პერსონალურ მონაცემთა
დაცვის სამსახური

ბენეფიციარებაღი ხელოვნური ინტელექტი და ევროკავშირის მონაცემთა დაცვის რეგულაცია ("EUDPR")

ევროკავშირის მონაცემთა დაცვის ზედამხედველის შეფასებები ბენეფიციარებაღი ხელოვნური ინტელექტის სისტემების გამოყენებისას მონაცემთა დაცვის შესაბამისობის უზრუნველყოფის შესახებ

გენერირებადი ხელოვნური ინტელექტი და ევროკავშირის მონაცემთა დაცვის რეგულაცია (“EUDPR”)

ევროკავშირის მონაცემთა დაცვის ზედამხედველის შეფასებები გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენებისას მონაცემთა დაცვის შესაბამისობის უზრუნველყოფის შესახებ

2024 წლის 3 ივნისი

დოკუმენტი შემუშავებულია ევროკავშირის მონაცემთა დაცვის ზედამხედველის (“EDPS”) მიერ, რომელიც არის პერსონალურ მონაცემთა დაცვის ევროკავშირის დამოუკიდებელი ორგანო

გენერირებადი ხელოვნური ინტელექტისა (“Generative Artificial Intelligence”) და პერსონალურ მონაცემთა დაცვის თემებზე ევროკავშირის მონაცემთა დაცვის ზედამხედველის (“EDPS”) მიერ შემუშავებული სახელმძღვანელო დასკვნების მიზანია გენერირებადი ხელოვნური ინტელექტის გამოყენებით, პერსონალური მონაცემების დამუშავების პროცესში ევროკავშირის ინსტიტუტებისთვის, ორგანოებისთვის, ოფისებისა და სააგენტოებისთვის (“EUIs”) პრაქტიკული რჩევებისა და ინსტრუქციების მიცემა. აღნიშნული ხელს შეუწყობს ევროკავშირის რეგულაციით (2018/1725) გათვალისწინებული ვალდებულებების შესრულებას. სახელმძღვანელო დასკვნები შემუშავდა სხვადასხვა სიტუაციის განსახილველად და შესაბამისად, მასში არ არის ჩამოყალიბებული კონკრეტული ტექნიკური ზომები. აღნიშნულის საპირწონედ, ყურადღებაა გამახვილებული მონაცემთა დაცვის ძირითად პრინციპებზე, რომლებიც ევროკავშირის ინსტიტუტებს, ორგანოებს, ოფისებსა და სააგენტოებს (“EUIs”) დაეხმარებათ, გაითვალისწინონ ევროკავშირის 2018/1725 რეგულაციით დადგენილი მოთხოვნები.

წინამდებარე დოკუმენტი წარმოადგენს პირველ წინგადადგმულ ნაბიჯს, რომელშიც დეტალურადაა ასახული: გენერირებადი ხელოვნური ინტელექტისა და ტექნოლოგიების განვითარება; გენერირებადი ხელოვნური ინტელექტისა და ტექნოლოგიების გამოყენება “EUIs”-ის მიერ; “EDPS”-ის მონიტორინგისა და ზედამხედველობასთან დაკავშირებული საქმიანობების შედეგები.

აღნიშნული დასკვნები ევროკავშირის მონაცემთა დაცვის ზედამხედველმა შეიმუშავა „ხელოვნური ინტელექტის აქტის“ შესაბამისად, საკუთარი უფლებამოსილების ფარგლებში, როგორც მონაცემთა დაცვის საზედამხედველო ორგანომ და არა ახალმა საზედამხედველო უწყებამ ხელოვნური ინტელექტის მიმართულებით.

სახელმძღვანელო დოკუმენტი არ ეწინააღმდეგება „ხელოვნური ინტელექტის აქტს“ (“Artificial Intelligence Act”).

სარჩევი

შესავალი და ზოგადი შინაარსი 4

1. რას წარმოადგენს გენერირებადი ხელოვნური ინტელექტი? 5

2. შეუძლიათ თუ არა ევროკავშირის ინსტიტუტებს, ორგანოებს, ოფისებსა და სააგენტოებს (“EUIs”) გენერირებადი ხელოვნური ინტელექტის გამოყენება? 9

3. როგორ უნდა დადგინდეს გენერირებადი ხელოვნური ინტელექტის სისტემის გამოყენებისას მუშავდება თუ არა პერსონალური მონაცემები? 11

4. რა როლი აქვთ მონაცემთა დაცვის ოფიცრებს (“DPOs”) გენერირებადი ხელოვნური ინტელექტის განვითარების ან გამოყენების პროცესში? 13

5. “EUI”-ის მხრიდან გენერირებადი ხელოვნური ინტელექტის სისტემების შემუშავების პროცესში როდის უნდა განხორციელდეს მონაცემთა დაცვაზე ზეგავლენის შეფასება (“DPIA”)? 15

6. კანონიერია თუ არა პერსონალური მონაცემების დამუშავება გენერირებადი ხელოვნური ინტელექტის შექმნის, განვითარებისა და გამოყენებისთვის? 17

7. როგორ მოქმედებს მონაცემთა მინიმიზაციის პრინციპი გენერირებადი ხელოვნური ინტელექტის გამოყენებისას? 21

8. შეესაბამება თუ არა ხელოვნური ინტელექტის სისტემები მონაცემთა სიზუსტის პრინციპს? 23

9. როგორ უნდა შევატყობინოთ მონაცემთა სუბიექტებს მონაცემთა ხელოვნური ინტელექტის სისტემის საშუალებით დამუშავების შესახებ? 25

10. რას გულისხმობს ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღება რეგულაციის 24-ე მუხლის კონტექსტში? 27

11. გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენებისას, როგორ უნდა იქნეს უზრუნველყოფილი კანონიერი დამუშავება მიკერძოების გარეშე? 29

12. რას გულისხმობს ფიზიკურ პირთა მიერ უფლებების განხორციელება? 31

13. რას გულისხმობს მონაცემთა უსაფრთხოება? 33

14. გსურთ იცოდეთ მეტი? 35

შესავალი და ზოგადი შინაარსი

სახელმძღვანელოს მიზანია ევროკავშირის ინსტიტუტებისთვის, ორგანოებისთვის, ოფისებისა და სააგენტოებისთვის (“EUIs”) გარკვეული პრაქტიკული რჩევების გაზიარება გენერირებადი ხელოვნური ინტელექტის გამოყენებისას პერსონალურ მონაცემთა დამუშავების შესახებ, რათა უზრუნველყოფილი იქნას ევროკავშირის 2018/1725 რეგულაციით (“EUDPR”)[1] დადგენილ ვალდებულებებთან შესაბამისობა. მიუხედავად იმისა, რომ რეგულაცია მკაფიოდ არ განმარტავს ხელოვნური ინტელექტის ცნებას, ამ სისტემების სასარგებლოდ გამოყენების მიზნებისთვის, არსებითად მნიშვნელოვანია მონაცემთა დაცვის პრინციპების სწორი ინტერპრეტაცია და გამოყენება. აღნიშნულის მთავარი ორიენტირი ფიზიკური პირების ძირითადი უფლებებისა და თავისუფლებების დაცვაა.

აღნიშნული სახელმძღვანელო ევროკავშირის მონაცემთა დაცვის ზედამხედველმა, „ხელოვნური ინტელექტის აქტის“ შესაბამისად, შეიმუშავა საკუთარი უფლებამოსილების ფარგლებში, როგორც მონაცემთა დაცვის საზედამხედველო ორგანომ და არა ახალმა საზედამხედველო ორგანომ ხელოვნური ინტელექტის დარგობრივი მიმართულებით.

სახელმძღვანელოს მიზანს არ წარმოადგენს გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენებით პერსონალური მონაცემების დამუშავებასთან დაკავშირებული ყველა მნიშვნელოვანი საკითხის დეტალურად განხილვა. აღნიშნული საკითხის შესწავლა მონაცემთა დაცვის საზედამხედველო ორგანოების პრეროგატივაა. გენერირებადი ხელოვნური ინტელექტის მუშაობასთან დაკავშირებით არაერთი კითხვა რჩება პასუხგაუცემელი და მათი მზარდი გამოყენების კვალდაკვალ, დამატებით ახალი კითხვები ჩნდება.

ხელოვნური ინტელექტის სწრაფად განვითარების გათვალისწინებით, მრავალფეროვანია მომსახურების უზრუნველსაყოფად გამოყენებული ინსტრუმენტები და საშუალებები, რაც შეიძლება სწრაფად შეიცვალოს. ამდენად, დოკუმენტის მიზანია სხვადასხვა შემთხვევებისა და გარემოებების განხილვა.

[1] Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC, < http://data.europa.eu/eli/reg/2018/1725/oj>.

სახელმძღვანელო შედგება შემდეგი სტრუქტურული ნაწილებისგან: მთავარი კითხვები; პირველადი პასუხები შესაბამის დასკვნებთან ერთად; დამატებითი განმარტებები და მაგალითები.

დოკუმენტში ასახული პირველადი მიგნებები წარმოადგენს საწყის ნაბიჯს მეტად დეტალიზებული და ყოვლისმომცველი სახელმძღვანელოს შესაქმნელად. პერიოდულად აღნიშნული დასკვნები განახლდება, გაუმჯობესდება და გაფართოვდება "EUIs"-ის სისტემების განვითარების ხელშესაწყობად. დოკუმენტი განახლდება მისი გამოქვეყნებიდან 12 თვეში.

1. რას წარმოადგენს გენერირებადი ხელოვნური ინტელექტი?

გენერირებადი ხელოვნური ინტელექტი წარმოადგენს ხელოვნური ინტელექტის ქვეჯგუფს, რომლის ფუნქციონირებისას დროს სპეციალიზებული სწავლების მოდელები გამოიყენება. აღნიშნულის მიზანია სხვადასხვა პროგრამის შექმნა, რომელიც ორიენტირებული იქნება რიგი დავალებების შესრულებაზე, მაგალითად, როგორიცაა: გენერირებადი ტექსტი, სურათი ან აუდიოჩაწერა. კონკრეტულად, მითითებული მოდელი ეყრდნობა ე. წ. „ძირითად მოდელებს", რომლებიც წარმოადგენს სხვა გენერირებადი ხელოვნური ინტელექტის საწყის მოდელებს და შემდგომ განხორციელდება მათი ფორმირება.

საწყისი მოდელი არის ძირითადი სტრუქტურა სხვა სპეციალიზებული მოდელების შესაქმნელად. ამგვარი მოდელები მრავალფეროვანია და დიდი მოცულობის მონაცემების საფუძველზე ფუნქციონირებს, მათ შორის, იმ მონაცემებზე დაყრდნობით, რომლებშიც ასახულია საჯაროდ ხელმისაწვდომი ინფორმაცია. აღნიშნული შეიძლება, წარმოადგენდეს კომპლექსურ სტრუქტურებს, მაგალითად, როგორებიცაა: სურათები, აუდიო, ვიდეო ან ენა, რაც ცვალებადია კონკრეტული ამოცანებისა და გამოყენების მიზნების გათვალისწინებით.

„მოცულობითი ენობრივი მოდელები" ("Large Language Models", "LLM") წარმოადგენს კონკრეტული სახეობის საწყის მოდელს, რომელიც დიდი რაოდენობის ტექსტობრივ მონაცემებთან მიმართებით გამოიყენება (სიტყვათა ოდენობა მილიონიდან მილიარდამდე). აღნიშნული სიტყვების საშუალებით შეიძლება, შემუშავდეს პასუხები ჩვეულებრივი ენის გამოყენებით ("Natural Language Responses", "NLR"), რომლებიც დაეფუძნება ნიმუშებს, აგრეთვე, სიტყვებსა და ფრაზებს შორის ურთიერთმიმართებას. მოდელის გასამართად დიდი მოცულობის ტექსტი შეიძლება,

გამოყენებული იქნას სხვადასხვა წყაროდან, მაგალითად, როგორებიცაა: ინტერნეტი, წიგნები და სხვა ხელმისაწვდომი მასალა. აღნიშნულ მოდელებზე დაყრდნობით შემუშავებულია არაერთი აპლიკაცია, როგორებიცაა: „წესების ფუნქციონირების სისტემები“ (“Code Generation Systems”); „ვირტუალური დამხმარეები“ (“Virtual Assistants”); „შინაარსის შექმნის ინსტრუმენტები“ (“Content Creation Tools”); „თარგმნის მექანიზმები“ (“Language Translation Engines”); „ენის ავტომატური ამომცნობი“ (“Automated Speech Recognition”); „სამედიცინო დიაგნოზის სისტემები“ (“Medical Diagnosis Systems”); „სამეცნიერო კვლევის ინსტრუმენტები“ (“Scientific Research Tools”) და სხვა.

ზემოაღნიშნულ ცნებებს შორის ურთიერთმიმართება იერარქიული ხასიათისაა. გენერირებადი ხელოვნური ინტელექტი წარმოადგენს ფართო კატეგორიას, რომელიც მოიცავს მოდელებს შინაარსის შესაქმნელად. „ენობრივი მოდელი“ გამოიყენება საწყის სტრუქტურად, რომლის საფუძველზეც სპეციალიზებული მოდელები იქმნება. ამ უკანასკნელის მიზანია კონკრეტული დავალების შესრულება საწყისი სტრუქტურის ცოდნისა და უნარების გამოყენებით.

გენერირებადი ხელოვნური ინტელექტის „სწავლების“ სრული პროცესი ფარავს სხვადასხვა ეტაპს, დაწყებული მოდელის ფარგლებისა და მისი გამოყენების განმარტებით. ზოგიერთ შემთხვევაში, საწყის ეტაპზე შესაძლებელია შესაბამისი „საწყისი მოდელის“ გამოყენება, ხოლო სხვა შემთხვევებში, შეიძლება შეიქმნას ახალი მოდელი. მომდევნო ფაზა მოიცავს კონკრეტულ მონაცემებთან მიმართებით მოდელის სწავლების დაწყებას შესაბამისი მონაცემების გამოყენებით, სამომავლოდ დახვეწის მიზნით. აღნიშნული მოიაზრებს, მათ შორის, სისტემაში მინიმალური სწორებების შეტანას სპეციალიზებული მონაცემების საშუალებით, რომლებიც შექმნილია მოდელის გამოსაყენებლად. სწავლების სრულყოფისთვის გამოიყენება სპეციფიკური მექანიზმები, რომლებიც მოითხოვს ფიზიკურ პირთა ჩართულობას, რათა უზრუნველყოფილი იქნას ინფორმაციის სიზუსტე და ქცევის კონტროლი. შემდგომი ფაზის დროს მოდელი ფასდება სპეციალური საზომი საშუალებების მეშვეობით, რათა მუდმივ რეჟიმში შეფასდეს ისეთი მიმართულებები, როგორებიცაა სიზუსტე და მოდელის კონკრეტულ შემთხვევასთან შესაბამისობა. საბოლოო ეტაპზე, იწყება მოდელების ფუნქციონირება და ხორციელდება მათ მიმართ მუდმივი მონიტორინგი და შეფასება იმ საზომი ინსტრუმენტების გამოყენებით, რომლებიც შემუშავდა წინა ეტაპებზე.

გენერირებადი ხელოვნური ინტელექტი გამოიყენება მომხმარებელზე ორიენტირებულ სხვადასხვა სფეროში (მაგალითად, “ChatGPT” და სხვა მსგავსი

სისტემები, რომლებიც შეიძლება, ხელმისაწვდომი იყოს სხვადასხვა ფორმით[2], მათ შორის, მობილური აპლიკაციების მეშვეობით). ამასთანავე, ბიზნეს მიზნებისთვის შესაძლებელია, გამოყენებული იქნას „დატრენინგებამდე“ არსებული მოდელები და შემდგომ განხორციელდეს მათში მცირე სწორებების შეტანა საქმიანობის სფეროს გათვალისწინებით.

გენერირებადი ხელოვნური ინტელექტი, როგორც სხვა ახალი ტექნოლოგიები, სხვადასხვა სფეროში საკითხების გადაწყვეტის საშუალებას იძლევა, რაც ხელს უწყობს ადამიანური შესაძლებლობების გაძლიერებას. თუმცა, ამავდროულად, წარმოშობს გარკვეულ გამოწვევებს, რაც გულისხმობს ადამიანის უფლებებსა და თავისუფლებებზე პოტენციურ გავლენას. აღნიშნულის მიზეზია არაიდენტიფიცირებული რისკი და შესაბამისი განხილვისა და შეფასების არარსებობა.

➲ „მოცულობითი ენობრივი მოდელის“ დატრენინგება (ზოგადად, ნებისმიერი მანქანური სწავლების მოდელი) წარმოადგენს მრავალეტაპიან და ინტენსიურ პროცესს, რაც მოიცავს შესაბამისი ტექნიკური საშუალებების გამოყენებას, რომლის მიზანია ადამიანის ენაზე მორგებული მოდელის შექმნა. აღნიშნული მოდელი იქმნება იმისათვის, რომ მომხმარებლის მოთხოვნებს/კითხვებს გაეცეს შესაბამისი პასუხი. წარმოდგენილი პროცესი იწყება დიდი ოდენობის მონაცემების შესახებ მოდელის „დატრენინგებით“. ამ მონაცემთა უმეტესი ნაწილი, როგორც წესი, არ არის დაკონკრეტებული და მოპოვებულია საჯარო წყაროებიდან ინტერნეტ საშუალებების (“Webscraping Technologies”) გამოყენებით (აღნიშნულთან დაკავშირებით პერსონალურ მონაცემთა დაცვის საზედამხედველო ორგანოებმა უკვე გამოხატეს შეშფოთება და დაასახელეს პირადი ცხოვრების ხელშეუხებლობასა და მონაცემთა დაცვასთან დაკავშირებული რისკები, რომლებიც შეეხება საჯაროდ ხელმისაწვდომი პერსონალური მონაცემების მოპოვებას). აღნიშნულის შემდეგ, ზოგიერთ შემთხვევაში, ხორციელდება „მოცულობითი ენობრივი მოდელის“ კორექტირება სასწავლო ან სხვა ტექნიკური საშუალებების გამოყენებით, რომლებიც მოიცავს ადამიანური რესურსის ჩართულობას (მაგალითად, საშუალებები, როგორიცაა: „გაძლიერებული სწავლება ფიზიკურ პირთა უკუკავშირით“ ან “Adversarial Testing via Domain Experts”). ამ უკანასკნელის მიზანია სისტემის მხარდაჭერა, რათა უკეთ დაამუშაოს ინფორმაცია და შინაარსი, ასევე, განსაზღვროს შესაბამისი პასუხები.

[2] „მოცულობითი ენობრივი მოდელის“ ფორმა, როგორც წესი, ფასდება სხვადასხვა პარამეტრის საშუალებით (სიტყვების რაოდენობა. „მოცულობითი ენობრივი მოდელის“ ზომა მნიშვნელოვანია რამდენადაც ზოგიერთი საკითხი მაშინ იკვეთება, როდესაც მოდელი შესაბამის ზღვარს გასცდება).

აღნიშნული მოიაზრებს სისტემის შემმუშავებელთა ღირებულებებთან შესაბამისობის უზრუნველყოფას შინაარსის გაფილტვრის საშუალებით (მაგალითად, საზიანო შედეგის თავიდან აცილება). ფუნქციონირების დაწყების შემდეგ, ზოგიერთი სისტემა ხელს უწყობს მოდელის გაუმჯობესებას იმ მონაცემების გამოყენებით, რომლებიც მოპოვებულია ახალი სასწავლო მასალებიდან და მომხმარებლებთან განხორციელებული ინტერაქციის საფუძველზე.

2. შეუძლიათ თუ არა ევროკავშირის ინსტიტუტებს, ორგანოებს, ოფისებსა და სააგენტოებს (“EUIs”) გენერირებადი ხელოვნური ინტელექტის გამოყენება?

ევროკავშირის ინსტიტუტებისთვის, ორგანოებისთვის, ოფისებისა და სააგენტოებისთვის (“EUIs”) არ არსებობს რაიმე შეზღუდვა საჯარო მომსახურებისთვის გენერირებადი ხელოვნური ინტელექტის შექმნისა და გამოყენების თვალსაზრისით. აღნიშნულ პროცესში მნიშვნელოვანია დადგენილი წესებისა და სამართლებრივი მოთხოვნების გათვალისწინება. განსაკუთრებით საყურადღებოა საჯარო სექტორის პასუხისმგებლობა ახალი ტექნოლოგიების გამოყენებისას ადამიანის უფლებებისა და თავისუფლების პატივისცემასთან მიმართებით.

ნებისმიერ შემთხვევაში, თუ პერსონალური მონაცემები მუშავდება გენერირებადი ხელოვნური ინტელექტის საშუალებით, რეგულაცია გამოიყენება სრული მოცულობით. რეგულაცია ვრცელდება პერსონალურ მონაცემთა დამუშავებასთან დაკავშირებულ ყველა საქმიანობაზე, მიუხედავად გამოყენებული ტექნოლოგიების სახეობისა. ასევე, რეგულაცია არ იკვეთება სხვა კანონმდებლობასთან, მაგალითად, როგორიცაა: „ხელოვნური ინტელექტის“ აქტი. ანგარიშვალდებულების პრინციპი მოითხოვს პასუხისმგებლობების განსაზღვრასა და პატივისცემას სხვა დაკავშირებულ საკითხებთან ერთობლიობაში.

“EUI”-ებს შეუძლიათ, შეიმუშაონ საკუთარი გენერირებადი ხელოვნური ინტელექტის სისტემები გადაწყვეტილების მისაღებად ან სანაცვლოდ გამოიყენონ ბაზარზე არსებული პრობლემის გადაჭრის გზები. ორივე შემთხვევაში, “EUI”-ებმა შეიძლება, დაიხმარონ მიმწოდებლები გენერირებადი ხელოვნური ინტელექტის სისტემებთან დაკავშირებული ინფორმაციის მოსაპოვებლად. ამ მიმართულებით აუცილებელია, რომ “EUI”-ებმა განსაზღვრონ დამუშავებისთვის პასუხისმგებელი პირის, დამუშავებაზე უფლებამოსილი პირისა და თანადამუშავებისთვის პასუხისმგებელი პირის სპეციალური უფლებამოსილებები. აღნიშნულის მიზანია მონაცემთა დამუშავების ოპერაციების რეგულაციით გათვალისწინებულ ვალდებულებებთან შესაბამისობის უზრუნველყოფა.

გამომდინარე იქიდან, რომ ხელოვნური ინტელექტის ტექნოლოგიები სწრაფად ვითარდება, “EUI”-ებმა უნდა იმსჯელონ, თუ რა შემთხვევაში და როგორ უნდა გამოიყენებოდეს გენერირებადი ხელოვნური ინტელექტი, რათა საზოგადოებამ სარგებელი მიიღოს. გენერირებადი ხელოვნური ინტელექტის ფუნქციონირების ყველა

ეტაპზე უნდა იქნას გათვალისწინებული შესაბამისი კანონმდებლობა, მათ შორის, რეგულაცია, როდესაც სისტემის მეშვეობით მუშავდება პერსონალური მონაცემები.

➲ სანდო ხელოვნური ინტელექტი უნდა შემუშავდეს ეთიკური და სამართლიანი გზით. ამ მიზნით, მხედველობაში უნდა იქნას მიღებული ხელოვნური ინტელექტის ტექნოლოგიების გამოყენებისას გაუთვალისწინებელი შემთხვევები და უზრუნველყოფილი იქნას რისკზე ორიენტირებული მიდგომა, რომელიც დაფარავს სისტემის ფუნქციონირების ყველა ეტაპს. აღნიშნული, ასევე, მოიცავს სატესტო მონაცემების გამოყენებასა და მის წყაროებს, რომლებიც შეეხება: ალგორითმების შექმნასა და დანერგვას; სისტემაში მიკერძოების საკითხებს; ფიზიკურ პირთა უფლებებსა და თავისუფლებებზე შესაძლო გავლენასთან ბრძოლას. ამ თვალსაზრისით, გენერირებადი ხელოვნური ინტელექტის სისტემები პერსონალური მონაცემების კანონიერად დასამუშავებლად უნდა იყოს: გამჭვირვალე; გასაგები; თანმიმდევრული; შემოწმებადი; ხელმისაწვდომი.

3. როგორ უნდა დადგინდეს გენერირებადი ხელოვნური ინტელექტის სისტემის გამოყენებისას მუშავდება თუ არა პერსონალური მონაცემები?

პერსონალური მონაცემები შეიძლება დამუშავდეს გენერირებადი ხელოვნური ინტელექტის სისტემის ფუნქციონირების სხვადასხვა ეტაპზე. ზოგჯერ საწყის ეტაპზე არ იკვეთება პერსონალური მონაცემების არსებობა. მაგალითად, აღსანიშნავია სატესტო მონაცემების შექმნის ეტაპი, რომელიც მოიაზრებს მოდელის გამოყენების მიმდინარე და მისი შექმნის შემდგომ ფაზებს. ამ ეტაპზე, მოდელის შექმნისა და გამოყენებისას, შეიძლება, აისახოს ახალი ან დამატებითი ინფორმაცია, ან გამოყენებული იქნას არსებული სისტემა.

როდესაც გენერირებადი ხელოვნური ინტელექტის სისტემის შემქმნელი ან მიმწოდებელი განაცხადებს, რომ მათი სისტემის მეშვეობით პერსონალური მონაცემები არ მუშავდება, არსებითად მნიშვნელოვანია, თუ რამდენად არსებობს შესაბამისი გარანტიები, რომ სამომავლოდ არ დამუშავდება პერსონალური მონაცემები. ძირითად შემთხვევებში, "EUI"-ებს შეიძლება სურდეთ, მიიღონ ინფორმაცია მასზედ, თუ რა ეტაპები ან მეთოდები გამოიყენება პერსონალური მონაცემების დამუშავების კანონიერების უზრუნველსაყოფად.

"EDPS" არ ემხრობა[3] პერსონალური მონაცემების მოსაპოვებლად ინტერნეტ საშუალებების გამოყენებას ("Web Scraping Techniques"), რამდენადაც ფიზიკურმა პირებმა შეიძლება, დაკარგონ კონტროლი საკუთარ პერსონალურ ინფორმაციაზე. აღნიშნულის მაგალითებს წარმოადგენს მონაცემების შეგროვება მონაცემთა სუბიექტების ცოდნის გარეშე; პირთათვის მოულოდნელად; თავდაპირველი დამუშავების მიზნისაგან განსხვავებული მიზნით. "EDPS", ასევე, აღნიშნავს, რომ საჯაროდ ხელმისაწვდომი პერსონალური მონაცემების დამუშავება ექვემდებარება ევროკავშირის მონაცემთა დაცვის კანონმდებლობას. ამ თვალსაზრისით, ვებგვერდებიდან, ინტერნეტ საშუალებების გამოყენებით პერსონალური მონაცემების შეგროვება და მათი გამოყენება „სწავლების" მიზნებისთვის შეიძლება, არ შეესაბამებოდეს მონაცემთა დაცვის პრინციპებს, მათ შორის, მონაცემთა მინიმიზაციასა და სიზუსტის პრინციპებს, რადგან არ ფასდება წყაროების სანდოობა.

[3] 2023 წლის 25 სექტემბრის 41/2023 მოსაზრება, რომელიც შეეხება შეთავაზებას რეგულაციის შექმნაზე, ევროკავშირის შრომითი ბაზრის სტატისტიკასთან დაკავშირებით.

ყოველ ეტაპთან მიმართებით რეგულარული მონიტორინგისა და კონტროლის განხორციელება ხელს შეუწყობს პერსონალური მონაცემების დამუშავების დასაბუთებას იმ შემთხვევებშიც კი, როდესაც მოდელი არ არის აღნიშნული მიზნით შექმნილი.

➲ "EUI-X" — ევროკავშირის პირობითი ინსტიტუტი, გეგმავს პროდუქტის შეძენას, ავტომატური ენის ამოცნობისა და ტრანსკრიფციის მიზნებისთვის. სხვადასხვა ვარიანტის შესწავლის შემდგომ, "EUI-X"-მა ყურადღება გაამახვილა დასახელებული ფუნქციის ხელშესაწყობად გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენების შესაძლებლობაზე. მოცემულ შემთხვევაში, აღნიშნული წარმოადგენს სისტემას, რომელიც უზრუნველყოფს „სწავლებამდე" მოდელს, ენის ამოცნობისა და თარგმნის მიზნებისათვის. იმის გათვალისწინებით, რომ აღნიშნული მოდელის გამოყენება ხმოვანი მასალების მეშვეობით შეხვედრების ჩაწერის მიზნებისთვის ხორციელდება, ინსტიტუტმა დაადგინა, რომ აღნიშნული მოდელის გამოყენება მოითხოვს პერსონალური მონაცემების დამუშავებას, რაც საჭიროებს რეგულაციასთან შესაბამისობას.

4. რა როლი აქვთ მონაცემთა დაცვის ოფიცრებს (“DPOs”) გენერირებადი ხელოვნური ინტელექტის განვითარების ან გამოყენების პროცესში?

რეგულაციის 45-ე მუხლი ითვალისწინებს მონაცემთა დაცვის ოფიცრის ფუნქცია-მოვალეობებს. მონაცემთა დაცვის ოფიცერთა საქმიანობის ერთ-ერთი მიმართულებაა მონაცემთა დაცვის ვალდებულებებთან დაკავშირებით ინფორმაციის მიწოდება და კონსულტაციების გაწევა. ამასთანავე, ოფიცერი ეხმარება დამუშავებისთვის პასუხისმგებელ პირებს, უზრუნველყონ მონაცემთა დამუშავების სტანდარტებთან შესაბამისობა; საჭიროების შემთხვევაში, აძლევს რჩევას მონაცემთა დაცვაზე ზეგავლენის შეფასების განხორციელებაზე (“DPIAs”); წარმოადგენს საკონტაქტო პირს მონაცემთა სუბიექტებთან და ევროკავშირის მონაცემთა დაცვის ზედამხედველთან.

“EUI”-ების მიერ პერსონალური მონაცემების დამუშავების მიზნით, გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენების კონტექსტში, მნიშვნელოვანია, რომ მონაცემთა დაცვის ოფიცრებმა, კომპეტენციის ფარგლებში, შესაბამის პირებს მისცენ რჩევები რეგულაციის გამოყენებასთან დაკავშირებით. დამატებით, ოფიცრებს უნდა გააჩნდეთ შესაბამისი ცოდნა გენერირებადი ხელოვნური ინტელექტის სისტემების ფუნქციონირების ყველა ეტაპის შესახებ, რაც მოიცავს შესყიდვას, შექმნას, დანერგვასა და მუშაობის სპეციფიკას. აღნიშნულში, ასევე, მოიაზრება ინფორმაციის შეგროვება, თუ როდის და როგორ მუშავდება აღნიშნული სისტემების საშუალებით პერსონალური მონაცემები. ასევე, გასათვალისწინებელია სხვადასხვა მექანიზმების (“Input and Output Mechanisms”) მუშაობა და მოდელის საშუალებით დანერგილი გადაწყვეტილების მიღების პროცესი. რეგულაციის თანახმად, მნიშვნელოვანია, დამუშავებისთვის პასუხისმგებელი პირებისთვის რჩევების მიცემა მონაცემთა დაცვაზე ზეგავლენის შეფასების განხორციელების პროცესში. დამუშავებისთვის პასუხისმგებელმა პირებმა უნდა აღრიცხონ მონაცემთა დამუშავებასთან დაკავშირებული მოქმედებები და უზრუნველყონ გამჭვირვალობა, მათ შორის, დამუშავებასთან დაკავშირებული ჩანაწერების განახლებითა და საუკეთესო პრაქტიკის შესაბამისად, აღწერონ გენერირებადი ხელოვნური ინტელექტის სისტემები და აპლიკაციები. დაბოლოს, შესაბამისობის საკითხების განხილვის მიზნით, მონაცემთა დაცვის ოფიცერი ჩართული უნდა იყოს მოდელების მიმწოდებლებთან მონაცემთა გაზიარებასთან დაკავშირებით დასადები ხელშეკრულებების ხელმოწერის პროცესში.

ორგანიზაციული პერსპექტივიდან გამომდინარე, რეგულაციასთან შესაბამისობა, გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენების პროცესში, ვერ

მიიღწევა მხოლოდ ერთი ინდივიდის ძალისხმევით. ამისათვის აუცილებელია ხელოვნური ინტელექტის ფუნქციონირების ყველა ეტაპზე ჩართულ, დაინტერესებულ პირებთან მუდმივი დიალოგი. ამდენად, დამუშავებისთვის პასუხისმგებელ პირებს ორგანიზაციაში უნდა ჰქონდეთ კომუნიკაცია: მონაცემთა დაცვის ოფიცერთან; იურიდიულ მიმართულებასთან; ინფორმაციული ტექნოლოგიების მიმართულებასთან; ადგილობრივი საინფორმაციო უსაფრთხოების ოფიცერთან ("Local Informatics Security Officer"), რათა უზრუნველყოფილი იქნას "EUI"-ის მიერ სანდო გენერირებადი ხელოვნური ინტელექტით მუშაობა, მონაცემთა შესაბამისი მმართველობა და რეგულაციასთან შესაბამისობა. ამ მიზნების მიღწევის ხელშემწყობი ფაქტორებია: ხელოვნური ინტელექტის სამუშაო ჯგუფის შექმნა, მათ შორის, მოიაზრება მონაცემთა დაცვის ოფიცრის ინსტიტუტი; სამოქმედო გეგმის შემუშავება; ორგანიზაციის ცნობიერების ამაღლება; შიდაორგანიზაციული გამოყენების სახელმძღვანელო დოკუმენტის შემუშავება.

➲ ხელშეკრულების პირობების საილუსტრაციოდ, ევროპულმა კომისიამ, ხელოვნური ინტელექტის სისტემების ("Procurement of AI Community") შესყიდვის ინიციატივის ფარგლებში, შეკრიბა "AI" სისტემების შესყიდვის მიმართულებით დაინტერესებული პირები. აღნიშნულის მიზანს ხელშეკრულების სტანდარტული პირობების შემუშავება წარმოადგენდა, რომელიც დაარეგულირებდა საჯარო ორგანიზაციების მიერ ხელოვნური ინტელექტის სისტემების შესყიდვასთან დაკავშირებულ საკითხებს. ამასთანავე, მნიშვნელოვანია, დამუშავებისთვის პასუხისმგებელ პირებსა და დამუშავებაზე უფლებამოსილ პირებს შორის გასაფორმებელი ხელშეკრულების სტანდარტული პირობების განხილვა რეგულაციის დებულებების[4] გათვალისწინებით.

[4] რეგულაციის 39(2)-ე მუხლი.

5. "EUI"-ის მხრიდან გენერირებადი ხელოვნური ინტელექტის სისტემების შემუშავების პროცესში როდის უნდა განხორციელდეს მონაცემთა დაცვაზე ზეგავლენის შეფასება ("DPIA")?

მონაცემთა დაცვის სტანდარტების გათვალისწინება ახალი პროდუქტის ან მომსახურების შექმნისას და მონაცემთა დაცვა პირველად პარამეტრად ("By Design and By Default")[5] წარმოადგენს მონაცემთა დაცვის პრინციპების უზრუნველყოფის ხელშემწყობ ფაქტორს მონაცემთა დამუშავების ყველა ეტაპზე, მათ შორის, ასევე, მოიაზრება შექმნის ფაზა. რეგულაციის პრინციპის გათვალისწინებით, რისკზე ორიენტირებულ მიდგომაზე დაყრდნობით, შესაძლებელია, განხილული იქნას გენერირებადი ხელოვნური ინტელექტის გამოყენების თანმხლები რისკები და საფრთხეები, რათა პროაქტიულად იქნას მიღებული შესაბამისი ზომები. სისტემის შემმუშავებლებსა და მის დამნერგავ პირებს შეიძლება, დასჭირდეთ რისკის შეფასება და მის შესამცირებლად მიღებული ღონისძიებების აღრიცხვა.

რეგულაციის შესაბამისად, მონაცემთა დაცვაზე ზეგავლენის შეფასება[6] უნდა განხორციელდეს იმ მონაცემების დამუშავებამდე, რომლებიც დიდი ალბათობით წარმოშობს მაღალ რისკს[7] ადამიანის უფლებებსა და თავისუფლებებთან მიმართებით. რეგულაციის თანახმად, ყურადღებაა გამახვილებული ახალი ტექნოლოგიების გამოყენებისას მონაცემთა დაცვაზე ზეგავლენის შეფასების საჭიროებაზე.

დამუშავებისთვის პასუხისმგებელი პირი ვალდებულია, მონაცემთა დაცვაზე ზეგავლენის შეფასების განხორციელებისას, კონსულტაცია გაიაროს მონაცემთა დაცვის ოფიცერთან. შეფასების განხორციელების შედეგად გამოვლენილი რისკის შესამცირებლად, ასევე, უნდა იქნას მიღებული შესაბამისი ტექნიკური და ორგანიზაციული ზომები.

გენერირებადი ხელოვნური ინტელექტის გამოყენებისას, მნიშვნელოვანია ინფორმაციის მიღება მათგან, ვისზეც სისტემებმა უარყოფითი გავლენა იქონია. გარდა მონაცემთა დაცვაზე ზეგავლენის შეფასებისა ("DPIA"), მნიშვნელოვანია რისკის შეფასების მუდმივი მონიტორინგი, რაც ხელს შეუწყობს გამოვლენილი რისკების

[5] რეგულაციის 27-ე მუხლი.
[6] რეგულაციის 39-ე და 89-ე მუხლები.
[7] "AI" აქტის თანახმად, "AI" სისტემების კლასიფიკაცია „მაღალ რისკად" ადამიანის უფლებებზე მისი გავლენის გათვალისწინებით, აპრიორი გულისხმობს "GDPR"-ით, "EUDPR"-ითა და "LED" დირექტივით დადგენილ „მაღალ რისკს".

ეფექტიან მართვას. აღნიშნული რისკები უკავშირდება როგორც პერსონალური მონაცემებს, ასევე სხვა ძირითად უფლებებსა და თავისუფლებებს.

მონაცემთა დაცვაზე ზეგავლენის შეფასების პროცესში ჩართული ყველა პირი ვალდებულია, ნებისმიერი გადაწყვეტილება და მოქმედება შესაბამისად აღრიცხოს, რაც მიემართება გენერირებადი ხელოვნური ინტელექტის ფუნქციონირების ყველა ეტაპს, მათ შორის, რისკის სამართავად მიღებულ ზომებს.

"EUI"-ების პასუხისმგებლობაა გენერირებადი ხელოვნური ინტელექტის გამოყენებისას რისკების შესაბამისად მართვა. აუცილებელია მონაცემთა დაცვასთან დაკავშირებული რისკების იდენტიფიცირება და მათი გათვალისწინება გენერირებადი ხელოვნური ინტელექტის ფუნქციონირების ყველა ეტაპზე. აღნიშნული მოიაზრებს გამოვლენილ რისკებზე მონიტორინგის განხორციელებას და ახალი რისკების პრევენციას. გენერირებადი "AI" სისტემების გამოყენებასთან დაკავშირებული რისკები მიმდინარე პროცესის ნაწილია, რომელთან მიმართებითაც მნიშვნელოვანია ფრთხილი მიდგომა. თუ გამოვლენილი რისკების შემცირება გონივრული საშუალებებით შეუძლებელია, აუცილებელია "EDPS"-სთან კონსულტაცია.

➲ "EDPS"-მა შეიმუშავა შაბლონი, რომელიც დაეხმარება დამუშავებისთვის პასუხისმგებელ პირებს იმის განსაზღვრაში, თუ როდის უნდა განახორციელონ მონაცემთა დაცვაზე ზეგავლენის შეფასება (ანგარიშვალდებულებასთან დაკავშირებული საკითხები, მე-6 დანართი, ნაწილი I). დამატებით, "EDPS"-მა შეიმუშავა დამუშავების ოპერაციებთან დაკავშირებული ე. წ. „ღია სია", რომელიც შეეხება მონაცემთა დაცვაზე ზეგავლენის შეფასებას. საჭიროების შემთხვევაში, დამუშავებისთვის პასუხისმგებელმა პირმა უნდა შეაფასოს მონაცემთა დამუშავების შესაბამისობა მონაცემთა დაცვაზე ზეგავლენის შეფასების სტანდარტებთან. თუ შეფასების შედეგად დამუშავებისთვის პასუხისმგებელ პირებს არ აქვთ შესაძლებლობა დარწმუნდნენ, რომ რისკები შემცირდა, მიზანშეწონილია, "EDPS"-სთან კონსულტაციის გავლა.

6. კანონიერია თუ არა პერსონალური მონაცემების დამუშავება გენერირებადი ხელოვნური ინტელექტის შექმნის, განვითარებისა და გამოყენებისთვის?

პერსონალური მონაცემების გენერირებადი ხელოვნური ინტელექტის საშუალებით მისსავე სისტემაში დამუშავება მოიცავს ყველა იმ აქტივობას, რომელიც მონაცემთა შეგროვებას, მთლიან პროცესს, სისტემასა და სისტემის მიზნობრივ გამოყენებას უკავშირდება. მონაცემთა შეგროვების პროცესი დაკავშირებულია ინტერნეტში საჯაროდ ხელმისაწვდომი წყაროებიდან, მესამე პირისგან ან საკუთარი სისტემური ფაილებიდან მონაცემთა მოპოვებასთან. მონაცემები, ასევე, შეიძლება მოპოვებული იქნეს ინფორმაციული სისტემის მომხმარებლებისგან, სისტემაში მათი აქტიურობისა და განხორციელებული მოქმედებების აღრიცხვის საფუძველზე. გენერირებადი ხელოვნური ინტელექტის საშუალებით მონაცემთა დამუშავება ხორციელდება პერსონალური მონაცემების სისტემური და ფართომასშტაბიანი მოპოვების გზით, რომლის შესახებ, ხშირ შემთხვევაში, მონაცემთა სუბიექტებს არ აქვთ ინფორმაცია.

ევროკავშირის ინსტიტუტების, ორგანოების, ოფისების და სააგენტოების (EUIs) მიერ პერსონალური მონაცემების დამუშავება კანონიერია, თუ არსებობს მონაცემთა დამუშავების სულ მცირე ერთ-ერთი საფუძველი მაინც.[8] განსაკუთრებული კატეგორიის პერსონალური მონაცემების დამუშავების კანონიერებისთვის უნდა არსებობდეს რეგულაციით განსაზღვრული ერთ-ერთი პირობა.[9] საჯარო ინტერესიდან[10] ან სამართლებრივი ვალდებულებიდან[11] გამომდინარე მონაცემთა დამუშავება უნდა შეესაბამებოდეს ევროკავშირის კანონმდებლობას. შესაბამისად, ევროკავშირის კანონმდებლობაში ნათლად, ზუსტად და განჭვრეტადი ფორმით უნდა განისაზღვროს პროცედურა, რომელიც, თავის მხრივ, ევროკავშირის ძირითად უფლებათა ქარტიასა და კონვენციას შეესაბამება.

მაშინ, როდესაც სამართლებრივი საფუძველი, მისი გამოყენებისას ძირითად უფლებასთან მიმართებით განსხვავებული განმარტების შესაძლებლობას ქმნის, მნიშვნელოვანია, არსებობდეს ნათელი და ზუსტი წესები პროცედურის მიზანთან დაკავშირებით. რაც უფრო დიდია სხვაობა, უფრო მეტად გასაგები უნდა იყოს წესი, რომელიც გამოიყენება უფლების დაცვის საკითხისათვის. ეროვნული კანონმდებლობა, შიდა წესები შესაძლებელია, განსაზღვრავდეს მიზანს, მონაცემთა კატეგორიებს, დამუშავებისთვის პასუხისმგებელ პირებსა და დამუშავებაზე

[8] რეგულაციის მე-5 მუხლი.
[9] რეგულაციის მე-10 მუხლის მე-2 პუნქტი.
[10] რეგულაციის მე-5 მუხლის პირველი პუნქტის (a) ქვეპუნქტი.
[11] რეგულაციის მე-5 მუხლის პირველი პუნქტის (b) ქვეპუნქტი.

უფლებამოსილ პირებს, შენახვის ვადებს, მონაცემთა დაცვის მინიმალურ სტანდარტს და მონაცემთა სუბიექტების უფლებების რეალიზების ხელშეწყობას.

აღნიშნული ფაქტობრივი მოცემულობა[12] გამოიყენება როგორც საფუძველი გენერირებადი ხელოვნური ინტელექტის საშუალებით მონაცემთა დამუშავების მიზნებისთვის. ფაქტობრივი მოცემულობა,[13] რომელსაც რეგულაციის მოქმედების ფარგლები მიემართება, საჭიროებს სამართლებრივი მოთხოვნების შესრულებას, უფლებამოსილების სათანადო გამოყენებას. სივრცე, სადაც გენერირებადი ხელოვნური ინტელექტის საშუალებით მუშავდება პერსონალური მონაცემები ყურადღებით უნდა შეირჩეს, ყველა იმ ფაქტობრივი მოცემულობის გათვალისწინებით, რომელთა მოთხოვნებსაც უნდა აკმაყოფილებდეს ამ სივრცეში მონაცემთა დამუშავება.

ყველა მოქმედება, რომელიც ფაქტობრივი სამართლებრივი მოცემულობის ფარგლებში ხორციელდება — რეგულაციასთან შესაბამისობაშია. სხვა შემთხვევაში, როდესაც არ იარსებებს რეგულაციასთან შესაბამისობა და სამართლებრივი საფუძვლები, მაშინ მონაცემთა დამუშავებისთვის პასუხისმგებელმა პირმა უნდა შეწყვიტოს მონაცემთა დამუშავება და წაშალოს ამ დროის განმავლობაში მოპოვებული მონაცემები.

„ინფორმაციული სისტემის მიმწოდებლების“ („პროვაიდერების“) ევროკავშირის „მონაცემთა დაცვის ძირითადი რეგულაციის“ ფარგლებში[14] მონაცემთა დამუშავების საფუძვლად ასახელებენ ლეგიტიმურ საჯარო ინტერესს, რომელიც გამოიხატება მონაცემების შეგროვებასა და მათ გამოყენებაში სისტემის განსავითარებლად, პროცესის უზრუნველსაყოფად. ევროკავშირის მართლმსაჯულების სასამართლოს განმარტების თანახმად,[15] ლეგიტიმური ინტერესის გამოყენება სამ კუმულაციურ პირობას უნდა აკმაყოფილებდეს, რომ აღნიშნული საფუძველი იყოს კანონიერი. პირველი, ლეგიტიმური ინტერესი უნდა გამოხატოს მონაცემთა დამუშავებისთვის პასუხისმგებელმა პირმა ან მესამე პირმა; მეორე, უნდა არსებობდეს მისი გამოყენების საჭიროება და მესამე, მონაცემთა დაცვის ინტერესი, უფლება ან თავისუფლება არ

[12] რეგულაციის მე-5 მუხლის პირველი პუნქტის (d) ქვეპუნქტი და მე-7 მუხლი.

[13] EDPB-ის გაიდლაინი 05/2020,
https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf

[14] ევროპარლამენტისა და საბჭოს 2016 წლის 27 აპრილის რეგულაცია (EU) 2016/679 პერსონალურ მონაცემთა დამუშავებისას ფიზიკურ პირთა დაცვისა და ასეთი მონაცემების თავისუფალი მიმოცვლის შესახებ, რომელიც აუქმებს 95/46/EC დირექტივას (მონაცემთა დაცვის ზოგადი რეგულაცია).

[15] 2023 წლის 4 ივლისის გადაწყვეტილება, Meta პლატფორმა და სხვები (სოციალური ქსელის გამოყენების ძირითადი პირობები), C-252/21, EU:C:2023:537, პარაგრაფი106.

უნდა იყოს ლეგიტიმურ ინტერესზე უპირატესი, იმ შემთხვევაში, როდესაც მონაცემები მუშავდება გენერირებადი ხელოვნური ინტელექტის მიერ, რომელზეც ბევრ მახასიათებელს შეუძლია გავლენის მოხდენა დამუშავებისთვის პასუხისმგებელი პირის მიერ განსაზღვრულ ან მოსალოდნელ შედეგებთან მიმართებით. სწორედ ამიტომ, ევროკავშირის ინსტიტუტებს, ორგანოებს, ფილიალებსა და სააგენტოებს აქვთ განსაკუთრებული (სპეციალური) პასუხისმგებლობა დარწმუნდნენ, რომ გენერირებადი ხელოვნური ინტელექტის „პროვაიდერები“ იცავენ იმ პირობებსა და მოთხოვნებს, რომლებსაც განსაზღვრავს ევროკავშირის კანონმდებლობა.

ევროკავშირის ინსტიტუტები, ორგანოები, ფილიალები და სააგენტოები (“EUI”), როგორც მონაცემთა დამუშავებისთვის პასუხისმგებელი პირები, ანგარიშვალდებულნი არიან პერსონალური მონაცემების ევროპული ეკონომიკური ზონის ფარგლებს მიღმა გადაცემაზე. მსგავსი გადაცემა მოხდება იმ შემთხვევაში, როდესაც ევროკავშირის ინსტიტუტები, ორგანოები, ფილიალები და სააგენტოები ახორციელებენ უფლებამოსილებას ევროკავშირის კანონმდებლობის ან წევრი სახელმწიფოს კანონმდებლობის საფუძველზე. ხელოვნური ინტელექტის განვითარების ან გამოყენების მიზნებისთვის, მონაცემთა გადაცემა სხვადასხვა ეტაპებს გულისხმობს — ეს მოიცავს შემთხვევას, როდესაც “EUI” იყენებს შესაბამის სერვერებს ან ისეთ შემთხვევას, როდესაც ხორციელდება სისტემის სწავლება, ფუნქციების შემოწმება და კონკრეტული მოდელის ვალიდაცია. ამ შემთხვევაში მონაცემთა გადაცემა რეგულაციის მეხუთე თავით განსაზღვრულ წესს უნდა შეესაბამებოდეს[16] და მონაცემთა დამუშავების თავდაპირველ მიზანთან უნდა იყოს შესაბამისი.

ხელოვნური ინტელექტით მართულ სისტემებში პერსონალური მონაცემების დამუშავება მოითხოვს შესაბამისი საფუძვლების არსებობას, რომლებიც რეგულაციიდან უნდა გამომდინარეობდეს. თუკი დამუშავება გამომდინარეობს სამართლებრივი ვალდებულების შესრულების ინტერესიდან ან საჯარო უწყების უფლებამოსილებიდან, დამუშავების საფუძველი და მიზანი უნდა იყოს ნათელი და ევროკავშირის კანონმდებლობასთან შესაბამისი.

[16] რეგულაციის 46-ე და 51-ე მუხლები.

➲ მაგალითად, „პირადი ცხოვრების ხელშეუხებლობის გლობალური ასამბლეის“ რეზოლუცია გენერირებადი ხელოვნური ინტელექტის სისტემებთან მიმართებით ადგენს, რომ შესაბამისი კანონმდებლობის ფარგლებში, გენერირებად ხელოვნურ ინტელექტთან დაკავშირებული პირები უნდა განსაზღვრავდნენ დამუშავების საფუძველს: ა) გენერირებადი ხელოვნური ინტელექტის სისტემის განსავითარებლად მონაცემთა შეგროვების გზით; ბ) პროცესის სატესტო რეჟიმში გადაყვანით, რისთვისაც საჭიროა მონაცემთა ბაზებზე წვდომა სისტემის ფუნქციონირების შესასწავლად; გ) ფიზიკური პირების მონაწილეობის, უშუალოდ პროცესში ჩართვის მასშტაბები; დ) მოქმედების შედეგების/შინაარსის ჩამოყალიბება ხელოვნური ინტელექტის სისტემის მიერ.

7. როგორ მოქმედებს მონაცემთა მინიმიზაციის პრინციპი გენერირებადი ხელოვნური ინტელექტის გამოყენებისას?

მონაცემთა მინიმიზაციის პრინციპის თანახმად, მონაცემთა დამუშავებისთვის პასუხისმგებელმა პირი უნდა დარწმუნდეს, რომ მონაცემთა დამუშავება შეესაბამება დამუშავების საფუძველს/წესებს და შეზღუდულია იმ თავდაპირველი მიზნით, რომლისთვისაც ხორციელდება დამუშავება. ზოგიერთი მოსაზრების თანახმად, მონაცემთა მინიმიზაციის პრინციპი არ ვრცელდება ხელოვნური ინტელექტის სისტემების მიმართ.[17] მიუხედავად ამისა, მონაცემთა დამუშავებისთვის პასუხისმგებელ პირებს აქვთ ვალდებულება, შეზღუდონ მონაცემთა დამუშავება იმ ფარგლებით, რომლებიც განსაზღვრავს დამუშავების მიზანს. ეს ვალდებულება მიემართება სისტემის მოქმედების ყველა ეტაპს, რაც მოიცავს ტესტირებას, ფუნქციონირებასა და საბოლოო პროდუქტის მიღებას. აღსანიშნავია, რომ პერსონალური მონაცემები არ უნდა დამუშავდეს ყოველთვის, ნებისმიერ შემთხვევაში. ევროკავშირის ინსტიტუტები, ორგანოები, ოფისები და სააგენტოები უნდა დარწმუნდნენ, რომ ხელოვნური ინტელექტის განვითარებაში ჩართული პერსონალი გამოიყენებს მონაცემთა მინიმიზაციის პრინციპს სისტემის ფუნქციონირების თითოეულ ეტაპზე.

ამ მიზნით, მათ შეუძლიათ განავითარონ და გამოიყენონ ისეთი მაღალი ხარისხის მონაცემთა ბაზები, რომლებსაც მონაცემთა მინიმიზაციის პრინციპის საშუალებით დაუკავშირებენ მონაცემთა მოპოვებისა და დამუშავების კერძო მიზანს. ამგვარად შესაძლებელი იქნება ერთგვარი ჩარჩოს შექმნა სისტემის მართვის პროცედურებისთვის, პროცესის პერიოდული და სისტემური კონტროლის მიზნით. მონაცემთა ბაზები და მოდელები უნდა ფუნქციონირებდეს შედეგების დოკუმენტირების გზით, სტრუქტურისა და საჭიროებების აღრიცხვით. მათ შორის, რომელიც შექმნილია და იმართება მესამე მხარის სერვის-პროვაიდერების მიერ, ევროკავშირის ინსტიტუტებმა, ორგანოებმა, ფილიალებმა და სააგენტოებმა უნდა გაითვალისწინონ მონაცემთა მინიმიზაციის პრინციპი.

ხელოვნური ინტელექტის სისტემების ფუნქციონირების შესამოწმებლად დიდი მოცულობის მონაცემების შეგროვებისა და დამუშავების პროცესი შეიძლება ეფექტიანობის გაზრდასა და შედეგების მიღებამდე, ითვალისწინებდეს კარგად

[17] რეგულაციის მე-5 მუხლის პირველი პუნქტის (c) ქვეპუნქტის თანახმად, პერსონალური მონაცემები უნდა იყოს ადეკვატური, რელევანტური და დამუშავდეს მხოლოდ იმ მოცულობით, რომელიც აუცილებელია დამუშავების მიზნების მისაღწევად (მონაცემთა მინიმიზაცია).

სტრუქტურირებული მონაცემთა ბაზების არსებობას, რომლებიც გამოყენებული იქნება პროცესის ხარისხის გასაუმჯობესებლად, ფუნქციონირების შემოწმებისა და კონტროლისთვის ფორმალური თვალსაზრისით. ასეთ შემთხვევაში მონაცემთა მინიმიზაციის პრინციპი სისტემის გამართული ფუნქციონირებისთვის საჭირო მნიშვნელობას იძენს.

➲ ხელოვნური ინტელექტის სისტემები შეიძლება გამოიყენებოდეს სისტემური პროგრამული სამუშაოებისთვის. ამ შემთხვევაში, საჭიროა მომხმარებლების მონაცემებზე ხელმისაწვდომობა, რათა სისტემამ შეძლოს მონაცემებისა და სისტემური დავალებების ერთმანეთთან დაკავშირება. სატესტო რეჟიმის ამოქმედებამდე, უნდა შემოწმდეს შერჩეული ალგორითმის შესაბამისობა დამუშავების მიზანთან. მაგალითად, სტატისტიკური ანალიზის ჩასატარებლად გამოყენებული იქნეს მცირე მოცულობის მონაცემები, რომლებითაც ის შედეგი მიიღწევა, რაც მეტი მოცულობის მონაცემების გამოყენებისას. აქვე, უნდა შემოწმდეს იყენებს თუ არა სისტემა განსაკუთრებული კატეგორიის პერსონალურ მონაცემებს. ამ შემთხვევაში შემოთავაზებული უნდა იყოს მონაცემთა დეპერსონალიზაციის, ფსევდონიმიზირების ფორმები. დასკვნის სახით, აუცილებელია ყველა რელევანტური ტექნიკური საშუალების და სამართლებრივი საფუძვლის გაერთიანება დამუშავების კანონიერების, გამჭვირვალობისა და სიზუსტის დასასაბუთებლად.

8. შეესაბამება თუ არა ხელოვნური ინტელექტის სისტემები მონაცემთა სიზუსტის პრინციპს?

გენერირებადი ხელოვნური ინტელექტი შეიძლება გამოიყენებოდეს სისტემის ფუნქციონირების ფარგლებში, დიდი მოცულობის ინფორმაციის, მათ შორის, პერსონალური მონაცემების გამოყენების ეტაპზე.

მონაცემთა სიზუსტის პრინციპის[18] თანახმად, მონაცემი უნდა იყოს ზუსტი, განახლებული, ხოლო მონაცემთა დამუშავებისთვის პასუხისმგებელმა პირმა უნდა განაახლოს ან წაშალოს არაზუსტი მონაცემები. მონაცემთა დამუშავებისთვის პასუხისმგებელი პირები უნდა დარწმუნდნენ, რომ მონაცემთა სიზუსტე უზრუნველყოფილია ხელოვნური ინტელექტის შემუშავებისა და გამოყენების ყველა ეტაპზე. მათ უნდა დანერგონ იმგვარი მექანიზმები, რომლებიც გაზრდის მონაცემთა სიზუსტეს სისტემის მოქმედების თითოეულ ეტაპზე.

აღნიშნული შეიძლება, მოიცავდეს მესამე მხარისგან მიღებულ მონაცემებსაც. თანაბრად მნიშვნელოვანია, რომ ხელოვნური ინტელექტის სისტემისთვის მიწოდებული მონაცემების დამუშავება მოწმდებოდეს ადამიანური რესურსის ჩართულობით. ამისთვის გამოყენებული უნდა იყოს პროგრამული უზრუნველყოფის მეთოდები,[19] რომლებიც შეაფასებს სისტემის მიერ მიღებული ინფორმაციის, ასევე, მონაცემებთან დაკავშირებული შედეგების სიზუსტეს. ეს შეიძლება იყოს მიღწეული არა მხოლოდ მონაცემთა დაცვის, არამედ სტატისტიკური სიზუსტის საზომი ერთეულების დანერგვით გზით. სიზუსტის დამდგენი კრიტერიუმი შეიძლება, დაინერგოს მიმდინარე პროცესებთან ან პროცესის შედეგად მისაღებ სამომავლო შედეგის განსაზღვრასთან დაკავშირებით.

როდესაც ევროკავშირის ინსტიტუტები გენერირებად ხელოვნურ ინტელექტს იყენებენ სატესტო რეჟიმში ან მესამე მხარის ჩართულობით მონაცემთა შეგროვების მიზნებისთვის, აუცილებელია, რომ ხელშეკრულებით განისაზღვროს პროცედურები, რომლებიც გამოყენებული იქნება ამ სისტემის ინტეგრირების, განვითარებისა და მოქმედებისთვის. ეს შეიძლება, აგრეთვე, მოიცავდეს მონაცემთა შეგროვების, მოსამზადებელ პროცედურებს, იმგვარი რისკების განსაზღვრას, რომელმაც შეიძლება იქონიოს გავლენა მონაცემთა სიზუსტეზე. ტექნიკური გამართულობა და

[18] რეგულაციის მე-4 მუხლის პირველი პუნქტის (d) ქვეპუნქტი.

[19] აღნიშნული პარამეტრები გამოიყენება სისტემის ძირითადი პარამეტრების დაზუსტებისა და მისი შესრულების ხარისხის შესაფასებლად.

დოკუმენტური მხარე მონაცემთა დამუშავებისთვის პასუხისმგებელ პირს შესაძლებლობას აძლევს, შეიმუშაოს მოდელი, რომლის გამოყენებითაც იგი შეძლებს პროცესსა და მიღებულ შედეგებზე დაკვირვებას. აღნიშნული იმაზე მეტად მნიშვნელოვანია, ვიდრე მაღალი ხარისხით მოქმედი სისტემა არაზუსტ მონაცემებს ამუშავებდეს. ე. ი. სისტემის გამართულ მუშაობაზე მეტად მნიშვნელოვანია, ის, თუ რამდენად არის უზრუნველყოფილი სისტემაში მონაცემთა სიზუსტის პრინციპის მოქმედების ფარგლები. არაზუსტი და ყალბი მონაცემების გამოყენებამ, შეიძლება გამოიწვიოს ე. წ. „ჰალუცინაციის“ ეფექტი.

მიუხედავად უკვე განხორციელებული მოქმედებებისა, ხელოვნური ინტელექტის სისტემებში მონაცემთა სიზუსტის პრინციპის გამოყენება კვლავ გამოწვევად რჩება, რაც იწვევს არაზუსტი მონაცემების საფუძველზე მონაცემთა სუბიექტების უფლებების შელახვას.

ხელოვნური ინტელექტის სისტემის პროვაიდერები სისტემის განვითარების პროცესში უნდა დარწმუნდნენ, რომ ამ სისტემას აქვს უნარი, დაიცვას მონაცემთა სიზუსტის პრინციპი, ხოლო, თავის მხრივ, ევროკავშირის ინსტიტუტებს, ორგანოებს, ოფისებსა და სააგენტოებსაც ეკისრებათ რიგი ვალდებულებები, შეაფასონ ხელოვნური ინტელექტის სისტემების დანერგვის პროცესში მათი მოქმედება მონაცემთა სიზუსტის პრინციპის გათვალისწინებით.

➲ “EUI-X” — ევროკავშირის პირობითი ინსტიტუტს, მონაცემთა დაცვის ოფიცრის რჩევით, წარედგინათ პოლიტიკის დოკუმენტი, რომელიც ითვალისწინებს ოფიციალური შეხვედრებისა და საჯარო მოსმენების თარგმნის სიზუსტის უზრუნველყოფის მექანიზმებს. აღნიშნული დოკუმენტის თანახმად, თუ მაღალი დონის შეხვედრა ან საჯარო მოსმენა ითარგმნება ხელოვნური ინტელექტის სისტემის გამოყენებით, მაშინ შედეგებს უნდა აკონტროლებდეს და ასწორებდეს ადამიანური პერსონალი, ხოლო შედარებით ნაკლები მნიშვნელობის მქონე შეხვედრებისა და საჯარო გამოსვლების თარგმნისას, უნდა მიეთითოს, რომ დოკუმენტი ხელოვნური ინტელექტის სისტემის საფუძველზეა შემუშავებული. ასევე, ამ მოქმედებების განსახორციელებლად ხელოვნური ინტელექტის სისტემაში ინტეგრირებულია ფუნქცია, რომელიც მას უფლებას ანიჭებს, ჩაიწეროს ხმოვანი შეტყობინებები და წერილობითად დააინტეგრიროს ისინი. პოლიტიკის დოკუმენტში სწორედ ამ საკითხების მნიშვნელობაზეა ხაზი გასმული.

9. როგორ უნდა შევატყობინოთ მონაცემთა სუბიექტებს მონაცემთა ხელოვნური ინტელექტის სისტემის საშუალებით დამუშავების შესახებ?

სათანადო ინფორმაცია და გამჭვირვალობის შესახებ პოლიტიკის დოკუმენტები უზრუნველყოფს მონაცემთა სუბიექტების უფლების დარღვევის რისკის შემცირებას და რეგულაციის მოთხოვნებთან შესაბამისობას, კერძოდ, დეტალურ ინფორმაციას, თუ როგორ, როდის და რატომ ამუშავებენ ევროკავშირის ინსტიტუტები, ორგანოები, ოფისები და სააგენტოები პერსონალურ მონაცემებს გენერირებადი ხელოვნური ინტელექტის გამოყენებით. ისინი უნდა დარწმუნდნენ, რომ აქვეყნებენ სანდო, რელევანტურ და განახლებულ ინფორმაციას დამუშავების შედეგების შესახებ. ცალკეული სისტემები შეიძლება, მოითხოვდეს გამჭვირვალობასთან დაკავშირებით დამატებით განმარტებებს, როდესაც ისინი ინფორმაციას გენერირებადი ფორმით ადგენენ, უშუალოდ მომხმარებელთან ინტერაქციით, ადამიანური რესურსების ჩარევის გარეშე.

ინფორმირების უფლება მოიცავს ვალდებულებას,[20] რათა ფიზიკურმა პირებმა პროფაილინგის ან მონაცემთა ავტომატიზებული ფორმით დამუშავებისას მიიღონ საჭირო ინფორმაცია იმ ალგორითმის შესახებ, რომელიც გამოიყენება შედეგების მისაღებად. ეს შეეხება მათი განახლების მნიშვნელობასაც, რა პრინციპით მუშაობს ალგორითმი და განახლდება თუ არა შუალედური შედეგები საბოლოო შედეგით. აღნიშნული ვალდებულება სასურველია, ასევე, გავრცელდეს ისეთ შემთხვევებზეც, როდესაც პროცესი არ არის სრულიად ავტომატური და მოიცავს ნაწილობრივ ავტომატურ დამუშავებას.

რეგულაციის თანახმად, ევროკავშირის ინსტიტუტებმა, ორგანოებმა, ოფისებმა და სააგენტოებმა მონაცემთა სუბიექტებს უნდა მიაწოდონ სათანადო ინფორმაცია. მონაცემთა სუბიექტებისთვის გაზიარებული ინფორმაცია უნდა განახლდეს ეტაპობრივად, როდესაც აღნიშნულის საჭიროება ცალსახად დგას.

➲ მაგალითად, ევროკავშირის ცალკეულ ვებგვერდზე ჩატ-ბოტის დასაინტეგრირებლად დაწყებულია გარკვეული სამუშაოები. მონაცემთა დაცვის ოფიცრის რჩევით, საიტზე განთავსდა აღნიშვნა იმის შესახებ, თუ რა ინფორმაციას ამუშავებს ჩატ-ბოტი, რა მიზნით და როგორ. მიეთითა მონაცემთა დაცვის ოფიცრის საკონტაქტო ინფორმაცია, მონაცემთა მიმღებები, მონაცემთა კატეგორია, რომელიც გროვდება. ასევე, განცხადებაში მითითებულია სისტემის მუშაობის პროცედურა.

[20] რეგულაციის მე-14 მუხლი.

ამგვარ აღნიშვნას დამუშავებისთვის პასუხისმგებელი პირი იყენებს, როგორც სამართლებრივ საფუძველს, რომელზე მონაცემთა სუბიექტის თანხმობა მას ანიჭებს უფლებას, აამოქმედოს ჩატ-ბოტი.

10. რას გულისხმობს ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღება რეგულაციის 24-ე მუხლის კონტექსტში?

გენერირებადი ხელოვნური ინტელექტის სისტემის გამოყენება არ გულისხმობს ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღებას რეგულაციის 24-ე მუხლის საფუძველზე.[21] მიუხედავად ამისა, არსებობს გენერირებადი ხელოვნური ინტელექტის სისტემები, რომლებშიც ინტეგრირებულია პროფაილინგისა და ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღების ფუნქციები. ევროკავშირის ინსტიტუტებმა, ორგანოებმა, ოფისებმა და სააგენტოებმა უფლებამოსილების განხორციელების ეტაპზე, ხელოვნური ინტელექტის სისტემის საფუძველზე გადაწყვეტილების მიღებისას, უნდა გაითვალისწინონ რეგულაციის 24-ე მუხლის მოთხოვნები, რომლის თანახმად, მონაცემთა სუბიექტს უფლება აქვს, არ დაექვემდებაროს ავტომატიზებული ფორმით მიღებულ გადაწყვეტილებას.

ხელოვნური ინტელექტის სისტემების ტექნიკური გამართულობისთვის, გათვალისწინებული უნდა იქნეს პროგრამული გადაწყვეტილების პროცესში ადამიანური პერსონალის მონაწილეობის საკითხები, მაგალითად, იმგვარი სისტემის შექმნისას, რომელიც დამოუკიდებლად მიიღებს გადაწყვეტილებებს და წარმართავს პროცესებს.

დამუშავებისთვის პასუხისმგებელი პირის საბოლოო გადაწყვეტილებას აქვს განსაკუთრებული მნიშვნელობა მაშინაც, როდესაც ავტომატიზებული ინდივიდუალური გადაწყვეტილებით ქვეყნდება საბოლოო დასკვნა. ამ დროს უნდა შეფასდნენ მონაცემთა სუბიექტების მოწყვლადი ჯგუფები და არასრულწლოვნები,[22] რომლებიც ექცევიან ხელოვნური ინტელექტის სისტემის მიერ მიღებული გადაწყვეტილების მოქმედების ფარგლებში, რა რისკები და საფრთხეები შეიძლება, წარმოიშვას ამ პირთა მიმართ საბოლოო გადაწყვეტილების მიღებისას.

ევროკავშირის ინსტიტუტებმა, ორგანოებმა, ოფისებმა და სააგენტოებმა ხელოვნური ინტელექტის მიერ ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღებისას უნდა განსაზღვრონ წესები, რომლებიც მათ დაეხმარება, გაითვალისწინონ სისტემის მიერ მონაცემთა დამუშავების კანონიერების დარღვევის, არაეთიკური და დისკრიმინაციული გადაწყვეტილებების მიღების თავიდან არიდებაში.

[21] რეგულაციის 24-ე მუხლი.

[22] პირადი ცხოვრების ხელშეუხებლობის ასამბლეა (GPA) (2023). რეზოლუცია გენერირებადი ხელოვნური ინტელექტის სისტემების შესახებ

➲ დასაქმების აპლიკაციაში შემოსული განაცხადების შეფასების მიზნით, შესაძლებელია, რომ “EUI-X” — ევროკავშირის პირობითი ინსტიტუტი იყენებდეს ხელოვნური ინტელექტის სისტემას. ხელოვნური ინტელექტის სისტემა უზრუნველყოფს ავტომატიზებული ფორმით ინდივიდუალური გადაწყვეტილების მიღებას, ფორმალური წინაპირობების შეფასებას, ქულათა მინიჭებას, რანჟირებასა და აპლიკანტის შემდეგ ეტაპზე გადაყვანას. მონაცემთა სიზუსტისა და მიუკერძოებლობის უზრუნველყოფის მიზნით, დამუშავებისთვის პასუხისმგებელმა პირმა გადაწყვიტა, არ გამოეყენებინა სისტემა, ვიდრე იარსებებდა გარკვეული რისკები მისი გამოყენების მიმართ.

ნებისმიერ შემთხვევაში სისტემა მიიჩნევა დამუშავების მიზანზე მორგებულად და რეგულაციასთან შესაბამისად, თუ იგი ითვალისწინებს რეგულაციის 24-ე მუხლის მე-2 პუნქტით გათვალისწინებულ გამონაკლისებსაც. ამ გამონაკლისების იმპლემენტირება უფლების დაცვას ემსახურება და გამოხატავს მონაცემთა სუბიექტის უფლებას, არ დაექვემდებაროს მხოლოდ ავტომატიზებული ინდივიდუალური ფორმით მიღებულ გადაწყვეტილებას.

რეგულაციის მე-15 მუხლის მე-2 პუნქტის (f) ქვეპუნქტის თანახმად, ალგორითმის შესახებ ინფორმაცია უნდა მიეწოდოს მონაცემთა სუბიექტს, როდესაც მონაცემი გროვდება ფიზიკური პირისგან და წარმოშობს შედეგებს მასზე.

ამ შემთხვევაში, მონაცემთა დამუშავებისთვის პასუხისმგებელი პირი უფლებამოსილია, გამოიყენოს გენერირებადი ხელოვნური ინტელექტის სისტემა ან ავტომატიზებული ინდივიდუალური გადაწყვეტილების მიღების მარტივი ფორმა. თითოეულ შემთხვევაში სწორად უნდა შეირჩეს ის მექანიზმი, ალგორითმი ან სისტემა, რომლითაც შეფასდება აპლიკანტის მახასიათებლები.

11. გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენებისას, როგორ უნდა იქნეს უზრუნველყოფილი კანონიერი დამუშავება მიკერძოების გარეშე?

ზოგადად, ხელოვნური ინტელექტის გადაწყვეტილებები ("AI solutions") ხშირად ზრდის არსებული ადამიანური მიკერძოებების ხარისხს, უფრო მეტიც, შესაძლოა, ახალი რისკიც კი გამოიწვიოს, რაც სამართლებრივი შესაბამისობის თვალსაზრისით, უკავშირდება გარკვეულ ეთიკურ გამოწვევებს. მიკერძოება შეიძლება წარმოიშვას გენერირებადი ხელოვნური ინტელექტის სისტემის განვითარების ნებისმიერ ეტაპზე, იქნება ეს „მონაცემთა სწავლებისას" ("data training"), ალგორითმების მომზადებისას თუ იმ პირთა მონაწილეობით, რომლებიც ამუშავებენ ან იყენებენ სისტემას. გენერირებად AI სისტემებში მიკერძოებამ, შესაძლოა, გამოიწვიოს უარყოფითი შედეგები, მათ შორის, ფიზიკური პირების ძირითადი უფლებებისა და თავისუფლების შელახვა, უკანონო დამუშავება, დისკრიმინაცია განსაკუთრებით ისეთ სფეროებში, როგორებიცაა: ადამიანური რესურსების მართვა, საზოგადოებრივი ჯანდაცვა, სამედიცინო დახმარება, სოციალური სერვისების მიწოდება, სამეცნიერო და საინჟინრო პრაქტიკა, პოლიტიკური და კულტურული პროცესები, ფინანსური სექტორი, გარემო და ეკოსისტემები და ასევე, საჯარო მმართველობა.

მიკერძოება, შეიძლება, მომდინარეობდეს სხვადასხვა წყაროებიდან, როგორებიცაა: „მონაცემთა სწავლების" შაბლონები, დაზარალებული ჯგუფების შესახებ არასრული ინფორმაცია, მეთოდოლოგიური შეცდომები და თუნდაც მიკერძოება, რომელიც დანერგილია სისტემის მონიტორინგის პროცესში.

გადამწყვეტი მნიშვნელობა აქვს მონაცემთა ნაკრების ("Dataset") გამოყენებას სატესტო მოდელების შექმნის პროცესში, რომელიც რეალურ სამყაროს სამართლიანად და მიკერძოების გარეშე წარმოაჩენს. უწყვეტი მონიტორინგი და ანგარიშვალდებულება აუცილებელია მიკერძოების აღმოსაჩენად და გამოსასწორებლად. ასევე, მნიშვნელოვანია ყურადღება მიექცეს, თუ როგორ მუშავდება მონაცემები და ხორციელდება მისი დოკუმენტირება.[23] ამასთან დაკავშირებით, მნიშვნელოვანია, რომ

[23] სატესტო მონაცემების აუდიტი გამოავლენს მიკერძოებულობასა და სხვა პრობლემურ საკითხებს, ასევე, მოიძიებს ინფორმაციას, თუ როგორ ხდება სატესტო მონაცემების შეგროვება, ეტიკეტირება და ანოტაცია. აუდიტის ხარისხი და მისი შედეგები დამოკიდებულია შესაბამისი ინფორმაციის ხელმისაწვდომობაზე, მათ შორის, სატესტო მონაცემთა ნაკრების, დოკუმენტაციისა და იმპლემენტაციის დეტალებზე.

“EUI”-ებმა მიიღონ და შეიმუშაონ ტექნიკური დოკუმენტაციის მოდელები, განსაკუთრებით მონაცემთა სხვადასხვა ნაკრებისა და წყაროს გაერთიანებისას.

გენერირებადი AI სისტემების პროვაიდერები ცდილობენ, აღმოაჩინონ და შეამცირონ მიკერძოების ხარისხი მათ სისტემებში. თუმცა “EUI”-ები ყველაზე უკეთ ფლობენ საკუთარი საქმის შესახებ ინფორმაციას, აქედან გამომდინარე, ვალდებულნი არიან შეამოწმონ და რეგულარულად აკონტროლონ “AI” სისტემის შედეგები.

“EUI”-ების, როგორც საჯარო ორგანოების ვალდებულებაა, დანერგონ გარანტიები ხელოვნური ინტელექტის შედეგებზე გადაჭარბებული დამოკიდებულების თავიდან აცილების მიზნით.

მიკერძოების მინიმიზაცია და საუკეთესო პრაქტიკის გამოყენება პრიორიტეტული უნდა იყოს გენერირებადი “AI” სისტემების სასიცოცხლო ციკლის ყველა ეტაპზე, რათა განხორციელდეს მონაცემთა კანონიერი დამუშავება და დისკრიმინაციული პრაქტიკის პრევენცია. ამისთვის, საჭიროა გავიაზროთ, თუ როგორ მუშაობს ალგორითმები და მონაცემები, რომლებიც გამოიყენება მოდელის ტესტირებისთვის.

➲ მაგალითისთვის, “EU-X” აფასებს, არსებობს თუ არა შერჩევითი მიკერძოება მეტყველების ამოცნობის ავტომატურ სისტემაში (“Automated Speech Recognition System”). სათარჯიმნო სამსახურებმა შეამჩნიეს, რომ ზოგიერთი მომხსენებლის სიტყვაში აღმოჩენილია გაცილებით მეტი შეცდომა, რაც მეტყველებს ინგლისური ენის აქცენტის გაგების სირთულეზე. სისტემის შემმუშავებლებთან (“Developer”) კონსულტაციის შედეგად დადგინდა, რომ ინგლისური ენის კონკრეტული აქცენტებისთვის არსებობს სატესტო მონაცემების დეფიციტი. განსაკუთრებით პრობლემურია, როდესაც მოსაუბრე არ საუბრობს თავის მშობლიურ ენაზე. საკითხის კომპლექსური ხასიათიდან გამომდინარე, “EU-X” განიხილავს მოდელის გაუმჯობესებას საკუთარი გენერირებული მონაცემთა ნაკრების მეშვეობით.

12. რას გულისხმობს ფიზიკურ პირთა მიერ უფლებების განხორციელება?

გენერირებადი ხელოვნური ინტელექტის სისტემების გამოყენებისას, ფიზიკური პირების უფლებების[24] განხორციელებამ შესაძლოა, წარმოშვას გარკვეული სირთულეები წვდომის, შესწორების წაშლისა და მონაცემთა დამუშავებაზე უარის გაცხადების უფლებებთან მიმართებით.

ერთ-ერთი გამოწვევაა სისტემის მიერ შენახული პერსონალური მონაცემების იდენტიფიცირება და წვდომა. მაგალითად, მსხვილ ენობრივ მოდელებში, ცალკეული სიტყვები, როგორებიცაა „კატა“ ან „ძაღლი“, არ ინახება ტექსტის სტრიქონებად. ამის ნაცვლად, ისინი წარმოდგენილია როგორც რიცხვითი ვექტორები, რომლებსაც ეწოდება „სიტყვათა ჩაშენება“ (“word embedding”). აღნიშნული მომდინარეობს დიდი რაოდენობის ტექსტური მონაცემების სატესტო მოდელიდან. აქედან გამომდინარე, ამ მოდელებში შენახულ მონაცემებზე წვდომა, განახლება ან წაშლა ძალიან რთულია. ამ თვალსაზრისით, მონაცემთა ბაზის სწორად მართვა რთულდება, რადგან ხელოვნური ინტელექტის სწავლება/დატესტვა უკონტროლოდ ხორციელდება საჯაროდ ხელმისაწვდომი წყაროებიდან. გარდა ამისა, წაშლის უფლებამ, შეიძლება გავლენა იქონიოს მთლიანად მოდელის ეფექტიანობაზე.

პერსონალური მონაცემების დამუშავების შესახებ მიკვლევადი (“Tracable”) ჩანაწერის შენახვა ხელს უწყობს ფიზიკურ პირთა უფლებების განხორციელებას, ხოლო მონაცემთა მინიმიზაცია ამცირებს დარღვევის რისკებს.

“EUI”-ები, როგორც მონაცემთა დამუშავებისთვის პასუხისმგებელი პირები, ანგარიშვალდებულნი არიან შესაბამისი ტექნიკური, ორგანიზაციული და პროცედურული ღონისძიებების განხორციელებაზე. მნიშვნელოვანია, რომ აღნიშნული ღონისძიებები შემუშავდეს და განხორციელდეს სისტემის განვითარების სასიცოცხლო ციკლის ადრეული ეტაპებიდან, რაც შესაძლებელს გახდის დამუშავების დეტალურ აღრიცხვასა და ჩანაწერების მიკვლევადობის უზრუნველყოფას.

➲ “EUI-X”-მა, “EUDPR”-ის შესაბამისად, ჩატბოტის მონაცემთა დაცვის შეტყობინებაში შეიტანა მითითება ფიზიკური პირის უფლებების რეალიზების თაობაზე, მათ შორის, წვდომის, გამოსწორების, წაშლის, დამუშავებაზე უარისა და შეზღუდვის შესახებ. შეტყობინება შეიცავს მონაცემთა დამუშავებისთვის პასუხისმგებელი პირისა და “EU-X”-ის მონაცემთა დაცვის ოფიცრის საკონტაქტო

[24] რეგულაციის მე-3 თავი.

ინფორმაციას, ასევე მითითებას “EDPS”-ში საჩივრის შეტანის შესაძლებლობის შესახებ. გარდა ამისა, დადასტურდა, რომ ჩატბოტთან საუბრების შინაარსი არ ინახება 30 დღეზე მეტი ვადით. გარდა ამისა, საუბრები არ გამოიყენება ხელოვნური ინტელექტის სწავლების/ტესტირების მიზნებისთვის.

13. რას გულისხმობს მონაცემთა უსაფრთხოება?

გენერირებადი AI სისტემების გამოყენებამ, შესაძლოა რისკის ქვეშ დააყენოს უსაფრთხოება. ტრადიციულ სისტემებთან შედარებით, გენერირებადი ხელოვნური ინტელექტისთვის დამახასიათებელი სპეციფიკური რისკები შეიძლება, მომდინარეობდეს „არასანდო სატესტო მონაცემების“ (“Unreliable Training Data”), სისტემების გაუმჭვირვალობიდან, სათანადო ტესტირების ჩატარების სირთულიდან, სისტემის უსაფრთხოების დაუცველობიდან და ა. შ. ხელოვნური ინტელექტის მოდელების შეზღუდულმა ხელმისაწვდომობამ, მაგალითად, ჯანდაცვის სექტორში, შეიძლება, გაზარდოს სისტემებში მოწყვლადობის ხარისხი. რეგულაციები ავალდებულებს სუბიექტებს, მიიღონ ადამიანების უფლებებისა და თავისუფლებების პოტენციურ რისკების შესაბამისი ზომები უსაფრთხოების[25] უზრუნველსაყოფად.

IT სისტემების დაცვის ტრადიციული ღონისძიებების გარდა, მონაცემთა დამუშავებისთვის პასუხისმგებელმა პირებმა უნდა განახორციელონ სპეციფიკური კონტროლის მექანიზმები, რომელიც შექმნილია კიბერუსაფრთხოების თანამედროვე გამოწვევების დასაძლევად, რომელთა შორისაა: მოდელის ინვერსიული შეტევები[26], სწრაფი ინექცია[27] და ე. წ. „ჯეილბრეიქი“[28]. მონაცემთა დამუშავებისთვის პასუხისმგებელ პირებს ურჩევენ, რომ გამოიყენონ მხოლოდ სანდო წყაროების მიერ მოწოდებული მონაცემთა ნაკრები და რეგულარულად აწარმოონ გადამოწმებისა და ვალიდაციის პროცედურები, მათ შორის შიდა მონაცემთა ნაკრებისთვის (“in-house datasets”). EUI ვალდებულია, გადაამზადოს პერსონალი გენერირებადი “AI” სისტემების გამოყენებასთან დაკავშირებული უსაფრთხოების რისკებთან გასამკლავებლად. იქიდან გამომდინარე, რომ ზემოხსენებული რისკები სწრაფად მზარდია, საჭიროა რეგულარული მონიტორინგი და რისკის შეფასება. უცნობ რისკებთან გამკლავების შესაძლო გზაა ე. წ. „წითელი გუნდის[29]“ ტექნიკის გამოყენება.

---

[25] რეგულაციის 33-ე მუხლი.

[26] Model Inversion Attacks - მოდელის ინვერსიული შეტევები ხორციელდება მაშინ, როდესაც თავდამსხმელი ამოიღებს ინფორმაციას მისგან უკუინჟინერიის (Reverse-engineering) გზით.

[27] Prompt Injection - თავდამსხმელი იყენებს სწრაფ „ინექციურ“ შეტევებს მავნე ინსტრუქციების დასანერგად, თითქოს ისინი უვნებელია.

[28] Jailbreak - პროგრამული შეზღუდვების წაშლის პროცესია, რომელიც მოწყობილობაზე მწარმოებლის მიერ იყო დაყენებული, მოცემულ შემთხვევაში, თავდამსხმელი იყენებს „ჯეილბრეიქის“ ტექნიკას, რომ გაანეიტრალოს მოდელის უსაფრთხოების ზომები.

[29] წითელი გუნდი (red team) იყენებს შეტევის მეთოდებს სისტემის დაუცველობის აღმოსაჩენად.

როდესაც გამოიყენებთ ე. წ. „Retrieval Augmented Generation“-ს[30] გენერირებადი “AI” სისტემებთან, გაითვალისწინეთ, რომ სისტემა უნებლიედ არ ამხელს პერსონალურ მონაცემებს, რომლებიც შეიძლება არსებობდეს მის ბაზაში.

გენერირებადი “AI” სისტემების გამოყენებასთან დაკავშირებული უსაფრთხოების რისკების შესახებ ინფორმაციის ნაკლებობა მოითხოვს განსაკუთრებულ სიფრთხილეს, რათა დაიგეგმოს IT უსაფრთხოებასთან დაკავშირებული ყველა ასპექტი მაქსიმალური სიფრთხილით, მათ შორის, უწყვეტი მონიტორინგი და სპეციალიზებული ტექნიკური მხარდაჭერა.

➲ “EUI-X”-მა, უსაფრთხოების შეფასების შემდეგ, გადაწყვიტა დანერგოს ASR სისტემა, API სერვისების გამოყენების ნაცვლად. EU-X გეგმავს, გადაამზადოს IT პერსონალი, პროვაიდერთან მჭიდრო თანამშრომლობით. გარდა ამისა, EU-X მიიღებს გარე აუდიტორის მომსახურებას, რათა შეამოწმოს სისტემის სათანადო იმპლემენტაცია, მათ შორის, უსაფრთხოების საკითხებშიც.

[30] ხელოვნური ინტელექტის სისტემები, როგორიცაა Large Language Models (LLMs), ქმნიან პასუხებს სისტემის მფლობელის მიერ მოწოდებული ცოდნის ბაზის საფუძველზე, შიდა წყაროების გამოყენებით და არა მხოლოდ LLM-ში შენახულ ცოდნაზე დაყრდნობით.

14. გსურთ იცოდეთ მეტი?

- EDPS- ის გაწეული საქმიანობა AI-სთან მიმართებით:

- *პირადი ცხოვრების ხელშეუხებლობის გლობალური ასამბლეის 45-ე დახურული სესია - რეზოლუცია გენერირებადი ხელოვნური ინტელექტის სისტემების შესახებ - 2023 წლის 20 ოქტომბერი*
- *EDPS TechDispatch #2/2023 - Explainable Artificial Intelligence*
- *EDPS: მონაცემთა დაცვა და ხელოვნური ინტელექტი*
- EDPB-EDPS ერთობლივი მოსაზრება 5/2021 ევროპარლამენტისა და საბჭოს რეგულაციის წინადადების შესახებ, რომელიც ადგენს ხელოვნური ინტელექტის ჰარმონიზებულ წესებს (ხელოვნური ინტელექტის აქტი)
- EDPS Opinion 44/2023on the Proposal for Artificial Intelligence Act in the light of legislative developments Large Language Models (EDPS website, part of the EDPS "TechSonar" report 2023-2024)

– *სხვა რელევანტური დოკუმენტები:*

- Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)
- CNIL: AI how-to-sheets
- Spanish Data Protection Authority: Artificial Intelligence: accuracy principle in the processing activity
- Italian Data Protection Authority: Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di Intelligenza Artificiale – September 2023 (Italian)
- The Hamburg Commissioner for Data Protection and Freedom of Information - Checklist for the use of LLM-based chatbots - 15/11/2023
- AI Security Concerns in a nutshell (DE Federal Office for Information Security, March 2023)
- Multilayer Framework for Good Cybersecurity Practices for AI (ENISA, June 2023)
- Ethics Guidelines for Trustworthy AI (EC High-Level Expert Group on AI, 2019)
- Living Guidelines on the responsible use of Generative AI in research (ERA Forum Stakeholders' document, March 2024)
- OECD AI Incidents Monitor (AIM)
- OECD Catalogue or tools and metrics for trustworthy AI

ნატო ვაჩნაძის ქუჩა N° 7, თბილისი

ბაქოს ქუჩა N° 48, ბათუმი

(+995 32) 242 1000

office@pdps.ge