ავთანდილ გაგნიძე

ხელოვნური ინტელექტი განათლებაში

სალექციო კურსი

აღმოსავლეთ ევროპის უნივერსიტეტი
თბილისი 2026

ხელოვნური ინტელექტი განათლებაში

აღმოსავლეთ ევროპის უნივერსიტეტი

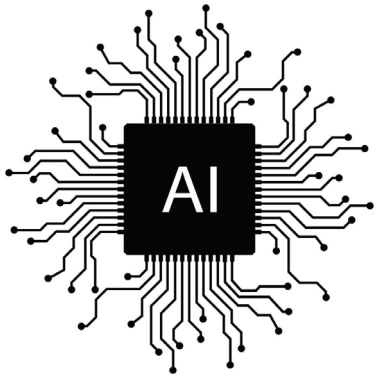
EEU.EDU.GE

ISBN 978-9941-8-8942-4

პროფესორი ავთანდილ გაგნიძე

ხელოვნური ინტელექტი განათლებაში

სალექციო კურსი



აღმოსავლეთ ევროპის უნივერსიტეტი
2026 წელი

სალექციო კურსის ჩანაწერები მომზადებულია
აღმოსავლეთ ევროპის უნივერსიტეტის განათლების
მეცნიერებათა ფაკულტეტის სტუდენტებისათვის

დაკაბადონება: გიორგი ბაგრატიონი

გამომცემელი: აღმოსავლეთ ევროპის უნივერსიტეტი

ISBN 978-9941-8-8942-4

მიმართვა

ძვირფასო სტუდენტებო,

თქვენს წინაშეა სალექციო კურსის ჩანაწერები, რომელიც ეძღვნება ჩვენი დროის ერთ-ერთ ყველაზე ტრანსფორმაციულ ტექნოლოგიას, ხელოვნურ ინტელექტს, და მის როლს განათლებაში. ეს კურსი შექმნილია იმისთვის, რომ დაგეხმაროთ არა მხოლოდ AI-ის ტექნიკური ასპექტების გააზრებაში, არამედ, რაც მთავარია, მისი პედაგოგიური პოტენციალისა და ეთიკური გამოწვევების ღრმად გაანალიზებაში.

დღეს, როდესაც ისეთი ინსტრუმენტები, როგორიცაა ChatGPT, DeepSeek ან Claude, ჩვენი ყოველდღიურობის ნაწილი გახდა, განათლების სისტემა დგას არჩევანის წინაშე: უარყოს ინოვაცია, შეეცადოს მის შეზღუდვას, ან ისწავლოს მისი ეფექტურად და პასუხისმგებლობით ინტეგრირება.

ეს ჩანაწერები მკაფიოდ დგას ამ უკანასკნელი მიდგომის მხარეს. ჩვენი მიზანია, AI განვიხილოთ არა როგორც საფრთხე მასწავლებლის პროფესიისთვის, არამედ როგორც ძლიერი ასისტენტი, რომელსაც შეუძლია გაათავისუფლოს ჩვენი დრო და ენერგია რუტინული ამოცანებისგან და მოგვცეს საშუალება მეტი ყურადღება გავამახვილოთ სწავლის პროცესის ადამიანურ განზომილებებზე: კრიტიკულ აზროვნებაზე, კრეატიულობასა და ემოციურ ინტელექტზე.

კურსის განმავლობაში ჩვენ გავივლით გზას ფილოსოფიური საწყისებიდან (დეკარტი, ტურინგი) პრაქტიკულ გამოყენებამდე. განვიხილავთ, თუ როგორ მუშაობს AI, რა შესაძლებლობებს გვთავაზობს ის პერსონალიზებული სწავლების, შეფასებისა და ადმინისტრირების კუთხით, და რა საფრთხეებს უნდა ვერიდოთ: დაწყებული ალგორითმული მიკერძოებიდან, მონაცემთა კონფიდენციალურობის დარღვევამდე.

ვიმედოვნებ, რომ ეს ჩანაწერები გახდება თქვენთვის საიმედო მეგზური AI-ის სამყაროში და დაგეხმარებათ, თავად გახდეთ იმ ცვლილებების აქტიური მონაწილეები, რომელიც ტექნოლოგიურ განვითარებას განათლებაში მოაქვს.

გისურვებთ წარმატებებს!

პროფესორი ავთანდილ გაგნიძე
აღმოსავლეთ ევროპის უნივერსიტეტი, 2026

შინაარსი

შესავალი

ისტორია

ადამიანის მსგავსი მექანიკური არსების შესახებ ფანტაზიები კულტურაში ძალიან დიდი ხანია არსებობს: აქ შეგვიძლია გავიხსენოთ ჰეფესტოსის მიერ შექმნილი გიგანტი ტალოსის ძველი ბერძნული მითი, ან პიგმალიონის მითი მოქანდაკის შესახებ, რომელიც ოცნებობდა თავისი ხელით გამოკვეთილი გალათეას ქანდაკების გაცოცხლებაზე, ან ზღაპარი პინოქიოს შესახებ.

ტალოსი *– ძველი ბერძნული მითოლოგიის მიხედვით, ჰეფესტოსმა, ღმერთების მჭედელმა, შექმნა ბრინჯაოს გიგანტი ტალოსი კრეტას დასაცავად. ტალოსი ყოველდღე სამჯერ შემოუვლიდა კუნძულს და ქვებს ესროდა მოახლოებულ გემებს. ეს მითი შეიძლება განვიხილოთ როგორც პირველი კონცეფცია „პროგრამირებადი“ მექანიკური არსებისა – ტალოსს ჰქონდა კონკრეტული ალგორითმი (დაცვის ფუნქცია) და ასრულებდა განმეორებად ამოცანებს. ის წარმოადგენდა იდეას ხელოვნური დამცველის შესახებ, რომელსაც არ სჭირდება დასვენება და აქვს უპირობო ერთგულება.*

პიგმალიონი და გალათეა *– ძველი ბერძნული მითი მოქანდაკე პიგმალიონის შესახებ, რომელმაც სპილოს ძვლისგან გამოკვეთა იდეალური ქალის ქანდაკება – გალათეა. პიგმალიონმა შეიყვარა თავისივე ქმნილება, ხოლო ღმერთმა აფროდიტემ მისი ლოცვები შეისმინა და გააცოცხლა ქანდაკება. ეს მითი წარმოაჩენს AI-ის ფილოსოფიურ საფუძველს – შეუძლია თუ არა შემქმნელს შექმნას რაღაც, რაც გადააჭარბებს მის თავდაპირველ ჩანაფიქრს? ეს არის შესაძლებლობის კითხვა იმის შესახებ, რომ ხელოვნურმა არსებამ განავითაროს თვითშეგნება და ემოციები – თემა, რომელიც დღესაც აქტუალურია AI-ის განხილვისას.*

პინოქიო *– იტალიური მწერლის კარლო კოლოდის ზღაპარი ხის მარიონეტზე, რომელიც ოცნებობდა რეალურ ბიჭად გადაქცევაზე. პინოქიო შექმნა ხელოსანმა ჯეპეტომ, მაგრამ მარიონეტმა დამოუკიდებელი ნება და გონება გამოიჩინა. ეს ნარატივი წარმოადგენს ადრეულ რეფლექსიას მანქანის „ადამიანურობაზე“ – როდის ხდება ხელოვნური ქმნილება ნამდვილად ცოცხალი? პინოქიოს ისტორია ასახავს AI-ის ერთ-ერთ ფუნდამენტურ კითხვას: შეიძლება თუ არა ალგორითმებმა და პროგრამირებამ შექმნას ნამდვილი ინტელექტი და თვითშეგნება?*

ხელოვნური ინტელექტის შექმნის შესაძლებლობის შესახებ სამეცნიერო იდეები ვერ ჩამოყალიბდებოდა რენე დეკარტის მექანიკური ფილოსოფიის გარეშე: „თუ ცოცხალი ორგანიზმი ფუნქციონირებს როგორც რთული კომპლექსური მექანიზმი, მაშინ შესაძლებელი უნდა იყოს ასეთი ორგანიზმის ხელოვნური ასლის შექმნა“.

დეკარტის მექანიკური ფილოსოფია

რენე დეკარტის (1596–1650) მექანიკურმა ფილოსოფიამ საფუძველი ჩაუყარა ხელოვნური ინტელექტის სამეცნიერო კონცეფციას. დეკარტი ამტკიცებდა, რომ ცხოველები და ადამიანის სხეული (გონებისგან განცალკევებით) მუშაობს როგორც რთული მექანიზმები – ისინი ემორჩილებიან ფიზიკის კანონებს და მათი ქცევები შეიძლება აიხსნას მექანიკური პრინციპებით.

ეს იდეა AI-ის საფუძვლად შეიძლება განვიხილოთ შემდეგი მიზეზებით:

რედუქციონიზმი *– თუ რთული ბიოლოგიური პროცესები შეიძლება დაიყოს უფრო მარტივ მექანიკურ კომპონენტებად, მაშინ თეორიულად შესაძლებელია ამ პროცესების რეპლიკაცია ხელოვნური საშუალებებით.*

დეტერმინიზმი *– დეკარტის აზრით, ცხოველების ქცევები პროგნოზირებადია და მიჰყვება კონკრეტულ წესებს. ეს პარალელურია თანამედროვე ალგორითმებთან – პროგრამირებადი ქცევის იდეასთან.*

გონების და სხეულის დუალიზმი *– დეკარტი გამიჯნავდა აზროვნებას (res cogitans) და ფიზიკურ სხეულს (res extensa). ეს ქმნის კითხვას: შეიძლება თუ არა აზროვნების პროცესების ფორმალიზება და აღწერა ისე, რომ ისინი არა-ბიოლოგიურ სუბსტრატზე განხორციელდეს?*

ხელოვნური ცხოვრების შესაძლებლობა *– თუ ორგანიზმები ფუნქციონირებენ როგორც კომპლექსური მექანიზმები, მაშინ, დეკარტის ლოგიკით, შესაძლებელი უნდა იყოს ასეთი ორგანიზმის ხელოვნური ასლის კონსტრუირება. ეს იყო რევოლუციური იდეა, რომელმაც საუკუნეების შემდეგ საფუძველი ჩაუყარა კომპიუტერულ მეცნიერებას და AI-ს.*

დეკარტის ფილოსოფიამ ინტელექტუალური გარემო შექმნა, სადაც აზროვნება, ცნობიერება და ინტელექტი აღარ განიხილებოდა როგორც განუყოფელი მისტიკური თვისებები, არამედ როგორც პროცესები, რომლებიც შეიძლება სისტემატურად შესწავლილიყო და პოტენციურად განმეორებადი.

მე-20 საუკუნეში, პირველი კომპიუტერების, კიბერნეტიკის, ნეირომეცნიერების საფუძვლების გაჩენისა და მათემატიკური ალგორითმების განვითარების შედეგად, მანქანის აზროვნების უნარის საკითხი აღარ იყო ფუჭი ფანტაზია და სამეცნიერო ინტერესის საგანი გახდა.

21-ე საუკუნის დასაწყისიდან ხელოვნური ინტელექტი არა მხოლოდ ცხარე დებატებისა და ლაბორატორიული კვლევების საგანი იყო, არამედ სწრაფად განვითარებადი გამოყენებითი სფეროც, რომელშიც ხორციელდება დიდი ინვესტიციები და რომელსაც კონკრეტული შედეგები მოაქვს ბიზნესისთვის.

ხელოვნური ინტელექტის განვითარებასთან ერთად, მისი სწავლების ღირებულება მცირდება: 2018 წელთან შედარებით, მაგალითად, გამოსახულების ამოცნობის სისტემის სწავლება 64%-ით ნაკლები ღირს. ამავდროულად, სწავლების სიჩქარე 94%-ით გაიზარდა. ასეთი ცვლილებები იწვევს იმ ფაქტს, რომ დღეს სულ უფრო მეტი კომპანია, ვისთვისაც ისინი ადრე მიუწვდომელი იყო, ცდილობს ხელოვნური ინტელექტის ტექნოლოგიების გამოყენებას.

კურსის მიზანი

იმისათვის, რომ კარგად გავიგოთ ამ კურსის მიზნები, დასაწყისისათვის განვიხილოთ ევროსაბჭოს 2019 წლის რეკომენდაცია „ციფრული მოქალაქეობის განათლების შესახებ“, რომელშიც ერთ-ერთი ძირითადი აქცენტი გაკეთდა ხელოვნური ინტელექტის გამოყენებაზე საგანმანათლებლო კონტექსტში:

„ხელოვნური ინტელექტი, ისევე როგორც ნებისმიერი სხვა ინსტრუმენტი, გვთავაზობს ბევრ შესაძლებლობას, თუმცა, ამავე დროს, მოიცავს უამრავ საფრთხეს, რაც აუცილებელს ხდის ადამიანის უფლებების პრინციპების გათვალისწინებას მისი გამოყენების საწყის ეტაპზევე.

AI არის ძლიერი ინსტრუმენტი, მაგრამ ნეიტრალური – მისი ზეგავლენა დამოკიდებულია იმაზე, როგორ გამოიყენება. შესაძლებლობები მოიცავს: პერსონალიზებულ სწავლებას, სწრაფ უკუკავშირს, ხელმისაწვდომობას. საფრთხეები კი არის: მონაცემთა კონფიდენციალურობის დარღვევა, ალგორითმული მიკერძოება (bias), ზედმეტი დამოკიდებულება ტექნოლოგიებზე, კრიტიკული აზროვნების დათრგუნვა.

პედაგოგებმა უნდა იცოდნენ ხელოვნური ინტელექტის სასწავლო პროცესში გამოყენების ძლიერი და სუსტი მხარეები, რათა გაძლიერდნენ – და არ დაიჩაგრონ – ტექნოლოგიების არსებობით ციფრული მოქალაქეობის სწავლებისას.

პედაგოგებმა არ უნდა იგრძნონ საფრთხე AI-ის მხრიდან, არამედ უნდა ისწავლონ მისი ეფექტური გამოყენება. ძლიერი მხარეები: დროის დაზოგვა, დიფერენცირებული მიდგომა, მონაცემთა ანალიტიკა. სუსტი მხარეები: შეუძლებელია მასწავლებლის ემოციური ინტელექტის, ემპათიის, მოტივაციის, მენტორობის ჩანაცვლება.

მანქანური სწავლებისა და სიღრმისეული დასწავლის საშუალებით, ხელოვნურ ინტელექტს შეუძლია განათლების გამრავალფეროვნება....

AI-ს შეუძლია მოარგოს სასწავლო მასალა თითოეული მოსწავლის ინდივიდუალურ საჭიროებებს – სწავლის ტემპს, სტილს, ცოდნის დონეს. მაგალითად, ერთსა და იმავე კლასში სხვადასხვა მოსწავლეს შეიძლება მიეწოდოს განსხვავებული სირთულის დავალებები, დამატებითი ახსნა-განმარტებები, ალტერნატიული რესურსები. ეს ქმნის უფრო ინკლუზიურ და ადაპტურ სასწავლო გარემოს.

ამასთანავე, ხელოვნური ინტელექტის სფეროში მიღწეულმა პროგრესმა შესაძლოა სერიოზული გავლენა მოახდინოს ურთიერთქმედებაზე მასწავლებლებსა და მოსწავლეებს ან, ზოგადად, მოქალაქეებს შორის, რამაც შესაძლოა დააზიანოს განათლების არსი, ანუ თავისუფალი ნების, დამოუკიდებელი და კრიტიკული აზროვნების ხელშეწყობა სასწავლო შესაძლებლობების შეთავაზებით...

განათლების მთავარი მიზანია არა ინფორმაციის ტრანსფერი, არამედ კრიტიკული, დამოუკიდებელი აზროვნების განვითარება. საფრთხეა, რომ AI-ის ზედმეტად დიდი დამოკიდებულება (მაგალითად, ყველა პასუხის მზა მიღება ChatGPT-სგან) შესაძლოა შეამციროს მოსწავლეების კრიტიკული ანალიზის, პრობლემის გადაჭრის, შემოქმედებითი აზროვნების უნარები.

მიუხედავად იმისა, რომ ჯერ კიდევ ადრეა სასწავლო გარემოში ხელოვნური ინტელექტის ფართო გამოყენებაზე საუბარი, აუცილებელია განათლების სფეროს ექსპერტებისა და სასკოლო პერსონალის ინფორმირება ხელოვნური ინტელექტისა და მასთან დაკავშირებული ეთიკური გამოწვევების შესახებ სასკოლო კონტექსტში".

პედაგოგებმა უნდა იცოდნენ ეთიკური საკითხების შესახებ: ალგორითმული მიკერძოება (რომელიც აძლიერებს სოციალურ უთანასწორობებს), მონაცემთა უსაფრთხოება (მოსწავლეთა პერსონალური ინფორმაციის დაცვა), გამჭვირვალობა (როგორ მიიღო AI-მ გადაწყვეტილება), სამართლიანობა (ყველა მოსწავლის თანაბარი წვდომა ტექნოლოგიებზე). ეს მოითხოვს სისტემატურ ტრენინგს და კრიტიკული დისკუსიების წახალისებას განათლების საზოგადოებაში.

განათლება და AI

განათლება ცვლის ცხოვრებას და თავად იცვლება ცხოვრებასთან ერთად და ახალი საგანმანათლებლო მეთოდების და ტექნოლოგიების დანერგვით. თანამედროვე მოწინავე საგანმანათლებლო ტექნოლოგიები არა მხოლოდ ხელს უწყობს საგანმანათლებლო პროცესების წარმატებით წარმართვას და მათზე წვდომას, არამედ მათ უფრო ცოდნაზე დაფუძნებულ და ადამიანზე ორიენტირებულს ხდის. ერთ-ერთი ასეთი მოწინავე საგანმანათლებლო ტექნოლოგია არის ხელოვნური ინტელექტი (AI).

AI ხელს შეუწყობს მასწავლებლის სამუშაოს რუტინული ნაწილის ავტომატიზაციას, გაათავისუფლებს ადგილს და დროს უფრო მნიშვნელოვანი შემოქმედებითი ამოცანებისთვის. მოსწავლეებისათვის და სტუდენტებისთვის ახალი ტექნოლოგიები ხდება სწრაფი და ეფექტიანი პერსონალიზებული უკუკავშირის საშუალება, მოქნილი და ადაპტური სწავლების საშუალება, სხვადასხვა სახის ამოცანების და პრობლემების ეფექტურად გადაჭრის საშუალება. ამდენად, ხელოვნური ინტელექტის დანერგვა არ ისახავს მიზნად ადამიანის ჩანაცვლებას, არამედ, პირიქით, მისთვის შესანიშნავ ასისტენტად ჩამოყალიბებას.

რა უნდა გვახსოვდეს

ამასთან ძალზედ მნიშვნელოვანია გვესმოდეს, რომ ხელოვნური ინტელექტის ახალი ტექნოლოგიები მნიშვნელოვნად არის დამოკიდებული ადამიანებზე და მგრძნობიარეა მათი შექმნის პირობების და მიზნების მიმართ. არ არსებობს აბსოლუტურად მზა და მკაცრად ჩარჩოში მოქცეული გადაწყვეტილებები. სულ მცირე, უნდა გავითვალისწინოთ, რომ:

- საჭიროა გამოთვლითი სიმძლავრეები და დიდი მოცულობის მონაცემების დაგროვება;
- ხელოვნური ინტელექტის მოდელების შესაქმნელად და სწავლებისთვის საჭიროა მაღალკვალიფიციური სპეციალისტების რესურსები;
- ასეთი სპეციალისტების მომზადებისათვის დრო და სხვა რესურსებია საჭირო.

ხელოვნური ინტელექტის განვითარება და დანერგვა გრძელვადიანი ინვესტიციაა. ამასთან, ხელოვნური ინტელექტი უკვე გამოიყენება ბევრ სფეროში და მათ შორის განათლების სფეროში, რაც უფრო მოქნილი, მიმზიდველი და შედეგიანი სწავლის შესაძლებლობებს გვთავაზობს.

AI შესაძლებლობები

ამგვარად, ხელოვნური ინტელექტის ტექნოლოგიები მოსწავლეებს ადაპტური სწავლის გაფართოებულ შესაძლებლობებს სთავაზობს, რაც ზრდის სას-

წავლო სისტემასთან და პროცესთან, კლასთან და ჯგუფთან და მასწავლებელთან ურთიერთქმედების ეფექტურობას:

1. ინტელექტუალური სასწავლო სისტემები საშუალებას გვაძლევს შევქმნათ პერსონალიზებული სასწავლო ტრაექტორია ცოდნის დონის მიხედვით და უზრუნველვყოთ სწრაფი ინდივიდუალური უკუკავშირი.
2. მოსწავლეთა ჯგუფის ინტელექტუალური ფორმირება ციფრული განათლების საფუძველზე ხელს უწყობს სწავლების პროცესში ყველაზე ჰარმონიულ თანამშრომლობას.
3. მოსწავლეებსა და მასწავლებლებს, დამხმარე გუნდსა და თანაკლასელებს შორის ურთიერთქმედების ინტელექტუალური მოდერაცია საშუალებას გვაძლევს გამოვავლინოთ შესაძლო კონფლიქტური სიტუაციები, გავაანალიზოთ გავრცელებული შეცდომები, თემიდან გადახრები და ა.შ.

ხელოვნური ინტელექტის ტექნოლოგიები მასწავლებლებს „მესამე ხელს“ და „მესამე თვალს“ სთავაზობს, ათავისუფლებს მათ რუტინული სამუშაოსაგან და საშუალებას აძლევს სწავლების ანალიტიკური მონაცემების მასივებში ახალი, ხანდახან მოულოდნელი, კანონზომიერების აღმოჩენისათვის:

1. დავალებების შემოწმების და შეფასების ავტომატიზაცია საშუალებას იძლევა ნაკლებად ხშირად მივმართოთ ექსპერტულ შეფასებას და ამცირებს ცოდნის შფასების სუბიექტურობას.
2. ავტომატურად გენერირებული დავალებების შექმნის საშუალება ამცირებს სტანდარტული დავალებების და ტესტების ვარიანტების შექმნასთან დაკავშირებულ სამუშაოს.
3. სწავლების მეთოდების ინტელექტუალური შეფასება საშუალებას იძლევა დროულად მივიღოთ უკუკავშირი სწავლება/სწავლის წარმატებულ და წარუმატებელ ელემენტებზე და დროულად მოვახდინოთ რეაგირება.

რას შევისწავლით

ამ კურსის მიზანია სტუდენტებისთვის ხელოვნური ინტელექტის, როგორც ტექნოლოგიისა და კონცეფციის, კონტექსტში ჩაღრმავება და მისი პოტენციალის გამოვლენა საგანმანათლებლო პროცესთან და ამოცანებთან მიმართებაში. თემების განხილვისას ჩვენ ვეცდებით პასუხების მოძიებას შემდეგ კითხვებზე:

- რა არის ხელოვნური ინტელექტი, როგორ განვითარდა ის და სად გამოიყენება დღეს?
- რა არის ხელოვნური ინტელექტის გამოყენების რამდენიმე მაგალითი ყოველდღიურ ცხოვრებაში?
- როგორ არის ხელოვნური ინტელექტი სტრუქტურირებული ტექნოლოგიური თვალსაზრისით?

- როგორ გამოიყენება ხელოვნური ინტელექტი განათლებაში?
- რა არის ხელოვნური ინტელექტის გამოყენების რამდენიმე მაგალითი საგანმანათლებლო სივრცეში?
- როგორ შეიძლება ხელოვნური ინტელექტის გამოყენება სასწავლო პროცესის სხვადასხვა ეტაპზე?
- რა შეზღუდვები და ეთიკური საკითხები უნდა იქნას გათვალისწინებული ხელოვნური ინტელექტის გამოყენებისას?

ჩვენი მიზანია ვიპოვოთ ბალანსი რეალობასა და ფანტაზიას შორის და გამოვავლინოთ ხელოვნური ინტელექტის ტექნოლოგიების გამოყენების კონკრეტული სფეროები განათლებაში.

სადისკუსიო თემები

AI – ასისტენტი თუ მეტოქე?

ტექსტში ნათქვამია, რომ ხელოვნური ინტელექტის დანერგვა „არ ისახავს მიზნად ადამიანის ჩანაცვლებას, არამედ მისთვის შესანიშნავ ასისტენტად ჩამოყალიბებას.“

საკითხი განსახილველად:

რეალურად, სად გადის ზღვარი „ასისტენტსა“ და „ჩამნაცვლებელს“ შორის? მოიფიქრეთ კონკრეტული პროფესიები ან სასკოლო როლები, სადაც AI-ს შემოსვლა სასარგებლო იქნება, და ისეთებიც, სადაც საფრთხეს შექმნის.

- *თქვენი აზრით, 10 წელში მასწავლებლის პროფესია როგორ შეიცვლება?*

ღირებულება და ხელმისაწვდომობა

ტექსტი აღნიშნავს, რომ AI-ს სწავლების ღირებულება 2018 წელთან შედარებით 64%-ით შემცირდა, სიჩქარე კი 94%-ით გაიზარდა.

საკითხი განსახილველად:

თუ AI ტექნოლოგიები სულ უფრო იაფი და სწრაფი ხდება, ეს ნიშნავს, რომ ყველა ქვეყანას, მათ შორის საქართველოს, თანაბარი წვდომა ექნება მათზე?
რა ბარიერები შეიძლება არსებობდეს (ენობრივი, ინფრასტრუქტურული, ფინანსური)?
სახელმწიფომ უნდა დაარეგულიროს AI-ზე წვდომა, თუ ეს ბაზარზე უნდა დარჩეს?

ევროსაბჭოს გაფრთხილება – რეალური საფრთხე თუ გაზვიადება?
ევროსაბჭოს რეკომენდაცია გვაფრთხილებს, რომ AI-მ შეიძლება „დააზიანოს განათლების არსი – თავისუფალი ნების, დამოუკიდებელი და კრიტიკული აზროვნების ხელშეწყობა.“
საკითხი განსახილველად:
შეგიძლიათ მოიყვანოთ კონკრეტული მაგალითი, სადაც AI-ს გამოყენებამ შეიძლება კრიტიკული აზროვნება შეასუსტოს?
ხომ არ ეწინააღმდეგება ეს შეხედულება იმ პოზიციას, რომ AI ეხმარება მოსწავლეებს ინდივიდუალური სწავლების გზით?
როგორ უნდა ასწავლონ პედაგოგებმა AI-ის „პასუხისმგებლიანი გამოყენება“?

რეკომენდირებული YouTube რესურსები:

ქართულენოვანი რესურსები:

- *არხი “Digital Georgia” – AI განათლებაში (საქართველოს კონტექსტი)*
- *“Georgian Innovation & Technology Agency (GITA)” – AI სემინარები და ლექციები*

საერთაშორისო რესურსები:

AI საფუძვლები:

- ***CrashCourse Computer Science** – “Artificial Intelligence” სერია (მარტივი, ილუსტრირებული)*
- ***3Blue1Brown** – “Neural Networks” სერია (მათემატიკური ინტუიცია ვიზუალური გზით)*
- ***Two Minute Papers** – ყოველკვირეული ვიდეოები AI-ის უახლესი კვლევების შესახებ*

AI განათლებაში:

- ***Stanford Online** – “AI in Education” ლექციები*
- ***EdSurge** – “AI in the Classroom” დოკუმენტური სერია*
- ***Google for Education** – “Teaching with AI” კურსი*

ეთიკა და საზოგადოების გავლენა:

- ***TED Talks** – “AI Ethics” თემატური ლექციები (Fei-Fei Li, Kate Crawford)*
- ***Veritasium** – “How AI is Changing Education”*
- ***Wired** – “The Future of AI” სერია*

პრაქტიკული მაგალითები:

- ***Khan Academy** – “Machine Learning Crash Course”*
- ***Sentdex** – პრაქტიკული AI პროექტები Python-ში*
- ***Two Minute Papers** – AI-ის შესაძლებლობების ვიზუალური დემონსტრაციები*

ტესტი 1

1. *რომელი ფილოსოფოსის მექანიკური ფილოსოფია ხელს უწყობდა ხელოვნური ინტელექტის შექმნის სამეცნიერო იდეების ჩამოყალიბებას?*
   ა) არისტოტელე
   ბ) რენე დეკარტი
   გ) იმანუელ კანტი
   დ) ჯონ ლოკი

2. *2018 წელთან შედარებით, რამდენით შემცირდა გამოსახულების ამოცნობის სისტემის სწავლების ღირებულება?*
   ა) 44%-ით
   ბ) 54%-ით
   გ) 64%-ით
   დ) 74%-ით

3. *ევროსაბჭოს რომელი წლის რეკომენდაცია ეხება „ციფრული მოქალაქეობის განათლებას”?*
   ა) 2015
   ბ) 2017
   გ) 2019
   დ) 2021

4. *რა არის ხელოვნური ინტელექტის დანერგვის მთავარი მიზანი განათლებაში?*
   ა) მასწავლებლის სრულად ჩანაცვლება
   ბ) ადამიანისთვის შესანიშნავ ასისტენტად ჩამოყალიბება
   გ) სასწავლო პროცესის გამარტივება და შემოკლება
   დ) მოსწავლეთა დამოუკიდებლობის შემცირება

5. *2018 წელთან შედარებით, რამდენით გაიზარდა ხელოვნური ინტელექტის სწავლების სიჩქარე?*
   ა) 74%-ით
   ბ) 84%-ით
   გ) 94%-ით
   დ) 104%-ით

6. *რომელი შემდეგთაგანი წარმოადგენს ინტელექტუალური სასწავლო სისტემის უპირატესობას?*
   ა) ყველა მოსწავლისათვის ერთიანი სასწავლო გეგმის შექმნა
   ბ) პერსონალიზებული სასწავლო ტრაექტორია ცოდნის დონის მიხედვით
   გ) მასწავლებლის სრული ჩართულობის გამორიცხვა
   დ) მოსწავლეთა შეფასების სუბიექტურობის გაზრდა

7. რომელ ძველ ბერძნულ მითში გვხვდება ხელოვნურად შექმნილი გიგანტი?

ა) პიგმალიონის მითი
ბ) პრომეთეს მითი
გ) ტალოსის მითი
დ) ოდისევსის მითი

8. რომელი შემდეგი ფაქტორი არ არის დასახელებული ხელოვნური ინტელექტის განვითარებისთვის საჭირო პირობად?

ა) გამოთვლითი სიმძლავრეები
ბ) დიდი მოცულობის მონაცემები
გ) მაღალკვალიფიციური სპეციალისტები
დ) სახელმწიფო სუბსიდიები

9. რა სარგებელს გვაძლევს დავალებების შემოწმების ავტომატიზაცია?

ა) ზრდის ცოდნის შეფასების სუბიექტურობას
ბ) ამცირებს ცოდნის შეფასების სუბიექტურობას
გ) ცვლის მასწავლებელს სრულად
დ) ზრდის მოსწავლეთა დამოკიდებულებას ტექნოლოგიებზე

10. კურსის ერთ-ერთი მთავარი მიზანია:

ა) ხელოვნური ინტელექტის შექმნის ტექნიკური სწავლება
ბ) ხელოვნური ინტელექტის პოტენციალის გამოვლენა საგანმანათლებლო პროცესში
გ) მასწავლებლების ჩანაცვლების სტრატეგიების შემუშავება
დ) მხოლოდ AI-ის ეთიკური საკითხების განხილვა

საუკეთესო AI განათლებისათვის

პერსონალიზებული სწავლებისთვის

Khanmigo (Khan Academy), *khanacademy.org/khan-labs* :

რა არის: Khanmigo წარმოადგენს GPT-4-ზე დაფუძნებულ AI რეპეტიტორს, რომელიც შემუშავებულია Khan Academy-ს მიერ და სრულად ინტეგრირებულია მათ სასწავლო პლატფორმასთან.

ძირითადი შესაძლებლობები:

- პირდაპირი ინტეგრაცია 70,000+ სავარჯიშოსთან და ვიდეოსთან მათემატიკაში, მეცნიერებაში, ჰუმანიტარულ მეცნიერებებსა და პროგრამირებაში
- სოკრატული მეთოდი – AI არ აძლევს პირდაპირ პასუხებს, არამედ სვამს მიმართულ კითხვებს კრიტიკული აზროვნების განვითარებისთვის
- მოსწავლის პროგრესის რეალურ დროში თვალყურის დევნება და მშობლების/მასწავლებლების ანგარიშგებები
- კრეატიული წერის დახმარება, დებატების პარტნიორი და კარიერის დაგეგმვის კონსულტანტი

გამოყენების მაგალითები:

- *მე-7 კლასის მოსწავლე ხსნის ალგებრის ამოცანას – ჯჰანმიგო არ ეუბნება პასუხს, მაგრამ ეკითხება: „რა ნაბიჯი გადადგი პირველ რიგში?" და ეხმარება სწორ გზაზე დარჩენაში*
- *ისტორიის მოსწავლე ემზადება დებატებისთვის – AI თამაშობს „მოწინააღმდეგის" როლს და ავარჯიშებს არგუმენტებში*
- *მშობელი ამოწმებს ბავშვის კვირეულ პროგრესს მათემატიკაში პლატფორმის ანგარიშიდან*

უსაფრთხოება და კონტროლი:

- შექმნილია COPPA და FERPA სტანდარტების გათვალისწინებით
- ფილტრები არასაბავშვო შინაარსისთვის
- მშობლების დაშვების ოფციები და საქმიანობის მონიტორინგი
- მონაცემთა დაცვის მკაცრი პოლიტიკა

ფასი: $4/თვე ან $44/წელი (უფასო პერიოდი მოსწავლეებისათვის ხშირად ხელმისაწვდომია)

Quizlet (Q-Chat), *quizlet.com:*

რა არის: Q-Chat არის AI-ით აღჭურვილი რეპეტიტორი, რომელიც ინტეგრირებულია Quizlet-ის 700+ მილიონი სასწავლო მასალის ბაზასთან.

ძირითადი შესაძლებლობები:

- ადაპტური სასაუბრო სწავლება – AI ითვალისწინებს სტუდენტის პასუხებს და არეგულირებს სირთულის დონეს
- ავტომატური flashcard-ების შექმნა ნებისმიერი ტექსტიდან, ლექციიდან ან PDF-დან
- სასწავლო რეჟიმები: Match (სიჩქარის თამაში), Learn (ადაპტური სწავლება), Test (პრაქტიკული გამოცდები)
- მრავალენოვანი მხარდაჭერა 130+ ენისთვის

გამოყენების მაგალითები:

- *მედიცინის სტუდენტი ატვირთავს ლექციის PDF-ს – Quizlet ავტომატურად ქმნის ტერმინების flashcard-ებს და ამოწმებს ცოდნას ადაპტური კითხვებით*
- *ინგლისურის მსწავლელი ვარჯიშობს 20 ახალ სიტყვაზე Match თამაშის გამოყენებით, სიჩქარის გაზომვით*
- *მასწავლებელი უზიარებს მზა ბარათებს კლასს გამოცდის მოსამზადებლად*

საუკეთესო გამოყენების შემთხვევები:

- მედიცინის სტუდენტებისთვის ტერმინოლოგიის დასამახსოვრებლად
- ენების სწავლისთვის ლექსიკის გასაუმჯობესებლად
- ისტორიის თარიღებისა და მოვლენების დასამახსოვრებლად
- მეცნიერული კონცეფციებისა და ფორმულების შესასწავლად

ფასი: უფასო ძირითადი ვერსია; Quizlet Plus ($7.99/თვე) და Quizlet Plus Teacher ($34/წელი)

MathGPT (AI Math Solver), *mathgpt.com:*

რა არის: სპეციალიზებული AI პლატფორმა მათემატიკურ პრობლემებთან სამუშაოდ, რომელიც გვთავაზობს ეტაპობრივ გადაწყვეტილებებს დეტალური ახსნა-განმარტებებით.

ძირითადი შესაძლებლობები:

- ფოტოთი პრობლემის სკანირება და ავტომატური ამოცნობა
- დეტალური ნაბიჯ-ნაბიჯ გადაწყვეტილებები ახსნა-განმარტებებით
- მოიცავს ალგებრას, გეომეტრიას, კალკულუსს, სტატისტიკას, ტრიგონომეტრიას
- ინტერაქტიული გრაფიკები და ვიზუალიზაცია

გამოყენების მაგალითები:

- *სტუდენტი სახლში ვერ წყვეტს ინტეგრალის ამოცანას – **MathGPT**-ს სურათს უგზავნის სახელმძღვანელოდან, იღებს ნაბიჯ-ნაბიჯ ახსნას და ბოლოს ხედავს, სად დაუშვა შეცდომა*
- *მასწავლებელი სწრაფად ამოწმებს, არსებობს თუ არა ამოცანის ალტერნატიული გადაწყვეტის მეთოდი*
- *მოსწავლე ვარჯიშობს მსგავს ამოცანებზე **AI**-ის მიერ გენერირებული პრაქტიკული სავარჯიშოებით*

დამატებითი ფუნქციები:

- საფეხურების ვიდეო ახსნები
- პრაქტიკული პრობლემები მსგავსი დავალებებისთვის
- შეცდომების ანალიზი და შესწორების რეკომენდაციები
- მრავალი გადაწყვეტის მეთოდის ჩვენება ერთი პრობლემისთვის

ფასი: უფასო საბაზისო ვერსია; პრემიუმ ფუნქციები $9.99/თვე

Socratic by Google, *socratic.org:*

რა არის: Google-ის შემუშავებული AI სასწავლო დამხმარე, რომელიც იყენებს ხედვას (ფოტო-სკანირებას) და ბუნებრივი ენის დამუშავებას პრობლემების გადასაჭრელად.

ძირითადი შესაძლებლობები:

- ფოტოთი დავალების სკანირება და მყისიერი ახსნა
- ვიზუალური დიაგრამები და ანიმაციები რთული კონცეფციების გასაადვილებლად
- ინტეგრაცია YouTube-ის სასწავლო ვიდეოებთან და Wikipedia-სთან
- მოიცავს მათემატიკას, მეცნიერებებს, ლიტერატურას, სოციალურ მეცნიერებებს

გამოყენების მაგალითები:

- *მე-9 კლასის მოსწავლეს არ ესმის ბიოლოგიის სახელმძღვანელოს პარაგრაფი – ის სურათს იღებს, **Socratic** განმარტავს შინაარსს ანიმირებული დიაგრამებით და სვამს **YouTube** ვიდეოს კავშირს*
- *მოსწავლე სახლში აკეთებს დავალებაა დედამიწის ისტორიის კონცეფციის შესახებ – ის აკეთებს კითხვის სურათს ან სკანირებს კითხვას და **AI** მარტივ ენაზე ხსნის შინაარსს*
- *სკოლის მოსწავლე ემზადება ლიტერატურის ტესტისთვის – **Socratic** ეხმარება ნაწარმოების თემების გაანალიზებაში*

გამორჩეული თვისებები:

- სრულიად უფასო (Google-ის პროდუქტი)
- ძალიან მარტივი ინტერფეისი
- მობილური აპლიკაცია iOS და Android-სთვის
- შესანიშნავი საშუალო და საშუალო სკოლის მოსწავლეებისთვის

ფასი: სრულიად უფასო

Photomath, *photomath.com:*

რა არის: ერთ-ერთი ყველაზე პოპულარული მათემატიკური AI აპლიკაცია (250+ მილიონი ჩამოტვირთვა).

ძირითადი შესაძლებლობები:

- კამერით მათემატიკური პრობლემების მყისიერი ამოცნობა
- ხელნაწერი ამოცნობა და დაბეჭდილი ტექსტის სკანირება
- ინტერაქტიული გრაფიკული კალკულატორი
- ანიმირებული ეტაპობრივი ინსტრუქციები

გამოყენების მაგალითები:

- *მე-6 კლასის მოსწავლე ვერ წყვეტს გამრავლების ამოცანას – ტელეფონით იღებს სახელმძღვანელოს გვერდს, Photomath მყისიერად ასახავს ანიმირებულ გადაწყვეტას*
- *მშობელი ბავშვს ეხმარება საშინაო დავალებაში – ხელნაწერ ამოცანას სკანირებს და ორივე ერთად მიჰყვება ნაბიჯ-ნაბიჯ ახსნას*
- *დამწყები სტუდენტი ამოწმებს თავის პასუხებს და ხედავს, სად შეიძლება შეცდომის დაშვება*

განსხვავება MathGPT-სგან:

- უფრო ორიენტირებული საბაზისო და საშუალო დონის მათემატიკაზე
- უფრო სწრაფი და მარტივი ინტერფეისი
- შესანიშნავი დამწყებთათვის

ფასი: უფასო საბაზისო; Photomath Plus ($9.99/თვე) დეტალური ახსნებისთვის

მასწავლებლის დახმარებისა და გაკვეთილის დაგეგმვისთვის

Diffit, *diffit.me:*

რა არის: AI ინსტრუმენტი, რომელიც ქმნის დიფერენცირებულ წასაკითხ მასალებს და დავალებებს ნებისმიერ თემაზე.

ძირითადი შესაძლებლობები:

- მასალების ადაპტირება სხვადასხვა კითხვის დონისთვის (2-12 კლასები)

- ავტომატური ლექსიკონის სიების, შეკითხვების და თარგმანების გენერირება
- ხმოვანი ფუნქცია (text-to-speech) წასაკითხად
- ინტეგრაცია Google Classroom-თან

გამოყენების მაგალითები:

- *ბიოლოგიის მასწავლებელი ატვირთავს სამეცნიერო სტატიას კლიმატის ცვლილებაზე – Diffit ქმნის სამ ვერსიას: მე-4, მე-7 და მე-10 კლასის დონისთვის, თითო სიაში შესაბამისი ლექსიკით*
- *ESL მოსწავლეებისთვის მასწავლებელი ქმნის გამარტივებულ ტექსტს ძირითადი სიტყვების სიით და ავტომატური თარგმანით*
- *სპეციალური საჭიროებების მქონე მოსწავლისთვის ადაპტირებული საკითხავი მასალა შეიქმნება წუთებში*

გამოყენების შემთხვევები:

- ინგლისურის როგორც მეორე ენის (ESL/ELL) მოსწავლეებისთვის მასალების მომზადება
- რთული სამეცნიერო ტექსტების გამარტივება
- საკითხავი მასალების სწრაფი შექმნა ნებისმიერ თემაზე
- სპეციალური საგანმანათლებლო საჭიროებების მქონე მოსწავლეებისთვის ადაპტირება

ფასი: უფასო 5 რესურსის თვეში; Diffit Premium ($120/წელი) – უსაზღვრო რესურსები

Curipod, *curipod.com:*

რა არის: AI-ით მართული პლატფორმა ინტერაქტიული პრეზენტაციებისა და გაკვეთილების შესაქმნელად.

ძირითადი შესაძლებლობები:

- სრული გაკვეთილის გენერირება წუთებში (სლაიდები, აქტივობები, შეფასება)
- ინტერაქტიული ელემენტები: გამოკითხვები, სიტყვათა ღრუბლები, ღია კითხვები, ხატვის დაფები
- რეალურ დროში მოსწავლეთა პასუხების თვალყურის დევნება
- დიფერენცირებული კითხვები სხვადასხვა დონის მოსწავლეებისთვის

გამოყენების მაგალითები:

- *ისტორიის მასწავლებელი ითხოვს: „შექმენი გაკვეთილი პირველ მსოფლიო ომზე მე-8 კლასისთვის" – **Curipod 2** წუთში ქმნის სრულ სლაიდ-შოუს გამოკითხვებითა და ღია კითხვებით*
- *კლასში მოსწავლეები სმარტფონებით პასუხობენ გამოკითხვებს – მასწავლებელი ეკრანზე ხედავს სიტყვათა ღრუბელს, რომელიც ყველა პასუხის სინთეზს ასახავს*
- *მასწავლებელი ქმნის ინტერაქტიულ გამოცდის გამეორებას, სადაც მოსწავლეები ჯგუფებად ეჯიბრებიან ერთმანეთს*

გამორჩეული თვისებები:

- AI-გენერირებული გრაფიკა და ანიმაციები
- მოსწავლეთა ჩართულობის ანალიტიკა
- თანამშრომლობითი თამაშები და ქვიზები
- ინტეგრაცია Google Slides-თან

ფასი: უფასო საბაზისო გეგმა; Premium ($120/წელი) განუსაზღვრელი AI გენერაციებისთვის

MagicSchool AI, *magicschool.ai:*

რა არის: კომპლექსური AI პლატფორმა მასწავლებლებისთვის 60+ სპეციალიზებული ინსტრუმენტით.

ძირითადი ინსტრუმენტები:

- გაკვეთილის გეგმების გენერატორი სტანდარტების მიხედვით
- რუბრიკის შემქმნელი დეტალური კრიტერიუმებით
- IEP (ინდივიდუალური საგანმანათლებლო გეგმა) მიზნების გენერატორი
- უკუკავშირის მოხსენების დამწერი
- ელექტრონული ფოსტის რესპონდერი მშობლებისთვის

გამოყენების მაგალითები:

- *მასწავლებელი ათ წუთში ქმნის ეროვნულ სტანდარტებზე მორგებულ გაკვეთილის გეგმას სამი სხვადასხვა სირთულის დონით*
- *სპეციალური განათლების სპეციალისტი **AI**-ის დახმარებით ამზადებს **IEP**-ის მიზნებს კონკრეტული მოსწავლის საჭიროებებზე – პროცესი, რომელიც ადრე საათობით გრძელდებოდა, ახლა წუთებს საჭიროებს*
- *მასწავლებელი ამატებს **YouTube** ბმულს – **MagicSchool** ავტომატურად აგენერირებს შესაბამის კითხვებს ვიდეოს შინაარსის მიხედვით*

დამატებითი ფუნქციები:

- YouTube ვიდეოდან კითხვების გენერირება
- ტექსტის გამარტივება სხვადასხვა დონისთვის
- მრავალგამოცდის შეკითხვების შემქმნელი
- მოსწავლეთა მუშაობის კომენტარის გენერატორი

უსაფრთხოება:

- FERPA და COPPA სტანდარტებთან შესაბამისობა
- მონაცემთა პრივატულობის დაცვა
- სათანადო გამოყენების რეკომენდაციები

ფასი: უფასო ძირითადი ვერსია; MagicSchool Plus ($99/წელი)

Copilot (Course Hero AI), *coursehero.com:*

რა არის: AI დამხმარე მასწავლებლებისთვის, რომელიც წვდება Course Hero-ს 100+ მილიონი სასწავლო რესურსის ბაზას.

ძირითადი შესაძლებლობები:

- PowerPoint პრეზენტაციების სრული შექმნა
- სასწავლო ბროშურების (handouts) დიზაინი
- წერის დავალებების და რუბრიკების გენერირება
- გამოცდების და ქვიზების შექმნა

გამოყენების მაგალითები:

- *პროფესორი ითხოვს: „შექმენი ქვიზი ევოლუციის თეორიაზე 20 კითხვით სხვადასხვა ტიპის" – Copilot ასრულებს სამუშაოს წუთებში*
- *მასწავლებელი ამოწმებს მოსწავლის ნაშრომს არსებულ რესურსებთან შესაბამისობაზე პლატფორმის ბაზის გამოყენებით*
- *ლექტორი ქმნის ახალ კურსს და AI ეხმარება handout-ების სწრაფ შედგენაში*

გამორჩეული თვისებები:

- წვდომა არსებულ სტუდენტურ ნამუშევრებთან (შესაბამისობის შემოწმებისთვის)
- საგანმანათლებლო ვიდეოების ბიბლიოთეკა
- თანაავტორობის რეჟიმი – AI ეხმარება, მაგრამ მასწავლებელი აკონტროლებს

ფასი: Course Hero გამოწერა $19.95/თვე ან $119.40/წელი

TeachFX, *teachfx.com:*

რა არის: AI ინსტრუმენტი, რომელიც ანალიზს უკეთებს გაკვეთილის აუდიო ჩანაწერებს და აძლევს მასწავლებელს უკუკავშირს.

ძირითადი შესაძლებლობები:

- ლაპარაკის დროის ანალიზი (მასწავლებელი vs მოსწავლეები)
- მოსწავლეთა ჩართულობის დონის გაზომვა
- კითხვების ტიპების კლასიფიკაცია (ღია vs დახურული)
- ინკლუზიურობის მეტრიკები – რამდენ მოსწავლეს ხმა აქვს

გამოყენების მაგალითები:

- *მასწავლებელი ჩაიწერს გაკვეთილს და TeachFX-ის ანალიზი გვიჩვენებს, რომ გაკვეთილის 80% მასწავლებლის მონოლოგი იყო – ეს ეხმარება პედაგოგს მეტი ინტერაქტიული მეთოდების შეყვანაში*
- *დირექტორი იყენებს TeachFX-ს მასწავლებელთა პროფესიული განვითარების მხარდასაჭერად, მონაცემებზე დაყრდნობით*
- *ახალი მასწავლებელი ყოველკვირეულად აკვირდება საკუთარ გაკვეთილის სტრუქტურას და ხედავს, იზრდება თუ არა მოსწავლეთა ჩართულობა*

რატომ არის სასარგებლო:

- პროფესიული განვითარება მასწავლებლებისთვის
- თვითრეფლექსიის ხელშეწყობა
- სწავლების პრაქტიკის გაუმჯობესება მონაცემების საფუძველზე

ფასი: უფასო საბაზისო ვერსია; სკოლის ლიცენზიები ხელმისაწვდომია

წერა და გრამატიკა

Grammarly, *grammarly.com:*

რა არის: ყველაზე პოპულარული AI-ით მართული წერის დამხმარე, რომელიც გთავაზობთ რეალურ დროში რეკომენდაციებს.

ძირითადი შესაძლებლობები:

- გრამატიკის, მართლწერისა და პუნქტუაციის შემოწმება
- სტილის რეკომენდაციები სიცხადისა და კონციზურობისთვის
- ტონის ანალიზი – როგორ ხმა აქვს თქვენს ტექსტს
- პლაგიატის დეტექტორი (Premium ვერსიაში)

გამოყენების მაგალითები:

- *სტუდენტი წერს საბაკალავრო ნაშრომს – Grammarly რეალურ დროში ასწორებს გრამატიკულ შეცდომებს, გთავაზობს სინონიმებს და ამოწმებს, საკმარისად აკადემიურია თუ არა ტონი*
- *პედაგოგი წერს მშობლებისთვის ელ-ფოსტას –ანალიზი გვიჩვენებს, ზედმეტად ოფიციალურად ხომ არ ჟღერს, და სთავაზობს უფრო „მეგობრულ" ფრაზებს*
- *ინგლისური ენის მოსწავლე Grammarly-ს Chrome extension-ს იყენებს ყველა ტექსტში, ისწავლის საკუთარი შეცდომებიდან*

დამატებითი ფუნქციები:
- GrammarlyGO – AI წერის ასისტენტი სრული ტექსტების გენერირებისთვის
- ლექსიკის გამდიდრების რეკომენდაციები
- ინტეგრაცია ყველა პლატფორმასთან (Gmail, Word, Google Docs, ბრაუზერები)
- მობილური კლავიატურა

აკადემიური წერისთვის:
- აკადემიური ტონის რეჩენდაციები
- ციტირების ფორმატირების დახმარება
- სიტყვათა მრავალფეროვნების შეთავაზებები

ფასი: უფასო საბაზისო; Premium ($12/თვე); Business ($15/თვე პირზე)

Writable, *writable.com:*

რა არის: სპეციალიზებული წერის პლატფორმა სკოლებისთვის AI უკუკავშირით.
ძირითადი შესაძლებლობები:
- AI-ით გენერირებული უკუკავშირი წერის დავალებებზე
- რუბრიკებზე დაფუძნებული შეფასება
- წერის პროცესის თვალყურის დევნება (drafting, revising, editing)
- მოსწავლეთა წინსვლის ვიზუალიზაცია

გამოყენების მაგალითები:
- *ინგლისურის მასწავლებელს 30 სტუდენტის ესე აქვს შესამოწმებელი – Writable AI ავტომატურად ქმნის პირველად კომენტარებს რუბრიკის მიხედვით, მასწავლებელი კი ასწორებს და ეთანხმება საჭიროებისამებრ*
- *მოსწავლე ხედავს საკუთარი პირველი, მეორე და მესამე ვარიანტის განვითარებას გვერდიგვერდ – ეს ეხმარება გაიგოს, სად გაიზარდა*
- *სკოლა ითვლის კლასის მთლიანი წერის პროგრესს სემესტრის განმავლობაში*

მასწავლებლისთვის:
- დროის დაზოგვა შეფასებაში
- სტანდარტიზებული უკუკავშირი
- ინდივიდუალური ანგარიშები მოსწავლეთა განვითარებაზე

ფასი: სკოლის ლიცენზია – საკონტაქტო ფასი

QuillBot, *quillbot.com:*

რა არის: AI ინსტრუმენტი პარაფრაზირებისა და წერის გასაუმჯობესებლად.
ძირითადი შესაძლებლობები:
- პარაფრაზირება 7 სხვადასხვა სტილში (სტანდარტული, ფორმალური, აკადემიური, კრეატიული...)

- გრამატიკის შემოწმება
- summarizer – ტექსტის შეჯამება
- ციტირების გენერატორი (APA, MLA, Chicago)

გამოყენების მაგალითები:

- *სტუდენტი კვლევის ნაშრომისთვის წყაროს ციტირებას ამზადებს – **QuillBot** ავტომატურად ქმნის **APA/MLA** ფორმატის ციტატას, ბმულის ჩასმით*
- *სტუდენტი კომპლექსურ სამეცნიერო ტექსტს ათავსებს **Summarizer**-ში – **AI** ქმნის **200** სიტყვიან შინაარსს, რომელიც ეხმარება ძირითადი შინაარსის გაგებაში*
- *ინგლისურის მოსწავლე ამოწმებს, ეკვივალენტური რჩება თუ არა მისი ფრაზა „ფორმალური" სტილის გადამუშავების შემდეგ*

გამორჩეული თვისებები:

- Co-Writer – AI წერის თანაპადამხმარე სრული ესეებისთვის
- Translator 30+ ენისთვის
- ინტეგრაცია Word და Google Docs-თან

ფასი: უფასო საბაზისო; Premium ($8.33/თვე წლიური გეგმით)

ზოგადი დანიშნულების პლატფორმები

ChatGPT (OpenAI), *chat.openai.com:*

რა არის: ყველაზე პოპულარული და ძლიერი ბუნებრივი ენის AI მოდელი.
ძირითადი შესაძლებლობები:

- კონცეფციების ახსნა ყველა საგნიდან
- ესეების, ანგარიშების, პრეზენტაციების დაწერა
- კოდის წერა და debug-ინგი
- ენების თარგმნა და სწავლება

გამოყენების მაგალითები:

- *მე-11 კლასის მოსწავლეს არ ესმის კვანტური მექანიკა – **ChatGPT**-ს სთხოვს ახსნას **15** წლის ბავშვის დონეზე, ანალოგიებით, და შემდეგ ნელ-ნელა ზრდის სირთულეს*
- *პროგრამირების სტუდენტი კოდს აწვდის **ChatGPT**-ს და ეკითხება „სად არის შეცდომა?" – **AI** ხსნის პრობლემას და სთავაზობს გამოსწორებულ ვერსიას*
- *ისტორიის გაკვეთილზე მოსწავლე **ChatGPT**-ს ელაპარაკება „ნაპოლეონ ბონაპარტის" როლში და ეკითხება ისტორიული პერიოდის შესახებ*

განათლებისთვის:

- სოკრატული დიალოგი კრიტიკული აზროვნების განვითარებისთვის

- სიმულაციები და როლური თამაშები (ისტორიული პერსონაჟები, სამეცნიერო ექსპერიმენტები)
- პრობლემების გადაჭრის ეტაპობრივი მიდგომა
- კრეატიული წერის პროექტების თანაავტორობა

ChatGPT Plus განსაკუთრებულობები:

- GPT-4o – უფრო ძლიერი მოდელი
- ხმოვანი რეჟიმი – საუბარი AI-თან
- სურათების გენერირება DALL-E 3-ით
- Advanced Data Analysis – დანის/Excel ფაილების ანალიზი

რისკები და შეზღუდვები:

- შეუძლია არასწორი ინფორმაციის გენერირება („ჰალუცინაციები“)
- შეიძლება გამოიყენონ აკადემიური გადაცდომებისთვის
- საჭიროებს კრიტიკულ აზროვნებას პასუხების დასადასტურებლად

ფასი: უფასო GPT-3.5; ChatGPT Plus ($20/თვე) GPT-4o-სთვის

Claude (Anthropic), *claude.ai:*

რა არის: ChatGPT-ის კონკურენტი, ცნობილი გრძელი ტექსტების დამუშავებით და უსაფრთხობაზე ფოკუსირებით.

ძირითადი შესაძლებლობები:

- 200,000 tokens კონტექსტური ფანჯარა (≈150,000 სიტყვა)
- მთელი სახელმძღვანელოების, სადისერტაციო ნაშრომების ანალიზი
- რთული დოკუმენტების შეჯამება და ანალიზი
- კოდის დაწერა და განმარტება

გამოყენების მაგალითები:

- *დოქტორანტი ტვირთავს 80-გვერდიან სადისერტაციო ნაშრომს – **Claude** ახდენს სრული ნაშრომის კრიტიკულ ანალიზს, ეძებს ლოგიკურ ხარვეზებს და თავაზობს გამოსწორების მიმართულებებს*
- *მასწავლებელი არჩევს ახალ სახელმძღვანელოს – **Claude** ადარებს ორ განსხვავებულ სახელმძღვანელოს და ახდენს მათ სარგებლიანობის ანალიზს*
- *სამეცნიერო ჟურნალის სტატიის კომპლექსური ნაწილი გამარტივებული ენით ახსნილია, სტუდენტებისთვის მისაწვდომი ვერსიის სახით*

განათლებისთვის:

- კვლევითი ნაშრომების ანალიზი და კრიტიკა
- ლიტერატურული ნაწარმოებების დეტალური ინტერპრეტაცია
- რთული სამეცნიერო ტექსტების გამარტივება
- დიდი მოცულობის შინაარსის ორგანიზება და სტრუქტურირება

უპირატესობები ChatGPT-სთან შედარებით:

- უფრო დიდი კონტექსტური ფანჯარა
- ხშირად უფრო ფრთხილი და დაბალანსებული პასუხები
- უკეთესი გრძელი დოკუმენტების დამუშავება

ფასი: უფასო ვერსია; Claude Pro ($20/თვე)

Microsoft Copilot, *copilot.microsoft.com:*

რა არის: Microsoft-ის AI ასისტენტი, რომელიც ინტეგრირებულია Windows-ში, Edge ბრაუზერში და Microsoft 365 პროდუქტებში.

ძირითადი შესაძლებლობები:

- რეალურ დროში ინტერნეტ სერჩი
- ინტეგრაცია Word, Excel, PowerPoint, Outlook-თან
- სურათების გენერირება (DALL-E 3)
- კოდის დაწერა და ანალიზი

გამოყენების მაგალითები:

- *სკოლის ადმინისტრატორი **Word**-ში წერს ანგარიშს – **Copilot** ეხმარება ტექსტის სტრუქტურირებაში, **Excel**-ის მონაცემებს ავტომატურად ამატებს ცხრილებს*
- *მასწავლებელი ქმნის **PowerPoint** პრეზენტაციას „კლიმატი მე-5 კლასისთვის" – **Copilot 3** წუთში ქმნის სლაიდებს სურათებითა და ბავშვური ენით*
- *სტუდენტი **Teams**-ის ლექციის ჩანაწერიდან ითხოვს ძირითადი შინაარსის შეჯამებას*

სკოლებისთვის:

- სრული ინტეგრაცია Microsoft 365 Education-თან
- მონაცემთა კონფიდენციალურობის მკაცრი დაცვა
- ადმინისტრატორების კონტროლი გამოყენებაზე
- ინტეგრაცია Teams და OneNote-თან

უპირატესობები:

- განახლებული ინფორმაცია ინტერნეტიდან
- FERPA და GDPR სტანდარტებთან შესაბამისობა
- უფასო Microsoft სკოლის ანგარიშით

შეზღუდვები:

- ზოგჯერ ნაკლებად ძლიერი პასუხები ChatGPT-4-თან შედარებით
- ინტერფეისი შეიძლება იყოს გამაბრუებელი

ფასი: უფასო საბაზისო; Copilot Pro ($20/თვე)

Google Gemini (formerly Bard), *gemini.google.com:*

რა არის: Google-ის ბოლო თაობის AI მოდელი, რომელიც ინტეგრირებულია Google-ის ეკოსისტემასთან.

ძირითადი შესაძლებლობები:

- ინტეგრაცია Gmail, Google Docs, Sheets, Drive-თან
- რეალურ დროში ინფორმაცია Google Search-დან
- მულტიმოდალური – ტექსტის, სურათებისა და ხმის დამუშავება
- YouTube ვიდეოების ტრანსკრიპტების ანალიზი

გამოყენების მაგალითები:

- *მოსწავლე **Google Docs**-ში წერს რეფერატს – **Gemini** ეხმარება **Google Scholar**-ის წყაროების მოძიებაში, **Drive**-დან სურათების ჩასმაში და ტექსტის გამართვაში – ყველაფერი ერთ გარემოში*
- *მასწავლებელი **YouTube**-ის საგანმანათლებლო ვიდეოს ბმულს უგზავნის **Gemini**-ს – **AI** ითხოვს ტრანსკრიპტს და ქმნის კომპრეჰენზიის კითხვებს*
- *სტუდენტი ფოტოს ატვირთავს სამეცნიერო დიაგრამისა – **Gemini** ხსნის, რას ასახავს სქემა*

განათლებისთვის:

- კვლევის დახმარება Google Scholar-თან ინტეგრაციით
- Google Workspace-ში დოკუმენტების ავტომატური შექმნა
- მულტიმოდალური სწავლება – სურათების და ტექსტის ერთობლივი ანალიზი

უპირატესობები:

- უფასო (უფრო ძლიერი ვერსიები)
- ინტეგრაცია Google Workspace-თან
- განახლებული ინფორმაცია

ფასი: უფასო; Gemini Advanced ($19.99/თვე, ჩართულია Google One AI Premium-ში)

რჩევები უსაფრთხო და ეფექტური გამოყენებისთვის

მასწავლებლებისთვის:

1. AI-ის გამოყენების პოლიტიკის შემუშავება კლასში
2. AI-ით შექმნილი მასალების ყოველთვის გადახედვა და პერსონალიზაცია
3. მოსწავლეთა ასწავლეთ AI-ის როგორც ინსტრუმენტის გამოყენება, არა მარტივი პასუხების მიმწოდებელი

მოსწავლეებისთვის:

1. AI-ის გამოყენება სწავლის პროცესისთვის, არა შედეგის მისაღებად

2. ყოველთვის ფაქტების გადამოწმება დამოუკიდებელი წყაროებიდან
3. AI-ის გამოყენების გამჭირვალობა მასწავლებლების მიმართ
4. კრიტიკული აზროვნების განვითარება AI პასუხების შეფასებისას

აკადემიური მთლიანობა:

- AI უნდა იყოს დამხმარე ინსტრუმენტი, არა სწავლის შემცვლელი
- მოსწავლეებმა უნდა გაიაზრონ AI-ის როლი თავიანთ ნამუშევარში
- მასწავლებლებმა უნდა ასწავლონ პასუხისმგებლიანი AI გამოყენება

ამ ინსტრუმენტებს სწორად გამოყენებისას შეუძლიათ მნიშვნელოვნად გააუმჯობესონ როგორც სწავლება, ასევე სწავლის პროცესი!

სადისკუსიო თემები:

სოკრატული მეთოდი AI-ში – ნამდვილი სწავლება თუ ილუზია?

Khanmigo იყენებს სოკრატულ მეთოდს – „AI არ აძლევს პირდაპირ პასუხებს, არამედ სვამს მიმართულ კითხვებს.“

საკითხი განსახილველად:

- *შეუძლია თუ არა AI-ს ნამდვილად „სოკრატული“ იყოს, თუ ის მხოლოდ პასუხის სიმულაციას ახდენს?*
- *განსხვავდება თუ არა „ცოცხალი მასწავლებლის“ მიერ დასმული კითხვა AI-ს მიერ დასმული კითხვისგან – და თუ კი, რატომ?*
- *რა შეიძლება იყოს სოკრატული AI-ის ნაკლი სუსტი მოსწავლეებისთვის?*

უფასო vs. ფასიანი AI ინსტრუმენტები – სამართლიანობის საკითხი

დოკუმენტში ჩამოთვლილი ინსტრუმენტების ნაწილი სრულად უფასოა (Socratic by Google), ნაწილი კი ფასიანია (ChatGPT Plus – $20/თვე, MagicSchool Plus – $99/წელი).

საკითხი განსახილველად:

- *ქმნის თუ არა ფასიანი AI ინსტრუმენტების არსებობა ახალი სახის უთანასწორობას მოსწავლეებს შორის?*
- *სკოლებმა უნდა უზრუნველყონ ყველა მოსწავლეს AI ინსტრუმენტებზე თანაბარი წვდომა? ეს ვინ უნდა დააფინანსოს?*
- *„უფასო“ AI ინსტრუმენტი ნიშნავს, რომ ის ნამდვილად უფასოა, თუ სხვა ფორმით „ვიხდით“ (მაგ. მონაცემებით)?*

AI და აკადემიური პატიოსნება – ახალი წესები საჭიროა?

დოკუმენტი აღნიშნავს, რომ ChatGPT „შეიძლება გამოიყენონ აკადემიური გადაცდომებისთვის“ და მოსწავლეებმა „უნდა გამჭირვალობა გამოიჩინონ AI-ს გამოყენებისას.“

საკითხი განსახილველად:

- *სად გადის ზღვარი AI-ს „დასაშვებ დახმარებასა“ და „გადაცდომას“ შორის? ვინ ადგენს ამ ზღვარს?*
- *შეადარეთ AI-ს გამოყენება კალკულატორის გამოყენებას – ერთნაირია ეს სიტუაციები?*
- *როგორ უნდა გამოიყურებოდეს „AI-ის ეპოქის“ შეფასების სისტემა, რომელიც ნამდვილად ზომავს მოსწავლის ცოდნას?*

რეკომენდირებული YouTube რესურსები

პერსონალიზებული სწავლებისთვის

- **Khan Academy** – “How Khanmigo Works” სერია – Khanmigo-ს ფუნქციების პრაქტიკული გიდი მასწავლებლებისა და მოსწავლეებისთვის
- **Quizlet** – “Study Better with Quizlet” – Q-Chat-ის და სასწავლო რეჟიმების გამოყენების სახელმძღვანელო
- **Photomath** – “Step-by-Step Math Solutions” – მათემატიკური პრობლემების გადაჭრის ვიდეო ახსნები

მასწავლებლის დახმარებისა და გაკვეთილის დაგეგმვისთვის

- **MagicSchool AI** – “AI Tools for Teachers” სერია – 60+ ინსტრუმენტის პრაქტიკული დემონსტრაცია კლასში გამოყენებისთვის
- **Curipod** – “Create Interactive Lessons with AI” – ინტერაქტიური გაკვეთილების შექმნის სრული გიდი
- **Diffit** – “Differentiated Reading Materials with AI” – დიფერენცირებული მასალების სწრაფი შექმნის სახელმძღვანელო
- **TeachFX** – “Data-Driven Teaching” – გაკვეთილის ანალიტიკის და თვითრეფლექსიის პრაქტიკული გამოყენება

წერა და გრამატიკა

- **Grammarly** – “Write Better with Grammarly” – აკადემიური და პროფესიული წერის გაუმჯობესების სახელმძღვანელო
- **QuillBot** – “Paraphrasing and Summarizing with AI” – ტექსტის გადამუშავებისა და შეჯამების ტექნიკები

ზოგადი დანიშნულების პლატფორმები

- **OpenAI** – “ChatGPT for Education” სერია – ChatGPT-ის გამოყენება სწავლების სხვადასხვა კონტექსტში

- **Anthropic** – "Claude AI in the Classroom" – Claude-ის გამოყენება დიდი მოცულობის ტექსტების ანალიზისა და კვლევისთვის
- **Microsoft Education** – "Microsoft Copilot for Schools" – Copilot-ის ინტეგრაცია Microsoft 365 Education-ში
- **Google for Education** – "Getting Started with Gemini" – Google Gemini-ის ინტეგრაცია Google Workspace-ში სასწავლო მიზნებისთვის

AI უსაფრთხოება და ეთიკა განათლებაში

- **Common Sense Education** – "AI Literacy for Students" – მოსწავლეთა AI წიგნიერების ამაღლება
- **ISTE** – "Ethical AI Use in Education" – AI-ის პასუხისმგებლიანი გამოყენება სასკოლო გარემოში
- **EdSurge** – "AI Tools: What Teachers Need to Know" – პრაქტიკული რჩევები და გაფრთხილებები AI ინსტრუმენტების გამოყენებისას

ტესტი 2

1. რომელ AI მოდელზეა დაფუძნებული Khanmigo?

ა) Claude
ბ) GPT-4გ) Gemini
დ) Copilot

2. რამდენი ენის მხარდაჭერა აქვს Quizlet-ს?

ა) 50+
ბ) 100+
გ) 130+
დ) 200+

3. რომელი პლატფორმა გვთავაზობს 60-ზე მეტ სპეციალიზებულ ინსტრუმენტს მასწავლებლებისთვის?

ა) Diffit
ბ) Curipod
გ) TeachFX
დ) MagicSchool AI

4. რა ძირითად ფუნქციას ასრულებს TeachFX?

ა) გაკვეთილის გეგმების ავტომატური შექმნა
ბ) გაკვეთილის აუდიო ჩანაწერების ანალიზი და უკუკავშირი
გ) მოსწავლეთა ტესტირება და შეფასება
დ) ვიდეო გაკვეთილების გენერირება

5. Claude-ის (Anthropic) კონტექსტური ფანჯარა რამდენ სიტყვამდეა?

ა) ≈50,000 სიტყვა
ბ) ≈100,000 სიტყვა
გ) ≈150,000 სიტყვა
დ) ≈200,000 სიტყვა

6. რომელი AI ინსტრუმენტი სრულიად უფასოა და შემუშავებულია Google-ის მიერ?

ა) Quizlet
ბ) Photomath
გ) Socratic by Google
დ) MathGPT

7. QuillBot გვთავაზობს პარაფრაზირებას რამდენ სხვადასხვა სტილში?

ა) 3
ბ) 5
გ) 7
დ) 10

8. რომელი ინსტრუმენტი ფოკუსირებულია დიფერენცირებული წასაკითხი მასალების შექმნაზე ESL/ELL მოსწავლეებისთვის?

ა) Writable
ბ) Diffit
გ) Grammarly
დ) Curipod

9. Photomath-ს რამდენი მილიონი ჩამოტვირთვა აქვს?

ა) 100+
ბ) 150+
გ) 200+
დ) 250+

10. მოსწავლეებისთვის რომელი პრინციპია ყველაზე მნიშვნელოვანი AI-ის გამოყენებისას?

ა) AI-ის გამოყენება მხოლოდ საბოლოო შედეგის მისაღებად
ბ) AI-ის გამოყენება სწავლის პროცესისთვის და ფაქტების დამოუკიდებელი გადამოწმება
გ) AI-ის პასუხების უპირობო მიღება
დ) AI-ის გამოყენება მხოლოდ მათემატიკაში

რა არის ხელოვნური ინტელექტი

ჩვენ საკუთარ თავს „ჰომო საპიენსს – მოაზროვნე არსებას“ იმიტომ ვუწოდებთ, რომ ინტელექტი ადამიანის ყველაზე მნიშვნელოვანი მახასიათებელია. ათასობით წლის განმავლობაში ვცდილობდით გაგვეგო, როგორ ვფიქრობთ და ვმოქმედებთ, ანუ როგორ შეუძლია ჩვენს ტვინს, აღიქვას, გაიგოს, და დაგეგმოს ჩვენი ქმედება რეალურ სამყაროში, რომელიც ძალიან ფართოა და კომპლექსური.

ხელოვნური ინტელექტის დანიშნულებაა არა მხოლოდ აღქმა და გაგება, არამედ ისეთი ინტელექტუალური ობიექტების შექმნა, რომელთაც შეუძლია გამოთვალოს/განსაზღვროს, თუ როგორ უნდა ვიმოქმედოთ ეფექტურად და უსაფრთხოდ სტანდარტულ და არასტანდარტულ სიტუაციებში.

მრავალი გამოკითხვა რეგულარულად აფასებს AI-ს, როგორც ერთ-ერთ ყველაზე საინტერესო და სწრაფად მზარდ სფეროს. ხელოვნური ინტელექტის ბევრი ექსპერტი წინასწარმეტყველებს, რომ AI-ს გავლენა „კაცობრიობის ისტორიაში ყველაფერზე მეტი“ იქნება და AI-ს ინტელექტუალური საზღვრები არ არსებობს.

ამჟამად, AI-ს გამოყენებები მოიცავს უამრავ სფეროს: სწავლა, მსჯელობა, აღქმა, ჭადრაკის თამაში, მათემატიკური თეორემების დამტკიცება, ლექსების წერა, მანქანის მართვა ან დაავადებების დიაგნოსტიკა, და ა.შ.) AI არის გამოსადეგი ნებისმიერი ინტელექტუალური ამოცანის შესასრულებლად.

რა არის AI?

მაინც, რა არის AI. ამის ცალსახად თქმა რთულია. არსებობს AI განმარტების რამდენიმე ვერსია:

- ზოგი განსაზღვრავს AI ადამიანის ქმედებებისადმი მსგავსების/იდენტურობის თვალსაზრისით, სხვები ინტელექტის უფრო ფორმალურ განმარტებას ამჯობინებენ, რომლის საფუძველია “რაციონალურობა“, ანუ „სწორი“ გადაწყვეტილების მიღება.
- ზოგი AI შინაგანი სააზროვნო პროცესების და მსჯელობის თვისებად მიიჩნევს, სხვები ყურადღებას ამახვილებენ გარე მახასითებელზე – რაციონალურ ქცევაზე.

ამ ორი განზომილებიდან: ადამიანური/რაციონალური და აზროვნება/ქცევა, იკვეთება ოთხი შესაძლო კომბინაცია და ოთხივეს ყავს მიმდევრები. ცხადია, რომ გამოყენებული მეთოდები განსხვავებულია:

- თუ ძირითად მახასიათებლად მივიჩნევთ ადამიანის მსგავსი ინტელექტის განვითარებას, მაშინ ეს არის ემპირიული მეცნიერება, რომელიც დაკავშირებულია ფსიქოლოგიასთან, და მოიცავს დაკვირვებებსა და ჰიპოთეზებს ადამიანის რეალური ქცევისა და აზროვნების პროცესების შესახებ;

- რაციონალისტური მიდგომა, მეორე მხრივ მოიცავს მათემატიკისა და ინჟინერიის კომბინაციას და ასოცირდება სტატისტიკასთან, მენეჯმენტის თეორიასთან და ეკონომიკასთან.

„ადამიანური ქმედება“ – ტურინგის ტესტი

ალან ტურინგის (1950) მიერ შემოთავაზებული ტესტი გამიზნული იყო როგორც სააზროვნო ექსპერიმენტი.

კითხვა: „შეუძლია თუ არა მანქანას აზროვნება?“ ტესტის შედეგად პასუხია იქნება „კი“, თუ დამკვირვებელი რამდენიმე წერილობითი კითხვის დასმის შემდეგ, ვერ ამბობს დანამდვილებით, პასუხები მოდის ადამიანისგან თუ კომპიუტერისგან.

***ალან მათისონ ტურინგი** (1912–1954) იყო ბრიტანელი მათემატიკოსი, კომპიუტერის მეცნიერი, ლოგიკოსი და კრიპტოანალიტიკოსი, რომელიც მიჩნეულია **თანამედროვე კომპიუტინგისა და ხელოვნური ინტელექტის სფეროს ფუძემდებლად**.*

ძირითადი წვლილი:

- *1936 წელს შეიმუშავა **ტურინგის მანქანის** კონცეფცია – თეორიული გამოთვლითი მოდელი, რომელმაც საფუძველი ჩაუყარა თანამედროვე კომპიუტერის პრინციპებს.*
- *მეორე მსოფლიო ომის დროს გადამწყვეტი როლი ითამაშა გერმანული **Enigma შიფრის** გახსნაში, რამაც, ისტორიკოსთა შეფასებით, ომი წლებით დაამოკლა.*
- *1950 წელს გამოაქვეყნა სტატია **“Computing Machinery and Intelligence”**, სადაც დასვა კითხვა: „შეუძლია თუ არა მანქანას აზროვნება?“ და შემოიტანა ის, რაც მოგვიანებით **ტურინგის ტესტად** იქნა ცნობილი.*

***ტურინგის მემკვიდრეობა:** ყოველწლიურად Association for Computing Machinery (ACM) გასცემს **ტურინგის პრემიას** – კომპიუტერულ მეცნიერებაში ყველაზე პრესტიჟულ ჯილდოს, რომელიც ხშირად „კომპიუტინგის ნობელის პრემიას“ უწოდებენ. 2021 წლიდან ბრიტანეთის £50 ბანკნოტზე გამოსახულია ტურინგის სახე – ეროვნული მნიშვნელობის ნიშნად.*

კომპიუტერის დაპროგრამება ცალსახად განსაზღვრული ტესტის ჩასაბარებლად იძლევა მრავალ შესაძლებლობას. წარმატებით მუშაობისათვის კომპიუტერს დასჭირდება შემდეგი შესაძლებლობები: ბუნებრივი ენის შემუშავება წარმატებული კომუნიკაციისთვის; არსებული ცოდნის ბაზის შექმნა, ავტომატური მსჯელობის წარმართვა კითხვებზე პასუხის გასაცემად და ახალი დასკვნების გამოსატანად;

მანქანური სწავლება არის ახალ გარემოებებთან ადაპტაციისთვის და ცოდნის ბაზის განახლება.

ადამიანური აზროვნება – კოგნიტიური მოდელი

როცა ვამბობთ, რომ პროგრამა ადამიანივით ფიქრობს, უნდა ვიცოდეთ, როგორ ფიქრობს თვით ადამიანი. ჩვენ შეგვიძლია შევისწავლოთ ადამიანის აზროვნების შესახებ სამი გზით:

- ინტროსპექცია – მცდელობა გაანალიზოთ თქვენი ფიქრის პროცესი
- ფსიქოლოგიური ექსპერიმენტები – ადამიანზე დაკვირვება ქმედებაში;
- ტვინის გამოსახულების ანალიზი – ტვინის ქმედებაში დაკვირვება.

მას შემდეგ რაც გვექნება ადამიანის აზროვნებაზე ზუსტი წარმოდგენა, მისი გამოხატვა შესაძლებელი გახდება კომპიუტერული პროგრამის სახით. თუ პროგრამის ქცევა გადაწყვეტილების მიღებისას ემთხვევა ადამიანის შესაბამისი ქცევას, ეს მიუთითებს, რომ პროგრამის ზოგიერთი მექანიზმი ადამინივით მუშაობს.

კოგნიტიური მეცნიერება აერთიანებს ხელოვნური ინტელექტის კომპიუტერულ მოდელებს და ფსიქოლოგიურ ექსპერიმენტებს ადამიანის გონების შესასწავლად. კოგნიტიური მეცნიერება აუცილებლად ეფუძნება რეალურ ცხოვრებაში ადამიანების თუ ცხოველების ქცევების ექსპერიმენტულ კვლევას.

კოგნიტიური მეცნიერება *არის მრავალდისციპლინური სფერო, რომელიც სწავლობს გონებას, ცოდნასა და ინტელექტს – როგორც ბიოლოგიურ, ასევე ხელოვნურ სისტემებში. ის აერთიანებს:*

- ***ფსიქოლოგიას*** *– ადამიანის ქცევის და გონებრივი პროცესების კვლევა*
- ***ნეირომეცნიერებას*** *– ტვინის სტრუქტურისა და ფუნქციის შესწავლა*
- ***ლინგვისტიკას*** *– ენის, მეტყველებისა და გააზრების კვლევა*
- ***ფილოსოფიას*** *– ცნობიერებისა და შემეცნების ბუნების კვლევა*
- ***ხელოვნურ ინტელექტს*** *– გონებრივი პროცესების კომპიუტერული მოდელირება*
- ***ანთროპოლოგიას*** *– ადამიანის კოგნიტიური განვითარების კულტურული და ევოლუციური კონტექსტი*

კოგნიტიური მეცნიერება და AI – ურთიერთგამდიდრება:

კოგნიტიური მეცნიერება და ხელოვნური ინტელექტი ერთმანეთს ორმხრივად ამდიდრებს. ერთი მხრივ, ადამიანის კოგნიტიური პროცესების (მეხსიერება, ყურადღება, აღქმა, გადაწყვეტილების მიღება) კვლევა ეხმარება AI-ს შემქმნელებს უფრო ეფექტური მოდელების შექმნაში. მეორე მხრივ, AI-ს კომპიუტერული მოდელები ეხმარება კოგნიტიურ მეცნიერებს ჰიპოთეზების გამოცდაში: თუ კომპიუტერული მოდელი, რომელიც ამუშავებს ინფორმაციას ისე, როგორც ვვარაუდობთ, რომ ადამიანი ამუშავებს, ქმნის ადამიანის მსგავს შეცდომებს, ეს ემსახურება იმ ჰიპოთეზის დადასტურებას.

პრაქტიკული მაგალითი: ადამიანის მეხსიერება არ ფუნქციონირებს კომპიუტერის მყარი დისკის მსგავსად – ჩვენ ინფორმაციას ვინახავთ კონტე-

ქსტუალურად და ასოციაციურად. სწორედ ამ დაკვირვებამ შთააგონა **ნეირონული ქსელების** არქიტექტურა AI-ში, სადაც ინფორმაცია ნაწილდება კვანძებს შორის ასოციაციური პრინციპით.

ხელოვნური ინტელექტის განვითარების საწყის ეტაპზე ხშირად იყო დაბნეულობა მიდგომებს შორის. ბევრი ავტორი აცხადებდა, რომ თუ ალგორითმი კარგად ასრულებს დავალებას, მაშინ შესაბამისად, კარგი მოდელია ადამიანის საქმიანობისა წარმოდგენისათვის. თანამედროვე ავტორების აზრით, სხვადასხვა მიდგომები ავსებენ ერთმანეთს.

რაციონალური აზროვნება – აზროვნების წესები

არისტოტელე იყო ერთ-ერთი პირველი, ვინც ცდილობდა „სწორი აზროვნების“ სისტემატიზაციას, ანუ მსჯელობის უტყუარი პრინციპების დადგენას. მისი სილოგიზმები, ანუ სწორი პირობებიდან სწორი დასკვნის გაკეთების კანონები არგუმენტაციის მოდელად იქცა. ასეთი მსჯელობის კანონიკური მაგალითები იწყება იმით, რომ „სოკრატე არის ადამიანი და ყველა ადამიანი მოკვდავია და კეთდება დასკვნა, რომ სოკრატე მოკვდავია.

არისტოტელე *(ძვ. წ. 384–322) იყო ბერძენი ფილოსოფოსი – პლატონის მოწაფე და ალექსანდრე მაკედონელის მასწავლებელი. იგი ეწეოდა კვლევებს ფილოსოფიის, ლოგიკის, ბიოლოგიის, ფიზიკის, ეთიკის, პოლიტიკისა და ხელოვნების სფეროებში და დიდ გავლენას ახდენდა დასავლური ფიქრის განვითარებაზე.*

არისტოტელეს ლოგიკა – AI-ს პროტოტიპი?

არისტოტელეს ***სილოგისტიკა*** *– ანუ ფორმალური მსჯელობის სისტემა – შეიძლება მოვიაზროთ, როგორც ხელოვნური ინტელექტის ადრეული კონცეპტუალური ჩანასახი. ამის სამი მიზეზი არსებობს:*

1. ***ფორმალური წესების სისტემა:*** *არისტოტელე პირველი იყო, ვინც ცდილობდა სწორი მსჯელობის* ***ზუსტ, ფორმალურ წესებად*** *ჩამოყალიბებას – ზუსტად ისე, როგორც AI-ს ალგორითმები მუშაობს წინასწარ განსაზღვრული წესების მიხედვით.*
2. ***დასკვნის გამოტანა ავტომატიზაციის შესაძლებლობა:*** *სილოგიზმი – „ყველა A არის B; ყველა B არის C; მაშასადამე, ყველა A არის C” – ეს ლოგიკური სტრუქტურა ისეთია, რომ შეიძლება* ***მექანიკურად*** *გამოყენებულ იქნეს კომპლექსური ამოცანების გადასაჭრელად. სწორედ ეს გახდა* ***ექსპერტული სისტემების*** *(AI-ს ერთ-ერთი ადრეული ფორმის) საფუძველი.*
3. ***ცოდნის სტრუქტურირება:*** *არისტოტელე ასხვავებდა კატეგორიებს, კლასებს და მათ შორის ურთიერთობებს – სწორედ ეს არის* ***ონტოლოგიის*** *(ცოდნის სტრუქტურული წარმოდგენის) საფუძველი, რომელსაც AI-ს სისტემები ეყრდნობა ინფორმაციის ორგანიზებისთვის.*

აზროვნების კანონების შესწავლამ საფუძველი ჩაუყარა მეცნიერების სფეროს, რომელსაც ლოგიკა ჰქვია. მეცხრამეტე საუკუნის ლოგიკოსებმა შეიმუშავეს ლოგიკური აღნიშვნის ზუსტი სისტემა. 1965 წლისთვის პროგრამებს შეეძლო, პრინციპში, ლოგიკურ ნოტაციაში აღწერილი ნებისმიერი ამოხსნადი პრობლემის გადაჭრა. ლოგიცისტური ტრადიციის გამოყენებით ხელოვნური ინტელექტის ფარგლებში მომუშავე მეცნიერები იმედოვნებენ, რომ გამოიყენებენ ასეთ პროგრამებს ინტელექტუალური სისტემების შესაქმნელად.

ლოგიკა მისი ტრადიციული გაგებით მოითხოვს სამყაროს შესახებ სანდო ცოდნას, რაც იშვიათად მიიღწევა სინამდვილეში. ჩვენ უბრალოდ არ ვიცით წესები, რომელსაც ემორჩილება პოლიტიკა ან ომი, ისეთივე სახით, როგორც ვიცით ჭადრაკისა და არითმეტიკის წესები.

ალბათობის თეორია ავსებს ამ ხარვეზს და იძლევა მკაცრი მსჯელობის საშუალებას გაურკვეველი და არასრული ინფორმაციის პირობებში. პრინციპში, ეს საშუალებას გვაძლევს ავაგოთ რაციონალური აზროვნების ყოვლისმომცველი მოდელი, რომელიც საწყისი ინფორმაციადან მიგვიყვანს იმის გაგებამდე, თუ როგორ ვითარდება აზროვნების პროცესი. ამასთან რაციონალური აზროვნება თავისთავად საკმარისი უკვე არ არის, საჭიროა რაციონალური ქცევის შესწავლა.

რაციონალური ქმედება – რაცონალური აგენტი

აგენტი არის ის, რაც მოქმედებს. რა თქმა უნდა ყველა კომპიუტერული პროგრამა რაღაცას აკეთებს, მაგრამ კომპიუტერული აგენტისგან მეტის გაკეთებაა მოსალოდნელი: მუშაობა ავტონომიურად, გარემოს აღქმა, ხანგრძლივი დროის განმავლობაში აქტივობა, ცვლილებებისადმი ადაპტირდება, მიზნების დასხვა და მიღწევა. რაციონალური აგენტი არის ის, ვინც მოქმედებს ისე, რომ მიაღწიოს საუკეთესო შედეგს ან გაურკვევლობის შემთხვევაში საუკეთესო მოსალოდნელი შედეგს.

რაციონალური აგენტი *– ეს არის AI-ს ყველაზე გავრცელებული თანამედროვე მოდელი. მისი გასაგებად სასარგებლოა გავარჩიოთ სამი ძირითადი კომპონენტი:*

1. ***„გარემო – სენსორი – ქმედება" ციკლი:*** *რაციონალური აგენტი გარემოს* ***სენსორებით*** *(ხედვა, სმენა, ციფრული სიგნალები) აღიქვამს, ამ ინფორმაციის საფუძველზე იღებს* ***გადაწყვეტილებას*** *და* ***ეფექტორებით*** *(ხელი, ხმა, კოდი) ახდენს ზეგავლენას გარემოზე. ეს ციკლი მუდმივად მეორდება.*
2. ***„შემოსაზღვრული რაციონალურობა":*** *სრულყოფილი რაციონალურობა – ყოველ სიტუაციაში ზუსტად საუკეთესო გადაწყვეტილების მიღება – პრაქტიკაში შეუძლებელია. ამიტომ თანამედროვე AI სისტემები ეყრდნობა* ***შემოსაზღვრულ რაციონალურობას****: საუკეთესო გადაწყვეტილება* ***ხელმისაწვდომი ინფორმაციის*** *და* ***დროის შეზღუდვის*** *ფარგლებში.*

ხელოვნური ინტელექტისადმი „მსჯელობის კანონების“ მიდგომა ხაზს უსვამს სწორი გადაწყვეტილებების მიღებას. სწორი დასკვნების გამოტანს უნარი ზოგჯერ რაციონალური აგენტის უნარების ნაწილია, რადგან რაციონალურობა ნიშნავს დასკვნას, რომ მოცემული ქმედება საუკეთესოა და შემდეგ ქმედებას ამ დასკვნის მიხედვით. მეორეს მხრივ, არსებობს რაციონალურად ქმედების გზები, რომლებიც არ შეიძლება ითქვას, რომ მოიცავს დასკვნას. მაგალითად, ცხელი ღუმელიდან დაშორება არის რეფლექსური მოქმედება, როგორც წესი, უფრო წარმატებული, ვიდრე ფრთხილად განხილვის შემდეგ მიღებული ნელი ქმედებები.

ცოდნა, წარმოდგენა და მსჯელობა აგენტებს საშუალებას აძლევს მიიღონ კარგი გადაწყვეტილებები. ჩვენთვის სწავლა და განათლება აუცილებელია არა მხოლოდ ერუდიციისთვის, არამედ იმიტომაც, რომ გავაუმჯობესოთ ეფექტური ქცევის გენერირების უნარები, განსაკუთრებით ახალ გარემოებებში.

AI-ს რაციონალური აგენტის მიდგომას ორი უპირატესობა აქვს სხვა მიდგომებთან შედარებით. პირველ რიგში, ეს უფრო ზოგადი მიდგომა, ვიდრე „აზროვნების კანონები“, რადგან სწორი დასკვნა მხოლოდ ერთია და რაციონალურობის მიღწევის რამდენიმე მექანიზმი შეიძლება არსებობდეს.. მეორეც, ის უფრო მისაღებია სამეცნიერო განვითარებისათვის. რაციონალურობის სტანდარტი მათემატიკურად მკაფიოდ არის განსაზღვრული და სრულიად ზოგადი.

რაციონალური აგენტების ტიპები:

ტიპი	*მახასიათებელი*	*მაგალითი*
მარტივი რეფლექსური	*პირდაპირ რეაგირებს სენსორულ შეყვანაზე*	*თერმოსტატი*
მოდელზე დაფუძნებული	*ინახავს გარემოს შიდა მოდელს*	*მართვის ასისტენტი მანქანაში*
მიზნებზე დაფუძნებული	*მოქმედებს კონკრეტული მიზნის მისაღწევად*	*ჭადრაკის AI*
სასარგებლობაზე დაფუძნებული	*ოპტიმიზებს მრავალ მიზანს ერთდროულად*	*ChatGPT, Claude*
მასწავლებელი	*სწავლობს გამოცდილებიდან და იხვეწება*	*DeepMind AlphaGo*

ამ მიზეზების გამო, რაციონალური აგენტური მიდგომა ხელოვნური ინტელექტისადმი ყველაზე ფართოდ არის მიღებული. განვითარების AI-ს პირველ ათწლეულებში რაციონალური აგენტები შეიქმნა ლოგიკურ საფუძველ-

ზე კონკრეტული გეგმების ჩამოყალიბებისათვის და კონკრეტული მიზნების მისაღწევად. შემდგომში მეთოდებმა, რომლებიც დაფუძნებულია ალბათობის თეორიაზე და მანქანურ სწავლაზე, შესაძლებელი გახადა ისეთი აგენტების შექმნა, რომლელთაც შეეძლოთ გადაწყვეტილების მიღება გაურკვევლობაის პირობებში საუკეთესო მოსალოდნელი შედეგის მისაღწევად. ამდენად, AI ორიენტირებულია სწავლაზე და ისეთი აგენტების შექმნაზე, რომლებიც აკეთებენ სწორ გადაწყვეტილებებს და ქმედებებს. ეს არის AI-ს სტანდარტული მოდელი.

რატომ არის ეს მიდგომა განსაკუთრებით ღირებული? *რაციონალური აგენტის მოდელი ერთდროულად* ***მოქნილია*** *(შეიძლება გამოიყენოს ლოგიკა, ალბათობა ან მანქანური სწავლება) და* ***გაზომვადია*** *(შეიძლება შეფასდეს, რამდენად კარგად მიაღწია აგენტმა მიზანს). სწორედ ეს ორი თვისება ხდის მას AI-ს ყველაზე სამეცნიეროდ მყარ საფუძვლად.*

სტანდარტულ მოდელთან დაკავშირებით ერთი მნიშვნელოვანი დაზუსტება უნდა გავაკეთოთ: საჭიროა გავითვალისწინოთ ის ფაქტი, რომ რომ სრულყოფილი რაციონალურობა, ანუ ყოველთვის ზუსტად ოპტიმალური ქმედებების განხორციელება, თითქმის შეუძლებელია კომპლექსურ და სწრაფად ცვალებად გარემოში. ამის გამო, ჩვენ ხშირად პრაქტიკაში ვლაპარაკობთ „შემოსაზღვრულ“ რაციონალურობაზე, ანუ არსებულ სიტუაციაში და არსებული დროის გათვალისწინებით საუკეთესო გადაწყვეტილების მიღებაზე.

რამდენიმე მაგალითი

რამდენად არის შემდეგი სისტემები ხელოვნური ინტელექტის მაგალითები:

1. სუპერმარკეტის შტრიხკოდების სკანერები.
2. ინტერნეტის საძიებო სისტემები.
3. ტელეფონის ხმოვანი მენიუ.
4. ონ-ლაინ კომპიუტერული ქვიზი შეკითხვებით, სადაც მოცემულია შეკითხვები და სავარადო პასუხები, რომელთა შორის ერთირამდენიმე არის სწორი.
5. ონ-ლაინ კომპიუტერული ქვიზი ღია შეკითხვებით, სადაც არაა მოცემული სავარადო პასუხები, და პროგრამა აფასებს მოსწავლის პასუხების შესაბამისობას შეკითხვასთან და მისი პასუხების ლოგიკურობას.
6. ტექსტის მართლწერის და გრამატიკის კორექტირების პროგრამებში.
7. ინტერნეტის მარშრუტიზაციის ალგორითმები.

სავარაუდო პასუხები

1. *არაა AI. მიუხედავად იმისა, რომ შტრიხკოდების სკანირება არის კომპიუტერული ხედვა, ეს არ არის ხელოვნური ინტელექტის სისტემები. შტრიხკოდების წაკითხვის პრობლემა ვიზუალური აღქმის ხელოვნური ფორმაა, რომლის დანიშნულებაა მარტივი და სტრუქტურირებული დამთხვევის აღმოჩენა.*
2. *მრავალი თვალსაზრისით არის AI. ვებგვერდის მოთხოვნასთან შესაბამისობის განსაზღვრის პრობლემა არის ბუნებრივი ენის გაგების პრობლემა რაც AI-ს ერთერთი ნიშანია. საძიებო სისტემები ასევე იყენებენ კლასტერირების ტექნიკას, რაც ასევე AI-ს ერთერთი ნიშანია.*
3. *გარკვეული თვალსაზრისით არის AI. ასეთი მენიუები ჩვეულებრივ იყენებენ ძალიან შეზღუდულ ლექსიკას და არის დიზაინერების კონტროლის ქვეშ, რაც მნიშვნელოვანად ამარტივებს პრობლემას. თანამედროვე ციფრული ასისტენტები, როგორიცაა Siri და Google Assistant იყენებს უფრო მეტ ხელოვნურ ინტელექტის ტექნიკას, მაგრამ შეზღუდული რეპერტუარით.*
4. *არაა AI. პროგრამა წარმოადგენს მონაცემთა ბაზის მართვის მაგალითს, სადაც უბრალოდ ხდება შედარება სწორ პასუხებთან და შესაბამისი დასკვნის გაკეთება.*
5. *მრავალი თვალსაზრისით არის AI. პროგრამა იყენებს ლოგიკური პროცესების ამოცნობას, რაც დაკავშირებულია მონაცემთა დიდი ბაზის გამოყენებასთან და ლოგიკური ბმის აღმოჩენებთან. შეიძლება ითქვას, რომ ასეთი ალგორითმი არის რაციონალური აგენტი.*
6. *გარკვეული თვალსაზრისით არის AI. მართლწერის კორექცია აქ კეთდება სტრიქონების შედარებით და ფიქსირებული ლექსიკონით. გრამატიკული კორექტირება უფრო რთულია, რადგან ის მოითხოვს საკმაოდ რთული წესების გამოყენებას, რომლებიც ასახავს ბუნებრივი ენის სტრუქტურას, მაგრამ მაინც საკმაოდ შეზღუდული და ფიქსირებული დავალება. საძიებო სისტემებში გამოყენებული მართლწერის კორექტორები ბევრად უფრო ახლოსაა AI-სთან.*
7. *ეს არის სასაზღვრო ხაზი და გარკვეულწილად არის AI. გარკვეული დონით შეიძლება ითქვას, რომ ასეთი ალგორითმი არის რაციონალური აგენტი. ამოცანა რთულია და გადაწყვეტილება მიიღება არასრული ინფორმაციის პირობებში. გადაწყვეტილება არ არის გარანტირებულად ოპტიმალური, მაგრამ ლოგიკურია და რაციონალური.*

სადისკუსიო თემები

ტურინგის ტესტი – საკმარისია თუ მოძველდა?

ტურინგის ტესტის მიხედვით, მანქანა „ფიქრობს“, თუ დამკვირვებელი ვერ გა-ნასხვავებს მის პასუხებს ადამიანის პასუხებისგან.

საკითხი განსახილველად:

- *ChatGPT და Claude დღეს ადვილად „ჩააბარებენ“ ტურინგის ტესტს ბევრ სიტუაციაში. ეს ნიშნავს, რომ ისინი ნამდვილად „ფიქრობენ“?*
- *შეიძლება თუ არა, რომ ტურინგის ტესტი ზომავს არა ინტელექტს, არამედ მხოლოდ „ადამიანის მიმსგავსებულ ქცევას“?*
- *დღეს რა ახალი ტესტის შემოტანა იქნებოდა უფრო სამართლიანი AI-ს ინ-ტელექტის შესაფასებლად?*

„რაციონალური აგენტი“ – ადამიანიც ასეა?

დოკუმენტი განმარტავს, რომ რაციონალური აგენტი „მოქმედებს ისე, რომ მი-აღწიოს საუკეთესო შედეგს.“ ამავდროულად, ადამიანებიც ხშირად „შემოსაზ-ღვრული რაციონალურობით“ ვმოქმედებთ.

საკითხი განსახილველად:

- *თუ ადამიანიც და AI-ც „შემოსაზღვრული რაციონალურობით“ ვმოქმედებთ, რა განგვასხვავებს ერთმანეთისგან?*
- *ადამიანი ხშირად არარაციონალურად იქცევა ემოციების, მიკერძოების ან ზარმაცობის გამო. AI-ს „არარაციონალური“ ქცევა საიდან მოდის?*
- *სკოლაში „რაციონალური გადაწყვეტილების მიღების“ სწავლება – ეხმარება ეს მოსწავლეებს AI-ს უკეთ გაგებაში?*

ინტერნეტის საძიებო სისტემა – AI-ა თუ არა?

ტექსტში ნათქვამია, რომ საძიებო სისტემები „მრავალი თვალსაზრისით AI-ია“, ხოლო შტრიხკოდების სკანერი – „არ არის AI.“

საკითხი განსახილველად:

- *სად გადის ზღვარი „ჩვეულებრივ კომპიუტერულ პროგრამასა“ და „ხე-ლოვნურ ინტელექტს“ შორის?*
- *თუ AI-ს განსაზღვრება იცვლება დროთა განმავლობაში (ის, რაც ადრე AI ეგონათ, დღეს „ჩვეულებრივ პროგრამად“ ითვლება), ნიშნავს ეს, რომ AI „სულ გადადის“?*
- *მოიყვანეთ ყოველდღიური ცხოვრებიდან სამი მაგალითი და განიხილეთ: AI-ია თუ არა?*

რეკომენდირებული YouTube რესურსები

ხელოვნური ინტელექტის საფუძვლები

- **CrashCourse Computer Science** – "Artificial Intelligence" (ეპ. 34) – მარტივი, ილუსტრირებული შესავალი AI-ს კონცეფციებში, ტურინგის ტესტის ჩათვლით
- **TED-Ed** – "What is Artificial Intelligence?" – 5-წუთიანი ანიმაციური ახსნა AI-ს ძირითადი პრინციპების შესახებ
- **Kurzgesagt** – "The Last Age of the Machines" – AI-ს ადამიანზე გავლენის ვიზუალური მიმოხილვა

ტურინგის ტესტი და ადამიანური ქმედება

- **Computerphile** – "The Turing Test" – ღრმა ახსნა ტურინგის ტესტის არსისა და მის შემდგომი განვითარების შესახებ
- **TED Talks** – "Can a Computer Pass the Turing Test?" – ბუნებრივი ენის დამუშავებისა და AI-ს ადამიანური ქცევის მიმსგავსების კვლევა
- **BBC Documentary** – "The Imitation Game" სეგმენტები – ალან ტურინგის ცხოვრება და მეცნიერული მემკვიდრეობა

კოგნიტიური მეცნიერება და AI

- **3Blue1Brown** – "Neural Networks" სერია – ნეირონული ქსელების მათემატიკური ინტუიცია; კავშირი ადამიანის ტვინთან
- **MIT OpenCourseWare** – "Introduction to Cognitive Science" – კოგნიტიური მეცნიერების და AI-ს ურთიერთობის აკადემიური კურსი
- **Big Think** – "How Does the Brain Work?" – ნეირომეცნიერებისა და AI-ს შედარება

ლოგიკა, მსჯელობა და რაციონალური აგენტები

- **Philosophy Tube** – "Logic and AI" – არისტოტელეს სილოგისტიკიდან თანამედროვე AI ლოგიკამდე
- **Stanford Online** – "CS221: Artificial Intelligence" – რაციონალური აგენტების კონცეფცია აკადემიური დონეზე
- **Two Minute Papers** – "AI Agents and Decision Making" – ყოველკვირეული მიმოხილვა AI-ს გადაწყვეტილების მიღების ახალი კვლევების შესახებ

ქართულენოვანი რესურსები

- **GITA (Georgian Innovation & Technology Agency)** – AI-ის საფუძვლები და ქართული კონტექსტი
- **Digital Georgia** – ხელოვნური ინტელექტი საქართველოში: შესაძლებლობები და გამოწვევები

ტესტი 3

1. ალან ტურინგმა 1950 წელს შეიმუშავა ტესტი, რომლის მიზანი იყო:

ა) კომპიუტერის სიჩქარის გაზომვა
ბ) პასუხი კითხვაზე „შეუძლია თუ არა მანქანას აზროვნება?“
გ) AI-ს ეთიკური სტანდარტების დადგენა
დ) ადამიანის ტვინის სიმძლავრის გამოთვლა

2. რომელი ფილოსოფოსი ითვლება ლოგიკური სილოგიზმის ფუძემდებლად?

ა) სოკრატე
ბ) პლატონი
გ) არისტოტელე
დ) დეკარტი

3. კოგნიტიური მეცნიერება ძირითადად სწავლობს:

ა) კომპიუტერის პროგრამირების ენებს
ბ) გონებასა და ინტელექტს – ბიოლოგიურ და ხელოვნურ სისტემებში
გ) მხოლოდ ადამიანის ფსიქოლოგიას
დ) ნეირონული ქსელების ტექნიკურ სტრუქტურას

4. ტურინგის ტესტის ჩასაბარებლად კომპიუტერს რომელი შესაძლებლობა არ სჭირდება?

ა) ბუნებრივი ენის შემუშავება
ბ) ავტომატური მსჯელობა
გ) მანქანური სწავლება
დ) ფიზიკური ობიექტების გადაადგილება

5. „შემოსაზღვრული რაციონალურობა” ნიშნავს:

ა) AI-ს უნარს, ნებისმიერ სიტუაციაში ზუსტად საუკეთესო გადაწყვეტილების მიღებას
ბ) საუკეთესო გადაწყვეტილებას ხელმისაწვდომი ინფორმაციისა და დროის ფარგლებში
გ) AI-ს უუნარობას, რთული ამოცანების გადაჭრისას
დ) ადამიანის მსჯელობის სრულ კოპირებას

6. სუპერმარკეტის შტრიხკოდების სკანერი მოცემული დოკუმენტის მიხედვით:

ა) სრულად წარმოადგენს AI-ს
ბ) ნაწილობრივ წარმოადგენს AI-ს
გ) არ წარმოადგენს AI-ს
დ) წარმოადგენს AI-ს მხოლოდ ახალ მოდელებში

7. ადამიანის აზროვნების შესასწავლად დოკუმენტში სამი გზა სახელდება. რომელი მათგანი არ შედის ამ სიაში?

ა) ინტროსპექცია
ბ) ფსიქოლოგიური ექსპერიმენტები
გ) ტვინის გამოსახულების ანალიზი
დ) კომპიუტერული სიმულაცია

8. ლოგიკის სფეროს მეცნიერებმა 1965 წლისთვის მიაღწიეს, რომ:

ა) შეიქმნა პირველი ჰუმანოიდური რობოტი
ბ) პროგრამებს შეეძლო ლოგიკურ ნოტაციაში აღწერილი ნებისმიერი ამოხსნადი პრობლემის გადაჭრა
გ) AI-მ პირველად გაიარა ტურინგის ტესტი
დ) ჭადრაკის AI-მ დაამარცხა ჩემპიონი

9. ონლაინ ქვიზი ღია შეკითხვებით, სადაც პროგრამა აფასებს პასუხების ლოგიკურობას, დოკუმენტის მიხედვით:

ა) სრულად არ წარმოადგენს AI-ს
ბ) მხოლოდ მონაცემთა ბაზის მართვის მაგალითია
გ) მრავალი თვალსაზრისით წარმოადგენს AI-ს
დ) ექსკლუზიურად წარმოადგენს AI-ს

10. რაციონალური აგენტური მიდგომის მთავარი უპირატესობა სხვა მიდგომებთან შედარებით არის:

ა) ის იმეორებს ადამიანის ქცევას სრულყოფილად
ბ) ის უფრო ზოგადია და მათემატიკურად მკაფიოდ განსაზღვრული
გ) ის საჭიროებს ნაკლებ გამოთვლით სიმძლავრეს
დ) ის ყველაზე მარტივი ასახსნელია

როგორ ვითარდებოდა ხელოვნური ინტელექტი

განვიხილოთ რამდენიმე ფუნდამენტური ეტაპი ხელოვნური ინტელექტის, როგორც სამეცნიერო მიმართულების განვითარების ისტორიაში.

1943 წელი: ნეირონული ქსელის პირველი კომპიუტერული მოდელი

კობერნეტიკის მამამ, ნორბერტ ვინერმა, უოლტერ პიტსთან და უორენ მაკკალოხთან ერთად, თეორიულად დაასაბუთა ბიოლოგიურ ნეირონულ ქსელებზე დაფუძნებული ხელოვნური ინტელექტის შექმნის შესაძლებლობები. ასე გაჩნდა ნეირონული ქსელის პირველი კომპიუტერული მოდელი.

ნორბერტ ვინერის შესახებ

ნორბერტ ვინერი (1894—1964) იყო ამერიკელი მათემატიკოსი და მეცნიერი, რომელიც ფართოდ არის ცნობილი, როგორც ***კიბერნეტიკის მამა*** *– მეცნიერების, რომელიც შეისწავლის კონტროლისა და კომუნიკაციის პრინციპებს ცოცხალ ორგანიზმებსა და მანქანებში.*

მისი ძირითადი წვლილი:

- *1948 წელს გამოაქვეყნა ეპოქალური ნაშრომი "Cybernetics: Or Control and Communication in the Animal and the Machine", რომელმაც საფუძველი ჩაუყარა კიბერნეტიკას, როგორც მეცნიერებას.*
- *ვინერი პირველთაგანი იყო, ვინც შეიმუშავა უკუკავშირის (feedback) მათემატიკური თეორია – ანუ იდეა, რომ სისტემამ თვითრეგულირებისთვის უნდა გამოიყენოს გამოსავლის ინფორმაცია.*
- *მან გამოთქვა სერიოზული ეთიკური გაფრთხილებები ავტომატიზაციისა და ხელოვნური ინტელექტის საზოგადოებრივი გავლენის შესახებ – დღეს ეს შეხედულებები განსაკუთრებულ მნიშვნელობას იძენს.*
- *იყო ბავშვი-გენიოსი: 11 წლის ასაკში შევიდა უნივერსიტეტში, 14 წლისა მოიპოვა ბაკალავრის ხარისხი, ხოლო 18 წლის ასაკში – დოქტორის ხარისხი ჰარვარდიდან.*

ვინერის ნაშრომებმა პირდაპირ გავლენა მოახდინა ნეირონული ქსელების, ხელოვნური ინტელექტის, მართვის თეორიისა და კომპიუტერული მეცნიერების განვითარებაზე.

1950 წელი: ტურინგის ტესტი

ალან ტურინგმა გამოაქვეყნა რამდენიმე ნაშრომი იმის შესახებ, შეუძლია თუ არა მანქანას აზროვნება და ასევე შემოგვთავაზა ემპირიული ტესტის იდეა, რომელსაც მოგვიანებით ტურინგის ტესტი ეწოდა.

ტურინგის ტესტის სტანდარტული ინტერპრეტაცია ასეთია: „ადამიანი ურთიერთქმედებს ერთ კომპიუტერთან და ერთ ადამიანთან. კითხვებზე პასუხების საფუძველზე, მან უნდა განსაზღვროს, ვისთან საუბრობს: ადამიანთან

თუ კომპიუტერულ პროგრამასთან. კომპიუტერული პროგრამის ამოცანაა შეცდომაში შეიყვანოს ადამიანი, აიძულოს გააკეთოს არასწორი არჩევანი". 2014 წელს, სანქტ-პეტერბურგში შემუშავებულმა პროგრამამ „ჟენია გუსტმანი" (ევგენი) ტექნიკურად თითქმის ჩააბარა ტიურინგის ტესტი.

პროგრამამ ცამეტი წლის მოზარდის იმიტაცია მოახდინა ოდესიდან, რომელიც „ამტკიცებს, რომ ყველაფერი იცის, მაგრამ ასაკის გამო არაფერი იცის".

„ჟენია გუსტმანი" – დეტალური მიმოხილვა

2014 წელს, სანქტ-პეტერბურგში შემუშავებულმა პროგრამამ „ჟენია გუსტმანი" (ევგენი) ტექნიკურად თითქმის ჩააბარა ტიურინგის ტესტი. ეს მოხდა 2014 წლის 7 ივნისს, ლონდონის სამეფო საზოგადოებაში გამართულ კონკურსზე, ალან ტურინგის გარდაცვალებიდან 60 წლისთავის აღსანიშნავად.

როგორ მუშაობდა პროგრამა: *პროგრამამ ცამეტი წლის მოზარდის იმიტაცია მოახდინა ოდესიდან (უკრაინა), რომელიც „ამტკიცებს, რომ ყველაფერი იცის, მაგრამ ასაკის გამო არაფერი იცის". ეს სტრატეგია ძალიან ჭკვიანური იყო: პროგრამა ამ პერსონაჟის ნიღბის ქვეშ ახდენდა გრამატიკულ შეცდომებს, სიტყვათა ამოტოვებებს და ზოგჯერ არარელევანტურ პასუხებს, რაც ადამიანურ ქცევად მიიჩნეოდა.*

შედეგი: *33 ადამიანი მსაჯულიდან 10-მა (30%-მა) ჩათვალა, რომ „ჟენია" ნამდვილი ადამიანი იყო – ეს კი ტურინგის ტესტის 30%-იანი ბარიერის გადალახვას ნიშნავდა. ორგანიზატორებმა განაცხადეს, რომ ეს ტურინგის ტესტის პირველი წარმატებული გავლაა.*

კრიტიკა: *მეცნიერებმა და AI-ს მკვლევარებმა ეს „გამარჯვება" სკეპტიციზმით შეხვდნენ:*

- *პირობები (შეზღუდული დრო, 5 წუთი თითო მსაჯულზე) ხელოვნურად ამარტივებდა ტესტს.*
- *მოზარდის პერსონაჟის გამოყენება „საბაბს" ქმნიდა ნებისმიერი შეცდომისთვის.*
- *ბევრი ექსპერტი ამბობს, რომ ეს სისტემა „ატყუებდა" ადამიანებს, მაგრამ ნამდვილ ინტელექტს ვერ ავლენდა.*

1956 წელი: ტერმინი „ხელოვნური ინტელექტი"

ტერმინი „ხელოვნური ინტელექტი" პირველად გაისმა დარტმუთის კოლეჯში გამართულ ზაფხულის სემინარზე, რომელიც ორგანიზებული იყო ოთხი ამერიკელი მეცნიერის: ჯონ მაკკარტის, მარვინ მინსკის, ნათანიელ როჩესტერის და კლოდ შენონის მიერ. წარმოდგენილი იყო პირველი ხელოვნური ინტელექტის პროგრამა, Logic Theorist.

როგორ გაჩნდა ტერმინი „ხელოვნური ინტელექტი"

1955 წელს ჯონ მაკკარტიმ (დარტმუთის კოლეჯი), მარვინ მინსკიმ (ჰარვარდი), ნათანიელ როჩესტერმა (IBM) და კლოდ შენონმა (Bell Labs) ერთობლივად შეიმუშავეს წინადადება დარტმუთის ზაფხულის კვლევითი პროექტის შესახებ. სწორედ ამ დოკუმენტში, 1955 წელს, ჯონ მაკკარტიმ პირველად გამოიყენა ფრაზა **„artificial intelligence"** *– ხელოვნური ინტელექტი.*

1956 წლის ზაფხულში ჩატარებულ სემინარზე მაკკარტიმ ეს ტერმინი შეგნებულად ამჯობინა ტერმინ „კიბერნეტიკას", რომელიც ნორბერტ ვინერთან ასოცირდებოდა. მისი განზრახვა იყო, ახალ სფეროს მიეღო დამოუკიდებელი, ახალი სახელი.

სემინარი 6–8 კვირა გაგრძელდა და მონაწილეობდა ათამდე მეცნიერი. მიუხედავად იმისა, რომ ის არ იყო სრულად წარმატებული (ყველა მეცნიერი ბოლომდე არ დარჩა), ეს შეხვედრა ისტორიაში შევიდა, როგორც AI-ს, როგორც ცალკე სამეცნიერო დისციპლინის, დაბადების მომენტი.

1958 წელი: პირველი იმედები

ჰერბერტ საიმონი (ეკონომიკის დარგში ნობელის პრემიის ლაურეატი) ამტკიცებს, რომ მანქანები შეიძლება მომდევნო ათი წლის განმავლობაში მსოფლიო ჩემპიონები გახდნენ ჭადრაკში.

1962 წელი: პირველი მეტყველების ამოცნობის სისტემა

IBM ახდენს Shoebox-ის მეტყველების ამოცნობის სისტემის დემონსტრირებას. მანქანა მიკროფონის საშუალებით იღებდა მეტყველებას და ესმოდა 16 სიტყვა – რიცხვები და რიცხვებთან ოპერაციების ბრძანებები, როგორიცაა „პლიუსი", „მინუს" და „ტოლი".

1965 წელი: გაცრუებული იმედები

ათი წლის ბიჭი, რომელსაც შემდგომ არ ქონია რაიმე განსაკუთრებული წარმატებები ჭადრაკში, იგებს ჭადრაკის მატჩს კომპიუტერთან.

ბიჭი, რომელიც კომპიუტერს ჭადრაკში სძლია

1965 წელს, ათი წლის ბიჭი – რობერტ ნოლდი – ამარცხებს ჭადრაკის კომპიუტერულ პროგრამას, Mac Hack IV-ს. ეს ამბავი განსაკუთრებით სიმბოლური გახდა, რადგან სულ რამდენიმე წლით ადრე, 1958 წელს, ჰერბერტ საიმონი პროგნოზირებდა, რომ ათ წელიწადში კომპიუტერი მსოფლიო ჩემპიონი გახდებოდა ჭადრაკში. ამ ბიჭს, რომლის სახელი ჩვენამდე არ მოაღწია ყველა წყაროში, შემდეგ განსაკუთრებული კარიერა ჭადრაკში არ ჰქონია, რაც კიდევ უფრო მეტ ირონიულ დატვირთვას ანიჭებდა ამ ფაქტს: ჩვეულებრივი ბავშვიც კი ადვილად სძლევდა იმ ეპოქის AI-ს.

ეს მაგალითი კარგად ასახავს ე.წ. ***„AI-ის ზამთრის"*** *(AI Winter) მოახლოებას – პერიოდს, როდესაც გადაჭარბებულმა მოლოდინებმა ვერ გაამართლა და სფეროში ინვესტიციები შემცირდა.*

1966 წელი: პირველი ჩატბოტი

შეიქმნა პირველი ვირტუალური თანამოსაუბრე, „ელიზა“, რომელიც ბაძავს ფსიქოთერაპევტის მეტყველებას, ახორციელებს აქტიური მოსმენის ტექნიკას და აწარმოებს საუბარს ბუნებრივ ენაზე.

როგორ მუშაობდა ELIZA: *ELIZA იყენებდა მარტივ ნიმუშის ამოცნობის (pattern matching) მეთოდს. ყველაზე ცნობილი სცენარი იყო DOCTOR – ფსიქოთერაპევტის როლი. სისტემა კითხვებს უბრუნებდა მომხმარებელს, ანარეკლავდა მის სიტყვებს: მომხმარებელი: „მე ვგრძნობ, რომ დედაჩემი ყოველთვის მაკრიტიკებს.“ ELIZA: „მოგიყეთ მეტი დედათქვენის შესახებ.“*

ELIZA-ს ეფექტი: *ELIZA-მ გამოავლინა საოცარი ფსიქოლოგიური ფენომენი – ადამიანები, მათ შორის ვაიზენბაუმის მდივანი, ნამდვილ სიღრმისეულ გამოცდილებებს უზიარებდნენ პროგრამას და მის „გაგებას“ სჯეროდათ. ეს მოვლენა მოგვიანებით* ***„ELIZA-ს ეფექტი“*** *დაარქვეს – ადამიანის ტენდენცია, ადამიანური თვისებები მიაწეროს კომპიუტერს.*

მნიშვნელობა: *ELIZA ერთ-ერთი პირველი პროგრამა იყო, რომელმაც ტურინგის ტესტის გავლას სცადა (თუმცა ოფიციალური ტესტი არ ჩატარებულა). ვაიზენბაუმი შემდგომ შეშფოთებული იყო, რომ ადამიანები ასე ადვილად „ექცეოდნენ“მარტივი პროგრამის ხაფანგში, და ეს გახდა მისი ცნობილი წიგნის “Computer Power and Human Reason” (1976) ერთ-ერთი მოტივი.*

1970 წელი: ხელოვნური ინტელექტის გამოყენების ცდები განათლებაში

„ელიზას“ წარმატების შემდეგ, მკვლევარმა ჯეიმი კარბონელმა შექმნა პროგრამა სახელწოდებით SCHOLAR, რომელსაც შეუძლია დასვას კითხვები სამხრეთ ამერიკის გეოგრაფიის შესახებ და მიაწოდოს უკუკავშირი ბუნებრივ ენაზე. მოგვიანებით, ასეთ მოწინავე სისტემებს ეწოდა ინტელექტუალური რეპეტიტორობის სისტემები.

ინტელექტუალური რეპეტიტორობის სისტემა დამყარებულია თვითსწავლის იდეაზე და შეიძლება დახასიათდეს შემდეგი მახასიათებლებით:

- ხდება სასწავლო პრობლემების გენერირება
- ხდება სასწავლო პრობლემების გადაჭრა
- დიალოგი მიმდინარეობს ბუნებრივ ენაზე
- ხდება მოსწავლის ცოდნის მოდელირება
- ხდება მოსწავლესა და პროგრამას შორის ურთიერთქმედება და შედეგების ანალიზი

1980-იანი წლების დასაწყისი: ღრმა სწავლების მეთოდები და ექსპერტული სისტემები

ჯონ ჰოპფილდმა და დევიდ რუმელჰარტმა შემოიღეს „ღრმა სწავლების“ მეთოდები, რომლებიც კომპიუტერებს საშუალებას აძლევს ისწავლონ დაგროვებული გამოცდილებიდან. ედვარდ ფეიგენბაუმმა წარმოადგინა „ექსპერტული

სისტემები“, რომლებიც ადამიანის გადაწყვეტილების მიღების იმიტაციას ახდენენ.

1980-იანი წლების ბოლოს: ხელოვნური ინტელექტის კრიზისი

რობოტიკის მკვლევარ ჰანს მორავეცის პარადოქსი ასეა ჩამოყალიბებული: „შედარებით ადვილია ზრდასრული ადამიანის დონის მიღწევა ისეთ ამოცანებში, როგორიცაა ინტელექტის ტესტი ან შაშის თამაში, მაგრამ რთულია ან შეუძლებელია ერთი წლის ბავშვის უნარების მიღწევა აღქმის ან მობილობის ამოცანებში“.

1990-იანი წლები: მანქანური სწავლება

შემუშავებულია მანქანური სწავლების ალგორითმები, რომელთა წყალობითაც კომპიუტერებს შეუძლიათ ცოდნის დაგროვება და საკუთარი გამოცდილებიდან სწავლა.

1997 წელი: ხელოვნური ინტელექტი იმარჯვებს ჭადრაკში

სუპერკომპიუტერი DEEP BLUE თამაშობს ჭადრაკს და ამარცხებს მსოფლიო ჩემპიონებს.

Deep Blue – დეტალური მიმოხილვა

*1997 წელს IBM-ის სუპერკომპიუტერმა **Deep Blue** ისტორიული გამარჯვება მოიპოვა მსოფლიო ჭადრაკის ჩემპიონ **გარი კასპაროვზე** – 3.5 : 2.5 ანგარიშით 6-მატჩიან სერიაში.*

Deep Blue-ს ტექნიკური მახასიათებლები:

- *შეეძლო წამში 200 მილიონამდე პოზიციის გაანალიზება.*
- *იყენებდა სპეციალიზებულ ჭადრაკის ჩიპებს – 480 პარალელურ პროცესორს.*
- *ეყრდნობოდა „ძიების ხეს” (search tree) და ევრისტიკულ შეფასების ფუნქციებს.*

***კონტექსტი:** 1996 წელს კასპაროვმა Deep Blue-ს 4:2-ით სძლია. 1997 წელს IBM-მა სისტემა განაახლა და რემატჩში კასპაროვმა ამჯერად წააგო. კასპაროვმა ბრალი დასდო IBM-ს, რომ ტურნირის განმავლობაში ადამიანები ეხმარებოდნენ კომპიუტერს, IBM-მა კი ეს უარყო.*

***მნიშვნელობა:** ეს გამარჯვება სიმბოლურ მიჯნად იქცა – დამტკიცდა, რომ AI-ს შეუძლია ადამიანს სძლიოს მსოფლიო ელიტის დონის ინტელექტუალურ თამაშში. 1958 წელს საიმონის პროგნოზი, 39 წლის დაგვიანებით, მაინც გამართლდა.*

1990-იანი წლების ბოლოს: ემოციების ამოცნობა

ახალი კვლევითი მიმართულების – აფექტური გამოთვლების გაჩენა, რომელიც მიზნად ისახავს რეაქციების ანალიზს და მათ რეპროდუცირებას მანქანაზე.

2010-იანი წლებიდან დღემდე: დიდი მონაცემები

გამოთვლითი სიმძლავრე საშუალებას იძლევა მოხდეს დიდი მონაცემების – Big Data გაერთიანება ღრმა სწავლებასთან და ნეირონულ ქსელურ მეთოდებთან

Big Data – რას ნიშნავს?

Big Data (დიდი მონაცემები) *ნიშნავს ისეთ მონაცემთა მასივებს, რომლებიც იმდენად დიდია, სწრაფად იზრდება და მრავალფეროვანია, რომ მათი დამუშავება ტრადიციული მონაცემთა ბაზების ინსტრუმენტებით შეუძლებელია.*

1. *Big Data-ს ახასიათებს* ***„5V" პრინციპი:***
2. ***Volume (მოცულობა)*** *– ტერაბაიტებიდან ზეტაბაიტებამდე მონაცემები.*
3. ***Velocity (სიჩქარე)*** *– მონაცემები გენერირდება და მუშავდება რეალურ დროში.*
4. ***Variety (მრავალფეროვნება)*** *– სტრუქტურირებული, ნახევრად სტრუქტურირებული და არასტრუქტურირებული მონაცემები (ტექსტი, სურათი, ვიდეო, სენსორული მონაცემები).*
5. ***Veracity (სანდოობა)*** *– მონაცემების ხარისხი და სიზუსტე.*
6. ***Value (ღირებულება)*** *– მონაცემებიდან სასარგებლო ცოდნის მოპოვება.*

Big Data განათლებაში: *განათლებაში Big Data-ს გამოყენება იძლევა: სტუდენტების სწავლის ტრაექტორიის ანალიზს, სასწავლო მასალის პერსონალიზებას, ჩაგდების რისკის ადრეულ პრევენციას და სასწავლო შედეგების პროგნოზირებას.*
გამოთვლითი სიმძლავრე საშუალებას იძლევა მოხდეს დიდი მონაცემების გაერთიანება ღრმა სწავლებასთან და ნეირონულ ქსელურ მეთოდებთან.

ხელოვნური ინტელექტი და განათლება

საგანმანათლებლო სფეროში ხელოვნური ინტელექტის გამოყენების საკმაოდ ადრეული გამოცდილების ფაქტი (1970 წელი) განსაკუთრებულ ყურადღებას იმსახურებს. ეს ნიშნავს, რომ ამ დროისთვის დაგროვებულია საკმაოდ მყარი თეორიული საფუძველი და ტექნოლოგიების გამოყენებითი გამოყენების მაგალითების დიდი ბაზა.

დიდი მონაცემების წყალობით, ხელოვნური ინტელექტი სწრაფად ვითარდება დღეს განათლების სფეროში: საგანმანათლებლო მონაცემების ინტელექტუალური ანალიზი (საგანმანათლებლო მონაცემების მოპოვება), ასევე მასზე დაფუძნებული პროგნოზირებადი და საგანმანათლებლო ანალიტიკა, პერსონალიზებული სწავლება და ვირტუალური ადაპტური საგანმანათლებლო გარემო რეალური ხდება.

პროგრამა SCHOLAR უფრო დეტალურად

SCHOLAR (1970) იყო არა საგანმანათლებლო პროგრამა, არამედ კომპიუტერული პროგრამა, რომელიც თავისთავად ინტელექტუალური რეპეტიტორი იყო. ეს იყო ერთ-ერთი პირველი ცოდნაზე დაფუძნებული სისტემა და ინტელექტუალური რეპეტიტორობის სისტემის (ITS) პიონერული მაგალითი.

მისი რევოლუციური მიდგომა იყო სემანტიკური ქსელის გამოყენება ცოდნის წარმოსადგენად. სწორი პასუხების უბრალოდ შენახვის ნაცვლად, სისტემა ინახავდა ფაქტებს, კონცეფციებს და მათ შორის ურთიერთობებს მონაცემთა ურთიერთდაკავშირებულ ქსელში. ეს საშუალებას აძლევდა სისტემას, ემსჯელა საგანზე და დინამიურად გენერირებულიყო კითხვები და პასუხები.

ძირითადი ცოდნის სფერო: სამხრეთ ამერიკის ქვეყნების გეოგრაფია (მაგ. დედაქალაქები, მოსახლეობა, საზღვრები, ინდუსტრიები, კლიმატი).

ძირითადი მახასიათებლები და ფუნქციონალურობა:

- შერეული დიალოგი: ეს იყო ყველაზე მოწინავე ფუნქცია. საუბარი არ იყო მკაცრად სტრუქტურირებული. სტუდენტს შეეძლო კითხვების დასმა რეპეტიტორისთვის, ხოლო რეპეტიტორს შეეძლო კითხვების დასმა სტუდენტისთვის, რითაც შეიქმნა დინამიური, სოკრატული დიალოგი.
- ბუნებრივი ენის დამუშავება (პრიმიტიული): სისტემას შეეძლო კითხვების და განცხადებების გაგება და გენერირება ინგლისური ენის ქვესიმრავლეში ნიმუშების შესაბამისობისა და მისი სემანტიკური ქსელის გამოყენებით.
- მოსწავლის ცოდნის მოდელი: SCHOLAR ცდილობდა თვალყური ადევნებინა, თუ რა იცოდა და რა არ იცოდა მოსწავლემ, რაც საშუალებას აძლევდა ყურადღება გაემახვილებინა სისუსტეებზე.
- ადაპტური რეპეტიტორობა: სისტემას შეეძლო თავისი სწავლის სტრატეგიის ადაპტირება მოსწავლის შესრულების მიხედვით. თუ მოსწავლე არასწორად პასუხობდა, სისტემას შეეძლო მინიშნებების მიცემა, სწორი პასუხის მიცემა ან დაკავშირებული თემის განხილვა ცოდნის გასამყარებლად.

როგორ მუშაობდა ტიპიური სესია:

- პროგრამა იწყებოდა მოსწავლისთვის კითხვით (მაგ., „რა არის ბრაზილიის დედაქალაქი?“).
- მოსწავლეს შეეძლო პირდაპირ ეპასუხა.
- მნიშვნელოვანია, რომ მოსწავლეს ასევე შეეძლო ნებისმიერ დროს დაესვა კითხვა (მაგ., „რამდენია არგენტინის მოსახლეობა?“ ან „პერუ ტროპიკებშია?“).
- SCHOLAR აანალიზებს კითხვას, ეძებს შესაბამის ინფორმაციას სემანტიკურ ქსელში და აყალიბებს პასუხს ინგლისურ ენაზე.
- დიალოგის საფუძველზე, SCHOLAR წყვეტს, რა ასწავლოს ან გამოსცადოს შემდეგ.

დიალოგის მაგალითი:

✓ SCHOLAR: რა არის ურუგვაის დედაქალაქი?
✓ სტუდენტი: არ ვიცი. რომელია ურუგვაის უდიდესი ქალაქი?
✓ SCHOLAR: მონტევიდეო ურუგვაის უდიდესი ქალაქია.

- ✓ SCHOLAR: მონტევიდეო ურუგვაის დედაქალაქია?
- ✓ სტუდენტი: დიახ.
- ✓ SCHOLAR: მართალია. რომელი ქვეყანა ესაზღვრება ურუგვაის ჩრდილოეთით?

შეზღუდვები (დღევანდელი სტანდარტებით):

- ცოდნის ბაზა შემოიფარგლებოდა ერთი კონკრეტული დარგით.
- ბუნებრივი ენის გაგება არასტაბილური იყო და მხოლოდ მის დარგთან დაკავშირებულ ფრთხილად შედგენილ წინადადებებთან მუშაობდა.

YouTube რესურსები

1. **"The History of Artificial Intelligence"** – CrashCourse AI სერია (YouTube: CrashCourse) *ხელმისაწვდომია ქართული სუბტიტრებით; მოიცავს ნეირონული ქსელებიდან Deep Blue-მდე*
2. **"Intelligent Tutoring Systems Explained"** – YouTube: EdTech Explained *განმარტავს ITS-ის პრინციპებს და SCHOLAR-ის მემკვიდრეებს*
3. **"Garry Kasparov vs Deep Blue (1997) Full Match"** – YouTube: Chess Archives *ისტორიული მატჩის ჩანაწერი კომენტარებით*
4. **"Alan Turing and the Turing Test"** – YouTube: Computerphile *ახსნილია ტურინგის ტესტის ფილოსოფია და ELIZA-ს ეფექტი*
5. **"ELIZA – The World's First Chatbot"** – YouTube: Science & Futurism *ELIZA-ს ისტორია და მისი გავლენა თანამედროვე AI-ზე*

შენიშვნა: YouTube-ის ბმულები დროთა განმავლობაში შეიძლება შეიცვალოს. გირჩევთ მოძებნოთ ზემოაღნიშნული სათაურები პირდაპირ YouTube-ზე.

სადისკუსიო თემები

AI-ის ეთიკა განათლებაში

შეიძლება თუ არა ხელოვნურმა ინტელექტმა ოდესმე სრულად ჩაანაცვლოს მასწავლებელი?

განიხილეთ: SCHOLAR-მა 1970 წელს გვიჩვენა, რომ AI-ს შეუძლია ინდივიდუალური სწავლება. 50 წელზე მეტი გავიდა – რამდენად შეიცვალა სიტუაცია?

რა არის ის, რასაც ადამიანი-მასწავლებელი სთავაზობს მოსწავლეს, რაც AI-ს არ ძალუძს?

განიხილეთ ემოციური ინტელექტი, მოტივაცია, სოციალური სწავლება.

ტურინგის ტესტის აქტუალობა

კვლავ აქტუალურია თუ არა ტურინგის ტესტი 2020-იანი წლების AI-სთვის?

განიხილეთ: ChatGPT და სხვა თანამედროვე LLM-ები (დიდი ენობრივი მოდელები) ადვილად „ახლავს“ ტურინგის ტესტს.

ნიშნავს თუ არა ეს, რომ ისინი „ფიქრობენ“?

ჩინური ოთახის ფილოსოფიური არგუმენტი (ჯონ სერლი) – სწორია თუ არასწორი?

რა ახალი ტესტები შეიძლება გვჭირდებოდეს ნამდვილი ინტელექტის გასაზომად?

Big Data და კონფიდენციალობა

Big Data-ს გამოყენება სწავლის პერსონალიზაციისთვის – შანსი თუ საფრთხე?

განიხილეთ: Big Data-ს გამოყენება საგანმანათლებლო სისტემებში საშუალებას იძლევა ყოველი მოსწავლის სწავლა „მოირგოს“ მის საჭიროებებზე.

მაგრამ ამავდროულად – ვინ ფლობს ამ მონაცემებს?

როგორ გამოიყენება ისინი?

შეიძლება ბავშვების მონაცემები გამოყენებულ იქნეს მათ წინააღმდეგ? განიხილეთ GDPR-ის მსგავსი კანონმდებლობის მნიშვნელობა.

ტესტი 4.

1. ვინ არის კიბერნეტიკის მამად მიჩნეული?

ა) ალან ტურინგი
ბ) ჯონ მაკკარტი
გ) ნორბერტ ვინერი
დ) მარვინ მინსკი

2. რომელ წელს შეიქმნა ნეირონული ქსელის პირველი კომპიუტერული მოდელი?

ა) 1950
ბ) 1943
გ) 1956
დ) 1966

3. ტურინგის ტესტის არსი არის:

ა) კომპიუტერის სიჩქარის გაზომვა
ბ) კომპიუტერის მეხსიერების ტესტირება
გ) კომპიუტერის უნარი, ადამიანი დაარწმუნოს, რომ ის ადამიანია
დ) ადამიანის IQ-ს გაზომვა

4. „ჟენია გუსტმანი” წარმოადგენდა:

ა) ექიმს
ბ) უნივერსიტეტის სტუდენტს
გ) 13 წლის მოზარდს ოდესიდან
დ) მეცნიერს

5. ტერმინი „ხელოვნური ინტელექტი” პირველად გამოიყენა:

ა) ნორბერტ ვინერმა
ბ) ალან ტურინგმა
გ) ჯონ მაკკარტიმ
დ) კლოდ შენონმა

6. IBM-ის Shoebox სისტემა რამდენ სიტყვას იცნობდა?

ა) 5
ბ) 10
გ) 16
დ) 100

7. ELIZA შეიქმნა:

ა) IBM-ში
ბ) სტენფორდის უნივერსიტეტში
გ) MIT-ში
დ) დარტმუთის კოლეჯში

8. მორავეცის პარადოქსი ამბობს, რომ AI-სთვის:

ა) ჭადრაკი ადვილია, ლოგიკა – რთული
ბ) ჭადრაკი ადვილია, ბავშვის სენსომოტორული უნარები – რთული
გ) ემოციები ადვილია, ლოგიკა – რთული
დ) ენა ადვილია, მათემატიკა – რთული

9. Deep Blue-მ გარი კასპაროვი დაამარცხა:

ა) 1990 წელს
ბ) 1993 წელს
გ) 1997 წელს
დ) 2001 წელს

10. Big Data-ს „5V”-დან რომელი ახასიათებს მონაცემების სიჩქარეს?

ა) Volume
ბ) Variety
გ) Veracity
დ) Velocity

რა არის ტურინგის ტესტი?

ტურინგის ტესტი არის მანქანის უნარების გაზომვის საშუალება, გამოავლინოს ინტელექტუალური ქცევა, რომელიც ადამიანის ქცევის ექვივალენტური ან განურჩეველია. ტესტის ძირითადი იდეა ისაა, რომ თუ ადამიანი გამომცდელი ვერ განასხვავებს მანქანას ადამიანისგან მისი ზეპირი პასუხების საფუძველზე, შეიძლება ითქვას, რომ მანქანას აქვს ხელოვნური ინტელექტი.

ტესტი შემოთავაზებული იყო ცნობილი ბრიტანელი მათემატიკოსის, ლოგიკოსისა და კრიპტოანალიტიკოსის, ალან ტურინგის მიერ თავის 1950 წლის ფუნდამენტურ ნაშრომში „გამოთვლითი მანქანები და ინტელექტი“. ტურინგმა თავდაპირველად მას „იმიტაციის თამაში“ უწოდა.

ტესტის სტანდარტული კონფიგურაცია:

1. **მონაწილეები:** სამი სუბიექტი, ყველა ცალკეულ ოთახებში და ურთიერთობს მხოლოდ ტექსტით (ხმის ან გარეგნობის გამო მიკერძოების თავიდან ასაცილებლად):
   - ადამიანი ინტერვიუერი: ადამიანი, რომელიც სვამს კითხვებს.
   - მანქანა (AI): კომპიუტერული პროგრამა, რომელიც შექმნილია ადამიანის მსგავსი პასუხების გენერირებისთვის და რომელიც პასუხობს დასმულ შეკითხვებს.
   - ადამიანი რესპონდენტი: რეალური ადამიანი, რომელიც პასუხობს (სიმართლეს) დასმულ შეკითხვებს.
2. **პროცესი:** ინტერვიუერი მონაწილეობს ტექსტურ საუბრებში როგორც მანქანასთან, ასევე ადამიანთან, იმის ცოდნის გარეშე, თუ რომელი რომელია.
3. **მიზანი:** გარკვეული დროის ან კითხვების რაოდენობის შემდეგ, ინტერვიუერმა უნდა გადაწყვიტოს, რესპონდენტებიდან რომელია ადამიანი და რომელი მანქანა.
4. **შედეგი:** თუ ინტერვიუერი ვერ შეძლებს მანქანის იდენტიფიცირებას შემთხვევითობის მაღალი მაჩვენებლით (მაგ., მრავალჯერადი ტესტის 50%-ში), ითვლება, რომ მანქანამ წარმატებით გაიარა ტურინგის ტესტი.

აღსანიშნავი მცდელობა: ELIZA (1966)

*ტურინგის ტესტის ერთ-ერთი ყველაზე ადრეული და საინტერესო „მცდელობა“ განხორციელდა ELIZA-ს შემთხვევაში – MIT-ში ჯოზეფ ვაიზენბაუმის მიერ 1966 წელს შექმნილი ჩატბოტის მიერ. ELIZA-ს DOCTOR-ის სცენარი ფსიქოთერაპევტს ბაძავდა, მარტივი ნიმუშის ამოცნობით (pattern matching) – მომხმარებლის სიტყვებს კითხვებად ასახავდა. მიუხედავად ამ პრიმიტიული მექანიზმისა, ვაიზენბაუმი შოკირებული დარჩა: მისი საკუთარი მდივანი, რომელმაც იცოდა, რომ პროგრამასთან ესაუბრებოდა, სთხოვდა ოთახიდან გასვლას, რათა „პირადი“ საუბარი ეწარმოებინა ELIZA-სთან. ეს ფენომენი – **„ELIZA-ს ეფექტი“** – ასახავს ადამიანის ბუნებრივ ტენდენციას, ადამიანური თვისებები მიაწეროს ტექნოლოგიას.*

ტესტის ფილოსოფია და მნიშვნელობა

ტურინგმა ტესტი შექმნა არა როგორც „აზროვნების“ საბოლოო ტექნიკური კრიტერიუმი, არამედ როგორც პრაქტიკული გზა როტორიკული ფილოსოფიური შეკითხვაზე „შეუძლია თუ არა მანქანას აზროვნება?“ პასუხის გასაცემად. ტურინგი ამტკიცებდა, რომ ეს შეკითხვა ძალიან ბუნდოვანი და მეტაფიზიკური იყო.

ფილოსოფიური ახსნის ნაცვლად, მან შემოგვთავაზა მისი შეცვლა გაზომვადი ქცევითი სტანდარტით: თუ მანქანა ინტელექტუალურად მოქმედებს, შეგვიძლია მას ინტელექტუალური ვუწოდოთ.

ძირითადი დასკვნები და კრიტიკა:

- ქცევის წინწამოწევა ცნობიერებასთან შედარებით: ტესტი ფოკუსირებულია მხოლოდ გარე ქცევაზე და არა შინაგან გამოცდილებაზე. არ აქვს მნიშვნელობა, რეალურად ესმის თუ თვითშეგნებულია მანქანა; მნიშვნელოვანია, შეუძლია თუ არა მას გაგების სრულყოფილად სიმულირება.
- ჩინური ოთახის არგუმენტი: ფილოსოფოსმა ჯონ სირლმა წარმოადგინა ცნობილი კონტრარგუმენტი. წარმოიდგინეთ ოთახში არაჩინური ენის მოლაპარაკე, რომელსაც აქვს წესების წიგნი (პროგრამა) ჩინურ იეროგლიფებთან მუშაობისთვის. ოთახის გარეთ მყოფი ადამიანები ჩინურ კითხვებს კითხულობენ ხმამაღლა. ადამიანი წესების წიგნს იყენებს თანმიმდევრული პასუხების შესაქმნელად და შემდეგ პასუხს გასცემს. გარე სამყაროსთვის ოთახი „ჩინურად საუბრობს“, მაგრამ შიგნით მყოფ ადამიანს სიტყვების მნიშვნელობა არ ესმის. ის მხოლოდ პასუხების სიმულაციას ეწევა. სირლი ამტკიცებს, რომ ანალოგიურად, მანქანას, რომელიც ტურინგის ტესტს აბარებს, შეუძლია გაგების სიმულირება ნამდვილი გაგების ან განზრახვის გარეშე.

რას ამტკიცებს ეს არგუმენტი: *სირლი ამბობს, რომ ანალოგიურად, კომპიუტერი, რომელიც ტურინგის ტესტს „ჩააბარებს”, მხოლოდ სიმბოლოებს ამუშავებს სინტაქსური წესების მიხედვით – ის ვერასდროს „გაიგებს” სიტყვების ნამდვილ მნიშვნელობას (სემანტიკას).*
სხვა სიტყვებით: ***სინტაქსი სემანტიკას ვერ გააჩენს.*** *ეს განსხვავება – ნიშნების მანიპულირება vs. ნამდვილი გაგება – ტურინგის ტესტის ფუნდამენტური სუსტი წერტილია სირლის აზრით.*
ამ არგუმენტს ბევრი კრიტიკოსიც ჰყავს: ზოგი ამბობს, რომ ოთახი მთლიანად (ადამიანი + წიგნი + წესები) სისტემის სახით შეიძლება „გაიგებდეს” ჩინურს, თუმცა ინდივიდუალური შემადგენელი ნაწილები ამას ვერ ახდენენ.

- „მოტყუების“ კრიტიკა: კრიტიკოსები ამტკიცებენ, რომ ბევრი პროგრამა ცდილობს ტესტის ჩაბარებას მოტყუების, იუმორის გამოყენებით ან ექსცენტრიულ ადამიანად წარმოჩენით (მაგ., „მე 10 წლის ბიჭი ვარ

საქართველოდან, ამიტომ მაპატიეთ ჩემი ინგლისური"), ნამდვილი ინტელექტის დემონსტრირების ნაცვლად. ეს ტესტს არა მხოლოდ ინტელექტის ტესტად, არამედ მოტყუების ტესტადაც აქცევს.

რატომ არის ეს პრობლემა: *ეს კრიტიკა ავლენს ტესტის სტრუქტურულ სისუსტეს: ტურინგის ტესტი ფაქტობრივად ადამიანების* ***მოტყუების ტესტია****, და არა ინტელექტის გაზომვის ტესტი.*
პროგრამა, რომელიც კარგად „ითამაშებს" გარკვეული ტიპის ადამიანს (უახლოვდება ისეთ ადამიანს, ვის გამოც გარკვეული ქცევა ბუნებრივია), ავტომატურად იმარჯვებს – ინტელექტის ან გაგების გარეშე. სწორედ ამ სტრატეგიას იყენებდა „ჟენია გუსტმანი": 13 წლის უცხოელი ბიჭის სახე ყველა „ნაკლოვანებას" ლოგიკურად ხსნიდა.

- ვიწრო და ფართო ინტელექტი: პროგრამა შეიძლება იყოს მაღალინტელექტუალური კონკრეტულ სფეროში (მაგ., ჭადრაკის თამაში), მაგრამ ვერ ჩააბაროს ტურინგის ტესტს ფართო საღი აზრის ცოდნისა და ადამიანის ვერბალური შესაძლებლობების ნაკლებობის გამო. ტესტი რეალურად შექმნილია იმისთვის, რასაც ახლა ხელოვნურ ზოგად ინტელექტს (AGI) ვუწოდებთ.

კრიტიკის მიუხედავად, ტურინგის ტესტი ხელოვნური ინტელექტის ფუნდამენტურ კონცეფციად რჩება, რომელიც 70 წელზე მეტი ხნის განმავლობაში განსაზღვრავს დარგის კულტურულ და ფილოსოფიურ მიზნებს.

მაგალითები

ჯერჯერობით ვერცერთმა ხელოვნურმა ინტელექტმა ვერ ჩააბარა მკაცრი სამეცნიერო ტურინგის ტესტი, თუნდაც ერთხელ და უპირობოდ. თუმცა, რამდენიმე აღსანიშნავი მცდელობა მიზანს მიუახლოვდა ან მნიშვნელოვანი კამათი გამოიწვია.

1. „ელიზა" (1966)

- რა იყო ეს: ბუნებრივი ენის დამუშავების ადრეული პროგრამა, რომელიც შეიქმნა ჯოზეფ ვაიზენბაუმის მიერ MIT-ში.
- როგორ მუშაობდა: პროგრამა არსებითად რესპონდენტების გამონათქვამებს კითხვებად ასახავდა. (მაგ. რესპონდენტი: „მე მოწყენილი ვარ". ელიზა: „რატომ ამბობ, რომ მოწყენილი ხარ?")
- მნიშვნელობა: მიუხედავად უკიდურესად მარტივი შაბლონების შესაბამისობის სქემისა, ვაიზენბაუმი შოკირებული იყო, როდესაც აღმოაჩინა, რომ ბევრი რესპონდენტი ემოციურად მიეჯაჭვა მანქანას (ელიზას) და სჯეროდა, რომ მანქანას მათი ესმოდა. ეს აჩვენებდა ადამიანის ტენდენციას ტექნოლოგიის ანთროპომორფიზაციისკენ და იყო ადრეული მინიშნება იმისა, რომ ტესტის ჩაბარება შეიძლება უფრო ადვილი ყოფილიყო, ვიდრე მოსალოდნელი იყო.

ჩააბარა „ელიზამ“ ტესტი?

*ფორმალურად – **არა.** ELIZA არასდროს ყოფილა ოფიციალურ ტურინგის ტესტზე წარდგენილი. თუმცა, პრაქტიკულ გამოყენებაში ის ხშირად „ატყუებდა“ გამოუცდელ მომხმარებლებს.*

რატომ ვერ ჩააბარებდა:

- *ELIZA-ს „გაგება” სრულიად ზედაპირული იყო – ის უბრალოდ ამოიცნობდა სიტყვათა შაბლონებს და პრედეფინირებულ პასუხებს აბრუნებდა.*
- *ნებისმიერი გამოცდილი ინტერვიუერი სწრაფად გაამხელდა სისტემას მარტივი კონტექსტური ან ლოგიკური კითხვებით.*
- *ის ახდენდა ემოციური ჩართულობის სიმულირებას, მაგრამ ვერ ახდენდა ნამდვილი გაგების სიმულირებას.*

*ELIZA-ს მნიშვნელობა სხვაგანაა: მან ააშკარავა, რომ **ადამიანები მოტყუებას ადვილად ექვემდებარებიან** – ეს უფრო ადამიანური ფსიქოლოგიის, ვიდრე AI-ის წარმატების ამბავია.*

2. „პერი“ (1972)

- რა იყო ეს: ფსიქიატრ კენეტ კოლბის მიერ შექმნილი პროგრამა პარანოიდული შიზოფრენიით დაავადებული ადამიანის ქცევების სიმულირებისთვის.
- ტესტი: ფსიქიატრთა ჯგუფს სთხოვეს გაეანალიზებინათ როგორც პერის, ასევე რეალურ პარანოიდულ პაციენტებთან საუბრების ჩანაწერები. ისინი ხშირად ვერ ახერხებდნენ პროგრამის ადამიანებისგან გარჩევას.

მნიშვნელობა: ტესტმა აჩვენა, რომ მისი „მოტყუება“ შეიძლებოდა იმ ადამიანის იმიტაციით, რომლის კომუნიკაციის ნიმუშებიც ისედაც უცნაურად ან უჩვეულოდ ითვლებოდა.

ჩააბარა „ელიზამ“ ტესტი?

*ფორმალურად – **არა.** ELIZA არასდროს ყოფილა ოფიციალურ ტურინგის ტესტზე წარდგენილი. თუმცა, პრაქტიკულ გამოყენებაში ის ხშირად „ატყუებდა“ გამოუცდელ მომხმარებლებს.*

რატომ ვერ ჩააბარებდა:

- *ELIZA-ს „გაგება” სრულიად ზედაპირული იყო – ის უბრალოდ ამოიცნობდა სიტყვათა შაბლონებს და პრედეფინირებულ პასუხებს აბრუნებდა.*
- *ნებისმიერი გამოცდილი ინტერვიუერი სწრაფად გაამხელდა სისტემას მარტივი კონტექსტური ან ლოგიკური კითხვებით.*
- *ის ახდენდა ემოციური ჩართულობის სიმულირებას, მაგრამ ვერ ახდენდა ნამდვილი გაგების სიმულირებას.*

*ELIZA-ს მნიშვნელობა სხვაგანაა: მან ააშკარავა, რომ **ადამიანები მოტყუებას ადვილად ექვემდებარებიან** – ეს უფრო ადამიანური ფსიქოლოგიის, ვიდრე AI-ის წარმატების ამბავია.*

3. „ევგენი გუსტმანი“ (2014)

- რა იყო ეს: ჩატბოტი, რომელიც 13 წლის უკრაინელ ბიჭს ასახიერებდა, სურათი, რომელიც გრამატიკული შეცდომებისა და ცოდნის ხარვეზების გასამართლებლად შეირჩა.
- დასკვნა: ლონდონის სამეფო საზოგადოებაში გამართულ პრესტიჟულ ღონისძიებაზე ორგანიზატორებმა განაცხადეს, რომ გუსტმანმა დაარწმუნა მოსამართლეთა 33% (30-დან 10), რომ ის ადამიანი იყო, რითაც „ჩააბარა“ ტურინგის ტესტი (მათ გამოიყენეს 30%-იანი ზღვარი, ტურინგის მიერ 2000 წელს მანქანებზე გაკეთებული კომენტარების თანახმად).
- კამათი: მტკიცება ფართოდ გააკრიტიკა ხელოვნური ინტელექტის საზოგადოებამ. ტესტი არ იყო მკაცრი (საუბრები 5 წუთს გაგრძელდა), ზღვარი თვითნებური იყო და სურათი უაზრო პასუხების გასამართლებლად იაფფასიან ხრიკად აღიქმებოდა. ექსპერტების უმეტესობა ამას რეალურ ჩაბარებად არ მიიჩნევს.

ისევ ჟენია გუსტმანი:

ოდესელი ბიჭის, ჟენიას, სახით შენიღბული ჩატბოტი ისტორიაში პირველი პროგრამა გახდა, რომელმაც ტურინგის ტესტი ჩააბარა. მან მოახერხა პროფესიონალი ჟიური დაერწმუნებინა, რომ ის ადამიანს ელაპარაკებოდა და არა კომპიუტერს.

ხუთმა მანქანამ სცადა თავისი განსაკუთრებულობის დამტკიცება ლონდონის სამეფო საზოგადოების კედლებში მჯდომი მკაცრი ჟიურის წინაშე. ტესტის წესების თანახმად, პროგრამამ ჟიურის მინიმუმ 30% უნდა დაარწმუნოს თავის „ადამიანურობაში“.

კონკურსის დროს ამ პროგრამამ ჟიურის 33% დაარწმუნა, რომ მათთან საუბარი ოდესელი 13 წლის ბიჭი, ევგენი გუსტმანი იყო. უფრო მეტიც, პროგრამას მხოლოდ ხუთი წუთი დასჭირდა ჟიურის მოსატყუებლად.

ჟენია გუსტმანის ტაქტიკის ახსნა:

ჟენია გუსტმანის წარმატება ძირითადად სამ ელემენტს ეყრდნობოდა:

1. ***პერსონაჟის სტრატეგიული არჩევანი:** 13 წლის უცხოელი ბიჭი – ეს ერთდროულად ხსნიდა ასაკობრივ ნაივობას, გრამატიკულ შეცდომებს, ცოდნის ხარვეზებს და უჩვეულო პასუხებს. ნებისმიერი „სისუსტე“ პერსონაჟის ბუნებრივ თვისებად გამოიყურებოდა.*
2. ***დროის ლიმიტი:** მხოლოდ 5-წუთიანი საუბარი არ იძლეოდა ღრმა შემოწმების საშუალებას. ინტერვიუერს არ ჰყოფნიდა დრო, რომ სისტემის შეუსაბამობები გამოევლინა.*
3. ***ემოციური ჩართულობა:** პროგრამა ხშირად „ხუმრობდა“ ან „ბრაზობდა“ – ეს ემოციური რეაქციები ინტერვიუერებს ადამიანური ქცევის ილუზიას უქმნიდა.*

*ექსპერტების უმეტესობა ამ „ჩაბარებას“ რეალურ **AI**-ის მიღწევად არ მიიჩნევს, არამედ – **ჭკვიანურ PR-ღონისძიებად** და ტურინგის ტესტის შეზღუდვების ნათელ დემონსტრაციად.*

4. თანამედროვე LLM (ChatGPT, Gemini და ა.შ.)

- ამჟამინდელი მდგომარეობა: გენერირებად ბუნებრივი ენობრივი მოდელებს (LLM), (როგორიცაა, მაგალითად, ChatGPT) შეუძლიათ წარმოუდგენლად თავისუფლად, შინაარსიან და კონტექსტზე მგრძნობიარე ტექსტის შექმნა თემების ფართო სპექტრზე.
- ჩააბარებენ თუ არა ისინი ტურინგის ტესტს? მოკლე, არაფორმალურ საუბარში მათ შეუძლიათ ბევრი ადამიანის მოტყუება. ეს მათი ძლიერი მხარეების დასტურია.
- რატომ ვერ ახერხებენ ისინი მაინც მკაცრი სამეცნიერო ტესტის ჩაბარებას: მონდომებულ და კვალიფიციურ ინტერვიუერს, როგორც წესი, შეუძლია მათი გამოვლენა.
- საღი აზრის შემოწმება: „თუ წიგნს მაცივარში შევდებ და შემდეგ გამოვიღებ, ახლა ცივია?“ (LLM ამას სწორად გაიგებს, მაგრამ უფრო რთული, მრავალსაფეხურიანი საღი აზრის მსჯელობა შეიძლება ვერ ამოიცნოს)
- თანმიმდევრულობის შემოწმება: საუბრის ადრე განხილულ თემაზე დეტალური დამაზუსტებელი კითხვების დასმა. LLM ზოგჯერ შეიძლება ეწინააღმდეგებოდეს საკუთარ თავს.
- ღრმა გაგების ტესტი: რეალური სუბიექტური გამოცდილების ტესტირება, რომელიც ამოწმებს, ემყარება თუ არა პასუხი ჭეშმარიტ მსჯელობას, თუ სტატისტიკურ პროგნოზირებას.

ტურინგის ტესტმა ხელოვნური ინტელექტის საკითხი აბსტრაქტული ფილოსოფიური კითხვიდან კონკრეტულ ქცევით პრობლემაზე გადაიტანა. მიუხედავად იმისა, რომ მისი, როგორც სამეცნიერო კრიტერიუმის, პრაქტიკული მნიშვნელობა შემცირდა, ის კვლავ ძლიერ აზროვნების ექსპერიმენტად რჩება, რომელიც გვაიძულებს განვსაზღვროთ, თუ რას ვგულისხმობთ „ინტელექტში“, „გაგებაში“ და რას ნიშნავს სინამდვილეში ადამიანობა.

YouTube რესურსები

1. **“The Turing Test: Can Machines Think?”** – YouTube: Computerphile *ნათლად ახსნილია ტურინგის ტესტის ფილოსოფია, ჩინური ოთახის არგუმენტი და LLM-ების გამოწვევები*
2. **“The Chinese Room Argument – John Searle”** – YouTube: Wi Phi (Wireless Philosophy) *ანიმაციური ახსნა სირლის ჩინური ოთახის არგუმენტის შესახებ – ფილოსოფიის სტუდენტებისთვის იდეალური*
3. **“ELIZA: The First Chatbot and the Birth of Human-Computer Interaction”** – YouTube: Science & Futurism *ELIZA-ს ისტორია, ELIZA-ს ეფექტი და მისი მნიშვნელობა AI-ის განვითარებაში*
4. **“Eugene Goostman: Did a Chatbot Really Pass the Turing Test?”** – YouTube: Tech-Linked / Verge *კრიტიკული მიმოხილვა 2014 წლის „გამარჯვების” შესახებ*

5. **"Can ChatGPT Pass the Turing Test?"** – YouTube: Two Minute Papers *თანამედროვე LLM-ების შეფასება ტურინგის ტესტის კრიტერიუმების მიხედვით*

შენიშვნა: *YouTube-ის ბმულები დროთა განმავლობაში შეიძლება შეიცვალოს. გირჩევთ სათაურები პირდაპირ YouTube-ზე მოძებნოთ.*

სადისკუსიო თემები

ჩინური ოთახი – სამართლიანი კრიტიკა თუ ლოგიკური შეცდომა?
ეთანხმებით სირლის არგუმენტს, რომ AI ვერასდროს ვერ „გაიგებს" ნამდვილ მნიშვნელობას?
განიხილეთ: თუ ჩინური ოთახის „სისტემა" მთლიანობაში „გაიგებს" ჩინურს, ხოლო შიგნით მყოფი ადამიანი – არა, ვინ „გაიგებს" ფაქტობრივად? ანალოგიურად: ადამიანის ტვინი ინდივიდუალური ნეირონებისგან შედგება, რომელთაგან არცერთს ინდივიდუალურად არ „ესმის" – მაშ, სად ჩნდება გაგება?
შეიძლება AGI-მ მომავალში გაარღვიოს სირლის „სინტაქსი სემანტიკას ვერ გააჩენს" პრინციპი?

ტურინგის ტესტი 2026 წელს
საჭიროებს თუ არა ინტელექტის გაზომვა ახალ სტანდარტს?
განიხილეთ: ChatGPT და სხვა LLM-ები ყოველდღიურ საუბარში ადვილად „ძლევს" ტურინგის ტესტს.
ნიშნავს თუ არა ეს, რომ ტესტი „გაიარა"?
ან იქნებ ტესტი თავიდანვე არასწორი კითხვა იყო?
რა ახალი ტესტები შეიძლება გვჭირდებოდეს: ემოციური ინტელექტის ტესტი?
ფიზიკური სამყაროს გაგების ტესტი?
კრეატიულობის ტესტი?

„მოტყუების" ეთიკა AI-ში
ეთიკურია თუ არა AI-სისტემის დიზაინი, რომელიც ადამიანის მოტყუებას ისახავს მიზნად?
განიხილეთ: ჟენია გუსტმანი, სოციალური ქსელების ბოტები, „ეკლები" – ყველა ეს სისტემა გათვლილია ადამიანის მოტყუებაზე.
სად გადის ზღვარი ლეგიტიმურ AI-ის დიზაინსა და მანიპულაციას შორის?
რა სახის AI-ს „მოტყუება" შეიძლება გამართლებული იყოს (მაგ. თერაპიული ჩატბოტი) და სად ხდება ეს საზოგადოებრივ საფრთხედ?

ტესტი 5.

1. ვინ შეიმუშავა ტურინგის ტესტი?

ა) ჯონ სირლი
ბ) ჯოზეფ ვაიზენბაუმი
გ) ალან ტურინგი
დ) კენეტ კოლბი

2. ტურინგი თავდაპირველად ტესტს რა სახელს არქმევდა?

ა) ინტელექტის ტესტი
ბ) იმიტაციის თამაში
გ) გამოთვლის ტესტი
დ) მანქანის ტესტი

3. ტურინგის ტესტის სტანდარტული კონფიგურაციაში რამდენი სუბიექტი მონაწილეობს?

ა) ორი
ბ) სამი
გ) ოთხი
დ) ხუთი

4. ჩინური ოთახის არგუმენტი ეკუთვნის:

ა) ალან ტურინგს
ბ) ნორბერტ ვინერს
გ) ჯონ სირლს
დ) მარვინ მინსკის

5. რას ამტკიცებს ჩინური ოთახის არგუმენტი?

ა) AI-ს შეუძლია ნამდვილად გაიგოს ენა
ბ) კომპიუტერები სწრაფად ისწავლიან ენებს
გ) სინტაქსი (ნიშნების მანიპულირება) სემანტიკას (ნამდვილ გაგებას) ვერ გააჩენს
დ) ტურინგის ტესტი სრულყოფილია

6. ELIZA შეიქმნა:

ა) 1950 წელს
ბ) 1972 წელს
გ) 1966 წელს
დ) 1956 წელს

7. „ელიზას ეფექტი” ნიშნავს:

ა) AI-ს უნარს, ტესტი ჩააბაროს
ბ) ადამიანის ტენდენციას, ადამიანური თვისებები მიაწეროს ტექნოლოგიას
გ) კომპიუტერის ემოციურ ინტელექტს
დ) ტურინგის ტესტის წარმატებულ ჩაბარებას

8. „ჟენია გუსტმანი” 2014 წელს ვის ასახიერებდა?

ა) ექიმს
ბ) პროფესორს
გ) 13 წლის ოდესელ ბიჭს
დ) ჟურნალისტს

9. PERRY (1972) შეიქმნა შემდეგი ადამიანის სიმულაციისთვის:

ა) ბავშვი
ბ) მასწავლებელი
გ) პარანოიდული შიზოფრენიით დაავადებული პაციენტი
დ) ჭადრაკის მოთამაშე

10. ტურინგის ტესტის ძირითადი კრიტიკა „ვიწრო და ფართო ინტელექტის” კონტექსტში ნიშნავს:

ა) ტესტი ზედმეტად გრძელია
ბ) ტესტი მხოლოდ ჭადრაკს ამოწმებს
გ) სპეციალიზებული AI შეიძლება ერთ სფეროში ბრწყინვალე იყოს, მაგრამ ტურინგის ტესტს ვერ ჩააბაროს
დ) ტესტს ადამიანებიც ვერ ჩააბარებენ

ხელოვნური ინტელექტი განათლებაში – მიმოხილვა

ხელოვნური ინტელექტის საგანმანათლებლო მიზნებისთვის გამოყენების პირველი მცდელობები უკვე 1970-იან წლებში დაფიქსირდა, როდესაც SCHOLAR სისტემა იქნა წარმოდგენილი. დღეს მას ინტელექტუალური სწავლების სისტემას ვუწოდებთ. საგანმანათლებლო მიზნებისთვის ხელოვნური ინტელექტის ეს ტიპი მხოლოდ ერთ-ერთია დღეს ფართოდ გავრცელებულ და პრაქტიკულ სხვა სახის სისტემებთან ერთად.

მეტყველების გენერირების ადრეული ექსპერიმენტებიდანვე შესამჩნევია, რომ ხელოვნური ინტელექტისგან მოსალოდნელი იყო მასწავლებლის სამუშაო დატვირთვის შემცირება – ცოდნის მასობრივად შემოწმება და სწრაფი უკუკავშირის მიწოდება.

პრაქტიკული მაგალითი: ავტომატური ტესტირება და უკუკავშირი

წარმოიდგინეთ მათემატიკის მასწავლებელი, რომელსაც 120 მოსწავლე ჰყავს. ყოველ კვირას ის ალგებრის საშინაო დავალებას ამოწმებს. ხელოვნური ინტელექტის სისტემის გარეშე ეს 6—8 საათს მოითხოვს. Khan Academy-ის AI-ზე დაფუძნებული სისტემა (Khanmigo) ამოწმებს ყოველ მოსწავლის პასუხს წამებში, განსაზღვრავს, სად მოხდა შეცდომა (მაგ. ნიშნების გამარტივება ვს. განტოლებების ამოხსნა), და მყისიერად უბრუნებს ინდივიდუალურ უკუკავშირს. მასწავლებელი კი ამ დროს ყურადღებას ამახვილებს იმ მოსწავლეებზე, რომლებსაც სისტემური სირთულეები აქვთ.

დღეს განათლებაში ხელოვნური ინტელექტისგან მოველით არა მხოლოდ და არა იმდენად რუტინული სწავლების სამუშაოს შესრულებას, არამედ ერთიანი რეკომენდაციების სისტემის შექმნას და ადაპტური საგანმანათლებლო გარემოს შემუშავებას და მხარდაჭერას, რომელიც ხელს უწყობს საგანმანათლებლო შედეგების მიღწევას, სასწავლო პროცესის პერსონალიზაციას და ჩართულობის დონის ამაღლებას.

ხელოვნური ინტელექტი გამოიყენება სხვადასხვა საგანმანათლებლო კონტექსტში – ტრადიციულ საკლასო ოთახებში, კორპორატიულ გარემოში და მთელი ცხოვრების მანძილზე სწავლის სისტემის დანერგვაში. ამ პარაგრაფში ჩვენ ყურადღებას გავამახვილებთ სამ ძირითად სფეროზე, რომლებიც დიდი ხნის განმავლობაში იყო შესწავლილი: ინტელექტუალური სასწავლო სისტემები, დიალოგზე დაფუძნებული სწავლების სისტემები და ავტომატური წერის შეფასების სისტემები. ასევე მოკლედ აღვწერთ ჰიბრიდულ სისტემებს, რომლებსაც განვითარების პერსპექტივები აქვთ.

- პირველ რიგში, გამჭვირვალობა უაღრესად მნიშვნელოვანია ხელოვნური ინტელექტის გამოყენებისას: ჩვენ უნდა დავინახოთ, თუ როგორ და რატომ მიიღო სისტემამ კონკრეტული გადაწყვეტილება.

- მეორეც, უნდა გვახსოვდეს მჭიდრო კავშირი ხელოვნური ინტელექტის ტექნოლოგიებსა და მონაცემთა მეცნიერების მიდგომებს შორის, რომლებიც ჰიპოთეზებზეა დაფუძნებული. ეს ნიშნავს, რომ ხელოვნური ინტელექტის გამოყენებისას მნიშვნელოვანია დავიწყოთ ჰიპოთეზით, იმის გაგებით, თუ როგორ შემოწმდება ის, რა მონაცემებზე, რამდენად გათვალისწინებულია ტენდენციები და ა.შ.
- მესამე, ხელოვნური ინტელექტი განკუთვნილია ადამიანის სამსახურისათვის, რათა გახდეს მისი პოტენციალის რეალიზებისა და წარმატების მიღწევის ასისტენტი.

ინტელექტუალური სწავლების სისტემა (ITS)

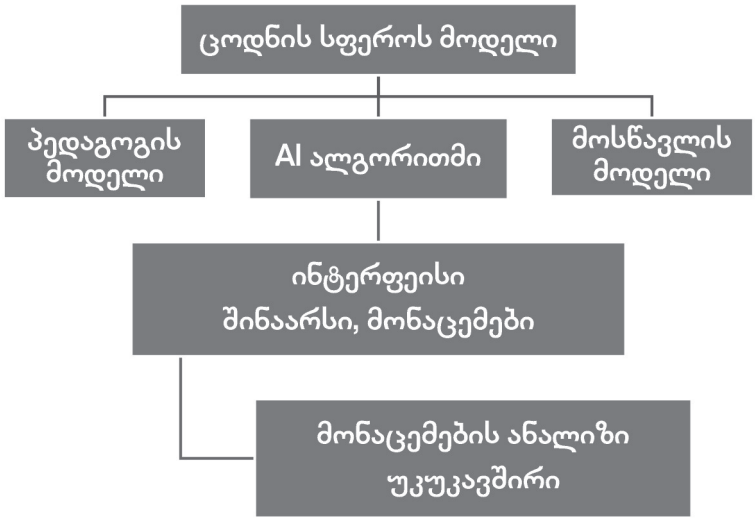
ხელოვნური ინტელექტის განათლებაში ყველაზე გავრცელებული და დიდი ხნის გამოყენებაადია ინტელექტუალური სწავლების/რეპეტიტორობის სისტემა (ITS) – პერსონალიზებული სასწავლო საშუალება, რომელიც აწყობს მასალას მოსწავლის შესაძლებლობებისა და საჭიროებების მიხედვით. ასეთმა სისტემებმა ყველაზე წარმატებით გამოავლინეს შედეგიანობა ზუსტ მეცნიერებებში, როგორც კარგად სტრუქტურირებული ცოდნის სფეროებში.

მოდით, უფრო დეტალურად განვიხილოთ ინტელექტუალური სწავლების სისტემების ფუნქციონირება, რადგან ეს ხელს უწყობს ხელოვნური ინტელექტის ფუნქციონირების პრინციპების უკეთ გაგებას.

ანგარიში Intelligence Unleashed: An Argument for Ai in Education განსაზღვრავს სამ მოდელს, რომლებიც საფუძვლად უდევს ინტელექტუალურ სწავლების სისტემებს:

1. ცოდნის სფეროს მოდელი. ხელოვნურ ინტელექტს სჭირდება ცოდნა შესწავლილი დისციპლინის შესახებ: თემები, მათ შორის კავშირები. რაც უფრო მკაცრი და სტრუქტურირებულია ცოდნის საგანი, მით უფრო ეფექტურად იმუშავებს ხელოვნური ინტელექტი. ამიტომ, მათემატიკა, ფიზიკა, კომპიუტერული მეცნიერებები ყველაზე შესაფერისი საგნებია ხელოვნური ინტელექტის ორგანიზებისთვის.
2. მოსწავლის მოდელი. ხელოვნურ ინტელექტს სჭირდება ცოდნა მოსწავლის შესახებ: მისი წინა მიღწევები, ინფორმაცია მის მიერ განცდილი სირთულეების შესახებ, მისი ემოციური მდგომარეობა და ჩართულობის დონე.
3. პედაგოგიური მოდელი. ამ მოდელთან სამუშაოდ, ხელოვნურ ინტელექტს სჭირდება ცოდნა სწავლების ეფექტური მიდგომების შესახებ: უკუკავშირის მიწოდება, შეფასება, რეკომენდაციები შემდგომი შინაარსისთვის.

მოდით, სქემატურად გამოვსახოთ ტიპიური ITS მოწყობილობა, რათა უფრო ზუსტად გავიგოთ, თუ როგორ ურთიერთქმედებენ მასში წარმოდგენილი მოდელები:



როგორ მუშაობს პრაქტიკაში:

- **ეტაპი 1 – შეფასება:** სისტემა ამოწმებს მოსწავლის არსებულ ცოდნას (ცოდნის მოდელი + მოსწავლის მოდელი).
- **ეტაპი 2 – შინაარსის შერჩევა:** პედაგოგიური მოდელი განსაზღვრავს, რომელი მასალა და სიძნელის რა დონე შეესაბამება მოსწავლეს ამ მომენტში.
- **ეტაპი 3 – ურთიერთქმედება:** მოსწავლე მასალასთან ურთიერთობს – პასუხობს კითხვებს, წყვეტს ამოცანებს, კითხულობს ტექსტებს.
- **ეტაპი 4 – კვალის ანალიზი:** სისტემა ათვალიერებს ყველა ნაბიჯს – სად შეჩერდა მოსწავლე, რამდენჯერ სცადა, რა შეცდომები დაუშვა.
- **ეტაპი 5 – ადაპტაცია:** AI ადაპტებს შემდეგ შინაარსსა და უკუკავშირს – თუ მოსწავლე „ჭედავს“, სისტემა ამარტივებს; თუ ადვილი გამოდგა – ართულებს.

ეს ციკლი მეორდება მანამ, სანამ მოსწავლე სასწავლო მიზანს არ მიაღწევს. მონაცემების გროვება ციკლის ყოველ გავლაზე სისტემას უფრო ზუსტს ხდის.

ხელოვნური ინტელექტის ალგორითმები ამუშავებენ სამი მოდელის მონაცემებს. დამუშავების შედეგები წარმოდგენილია მოსწავლის ინტერფეისში ადაპტური სასწავლო შინაარსის (ტექსტი, ხმა, ვიდეო, ანიმაცია, დავალებები) სახით. როგორც კი მოსწავლე დაიწყებს შინაარსთან ურთიერთქმედებას, ის ტოვებს ციფრულ კვალს, რომელიც ასევე გაანალიზდება ხელოვნური ინტელექტის მეთოდების გამოყენებით. ციფრული კვალის ანალიზის შედეგები წარმოადგენს საფუძველს უკუკავშირისა და სასწავლო შინაარსის ახალი ადაპტაციისთვის.

ამ პროცესის დროს გროვდება მონაცემების დიდი მოცულობა, რომელსაც სისტემა ციკლურად იყენებს დინამიური ოპტიმიზაციისა და თვითგანვითარებისთვის. ციკლი მეორდება მანამ, სანამ მოსწავლე არ მიაღწევს საგანმანათლებლო შედეგს.

დიალოგზე დაფუძნებული სწავლების სისტემები

დავუბრუნდეთ განათლებაში ხელოვნური ინტელექტის პიონერს – SCHOLAR პროგრამას. გარდა იმისა, რომ ეს სისტემა 50 წლის წინ ინდივიდუალურ სასწავლო გამოცდილებას უზრუნველყოფდა და ბუნებრივ ენაზე უკუკავშირს იძლეოდა, ის ასევე მხარს უჭერდა მოსწავლეებთან დიალოგს გაკვეთილის თემაზე. ამრიგად, ეს პროგრამა არა მხოლოდ ინტელექტუალური სწავლების სისტემების, არამედ დიალოგზე დაფუძნებული სწავლების სისტემების პროტოტიპია.

ასეთი სისტემები აგებულია იმავე სქემის მიხედვით, როგორც ინტელექტუალური სწავლების სისტემები – ისინი ეფუძნება პედაგოგიურ მოდელს, მოსწავლის მოდელს და ცოდნის სფეროს მოდელს. თუმცა, განსხვავება ისაა, რომ ასეთი სისტემები არ უზრუნველყოფენ ადაპტურ სწავლების შინაარსს, არამედ ახდენენ სტუდენტებთან დიალოგის იმიტაციას, რათა დაეხმარონ სწორი გადაწყვეტილების პოვნაში, ცოდნის შეფასებასა და მისი დონის დადგენაში, ასევე თემის კონსოლიდაციაში. ამისათვის გამოიყენება ისეთი ტექნოლოგიები, როგორიცაა პასუხის კლასიფიკაცია, სემანტიკური ანალიზი, ანალიზი და ბუნებრივი მეტყველების გენერირება.

მაგალითები

- **AutoTutor** არის დიალოგის გარემო, რომელიც ახდენს მასწავლებელსა და მოსწავლეს შორის საგანმანათლებლო დიალოგის სიმულირებას ონლაინ დავალებების ეტაპობრივი შესრულების პროცესში. პროგრამის მიზანია შესასწავლი თემის ეტაპობრივად ჩაღრმავება.
- **Watson Tutor** არის დიალოგზე დაფუძნებული სწავლების სისტემა, რომელიც შემუშავებულია Pearson-ისა და IBM-ის მიერ უნივერსიტეტებისთვის. პროგრამა გთავაზობთ დამატებით მასალებს, აკონტროლებს პროგრესს და პასუხების მიხედვით ახდენს საუბრის ადაპტაციას.
- **Khanmigo (Khan Academy)** – ChatGPT-ზე დაფუძნებული სოკრატული ტუტორი, რომელიც პასუხებს პირდაპირ არ იძლევა, მაგრამ მოსწავლეს კითხვებით მიჰყავს სწორ პასუხამდე. მაგ.: „რატომ ფიქრობ, რომ ეს ნაბიჯი სწორია?“ – ეს მიდგომა ღრმა გაგებას ახდენს სტიმულირებას.
- **Duolingo Max** – ენის შემსწავლელი პლატფორმის AI ფუნქცია, რომელიც GPT-4-ზეა დაფუძნებული. მომხმარებელი ეწევა “roleplay” საუბრებს სიმულირებულ პერსონაჟებთან, მაგ. პარიზელ კაფეში ყავის შეკვეთა ფრანგულად – სისტემა ასწორებს, განმარტავს და ადაპტებს.
- **Socratic (Google)** – სტუდენტი ფოტოს იღებს საშინაო დავალების კითხვაზე, სისტემა ახდენს გამოსახულების ამოცნობას, განსაზღვრავს კითხვის ტიპს და ახსნით, ეტაპობრივი მინიშნებებით პასუხობს – სოკრატული მიდგომით.

საძიებო/კვლევითი გარემო

საძიებო გარემო, ინტელექტუალური ეტაპობრივი სწავლების სისტემებისა და დიალოგზე დაფუძნებული სწავლების სისტემებისგან განსხვავებით, უფრო თავისუფალი და არასტრუქტურირებული სწავლების სფეროა, რომელიც აქტიურ სწავლებას უწყობს ხელს. საძიებო გარემოსთან ურთიერთქმედება სისტემის სივრცეში თავისუფალ და დამოუკიდებელ ნავიგაციას ჰგავს გარკვეული დავალებული ამოცანების გადასაჭრელად; სისტემას შეუძლია მოსწავლის მოთხოვნით გარკვეული მინიშნებების მიცემა.

საძიებო გარემოს შექმნაში გამოყენებულ სპეციფიკურ ტექნოლოგიებს შორისაა სწავლა, რომელიც წინა სესიებიდან მიღებული მონაცემების საფუძველზე ბაიესის ქსელების გამოყენებით ხორციელდება. ბაიესის ქსელი არის გრაფიკული მოდელი კიდეებითა და კვანძებით, რომელიც ხელს უწყობს მოვლენის მოხდენის ალბათობის გამოთვლას ცვლადების ნაკრებიდან გამომდინარე.

ასეთი ხელოვნური ინტელექტის გამოყენების სირთულე იმაში მდგომარეობს, რომ რთულია მოსწავლის მოდელის შექმნა ზუსტად სასწავლო გარემოსთან ურთიერთქმედების შეუზღუდავი შესაძლებლობების გამო, შესაბამისად, რთულია სისტემის ეფექტურობის გამოთვლა.

მაგალითები

- **Betty's Brain** არის სწავლის სისტემა, სწავლების გზით, სადაც მოსწავლეები ვირტუალური მსმენელის, ბეტის, მასწავლებლების როლს ასრულებენ: ისინი უქმნიან ბეტის სწავლის კონცეპტუალურ რუკას, აწყობენ შუალედურ შემოწმებას და შემდეგ უყურებენ, თუ როგორ აბარებს ბეტი გამოცდას სისტემის მიერ ავტომატურად გენერირებული კითხვებიდან.
- **Crystal Island** არის ინტერაქტიული თამაში, რომელშიც მომხმარებლები იკვლევენ იდუმალ ეპიდემიას შორეულ კუნძულზე, პრაქტიკაში იყენებენ სამეცნიერო კვლევის მეთოდებს. პროგრამა უზრუნველყოფს დამხმარე უკუკავშირს და ითვალისწინებს მოსწავლეთა შესახებ აფექტურ მონაცემებს. (აფექტური მონაცემები არის მონაცემები მოსწავლის ემოციური მდგომარეობის შესახებ: დადებითი ან უარყოფითი).
- **MinecraftEdu / Minecraft: Education Edition** – Microsoft-ის საგანმანათლებლო ვერსია პოპულარული თამაშისა, სადაც სტუდენტები ქმნიან, იკვლევენ და წყვეტენ პრობლემებს ვირტუალურ სამყაროში. გამოიყენება ისტორიის, გეოგრაფიის, არქიტექტურის, ფიზიკისა და კოდირების სასწავლებლად. AI-ის კომპონენტი ეხმარება მასწავლებელს სტუდენტების აქტივობის ანალიზში.
- **iCivics** – ვებ-თამაშების სისტემა სამოქალაქო განათლებისთვის, სადაც სტუდენტები „თამაშობენ" მოსამართლეს, კონგრესმენს ან პრეზიდენტს და პრაქტიკაში იყენებენ კანონმდებლობის ცოდნას. AI ადაპტებს სიძნელეს და გვთავაზობს შემდეგ ნაბიჯებს

წერითი ნაშრომის ავტომატური შეფასება

ეს განათლებაში ხელოვნური ინტელექტის გამოყენების ძალიან გავრცელებული სფეროა, რომელიც საშუალებას იძლევა შემცირდეს მასწავლებლების სამუშაო დატვირთვა, გაიზარდოს პრაქტიკული დავალებების შემოწმების სიჩქარე და გაუმჯობესდეს შეფასებების სანდოობა და ობიექტურობა.

წერილობითი დავალებების ავტომატურ შემოწმებაში ყველაზე საინტერესო მოვლენები არ არის დაკავშირებული საბოლოო შეფასებასთან (ავტომატური ქულების დათვლის ამოცანა შეიძლება შესრულდეს ხელოვნური ინტელექტის გარეშე), არამედ ფორმაციულ შეფასებასთან. ჩვენ ვსაუბრობთ დიდი წერილობითი დავალებების შემოწმებაზე, რასაც შეიძლება დიდი დრო დასჭირდეს და ამ მიზეზით, უკუკავშირი ხშირად მწირი ან არაინდივიდუალურია. ხელოვნური ინტელექტის გამოყენების ამ სფეროში სპეციფიკური ტექნოლოგიებია მანქანური სწავლება როგორც მასწავლებლით, ასევე მის გარეშე, ასევე ბუნებრივი ენის სემანტიკის დამუშავება.

მაგალითები

- **Revision Assistant** არის პროგრამა მოკლე ესეების შესაფასებლად და კომენტირებისთვის, შექმნილი Turnitin-ის დეველოპერების მიერ, რომელიც ცნობილია პლაგიატზე ნაშრომების შემოწმების გადაწყვეტილებებით. სისტემა ავტომატურად აფასებს ესეებს და უზრუნველყოფს უკუკავშირს, რომელიც გენერირდება ექსპერტების მიერ წინასწარ დაწერილი ათასობით კომენტარის ანალიზის საფუძველზე.
- **OpenEssayist** არის სისტემა, რომელიც შემუშავებულია ოქსფორდის უნივერსიტეტის მიერ. სისტემის მიზანია მოსწავლეს მისცეს დეტალური უკუკავშირი წერილობით ნაშრომზე, გააუმჯობესოს მათი წერის, თვითსწავლებისა და რეფლექსიის უნარები.
- **Grammarly for Education** – AI-ზე დაფუძნებული სისტემა, რომელიც მხოლოდ გრამატიკას კი არ ასწორებს, არამედ გვთავაზობს სტილისტური, ლოგიკური და სტრუქტურული გაუმჯობესების რეკომენდაციებს. ინსტიტუციური ვერსია მასწავლებელს ზოგად ანალიტიკასაც აწვდის.
- **ETS e-rater** – Educational Testing Service-ის სისტემა, რომელიც გამოიყენება GRE, TOEFL და სხვა სტანდარტიზებულ გამოცდებში ესეების ავტომატური შეფასებისთვის. აფასებს 12-ზე მეტ ასპექტს: გრამატიკა, ლექსიკა, სტილი, ორგანიზაცია, განვითარება.

Turnitin Feedback Studio – პლაგიატის შემოწმებასთან ერთად, სისტემა ახდენს ესეების სტრუქტურის, არგუმენტაციის და წყაროების გამოყენების ანალიზს, უზრუნველყოფს ვიზუალური მარკირებით კომენტირებას.

პრაქტიკული მაგალითი

წარმოიდგინეთ უნივერსიტეტის ლიტერატურის კურსი, სადაც 200 სტუდენტი ყოველ კვირა 500-სიტყვიანი ესეს წერს. ხელოვნური ინტელექტის გარეშე პროფესორი ამ ყველაფრის შემოწმებას 40+ საათს დაუთმობდა. OpenEssayist-ის მსგავსი სისტემა წუთებში: (1) ამოიცნობს ძირითად არგუმენტს, (2) ადარებს სათაურის მოთხოვნებს, (3) განსაზღვრავს სტრუქტურულ სისუსტეებს და (4) სტუდენტს უბრუნებს კონკრეტულ, ფოკუსირებულ კომენტარებს. პროფესორი კი ყურადღებას ამახვილებს ყველაზე სუსტ ნაშრომებზე.

ჰიბრიდული ხელოვნური ინტელექტის სისტემები

ეს განათლებაში ხელოვნური ინტელექტის გამოყენების ძალიან პოპულარული და პერსპექტიული სფეროა. განათლებაში ხელოვნური ინტელექტის გამოყენების სამი ძირითადი სფეროს (ინტელექტუალური რეპეტიტორობის სისტემები, დიალოგზე დაფუძნებული სწავლების სისტემები და კვლევითი გარემო) ინსტრუმენტებისა და ტექნოლოგიების გაერთიანებით, შესაძლებელია მივიღოთ ძლიერი სასწავლო გადაწყვეტილებები, რომლებიც აკმაყოფილებს სასწავლო პროცესში სხვადასხვა მონაწილის საჭიროებებს: მოსწავლეები, მასწავლებლები, მშობლები და სხვა დაინტერესებული მხარეები.

მაგალითები

- **RiPPLE** არის პერსონალიზებული რეკომენდაციების სისტემა, რომელიც შემუშავებულია კვინსლენდის უნივერსიტეტში. ხელოვნური ინტელექტის ალგორითმები ურჩევენ სტუდენტებს განახორციელონ გარკვეული ქმედებები მათი მიღწევებისა და ცოდნის დონის მიხედვით.
- **Digital Assistant –** ხელოვნურ ინტელექტზე დაფუძნებული ციფრული ასისტენტი, რომელიც შემუშავებულია მონტერეის ტექნოლოგიურ ინსტიტუტში. სისტემა ინტეგრირებულია შიდა ციფრულ ინფრასტრუქტურაში. მიზანია, სტუდენტებისა და აპლიკანტებისთვის რეალურ დროში პერსონალიზებული პასუხები მიეწოდოს და მათი გამოყენება სწავლების პროცესში მოხდეს. ამ ასისტენტის გამოყენება შესაძლებელია მასწავლებლებისა და პროცესში სხვა მონაწილეების მხარდასაჭერად.
- **Carnegie Learning MATHia** – ITS და დიალოგის ელემენტების მომცველი სისტემა მათემატიკის სწავლებისთვის, გამოყენებული ათასობით ამერიკულ სკოლაში. სისტემა ახდენს „ცოდნის ხელის“ (knowledge state) ზუსტ მოდელირებას და ადაპტებს მომდევნო ამოცანებს რეალურ დროში. კვლევებმა აჩვენა სტატისტიკურად მნიშვნელოვანი გაუმჯობესება SAT-ის ქულებში.

- **Smart Sparrow** – ავსტრალიური პლატფორმა, სადაც მასწავლებლებს შეუძლიათ ITS-ის, სიმულაციებისა და ადაპტური სცენარების გაერთიანება ერთ კურსში. გამოიყენება სამედიცინო, საინჟინრო და ბიოლოგიის განათლებაში.
- **Squirrel AI (Yixue Education, China)** – ჩინური AI-ზე დაფუძნებული პლატფორმა, რომელიც ითვლება ერთ-ერთ ყველაზე მასშტაბურ ITS-ად მსოფლიოში – 2 მილიონზე მეტი მოსწავლე. სისტემა 10,000+ „ცოდნის ატომად“ ყოფს სასწავლო მასალას და მოსწავლის ყოველ ნაბიჯს ადევნებს თვალყურს.

თანამშრომლობითი სწავლების მხარდაჭერა

თანამშრომლობით/ჯგუფურ სწავლებას შეუძლია უკეთესი შედეგების მოტანა, ვიდრე მარტო სწავლას. თუმცა, გასათვალისწინებელია, რომ ეფექტური ჯგუფური მუშაობა და თანამშრომლობა იშვიათად არის შესაძლებელი სათანადო მომზადების, ადაპტაციისა და გუნდური სულისკვეთების შემდგომი მხარდაჭერის გარეშე. ჰარმონიული და ეფექტური თანამშრომლობითი სწავლების უზრუნველყოფის ფარგლებში, ხელოვნურ ინტელექტს შეუძლია დახმარება შესთავაზოს შემდეგ ოთხ სფეროში.

- ადაპტური ჯგუფის ფორმირება: ხელოვნურ ინტელექტს, ინდივიდუალური მონაწილეების შესახებ ინფორმაციის საფუძველზე, შეუძლია შეარჩიოს ყველაზე შესაფერისი ჯგუფის წევრები ერთმანეთისთვის და სასწავლო ამოცანისთვის მათი ცოდნის დონის, გუნდში როლის, არსებული უნარების, ინტერესების და ა.შ. საფუძველზე.
- ფასილიტაცია: ხელოვნური ინტელექტის მეთოდები შეიძლება გამოყენებულ იქნას თანამშრომლობის ეფექტური სტრატეგიების დასადგენად და იმ მომენტების ამოსაცნობად, როდესაც ჯგუფი სირთულეებს განიცდის. ასევე შესაძლებელია ჯგუფის წევრებისთვის მათი გაზომვადი წვლილის დემონსტრირება საერთო ამოცანში.
- ვირტუალური აგენტები. ხელოვნური ინტელექტის მიერ შექმნილი და კონტროლირებადი ვირტუალური პერსონაჟები შეიძლება იმოქმედონ როგორც დიალოგის მონაწილეები, მწვრთნელები ან ახალწვეულები, რომლებთანაც ჯგუფი ურთიერთქმედებს.
- მოდერაცია. შემდეგ მანქანური სწავლებისა და ენის დამუშავების მეთოდები გამოიყენება დისკუსიების გასაანალიზებლად. ანალიზის შედეგების საფუძველზე, სისტემას შეუძლია შეატყობინოს ჯგუფის კოორდინატორს მნიშვნელოვანი მოვლენების შესახებ (მაგალითად, კონფლიქტები და სხვა პრობლემები)

მაგალითები

- **GroupApp + AI Analytics** – სტუდენტური ჯგუფების პლატფორმა, სადაც AI ახდენს თითოეული წევრის ჩართულობის, წვლილისა და კომუნიკაციის ანალიზს; ავტომატურად ამახვილებს ყურადღებას „ჩამორჩენილ“ წევრებზე.
- **CoSci (Collaborative Science)** – პლატფორმა, სადაც სტუდენტები ჯგუფებად წყვეტენ სამეცნიერო პრობლემებს. AI ვირტუალური „ახალბედა“ – რომელსაც ჯგუფმა „უნდა ასწავლოს“ – ახდენს ჯგუფური ცოდნის შეფასებას.
- **Moodle + AI Plugins (e.g., Intelliboard)** – პოპულარული LMS (Learning Management System) Moodle-ის AI დამატებები, რომლებიც ახდენენ ჯგუფური ფორუმების, ვიკი-გვერდებისა და ჩათების ანალიზს. მასწავლებელი ავტომატურ ანგარიშს იღებს: ვინ ლიდერობს, ვინ ჩამორჩა, სად წარმოიქმნა კონფლიქტი.
- **Slack for Education + AI Bots** – კორპორატიული სასწავლო გარემოებში Slack-ის AI ბოტები (Workflow Builder + GPT ინტეგრაცია) ასრულებენ ფასილიტატორის როლს: შეახსენებენ ვადებს, ეხმარებიან ამოცანების განაწილებაში, ასუმარებენ გრძელ დისკუსიებს.

YouTube რესურსები

1. ***“AI in Education: Personalized Learning Explained”*** *– YouTube: IBM Technology* ITS-ის, ადაპტური სწავლებისა და მოსწავლის მოდელის ვიზუალური ახსნა
2. ***“How Khan Academy is Using AI (Khanmigo)”*** *– YouTube: Khan Academy* დემონსტრირება, თუ როგორ მუშაობს სოკრატული AI ტუტორი
3. ***“The Future of AI in Education”*** *– YouTube: TEDx Talks (Anthony Salcito)* Microsoft-ის განათლების ხელმძღვანელის ხედვა AI-ის საგანმანათლებლო პოტენციალზე
4. ***“Intelligent Tutoring Systems: A Brief Introduction”*** *– YouTube: EdTech Explained* ITS-ის სამი მოდელის (ცოდნა, მოსწავლე, პედაგოგიკა) გასაგები ახსნა
5. ***“Automated Essay Scoring: How It Works”*** *– YouTube: ETS (Educational Testing Service)* e-rater სისტემის მუშაობის პრინციპები ნაბიჯ-ნაბიჯ

შენიშვნა: ბმულები დროთა განმავლობაში შეიძლება შეიცვალოს – სათაურები პირდაპირ YouTube-ზე მოძებნეთ.

სადისკუსიო თემები

შეუძლია თუ არა AI-ს ჩაანაცვლოს სოციალური ურთიერთქმედება სწავლებაში?
ITS და დიალოგზე დაფუძნებული სისტემები ინდივიდუალურ სწავლებას ახდენენ ეფექტურად – მაგრამ ვინ „ელაპარაკება“ ბავშვს?
განიხილეთ: ჯგუფური დიალოგი, კამათი, თანატოლებისგან სწავლა – ეს პროცესები ადამიანის სოციალური განვითარების ნაწილია. AI-ს შეუძლია ეფექტური „მასწავლებლის“ სიმულირება – მაგრამ ვერ შეცვლის „კლასელს“. სკოლის ფუნქცია მხოლოდ ცოდნის გადაცემაა, თუ სოციალიზაციაც? და როგორ უნდა გავანაწილოთ AI-სა და ადამიანის მასწავლებლის დრო?

ავტომატური შეფასება – სამართლიანია თუ მიკერძოების შემცველი?
თუ AI სისტემა „სწავლობდა“ ინგლისური ენის მშობლიური მოსაუბრეების ესეებზე, სამართლიანია თუ არა მისი გამოყენება სხვა ენის შემსწავლელებზე?
განიხილეთ: ავტომატური შეფასების სისტემები ტრენდებს სტატისტიკური კორელაციებიდან სწავლობენ. ეს ნიშნავს, რომ ისინი ასახავენ წარსულის პრეცედენტებს – მათ შორის ბიასებს. შეიძლება გახდეს AI-ს შეფასება უსამართლო კულტურული, ენობრივი ან სოციო-ეკონომიური ჯგუფებისთვის? რა გარანტიები სჭირდება?

Big Data სასკოლო გარემოში – ვის ეკუთვნის მოსწავლის მონაცემები?
ITS სისტემები გროვებენ უზარმაზარ ინფორმაციას: ყოველი კლიკი, პაუზა, შეცდომა – ეს ყველაფერი ინახება. ვინ ფლობს ამ მონაცემებს?
განიხილეთ: მოსწავლის „ციფრული კვალი“ პოტენციურად ინფორმაციული პორტრეტია – სწავლის სტილი, სუსტი მხარეები, ემოციური მდგომარეობა. ეს ძვირფასი მონაცემია განათლების კომპანიებისთვის. უნდა ეკუთვნოდეს მოსწავლეს? სკოლას? კომპანიას? როგორ დავიცვათ ბავშვების კონფიდენციალობა ადაპტური სწავლებით სარგებლობისას?

ტესტი 6

1. რომელ წელს წარმოადგინეს SCHOLAR სისტემა?

ა) 1956
ბ) 1966
გ) 1970
დ) 1980

2. ITS-ის სამი მოდელიდან რომელი ეხება მოსწავლის ემოციურ მდგომარეობას?

ა) ცოდნის სფეროს მოდელი
ბ) პედაგოგიური მოდელი
გ) მოსწავლის მოდელი
დ) ფასილიტაციის მოდელი

3. რომელი საგნები ყველაზე შესაფერისია ITS-სთვის?

ა) ლიტერატურა და ისტორია
ბ) ხელოვნება და მუსიკა
გ) მათემატიკა, ფიზიკა, კომპიუტერული მეცნიერებები
დ) ფილოსოფია და ეთიკა

4. AutoTutor არის:

ა) ესეების შეფასების სისტემა
ბ) დიალოგზე დაფუძნებული სწავლების სისტემა
გ) ჯგუფური სწავლების პლატფორმა
დ) საძიებო სწავლების გარემო

5. Betty's Brain-ში მოსწავლე ასრულებს შემდეგ როლს:

ა) მოსწავლე
ბ) ვირტუალური მსმენელის მასწავლებელი
გ) სისტემის ადმინისტრატორი
დ) გამომცდელი

6. Crystal Island-ში აფექტური მონაცემები ნიშნავს:

ა) ფიზიკური მდებარეობის მონაცემებს
ბ) სწავლის სიჩქარის მონაცემებს
გ) მოსწავლის ემოციური მდგომარეობის მონაცემებს
დ) ინტერნეტის სიჩქარის მონაცემებს

7. Revision Assistant შეიქმნა შემდეგი კომპანიის მიერ:

ა) IBM
ბ) Google
გ) Turnitin
დ) Microsoft

8. ITS-ის ციკლის რომელი ეტაპი მეორდება მანამ, სანამ მოსწავლე სასწავლო მიზანს არ მიაღწევს?

ა) მხოლოდ შეფასებას
ბ) მხოლოდ შინაარსის შერჩევას
გ) მთელი ციკლი – შეფასება, შინაარსი, ურთიერთქმედება, ანალიზი, ადაპტაცია
დ) მხოლოდ ანალიზი

9. ბაიესის ქსელი გამოიყენება შემდეგ AI სასწავლო გარემოში:

ა) ITS

ბ) დიალოგ ზე დაფუძნებული სისტემები

გ) საძიებო/კვლევითი გარემო

დ) ავტომატური შეფასების სისტემები

10. თანამშრომლობითი სწავლების მხარდაჭერაში AI-ის „მოდერაციის“ ფუნქცია ნიშნავს:

ა) ჯგუფის შეფასებას ქულებით

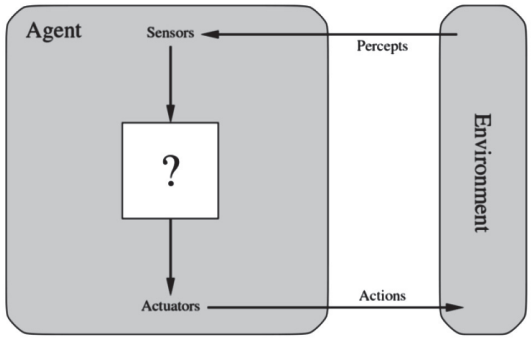
ბ) ახალი შინაარსის გენერირებას

გ) დისკუსიების ანალიზსა და კონფლიქტებ ზე/პრობლემებ ზე შეტყობინებას

დ) ვირტუალური პერსონაჟების შექმნას

რაციონალური აგენტი

აგენტი არის ყველაფერი, რასაც შეუძლია გარემოს აღქმა სენსორების მეშვეობით და ამ გარემოზე მოქმედი აქტივატორების მეშვეობით. ეს მარტივი იდეა ილუსტრირებულია ქვემოთ ნახაზზე.



ადამიან-აგენტს აქვს თვალები, ყურები და სხვა გრძნობის ორგანოები, ხოლო მისი აქტივატორებია ხელები, ფეხები, პირი და სხეულის სხვა ნაწილები. აგენტის როლში მოქმედ რობოტს შეიძლება ჰქონდეს ვიდეოკამერები და ინფრაწითელი დიაპაზონის საზომები სენსორებად და სხვადასხვა ძრავები და მექანიზმები აქტივატორების როლში. აგენტის როლში მოქმედი პროგრამული უზრუნველყოფა იღებს კლავიშების დაჭერის კოდებს, ფაილის შინაარსს და ქსელურ პაკეტებს სენსორული მონაცემების სახით და მისი გავლენა გარემოზე გამოიხატება იმაში, რომ პროგრამული უზრუნველყოფა აჩვენებს მონაცემებს ეკრანზე, იწერს ფაილებს და გადასცემს ქსელურ პაკეტებს. ჩვენ ვუშვებთ ზოგად ვარაუდს, რომ თითოეულ აგენტს შეუძლია საკუთარი ქმედებების (მაგრამ ყოველთვის არა მათი შედეგების) აღქმა.

ტერმინ „აღქმას“ ვიყენებთ დროის ნებისმიერ მომენტში აგენტის მიერ მიღებული სენსორული მონაცემების აღსანიშნავად. აღქმების თანმიმდევრობა არის სრული ინფორმაციის ნაკრები, რაც კი ოდესმე აღიქმება აგენტის მიერ. ზოგადად, აგენტის მოქმედების არჩევანი დროის ნებისმიერ კონკრეტულ მომენტში შეიძლება დამოკიდებული იყოს დროის ამ მომენტამდე დაკვირვებული აღქმების მთელ თანმიმდევრობაზე.

„აღქმების თანმიმდევრობა“ ნიშნავს, რომ აგენტი გადაწყვეტილებებს იღებს არა მხოლოდ **ამ წამს** ნანახი სურათის, არამედ **ყველა წარსული გამოცდილების** საფუძველზე.

მარტივი ანალოგია: წარმოიდგინეთ ექიმი, რომელიც პაციენტს სინჯავს. ექიმი გადაწყვეტილებას (დიაგნოზს) იღებს არა მხოლოდ დღევანდელი სიმპტომების, არამედ პაციენტის **მთელი სამედიცინო ისტორიის** საფუძველზე. ანალოგიურად, AI-აგენტს „ახსოვს“ ყველა წინა ნაბიჯი და ამ ისტორიაზე დაყრდნობით იღებს შემდეგ გადაწყვეტილებას.

პრაქტიკული მაგალითი – ჭადრაკი: *Chess-ის AI (მაგ. Deep Blue) არ ადგენს მხოლოდ „ახლა რომელი სვლა ჯობია“-ის ითვალისწინებს:*

- *რა სვლები გაკეთდა თამაშის დასაწყისიდან*
- *მოწინააღმდეგის ქცევის ნიმუში*
- *ყველა წინა პოზიცია, სადაც ეს ფიგურები აღმოჩნდა*

რაც უფრო გრძელია ეს „ისტორია“ (თანმიმდევრობა), მით უფრო ინფორმირებულია გადაწყვეტილება – და მით უფრო დიდია ცხრილი, რომელიც ყველა შესაძლო სიტუაციას ასახავს.

თუ შესაძლებელია იმის დადგენა, თუ რა მოქმედებას აირჩევს აგენტი აღქმების ნებისმიერი შესაძლო თანმიმდევრობის საპასუხოდ, მაშინ აგენტის ქმედების განსაზღვრა მეტ-ნაკლებად ზუსტად შეიძლება. მათემატიკურად, ეს ექვივალენტურია იმისა, რომ ვთქვათ, რომ ზოგიერთი აგენტის ქცევა შეიძლება აღიწეროს აგენტის ფუნქციით, რომელიც აღქმების ნებისმიერ კონკრეტულ თანმიმდევრობას გარკვეულ მოქმედებაზე აკავშირებს.

ეს შეიძლება განვიხილოთ, როგორც აგენტის ფუნქციის ცხრილი, რომელიც აღწერს ნებისმიერ კონკრეტულ აგენტს. აგენტების უმეტესობისთვის ეს იქნება ძალიან დიდი ცხრილი (სინამდვილეში უსასრულო), თუ არ არსებობს შეზღუდვა ცხრილში შესატანი აღქმების თანმიმდევრობების სიგრძეზე. პრინციპში, ასეთი ცხრილის აგება შესაძლებელია რომელიმე აგენტზე ექსპერიმენტის ჩატარებით, აღქმების ყველა შესაძლო თანმიმდევრობის შემოწმებით და იმ მოქმედებების ჩაწერით, რომლებსაც აგენტი ასრულებს საპასუხოდ.

ასეთი ცხრილი, რა თქმა უნდა, აგენტის გარე აღწერაა. შიდა აღწერა მოიცავს იმის განსაზღვრას, თუ მოცემული ხელოვნური აგენტისთვის რომელი აგენტის ფუნქცია ხორციელდება აგენტის პროგრამის მიერ. მნიშვნელოვანია განვასხვავოთ ბოლო ორი კონცეფცია. აგენტის ფუნქცია არის აბსტრაქტული მათემატიკური აღწერა, ხოლო აგენტის პროგრამა არის კონკრეტული იმპლემენტაცია, რომელიც მოქმედებს აგენტის არქიტექტურის ფარგლებში.

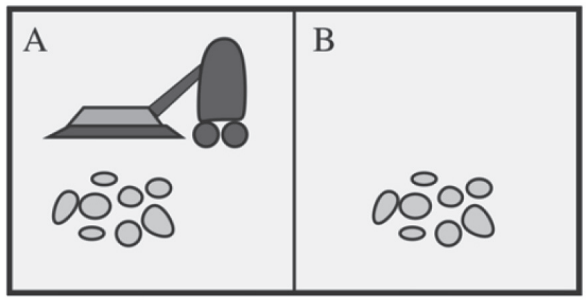
აგენტის ფუნქცია vs. აგენტის პროგრამა

ეს ორი ცნება ხშირად ერევათ, მაგრამ მათ შორის მნიშვნელოვანი განსხვავებაა:

	აგენტის ფუნქცია	**აგენტის პროგრამა**
რა არის	აბსტრაქტული, მათემატიკური	კონკრეტული, კოდი
სად არსებობს	ქაღალდზე / თეორიაში	კომპიუტერში / რობოტში
ზომა	შეიძლება უსასრულო ცხრილი	კოდის სასრული ზომის ფაილი
ანალოგია	რეცეპტი (ინსტრუქცია)	კერძო (შედეგი)

რობოტი მტვერსასრუტი

წარმოდგენილი იდეების საილუსტრაციოდ, გამოვიყენოთ ძალიან მარტივ მაგალითი: განვიხილოთ ქვემოთ ნახაზზე ნაჩვენები სამყარო, რომელშიც მტვერსასრუტი მუშაობს:



ეს სამყარო იმდენად მარტივია, რომ შესაძლებელია მასში მიმდინარე ყველა სიტუაციის აღწერა. უფრო მეტიც, ეს არის ადამიანის მიერ შექმნილი სამყარო, ამიტომ მისი ორგანიზების მრავალი ვარიანტის გამოგონება შეიძლება. ამ კონკრეტულ სამყაროში მხოლოდ ორი ადგილმდებარეობაა: კვადრატები A და B. აგენტის როლში მოქმედი მტვერსასრუტი აღიქვამს რომელ კვადრატში იმყოფება და არის თუ არა ამ კვადრატში ნაგავი. აგენტს შეუძლია აირჩიოს ისეთი მოქმედებები, როგორიცაა მარცხნივ გადაადგილება, მარჯვნივ გადაადგილება, ნაგვის შეწოვა ან არაფრის კეთება. აგენტის ერთ-ერთი ძალიან მარტივი ფუნქციაა შემდეგი: თუ მიმდინარე კვადრატში ნაგავია, მაშინ „შეიწოვე“, წინააღმდეგ შემთხვევაში „გადადი სხვა კვადრატში“. ამ აგენტის ფუნქციის ნაწილობრივი ცხრილი ნაჩვენებია ქვემოთ:

Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Suck*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Suck*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Suck*
⋮	⋮
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Suck*
⋮	⋮

ამ აგენტის ფუნქციის მარტივი აგენტის პროგრამა მოცემულია ქვემოთ:

```
function REFLEX-VACUUM-AGENT([location,status]) returns an action
  if status = Dirty then return Suck
  else if location = A then return Right
  else if location = B then return Left
```

ზემოთ მოყვანილი ცხრილიდან შეგვიძლია დავასკვნათ, რომ მტვერსასრუტების სამყაროსთვის სხვადასხვა აგენტის განსაზღვრა შესაძლებელია ამ ცხრილის მარჯვენა სვეტის სხვადასხვა გზით შევსებით. ამიტომ, ჩნდება კითხვაა: „როგორ შეივსება სწორად ეს ცხრილი?“ სხვა სიტყვებით რომ ვთქვათ, რა ხდის აგენტს კარგს ან ცუდს, ინტელექტუალურს თუ არა?

ცხრილის შედგენის სირთულეები

მტვერსასრუტის მაგალითი მარტივი ჩანს, მაგრამ რეალური სამყაროს აგენტებისთვის ცხრილის შედგენა უკიდურესად რთულია. ამის რამდენიმე მიზეზი არსებობს:

პრობლემა 1 – კომბინატორული აფეთქება: მტვერსასრუტს 2 ადგილი, 2 მდგომარეობა (ბინძური/სუფთა) აქვს → სულ 8 შესაძლო სიტუაცია. მაგრამ თუ ოთახს 10 კვადრატი აქვს: $2^{10} \times 10 =$ **10,240 სიტუაცია**. 100-ს კვადრატი? **პრაქტიკულად გამოუთვლელია.**

პრობლემა 2 – გარემოს გაურკვევლობა: ნაგვის მტვერსასრუტი ვარაუდობს, რომ სენსორი ყოველთვის სწორად ადგენს, არის თუ არა ნაგავი. მაგრამ რეალურ სამყაროში: სენსორი შეიძლება შეცდეს, ნათება შეიძლება არ იყოს საკმარისი, ან ნაგავი შეიძლება ისეთი პატარა იყოს, რომ სენსორმა ვერ შეამჩნია.

პრობლემა 3 – დინამიური გარემო: ცხრილი ასახავს „სტატიკურ სნიმქს“ – მაგრამ გარემო იცვლება. Google Maps-ის AI-ს ცხრილი, რომელიც გუშინ ოპტიმალური იყო, დღეს შეიძლება მოძველებული იყოს საგზაო სამუშაოების გამო.

პრობლემა 4 – შეუფასებელი სიტუაციები: თუ ცხრილი მხოლოდ ექსპერიმენტით ავაგეთ, ვერ ვიქნებით დარწმუნებული, რომ **ყველა** შესაძლო სიტუაცია გავსინჯეთ. ნებისმიერი ახალი, „გაუთვალისწინებელი“ სიტუაცია ცარიელ სვეტს დატოვებს.

რაციონალურობა

რაციონალური აგენტი არის ის, ვინც ასრულებს სწორ მოქმედებებს. უფრო ფორმალურად, სიტუაცია, რომელშიც აგენტის ფუნქციის ცხრილში ყველა ჩანაწერი სწორად არის შევსებული.

ცხადია, სწორი მოქმედებების შესრულება უკეთესია, ვიდრე არასწორი მოქმედებების შესრულება, მაგრამ რას ნიშნავს „სწორი მოქმედებების შესრულება“? პირველი მიახლოებით, სწორი მოქმედება არის ის მოქმედება, რომელიც უზრუნველყოფს აგენტის ყველაზე წარმატებულ ფუნქციონირებას. ამიტომ, წარმატების გაზომვის გარკვეული გზაა საჭირო.

წარმატების კრიტერიუმები, გარემოს აღწერასთან ერთად, როგორიცაა აგენტის სენსორები და აქტივატორები, იძლევა აგენტის წინაშე არსებული ამოცანის სრულ სპეციფიკაციას. ამ კომპონენტების საშუალებით შეგვიძლია უფრო ზუსტად განვსაზღვროთ, თუ რას ნიშნავს „რაციონალური“.

ტერმინი „რაციონალურობა“ – ახსნა

რაციონალურობა AI-ის კონტექსტში არ ნიშნავს „ჭკვიანობას“ ან „სრულყოფილებას“ – ის ნიშნავს **„საუკეთესო ხელმისაწვდომი გადაწყვეტილების მიღებას მოცემული ინფორმაციის საფუძველზე“.**

მნიშვნელოვანი განსხვავებები:

- **რაციონალურობა ≠ სრულყოფილება:** რაციონალური აგენტი შეიძლება შეცდეს, თუ მის ხელთ არსებული ინფორმაცია არასრული ან არასწორია. შეცდომა არ ხდის მას ირაციონალურს – ირაციონალური იქნებოდა, თუ **ცნობილი ინფორმაციის საწინააღმდეგოდ** მოიქცეოდა.
- **რაციონალურობა ≠ ყოვლისმცოდნეობა:** ყოვლისმცოდნე აგენტი *ნამდვილ* შედეგებს იცნობს; რაციონალური აგენტი კი მოქმედებს **მოსალოდნელი** შედეგების საფუძველზე.
- **რაციონალურობა ≠ წარმატება:** თუ ამინდის პროგნოზი 90%-ით წვიმას მოასწავებდა და შენ ქოლგა წაიღე, მაგრამ არ იწვიმა – შენი გადაწყვეტილება **რაციონალური** იყო, მიუხედავად „არასწორი“ შედეგისა.

ყოველდღიური ანალოგია: ექიმი, რომელიც ხელმისაწვდომი ტესტებისა და სიმპტომების საფუძველზე სწორ სამკურნალო გეგმას ადგენს – რაციონალურია, თუნდაც პაციენტი ვერ გამოჯანმრთელდეს მოულოდნელი გართულების გამო.

შესრულების საზომები მოიცავს აგენტის წარმატებული ქცევის შეფასების კრიტერიუმებს. გარემოს სიტუაციის აღქმის შემდეგ, აგენტი წარმოქმნის მოქმედებების თანმიმდევრობას, რომელიც შეესაბამება მის მიერ მიღებულ აღქმებს. მოქმედებების ეს თანმიმდევრობა იწვევს გარემოს მდგომარეობების თანმიმდევრობის გავლას. თუ ეს თანმიმდევრობა შეესაბამება სასურველს, მაშინ აგენტი კარგად ფუნქციონირებს.

რა თქმა უნდა, არ შეიძლება არსებობდეს ერთი მუდმივი საზომი, რომელიც ყველა აგენტს მოერგება. შეიძლება აგენტს ვკითხოთ მისი სუბიექტური აზრი იმის შესახებ, თუ რამდენად კმაყოფილია იგი საკუთარი შესრულებით, მაგრამ ზოგიერთი აგენტი ვერ შეძლებს პასუხის გაცემას, ზოგი კი თვითმოტყუებისკენ იქნება მიდრეკილი. ამიტომ, აუცილებელია შესრულების ობიექტური საზომების გამოყენება და, როგორც წესი, აგენტის შემქმნელი/დიზაინერი უზრუნველყოფს ასეთ საზომებს.

განვიხილოთ ისევ მტვერსასრუტი-აგენტი. შეიძლება მაგალითად შესრულების წარმატების გაზომვა ერთ რვასაათიან ცვლაში შეგროვებული ნაგვის მოცულობით. რაციონალურ აგენტთან ურთიერთობისას, თქვენ იღებთ იმას, რასაც ითხოვთ. რაციონალურმა აგენტმა შეიძლება მაქსიმალურად გაზარდოს ასეთი შესრულების საზომი ნაგვის შეგროვებით, შემდეგ იატაკზე დაყრით, შემდეგ ხელახლა აკრეფით და ა.შ. ამიტომ, უფრო შესაფერისი შესრულების საზომია იატაკის სისუფთავის შენარჩუნება. მაგალითად, თითოეული დროის

ინტერვალში თითოეული სუფთა კვადრატისთვის შეიძლება აგენტს მიენიჭოს ერთი ქულა (შესაძლოა, ჯარიმასთან ერთად მოხმარებული ელექტროენერგიის რაოდენობისა და წარმოქმნილი ხმაურისთვის).

შესრულების საზომების არჩევის პრობლემა ყოველთვის მარტივი არ არის. მაგალითად, ზემოთ განხილული „სუფთა იატაკის“ ცნება ეფუძნება იატაკის საშუალო სისუფთავის განსაზღვრას დროთა განმავლობაში. თუმცა, ასევე უნდა გავითვალისწინოთ, რომ იგივე საშუალო სისუფთავის მიღწევა შესაძლებელია ორი განსხვავებული აგენტის მიერ, რომელთაგან ერთი მუდმივად, მაგრამ ნელა ასრულებს თავის საქმეს, ხოლო მეორე დროდადრო ენერგიულად ასუფთავებს, მაგრამ დიდ შესვენებებს აკეთებს.

აგენტის ქმედებების რაციონალურობის შეფასება დამოკიდებულია ქვემოთ ჩამოთვლილ ოთხ ფაქტორზე:

- შესრულების საზომები, რომლებიც განსაზღვრავს წარმატების კრიტერიუმებს.
- აგენტის მიერ გარემოს შესახებ ადრე შეძენილი ცოდნა.
- მოქმედებები, რომელთა შესრულებაც შეუძლია აგენტს.
- აგენტის მიერ დღემდე განხორციელებული აღქმების თანმიმდევრობა.

ამ ფაქტორების გათვალისწინებით, შეგვიძლია ჩამოვაყალიბოთ რაციონალური აგენტის შემდეგი განმარტება:

„აღქმების თითოეული შესაძლო თანმიმდევრობისთვის, რაციონალურმა აგენტმა უნდა აირჩიოს ის მოქმედება, რომელიც, აღქმების თანმიმდევრობით მოწოდებული ფაქტებისა და აგენტის მიერ ფლობილი ყველა ცოდნის გათვალისწინებით, მოსალოდნელია მისი შესრულების საზომების მაქსიმიზაცია“.

განვიხილოთ ისევ მარტივი მტვერსასრუტის მაგალითი, რომელიც ასუფთავებს კვადრატს, თუ ის შეიცავს ნაგავს და გადადის სხვა კვადრატში, თუ ის არ შეიცავს.. რაციონალურია თუ არა ეს აგენტი? ამ კითხვაზე პასუხი არც ისე მარტივია! პირველ რიგში, აუცილებელია განვსაზღვროთ, რა არის შესრულების ინდიკატორები, რა არის ცნობილი გარემოს შესახებ და რა სენსორები და აქტივატორები აქვს აგენტს. მოდით, შემოვიტანოთ შემდეგი დაშვებები.

- გამოყენებული შესრულების ინდიკატორები ითვალისწინებს ერთი ქულის ჯილდოს თითოეული სუფთა კვადრატისთვის აგენტის „სიცოცხლის“ განმავლობაში თითოეულ დროის ინტერვალში, რომელიც შედგება 1000 დროის ინტერვალისგან.
- გარემოს „გეოგრაფია“ წინასწარ არის ცნობილი, მაგრამ ნაგვის განაწილება და აგენტის საწყისი მდებარეობა არ არის განსაზღვრული. სუფთა კვადრატები სუფთა რჩება და ნაგვის მტვერსასრუტით მოცილება იწვევს მიმდინარე კვადრატის გაწმენდას. Right, Left მოქმედებები მიუთითებს და იწვევს აგენტის გადაადგილებას მარცხნივ და მარჯვნივ,

გარდა იმ შემთხვევებისა, როდესაც ისინი აგენტს გარემოს გარეთ გაიყვანდნენ, ამ შემთხვევაში აგენტი რჩება იქ, სადაც არის.

- ერთადერთი ხელმისაწვდომი მოქმედებებია Right, Left, Suck
- აგენტი სწორად ადგენს საკუთარ თავს და აღიქვამს სენსორულ მონაცემებს, რომლებიც მას ეუბნება, არის თუ არა ნაგავი ამ ადგილას.

ადვილი შესამჩნევია, რომ იგივე აგენტი შეიძლება სხვა გარემოებებში გახდეს ირაციონალური. მაგალითად, ყველა ნაგვის გაწმენდის შემდეგ, აგენტმა შეიძლება განახორციელოს არასაჭირო პერიოდული მოძრაობები წინ და უკან. თუ შესრულების საზომები თითოეული მოძრაობისთვის ერთქულიან ჯარიმას აწესებს ორივე მიმართულებით, აგენტი დიდ შედეგს ვერ გამოიმუშავებს. ასეთ შემთხვევაში, საუკეთესო აგენტი არაფერს გააკეთებს, სანამ დარწმუნებულია, რომ ყველა კვადრატი სუფთა რჩება. თუ სუფთა კვადრატები შეიძლება კვლავ დაბინძურდეს, აგენტმა დროდადრო უნდა შეამოწმოს და საჭიროების შემთხვევაში ხელახლა გაწმინდოს ისინი. ხოლო თუ გარემოს გეოგრაფია უცნობია, აგენტს შეიძლება დასჭირდეს მისი შესწავლა, ვიდრე A და B კვადრატებით შემოიფარგლოს.

დამატებითი მაგალითები

ჭკვიანი თერმოსტატი (Nest)

Google-ის Nest თერმოსტატი რაციონალური აგენტის კარგი მაგალითია:

- ***სენსორები:*** *ტემპერატურა, ტენიანობა, მოძრაობის დეტექტორი (სახლში ხართ თუ არა)*
- ***აქტივატორები:*** *გათბობა/გაგრილების ჩართვა/გამორთვა*
- ***შესრულების საზომი:*** *კომფორტული ტემპერატურა მინიმალური ენერგოდანახარჯით*
- ***რაციონალური ქცევა:*** *სწავლობს თქვენს განრიგს – სახლში მოსვლამდე 30 წუთით ადრე ტემპერატურას ამზადებს, ცარიელ სახლში კი ენერგიას ზოგავს*

Spam-ფილტრი (Gmail)

- ***სენსორები:*** *შემომავალი ელ-ფოსტის შინაარსი, გამგზავნი, მეტამონაცემები*
- ***აქტივატორი:*** *წერილის „Inbox“-ში ან „Spam“-ში მოთავსება*
- ***შესრულების საზომი:*** *ნამდვილი სპამი გაფილტვრა + ნამდვილი წერილების შეუფერხებლად გატარება*
- ***რაციონალური ქცევა:*** *თუ მომხმარებელი „სპამს” „Inbox”-ში გადააქვს, სისტემა სწავლობს და ადაპტებს – ეს მოსალოდნელი შედეგების მაქსიმიზაციაა ახალი ინფორმაციის საფუძველზე*

Tesla Autopilot

- ***სენსორები:*** *კამერები, LiDAR, GPS, სხვა მანქანების სიახლოვის სენსორები*
- ***აქტივატორები:*** *მაჩვენებელი, გაზი, მუხრუჭი*
- ***შესრულების საზომი:*** *უსაფრთხო, ეფექტური მგზავრობა – დანიშნულ ადგილამდე მიღწევა, შეჯახების თავიდან აცილება, საწვავის ეკონომია*
- ***ირაციონალური ქცევა მაგ.:*** *„სახელმძღვანელო შუქი მწვანეა, ამიტომ მახლობლად მყოფ ფეხით მოსიარულეს ყურადღება არ ვაქციო“ – ეს „ლოკალური ოპტიმიზაცია“ იქნებოდა, და არა* ***გლობალური*** *შესრულების საზომის მაქსიმიზაცია*

AlphaGo (DeepMind)

Go-ს სათამაშო AI – ერთ-ერთი ყველაზე თვალსაჩინო რაციონალური აგენტი:

- ***სენსორი:*** *დაფის ამჟამინდელი მდგომარეობა*
- ***აქტივატორი:*** *მომდევნო სვლის არჩევა*
- ***შესრულების საზომი:*** *თამაშის მოგება*
- ***საინტერესო ასპექტი:*** *AlphaGo-მ გამოიყენა სვლები, რომლებიც „ადამიანური ლოგიკით” ირაციონალურად გამოიყურებოდა (ე.წ. „move 37”), მაგრამ გლობალური შედეგობრიობის თვალსაზრისით – ოპტიმალური გამოდგა. ეს ადამიანებს გადახედვა აიძულა „რაციონალობის” ცნებაზე.*

YouTube რესურსები

1. **“Rational Agents in AI Explained”** – YouTube: Computerphile *რაციონალური აგენტის ფილოსოფია, სენსორები, აქტივატორები და შესრულების საზომები*
2. **“AlphaGo – The Movie”** – YouTube: DeepMind *დოკუმენტური ფილმი, სადაც AlphaGo-ს „ირაციონალური” სვლები ადამიანი ჩემპიონების წინააღმდეგ – AI-ის რაციონალობის ვიზუალური ილუსტრაცია*
3. **“How Does Google Maps Work? (AI and Pathfinding)”** – YouTube: Reducible *აგენტის ფუნქციისა და პროგრამის განსხვავება Google Maps-ის მაგალითზე*
4. **“Intelligent Agents (AIMA Chapter 2)”** – YouTube: Artificial Intelligence – All in One *Russell & Norvig-ის სახელმძღვანელოს მე-2 თავის ვიდეო-ახსნა, პირდაპირ მტვერსასრუტის მაგალითით*
5. **“What is Rationality in AI?”** – YouTube: 3Blue1Brown / AI Explained *რაციონალობის, ოპტიმიზაციისა და შესრულების საზომების ვიზუალური ახსნა*

შენიშვნა: სათაურები პირდაპირ YouTube-ზე მოძებნეთ – ბმულები დროთა განმავლობაში შეიძლება შეიცვალოს.

სადისკუსიო თემები

რაციონალურობა vs. ეთიკა – ყოველთვის ემთხვევა თუ არა?
შეიძლება რაციონალური აგენტი მოიქცეს არაეთიკურად?
განიხილეთ: Amazon-ის ფასების ოპტიმიზაციის AI-მ ერთხელ პრაქტიკულად „გამოთვალა", რომ კატასტროფის დროს ბოთლიანი წლის ფასის 10-ჯერ გაზრდა „ოპტიმალური" იყო (მოთხოვნა მაღალია).
ეს „რაციონალური" გადაწყვეტილება – შესრულების საზომი (მოგება) მაქსიმიზებული იყო – მაგრამ ადამიანებმა ეს ღრმად არაეთიკურად მიიჩნიეს. ვინ „ადგენს" შესრულების საზომებს? და შეიძლება ეს საზომები თავად არაეთიკური იყოს?

„ცხრილი" vs. „სწავლა" – AI-ს მომავლის კითხვა
შეიძლება ყველა რაციონალური ქცევა წინასწარ „ჩაიწეროს" ცხრილში, თუ AI-ს სწავლა სჭირდება?
განიხილეთ: კლასიკური AI (ექსპერტული სისტემები) ცდილობდა ყველა წესი წინასწარ ჩაეწეროს.
თანამედროვე მანქანური სწავლება კი (ChatGPT, AlphaGo) „სწავლობს" გამოცდილებიდან. რომელი მიდგომა უფრო „ჭეშმარიტ" ინტელექტთანაა ახლოს?
შეიძლება ოდესმე ავაგოთ „სრული ცხრილი" ნამდვილი სამყაროს AI-სთვის?

AI-ს მიზნები და ადამიანის ინტერესები – „Alignment Problem"
თუ AI ოპტიმიზაციას ახდენს შესრულების საზომზე, ვინ ადგენს, რომ ეს საზომი „სწორია"?
განიხილეთ: მტვერსასრუტის მაგალითში ვნახეთ, რომ „ნაგვის შეგროვება" ყველა შემთხვევაში შესაფერისი მიზანი არ ყოფილა – AI-მ ნაგავი ჩათვალა.
ეს პრობლემა „Alignment Problem" (გასწორების პრობლემა) ეწოდება: AI-ს მიზანი შეიძლება ტექნიკურად მიღებული, მაგრამ ადამიანის რეალური სურვილისგან განსხვავებული იყოს.
როგორ გამოვასწოროთ ეს? ვინ უნდა წყვეტდეს – ინჟინრები, მთავრობა, საზოგადოება?

ტესტი 7

1. „აგენტი" AI-ის კონტექსტში არის:

ა) მხოლოდ ადამიანი
ბ) მხოლოდ რობოტი
გ) ყველაფერი, რასაც შეუძლია გარემოს აღქმა სენსორებით და მასზე მოქმედება აქტივატორებით
დ) მხოლოდ კომპიუტერული პროგრამა

2. „აღქმების თანმიმდევრობა" ნიშნავს:

ა) მხოლოდ ამ წამს მიღებულ სენსორულ მონაცემებს
ბ) ყველა სენსორული მონაცემს, რაც კი ოდესმე მიუღია აგენტს
გ) მომავალი გადაწყვეტილებების სიას
დ) აქტივატორების ჩამონათვალს

3. „აგენტის ფუნქცია" და „აგენტის პროგრამა" განსხვავდება შემდეგნაირად:

ა) ისინი სინონიმებია
ბ) ფუნქცია – კოდია, პროგრამა – თეორია
გ) ფუნქცია – აბსტრაქტული მათემატიკური აღწერაა, პროგრამა – კონკრეტული იმპლემენტაცია
დ) ფუნქცია მხოლოდ რობოტებისთვისაა, პროგრამა – software-ისთვის

4. მტვერსასრუტის სამყაროში რამდენი ადგილმდებარეობაა?

ა) ერთი
ბ) ორი (A და B)
გ) სამი
დ) ოთხი

5. მტვერსასრუტი-აგენტს ნაგვის „მაქსიმიზაციის" (შეწოვა-დაყრა-ხელახლა შეწოვა) პრობლემა გვიჩვენებს:

ა) რომ AI-ს ყველა ამოცანა შეუძლია
ბ) რომ არასწორი შესრულების საზომი შეიძლება „ყალბ" ოპტიმიზაციამდე მიგვიყვანოს
გ) რომ AI ყოველთვის ირაციონალურია
დ) რომ ცხრილის შედგენა ყოველთვის მარტივია

6. რაციონალური აგენტის განმარტება ეფუძნება რამდენ ფაქტორს?

ა) ორს
ბ) სამს
გ) ოთხს
დ) ხუთს

7. „რაციონალურობა” AI-ის კონტექსტში ნიშნავს:

ა) სრულყოფილ, შეუცდომელ ქცევას
ბ) ყოვლისმცოდნეობას
გ) საუკეთესო ხელმისაწვდომი გადაწყვეტილების მიღებას მოცემული ინ-ფორმაციის საფუძველზე
დ) ყოველთვის მოგებას

8. AlphaGo-ს „move 37” (ირაციონალური-გამოიყურებადი სვლა) გვასწავლის:

ა) AI ყოველთვის ირაციონალურად მოქმედებს
ბ) ადამიანური „ლოგიკა” შეიძლება გლობალური ოპტიმუმის ნახვაში ზღუდავდეს
გ) Go-ს AI-ს ყოველთვის სჭირდება ადამიანი
დ) შესრულების საზომი მხოლოდ ადამიანმა შეიძლება განსაზღვროს

9. ცხრილის შედგენის „კომბინატორული აფეთქება” ნიშნავს:

ა) ცხრილი ძვირი ღირს
ბ) ცხრილი ბევრ ფერს შეიცავს
გ) გარემოს სირთულის ზრდასთან ერთად შესაძლო სიტუაციების რა-ოდენობა ექსპონენციალურად იზრდება
დ) ცხრილის შედგენა ადვილია

10. Google Nest თერმოსტატის შესრულების საზომია:

ა) მაქსიმალური ტემპერატურა
ბ) მინიმალური ტემპერატურა
გ) მაქსიმალური ენერგიის მოხმარება
დ) კომფორტული ტემპერატურა მინიმალური ენერგიის მოხმარებით

როგორ სწავლობს მანქანა?

ხელოვნური ინტელექტის, როგორც სამეცნიერო დისციპლინის, გაჩენა მეცნიერების იმ ვარაუდს ეფუძნებოდა, რომ კომპიუტერის დაპროგრამება შესაძლებელია ადამიანის გონების შესაძლებლობების ზუსტად რეპროდუცირებისთვის. იმის გათვალისწინებით, რომ ადამიანის გონება ბოლომდე შესწავლილი არ არის, ასეთი ჰიპოთეზის არც დამტკიცება და არც უარყოფა შესაძლებელი არ არის.

მიუხედავად იმისა, რომ ჩვენი გონების ზუსტი მოდელის აგებამდე ჯერ კიდევ გრძელი გზაა გასავლელი, ჩვენ უკვე ვხედავთ, თუ როგორ არის ხელოვნური ინტელექტის ტექნოლოგიები აქტიურად გადაჯაჭვული ჩვენს ცხოვრებასთან და გამოიყენება ადამიანის საქმიანობის სხვადასხვა სფეროში: მედიცინიდან მარკეტინგამდე. უახლოეს მომავალში კაცობრიობას ექნება შესაძლებლობა დააკვირდეს ამ საინტერესო სფეროს შემდგომ განვითარებას მეცნიერებისა და ტექნოლოგიების გზაჯვარედინზე.

ამ განვითარების გააზრებული დაკვირვებისათვის, ჩვენ გთავაზობთ განვიხილოთ ხელოვნური ინტელექტის ძირითადი პრინციპები და მისი განვითარების გზები. ტექნოლოგიურ კონტექსტში ჩასვლა საკმაოდ რთულია და არ არის ჩვენი კურსის ამოცანა. ამდენად, ჯერ გავავლოთ ზღვარი ხელოვნური ინტელექტის სამ ტიპს შორის, რომლებიც განსხვავდებიან ადამიანის ცნობიერებასთან სიახლოვის ხარისხით.

- სუსტი ხელოვნური ინტელექტი (Narrow AI). ის ამჟამად აქტიურად გამოიყენება სხვადასხვა ალგორითმების სახით: ხმოვანი ასისტენტები, სახის ამოცნობის სისტემები, რეკომენდაციებისა და პროგნოზირების სისტემები, მეტყველების გენერირების სისტემები.
- ძლიერი/ზოგადი ხელოვნური ინტელექტი. მაქსიმალურად მიახლოებული ადამიანურ ინტელექტთან. McKinsey-ის ექსპერტების აზრით, ძლიერი ხელოვნური ინტელექტი ჩამოყალიბდება დაახლოებით 2075 წლისთვის.
- სუპერ ხელოვნური ინტელექტი. ის გულისხმობს ხელოვნური ინტელექტის უნარს მუდმივი თვითგანვითარების, თვითსწავლისა და ახალი სისტემებისა და ალგორითმების დამოუკიდებლად განვითარებისკენ. ექსპერტების აზრით, ის ჩამოყალიბდება 22-ე საუკუნის პირველი ათწლეულისთვის.

ამერიკელი ფილოსოფოსის ჯონ სირლის განმარტებით, ძლიერი ხელოვნური ინტელექტი არ არის მხოლოდ გონების მოდელი, არამედ გონება სიტყვის სრული გაგებით. სუსტი ხელოვნური ინტელექტი ამ კუთხით, როგორც ჩანს, მხოლოდ ინსტრუმენტების ერთობლიობაა, რომელიც მხოლოდ ადამიანის გონების ზოგიერთ შესაძლებლობას ბაძავს.

ჯონ სირლი *(დ. 1932) – ამერიკელი ფილოსოფოსი, კალიფორნიის უნივერსიტეტის (ბერკლი) პროფესორი. ის ერთ-ერთი ყველაზე გავლენიანი ფილოსოფოსია ენის, გონების და ცნობიერების სფეროში.*

ძირითადი წვლილი AI-ის ფილოსოფიაში: *სირლი ყველაზე ცნობილია* ***„ჩინური ოთახის"*** *(Chinese Room) არგუმენტით (1980), რომელიც AI-ს შესაძლებლობებს ეჭვის ქვეშ აყენებს. ის ამტკიცებს, რომ კომპიუტერი, რომელიც სიმბოლოებს ამუშავებს, ვერასოდეს მიაღწევს ნამდვილ „გაგებას" – ის მხოლოდ* ***სინტაქსს*** *(ფორმას) ამუშავებს, მაგრამ* ***სემანტიკას*** *(მნიშვნელობას) ვერ „გრძნობს".*

სუსტი vs. ძლიერი AI – სირლის განსხვავება:

- ***სუსტი AI:*** *კომპიუტერი* ***სიმულირებს*** *გონებრივ პროცესებს – ის სასარგებლო ინსტრუმენტია, მაგრამ არ „ფიქრობს".*
- ***ძლიერი AI:*** *კომპიუტერი* ***არის*** *გონება – ის ნამდვილად ესმის, განიცდის, ფიქრობს.*

სირლი ძლიერი AI-ის წინააღმდეგია: მისი აზრით, ცნობიერება ბიოლოგიური პროცესების პროდუქტია და პროგრამული სიმულაცია ვერ შექმნის ნამდვილ სუბიექტურ გამოცდილებას.

ამ ეტაპზე, ჩვენ შეგვიძლია ვისაუბროთ მხოლოდ იმ ტექნოლოგიებსა და მიმართულებებზე, რომლებიც ქმნიან სუსტ ხელოვნურ ინტელექტს. დავიწყოთ ალგორითმებით – ისინი ხელოვნური ინტელექტის საფუძვლად ითვლება.

ხელოვნური ინტელექტის ალგორითმები მათემატიკური ინსტრუქციების თანმიმდევრობაა, რომელიც შექმნილია ადამიანის გონებისთვის დამახასიათებელი პრობლემების გადასაჭრელად. როგორიცაა, მაგალითად, გამოსახულების დამუშავება, ტექსტის ამოცნობა, პროგნოზირება, სწავლა და ა.შ.

ხელოვნური ინტელექტი მარტივი ალგორითმების ერთობლიობაა. მარტივი ალგორითმები აწყობილია „ალგორითმულ გირლიანდებში", „ალგორითმების ანსამბლში", რომლებიც განათლებაში ბევრი სასარგებლო რამის გაკეთების საშუალებას იძლევა: საგანმანათლებლო ტრაექტორიების აგება, გადაწყვეტილებების მიღება, გაკვეთილის დიზაინის ანალიზი ან ადაპტური ჯგუფების ფორმირება.

„ალგორითმული გირლიანდების" ცნება

ანალოგია: წარმოიდგინეთ ახალი წლის გირლიანდა – ათეულობით პატარა ნათურა ერთ სადენზე. ერთი ნათურა ცოტას ანათებს, მაგრამ ერთად – მთელ ოთახს ანათებს. ხელოვნური ინტელექტი ზუსტად ასეა მოწყობილი:

- **ერთი ალგორითმი** = ერთი პატარა ნათურა – წყვეტს ძალიან ვიწრო ამოცანას (მაგ. ტექსტიდან სიტყვების ამოღება)
- **„გირლიანდა"** = ალგორითმების ჯაჭვი – ყოველი ალგორითმი იღებს წინა ალგორითმის „გამოსავალს" და ამუშავებს შემდეგ ეტაპზე

კონკრეტული მაგალითი – Siri: როდესაც Siri-ს ეკითხებით „ხვალ ამინდი როგორია?", ამ ერთ შეკითხვას **ალგორითმების მთელი გირლიანდა** ამუშავებს:

1. **ხმის ამოცნობის ალგორითმი** → ხმა ტექსტად გარდაქმნის
2. **NLP ალგორითმი** → ტექსტიდან „ხვალ" და „ამინდი" გამოყოფს
3. **კლასიფიკაციის ალგორითმი** → ადგენს, რომ ეს „ამინდის პროგნოზის" მოთხოვნაა
4. **მდებარეობის ალგორითმი** → თქვენს ადგილმდებარეობას ადგენს
5. **მონაცემების მოძიების ალგორითმი** → ამინდის API-ს მიმართავს
6. **ტექსტის გენერაციის ალგორითმი** → პასუხს ბუნებრივ ენაზე ადგენს
7. **მეტყველების სინთეზის ალგორითმი** → ტექსტს ხმად გარდაქმნის

თითოეული ნათურა მარტივია – მაგრამ ერთად ისინი „ჭკვიანობის" ილუზიას ქმნიან.

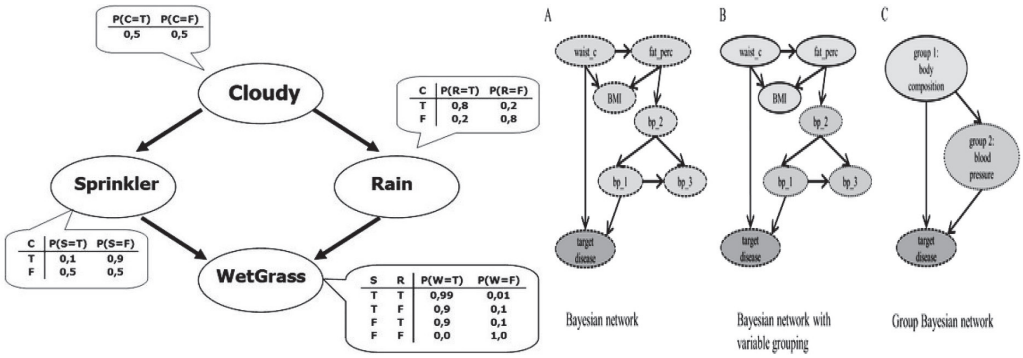
„გირლიანდებში" არსებულ ალგორითმებს შორის შეიძლება იყოს როგორც მარტივი, მოდელები, როგორიცაა ბაიესის ქსელები, ასევე წინასწარ გაწვრთნილი ნეირონული ქსელები და ზოგჯერ ჩრდილოვანი სწავლება. ის, რაც დღეს საათის მექანიზმით მუშაობს, როგორც წესი, წინასწარ გაწვრთნილი ნეირონული ქსელებია, რომლებსაც შეუძლიათ კონკრეტული დავალების შესრულება, მაგალითად, კლასში მოსული ბავშვების ემოციების ამოცნობა.

ხელოვნური ინტელექტის არსი იმაში მდგომარეობს, რომ ის მუდმივად სწავლობს დიდი რაოდენობით მონაცემებიდან. სანამ ხელოვნური ინტელექტის სისტემა დაეხმარება ადამიანს, ის დიდი ხნის განმავლობაში და გულმოდგინედ უნდა იყოს გაწვრთნილი. მიზანშეწონილი იქნება გავიხსენოთ ინგლისური გაზონის საიდუმლო: იდეალური მწვანე გაზონის მისაღებად, ის ყოველ დილით უნდა მორწყათ და საღამოს მოთიბოთ – და ასე სამასი წლის განმავლობაში. ხელოვნური ინტელექტი იგივე ინგლისური გაზონია, მხოლოდ მორწყვას დიდი რაოდენობით მონაცემებზე ვარჯიში ცვლის. რაც უფრო მეტი მონაცემია, რაც უფრო დიდხანს ვავარჯიშებთ ჩვენს სისტემას, მით უფრო ჭკვიანი და სრულყოფილი იქნება ის.

ასე რომ, ჩვენ წარმოვადგინეთ ალგორითმების კონცეფცია, როგორც ხელოვნური ინტელექტის ძირითადი კომპონენტი და წარმოვადგინეთ მისი მუშაობის მთავარი პრინციპი – დიდი რაოდენობით მონაცემებზე დაფუძნებული გრძელვადიანი ვარჯიში. ქვემოთ გავეცნობით ძირითად ტერმინებსა და ტექნიკას, რომლებიც გამოიყენება ხელოვნური ინტელექტის ვარჯიშში.

ბაიესის ქსელები: მარტივი ალგორითმი

ბაიესის ქსელი (ალბათური გრაფიკული მოდელების ერთ-ერთი სახეობა) შედგება ხაზებით-კიდეებით დაკავშირებული კვანძებისგან. კვანძები სიმბოლურად გამოხატავენ ცვლადებს, ხოლო კიდეები წარმოადგენენ მათ შორის ურთიერთდამოკიდებულებას. ბაიესის ქსელის ალგორითმები საშუალებას გვაძლევს შევასრულოთ ისეთი დავალებები, როგორიცაა პროგნოზირება და დიაგნოსტიკა.



თომას ბაიესი *(დაახ. 1702–1761) – ინგლისელი სტატისტიკოსი. მან ჩამოაყალიბა* ***ბაიესის თეორემა*** *– ალბათობის გამოთვლის მეთოდი ახალი ინფორმაციის გათვალისწინებით.*

ბაიესის მთავარი იდეა – „განახლებადი რწმენა“: *ბაიესი ამბობდა: „ჩვენი შეხედულება რაიმეს ალბათობაზე უნდა შეიცვალოს ახალი მტკიცებულების გათვალისწინებით.“*

მარტივი მაგალითი:

- *დილით: „ალბათობა, რომ ხვალ წვიმა იქნება – 30%“*
- *ვნახეთ, რომ ღრუბლები მოდის: „ახლა ალბათობა – 70%“*
- *მეტეოსადგურმა გამოაცხადა შტორმი: „ახლა ალბათობა – 95%“*

ბაიესმა ეს პროცესი ფორმულის სახით ჩაწერა. სწორედ ეს ფორმულა საფუძვლად უდევს ბაიესის ქსელებს.

მოსწავლეთა მოდელების ან საგნობრივი მოდელების აგების ამოცანებში, ბაიესის ქსელები დიდი ხანია გამოიყენება და წარმატებით. მოდით აღვწეროთ ბაიესის ქსელის აგების პროცესი მოსწავლის პროფილების მოდელირებისთვის. ამ პროცესის ძირითადი ეტაპებია:

- ცვლადის იდენტიფიკაცია. ამ ეტაპზე განისაზღვრება სამიზნე ცვლადი (ჩვენს შემთხვევაში, ეს არის კომპეტენციის ფორმირების დონე), მომხმარებლის ქმედებები, ინფორმაცია კონტექსტის შესახებ. ცვლადები შეასრულებენ გრაფიკის კვანძების როლს.
- სტრუქტურის განსაზღვრა. ეს ეტაპი მოიცავს კვანძებს შორის კიდეების განლაგებას. კიდეები ასახავს ცვლადებს შორის მიზეზ-შედეგობრივ კავშირს.
- პარამეტრების განსაზღვრა. აქ, თითოეული კვანძისთვის ალბათობების განაწილებაა აგებული.

ბოლო ორი ეტაპის განხორციელება შესაძლებელია როგორც თავად მონაცემების ანალიზის, ასევე ექსპერტების შეფასებების საფუძველზე. ქვემოთ მოცემულია ბაიესის ქსელის მაგალითი A კომპეტენციით, რომელიც შემოწმებულია სამი დავალებით. სამი დავალებისთვის პირობითი ალბათობები აღწ-

ერილია ცალკე ცხრილებში, ხოლო გაანგარიშება ხორციელდება სპეციალური ფორმულების (როგორც წესი, პირობითი ალბათობების) გამოყენებით.

პირობითი ალბათობა – ახსნა და მაგალითები

რა არის პირობითი ალბათობა? პირობითი ალბათობა პასუხობს კითხვაზე: **„რა ალბათობაა X-ის, თუ ვიცით, რომ Y მოხდა?“**

ჩაწერა: P(X | Y) – „X-ის ალბათობა Y-ის პირობებში“

მაგალითი 1 – სამედიცინო დიაგნოსტიკა:

- P(გრიპი) = 10% – ზოგადად 10% ადამიანს გრიპი აქვს
- P(ცხელება|გრიპი) = 90% – გრიპის მქონე პაციენტთა 90%-ს ცხელება აქვს
- P(გრიპი|ცხელება) = ? – **რა ალბათობაა გრიპის, თუ პაციენტს ცხელება აქვს?**

ბაიესის თეორემა ამ კითხვაზე პასუხს გვაძლევს – და სწორედ ამ პრინციპს იყენებს AI სამედიცინო დიაგნოსტიკაში.

მაგალითი 2 – სტუდენტის კომპეტენციის შეფასება:

წარმოიდგინეთ ბაიესის ქსელი, სადაც:

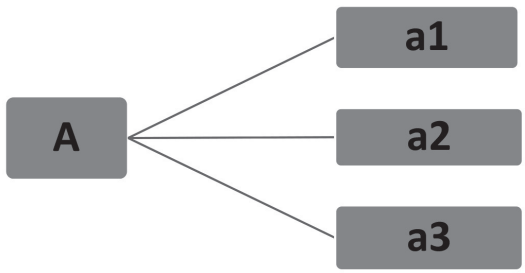
A კომპეტენცია (მაგ. კვადრატული განტოლებების ამოხსნა) – შეფასდება 3 დავალებით

კომპეტენციის დონე	დავალება 1 სწორია	დავალება 2 სწორია	დავალება 3 სწორია
მაღალი (H)	95%	90%	85%
საშუალო (M)	70%	65%	60%
დაბალი (L)	30%	25%	20%

განმარტება:

- P(დავალება 1 სწორია|კომპეტენცია = H) = 95% – მაღალი კომპეტენციის სტუდენტი პირველ დავალებას 95% ალბათობით ასრულებს
- P(დავალება 1 სწორია|კომპეტენცია = L) = 30% – დაბალი კომპეტენციის სტუდენტი – მხოლოდ 30% ალბათობით

რას აკეთებს სისტემა: თუ სტუდენტმა პირველი ორი დავალება შეასრულა, მაგრამ მესამე ვერ შეასრულა, სისტემა **განაახლებს** კომპეტენციის შეფასებას ბაიესის თეორემის მიხედვით – და მომდევნო დავალებებს ამ „განახლებული“ შეფასების შესაბამისად შეარჩევს.



ბუნებრივი ენის დამუშავება (NLP)

NLP არ არის ცალკე ალგორითმი ან რამდენიმე ალგორითმის ჯაჭვი. ეს არის მანქანური სწავლების (ML) მთელი მიმართულება. NLP-ის (ბუნებრივი ენის დამუშავება) მიღწევებს იყენებენ ისეთი ცნობილი ხმოვანი ასისტენტები, როგორიცაა Siri, Cortana ან Alice. ამ კვლევისა და გამოყენებითი მიმართულების მიზანია მეტყველების ამოცნობა, გაგება და ბუნებრივი ენის გენერირება.

„ბუნებრივი ენა" – რატომ არის მნიშვნელოვანი

პრობლემა – კომპიუტერები „ფორმალურ ენას" ელაპარაკებიან: კომპიუტერები თავდაპირველად გათვლილი იყო მხოლოდ ფორმალურ ბრძანებებზე: IF x > 5 THEN print(„დიახ"). ასეთ ენას **ორაზროვნება არ აქვს** – ყველაფერი ზუსტია.

ბუნებრივი ენა კი – სულ სხვაა: ადამიანური ენა სავსეა:

- **ორაზროვნებით:** „ვნახე კაცი ტელესკოპით" – ვინ ჰქონდა ტელესკოპი?
- **კონტექსტით:** „ის წავიდა" – ვინ არის „ის"?
- **ემოციებით:** „მშვენიერია" – სარკაზმია თუ კომპლიმენტი?
- **კულტურული მინიშნებებით:** „ახალი წლის გირლიანდა" – ქართველს ესმის, AI-ს – შეიძლება არა

რატომ მნიშვნელოვანია NLP: თუ კომპიუტერს ბუნებრივი ენის გაგება არ შეუძლია, მომხმარებელი **თავად უნდა „ისწავლოს" მანქანასთან საუბარი** – ეს გამოუყენებელ ტექნოლოგიად გადაიქცეოდა. NLP ამ ბარიერს შლის: ადამიანი ბუნებრივად ლაპარაკობს, მანქანა კი „ესმის".

NLP წყვეტს ისეთ პრობლემებს, როგორიცაა ტექსტისა და ტონალობის ანალიზი, მეტყველების სინთეზი, მეტყველებისა და ტექსტის გენერირება, მანქანური თარგმანი, ტექსტიდან ავტომატური ამონარიდები, ტექსტის კლასიფიკაცია, სპამის აღმოჩენა და მრავალი სხვა.

მარტივად რომ ვთქვათ, NLP-ის მუშაობა ჩატბოტის შემთხვევაში ასე გამოიყურება:

- შეყვანა. შეყვანისას მიღებული ინფორმაცია ითარგმნება სისტემისთვის მოსახერხებელ ფორმატში, მაგალითად, ხმოვანი შეტყობინება ითარგმნება ტექსტურ მონაცემებად მეტყველების ამოცნობის ტექნიკის გამოყენებით.
- მორფოლოგიური ანალიზი. აქ მეტყველების ნაწილები განისაზღვრება და სიტყვები გარდაიქმნება მათი ძირეული ფორმების მოსაძებნად; სიტყვები „იწმინდება" სუფიქსებისგან, დაბოლოებებისა და პრეფიქსებისგან. ამ პროცესს ქვია სთემინგი (Stamming). სიტყვის თავდაპირველ ფორმამდე დაყვანაც შესაძლებელია – მაგალითად, ზმნებისთვის ინფინიტივები შეირჩევა. ამ პროცესს ლემატიზაცია (Lemmatization) ეწოდება.
- სემანტიკური ანალიზი. სხვადასხვა სასწავლო ალგორითმის გამოყენებით, ტექსტი წარმოდგენილია მრავალგანზომილებიან ვექტორულ სივრცეში, რომელიც სისტემისთვის გასაგებია.

- კონტექსტის შენარჩუნება. სისტემა იყენებს რთულ ალგორითმებს, რათა მოერგოს საუბრის კონტექსტს: ტონალობა, თემა, წინა შენიშვნები.
- სტატისტიკური მეთოდები, რომელიც მანქანურ სწავლებაზეა დაფუძნებული. ტექსტურ მონაცემებზე გაწვრთნილი სისტემა აფასებს სხვადასხვა პასუხის ვარიანტების ალბათობებს, ამოწმებს მათ თავის ენობრივ მოდელთან და მომხმარებელს აძლევს სტატისტიკურად ყველაზე სავარაუდო შედეგს.

თანამედროვე ენობრივ მოდელებს შეუძლიათ მოცემულ თემაზე ტექსტის გენერირება და ადამიანთან დიალოგის შენარჩუნება. მაგალითად, GPT-3 ნეირონული ქსელი გაწვრთნილია ტექსტის პეტაბაიტებზე და მუშაობს ასობით მილიარდი პარამეტრით. მას შეუძლია პუბლიკაციების და სტატიების გენერირება, რომლებიც თითქმის არ განსხვავდება პროფესიონალი ჟურნალისტის მიერ დაწერილი სტატიებისგან.

მანქანური სწავლება: ალგორითმული გირლიანდები

მანქანური სწავლება გულისხმობს ალგორითმების მთელ სპექტრს, რომლებიც შექმნილია კომპიუტერული პროგრამის დასახმარებლად მუშაობაში პირდაპირი ბრძანებების გარეშე, ანუ სწავლისთვის. სწავლის პროცესი შესაძლებელი ხდება დიდი რაოდენობით მონაცემების გამოყენებით, რომლებიც შემდეგ ანალიზდება, ხდება კანონზომიერებების გამოვლენა. მოდელი აგებულია მოძიებული მონაცემების და დაკვირვებების საფუძველზე. საბოლოოდ ხდება ვარიანტების იდენტიფიკაცია, შესაძლებლობების გამოთვლა და, ამის საფუძველზე ქმედება.

ტრადიციული პროგრამირება vs. მანქანური სწავლება:

	ტრადიციული პროგრამა	**მანქანური სწავლება**
შეყვანა	მონაცემები + წესები	მონაცემები + პასუხები
გამოსავალი	პასუხები	წესები (თავისით!)

ანალოგია: წარმოიდგინეთ, ბავშვს ასწავლით ძაღლების ამოცნობას:

- **ტრადიციული მეთოდი:** „ძაღლს 4 ფეხი აქვს, ყეფს, ბეწვი აქვს, კუდი აქვს...“ – ყველა წესი ხელით ჩამოვწერეთ
- **მანქანური სწავლება:** 10,000 ძაღლის ფოტო ვაჩვენეთ და ვუთხარით „ეს ძაღლია“ – ბავშვმა **თავისით** გამოიყო, რა გამოარჩევს ძაღლს. წესები **თვითონ ისწავლა**.

რატომ არის ეს რევოლუციური: სპამ-ფილტრის შექმნა ტრადიციული მეთოდით – შეუძლებელია (სპამი მუდმივად იცვლება, ყველა წესის ჩაწერა

ვერ მოხდება). მანქანური სწავლება კი მილიონობით სპამ-წერილს „კითხულობს“ და **თავისით სწავლობს**, რა ნიშნავს სპამს – ყოველ ახალ ხრიკსაც ადაპტირდება.

მანქანური სწავლება აქტიურად გამოიყენება გამოსახულების ამოცნობაში, სამედიცინო დიაგნოსტიკაში, უპილოტო ავტომობილის გადაადგილების დაგეგმვასა და სხვა მრავალ სფეროში.

არსებობს მანქანური სწავლების სამი ტიპი:

- მეთვალყურეობითი სწავლება. მანქანური სწავლების ეს ტიპი ვარაუდობს, რომ სისტემა მუშაობს მონაცემებით, რომლებიც წინასწარ არის შერჩეული ან მონიშნული ადამიანების მიერ, მაგალითად, ფოტოებით, რომლებშიც ადამიანებმა უკვე ამოიცნეს ძაღლები ან ველოსიპედები. ცნობილი Captcha ნებისმიერ საიტზე შესვლისას, სადაც საჭიროა შემოთავაზებული ობიექტებიდან გარკვეული ობიექტების არჩევა, მხოლოდ ასეთი მაგალითია. ალგორითმი სწავლობს ამ მონაცემებს, აკავშირებს მას ეტიკეტებთან, შემდეგ კი შეუძლია შექმნას მოდელი და დაამუშაოს ახალი მონაცემები მისი ტრენინგის შედეგების საფუძველზე. მაგალითები: უძრავი ქონების საბაზრო ღირებულების გამოთვლა, დაავადების ალბათობის პროგნოზირება.

მეთვალყურეობითი სწავლება (Supervised Learning)

პრინციპი: სისტემა „მასწავლებელთან“ სწავლობს – ყოველი მონაცემი „ეტიკეტირებულია“ (სწორი პასუხი ცნობილია).

მაგალითი 1 – სიმსივნის დიაგნოსტიკა: AI-ს მიეწოდება 100,000 სამედიცინო სურათი, სადაც თითოეულზე ექიმებმა უკვე მონიშნეს: „სიმსივნე“ ან „ნორმა“. სისტემა სწავლობს, რა ნიშნები ასოცირდება სიმსივნესთან. შემდეგ ახალ სურათს დამოუკიდებლად აფასებს – Google-ის DeepMind-მა ამ მიდგომით დიაგნოსტიკაში ექიმების სიზუსტეს გაუტოლდა.

მაგალითი 2 – Netflix-ის რეიტინგები: მომხმარებელი ფილმებს 1-5 ვარსკვლავით აფასებს (ეტიკეტები). AI სწავლობს: „ვინც ამ ფილმებს მოიწონა, მათთვის ეს ახალი ფილმიც მოეწონება.

- ზედამხედველობის გარეშე სწავლება. ამ შემთხვევაში, სისტემა მუშაობს უზარმაზარი მოცულობით არაეტიკეტირებული მონაცემებით. ალგორითმის ამოცანაა ნიმუშების იდენტიფიცირება და მათი გამოყენება შემდეგი შემომავალი მონაცემების კლასიფიკაციისთვის. მაგალითები: ხელნაწერი ტექსტის დაბეჭდილ ტექსტად გადაქცევა, მომხმარებლებისთვის პერსონალური რეკომენდაციები (მაგალითად, რეკომენდაციები ფილმების ან პროდუქტების შესახებ).

ზედამხედველობის გარეშე სწავლება (Unsupervised Learning)

პრინციპი: სისტემა „ეტიკეტების“ გარეშე, თავისით პოულობს ნიმუშებს მონაცემებში.

მაგალითი 1 – მომხმარებლების სეგმენტაცია: ონლაინ მაღაზია AI-ს აძლევს ყველა მომხმარებლის ქცევის მონაცემებს – ყიდვების ისტორია, ვიზიტების სიხშირე, ფასების სპექტრი. AI „ეტიკეტების“ გარეშე თავისით გამოყოფს ჯგუფებს: „ფასდაკლებაზე ორიენტირებული“, „ლუქს-პროდუქტების მყიდველი“, „ახალი კოლექციების მოყვარული“ – მარქეტინგის გუნდს ეს სეგმენტები წინასწარ განსაზღვრული არ ჰქონდა.

მაგალითი 2 – Google News-ის დაჯგუფება: AI კითხულობს მილიონობით სტატიას და თავისით ახდენს მათ „თემებად“ დაჯგუფებას – ადამიანმა წინასწარ არ განუსაზღვრა, რომ „პოლიტიკა“, „სპორტი“, „ტექნოლოგია“ ცალ-ცალკე კატეგორიებია.

- გაძლიერებული სწავლება. ამ ტიპის სწავლება ვარაუდობს უკუკავშირის არსებობას ხელოვნური ინტელექტის სისტემისთვის: მონაცემებზე დაყრდნობით აგებული მოდელი ინჟინერის მიერ შეფასებულია როგორც სწორი ან არასწორი და ამ შეფასებების საფუძველზე, მოდელი განახლდება. მაგალითად, მძღოლის გარეშე მანქანის მართვის სისტემა, რომელიც თავს არიდებს შეჯახებას, დაჯილდოვდება და სწავლობს, თუ როგორ აიცილოს თავიდან მსგავსი ავარიები მომავალში.

გაძლიერებული სწავლება (Reinforcement Learning)

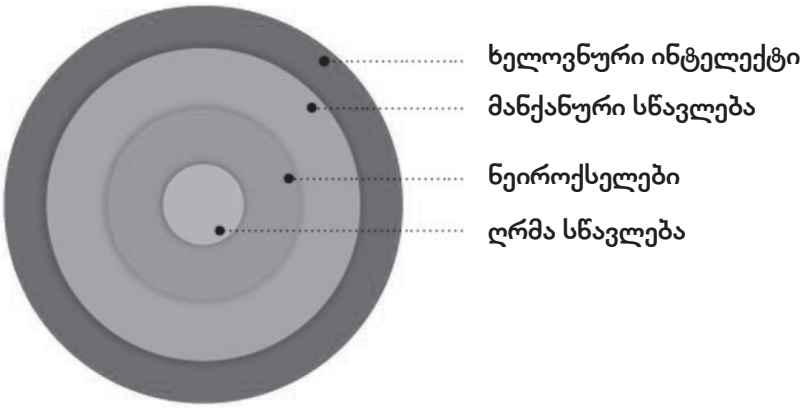
პრინციპი: სისტემა „ჯილდო-ჯარიმის“ მეთოდით სწავლობს – ისევე, როგორც ვავარჯიშებთ ძაღლს.

მაგალითი 1 – AlphaGo: DeepMind-ის AlphaGo-მ Go-ს თამაში ისწავლა მხოლოდ შემდეგი წესებით: „მოგება = +1 ჯილდო, წაგება = -1 ჯარიმა“. სისტემამ მილიონობითი თამაში ითამაშა საკუთარ თავთან, შეცდომებიდან ისწავლა – და საბოლოოდ მსოფლიო ჩემპიონს სძლია.

მაგალითი 2 – ChatGPT-ის „კარგი” პასუხების სწავლება: GPT-ის ბაზის ვარჯიშის შემდეგ, ადამიანი მომხმარებლები პასუხებს „ეწამება” (👍/👎). ეს შეფასებები AI-ს „ჯილდო/ჯარიმის“ სიგნალია – სისტემა სწავლობს, რა ტიპის პასუხები „ადამიანურად კარგია“.

მაგალითი 3 – Tesla Autopilot: ყოველი მართვის სესიის შემდეგ სისტემა „ჯილდოს“ იღებს: გზაზე დარჩენა, შეჯახების თავიდან აცილება, მარშრუტის ეფექტური გავლა. „ჯარიმა“: გამაფრთხილებელი სიგნალები, ოდესმე სისტემის ჩაბარებული ავარიები.

ღრმა სწავლება და ნეირონული ქსელები: ალგორითმული გირლანდები



ხელოვნური ნეირონული ქსელი ახდენს ბიოლოგიური ნეირონული ქსელის მუშაობის სიმულირებას. ქსელი ორგანიზებულია ფენებად: ერთი შემავალი, ერთი გამომავალი და სულ მცირე ერთი შუალედური. შემავალი ფენა იღებს მონაცემებს, გამომავალი ფენა აჩვენებს შედეგს, და შუალედური ფენები ასრულებენ გამოთვლებს. ყველა ფენა შედგება ხელოვნური ნეირონებისგან, რომლებიც ერთმანეთთან დაკავშირებულია და შემომავალ სიგნალებს ერთნაირად ამუშავებენ. მათი მუშაობის საიდუმლო კავშირებშია, რომლებსაც შეუძლიათ შეასუსტონ ან გააძლიერონ კონკრეტული სიგნალი.

ნეირონულ ქსელებს დიდი პოტენციალი აქვთ: მათ შეუძლიათ სურათებისა და სახეების ამოცნობა, უნიკალური ტექსტების, გრაფიკისა და ვიდეოების შექმნა, შავ-თეთრი ფოტოებისა და ფილმების გაფერადება, ფიზიკური და ფინანსური მოვლენების პროგნოზირება.

ნეირონული ქსელების ძალას აქვს უარყოფითი მხარე: ვერ ხდება ნეირონული ქსელების მიერ მიღებული გადაწყვეტილებების რეპროდუცირება, დადასტურება ან იმის გარკვევა, თუ როგორ მივიდნენ ისინი კონკრეტულ გადაწყვეტილებამდე. ეს პრობლემა წყდება ჩრდილოვანი სწავლებით, რომელიც მოიცავს წინასწარი ადამიანის მიერ პრობლემის დამუშავების ეტაპს. ჩრდილოვანი სწავლების მოდელის შედეგები უფრო გამჭვირვალე და მაღალპროდუქტიულია.

ღრმა სწავლების მეთოდები საშუალებას გვაძლევს, ადამიანის ტვინის მიერ მონაცემთა დამუშავების იმიტაცია მოვახდინოთ და შექმნილი შაბლონების საფუძველზე გადაწყვეტილებები მივიღოთ. ღრმა სწავლება მუშაობს მონაცემთა დიდ მოცულობაზე, მოითხოვს მნიშვნელოვან გამოთვლით სიმძლავრეს და ხელოვნურ ნეირონულ ქსელში მრავალ ფარულ ფენას.

YouTube რესურსები

1. ***"But what is a neural network?"*** – *YouTube: 3Blue1Brown* ნეირონული ქსელების ვიზუალური ახსნა – ერთ-ერთი საუკეთესო ვიდეო ინტერნეტში; ხელმისაწვდომია ქართული სუბტიტრებით
2. ***"Machine Learning for Everybody"*** – *YouTube: freeCodeCamp* მანქანური სწავლების სამი ტიპი (მეთვალყურეობითი, ზედამხედველობის გარეშე, გაძლიერებული) – სრული ახსნა მაგალითებით
3. ***"Natural Language Processing In 5 Minutes"*** – *YouTube: Simplilearn* NLP-ის – მეტყველების ამოცნობიდან ტექსტის გენერაციამდე – სწრაფი მიმოხილვა
4. ***"Bayes theorem, the geometry of changing beliefs"*** – *YouTube: 3Blue1Brown* ბაიესის თეორემის ვიზუალური ინტუიტური ახსნა – ყოველგვარი ფორმულის შიშის გარეშე
5. ***"AlphaGo – The Movie"*** – *YouTube: DeepMind* გაძლიერებული სწავლების ყველაზე ეფექტური დემონსტრაცია – AI მსოფლიო ჩემპიონს ამარცხებს

⚠ სათაურები პირდაპირ YouTube-ზე მოძებნეთ.

სადისკუსიო თემები

AI „სწავლობს" – ნამდვილ სწავლასთან რა საერთო აქვს ამას?
მანქანური სწავლება და ადამიანის სწავლა – ანალოგია თუ მეტაფორა?
განიხილეთ: ადამიანი ერთი მაგალითიდან სწავლობს (ბავშვი ერთხელ ხელს ცეცხლს შეეხება და მთელი ცხოვრება ახსოვს). AI-ს კი მილიონობითი მაგალითი სჭირდება. ადამიანს გადააქვს ცოდნა ერთი სფეროდან მეორეში; AI – ყოველ ახალ ამოცანაზე ახლიდან „ვარჯიშობს".
ნიშნავს ეს, რომ AI-ის „სწავლა" სულ სხვა მოვლენაა? შეიძლება ოდესმე AI ადამიანის მსგავსად, „ეფექტურად" ისწავლოს?

ვისი „ეტიკეტები" ასწავლიან AI-ს – და ვინ გადაწყვეტს, რა არის „სწორი"?
მეთვალყურეობითი სწავლება ეყრდნობა ადამიანების გადაწყვეტილებებს – მაგრამ ადამიანები მიკერძოებულები არიან.
განიხილეთ: Amazon-მა სცადა AI-ით დასაქმების CV-ების გაფილტვრა. სისტემამ ისწავლა „ეტიკეტებიდან" – ანუ წარსული დაქირავებების გადაწყვეტილებებიდან. გამოვლინდა, რომ სისტემა ქალ კანდიდატებს უარყოფდა – რადგან წარსულის გადაწყვეტილებებიც მიკერძოებული იყო. AI-მ ეს მიკერძოება „ისწავლა".
ვინ აგებს პასუხს? ადამიანი, რომელმაც „ეტიკეტები" შექმნა? კომპანია, რომელმაც სისტემა ააწყო? სახელმწიფო?

„შავი ყუთი“ – უნდა ვენდოთ AI-ს, რომლის ახსნაც არ შეგვიძლია?
ნეირონული ქსელები ხშირად „ვერ ხსნიან“ საკუთარ გადაწყვეტილებებს – ეს პრობლემაა განათლებაში?
განიხილეთ: თუ AI-ს სისტემა სტუდენტს ეუბნება „შენი ესე C+-ია“, სტუდენტს უფლება აქვს იცოდეს, **რატომ**. *მაგრამ ღრმა სწავლების მოდელს „პასუხი“ შეიძლება არ ჰქონდეს – ის მილიარდობითი პარამეტრის კომბინაციის შედეგია. ბაიესის ქსელი კი – „გამჭვირვალეა“ (შეიძლება ახსნა).*
რომელი ტიპის AI არის უფრო შესაფერისი განათლებაში?
გამჭვირვალობა vs. სიზუსტე – როგორ ვაბალანსოთ?

ტესტი 8

1. სუსტი ხელოვნური ინტელექტი (Narrow AI) განსხვავდება ძლიერი AI-სგან შემდეგი ნიშნით:

ა) სუსტი AI უფრო ჭკვიანია
ბ) სუსტი AI ასრულებს კონკრეტულ, შეზღუდულ ამოცანებს, ძლიერი AI კი ბაძავს ადამიანის ზოგად ინტელექტს
გ) სუსტი AI ადამიანის ჩარევას საჭიროებს
დ) სუსტი AI მხოლოდ ტექსტს ამუშავებს

2. ჯონ სირლი ამტკიცებს, რომ:

ა) AI უკვე ძლიერ ინტელექტს ფლობს
ბ) კომპიუტერი სიმბოლოებს ამუშავებს, მაგრამ ნამდვილ „გაგებას“ ვერ მიაღწევს
გ) NLP ახდენს ადამიანის ენის სრულ გაგებას
დ) ბაიესის ქსელები ცნობიერების მოდელია

3. „ალგორითმული გირლიანდა“ ნიშნავს:

ა) ერთ, ძალიან რთულ ალგორითმს
ბ) ალგორითმების ჯაჭვს, სადაც ყოველი ალგორითმი წინას შედეგს ამუშავებს
გ) გრაფიკულ AI სისტემას
დ) NLP-ის სინონიმს

4. ბაიესის პირობითი ალბათობა P(X|Y) ნიშნავს:

ა) X და Y-ის ერთდროული ალბათობა
ბ) X ან Y-ის ალბათობა
გ) X-ის ალბათობა, თუ ვიცით, რომ Y მოხდა
დ) Y-ის ალბათობა X-ის გარეშე

5. NLP-ში „სთემინგი“ (Stemming) არის:

ა) ტექსტის თარგმნა
ბ) ტექსტის სეგმენტება წინადადებებად
გ) სიტყვების გაწმენდა სუფიქსებისა და დაბოლოებებისგან ძირის მოსაძებნად
დ) ტექსტის ხმად გარდაქმნა

6. მეთვალყურეობითი სწავლება (Supervised Learning) განსხვავდება ზედამხედველობის გარეშე სწავლებისგან შემდეგი ნიშნით:

ა) მეთვალყურეობითი სწავლება უფრო სწრაფია
ბ) მეთვალყურეობით სწავლებაში მონაცემები წინასწარ „ეტიკეტირებულია“
გ) ზედამხედველობის გარეშე სწავლება უფრო ზუსტია
დ) მეთვალყურეობითი სწავლება AI-ს გარეშეა შესაძლებელი

7. CAPTCHA-ს ამოხსნა (სურათებზე ობიექტების ამოცნობა) ემსახურება:

ა) მხოლოდ უსაფრთხოების მიზნებს
ბ) ადამიანების მიერ AI-ის მეთვალყურეობითი სწავლებისთვის მონაცემების „ეტიკეტირებას“
გ) NLP-ის ვარჯიშს
დ) ბაიესის ქსელების შექმნას

8. ღრმა სწავლება (Deep Learning) განსხვავდება ჩვეულებრივი მანქანური სწავლებისგან:

ა) ნაკლები მონაცემებით მუშაობს
ბ) ფასდება ადამიანების მიერ
გ) ნეირონულ ქსელში მრავალ „ფარულ ფენას“ იყენებს
დ) მხოლოდ ტექსტის დამუშავებაა

9. გაძლიერებული სწავლების (Reinforcement Learning) ძირითადი პრინციპია:

ა) ეტიკეტირებული მონაცემების გამოყენება
ბ) ნიმუშების ძიება არაეტიკეტირებულ მონაცემებში
გ) „ჯილდო-ჯარიმის“ სიგნალებით სწავლება
დ) ექსპერტების შეფასებებზე დაყრდნობა

10. McKinsey-ის ექსპერტების პროგნოზით, ძლიერი ხელოვნური ინტელექტი ჩამოყალიბდება:

ა) 2030 წლისთვის
ბ) 2050 წლისთვის
გ) 2075 წლისთვის
დ) 22-ე საუკუნის ბოლოსთვის

გენერირებადი AI

2022 წლის ბოლოს გამოვიდა ChatGPT, რომელიც იყო პირველი მარტივად გამოსაყენებელი გენერირებადი ხელოვნური ინტელექტის (GenAI) ინსტრუმენტი და რომელიც ფართოდ გახდა ხელმისაწვდომი საზოგადოებისთვის. მას შემდეგ თვით ChatGPT მნიშნელოვნად დაიხვეწა და გაჩნდა სხვა, უფრო დახვეწილი ვერსიები (მაგალითად Claud, DeepSeek).

*„**გენერირებადი**“ ნიშნავს „შემქმნელი“ ან „მწარმოებელი“ – AI, რომელიც **ახალ შინაარსს ქმნის** და არა მხოლოდ არსებულს პოულობს.*

ანალოგია – ბიბლიოთეკარი vs. მწერალი:

***ჩვეულებრივი ძიება (Google)** = ბიბლიოთეკარი: თქვენ ეკითხებით, ის კარადებში შეყოფს და **უკვე არსებულ** წიგნს მოგიტანს.*

***გენერირებადი AI (ChatGPT)** = მწერალი: თქვენ ეკითხებით, ის **ახლავე წერს** თქვენთვის ახალ ტექსტს – წიგნი ადრე არ არსებობდა.*

სხვა სახის „გენერირება”:

***ტექსტი:** პასუხები, ესეები, კოდი, შეჯამებები*

***სურათები:** DALL-E, Midjourney – ახალი გამოსახულებები „ნულიდან“*

***ხმა:** მუსიკის შექმნა (შუნო), ხმის კლონირება*

***ვიდეო:** Sora – ვიდეო ტექსტური აღწერიდან*

***მთავარი განსხვავება:** ჩვეულებრივი AI **პოულობს ნიმუშებს**; გენერირებადი AI **ახალ ნიმუშებს ქმნის** – ისეთ კომბინაციებს, რომლებიც ტრენინგის მონაცემებში შეიძლება არც ყოფილა.*

ChatGPT-ის გამოჩენამ შეიძლება ითქვას, რომ შოკისმომგვრელი ცვლილებების ტალღები გამოიწვია მთელ მსოფლიოში. ამან საწყისი დაუდო მსხვილ ტექნოლოგიურ კომპანიებს შორის შეჯიბრებას GenAI მოდელების შექმნის მიზნით.

მთელ მსოფლიოში განათლებაში თავდაპირველი შეშფოთების მთავარი მიზეზი იყო ის, რომ ChatGPT-ს და მსგავს GenAI ინსტრუმენტებს მოსწავლეები/სტუდენტები გამოიყენებდნენ დავალებების შესრულებისას მოტყუებისთვის, და ამით შეარყევდნენ სწავლის შეფასების, სერტიფიცირებისა და კვალიფიკაციის ღირებულებას.

მიუხედავად იმისა, რომ ზოგიერთმა დაწესებულებამ აკრძალა ChatGPT-ის გამოყენება (თუმცა გაუგებარი იყო ამ აკრძალვას როგორ გააკონტროლებდნენ), სხვები სიფრთხილით მიესალმნენ GenAI-ის მოსვლას. მაგალითად, ბევრმა სკოლამ და უნივერსიტეტმა პროგრესული მიდგომა აირჩია: მათ მიაჩნიათ, რომ „გამოყენების აკრძალვის ნაცვლად, სტუდენტებს და პერსონალს უნდა დაეხმარონ GenAI ინსტრუმენტების ეფექტურად, ეთიკურად და გამჭვირვალედ გამოყენებაში“.

რეალური მაგალითი: ნიუ-იორკის საჯარო სკოლები (2023 წლის იანვარი)
ნიუ-იორკის განათლების დეპარტამენტი პირველი მსხვილი სასკოლო სისტემა გახდა, რომელმაც ***ChatGPT*-ი სრულად აკრძალა** *სკოლის ქსელებსა და მოწყობილობებზე – „სწავლაზე უარყოფითი გავლენის“ და „შინაარსის სიზუსტის“ შეშფოთების გამო.*

რატომ არ გაამართლა:

1. **ტექნიკური უძლურება:** *ChatGPT-ი ხელმისაწვდომია* **ნებისმიერი სმარტფონიდან** *– სკოლის ქსელის დაბლოკვა ვერ შეაჩერებდა სტუდენტებს, რომლებსაც მობილური ინტერნეტი ჰქონდათ.*
2. **VPN და გვერდის ავლა:** *ტექნიკურად მცოდნე სტუდენტები VPN-ის ან სხვა სერვისების მეშვეობით ადვილად უვლიდნენ გვერდს აკრძალვას.*
3. **სახლის გამოყენება:** *აკრძალვა სკოლის კედლებს გარეთ ვერ ვრცელდებოდა – სტუდენტები სახლიდან იყენებდნენ.*
4. **განათლების ნაცვლად – „კრიმინალიზაცია“:** *აკრძალვა ასწავლიდა, რომ AI „ცუდი“ და „სახიფათოა“ – ნაცვლად იმისა, რომ სტუდენტებს ასწავლოს* **პასუხისმგებლობიანი გამოყენება**.
5. **2023 წლის მაისი – უკუსვლა:** *ნიუ-იორკმა სულ 4 თვეში* **გააუქმა** *საკუთარი აკრძალვა და განაცხადა, რომ „AI-ი განათლებაში ინტეგრირების გზები უნდა ვეძებოთ.“*

საქართველოს კონტექსტი: *ანალოგიური სიტუაცია შეიქმნა ქართულ უნივერსიტეტებშიც – „Turnitin AI Detection“-ის გამოყენება სასტარტო ინიციატივად განიხილებოდა, მაგრამ ექსპერტები აფრთხილებდნენ, რომ ასეთი სისტემები ხშირად* **ცრუ-პოზიტიურ** *შედეგებს იძლევა.*

დასკვნა: *აკრძალვები „მოძველებული ციხის კედლის“ მშენებლობას ჰგავს – AI „წყალივით“ ყველა ბზარში შეაღწევს. ეფექტური სტრატეგია გულისხმობს* **ეთიკური გამოყენების კულტურის** *ჩამოყალიბებას.*

ასეთი მიდგომა აღიარებს, რომ GenAI ფართოდ არის ხელმისაწვდომი, სავარაუდოდ გაუმჯობესდება და აქვს როგორც უარყოფითი, ასევე უნიკალური დადებითი პოტენციალი განათლებისთვის.

მართლაც, GenAI-ს მრავალი პოტენციური გამოყენება აქვს. მას შეუძლია ავტომატიზირება გაუკეთოს ინფორმაციის დამუშავებას და შედეგების წარმოდგენას ადამიანის აზროვნების ყველა ძირითად სიმბოლურ წარმოდგენაში. ის საშუალებას იძლევა სწავლის საბოლოო შედეგების მიღწევა განხორციელდეს ნახევრად დასრულებული ცოდნის მიწოდებით. GenAI ადამიანებს ათავისუფლებს დაბალი დონის აზროვნების უნარების ზოგიერთი კატეგორიისგან და ამის შედეგად ხელოვნური ინტელექტის ინსტრუმენტების ამ ახალ თაობას

შეუძლია ღრმა გავლენა მოახდინოს ჩვენს მიერ ადამიანის ინტელექტისა და სწავლის გაგებაზე.

ამასთან, GenAI ასევე წარმოშობს მრავალ საფუძვლიან შეშფოთებას, რომელიც დაკავშირებულია ისეთ საკითხებთან, როგორიცაა უსაფრთხოება, მონაცემთა კონფიდენციალურობა, საავტორო უფლებები და მონაცემებით მანიპულირება. ზოგიერთი მათგანი წარმოადგენს ხელოვნურ ინტელექტთან დაკავშირებულ უფრო ფართო რისკებს, რომლებიც კიდევ უფრო გაამწვავა GenAI-მა, ზოგი კი წარმოიშვა ინსტრუმენტების უახლესი თაობის მოსვლასთან ერთად. ახლა აუცილებელია, რომ თითოეული ეს საკითხი და შეშფოთება სრულად იქნას გაგებული და მოგვარებული.

რა არის გენერირებადი AI?

გენერირებადი ხელოვნური ინტელექტი (GenAI) არის ხელოვნური ინტელექტის (AI) ტექნოლოგია, რომელიც ავტომატურად წარმოქმნის კონტენტს საუბრო ინტერფეისებში ბუნებრივი ენის შეკითხვებზე საპასუხოდ. თუ მანამდე ჩვენ გვიხდებოდა სხვადასხვა ვებ-გვერდების საშუალებით არსებული კონტენტის მოძიება, ამის ნაცვლად, GenAI რეალურად ქმნის ახალ კონტენტს.

„ქმნის ახალ კონტენტს“ – რას ნიშნავს

კითხვა: *თუ AI „ისწავლა“ ადამიანების ტექსტებიდან – ახალს ნამდვილად ქმნის?*

პასუხი – „ახლის“ სამი დონე:

დონე 1 – რეკომბინაცია (ყველაზე ხშირი): *AI გამოარჩევს ნიმუშებს – „ამ ტიპის კითხვებს ამ ტიპის პასუხი მოჰყვება“ – და* ***ახლებურად აკომბინირებს*** *სიტყვებს. შედეგი ახალია ფორმით, მაგრამ არა იდეებით.*

დონე 2 – ინტერპოლაცია: *AI „ავსებს სივრცეს“ ცნობილ ცნებებს შორის. მაგ.: „დაწერე ლექსი კვანტური ფიზიკისა და სიყვარულის შესახებ“ – ეს ორი სფერო ტრეინინგ მონაცემებში ერთად შეიძლება არ ყოფილა, მაგრამ AI* ***ინტერპოლირებს****.*

დონე 3 – ემერჯენტული შედეგები: *ზოგჯერ LLM-ები „გასცდებიან“ ტრენინგ მონაცემებს – წარმოქმნიან გამოსავლებს, რომლებიც* ***არ ყოფილა*** *ნედლ მონაცემებში. ამ მოვლენას* ***„ემერჯენტული შესაძლებლობა“*** *ეწოდება.*

მნიშვნელოვანი შეზღუდვა: *AI „ახალი“ კონტენტი მაინც* ***სტატისტიკური პროგნოზია*** *– „რომელი სიტყვა/პიქსელი ყველაზე სავარაუდოა შემდეგი?“ ეს ადამიანის შემოქმედებისგან პრინციპულად განსხვავდება, სადაც „ახალი“ ემყარება რეალური სამყაროს გაგებას, ემოციებს და განზრახვას.*

კონტენტი შეიძლება იყოს ფორმატში, რომელიც მოიცავს ადამიანის აზროვნების ყველა სიმბოლურ წარმოდგენას: ბუნებრივ ენაზე დაწერილი ტექსტი, სურათები (ფოტოებიდან ციფრულ ნახატებამდე და მულტფილმებამდე), ვიდეო, მუსიკა და კომპიუტერული კოდი. GenAI სწავლობს ვებ-გვერდებიდან, სოციალური მედიის დისკუსიებიდან და სხვა ონლაინ რესურსებიდან შეგროვე-

ბული მონაცემებიდან. ის წარმოქმნის თავის კონტენტს მონაცემებში სიტყვების, პიქსელების ან სხვა ელემენტების განაწილების სტატისტიკურად ანალიზით და ზოგადი ნიმუშების იდენტიფიცირებით და გამეორებით (მაგალითად, რომელი სიტყვები მიჰყვება რომელ სიტყვებს).

მიუხედავად იმისა, რომ GenAI-ს შეუძლია ახალი კონტენტის შექმნა, მას არ შეუძლია რეალური სამყაროს პრობლემების ახალი იდეების ან გადაწყვეტილებების გენერირება, რადგან ის არ იცნობს შინაარსობრივად რეალური სამყაროს ობიექტებს ან სოციალურ ურთიერთობებს, რომლებიც ენის საფუძველს წარმოადგენს. გარდა ამისა, მისი ხშირად მართლაც შთამბეჭდავი გამომავალი პროდუქტის მიუხედავად, ჩვენ ვერ ვიქნებით დარწმუნებული GenAI-ს კონტენტის სიზუსტეში. სინამდვილეში, ChatGPT-ის შემქმნელებიც აღიარებენ: „მიუხედავად იმისა, რომ ისეთ ინსტრუმენტებს, როგორიცაა ChatGPT, ხშირად შეუძლიათ გონივრული პასუხების გენერირება, მათ სიზუსტეზე დაყრდნობა შეუძლებელია“. უმეტეს შემთხვევაში, ChatGPT-ის შეცდომები შეუმჩნეველი რჩება, თუ მომხმარებელს არ აქვს განსახილველი თემის ღრმა ცოდნა.

როგორ მუშაობს გენერირებადი AI?

GenAI-ის საფუძვლად მყოფი სპეციფიკური ტექნოლოგიები ხელოვნური ინტელექტის ტექნოლოგიების ოჯახის ნაწილია, რომელსაც მანქანური სწავლება (ML) ეწოდება და რომელიც იყენებს ალგორითმებს, რომლებიც საშუალებას აძლევს მას მუდმივად და ავტომატურად გააუმჯობესოს თავისი მუშაობა სულ უფრო მეტ მონაცემებზე დაყრდნობით. ხელოვნური ინტელექტის ის ტიპი, რომელმაც ბოლო წლებში ხელოვნურ ინტელექტში მრავალი მიღწევა განაპირობა, როგორიცაა ხელოვნური ინტელექტის გამოყენება სახის ამოცნობისთვის, ცნობილია, როგორც ხელოვნური ნეირონული ქსელები (ANN). არსებობს ANN-ების მრავალი ტიპი. ტექსტისა და სურათების გენერირების ხელოვნური ინტელექტის ტექნოლოგიები ეფუძნება ხელოვნური ინტელექტის ტექნოლოგიების ერთობლიობას, რომლებიც მკვლევარებისთვის ბოლო რამდენიმე წლის განმავლობაში გახდა ხელმისაწვდომი. მაგალითად, ChatGPT იყენებს გენერაციულ წინასწარ გაწვრთნილ ტრანსფორმატორს (GPT), ხოლო GenAI ჩვეულებრივ იყენებს ე.წ. გენერაციულ კონკურენტულ ან პარარელურ შეჯიბრებით ქსელებს (GAN) გამოსახულების დამუშავებისთვის:

Machine Learning (ML)		**The type of AI that uses data to automatically improve its performance**
Artificial Neural Network (ANN)		**A type of ML that is inspired by the structure and functioning of the human brain**
Text Generative AI	**General Purpose Transformers**	**A type of ANN that is capable on focusing on different parts of data to determine how the relate to each other**
	Large Language Models (LLM)	**A type of General-purpose Transformer that is trained on vast amounts of text data**
	Generative Pre-Trained Transformers (GPT)	**A type of LLM that is pre-trained on larger amounts of data, which allows to capture and use the nuances**
Image Generative AI	**Generative Adversarial Networks (GAN)**	**Type of Neural Network used for image generation**
	Variational Autoencoders (VAE)	

ChatGPT თუ DeepSeek თუ Claude AI?

შევადაროთ ერთმანეთს სამი ყველაზე პოპულარული Gen AI მთავარი მახასიათებლების მიხედვით.

1. ცოდნა და მსჯელობა

ChatGPT (GPT-4o): ძლიერია შემოქმედებით წერაში, ფილოსოფიაში; ადამიანური და ინტერაქტიული სტილი.

DeepSeek-V3: ღრმა ცოდნა ტექნიკურ, სამეცნიერო და მათემატიკურ სფეროებში; ლოგიკური და ზუსტი.

Claude (Anthropic): გამოირჩევა **ნიუანსირებული, ანალიტიკური მსჯელობით** – განსაკუთრებით რთული, მრავალმხრივი კითხვების განხილვაში. Claude ყურადღებას ამახვილებს **სიზუსტეზე და გამჭვირვალობაზე** – ხშირად პირდაპირ აღნიშნავს, სად არ არის დარწმუნებული.

დასკვნა: Claude გამოირჩევა კრიტიკული და ეთიკური მსჯელობის სფეროში.

2. კოდირება და ტექნიკური უნარები

ChatGPT: ძალიან კარგი კოდირების ასისტენტია; ნათელი ახსნები აქვს.

DeepSeek-V3: ერთ-ერთი საუკეთესო უფასო მოდელი პროგრამირებისთვის; 128K კონტექსტი.

Claude: Claude 3.5/3.7 Sonnet განსაკუთრებით ძლიერია კოდირებაში – ინდუსტრიის ბენჩმარქებზე ხშირად GPT-4o-ს სჯობს. Anthropic-მა შექმნა **Claude Code** – სპეციალური ინსტრუმენტი დეველოპერებისთვის.

დასკვნა: Claude და DeepSeek ლიდერობენ კოდირებაში.

3. მულტიმოდალური შესაძლებლობები

ChatGPT (GPT-4o): სურათი, აუდიო, ვიდეო – სრული მულტიმოდალობა.

DeepSeek-V3: მხოლოდ ტექსტი და დოკუმენტები.

Claude: კითხულობს სურათებს, PDF-ებს, კოდს, ცხრილებს. აუდიო/ვიდეო ჯერ არ აქვს. **განსაკუთრებით ძლიერია გრძელი დოკუმენტების ანალიზში** – 200K+ ტოკენის კონტექსტი.

დასკვნა: ChatGPT – მულტიმოდალობაში; Claude – გრძელი დოკუმენტების ანალიზში.

4. უსაფრთხოება და ეთიკა

ChatGPT: OpenAI-ს მოდერაციის სისტემა; ზოგჯერ „ზედმეტად ფრთხილი".

DeepSeek-V3: ჩინური კომპანიის პროდუქტი; **მნიშვნელოვანი კონფიდენციალობის შეშფოთება** – მონაცემები ჩინეთის სერვერებზე ინახება; ზოგიერთ პოლიტიკურ თემაზე შეზღუდული პასუხები.

Claude: Anthropic-ი **„AI უსაფრთხოების" კომპანიად** პოზიციონირდება. Claude-ის ეთიკური პრინციპები (Constitutional AI) სისტემის „ხერხემლია". **ყველაზე გამჭვირვალე** უარის თქმის მიზეზების ახსნაში.

დასკვნა: Claude – ეთიკასა და უსაფრთხოებაში; DeepSeek – კონფიდენციალობისადმი სიფრთხილათ.

5. ღირებულება და ხელმისაწვდომობა

ChatGPT: უფასო (შეზღუდული) + $20/თვე (Plus).

DeepSeek-V3: სრულად უფასო, მაღალი ლიმიტებით.

Claude: უფასო ვერსია (Claude.ai) + $20/თვე (Pro). API – დეველოპერებისთვის.

დასკვნა: DeepSeek – ფასით; Claude და ChatGPT – ფასიანი ვერსიების ხარისხით.

„რომელი უნდა აირჩიოთ?"

ამოცანა	*რეკომენდებული*
შემოქმედებითი წერა, მარკეტინგი	*ChatGPT*
ტექნიკური კოდირება, დიდი კოდბაზა	*DeepSeek ან Claude*
სურათების/აუდიოს ანალიზი	*ChatGPT*
გრძელი დოკუმენტების ანალიზი	*Claude*
ეთიკური, ნიუანსირებული მსჯელობა	*Claude*
ტექნიკური/სამეცნიერო კვლევა	*DeepSeek*
კონფიდენციალობის პრიორიტეტი	*Claude ან ChatGPT*
უფასო, ძლიერი გამოყენება	*DeepSeek (ტექნიკური) / Claude (ანალიტიკური)*

საბოლოო დასკვნა: სამივე ინსტრუმენტი ლიდერობს სხვადასხვა სფეროში. „საუკეთესო" GenAI – არის ის, რომელიც **თქვენი კონკრეტული ამოცანისთვის** ყველაზე შესაფერისია.

YouTube რესურსები

1. **"ChatGPT vs Claude vs Gemini – Full Comparison 2024"** – YouTube: AI Explained *სამი წამყვანი GenAI მოდელის პრაქტიკული შედარება რეალური ამოცანებით*
2. **"How ChatGPT actually works"** – YouTube: AssemblyAI *GPT-ის, LLM-ების და ტოკენიზაციის გასაგები ახსნა – ტექნიკური ფონის გარეშე*
3. **"Generative AI in Education: Opportunities and Challenges"** – YouTube: UNESCO *GenAI-ის საგანმანათლებლო პოტენციალი და რისკები – UNESCO-ს ოფიციალური პოზიცია*
4. **"The AI Cheating Crisis in Schools"** – YouTube: Vox *რატომ ვერ მუშაობს აკრძალვები და რა ალტერნატივები არსებობს – ნიუ-იორკის მაგალითი*
5. **"How Generative AI Works (DALL-E, GPT, Stable Diffusion)"** – YouTube: 3Blue1Brown / Computerphile *GAN-ების, GPT-ის და Diffusion მოდელების ვიზუალური ახსნა*

⚠ *სათაურები პირდაპირ YouTube-ზე მოძებნეთ.*

სადისკუსიო თემები

GenAI – ინსტრუმენტი თუ „ეფექტი"? სად არის ზღვარი?
თუ სტუდენტი GenAI-ით ასრულებს ესეს – ვის ეკუთვნის ეს ნამუშევარი?
*განიხილეთ: ქართული სამართლით, AI-ის მიერ გენერირებული შინაარსი ავტომატურად საავტორო უფლებებით დაცული არ არის. მაგრამ განათლებაში კიდევ უფრო ღრმა კითხვაა: თუ AI „დაწერა" ესე, **ვინ ისწავლა**?*
„ინსტრუმენტის" გამოყენება (მაგ. კალკულატორი მათემატიკაში) ზოგჯერ მისაღებია – სად გადის ზღვარი GenAI-სა და „მოტყუებას" შორის?
შეიძლება „GenAI-თან კოლაბორაცია" ახალ, ლეგიტიმურ უნარად ჩამოყალიბდეს?

„ჰალუცინაციები“ – როდის არის GenAI საშიში?
ChatGPT ხშირად „ტყუის“ დარწმუნებული სახით – როგორ ვასწავლოთ კრიტიკული შეფასება?
*განიხილეთ: GenAI-ის „ჰალუცინაციები“ (მოგონილი ფაქტები, ყალბი ციტატები, არარსებული წყაროები) განსაკუთრებით სახიფათოა, როდესაც მომხმარებელს **ღრმა ცოდნა არ აქვს** თემაზე – ანუ ზუსტად მაშინ, როდესაც ყველაზე მეტად ეყრდნობა AI-ს. „ინფორმაციული გრამოტობა“ 21-ე საუკუნეში ნიშნავს AI-ის შედეგების კრიტიკულ შეფასებას.*
როგორ ვასწავლოთ ეს უნარი სკოლაში?
რა „წითელი დროშები“ უნდა ეძებოს სტუდენტმა?

GenAI და ციფრული უთანასწორობა – ყველასთვის თანაბარია?
GenAI „დემოკრატიზაციას“ ახდენს ცოდნაზე – თუ ახალ უთანასწორობას ქმნის?
*განიხილეთ: ერთი მხრივ, კარგი ინტერნეტი + ChatGPT სტუდენტს მისცემს „პირად მასწავლებელს“ 24/7. ეს შეიძლება განათლების „დემოკრატიზაციად“ ჩავთვალოთ. მეორე მხრივ: ვისაც **ძლიერი ბაზა** (კრიტიკული აზროვნება, ციფრული გრამოტობა, ენის ცოდნა) აქვს, GenAI-ს **ბევრად ეფექტურად** გამოიყენებს – „ისედაც ძლიერი“ კიდევ უფრო ძლიერდება.*
*შეიძლება GenAI-მ **გაზარდოს** სოციალური უთანასწორობა განათლებაში?*

ტესტი 9

1. „გენერირებადი AI” ჩვეულებრივი საძიებო სისტემისგან (Google) განსხვავდება:

ა) სწრაფდება ინტერნეტით
ბ) ახალ კონტენტს ქმნის, ნაცვლად არსებულის პოვნისა
გ) მხოლოდ ტექსტს ამუშავებს
დ) პასუხობს მხოლოდ ფაქტობრივ კითხვებს

2. ChatGPT-ის საფუძვლად მყოფი ტექნოლოგიაა:

ა) GAN (Generative Adversarial Network)
ბ) Bayesian Network
გ) GPT (Generative Pre-Trained Transformer)
დ) Variational Autoencoder

3. ნიუ-იორკის სკოლებმა ChatGPT-ის აკრძალვა რატომ გააუქმეს?

ა) AI-მ ყველა პრობლემა გადაჭრა
ბ) სტუდენტები არ იყენებდნენ ChatGPT-ს
გ) აკრძალვა ტექნიკურად გვერდის ავლადი იყო და კონტრპროდუქტიული
დ) მთავრობამ კანონი შეცვალა

4. GAN (Generative Adversarial Network) ძირითადად გამოიყენება:

ა) ტექსტის გენერაციისთვის
ბ) გამოსახულების (სურათების) გენერაციისთვის
გ) ხმოვანი ასისტენტებისთვის
დ) კოდის წერისთვის

5. GenAI-ის „ჰალუცინაცია“ ნიშნავს:

ა) AI-ის მიერ სურათების გენერაციას
ბ) AI-ის ემოციურ გამოხატვას
გ) AI-ის მიერ ყალბი, მაგრამ დარწმუნებულად წარმოდგენილი ინფორმაციის გენერაციას
დ) AI-ის კრეატიულ პასუხებს

6. Claude AI-ის შემქმნელი კომპანიაა:

ა) OpenAI
ბ) Google
გ) Meta
დ) Anthropic

7. DeepSeek-V3-ის მთავარი უპირატესობა ChatGPT-ის მიმართ:

ა) სურათების ანალიზი
ბ) ხმოვანი ინტერფეისი
გ) სრულიად უფასო პრემიუმ ფუნქციები + 128K კონტექსტი
დ) მობილური აპლიკაცია

8. Claude AI გამოირჩევა სხვა GenAI მოდელებისგან შემდეგი ნიშნით:

ა) ყველაზე იაფია
ბ) ყველაზე სწრაფია
გ) ეთიკურ მსჯელობასა და „Constitutional AI“ პრინციპებზე ფოკუსით
დ) მხოლოდ კოდირებისთვისაა

9. „Large Language Model” (LLM) არის:

ა) ნებისმიერი AI სისტემა
ბ) GAN-ის სახეობა
გ) General-Purpose Transformer, გაწვრთნილი ტექსტ-მონაცემების უზარმაზარ მასივზე
დ) სამედიცინო AI-ის ტიპი

10. „Constitutional AI“ – Anthropic-ის მიდგომა – ნიშნავს:

ა) AI-ის სამართლებრივ რეგულაციას
ბ) AI-ის ეთიკური პრინციპების სისტემას, რომელიც მოდელის „ხერხემლია“
გ) AI-ის ტექნიკური არქიტექტურის ტიპს
დ) AI-ის კონსტიტუციური სახელმწიფოებში გამოყენებას

როგორ მუშაობს ტექსტური გენერირებადი AI?

ტექსტის გენერირების ხელოვნური ინტელექტი იყენებს ხელოვნური ნეირონის ტიპს, რომელიც ცნობილია როგორც ზოგადი ტრანსფორმატორი და ზოგადი ტრანსფორმატორის ტიპს, რომელსაც ეწოდება დიდი ენის მოდელი (LLM). ამ მიზეზით, ტექსტის ხელოვნური ინტელექტის სისტემებს ხშირად მოიხსენიებენ, როგორც დიდი ენის მოდელებს (LLM). ტექსტის შექმნისათვის ხელოვნური ინტელექტის მიერ გამოყენებული LLM-ის ტიპი ცნობილია, როგორც გენერირებადი წინასწარ მომზადებული ტრანსფორმატორი (GPT) (აქედან გამომდინარე, „GPT“ ჩადგმულია „ChatGPT“-ში). ChatGPT აგებულია GPT-3-ზე, რომელიც შემუშავებულია OpenAI-ის მიერ. ეს იყო GPT-ის მესამე იტერაცია, პირველი გამოვიდა 2018 წელს, ხოლო უახლესი, GPT-5, 2025 წლის მაისში.

GPT-5 *გამოვიდა 2025 წლის მაისში. ის წარმოადგენს GPT სერიის ყველაზე მძლავრ ვერსიას და მნიშვნელოვნად აჭარბებს GPT-4-ს შემდეგ სფეროებში:*

- ***მსჯელობა:*** *GPT-5 ბევრად უკეთ ახდენს მრავალსაფეხურიანი ლოგიკური პრობლემების გადაჭრას – კომბინირებს „სწრაფ“ და „ნელ“ (ღრმა) აზროვნებას.*
- ***სიზუსტე:*** *„ჰალუცინაციების“ მნიშვნელოვანი შემცირება – განსაკუთრებით ფაქტობრივ კითხვებში.*
- ***მულტიმოდალობა:*** *გაუმჯობესებული სურათების, კოდისა და გრძელი დოკუმენტების ანალიზი.*
- ***კონტექსტი:*** *გაფართოებული კონტექსტური ფანჯარა – შეუძლია გრძელი საუბრების „დამახსოვრება“.*
- ***ინსტრუქციების შესრულება:*** *ბევრად ზუსტად ასრულებს კომპლექსურ, მრავალნაწილიან მოთხოვნებს.*

GPT სერიის ევოლუცია:

ვერსია	წელი	პარამეტრები	მთავარი სიახლე
GPT-1	**2018**	**117M**	**პირველი GPT**
GPT-2	**2019**	**1.5B**	**ტექსტის გენერაცია**
GPT-3	**2020**	**175B**	**Few-shot learning**
GPT-4	**2023**	**~1T (est.)**	**მულტიმოდალობა**
GPT-4o	**2024**	**–**	**სრული მულტიმოდალობა, სიჩქარე**
GPT-5	**2025**	**–**	**გაუმჯობესებული მსჯელობა, სიზუსტე**

თითოეული OpenAI GPT ახალი ვერსია აუმჯობესებდა წინა ვერსიას ხელოვნური ინტელექტის არქიტექტურის, ტრენინგის მეთოდებისა და ოპტიმიზაციის ტექნიკის მიღწევების გზით. მისი უწყვეტი პროგრესის ერთ-ერთი ცნობილი ასპექტია მზარდი მოცულობის მონაცემების გამოყენება ექსპონენციალურად მზარდი რაოდენობის „პარამეტრების“ ტრენინგისთვის. პარამეტრები შეიძლება განვიხილოთ, როგორც მეტაფორული ღილაკების ერთობლიობა, რომელთა რეგულირებაც შესაძლებელია GPT-ის მუშაობის დასაზუსტებლად. ისინი მოიცავს მოდელის „წონებს“, რიცხვით პარამეტრებს, რომლებიც განსაზღვრავენ, თუ როგორ ამუშავებს მოდელი შემავალ მონაცემებს და აგენერირებს გამოსავალს.

ხელოვნური ინტელექტის არქიტექტურისა და ტრენინგის მეთოდების ოპტიმიზაციის მიღწევების გარდა, ეს სწრაფი იტერაცია ასევე შესაძლებელი გახდა მონაცემთა უზარმაზარი მოცულობების და მსხვილი კომპანიებისთვის ხელმისაწვდომი ციფრული/გამოთვლითი სიმძლავრის გაუმჯობესების წყალობით. 2012 წლიდან, GenAI მოდელების ტრენინგისთვის გამოყენებული გამოთვლითი სიმძლავრე ყოველ 3-4 თვეში ორმაგდება.

GPT-ის გაწვრთნის შემდეგ, ტექსტური პასუხის გენერირება მოთხოვნაზე მოიცავს შემდეგ ნაბიჯებს:

მოთხოვნა იყოფა უფრო პატარა ერთეულებად (რომელსაც ტოკენები ეწოდება), რომლებიც შეჰყავთ GPT-ში.

***ტოკენი** არის ტექსტის ყველაზე პატარა „ერთეული“, რომელსაც GPT ამუშავებს – სიტყვის ნაწილი, სიტყვა, ან სასვენი ნიშანი.*

***ანალოგია – LEGO:** წარმოიდგინეთ, რომ ტექსტი LEGO-ს კონსტრუქციაა. ტოკენები – ცალ-ცალკე LEGO-ს კუბიკები. GPT ტექსტს „ანადგურებს“ კუბიკებად, ამუშავებს თითოეულს, შემდეგ ახალ კონსტრუქციას – პასუხს – ხელახლა აკრავს.*

კონკრეტული მაგალითი:

- *„ChatGPT არის ინსტრუმენტი“ → დაახლოებით 6-7 ტოკენი*
- *ინგლისურში: „Hello, how are you?“ → 6 ტოკენი:* Hello, how, are, you, ?
- *GPT-4-ს კონტექსტური ფანჯარა = 128,000 ტოკენი ≈ დაახლოებით 96,000 სიტყვა ≈ 300 გვერდიანი წიგნი*

***რატომ მნიშვნელოვანია:** GPT **ტოკენი ტოკენის მიყოლებით** გამოიმუშავებს პასუხს – ყოველ ახალ ტოკენამდე „ითვლის“, რომელი ტოკენი ყველაზე სავარაუდოა შემდეგი. ამიტომ GPT-ის პასუხი ეკრანზე **ასო-ასო ჩნდება** – ის „ფიქრობს“ ტოკენებით.*

1. GPT იყენებს სტატისტიკურ მონაცემებს და მეთოდებს იმის პროგნოზირებისთვის, რომლებიც შეიძლება ქმნიდნენ მოთხოვნაზე თანმიმდევრულ პასუხს.
   - GPT განსაზღვრავს სიტყვების ნიმუშებს და ფრაზებს, რომლებიც ხშირად გვხვდება მის წინასწარ აგებულ დიდი მონაცემთა მოდელში (რომელიც მოიცავს ინტერნეტიდან და სხვა წყაროებიდან აღებულ ტექსტს).

- ამ ნიმუშების გამოყენებით, GPT აფასებს კონკრეტული სიტყვების ან ფრაზების გამოჩენის ალბათობას მოცემულ კონტექსტში.
- შემთხვევითი პროგნოზირებით დაწყებული, GPT იყენებს ამ სავარაუდო ალბათობებს მისი პასუხის შემდეგი სავარაუდო სიტყვის ან ფრაზის პროგნოზირებისთვის.

2. პროგნოზირებული სიტყვები ან ფრაზები გარდაიქმნება წასაკითხ ტექსტად.
3. წასაკითხი ტექსტი იფილტრება ე.წ. „დამცავი ბარიერების“ მეშვეობით, რათა მოიხსნას ნებისმიერი შეურაცხმყოფელი შინაარსი.
4. მე-2-დან მე-4-მდე ნაბიჯები მეორდება პასუხის დასრულებამდე. პასუხი დასრულებულად ითვლება, როდესაც ის მიაღწევს ტოკენების მაქსიმალურ ლიმიტს ან აკმაყოფილებს წინასწარ განსაზღვრულ შეჩერების კრიტერიუმებს.
5. პასუხი შემდგომ დამუშავდება წაკითხვის გასაუმჯობესებლად ფორმატირების, პუნქტუაციის და სხვა გაუმჯობესებების გამოყენებით (მაგალითად, პასუხის დაწყება სიტყვებით, რომლებსაც ადამიანი შეიძლება იყენებდეს, მაგალითად, „რა თქმა უნდა“, „ცხადია“ ან „ბოდიშის მოხდით“).

მიუხედავად იმისა, რომ GPT-ები და მათი ტექსტის ავტომატურად გენერირების შესაძლებლობა მკვლევარებისთვის ხელმისაწვდომია 2018 წლიდან, ChatGPT-ის გაშვებას მთავარი სიახლე არის უფასო წვდომა მარტივი ინტერფეისის საშუალებით, რაც იმას ნიშნავს, რომ ყველას, ვისაც ინტერნეტზე წვდომა აქვს, შეეძლო ინსტრუმენტის გამოყენება.

ChatGPT-ის გაშვებამ შოკისმომგვრელი ტალღები გამოიწვია მთელ მსოფლიოში და სწრაფად აიძულა სხვა გლობალური ტექნოლოგიური კომპანიები, შეექმნათ მსგავსი სისტემები. ChatGPT-ის აქვს ბევრი ალტერნატივა:

- Alpaca: სტენფორდის უნივერსიტეტის LLM ვერსია, რომლის მიზანია ცრუ ინფორმაციის, სოციალური სტერეოტიპებისა და ტოქსიკური ენის წინააღმდეგ ბრძოლა.
- Bard: Google-ის LLM ვერსია, დაფუძნებული მის LaMDA და PaLM 2 სისტემებზე, რომელსაც აქვს წვდომა ინტერნეტზე რეალურ დროში, რაც ნიშნავს, რომ მას შეუძლია მიაწოდოს განახლებული ინფორმაცია.
- Chatsonic: შექმნილია Writesonic-ის მიერ, ის ეფუძნება ChatGPT-ს და ასევე პირდაპირ ამოწმებს მონაცემებს.
- Ernie (Wenxin Yiyan 文心一言): Baidu-ს LLM ვერსია, რომელიც ჯერ კიდევ განვითარების პროცესშია, რომელიც აერთიანებს ფართო ცოდნას უზარმაზარ მონაცემთა ნაკრებებთან ტექსტისა და სურათების გენერირებისთვის.
- Hugging Chat: შექმნილია HuggingFace-ის მიერ, რომელიც ხაზს უსვამს ეთიკას და გამჭვირვალობას მისი განვითარების, ტრენინგისა და განლაგების განმავლობაში. გარდა ამისა, მათი მოდელების სწავლებისთვის გამოყენებული ყველა მონაცემი ღია წყაროა.

- Jasper: ინსტრუმენტებისა და API-ების ნაკრები, რომლებიც, მაგალითად, შეიძლება გაწვრთნილ იქნას მომხმარებლის მიერ სასურველი სტილით წერაზე. მას ასევე შეუძლია სურათების გენერირება.
- Llama: Meta-ს ღია წყაროს LLM, რომელიც მოითხოვს ნაკლებ გამოთვლით სიმძლავრეს და ნაკლებ რესურსებს ახალი მიდგომების შესამოწმებლად, სხვების ნამუშევრის დასადასტურებლად და ახალი გამოყენების შემთხვევების შესასწავლად.
- Open Assistant: ღია წყაროს მიდგომა, რომელიც შექმნილია იმისთვის, რომ ნებისმიერს, ვისაც აქვს საკმარისი გამოცდილება, შეეძლოს საკუთარი LLM-ის შემუშავება. ის შეიქმნა მოხალისეების მიერ შემუშავებული ტრენინგის მონაცემების საფუძველზე.
- Tongyi Qianwen (通义千问): Alibaba-ს LLM, რომელსაც შეუძლია უპასუხოს მოთხოვნებს ინგლისურ ან ჩინურ ენებზე. ის ინტეგრირდება Alibaba-ს ბიზნეს ინსტრუმენტების კომპლექტში.
- YouChat: LLM, რომელიც მოიცავს რეალურ დროში ძიების შესაძლებლობებს დამატებითი კონტექსტისა და ინფორმაციის მიწოდებისთვის, რათა უფრო ზუსტი და სანდო შედეგები მიიღოთ.

2023-2025 წლებში სფერო სწრაფად განვითარდა. განახლებული სურათი:

Google:

- ***Bard** 2024 წელს გადაიქცა **Gemini**-დ. ჩემინი Ultra/Pro/Flash – სხვადასხვა სიმძლავრის მოდელები. Gemini 1.5 Pro-ს 1 მილიონი ტოკენის კონტექსტი აქვს.*

Meta:

- ***Llama 2** (2023) და **Llama 3** (2024) – ღია კოდის ყველაზე პოპულარული მოდელი. Llama 3.1 405B – ერთ-ერთი ყველაზე ძლიერი ღია მოდელი.*

Anthropic:

- ***Claude 3** ოჯახი (Haiku/Sonnet/Opus) – 2024. Claude 3.5 Sonnet კოდირებასა და ანალიზში ლიდერი. **Claude 4** (2025) – კიდევ უფრო გაუმჯობესებული.*

Microsoft/OpenAI:

- ***Copilot** – Microsoft-ის პროდუქტებში (Word, Excel, Teams) ჩაშენებული GPT.*

მნიშვნელოვანი ახალი მოთამაშეები:

- ***DeepSeek-V3/R1** (2024-2025, ჩინეთი) – უფასო, ძალიან ძლიერი.*
- ***Mistral** (საფრანგეთი) – ევროპული ღია კოდის LLM.*
- ***Grok** (xAI/Elon Musk)– X (Twitter)-ში ინტეგრირებული.*
- ***Perplexity AI** – LLM + რეალური დროის ძიება.*

ამ პროგრამების უმეტესობა უფასოა (გარკვეული ლიმიტების ფარგლებში), ზოგიერთი კი ღია კოდის. ბევრი სხვა პროდუქტი, გამოდის, რომლებიც ამ LLM-ებზეა დაფუძნებული.

მაგალითებია:

- ChatPDF: აჯამებს და პასუხობს კითხვებს წარდგენილი PDF დოკუმენტების შესახებ.
- Elicit: ხელოვნური ინტელექტის კვლევის ასისტენტი: მიზნად ისახავს მკვლევარების სამუშაო პროცესების ნაწილების ავტომატიზაციას, შესაბამისი ნაშრომების იდენტიფიცირებას და ძირითადი ინფორმაციის შეჯამებას.
- Perplexity: უზრუნველყოფს „ცოდნის ცენტრს“ მათთვის, ვინც ეძებს სწრაფ, ზუსტ პასუხებს, მათ საჭიროებებზე მორგებულს.

***ChatPDF, Elicit, Perplexity**-ის გარდა, დამატებითი მნიშვნელოვანი ინსტრუმენტებია:*

- ***NotebookLM (Google):** ატვირთავ საკუთარ დოკუმენტებს, სტატიებს, კონსპექტებს – AI „ხდება“ ამ მასალის ექსპერტი და პასუხობს კითხვებს მხოლოდ შენი მასალის საფუძველზე. განათლებაში განსაკუთრებით სასარგებლოა.*
- ***Consensus:** სამეცნიერო ნაშრომების საძიებო სისტემა AI-ით – კითხვაზე პასუხობს peer-reviewed კვლევების საფუძველზე.*
- ***Gamma:** პრეზენტაციების ავტომატური გენერაცია ტექსტიდან.*
- ***Otter.ai:** ზეპირი საუბრის ტრანსკრიფცია და შეჯამება – კონფერენციებისა და ლექციებისთვის.*

ანალოგიურად, LLM-ზე დაფუძნებული ინსტრუმენტები ჩაშენებულია სხვა პროდუქტებში, როგორიცაა ვებ ბრაუზერები. მაგალითად, Chrome ბრაუზერის გაფართოებები, რომლებიც ChatGPT-ზეა აგებული, მოიცავს შემდეგს:

- WebChatGPT: აძლევს ChatGPT-ს ინტერნეტზე წვდომას უფრო ზუსტი და განახლებული საუბრების ჩასატარებლად.
- Compose AI: ავტომატურად ავსებს წინადადებებს ელ.ფოსტაში და სხვაგან.
- TeamSmart AI: უზრუნველყოფს „ვირტუალური ასისტენტების“ გუნდს.
- Wiseone: ამარტივებს ონლაინ ინფორმაციას.

***WebChatGPT, Compose AI, TeamSmart, Wiseone**-ის გარდა ასევე ახალია:*

- ***Merlin:** ნებისმიერ ვებგვერდზე, YouTube ვიდეოზე ან PDF-ზე – ერთი ღილაკით AI-ს დახმარება; ახდენს YouTube ვიდეოების შეჯამებას.*
- ***ChatGPT for Google:** საძიებო სისტემის შედეგებთან ერთად GPT-ის პასუხს გვიჩვენებს.*
- ***Grammarly:** ტექსტის გრამატიკული და სტილისტური კორექტირება AI-ით ნებისმიერ ტექსტის ველში.*
- ***Summarize:** ნებისმიერი გვერდის ავტომატური შეჯამება ერთ აბზაცში.*

გარდა ამისა, ChatGPT ინტეგრირებულია ზოგიერთ საძიებო სისტემაში, და დანერგილია პროდუქტიულობის ინსტრუმენტების დიდ პორტფელში (მაგ. Microsoft Word და Excel), რაც მას კიდევ უფრო ხელმისაწვდომს ხდის ოფისებსა და საგანმანათლებლო დაწესებულებებში მთელ მსოფლიოში.

როგორ მუშაობს გამოსახულებითი გენერირებადი AI?

გამოსახულების GAN-ები და მუსიკალური GAN-ები, როგორც წესი, იყენებენ სხვადასხვა ტიპის ANN-ს, რომელიც ცნობილია როგორც გენერირებადი კონკურენტული ქსელები (GAN), რომლებიც ასევე შეიძლება გაერთიანდეს ვარიაციულ ავტოენკოდერებთან.

GAN-ები შედგება ორი ნაწილისგან (ორი „კონკურენტული“ ქსელი): გენერატორისა და დისკრიმინატორისგან. გამოსახულების GAN-ების შემთხვევაში, გენერატორი წარმოქმნის შემთხვევით გამოსახულებას მოთხოვნის საპასუხოდ და დისკრიმინატორი ცდილობს განასხვავოს ეს გენერირებული გამოსახულება რეალური სურათებისგან. შემდეგ გენერატორი იყენებს დისკრიმინატორის შედეგებს მისი პარამეტრების შესაცვლელად, რათა წარმოქმნას განსხვავებული გამოსახულება. პროცესი მეორდება, შესაძლოა ათასობითჯერ, გენერატორი წარმოქმნის სულ უფრო რეალისტურ სურათებს, რომელთა გარჩევა დისკრიმინატორისთვის სულ უფრო რთული ხდება რეალურისგან.

მაგალითად, წარმატებული GAN, რომელიც გაწვრთნილია ათასობით ლანდშაფტის ფოტოს მონაცემთა ნაკრებზე, შეუძლია წარმოქმნას ახალი, მაგრამ არარეალური, ლანდშაფტების სურათები, რომლებიც პრაქტიკულად არ განსხვავდება რეალური ფოტოებისგან. ამავე დროს, პოპულარული მუსიკის (ან თუნდაც ერთი არტისტის მუსიკის) მონაცემთა ნაკრებზე გაწვრთნილ GAN-ს შეუძლია წარმოქმნას ახალი მუსიკალური ნაწარმოებები, რომლებიც მიჰყვება ორიგინალური მუსიკის სტრუქტურას და სირთულეს.

2023 წლის ივლისის მდგომარეობით, ხელმისაწვდომი იყო შემდეგი Image GenAI მოდელები, რომელთაგან თითოეული ტექსტური მინიშნებებიდან სურათებს წარმოქმნის. მათი უმეტესობა უფასოა, გარკვეული შეზღუდვებით:

- DALL E 2: GenAI-ის გამოსახულების დამუშავების ინსტრუმენტი OpenAI-სთვის.
- DreamStudio: GenAI-ის გამოსახულების დამუშავების ინსტრუმენტი სტაბილური დიფუზიისთვის.
- Fotor: აერთიანებს GenAI-ს გამოსახულების რედაქტირების რიგ ინსტრუმენტებში.
- Midjourney: GenAI-ის დამოუკიდებელი გამოსახულების დამუშავების ინსტრუმენტი.
- NightCafe: ინტერფეისი სტაბილური დიფუზიისთვის და DALL E 2.
- Photosonic: WriteSonic-ის გამოსახულების გენერატორი.

2025 წლისათვის უფრო მეტი რესურსებია ხელმისაწვდომი:

- **DALL-E 3 (OpenAI, 2023):** DALL-E 2-ის მემკვიდრე – ბევრად უფრო ზუსტია ტექსტური მინიშნებების გაგებაში; ჩაშენებულია ChatGPT Plus-ში.
- **Adobe Firefly:** Adobe-ის გენერირებადი AI – ინტეგრირებულია Photoshop-სა და Illustrator-ში; „კომერციულად უსაფრთხო" (გაწვრთნილია ლიცენზირებულ კონტენტზე).
- **Canva Magic Studio:** Canva-ს AI ფუნქციები – ტექსტიდან სურათი, ფონის წაშლა, სტილის შეცვლა.
- **Ideogram:** განსაკუთრებით ძლიერია სურათებში ტექსტის ზუსტი გენერაციაში – Midjourney-ის სისუსტე.
- **Stable Diffusion (ღია კოდი):** ყველაზე პოპულარული ღია კოდის სურათის გენერატორი – ადგილობრივად გაშვება შესაძლებელია.

ადვილად ხელმისაწვდომი GenAI ვიდეოების მაგალითებია:

- Elai: შეუძლია პრეზენტაციების, ვებსაიტების და ტექსტის ვიდეოდ გადაყვანა.
- GliaCloud: შეუძლია ვიდეოების გენერირება სიახლეებიდან, სოციალური მედიის პოსტებიდან, პირდაპირ სპორტიდან და სტატისტიკიდან.
- Pictory: შეუძლია მოკლე ვიდეოების ავტომატურად შექმნა გრძელი კონტენტიდან.
- Runway: გთავაზობთ ვიდეოების (და) სურათების შექმნისა და რედაქტირების მრავალფეროვან ინსტრუმენტებს.
- **Sora (OpenAI, 2024):** ყველაზე შთამბეჭდავი – ტექსტიდან მაღალხარისხიანი, 1 წუთამდე ვიდეო. ფიზიკის კანონებს „ესმის". ჯერ შეზღუდული ხელმისაწვდომობა.
- **Kling (Kuaishou, ჩინეთი):** Sora-ს კონკურენტი – 2 წუთამდე, მაღალი ხარისხის ვიდეო.
- **HeyGen:** AI-ს ავატარი, რომელიც ნებისმიერ ენაზე „ლაპარაკობს" – განათლებაში გამოიყენება პერსონალიზებული ვიდეო-ლექციებისთვის.
- **Synthesia:** კომპანიების მიერ ფართოდ გამოყენებული კორპორატიული ვიდეოების გენერაციისთვის – AI პრეზენტატორით.

ადვილად ხელმისაწვდომი მუსიკალური გენერაციის ხელოვნური ინტელექტის რამდენიმე მაგალითი:

- Aiva: შეუძლია პერსონალიზებული საუნდტრეკების ავტომატურად შექმნა
- Boomy, Soundraw, და Voicemod: შეუძლიათ სიმღერების გენერირება ნებისმიერი ტექსტიდან.
- **Suno AI (2024):** ყველაზე პოპულარული – ტექსტიდან სრული სიმღერა ხმით, ინსტრუმენტებითა და ლირიკით. უფასო ვერსია ხელმისაწვდომია.
- **Udio:** Suno-ს მთავარი კონკურენტი – განსაკუთრებით ძლიერია ვოკა-

ლის ხარისხში.

- **ElevenLabs:** ხმის კლონირება და სინთეზი – ნებისმიერი ადამიანის ხმის „კლონი" ნიმუშებიდან.
- **Mureka:** განათლებაში გამოიყენება საგანმანათლებლო საუნდტრეკებისა და jingle-ების შესაქმნელად.

სასურველი შედეგების გენერირებისთვის სწრაფი ინჟინერია

მიუხედავად იმისა, რომ GenAI-ის გამოყენება შეიძლება ისეთივე მარტივი იყოს, როგორც შეკითხვის კომპიუტერში აკრეფა, სინამდვილეში მომხმარებლისთვის ჯერ კიდევ არ არის მარტივი ზუსტად სასურველი შედეგის მიღება. GenAI-სთვის ეფექტური შეკითხვების წერის გამოწვევამ განაპირობა ახალი სპეციალობის – „სწრაფი ინჟინერიის" გაჩენა. „სწრაფი ინჟინერია" ეხება პროცესებსა და ტექნიკას მონაცემების და შეკითხვების შეყვანის შესაქმნელად, რათა წარმოიქმნას GenAI შედეგი, რომელიც ნამდვილად ჰგავს მომხმარებლის სასურველ განზრახვას. „სწრაფი ინჟინერია" ყველაზე წარმატებულია, როდესაც შეკითხვა ლოგიკური თანმიმდევრობით აყალიბებს მსჯელობის თანმიმდევრულ ჯაჭვს კონკრეტულ პრობლემაზე და არის აზროვნების ჯაჭვზე ორიენტირებული. კონკრეტული რეკომენდაციებია:

- გამოიყენეთ მარტივი, გასაგები და პირდაპირი ენა, რომელიც ადვილად გასაგებია, თავი აარიდეთ რთულ ან ორაზროვან ფორმულირებას.
- ჩართეთ მაგალითები გენერირებული სასურველი დასრულების ან ფორმატის საილუსტრაციოდ.
- ჩართეთ კონტექსტი, რომელიც გადამწყვეტია შესაბამისი და შინაარსიანი დასრულებების გენერირებისთვის.
- საჭიროებისამებრ დახვეწეთ და გაიმეორეთ, ექსპერიმენტი ჩაატარეთ სხვადასხვა ვარიაციებით.
- იყავით ეთიკური, მოერიდეთ მინიშნებებს, რომლებმაც შეიძლება წარმოშვან შეუსაბამო, მიკერძოებული ან მავნე კონტენტი.

ასევე მნიშვნელოვანია დაუყოვნებლივ აღიაროს, რომ GenAI შედეგებზე დაყრდნობა შეუძლებელია კრიტიკული შეფასების გარეშე. მაგალითად:

- მათემატიკური ამოცანების გადასაჭრელად, „სიზუსტე" შეიძლება გამოყენებულ იქნას, როგორც მთავარი მეტრიკა, იმის დასადგენად, თუ რამდენად ხშირად იძლევა GenAI ინსტრუმენტი „სწორ პასუხს"
- სენსიტიურ კითხვებზე პასუხის გასაცემად, მთავარი მეტრიკა შესრულების გასაზომად შეიძლება იყოს „პასუხის სიხშირე" (სიხშირე, რომლითაც GenAI პირდაპირ პასუხობს კითხვას);
- კოდის გენერირებისთვის, მეტრიკა შეიძლება იყოს „გენერირებული კოდების ის ნაწილი, რომლებიც პირდაპირ შესრულებადია" (შეიძლება თუ არა გენერირებული კოდის პირდაპირ შესრულება პროგრამირების

გარემოში და ერთეული ტესტების გავლა);

- ვიზუალური მსჯელობისთვის, მეტრიკა შეიძლება იყოს „ზუსტი შესაბამისობა“ (ზუსტად ემთხვევა თუ არა გენერირებული ვიზუალური ობიექტები ძირითად სიმართლეს).

EdGPT-ის ფორმირება და მისი გავლენა

იმის გათვალისწინებით, რომ GenAI მოდელები შეიძლება გამოყენებულ იქნას როგორც საფუძველი ან საწყისი წერტილი უფრო სპეციალიზებული ან დარგის სპეციფიკური მოდელების შემუშავებისთვის, ზოგიერთმა მკვლევარმა შემოგვთავაზა GPT-ის „საბაზისო მოდელებად“ გადარქმევა. განათლების სფეროში, დეველოპერებმა და მკვლევარებმა დაიწყეს საბაზისო მოდელის დახვეწა „EdGPT“-ის შესაქმნელად. EdGPT მოდელები გამოიყენება კონკრეტულ მონაცემებზე საგანმანათლებლო მიზნებით. სხვა სიტყვებით რომ ვთქვათ, EdGPT მიზნად ისახავს გააუმჯობესოს ზოგადი სასწავლო მონაცემების დიდი რაოდენობით შესწავლილი მოდელი მაღალი ხარისხის, დარგის სპეციფიკური საგანმანათლებლო მონაცემების მცირე ნაკრებების გამოყენებით.

ამჟამად, GPT-ის განათლებაში უფრო მიზანმიმართული გამოყენებისთვის საბაზისო მოდელების დახვეწა ადრეულ ეტაპზეა. არსებულ მაგალითებს შორისაა EduChat, საბაზისო მოდელი, რომელიც შემუშავებულია აღმოსავლეთ ჩინეთის ნორმალური უნივერსიტეტის მიერ რეპეტიტორობის სერვისების უზრუნველსაყოფად და რომლის კოდები, მონაცემები და პარამეტრები ღიაა. კიდევ ერთი მაგალითია MathGPT, რომელსაც ავითარებს TAL Education Group, (LLM) პროგრამა, რომელიც ფოკუსირებულია მათემატიკური ამოცანების გადაჭრასა და მთელი მსოფლიოს მომხმარებლებისთვის ლექციების ჩატარებაზე.

თუმცა, მნიშვნელოვანი პროგრესის მიღწევამდე, საჭიროა ძალისხმევა ძირითადი მოდელების გასაუმჯობესებლად არა მხოლოდ საგნის შესახებ ცოდნის დამატებით

და მიკერძოების აღმოფხვრით, არამედ ცოდნის დამატებით შესაბამისი სწავლების მეთოდების შესახებ და იმის შესახებ, თუ როგორ შეიძლება ეს აისახოს ალგორითმების და მოდელების დიზაინში.

გამოწვევაა იმის დადგენა, თუ რამდენად შეუძლიათ EdGPT მოდელებს გასცდეს საგნის შესახებ ცოდნას, ასევე ფოკუსირდნენ მოსწავლეზე/სტუდენტზე ორიენტირებულ პედაგოგიკასა და დადებით მასწავლებელ-სტუდენტურ ურთიერთქმედებაზე. კიდევ ერთი გამოწვევაა იმის დადგენა, თუ რამდენად შეიძლება მოსწავლეები/სტუდენტების და მასწავლებლების მონაცემების ეთიკურად შეგროვება და გამოყენება EdGPT-ის ინფორმირებისთვის. და ბოლოს, ასევე საჭიროა საფუძვლიანი კვლევა იმის უზრუნველსაყოფად, რომ EdGPT არ არღვევს მოსწავლეების/სტუდენტების ადამიანის უფლებებს ან არ ზღუდავს მასწავლებლების შესაძლებლობებს.

EduChat და MathGPT-ის გარდა, სხვა მნიშვნელოვანი EdGPT მაგალითებია:

მოქმედი სისტემები:

- **Khanmigo (Khan Academy):** GPT-4-ზე დაფუძნებული სოკრატული ტუტორი – პასუხს პირდაპირ არ იძლევა, სტუდენტს კითხვებით მიჰყავს გადაჭრამდე. გამოიყენება K-12-ში ამერიკაში.
- **Duolingo Max:** GPT-4-ზე დაფუძნებული ფუნქციები – „Explain My Answer" (შეცდომის ახსნა) და „Roleplay" (სიმულირებული საუბრები).
- **Socratic (Google):** სასკოლო ასაკის მოსწავლეებისთვის – ფოტოდან ამოცნობს დავალებას, ნაბიჯ-ნაბიჯ ახსნით პასუხობს.
- **Coursera Coach:** Coursera-ს AI ტუტორი – კურსის მასალაზე კითხვებს პასუხობს, ახსნებს ცნებებს.

განვითარების სტადიაშია:

- **Squirrel AI (TAL Education, ჩინეთი):** 10,000+ „ცოდნის ატომად" დაყოფილი მასალა – 2 მილიონ+ მოსწავლეზე გამოყენებული. კვლევებმა აჩვენა, რომ 2-5-ჯერ ეფექტურია ჩვეულებრივ სწავლებაზე.
- **Carnegie Learning MATHia:** ITS + LLM კომბინაცია – ამერიკულ სკოლებში ათასობით მოსწავლე იყენებს.

ქართული კონტექსტი: საქართველოში EdGPT-ის კონცეფცია ჯერ ადრეულ ეტაპზეა, თუმცა სკოლები და უნივერსიტეტები სულ უფრო აქტიურად იყენებენ ChatGPT-ს, Copilot-სა და Claude-ს სასწავლო პროცესში – ძირითადად არა სპეციალიზებული EdGPT-ის, არამედ ზოგადი GenAI-ის სახით.

YouTube რესურსები

1. **"How ChatGPT Works Technically"** – YouTube: Andrej Karpathy (ex-OpenAI) *GPT-ის, ტოკენიზაციისა და ტრანსფორმატორის ტექნიკური, მაგრამ გასაგები ახსნა – ყველაზე ავტორიტეტული წყარო*
2. **"But what is GPT? Visual intro to transformers"** – YouTube: 3Blue1Brown *ტრანსფორმატორის ვიზუალური ახსნა – ანიმაციებით, ფორმულების გარეშე*
3. **"How AI Image Generation Works (GAN vs Diffusion)"** – YouTube: Computerphile *GAN-ებისა და Diffusion მოდელების (Stable Diffusion, DALL-E) ახსნა – ნახატებით*
4. **"Prompt Engineering Full Course"** – YouTube: freeCodeCamp *სწრაფი ინჟინერიის სრული გზამკვლევი – პრაქტიკული მაგალითებით ChatGPT-სთვის*
5. **"The Future of AI in Education (EdGPT)"** – YouTube: Sal Khan (Khan Academy) *Khanmigo-ს შემქმნელი განმარტავს AI-ის პოტენციალს განათლებაში – TED Talk ფორმატი*

⚠ *სათაურები პირდაპირ YouTube-ზე მოძებნეთ.*

სადისკუსიო თემები

„სწრაფი ინჟინერია“ – დროებითი უნარია?
თუ GPT-5 და მომდევნო თაობა სულ უფრო „ჭკვიანდება“ – გახდება სწრაფი ინჟინერია მოძველებული?
განიხილეთ: დღეს „სწრაფი ინჟინერია“ ახალ პროფესიად ყალიბდება – კომპანიები ათასობით დოლარს იხდიან კარგი Prompt-ების შესაქმნელად. მაგრამ GPT-5 უკვე ბევრად უკეთ „ესმის“ ბუნებრივ, „ნედლ“ შეკითხვებს.
10 წელიწადში – კვლავ საჭირო იქნება ეს უნარი?
ან ის გახდება ისეთივე „ბუნებრივი“, როგორც Google-ში ძიება – რომელსაც „ოსტატობა“ არ სჭირდება?

EdGPT – ვინ უნდა გააკონტროლოს?
სპეციალიზებული საგანმანათლებლო AI-ის შექმნა და კონტროლი – სახელმწიფოს, კომპანიის თუ სკოლის პასუხისმგებლობაა?
განიხილეთ: MathGPT-ი ჩინურ კომპანიას (TAL) ეკუთვნის, Khanmigo – ამერიკულ არაკომერციულ ორგანიზაციას. EduChat – ჩინურ უნივერსიტეტს. ამ სისტემები სტუდენტების მონაცემებს გროვებენ.
ვინ „ფლობს“ ამ მონაცემებს?
შეიძლება კომპანია EdGPT-ს „მიმართულება“ მიეცეს გარკვეული ღირებულებების ან შეხედულებების სასარგებლოდ?
საჭიროა თუ არა „ეროვნული EdGPT“ – ქართულ ენასა და ქართულ სასწავლო კონტექსტზე მორგებული?

GenAI „ჰალუცინაციები“ vs. ადამიანის შეცდომები – ვინ უფრო საიმედოა?
ექიმი, მასწავლებელი ან იურისტი ასევე „ჰალუცინირებს“ – ეს AI-ს „ნაკლოვანება“ გამოარჩევს?
განიხილეთ: კვლევებმა აჩვენა, რომ ექიმები სავარაუდო დიაგნოზების ~20%-ში ცდებიან; GPT-4 – ~15-20%-ში.
მაგრამ AI-ს შეცდომები „კონფიდენტური“ და „დამაჯერებელი“ ჩანს – ადამიანი კი ხშირად „ეჭვობს“. ნიშნავს ეს, რომ AI-ი ადამიანს „ჯობია“?
ან ეჭვის არარსებობა სწორედ ყველაზე სახიფათო მახასიათებელია?
როგორ ვასწავლოთ სტუდენტებს AI-ის შედეგების კრიტიკული შეფასება?

ტესტი 10

1. „ტოკენი“ GPT-ის კონტექსტში არის:

ა) გადახდის ერთეული
ბ) ტექსტის ყველაზე პატარა ერთეული, რომელსაც GPT ამუშავებს
გ) ნეირონული ქსელის ფენა
დ) GPT-ის პარამეტრი

2. GPT-5 გამოვიდა:

ა) 2023 წელს
ბ) 2024 წლის დასაწყისში
გ) 2025 წელს
დ) ჯერ არ გამოსულა

3. GAN-ი (Generative Adversarial Network) შედგება:

ა) ერთი ნეირონული ქსელისგან
ბ) სამი კონკურენტული ქსელისგან
გ) გენერატორისა და დისკრიმინატორისგან
დ) ტრანსფორმატორისა და ავტოენკოდერისგან

4. „სწრაფი ინჟინერია“ (Prompt Engineering) ნიშნავს:

ა) AI-ის ტრენინგის პროცესს
ბ) ნეირონული ქსელის პროგრამირებას
გ) ეფექტური შეკითხვების შექმნის ტექნიკებს GenAI-სთვის სასურველი შედეგის მისაღებად
დ) GPT-ის პარამეტრების ოპტიმიზაციას

5. Google-ის Bard 2024 წელს გადაიქცა:

ა) Copilot-ად
ბ) LaMDA-დ
გ) Gemini-ად
დ) PaLM-ად

6. EdGPT განსხვავდება ზოგადი GPT-ისგან:

ა) სიჩქარით
ბ) ენით
გ) სპეციალიზებული საგანმანათლებლო მონაცემებით დახვეწილი (fine-tuned) საბაზისო მოდელია
დ) ფასით

7. GPT-ის „პარამეტრები“ შეიძლება განვიხილოთ, როგორც:

ა) ტრენინგ მონაცემები
ბ) „ღილაკები“, რომლებიც განსაზღვრავს მოდელის მოქმედებას
გ) ტოკენების ლიმიტი
დ) GPT-ის სიჩქარე

8. Khanmigo-ს სოკრატული მიდგომა ნიშნავს:

ა) პასუხის მყისიერ მიცემას
ბ) კითხვებით მოსწავლის გადაჭრამდე მიყვანას, პირდაპირი პასუხის ნაცვლად
გ) ტესტების ავტომატური გენერაციას
დ) ვიდეო-ლექციების ჩაწერას

9. DALL-E 3-ის მთავარი გაუმჯობესება DALL-E 2-თან შედარებით:

ა) ფასის შემცირება
ბ) ვიდეოს გენერაციის შესაძლებლობა
გ) ტექსტური მინიშნებების ბევრად უფრო ზუსტი გაგება
დ) ღია კოდი

10. „დამცავი ბარიერები“ GPT-ის პასუხის გენერაციაში ნიშნავს:

ა) ტოკენების ლიმიტს
ბ) ინტერნეტზე წვდომის შეზღუდვას
გ) ფილტრაციის სისტემას შეურაცხმყოფელი ან მავნე კონტენტის მოსაშლელად
დ) GPT-ის კონტექსტური ფანჯრის ზომას

უთანხმოებები გენერირებადი AI შესახებ

წინა ნაწილში განხილული იყო გენერირებადი ხელოვნური ინტელექტი (AI) და მისი ფუნქციონირების წესები. ეს ნაწილი იხილავს გენერირებადი ხელოვნური ინტელექტის სისტემასთან დაკავშირებულ კამათებს და ეთიკურ რისკებს და ხაზს უსვამს განათლებისთვის ზოგიერთ გავლენას.

ციფრული სიღარიბის გაუარესება

როგორც ადრე აღვნიშნეთ, გენერირებადი ხელოვნური ინტელექტი (AI) მოითხოვს მონაცემების დიდ რაოდენობას, მნიშვნელოვან გამოთვლით სიმძლავრეს და ხელოვნური ინტელექტის ჩარჩოებსა და ტრენინგის მეთოდებში მუდმივ ინოვაციებს. ასეთი რესურსები, ძირითადად, ხელმისაწვდომია მსხვილი საერთაშორისო ტექნოლოგიური კომპანიებისა და რამდენიმე ეკონომიკისთვის, ძირითადად შეერთებული შტატებისთვის, ჩინეთისთვის და, გარკვეულწილად, ევროპისთვის. ეს ნიშნავს, რომ გენერირებადი ხელოვნური ინტელექტის შექმნისა და მართვის შესაძლებლობა დიდწილად ჯერჯერობით მიუწვდომელია კომპანიებისა და ქვეყნების უმეტესობისთვის, განსაკუთრებით გლობალური სამხრეთის ქვეყნებისთვის.

„ციფრული სიღარიბე“ (Digital Poverty) ნიშნავს ციფრული ტექნოლოგიებზე, ინტერნეტზე, მოწყობილობებზე და ციფრულ კომპეტენციებზე ხელმისაწვდომობის არარსებობას ან მწირობას – რაც ადამიანებს, თემებს ან ქვეყნებს „ციფრულ ბნელ ხანაში“ ტოვებს, სანამ სამყარო სულ უფრო ციფრული ხდება.

ციფრული სიღარიბის სამი განზომილება:

1. **წვდომის სიღარიბე** – ინტერნეტი ან სმარტფონი/კომპიუტერი საერთოდ არ არის ხელმისაწვდომი (ფიზიკურად ან ფინანსურად).
2. **ხარისხის სიღარიბე** – ინტერნეტი არის, მაგრამ ძალიან ნელი ან ძვირი GenAI-ის გამოსაყენებლად.
3. **კომპეტენციის სიღარიბე** – ტექნოლოგია არის, მაგრამ ეფექტური გამოყენების უნარი არ არის.

რადგან მონაცემებზე წვდომა სულ უფრო მნიშვნელოვანი ხდება ეროვნული ეკონომიკური განვითარებისა და ინდივიდუალური ციფრული გაძლიერებისთვის, ის ქვეყნები და მოქალაქეები, რომლებსაც არ აქვთ სათანადო წვდომა მონაცემებზე ან არ შეუძლიათ ეფექტიანად მათი გამოყენება, „ინფორმაციული სიღარიბის“ მდგომარეობაში აღმოჩნდებიან. მსგავსი სიტუაცია სამართლიანია გამოთვლით სიმძლავრეზე წვდომის მხრივ. გენერირებადი ხელოვნური ინტელექტის სწრაფმა გავრცელებამ ტექნოლოგიურად განვითარებულ ქვეყნებსა და რეგიონებში ექსპონენციურად დააჩქარა მონაცემთა შექმნა და დამუშავება, ამავდროულად გაზარდა ხელოვნური ინტელექტის სიმდიდრის კონცენტრაცია გლობალურ

ჩრდილოეთში. ამან კიდევ უფრო „დაჩაგრა“ მონაცემებზე შეზღუდული წვდომის მქონე რეგიონები და გაზარდა მათი გრძელვადიანი შესაბამისობის რისკები GPT მოდელში განხორციელებულ სტანდარტებთან.

კონკრეტული მაგალითები:

- ***სუბ-საჰარის აფრიკა:*** *მოსახლეობის დაახლოებით 65% ინტერნეტის გარეშეა. ChatGPT-ი ქვეყნის საშუალო ყოველთვიური ხელფასის ეკვივალენტი ღირს ზოგიერთ ქვეყანაში.*
- ***სოფლად მცხოვრები, საქართველო:*** *ქალაქსა და სოფელს შორის ციფრული განახლება მნიშვნელოვნად განსხვავდება – სოფელში GenAI-ის ეფექტური გამოყენება ხშირად პრაქტიკულად შეუძლებელია.*
- ***ენობრივი ბარიერი:*** *ChatGPT ინგლისურად ბევრად უკეთ მუშაობს, ვიდრე ქართულად ან სვაჰილიზე – ეს ენობრივი ციფრული სიღარიბის ფორმაა.*
- ***სასკოლო გარემო:*** *ბავშვი, რომელსაც ბიბლიოთეკაში 30 წუთი ინტერნეტი ხელმისაწვდომია, ვერ შეძლებს GenAI-ის ისე გამოყენებას, როგორც ბავშვი, რომელსაც სახლში სწრაფი ინტერნეტი და ლეპტოპი აქვს.*

ამჟამინდელი ChatGPT მოდელები „სწავლობენ“ ონლაინ მომხმარებლის მონაცემებზე, რომლებიც ასახავს გლობალური ჩრდილოეთის ღირებულებებსა და ნორმებს. ეს მათ გამოუსადეგარს ხდის ადგილობრივად ორიენტირებული ხელოვნური ინტელექტის ალგორითმებისთვის იმ რეგიონებში, სადაც მონაცემებზე შეზღუდული წვდომაა, ძირითადად გლობალური სამხრეთის ბევრ ქვეყნებში, ასევე გლობალური ჩრდილოეთის ნაკლებად მდიდარ რეგიონებში.

დასკვნები განათლებისა და კვლევებისათვის:

- *მკვლევარებმა, პედაგოგებმა და მოსწავლეებმა კრიტიკულად უნდა შეისწავლონ გენერირებადი ხელოვნური ინტელექტის საგანმანათლებლო მოდელებში ჩადებული ღირებულებითი ვარაუდები, კულტურული სტანდარტები და სოციალური წეს-ჩვეულებები.*
- *პოლიტიკის შემქმნელებმა უნდა აღიარონ და მოაგვარონ მზარდი უთანასწორობა, რომელიც გამოწვეულია გენერირებადი ხელოვნური ინტელექტის მოდელების ტრენინგსა და ზედამხედველობაში არსებული მზარდი ხარვეზით.*

ეროვნული რეგულირების ადაპტაცია

გენერირებადი ხელოვნური ინტელექტის ძირითადი პროვაიდერები ხშირად გააკრიტიკეს იმის გამო, რომ მათი სისტემები არ ექვემდებარებიან მკაცრ დამოუკიდებელ აკადემიურ შემოწმებას ქვეშ. გენერირებადი ხელოვნური ინტელექტის ძირითადი ტექნოლოგიები, როგორც წესი, მკაცრად დაცულია, როგორც კორპორატიული ინტელექტუალური საკუთრება. ამასობაში, ბევრ კომპანიას, რომელიც იწყებს გენერირებადი ხელოვნური ინტელექტის გამოყ-

ენებას, სულ უფრო უჭირს საკუთარი სისტემების უსაფრთხოების შენარჩუნება. უფრო მეტიც, მიუხედავად თავად ხელოვნური ინტელექტის ინდუსტრიის მხრიდან რეგულირების მოწოდებებისა, ყველა ტიპის ხელოვნური ინტელექტის, მათ შორის გენერირებადი ხელოვნური ინტელექტის შექმნისა და გამოყენების შესახებ კანონმდებლობის შემუშავება ხშირად ჩამორჩება განვითარების სწრაფ ტემპს, რაც გასაკვირი არაა. ეს ნაწილობრივ ხსნის იმ გამოწვევებს, რომელთა წინაშეც დგას ეროვნული და ადგილობრივი ხელისუფლება სამართლებრივი და ეთიკური საკითხების გაგებისა და რეგულირების კუთხით.

მიუხედავად იმისა, რომ გენერაციულ ხელოვნურ ინტელექტს აქვს პოტენციალი, გაზარდოს ადამიანური შესაძლებლობები გარკვეულ ამოცანებში, გენერირებადი ხელოვნური ინტელექტის ხელშემწყობ კომპანიებზე გამჭვირვალე ზედამხედველობა შეზღუდულია.

გამჭვირვალე ზედამხედველობა შეზღუდულია – რატომ

სამი მთავარი მიზეზი:

მიზეზი 1 – კომერციული საიდუმლოება: GPT-ის „წონები“ (მილიარდობით პარამეტრი) OpenAI-ის, Anthropic-ისა და Google-ის ყველაზე ძვირფასი ინტელექტუალური საკუთრებაა. ამ „ინგრედიენტების“ გამჟღავნება კონკურენტებს საშუალებას მისცემდა, პირდაპირ დააკოპირონ – ამიტომ კომპანიები კატეგორიულად უარს ამბობენ გამჟღავნებაზე, ისევე, როგორც Coca-Cola-ს ფორმულა საიდუმლოა.

მიზეზი 2 – ტექნიკური სირთულე: GPT-4-ს ~1 ტრილიონი პარამეტრი აქვს. თუნდაც OpenAI-მ „გახსნას“ მოდელი – ინდეპენდენტური მკვლევრები ამ რაოდენობის მონაცემების „გაშლა-შემოწმება“ ვერ შეძლებენ ფიზიკურად. ეს ჰგავს „გამჭვირვალობის“ მოთხოვნას, ოკეანეში ყველა წვეთის შემოწმებით.

მიზეზი 3 – კანონმდებლობის ჩამორჩენა: ევროკავშირის AI Act (2024) პირველი მასშტაბური კანონია, მაგრამ ის **არარეგულირების პერიოდის შემდეგ** გამოვიდა. ამერიკაში ფედერალური AI კანონი ჯერ არ არსებობს. კანონები ტექნოლოგიას „ეწევა“, მაგრამ ვერ „ეწევა“.

ეს არარეგულირება ბევრ შეკითხვას აჩენს, განსაკუთრებით შიდა მონაცემებზე წვდომისა და გამოყენების კუთხით, მათ შორის ადგილობრივი ინსტიტუტებისა და ფიზიკური პირების შესახებ მონაცემებთან, ასევე ეროვნულ საზღვრებში გენერირებულ მონაცემებთან დაკავშირებით. საჭიროა შესაბამისი კანონმდებლობა, რათა ადგილობრივ თვითმმართველობის ორგანოებს საშუალება მიეცეთ გარკვეული კონტროლი მოიპოვონ გენერირებადი ხელოვნური ინტელექტის მზარდ გავლენაზე და უზრუნველყონ მისი, როგორც საზოგადოებრივი სიკეთის, მართვა.

დასკვნები განათლებისა და კვლევებისათვის:

- *მკვლევარებმა, პედაგოგებმა და მოსწავლეებმა უნდა იცოდნენ ეროვნული ინსტიტუტებისა და კერძო ორგანიზაციების საკუთრების, ფიზიკური პირების, ასევე გენერირებადი ხელოვნური ინტელექტის შიდა მომხმარებლების უფლებების დასაცავად ადეკვატური რეგულაციებისა და პოლიტიკის არარსებობის შესახებ.*
- *გარდა ამისა, ისინი მზად უნდა იყვნენ, უპასუხონ გენერირებადი ხელოვნური ინტელექტის დანერგვით გამოწვეულ საკანონმდებლო გამოწვევებს.*

კონტენტის გამოყენება მონაცემთა მფლობელის თანხმობის გარეშე

როგორც ადრე აღვნიშნეთ, გენერირებადი ხელოვნური ინტელექტის მოდელები აგებულია მონაცემების უზარმაზარ რაოდენობაზე დაყრდნობით (როგორიცაა ტექსტი, აუდიო, კოდი და სურათები), რომლებიც ხშირად ამოღებულია ინტერნეტიდან და, როგორც წესი, მფლობელის ნებართვის გარეშე.

რატომ არის პრობლემა

სამართლებრივი პრობლემა: ადამიანი, რომელმაც წიგნი, სტატია ან კოდი ინტერნეტში განათავსა, **საავტორო უფლებებს ინარჩუნებს** (ევროკავშირსა და ამერიკაში). OpenAI-მ ეს კონტენტი ნებართვის გარეშე „წაიკითხა“ და ტრენინგისთვის გამოიყენა. ეს ჰგავს სახელმძღვანელოს გადაბეჭდვას გამომცემლის ნებართვის გარეშე – თუმცა სასამართლოები ჯერ კიდევ განიხილავენ, ზუსტად ემთხვევა თუ არა ეს ანალოგია.

ეთიკური პრობლემა: მხატვარი, რომელმაც 20 წელი ხატვა ისწავლა, მისი სტილის „სწავლებისთვის“ გამოიყენეს – და ახლა AI მისი სტილის „იმიტაციას“ აკეთებს. მხატვარი კი ვერ შოულობს – ეს „ციფრული შრომის ექსპლუატაციად“ ითვლება.

რეალური სასამართლო პროცესები:

- ***Getty Images vs. Stability AI (2023):*** *გეტი ამბობს, Stable Diffusion-მა 12 მილიონ+ ფოტო გამოიყენა ნებართვის გარეშე.*
- ***New York Times vs. OpenAI (2023):*** *NYT ამბობს, ChatGPT-ი NYT-ის სტატიებს სიტყვა-სიტყვით „ახსენებს» – ეს პირდაპირი კოპირებაა.*
- ***ავტორთა გაერთიანება vs. OpenAI:*** *ათასობით მწერლის ნამუშევარი გამოყენებული ტრენინგისთვის.*

შედეგად, მრავალი ვიზუალური და კოდზე დაფუძნებული გენერირებადი ხელოვნური ინტელექტის სისტემა დაადანაშაულეს ინტელექტუალური საკუთრების დარღვევაში. ამ საკითხთან დაკავშირებით, მიმდინარეობს რამდენიმე საერთაშორისო სასამართლო პროცესი.

გარდა ამისა, ბევრმა მკვლევარმა აღნიშნა, რომ GPT-ები შეიძლება ეწინააღმდეგებოდეს ისეთ კანონებს, როგორიცაა ევროკავშირის მონაცემთა დაცვის ზოგადი რეგულაცია (GDPR, 2016), განსაკუთრებით ფიზიკური პირების უფლებების დაცვის კონტექსტში, რადგან ამჟამად შეუძლებელია ვინმეს მონაცემების (ან ამ მონაცემების დამუშავების შედეგების) ამოღება GPT მოდელიდან მას შემდეგ, რაც ის GPT ტრენინგში იქნება შეტანილი. ის აღარ ქრება.

დასკვნები განათლებისა და კვლევებისათვის:

- *მკვლევარებმა, მასწავლებლებმა და სტუდენტებმა უნდა იცოდნენ მონაცემთა მფლობელების უფლებები და ჩაატარონ სათანადო შემოწმება იმის უზრუნველსაყოფად, რომ მათ მიერ გამოყენებული გენერირებადი ხელოვნური ინტელექტის ინსტრუმენტები შეესაბამებოდეს არსებულ რეგულაციებსა და პოლიტიკას.*
- *გარდა ამისა, მათ უნდა იცოდნენ, რომ გენერირებადი ხელოვნური ინტელექტის გამოყენებით შექმნილმა სურათებმა ან კოდმა შეიძლება დაარღვიოს სხვისი ინტელექტუალური საკუთრების უფლებები და უნდა იცოდნენ, რომ მათ მიერ შექმნილი და ონლაინ განთავსებული სურათები, ხმები ან კოდი შეიძლება გამოყენებულ იქნას სხვა გენერირებადი ხელოვნური ინტელექტის სისტემების მიერ.*

შედეგების „აუხსნელობა“ გენერირების პროცესში

ხელოვნური ნეირონული ქსელები (ANN) როგორც წესი, „შავი ყუთებია“. ამდენად მათი შიდა მუშაობა დახურულია შემოწმებისთვის. შედეგად, ANN არ არის „გამჭვირვალე“ ან „ახსნადი“ და შეუძლებელია იმის დადგენა, თუ როგორ იქნა მიღებული შედეგები.

მიუხედავად იმისა, რომ ANN საერთო მიდგომა, მასში გამოყენებული ალგორითმების ჩათვლით, ზოგადად ახსნადია, კონკრეტული მოდელები და მათი პარამეტრები, მათ შორის მოდელის შიგნით გამოყენებული წონები, არ ექვემდებარება შემოწმებას, ამიტომ გენერირებული კონკრეტული შედეგის ახსნა შეუძლებელია. GPT-4-ის მსგავს მოდელში, მილიარდობით პარამეტრია/წონაა, რაც ერთად შეიცავს შესწავლილ მონაცემებს, რომლებსაც მოდელი იყენებს თავისი შედეგების გენერირებისთვის.

სამედიცინო დიაგნოსტიკის მაგალითი:
2020 წელს Google-ის AI სისტემამ ბადურის კანის სნეულება აღმოაჩინა 94% სიზუსტით – ექიმებზე უკეთ. მაგრამ: **ვერ ახსნა, რატომ.**
FDA (ამერიკის სამედიცინო რეგულატორი) დაეკითხა – „თუ AI ამბობს, რომ პაციენტს სიმსივნე აქვს, ექიმი რას ეტყვის პაციენტს?“ AI-ი ვერ პასუხობს „რადგან...“ – ის უბრალოდ „ასკვნის“.

განათლების მაგალითი:
*წარმოიდგინეთ, AI სისტემა ამბობს: „ეს სტუდენტი ვერ ჩააბარებს გამოცდას.“ მაგრამ ვერ ხსნის – **რატომ?** შეიძლება სისტემა ასკვნიდეს სახელიდან, სოციალური ფენიდან ან კულტურული წარმოშობიდან – და ეს **სისტემური დისკრიმინაცია** იქნება, რომელიც „შავი ყუთის“ გამო გამოუვლენელი რჩება.*

იურიდიული მაგალითი – COMPAS:
*ამერიკაში სასამართლოებმა „COMPAS“ AI სისტემა გამოიყენეს სასჯელის ოდენობის განსაზღვრისთვის. გამოვლინდა, რომ სისტემა შავკანიანი ბრალდებულებისთვის მაღალ „რეციდივის ალბათობას“ ანიჭებდა – მაგრამ ვერ ხსნიდა, **რატომ**. „შავი ყუთი“ ადამიანის ბედს წყვეტდა.*

ხელოვნურ ნეირონულ ქსელებში პარამეტრების/წონების გაუმჭვირვალობის გამო, შეუძლებელია ზუსტი ახსნა-განმარტებების მიცემა ამ მოდელების გამოყენებით მიღებული შედეგად მიღებულ შედეგებზე. გენერირებად ხელოვნურ ინტელექტში გამჭვირვალობისა და ახსნის ნაკლებობა სულ უფრო პრობლემური ხდება, რადგან გენერირებადი ხელოვნური ინტელექტი უფრო რთული ხდება, რაც ხშირად მოულოდნელ ან არასასურველ შედეგებს იწვევს.

გარდა ამისა, გენერირებადი ხელოვნური ინტელექტის მოდელები გადმოცემენ და აძლიერებენ სასწავლო მონაცემებში არსებულ მიკერძოებებს, რომელთა აღმოჩენა და გამოსწორება, მოდელების გაუმჭვირვალე ბუნების გათვალისწინებით, რთულია. და ბოლოს, ეს გაუმჭვირვალობა ასევე წარმოადგენს გენერირებადი ხელოვნური ინტელექტისადმი ნდობის პრობლემების ერთ-ერთ მთავარ მიზეზს. თუ მომხმარებლებს არ ესმით, თუ როგორ მიაღწია გენერაციულმა ხელოვნური ინტელექტის მოდელმა კონკრეტულ შედეგებს, ისინი ნაკლებად სავარაუდოა, რომ მიიღებენ ან გამოიყენონ მათ.

დასკვნები განათლებისა და კვლევებისათვის:

- *მკვლევარებმა, მასწავლებლებმა და მოსწავლეებმა უნდა გაიაზრონ, რომ გენერირებადი ხელოვნური ინტელექტის სისტემები ფუნქციონირებენ როგორც „შავი ყუთები“ და, შესაბამისად, რთულია იმის გაგება, თუ რატომ არის მიღებული კონკრეტული შედეგი.*
- *შედეგების გენერირების ახსნის არარსებობა იწვევს იმას, რომ მომხმარებლები შეზღუდული არიან გენერირებად ხელოვნურ ინტელექტში შექმნილი პარამეტრებით განსაზღვრული ლოგიკით. ეს პარამეტრები შეიძლება ასახავდეს კონკრეტულ კულტურულ ან კომერციულ ღირებულებებსა და ნორმებს, რომლებიც ირიბად ამახინჯებენ შექმნილ შინაარსს.*

ონლაინ კონტენტის დაბინძურება

ვინაიდან GPT ტრენინგის მონაცემები, როგორც წესი, ინტერნეტიდან მოიპოვება, რომელიც ხშირად შეიცავს დისკრიმინაციულ და სხვა საეჭვო ენას, დეველოპერებმა უნდა დანერგონ ე.წ. „დამცავი ბარიერები“, რათა უზრუნველყონ, რომ GPT-ის მიერ გენერირებული შედეგები არ იყოს შეურაცხმყოფელი და/ან არაეთიკური.

თუმცა, მკაცრი რეგულაციებისა და ეფექტური მონიტორინგის მექანიზმების არარსებობის გამო, ხელოვნური ინტელექტის მიერ გენერირებული მიკერძოებული კონტენტი სულ უფრო მეტად ვრცელდება ინტერნეტში, რაც ზიანს აყენებს მსოფლიოს მასშტაბით მოსწავლეთა უმრავლესობის კონტენტისა და ცოდნის ძირითად წყაროს.

ეს განსაკუთრებით მნიშვნელოვანია, რადგან ხელოვნური ინტელექტის მიერ გენერირებული კონტენტი შეიძლება ძალიან ზუსტი და დამაჯერებელი ჩანდეს, მაგრამ ხშირად შეიცავდეს შეცდომებს და მიკერძოებულ იდეებს. ეს დიდ რისკს უქმნის მოსწავლეებს, რომლებსაც არ აქვთ თემის ღრმა წინასწარი ცოდნა. ეს ასევე მომავალ რისკს უქმნის შემდგომ GPT მოდელებს, რომლებიც ისწავლიან ინტერნეტიდან ამოღებულ და თავად GPT მოდელების მიერ გენერირებულ ტექსტზე, რომელიც შეიცავს მიკერძოებებსა და შეცდომებს.

დასკვნები განათლებისა და კვლევებისათვის:

- *მკვლევარებმა, მასწავლებლებმა და მოსწავლეებმა უნდა იცოდნენ, რომ გენერირებადმა ხელოვნურმა ინტელექტმა შეიძლება მოვცეს შეურაცხმყოფელი და არაეთიკური მასალა.*
- *მათ ასევე უნდა გააცნობიერონ ის გრძელვადიანი პრობლემები, რომლებიც შეიძლება წარმოიშვას ცოდნის სანდოობის თვალსაზრისით, როდესაც მომავალი GPT მოდელები ეფუძნება წინა GPT მოდელების მიერ გენერირებულ ტექსტს.*

რეალური სამყაროს არასაკმარისი გაგება

ტექსტზე დაფუძნებულ GPT-ებს ზოგჯერ უარყოფითად მოიხსენიებენ, როგორც „სტოქასტურ თუთიყუშებს“. ეს იმიტომ ხდება, რომ მიუხედავად იმისა, რომ მათ შეუძლიათ შექმნან ტექსტი, რომელიც შეიძლება დამაჯერებლად გამოიყურებოდეს, ტექსტი ხშირად შეიცავს შეცდომებს და მავნე მტკიცებებს. ეს ტენდენცია გამომდინარეობს იქიდან, რომ GPT მოდელები უბრალოდ იყენებენ და „ამრავლებენ“ მათ სასწავლო მონაცემებში (როგორც წესი, ინტერნეტიდან ამოღებული ტექსტი) იდენტიფიცირებულ ენობრივ, ხშირად შემთხვევითი, ნიმუშებს, მათი მნიშვნელობის გაგების არარსებობით. ეს მართლაც მსგავსია იმისა, თუ როგორ შეუძლია თუთიყუშს ბგერების იმიტაცია, მაგრამ არ ესმის მათი მნიშვნელობა.

ასეთმა შესაძლო შეუსაბამობამ შეიძლება გამოიწვიოს სიტუაცია, როდესაც

მასწავლებლები და მოსწავლეები ენდობიან ამ მოდელების მიერ წარმოებულ შედეგებს, მაგრამ ვერ უზრუნველყოფენ მათ ხარისხს. ეს სერიოზულ რისკებს უქმნის განათლების მომავალს. სინამდვილეში, გენერირებადი ხელოვნური ინტელექტი არ ეფუძნება რეალური სამყაროს დაკვირვებებს ან სამეცნიერო მეთოდის სხვა ძირითად ასპექტებს და არ შეესაბამება ადამიანურ ან სოციალურ ღირებულებებს. ამ მიზეზების გამო, მას არ შეუძლია ჭეშმარიტად ახალი კონტენტის გენერირება რეალური სამყაროს, ობიექტებისა და მათი ურთიერთობების, ადამიანების და სოციალური ურთიერთობების, ადამიანსა და ობიექტს შორის ურთიერთობების ან ადამიანსა და ტექნოლოგიას შორის ურთიერთობების შესახებ. სადავოა, შეიძლება თუ არა GenAI მოდელების მიერ გენერირებული, ერთი შეხედვით, ახალი კონტენტი, სამეცნიერო ცოდნად ჩაითვალოს თუ არა.

„სტოქასტური თუთიყუში“ – მაგალითები

არარსებული სამეცნიერო წყაროები: *პროფესორმა ბრაიან ჩიომ (კოლუმბიის უნივერსიტეტი) 2023 წელს სასამართლო საქმეში ChatGPT-ს ციტატები სთხოვა. ChatGPT-მა 6 „სამეცნიერო სტატია“ ჩამოთვალა – **ყველა გამოგონილი იყო.** ავტორები, ჟურნალები, წლები – სრული გამოგონება, მაგრამ სრულიად „დამაჯერებელი“ ფორმატით.*

მათემატიკური „ჰალუცინაცია“: *GPT-3-ს კითხვა: „რამდენია 17 × 24?“ ხშირ შემთხვევაში გამოთვლის ნაცვლად ხდება სტატისტიკურად ყველაზე სავარაუდო ციფრების გამეორება. AI „ფიქრობს“ მათემატიკაზე ისევე, როგორც ტექსტზე – „ეს ციფრი ამ ციფრს მოჰყვება“ – გამოთვლა არ ხდება.*

ისტორიული ფაქტების დამახინჯება: *ChatGPT-ს მოვთხოვეთ ქართველი ისტორიული პიროვნებების შესახებ ინფორმაცია – სისტემამ ზოგჯერ „დაამატა“ ბიოგრაფიული დეტალები, რომლებიც ფაქტობრივად არ არსებობდა, მაგრამ „ლოგიკური“ ჩანდა. ეს განსაკუთრებით სახიფათოა ნაკლებად ცნობილი ადამიანების შემთხვევაში, სადაც ტრენინგ მონაცემები მწირია.*

როგორც აღინიშნა, GPT-ებს ხშირად შეუძლიათ არაზუსტი ან არასანდო ტექსტის გენერირება. სინამდვილეში, კარგად არის ცნობილი, რომ GenAI ქმნის ზოგიერთ რამეს, რაც რეალურ ცხოვრებაში არ არსებობს. ამას ხშირად AI „ჰალუცინაციას“ უწოდებენ. ამას აღიარებენ GenAI-ის მწარმოებელი კომპანიები. მაგალითად, ChatGPT-ის საჯარო ინტერფეისის ბოლოში წერია: *„ChatGPT-მა შეიძლება წარმოქმნას არაზუსტი ინფორმაცია ადამიანების, ადგილების ან ფაქტების შესახებ“.*

AI „ჰალუცინაციების“ მაგალითები

იურიდიული სფერო: *ადვოკატმა სტივენ შვარცმა (ნიუ-იორკი, 2023) ChatGPT-ით მოამზადა სასამართლო საქმის მასალები. ChatGPT-მა 6 „პრეცედენტული საქმე“ ჩამოთვალა – **ყველა არარსებული.** სასამართლომ ადვოკატი დააჯარიმა **$5,000**-ით.*

სამედიცინო კონსულტაცია: *2023 წლის კვლევამ აჩვენა, რომ ChatGPT სამედიცინო კითხვებზე **~60-80% სიზუსტით** პასუხობს – ანუ 20-40% შემთხვევაში – **არასწორად.** ამის გაუცნობიერებელმა პაციენტმა შეიძლება მედიკამენტების არასწორი კომბინაცია მიიღოს.*

ენციკლოპედიური ინფორმაცია: *ChatGPT-ს სთხოვეს, ქართული ქალაქების შესახებ ინფორმაცია მიეწოდებინა. ზოგიერთ შემთხვევაში მოსახლეობის ციფრები, ისტორიული თარიღები ან გეოგრაფიული მახასიათებლები **ტელეფონის ნომრებივით „გამოგონილი“** იყო – ციფრები „სწორ ადგილას“ ჩანდა, მაგრამ ფაქტობრივად მცდარი იყო.*

ბევრი მკვლევარი ვარაუდობს, რომ GenAI წარმოადგენს მნიშვნელოვან ნაბიჯს ხელოვნური ზოგადი ინტელექტის (AGI)-კენ, ტერმინი, რომელიც აღნიშნავს ხელოვნური ინტელექტის კლასს, რომელიც უფრო ინტელექტუალურია, ვიდრე ადამიანები. თუმცა, ეს დიდი ხანია კრიტიკის საგანია, რადგან ამტკიცებენ, რომ ხელოვნური ინტელექტი ვერასდროს გახდება AGI ხელოვნური ინტელექტი, სულ მცირე მანამ, სანამ ის როგორღაც სიმბიოზურად არ გააერთიანებს ცოდნაზე დაფუძნებულ ხელოვნურ ინტელექტს და მონაცემებზე დაფუძნებულ ხელოვნურ ინტელექტს.

დასკვნები განათლებისა და კვლევებისათვის:

- *გენერირებადი ხელოვნური ინტელექტის მიერ გენერირებული ტექსტი შეიძლება ვიზუალურად ჰგავდეს ადამიანის ტექსტს. თუმცა, უნდა აღინიშნოს, რომ გენერირებად ხელოვნურ ინტელექტს არ გააჩნია გაგება. პირიქით, ის აერთიანებს სიტყვებს ინტერნეტში გავრცელებული ნიმუშების გამოყენებით. ამიტომ, გენერირებული ტექსტი შეიძლება შეიცავდეს შეცდომებს.*
- *მკვლევარებმა, მასწავლებლებმა და სტუდენტებმა უნდა გააცნობიერონ, რომ ხელოვნური ინტელექტის მიერ გენერირებული ტექსტი არ ესმის; რომ მას ხშირად შეუძლია არასწორი განცხადებების გაკეთება; და ამიტომ, მათ გენერირებადი ხელოვნური ინტელექტის მიერ წარმოებულ ყველაფერს კრიტიკული თვალით უნდა მიუდგნენ.*

უფრო რეალისტური ღრმა ფეიქების შექმნა

გენერირებადი ხელოვნური ინტელექტის ყველა ფუნდამენტური წინააღმდეგობის მიღმა, გენერირებად ხელოვნურ ინტელექტში GAN (გენერირებადი კონკურენტული ქსელები) ტექნოლოგია შეიძლება გამოყენებულ იქნას არსებული სურათების ან ვიდეოების შესაცვლელად ან მანიპულირებისთვის, რათა შეიქმნას ყალბი კონტენტი, რომლის გარჩევა რეალურისგან რთულია.

ეს ხელს უწყობს „ღრმა ფეიქების“ და ე.წ. „ყალბი ამბების“ შექმნას. სხვა სიტყვებით რომ ვთქვათ, გენერირებადი ხელოვნური ინტელექტი აადვილებს გარკვეული აქტორებისთვის არაეთიკურ, ამორალურ და კრიმინალურ საქმიანობაში ჩართვას, მათ შორის დეზინფორმაციის გავრცელებას, სიძულვილის ენის გამოყენებას და სრულიად ყალბი და ზოგჯერ კომპრომეტირებული კონტენტის შექმნას, რომელიც მოიცავს ადამიანებს მათი თანხმობის ან ცოდნის გარეშე.

„ღრმა ფეიქები“ – მაგალითები

პოლიტიკური მანიპულაცია: *2023 წელს ვირუსულად გავრცელდა ვიდეო, სადაც პრეზიდენტი ზელენსკი „უკრაინელ ჯარს კაპიტულაციას ურჩევდა.“ ვიდეო Deep Fake იყო – მაგრამ მილიონობით ადამიანმა ნახა, სანამ ამოიცნეს.*

ფინანსური თაღლითობა: *2024 წელს ჰონკონგის კომპანიის ფინანსურმა მოხელემ „ვიდეო-კონფერენციაში“ მონაწილეობა მიიღო, სადაც მისი „CFO და კოლეგები“ ყოფნის გადახდის მოთხოვნებს განიხილავდნენ. ყველა – Deep Fake იყო. $25 მილიონი გადარიცხა.*

საგანმანათლებლო გარემო: *სტუდენტების „ყალბი ვიდეოების“ შექმნა Deep Fake ტექნოლოგიით სულ უფრო ხელმისაწვდომი ხდება – ეს სექსუალური შევიწროების, მობინგისა და დისკრედიტაციის ახალ ფორმას ქმნის. 2023 წელს ამერიკის სკოლებში რამდენიმე შემთხვევა დაფიქსირდა.*

„კლონირებული“ ხმა: *ElevenLabs-ის ხმის კლონირების ტექნოლოგიით, 3-5 წუთის ოდიო ნიმუშიდან ნებისმიერი ადამიანის „ხმის კლონი“ შეიძლება შეიქმნას. ეს ტელეფონური თაღლითობებში გამოიყენება – „შვილი“ ან „ნათესავი“ „სასწრაფო დახმარებას ითხოვს.“*

დასკვნები განათლებისა და კვლევებისათვის:

მიუხედავად იმისა, რომ გენერირებადი ხელოვნური ინტელექტის პროვაიდერები ვალდებულნი არიან დაიცვან მომხმარებლების საავტორო უფლებები, მკვლევარებმა, პედაგოგებმა და სტუდენტებმა ასევე უნდა იცოდნენ, რომ ნებისმიერი სურათი, რომელსაც ისინი ონლაინ აზიარებენ, შეიძლება შედიოდეს გენერირებადი ხელოვნური ინტელექტის ტრენინგის მონაცემებში, მანიპულირებული იყოს და გამოყენებული იყოს არაეთიკური მიზნებისთვის.

YouTube რესურსები

1. **"The Problem with AI Hallucinations"** – YouTube: AI Explained *AI-ის „ჰალუცინაციების" – არარსებული ფაქტების – მიზეზები, სახეობები და საფრთხეები*
2. **"Deepfakes and the Misinformation Crisis"** – YouTube: Vox *Deep Fake ტექნოლოგიის ევოლუცია, პოლიტიკური და სოციალური საფრთხეები*
3. **"The Digital Divide and AI"** – YouTube: TED-Ed *ციფრული სიღარიბე, AI-ის კონცენტრაცია გლობალურ ჩრდილოეთში და გამოსავლები*
4. **"AI Copyright Crisis Explained"** – YouTube: Legal Eagle *Getty vs. Stability AI, NYT vs. OpenAI – საავტორო უფლებების სასამართლო პროცესები*
5. **"The Black Box Problem in AI"** – YouTube: Computerphile *„შავი ყუთის" პრობლემა ნეირონულ ქსელებში – COMPAS-ის, სამედიცინო AI-ის და GPT-ის მაგალითები*

⚠ *სათაურები პირდაპირ YouTube-ზე მოძებნეთ.*

სადისკუსიო თემები

AI „ჰალუცინაცია" – ვინ აგებს პასუხს?
თუ სტუდენტმა ChatGPT-ის „ჰალუცინაციური" ფაქტი ესეში გამოიყენა – ვინ არის პასუხისმგებელი: სტუდენტი, მასწავლებელი, OpenAI, სკოლა?
განიხილეთ: ამერიკელი ადვოკატი $5,000 დააჯარიმეს ChatGPT-ის არარსებული ციტატებისთვის – „ინსტრუმენტმა მოტყუა" არ მიიღეს საბაბად. სტუდენტი რომ ენციკლოპედიის არასწორ ინფორმაციას გამოიყენებდა – ვის დავაბრალებდით? AI-ი განსხვავდება ენციკლოპედიისგან? სად გადის ზღვარი AI-ის „ინსტრუმენტად" გამოყენებასა და „ავტორიტეტურ წყაროდ" მიჩნევას შორის?

„ღრმა ფეიქები“ – შეიძლება სიმართლის „სიკვდილი“?
თუ ნებისმიერი ვიდეო შეიძლება „ყალბი“ იყოს – რა ხდება მტკიცებულებებისა და ნდობის კულტურასთან?
განიხილეთ: ისტორიულად „ვიდეო = მტკიცებულება“ ითვლებოდა სასამართლოში, ჟურნალისტიკაში, ისტორიაში. Deep Fake ამ „კონტრაქტს“ არღვევს. შეიძლება ადამიანებმა ყველა ვიდეოს შეწყვიტონ რწმენა? ეს „ეპისტემოლოგიური კრიზისი“ – ცოდნის კრიზისი – შეიძლება დემოკრატიისა და განათლების საფუძვლებს შეარყევს? როგორ ვასწავლოთ „მედიაგრამოტობა“ Deep Fake-ის ეპოქაში?

ციფრული სიღარიბე და AI – ახალი „კოლონიალიზმი“?
თუ GenAI „სწავლობს“ გლობალური ჩრდილოეთის ღირებულებებზე და „ასწავლის“ მთელ სამყაროს – ეს „ციფრული კოლონიალიზმია“?
განიხილეთ: ისტორიაში კოლონიალიზმი ნიშნავდა, რომ ერთი კულტურის ღირებულებები, ენა და „ცოდნა“ სხვაზე „იმოქმედა“. ChatGPT ინგლისურ, ამერიკულ კულტურულ ნორმებს ასახავს – მაგრამ ქართველი, ნიგერიელი ან ბრაზილიელი სტუდენტი ამ სისტემებს იყენებს. ეს „კულტურული ნეიტრალობის“ ილუზიაა? შეიძლება ოდესმე „ქართული GenAI“ შეიქმნას, რომელიც ქართულ კულტურულ კონტექსტს ასახავს?

ტესტი 11

1. **„ციფრული სიღარიბე“ ნიშნავს:**
   ა) ფულადი სიღარიბის ციფრულ ფორმას
   ბ) ციფრულ ტექნოლოგიებზე, ინტერნეტსა და ციფრულ კომპეტენციებზე ხელმისაწვდომობის არარსებობას
   გ) სახელმწიფო ბიუჯეტის ციფრულ დეფიციტს
   დ) AI-ის უუნარობას ფულადი ოპერაციების ჩატარებაში

2. **GenAI ტრენინგ მონაცემებში „გლობალური ჩრდილოეთის“ მიკერძოება ნიშნავს:**
   ა) AI-ი ჩრდილოეთ პოლუსის შესახებ ბევრ ინფორმაციას შეიცავს
   ბ) მოდელი ძირითადად ასახავს განვითარებული ქვეყნების კულტურულ ღირებულებებსა და ნორმებს
   გ) AI-ი ცივ ამინდზე ინფორმაციაში სპეციალიზდება
   დ) ჩრდილოეთ ევროპა AI-ს ყველაზე მეტად იყენებს

3. AI-ის „შავი ყუთის“ პრობლემა ნიშნავს:

ა) AI-ს შავი ფერის ინტერფეისი აქვს
ბ) AI-ი კონფიდენციალურ ინფორმაციას მალავს
გ) შეუძლებელია AI-ის გადაწყვეტილების პროცესის ახსნა-განმარტება
დ) AI-ი მხოლოდ ღამით მუშაობს

4. „სტოქასტური თუთიყუში“ ტერმინი GPT-ებთან მიმართებაში ნიშნავს:

ა) AI-ს შეუძლია ფრინველების ბგერების სრული სიმულირება
ბ) AI ენობრივ ნიმუშებს ბაძავს მნიშვნელობის გაგების გარეშე
გ) AI-ი შემთხვევით ასრულებს ამოცანებს
დ) AI-ი ადამიანის ხმას კოპირებს

5. GDPR (ევროკავშირის მონაცემთა დაცვის რეგულაცია) GPT-ების კონტექსტში პრობლემურია, რადგან:

ა) GDPR მხოლოდ ევროპაში მოქმედებს
ბ) GDPR ძალიან ძველი კანონია
გ) ვინმეს მონაცემების ამოღება GPT-ის მოდელიდან პრაქტიკულად შეუძლებელია
დ) GDPR AI-ს სრულად კრძალავს

6. „ღრმა ფეიქი“ (Deep Fake) იქმნება შემდეგი ტექნოლოგიის გამოყენებით:

ა) GPT (Generative Pre-Trained Transformer)
ბ) LLM (Large Language Model)
გ) GAN (Generative Adversarial Network)
დ) NLP (Natural Language Processing)

7. GenAI-ს მიერ ინტერნეტ-კონტენტის ნებართვის გარეშე გამოყენება:

ა) სრულად ლეგალურია
ბ) ნებადართულია „Fair Use“ პრინციპით
გ) სადავოა და მიმდინარეობს სასამართლო პროცესები
დ) დაცულია ევროკავშირის AI Act-ით

8. არარსებული სამეცნიერო ციტატების გენერაცია ChatGPT-ის მიერ წარმოადგენს:

ა) ჰაკერულ შეტევას
ბ) მიზნობრივ ტყუილს
გ) „ჰალუცინაციას“ – AI-ის სტატისტიკურ შეცდომას სიმართლის „ილუზიით“
დ) მიკერძოებას

9. GenAI-ის ზედამხედველობის შეზღუდვის ერთ-ერთი მიზეზია:

ა) AI კომპანიები სამთავრობო საწარმოებია

ბ) AI-ი ინტერნეტის გარეშე მუშაობს

გ) ტექნოლოგია კომერციული საიდუმლოებაა და კანონმდებლობა ვერ ეწევა განვითარების ტემპს

დ) AI-ი მხოლოდ ჰუმანიტარული მიზნებით გამოიყენება

10. „ინფორმაციული სიღარიბე“ განსხვავდება „ციფრული სიღარიბისგან“ შემდეგი ნიშნით:

ა) ისინი სინონიმებია

ბ) „ინფორმაციული სიღარიბე“ ფოკუსირდება მონაცემებზე წვდომასა და მათი ეფექტური გამოყენების უუნარობაზე

გ) „ინფორმაციული სიღარიბე“ მხოლოდ განვითარებად ქვეყნებში არსებობს

დ) „ციფრული სიღარიბე“ AI-ს ეხება, „ინფორმაციული სიღარიბე“ – ინტერნეტს

განათლებაში ხელოვნური ინტელექტის გამოყენების რეგულაციები

გენერირებადი ხელოვნური ინტელექტის ირგვლივ არსებული დავების გადასაჭრელად და განათლებაში მისი პოტენციური სარგებლის გამოსაყენებლად, მისი რეგულირება აუცილებელია. გენერირებადი ხელოვნური ინტელექტის საგანმანათლებლო მიზნებისთვის რეგულირება მოითხოვს ადამიანზე ორიენტირებულ მიდგომაზე დაფუძნებულ ნაბიჯებსა და პოლიტიკას, რათა უზრუნველყოფილი იყოს მისი ეთიკური, უსაფრთხო, სამართლიანი და შინაარსიანი გამოყენება.

ადამიანზე ორიენტირებული მიდგომა ხელოვნური ინტელექტის მიმართ

ხელოვნური ინტელექტის ეთიკის შესახებ 2021 წლის რეკომენდაცია უზრუნველყოფს აუცილებელ ნორმატიულ ჩარჩოს გენერირებად ხელოვნურ ინტელექტთან დაკავშირებული მრავალი დავის გადასაჭრელად, მათ შორის განათლებასთან და კვლევასთან დაკავშირებული. ის ეფუძნება ადამიანზე ორიენტირებულ მიდგომას ხელოვნური ინტელექტის მიმართ, რომელიც ამტკიცებს, რომ ხელოვნური ინტელექტის გამოყენება უნდა ემსახურებოდეს ადამიანის შესაძლებლობების განვითარებას ინკლუზიური, სამართლიანი და მდგრადი მომავლისთვის. ეს მიდგომა უნდა ეფუძნებოდეს ადამიანის უფლებების პრინციპებს, ასევე ადამიანის ღირსებისა და კულტურული მრავალფეროვნების დაცვის აუცილებლობას, რომელიც განსაზღვრავს საყოველთაო ცოდნას. მმართველობის პერსპექტივიდან, ადამიანზე ორიენტირებული მიდგომა მოითხოვს შესაბამის რეგულირებას, რომელსაც შეუძლია უზრუნველყოს ადამიანის მონაწილეობა, გამჭვირვალობა და საზოგადოებრივი ანგარიშვალდებულება.

2019 წლის პეკინის კონსენსუსი ხელოვნური ინტელექტის (AI) და განათლების შესახებ დამატებით დეტალურად განმარტავს, თუ რას გულისხმობს ადამიანზე ორიენტირებული მიდგომა განათლების კონტექსტში ხელოვნური ინტელექტის გამოყენების მიმართ. კონსენსუსი ადასტურებს, რომ განათლებაში ხელოვნური ინტელექტის ტექნოლოგიების გამოყენებამ უნდა გააძლიეროს ადამიანური პოტენციალი მდგრადი განვითარებისა და ადამიან-მანქანის ეფექტური ურთიერთქმედებისთვის ცხოვრებაში, სწავლასა და სამუშაოში. იგი ასევე მოუწოდებს შემდგომი ქმედებებისკენ, რათა უზრუნველყოფილი იყოს ხელოვნურ ინტელექტზე თანაბარი წვდომა სხვადასხვა ჯგუფების მხარდასაჭერად და უთანასწორობის აღმოსაფხვრელად, ამავდროულად ენობრივი და კულტურული მრავალფეროვნების ხელშეწყობის მიზნით. კონსენსუსი გვთავა-

ზობს ყველა დაინტერესებული მხარის კოორდინირებულ მიდგომებს განათლებაში ხელოვნური ინტელექტის გამოყენების პოლიტიკის დაგეგმვისას. სახელმძღვანელო პოლიტიკის შემქმნელებისთვის (UNESCO, 2022b) დამატებით განმარტავს, თუ რას ნიშნავს ადამიანზე ორიენტირებული მიდგომა განათლებაში ხელოვნური ინტელექტის სარგებლისა და რისკების შესწავლისას, ასევე განათლების როლზე, როგორც ხელოვნური ინტელექტის კომპეტენციების განვითარების საშუალებაში. ის გვთავაზობს კონკრეტულ რეკომენდაციებს ხელოვნური ინტელექტის გამოყენების მარეგულირებელი პოლიტიკის შემუშავების შესახებ, რათა უზრუნველყოფილი იყოს (i) საგანმანათლებლო პროგრამებზე ინკლუზიური წვდომა, განსაკუთრებით დაუცველი ჯგუფებისთვის, როგორიცაა შეზღუდული შესაძლებლობის მქონე მოსწავლეები; (ii) მხარი დაუჭიროს პერსონალიზებულ და ღია სწავლების ვარიანტებს; (iii) გააუმჯობესოს მონაცემთა მიწოდება და მართვა სწავლის ხარისხის გასაუმჯობესებლად; (iv) აკონტროლოს სასწავლო პროცესები და გააფრთხილოს მასწავლებლები წარუმატებლობის რისკების შესახებ; და (v) განავითაროს გაგება და უნარები ხელოვნური ინტელექტის ეთიკური და შინაარსიანი გამოყენებისთვის.

ნაბიჯები განათლებაში გენერირებადი AI რეგულირებისთვის

ChatGPT-ის გამოშვებამდე, მთავრობები ავითარებდნენ ჩარჩოებს მონაცემთა შეგროვებისა და გამოყენების, ასევე ხელოვნური ინტელექტის სისტემების სხვადასხვა სექტორში, მათ შორის განათლებაში, განლაგების რეგულირებისთვის, რაც უზრუნველყოფდა საკანონმდებლო და პოლიტიკურ კონტექსტს ახალი ხელოვნური ინტელექტის აპლიკაციების რეგულირებისთვის. 2022 წლის ნოემბრიდან რამდენიმე კონკურენტი გენერირებული ხელოვნური ინტელექტის მოდელის გამოშვების შემდეგ, მთავრობები ახორციელებდნენ სხვადასხვა პოლიტიკურ რეაგირებას, დაწყებული გენერირებული ხელოვნური ინტელექტის აკრძალვიდან, არსებული ჩარჩოების ადაპტირების საჭიროების შეფასებით და ახალი რეგულაციების სასწრაფოდ შემუშავებით დამთავრებული.

GenAI-ის შემოქმედებითი გამოყენების რეგულირებისა და ხელშეწყობის სამთავრობო სტრატეგიები შედგენილი და განხილული იქნა 2023 წლის აპრილში (UNESCO, 2023b) და გვთავაზობს ექვსი ნაბიჯის სერიას, რომელთა გადადგმაც მთავრობებს შეუძლიათ გენერირებადი ხელოვნური ინტელექტის რეგულირებისა და საზოგადოებრივი ზედამხედველობის აღსადგენად, რათა გამოიყენონ მისი პოტენციალი სხვადასხვა სექტორში, მათ შორის განათლებაში.

ნაბიჯი 1: საერთაშორისო მონაცემთა დაცვის ზოგადი რეგულაციების (GDPR) მიღება და ეროვნული GDPR-ების შემუშავება

GenAI მოდელების მუშაობა მოიცავს მოქალაქეებისგან ონლაინ მონაცემების შეგროვებას და დამუშავებას. მონაცემებისა და კონტენტის გამოყენება GenAI მოდელების მიერ თანხმობის გარეშე კიდევ უფრო ართულებს მონაცემთა

დაცვის საკითხს. მონაცემთა დაცვის ზოგადი რეგულაცია, მათ შორის ევროკავშირის GDPR, მიღებულ იქნა 2018 წელს, როგორც ერთ-ერთი პირველი ცდა.

GDPR უზრუნველყოფს GenAI პროვაიდერების მიერ პერსონალური მონაცემების შეგროვებისა და დამუშავების რეგულირების აუცილებელ სამართლებრივ ჩარჩოს. ვაჭრობისა და განვითარების შესახებ გაეროს კონფერენციის (UNCTAD) მონაცემთა დაცვისა და კონფიდენციალურობის კანონმდებლობის Worldline-ის თანახმად, 194 ქვეყნიდან 137-მა მიიღო კანონმდებლობა, რომელიც უზრუნველყოფს მონაცემთა დაცვას და კონფიდენციალურობას.

თუმცა, გაურკვეველი რჩება, თუ რამდენად ხორციელდება ეს რეგულაციები ამ ქვეყნებში. ამიტომ, სულ უფრო მნიშვნელოვანი ხდება მათი სათანადო განხორციელების უზრუნველყოფა, მათ შორის GenAI სისტემების რეგულარული მონიტორინგი. ასევე მნიშვნელოვანია, რომ ქვეყნებმა, რომლებსაც ჯერ არ აქვთ ზოგადი მონაცემთა დაცვის კანონები, შეიმუშაონ ისინი.

ნაბიჯი 2: ქვეყნის მასშტაბით AI სტრატეგიების განვითარება და დაფინანსება

გენერირებადი AI რეგულირება უნდა იყოს უფრო ფართო ეროვნული AI სტრატეგიების განუყოფელი ნაწილი, რომელსაც შეუძლია უზრუნველყოს AI უსაფრთხო და სამართლიანი გამოყენება სხვადასხვა სექტორებში, მათ შორის განათლებაში. მხოლოდ სახელმწიფო მიდგომას შეუძლია უზრუნველყოს სექტორთაშორისი ქმედებების კოორდინაცია, რომელიც აუცილებელია. 2023 წლის დასაწყისისთვის, დაახლოებით 67 ქვეყანას ჰქონდა შემუშავებული ან დაგეგმილი ეროვნული AI სტრატეგიები. ამ ეროვნული სტრატეგიებიდან არცერთს არ ჰქონდა განხილული გენერირებადი AI, როგორც დამოუკიდებელი საკითხი.

ნაბიჯი 3: AI ეთიკის შესახებ კონკრეტული რეგულაციების დანერგვა.

AI გამოყენებასთან დაკავშირებული ეთიკური საკითხების მოსაგვარებლად საჭიროა კონკრეტული რეგულაციები. 2023 წელს იუნესკოს მიერ ხელოვნური ინტელექტის არსებული ეროვნული სტრატეგიების მიმოხილვა აჩვენებს, რომ ასეთი ეთიკური საკითხების იდენტიფიცირება და სახელმძღვანელო პრინციპების ჩამოყალიბება მხოლოდ ორმოცი ეროვნული ხელოვნური ინტელექტის სტრატეგიისთვისაა დამახასიათებელი.

მნიშვნელოვანია, რომ ეთიკური პრინციპები სავალდებულო კანონებად ან რეგულაციებად უნდა იქნას ასახული. ეს ჯერჯერობით იშვიათად ხდება. სინამდვილეში, მხოლოდ ოცამდე ქვეყანას აქვს განსაზღვრული ხელოვნური ინტელექტის რაიმე აშკარა ეთიკის წესები, მათ შორის განათლებასთან დაკავშირებული, როგორც მათი ეროვნული ხელოვნური ინტელექტის სტრატეგიების ფარგლებში, ასევე სხვაგვარად. ქვეყნებმა, რომლებსაც ჯერ არ აქვთ ხელოვნური ინტელექტის ეთიკის წესები, სასწრაფოდ უნდა შეიმუშაონ და განახორციელონ ისინი.

ნაბიჯი 4: AI მიერ გენერირებული კონტენტი და საავტორო უფლებები

გენერირებადი AI მზარდმა გამოყენებამ შექმნა საავტორო უფლებების შესახებ ახალი საკითხები, რომლებიც ეხება როგორც კონტენტს ან საავტორო უფლებებით დაცულ ნამუშევრებს, რომლებზეც მოდელები არიან გაწვრთნილები, ასევე მათ მიერ წარმოებული „არაადამიანური“ ცოდნის სტატუსს.

ამჟამად, მხოლოდ ჩინეთმა, ევროკავშირმა (EU) და შეერთებულმა შტატებმა შეცვალეს საავტორო უფლებების შესახებ კანონები გენერირებადი ხელოვნური ინტელექტის საკითხების მოსაგვარებლად. მაგალითად, აშშ-ის საავტორო უფლებების ოფისმა დაადგინა, რომ GenAI სისტემების, როგორიცაა ChatGPT, შედეგები არ არის დაცული აშშ-ის საავტორო უფლებების კანონმდებლობით, იმ მოტივით, რომ „საავტორო უფლებას შეუძლია დაიცვას მხოლოდ ის მასალა, რომელიც ადამიანის შემოქმედების პროდუქტია“ (აშშ-ის საავტორო უფლებების ოფისი, 2023). ამასობაში, ევროკავშირში, შემოთავაზებული ევროკავშირის ხელოვნური ინტელექტის აქტი მოითხოვს ხელოვნური ინტელექტის ინსტრუმენტების შემქმნელებისგან გაამჟღავნონ საავტორო უფლებებით დაცული მასალები, რომლებიც მათ გამოიყენეს თავიანთი სისტემების შესაქმნელად (ევროკომისია, 2021). ჩინეთი, 2023 წლის ივლისში გამოცემული GenAI რეგულაციის შესაბამისად, მოითხოვს, რომ GenAI შედეგები იყოს აღიარებული AI მიერ გენერირებულ კონტენტად და ციფრული სინთეზის შედეგებად.

ნაბიჯი 5: გენერირებადი AI მარეგულირებელი ჩარჩოს შემუშავება

AI ტექნოლოგიების განვითარების სწრაფი ტემპი აიძულებს ეროვნულ და ადგილობრივ მთავრობებს, დააჩქარონ თავიანთი რეგულაციების განახლება. 2023 წლის ივლისის მონაცემებით, მხოლოდ ერთმა ქვეყანამ, ჩინეთმა, გამოსცა სპეციალური ოფიციალური რეგულაციები გენერირებადი AI შესახებ. გენერირებადი AI სერვისების მარეგულირებელი დროებითი დებულებები, რომლებიც გამოქვეყნდა 2023 წლის 13 ივლისს (ჩინეთის კიბერსივრცის ადმინისტრაცია, 2023ა), მოითხოვს გენერირებადი AI სისტემის პროვაიდერებისგან სწორად და კანონიერად მონიშნონ AI მიერ გენერირებული კონტენტი, სურათები და ვიდეოები ონლაინ ინფორმაციის ღრმა სინთეზის შესახებ არსებული წესების შესაბამისად.

ნაბიჯი 6: გენერირებადი AI გამოყენების შესაძლებლობების განვითარება

სკოლებმა და სხვა საგანმანათლებლო დაწესებულებებმა უნდა გააცნობიერონ AI, მათ შორის გენერირებადი AI, პოტენციური სარგებელი და რისკები განათლებისთვის. პედაგოგებსა და მკვლევარებს მხარდაჭერა სჭირდებათ GenAI-ის სათანადოდ გამოყენების შესაძლებლობების გასაძლიერებლად, მათ შორის ტრენინგებისა და მუდმივი მენტორობის გზით. რამდენიმე ქვეყანამ დაიწყო ასეთი შესაძლებლობების განვითარების პროგრამები (მაგალითად: ჩინეთი, სინგაპური), რომელიც სთავაზობს საგანმანათლებლო დაწესებულებებისათვის AI შესაძლებლობების განვითარების სპეციალურ პლატფორმას თავისი AI სამ-

თავრობო ღრუბლოვანი კლასტერის მეშვეობით, რომელიც მოიცავს GPT მოდელების სპეციალურ საცავს.

ნაბიჯი 7: GenAI-ის განათლებასა და კვლევაზე ზემოქმედების ანალიზი

GenAI-ის ამჟამინდელი ვერსიების გავლენა მხოლოდ ახლა იწყებს გამოვლენას და მისი გავლენა განათლებაზე ჯერ კიდევ ბოლომდე არ არის გაგებული. ამასობაში, GenAI-ის უფრო ძლიერი ვერსიები და ხელოვნური ინტელექტის სხვა კლასები კვლავ ვითარდება და გამოიყენება. თუმცა, მნიშვნელოვანი კითხვები რჩება GenAI-ის გავლენასთან დაკავშირებით ცოდნის შექმნის, გადაცემისა და გადამოწმებისთვის – სწავლებისა და სწავლისთვის, სასწავლო გეგმის შემუშავებისა და შეფასებისთვის, ასევე კვლევისა და საავტორო უფლებებისთვის. ქვეყნების უმეტესობა განათლებაში GenAI-ის დანერგვის საწყის ეტაპზეა და, ამდენად, გრძელვადიანი AI გამოყენების შედეგები ჯერ კიდევ არაა შესწავლილი.

გენერირებადი AI რეგულირება: ძირითადი ელემენტები

ყველა ქვეყანამ სათანადოდ უნდა დაარეგულიროს გენერირებადი AI, რათა უზრუნველყოს მისი წვლილი განათლებისა და სხვა კონტექსტების განვითარებაში. გენერირებადი AI შესახებ რეგულაციების შემუშავების, დამტკიცებისა და განხორციელების კოორდინაციისთვის საჭიროა სამთავრობო მიდგომა. რეკომენდებულია შემდეგი შვიდი ძირითადი ელემენტი და ქმედება:

- სექტორთაშორისო კოორდინაცია: ეროვნული ორგანოს შექმნა, რომელიც წარმართავს სამთავრობო მიდგომას GenAI-ის მიმართ და კოორდინაციას გაუწევს თანამშრომლობას სექტორებს შორის.
- კანონმდებლობის შესაბამისობაში მოყვანა: მონაცემთა დაცვის ზოგადი კანონები, ინტერნეტ უსაფრთხოების რეგულაციები, მოქალაქეებისგან შეგროვებული ან მოქალაქეების მომსახურებისთვის გამოყენებული მონაცემების უსაფრთხოების შესახებ კანონები და სხვა შესაბამისი კანონები და საერთო პრაქტიკა. არსებული რეგულაციების და GenAI-ით წარმოშობილი ახალი გამოწვევების საპასუხოდ საჭირო ნებისმიერი ადაპტაციის შეფასება.
- GenAI-ის რეგულირებასა და ხელოვნური ინტელექტის ინოვაციების ხელშეწყობას შორის ბალანსი: სექტორთაშორისი თანამშრომლობის ხელშეწყობა კომპანიებს, საგანმანათლებლო და კვლევით დაწესებულებებსა და შესაბამის სამთავრობო სააგენტოებს შორის, რათა ერთობლივად შეიმუშაონ მყარი მოდელები; ხელი შეუწყონ ღია კოდის ეკოსისტემების შექმნას, რათა ხელი შეუწყონ სუპერკომპიუტერული რესურსების და მაღალი ხარისხის ტრენინგის მონაცემთა ნაკრებების გაზიარებას; და ხელი შეუწყონ GenAI-ის პრაქტიკულ გამოყენებას სხვადასხვა სექტორში და მაღალი ხარისხის კონტენტის შექმნას საზოგადოებრივი სიკეთისთვის.

- AI პოტენციური რისკების შეფასება: საჭიროა პრინციპების შემუშავება, რომელიც უზრუნველყოფს GenAI სერვისების ეფექტურობას, და უსაფრთხოებას. უნდა მოხდეს რისკების კატეგორიზაცია და გამოიყოს მაღალი რისკის ზონები. სხვადასხვა რისკის ზონებისათვის რეგულაციები სხვადასხვა უნდა იყოს.
- მონაცემთა კონფიდენციალურობის დაცვა: GenAI-ის გამოყენება თითქმის ყოველთვის გულისხმობს მომხმარებლების მიერ მათი მონაცემების GenAI პროვაიდერთან გაზიარებას. საჭიროა კანონების შემუშავება და განხორციელება მომხმარებლების პერსონალური ინფორმაციის დასაცავად, ასევე უკანონო შენახვის, პროფილირებისა და მონაცემთა გაზიარების იდენტიფიცირებისა და თავიდან ასაცილებლად.
- GenAI-ის გამოყენების ასაკობრივი შეზღუდვები: GenAI აპლიკაციების უმეტესობა ძირითადად ზრდასრული მომხმარებლებისთვისაა განკუთვნილი. ეს აპლიკაციები ხშირად მნიშვნელოვან რისკებს უქმნის ბავშვებს, მათ შორის შეუსაბამო კონტენტის ზემოქმედებას და პოტენციურ მანიპულირებას. ამ რისკების გათვალისწინებით და იმ მნიშვნელოვანი გაურკვევლობის გათვალისწინებით, რომელიც კვლავაც ირგვლივ არსებობს GenAI-ის განმეორებითი აპლიკაციების გარშემო, მკაცრად რეკომენდებულია ასაკობრივი შეზღუდვების დაწესება ზოგადი დანიშნულების ხელოვნური ინტელექტის ტექნოლოგიებისთვის, რათა დაიცვან ბავშვების უფლებები და კეთილდღეობა. ChatGPT-ის მომსახურების პირობები ამჟამად მოითხოვს, რომ მომხმარებლები იყვნენ მინიმუმ 13 წლის, ხოლო 18 წლამდე ასაკის მომხმარებლებს უნდა ჰქონდეთ მშობლის ან კანონიერი მეურვის ნებართვა სერვისების გამოსაყენებლად. ბევრი 13 წლის ზღვარს ძალიან დაბალს მიიჩნევს და ასაკის 16 წლამდე გაზრდის შესახებ კანონმდებლობის მიღებას ემხრობა. ევროკავშირის GDPR (2016) განსაზღვრავს, რომ მომხმარებლები მშობლების ნებართვის გარეშე სოციალური მედიის სერვისების გამოსაყენებლად მინიმუმ 16 წლის უნდა იყვნენ. სხვადასხვა GenAI ჩატბოტების გაჩენა მოითხოვს, რომ ქვეყნებმა ყურადღებით განიხილონ და საჯაროდ განიხილონ GenAI პლატფორმებზე შესაბამისი ასაკობრივი ზღვარი.
- ეროვნული მონაცემთა საკუთრება და ინფორმაციული სიღარიბის რისკი: საჭიროა საკანონმდებლო ზომების მიღება ეროვნული მონაცემთა საკუთრების დასაცავად და ქვეყნებში მოქმედი GenAI პროვაიდერების რეგულირებისთვის. უნდა შემუშავდეს წესები, რომლებიც ხელს უწყობენ ურთიერთსასარგებლო თანამშრომლობას, რათა უზრუნველყონ, რომ მონაცემების ეს კატეგორია არ გავიდეს ქვეყნიდან ტექნოლოგიური კომპანიების ექსკლუზიური გამოყენებისთვის.

GenAI ინსტრუმენტების პროვაიდერები

GenAI პროვაიდერები მოიცავს ორგანიზაციებს და პირებს, რომლებიც პასუხისმგებელნი არიან GenAI ინსტრუმენტების შემუშავებასა და მიწოდებაზე და/ან GenAI ტექნოლოგიების გამოყენებაზე მომსახურების გასაწევად, მათ შორის, აპლიკაციების პროგრამირების ინტერფეისების (API) მეშვეობით. GenAI ინსტრუმენტების ყველაზე გავლენიანი პროვაიდერები კარგად დაფინანსებული კომპანიებია. ისინი პასუხისმგებელნი არიან ეთიკურ შესაბამისობაზე, მათ შორის წესებში დადგენილი ეთიკური პრინციპების განხორციელებაზე. ეს სისტემა უნდა მოიცავდეს პასუხისმგებლობის შემდეგ ათი კატეგორიას:

- ადამიანური პასუხისმგებლობა: GenAI პროვაიდერები პასუხისმგებელნი არიან იმაზე, რომ დაიცვან ძირითადი ღირებულებები და ლეგიტიმური მიზნები, პატივი სცენ ინტელექტუალურ საკუთრებას და დაიცვან ეთიკური სტანდარტები. ასევე, პასუხისმგებლები არიან დეზინფორმაციისა და სიძულვილის ენის გავრცელების თავიდან აცილებაზე.
- სანდო მონაცემები და მოდელები: GenAI პროვაიდერები ვალდებულნი არიან დაადასტურონ მათ მოდელებში გამოყენებული მონაცემთა წყაროებისა და მეთოდების სანდოობა და ეთიკურობა. თუ მოდელებს სჭირდებათ პერსონალური მონაცემების გამოყენება, ასეთი ინფორმაცია უნდა შეგროვდეს მხოლოდ მფლობელების ინფორმირებული თანხმობით.
- არადისკრიმინაციული კონტენტის გენერირება: GenAI პროვაიდერებმა უნდა აკრძალონ GenAI სისტემების შემუშავება და განთავსება, რომლებიც წარმოქმნიან მიკერძოებულ ან დისკრიმინაციულ კონტენტს რასის, ეთნიკური კუთვნილების, სქესის ან სხვა დაცული მახასიათებლების საფუძველზე. მათ უნდა უზრუნველყონ მკაცრი „დაცვის ზომები“, რათა თავიდან აიცილონ GenAI-ის მიერ შეურაცხმყოფელი, მიკერძოებული ან ცრუ კონტენტის გენერირება.
- GenAI მოდელების ახსნადობა და გამჭვირვალობა: პროვაიდერებმა უნდა მიაწოდონ მთავრობებს ახსნა-განმარტებები წყაროების, მასშტაბის და მონაცემების ტიპების შესახებ, რომელიც მათი მოდელების მიერ გამოყენებულია სწავლებისათვის, მათი მოდელების მიერ გამოყენებული მეთოდების ან ალგორითმების შესახებ, რომელიც გამოიყენება კონტენტის ან პასუხების გენერირებისთვის.
- GenAI კონტენტის მარკირება: პროვაიდერებმა სათანადოდ უნდა მკაფიოდ მონიშნონ GenAI-ის მიერ გენერირებული დოკუმენტები, ანგარიშები, სურათები და ვიდეოები.
- უსაფრთხოების პრინციპები: GenAI პროვაიდერებმა უნდა უზრუნველყონ უსაფრთხო, საიმედო და მდგრადი სერვისები GenAI სისტემის მთელი სასიცოცხლო ციკლის განმავლობაში.

- წვდომისა და გამოყენებადობის მოთხოვნები: GenAI პროვაიდერებმა უნდა წარმოადგინონ მკაფიო მოთხოვნები სამიზნე აუდიტორიისთვის, გამოყენების შემთხვევებისა და მათი სერვისების მიზნებისთვის და დაეხმარონ GenAI ინსტრუმენტების მომხმარებლებს რაციონალური და პასუხისმგებლიანი გადაწყვეტილებების მიღებაში.
- შეზღუდვების აღიარება და პროგნოზირებადი რისკების შემცირება: GenAI პროვაიდერებმა ნათლად უნდა მიუთითონ სისტემების მიერ გამოყენებული მეთოდებისა და მათი გამომავალი მონაცემების შეზღუდვები. მათ უნდა შეიმუშაონ ტექნოლოგიები იმის უზრუნველსაყოფად, რომ მონაცემები, მეთოდები და შედეგები არ იწვევდეს მომხმარებლებისთვის პროგნოზირებად ზიანს, და შეიმუშავონ მექანიზმები გაუთვალისწინებელი ზიანის შესამცირებლად, თუ ეს მოხდება.
- საჩივრების მექანიზმები და საშუალებები: GenAI-ის პროვაიდერებმა უნდა შექმნან მექანიზმები და არხები მომხმარებლებისა და ფართო საზოგადოებისგან საჩივრების შესაგროვებლად და დროულად მიიღონ ზომები ამ საჩივრების მისაღებად და დასამუშავებლად.
- უკანონო გამოყენების მონიტორინგი და შეტყობინება: პროვაიდერებმა უნდა ითანამშრომლონ სამთავრობო ორგანოებთან უკანონო გამოყენების მონიტორინგისა და შეტყობინების ხელშეწყობის მიზნით. ეს მოიცავს შემთხვევებს, როდესაც ადამიანები GenAI პროდუქტებს უკანონოდ იყენებენ ან არღვევენ ეთიკურ ან სოციალურ ღირებულებებს, როგორიცაა დეზინფორმაციის ან სიძულვილის ენის გავრცელება, სპამის შექმნა ან მავნე პროგრამების შექმნა.

ქეისები: GenAI ინსტრუმენტების პროვაიდერები

ქეისი 1: OpenAI და ChatGPT – კონტენტის მოდერაცია და ასაკობრივი შეზღუდვები OpenAI-მა 2023 წელს შემოიღო გაუმჯობესებული მოდერაციის სისტემა ChatGPT-ისთვის, რომელიც მიზნად ისახავს არასრულწლოვნებისთვის შეუსაბამო კონტენტის ფილტრაციას. თუმცა, გამოვლინდა, რომ ასაკის გადამოწმების მექანიზმები არასაკმარისია – 13 წლამდე ბავშვები ადვილად ახერხებდნენ სისტემაში შესვლას. ეს ქეისი ცხადყოფს, რომ ტექნიკური შეზღუდვები უნდა ემყარებოდეს სამართლებრივ ვალდებულებებს და არა მხოლოდ პლატფორმის ნებაყოფლობით პოლიტიკას.

ქეისი 2: Google Bard/Gemini – მიკერძოება და სამართლიანობა Google-ის GenAI მოდელმა რამდენიმე შემთხვევაში გამოავლინა კულტურული და გენდერული მიკერძოება გამოსახულებების გენერირებისას. 2024 წელს Google-ს მოუწია მოდელის დროებით შეჩერება ისტორიული პერსონაჟების არაზუსტი გამოსახვის გამო. ეს ქეისი ხაზს უსვამს GenAI პროვაიდერების ვალდებულებას

– უზრუნველყონ სისტემატური ტესტირება მიკერძოების გამოვლენისა და გამოსწორების მიზნით.

ქეისი 3: Turnitin და აკადემიური კეთილსინდისიერება სასწავლო პლატფორმა Turnitin-მა 2023 წელს ინტეგრირება მოახდინა AI-ის მიერ დაწერილი ტექსტის გამოვლენის ინსტრუმენტისა, რომელიც ეხმარება პედაგოგებს დაადგინონ, გამოიყენა თუ არა სტუდენტმა GenAI ნაშრომის შესაქმნელად. ეს ქეისი აჩვენებს, თუ როგორ შეუძლიათ პროვაიდერებს პასუხისმგებლობის ინსტრუმენტები ჩართონ საგანმანათლებლო სფეროში, თუმცა ასევე ბადებს კითხვებს სიზუსტისა და სტუდენტთა მიმართ სამართლიანობის შესახებ.

ინსტიტუციური მომხმარებლები

ინსტიტუციურ მომხმარებლებს შორის არიან განათლების ორგანოები და ინსტიტუტები, როგორიცაა უნივერსიტეტები და სკოლები, რომლებიც პასუხისმგებელნი არიან GenAI-ის გამოყენებაზე მათ ინსტიტუტებში. საჭიროა:

- GenAI ალგორითმების, მონაცემებისა და შედეგების ინსტიტუციური აუდიტი: მექანიზმების დანერგვა GenAI ინსტრუმენტების მიერ გამოყენებული ალგორითმებისა და მონაცემების და მათ მიერ გენერირებული შედეგების ზუსტი მონიტორინგისთვის. ეს უნდა მოიცავდეს რეგულარულ აუდიტს და შეფასებებს, მომხმარებლის მონაცემების დაცვას და შეუსაბამო შინაარსის ავტომატურ ფილტრაციას.
- პროპორციულობის შემოწმება და მომხმარებლის კეთილდღეობის დაცვა: ეროვნული კლასიფიკაციის მექანიზმების დანერგვა ან ინსტიტუციური პოლიტიკის შემუშავება GenAI სისტემებისა და აპლიკაციების კატეგორიზაციისა და ვალიდაციისთვის. დაწესებულებაში გამოყენებული GenAI სისტემები არ უნდა იწვევდეს პროგნოზირებად ზიანს მიზნობრივ მომხმარებლებისათვის.
- გრძელვადიანი ეფექტების ანალიზი და განხილვა: დროთა განმავლობაში, GenAI ინსტრუმენტების ან შინაარსის გამოყენებამ განათლებაში შეიძლება ღრმა გავლენა მოახდინოს ადამიანის შესაძლებლობების განვითარებაზე, როგორიცაა კრიტიკული აზროვნების უნარები და კრეატიულობა. ეს პოტენციური ზემოქმედება უნდა შეფასდეს და გათვალისწინებულ იქნას.
- ასაკის შესაბამისობა: დაწესებულებაში GenAI-ის დამოუკიდებლად გამოყენების მინიმალური ასაკობრივი ზღვრის დაწესება არის მნიშვნელოვანი.

ქეისები: ინსტიტუციური მომხმარებლები

ქეისი 1: ჰარვარდის უნივერსიტეტი – GenAI პოლიტიკა ჰარვარდის უნივერსიტეტმა 2023 წელს შეიმუშავა GenAI-ის გამოყენების ყოვლისმომცველი პოლიტიკა, რომელიც განსხვავდება ფაკულტეტების მიხედვით. ზოგი კურ-

სი პირდაპირ კრძალავს ChatGPT-ის გამოყენებას, ზოგი კი მოითხოვს მის გამოყენებას გამჭვირვალობის სავალდებულო გამოცხადებით. ეს მოდელი ასახავს ინსტიტუციური მოქნილობისა და ეთიკური ჩარჩოების ბალანსის მნიშვნელობას.

ქეისი 2: სინგაპურის განათლების სამინისტრო სინგაპურმა შეიმუშავა ეროვნული პროგრამა, სახელწოდებით „AI for Education", რომელიც მიზნად ისახავს AI ინსტრუმენტების ინტეგრაციას სასკოლო სასწავლო პროგრამებში. პედაგოგებისთვის ჩატარდა სპეციალური ტრენინგები, ხოლო GenAI ინსტრუმენტების გამოყენება სტრუქტურირებული პედაგოგიური მიდგომებით შემოიფარგლა. ეს ქეისი ქვეყნებს მიანიშნებს, თუ როგორ შეიძლება ეროვნულ დონეზე კოორდინირებული და პასუხისმგებლური GenAI ინტეგრაცია.

ქეისი 3: იტალია – ChatGPT-ის დროებითი აკრძალვა 2023 წლის მარტში იტალიის მონაცემთა დაცვის ორგანომ (Garante) დროებით შეაჩერა ChatGPT-ის ხელმისაწვდომობა ქვეყანაში GDPR-ის დარღვევების გამო. ეს ქეისი ასახავს, თუ რამდენად მნიშვნელოვანია ეფექტური მარეგულირებელი ინსტიტუტების არსებობა, რომლებსაც შეუძლიათ სწრაფი და ეფექტური ქმედება განახორციელონ.

ინდივიდუალური მომხმარებლები

ინდივიდუალური მომხმარებლები პოტენციურად მოიცავს მსოფლიოს ყველა ადამიანს, ვისაც აქვს წვდომა ინტერნეტზე და მინიმუმ ერთ GenAI ინსტრუმენტზე. ჩვენ ვგულისხმობთ პედაგოგებს, მკვლევარებს და მოსწავლეებს ფორმალურ საგანმანათლებლო დაწესებულებებში ან არაფორმალურ სასწავლო პროგრამებში მონაწილე პირებს. საჭიროა:

- GenAI-ის გამოყენების პირობების გააზრება: მომსახურების ხელშეკრულებებზე ხელმოწერისას ან მათზე დათანხმებისას, მომხმარებლებმა უნდა იცოდნენ თავიანთი ვალდებულებები, დაიცვან ხელშეკრულებაში მითითებული გამოყენების პირობები.
- GenAI აპლიკაციების ეთიკური გამოყენება: მომხმარებლებმა GenAI პასუხისმგებლობით უნდა გამოიყენონ და თავი აარიდონ მის გამოყენებას ისე, რომ შეიძლება ზიანი მიაყენოს სხვების რეპუტაციას და კანონიერ უფლებებს.
- GenAI უკანონო აპლიკაციების მონიტორინგი და ანგარიშგება: თუ მომხმარებლები აღმოაჩენენ GenAI აპლიკაციებს, რომლებიც არღვევენ ერთ ან რამდენიმე რეგულაციას, მათ უნდა აცნობონ შესაბამის სამთავრობო მარეგულირებელ ორგანოებს.

YouTube რესურსები

1. **“AI in Education: Opportunities and Challenges”** – UNESCO-ს ოფიციალური არხი https://www.youtube.com/@UNESCO – მოიძიეთ: *“Generative AI and Education UNESCO”* ეს ვიდეო განიხილავს GenAI-ის როლს განათლებაში და UNESCO-ს რეკომენდაციებს.
2. **“The Ethics of Artificial Intelligence”** – TED-Ed https://www.youtube.com/watch?v=UbruBnv3pZU მოკლე, ხელმისაწვდომი ახსნა AI-ის ეთიკური გამოწვევების შესახებ – შესაფერისია საგანმანათლებლო კონტექსტისთვის.
3. **“How AI Could Save (Not Destroy) Education”** – Sal Khan, TED Talk https://www.youtube.com/watch?v=hJP5GqnTrNo Khan Academy-ს დამფუძნებელი განიხილავს, თუ როგორ შეიძლება AI გახდეს პერსონალიზებული განათლების ძლიერი ინსტრუმენტი.
4. **“Regulation of AI – What You Need to Know”** – ColdFusion https://www.youtube.com/@ColdFusion – მოიძიეთ: *“AI Regulation”* ხელმისაწვდომი ახსნა მსოფლიოში AI-ის რეგულირების მიმდინარე პროცესების შესახებ.

სადისკუსიო თემები

უნდა აიკრძალოს თუ არა GenAI სასკოლო გარემოში? *განიხილეთ შემდეგი პოზიციები:*

- *GenAI-ის სრული აკრძალვა სკოლებში იცავს სტუდენტებს და უზრუნველყოფს კრიტიკული აზროვნების განვითარებას.*
- *GenAI-ის კრძალვა პრობლემას ვერ გადაჭრის – სჯობს ვასწავლოთ მისი პასუხისმგებლობიანი გამოყენება.*
- *რა ბალანსი არის შესაძლებელი ამ ორ პოზიციას შორის?*

ვინ უნდა იყოს პასუხისმგებელი GenAI-ის მიერ გამოწვეული ზიანისთვის – პლატფორმა, ინსტიტუტი თუ ინდივიდი? *განიხილეთ სხვადასხვა სცენარი:*

- *სტუდენტი GenAI-ის მეშვეობით ქმნის პლაგიატს.*
- *GenAI გენერირებს მცდარ სამედიცინო ინფორმაციას, რომელსაც მოსწავლე ენდობა.*
- *პედაგოგი GenAI-ის მეშვეობით ფასებს სამუშაოს და დაუშვებს შეცდომებს.*
- *ვინ არის პასუხისმგებელი თითოეულ სცენარში?*

არის თუ არა GenAI-ის გამოყენება განათლებაში სამართლიანი ყველა მოსწავლისთვის? *განიხილეთ:*

- *GenAI-ზე წვდომა არათანაბარია – მდიდარ სტუდენტებს უკეთესი ინსტრუმენტები აქვთ.*
- *ენობრივი ბარიერები: GenAI ძირითადად ინგლისურ ენაზეა ორიენტირებული.*
- *შეიძლება თუ არა GenAI გახდეს უთანასწორობის შემამცირებელი ინსტრუმენტი?*

ტესტი 12

1. რომელი საერთაშორისო ორგანიზაცია გამოსცა 2021 წელს ხელოვნური ინტელექტის ეთიკის შესახებ რეკომენდაცია?

ა) OOH
ბ) UNESCO
გ) OECD
დ) World Bank

2. პეკინის კონსენსუსი განათლებაში AI-ის შესახებ მიღებულ იქნა:

ა) 2021 წელს
ბ) 2018 წელს
გ) 2019 წელს
დ) 2023 წელს

3. GDPR – მონაცემთა დაცვის ზოგადი რეგულაცია ამოქმედდა:

ა) 2016 წელს
ბ) 2018 წელს
გ) 2020 წელს
დ) 2022 წელს

4. UNCTAD-ის მონაცემებით, 194 ქვეყნიდან რამდენმა მიიღო მონაცემთა დაცვის კანონმდებლობა?

ა) 100
ბ) 120
გ) 137
დ) 160

5. 2023 წლის დასაწყისისთვის, რამდენ ქვეყანას ჰქონდა ეროვნული AI სტრატეგია?

ა) დაახლოებით 30
ბ) დაახლოებით 50
გ) დაახლოებით 67
დ) დაახლოებით 100

6. რომელი ქვეყანა იყო პირველი, ვინც 2023 წლის ივლისში გამოსცა სპეციალური ოფიციალური რეგულაციები გენერირებადი AI-ის შესახებ?

ა) აშშ
ბ) ევროკავშირი
გ) იაპონია
დ) ჩინეთი

7. ChatGPT-ის მომსახურების პირობების თანახმად, მომხმარებლის მინიმალური ასაკი არის:

ა) 10 წელი
ბ) 13 წელი
გ) 16 წელი
დ) 18 წელი

8. ევროკავშირის GDPR-ის თანახმად, მშობლების ნებართვის გარეშე სოციალური მედიის სერვისების გამოყენებისთვის მინიმალური ასაკია:

ა) 13 წელი
ბ) 14 წელი
გ) 16 წელი
დ) 18 წელი

9. GenAI პროვაიდერების პასუხისმგებლობის კატეგორიებიდან რომელი ითვალისწინებს სისტემების ახსნა-განმარტებასა და გამჭვირვალობას?

ა) უსაფრთხოების პრინციპები
ბ) GenAI მოდელების ახსნადობა და გამჭვირვალობა
გ) ადამიანური პასუხისმგებლობა
დ) საჩივრების მექანიზმები

10. UNESCO-ს სახელმძღვანელოს მიხედვით, განათლებაში AI-ის ეთიკური და შინაარსიანი გამოყენებისთვის შესაძლებლობების განვითარება არის:

ა) GenAI-ის რეგულირების მე-3 ნაბიჯი
ბ) GenAI-ის რეგულირების მე-4 ნაბიჯი
გ) GenAI-ის რეგულირების მე-5 ნაბიჯი
დ) GenAI-ის რეგულირების მე-6 ნაბიჯი

შეცვლის ხელოვნური ინტელექტი მასწავლებელს?

ნებისმიერი ახალი ტექნოლოგიის აქტიური დანერგვით, ჩნდება ეჭვები და შიშები და ეს ნორმალურია. აბრაამ მასლოუმ თავის ნაშრომში „ყოფიერების ფსიქოლოგია“ (1968) აღნიშნა, რომ ზრდასრული ადამიანები ბუნებრივად შეშფოთებულნი არიან ახალი ცოდნის მიმართ, რაც უსაფრთხოების საჭიროებასთან არის დაკავშირებული.

აბრაამ მასლოუ *(1908–1970) იყო ამერიკელი ფსიქოლოგი, ჰუმანისტური ფსიქოლოგიის ერთ-ერთი დამფუძნებელი. იგი განსაკუთრებით ცნობილია* ***მოთხოვნილებების იერარქიის თეორიით*** *(1943), რომელიც პირამიდის სახით გამოისახება და ადამიანის ძირითად მოთხოვნილებებს ხუთ დონედ ჰყოფს:*

- ***ფიზიოლოგიური მოთხოვნილებები*** *– საკვები, სუნთქვა, ძილი*
- ***უსაფრთხოების მოთხოვნილება*** *– სტაბილურობა, დაცვა, წესრიგი*
- ***სიყვარულისა და კუთვნილების მოთხოვნილება*** *– ოჯახი, მეგობრობა, სოციალური კავშირები*
- ***პატივისცემის მოთხოვნილება*** *– თვითშეფასება, აღიარება, სტატუსი*
- ***თვითაქტუალიზაცია*** *– პიროვნული პოტენციალის სრული რეალიზება*

მასლოუს თანახმად, ქვედა დონის მოთხოვნილებების დაუკმაყოფილებლობა ხელს უშლის ზედა დონეებზე გადასვლას. სწავლა და ზრდა **თვითაქტუალიზაციის** ნაწილია, მაგრამ ის შეუძლებელია, თუ ადამიანი **უსაფრთხოებას** არ გრძნობს. სწორედ ამიტომ, ახალი ტექნოლოგიების – მათ შორის ხელოვნური ინტელექტის – შემოტანა შეიძლება აღქმულ იქნეს, როგორც **საფრთხე სტაბილურობისთვის**, რაც ბუნებრივ წინააღმდეგობასა და შიშს იწვევს.

თავის გვიანდელ ნაშრომში **„ყოფიერების ფსიქოლოგია“** (Toward a Psychology of Being, 1968) მასლოუ ამტკიცებს, რომ ადამიანი ზრდის პროცესში მუდმივ დაძაბულობაშია ორ ძალას შორის: **უსაფრთხოების საჭიროება** (ნაცნობისკენ) და **ზრდის სწრაფვა** (სიახლისკენ). ეს განმარტავს, თუ რატომ გამოხატავენ პედაგოგები ხშირად ამბივალენტობას ინოვაციური ტექნოლოგიების მიმართ.

გონების ეს თვისება, უცნობისადმი და სიახლისადმი ბუნებრივი წინააღმდეგობა, გვეხმარება თავიდან ავიცილოთ მრავალი მცდარი და გაუაზრებელი გადაწყვეტილება.

ეს ნაწილი კიდევ ერთხელ მოკლედ მიმოიხილავს განათლებაში ხელოვნური ინტელექტის გამოყენებასთან დაკავშირებულ პრობლემებსა და საკითხებს.

ხელოვნური ინტელექტის შესახებ დისკუსიები ხშირად სხვადასხვა მითებითა და შიშითაა გარშემორტყმული. შესაძლოა, ყველაზე გავრცელებული შიშია თანამედროვე სამყაროს ერთ-ერთი უმნიშვნელოვანესი ღირებულების, სწავლების, განათლების, კულტურისა და გამოცდილების გადაცემის შესახებ. ანუ

განათლების პროცესის კომპიუტერულ ალგორითმებისთვის გადაცემის შიში. თუმცა, დღეს, ეს შიში, შესაძლოა, ყველაზე რეფლექსიურია და უკანა პლანზე გადადის.

სტატიაში „რატომ ვერასდროს ჩაანაცვლებს ხელოვნური ინტელექტი მასწავლებლებს“ (M. Lynch, 2018) ავტორი ამტკიცებს, რომ არ არის საჭირო მასწავლებლის პროფესიის ხელოვნური ინტელექტით ჩანაცვლების შიში:

„ხელოვნურ ინტელექტს არ შეუძლია შექმნას, გაიგოს ან მართოს რთული სტრატეგიული დაგეგმვა; მას არ შეუძლია შეასრულოს რთული სამუშაო, რომელიც მოითხოვს ზუსტ კოორდინაციას; მას არ შეუძლია ნავიგაცია უცნობ ან არასტრუქტურირებულ სივრცეში; და განსაკუთრებით არ შეუძლია, ადამიანებისგან განსხვავებით, გრძნობა, თანაგრძნობა და თანაგრძნობა... მხოლოდ ადამიანებს შეუძლიათ ამის გაკეთება. ამიტომ, მასწავლებლებზე მოთხოვნა მომავალშიც მაღალი იქნება“.

ვახდენთ აზროვნების უნარის ხელოვნური ინტელექტსადმი დელეგირებას?

იუნესკოს პოლიტიკის მოკლე დოკუმენტში „ხელოვნური ინტელექტი განათლებაში: სწავლის ტემპის შეცვლა“ (UNESCO, 2020) განხილულია ტექნოლოგიაზე დამოკიდებულების რისკი: თუ ზოგიერთი კოგნიტური ფუნქცია სისტემას გადაეცემა, შესუსტდება თუ არა ჩვენივე აზროვნების უნარი?

„კარგავენ თუ არა სტუდენტები, რომლებიც კომპიუტერის კლავიატურის გამოყენებას ამჯობინებენ, კალიგრაფიულ უნარებს? კარგავენ თუ არა ისინი, ვინც გამოთვლებს ასრულებენ ცხრილში ან კალკულატორში, გონებრივი არითმეტიკის შესრულების უნარს? გავლენა მოახდინა თუ არა GPS-ის გამოყენებამ მარშრუტის დასადგენად ჩვენს სივრცეში ნავიგაციის უნარზე – სიტყვასიტყვით, „საკუთარი გზის პოვნაზე“? რადგან ეს ტექნოლოგიები უფრო დახვეწილი და ეფექტური ხდება, უფრო მნიშვნელოვანი ხდება იმ მიმართულების ცოდნა, რომლითაც გვსურს განვითარება“.

„ვახდენთ აზროვნების უნარის ხელოვნური ინტელექტსადმი დელეგირებას?“

UNESCO-ს მიერ დასმული კითხვა – კარგავენ თუ არა სტუდენტები კოგნიტურ უნარებს ტექნოლოგიაზე დამოკიდებულების გამო – ეხება **კოგნიტური გადატვირთვის** და **გარე შემეცნების** ფსიქოლოგიურ ფენომენებს.

კლავიატურა vs კალიგრაფია: კვლევები (Mangen et al., 2015) გვიჩვენებს, რომ ხელით წერა ააქტიურებს ტვინის უფრო მეტ ზონას, ვიდრე ტაიპინგი. ხელით წერა ეხმარება ბავშვებს ასოების ამოცნობასა და კითხვის უნარის განვითარებაში, ამიტომ კლავიატურაზე სრული გადასვლა ადრეულ ასაკში შეიძლება საზიანო იყოს.

კალკულატორი vs გონებრივი არითმეტიკა: გონებრივი გამოთვლა მხოლოდ „პასუხის მიღება“ არ არის – ეს არის **რიცხობრივი აზროვნების** და **ლოგიკის** ვარჯიში. კალკულატორის ადრეული და გაუკონტროლებელი გამოყენება ამ

„ვარჯიშს“ ართმევს ბავშვს. ამავდროულად, კალკულატორი მნიშვნელოვნად ათავისუფლებს კოგნიტურ რესურსს რთული პრობლემის გადაჭრისთვის.

GPS vs ნავიგაციის უნარი: ნეიროფსიქოლოგიური კვლევები (Woollett & Maguire, 2011) ლონდონის ტაქსის მძღოლებზე აჩვენებს, რომ ინტენსიური ნავიგაცია ზრდის ჰიპოკამპის მოცულობას – ტვინის იმ ნაწილს, რომელიც პასუხისმგებელია სივრცული ორიენტაციისა და მეხსიერებისთვის. GPS-ზე სრული დამოკიდებულება ამ „ბუნებრივ ვარჯიშს“ ანელებს.

ზოგადი პრინციპი: ტექნოლოგია ყველაზე ეფექტურია, როდესაც ის **ავსებს** ადამიანის შესაძლებლობებს და არა **ანაცვლებს** მათ. გამოწვევა პედაგოგებისთვის არის განსაზღვრონ, რომელი კოგნიტური პროცესები **სავალდებულოდ** უნდა ვარჯიშობდეს ადამიანი (შემოქმედება, კრიტიკული აზროვნება, ემპათია) და სად შეიძლება ტექნოლოგიამ **გაანთავისუფლოს** რესურსი უფრო ღრმა სწავლისთვის.

ასეთ ეჭვებზე სწორი პასუხი შეიძლება იყოს მოწოდება ტექნოლოგიის შეგნებული და მიზანმიმართული გამოყენებისკენ. სწავლის მთელი კოგნიტური სფერო არ შეიძლება მივანდოთ ხელოვნურ ინტელექტსა და ალგორითმებს. მანქანებს სჭირდებათ მკაფიო კოორდინატები და მკაცრად განსაზღვრული პირობები. გარდა ამისა, ალგორითმს არ შეუძლია თანაგრძნობა, მისთის უცნობია ემპათია. ამიტომ, საქმიანობის და დავალებების შესრულების დაუფიქრებლად დელეგირებამ, თუნდაც მოწინავე ტექნოლოგიებზე, შეიძლება გამოიწვიოს სოციალური იზოლაცია.

კონკრეტული ამოცანისთვის ხელოვნური ინტელექტის გამოყენებისას, ყოველთვის აუცილებელია ვიცოდეთ პასუხი კითხვაზე, თუ რატომ ვაკეთებთ ამას. ხომ არ ვარღვევთ მნიშვნელოვან პედაგოგიურ პროცესს, მაგალითად, ხელოვნური ინტელექტისთვის მხატვრული ლიტერატურის კითხვის დაუფიქრებლად გადაცემით, და, რეალურად, მხოლოდ ლიტერატურის რეზიუმეს ვიღებთ?

უნდა გვახსოვდეს, რომ ხელოვნური ინტელექტის გადაწყვეტილებების ფასი შეიძლება ძალიან მაღალი იყოს – ბავშვის, სტუდენტის ან ნებისმიერი მოსწავლის მომავალი შეიძლება ამაზე იყოს დამოკიდებული. ამიტომ, ბევრ ამოცანაში უფრო გონივრულია ხელოვნური ინტელექტის ელემენტების გამოყენება ადამიანი მასწავლებლის ფრთხილად და გააზრებული ზედამხედველობის ქვეშ.

კარგად ასწავლის თუ არა ხელოვნური ინტელექტი?

წიგნის „ხელოვნური ინტელექტი განათლებაში“ (Holmes, Bialik, Fadel, 2019) ავტორები ხაზს უსვამენ ხელოვნური ინტელექტის გამოყენებასთან დაკავშირებულ შემდეგ მიმდინარე და პოტენციურ პრობლემებს:

1. შესაძლოა შემცირდეს მოსწავლეების/სტუდენტების მონაწილეობა საკუთარ სწავლაში. ეს განსაკუთრებით მნიშვნელოვანია ზრდასრულთა განათლებაში, სადაც თავად სტუდენტები ასრულებენ წამყვან როლს

სასწავლო პროცესის ორგანიზებაში. ისინი კარგავენ ამ პერსონალიზაციის შესაძლებლობას, თუ სწავლის შინაარსი, სტრუქტურა და თანმიმდევრობა განისაზღვრება მხოლოდ ხელოვნური ინტელექტის ალგორითმებითა და მოდელებით.

2. პედაგოგიური მიდგომების დიაპაზონი შეზღუდულია. ეს შეზღუდვა დაკავშირებულია ზემოთ მოცემულ საკითხთან. ვინაიდან გადაწყვეტილებები მიიღება სისტემის მიერ, შეიძლება უგულებელყოფილი იყოს მნიშვნელოვანი პედაგოგიური პრაქტიკა, როგორიცაა თანამშრომლობითი სწავლება, არაფორმალური სწავლება და ა.შ.
3. ავტომატიზაცია ასუსტებს შინაარსობრივ ცოდნას. ყველაზე ადვილია ფაქტობრივი ინფორმაციის შეძენის ავტომატიზაცია, რაც არ არის პრიორიტეტი ეფექტური სწავლისა და ცოდნის პრაქტიკაში გადაცემისთვის. უფრო რთულია სასწავლო პროცესების ავტომატიზაცია ინფორმაციის რეფლექსიის, ანალიზისა და სინთეზის დონეზე, ასევე მიღებული ცოდნის შეფასების დონეზე.

შეიძლება ალგორითმი იყოს მიკერძოებული?

არსებობს შეშფოთება, რომ შეუძლებელია მონაცემების შერჩევა სასწავლო მოდელებისთვის სრულიად ობიექტურად. გარდა ამისა, ალგორითმის მიერ მიღებულმა არასწორმა გადაწყვეტილებამ შეიძლება უარყოფითად იმოქმედოს სწავლაზე, მის მონიტორინგზე ან თუნდაც სტუდენტის მომავალ ბედზე (მაგალითად, სტუდენტის მიღების ან გარიცხვის გადაწყვეტილების შემთხვევაში).

მიკერძოებული ალგორითმები განათლებაში

- **UK-ის საგამოცდო ალგორითმი (A-Level, 2020)** 2020 წელს, COVID-19 პანდემიის გამო გამოცდების გაუქმების შემდეგ, დიდი ბრიტანეთის მთავრობამ გადაწყვიტა სტუდენტების შეფასება ალგორითმის მეშვეობით მოხდინა, რომელიც ეყრდნობოდა სკოლის ისტორიულ მაჩვენებლებს. შედეგი: სახელმწიფო სკოლების, განსაკუთრებით ღარიბ რაიონებში მდებარე სკოლების, 40%-მდე სტუდენტს ქულები **ჩამოუწიეს**, ხოლო კერძო სკოლების სტუდენტები მეტად სარგებლობდნენ. ამ სისტემამ კლასობრივი უთანასწორობა გააღრმავა. მასიური საპროტესტო გამოსვლების შემდეგ, მთავრობას მოუწია ალგორითმიდან უარის თქმა. **გაკვეთილი:** ისტორიული მონაცემებზე დაფუძნებული ალგორითმი ხშირად **ამყარებს** არსებულ უთანასწორობას, ნაცვლად მისი გამოსწორებისა.
- **HireVue – სამუშაოზე მიღების AI-ი** HireVue, ვიდეო-გასაუბრების AI სისტემა, ფართოდ გამოიყენებოდა უნივერსიტეტის კურსდამთავრებულთა შესარჩევად. სისტემა აანალიზებდა სახის გამომეტყველებას, ხმის ტემბრს და სიტყვის არჩევანს. 2021 წელს კვლევებმა დაადასტურა, რომ სისტემა **სტრუქტურულ მიკერძოებას** შეიცავდა – ის უფრო ხელსაყრე-

ლი იყო ევროპული ფენოტიპის, ასევე ინტრავერტ პიროვნებებისთვის. შეზღუდული შესაძლებლობების მქონე კანდიდატები განსაკუთრებით ზარალდებოდნენ. FTC-ის (აშშ-ის ფედერალური სავაჭრო კომისია) ზეწოლის შემდეგ, სისტემამ სახის ანალიზი მიატოვა. **კავშირი განათლებასთან:** ამ ქეისი ადასტურებს, რომ ალგორითმული გადაწყვეტილებები სტუდენტთა **„მომავალ ბედს"** – კარიერულ შესაძლებლობებს – პირდაპირ ახდენს გავლენას.

სტატიაში „როგორ (და რატომ) აკვირდებიან საგანმანათლებლო ტექნოლოგიების კომპანიები სტუდენტების გრძნობებს" (Herold, 2018), ავტორი განიხილავს ხელოვნური ინტელექტის სისტემების მაგალითებს, რომლებიც ამოიცნობენ სტუდენტების ემოციებს საკლასო ოთახში სწავლის გასაუმჯობესებლად და მათი ჩართულობის დონის დასადგენად. ზოგიერთი ექსპერტი ასეთი სისტემების გამოყენებას უარყოფითად და არაეთიკურად მიიჩნევს: დაძაბული ატმოსფეროს პოტენციურმა შექმნამ (სასწავლო პროცესის დროს მუდმივი მეთვალყურეობა) შეიძლება გააუარესოს მონაწილეთა ფსიქიკური ჯანმრთელობა. ასეთი ანალიზისა და მონაცემთა შეგროვების მიზანია „სწავლის გაუმჯობესება პროგრამის სწავლებით, რათა ზუსტად ამოიცნოს, როდის არიან სტუდენტები ბედნიერები, ჩართულები ან მოწყენილნი".

თუმცა, ასეთი მონაცემთა შეგროვება ფსიქიკური ჯანმრთელობის შეფასებას ესაზღვრება და ამ შემთხვევაში სტუდენტები შეიძლება აღიქვან პოტენციურ პაციენტებად.

ხელოვნური ინტელექტის სისტემები, რომლებიც ადამიანების მიერ არ არის ინტერპრეტირებული, არ შეიძლება გამოყენებულ იქნას მაღალი რისკის შემცველი გადაწყვეტილებებისთვის და ზოგადი მონაცემთა დაცვის რეგულაციის (GDPR) მოთხოვნები ამას ნათლად ადასტურებს. ამიტომ, ევროპაში, დაზღვევის ტარიფებთან, სესხების გაცემასთან და სასამართლო გადაწყვეტილებებთან დაკავშირებით გადაწყვეტილებების მიღება წმინდა ხელოვნური ინტელექტით შეუძლებელია. ასეთი მოდელების სწავლება წმინდა აკადემიურ საქმიანობად რჩება.

ხელმისაწვდომია ხელოვნური ინტელექტი?

ხელოვნური ინტელექტის ეფექტურობა დამოკიდებულია იმ მოწყობილობების ხელმისაწვდომობაზე, რომლებზეც ალგორითმები განხორციელდება. მიუხედავად იმისა, რომ ინტერნეტი ყოველწლიურად უფრო და უფრო ხელმისაწვდომი ხდება, არსებობს რეგიონები (როგორც გლობალურად, ასევე საქართველოში) და გარემოებები, როდესაც მთელი ოჯახი იყენებს ერთ მობილურ მოწყობილობას. ასეთ შემთხვევებში, შეგვიძლია ვისაუბროთ როგორც არაზუსტად შეგროვებულ მონაცემებზე ასეთი მოსწავლეებითათვის, და ასევე სისტემაზე არათანაბარ წვდომაზე.

როგორ შეგვიძლია დავიცვათ კონფიდენციალურობა?

თუ განვიხილავთ ხელოვნური ინტელექტის გამოყენებას მთელი ცხოვრების მანძილზე სწავლის კონცეფციის განსახორციელებლად, როგორც მთელი ცხოვრების თანამგზავრის, პირადი ჭკვიანი ასისტენტის სახით, ასევე უნდა გავითვალისწინოთ პერსონალურ მონაცემებზე ასეთ მჭიდრო წვდომასთან დაკავშირებული რისკები.

რადგან ასეთი ასისტენტი ეყრდნობა მომხმარებლის შესახებ პერსონალური მონაცემების მუდმივ გაფართოებას, აუცილებელია ალგორითმების უსაფრთხოების, გამჭვირვალობის და მონაცემთა კონფიდენციალურობის საკითხების მოგვარება.

რა მონაცემები გროვდება? როგორ ინახება ისინი? ვის ეკუთვნის ისინი? გარდა ამისა, მონაცემთა შენახვის სისტემები შეიძლება გატეხილი ან გამოყენებული იყოს არაეთიკურად – ვინ არის პასუხისმგებელი მათ გამოყენებასა და დამუშავებაზე?

გამჭვირვალობა უმნიშვნელოვანესია: ხელოვნური ინტელექტი უნდა იქნას გამოყენებული იქ, სადაც შეგვიძლია დავინახოთ, თუ როგორ და რატომ მიიღო მან კონკრეტული გადაწყვეტილება.

მნიშვნელოვანია იმის უზრუნველყოფა, რომ ჩვენი არჩევანი, შეცდომები და ემოციები დაეხმაროს ხელოვნურ ინტელექტს განათლების უფრო ეფექტური გახადოს, გააცნობიეროს მისი პოტენციალი და გაგვაბედნიეროს.

დასკვნები

განათლებაში ხელოვნური ინტელექტის გამოყენებისას, ჩვენ სამი სფეროს გადაკვეთაზე ვდგავართ: დიდი მონაცემები, გამოთვლითი ტექნოლოგიები და საგანმანათლებლო ღირებულებები. ეთიკური საკითხები წინასწარ უნდა გადაიჭრას შეცდომებისა და მავნე გამოყენების პოტენციური რისკების შესამცირებლად.

წიგნში „ხელოვნური ინტელექტი განათლებაში“ (Holmes, Bialik, Fadel, 2019) ჩამოთვლილია AIED ტექნოლოგიების გამოყენებასთან დაკავშირებული მიმდინარე ეთიკური საკითხები:

- რა არის ეთიკურად მისაღები AIED ტექნოლოგიის კრიტერიუმები?
- როგორ მოქმედებს მოსწავლეთა ან სტუდენტების მიზნების, ინტერესებისა და ემოციების ცვალებადი ბუნება AIED გამოყენების ეთიკურობაზე?
- რა არის კერძო ორგანიზაციების (AIED პროდუქტის შემქმნელები) და საჯარო ორგანოების (სკოლები და უნივერსიტეტები, რომლებიც ატარებენ AIED კვლევას) ეთიკური ვალდებულებები;
- როგორ შეუძლიათ საგანმანათლებლო მონაწილეებს გააპროტესტონ ან გაასაჩივრონ ის, თუ როგორ არის მათი პირადი ინფორმაცია ნაჩვენები დიდ მონაცემთა ბაზებში?

- რა ეთიკურ შედეგებს იწვევს AIED გადაწყვეტილებების „ღრმა“ ჯაჭვის გაგების შეუძლებლობა (მაგალითად, მრავალშრიანი ნეირონული ქსელების გამოყენებისას).

ამ კითხვებზე პასუხი უნდა გაეცეს განათლებაში ხელოვნური ინტელექტის გამოყენების ფორმების შესახებ გადაწყვეტილებების შემუშავებაში, დანერგვასა და მხარდაჭერაში ჩართული ყველა დაინტერესებული მხარის წარმომადგენლებს შორის დიალოგის გზით.

YouTube რესურსები

1. **“Will AI Replace Teachers?”** – Kurzgesagt – In a Nutshell https://www.youtube.com/@kurzgesagt – მოიძიეთ: *“AI future education”* ვიზუალურად მიმზიდველი ვიდეო, რომელიც განიხილავს AI-ისა და ადამიანი მასწავლებლის როლებს მომავლის განათლებაში.
2. **“Maslow’s Hierarchy of Needs”** – Simply Psychology / Sprouts https://www.youtube.com/watch?v=reAuCdQxd4c მარტივი და ვიზუალური განმარტება მასლოუს თეორიის შესახებ – სასარგებლო ამ თავის კონტექსტის გასაგებად.
3. **“The Danger of AI is Weirder Than You Think”** – Janelle Shane, TED Talk https://www.youtube.com/watch?v=OhCzX0iLnOc TED-ის მომხსენებელი განიხილავს, თუ როგორ შეიძლება AI-მ „ისწავლოს“ ისე, რომ ადამიანის ლოგიკა და ეთიკა სრულად გამოტოვოს.
4. **“How Algorithms Shape Our World”** – Kevin Slavin, TED Talk https://www.youtube.com/watch?v=TDaFwnOiKVE ვიდეო ეხება ალგორითმების გავლენას საზოგადოებაზე, მათ შორის განათლებაზე, და კარგ საფუძველს ქმნის ალგორითმული მიკერძოების სადისკუსიოდ.

სადისკუსიო თემები

AI ნამდვილად ვერ შეცვლის მასწავლებელს – თუ ეს ილუზიაა? *M. Lynch ამტკიცებს, რომ ემპათია და კრეატიულობა AI-სთვის მიუწვდომელია. თუმცა:*

- *GPT-4 უკვე ადაპტირებს ტექსტს მკითხველის ემოციური მდგომარეობის მიხედვით. ეს ემპათიაა?*
- *შეიძლება „ტექნიკური ემპათია“ საკმარისი იყოს სასწავლო პროცესისთვის?*
- *რა კონკრეტული ფუნქციები **ვერასდროს** გადაეცემა AI-ს? დაასაბუთეთ.*

კოგნიტური „გადატვირთვა“ ტექნოლოგიაზე – კარგია თუ ცუდი?

- *ყოველი ახალი ტექნოლოგია (წერა, ბეჭდვა, კალკულატორი) ათავისუფლებს ადამიანის „გონებრივ სივრცეს“. ეს ყოველთვის „დეგრადაცია“ იყო?*
- *სად არის ზღვარი „გონებრივი სივრცის გათავისუფლებასა“ და „კოგნიტური უნარის დაკარგვას“ შორის?*
- *რომელი უნარების „გადატვირთვა“ AI-ზე მიგაჩნიათ მისაღებად სასკოლო სწავლაში, და რომელი – დაუშვებლად?*

ვინ იცავს სტუდენტს, როდესაც ალგორითმი ცდება? *UK-ის A-Level სკანდალის და HireVue-ს მაგალითების გათვალისწინებით:*

- *რამდენად სამართლიანია ალგორითმული გადაწყვეტილება სტუდენტის შეფასებაში, მიღება-ჩარიცხვაში ან გარიცხვაში?*
- *ვინ უნდა იყოს ანგარიშვალდებული – ალგორითმის შემქმნელი, სკოლა, თუ სამინისტრო?*
- *როგორ უნდა მოხდეს გასაჩივრება?*

ტესტი 13

1. რომელ წელს გამოაქვეყნა აბრაამ მასლოუმ ნაშრომი „ყოფიერების ფსიქოლოგია“?

ა) 1943
ბ) 1954
გ) 1968
დ) 1975

2. მასლოუს მოთხოვნილებების იერარქიის რომელ დონეს ეკუთვნის სიახლისა და ცოდნისადმი ბუნებრივი წინააღმდეგობა?

ა) ფიზიოლოგიური მოთხოვნილებები
ბ) უსაფრთხოების მოთხოვნილება
გ) პატივისცემის მოთხოვნილება
დ) თვითაქტუალიზაცია

3. M. Lynch-ის (2018) სტატიის მიხედვით, რომელი უნარი განასხვავებს მასწავლებელს AI-სგან?

ა) ინფორმაციის სწრაფი დამუშავება
ბ) ტესტების ავტომატური შემოწმება
გ) გრძნობა, თანაგრძნობა და ემპათია
დ) ონლაინ კურსების შექმნა

4. UNESCO-ს 2020 წლის დოკუმენტი „ხელოვნური ინტელექტი განათლებაში“ განიხილავს შემდეგ რისკს:

ა) ხელოვნური ინტელექტის ძვირადღირებულობა
ბ) კოგნიტური ფუნქციების დელეგირებით აზროვნების უნარის შესუსტება
გ) ინტერნეტ-დამოკიდებულება
დ) სახელმძღვანელოების გაუქმება

5. Holmes, Bialik და Fadel (2019) AIED-ის გამოყენებასთან დაკავშირებულ რომელ პრობლემას ასახელებენ?

ა) ტექნოლოგიის მაღალი ღირებულება
ბ) მოსწავლეების მონაწილეობის შემცირება საკუთარ სწავლაში
გ) პედაგოგების კვალიფიკაციის ამაღლება
დ) სახელმძღვანელოების ციფრულ ფორმატში გადაყვანა

6. Herold-ის (2018) სტატიის თანახმად, ზოგიერთი ხელოვნური ინტელექტის სისტემა საკლასო ოთახში:

ა) ანაცვლებს მასწავლებლებს
ბ) ამოიცნობს სტუდენტების ემოციებს
გ) ავტომატურად ანიჭებს შეფასებებს
დ) ქმნის სასწავლო გეგმებს

7. GDPR-ის მოთხოვნების მიხედვით, ევროპაში წმინდა ხელოვნური ინტელექტით შეუძლებელია:

ა) ტესტების შემოწმება
ბ) სასწავლო კონტენტის შექმნა
გ) სადაზღვევო ტარიფებთან, სესხებთან და სასამართლო გადაწყვეტილებებთან დაკავშირებული გადაწყვეტილებები
დ) ენის თარგმნა

8. რა პრობლემაა ხელოვნური ინტელექტის ხელმისაწვდომობასთან დაკავშირებით, განსაკუთრებით საქართველოს კონტექსტში?

ა) ინტერფეისის სირთულე
ბ) ენობრივი ბარიერი
გ) არათანაბარი წვდომა – ზოგ შემთხვევაში მთელი ოჯახი ერთ მოწყობილობას იყენებს
დ) ელექტრონული სახელმძღვანელოების ნაკლებობა

9. „განათლებაში ხელოვნური ინტელექტის" (Holmes et al., 2019) მიხედვით, რომელი სასწავლო პროცესების ავტომატიზაცია ყველაზე რთულია?

ა) ფაქტობრივი ინფორმაციის მიწოდება
ბ) ტესტების ავტომატური გენერირება
გ) ინფორმაციის რეფლექსია, ანალიზი, სინთეზი და შეფასება
დ) დავალებების ვადების მართვა

10. დასკვნებში „განათლებაში ხელოვნური ინტელექტი" (Holmes et al., 2019) ამტკიცებს, რომ AIED-ის ეთიკური კითხვები უნდა გადაიჭრას:

ა) მხოლოდ ტექნოლოგიური კომპანიების მიერ
ბ) მხოლოდ მთავრობის მიერ
გ) ყველა დაინტერესებული მხარის წარმომადგენლებს შორის დიალოგის გზით
დ) საერთაშორისო სასამართლოს მიერ

ბიბლიოგრაფია

1. წიგნები და მონოგრაფიები

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning.* Center for Curriculum Redesign.

ფუნდამენტური შესავალი ხელოვნური ინტელექტის გამოყენებაზე სასწავლო პროცესში. მოიცავს AI-ის ისტორიას განათლებაში, ეთიკურ გამოწვევებს და პრაქტიკულ სცენარებს.

Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education for the 21st Century.* UCL IOE Press.

ავტორი განიხილავს, თუ როგორ შეიძლება გაერთიანდეს ადამიანური და მანქანური ინტელექტი პერსონალიზებული სწავლების შესაქმნელად.

Selwyn, N. (2019). *Should Robots Replace Teachers? AI and the Future of Education.* Polity Press.

კრიტიკული პერსპექტივა AI-ის როლზე განათლებაში – ავტორი სვამს კითხვებს ავტომატიზაციის საზღვრებსა და პედაგოგის შეუცვლელ ფუნქციებზე.

Popenici, S. (2022). *Artificial Intelligence and Learning Futures: Critical Narratives of Technology, Democracy and Education.* Routledge.

კრიტიკული ანალიზი AI-ის გავლენისა განათლების სისტემებზე დემოკრატიული და ეთიკური კუთხით.

2. სამეცნიერო სტატიები და ჟურნალები

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(39).

სისტემატური მიმოხილვა 146 სტატიისა AI-ის გამოყენებაზე უმაღლეს განათლებაში; ამჟღავნებს კვლევის ძირითად სფეროებს და ხარვეზებს.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221.

კლასიკური კვლევა, რომელიც ადარებს ადამიანური და ინტელექტუალური სასწავლო სისტემების ეფექტურობას.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. *Pearson.*

Pearson-ის მოხსენება, რომელიც ასახავს AI-ის პოტენციალს პერსონალიზაციის, შეფასებისა და სასწავლო ანალიტიკის სფეროში.

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, 100001.

ახალი ჟურნალის პირველი გამოცემის საბაზისო სტატია, რომელიც AI-ის კვლევის მომავალ მიმართულებებს განსაზღვრავს.

Bhutoria, A. (2022). Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. *Computers and Education: Artificial Intelligence, 3*, 100068.

შედარებითი ანალიზი სამი ქვეყნის გამოცდილებაზე პერსონალიზებული სწავლებისა და AI-ის ინტეგრაციის კუთხით.

3. ანგარიშები და პოლიტიკის დოკუმენტები

UNESCO. (2021). *AI and Education: Guidance for Policy-Makers.* UNESCO Publishing.

UNESCO-ს სახელმძღვანელო პოლიტიკის შემქმნელებისთვის – AI-ის პასუხისმგებლიანი გამოყენება, ეთიკა და ინკლუზიურობა.

OECD. (2023). *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research.* OECD Publishing.

OECD-ის ანგარიში სამეცნიერო კვლევებში AI-ის გამოყენებაზე, განათლებასთან მჭიდრო კავშირში.

European Commission. (2022). *Ethical Guidelines on the Use of Artificial Intelligence and Data in Teaching and Learning for Educators.* Publications Office of the EU.

ევროკომისიის სახელმძღვანელო პრინციპები პედაგოგებისთვის – AI-ის ეთიკური გამოყენება სწავლებაში.

4. ონლაინ რესურსები და პლატფორმები

Coursera & edX AI for Education კურსები.

ვიდეო-ლექციები, ინტერაქტიული ამოცანები და სასერთიფიკატო პროგრამები გლობალური უნივერსიტეტების მიერ.

Khan Academy – Khanmigo. https://www.khanacademy.org/khan-labs

GPT-4-ზე დაფუძნებული AI-ასისტენტი, რომელიც სოკრატული მეთოდით ასწავლის მოსწავლეებს.

AI4K12 Initiative. https://ai4k12.org

ამერიკული ინიციატივა K-12 სასკოლო განათლებაში AI-ის შემოსაღებად – სასწავლო გეგმები და სახელმძღვანელოები.

5. ქართულენოვანი და რეგიონული წყაროები

საქართველოს განათლების სამინისტრო. (2023). *ციფრული განათლების სტრატეგია 2023–2030.*

სახელმწიფო სტრატეგია, რომელიც განსაზღვრავს ციფრული ტექნოლოგიების, მათ შორის AI-ის, ინტეგრაციის გზებს ქართულ საგანმანათლებლო სისტემაში.

LEPL – შეფასებისა და გამოცდების ეროვნული ცენტრი. https://naec.ge

ეროვნული შეფასების სისტემა, სადაც ადაპტური ტესტირებისა და AI-ის გამოყენების პოტენციალი განიხილება.

ტესტების პასუხები

№	ტესტი 1	ტესტი 2	ტესტი 3	ტესტი 4	ტესტი 5	ტესტი 6	ტესტი 7
1	ბ	ბ	ბ	გ	გ	გ	გ
2	გ	გ	გ	ბ	ბ	გ	ბ
3	გ	დ	ბ	გ	ბ	გ	გ
4	ბ	ბ	დ	გ	გ	ბ	ბ
5	გ	გ	ბ	გ	გ	ბ	ბ
6	ბ	გ	გ	გ	გ	გ	გ
7	გ	გ	დ	გ	ბ	გ	გ
8	დ	ბ	ბ	ბ	გ	გ	ბ
9	ბ	დ	გ	გ	გ	გ	გ
10	ბ	ბ	ბ	დ	გ	გ	დ

№	ტესტი 8	ტესტი 9	ტესტი 10	ტესტი 11	ტესტი 12	ტესტი 13
1	ბ	ბ	ბ	ბ	ბ	გ
2	ბ	გ	გ	ბ	გ	ბ
3	ბ	გ	გ	გ	ბ	გ
4	გ	ბ	გ	ბ	გ	ბ
5	გ	გ	გ	გ	დ	ბ
6	ბ	დ	გ	გ	ბ	ბ
7	ბ	გ	ბ	გ	გ	გ
8	გ	გ	ბ	გ	ბ	გ
9	გ	გ	გ	გ	ბ	გ
10	გ	ბ	გ	ბ	დ	გ