საქართველო და გლობალიზაცია

გლობალური საფრთხეები

ხელოვნური ინტელექტის უარყოფითი შედეგები

მოზამზადა ნინო გვარამაძემ

მსოფლიო ეკონომიკური ფორუმის გლობალური რისკების 2024 წლის ანგარიშში, სხვებთან ერთად, საფრთხედ ნახსენებია ხელოვნური ინტელექტის უარყოფითი შედეგები. რაც უფრო განვითარებულია ქვეყნები ამ მიმართულებით, მით უფრო უკეთ იკვლევენ მოსალოდნელ საფრთხეებს. ქვემოთ გთავაზობთ ამონარიდებს მაიკ თომასის სტატიიდან „ხელოვნური ინტელექტის თორმეტი საფრთხე“.

რაც უფრო იხვეწება და ფართოვდება ხელოვნური ინტელექტი, მით უფრო მეტი გაფრთხილება ისმის ხელოვნური ინტელექტის პოტენციურ საფრთხეებთან დაკავშირებით.

ხელოვნური ინტელექტი შეიძლება ჩვენზე უფრო ინტელექტუალური გახდეს და გადაწყვიტოს ჩვენი მართვა და ჩვენ ახლა უნდა ვიფიქროთ იმაზე, თუ როგორ ავიცილოთ ეს თავიდან“, - თქვა ჯეფრი ჰინტონმა, რომელიც ცნობილია როგორც „ხელოვნური ინტელექტის ნათლია“ მანქანური სწავლებისა და ნეირონული ქსელების ალგორითმების შესახებ ფუნდამენტური ნაშრომისთვის. 2023 წელს ჰინტონმა დატოვა თანამდებობა Google-ში, რათა შეეძლო „ესაუბრა ხელოვნური ინტელექტის საშიშროებაზე“ და აღნიშნა, რომ მისი ნაწილი ნანობს მის მიერვე გაწეულ შრომას.

ცნობილი კომპიუტერული მეცნიერი მარტო ერთადერთი არ არის, ვინც შეშფოთებას გამოხატავს.

Tesla-ს და SpaceX-ის დამფუძნებელმა, ილონ მასკმა, 1000-ზე მეტ ტექნიკურ ლიდერთან ერთად, 2023 წლის ღია წერილში მოუწოდა, შეჩერებულიყო ხელოვნური ინტელექტის დიდი ექსპერიმენტები და მოჰყვა, რომ ტექნოლოგიამ შეიძლება „ღრმა საფრთხე შეუქმნას საზოგადოებას და კაცობრიობას“.

ხელოვნური ინტელექტის 12 საფრთხე

კითხვები იმის შესახებ, თუ ვინ ავითარებს ხელოვნურ ინტელექტს და რა მიზნებისთვის, მნიშვნელოვანია მისი პოტენციური უარყოფითი მხარეების გასაგებად.

საშიშია თი არა ხელოვნური ინტელექტი (Artificial Intellect – AI)?

ტექნიკური საზოგადოება დიდი ხანია განიხილავს ხელოვნური ინტელექტის საფრთხეებს. სამუშაოების ავტომატიზაცია, ყალბი ამბების გავრცელება და ხელოვნური ინტელექტის მქონე იარაღის სახიფათო შეიარაღებისთვის რბოლა ნახსენებია, როგორც ხელოვნური ინტელექტის გამომწვევი ყველაზე დიდი საფრთხე.

1. ხელოვნური ინტელექტის გამჭვირვალობისა და განმარტების ნაკლებობა

ხელოვნური ინტელექტისა და ღრმა სწავლის მოდელები შეიძლება იყოს რთული გასაგები, მათთვისაც კი, ვინც უშუალოდ მუშაობს ტექნოლოგიასთან. ეს იწვევს გამჭვირვალობის ნაკლებობას, თუ როგორ და რატომ მიდის AI დასკვნამდე, ქმნის ახსნა-განმარტების ნაკლებობას, თუ რა მონაცემებს იყენებს AI ალგორითმები, ან რატომ შეიძლება მიიღონ მიკერძოებული ან სახიფათო გადაწყვეტილებები. ამ შეშფოთებამ გამოიწვია ახსნადი ხელოვნური ინტელექტის გამოყენება, მაგრამ ჯერ კიდევ დიდი გზაა, სანამ გამჭვირვალე AI სისტემები ჩვეულებრივ პრაქტიკად იქცევა.

2. AI ავტომატიზაციის გამო სამუშაოს დაკარგვა

AI-ზე მომუშავე სამუშაოს ავტომატიზაცია მწვავე შეშფოთებას იწვევს, რადგან ტექნოლოგია მიღებულია ისეთ ინდუსტრიებში, როგორიცაა მარკეტინგი, წარმოება და ჯანდაცვა. McKinsey[1]-ის თანახმად, 2030 წლისთვის, საქმიანობები, რომლებიც ამჟამად სამუშაო საათების 30 პროცენტს შეადგენს აშშ-ს ეკონომიკაში, შეიძლება ავტომატიზირებული იყოს. შედეგად შავკანიანი და ესპანურენოვანი თანამშრომლები დარჩებიან განსაკუთრებით დაუცველები ცვლილებების მიმართ. Goldman Sachs[2] კი აცხადებს, რომ 300 მილიონი სრული განაკვეთის სამუშაო ადგილი შეიძლება დაიკარგოს ხელოვნური ინტელექტის ავტომატიზაციის გამო.

”მიზეზი იმისა, რომ ჩვენ გვაქვს უმუშევრობის დაბალი მაჩვენებელი, ძირითადად არის ის, რომ დაბალი ხელფასის სერვისის სექტორის სამუშაო ადგილები საკმაოდ მტკიცედ შეიქმნა არსებული ეკონომიკის მიერ,” - განუცხადა Built In-ს ფუტურისტმა[3] მარტინ ფორდმა. მიუხედავად იმისა, რომ ხელოვნური ინტელექტი იზრდება, ”მე არ ვფიქრობ, რომ ეს გაგრძელდება.”

როდესაც ხელოვნური ინტელექტის რობოტები უფრო ჭკვიანები და მოხერხებულები ხდებიან, იგივე ამოცანები მოითხოვს ნაკლებ ადამიანს. და მიუხედავად იმისა, რომ ხელოვნური ინტელექტი 2025 წლისთვის 97 მილიონ ახალ სამუშაო ადგილს შექმნის, ბევრ თანამშრომელს არ ექნება ამ ტექნიკური როლებისთვის საჭირო უნარები და შეიძლება დარჩეს უმუშევრად, თუ კომპანიები არ იზრუნებენ თავიანთი სამუშაო ძალის განვითარებაზე.

ავტომატიზაცია ჩაანაცვლებს მარტივ შრომას, ხოლო ახალი სამუშაო ადგილები მოითხოვს კარგ განათლებას ან ტრენინგს ან, შესაძლოა, შინაგან ნიჭს - მართლაც ძლიერ ინტერპერსონალურ უნარებს ან კრეატიულობას - რაც შეიძლება გამოთავისუფლებულ სამუშაო ძალას არ ჰქონდეს. ეს ის უნარებია, რომლებშიც, ყოველ შემთხვევაში, ჯერჯერობით, კომპიუტერები არც თუ ისე ძლიერია.

---

[1] Global management consulting | McKinsey & Company

[2] Goldman Sachs Group, Inc. არის წამყვანი გლობალური ფინანსური ინსტიტუტი, რომელიც აწვდის ფინანსურ მომსახურებას ფართო სპექტრს დიდ და დივერსიფიცირებულ კლიენტთა ბაზაზე, რომელიც მოიცავს კორპორაციებს, ფინანსურ ინსტიტუტებს, მთავრობებსა და ინდივიდებს.

[3] ფუტურისტები (ასევე ცნობილი როგორც ფუტუროლოგები, პერსპექტივისტები, შორსმჭვრეტელნი პრაქტიკოსები და ჰორიზონტის სკანერები) არიან ადამიანები, რომელთა სპეციალობა ან ინტერესია ფუტუროლოგია ან მცდელობა სისტემატიურად გამოიკვლიონ პროგნოზები და შესაძლებლობები მომავლის შესახებ და როგორ შეიძლება გამოვიდნენ ისინი აწმყოდან, იქნება ეს კონკრეტულად ადამიანთა საზოგადოებასთან, თუ ზოგადად დედამიწაზე სიცოცხლესთან კავშირში.

ის პროფესიებიც კი, რომლებიც საჭიროებენ სამაგისტრო ხარისხს და დამატებით შემდგომ ტრენინგს, არ არიან დაცული ხელოვნური ინტელექტის ინტერვენციისგან.

როგორც ტექნოლოგიების სტრატეგმა კრის მესინამ აღნიშნა, ისეთი სფეროები, როგორიცაა სამართალი და ბუღალტერია, მზად არის ხელოვნური ინტელექტის „ხელში ჩასავარდნად“. სინამდვილეში, მესინამ თქვა, რომ ზოგიერთი მათგანი შესაძლოა განადგურდეს. ხელოვნური ინტელექტი უკვე მნიშვნელოვან გავლენას ახდენს მედიცინაზე. სამართალი და ბუღალტრული აღრიცხვა შემდეგია, თქვა მესინამ, პირველი მზად არის "მასიური რყევისთვის".

”იფიქრეთ კონტრაქტების სირთულეზე და ნამდვილად ჩავუღრმავდით და გაიგეთ, რა არის საჭირო სრულყოფილი გარიგების სტრუქტურის შესაქმნელად,” - თქვა მან იურიდიულ სფეროსთან დაკავშირებით. „ბევრი ადვოკატი კითხულობს უამრავ ინფორმაციას - ასობით ან ათასობით გვერდის მონაცემებსა და დოკუმენტებს. მართლაც ადვილია რაღაცის გამოტოვება. ასე რომ, ხელოვნური ინტელექტი, რომელსაც აქვს შესაძლებლობა მოაგვაროს და სრულყოფილად მიაწოდოს საუკეთესო შესაძლო კონტრაქტი იმ შედეგისთვის, რომლის მიღწევასაც ცდილობთ, სავარაუდოდ, ჩაანაცვლებს უამრავ კორპორატიულ ადვოკატს.” სავარაუდოდ, ჩაანაცვლებს მას, ვინც ვერ გამოიყენებს ხელოვნური ინტელექტის შესაძლებლობას.

3. სოციალური მანიპულირება ხელოვნური ინტელექტის ალგორითმების მეშვეობით

სოციალური მანიპულირება ასევე წარმოადგენს ხელოვნური ინტელექტის საშიშროებას. ეს შიში რეალობად იქცა, რადგან პოლიტიკოსები ეყრდნობიან პლატფორმებს თავიანთი შეხედულებების პოპულარიზაციისთვის, მაგალითად, ფერდინანდ მარკოს უმცროსი[4], რომელიც იყენებდა TikTok ტროლების არმიას ახალგაზრდა ფილიპინელების ხმების მოსაპოვებლად ფილიპინების 2022 წლის არჩევნებზე.

TikTok, რომელიც სოციალური მედიის პლატფორმის მხოლოდ ერთი მაგალითია, რომელიც ეყრდნობა AI ალგორითმებს, ავსებს მომხმარებლის არხს კონტენტით, რომელიც დაკავშირებულია წინა მედიასთან, რომელიც მათ ნახეს პლატფორმაზე. აპლიკაციის კრიტიკა მიმართულია ამ პროცესისა და ალგორითმის უუნარობის წინააღმდეგ მავნე და არაზუსტი შინაარსის გაფილტვრაში, რაც იწვევს შეშფოთებას TikTok-ის შესაძლებლობებთან დაკავშირებით, დაიცვას თავისი მომხმარებლები შეცდომაში შემყვანი ინფორმაციისგან.

ონლაინ მედია და სიახლეები კიდევ უფრო ბუნდოვანი გახდა ხელოვნური ინტელექტის მიერ გენერირებული სურათებისა და ვიდეოების, ხელოვნური ინტელექტის ხმის შემცვლელების, ასევე ღრმა ფეიქების ფონზე, რომლებმაც შეაღწიეს პოლიტიკურ და სოციალურ სფეროებში. ეს ტექნოლოგიები აადვილებს რეალისტური ფოტოების, ვიდეოების, აუდიო კლიპების შექმნას ან არსებული სურათის ან ვიდეოს ერთი ფიგურის მეორე გამოსახულებით შეცვლას. შედეგად, ჩნდება დეზინფორმაციისა და ომის პროპაგანდის გაზიარების კიდევ ერთი გზა, რაც ქმნის კოშმარული სცენარს, სადაც თითქმის შეუძლებელი იქნება განასხვავო სანდო და ყალბი ინფორმაცია.

”არავინ იცის, რა არის რეალური და რა არა”, - თქვა ფუტურისტმა მარტინ ფორდმა. „ასე რომ, ეს ნამდვილად მიგვიყვანს ისეთ სიტუაციამდე, როდესაც ფაქტიურად არ უჯერებ საკუთარ თვალებს

[4] ფერდინანდ მარკოს უმცროსი - ფილიპინების პრეზიდენტი 2022 წლიდან

და ყურებს; თქვენ ვერ შეძლებთ დაეყრდნოთ იმას, რაც, ისტორიულად, მიგვაჩნია საუკეთესო შესაძლო მტკიცებულებად... ეს იქნება უზარმაზარი პრობლემა. ”

4. სოციალური მეთვალყურეობა AI ტექნოლოგიით

გარდა მისი უფრო ეგზისტენციალური საფრთხისა, მარტინ ფორდი ორიენტირებულია იმაზე, თუ როგორ იმოქმედებს ხელოვნური ინტელექტი კონფიდენციალურობაზე და უსაფრთხოებაზე. მთავარი მაგალითია ჩინეთის მიერ სახის ამოცნობის ტექნოლოგიის გამოყენება ოფისებში, სკოლებში და სხვა ადგილებში. გარდა იმისა, რომ თვალყურს ადევნებს პირის მოძრაობას, ჩინეთის მთავრობას შეუძლია შეაგროვოს საკმარისი მონაცემები პირის საქმიანობის, ურთიერთობებისა და პოლიტიკური შეხედულებების მონიტორინგისთვის.

კიდევ ერთი მაგალითია აშშ-ს პოლიციის დეპარტამენტები, რომლებიც იყენებენ პოლიციის პროგნოზირებად ალგორითმებს, რათა წინასწარ განსაზღვრონ, თუ სად მოხდება დანაშაული. პრობლემა ის არის, რომ ამ ალგორითმებზე გავლენას ახდენს დაკავების მაჩვენებლები, რაც არაპროპორციულად მოქმედებს შავკანიანთა თემებზე. პოლიციის განყოფილებები შემდეგ ორჯერ ზრდიან ზეწოლას ამ თემებზე, რაც იწვევს ზედმეტ პოლიციას და კითხვებს იმის შესახებ, შეუძლიათ თუ არა თვითგამოცხადებულ დემოკრატიებს წინააღმდეგობა გაუწიონ ხელოვნური ინტელექტის ავტორიტარულ იარაღად გადაქცევას.

”ავტორიტარული რეჟიმები იყენებენ ან აპირებენ მის გამოყენებას”, - თქვა მარტინ ფორდმა. „საკითხავია, რამდენად შემოიჭრება ის დასავლეთის ქვეყნებში, დემოკრატიულ ქვეყნებში და რა შეზღუდვებს ვუყენებთ მას?

5. მონაცემთა კონფიდენციალურობის ნაკლებობა AI ინსტრუმენტების გამოყენებით

თუ თქვენ თამაშობდით AI ჩეთბოტით ან გამოსცადეთ AI სახის ფილტრი ონლაინ, თქვენი მონაცემები გროვდება – მაგრამ სად მიდის და როგორ გამოიყენება? ხელოვნური ინტელექტის სისტემები ხშირად აგროვებენ პერსონალურ მონაცემებს მომხმარებლის გამოცდილების მოსარგებად ან იმისთვის, რომ დაეხმარონ AI მოდელების მომზადებას, რომელსაც იყენებთ (განსაკუთრებით თუ AI ინსტრუმენტი უფასოა). მონაცემები არ შეიძლება ჩაითვალოს დაცულად სხვა მომხმარებლებისგან, როდესაც მიეწოდება AI სისტემას. მაგალითად, ერთი ინციდენტი, რომელიც მოხდა ChatGPT-თან 2023 წელს, ზოგიერთ მომხმარებელს საშუალება ეძლეოდა ენახა სათაურები სხვა აქტიური მომხმარებლის ჩატის ისტორიიდან. მიუხედავად იმისა, რომ არსებობს კანონები პერსონალური ინფორმაციის დასაცავად ზოგიერთ შემთხვევაში შეერთებულ შტატებში, არ არსებობს მკაფიო ფედერალური კანონი, რომელიც იცავს მოქალაქეებს ხელოვნური ინტელექტის მიერ გამოწვეული მონაცემების კონფიდენციალურობის ზიანისგან.

6. მიკერძოება ხელოვნური ინტელექტის გამო

AI მიკერძოების სხვადასხვა ფორმა ასევე საზიანოა. New York Times-თან საუბრისას, პრინსტონის კომპიუტერული მეცნიერების პროფესორმა ოლგა რუსაკოვსკიმ თქვა, რომ ხელოვნური ინტელექტის მიკერძოება სცილდება სქესისა და რასის ჩარჩოებს. გარდა მონაცემებისა და ალგორითმული მიკერძოებისა (რომელთაგან უკანასკნელს შეუძლია პირველის „გაძლიერება“), AI შემუშავებულია ადამიანების მიერ - და ადამიანები არსებითად მიკერძოებულნი არიან.

„A.I. მკვლევარები ძირითადად არიან მამრობითი სქესის წარმომადგენელი ადამიანები, რომლებიც მოდიან გარკვეული რასობრივი დემოგრაფიიდან, გაზრდილი მაღალ სოციალურ-

ეკონომიკურ სფეროებში, პირველ რიგში, შეზღუდული შესაძლებლობის არმქონე პირები“, - თქვა რუსაკოვსკიმ. ”ჩვენ საკმაოდ ერთგვაროვანი მოსახლეობა ვართ, ამიტომ გამოწვევაა მსოფლიო საკითხებზე ფართო ფიქრი.”

ხელოვნური ინტელექტის შემქმნელების შეზღუდულმა გამოცდილებამ შეიძლება ახსნას, თუ რატომ ვერ ახერხებს მეტყველების ამომცნობი ხელოვნური ინტელექტი გარკვეული დიალექტებისა და აქცენტების გაგებას, ან რატომ არ იღებენ კომპანიები მხედველობაში კაცობრიობის ისტორიაში ცნობილი ფიგურების იმიტირებული ჩატბოტის შედეგებს. დეველოპერებმა და ბიზნესებმა მეტი ყურადღება უნდა გამოიჩინონ, რათა თავიდან აიცილონ ძლიერი მიკერძოება და ცრურწმენები, რომლებიც საფრთხეს უქმნის უმცირესობებს.

7. სოციო-ეკონომიკური უთანასწორობა ხელოვნური ინტელექტის შედეგად

თუ კომპანიები უარს ამბობენ ხელოვნური ინტელექტის ალგორითმებში ჩადებული თანდაყოლილი მიკერძოებების აღიარებაზე, მათ შეუძლიათ კომპრომეტირება გაუკეთონ თავიანთი DEI (Diversity, Equity and Inclusion - მრავალფეროვნება თანასწორობა და ინკლუზია ინიციატივებს ხელოვნური ინტელექტის ფუნქციონირებით. იდეა, რომ AI-ს შეუძლია კანდიდატის თვისებების გაზომვა სახის და ხმის ანალიზის საშუალებით, ჯერ კიდევ დაბინძურებულია რასობრივი მიკერძოებით და ახდენს იგივე დისკრიმინაციული რეკრუტირების პრაქტიკის რეპროდუცირებას, რომლის აღმოფხვრასაც ბიზნესი თითქოსდა ცდილობს.

სოციო-ეკონომიკური უთანასწორობის გაფართოება, რომელიც გამოწვეულია ხელოვნური ინტელექტის საფუძველზე სამუშაოს დაკარგვით, შეშფოთების კიდევ ერთი მიზეზია, რომელიც ავლენს კლასობრივ მიკერძოებას იმის შესახებ, თუ როგორ გამოიყენება ხელოვნური ინტელექტი. მუშები, რომლებიც ასრულებენ უფრო მეტ მექანიკურ, განმეორებით დავალებებს, განიცდიან ხელფასების 70 პროცენტამდე შემცირებას ავტომატიზაციის გამო, ოფისისა მუშაკები ძირითადად ხელუხლებელი რჩებიან ხელოვნური ინტელექტის ადრეულ ეტაპზე. თუმცა, გენერაციული AI გამოყენების ზრდა უკვე გავლენას ახდენს საოფისე სამუშაოებზე, რაც ქმნის როლების ფართო სპექტრს, რომლებიც შეიძლება უფრო დაუცველი იყოს ხელფასის ან სამუშაოს დაკარგვის მიმართ, ვიდრე სხვები.

მოსაზრება, რომ ხელოვნურმა ინტელექტმა როგორღაც გადალახა სოციალური საზღვრები ან შექმნა მეტი სამუშაო ადგილი, ვერ ასახავს სრულ სურათს მისი ეფექტის შესახებ. გადამწყვეტი მნიშვნელობა აქვს განსხვავებების გათვალისწინებას რასის, კლასისა და სხვა კატეგორიების მიხედვით. წინააღმდეგ შემთხვევაში, უფრო რთული ხდება იმის დადგენა, თუ როგორ მოაქვს სარგებელი ხელოვნურ ინტელექტს და ავტომატიზაციას გარკვეულ ინდივიდებსა და ჯგუფებისთვის სხვების ხარჯზე.

8. ეთიკისა და კეთილგანწყობის შესუსტება ხელოვნური ინტელექტის გამო

ტექნოლოგებთან, ჟურნალისტებთან და პოლიტიკურ ფიგურებთან ერთად, რელიგიური ლიდერებიც კი აფრთხილებენ ხელოვნური ინტელექტის პოტენციურ მახეებზე. 2023 წლის ვატიკანის შეხვედრაზე და 2024 წლის მშვიდობის მსოფლიო დღისადმი მიძღვნილ გზავნილში, რომის პაპმა ფრანცისკემ მოუწოდა ერებს შექმნან და მიიღონ სავალდებულო საერთაშორისო ხელშეკრულება, რომელიც არეგულირებს ხელოვნური ინტელექტის განვითარებას და გამოყენებას.

რომის პაპმა ფრანცისკემ გააფრთხილა ხელოვნური ინტელექტის ბოროტად გამოყენების შესაძლებლობა და თქვა, რომ შექმნილია განცხადებები, რომლებიც ერთი შეხედვით დამაჯერებლად გამოიყურება, მაგრამ უსაფუძვლოა ან ღალატობს მიუკერძოებულობას. მან ხაზგასმით აღნიშნა, თუ როგორ შეიძლება ამან გააძლიეროს დეზინფორმაციის კამპანიები, უნდობლობა საკომუნიკაციო მედიის მიმართ, არჩევნებში ჩარევა და სხვა - საბოლოოდ გაზარდოს „კონფლიქტების გაღვივება და მშვიდობის შეფერხება“.

გენერაციული AI ინსტრუმენტების სწრაფი ზრდა ამ შეშფოთებას უფრო მეტ მნიშვნელობას ანიჭებს. ბევრმა მომხმარებელმა გამოიყენა ტექნოლოგია დავალების წერისგან თავის დასაღწევად, რაც საფრთხეს უქმნის აკადემიურ მთლიანობასა და კრეატიულობას. გარდა ამისა, მიკერძოებული ხელოვნური ინტელექტის გამოყენება შეიძლება იმის დასადგენად, არის თუ არა ინდივიდი შესაფერისი სამუშაოსთვის, იპოთეკური სესხისთვის, სოციალური დახმარებისთვის ან პოლიტიკური თავშესაფრისთვის, რაც გამოიწვევს შესაძლო უსამართლობას და დისკრიმინაციას, აღნიშნა პაპმა ფრანცისკემ.

„ადამიანის უნიკალური უნარი მორალური განსჯისა და ეთიკური გადაწყვეტილების მიღებისთვის უფრო მეტია, ვიდრე ალგორითმების რთული კოლექცია“, - თქვა მან. ”და ეს უნარი არ შეიძლება დაიყვანო მანქანურ პროგრამებამდე.”

9. ავტონომიური იარაღი, რომელსაც ამუშავებს AI

როგორც ძალიან ხშირად ხდება, ტექნოლოგიური მიღწევები გამოიყენება ომის მიზნით. რაც შეეხება AI-ს, ზოგიერთს სურს რაიმე გააკეთოს ამასთან დაკავშირებით, ვიდრე არ არის ძალიან გვიან: 2016 წლის ღია წერილში 30000-ზე მეტმა ადამიანმა, მათ შორის ხელოვნური ინტელექტისა და რობოტიკის მკვლევარებმა, უარი თქვა ინვესტიციებზე ხელოვნური ინტელექტის მქონე ავტონომიურ იარაღში.

„დღეს კაცობრიობის მთავარი კითხვაა, დაიწყოს თუ არა ხელოვნური ინტელექტის გლობალური შეიარაღების რბოლა, თუ თავიდან აიცილოს მისი დაწყება“, - წერენ ისინი. „თუ რომელიმე მსხვილი სამხედრო ძალა წინ წაიწევს ხელოვნური ინტელექტის იარაღის განვითარებაში, გლობალური შეიარაღების რბოლა პრაქტიკულად გარდაუვალია და ამ ტექნოლოგიური ტრაექტორიის საბოლოო წერტილი აშკარაა: ავტონომიური იარაღი ხვალინდელი კალაშნიკოვი გახდება“.

ეს პროგნოზი შესრულდა სასიკვდილო ავტონომიური იარაღის სისტემების სახით, რომლებიც დამოუკიდებლად ადგენენ და ანადგურებენ სამიზნეებს რამდენიმე რეგულაციის დაცვით. ძლიერი და რთული იარაღის გავრცელების გამო, მსოფლიოს ზოგიერთმა ყველაზე ძლევამოსილმა ქვეყანამ თავი დააღწია შფოთვას და ხელი შეუწყო ტექნიკურ ცივ ომს.

ამ ახალი იარაღიდან ბევრი ქმნის დიდ რისკს მშვიდობიანი მოსახლეობისთვის ადგილზე, მაგრამ საფრთხე ძლიერდება, როდესაც ავტონომიური იარაღი არასწორ ხელში მოხვდება. ჰაკერებმა აითვისეს კიბერშეტევების სხვადასხვა სახეობა, ამიტომ ძნელი წარმოსადგენი არაა მავნე მოქმედ ავტონომიურ იარაღში შეღწევა და აბსოლუტური არმაგედონის სტიმულირება.

თუ პოლიტიკური დაპირისპირება და მეომარი ტენდენციები არ იქნება კონტროლირებადი, ხელოვნური ინტელექტი შეიძლება დასრულდეს ყველაზე ცუდი განზრახვებით. ზოგიერთი შიშობს, რომ, რაც არ უნდა ბევრი ძლიერი ფიგურა მიუთითებდეს ხელოვნური ინტელექტის

საშიშროებაზე, მასზე მუშაობა მაინც გაგრძელდება თუ ეს ფულის გამომუშავების საშუალება იქნება.

„მენტალიტეტი ასეთია: „თუ ჩვენ შეგვიძლია ამის გაკეთება, უნდა ვცადოთ; ვნახოთ, რა მოხდება“, - თქვა კრის მესინამ (ტექნოლოგიებს სტრატეგმა).” და თუ ჩვენ შეგვიძლია ფულის გამომუშავება, ჩვენ ბევრს გამოვიმუშავებთ“. მაგრამ ეს არ არის უნიკალური ტექნოლოგიებისთვის. ეს სხვა სფეროებშიც ყოველთვის ხდებოდა და ხდება.”

10. AI ალგორითმების მიერ გამოწვეული ფინანსური კრიზისები

ფინანსური ინდუსტრია უფრო მგრძნობიარე გახდა ხელოვნური ინტელექტის ტექნოლოგიის ჩართულობის მიმართ ყოველდღიურ ფინანსურ და სავაჭრო პროცესებში. შედეგად, ალგორითმული ვაჭრობა შეიძლება იყოს პასუხისმგებელი ბაზრებზე ჩვენს მომავალ დიდ ფინანსურ კრიზისზე.

გარდა იმისა, რომ ხელოვნური ინტელექტის ალგორითმები არ არის დაბინდული ადამიანის განსჯით ან ემოციებით, ისინი ასევე არ ითვალისწინებენ კონტექსტს, ბაზრების ურთიერთდაკავშირებას და ფაქტორებს, როგორიცაა ადამიანის ნდობა და შიში. ეს ალგორითმები შემდეგ ახორციელებენ ათასობით გარიგებას ბუნდოვანი ტემპით, რომლის მიზანია გაყიდონ რამდენიმე წამის შემდეგ მცირე მოგებისთვის. ათასობით გარიგების გაყიდვამ შეიძლება შეაშინოს ინვესტორები იგივე საქმის გაკეთებაში, რაც გამოიწვევს უეცარ კრახებს და ბაზრის ექსტრემალურ არასტაბილურობას.

ისეთი შემთხვევები, როგორიც არის 2010 Flash Crash[5] და Knight Capital Flash Crash[6], ემსახურება როგორც შეხსენება იმის შესახებ, თუ რა შეიძლება მოხდეს, როდესაც ვაჭრობით ბედნიერი ალგორითმები გაბრაზდებიან, მიუხედავად იმისა, არის თუ არა სწრაფი და მასიური ვაჭრობა მიზანმიმართული.

ეს არ ნიშნავს იმას, რომ AI-ს არაფერი აქვს შესათავაზებელი ფინანსური სამყაროსთვის. სინამდვილეში, ხელოვნური ინტელექტის ალგორითმები შეიძლება დაეხმაროს ინვესტორებს, მიიღონ უფრო ჭკვიანი და ინფორმირებული გადაწყვეტილებები ბაზარზე. მაგრამ ფინანსური ორგანიზაციები უნდა დარწმუნდნენ, რომ მათ ესმით თავიანთი AI ალგორითმები და როგორ იღებენ გადაწყვეტილებებს ეს ალგორითმები. კომპანიებმა უნდა განიხილონ, ამაღლებს თუ ამცირებს ხელოვნური ინტელექტი ნდობას ტექნოლოგიის დანერგვამდე, რათა თავიდან აიცილონ ინვესტორებში შიშის გაღვივება და ფინანსური ქაოსი.

11. ადამიანის გავლენის დაკარგვა

ხელოვნური ინტელექტის ტექნოლოგიაზე გადაჭარბებულმა დამოკიდებულებამ შეიძლება გამოიწვიოს ადამიანის გავლენის დაკარგვა და ადამიანის ფუნქციონირების ნაკლებობა

[5] 2010 Flash Crash - 2010 წლის 6 მაისის კრახი, ასევე ცნობილი როგორც 2:45 კრახი ან უბრალოდ ფლეშ კრახი, იყო შეერთებული შტატების ტრილიონ დოლარის ფლეშ კრახი (საბირჟო კრახი. ბაზრის კრახი) რომელიც 14:32 საათზე დაიწყო (აღმოსავლეთის ზაფხულის დროით) და გაგრძელდა დაახლოებით 36 წუთი.

[6] Knight Capital Group - ამერიკული გლობალური ფინანსური მომსახურების ფირმა, რომელმაც 2012 წლის აგვისტოში სავაჭრო შეცდომის გამო $ 460 მილიონი დოლარი დაკარგა.

საზოგადოების ზოგიერთ ნაწილში. მაგალითად, ხელოვნური ინტელექტის გამოყენებამ ჯანდაცვის სფეროში შეიძლება გამოიწვიოს ადამიანის თანაგრძნობისა და მსჯელობის შემცირება. გენერაციული ხელოვნური ინტელექტის გამოყენებამ შემოქმედებითი მცდელობისთვის შეიძლება შეამციროს ადამიანის შემოქმედებითობა და ემოციური გამოხატულება. ხელოვნური ინტელექტის სისტემებთან ზედმეტმა ურთიერთობამ შეიძლება გამოიწვიოს თანატოლებთან კომუნიკაციისა და სოციალური უნარების შემცირება. ასე რომ, მიუხედავად იმისა, რომ ხელოვნური ინტელექტი შეიძლება იყოს ძალიან გამოსადეგი ყოველდღიური ამოცანების ავტომატიზაციისთვის, კითხვის ნიშნის ქვეშ დგას, შეიძლება თუ არა მან შეაფერხოს ადამიანის საერთო ინტელექტი, შესაძლებლობები და საზოგადოების საჭიროება.

12. უკონტროლო თვითშემეცნება AI

ასევე არსებობს შეშფოთება, რომ ხელოვნური ინტელექტი იმდენად სწრაფად განვითარდება, რომ ის გახდება მგრძნობიარე და იმოქმედებს ადამიანების კონტროლის მიღმა - შესაძლოა, მავნე გზით. ამ გრძნობის შესახებ სავარაუდო ცნობები უკვე არსებობდა, მაგალითად Google-ის ყოფილმაი ინჟინერმა განაცხადა, რომ AI ჩეთბოტი LaMDA მგრძნობიარე იყო და ესაუბრებოდა მას ისევე, როგორც ადამიანი. ვინაიდან ხელოვნური ინტელექტის შემდეგი დიდი ეტაპები მოიცავს ხელოვნური ზოგადი ინტელექტის მქონე სისტემების შექმნას და, საბოლოოდ, ხელოვნური სუპერინტელექტის მქონე სისტემების შექმნას, ამ მოვლენების სრულად შეჩერების მოთხოვნები კვლავ იზრდება.

როგორ შევამციროთ ხელოვნური ინტელექტის რისკები

AI-ს ჯერ კიდევ აქვს მრავალი სარგებელი, როგორიცაა ჯანმრთელობის მონაცემების ორგანიზება და თვითმართვადი მანქანების უზრუნველყოფა. თუმცა, ამ პერსპექტიული ტექნოლოგიიდან მაქსიმალური სარგებლობისთვის, ზოგი ამტკიცებს, რომ საჭიროა უამრავი რეგულაცია.

ჯეფრი ჰინტონის განცხადებით არსებობს სერიოზული საშიშროება იმისა, რომ ხელოვნური ინტელექტის განვითარება გამოიწვევს მოვლენებს ადამიანის კონტროლს მიღმა. „ეს არ არის მხოლოდ სამეცნიერო ფანტასტიკის პრობლემა. ეს არის სერიოზული პრობლემა, რომელიც, ალბათ, საკმაოდ მალე დადგება და პოლიტიკოსებმა უნდა იფიქრონ იმაზე, თუ რა უნდა გააკეთონ ამაზე ახლა“.

სამართლებრივი რეგულაციების შემუშავება

ხელოვნური ინტელექტის რეგულირება იყო ათობით ქვეყნის მთავარი აქცენტი და ახლა აშშ და ევროკავშირი ქმნიან უფრო მკაფიო ზომებს ხელოვნური ინტელექტის მზარდი დახვეწილობის სამართავად. ფაქტობრივად, თეთრი სახლის მეცნიერებისა და ტექნოლოგიების პოლიტიკის ოფისმა (OSTP) 2022 წელს გამოაქვეყნა AI უფლებების ბილი, დოკუმენტი, რომელიც ასახავს პასუხისმგებლობით ხელმძღვანელობას ხელოვნური ინტელექტის გამოყენებასა და განვითარებაში. გარდა ამისა, აშშ-ში პრეზიდენტის მიერ 2023 წელს გამოიცა ბრძანება, რომელიც ავალდებულებს ფედერალურ სააგენტოებს შეიმუშაონ ახალი წესები და მითითებები ხელოვნური ინტელექტის უსაფრთხოებისა და უსაფრთხოებისთვის.

მიუხედავად იმისა, რომ საკანონმდებლო რეგულაციები ნიშნავს, რომ ხელოვნური ინტელექტის გარკვეული ტექნოლოგიები შეიძლება საბოლოოდ აიკრძალოს, ეს ხელს არ უშლის საზოგადოებებს ამ სფეროს შესწავლაში.

ხელოვნური ინტელექტი აუცილებელია იმ ქვეყნებისთვის, რომლებიც ცდილობენ ინოვაციების შექმნას. თუმცა არსებობს საშიშროება იმისა, რომ სხვადასხვა ქვეყანა ამ საკითხს სხვადასხვაგვარად მიუდგება.

ორგანიზაციული AI სტანდარტების ჩამოყალიბება და დისკუსიების ორგანიზება

კომპანიის დონეზე, ბიზნესს შეუძლია გადადგას მრავალი ნაბიჯი, როდესაც AI ინტეგრირებულია მათ ოპერაციებში. ორგანიზაციებს შეუძლიათ განავითარონ პროცესები ალგორითმების მონიტორინგისთვის, მაღალი ხარისხის მონაცემების შედგენისა და AI ალგორითმების დასკვნების ახსნისთვის. ლიდერებს შეუძლიათ ხელოვნური ინტელექტი მათი კომპანიის კულტურისა და რუტინული საქმიანი დისკუსიების ნაწილად აქციონ, სტანდარტების დაწესება მისაღები AI ტექნოლოგიების გამოსაყენებლად.

ტექნოლოგიების სახელმძღვანელო ჰუმანიტარული პერსპექტივებით

ხელოვნური ინტელექტის შემქმნელებმა უნდა მოიძიონ ეროვნების, სქესის, კულტურისა და სოციალურ-ეკონომიკური ჯგუფის ადამიანების შეხედულებები, გამოცდილება და შეშფოთება, ისევე როგორც სხვა სფეროებიდან, როგორიცაა ეკონომიკა, სამართალი, მედიცინა, ფილოსოფია, ისტორია, სოციოლოგია, კომუნიკაციები, ადამიანისა და კომპიუტერის ურთიერთქმედება, ფსიქოლოგია და მეცნიერებისა და ტექნოლოგიების კვლევები.

მაღალტექნოლოგიური ინოვაციების დაბალანსება ადამიანზე ორიენტირებულ აზროვნებასთან იდეალური მეთოდია პასუხისმგებელი ხელოვნური ინტელექტის ტექნოლოგიის წარმოებისთვის და იმის უზრუნველსაყოფად, რომ ხელოვნური ინტელექტის მომავალი იმედიანი იყოს მომავალი თაობისთვის. ხელოვნური ინტელექტის საშიშროება ყოველთვის უნდა იყოს განხილვის საგანი, რათა ლიდერებმა გაარკვიონ ტექნოლოგიების კეთილშობილური მიზნებისთვის გამოყენების გზები.

”ვფიქრობ, ჩვენ შეგვიძლია ვისაუბროთ ყველა ამ რისკზე და ისინი ძალიან რეალურია”, - თქვა მარტინ ფორდმა. ”მაგრამ AI ასევე იქნება ყველაზე მნიშვნელოვანი ინსტრუმენტი ჩვენს ინსტრუმენტთა ყუთში ყველაზე დიდი გამოწვევების გადასაჭრელად.”

ბიბლიოგრაფია

1. Tomas, M., 12 Dangers of Artificial Intelligence (AI) | Built In
2. Hunt, T., Here's Why AI May Be Extremely Dangerous--Whether It's Conscious or Not | Scientific American