

ფსიქოლოგიაში ხელოვნური ინტელექტის გამოყენების 10 ძირითადი პრინციპი

დოკუმენტი მომზადებულია ფსიქოლოგთა გლობალური ალიანსის მიერ (იხ.
ორიგინალი ინგლისურ ენაზე)

თარგმანი მომზადებულია საქართველოს ფსიქოლოგთა პროფესიული კავშირის
მიერ

სარჩევი

<i>ფსიქოლოგიაში ხელოვნური ინტელექტის (AI) გამოყენების 10 ძირითადი პრინციპი.....</i>	<i>3</i>
ეთიკა, გამჭვირვალობა და ხელოვნური ინტელექტი	3
პროფესიული კომპეტენცია და ხელოვნური ინტელექტი.....	3
ხელმისაწვდომობა, მიკერძოება და ინკლუზია	4
<i>შესავალი.....</i>	<i>5</i>
<i>მეთოდოლოგია.....</i>	<i>5</i>
<i>პრინციპების განმარტება</i>	<i>8</i>
ეთიკა, გამჭვირვალობა და ხელოვნური ინტელექტი	8
პროფესიული კომპეტენცია და ხელოვნური ინტელექტი.....	13
ხელმისაწვდომობა, მიკერძოება და ინკლუზია	17
<i>გამოყენებული ტერმინები.....</i>	<i>23</i>

ფსიქოლოგიაში ხელოვნური ინტელექტის (AI) გამოყენების 10 ძირითადი პრინციპი

ეთიკა, გამჭვირვალობა და ხელოვნური ინტელექტი

პრინციპი 1: ეთიკური პასუხისმგებლობა - ფსიქოლოგები, ხელოვნური ინტელექტის გამოყენებისას, უნდა ხელმძღვანელობდნენ პროფესიაში აღიარებული ეთიკური პრინციპების დაცვით.

პრინციპი 2: გამჭვირვალობა - ხელოვნური ინტელექტის ინსტრუმენტების შერჩევისას, ფსიქოლოგებმა უნდა გამოიყენონ მხოლოდ ისეთი AI ინსტრუმენტები, რომელთა წარმომადგენლობა და მოქმედების მექანიზმი ნათელია და ცხადად აწვდის მომხმარებელს ინფორმაციას იმასთან დაკავშირებით, რომ ის ხელოვნური ინტელექტთან (არა ადამიანთან) ურთიერთობს. თუ მსგავსი გამჭვირვალობის უზრუნველყოფა ვერ ხერხდება, ფსიქოლოგმა უნდა მოიძიოს ყველა რელევანტური დეტალი და შეზღუდვები დოკუმენტურად აღწეროს.

პრინციპი 3: ინფორმირებული თანხმობა - პროფესიული საქმიანობის ფარგლებში AI ინსტრუმენტების გამოყენებისას, ფსიქოლოგებმა უნდა აცნობონ მომსახურების მიმღებ პირებს შესაძლო რისკებისა და სარგებლის შესახებ. მხარეების მიერ გაჟღერებული თანხმობა უნდა იყოს მკაფიო და ნებაყოფლობითი (opt-in). დაუშვებელია, თანხმობა მანიპულაციური გზებით იქნას მოპოვებული.

პრინციპი 4: ანგარიშვალდებულება და პროფესიული პასუხისმგებლობა - ფსიქოლოგებს, რომლებიც იყენებენ ან ქმნიან AI სისტემებს, ეკისრებათ პასუხისმგებლობა ამ ინსტრუმენტების დანიშნულებისამებრ გამოყენებასა და გამოყენების შედეგებზე. შესაბამისად, მათ ხელი უნდა შეუწყონ ინსტრუმენტების ცხადი და პასუხისმგებლიანი მოხმარების პრაქტიკას.

პროფესიული კომპეტენცია და ხელოვნური ინტელექტი

პრინციპი 5: მტკიცებულებაზე დაფუძნება - ფსიქოლოგებს, რომლებიც პროფესიულ პრაქტიკაში ხელოვნური ინტელექტის ინტეგრირებას გეგმავენ, უნდა ჰქონდეთ საკმარისი ცოდნა AI ინსტრუმენტების ფუნქციონირების შესახებ. მათ შორის, აუცილებელია, იცოდნენ ზოგადი მეთოდოლოგიური საკითხები, მონაცემთა გამოყენებასთან დაკავშირებული მოთხოვნები და შეზღუდვები, რომლებიც თან ახლავს ხელოვნური ინტელექტის ფსიქოლოგიური მიზნებისთვის გამოყენებას.

პრინციპი 6: ზედამხედველობა ადამიანის მხრიდან და პროფესიული მსჯელობა - ფსიქოლოგებს, რომლებიც თავიანთ პროფესიულ საქმიანობაში იყენებენ

ხელოვნური ინტელექტის (AI) ინსტრუმენტებს, უნდა ჰქონდეთ საკმარისი კომპეტენცია და პროფესიული განსჯის უნარი, რათა სათანადოდ განკარგონ AI-ის მხარდაჭერით მიღებული მონაცემები.

პრინციპი 7: ფსიქოლოგიური უსაფრთხოება და ზიანის პრევენცია - ფსიქოლოგებმა, რომლებიც იყენებენ ხელოვნური ინტელექტის (AI) სისტემებს, პრიორიტეტი უნდა მიანიჭონ მაქსიმალური სარგებლის მიღწევას, ფსიქოლოგიურ უსაფრთხოებასთან ერთად და აქტიურად იმუშაონ შესაძლო ზიანის პრევენციაზე.

ხელმისაწვდომობა, მიკერძოება და ინკლუზია

პრინციპი 8: ვალიდაცია და სამეცნიერო მტკიცებულება - ფსიქოლოგებმა, რომლებიც იყენებენ ხელოვნური ინტელექტის (AI) ინსტრუმენტებს, კრიტიკულად უნდა შეაფასონ, ეფუძნება თუ არა ასეთი სისტემები მკაცრ, გამჭვირვალე და კონტექსტის შესაბამის სამეცნიერო მტკიცებულებებს, რომლებიც ადასტურებს მათ სანდოობას, ვალიდობას და შესაბამისობას იმ კონტექსტთან და მიზნებთან, რისთვისაც ისინი გამოიყენება.

პრინციპი 9: ალგორითმული მიკერძოება და სამართლიანობა - ფსიქოლოგებმა, რომლებიც იყენებენ ხელოვნური ინტელექტის (AI) სისტემებს, პროაქტიულად უნდა გამოავლინონ, შეაფასონ და შეამცირონ ალგორითმული მიკერძოება, რათა დარწმუნდნენ, რომ AI-ის მხარდაჭერით მიღებული შედეგები არ ახდენს ინდივიდებისა თუ ჯგუფების დისკრიმინაციას დემოგრაფიული, კულტურული, სოციოეკონომიკური ან კონტექსტთან დაკავშირებული სხვა ფაქტორების საფუძველზე.

პრინციპი 10: ხელმისაწვდომობა, თანასწორობა და ინკლუზია - ფსიქოლოგები, რომლებიც პროფესიულ საქმიანობაში იყენებენ AI ინსტრუმენტებს, უნდა დარწმუნდნენ, რომ მომსახურება თანაბრად ხელმისაწვდომი და ინკლუზიურია ყველასთვის და არ აღრმავებს უკვე არსებულ უთანასწორობას, არ რიყავს ინდივიდებსა და ჯგუფებს სტრუქტურული, ეკონომიკური, ენობრივი ან ტექნოლოგიური ბარიერების გამო.

შესავალი

COVID-19-ის პანდემიის შემდეგ, ფსიქიკური ჯანმრთელობის მიმართ გაზრდილმა ინტერესმა და ასევე, ციფრული ინსტრუმენტების მოხმარების სიხშირემ, განაპირობა ფსიქოლოგიაში ტექნოლოგიის გამოყენებასთან დაკავშირებული ადამიანური ფაქტორების გააზრების საჭიროება.

ტექნოლოგიების მზარდი გავრცელება და ფსიქოლოგიურ კეთილდღეობასა და ფსიქიკურ ჯანმრთელობაზე ორიენტირებული ციფრული პლატფორმების, ინსტრუმენტებისა და აპლიკაციების მომრავლება კიდევ ერთხელ აჩვენებს, რომ უზრუნველყოფილი უნდა იყოს როგორც ამ ინსტრუმენტების ეფექტიანობა, ისე მათი უსაფრთხოება. ციფრულ სამყაროსთან მოქარბებულმა ინტერაქციამ, გააღრმავა დისკუსიები ფსიქოლოგიური კეთილდღეობის შესახებ და წარმოშვა ახალი ეთიკური გამოწვევები ახალი ტექნოლოგიების, განსაკუთრებით ხელოვნური ინტელექტის (AI) პასუხისმგებლიანი გამოყენების შესახებ.

მოცემული დოკუმენტი წარმოადგენს პასუხს AI-ის სწრაფ განვითარებასა და ფსიქოლოგიის სფეროში შემოსვლაზე, რაც სახეს უცვლის ფსიქოლოგების საქმიანობას. ეს გზამკვლევი მოიცავს კრიტიკულად მნიშვნელოვანი ინფორმაციისა და პროფესიული ვალდებულებების მიმოხილვას იმ ფსიქოლოგებისთვის, რომლებსაც სურთ AI-ის ინტეგრირება ფსიქიკური ჯანმრთელობის სერვისებში. ის აერთიანებს ხელოვნური ინტელექტის მოხმარებასთან დაკავშირებულ ეთიკის პრინციპებს, კლინიკურ პრაქტიკაში მისი გამოყენების საკითხებს, ასევე, ეხება ადამიანის მიერ ხელოვნური ინტელექტის განკარგვასა და თანაბარი ხელმისაწვდომობის ხელშეწყობის პრინციპებსა და მიკერძოებისა თუ პროფესიული კომპეტენციის გააზრების მნიშვნელობას.

მნიშვნელოვანია, ასევე, აღინიშნოს, რომ ხელოვნური ინტელექტის განვითარების, ხელმისაწვდომობისა თუ კულტურული რელევანტურობის თვალსაზრისით, არსებობს გარკვეული უთანასწორობა. ფსიქოლოგები აღიარებენ, რომ უთანასწორო ციფრული ინფრასტრუქტურა, ეკონომიკური ფაქტორები და კულტურული კონტექსტები შეიძლება ზღუდავდეს საზოგადოების დიდი ნაწილის შესაძლებლობას, ისარგებლოს AI-ზე დაფუძნებული ფსიქოლოგიური ინსტრუმენტებით. შესაბამისად, ფსიქოლოგიურ პრაქტიკაში AI-ის ეთიკური ინტეგრაციის თვალსაზრისით, ცენტრალურ საკითხს თანაბარად ხელმისაწვდომი და კულტურულად რელევანტური მიდგომის გამოყენება წარმოადგენს.

მეთოდოლოგია

მოცემული დოკუმენტი ეყრდნობა კონცეპტუალურ და ნორმატიულ მიდგომას და არა ემპირიულ მონაცემებს. მასში გაერთიანებულია მოქმედი ეთიკური კოდექსები

(მაგ., American Psychological Association, 2017; British Psychological Society, 2021; Canadian Code of Ethics for Psychologists, 2017), ადამიანის უფლებებისა და მონაცემთა დაცვის საერთაშორისო სტანდარტები (მაგ., Europe's General Data Protection Regulation, U.S. Health Insurance Portability and Accountability Act, World Health Organization guidance) და ფსიქოლოგიაში, ციფრული ჯანმრთელობისა და ხელოვნური ინტელექტის შესახებ არსებული სამეცნიერო ლიტერატურა.

დოკუმენტში განხილული 10 პრინციპი, რომელიც სამ კატეგორიაში ერთიანდება (ეთიკა; პროფესიული კომპეტენცია; ხელმისაწვდომობა, მიკერძოება და ინკლუზია), ეყრდნობა ფსიქოლოგთა გლობალური ალიანსის წევრი ფსიქოლოგების მიერ დაფიქსირებულ მოსაზრებათა თემატურ ანალიზს (მსოფლიოს სხვადასხვა რეგიონში მცხოვრები 20 ფსიქოლოგის მიერ შევსებული კითხვარი) და პასუხობს ციფრული ტექნოლოგიების განვითარების კვალდაკვალ, სწრაფად ცვალებად გარემოში არსებულ გამოწვევებს.

მოცემული დოკუმენტი რამდენიმე მიმართულებას აერთიანებს: ის ხაზს უსვამს AI-ზე დაფუძნებული და ციფრული ფსიქოლოგიური ინსტრუმენტების მკაცრი და კულტურულად რელევანტური ვალიდაციის საჭიროებას, რათა თავიდან ავიცილოთ გადაუმოწმებელი ან მიკერძოებული ალგორითმებით გამოწვეული ზიანი; ამასთან, ის ხაზს უსვამს ფსიქოლოგთა საბაკალავრო, სამაგისტრო თუ სადოქტორო საქმიანობასა და უწყვეტი პროფესიული განვითარების პროცესში ციფრული ეთიკის, AI-ის ცოდნის და კრიტიკული აზროვნების ინტეგრაციის აუცილებლობას; დოკუმენტში ასევე აღნიშნულია, რომ ფსიქიკური ჯანმრთელობის სერვისების მიმწოდებლებმა ხელოვნური ინტელექტი არ უნდა გამოიყენონ მხოლოდ როგორც მომხმარებლებმა, არამედ აქტიურად უნდა აჩვენონ ციფრული ინსტრუმენტების ეთიკური გამოყენების მაგალითი, ინფორმირებული თანხმობის, კონფიდენციალურობისა და ადამიანის ღირსების დაცვის უზრუნველყოფით.

ფსიქოლოგიის მიღმა, მოცემული დოკუმენტი შეიძლება გამოყენებულ იქნას საერთაშორისო სახელმძღვანელოების, ეროვნული რეგულაციებისა და ციფრული პლატფორმების დიზაინის ფორმირებისთვის, რაც საბოლოოდ, ხელს შეუწყობს ფსიქიკური ჯანმრთელობის მართვის თანმიმდევრული ციფრული სისტემის ჩამოყალიბებას.

მიუხედავად იმისა, რომ ფსიქიკური ჯანმრთელობის სერვისებში ხელოვნური ინტელექტის გამოყენების შესახებ კვლევების რაოდენობა გაიზარდა, ისეთი სახელმძღვანელოების დეფიციტი კვლავაც არსებობს, რომლებიც ფსიქოლოგებს ეთიკური პრაქტიკის, შესაძლო მიკერძოებისგან თავის არიდების, ინკლუზიური მიდგომების დანერგვისა და პროფესიული კომპეტენციის გაზრდის საკითხებში დაეხმარებოდა. აღნიშნული დოკუმენტი ავსებს ამ სიცარიელეს, მოიცავს რა

პრინციპებს, რომლებიც შექმნილია ფსიქოლოგთა გლობალური ალიანსის ფარგლებში გაერთიანებული ფსიქოლოგების მიერ, 80-ზე მეტი ქვეყნიდან.

შემაჯამებელი შენიშვნები

ფსიქოლოგიაში ხელოვნური ინტელექტის (AI) გამოყენების საუკეთესო პრაქტიკა მოითხოვს, რომ ფსიქოლოგები მოქმედებდნენ კანონმდებლობის, პროფესიული რეგულაციების, ეთიკის კოდექსებისა და მათ პროფესიულ კონტექსტში არსებული საუკეთესო პრაქტიკის გათვალისწინებით. გზამკვლევი მიზნად ისახავს აღნიშნული სტანდარტების გაძლიერებას და არა მათ ჩანაცვლებას ან გადაფარვას. ფსიქოლოგიაში AI-ის უსაფრთხო, ეთიკური და ეფექტიანი გამოყენება დამოკიდებულია, როგორც ინდივიდუალურ პრაქტიკოსებზე, ასევე, დაინტერესებულ მხარეებზე (მათ შორის, და არა მხოლოდ, AI დეველოპერებზე, ინსტიტუციებსა და მარეგულირებელ მხარეებზე). ეს პროცესი იმთავითვე მრავალი დაინტერესებული მხარის არსებობას მოიაზრებს და მათ კოორდინირებულ მოქმედებას საჭიროებს. ერთიანი მოქმედება, ინტერდისციპლინური თანამშრომლობა და უწყვეტი დიალოგი აუცილებელია AI სისტემების დაგეგმვის, დანერგვის, მონიტორინგისა და განახლებისთვის საჭირო ისეთი გზების შემუშავებისას, რომლებიც იცავს კლიენტებს, ქმნის ფსიქოლოგებისთვის პროფესიულ ფარს და ამყარებს პროფესიულ ღირებულებებს, ზრდის ფსიქოლოგიის მიმართ საზოგადოების ნდობას და ამავედროულად, შესაძლებელს ხდის ფსიქოლოგიურ პრაქტიკაში ინოვაციების ინტეგრირებას.

რედაქტორის კომენტარი. დოკუმენტის დასაწყისში შეჯამებულია ფსიქოლოგიაში ხელოვნური ინტელექტის გამოყენების 10 ძირითადი პრინციპი. მომდევნო გვერდებზე შეხვდებით თითოეული პრინციპის დეტალურ განმარტებასა და განხილვას.

გამოყენებული ლიტერატურა

- American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>
- British Psychological Society. (2021). Code of ethics and conduct. <https://explore.bps.org.uk/content/report-guidelines/bpsrep.2021.inf94>
- Canadian Psychological Association. (2017). Canadian code of ethics for psychologists. https://cpa.ca/doc/File/Ethics/CPA_Code_2017_4thEd.pdf

პრინციპების განმარტება

ეთიკა, გამჭვირვალობა და ხელოვნური ინტელექტი

პრინციპი 1 - ეთიკური პასუხისმგებლობა - ხელოვნური ინტელექტის გამოყენებისას, ფსიქოლოგები უნდა ხელმძღვანელობდნენ პროფესიაში აღიარებული ეთიკური პრინციპების დაცვით.

განმარტება

ხელოვნური ინტელექტის (AI) განვითარება და გამოყენება უნდა ეფუძნებოდეს ეთიკურ პრინციპებს, რომლებიც პრიორიტეტს ანიჭებს პიროვნებისა და საზოგადოების უფლებებისა და ღირსების პატივისცემას, ადამიანის კეთილდღეობას, სამართლიანობასა და პასუხისმგებლობას. ფსიქოლოგები უნდა იცნობდნენ და პრაქტიკაში იყენებდნენ კვლევებს, რომლებიც გვაწვდის ინფორმაციას ადამიანის მიერ გადაწყვეტილების მიღების პროცესის, კოგნიტური მიკერძოებების და სოციალური ქცევის შესახებ, რამდენადაც ეს ყველაფერი დაკავშირებულია ხელოვნურ ინტელექტთან და მისი გამოყენების მიზანშეწონილობასა თუ დაუშვებლობასთან. ფსიქოლოგებს უნდა ესმოდეთ, თუ როგორ შეიძლება AI სისტემები შემუშავდეს ისე, რომ ისინი შეესაბამებოდეს სოციალურ ღირებულებებს და თავიდან იქნას აცილებული უნებლიე ზიანი.

კავშირი ფსიქოლოგიასთან

ფსიქოლოგებისთვის ეს პრინციპი რელევანტურია, რადგან ადამიანის ქცევისა და ეთიკის შესახებ მათი ექსპერტული ცოდნა საშუალებას იძლევა არა მხოლოდ ამოიცნონ და გამოიყენონ, არამედ მონაწილეობაც მიიღონ ისეთი AI სისტემების შემუშავებაში, რომლებიც პატივს სცემს ადამიანის ღირსებას და ამავედროულად, აძლიერებს ფსიქიკური ჯანმრთელობის გაუმჯობესებისკენ მიმართულ ძალისხმევას. ამისათვის, ფსიქოლოგებმა, პირველ რიგში, უნდა გააცნობიერონ, მათსავე პროფესიულ კონტექსტში AI-ის გამოყენებასთან დაკავშირებული ეთიკური გამოწვევების მნიშვნელობა. ფსიქოლოგებს შეუძლიათ გასცენ რეკომენდაციები ხელოვნური ინტელექტის გამოყენებასთან დაკავშირებული ისეთი პოტენციური ფსიქოლოგიური გავლენების შესახებ, როგორებიცაა მომხმარებლის ნდობაზე, დამოკიდებულებაზე ან გადაწყვეტილების მიღებაზე ზემოქმედება და დარწმუნდნენ, რომ ეთიკური დამცავი მექანიზმები ასახავს რეალურ სამყაროში არსებულ ამ გამოწვევებს. ჯანდაცვის სფეროში ხელოვნური ინტელექტის გამოყენება უნდა ეფუძნებოდეს როგორც უშუალოდ ხელოვნურ ინტელექტთან დაკავშირებულ ეთიკურ სახელმძღვანელო პრინციპებს, ისე ფსიქოლოგიური პრაქტიკის ტრადიციულ გზამკვლევებს. ფსიქოლოგებს პასუხისმგებლობა

ეკისრებათ, ინფორმირებულები იყვნენ იმ ეთიკური გამოწვევებისა და შეზღუდვების შესახებ, რაც შესაძლოა თან ახლდეს ხელოვნური ინტელექტის გამოყენებას და გადადგან ნაბიჯები მათ აღმოსაფხვრელად.

გამოყენებული ლიტერატურა:

- American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. <https://www.apa.org/topics/artificial-intelligence-machine-learning/ethicalguidance-professional-practice.pdf>
- European Federation of Psychologists' Associations. (n.d.) Digitalisation priorities. <https://www.efpa.eu/digitalisation>
- Resnik, D.B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI and Ethics*, 5, 1499–1521. <https://doi.org/10.1007/s43681-024-00493-8>
- Wu, J. (2024, November 4). Social and ethical impact of emotional AI advancement: The rise of pseudo-intimacy relationships and challenges in human interactions. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1410462>

პრინციპი 2 - გამჭვირვალობა ხელოვნური ინტელექტის ინსტრუმენტების შერჩევისას

ფსიქოლოგებმა უნდა გამოიყენონ ხელოვნური ინტელექტის მხოლოდ ისეთი ინსტრუმენტები, რომელთა წარმომადგენლობა და მოქმედების მექანიზმი ნათელია და მომხმარებელს აწვდის მკაფიო და უწყვეტ ინფორმაციას იმის თაობაზე, რომ ის ხელოვნურ ინტელექტთან (არა ადამიანთან) ურთიერთობს. თუ მსგავსი გამჭვირვალობის უზრუნველყოფა ვერ ხერხდება, ფსიქოლოგებმა უნდა აღრიცხონ მათთვის ცნობილი შეზღუდვები და მოიძიონ ხელმისაწვდომი ტექნიკური დოკუმენტაცია მონაცემთა დამუშავების პრაქტიკის შესახებ.

განმარტება

გამჭვირვალობის პრინციპი ხელოვნური ინტელექტის (AI) გამოყენებაში, გულისხმობს, რომ სისტემამ უნდა წარმოადგინოს მკაფიო ახსნა-განმარტება იმის შესახებ, თუ რა მონაცემებს იყენებს, როგორ გროვდება ეს მონაცემები და რა პროცესები უდევს საფუძვლად ალგორითმულ გადაწყვეტილებებს. ასეთი სიცხადის გარეშე, AI ინსტრუმენტების გამოყენება ქმნის რისკს, რომ მომხმარებლებისა და პროფესიონალებისთვის გართულდეს სიზუსტის, სამართლიანობისა და ანგარიშვალდებულების შეფასება. ხელოვნური ინტელექტის გამჭვირვალობის უზრუნველყოფა ინსტრუმენტებს შეუძლია შეამციროს არსებული განსვლა რთულ ტექნიკურ პროცესებსა და მომხმარებლის მიერ მათ გაგებას შორის.

კავშირი ფსიქოლოგიასთან

ფსიქოლოგებისთვის გამჭვირვალობა კრიტიკულად მნიშვნელოვანია როგორც კვლევით, ისე გამოყენებით საქმიანობაში. ფსიქოლოგებს უნდა შეეძლოთ იმის გაგება, თუ როგორ წარმოქმნის ხელოვნური ინტელექტი შედეგებს, რათა შეაფასონ მათი ვალიდურობა და სანდოობა კლინიკურ, ორგანიზაციულ თუ სამეცნიერო კონტექსტში. გარდა ამისა, გამჭვირვალობა ფსიქოლოგს საშუალებას აძლევს, ხელოვნური ინტელექტის მიერ წარმოებული პროდუქტის არსი და წარმომავლობა ნათლად აუხსნას კლიენტს, რაც ხელს უწყობს პროფესიონალსა და კლიენტს შორის ნდობისა და პროფესიული კეთილსინდისიერების შენარჩუნებას.

გამოყენებული ლიტერატურა:

Resnik, D. B., & Hosseini, M. (2026). Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? *Accountability in Research*, 33(2).

პრინციპი 3. ინფორმირებული თანხმობა

ფსიქოლოგებმა, რომლებიც პროფესიულ საქმიანობაში ხელოვნურ ინტელექტზე დაფუძნებულ ინსტრუმენტებს იყენებენ, მკაფიოდ უნდა აცნობონ მომსახურების მიმღებ მხარეებს პოტენციური რისკებისა და სარგებლის შესახებ. მომსახურების მიმღები პირების მიერ გაჟღერებული თანხმობა, AI ინსტრუმენტების გამოყენების თანმხლები შეზღუდვებისა თუ ინდივიდუალური მონაცემების გამოყენების შესახებ, უნდა იყოს მკაფიო და ნებაყოფლობითი (opt-in). თანხმობა არ უნდა იყოს მანიპულაციური გზებით მოპოვებული.

განმარტება

ხელოვნური ინტელექტის ინსტრუმენტების დანერგვა კვლევით, სასწავლო, საგანმანათლებლო გარემოში ან ფსიქოლოგიურ პრაქტიკაში საჭიროებს ინფორმირებული თანხმობის მიღებას იმ შემთხვევაში, როდესაც ხდება მონაცემების შეგროვება, შენახვა ან ანალიზი. ამასთან, მომსახურების მიმღებ პირებს (კლიენტებს, სტუდენტებს, თუ სხვა მხარეებს) უნდა ჰქონდეთ ინფორმაცია იმ AI ინსტრუმენტების შესახებ, რომლებსაც ფსიქოლოგი გადაწყვეტილების მიღებისას თუ მომსახურების მიწოდების გასაიოლებლად იყენებს.

ხელოვნური ინტელექტის გამოყენებისას, ინფორმირებული თანხმობა გულისხმობს იმის უზრუნველყოფას, რომ ვისი მონაცემებიც გამოიყენება იმ ინდივიდებმა, ისევე როგორც მათ, ვინც იყენებს AI სისტემებს, იცოდნენ, თუ როგორ ხდება ინფორმაციის შეგროვება, დამუშავება და გამოყენება. გარდა ამისა, როდესაც AI ინსტრუმენტები ინერგება პროფესიულ პრაქტიკაში, აუცილებელია კლიენტების

ინფორმირება გამჭვირვალობისა და ავტონომიის შესანარჩუნებლად. ეს პრინციპი უკავშირდება ინდივიდის უფლებებისა და თავისუფლების პატივისცემას.

კავშირი ფსიქოლოგიასთან

ფსიქოლოგებისთვის ინფორმირებული თანხმობა თავისთავად წარმოადგენს ეთიკური პრაქტიკის ქვაკუთხედს, ხოლო ხელოვნური ინტელექტი (AI) ამ პასუხისმგებლობას ახალ განზომილებებს სძენს. ფსიქოლოგებმა კლიენტებთან მკაფიოდ უნდა განმარტონ AI ინსტრუმენტების როლი მათ მიერ განხორციელებული შეფასების, ინტერვენციისა თუ მონაცემთა დამუშავების პროცესში, რათა დარწმუნდნენ, რომ კლიენტებს შეუძლიათ მონაწილეობის შესახებ გააზრებული გადაწყვეტილების მიღება. ეს პრინციპი უზრუნველყოფს კლიენტის ნდობის შენარჩუნებას და ეხმარება ფსიქოლოგებს დარჩნენ პროფესიული ეთიკის სტანდარტებთან შესაბამისობაში.

გამოყენებული ლიტერატურა:

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>
European Federation of Psychologists' Associations AISBL. (2025). Meta-code of ethics. <https://www.efpa.eu/meta-code-ethic>

პრინციპი 4. ანგარიშვალდებულება და პროფესიული პასუხისმგებლობა

ფსიქოლოგებს, რომლებიც იყენებენ ან ქმნიან AI სისტემებს, ეკისრებათ პასუხისმგებლობა ამ ინსტრუმენტების დანიშნულებისამებრ გამოყენებასა და გამოყენების შედეგებზე. შესაბამისად, მათ ხელი უნდა შეუწყონ ინსტრუმენტების ცხადი და პასუხისმგებლიანი მოხმარების პრაქტიკას.

განმარტება

ფსიქოლოგიაში AI-ის გამოყენება არ ნიშნავს, რომ ეთიკური ან პროფესიული პასუხისმგებლობა ხელოვნურ ინტელექტზე გადადის. რა ხარისხითაც არ უნდა ხდებოდეს ხელოვნური ინტელექტის გამოყენება, ფსიქოლოგები ინარჩუნებენ პასუხისმგებლობას იმ გადაწყვეტილებებზე, რომლებიც AI ინსტრუმენტების გამოყენებით მიიღეს.

AI სისტემების შექმნასა და ფუნქციონირებაში ხშირად მრავალი დაინტერესებული მხარეა ჩართული, მათ შორის დეველოპერები, მომწოდებლები, ინსტიტუციები და პოლიტიკის განმსაზღვრელები, რაც ზიანის შემთხვევაში პასუხისმგებელი პირის იდენტიფიცირებას ართულებს. შესაბამისად, ფსიქოლოგების მიზანი უნდა იყოს იმის გარკვევა, თუ როგორ მუშაობს AI სისტემა, ვინ აკონტროლებს მის გამოყენებას

და როგორ ხდება გადაწყვეტილებების მიღება და მათი გადახედვა. ეს მოითხოვს არა მხოლოდ იმის ცოდნას, თუ რას აკეთებს AI სისტემა, არამედ იმის გააზრებასაც, თუ როგორ და რატომ წარმოქმნის ის კონკრეტულ შედეგებს.

ფსიქოლოგებმა უნდა წახალისონ ისეთი სისტემების ფორმირება, რომლებიც გამჭვირვალეა, იძლევა დოკუმენტაციის წარმოების, აუდიტისა და ზედამხედველობის შესაძლებლობას და სადაც არსებობს გასაჩივრების მექანიზმები. ეს ასევე გულისხმობს მზადყოფნას, რომ აუხსნან AI-ით მხარდაჭერილი გადაწყვეტილებები კლიენტებს, მონაწილეებს ან სხვა დაინტერესებულ მხარეებს, და საჭიროების შემთხვევაში, თუკი პროფესიული მსჯელობა პრობლემის არსებობაზე მიუთითებს, შეცვალონ AI-ის მიერ მხარდაჭერილი გადაწყვეტილებები.

ინდივიდუალურ და ორგანიზაციულ დონეზე, პასუხისმგებლობის განსაზღვრა ხელს უწყობს ხელოვნურ ინტელექტზე გადაჭარბებული დამოკიდებულების თავიდან აცილებას და ამყარებს პროფესიული პასუხისმგებლობის მნიშვნელობას.

კავშირი ფსიქოლოგიასთან

პასუხისმგებლობა უმნიშვნელოვანესია ფსიქოლოგიური მეცნიერებისა და პრაქტიკის მიმართ საზოგადოებრივი ნდობის შესანარჩუნებლად. როდესაც AI ინსტრუმენტები მაკონტროლებელი მექანიზმების გარეშე ინერგება, პასუხისმგებელი პირის იდენტიფიცირება რთულდება, რაც ზრდის ისეთი ეთიკური დარღვევების, შეცდომების ან ზიანის რისკს, რომელსაც შესაბამისი რეაგირება არ მოჰყვება.

ფსიქოლოგებისთვის, რომლებიც სხვადასხვა გარემოში მუშაობენ (სადაც განსხვავებული მარეგულირებელი მექანიზმები მოქმედებს), პასუხისმგებლობაზე ფოკუსი უზრუნველყოფს პროფესიულ სტანდარტებთან, სამართლებრივ ვალდებულებებსა და ეთიკურ კოდექსებთან შესაბამისობას. ეს პრინციპი ხაზს უსვამს, რომ AI წარმოადგენს ინსტრუმენტს ფსიქოლოგიურ პრაქტიკაში და არა პროფესიული პასუხისმგებლობის, შეფასებისა თუ ეთიკური მსჯელობის ჩამნაცვლებელ მექანიზმს.

გამოყენებული ლიტერატურა:

- American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>
- European Commission. (2019). Ethics guidelines for trustworthy AI.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society.

პროფესიული კომპეტენცია და ხელოვნური ინტელექტი

პრინციპი 5. მტკიცებულებაზე დაფუძნება

ფსიქოლოგებს, რომლებიც პროფესიულ პრაქტიკაში ხელოვნური ინტელექტის ინტეგრირებას გეგმავენ, უნდა ჰქონდეთ საკმარისი ცოდნა AI ინსტრუმენტების ფუნქციონირების შესახებ, მათ შორის უნდა იცოდნენ ზოგადი მეთოდოლოგიური საკითხები, მონაცემებთან დაკავშირებული მოთხოვნები და შეზღუდვები, რომელიც ფსიქოლოგიური მიზნებისთვის ხელოვნური ინტელექტის გამოყენებას ახლავს თან.

განმარტება

ფსიქოლოგებს უნდა ჰქონდეთ საკმარისი საბაზისო ცოდნა, რათა კრიტიკულად შეაფასონ AI ინსტრუმენტების ფუნქციონირების მექანიზმი; მათ შორის, ასეთი სისტემების საფუძვლად მდებარე თეორიული დაშვებები, მონაცემთა წყაროები და მეთოდოლოგიური შეზღუდვები. აქ ასევე იგულისხმება იმ AI სისტემებთან დაკავშირებული ფსიქოლოგიური, გამოთვლითი და სტატისტიკური პრინციპების ცოდნა, რომლებსაც ისინი ირჩევენ და იყენებენ პროფესიულ საქმიანობაში.

ამ კონტექსტში, ცოდნა არ ნიშნავს AI სისტემების შექმნის ტექნიკურ ექსპერტიზას, არამედ გულისხმობს ინფორმირებული და გააზრებული გადაწყვეტილებების მიღების უნარს იმასთან დაკავშირებით, თუ როდის და როგორ არის საჭირო ან საერთოდ არის თუ არა AI ინსტრუმენტების გამოყენება საჭირო და რელევანტური კონკრეტული პროფესიული მიზნებისთვის. ფსიქოლოგებმა ყურადღება უნდა მიაქციონ AI-ით მხარდაჭერილი პრაქტიკის ეფექტიანობის, რისკებისა და შეზღუდვების შესახებ ჩატარებული თანამედროვე ემპირიული კვლევების შედეგებს და ჩართული უნდა იყვნენ უწყვეტი პროფესიული განვითარების პროცესში, რათა მათი მხრიდან AI-ის გამოყენება შეესაბამებოდეს განვითარებად სამეცნიერო ცოდნასა და პროფესიულ სტანდარტებს.

კავშირი ფსიქოლოგიასთან

ადამიანური გამოცდილების შემსწავლელი ფსიქოლოგი მკვლევრებისა და პრაქტიკოსების მთავარი მიზანი ადამიანის ცხოვრების ხარისხის გაუმჯობესებაა. AI-ს შეუძლია დაეხმაროს ფსიქოლოგებს უკეთ შეისწავლონ ინდივიდის, ჯგუფისა თუ საზოგადოების ფუნქციონირების მექანიზმები მხოლოდ იმ შემთხვევაში, თუ

თავად ფლობენ ამ ტექნოლოგიების გამოყენებისთვის საჭირო პროფესიულ კომპეტენციას. ფსიქოლოგებმა, რომლებიც კომპეტენტურად იყენებენ AI ინსტრუმენტებს, როგორც სამეცნიერო და პრაქტიკული საქმიანობის დამატებით რესურსს, უნდა გააცნობიერონ ამ ინსტრუმენტების დიზაინი, დანიშნულება და ეფექტიანობა, რათა დაიცვან საზოგადოება პოტენციური ზიანისგან და მაქსიმალურად გაზარდონ სარგებელი.

გამოყენებული ლიტერატურა:

- American Psychological Association. (2023, July 1). AI is changing every aspect of psychology. Here's what to watch for. *Monitor on Psychology*, 54(5). <https://www.apa.org/monitor/2023/07/psychology-embracing-ai>
- American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. <https://www.apa.org/topics/artificial-intelligence-machine-learning/ethicalguidance-ai-professional-practice>
- Cruz-Gonzalez, P., He, A. W.-J., Lam, E. P., Ng, I. M. C., Li, M. W., Hou, R., ... Sánchez Vidaña, D. I. (2025). Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine*, 55, e18. doi:10.1017/S0033291724003295

პრინციპი 6. ზედამხედველობა ადამიანის მხრიდან და პროფესიული მსჯელობა

ფსიქოლოგებს, რომლებიც თავიანთ პროფესიულ საქმიანობაში იყენებენ ხელოვნური ინტელექტის (AI) ინსტრუმენტებს, უნდა ჰქონდეთ საკმარისი კომპეტენცია და პროფესიული განსჯის უნარი, რათა უზრუნველყონ AI-ის მხარდაჭერით მიღებული მონაცემების სათანადოდ განკარგვა.

განმარტება

ფსიქოლოგიაში AI-ის გამოყენება არ უნდა ანაცვლებდეს პროფესიულ მსჯელობას ან ადამიანურ პასუხისმგებლობას. მაშინაც, როცა AI ინსტრუმენტები გამოიყენება ფსიქოლოგიური საქმიანობის მხარდასაჭერად (მაგალითად, გადაწყვეტილების მიღებისას, შეფასების პროცესში, ინტერვენციის დაგეგმვისას და ა.შ.), ფსიქოლოგები პასუხს აგებენ AI-ის მიერ გენერირებული შედეგების ინტერპრეტაციაზე, კონტექსტუალიზაციასა და მათზე რეაგირებაზე.

ადამიანური ზედამხედველობა აუცილებელია იმის უზრუნველსაყოფად, რომ AI-ით მხარდაჭერილი პროცესები გამოყენებულ იქნას სწორად, ეთიკურად და ადამიანის ქცევისა და გამოცდილების კომპლექსურობის გათვალისწინებით. ეს გულისხმობს AI-ის მიერ გენერირებული შედეგების შემოწმებას მათი დამაჯერებლობის, რელევანტურობისა და თანხმობის რისკების კუთხით. ასევე, ჩარევას იმ შემთხვევაში, როდესაც ხელოვნური ინტელექტის მიერ წარმოებული

რეკომენდაციები ეწინააღმდეგება პროფესიულ მსჯელობას, ეთიკურ სტანდარტებს ან მოცემულ კონტექსტს.

ეს პრინციპი ვრცელდება სხვადასხვა პროფესიულ გარემოსა და პოპულაციაზე, იმის გათვალისწინებით, რომ AI შეიძლება გამოყენებულ იქნას არა მხოლოდ კლინიკურ პრაქტიკაში, არამედ საგანმანათლებლო პროცესში, ორგანიზაციული გადაწყვეტილებების მიღებისას, კვლევის პროცესებში და ფსიქოლოგიური პრაქტიკის სხვა სფეროებში.

ადამიანური ზედამხედველობის შენარჩუნება ხელს უწყობს ფსიქოლოგიური საქმიანობის ხარისხის, სამართლიანობისა და სანდოობის დაცვას და ამავე დროს ამცირებს ავტომატიზებულ სისტემებზე გადაჭარბებული დამოკიდებულების რისკს.

კავშირი ფსიქოლოგიასთან

ფსიქოლოგიურ კვლევასა და პრაქტიკაში, განსაკუთრებით გამოყენებითი მიმართულებით, ადამიანური ზედამხედველობა აუცილებელია ეთიკური პრინციპების დაცვისა და განხორციელებისთვის. ეს ასპექტი განსაკუთრებით მნიშვნელოვანია კვლევის მონაცემების ანალიზისას, რადგან ადამიანური ზედამხედველობის გარეშე გამოტანილი დასკვნები შესაძლებელია დაეყრდნოს ავტომატურ პასუხებს, რომლებიც უგულებელყოფს კონკრეტული კვლევის კონტექსტს.

კლინიკურ პრაქტიკაში, ადამიანური ზედამხედველობა საჭიროა „ავტომატიზაციის მიკერძოების“ (automation bias) შესამცირებლად, როდესაც კლიენტის მიერ მოწოდებული ინფორმაციის ცალკეული ნიუანსების იგნორირებამ შეიძლება გამოიწვიოს ისეთი შედეგები, რომლებიც მიკერძოებულ რეკომენდაციებს და კლიენტებისთვის პოტენციურ ზიანს წარმოქმნის.

ადამიანური ზედამხედველობა გულისხმობს, რომ საბოლოო გადაწყვეტილებას ფსიქოლოგი იღებს, რაც ამცირებს ტექნოლოგიაზე გადაჭარბებულ დამოკიდებულებას და განამტკიცებს პრინციპს, რომ AI მხოლოდ ეხმარება პროფესიულ მსჯელობას და არ ანაცვლებს მას. ასეთი ზედამხედველობა საშუალებას აძლევს ფსიქოლოგებს შეინარჩუნონ პასუხისმგებლობა, ეთიკური სტანდარტები და ფსიქოლოგიური პროცესის მთლიანობა.

გამოყენებული ლიტერატურა:

American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. <https://www.apa.org/topics/artificial->

- Babu, A., & Joseph, A. P. (2024). Artificial intelligence in mental healthcare: transformative potential vs. the necessity of human interaction. *Frontiers in Psychology*, 15, 1378904. <https://doi.org/10.3389/fpsyg.2024.1378904>
- European Commission. (n.d.) Article 14: Human oversight | EU Artificial Intelligence Act. <https://artificialintelligenceact.eu/article/14/>
- Özer M. (2024). Is artificial intelligence hallucinating? *Turkish Journal of Psychiatry*, 35(4), 333–335. Advance online publication. <https://doi.org/10.5080/u27587>

პრინციპი 7. ფსიქოლოგიური უსაფრთხოება და ზიანის პრევენცია

ფსიქოლოგებმა, რომლებიც იყენებენ ხელოვნური ინტელექტის (AI) სისტემებს, ფსიქოლოგიურ უსაფრთხოებასთან ერთად, პრიორიტეტი უნდა მიანიჭონ სარგებლის მაქსიმალურად გაზრდასა და პოტენციური ზიანის თავიდან აცილებას.

განმარტება

ხელოვნური ინტელექტის ინტეგრირება ფსიქოლოგიურ საქმიანობაში წარმოშობს ახალი ტიპის რისკებს, რომლებმაც შესაძლოა გავლენა მოახდინოს ინდივიდის კეთილდღეობაზე, ავტონომიასა და ქმედითობის განცდაზე. ფსიქოლოგებს ეკისრებათ პასუხისმგებლობა წინასწარ განსაზღვრონ, გამოავლინონ და შეამცირონ ხელოვნური ინტელექტის სისტემების გამოყენებასთან დაკავშირებული პოტენციური ფსიქოლოგიური ზიანი, მათ შორის ემოციური დისტრესი, დაბნეულობა, სტიგმატიზაცია ან არასათანადო ზეგავლენა აზრებზე, ქცევებსა თუ გადაწყვეტილებებზე.

ეს პრინციპი ვრცელდება ყველა იმ კონტექსტზე, სადაც AI-მ შესაძლოა გავლენა მოახდინოს ფსიქოლოგიურ გამოცდილებებზე (მაგ., შეფასებები, რეკომენდაციები, უკუკავშირის სისტემები, სასწავლო და ორგანიზაციული გარემო ან კვლევებში მონაწილეობა). ფსიქოლოგებმა ყურადღება უნდა გაამახვილონ იმაზე, თუ როგორ არის ფორმულირებული, მიწოდებული და გამოყენებული ხელოვნური ინტელექტის მიერ წარმოქმნილი შედეგები, განსაკუთრებით მოწყვლად ჯგუფებთან ურთიერთობისას, ან როდესაც AI სისტემები წარმოქმნიან შინაარსს, რომელიც შესაძლოა აღქმული იყოს როგორც ავტორიტეტული ან პერსონალიზებული (მომხმარებლის უნიკალურ გამოცდილებაზე მორგებული).

ზიანის პრევენცია მოითხოვს AI-ზე დაფუძნებული პროცესების მუდმივ მონიტორინგს, გაუთვალისწინებელ შედეგებზე რეაგირებას და მზაობას შეიცვალოს ან შეწყდეს AI-ს გამოყენება თუ გამოყენებასთან დაკავშირებული რისკები აჭარბებს სარგებელს. ფსიქოლოგიური უსაფრთხოების პრიორიტეტად დასახვით, ფსიქოლოგები ხელს უწყობენ ხელოვნური ინტელექტის ადამიანის საკეთილდღეოდ გამოყენებას.

კავშირი ფსიქოლოგიასთან

ფსიქოლოგებს ევალუატი ადამიანის ქცევის შესწავლა ყველაზე ფაქიზ და უნიკალურ კონტექსტში. ფსიქოლოგიის მეცნიერებასა და პრაქტიკაში ექსპერტული ცოდნის მიღწევა წლების განმავლობაში მიმდინარე მკაცრი მომზადებისა და სუპერვიზიის ქვეშ მუშაობის შედეგია. ეფექტური თერაპიული ინტერვენცია ეყრდნობა მტკიცებულებებზე დაფუძნებულ მიდგომებს და კონტექსტის შესაბამისად მათ გამოყენებას, რაც აგებულია ზრუნვისა და ნდობის პრინციპებზე. ფსიქოლოგები, რომლებიც კლიენტებთან მუშაობის პროცესში AI-ს ინტეგრირებას გეგმავენ, კარგად უნდა იცნობდნენ როგორც სწრაფად განვითარებად სამეცნიერო ლიტერატურას, ისე კლინიკურ შეხედულებებს AI ინსტრუმენტებისა და თერაპიული საშუალებების პოტენციალისა და შეზღუდვების შესახებ. ფსიქოლოგი პასუხისმგებელია AI-ს მიერ მიწოდებული ნებისმიერი ინფორმაციის სისწორეზე, რომელიც გამოიყენება პაციენტის მკურნალობისა თუ დიაგნოსტიკის პროცესში [მაგ.; „ჰალუცინაციები“ ვირტუალურ სივრცეში - შემთხვევები, როდესაც ჩანაწერების შემქმნელი პროგრამა იგონებს ფაქტებს; პაციენტის ხმის ტონისა თუ საუბრის თემატიკის ინტერპრეტაცია; AI-ს მიერ დაფიქსირებული ნიშნები, რომლებიც მიუთითებს პაციენტის მდგომარეობის გაუარესებაზე (რეციდივზე) ან სუიციდურ აზრებზე]. ფსიქოლოგი, რომელიც კარგად ერკვევა ხელოვნური ინტელექტის მუშაობის პრინციპებში, უკეთ ახერხებს პაციენტისა და საკუთარი პროფესიული იმიჯის დაცვას. AI ინსტრუმენტების გონივრული გამოყენება სპეციალისტს საშუალებას აძლევს პრიორიტეტი მიანიჭოს ადამიანის კეთილდღეობას, უზრუნველყოს სიზუსტე, შეინარჩუნოს ნდობა და იზრუნოს ზიანის თავიდან აცილებაზე.

გამოყენებული ლიტერატურა:

- Abrams, Z. (2025, June 26). Artificial intelligence is reshaping how psychologists work. American Psychological Association Services, Inc. <https://www.apaservices.org/practice/news/artificial-intelligence-psychologists-work>
- American Psychological Association. (2025, November). Health advisory: Use of generative AI chatbots and wellness applications for mental health. <https://www.apa.org/topics/artificial-intelligence-machinelearning/health-advisory-chatbots-wellness-apps>
- British Psychological Society. (2024, October 9). AI and the work of psychologists: Practical applications and ethical considerations. <https://www.bps.org.uk/blog/ai-and-work-psychologists-practical-applicationsand-ethical-considerations>
- Putica, A., Khanna, R., Bosl, W., Saraf, S., & Edgcomb, J. (2025, July 24). Ethical decision-making for AI in mental health: The Integrated Ethical Approach for Computational Psychiatry (IEACP) framework. *Psychological Medicine*, 55, e213. <https://doi.org/10.1017/S0033291725101311>

ხელმისაწვდომობა, მიკერძოება და ინკლუზია

პრინციპი 8. ვალიდაცია და მეცნიერული მტკიცებულება. ფსიქოლოგებმა, რომლებიც ხელოვნური ინტელექტის ინსტრუმენტებს იყენებენ, კრიტიკულად უნდა შეაფასონ, ეყრდნობა თუ არა ეს სისტემები გამჭვირვალე და ვალიდურ მეცნიერულ მტკიცებულებებს. ისინი უნდა დარწმუნდნენ, რომ AI-ს სანდოობა და ვალიდურობა დადასტურებულია სწორედ იმ ჯგუფებისა და მიზნებისთვის, რომლებისთვისაც მისი გამოყენება იგეგმება.

განმარტება

ფსიქოლოგიურ კონტექსტში გამოყენებული ხელოვნური ინტელექტის (AI) სისტემები უნდა ეყრდნობოდეს მყარ სამეცნიერო მტკიცებულებებს, რომლებიც შეესაბამება ფსიქოლოგიური კვლევისა და პრაქტიკის დადგენილ სტანდარტებს. ფსიქოლოგებს ეკისრებათ ეთიკური და პროფესიული პასუხისმგებლობა, უზრუნველყონ, რომ AI ინსტრუმენტები არ აღიქმებოდეს როგორც ნეიტრალური ან თავისთავად ვალიდური ტექნოლოგიები. პირიქით, ისინი უნდა განიხილებოდეს როგორც ინსტრუმენტები, რომელთა შედეგებიც დამოკიდებულია ისეთ ფუნდამენტურ მოდელებზე, მონაცემებსა და დაშვებებზე, რაც სისტემურ შეფასებას მოითხოვს.

ვალიდაცია გულისხმობს იმის შეფასებას, თუ რამდენად ავლენს AI სისტემები საკმარის სანდოობას, ვალიდურობასა და სიზუსტეს მათი დანიშნულებისამებრ გამოყენებისას. ასევე, მნიშვნელოვანია შეფასდეს, რამდენად შესაძლებელია AI სისტემებით მოპოვებული ინფორმაცია განზოგადდეს კულტურულ, ლინგვისტურ, დემოგრაფიულ და კონტექსტთან დაკავშირებულ ვარიაციებზე. ერთი პოპულაციიდან ან გარემოდან მიღებული მტკიცებულება შესაძლოა არ იყოს რელევანტური სხვების შემთხვევაში, განსაკუთრებით გლობალურ ან მცირერესურსიან კონტექსტში.

ფსიქოლოგებმა კრიტიკულად უნდა შეაფასონ ხელმოსაწვდომი დოკუმენტაცია, რეცენზირებული მტკიცებულებები და ვალიდაციის კვლევები; მათ უნდა გამოიჩინონ სიფრთხილე, როდესაც მტკიცებულებები შეზღუდული, დახურული ან მიუწვდომელია. როდესაც ემპირიული მხარდაჭერა არასაკმარისია, AI ინსტრუმენტები უნდა იქნას გამოყენებული მხოლოდ შესწავლის, მხარდაჭერის ან კვლევაზე ორიენტირებულ და არა, როგორც გადაწყვეტილების მიმღების როლში, რამაც შესაძლოა გავლენა მოახდინოს ინდივიდის უფლებებზე, კეთილდღეობასა თუ შესაძლებლობებზე.

ასეთი მიდგომით ფსიქოლოგები ხელს უწყობენ ფსიქოლოგიური მეცნიერების მთლიანობის შენარჩუნებას და იცავენ ინდივიდებს გადაუმოწმებელი, შეცდომაში შემყვანი ან გადაჭარბებული ტექნოლოგიური მტკიცებულებისგან.

კავშირი ფსიქოლოგიასთან

ხელოვნური ინტელექტის გლობალურად გავრცელებამ ფსიქიკური ჯანმრთელობის, განათლებისა და ორგანიზაციულ სფეროში გაუსწრო საყოველთაოდ მიღებული ვალიდაციის სტანდარტების შემუშავების პროცესს. ბევრ, კომერციულად ხელმისაწვდომ ინსტრუმენტს აკლია დამოუკიდებელი, ობიექტური შეფასება, ეყრდნობა არარეპრეზენტაციულ მონაცემებს ან არ გააჩნია საკმარისი დოკუმენტაცია, რათა მომხმარებელმა შეაფასოს მისი ადეკვატურობა.

მკაცრი ვალიდაციის გარეშე, AI ინსტრუმენტებმა შესაძლოა წარმოქმნას არაზუსტი ან არათანმიმდევრული შედეგები, შექმნას მიკერძოებული პროგნოზები, რომლებიც აზიანებს გარკვეულ ჯგუფებსა თუ კლინიკური თუ საგანმანათლებლო გადაწყვეტილებების მიღების პროცესს და შეამციროს საზოგადოების ნდობა ფსიქოლოგიური მეცნიერებისა და ტექნოლოგიებით მხარდაჭერილი ზრუნვის მიმართ. შესაბამისად, ვალიდაციის უზრუნველყოფა მნიშვნელოვანია ინდივიდებისა და საზოგადოების დასაცავად, სამართლიანი შედეგების მისაღებად და ფსიქოლოგიის მეცნიერული მთლიანობის შესანარჩუნებლად.

გამოყენებული ლიტერატურა:

Terry, J., Strait, G., Alsarraf, S., Weinmann, E., & Waychoff, A. (2025). Artificial intelligence in scale development: Evaluating AI-generated survey items against gold-standard measures. *Current Psychology*, 44, 16339–16350. <https://doi.org/10.1007/s12144-025-08240-w>

პრინციპი 9. ალგორითმული მიკერძოება და სამართლიანობა

ფსიქოლოგებმა, რომლებიც იყენებენ ხელოვნური ინტელექტის (AI) სისტემებს, პროაქტიულად უნდა გამოავლინონ, შეაფასონ და შეამცირონ ალგორითმული მიკერძოება, რათა დარწმუნდნენ, რომ AI-ის მხარდაჭერით მიღებული შედეგები არ ახდენს ინდივიდებისა თუ ჯგუფების დისკრიმინაციას დემოგრაფიული, კულტურული, სოციოეკონომიკური ან კონტექსტთან დაკავშირებული ფაქტორების საფუძველზე.

განმარტება

ფსიქოლოგიურ კონტექსტში გამოყენებულმა ხელოვნური ინტელექტის სისტემებმა შესაძლოა ასახოს, გააძლიეროს ან შენიღბოს ის სოციალური და სტრუქტურული უთანასწორობა, რომელიც უკვე არსებობს სატრენინგო მონაცემებში, მოდელის დიზაინში ან მისი გამოყენების კონტექსტში. ალგორითმული მიკერძოება შეიძლება წარმოიქმნას ხელოვნური ინტელექტის “სასიცოცხლო ციკლის” სხვადასხვა ეტაპზე,

მათ შორის: მონაცემთა შეგროვების, მახასიათებლების შერჩევის, მოდელის ფორმირების, ვალიდაციისა და რეალურ სამყაროში დანერგვის დროს.

ფსიქოლოგებს ეკისრებათ პასუხისმგებლობა გააცნობიერონ, რომ მიკერძოებული შედეგები არ არის მხოლოდ ტექნიკური შეცდომა, არამედ შესაძლოა ასახავდეს ისტორიულ უთანასწორობას რასის, გენდერის, ეთნიკური წარმომავლობის, ენის, შეზღუდული შესაძლებლობის, ასაკის, სოცოეკონომიკური სტატუსის ან გეოგრაფიული მდებარეობის მიხედვით. ამ მიკერძოებებმა შესაძლოა გამოიწვიოს უსამართლო შეფასებები, სერვისებზე არათანაბარი წვდომა ან არაპროპორციული ზიანი ისედაც მარგინალიზებული პოპულაციებისთვის.

სამართლიანობის ხელშესაწყობად, ფსიქოლოგებმა კრიტიკულად უნდა გამოიკვლიონ, თუ რამდენად თანაბრად მუშაობს AI სისტემები სხვადასხვა პოპულაციისა და კონტექსტში და სისტემატურად განსხვავდება თუ არა შედეგები სხვადასხვა ჯგუფისთვის. აღნიშნული გულისხმობს შეკითხვების დასმას იმის შესახებ, თუ ვისი მონაცემებია წარმოდგენილი, ვისი გამოცდილებაა გამოტოვებული და როგორ შეიძლება ურთიერთქმედებდეს ალგორითმული გადაწყვეტილებები ფსიქოლოგიურ პრაქტიკაში, კვლევასა თუ პოლიტიკაში არსებულ ძალაუფლებრივ ასიმეტრიასთან.

მიკერძოების გამოვლენის ან მასზე ეჭვის გაჩენის შემთხვევაში, ფსიქოლოგებმა მხარი უნდა დაუჭირონ ზიანის შემცირების ისეთ სტრატეგიებს, როგორებიცაა მონაცემთა დივერსიფიკაცია, მოდელის ხელახლა გამართვა (model recalibration), ხელოვნური ინტელექტის შეზღუდვების შესახებ გამჭვირვალობის უზრუნველყოფა ან მისი გამოყენების შეწყვეტა. ხელოვნური ინტელექტის სამართლიანი გამოყენება ერთჯერადი ძალისხმევის შედეგი არ არის, ეს უწყვეტი პროცესია, რომელიც მუდმივ მონიტორინგს, რეფლექსიასა და ინტერდისციპლინურ თანამშრომლობას მოითხოვს.

კავშირი ფსიქოლოგიასთან

ალგორითმული მიკერძოება პირდაპირ საფრთხეს უქმნის ფსიქოლოგიის ბაზისურ ეთიკურ პრინციპებს, მათ შორის სამართლიანობას, სარგებლიანობას და პიროვნებისა თუ საზოგადოების პატივისცემას. პრაქტიკაში (მაგ. შეფასების, დიაგნოსტიკის, ინტერვენციის დაგეგმვის, რისკების შეფასებისა თუ ციფრული ფსიქიკური ჯანმრთელობის სისტემების გამართვის პროცესში), მიკერძოებულმა AI სისტემებმა შესაძლოა შემცირების ნაცვლად, გააძლიერონ არსებული უთანასწორობა.

ფსიქოლოგებისთვის, რომლებიც მუშაობენ გლობალურ, მულტიკულტურულ ან მცირერესურსიან კონტექსტში, გადაუმოწმებელმა ალგორითმულმა მიკერძოებამ

შეიძლება გამოიწვიოს შეუსაბამო განზოგადებები, არასწორი კულტურული ინტერპრეტაციები ან იმ პოპულაციების გარიყვა, რომელთა მონაცემებიც ნაკლებადაა წარმოდგენილი დომინანტურ AI სისტემებში. შესაბამისად, მიკერძოებაზე რეაგირება არსებითია იმის უზრუნველსაყოფად, რომ AI-ით მხარდაჭერილი ფსიქოლოგიური პრაქტიკა დარჩეს სამართლიანი, კულტურულად მგრძნობიარე და პროფესიული ეთიკური სტანდარტების შესაბამისი.

გამოყენებული ლიტერატურა:

- American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning. <https://fairmlbook.org/>
- Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity Press.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1–35. <https://doi.org/10.1007/s12144-025-08240>

პრინციპი 10. წვდომა, თანასწორობა და ინკლუზია

ფსიქოლოგები, რომლებიც პროფესიულ საქმიანობაში იყენებენ AI ინსტრუმენტებს, უნდა დარწმუნდნენ, რომ მომსახურება თანაბრად ხელმისაწვდომი და ინკლუზიურია ყველასთვის და არ აღრმავებს უკვე არსებულ უთანასწორობას, არ რიყავს ინდივიდებსა და ჯგუფებს სტრუქტურული, ეკონომიკური, ენობრივი ან ტექნოლოგიური ბარიერების გამო.

განმარტება

ფსიქოლოგიურ პრაქტიკაში AI-ის ინტეგრაციას აქვს პოტენციალი, გააფართოოს წვდომა სერვისებზე, ცოდნასა და ინტერვენციებზე, თუმცა, თანასწორობისა და ინკლუზიისადმი მიზანმიმართული ყურადღების გარეშე, AI სისტემებმა შეიძლება გააძლიეროს ან გააღრმავოს არსებული უთანასწორობა რეგიონებს, პოპულაციებსა და სოციალურ ჯგუფებს შორის.

ხარჯებთან, ინფრასტრუქტურასთან, ციფრულ წიგნიერებასთან, ენის ცოდნასთან, შეზღუდულ შესაძლებლობებსა და ინტერნეტის ხელმისაწვდომობასთან დაკავშირებულმა ბარიერებმა შეიძლება შეზღუდოს იმ ადამიანთა წრე, ვინც სარგებელს იღებს AI-ზე დაფუძნებული ფსიქოლოგიური ინსტრუმენტებით. გლობალურ კონტექსტში, განსაკუთრებით დაბალი და საშუალო შემოსავლის მქონე ქვეყნებში, სოფლად მცხოვრებ ან მოწყვლად თემებში, ამ ბარიერებმა შეიძლება ყურადღების მიღმა დატოვოს ის პირები, რომლებსაც ფსიქოლოგიური მხარდაჭერა სჭირდებათ.

ფსიქოლოგებმა კრიტიკულად უნდა შეაფასონ, არის თუ არა AI ინსტრუმენტები ხელმისაწვდომი, გამოსაყენებლად ვარგისი და კულტურულად რელევანტური იმ

ადამიანებისთვის, ვისაც ისინი ემსახურებიან. ეს მოიცავს ენობრივი მრავალფეროვნების, შეზღუდული შესაძლებლობის მქონე პირთათვის წვდომის, კონტექსტური შესაბამისობისა და სხვა პირობების გათვალისწინებას, რომელშიც AI სისტემები ინერგება. AI-ში თანასწორობა მიიღწევა არა მხოლოდ ტექნიკური მახასიათებლების უზრუნველყოფით, არამედ იმგვარი დიზაინის დანერგვით, რომელიც ითვალისწინებს არათანაბარ საწყის პირობებს.

ფსიქოლოგების მიზანია, წახალისონ ინკლუზიური განვითარების პრაქტიკები, გამჭვირვალობა ხელმისაწვდომობის თვალსაზრისით არსებულ შეზღუდვებთან დაკავშირებით და ხელოვნური ინტელექტისგან განსხვავებული სტრატეგიების დანერგვა იმ შემთხვევებში, როდესაც ტექნოლოგიების გამოყენება არ არის მიზანშეწონილი ან რელევანტური. თანაბარი წვდომის უზრუნველყოფა მოითხოვს მუდმივ რეფლექსიას იმაზე, თუ ვინ გვყავს ჩართული, ვინ გვრჩება ყურადღების მიღმა და რატომ.

კავშირი ფსიქოლოგიასთან

თანაბარი ხელმისაწვდომობა და ინკლუზია ფსიქოლოგიაში ფუნდამენტური ეთიკური საკითხებია. AI სისტემები, რომლებიც არ არის ყველასთვის ხელმისაწვდომი ან კონტექსტთან შესაბამისი, შეიცავს რისკს, რომ პრივილეგია მიაწოდოს ისედაც უპირატეს მდგომარეობაში მყოფ ჯგუფებს და, ამავდროულად, მოახდინოს სხვების მარგინალიზაცია. თანაბარ ხელმისაწვდომობაზე ყურადღების გამახვილება აუცილებელია სამართლიანობისა და სოციალური პასუხისმგებლობის პრინციპების დასაცავად.

პრაქტიკაში, კვლევაში, განათლებასა და პოლიტიკის კონტექსტში, ფსიქოლოგები ასრულებენ საკვანძო როლს იმის შეფასებაში, AI-ით მხარდაჭერილი ინსტრუმენტები ნამდვილად აფართოებს წვდომას თუ დამატებით აწვდის სერვისებს იმ ჯგუფებს, რომლებსაც ისედაც მეტი ტექნოლოგიური რესურსი აქვთ. ხელმისაწვდომობასა და ინკლუზიის საკითხებზე ფოკუსი, გვეხმარება შევამციროთ ფსიქოლოგიური ზრუნვის თვალსაზრისით არსებული გლობალური და ლოკალური უთანასწორობა.

გამოყენებული ლიტერატურა:

- American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>
- D'Ignazio, C., & Klein, L. F. (2020). Data feminism. MIT Press.
- United Nations. (2015). Sustainable Development Goals. <https://sdgs.un.org/goals>
- World Health Organization. (2021). Ethics and governance of artificial intelligence for health. <https://www.who.int/publications/i/item/9789240029200>

გამოყენებული ტერმინები

ალგორითმული მიკერძოება (Algorithmic bias) - ხელოვნური ინტელექტის სისტემაში არსებული სისტემატური და განმეორებადი პატერნები, რომლებიც წარმოქმნიან უსამართლო/მიკერძოებულ შედეგებს. მიკერძოების მიზეზი ძირითადად არის სისტემაში წარმოდგენილი მონაცემები, არჩეული დიზაინი ან გამოყენების კონტექსტი, რომელიც, როგორც წესი არ ასახავს უმცირესობებს სრულად და ობიექტურად.

ხელოვნური ინტელექტი (AI - Artificial intelligence) - ტექნოლოგია, რომელიც კომპიუტერებსა და მანქანებს ადამიანის მსგავსად სწავლის, აღქმის, პრობლემის გადაჭრის, გადაწყვეტილების მიღების, შემოქმედებისა და ავტონომიურობის სიმულაციის შესაძლებლობას აძლევს და ეფუძნება ნეირონულ კავშირებსა და ალგორითმებს.

ავტომატიზაციის მიკერძოება (Automation bias) - ტენდენცია გადაჭარბებულად დაეყრდნო ავტომატიზირებულ რეკომენდაციებს, რომელიც რიგ შემთხვევებში შესაძლოა ეწინააღმდეგებოდეს ადამიანურ ლოგიკას ან კონტექსტის შესაბამის ინფორმაციას.

შავი ყუთის ტიპის AI (Black box AI): სისტემები, რომელთა გადაწყვეტილების მიღების პროცესებიც არ არის ინტერპრეტირებადი (გაშიფრვადი/გასაგები) ადამიანისთვის.

ციფრული ინსტრუმენტები (Digital tools): ელექტრონული მოწყობილობები, ვებსაიტები, პროგრამული უზრუნველყოფა და აპლიკაციები, რომლებიც ეყრდნობა ტექნოლოგიას ამოცანების შესასრულებლად, ინფორმაციის ორგანიზებისთვის, პროცესების გასამარტივებლად ან/და ეფექტური კომუნიკაციისთვის. ციფრული ინსტრუმენტები ეყრდნობა ელექტრონიკასა და ციფრულ სისტემებს, მაგალითად, აპარატურული ინსტრუმენტები (სმარტფონი, ლეპტოპი) და პროგრამული ინსტრუმენტები (პროგრამები, როგორებიცაა Word ან Excel და ონლაინ სერვისები, როგორიცაა საძიებო სისტემები).

გაციფრულება (digitalization) - სისტემებისა და პროცესების ისე გარდაქმნა და ადაპტაცია, რომ კომპიუტერისა და ინტერნეტის გამოყენებით მოქმედება შევძლოთ.

ახსნადობა/ინტერპრეტირებადობა (Explainability/interpretability) - ადამიანისთვის ხელოვნური ინტელექტის მიერ მიღებული გადაწყვეტილების გაგების ხარისხი.

გენერაციული ხელოვნური ინტელექტი (Generative AI) - ხელოვნური ინტელექტის სისტემები, რომლებიც წარმოქმნიან ახალ პროდუქტებს (ტექსტს, სურათს და ა.შ).

ჰალუცინაცია (Hallucination - ციფრულ სამყაროში) - ფენომენი, როდესაც AI მოდელი წარმოქმნის დამაჯერებელ, მაგრამ მცდარ ან უსაფუძვლო ინფორმაციას.

ინტერვენცია (Intervention) - განსაზღვრული ტექნიკა ან სტრუქტურა, რომელიც გამოიყენება ადამიანის აზროვნებაში, ქცევაში, ემოციებში, ურთიერთობებში ან საერთო კეთილდღეობასა და ჯანსაღ ცხოვრებაში პოზიტიური ცვლილებების შესატანად. ფსიქოლოგიური ინტერვენცია ეფუძნება თეორიულ და ემპირიულ მეცნიერებას.

დიდი ენობრივი მოდელი (LLM; Large language model) - რომელიც გაწვრთნილია ტექსტური და გამოსახულების მონაცემების დიდ რაოდენობაზე, რათა გაიგოს, შეაჯამოს და შექმნას ადამიანის ენის მსგავსი ტექსტი.

მანქანური დასწავლა (Machine learning) - AI-ის ქვედარგი, სადაც სისტემები უმჯობესდება მონაცემებთან შეხების (მონაცემების დამუშავების) პროცესში.

ნეირონული ქსელი (Neural network - ციფრულ სამყაროში) - კომპიუტერული არქიტექტურა, რომელიც შთაგონებულია ბიოლოგიური ნეირონებით.

ფსიქოლოგიური სერვისები (Psychological services): პროფესიონალი ფსიქოლოგების მიერ განხორციელებული ინტერვენციებისა და აქტივობების სრული სპექტრი, რომელიც მოიცავს ფსიქოლოგიის ყველა სფეროს სხვადასხვა გარემოში. ეს სერვისები მოიცავს (მაგრამ არ შემოიფარგლება) კვლევაზე დაფუძნებულ მידგომებს, კლინიკურ/საკონსულტაციო/საგანმანათლებლო მომსახურებას, პროფესიულ ორიენტაციასა და საზოგადოებრივ ჯანდაცვას და მიზნად ისახავს ინდივიდების, წყვილების, ოჯახების, ჯგუფების, საზოგადოებებისა და ტექნოლოგიების ფუნქციონირების მექანიზმის გაგებას და მათთან მუშაობას.

სატრენინგო მონაცემები (Training data): საწყის მონაცემთა ბაზა (dataset), რომელიც გამოიყენება მანქანური დასწავლის ალგორითმისთვის ინფორმაციის დამუშავებისა და პროგნოზების გაკეთების სწავლებისთვის.