

*Тања Казих**

Центар за стратешке прогнозе, Београд, Република Србија

ДИГИТАЛНА ПРОПАГАНДА И ДЕЗИНФОРМАЦИЈЕ ГЕНЕРИСАНЕ ВЕШТАЧКОМ ИНТЕЛИГЕНЦИЈОМ: СТУДИЈЕ СЛУЧАЈА ИЗРАЕЛСКО- ПАЛЕСТИНСКОГ СУКОБА И ПАДА БАШАРА АЛ-АСАДА У СИРИЈИ

Сажетак

У овом раду, ауторка ће анализирати улогу дигиталне пропаганде и дезинформација генерисаних вештачком интелигенцијом (ВИ) у текућим блискоисточним конфликтима, са фокусом на ескалацију сукоба на релацији Израел-Палестина, те на свргавање председника Сиријске Републике Башара ал-Асада (*Bashar al-Assad*). У овом географском оквиру, ауторка ће истражити на који начин технологије засноване на ВИ, а номинално генерисање дипфејк (*deepfake*) материјала, има употребну вредност у зонама конфликта, трасирајући прве кораке у коришћењу оваквих тактика у ратне сврхе у актуелном руско-украјинском сукобу. Потврђујући присутност и утицај овог типа сајбер операција, ауторка ће настојати да, спровођењем анализе две студије случаја, потврди хипотезу да савремена дигитална пропаганда, потпомогнута и генерисана ВИ, значајно усложњава ток текућих дешавања, доводећи до неминовне ескалације сукоба и отежавајући разграничење истине од лажи, како у војном, тако и у политичком контексту. Користећи се методом анализе садржаја, са фокусом на идентификовање инстанци пласирања дигиталне

* Имејл адреса: tanja.z.milosevic@gmail.com; ORCID: 0009-0007-1091-6695.

пропаганде генерисане ВИ, приступиће се свеобухватној анализи и студији појединачних случаја Израела/Палестине и Сиријске Републике, ради креирања оквира могућности и утицаја таквог садржаја на класичне оружане сукобе, али и на политичке кризе на нивоу Блиског истока.

Кључне речи: вештачка интелигенција, дигитална пропаганда, дезинформације, машинско учење, дубоко учење, *deepfake*, Блиски исток, Израел/Палестина, Сирија

УВОД

Сагледавајући историју човечанства и токове ратова које су обележили свет, приметно је да свако ратно стање наводи на изналажење нових решења за старе проблеме, на тај начин иницирајући развој технологије и рађање нових идеја. Примера ради, Први светски рат обележиле су три кључне промене парадигме ратовања, и то: индустријализација, увођење треће димензије ратовања кроз употребу летелица, те омогућавање дејствовања по непријатељу по дубини (Zabecki 2015, 2–3). Насупрот томе, Други светски рат заувек ће бити обележен развојем и употребом атомске бомбе, али и радарске технологије, која је осигурала могућност ране детекције, а самим тим и раног одговора на потенцијални напад непријатељских снага (Genuth 1988, 276). Након два светска рата и заједничких напора међународне заједнице ка очувању мира, поларизација света и присуство два снажна блока – Западног и Источног, условила је развијање техника у другом правцу, фокусирајући се на субверзивно деловање, унапређење обавештајног рада, али и методологију прикривеног учешћа у сукобу, као што су ратови преко посредника (*proxy wars*). У оваквој констелацији одржавања привидног мира, начињени су и први помаи у домену психолошког ратовања, а нарочито електронског и сајбер ратовања, усмереног на континуирану дестабилизацију непријатељских снага. Стога, природно можемо закључити да су пионири академске заједнице, мотивисани идејом прејудицирања појавног облика наредног великог рата, потенцијални Трећи светски рат конципирали управо у оквирима сукобљавања „из даљине”, кроз употребу напредне технологије из домена информационаих технологија. У складу са

тим, седамдесте године двадесетог века, канадски теоретичар медија Маршал Маклуан (*Marshal McLuhan*) даје пророчански суд о форми Трећег светског рата, предвидевши да ће то бити „герилски информациони рат, у којем неће постојати подела између војних и цивилних учесника” (Forest 2021, 13).

У духу претходног идентификовања кључних промена парадигми ратовања током прва два светска рата, а сагледавајући текући руско-украјински сукоб, који су савремени политиколози окарактерисали као „критичну раскрсницу”, „тачку без повратка” и „увртиру у Трећи светски рат”, потврђујемо идеју да ратно стање отвара простор и мотивише струку ка проналажењу нових решења за старе проблеме. Оно што знатно доминира руско-украјинским бојиштем, а уједно и руско-украјинским небом, јесу свакако беспилотне летелице (БПЛ), употреба сателита, хибридно ратовање и извођење сајбер напада, као и употреба алата вештачке интелигенције, што све потпада под ширу категорију „информационих технологија”, у складу са пророчанством академских промишљања с краја 20. века. Стога, уколико бисмо руско-украјински сукоб заиста третирали као потенцијалну увртиру у Трећи светски рат, могли бисмо да потврдимо да се Маклуаново пророчанство остварује, те у овом „информационом рату” примећујемо и нове факторе дејствовања, попут присуства напредних БПЛ и употребе алата ВИ. Наведена два, сходно томе, идентификујемо као сегменте вредне детаљнијег изучавања, при чему главни фокус у овом истраживању стављамо на перспективу употребе алата ВИ.

Примат изучавању примене ВИ уместо БПЛ у овом раду нипошто не значи да наведена технологија има мању вредност. Напротив, као сектор у развоју, производња БПЛ и даље доживљава значајна унапређења, тиме пружајући учесницима у сукобима разнолике погодности. Међутим, постојање знатне грађе о овој теми, као и досадашње анализе праксе и искустава употребе БПЛ на ратиштима, већ попуњавају академски вакуум који је раније постојао у овој области. Наиме, изразита употреба БПЛ на територијама Украјине и Руске Федерације (РФ) очекивана је и предвидљива, имајући у виду да је у пост-совјетском географском простору, којем припадају и ове две државе, већ забележена таква пракса током Другог рата у Нагорно Карабаху током јесени 2020. године. Раније присутне у Либији,¹ на сиријском ратишту, у редовима „Хезболаха”

¹ Присутан је став да је турска наменска индустрија на случају Либије, који се врло често у академској и стручној заједници назива и „ратом дронова”, тестирала

(*Hezbollah*), као и у рукама побуњеничке групе Хути (*Houthi*) у Јемену (видети Ђогић и Казич 2021), и др., за дејствовање БПЛ током рата у Нагорно Карабаху заиста није била новина у сфери ратовања, али је, ипак, представљало својеврсну прекретницу у модалитету употребе овог типа борбеног система у зони сукоба. Наиме, током овог сукоба, забележена је употреба БПЛ у пропагандне сврхе, имајући у виду да је нарочито Баку јавно приказивао снимке успешног дејствовања својих БПЛ по јерменским циљевима, подижући тиме, пре свега, морал свог становништва. Сличне технике пропагандног деловања бележе се и на руско-украјинском ратишту, уз чињеницу да су Украјина и Русија у том контексту начиниле и корак даље, пласирањем *deepfake* видео записа председника непријатељских снага у пропагандне сврхе, привлачећи пажњу на ову нову праксу.

Само 20 дана након отпочињања специјалне војне операције РФ у Украјини, 16. марта 2022. године, на хакованој страници украјинске телевизијске станице *Ukrayina 24*, објављен је видео запис председника Украјине Владимира Зеленског (*Володимир Зеленски*) у обраћању украјинској нацији (Pearson and Zinets March 17, 2022; Majchrzak 2023, 49). На први поглед, приказ Зеленског делује уверљиво, обележен детаљима карактеристичним за обраћања Зеленског од почетка рата – војничка мајица, минималистичка позадина, емотивни говори, директно обраћање народу. Ипак, након слушања, већина Украјинаца одбацила је наведени видео запис као лажан, што је потврдило и Председништво Украјине у рекордном року. Имајући у виду садржај говора – који је Украјинце позивао на предају и капитулацију, индикативно је то да је брза детекција лажираног записа спречила знатне губитке на украјинском ратишту, с обзиром на то да би поштовање наређења врховног команданта оружаних снага Украјине, у најмању руку, проузроковало напуштање положаја и отварање пута руским снагама ка освајању стратешких позиција. Упркос неуспеху овог пропагандног напада, наведени *deepfake* указао је на нову употребну вредност вештачке интелигенције у сукобима, прогнозирајући њихову све чешћу појаву у сукобима који следе. У том контексту, стручњаци у домену теорије примене ВИ истичу да, иако је руско-украјински сукоб интензивирао развој ВИ

операбилност своје БПЛ *Bayraktar TB2*, а током Другог рата у Нагорно Карабаху додатно усавршила технике ратовања уз помоћ ове БПЛ, постижући ненадмашив успех и привлачећи интересовање многих држава широм света за овај производ. Видети Al-Dhahab 2019.

у ратне сврхе, њихов пуни капацитет биће достигнут у наредним сукобима, а нарочито у израелско-палестинском сукобу (Eisele and Steinwehr 2023).

У овом контексту, значајно је нагласити да је наведени *deepfake* видео запис креиран као одговор на украјински покушај подривања морала руских снага задејствованих на непријатељској територији нешто раније, а у оквиру прве две недеље сукоба.² Наведени *deepfake* видео запис, који је добио медијску пажњу тек након лажираног видео записа Зеленског, приказује председника Руске Федерације Владимира Путина (*Владимир Владимирович Путин*) у обраћању руској нацији, у којем проглашава прекид размене ватре, објављује почетак мировних преговора са Украјином, те проглашава независност Крима као републике унутар граница Украјине (Baig 2022). Појавивши се на друштвеној мрежи *Twitter*, иницијално није привукао значајну пажњу услед чињенице да је сам материјал био видно оштећен манипулацијама, али и да је сам видео запис визуелно био познат широј јавности, насупрот видео запису Зеленског, који је опонашао и био идентичан униформисаним појављивањима украјинског председника с почетка сукоба.

Упркос неуспесима ове две иницијативе, с правом можемо тврдити да је напредак у развоју ВИ, а нарочито машинског учења као алата из спектра аналитичке вештачке интелигенције (видети Đorić i Kazić 2021), који омогућава конструисање *deepfake* садржаја, уврстио овај метод манипулације у будуће алате информационог ратовања и психолошких операција. То нас наводи да овом новом феномену посветимо нарочиту академску пажњу, не само у контексту разумевања технике конструисања оваквих материјала, већ и ради разумевања и идентификације њиховог присуства у другим актуелним сукобима и политичким кризама, као што су текући сукоб Израела и Палестине, те свргавање председника Сиријске Републике Башара ал-Асада и последично интензивирање политичко-безбедносне кризе у земљи, који ће у овом раду бити коришћени као студије случаја. Одабир наведених случајева оправдавам чињеницом да је прва студија војног карактера, док друга презентује елементе

² Процењује се да је наведени видео запис настао након 21. фебруара 2022. године, с обзиром на то да је техником *deepfake* измењен садржај говора Путина, док је видео запис остао нетакнут, и представља снимак обраћања Владимира Путина након признавања независности Доњецка и Луганска, два сепаратистичка региона на југу Украјине. Видети Baig 2022.

политичког карактера, омогућавајући генерисање студија случајева из истог географског региона, али у различитим контекстима, а све у циљу потврђивања хипотезе да такав вид пропаганде има тенденцију да ескалира постојеће сукобе и кризе, отежавајући разграничење истине од лажи.

За потребе израде теоријског оквира истраживања, кључни термини – пропаганда, дигитална пропаганда, информациони рат, дезинформација, вештачка интелигенција и синтетички медији, биће дефинисани на следећи начин. Под пропагандом, подразумева се ширење ставова, чињеница, аргумената и детаља, подразумевајући и нетачне и лажиране, ради креирања става јавног мњења о одређеном питању, при чему дигитална пропаганда подразумева спровођење наведене радње путем дигиталних медија. Пропагандно деловање може бити константно, спроводи се и у мирнодопско и у ратно доба, углавном ради креирања позитивне слике о питању од значаја за сопствено друштво, док се информациони рат, који подразумева и израду и пласирање пропаганде, спроводи непосредно пре, а нарочито током „конвенционалног, хибридног или хладног рата” (Kotelenets and Barabash 2019, 369), са циљем дискредитације непријатеља. У оквиру пропагандног деловања, а нарочито информационог рата, фокус је на пласирању дезинформација, при чему се дезинформација дефинише као с намером и свесно креирана лажна информација, усмерена на обману примаоца поруке (Bereskin 2023, 6), са крајњим циљем стицања политичке, економске или војне користи (Majchrzak 2023, 43). У ланцу радњи, дезинформација је најмања јединица, која се користи као средство ширења пропаганде, док пропаганда може, али и не мора бити део веће операције – информационог рата. На крају, вештачку интелигенцију дефинишемо као способност машине да реализује задатке и друге функције које, по природи ствари, подразумевају људски фактор, попут резоновања, планирања, учења и креативности (Majchrzak 2023, 44), док синтетичке медије описујемо као медијски садржај (текст, видео запис, аудио запис, фотографија, визуелни приказ) генерисан употребом вештачке интелигенције (Bereskin 2023, 25).

ПРОПАГАНДА НЕКАДА И САДА: ФАКТОР ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ

Пропагандно деловање према непријатељу није конструкт 20. века, нити новина у оквиру техника манипулација (Sayler and Harris 2022, 1), већ постоји од како постоји и човечанство. Сагледавши примере пропагандних делатности од значаја од времена древних цивилизација до данас, предлажемо следећу историјску поделу хронологије пропаганде:

Фаза митолошке пропаганде, окарактерисана креирањем предања, митова и легенди о владару, историјском догађају, или пак истакнутим личностима. Претежно базиране на оралном предању, репрезентацији кроз споменике, уметничке слике и друга културна добра, овакве пропагандне кампање пласиране су претежно локалном становништву, с обзиром на немогућност брзог одашиљања поруке у удаљене крајеве.³

Фаза механичке пропаганде отпочиње 1440. године механизацијом израде пропагандних материјала уз помоћ Гутенбергове штампарске пресе. Од половине 15. века, штампарска преса и други алати, претежно настали у 20. веку, као што су радио станице, филмска продукција, а затим и телевизија, омогућили су брже ширење пропаганде и допирање до ширих народних маса. Свој врхунац пропаганда доживљава у предратно и ратно доба, када се користи не само за обмањивање непријатеља, већ и за повећање регрутних капацитета кроз јачање морала, како борачког, тако и неборачког становништва (видети Sostarić 2019).

Фаза дигиталне пропаганде отпочиње са комерцијализацијом Интернета, а нарочито с појавом друштвених мрежа и платформи *MySpace*, *Facebook*, *YouTube* и *Twitter* током прве деценије 21. века. Оно што карактерише ову фазу није искључиво бржи приступ таргетираној групи и децентрализовано деловање на широке масе, већ и прелазак пропаганде из сфере државног у приватно, чиме се отвара пут за утицај недржавних актера на светску политику. Најјаснији пример ефектног деловања цивила у пропагандне сврхе видимо у покретима који су иницирали побуне у оквиру Арапског пролећа на

³ Дobar пример проналазимо у деловању Рамзеса II, који је након битке у Кадешу 1274. године пре нове ере, пласирао дезинформацију која је исход битке приказала као победу Египта инструисањем клесара да у храмовима уклешу сведочанства о овој победи. Видети Ingenia 2022.

Блиском истоку, али и у пропагандном деловању аутоматизованих руских ботова у оквиру председничких избора у САД 2016. године.⁴

Фаза синтетичке пропаганде надовезује се на фазу дигиталне пропаганде, са којом је у каузалном односу, имајући у виду да синтетички пропагандни материјали користе дигиталне медије као платформу преко којих се даље одашиљају у етар. Ова фаза почиње са првим помацима у области машинског учења, алата из спектра аналитичке вештачке интелигенције (видети Kazić 2023b), а нарочито са пробојем у домену дубоког учења (*deep learning*), захваљујући којем је 2014. године конструисана комплексна техника названа „генеративна адверсаријална мрежа” (*Generative Adversarial Network – GAN*). Уз помоћ ове технике, систем вештачке интелигенције учи из доступних аудио, видео записа и фотографија, након чега стиче способност да креира сасвим нови садржај. Употребом две алгоритамске форме – генератора (*generator*), који креира нови садржај, и дискриминатора (*discriminator*), који оцењује веродостојност и поверљивост истог, овом техником је могуће креирати потпуно лажне садржаје који веродостојно приказују појам или личност на које се оне односе (Busch and Ware 2023, 2). Услед чињенице да су у питању садржаји генерисани методом дубоког учења, то јест, *deeplearning*, производи генеративних адверсаријалних мрежа именовани су (*deepfake*), у буквалном преводу: дубоко лажирање, то јест, синтетички материјал.

Садржај *deepfake*, стога, односи се на синтетички видео, аудио или фотографски садржај генерисан методом дубоког учења, са циљем лажирања изворног материјала и његовог преиначења у садржај који приказује нешто сасвим друго. У контексту видео садржаја, *deepfake* се односи на садржаје који представљају измењене видео материјале, при чему се процесу обраде снимка реализује визуелна замена записа (примера ради, замена лица говорника, како би деловало да радњу приказану у снимку врши друго лице), или пак измена аудио записа, којом се актеру оригиналног снимка приписују друге изјаве. На крају, визуелни синтетички садржај односи се на фотографије креиране уз помоћ дубоког учења, које могу бити

⁴ У овом случају, руски ботови су учили председничких избора у САД 2016. године успели да поларизују гласачко тело ширењем дезинформација о демократском кандидату Хилари Клинтон (*Hillary Clinton*), истовремено промовишући републиканског кандидата Доналда Трампа (*Donald Trump*). Видети Ross, Vaccari and Chadwick 2022.

измењена варијанта оригиналних фотографија или пак потпуно нови производи, креирани у складу са препознатим сентиментом, ради изазивања нових осећања код циљне групе.

ТЕХНОЛОГИЈА *DEEPFAKE* У РАТНИМ ЗОНАМА И ПОЛИТИЧКИМ КРИЗАМА: СТУДИЈЕ СЛУЧАЈА

Употребна вредност све три варијанте синтетичких производа широка је, а садржаји ове врсте пенетрирали су не само свет политике, већ и ратне зоне, о чему ће сведочити одабране студије случаја. Међутим, пре саме анализе случаја, нужно је идентификовати потенцијале појавних облика ових производа, ради адекватне процене релевантности њихове употребе у политичким и ратним пропагандама данашњице, а стога и будућности.

Синтетички видео записи, то јест, *deepfake* видео записи, присутни су у свету од 2017. године (Sayler and Haggis 2022, 1), при чему од 2018. године бележе редовно појављивање у политичким круговима широм света. Од тада, израда синтетичких видео записа у политичке сврхе постала је правило, а не изузетак од праксе, чиме је само појачана постојећа пропагандна машинерија, а нарочито отворен пут ка ширењу пропаганде међу најмлађим политичким актерима, који се претежно информишу путем друштвених мрежа. Од 2022. године, присутност синтетичких видео записа бележи се и у ратној зони, почев од помињаних говора Путина и Зеленског, ширећи се и у друге зоне сукоба.

Синтетички аудио записи до сада су углавном бивали инкорпорирани у *deepfake* видео записе, о чему сведоче потоњи случајеви лажирања говора Путина и Зеленског, при чему се подразумева да је у овим случајевима лажиран ток дискурса, али задржан оригиналан глас претпостављених говорника. Најрепрезентативнији пример синтетичког аудио материјала представља лажирано обраћање бившег председника Сиријске Републике Башара ал-Асада, који је добио значајну популарност у данима његовог свргавања, о којем ће касније бити више речи. Стога, нужно је нагласити да и сами синтетички аудио записи имају своју потенцијалну употребну вредност као синтетички генерисан пропагандни материјал и део информационог рата, с обзиром на то да имају знатан потенцијал као средство обмане непријатеља, а нарочито непријатељских снага у зонама конфликта.

На крају, синтетичке фотографије представљају производ генерисан употребом ВИ који је најприсутнији у дигиталном простору, с обзиром на широку распрострањеност платформи и апликација које омогућавају израду тих материјала. Имајући у виду да фотографија као метод ширења пропаганде има најчешће за циљ изазивање емоције код посматрача, најрепрезентативнији пример *deepfake* фотографија представљају лажиране фотографије страдања деце у израелско-палестинском сукобу, при чему је таква методологија пропагандног деловања присутнија на палестинској страни.

Све три категорије лажираног садржаја, стога, имају значајан потенцијал за измену перцепције одређене политичке кризе или сукоба, при чему је важно идентификовати модалитете и ситуације у којима се оне најчешће појављују. По питању лажираних видео записа, из анализираног корпуса снимака можемо закључити да је главни циљ израде таквих материјала дискредитација противника (Вуман et al. 2023, 6–8), али и лажирање наређења. Ово се потврђује претходно описаним инстанцама лажирања видео записа Путина и Зеленског, којима се инсинуира неопходност обуставе ратних дејстава, полагање оружја и, у случају Украјине, предаја, те изазивање конфузије у оквиру матичног становништва и иницирање сукоба, као што је случај са Габоном,⁵ у којем је садржај лажираног видео записа послужио као легитимизација извођења пуча (7). Синтетички аудио записи имају исту ударну моћ попут синтетичких видео записа, при чему је значајно истаћи да се такви материјали у ратно доба могу користити и за преношење лажираних аудио записа путем радио средстава трупамa на терену, чиме се првенствено могу издати лажна наређења, изазвати конфузија, а потенцијално и дефетизам међу припадницима јединице. На крају, лажиране фотографије израђују се ради поларизације друштва и умањивања подршке непријатељским снагама (8), а нарочито ради изазивања жељених сентимената и окретања јавног мњења на страну актера приказаних на фотографији, као што је случај са фотографијама страдале деце у Палестини.

⁵ Сматра се да је овај *deepfake* видео запис иницирао извођење неуспелог пуча у држави 2019. године, с обзиром на то да је приказ председника Алија Бонга (*Ali Bongo*) унео конфузију у друштво и довео до пораста сумње у његову способност да влада државом. Видети Вуман et al. 2023.

Употреба синтетичког садржаја у зонама сукоба: израелско-палестински сукоб

Иако је синтетички садржај као средство ширења дезинформација у зонама сукоба дебитовао почетком руско-украјинског сукоба 2022. године, лажирани видео записи, аудио записи и фотографије приказали су свој пуни потенцијал у оквиру текућег израелско-палестинског сукоба (Mustafa 2021, 213), при чему је фокус стављен претежно на израду лажираних фотографија (Klepper 2023).

Анализирајући постојеће синтетичке видео записе, приметно је да је овај тип лажираних материјала најређе присутан, а у случају појаве, подразумева претежно манипулације и лажирање аудио записа оригиналног видео записа из савременог периода, или пак из протеклих неколико година. По питању идентификованих лажираних видео записа, приметно је да се такви материјали претежно израђују у оквиру палестинске пропаганде, са циљем провокације Израела, изазивајући страх међу локалним становништвом, смањујући подршку ратном кабинету, али и умањујући морал борбених ефектива. Последњи забележени такав случај у тренутку израде овог истраживања подразумевао је пласирање два видео записа напада на Израел који наводно изводи седам држава, објављеног на друштвеној мрежи *TikTok* 6. октобра 2024. године, дан пре годишњице отпочињања сукоба (Zhang and Wang 2024). Детаљнијом анализом порекла видео записа, утврђено је да су сами видео записи заиста лажирани, али и да су објављени нешто раније, а затим поново објављени дан пред годишњицу ради присећања на напад Хамаса, са циљем уношења немира и конфузије међу локалним становништвом.

С друге стране, корпус детектованих лажираних аудио записа који третирају питање израелско-палестинског сукоба углавном делује као оруђе дискредитације непријатеља и урушавања међународне подршке, при чему је приметно да се такав материјал претежно јавља у оквиру палестинске пропаганде. Наведено оправдавамо чињеницом да Палестина, као „слабији играч” у сукобу, настоји да искористи све постојеће адуте, а стога и нове технологије, попут генерисања синтетичких материјала, како би допринела борби и ослабила „јачи” Израел. У том контексту, као најрелевантнији пример такве праксе истичемо видео запис у којем је извршена синтетизација аудио записа, како би се припаднику ОС Израела приписало пласирање

регрутне пропаганде ка украјинским држављанима, обећавајући им држављанство, решење стамбеног питања, новац и пословне прилике (Center on Extremism 2023).

Израда лажираних фотографија је већ на самом почетку сукоба идентификована као „врло моћан алат” за пласирање дезинформација (Eisele and Steinwehr 2023) са циљем истицања палестинског страдања, или пак јачања морала палестинског/израелског становништва. Наиме, још током првих дана ескалације израелско-палестинског сукоба, у медијском простору се јављају прве лажне фотографије из Палестине, у којима се претежно приказује страдање деце, с обзиром на то да се бомбардовани циљеви препознају као идеална позадина за креирање садржаја која интензивира емоције посматрача (Eisele and Steinwehr 2023). Једна таква фотографија, објављена само пет дана након ескалације сукоба – 12. октобра 2023. године, имала је за циљ да приказом напуштеног одојчета у рушевинама у Гази (Khatsenkova 2023) генерише жаљење због страдања палестинског народа,⁶ те је идентификована као један од најранијих синтетичких материјала који је обележио овај сукоб. Пропагандни материјал усмерен на јачање морала палестинског народа од почетка сукоба појачан је и материјалима генерисаним уз помоћ вештачке интелигенције, у оквиру којих се углавном јављају слични мотиви – наставак породичног живота у рушевинама након израелског дејствовања, али и након заточеништва у израелским притворним центрима (Golshiri 2024), што је карактеристично за палестинску пропаганду. Поред материјала овог карактера, присутне су и синтетичке фотографије које приказују дејствовање ОС Израела и тенкове међу рушевинама, које су, по природи ствари, вероватно креиране у склопу израелске пропаганде (Eisele and Steinwehr 2023; Klepper 2023).

Увидом у наведене материјале, приметно је да су аутори ових пропагандних материјала претежно патриотски настројена цивилна лица, која на својим приватним налозима на друштвеним мрежама деле синтетички садржај, а нарочито синтетичке фотографије, које се користе и у новинарске сврхе (Eisele and Steinwehr 2023). На овај начин, истиче се социјални карактер овакве пропагандне активности, који омогућава деловање у гомили, али из удобности свог дома, чиме се утиче на морал локалног становништва, емоције посматрача, и подршку међународне заједнице.

⁶ Касније је утврђено да је у питању фотографија лажирана осам месеци раније, као приказ страдања деце у земљотресима.

Међутим, значајно је истаћи и да су истраживачи који се баве дезинформацијама у првим недељама сукоба утврдили да је број синтетичких садржаја ограничен, а већина није уверљива, одбацујући њихов реални утицај на ширење пропаганде, али потврђујући њихову моћ да изазову јаке емоције код посматрача. У том контексту, истраживачи истичу да су синтетички материјали довели и до креирања отклоне код одређене групе лица, која сада и реалне медијске садржаје карактерише као синтетичке, штитећи се на тај начин од манипулација, али и од истине (Hsu and Thompson 2023). Као последица тога, учесницима у конфликту је отворена могућност имплементације принципа „дивиденде лажљивца” (*liar's dividend*), у оквиру којег могу негирати аутентичност деликатног медијског садржаја негативне конотације, називајући га синтетичким конструктом, иако он то у бити није (Busch and Ware 2023, 4; Hsu and Thompson 2023), тиме доводећи међународну заједницу у недоумицу и изазивајући све дубљу ерозију поверења у истинитост доступних информација о дешавању на терену.

На крају, индикативно је да се корисници друштвених мрежа и даље ослањају на употребу старих снимака, монтирање визуелних материјала и приписивање истих дешавањима у Израелу/Палестини (Hsu and Thompson 2023), мада фреквентност појаве синтетичких садржаја указује да ће се употреба *deepfake* материјала интензивирати у даљем току овог, али и у будућим сукобима. Стога, можемо закључити да је палестинска пропагандна машинерија, која је доминантнија у овом сукобу, у процесу тестирања могућности алата за генерисање синтетичких материјала, што ће довести до унапређења техника, а самим тим и повећања квалитета садржаја, који ће преузети примат над квантитетом и на тај начин додатно замаглити линију између истине и лажи.

У пару са другим алатима које нуди аналитичка ВИ, попут алгоритамаске амплификације, обраде природних језика, па чак и манипулисања претраживачима (видети Kazić 2023a), могуће је произвести још деструктивнију манипулацију наративима присутним на друштвеним мрежама, не само кроз генерисање лажних садржаја, већ и кроз селектовање наратива који ће бити најприсутнији и тиме диктирати сентименте општег јавног мњења о одређеним питањима, чиме је могуће изградити потпуно нову слику о актуелним, али и будућим ратним дешавањима на нивоу региона.

Употреба синтетичког садржаја у политичким кризама: пад Владе у Сирији

Одолевајући протестима, грађанском рату, терористичком деловању и страном утицају скоро 14 година, свргавање председника Сиријске Републике, Башара ал-Асада, одиграло се брзопотезно, нападом „Покрета за ослобођење Леванта” (*Hayat Tahrir al-Sham*), у трајању од 27. новембра до 8. децембра 2024. године (Loft and Mills 2024, 10). Оставком ал-Асада и напуштањем државе тог 8. децембра, стављена је тачка на 54 године владавине породице ал-Асад Сиријом, и отворен је простор за деловање разуђеној опозицији побуњеничког карактера, која тренутно управља државом.

Иако се у овој студији случаја анализира период од нешто мање од два месеца, наведени догађај значајан је утолико због појаве синтетичког садржаја који третирају ову политичку кризу у врло кратком временском периоду. Говорећи о синтетичким видео записима, у задатом временском року није идентификован ниједан лажирани видео запис, што не значи нужно да такав материјал није израђен. Ипак, прва ноћ у Сирији без породице ал-Асад окарактерисана је појавом лажног наратива о страдању чланова некада владајуће породице након одласка Башара ал-Асада, алудирајући на наводно страдање рођака свргнутог председника, извесног Сулејмана Хилала ал-Асада (*Sulaiman Hilal al-Assad*), који је обешен на тргу у Латакији. По том питању, сиријски медији посведочили су ширењу видео записа јавне егзекуције бешењем лица мушког пола на тргу, уз скандирање окупљене масе. Детаљнијом анализом, утврђено је да је у питању егзекуција другог лица, по имену Амар ал-Асад (*Ammar al-Assad*), који није у сродству са председником, и то на другој локацији, у граду Кирбет Газале. У тренутку израде ове студије, постоје индиције да је Сулејман ал-Асад, у крвном сродству са Башаром ал-Асадом, заиста убијен у Латакији (Bahl 2024), али оне нису потврђене.

Најзначајнији узорак синтетичког аудио записа појавио се у ноћи након пада сиријске Владе, 8/9. децембра 2024. године. Наведени аудио запис, визуелно идентификован фотографијом ал-Асада, представља монолог у којем свргнути председник објављује да одступа са дужности и предаје власт у руке председника Народне скупштине, уз услов да именовани прогласи свеобухватно национално помирење и организује демократске изборе у року од годину дана

(Bahl 2024). Имајући у виду да је бег ал-Башара из земље био већ позната чињеница, наведени снимак се исте ноћи проширио друштвеним мрежама, те је подељен преко 200.000 пута (Bahl 2024). Међутим, оно што је народ нагнало на сумњу у веродостојност истог није била „објава капитулације”, већ јавно извињење ал-Асада сиријском народу за дела целокупне породице ал-Асад, наређење повлачења свих страних снага и милиција са сиријске територије, те раскидање свих споразума са РФ и Исламском Републиком Иран (Bahl 2024), посебно имајући у виду да је Башар ал-Асад уточиште пронашао у Москви (Woollacott 2024; Muhammad 2024). Касније је утврђено да је овај синтетички аудио запис генерисан раније, у септембру 2023. године, када је објављен на друштвеној платформи *YouTube*; међутим, штета је већ била начињена, информација је подељена стотинама хиљада пута, а поједини корисници у њу су чак и поверовали.

Синтетичке фотографије страдања локалног становништва генерисане су и пласиране од почетка грађанског рата у земљи 2011. године (Sidharth 2018), а нарочито током примата „Исламске државе” на територији Сирије и Ирака (Кажо 2024), па све до данашњих дана, при чему се доминантна тематика мењала сходно потребама и ситуацији. У том контексту, од свргавања Башара ал-Асада, сиријским сајбер простором доминирају синтетичке, али и реалне фотографије из затвора Сендаја (*Sednaya*), стварајући конфузију и приказујући стање сада ослобођених затвореника породице ал-Асад горим него што оно заиста јесте (Pars Today 2024). Идеја иза ове кампање односи се на додатно урушавање култа Башара ал-Асада, приказ његове репресивне владавине, али и креирање симпатије за актуелну транзициону владу из редова побуњеника, која је овим чином окончала патњу ових претежно политичких затвореника, ради спречавања ал-Асадовог евентуалног повратка на чело државе. Наиме, фотографија која је привукла највише пажње, а која приказује затвореника овог затвора, заправо је лажирана фотографија из периода пре пада Башара ал-Асада, те нема везе са тренутном ситуацијом (Woollacott 2024).

Приказ синтетичких материјала који су освојили медијски простор након свргавања Башара ал-Асада указују на чињеницу да овако кратак период није довољан за креирање оригиналног и новог синтетичког садржаја, али свакако оставља простора за ревитализацију старих материјала у нове сврхе. Уз то, постојање

адекватних материјала указује на тенденцију креирања пропаганде генерисане ВИ, а нарочито аудио записа и фотографија, чему иду у прилог лажирано обраћање свргнутог председника и фотографије политичких затвореника. Имајући у виду постојање интереса озбиљних играча у свету информационог ратовања, попут РФ и ИР Иран (Loft and Mills 2024, 24–29), за сиријски сајбер (али и географски) простор, можемо очекивати интензивирање појаве инстанци синтетичких садржаја, који ће, с обзиром на ток времена и напредак технологије, потенцијално бити веродостојнији, емотивнији, а самим тим – и убитачнији.

ЗАКЉУЧАК

Појава генеративних адверсаријалних мрежа и њиховог производа – *deepfake* материјала, додатно је усложнила идентификацију истине у мору информативног садржаја, а по питању конфликта и политичких криза, указала на потенцијал вештачке интелигенције да ескалира сукобе и кризе, али и додатно замагли линију која дели истину од лажи.

Иако је модалитет ширења дигиталне пропаганде и дезинформација знатно усложен појавом ВИ у овим процесима, академска заједница још увек не идентификује овај сегмент као нову фазу пропагандне активности. Овим радом, а нарочито овим завршним речима, настојимо да им променимо мишљење. Наиме, индикативно је да су израда и ширење пропаганде знатно олакшане појавом Интернета и друштвених мрежа, али је сама израда пропагандног материјала, до појаве вештачке интелигенције, остала у домену мануалне изrade, те је изискивала посвећеност, рад, труд, а нарочито – идеју. Ера синтетичке пропаганде и уплив ВИ ослободила је израђиваче тог мануелног рада, стављајући у њихове руке „готов производ”, израђен према датим спецификацијама. Ово се не односи само на садржаје из спектра *deepfake*, већ и на пропагандне текстове, који се сада могу израђивати уз помоћ алата за генерисање текста, али и алата за обраду природних језика (*natural language processing* – *NLP*). Други алати из спектра ВИ, а нарочито машинског учења, попут анализе сентимената, омогућавају хиперперсонализацију садржаја, што је приметно у случају Палестине/Израела. Напредни алати ВИ такође пружају могућност обраде великог скупа података, ради идентификовања емоција припадника циљне групе, чиме

се омогућава и идентификација друштвених подела и креирање садржаја којим се додатно ескалира сукоб или политичка криза, доводећи учеснике у сукобу у пат-позицију.

Сагледавајући презентоване студије случаја израелско-палестинског сукоба и свргавања Башара ал-Асада у Сирији, индикативно је да се циљеви пропагандног деловања не мењају, већ увек делују у сврху манипулације сентиментима, оправдавања сопственог деловања, те осуђивања активности непријатеља. Ипак, општа доступност алата овог карактера, а нарочито алата базираних на дубоком учењу, који омогућавају масовну производњу синтетичких садржаја, знатно олакшавају и убрзавају пропагандно деловање у оквиру дигиталног простора. Из сукоба у сукоб, од кризе до кризе, алати аналитичке вештачке интелигенције, сходно принципима машинског учења који захтевају константно тренирање машина ради достизања напретка, постижу све боље резултате, те пракса спровођења пропаганде уз овакве материјале бележи све веће успехе. Сходно томе, природно је – и очекивано, у наредним годинама, наредним сукобима, и наредним политичким кризама, очекивати свестранију, свеобухватнију и целисходнију употребу ових нових алата пропагандног деловања.

РЕФЕРЕНЦЕ

- Al-Dhahab, 'Alī. 2019. "Al-tā'irātu dūna ṭayār: al-taqniyya wa al-'āthar al-'askariyy wa al-'istrātījiyy." *Al Jazeera Centre for Studies*. Poslednji pristup 7. novembar 2024. https://studies.aljazeera.net/sites/default/files/articles/reports-ar/documents/b2af23cf5d894ed7a9ce3e2d1556b382_100.pdf.
- Bahl, Vedika. 2024. "Viral audio deepfake falsely claims Bashar al-Assad 'apologised' to Syrian people." *France24*. 12 December, 2024. <https://www.france24.com/en/tv-shows/truth-or-fake/20241211-viral-audio-deepfake-falsely-claims-bashar-al-assad-apologised-to-syrian-people>.
- Baig, Rachel. 2022. "The deepfakes in the disinformation war." *Deutsche Welle*. 8 March, 2022. <https://www.dw.com/en/fact-check-the-deepfakes-in-the-disinformation-war-between-russia-and-ukraine/a-61166433>.
- Busch, Ella, and Jacob Ware. 2023. "The Weaponisation of Deepfakes: Digital Deception by the Far-Right." *International Centre for Counter-Terrorism*. DOI: 10.19165/2023.2.07.

- Byman, Daniel L., Chongyang Gao, Chris Meserole, and V. S. Subrahmanian. 2023. "Deepfakes and International Conflict." *Foreign Policy at Brookings*. Poslednji pristup 7. novembar 2024. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflicts.
- Center on Extremism. 2023. "ADL Debunk: Myths and False Narratives About the Israel-Hamas War." *ADL*. November 10, 2023. <https://www.adl.org/resources/article/adl-debunk-myths-and-false-narratives-about-israel-hamas-war>.
- Dorić, Marija, i Tanja Milošević. 2021. „Zloupotreba veštačke inteligencije u terorističkim i ekstremističkim grupama.” *Srpska politička misao* 71 (1): 201–221. DOI: 10.22182/spm.7112021.9.
- Eisele, I. and Steinwehr, Uta. 2023. "Fact check: AI fakes in Israel's war against Hamas." *Deutsche Welle*. October 11, 2023. <https://www.dw.com/en/fact-check-ai-fakes-in-israels-war-against-hamas/a-67367744>.
- Forest, James J. F. 2021. "Political Warfare and Propaganda: An Introduction." *Journal of Advanced Military Studies* 12 (1): 13–33. DOI: 10.21140/mcu.j.20211201001.
- Genuth, Joel. 1988. "Microwave Radar, the Atomic Bomb, and the Background to U.S. Research Priorities in World War II." *Science, Technology, and Human Values* 13 (3–4): 276–289. DOI: 10.1177/016224398801303-4.
- Golshiri, Ghazal. 2024. "Israel frees director of Gaza hospital, sparking internal political backlash." *Le Monde*. 2 July, 2024. https://www.lemonde.fr/en/international/article/2024/07/02/israel-frees-director-of-gaza-hospital-sparking-internal-political-backlash_6676426_4.html.
- Hsu, Tiffany, and Stuart A. Thompson. 2023. "A.I. Muddies Israel-Hamas War in Unexpected Way." *The New York Times*. 30 October, 2023. <https://www.nytimes.com/2023/10/28/business/media/ai-muddies-israel-hamas-war-in-unexpected-way.html>.
- Ingenia. 2022. "The Battle of Kadesh: Fake News and Diplomacy in the New Kingdom." *Ingenia*. 16 September, 2022. <https://www.ingeniahistory.com/post/the-battle-of-kadesh>.
- Kajjo, Sirwan. 2024. "IS turns to artificial intelligence for advanced propaganda amid territorial defeats." *Voice of America*. 23 May, 2024. <https://www.voanews.com/a/is-turns-to-artificial-intelligence-for-advanced-propaganda-amid-territorial-defeats/7624397.html>.

- Kazić, Tanja. 2023a. "Perspectives of Implementation of Artificial Intelligence Tools in Researching Coups d'Etat." *The Policy of National Security* 25 (2): 117–137. DOI: 10.5937/pnb25-47012.
- Kazić, Tanja. 2023b. "The Use of Predictive Analytics in Political Science: Africa between the East and the West." In *Proceedings of the 5th Virtual International Conference – Path to a Knowledge Society – Managing Risks and Innovation*, eds. Miomir Stanković, and Vesna Nikolić, 319–326. Niš: Complex System Research Center; Belgrade: Mathematical Institute of the Serbian Academy of Sciences and Arts.
- Khatsenkova, Sophia. 2023. "Israel-Hamas War: This viral image of a baby trapped under rubble turned out to be fake." *Euro News*. 24 October, 2023. <https://www.euronews.com/my-europe/2023/10/24/israel-hamas-war-this-viral-image-of-a-baby-trapped-under-rubble-turned-out-to-be-fake>.
- Klepper, David. 2023. "Fake babies, real horror: Deepfakes from the Gaza war increase fears about AI's power to mislead." *Associated Press*. 28 November, 2023. <https://apnews.com/article/artificial-intelligence-hamas-israel-misinformation-ai-gaza-a1bb303b637ffb9cbc3aale000db47>.
- Kotelenets, Elena, and Viktor Barabash. 2019. "Propaganda and Information Warfare in Contemporary World: Definition Problems, Instruments and Historical Context." *Advances in Social Science, Education and Humanities Research* 374: 368–371. DOI: 10.2991/mplg-ia-19.2019.69.
- Loft, P. and Mills, Claire. 2024. *Syria after Assad: Consequences and next steps in 2024/25*. House of Commons Library, Research Briefing no. 10161.
- Majchrzak, Adam. 2023. "Russian disinformation and the use of images generated by artificial intelligence (deepfake) in the first year of the invasion of Ukraine." *Media Business Culture* 14 (1): 43–55.

- Muhammad, Mustafa. December 20, 2024. "Limādhā 'ikhtāra al-'asad rūsiyā wiġhatan lilfirār.. wa hal 'aṣbaḥa bi'aman?" *Al Modon*. <https://www.almodon.com/arabworld/2024/12/14/%D9%84%D9%85%D8%A7%D8%B0%D8%A7-%D8%A7%D8%AE%D8%AA%D8%A7%D8%B1-%D8%A7%D9%84%D8%A3%D8%B3%D8%AF-%D8%B1%D9%88%D8%B3%D9%8A%D8%A7-%D9%88%D8%AC%D9%87%D8%A9-%D9%84%D9%84%D9%81%D8%B1%D8%A7%D8%B1-%D9%88%D9%87%D9%84-%D8%A3%D8%B5%D8%A8%D8%AD-%D8%A8%D8%A3%D9%85%D8%A7%D9%86>
- Mustafa, Hebat Allah Nasr Hassan. 2021. "Dawru wikalā' i al-dhikā' i al-i'ṣṭanā' iyyi fī taḥqīqi al-a'mani al-ṣuḥufiyyi fī al-ḥurūbi al-haġīnati bi-ttaṭbīqi 'alā 'amaliyyat ḥāris al-ġidrāni al-i'srā' īliyyati." *Al-maġalla al-miṣriyya libuḥūth al-i'ttiṣāl al-ġamāhīriyy*. https://mebp.journals.ekb.eg/article_310097_b1b4e20cc38fe46ebe3e1818c8dcefaf.pdf.
- Pars Today. 2024. "Social media users' reaction to the use of fake images for Syrian prisons." *Pars Today*. 15 December, 2024. https://parstoday.ir/en/news/west_asia-i231256-social_media_users'_reaction_to_the_use_of_fake_images_for_syrian_prisons.
- Pearson, James, and Zinets, Natalia. "Deepfake footage purports to show Ukrainian president capitulating." *Reuters*. 17 March, 2022. <https://www.reuters.com/world/europe/deepfake-footage-purports-show-ukrainian-president-capitulating-2022-03-16/>.
- Ross, R. N. Andrew, Cristian Vaccari, and Andrew Chadwick. 2022. "Russian Meddling in U.S. Elections: How News of Disinformation's Impact Can Affect Trust in Electoral Outcomes and Satisfaction with Democracy." *Mass Communication and Society* 25 (6): 786–811. DOI: 10.1080/15205436.2022.2119871.
- Sayler, Kelley M., and Laurie Harris. 2022. "Deepfakes and National Security." *Congressional Research Service*. 3 June, 2022. <https://crsreports.congress.gov/product/pdf/IF/IF11333/5>.
- Sidharth, Arjun. 2018. "Amidst human tragedy, fake pictures of Syrian conflict flood social media." *ATL News*. 8 March, 2018. <https://www.altnews.in/amidst-human-tragedy-fake-pictures-of-syrian-conflict-flood-social-media/>.
- Sostaric, Mia. 2019. "The American Wartime Propaganda During World War II." *Australasian Journal of American Studies* 38 (1): 17–44.

- Zabecki, David T. 2015. "Military Developments of World War I." *International Encyclopedia of the First World War*. 7 May, 2015. <https://encyclopedia.1914-1918-online.net/article/military-developments-of-world-war-i/?format=pdf>.
- Woollacott, Emma. 2024. "Fog Of War Fuels Misinformation About Syria." *Forbes*. 10 December, 2024. <https://www.forbes.com/sites/emmawoollacott/2024/12/10/fog-of-war-fuels-misinformation-about-syria/>.
- Zhang, Joyce, and Tommy Wang. 2024. "AI-generated video falsely shared as 'Israel attack by Middle East countries.'" *AFP*. 18 December, 2024. <https://factcheck.afp.com/doc.afp.com.36Q89Q2>.

*Tanja Kazic**

Center for Strategic Prognosis, Belgrade, Republic of Serbia

DIGITAL PROPAGANDA AND DISINFORMATION GENERATED VIA ARTIFICIAL INTELLIGENCE: THE ISRAEL/ PALESTINE CONFLICT AND THE FALL OF BASHAR AL-ASSAD IN SYRIA CASE STUDIES

Resume

In this paper, the author will analyze the role of digital propaganda and disinformation generated via artificial intelligence (AI) in the ongoing Middle Eastern conflicts, with a focus on the escalation of conflict between Israel and Palestine, as well as the overthrow of the President of the Syrian Republic Bashar al-Assad. In this geographical framework, the author will examine how technologies based on AI, and primarily the generation of deepfake materials, have a use value in conflict zones, tracing the first steps of use of such tactics for war purposes in the ongoing Russia-Ukraine conflict. By confirming the presence and influence of this type of cyber operations through the analysis of two case studies, the author will confirm the hypothesis, which states that contemporary digital propaganda, aided and generated via AI, significantly complicates the course of ongoing events, thus leading to inevitable conflict escalation, making it difficult to distinguish the truth from lies, in both military and political context. With the use of the content analysis method, focused on the identification of instances of placement of AI-generated digital propaganda, this paper will present a comprehensive analysis and study of individual cases of Israel/Palestine and the Syrian Republic to create a framework of possibilities and influence of such content on classical armed conflicts, as well as on political crises in the Middle East.

Keywords: artificial intelligence, digital propaganda, disinformation, machine learning, deep learning, deepfake, Middle East, Israel/Palestine, Syria.

* E-mail address: tanja.z.milosevic@gmail.com; ORCID: 0009-0007-1091-6695.

** Овај рад је примљен 30. јануара 2025. године, а прихваћен за штампу на састанку Редакције 12. фебруара 2025. године.