AI halucinacija zahtijeva hitne kampanje javne svijesti. Ove kampanje edukuju građane, naročito mlade, o karakteristikama sintetičkih medija, mehanizmima AI manipulacije i rizicima algoritamske pristrasnosti. Korištenje formata prilagođenih mladima, interaktivnih poput sadržaja na društvenim mrežama, edukativnih videa i praktičnih radionica maksimizira domet i uticaj. Kampanje trebaju istaknuti primjere deepfake sadržaja u kontekstu rodno zasnovanog nasilja, kao što je zabilježeno u Telegram mrežama širom Balkana, kako bi se naglasile stvarne štete. Ova strategija inokulacije, upoznavajući pojedince s taktikama dezinformacija, pruža "kognitivni imunitet" prilikom susreta s lažnim informacijama.⁶⁶

Faza 2 (Srednjoročno): Razviti specijalizovane obuke za novinare o AI dezinformacijama: Oprema medijskih profesionalaca za istraživanje i izvještavanje o složenim štetama uzrokovanim AI-jem predstavlja

⁶² "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.3, stranica 29.

⁶³ Ibid., str. 32.

⁶⁴ Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

⁶⁵ Ibid., Sekcija II-C, Stav 2.

⁶⁶ "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective", European Western Balkans, 4. april 2025.

srednjoročni prioritet. To uključuje razvoj specijalizovanih programa obuke za novinare širom BiH, fokusiranih na ulogu AI u dezinformacijama, detekciju deepfake sadržaja, analizu algoritamske pristrasnosti i jedinstvene izazove "Jezikog šava". "Reporting Diversity Network 2.0" Media Diversity Institute Western Balkans nudi model za promociju tačnog i inkluzivnog novinarstva, koji se može proširiti uključivanjem AI-specifičnih obuka. Obuke trebaju naglasiti "lateralno čitanje" i provjeru izvora umjesto površnih pokazatelja pouzdanosti. Ovo povećava kapacitet medija za suprotstavljanje AI-pojačanim "Ratovima digitalnog sjećanja" i očuvanje novinarskog integriteta.

Faza 2 (Srednjoročno): Zagovarati veću transparentnost platformi: Organizacije civilnog društva i mediji moraju pojačati zagovaranje veće transparentnosti platformi. Ovo vrši pritisak na tehnološke kompanije da adresiraju jezičke probleme i unaprijede moderaciju. Uključuje zahtjeve za detaljnim podacima o ulaganjima platformi u lokalizovanu moderaciju sadržaja za B/H/S jezike, izvještavanje o neuspjesima u moderaciji i transparentnu komunikaciju procesa algoritamskog donošenja odluka. Korištenjem okvira "Acquis Digitalis", zagovaračke aktivnosti prisiljavaju tehnološke kompanije da ravnomjerno raspodijele resurse među jezičnim grupama, pretvarajući "Pustinjsku zonu moderacije" u regulisaniji prostor. Ovo direktno izaziva opreznu reakciju međunarodnih tehnoloških kompanija i njihovu "algoritamsku indiferentnost" prema jedinstvenim jezičkim i kulturnim nijansama Zapadnog Balkana.

Faza 3 (Dugoročno): Uspostaviti Centar za digitalnu otpornost mladih: Dugoročni cilj je uspostavljanje sveobuhvatnog Centra za digitalnu otpornost mladih u BiH. Ovaj centar centralizuje osnaživanje mladih i inicijative za kontranarative, integriše elemente "Mreža za provjeru činjenica koje vode mladi", "Kurikulum za algoritamsku svijest i kritičko razmišljanje" i "Digitalnog pripovijedanja sa osvrtom na traumu za pomirenje". Centar podstiče sigurne online prostore i mentorski program, razvija kritički digitalni angažman. Osnažuje mlade kao zagovornike digitalnih prava, omogućavajući im učešće u političkim forumima i procesima odgovornosti platformi. Ova inicijativa, temelj "Digitalnog imunog sistema", osigurava da mladi nisu samo zaštićeni od šteta uzrokovanih AI-jem, već aktivno doprinose otpornijoj, demokratskoj digitalnoj budućnosti.⁶⁷

9.3 Strateški plan: Regionalni koordinacioni mehanizmi za Zapadni Balkan⁶⁸⁶⁹

Fragmentirani nacionalni naponi nisu dovoljni protiv kontinuiranog toka štetnog sadržaja koji iskorištava zajedničke jezičke i kulturne prostore. Sljedeće preporuke podstiču kolektivnu otpornost, stvaraju jedinstveni front protiv malignog uticaja i sprječavajući da štete uzrokovane AI-jem iskorištavaju državne granice.⁷⁰

Faza 1 (Kratkoročno): RCC organizuje redovne dijaloge o digitalnoj politici: Regionalni savjet za saradnju (RCC) uspostavlja i organizuje redovne dijaloge o digitalnoj politici. Ovi dijalozi podstiču kontinuiranu komunikaciju i zajedničko razumijevanje šteta uzrokovanih AI-jem među kreatorima politika, regulatornim tijelima i stručnjacima za sajber sigurnost iz svih zemalja Zapadnog Balkana. Ovaj proaktivni angažman rješava "Fragmentirani digitalni prostor Zapadnog Balkana" promoviše kolektivno razumijevanje prijatelji. Integriše nacionalne odgovore u širu regionalnu odbranu, osigurava da sveprisutni uticaj AI-ja naiđe na koordinisani, a ne razjedinjeni pristup. Uloga RCC-a je ključna u facilitiranju ove razmjene, pretvarajući pojedinačne nacionalne zabrinutosti u kohezivnu regionalnu agendu.⁷¹

Faza 1 (Kratkoročno): Pokrenuti prekogranične programe obuke za regulatore: Izgradnja usklađenog stručnog znanja za rješavanje složenih izazova šteta uzrokovanih AI-jem zahtijeva pokretanje prekograničnih programa obuke za regulatore širom Zapadnog Balkana. Ovi programi se fokusiraju na upravljanje AI, digitalnu forenziku i medijsku i informacijsku pismenost, osiguravajući dosljedan nivo tehničkog i pravnog razumijevanja u regionu. Obuka se bavi pitanjima poput jezika, osposobljava regulatore za ra-

⁶⁷ <https://osis.bg/?p=3750&lang=en>

⁶⁸ <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>

⁶⁹ [https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/Publications/2023/Digital_Development_Country_Profile_Bosnia_and_Herzegovina_\[final- Mart 2023\].pdf](https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/Publications/2023/Digital_Development_Country_Profile_Bosnia_and_Herzegovina_[final- Mart 2023].pdf)

⁷⁰ <https://www.publicmediaalliance.org/how-western-balkans-public-media-empowered-youth/>

⁷¹ Isto.

zumijevanje i odgovor na nijansirani govor mržnje na bosanskom, srpskom i hrvatskom jeziku. Ova kolaborativna izgradnja kapaciteta jača "Digitalni imuni sistem" poboljšava institucionalnu spremnost i ublažava "Pustinju moderacije" nastalu nedovoljnim ulaganjima platformi.⁷²

Faza 2 (Srednjoročno): Razviti zajednički regionalni stav o upravljanju AI: Zapadni Balkan razvija zajednički regionalni stav o upravljanju AI. Ovaj jedinstveni stav predstavlja kohezivni front prema tehnološkim kompanijama i međunarodnim tijelima, povećavajući pregovaračku snagu regiona u zahtjevima za pravednu raspodjelu resursa za moderaciju sadržaja i zaštitu podataka. Ovaj zajednički stav, informisan EU standardima i "Acquis Digitalis" kao primarnim modelom, olakšava usklađivanje nacionalnih zakonodavstava i regulatornih praksi. Direktno se suprotstavlja "Digitalnom dvostrukom vezivanju" jer podstiče jedinstvo u digitalnoj politici, osigurava da AI nenamjerno ne pojačava političku fragmentaciju unutar regiona. Ova strateška usklađenost je ključna za suprotstavljanje prekograničnim kampanjama dezinformacija i promociju kolektivne digitalne sigurnosti.⁷³

Faza 2 (Srednjoročno): Uspostaviti Regionalni savjetodavni odbor za etiku AI: Regionalni savjetodavni odbor za etiku AI pruža nezavisne stručne smjernice o odgovornoj primjeni AI širom Zapadnog Balkana. Odbor čine akademici, predstavnici civilnog društva, tehnički stručnjaci i zagovornici mladih, osiguravajući perspektivu više zainteresovanih strana. Njegov mandat uključuje savjetovanje o etičkim implikacijama AI sistema, posebno u vezi sa istorijskom traumom, etničkim osjetljivostima i osnovnim pravima. Odbor takođe procjenjuje potencijalne "Algoritamske slipe tačke" koje bi mogle perpetuirati diskriminaciju ili pojačavati podjeljujuće narative. Ova proaktivna mjera osigurava da razvoj i primjena AI budu u skladu s principima usmjerenim na čovjeka.

Faza 3 (Dugoročno): Kreirati okvir za Digitalno jedinstveno tržište Zapadnog Balkana: Dugoročni cilj je kreiranje okvira za Digitalno jedinstveno tržište Zapadnog Balkana. Ovaj okvir, koristeći EU paradigmu za digitalnu integraciju, podstiče sigurnu i integrisanu digitalnu ekonomiju u regionu. Usklađuje regulative o digitalnim uslugama, protoku podataka i sajber sigurnosti, usklađujući ih sa EU *Digital Single Market*. Ova inicijativa promoviše ekonomski rast i inovacije te jača kolektivnu digitalnu sigurnost uspostavljanjem zajedničkih standarda za odgovornost platformi i zaštitu potrošača. Osigurava da sveprisutna prekogranična priroda AI-pokretanih dezinformacija naiđe na jedinstveno regulatorno okruženje, te tako jača "Digitalni imuni sistem" protiv malignog uticaja i olakšavajući put Zapadnog Balkana ka integraciji u EU. Dalje, potrebno je:

Narediti procjene uticaja AI za javni sektor: Nalaženje procjena uticaja AI (AIAs) za sve primjene AI u javnom sektoru širom Zapadnog Balkana je ključno za smanjenje rizika. To uključuje AI sisteme u javnim uslugama, izbornim procesima i širenju informacija. AIA identifikuje potencijalne algoritamske pristrasnosti, rizike diskriminacije i pojačavanje podjeljujućih narativa, posebno onih "prožetih" istorijskom traumom. Procjenjuje uticaj na osnovna prava, uključujući slobodu izražavanja i nediskriminaciju, te predlaže mjere ublažavanja prije implementacije. Ovaj korak uključuje ljudska prava u dizajn i rad AI sistema, sprječavajući nenamjerno perpetuiranje "Paradoksa političkog života."⁷⁴

Razviti standarde za AI reviziju i certifikaciju: Uspostavljanje čvrstih standarda za AI reviziju i certifikaciju širom Zapadnog Balkana stvara jasne smjernice za pouzdanu umjetnu inteligenciju. Ovi standardi usklađeni su s principima iz EU AI Act, pružajući okvir za nezavisnu reviziju AI sistema od strane trećih strana, posebno onih klasifikovanih kao visokorizični. Certifikacija osigurava da AI sistemi ispunjavaju određene kriterije za kvalitet podataka, tačnost, transparentnost i odgovornost. To uključuje rigorozno testiranje algoritamske pristrasnosti, naročito u pogledu etničkih i vjerskih osjetljivosti, kao i procjenu efikasnosti moderacije sadržaja za bosanski, srpski i hrvatski jezik. Takvi standardi suprotstavljaju se „zamagljivanju“ uticaja algoritama, nudeći provjerljivu garanciju odgovornog korištenja AI tehnologija.^{75 76 77}

⁷² <https://arxiv.org/html/2506.09992v1>

⁷³ <https://europeanwesternbalkans.com/2025/04/04/why-the-digital-services-act-is-needed-in-the-western-balkans-an-institutional-and-market-perspective/>

⁷⁴ Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

⁷⁵ Communications Regulatory Agency (CRA) Bosnia and Herzegovina, "GSR-25 Contribution: Best Practice Guidelines," predstavljeno na GSR-25 Konsultacijama, pristupljeno putem https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf

⁷⁶ Isto.

⁷⁷ Bojana Kostić and Caroline Sindors, "Responsible Artificial Intelligence: An overview of human rights' challenges of Artificial Intelligence and media literacy perspectives in the context of Bosnia and Herzegovina" (Council of Europe, juni 2022), pristupljeno putem <https://rm.coe.int/mil-study-3-artificial-intelligence-final-2759-3738-4198-2/1680a7cdd9>, Sekcija V.

Implementirati zahtjeve za algoritamsku transparentnost platformi: Zahtijevanje od tehnoloških kompanija da otkriju unutrašnje funkcionisanje svojih sistema predstavlja neophodan korak ka odgovornosti. Ova preporuka uspostavlja stroge zahtjeve za algoritamsku transparentnost platformi koje djeluju na Zapadnom Balkanu. Ti zahtjevi nalažu javno objavljivanje načina na koji sistemi za preporuku sadržaja rangiraju informacije, kako algoritmi za moderaciju detektuju i djeluju na štetni sadržaj (posebno za B/H/S jezike), kao i mehanizme ljudskog nadzora. Ova transparentnost omogućava istraživačima, civilnom društvu i regulatorima da identifikuju „algoritamske slijepe tačke“ i dijagnostikuju „jezične pukotine“, podstičući tako informisanu javnu raspravu i ciljane intervencije protiv AI-pokretane dezinformacije.⁷⁸⁷⁹

Uspostaviti nezavisne odbore za etiku AI: Uspostavljanje nezavisnih odbora za etiku AI na nacionalnom i regionalnom nivou pruža nepristrasan nadzor i smjernice. Ovi odbori, sastavljeni od multidisciplinarnih stručnjaka (uključujući etičare, pravne stručnjake, tehničke specijaliste i predstavnike mladih), nude nezavisnu reviziju i savjete o politici AI i specifičnim implementacijama. Njihov mandat uključuje procjenu društvenog uticaja AI, naročito na mlade i ranjive grupe, te savjetovanje o mjerama za sprječavanje „podmukle“ eksploatacije postkonfliktne traume i etničkih podjela. Ovaj nadzor jača povjerenje javnosti u AI tehnologije, posebno u regionu gdje je povjerenje u institucije često „oprezno“.⁸⁰

Kreirati mehanizme za pravnu zaštitu od algoritamskih šteta: Građani trebaju jasne puteve za osporavanje i ispravljanje „podmuklih“ AI odluka. Ova preporuka uspostavlja pristupačne mehanizme za pravnu zaštitu od algoritamskih šteta, uključujući pojednostavljene procedure za podnošenje žalbi, nezavisne usluge medijacije i pravne puteve kojima pojedinci mogu tražiti naknadu štete prouzrokovane pristrasnim ili diskriminatornim AI sistemima. Ovo je naročito važno za mlade, koji mogu biti disproportionalno pogođeni govorom mržnje ili dezinformacijama pokretanim AI, a nemaju adekvatne pravne mogućnosti. Time se osigurava da „regulatorna nevoljnost“ postojećih pravnih okvira ne ostavi građane nezaštićenim. Ovaj pristup gradi povjerenje javnosti i potvrđuje princip da AI tehnologije moraju odgovorno služiti društvu, posebno u kontekstima „prožetim“ istorijskim nepovjerenjem i sistemskim nepravdama.⁸¹

9.5 Dugoročne mjere za društvenu otpornost (fokus na mlade)

Podsticati kulturu kritičkog digitalnog angažmana od ranog obrazovanja: Dugoročna društvena otpornost na Zapadnom Balkanu zahtijeva integraciju sveobuhvatne digitalne i medijske pismenosti kao ključne kompetencije od ranog obrazovanja. Ovaj kurikulum prevazilazi osnovnu sigurnost na internetu, naglašava „algoritamsku svijest“ kako bi se mlade poučilo kako AI algoritmi personalizuju sadržaj i kako spoljašnji akteri iskorištavaju te algoritme za manipulaciju. Ključni elementi uključuju lateralno čitanje, imunizaciju protiv uobičajenih manipulativnih taktika, provjeru izvora i praktične vježbe triangulacije izvora. Ovaj pristup razvija promišljeno građanstvo, smanjuje podložnost „kognitivnom zarobljavanju“ AI-pokretanom dezinformacijom. On adresira „digitalnu gvozdenu zavjesu“ jer osnažuje mlade da kritički rastave algoritamski uticaj, pretvarajući ih u aktivne digitalne stražare.

Podržati istraživanja o uticaju AI na mentalno zdravlje mladih i polarizaciju: Dinamičan ekosistem AI-pokretane dezinformacije i „balkanske algoritamske pukotine“ zahtijeva ciljano istraživanje uticaja AI na mentalno zdravlje mladih i polarizaciju. Ova istraživanja pružaju empirijsku osnovu za intervencije. Trebalo bi ispitati kako algoritamsko pojačavanje podijeljenih narativa, naročito onih koji prizivaju istorijsku traumu i etničke tenzije, utiče na psihološko blagostanje, podstiče eho-komore i doprinosi radikalizaciji. Uključuju longitudinalne studije o izloženosti sintetičkim medijima i psihološkoj otpornosti mladih u multietničkim društvima. Takvi dokazi informišu razvoj „programa otpornosti na postkonfliktne traume pojačane AI“, osig-

78 <https://www.unesco.org/en/articles/unesco-supports-launch-coalition-freedom-expression-and-content-moderation-bosnia-and-herzegovina>

79 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 5. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

80 "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.5, stranica 34.

81 Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Patus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

uravajuci da su intervencije empirijski utemeljene i kulturološki osjetljive.

Promovisati međunacionalni i međuvjerski dijalog u digitalnim prostorima: Suočavanje s pojačavanjem podjela od strane AI algoritama zahtijeva promociju međunacionalnog i međuvjerskog dijaloga u digitalnim prostorima. Ova mjera razvija „Digitalni forum mladih“, konceptualno mjesto susreta mladih širom Zapadnog Balkana. Forum, olakšan tehnologijom, uključuje mlade u dijalog, kritičku analizu i zajedničko kreiranje kontranarativa. Podstiče međunacionalno razumijevanje i otpornost na štete koje uzrokuje AI. Forum podržava „Digitalnog diplomatu“, AI-pokretanu platformu dizajniranu da identifikuje i premosti informacijske podjele. Preporučuje različite perspektive i olakšava moderisane, međuetničke razgovore. Ovaj AI, eksplicitno treniran na „pro-socijalnim“ metrikama, prioritizuje sadržaj koji promoviše empatiju, kritičko razmišljanje i međusobno razumijevanje, aktivno tretira „balkansku algoritamsku pukotinu“ i suprotstavlja se „digitalnim ratovima sjećanja“.⁸²

Razviti programe otpornosti na postkonfliktne traume pojačane AI: Upornost uticaja postkonfliktnih trauma, pojačanih AI-pokretanim istorijskim revizionizmom, zahtijeva specijalizovane programe otpornosti. Ovi programi adresiraju psihološke posljedice AI-pojačanih podijeljenih narativa, naročito onih koji ciljaju etničke i vjerske osjetljivosti. Oni integrišu psihosocijalnu podršku s kritičkim digitalnim angažmanom, nude sigurne prostore i mentorski rad za mlade kako bi procesuirali traumu i razvili mehanizme suočavanja. Takve inicijative, ključni dio „digitalnog imunog sistema“, osnažuju mlade da kritički analiziraju istorijske narative na internetu, podstičući iscjeljenje umjesto perpetuiranja „digitalnih ratova sjećanja“. Inicijativa „Trauma-Informed Digital Storytelling for Reconciliation“ je ključni element, jer usmjerava mlade da kreiraju narative koji promovišu empatiju, međunacionalno razumijevanje i kritički pristup istorijskim prikazima.⁸³

Njegovati digitalno pripovijedanje za izgradnju mira među mladima: Osnaživanje mladih da kreiraju pozitivne kontranarative predstavlja ključnu dugoročnu mjeru za jačanje trajne društvene kohezije. To uključuje njegovanje digitalnog pripovijedanja za izgradnju mira, inicijativu koja oprema mlade vještinama za proizvodnju kratkih videozapisa, podcasta i interaktivnih web iskustava koja promovišu pomirenje i međunacionalno razumijevanje. Ovi programi su osjetljivi na traumu, prepoznajući dugotrajan uticaj prošlih sukoba. Oni vode mlade da konstruiraju narative koji podstiču empatiju i iscjeljenje, umjesto da perpetuiraju istorijski revizionizam ili nacionalističku propagandu. Ovaj napor gradi snažan rezervoar pozitivnog, ujedinjujućeg sadržaja koji direktno suprotstavlja „digitalnim ratovima sjećanja“ potpomognutim „balkanskom algoritamskom pukotinom“. Pretvarajući mlade u „digitalne graditelje mira“, ova mjera jača „digitalni imuni sistem“ proaktivnim, empatičnim sadržajem, osiguravajući da dinamičan sistem AI-pokretane dezinformacije bude odgovorena trajnim narativima jedinstva.⁸⁴

9.6 Sinteza nalaza i prioritizacija akcija⁸⁵⁸⁶

Uticaj AI na pojačavanje govora mržnje, ekstremizma i dezinformacija u Bosni i Hercegovini i širem Zapadnom Balkanu zahtijeva sažetu sintezu nalaza i rangirani set akcija. Izvještaj dokumentuje kako AI-pokretani procesi - manifestovani kao ono što nazivamo „balkanskom algoritamskom pukotinom“ i „ephemeralnom prirodom istine“ kreiranom sintetičkim medijima - iskorištavaju postojeće etničke podjele i istorijske zamjerke. Analiza identifikuje mlade kao visokorizičnu grupu: oni su disproporcionalno izloženi online sadržaji-

82 Namanja Sladaković i Milica Novaković, „Bosnia and Herzegovina,“ u IBA Alternative and New Law Business Structures Committee, juli 2024, pristupljeno putem <https://www.ibanet.org/medias/anlbs-ai-working-group-report-july-2024-4-bosnia-herzegovina.pdf?context=bWFzdGVyFB1YmXpY2F0aW9uUmVwb3J0c3w1NTQ3OHxhcHBsaWNhdGlvbi9wZGZ8YURJeUwyaGlaQzg1TVRN-NE1qQTJOREE0TnpNMEwyRnViR0p6TFdGcEYyZHZjbXRwYm1jdFozSnZkWEF0Y21Wd2IzSjBMV3AxYkhrdE1qQXIOQzAwTF-dKdmMyNXBZUzFvWihKNlpXZHkZkbWx1WVM1d1pHWXw4M2UzMmYzYTRIOGEwZWJiYTk1ZTkyYThkN2MxNGI1NmU4ZDF-hODUxOGY5ODQ0OTM1NWlZYTc1MmFODkwZGZlZlF0>

83 Communications Regulatory Agency (CRA) Bosnia and Herzegovina, „GSR-25 Contribution: Best Practice Guidelines,“ predstavljeno na GSR-25 Konsultaciji, pristupljeno putem https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf

84 Isto.

85 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 5. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

86 <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Looking Ahead: Generative AI.

ma i od njih zavisi kapacitet regiona za dugoročnu otpornost.⁸⁷⁸⁸⁹

Efektivni odgovor zahtijeva višeslojnu strategiju. Prvo, uskladiti zakonodavstvo na državnom i entitetskom nivou s evropskim standardima, naročito Digital Services Act (DSA), kako bi se smanjila pravna fragmentacija opisana u Poglavlju 3.1. Drugo, uspostaviti koordinirano tijelo za digitalno upravljanje BiH radi usklađivanja politika između institucija i postavljanja digitalno-specifičnih standarda za suzbijanje govora mržnje na internetu. Treće, ojačati pravne definicije i mehanizme sprovođenja za govor mržnje na internetu kako bi pravni lijekovi adresirali štete posredovane platformama.⁹⁰⁹¹⁹²

Ova analiza takođe dokumentuje praznine u odgovorima platformi i kapacitetima za moderaciju. Međunarodne tehnološke firme pokazuju ograničenu osjetljivost na jezičke i kontekstualne specifičnosti regiona, stvarajući prazninu u moderaciji za bosanski/srpski/hrvatski sadržaj. Politički i operativni odgovori moraju zatvoriti tu prazninu kroz ciljanu saradnju s dobavljačima, nabavku lokalizovanih kapaciteta za moderaciju i ulaganje u alate za detekciju specifične za jezik.⁹³

Prioriteti, poredani po vremenskom okviru, su sljedeći.⁹⁴⁹⁵

Hitno (0–12 mjeseci)

Usaglasiti zakonodavstvo i usvojiti digitalno-specifične izmjene zakona o govoru mržnje i medijima.

Izraditi Zakon o AI u BiH koji prilagođava ključne principe iz EU AI Act nacionalnim i entitetskim kontekstima.⁹⁶

Ojačati Agenciju za komunikacije (CRA) BiH tehničkim timovima za nadzor platformi i digitalne istrage.⁹⁷

Srednjoročno (12–36 mjeseci)⁹⁸

Pokrenuti Forum za digitalnu saradnju Zapadnog Balkana radi koordinacije prekograničnih politika, razmjene obavještajnih podataka o prijetnjama i usklađivanja standarda.

Razviti Regionalni sistem ranog upozoravanja za AI-pokretanu dezinformaciju kako bi se podržalo brzo otkrivanje i koordinisani odgovori.

Proširiti programe usmjerene na mlade: mreže za provjeru činjenica koje vode mladi, moduli za medijsku i informatičku pismenost, te kurikulum za svijest o algoritmima integrisan u škole i obuke civilnog društva.

87 <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Studija slučaja 2: Media Literacy Education, i Looking Ahead: Generative AI.

88 <https://www.idea.int/news/ethical-conundrum-electoral-ai-3>, "The ethical problem of electoral AI #3," International IDEA, 02. april 2025.

89 <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Studija slučaja 2: Media Literacy Education, Koliko se čini efikasnim?

90 Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

91 Council of Europe, Programme of the World Forum for Democracy 2024: Democracy and Diversity, 6-8 November 2024, Strasbourg, Stranica 16.

92 <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Studija slučaja 2: Media Literacy Education, Koliko se čini efikasnim?

93 Ibid., Studija slučaja 2: Media Literacy Education, Ključni zaključci.

94 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 5. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicali-

96 "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024," Regional Cooperation Council, decembar 2024, Sekcija 2.3.3, stranica 29.

97 Amel Muminovic and Amela Kadric Muminovic, "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages," arXiv preprint, 2025, Sekcija I-A, Stav 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

98 "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective", European Western Balkans, 4. april 2025.

Dugoročno (36+ mjeseci)⁹⁹

Osigurati održivo finansiranje putem EU Pre-Accession Instruments (IPA III) i ciljnih međunarodnih donatorskih programa za podršku otpornosti i izgradnji kapaciteta.¹⁰⁰

Podsticati javno-privatna partnerstva za tehnička rješenja, uključujući modele jezika i alate za moderaciju prilagođene lokalnim jezicima i dijalektima.¹⁰¹

Institucionalizovati razvoj radne snage za digitalno upravljanje, nadzor platformi i monitoring civilnog društva.¹⁰²¹⁰³

Strateška upotreba poluga EU integracije nudi praktičan put ka reformi na sistemskom nivou: usvajanje odredbi DSA i AI Act uskladiće tržišne podsticaje, povećati odgovornost platformi i pružiti tehničke standarde za provođenje. Dopunjavanje pravnih promjena operativnim poboljšanjima – lokalizovani kapacitet za moderaciju, unapređenje CRA, regionalna razmjena prijetnji i programi otpornosti mladih – pretvara regulaciju u mjerljive rezultate.

9.7 Mladi, saradnja i „neizbježan“ put naprijed

Mladi, uprkos izloženosti AI-pokretanoj manipulaciji, nisu samo pasivni primaoci digitalnih šteta. Oni služe kao agenti društvene otpornosti, sposobni formirati „Digitalni imuni sistem“ kroz namjerno osnaživanje. Ovaj pristup prepoznaje da mladi, duboko integrisani u „sveprisutni“ digitalni sistem, mogu aktivno prepoznavati, suprotstavljati se i uspostavljati odbrane protiv AI-pokretane dezinformacije i ekstremizma. Njihovo aktivno učešće je „neizbježan“ sastavni dio trajne demokratske budućnosti.¹⁰⁴¹⁰⁵

„Digitalni imuni sistem“ funkcioniše kroz međusobno povezane mreže mladih, obrazovanje vršnjaka i direktno uključivanje u kreiranje politika. Inicijative poput „Youth-Led Fact-Checking Networks“ omogućavaju aktivni nadzor online sadržaja, identifikaciju AI-pokretane dezinformacije i proizvodnju provjerenih kontra-narativa na lokalnim jezicima. Ove mreže direktno izazivaju „Digitalnu gvozdenu zavjesu“ jer podstiču kritičku procjenu algoritamski isporučenih informacija. „Algorithmic Awareness and Critical Thinking Curriculum“ obrazuje mlade o funkcionisanju AI algoritama, personalizaciji sadržaja i taktikama manipulacije, čime se jača otpornost protiv „kognitivnog zarobljavanja“.¹⁰⁶

„Linguistic Seam-Rip“ predstavlja značajnu ranjivost, jer AI sistemi za moderaciju često ne uspijevaju detektovati nijansirani govor mržnje na bosanskom, srpskom i hrvatskom jeziku. Ovaj nedostatak zahtijeva „Localized AI Moderation and Language Data Initiative“, vođenu od strane „The Acquis Digitalis“. Ova inicijativa nalaže tehnološkim kompanijama ulaganje u sveobuhvatne, kulturno nijansirane skupove podataka i ljudsku jezičku ekspertizu, direktno adresirajući prazninu u moderaciji koja ostavlja mlade ranjivim na suptilan, kulturološki ukorijenjen govor mržnje.

Snazna regionalna saradnja i povećana međunarodna podrška su „neizbježni“ za suprotstavljanje

⁹⁹ <https://www.kdz.eu/system/files/downloads/2025-04/AI%20at%20local%20level.pdf>, „Artificial Intelligence in Local Government: Driving Innovation, Bridging Gaps, and Shaping the Digital Transition in the Western Balkans and Moldova“, KDZ – Centre for Public Administration Research, NALAS, april 2025, str. 7.

¹⁰⁰ Ibid., str. 10.

¹⁰¹ Amel Muminovic and Amela Kadric Muminovic, „Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages“, arXiv preprint, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

¹⁰² Communications Regulatory Agency (CRA) Bosnia and Herzegovina, „GSR-25 Contribution: Best Practice Guidelines“, predstavljeno na GSR-25 Konsultacijama, pristupljeno putem https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf

¹⁰³ Isto.

¹⁰⁴ <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Studija slučaja 2: Obrazovanje o medijskoj pismenosti, i Pogled unaprijed: Generativna AI.

¹⁰⁵ Isto.

¹⁰⁶ Conclusion Paper: Online radicalisation and P/CVE approaches in the Western Balkans, published by the Radicalisation Awareness Network (RAN), European Commission, based on a paper prepared by Nejra Veljan after consultation with Fenna Canters and Rositsa Dzhekova (RAN Staff). https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf (Stranica 4)

„razuzdanoj“ digitalnoj plimi. Kolektivna akcija, osjetljiva na jezičke nijanse i istorijske traume, predstavlja jedini održivi bedem. „Regional Digital Diplomacy and Counter-Narrative Hub“ funkcioniše kao kolaborativna platforma koja okuplja omladinske aktiviste, medijske profesionalce i digitalne eksperte širom Zapadnog Balkana. Ovaj centar razvija i širi pozitivne, ujedinjujuće kontra-narative koji direktno izazivaju AI-pojačani govor mržnje i ekstremizam, posebno sadržaje „prožete“ istorijskim revizionizmom.¹⁰⁷

Inicijative „Trauma-Informed Digital Storytelling for Reconciliation“ osnažuju mlade da kreiraju digitalne priče koje promovišu empatiju i međunacionalno razumijevanje, direktno se suprotstavljajući „Digitalnim ratovima sjećanja“ koje potiče algoritamska pristrasnost. Ovi programi, svjesni trajnog uticaja prošlih sukoba, vode mlade da grade narative koji podstiču iscjeljenje. Sveobuhvatni okvir povezan je sa vizijom o ojačanom Zapadnom Balkanu, gdje su demokratski procesi otporni, a građani, posebno mladi, opremljeni za sigurno snalaženje u „sveprisutnom“ digitalnom pejzažu. Nepokolebljiva posvećenost EU integraciji pruža „neizbježnu“ polugu za provođenje ovih reformi, osiguravajući da regija pređe iz ranjive granice u model digitalne otpornosti.¹⁰⁸

Dodatak A: Rječnik pojmova

A.1 Definicija ključnih pojmova

Acquis Digitalis: Ovaj koncept tvrdi da proces pristupanja EU pruža inherentnu vanjsku polugu za sveobuhvatnu reformu digitalnog upravljanja širom Zapadnog Balkana. Zalaže se za strateško usvajanje i harmonizaciju sa Digital Services Act (DSA) i nadolazećim AI Act, smatrajući ih ključnim alatima za nacionalnu sigurnost i konsolidaciju demokratije.

Algorithmic Amplification: Mehanizam putem kojeg AI sistemi, kroz mehanizme preporuke sadržaja i personalizaciju, povećavaju vidljivost i doseg određenog sadržaja. Ovo često nenamjerno promovise podijeljene ili štetne narative.

Algorithmic Bias: Sistematske i ponavljajuće greške u izlazu AI sistema koje proizlaze iz pogrešnih pretpostavki u procesu mašinskog učenja. Do ovoga dolazi kada podaci za obuku pogrešno predstavljaju populaciju ili sadrže ugrađene društvene predrasude, što vodi do iskrivljenih ili nepravednih rezultata.

AI Hallucinations: Netočne informacije generisane od strane AI chatbota ili velikih jezičkih modela, predstavljene kao činjenice iako su izmišljene. One utiču na povjerenje javnosti i demokratske procese.¹⁰⁹

Balkan Algorithmic Fracture: Ovaj koncept opisuje kako sistemi za preporuku sadržaja i algoritmi personalizacije AI-a iskorištavaju i intenziviraju postojeće etničke, vjerske i političke podjele unutar BiH i Zapadnog Balkana. Algoritmi, optimizovani za angažman, stvaraju tendenciozne eho-komore koje jačaju već postojeće predrasude i istorijske zamjerke.¹¹⁰

Code-Switching: Praksa tečnih govornika da izmjenjuju dva ili više jezika ili dijalekata unutar jedne konverzacije ili teksta. Ovo mijenja kontekst i namjeru naizgled bezazlenih fraza, predstavljajući izazove za AI moderaciju.¹¹¹

¹⁰⁷ Council of Europe, Programme of the World Forum for Democracy 2024: Democracy and Diversity, 6-8 November 2024, Strasbourg, Lab 1, str. 14.

¹⁰⁸ Isto.

¹⁰⁹ “‘Perfidious’ priroda manipulacije vođene AI leži u njenoj sposobnosti da oponaša autentičnost dok sije duboko ukorijenjeno nepovjerenje.”

¹¹⁰ “Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective”, European Western Balkans, 4. april 2025.

¹¹¹ Amel Muminovic and Amela Kadric Muminovic, “Large Language Models for Toxic Language Detection in Low-Resource Balkan

Content Recommendation Engines: AI sistemi koji korisnicima predlažu sadržaj na osnovu njihove prethodne aktivnosti, preferencija i interakcija. Ovi sistemi prioritet daju angažmanu, što može stvoriti „filter bubbles“ pojačavajući postojeća uvjerenja.

Deepfakes: Medijski sadržaji, uključujući slike, audio ili video, generisani ili manipulirani pomoću AI-a kako bi se stvorile realistične, ali lažne reprezentacije. To uključuje mijenjanje lica u video zapisima, generisanje realističnih glasovnih klonova ili kreiranje potpuno umjetnih događaja.

Digital Services Act (DSA): Sveobuhvatna regulativa EU koja uspostavlja pravni okvir za online platforme i digitalne usluge. Stvara sigurniji digitalni prostor, štiti temeljna prava korisnika i osigurava odgovornost platformi.

Digital Divide of Governance: Izazovi koji proizlaze iz fragmentiranih pravnih i institucionalnih struktura Bosne i Hercegovine, što rezultira razlikama i neefikasnostima u razvoju i implementaciji politika digitalnog upravljanja, posebno u vezi sa AI.

Digital Immune System: Ovaj koncept pozicionira mlade ne samo kao primaoc digitalne pismenosti, već kao aktivne, samoodržive komponente društvene odbrane protiv šteta koje uzrokuje AI. Osnaživanje mladih naprednim medijskim pismenostima i kritičkim razmišljanjem pretvara ih u prvu liniju odbrane protiv manipulacija vođenih AI-jem, sposobne da generišu kontranarative i podstiču kolektivnu otpornost.¹¹²

Digital Iron Curtain: Ovaj koncept ističe akutnu ranjivost mladih, primarne demografske grupe izložene štetama koje uzrokuje AI. Sveprisutnost društvenih mreža, u kombinaciji sa niskom medijskom pismenošću, stvara „digitalnu željeznu zavjesu“ koja ograničava izloženost različitim perspektivama i povećava podložnost kognitivnom zarobljavanju ekstremističkim narativima.

Disinformation: Lažne, netačne ili obmanjujuće informacije namjerno dizajnirane, predstavljene i promovisane da izazovu štetu javnosti ili ostvare profit.¹¹³

Entity-Level: Odnosi se na administrativne jedinice unutar Bosne i Hercegovine (npr. Republika Srpska, Federacija BiH, Brčko Distrikt) i njihove odgovarajuće vladine i pravne nadležnosti.

EU AI Act: Pionirska regulativa EU koja osigurava sigurnu upotrebu AI sistema i uspostavlja jasan pravni okvir. Koristi pristup zasnovan na riziku, uvodeći proporcionalna, efikasna i obavezujuća pravila za AI sisteme.¹¹⁴

Extremism: Zalaganje ili praksa ekstremnih političkih ili vjerskih stavova, često karakterizirana netolerancijom prema suprotnim stavovima i korištenjem nasilja ili nezakonitih sredstava za postizanje ciljeva.

Hate Speech: Javno poticanje mržnje, diskriminacije ili nasilja prema pojedincu ili grupi na osnovu karakteristika kao što su porijeklo, religija ili etnička pripadnost.¹¹⁵

Hybrid Threats: Sofisticirane kampanje koje kombinuju konvencionalne i nekonvencionalne taktike, poput sajber-napada, kampanja dezinformacija i ekonomske prinude, s ciljem destabilizacije vlada i iskorištavanja društvenih podjela.

Languages,” arXiv preprint arXiv:2506.09992, 2025, Sekcija I-A, Pasus 4. Dostupno na: <https://arxiv.org/html/2506.09992v1>

112 Namanja Sladaković i Milica Novaković, “Bosnia and Herzegovina,” u IBA Alternative and New Law Business Structures Committee, juli 2024, pristupljeno putem <https://www.ibanet.org/medias/anlbs-ai-working-group-report-july-2024-4-bosnia-herzegovina.pdf?context=bWFzdGVyfFB1YmtpY2F0aW9uUmVwb3J0c3w1NTQ3OHxhcHBsaWNhdGlvbi9wZGZ8YURJeUwyaGlaQzg1TVRN-NE1qQTJOREE0TnpNMEwyRnViR0p6TFdGcExYZHZjbXRwYm1jdFozSnZkWEF0Y21Wd2lZSjBMV3AxYkhrdE1qQXIOQzAwTF-dKdmMyNXBZUzFvWlhKNpXZHkZkbWx1WVM1d1pHWXw4M2UzMmYzYTROGEwZWJiYTk1ZTkYThkN2MxNGI1NmU4ZDF-hODUxOGY5ODQ0OTM1NWlZYTc1MmFiodKwZGZlQ==>

113 “From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024,” Regional Cooperation Council, decembar 2024, Sekcija 2.3.3, stranica 29.

114 Radicalisation Awareness Network (RAN). (2022). Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper. European Commission. Stranica 5. https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf

115 “From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024,” Regional Cooperation Council, decembar 2024, Sekcija 2.3.5, stranica 34.

Linguistic Seam-Rip: Kritičan nedostatak u sposobnostima AI za moderaciju sadržaja na bosanskom/srpskom/hrvatskom jeziku. Ovo proizlazi iz jezičkih složenosti, kulturnih nijansi i statusa ovih jezika kao „low-resource languages“, što omogućava AI sistemima da ne prepoznaju sofisticirani govor mržnje i dezinformacije.

Low-Resource Languages: Jezici za koje postoji oskudnost digitalnih podataka (npr. tekstualni korpusi, označeni skupovi podataka) dostupnih za obuku modela umjetne inteligencije, čime se otežava efikasnost automatiziranih alata za obradu jezika.

Obrada Prirodnog Jezika (NLP) za Moderaciju: Grana AI koja omogućava računarima da razumiju, interpretiraju i generišu ljudski jezik. U moderaciji sadržaja, NLP modeli analiziraju tekst radi štetnih ključnih riječi, sentimenta i konteksta, ali često imaju poteškoće sa nijansama jezika sa ograničenim resursima, žargonom ili kulturološki specifičnim pogrdnim izrazima.

Odgovornost Platforme: Pravna i etička odgovornost online platformi (npr. kompanija društvenih medija) da se bave štetnim sadržajem, osiguraju transparentnost u svom radu i ublaže sistemske rizike na svojim uslugama.¹¹⁶

Postkonfliktno Društvo: Društvo koje prolazi kroz oporavak i rekonstrukciju nakon perioda oružanog sukoba. Takva društva često se suočavaju sa neriješenim istorijskim pritužbama, etničkim podjelama i krhkim demokratskim institucijama, što ih čini posebno ranjivim na manipulaciju informacijama.¹¹⁷¹¹⁸

Reference

Atlantic Council. "How the US and Europe Can Counter Russian Information Manipulation About Nonproliferation." *Atlantic Council*, October 4, 2024.

<https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/how-the-us-and-europe-can-counter-russian-information-manipulation-about-nonproliferation/>.

Atlantic Council, in collaboration with the US Department of State's Office of Cooperative Threat Reduction. "Issue Brief: How the US and Europe Can Counter Russian Information Manipulation About Nonproliferation." *Atlantic Council*, October 4, 2024.

<https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/how-the-us-and-europe-can-counter-russian-information-manipulation-about-nonproliferation/>.

Bradshaw, Samantha, and Philip N. Howard. "The Global Disinformation Disorder: 2019 Global Inventory of Organised Social Media Manipulation." *Working Paper 2019.2*. Oxford, UK: Project on Computational Propaganda, Oxford Internet Institute, University of Oxford, 2019.

<https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2019/09/CyberTroop-Report19.pdf>.

Carnegie Endowment for International Peace. "Countering Disinformation Effectively: An Evidence-Based Policy Guide." January 31, 2024.

<https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>.

¹¹⁶ Conclusion Paper: Online radicalisation and P/CVE approaches in the Western Balkans, published by the Radicalisation Awareness Network (RAN), European Commission, based on a paper prepared by Nejra Veljan after consultation with Fenna Canters and Rositsa Dzhekova (RAN Staff). https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf (Stranica 3)

¹¹⁷ Ibid., Stranica 5.

¹¹⁸ <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en>, Studija slučaja 2: Media Literacy Education, i Looking Ahead: Generative AI.

Communications Regulatory Agency (CRA) Bosnia and Herzegovina. "GSR-25 Contribution: Best Practice Guidelines." Presented at the GSR-25 Consultation. Accessed [Date of Access].

https://www.itu.int/itu-d/meetings/gsr-25/wp-content/uploads/sites/33/2025/06/GSR-25_Contribution_Best-Practice-Guidelines_RAK-Bosnia-and-Herzegovina.pdf.

Council of Europe. "Programme of the World Forum for Democracy 2024: Democracy and Diversity." Strasbourg, 6-8 November 2024.

<https://rm.coe.int/programme-of-the-world-forum-for-democracy-2024-democracy-and-diversit/1680b-11b3a>.

European Commission. "Horizon Europe Work Programme 2023-2025 - 5. Culture, Creativity and Inclusive Society." Decision C(2024) 2371 of 17 April 2024.

https://research-and-innovation.ec.europa.eu/document/download/6e3460b4-dd63-4555-8acb-3f999f638624_en.

European External Action Service (EEAS). "3rd EEAS Report on Foreign Information Manipulation and Interference Threats: Exposing the architecture of FIMI operations." March 2025.

<https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3nd-ThreatReport-March-2025-05-Digital-HD.pdf>.

European Parliament. "Resolution of 1 June 2023 on foreign interference in all democratic processes in the European Union, including disinformation (2022/2075(INI))." P9_TA(2023)0219.

https://www.europarl.europa.eu/doceo/document/TA-9-2023-0219_EN.pdf.

European Western Balkans. "Why the Digital Services Act is needed in the Western Balkans: An institutional and market perspective." April 4, 2025.

<https://europeanwesternbalkans.com/2025/04/04/why-the-digital-services-act-is-needed-in-the-western-balkans-an-institutional-and-market-perspective/>.

International Bar Association (IBA). "Bosnia and Herzegovina." In *IBA Alternative and New Law Business Structures Committee*, July 2024.

<https://www.ibanet.org/medias/anlbs-ai-working-group-report-july-2024-4-bosnia-herzegovina.pdf?context=bWFzdGVyfFB1YmV3J0c3w1NTQ3OHxhcHBsaWNhdGlvbi9wZGZ8YURJeUwyaGlaQzg1TVRNNE1qQTJOREE0TnpNMEwyRnViR0p6TFdGcExYZHZjbXRwYm1jdFozSnZkWEF0Y-21Wd2lzSjBMV3AxYkhrdE1qQXIOQzAwTFdKdmMyNXBZUzFvWlhKNlpXZHZkbWx1WVM1d1pHWXw-4M2UzMmYzYTRIOGEwZWJiYTk1ZTkYThkN2MxNGI1NmU4ZDFhODUxOGY5ODQ0OTM1NWlZYTc-1MmFIODkwZGZlZQ==>.

International IDEA. "The ethical question of electoral AI #3." April 02, 2025.

<https://www.idea.int/news/ethical-conundrum-electoral-ai-3>.

International Telecommunication Union (ITU) Office for Europe. "Digital Development Country Profile Bosnia and Herzegovina" (final - March 2023).

<https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/Publications/2023/Digital%20Development%20Country%20Profile%20Bosnia%20and%20Herzegovina%20%5Bfinal-%20March%202023%5D.pdf>.

International Telecommunication Union (ITU) Europe Region. "ITU Europe Region Initiatives 2023-2025."

<https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Pages/Regional%20Initiatives/2022/ITU-Europe-Region-Initiatives-2023-2025.aspx>.

Kostić, Bojana, and Caroline Sindere. "Responsible Artificial Intelligence: An overview of human rights' challenges of Artificial Intelligence and media literacy perspectives in the context of Bosnia and Herzegovina." *Council of Europe*, June 2022. Accessed via

<https://rm.coe.int/mil-study-3-artificial-intelligence-final-2759-3738-4198-2/1680a7cdd9>.

KDZ – Centre for Public Administration Research, and NALAS (Network of Associations of Local Authorities of South-East Europe). "Artificial Intelligence in Local Government: Driving Innovation, Bridging Gaps, and Shaping the Digital Transition in the Western Balkans and Moldova." April 2025.

<https://www.kdz.eu/system/files/downloads/2025-04/AI%20at%20local%20level.pdf>.

Muminovic, Amel, and Amela Kadric Muminovic. "Large Language Models for Toxic Language Detection in Low-Resource Balkan Languages." *arXiv preprint arXiv:2506.09992*, 2025. Available at:

<https://arxiv.org/html/2506.09992v1>.

Open Society Institute – Sofia. "Media Literacy Index 2021." March 14, 2021.

<https://osis.bg/?p=3750&lang=en>.

Public Media Alliance. "How Western Balkans public media empowered youth to tell fake from fact." April 2, 2025.

<https://www.publicmediaalliance.org/how-western-balkans-public-media-empowered-youth/>.

Radicalisation Awareness Network (RAN), European Commission. "Online radicalisation and P/CVE approaches in the Western Balkans: Conclusion Paper." June 16, 2022.

https://home-affairs.ec.europa.eu/system/files/2023-05/ran_paper_online_radicalisation_p-cve_approaches_in_wb_16062022_en.pdf.

Regional Cooperation Council (RCC). "Common Regional Market Action Plan 2025-2028 (CRM 2.0)."

<https://www.rcc.int/files/user/docs/0e5e72bb8334509a1feb954cdc7a3e54.pdf>.

Regional Cooperation Council (RCC) Secretariat. "DRAFT REPORT ON THE ACTIVITIES OF THE REGIONAL COOPERATION COUNCIL SECRETARIAT For the period 01 October 2024 – 28 February 2025."

<https://www.rcc.int/files/user/docs/a9b1d91ff97948db2d1e8adf2d6a2662.pdf>.

Regional Cooperation Council (RCC). "From Dialogue to Action: Working Group Outcomes and Recommendations from the 9th Regional Security Coordination Conference 2024." December 2024.

<https://www.rcc.int/download/docs/WWG-Outcomes-and-Recommendations-from-the-9th-Regional-Security-Coordination-Conference-2024-2.pdf/3121b2bcb36c0dd085d11d1826b961c.pdf>.

Safer Internet Center Bosnia and Herzegovina. "Sigurnodijete.ba - Home Page." Accessed June 12, 2024.

<https://www.sigurnodijete.ba/en/>.

UNESCO. "UNESCO Supports Launch of Coalition for Freedom of Expression and Content Moderation in Bosnia and Herzegovina." June 20, 2023.

<https://www.unesco.org/en/articles/unesco-supports-launch-coalition-freedom-expression-and-content-moderation-bosnia-and-herzegovina>.

World Economic Forum. "The Global Risks Report 2024." January 2024.

https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf.

World Economic Forum. "The Global Risks Report 2025." January 2025.

https://reports.weforum.org/docs/WEF_Global_Risks_Report_2025.pdf.

Other Useful Links and Research Guides

• **Council of Europe Resources on AI and Human Rights:** Explore the Council of Europe's work on artificial intelligence and its impact on human rights, democracy, and the rule of law for a broader understanding of the European regulatory context.

<https://www.coe.int/en/web/artificial-intelligence>.

• **EU Digital Services Act (DSA) Information:** Access official EU documentation and explanatory materials on the Digital Services Act to understand its scope, obligations, and enforcement mechanisms.

<https://digital-strategy.ec.europa.eu/en/policies/digital-services-act>.

• **EU AI Act Information:** Consult official EU resources detailing the Artificial Intelligence Act, its risk-based approach, and its implications for AI deployment across member states and potential candidate countries.

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689.

• **Open Society Foundations - Media Literacy Initiatives:** Explore the work of the Open Society Foundations on media literacy, disinformation, and digital rights, which often includes regional analyses and reports relevant to the Western Balkans.

<https://www.opensocietyfoundations.org/policy-and-advocacy/media-and-internet>.

• **Regional Cooperation Council (RCC) Digital Agenda:** Review the RCC's ongoing initiatives and strategy documents concerning the Digital Agenda for the Western Balkans to understand regional cooperation efforts and alignment with EU digital policies.

https://www.rcc.int/priority_areas/65/digital-transformation.

• **ITU Child Online Protection (COP) Guidelines:** Access the International Telecommunication Union's global guidelines and resources for child online protection, which provide a framework for national adaptation and implementation.

<https://www.itu-cop-guidelines.com/>.

- **Global Risks Report (World Economic Forum):** Consult the World Economic Forum's annual Global Risks Report for broad insights into emerging societal, technological, and geopolitical risks, including those related to AI and disinformation.

<https://www.weforum.org/reports/the-global-risks-report-2025>.

- **arXiv.org - NLP and AI Research:** Explore arXiv for the latest pre-print research in Natural Language Processing (NLP) and AI, particularly studies focusing on low-resource languages and toxic language detection, such as the work by Muminovic and Kadric Muminovic.

<https://arxiv.org/list/cs.CL/recent>.

- **Atlantic Council - Digital Forensic Research Lab (DFRLab):** Review DFRLab's analyses for insights into disinformation campaigns, manipulated media, and the operational tactics of state and non-state actors in the digital space.

<https://www.atlanticcouncil.org/programs/digital-forensic-research-lab/>.



Ovo istraživanje je provedeno u sklopu projekta AIKoristiSPaznjom, uz finansijsku podršku Vlade Kraljevine Nizozemske kroz program Matra. Izneseni stavovi i zaključci pripadaju isključivo udruženju JaBiHEU i autoru istraživanja i ne odražavaju nužno stavove donatora.