

ქალების კუთვნილი ციფრული სივრცის მიტაცების უხილავი საფრთხეები აშშ-ში

ქეთევან ხუტურიძე

რედაქტორი: პროფესორი მაგდა ცინცაძე

აბსტრაქტი

ნაშრომის მიზანია გამოააშკაროს ე.წ. სიდრმისეული გაყალბების (დიფ ფეიკის) ტექნოლოგიების ქალების წინააღმდეგ გამოყენების შემთხვევების საგანგაშო მატება და მისი შედეგები. ხელოვნური ინტელექტის დახვეწასთან ერთად, „დიფ ფეიკი“ ინდივიდთა სახეებისა და ხმების მანიპულირების ხელსაწყო გახდა. მას მსხვერპლის თანხმობის გარეშე, შეურაცხყოფელი, საზიანო კონტენტის შესაქმნელად და გასავრცელებლად იყენებენ, რითაც მაღალი რისკის ქვეშ აყენებენ პიროვნების უსაფრთხოებასა და კონფიდენციალურობას ინტერნეტ სივრცეში.

შესავალი

ჩვენ მუდმივად ვვითარდებით, როგორც საზოგადოება, რომელმაც შექმნა, განავითარა და მუდმივად თანაარსებობს სხვადასხვა სახის ტექნოლოგიებთან. თუმცა, ამასთანავე ვაწყდებით შედეგებს, რომლებიც მიზანმიმართული და დამაზიანებელია, განსაკუთრებით მარგინალური ჯგუფებისთვის. ტექნოლოგიებს სააშკარაოზე გამოაქვს ჩვენს საზოგადოებაში ფესვგადგმული სტერეოტიპები და დისკრიმინაციული პრაქტიკები. არცერთ სხვა ტექნოლოგიურ წინსვლას არ გამოუმჯღავნებია ღრმად დამკვიდრებული ქალთმობილუობა და პატრიარქატი ისე ეფექტურად, როგორც ეს ხელოვნური ინტელექტის ტექნოლოგიის განვითარების შედეგად მოხდა.

ამ ეტაპზე ჩვენ ვდგავართ ხელოვნური ინტელექტის (AI) უშუალო და დიდი საფრთხის წინაშე: ეგრეთწოდებული „დიფ ფეიკის“ სახით, რომელიც ხელოვნური ინტელექტის (AI) მიერ გენერირებული სინთეზური მედიაა და ადამიანის მსგავს გამოსახულებებს ქმნის. ხელოვნური ინტელექტის ტექნოლოგიის გამოყენებით შესაძლებელია ნებისმიერი ადამიანის სახის გამოსახულება ინტეგრირდეს ისეთ ვიდეო ან ფოტო მასალასთან, რომელშიც ამ გამოსახულების მქონე პირებს არასოდეს ჰქონიათ შეხება, თუმცა ეს ფოტო და ვიდეო მასალა ძალიან რეალისტურად და დამაჯერებლად გამოიყურება.

პრობლემები და საფრთხეები

ე.წ. „დიფ ფეიკის“ შესახებ საზოგადოების დიდი წუხილი გამოწვეულია შიშით, რომ მათი გამოყენება შესაძლებელია დემოკრატიის ძირის გამოთხრის, დეზინფორმაციის გავრცელებისა და პოლიტიკური არეულობის შექმნის მიზნით. 2019 წლის გაერთიანებული სამეფოს არჩევნების დროს ბორის ჯონსონი გამოჩნდა კომპანია Future Advocacy-ის „დიფ ფეიკში“, რომელშიც ის მხარს უჭერდა თავის ოპონენტს; 2020 წელს, ბრიტანეთის მე-4 არხმა შექმნა ალტერნატიული სამობაო გზავნილი დედოფლისგან, რომელშიც მან უჩვეულო

კომენტარები გააკეთა სამეფო ოჯახისა და სხვადასხვა მნიშვნელოვან საკითხზე მისი პოზიციის შესახებ. ორივე ვიდეოს მიმდინარეობისას ეკრანზე მითითებული იყო, რომ ისინი „დიფ ფეიკი“ იყო და მიზნად ისახავდა საზოგადოების ცნობიერების ამაღლებას მსგავსი ტექნოლოგიის პოტენციური ბოროტად გამოყენების საფრთხის შესახებ.

კვლევების თანახმად, ონლაინ სივრცეში გავრცელებული „დიფ ფეიკების“ აბსოლუტური უმრავლესობა ქალებისადმია არის მიმართული. გენდერის მკვლევრები აღნიშნავენ, რომ ხელოვნური ინტელექტის მოდელირება გენდერული უთანასწორობითა და მასკულინური ნორმებითაა ნასაზრდოები და არა მხოლოდ ასახავს არსებულ უთანასწორობას, არამედ ასევე, აძლიერებს მას. ხელოვნური ინტელექტის განვითარება ძირითადად დასავლეთის ქვეყნების მდიდარი, პრივილეგირებული თეთრკანიანი კაცების მიერ ხორციელდება, რამაც განაპირობა ხელოვნური ინტელექტის სისტემებში სექსიზმის, გენდერული მიკერძოების და რასიზმის ასპექტების ინკლუზია. ეს საკითხი უფრო ღრმა პრობლემას ასახავს, რომელსაც ძალაუფლების არაპროპორციული დინამიკა და საზოგადოებაში კაცების დომინირება იწვევს. ამიტომ, ხელოვნური ინტელექტის ანალიზთან ერთად, აუცილებელია ხელოვნური ინტელექტის სფეროში მასკულინური ნორმების გაანალიზება.

კვლევის შედეგები

აშშ-ში განხორციელებული კვლევების თანახმად, მიუხედავად იმისა, რომ ამერიკელ კაც რესპონდენტებს აწუხებთ მათ შესახებ ყალბი ვიდეოების ინტერნეტში გავრცელება, 86%-დან მხოლოდ 13%-ს ფიქრობს, რომ ვიდეო შეიძლება სექსუალური ხასიათის იყოს. ქალების შემთხვევაში კი, თითქმის ყველა რესპონდენტი ყალბი ვიდეოების სექსუალურ ხასიათს ძალიან სავარაუდოდ შესაძლებლობად მიიჩნევს. ეს მიგნება ხაზს უსვამს მნიშვნელოვან განსხვავებებს, თუ როგორ აღიქვამენ კაცები და ქალები ხელოვნური ინტელექტის განვითარებასა და ზეგავლენასთან დაკავშირებულ რისკებს.

ეგრეთწოდებული „დიფ ფეიკის“ ტექნოლოგიების გავლენა ქალებზე ღრმად ნეგატიური და დამაზიანებელია. განსაკუთრებით საგანგაშოა ამ ფენომენის მასშტაბები და სიხშირე, მითუმეტეს, რომ „დიფ ფეიკის“ გენერირებასა და გავრცელებას ბევრი ვებგვერდი გვთავაზობს.

ხელოვნურ ინტელექტზე ფართო ხელმისაწვდომობა განუხრელად ზრდის რისკებს ქალებისთვის, რადგან ეს ინსტრუმენტი შეიძლება გამოყენებულ იქნას არაკონსენსუალური პორნოგრაფიული შინაარსის შესაქმნელად და უპრეცედენტო მასშტაბით უთხრის ძირს ქალთა მიმართ საზოგადოებრივ ნდობას. საფრთხეები ძალიან რეალურია და მილიონობით ქალია აღნიშნული ტექნოლოგიური თავდასხმის მსხვერპლი. 2024 წლის მაისში, Channel 4 News-ის ანალიზმა დაადგინა, რომ თითქმის 4000 ცნობილი ადამიანი იყო ჩამოთვლილი ყველაზე ნახვად „დიფ ფეიკების“ საიტებზე. 2024 წლის იანვარში, ტელიფონ სვიფტი გახდა არაკონსენსუალური „დიფ ფეიკის“ სურათების ობიექტი, რადგან ხელოვნური ინტელექტის მიერ გენერირებული სვიფტის სექსუალური გამოსახულებების კოლექცია ვირუსულად გავრცელდა მრავალ სოციალურ მედია პლატფორმაზე, მათ შორის X-სა და Meta-ზე და სულ რამდენიმე საათში 47 მილიონზე მეტ ნახვას მიაღწია.

„დიფ ფეიკის“ სამიზნე უმეტესად ცნობადი და საზოგადო მოღვაწე ქალები ხდებიან:

- კოორდინირებული ონლაინ სიძულვილის კამპანიის შედეგად შექმნილი „დიფ ფეიკის“ ვიდეოს მსხვერპლი გახდა ინდოელი ჟურნალისტი რანა აიუბი, რომელიც მთავრობის მიერ რვა წლის გოგონას გაუპატიურების არასათანადო გამოძიებას აპროტესტებდა.

- ნოელ მარტინი, ახალგაზრდა ავსტრალიელი ქალი, რომელიც მხარს უჭერდა იმიჯზე დაფუძნებული სექსუალური ძალადობის წინააღმდეგ მოძრაობას, ასევე გახდა „დიფ ფეიკის“ ტექნოლოგიით გენერირებული სექსუალური ფოტოებისა და ვიდეოების სამიზნე.
- 2024 წლის ივნისში ინტერნეტში 50-მდე სკოლის გოგოს პორნოგრაფიული სურათები გავრცელდა, რომლებიც ხელოვნური ინტელექტის აპლიკაციების გამოყენებით შეიქმნა.

როგორ მოვაწესრიგოთ

„დიფ ფეიკის“ ტექნოლოგიის გავრცელებისა და ხელოვნური ინტელექტის განვითარების დინამიკა აუცილებელს ხდის მკაცრი რეგულაციებისა და ეთიკის ნორმების დამკვიდრებას რისკების შემცირების მიზნით. მიუხედავად იმისა, რომ პორნოგრაფიული ხასიათის „დიფ ფეიკი“ პირველად 2017 წელს გამოჩნდა ამერიკულ პლატფორმა Reddit-ზე, აშშ-ში ჯერ კიდევ არ არსებობს ფედერალური კანონმდებლობა, რომელიც ეფექტურად დაარეგულირებს აღნიშნულ საკითხს. აღსანიშნავია, რომ ტელიორ სვიფტთან დაკავშირებით მომხდარი ინციდენტის შემდეგ, ამერიკელმა პოლიტიკოსებმა მოითხოვეს ახალი კანონების მიღება „დიფ ფეიკის“ ფოტოების შექმნის კრიმინალიზაციის მიზნით. იყო სხვა მცდელობები, როგორცაა გაერთიანებული სამეფოს ონლაინ უსაფრთხოების აქტი და ევროკავშირის ხელოვნური ინტელექტის აქტი, მაგრამ ვერც ერთმა ვერ შეაჩერა „დიფ ფეიკების“ შექმნის დინამიკა, რადგან ეს კანონმდებლობა მხოლოდ გარკვეულ შემთხვევებს არეგულირებს.

იმის გათვალისწინებით, რომ „დიფ ფეიკი“ ძირითადად AI-ზეა დაფუძნებული, სწორედ AI განიხილება ამ პრობლემის გადაწყვეტის საშუალებად. მაგალითად, მკვლევრებმა შექმნეს „დიფ ფეიკის“ იდენტიფიცირების სისტემები, რომლებიც აანალიზებს განათებას, ჩრდილებს, სახის მოძრაობებს და სხვა მახასიათებლებს გაყალბების აღმოსაჩენად. უსაფრთხოების კიდევ ერთი ინოვაციური მეთოდია ფოტოს ფაილში ფილტრის დამატება, რაც შეუძლებელს გახდის გამოსახულების გამოყენებას „დიფ ფეიკის“ შესაქმნელად. ასევე გაჩნდა რამდენიმე სტარტაპი, რომლებიც გთავაზობს „დიფ ფეიკის“ აღმოჩენის პროგრამულ უზრუნველყოფას, მათ შორისაა პროგრამები Truepic და Deeptrace.

დასკვნა

მთავრობების და მკვლევრების მიერ საკითხის შესწავლისა და დარეგულირების პროცესში ძალისხმევა მიმართული უნდა იყოს, როგორც „დიფ ფეიკის“ შექმნისა და გავრცელების მავნე პრაქტიკის წინააღმდეგ, ასევე იმ ქალებისა და მარგინალური ჯგუფების დასაცავად, რომლებიც ყველაზე მეტად განიცდიან „დიფ ფეიკის“ ზემოქმედებას. „დიფ ფეიკის“ შემთხვევებზე რეაგირების მეთოდების შემუშავების დროს, გენდერული ასპექტები არა მეორეხარისხოვანი თემა, არამედ დისკუსიისა და რეაგირების უმთავრესი საკითხი უნდა გახდეს.

გამოყენებული ლიტერატურა:

1. Statista: U.S. Adults' Fear of AI and Deep Fakes by Gender
2. GOV.UK: A Guide to the Online Safety Bill
3. European Commission: Regulatory Framework on AI
4. https://www.taylorfrancis.com/books/edit/10.4324/9781003193579/routledge-handbook-men-masculinities-organizations-jeff-hearn-kadri-aavik-david-collinson-anika-thym?_ga=302483088.1718205760&_gl=1*1mobxg9*_ga*MzAyNDgzMDg4LjE3MTgyMDU3NjA.*_ga_0HYE8YG0M6*MTcxODUyOTkxMy4yLjAuMTcxODUyOTkyMi41MS4wLjA.*_gcl_au*NjYyMzA2MDQwLjE3MTgyMDU3NTk
5. Routledge: Handbook on Men, Masculinities and Organizations