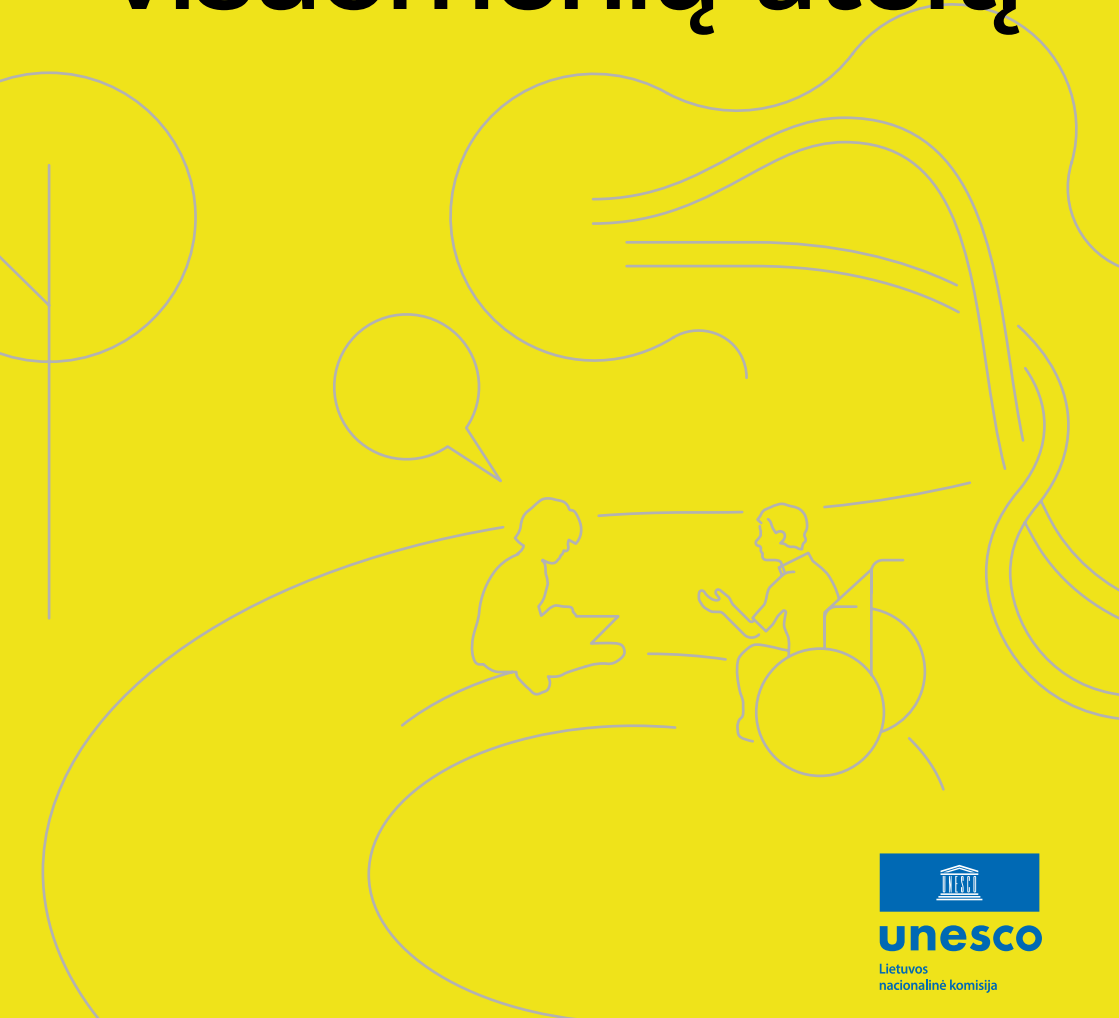


UNESCO

rekomendacija dėl

dirbtinio intelekto etikos: keičiant mūsų visuomenių ateitį



unesco

Lietuvos
nacionalinė komisija

UNESCO
rekomendacija dėl
dirbtinio intelekto
etikos: keičiant mūsų
visuomenių ateitį

UNESCO rekomendacija trumpai

UNESCO rekomendacija dėl dirbtinio intelekto (DI) etikos – tarptautiniu mastu sutartas teisinis tekstas, skirtas užtikrinti, kad kiekviena organizacija, įstaiga ar privatus asmuo, plėtojantis ar taikantis DI, veiktų etiškai ir gerbtų žmogaus teises;

Etika DI srityje reiškia, jog turime užtikrinti, kad mūsų sąveikavimas su DI sistemomis nesukeltų žalos, bet prisidėtų prie taikos, žmogaus orumo, darnumo ir saugumo skatinimo;

Tai bus įgyvendinta tik tuomet, kai kiekvienam vaikui ir suaugusiam žmogui visame pasaulyje bus suteikta galimybė pasiekti atitinkamą DI raštingumo lygį;

Todėl Rekomendacija veikia kaip gairės visoms vyriausybėms kuriant įstatymus ir strategijas DI srityje;

UNESCO reguliariai leidžia rekomendacijas. Tai norminės priemonės, dėl kurių kiekvieno įrašyto žodžio turi visuotinai sutarti 193 UNESCO valstybės narės;

Rekomendaciją priėmus, valstybės narės privalo pateikti pasaulinei bendruomenei ataskaitas apie veiksmus, kurių ėmėsi, kad sutarta politinė programa būtų įgyvendinama jų valstybėje.

keičiant mūsų visuomenių ateitį

UNESCO rekomendacija dėl dirbtinio intelekto etikos:

Kam reikalingas pasaulinis susitarimas dėl į žmogų orientuoto DI?

→ Dirbtinis intelektas tapo mūsų kasdienio gyvenimo dalimi. Jis turi milžinišką potencialą pasitarnauti visuomenės gerovei ir prisidėti prie darnaus vystymosi tikslų įgyvendinimo (pavyzdžiui, DI gali būti pritaikomas derliui prognozuoti ir taip užtikrinti aprūpinimą maistu pasaulyje). Tačiau DI naudojimas taip pat kelia grėsmių. Keletu atvejų DI taikymas iškėlė etinių klausimų: jis gali prisidėti prie didesnės lyčių nelygybės, pažeisti žmonių privatumą ar paskatinti manipuliavimą medijų turiniu. Kadangi beveik visos šalys susiduria su šiais iššūkiais ir beveik visos DI sistemos yra taikomos pasauliniu mastu, mums reikia tarpvalstybinio požiūrio, kad DI plėtra ir taikymas būtų etiški. 2021 m. lapkritį 193 UNESCO valstybės narės priėmė pirmą pasaulinį susitarimą dėl į žmogų orientuoto DI – Rekomendaciją dėl DI etikos.

Kokie esminiai Rekomendacijos uždaviniai?

→ Rekomendacija sukurta kaip gairės taip plėtoti DI sistemas, kad jos tarnautų žmonijos gerovei ir aplinkai bei būtų užkirstas kelias tokių sistemų žalai. Tai tarptautiniu mastu sutartas teisinis tekstas, kuriame išsakomos vertybės ir principai bei nurodymai, kaip juos praktiškai įgyvendinti. Siūlomos konkrečios užduotys ar veiksmai, kuriuos turi atlikti valstybės narės 11-oje politikos sričių, įskaitant sveikatą, aplinkosaugą ir lyčių lygybę. Šiuo požiūriu Rekomendacija siekia apsaugoti ir skatinti žmogaus teises ir pagrindines laisves, taip pat žmogaus orumą ir lygiateisiškumą. Etiniai klausimai, išskylantys plėtojant DI sistemas, apima visas ciklo stadijas – nuo tyrimų, projektavimo, sukūrimo iki diegimo ir naudojimo, taip pat priežiūrą, eksploatavimą, stebėseną, vertinimą ir eksploatacijos nutraukimą.

- Rekomendacija priimta visų 193 UNESCO valstybių narių konsensusu. Ypatingas dėmesys joje skiriamas mažų ir vidutinių pajamų šalių (MVPŠ) poreikiams ir indėliui į etiško DI plėtrą ir taikymą. Kartu pabrėžiama, kad į žmogų orientuoto DI plėtra yra visos visuomenės uždavinys, daug platesnis nei paprastas technologijos reguliavimas. Rekomendacijoje pabrėžiama, kad pagarba žmogaus teisėms ir jų skatinimas yra bet kokio DI reguliavimo pagrindas.
- Rekomendacija išsiskiria savo apimtimi, ji visapusiškai nukreipta į visas žmogaus gyvenimo sritis, veikiamas DI, ypač tas, kurios iki tol nebuvo įvardintos kaip susijusios su DI, t. y. ...

keičiant mūsų visuomenių ateitį

... skirtingas valstybių išsivystymo lygis,

... socialiai atskirtų, pažeidžiamų grupių poreikiai,

... globaliųjų Pietų poreikiai,

... lyčių nelygybė ir diskriminacija

... ir darnumas.

UNESCO rekomendacija dėl dirbtinio intelekto etikos:



Rekomendacija dėl
dirbtinio intelekto etikos
priimta 2021 m.
lapkričio 23 d.

Kas yra DI?

→ DI sistemos geba apdoroti duomenis ir informaciją tokiu būdu, kuris panašus į protingą elgesį ir paprastai apima samprotavimo, mokymosi, suvokimo, numatymo, planavimo ar valdymo aspektus.

Jų veikimas iš dalies galėtų būti apibūdinamas kaip mašininis mokymasis arba mašinų atliekama loginė analizė.

keičiant mūsų visuomenių ateitį

Kokie svarbiausi Rekomendacijos reikalavimai?

Siekiant užtikrinti etišką DI plėtrą ir naudojimą, Rekomendacijoje įtvirtinami reikalavimai 11-oje strateginių sričių. Toliau pateiktos nuostatos atspindi Rekomendacijos apimtį:

- 1 Įstatymai ir duomenų valdymo strategijos turi užtikrinti teisę į privatumą, spręsti tokius klausimus, kaip DI grindžiamas sekimas, ir išsaugoti žmonių teisę į jų asmens duomenis.
- 2 Tarptautinis bendradarbiavimas DI srityje turi būti intensyvinamas siekiant spręsti aktualius vystymosi iššūkius ir sumažinti skaitmeninę atskirtį tarp globaliųjų Pietų ir globaliosios Šiaurės.
- 3 DI sistemų etinio poveikio vertinimas turi būti vykdomas visuose DI sistemos gyvavimo ciklo etapuose, kad būtų išvengta žalos žmonėms ir užtikrintos žmogaus teisės.
- 4 Valdymo mechanizmai, pavyzdžiui, DI etikos pareigūnų skyrimas, yra skatintini kaip priemonė galimam DI sistemų sukeltam žalingam poveikiui tirti, apsaugos, užkardymo ir žalos atlyginimo veiksmams numatyti.
- 5 Turi būti sukurtos paskatos kurti DI taikomąsias programas aplinkos ir ekosistemų stebėsenai, apsaugai ir regeneracijai.
- 6 DI turi būti vystomas ir taikomas taip, kad užtikrintų didžiausią įmanomą naudą įvairovei ir įtraukčiai, įskaitant kultūrų įvairovės, nediskriminavimo, išraiškos laisvės ir lyčių lygybės puoselėjimą.

- 7 Turi būti diegiamos visuomenės raštingumo DI srityje programos, kurios suteiktų žmonėms būtinas kompetencijas ir mažintų skaitmeninę atskirtį.
- 8 DI sistemos turi pagerinti informacijos ir žinių prieigą.
- 9 Vyriausybės ir privatusis sektorius turi užtikrinti sąžiningą rizikos grupių darbuotojų prisitaikymą perduodant darbo vietas DI technologijoms ir sistemoms.
- 10 Būtina parengti gaires žmonių ir robotų sąveikai bei santykiams tarp žmonių, ypač atsižvelgiant į žmonių psichinę ir fizinę sveikatą.

keičiant mūsų visuomenių ateitį



UNESCO rekomendacija dėl dirbtinio intelekto etikos:

Kvietimas veikti penkiose strateginėse politikos srityse

Rekomendacijoje nubrėžiamos aiškios etinės gairės nacionalinėms vyriausybėms imtis konkrečių politinių veiksmų jai įgyvendinti. Perkeliant šią tarptautinę priemonę į nacionalinės teisės ir programų kalbą, esminis vaidmuo skiriamas tiek pilietinei visuomenei, tiek privačiajam sektoriui. Toliau išskiriamos penkios iš vienuolikos Rekomendacijoje aptariamų visuomenės gyvenimo sričių, konkrečiai įvardijant, koks politinis mandatas reikalingas siekiant etiško DI vystymo ir taikymo.



● Sveikata

Sveikatos priežiūros srityje tiek DI grėsmės, tiek galimybės yra labiau apčiuopiamos nei kitose srityse. DI plėtra lėmė revoliucinį proveržį anksti aptinkant ligas, plečiant gydymo galimybes, taip pat tai gali atverti galimybę spręsti sveikatos priežiūros sistemos darbuotojų trūkumo problemą. Tačiau sektoriuje kyla tokių svarbių etinių klausimų, kaip asmens sveikatos duomenų apsauga ir smegenų ir technologų sąveikos reguliavimas. Tai tiesiogiai paveiks gydymo įstaigų administraciją, pacientus ir gydytojus.

Vyriausybės ir sveikatos priežiūros srities sprendimų priėmėjai kviečiami imtis šių veiksmų:

- rengti gaires, reglamentuojančias žmogaus ir roboto sąveiką ir jos poveikį žmonių santykiams, ypatingą dėmesį skiriant psichinei ir fizinei žmonių sveikatai, siekiant, kad ši sąveika nepažeistų žmogaus teisių, pagrindinių laisvių, įvairovės skatinimo, pažeidžiamų žmonių apsaugos, ypač DI taikymo neurotechnologijose bei smegenų ir kompiuterio sąveikoje;
- DI sistemas pasitelkti žmonių sveikatai – ligų protrūkiams stebėti ir duomenimis dalytis pasauliniu mastu;
- skatinti bendradarbiavimu grindžiamus ilgalaikės žmonių ir DI sistemų sąveikos poveikio mokslinius tyrimus, ypač daug dėmesio skiriant psichologinei ir kognityvinei įtakai, kurią šios sistemos gali daryti vaikams ir jaunimui.

UNESCO rekomendacija dėl dirbtinio intelekto etikos:

keičiant mūsų visuomenių ateitį

● Kultūra

DI turi reikšmingos įtakos tam, kaip mes *gaminame* kultūrą, kaip (kiek) vartotojai moka už kultūros produktus, kaip menininkai gali užsidirbti iš savo darbo. Algoritmai jau daro įtaką ir lemia kultūros *vartojimą*, netgi formuoja, pavyzdžiui, mūsų klausomą muziką.

Kelios svarbiausios valstybių narių pareigos:

- skatinti technologijų bendroves ir kitus suinteresuotuosius subjektus propaguoti įvairialypę kultūros raiškos formų pasiūlą ir daugialypę prieigą prie jų, ypač siekiant užtikrinti vietinio turinio matomumą ir aptinkamumą;
- skatinti naujus tyrimus, kurie nagrinėtų DI ir intelektinės nuosavybės (IN) sankirtą, pavyzdžiui, siekiant nustatyti, ar remiantis IN teisėmis apsaugoti kūrinius, sukurtus naudojant DI technologijas, ir kaip tai padaryti;
- nagrinėti DI sistemų, ypač natūraliosios kalbos apdorojimo taikomųjų programų, pavyzdžiui, mašininio vertimo ir balso asistentų, kultūrinį poveikį žmogaus kalbos ir raiškos niuansams ir spręsti su juo susijusius klausimus.





● Lytis

Taikant DI sistemas gali atsirasti šališkumą arba būti sustiprinti jau esami. Organizacijų naudojamos vertimo programos gali lyčiai neutralų terminą (gydytojas, -a) išversti į lytį nurodantį (gydytojas – *red. past.* – *šiuo atveju tik vyrišką*). Panašiai, pavyzdžiui, vyrų dominuojamame sektoriuje kompanijoms darbu renkantis naujus kandidatus iš esamų CV duomenų bankų, labiau tikėtina, kad įdarbinimo programos pirmenybę teiks vyriškos lyties pretendents. Svarbus Rekomendacijos principas yra siekti kiek įmanoma sumažinti ir išvengti diskriminuojančių ar šališkų programų įtvirtinimo ir jų pasekmių visuose DI sistemos gyvavimo ciklo etapuose.

Sprendimų priėmėjai, darbdaviai DI tyrimų ir plėtros sektoriuje ir kiti DI sistemos veikėjai turi siekti:

- įtraukti lyčių lygybės veiksmų planus į nacionalines skaitmenines strategijas;
- didinti investicijas į tokias strategines sritis, kaip profesinis mokymas, pavyzdžiui, siekiant skatinti moterų karjerą STEM (angl. *science, technology, engineering, mathematics*, liet. gamtos mokslų, technologijų, inžinerijos ir matematikos) ir IRT (informacinės ir ryšių technologijos; angl. *information and communication technologies, ICT*) disciplinose;
- užtikrinti, kad į DI sistemas nebūtų perkeltami lyčių stereotipai ir diskriminacinės tendencijos (pavyzdžiui, padedant kurti įvairias atviras ir patikimas aukso standarto duomenų bases).

UNESCO rekomendacija dėl dirbtinio intelekto etikos:

keičiant mūsų visuomenių ateitį

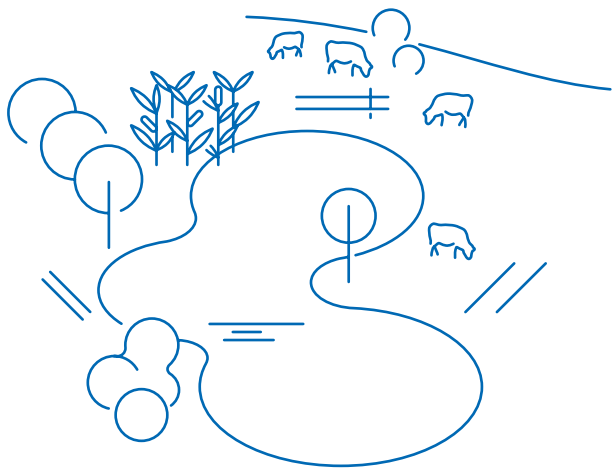


● Švietimas ir moksliniai tyrimai

Mokyklos ir neformaliojo švietimo aplinka yra esminės informuotumui apie DI etiką ir raštingumui šioje srityje skatinti ir prisideda prie siekio panaikinti skaitmeninę atskirtį. Rekomendacija teigia švietimą esant svarbią priemonę užtikrinant, kad vaikai ir suaugusieji visame pasaulyje lavintųsi DI raštingumo įgūdžius, kurių reikia tiek technologiniams, tiek etiniams DI aspektams suvokti. Tuo pat metu e. mokymasis mokyklose turi savų etinių dilemų (pavyzdžiui, mokinių asmens duomenų naudojimas turi atitikti galiojančius asmens duomenų apsaugos standartus).

Valstybės narės, taip pat švietimo sistemos administratoriai, mokytojai ir NVO skatinami:

- propaguoti bendras informavimo apie DI plėtrą programas, kurių pagrindu taip pat būtų informuojama apie DI poveikį žmogaus teisėms, taip pat vaikų teisėms, ir jų sukeliamas pasekmes;
- derinti DI techninių įgūdžių ir DI etikos mokymą, skatinti tarpdisciplininius DI mokslinius tyrimus, į kuriuos būtų įtrauktos ne tik gamtos mokslų, technologijų, inžinerijos ir matematikos (STEM), bet ir kitos disciplinos;
- skatinti mokslo ir politikos konsultacijas kuriant DI politikos strategijas ir auginti DI etikos suvokimo kultūrą.



● Aplinka ir ekosistemos

DI jau prisideda prie aplinkosaugos, pavyzdžiui, jis naudojamas bioįvairovei stebėti ir gerinti, naudojantis dronų ir palydovų padarytų nuotraukų analize. Vis dėlto DI gali būti ir žalingas aplinkai, pavyzdžiui, dėl ženkliausios energijos eikvojimo.

Rekomendacijoje, be kita ko, raginama imtis veiksmų šiais klausimais:

- vertinti tiesioginį ir netiesioginį poveikį aplinkai visuose DI sistemos gyvavimo ciklo etapuose ir užtikrinti, kad jis neprieštaruoja aplinkosaugos reikalavimams;
- užtikrinti teisėmis grindžiamų ir etiškų DI sistemų vystymą stichinių nelaimių prevencijos, ekosistemų stebėsenos, apsaugos bei atkūrimo ir planetos išsaugojimo tikslais;
- uždrausti DI sistemas, kurių keliama grėsmė gamtai yra neproporcingai didelė.

Laikas palaikyti etišką ir į žmogų orientuotą DI plėtrą

UNESCO rekomendacija siekiama užtikrinti, kad DI būtų vystomas visuomenių labui. Tam, kad į žmogų orientuota DI plėtra taptų ir įsitvirtintų kaip norma, reikia visų suinteresuotųjų subjektų – ne tik sprendimų priėmėjų, bet taip pat pilietinės visuomenės organizacijų, privačiojo sektoriaus ir plačiosios visuomenės indėlio. Šalia kvietimų imtis konkrečių veiksmų Rekomendacijoje pabrėžiamas kolektyvinių pastangų būtinumas. Visi DI sistemos veikėjai, ne tik dirbantys sveikatos priežiūroje, kultūroje, švietimo sektoriuose, turi pasinaudoti galimybe kartu kurti sistemas, kurios užtikrintų etišką ir į žmogų orientuotą DI vystymą ir taikymą. Laukia daug darbo!

Leidimo duomenys

Išleido

Vokietijos nacionalinė UNESCO komisija
Martin-Luther-Allee 42
53175 Bonn, Germany

Nyderlandų nacionalinė UNESCO komisija
Prins Willem-Alexanderhof 20
2595 BE Den Haag, Netherlands

Slovėnijos nacionalinė UNESCO komisija
Kotnikova 38
1000 Ljubljana, Slovenia

Lietuvos nacionalinė UNESCO komisija
Radvilų g. 6
01129 Vilnius, Lietuva

Šio leidinio originalią versiją išleido Vokietijos, Nyderlandų ir Slovėnijos nacionalinės UNESCO komisijos.

Ją galima rasti adresu: https://ircai.org/wp-content/uploads/2023/07/Recommendation_brochure_A5_web_doublepage.pdf

Kontaktai

German Commission for UNESCO e.V.
T +49 (0) 228 60 497-165
E zukunft@unesco.de
www.unesco.de

Netherlands National Commission for UNESCO
T +31 (0) 70 - 3315484
info@unesco.nl
www.unesco.nl

Slovenian National Commission for UNESCO
T +386 (0) 1-4784799
unesco.mizs@gov.si
www.gov.si/en/state-authorities/ministries/ministry-of-higher-education-science-and-innovation/about-the-ministry/urad-za-unesco/

Lietuvos nacionalinė UNESCO komisija
Tel. +370 5 2107340
Lietuva@unesco.lt
www.unesco.lt

Redaktoriai

Anne Diessner, Jeannine Hausmann (atsakingasis red.), Maximilian Müngersdorff, Davor Orlic, Jon Verriet, Liliana Bugailiškytė-Lideikienė, Margarita Gaubytė

Išleista

2024 m. sausis

ISBN

978-3-947675-35-7

Autorinės teisės

Šio leidinio tekstams taikoma licencija *Creative Commons License Attribution ShareAlike 4.0 International*. www.creativecommons.org/licenses/by-sa/4.0/

Dizainas ir maketas

Panatom
Leidimą lietuvių k. maketavo Jurga Dovydenaitė

Parėmė



Federal Foreign Office

Parémé



unesco

Centre
Under the auspices
of UNESCO



International Research Centre
of Artificial Intelligence
under the auspices of UNESCO