



AI@MEDIATORS

**Dirbtinis intelektas ir skaitmeniniai ugdymo įrankiai
patyčioms internete prevencijai**



**Co-funded by
the European Union**



Universidad de Alicante



ERUDITO LICÉJUS





Contents

1	Dirbtinio intelekto švietime vertinimo schema, užduoties pristatymas ir metodika	3
1.1.	Ivadas ir bendras tikslas	3
1.2.	Konkretūs tikslai	3
2	Etinio lygmens vertinimas	5
2.1.	Duomenų privatumas	5
2.2.	Etika	6
2.3.	ES rekomendacijos	7
3	Poveikio efekto vertinimas	8
3.1.	Gebėjimas	8
3.2.	Naudingumas	8
3.3.	Pritaikymas	9
4	Vertinimo veiklos	1
5	Išvados ir Įspėjimas	1
6	Bibliografinės nuorodos	2

1 Dirbtinio intelekto švietime vertinimo schema, užduoties pristatymas ir metodika

1.1. Įvadas ir bendras tikslas

Šiuolaikiniame skaitmeniniame kraštovaizdyje, kuriam būdinga visa apimanti technologinė integracija, **atsakingas dirbtinio intelekto (DI) diegimas švietimo sektoriuje tampa esmine problema** tiek galimybių supratimo, tiek susijusios rizikos atidaus stebėjimo prasme.

Projektas „AI@Mediator“, kurio tikslas – ištirti galimybes spręsti kibernetinio priekabiavimo problemas diegiant dirbtinio intelekto pagrįstą pokalbių botą, **atsiduria šio sudėtingo klausimo centre**. Kaip „AI@Mediator“ konsorciumo nariai, esame raginami ne tik prisidėti prie diskusijų, bet ir pateikti naudingus įrankius bei metodikas, kurie galėtų padėti kitiems, kurie vėliau pradėjo gilintis į šią temą ir žengti šiuo keliu, siekdami diegti dirbtinį intelektą švietimo srityje, ypač dirbant su pažeidžiama nepilnamečių grupe.

Šios ambicingos iniciatyvos kontekste itin svarbi reikšmė tenka kruopščios DI vertinimo švietime schemos kūrimui. **Šis dokumentas atspindi mūsų požiūrį, mūsų priemonę, padedančią orientuotis** skaitytojams, siekiant suprasti, ar DI įrankiai ar praktika, kuriuos jie taiko ar ketina taikyti, viena vertus, atitinka minimalius reikalavimus, kad būtų sėkmingai įgyvendinti, kita vertus, ar jie atitinka aukščiausius etikos standartus konfidencialumo ir Europos vertybių požiūriu.

1.2. Konkretūs tikslai

Vertinimo schema siekia tapti pagrindiniu orientyru, padedančiu pagrįstai įvertinti ir etiškai integruoti AI technologijas švietimo kontekste. Nustatydamas aiškius etiško AI naudojimo parametrus, schema prisideda prie bendro projekto tikslo – kurti saugią internetinę aplinką ir skatinti įvairių tikslinės grupės narių skaitmeninį atsparumą. Be to, schema tampa veiksminga priemonė, padedančia įgyvendinti projekto misiją – suteikti pedagogams galimybes veiksmingai kovoti su kibernetiniu priekabiavimu. **Teikdama išsamias gaires ir vertinimo sistemas, schema suteikia švietimo suinteresuotosioms šalims būtinas priemones, leidžiančias aktyviai reaguoti į kibernetinio priekabiavimo atvejus ir ugdyti įtraukties bei skaitmeninės pilietybės kultūrą švietimo ekosistemose.**

- **Pasiūlyti vertinimo strategiją, skirtą nustatyti dirbtinio intelekto technologijų teorinį aspektą švietimo srityje**, ypatingą dėmesį skiriant duomenų privatumui, etiniams aspektams ir suderinamumui su Europos Sąjungos vertybėmis bei politika. Šis požiūris strategiškai skirtas sukurti pagrindą, leidžiantį įvertinti atsakingą dirbtinio intelekto diegimą, ir taip sklandžiai dera su bendrais „AI@Mediator“ projekto tikslais. Iš tiesų, nereikėtų nuvertinti šio dokumento bendros svarbos projekto sėkmei.
- **Vertinimo kriterijų rinkinio nustatymas, siekiant įvertinti diegtų AI įrankių ir metodikų poveikio aspektą, remiantis trimis kriterijais: pajėgumu, naudingumu ir įsisavinimu.** Šis požiūris skirtas padėti suprasti, ar diegiamas metodas yra tinkamiausias, atsižvelgiant į alternatyvias sąnaudas, reikalavimus, tikslinės grupės poreikius, sistemą ir socialinę bei kultūrinę aplinką. Iš tiesų, net ir perspektyvi priemonė ar metodika gali baigtis nesėkme

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Project No: 2023-1-ES01-KA220-SCH-000160237



įgyvendinimo etape, jei pasirinkimas nebus padarytas atsižvelgiant į konkrečius reikalavimus ir poreikius tikslinės grupės, kuriai ji skirta.

2 Etinio lygmens vertinimas

Šioje išsamioje analizėje nagrinėjama rizika ir galimybės, susijusios su trimis aspektais: duomenų konfidencialumu, etika ir ES rekomendacijomis. Šia veikla siekiama pabrėžti jų esminį vaidmenį sėkmingai įgyvendinant DI iniciatyvas, ypač patyčių internete prevencijos srityje. Tirdamos duomenų privatumo, etikos ir ES vertybių sąveiką, švietimo įstaigos gali veiksmingai atkreipti dėmesį į rizikas, įgyvendinti rizikos mažinimo priemonės ir puoselėti atsakingo AI naudojimo kultūrą. Šis tyrimas tarnauja kaip pagrindinis pagrindas, padedantis orientuoti strateginį sprendimų priėmimą ir užtikrinti, kad „AI@Mediators“ atrinktos AI gerosios praktikos atitiktų etikos principus, reguliavimo reikalavimus ir visuomenines vertybes švietimo ekosistemoje.

Šiuo atžvilgiu svarbu suprasti, kad bet kokią etinį vertinimą gali paveikti asmeniniai ar kultūriniai šališkumai, dėl kurių išvados gali būti iš dalies netinkamos. Taigi, kad etinis vertinimas būtų prasmingas, jis turi būti išsamus, aiškiai apibrėžtas ir pakomentuotas, o bet kokia analizė turi būti periodiškai atnaujinama ir peržiūrima atsižvelgiant į besikeičiančią situaciją ir poreikius.

Toliau pateikiamas siūlomas vertinimo modelis.

Duomenų privatumas

- **Pagrindinės rizikos:** Viena iš pagrindinių rizikų, susijusių su duomenų privatumu ir dirbtiniu intelektu, yra neteisėtos prieigos, nesąmoningo jautrios informacijos perdavimo ar nutekėjimo galimybė, dėl kurios gali būti pažeistas konfidencialumas ir netgi atskleista nepilnamečių asmeninė informacija. Nepakankamos duomenų apsaugos priemonės ir bendrų žinių apie duomenų tvarkymą trūkumas gali sukelti nerimą keliančias situacijas, pažeidžiančias studentų ir pedagogų teisę į privatumą ir netgi turinčias įtakos jų kasdieniam gyvenimui. Be to, netinkamas duomenų tvarkymas gali pažeisti reguliavimo reikalavimus, pavyzdžiui, Bendrąjį duomenų apsaugos reglamentą (GDAR) ar kitus susijusius nacionalinius įstatymus, dėl ko švietimo įstaigos gali patirti teises ir finansines pasekmes.
- **Rizikos mažinimo priemonės:**
 - Įgyvendinti patikimas duomenų apsaugos priemonės: Švietimo įstaigos turėtų nustatyti griežtus duomenų rinkimo, saugojimo ir tvarkymo protokolus, siekdamos užtikrinti studentų privatumą. Tai apima šifravimą, prieigos kontrolę ir reguliarius saugumo auditus, skirtus užkirsti kelią neteisėtai prieigai ir duomenų saugumo pažeidimams.
 - Gauti informuotą sutikimą: Prieš renkant bet kokius studentų duomenis, įstaigos turėtų gauti aiškų studentų arba jų globėjų sutikimą, kuriame būtų išdėstyti duomenų naudojimo tikslas ir apimtis. Skaidrumas duomenų tvarkymo srityje stiprina pasitikėjimą ir suteikia asmenims galimybę priimti informuotus sprendimus dėl savo privatumo.
 - Duomenų saugojimo trukmės sutrumpinimas: siekiant sumažinti duomenų nutekėjimo riziką, įstaigos turėtų taikyti minimalaus duomenų saugojimo politiką, t. y. saugoti duomenis tik tiek, kiek būtina švietimo tikslams pasiekti. Reikėtų įgyvendinti reguliarius duomenų valymo ir ištrynimo protokolus, kad būtų pašalinta pasenusi ar nereikalinga informacija.
- **Siūlomi tyrimo klausimai:**
 - Kaip DI pagrįstame įrankyje ar švietimo praktikoje renkami, saugomi ir tvarkomi mokinių duomenys?
 - Ar taikomos tinkamos saugumo priemonės mokinių duomenims apsaugoti nuo neteisėtos prieigos ar pažeidimų?

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Project No: 2023-1-ES01-KA220-SCH-000160237

- Ar prieš renkant ir naudojant mokinių duomenis gaunamas aiškus mokinių ar jų globėjų sutikimas?
- Kokios procedūros taikomos duomenų saugojimui ir naikinimui, siekiant sumažinti atskleidimo riziką?
- Kaip DI įrankis ar praktika užtikrina duomenų anonimizavimą ir privatumo apsaugą per visą jo gyvavimo ciklą?

2.1. Etika

- **Pagrindinė rizika:** Etiniu požiūriu dirbtinio intelekto (AI) naudojimas švietime kelia rimtų susirūpinimų dėl teisingumo, skaidrumo ir atskaitingumo. AI algoritmai gali rodyti šališkumą ar duoti diskriminacinius rezultatus, taip įtvirtindami nelygybę ir sustiprindami esamas galios dinamikas (tai ypač jautrios temos, jei kalbama apie patyčių dinamiką). Be to, AI sprendimų priėmimo procesų neskaidrumas gali pakenkti pasitikėjimui ir susilpninti žmogaus veikimo laisvę, ypač tokiose jautriose srityse kaip vertinimas ir parama mokiniams. Etiniai svarstymai taip pat apima sutikimo, autonomijos klausimus bei galimybę, kad AI gali pažeisti asmens teises ir laisves. Siekiant sumažinti šiuos pavojus, būtina priimti etines gaires ir standartus, kuriuose AI kūrimo ir diegimo procese pirmenybė būtų teikiama teisingumui, skaidrumui ir į žmogų orientuotiems projektavimo principams.
- **Rizikos mažinimo priemonės:**
 - Atlikti šališkumo vertinimus: prieš diegdamos AI sistemas, švietimo įstaigos turėtų atlikti išsamius vertinimus, siekdamos nustatyti ir sumažinti algoritmų bei duomenų šaltinių šališkumą. Tai gali apimti mokymo duomenų auditą, nevienodo poveikio tikrinimą ir šališkumo mažinimo strategijų įgyvendinimą, siekiant užtikrinti teisingus ir lygiateisius rezultatus.
 - Didinti skaidrumą: įstaigos turėtų teikti pirmenybę skaidrumui AI sprendimų priėmimo procesuose, pateikdamos aiškius paaiškinimus apie algoritmų veikimą ir veiksnius, darančius įtaką jų rezultatams. Skaidrios AI sistemos suteikia suinteresuotosioms šalims galimybę suprasti sprendimus ir juos kritiškai vertinti, taip skatindamos atskaitingumą ir pasitikėjimą.
 - Įsteigti etikos vertinimo komisijas: Švietimo įstaigos turėtų įsteigti tarpdisciplinines vertinimo komisijas, kurių užduotis – vertinti AI iniciatyvų etines pasekmes. Šios komisijos gali teikti rekomendacijas dėl etinių dilemų, vertinti siūlomus AI taikymo atvejus ir užtikrinti atitiktį etiniams standartams bei taisyklėms.
- **Siūlomi tyrimo klausimai:**
 - Kaip dirbtinio intelekto pagrįsta mokymo priemonė ar praktika sprendžia galimų šališkumų problemą savo algoritmuose ir duomenų šaltiniuose?
 - Ar užtikrinamas AI priemonės sprendimų priėmimo procesų skaidrumas, įskaitant veiksnius, darančius įtaką jos rezultatams?
 - Ar yra įdiegtos priemonės, užtikrinančios teisingumą, atskaitingumą ir žmogiškąją priežiūrą AI pagrįstame sprendimų priėmime?
 - Kaip AI priemonė ar praktika skatina tokius etikos principus kaip autonomija, pagarba asmens teisėms ir nediskriminavimas?
 - Ar AI priemonė buvo patikrinta ar įvertinta etikos požiūriu, siekiant įvertinti jos atitiktį etikos standartams ir gairėms?

2.2. ES rekomendacijos

- **Pagrindinė rizika:** Atsižvelgiant į ES rekomendacijas, dirbtinio intelekto naudojimas švietime turi atitikti Europos Sąjungos nustatytus reguliavimo pagrindus ir politikos gaires. ES rekomendacijose pabrėžiama pagrindinių teisių ir vertybių, įskaitant duomenų apsaugą, privatumą ir nediskriminavimą, laikymosi svarba, taip pat konkretesni ir techniniai aspektai, susiję su numatomu Dirbtinio intelekto aktu, kuris, kaip tikimasi, įsigalios 2024 m. Apskritai tikimasi, kad švietimo įstaigos laikysis tokių teisės aktų kaip Bendrasis duomenų apsaugos reglamentas (BDAR) ir laikysis etikos principų, išdėstytų tokiuose dokumentuose kaip „Etikos gairės dėl patikimo dirbtinio intelekto“. Labai svarbu atkreipti dėmesį į tai, kad šios pastangos turėtų būti laikomos nuolatinėmis, nes ES dirbtinio intelekto reguliavimo sistema sparčiai keičiasi. Šių standartų nesilaikymas gali lemti teisinių reikalavimų pažeidimą, reputacijos sugadinimą ir suinteresuotųjų šalių pasitikėjimo praradimą. Todėl švietimo srities suinteresuotosios šalys privalo teikti pirmenybę ES rekomendacijoms ir integruoti jas į savo AI valdymo sistemas, siekdamos užtikrinti atsakingą ir etišką AI naudojimą švietime.
- **Rizikos mažinimo priemonės:**
 - Laikytis BDAR reikalavimų: Švietimo įstaigos privalo užtikrinti, kad būtų laikomasi Bendrojo duomenų apsaugos reglamento (BDAR) ir kitų susijusių duomenų apsaugos įstatymų. Tai apima tokių priemonių kaip duomenų anonimizavimas, privatumo poveikio vertinimai ir duomenų saugumo pažeidimų tvarkymo procedūros įgyvendinimą pagal BDAR reikalavimus.
 - Laikytis Europos dirbtinio intelekto etikos gairių: įstaigos turėtų suderinti savo dirbtinio intelekto praktiką su Europos dirbtinio intelekto etikos gairėmis, kuriose išdėstyti etiniai principai ir rekomendacijos dėl dirbtinio intelekto kūrimo ir diegimo. Laikydamosi šių gairių, įstaigos parodo savo įsipareigojimą etiškam dirbtinio intelekto valdymui ir atitikimą ES vertybėms.
- Nuolatinis mokymas ir sąmoningumo ugdymas: Švietimo sektoriaus suinteresuotosios šalys turėtų investuoti į mokymo programas ir sąmoningumo ugdymo iniciatyvas, siekdamos užtikrinti, kad darbuotojai ir studentai būtų informuoti apie savo teises ir pareigas, susijusias su AI naudojimu. Ugdydamos etinio sąmoningumo ir reikalavimų laikymosi kultūrą, įstaigos gali laikytis ES rekomendacijų ir skatinti atsakingą AI naudojimą švietime.
- **Siūlomi tyrimo klausimai:**
 - Ar dirbtinio intelekto pagrįsta mokymo priemonė ar praktika atitinka atitinkamus duomenų apsaugos teisės aktus, pavyzdžiui, BDAR?
 - Ar dirbtinio intelekto priemonės ar praktikos kūrimas ir diegimas atitinka Europos dirbtinio intelekto etikos gaires?
 - Ar taikomos priemonės, skirtos užtikrinti ES rekomendacijose išdėstytas pagrindines teises ir vertybes, įskaitant duomenų apsaugą, privatumą ir nediskriminavimą?
 - Kaip dirbtinio intelekto priemonė ar praktika užtikrina atskaitingumą ir skaidrumą pagal ES reguliavimo reikalavimus?
 - Ar įstaiga vykdo nuolatinius mokymus ir informavimo iniciatyvas, siekdama užtikrinti atitiktį ES rekomendacijoms ir skatinti atsakingą dirbtinio intelekto naudą

3 Poveikio efekto vertinimas

Plėtojantis dirbtinio intelekto (AI) sprendimų taikymo sričiai, dirbtinio intelekto integravimas tapo perspektyvia priemone ir strategija, skirta gerinti mokinių gerovę ir saugumą, ypač sprendžiant plačiai paplitusią kibernetinio priekabiavimo problemą. Kadangi švietimo įstaigos siekia pasinaudoti AI įrankiais, siekiamos sukurti saugesnę internetinę aplinką ir skatinti teigiamas mokymosi patirtis, **tampa būtina įvertinti šių technologijų veiksmingumą, galimą poveikį ir etines pasekmes**. Šiuo tikslu vertinimo schema apibrėžia struktūrizuotą sistemą, susidedančią iš **trijų pagrindinių komponentų: pajėgumo, naudingumo ir įdiegimo**.

Šis suskirstymas atspindi daugialypį dirbtinio intelekto technologijų vertinimo pobūdį etiškai jautriose aplinkose. **Kiekviena sudedamoji dalis atitinka atskirą, tačiau tarpusavyje susijusį rodiklį, skirtą dirbtinio intelekto priemonių etiško diegimo ir praktinio veiksmingumo vertinimui švietimo sistemoje**.

Apibrėždama šiuos komponentus, vertinimo schema suteikia išsamų pagrindą, leidžiantį įvertinti dirbtinio intelekto technologijų etišką naudojimą ir praktinį poveikį švietimo srityje. Šis struktūruotas požiūris leidžia švietimo srities suinteresuotosioms šalims sistemingai įvertinti technines galimybes, praktinę naudą ir etinius aspektus, susijusius su dirbtinio intelekto diegimu, taip skatindamas pagrįstą sprendimų priėmimą ir atsakingą dirbtinio intelekto priemonių integravimą į švietimo ekosistemas.

3.1. Gebėjimas

Vertinimo schemos „gebėjimų“ komponentas skirtas **švietimo aplinkose diegiamų AI sistemų techninio veiksmingumo vertinimui**. Tai apima AI gebėjimų aptikti ir veiksmingai spręsti kibernetinio priekabiavimo atvejus vertinimą. Pagrindiniai nagrinėjami veiksniai apima AI algoritmų tikslumą, patikimumą ir bendrą našumą nustatant bei mažinant kibernetinio priekabiavimo incidentus. **Šis vertinimas užtikrina, kad AI sistemos turėtų reikiamus techninius gebėjimus veiksmingai veikti kaip priemonės, skatinančios mokinių saugumą ir gerovę internete**.

Be to, **pajėgumų komponentas nagrinėja AI sistemų suderinamumą su esama švietimo infrastruktūra ir platformomis**. Tai apima vertinimą, ar AI technologijos gali sklandžiai integruotis į mokyklos skaitmeninę ekosistemą ir ją papildyti. Suderinamumo aspektai apima tokius dalykus kaip duomenų sąveika, sistemos mastelio keitimas ir diegimo paprastumas. **Įvertindamos AI sistemų techninį suderinamumą, švietimo įstaigos gali užtikrinti sklandžią šių technologijų integraciją ir optimalų jų panaudojimą esamose sistemose**.

3.2. Naudingumas

Vertinimo schemos naudingumo komponentas **nagrinėja AI taikomųjų programų praktinę naudą švietimo kontekste**. Šis vertinimas sutelkia dėmesį į apčiuopiamų privalumų ir rezultatų, gaunamų diegiant AI įrankius, įvertinimą, ypač kibernetinio priekabiavimo prevencijos ir mokinių gerovės

sirtyse. Pagrindinis šio vertinimo tikslas – ištirti, kaip veiksmingai AI taikomosios programos prisideda prie kibernetinio priekabiavimo atvejų skaičiaus mažinimo ir saugesnės internetinės aplinkos mokiniams kūrimo.

Be kibernetinio priekabiavimo prevencijos, **naudingumo aspektas taip pat apima išsamią AI technologijų teikiamo platesnio masto poveikio švietimui analizę.** Tai apima jų veiksmingumo vertinimą teikiant individualizuotas mokymosi patirtis, didinant mokinių įsitraukimą ir gerinant akademinius rezultatus. Įvertindami AI taikomųjų programų naudingumą ne tik kibernetinio priekabiavimo prevencijos srityje, švietimo srities suinteresuotosios šalys gali gauti įžvalgų apie visapusišką naudą ir galimus trūkumus, susijusius su AI integravimu į švietimo praktiką

3.3. Pritaikymas

Vertinimo schemos dalis, skirta diegimui, yra skirta įvertinti, **koku mastu dirbtinio intelekto technologijos yra priimanamos ir integruojamos į švietimo praktiką.** Šis vertinimas tiria, kaip mokytojai, mokiniai ir kiti mokyklos bendruomenės nariai priima ir naudoja dirbtinio intelekto priemones. **Juo siekiama nustatyti dirbtinio intelekto technologijų diegimo lygį ir veiksnius, darančius įtaką jų įsisavinimui švietimo aplinkoje.**

Be to, **diegimo vertinimo komponentas taip pat apima atitikties ES vertybėms, etikos standartams ir duomenų privatumo bei apsaugos reguliavimo sistemoms vertinimą.** Tai apima vertinimą, ar dirbtinio intelekto taikomosios programos, jas diegiant ir naudojant, atitinka sąžiningumo, skaidrumo ir atskaitingumo principus. Vertindamos dirbtinio intelekto technologijų diegimą atsižvelgdamos į etinius ir reguliavimo aspektus, švietimo įstaigos gali užtikrinti atsakingą ir tvarų šių technologijų integravimą į savo veiklą.

4 Vertinimo veiklos

TEORINĖ DIMENSIJA			
Duomenų konfidencialumo atitikties vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs
Duomenų apsaugos minimalių reikalavimų vertinimas	Ši veikla apima išsamų dirbtinio intelekto pagrįstos mokymo priemonės ar praktikos vertinimą, siekiant užtikrinti, kad būtų laikomasi minimalių duomenų privatumo reikalavimų. Ji apima duomenų rinkimo, saugojimo, tvarkymo praktikų bei saugumo priemonių, įgyvendintų siekiant apsaugoti mokinių duomenis, vertinimą.	Išsami ataskaita, kurioje apibūdinama dabartinė duomenų privatumo padėtis dirbant su dirbtinio intelekto įrankiu ar taikant dirbtinio intelekto praktiką, pabrėžiant atitikties reikalavimams sritis bei galimas spragas ar rizikas.	• Duomenų privatumo reikalavimų, kuriuos atitinka AI įrankis ar praktika pagal aukštesnius apsaugos standartus, procentinė dalis. • Nustatytų duomenų privatumo rizikos veiksnių arba pažeidžiamumų skaičius. • Rekomenduotų priemonių, skirtų duomenų privatumo spragoms pašalinti, įgyvendinimo būklė.



Etikos reikalavimų laikymosi vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs
Etinės rizikos potencialių veiksmų nustatymas ir analizė	Ši veikla apima galimų etinių rizikų, susijusių su dirbtinio intelekto naudojimu švietimo aplinkoje, nustatymą ir analizę. Ji apima galimų šališkumų, diskriminacijos, skaidrumo trūkumo ar kitų etinių problemų, būdingų dirbtinio intelekto pagrįstoms švietimo priemonėms ar praktikoms, identifikavimą.	Išsami etinio rizikos vertinimo ataskaita, kurioje išskiriamos galimos etinės dilemos, šališkumo atvejai ir rizika, susijusi su dirbtinio intelekto įrankiu ar praktika, taip pat pateikiamos rekomendacijos dėl rizikos mažinimo strategijų.	• Nustatytų etinių rizikų ar problemų skaičius. • Nustatytų etinių rizikų rimtumo įvertinimas. • Rekomenduotų rizikos mažinimo strategijų įgyvendinimo būklė.
ES rekomendacijų vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs
Įgyvendintų sprendimų ir naujausių ES politikos rekomendacijų palyginimas	Ši veikla apima švietimo įstaigose įdiegtų dirbtinio intelekto sprendimų savybių ir funkcionalumų palyginimą su naujausiomis Europos Sąjungos (ES) pateiktomis politikos rekomendacijomis. Ji apima atitikties ES reglamentams, gairėms ir etikos principams vertinimą.	Lyginamoji analizės ataskaita, kurioje pabrėžiamas įgyvendintų dirbtinio intelekto sprendimų atitikimo naujausioms ES politikos rekomendacijoms laipsnis, nurodant sritis, kuriose reikalavimai yra įvykdyti, ir sritis, kuriose galima tobulėti.	• Atitikties ES politikos rekomendacijoms procentinė dalis. • Sričių, kuriose įgyvendinti sprendimai nukrypsta nuo ES rekomendacijų, nustatymas. • Rekomendacijų, skirtų didesnei atitikčiai ES politikai užtikrinti, įgyvendinimo būklė.

POVEIKIO MASTAS			
Gebėjimo vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs
Dirbtinio intelekto sistemų techninis vertinimas	Atlikti išsamų švietimo įstaigose diegtų dirbtinio intelekto sistemų techninį vertinimą. Tai apima dirbtinio intelekto įrankių algoritmų, tikslumo rodiklių ir veiklos rodiklių vertinimą, susijusių su kibernetinio priekabiavimo atvejų nustatymu ir sprendimu.	Stebėjimo ataskaita, kurioje apibūdinami dirbtinio intelekto sistemų techniniai pajėgumai ir apribojimai, taip pat pateikiamos rekomendacijos dėl tobulinimo.	• Dirbtinio intelekto algoritmų tikslumas nustatant kibernetinio priekabiavimo atvejus. • Veiklos rodikliai, tokie kaip klaidingų teigiamų ir klaidingų neigiamų rezultatų rodikliai. • Suderinamumas su esama švietimo infrastruktūra ir platformomis.
Naudingumo vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs

Dirbtinio intelekto taikomųjų programų poveikio vertinimas	Jvertinti dirbtinio intelekto taikomųjų programų praktinę naudą ir poveikį švietimo kontekste. Tai apima jų veiksmingumo vertinimą skatinant mokinių gerovę, mažinant kibernetinio priekabiavimo atvejus ir gerinant mokymosi rezultatus.	Vertinimo ataskaita, kurioje išskiriamos konkrečios naudos ir rezultatai, gauti įdiegus dirbtinio intelekto priemones, taip pat pateikiamos rekomendacijos, kaip optimizuoti jų naudojimą.	• Kibernetinio priekabiavimo atvejų skaičiaus sumažėjimas įdiegus dirbtinio intelekto programas. • Mokinių gerovės ir saugumo pagerėjimas. • Akademinių rezultatų ir aktyvumo pagerėjimas.
Pritaikymo vertinimo kriterijai			
Veikla	Aprašymas	Laukiamas rezultatas	KPIs
Suinteresuotųjų šalių nuomonės tyrimas	Atlikite apklausas tarp pedagogų, studentų ir kitų suinteresuotųjų šalių, siekdami įvertinti jų nuomonę apie dirbtinio intelekto technologijas švietimo aplinkoje ir jų priimtinumą. Tai apima požiūrio į dirbtinio intelekto diegimą įvertinimą, diegimo kliūčių nustatymą ir atsiliepimų apie naudojimo patogumą rinkimą.	Apklausos analizės ataskaita, kurioje apibendrinami suinteresuotųjų šalių įžvalgos ir rekomendacijos, kaip skatinti dirbtinio intelekto technologijų diegimą.	• Švietimo specialistų ir studentų pritarimo lygis bei pasiryžimas diegti AI technologijas. • Nustatytų pagrindinių kliūčių, trukdančių diegti ir įgyvendinti šias technologijas, skaičius. • Atsiliepimai apie AI programų patogumą ir vartotojo patirtį.

5 Išvados ir Įspėjimas

Apibendrinant galima teigti, kad šiame dokumente pateikta „Dirbtinio intelekto švietime vertinimo schema“ yra pagrindinis pagrindas, turintis lemiamą reikšmę vertinant dirbtinio intelekto pagrįstų švietimo priemonių ir praktikų etinius, privatumo ir atitikties reikalavimams aspektus. **Nepaisant šio dokumento aktualumo, jame yra tam tikrų trūkumų, į kuriuos skaitytojas turi atsižvelgti.**

Nors „Dirbtinio intelekto švietime vertinimo schema“ siekia pateikti išsamią sistemą, būtina pripažinti švietimo srities būdingą sudėtingumą ir įvairovę. Suprasdamos, kad nė viena schema negali visapusiškai apimti visų niuansų, įstaigos raginamos taikyti lankstų požiūrį. Jos turėtų pritaikyti ir integruoti siūlomas vertinimo veiklas, kad jos atitiktų jų unikalią aplinką, institucinę kultūrą ir įvairių tikslinės grupės narių konkrečius poreikius. Taip elgdamasi, įstaigos gali užtikrinti, kad vertinimo procesas išliktų aktualus, prasmingas ir veiksmingas jų konkrečioje švietimo aplinkoje.

Atsižvelgiant į etinių dilemų iš esmės kokybinį pobūdį, etiniai aspektai yra bet kokios vertinimo schemos pagrindas. Dirbtinio intelekto taikymo švietimo srityje etinė analizė reikalauja kruopštaus tyrimo, atsižvelgiant į galimus šališkumus, kultūrinius niuansus ir kintančius visuomenės jautrumus. Tai reikalauja išsamaus etinių pasekmių, būdingų dirbtinio intelekto technologijų projektavimui, diegimui ir naudojimui švietimo aplinkoje, tyrimo. Vykdydamos apgalvotą ir refleksyvią etinį tyrimą, institucijos gali orientuotis sudėtingoje etinėje srityje, laikytis teisingumo ir lygybės principų bei skatinti etišką sprendimų priėmimą, kuriuo pirmiausia atsižvelgiama į visų suinteresuotųjų šalių gerovę ir teises.

Šioje schemoje apibūdintas poveikio vertinimas skirtas padėti įstaigoms įgyvendinti projektus ekonomiškai efektyviai, atsižvelgiant į išteklių apribojimus, su kuriais dažnai susiduriama švietimo įstaigose. Tačiau įstaigoms, turinčioms galimybių skirti papildomų išteklių, atsiveria naujovių diegimo ir eksperimentavimo galimybės. Investuodamos į AI technologijas ir švietimo praktikas, šios įstaigos gali tyrinėti naujus požiūrius, bandyti novatoriškus sprendimus ir diegti naujas metodikas, kurios gali žymiai pagerinti švietimo aplinką. Strateginės investicijos ir apdairus išteklių paskirstymas leidžia įstaigoms atverti naujas galimybes, skatinti teigiamus pokyčius ir puoselėti inovacijų kultūrą dirbtinio intelekto taikymo švietime srityje.

Labai svarbu suprasti, kad šis dokumentas yra pirmasis žingsnis tęsiamame „AI@Mediators“ projekto kelyje. Kadangi technologijos ir reguliavimo aplinka toliau sparčiai keičiasi, suinteresuotosios šalys raginamos nuolat sekti naujienas ir būti informuotos apie projekto eigą. Be to, suinteresuotosios šalys raginamos remtis kitais, naujesniais projekto rezultatais, kuriuose gali būti pateikta išsamesnė informacija, atnaujintos gairės ir patobulintos metodikos. Aktyviai dalyvaujant ir reaguojant į naujas tendencijas bei pokyčius, suinteresuotosios šalys gali užtikrinti, kad jų požiūris į AI švietime išliktų aktualus, veiksmingas ir etiškai pagrįstas, o tai galiausiai prisidės prie švietimo praktikos tobulinimo ir besimokančiųjų gerovės visame pasaulyje.

6 Bibliografines nuorodos

- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>
- JISC (2022) AI in tertiary Education. A summary of the current state of play. Retrieved from <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>
- Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocaranza-Prado, I., & Kreibel, D. (2023). Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: revisión sistemática de la literatura. *Revista de Psicodidáctica*, 28(2), 93-103. <https://doi.org/10.1016/j.psicod.2023.06.001>
- UNESCO (2023). ChatGPT and artificial intelligence in higher education. Retrieved from https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf
- UNESCO (2023). Guidance for generative AI in education and research. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Van Laar, E., Van Deursen, A. J., Van Dijk, J. A., & De Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in human behavior*, 72, 577-588. <https://doi.org/10.1016/j.chb.2017.03.010>
- Hinduja, S., & Patchin, J. W. (2015). Cyberbullying prevention and response: Expert perspectives. Routledge. https://www.researchgate.net/publication/345269061_Patchin_J_W_Hinduja_S_Cyberbullying_Prevention_and_Response_Expert_Perspectives_New_York_Routledge_2012_204_pages
- Smith, P. K., & Steffgen, G. (Eds.). (2013). Cyberbullying through the new media: Findings from an international network. Psychology Press. <https://psycnet.apa.org/record/2013-37310-000>
- Cénat, J. M., Hébert, M., Blais, M., Lavoie, F., Guerrier, M., & Derivois, D. (2014). Cyberbullying, psychological distress and self-esteem among youth in Quebec schools. *Journal of affective disorders*, 169, 7-9. https://www.researchgate.net/publication/264084590_Cyberbullying_psychological_distress_and_self-esteem_among_youth_in_Quebec_Schools
- Juvonen, J., & Gross, E. F. (Eds.). (2008). Extending the school grounds?—Bullying experiences in cyberspace. *Journal of School Health*, 78(9), 496-505. https://www.researchgate.net/publication/23252828_Extending_the_School_Grounds-Bullying_Experiences_in_Cyberspace
- Patchin, J. W., & Hinduja, S. (2011). Cyberbullying prevention and response: Expert perspectives. Routledge. <https://www.routledge.com/Cyberbullying-Prevention-and-Response-Expert-Perspectives/Patchin-Hinduja/p/book/9780415892377>
- Belsey, B. (2004). Cyberbullying: An emerging threat to the “always on” generation. Cyberbullying Conference, 2-3. <https://billbelsey.com/?p=1827>
- Spears, B., Slee, P., Owens, L., & Johnson, B. (2009). Behind the scenes and screens: Insights into the human dimension of covert and cyberbullying. *Social Inclusion*, 35-47. https://www.researchgate.net/publication/232459474_Behind_the_Scenes_and_Screens_In_sights_into_the_Human_Dimension_of_Covert_and_Cyberbullying