

Biuletenio tematika ir tema

Kultūra ir kultūros politika

Biuletenio laidos antraštė, probleminis klausimas

Lietuvių kalba ir dirbtinis intelektas

Esminiai žodžiai

Dirbtinis intelektas, kalbos technologijos, lietuvių kalba, lietuvių kalbos ištekliai, Europos Sąjunga, Lietuva, reglamentavimas

Serija ir registracijos numeris

R-2026-5

Leidimo data

2026-05-28

Leidimo vieta

Vilnius

Žanras

Analitinė apžvalga Kita

Šaltiniai: kategorijos

- Teisės aktai Politinė komunikacija
- Analitinių centrų kūriniai / leidiniai
- Žiniasklaidos turinys Socialinių tinklų turinys
- Statistiniai duomenys Mokslo darbai
- Metainformaciniai produktai
- Išviešinti slapti / privatūs duomenys

Šaltiniai: nuo - iki

1998 07 23 -
2026 05 05

Šaltiniai: kalbos

- Lietuvių k. Lenkų k.
- Anglų k. Kitos ES kalbos
- Rusų k. Kitos

Citavimui (APA stiliumi)

Nacionalinė biblioteka, Informacijos analitikos skyrius (2026). *Lietuvių kalba ir dirbtinis intelektas* (R-2026-5). Vilnius.

Kontaktiniai duomenys

Informacijos analitikos skyrius; analitika@lnb.lt. Nacionalinė biblioteka, Gedimino pr. 51, 01109 Vilnius.

Turinio apžvalga

Analitinėje apžvalgoje lietuvių kalbos integracija į dirbtinio intelekto technologijas analizuojama šiais aspektais:

- Europos Sąjungos pozicija ir reglamentacija dėl kalbų plėtros dirbtinio intelekto modeliuose;
- nacionalinis reguliavimas ir politika dirbtinio intelekto bei kalbos technologijų srityje; subjektai;
- lietuvių kalba ir dirbtinis intelektas: būklė, problematika ir perspektyvos.

1. Įžanga

Dirbtinis intelektas (DI) ir kalba – sritys, kurios šiuolaikiniame pasaulyje vis glaudžiau sąveikauja. Sparčiai tobulėjant technologijoms, DI geba ne tik apdoroti didžiulius kalbinių duomenų kiekius, bet ir suprasti, generuoti bei interpretuoti žmogaus kalbą. Toks sąlytis atveria naujų galimybių komunikacijai, švietimui, verslui ir kūrybai, tačiau kyla ir kai kurių svarbių klausimų, pavyzdžiui, kaip užtikrinti kalbų lygybę, kalbos autentiškumą, etikos standartus, kas atsakingas už klaidas ir t. t. Kalbant apie DI ir kalbos sąsają, akivaizdu, kad technologijos keičia ne tik visuomenės komunikavimo būdus, bet veikia ir kalbos raidą. Be to, šis ryšys yra abipusis: kalba taip pat turi įtakos, kiek efektyvus bus DI – jo sistemų tikslumas tiesiogiai priklauso nuo kalbinių duomenų kokybės, įvairovės ir konteksto.

Analitinių apžvalgų archyvas: <https://lnb.lt/istekliai/kiti-istekliai/analitines-apzvalgos>

Dirbtinio intelekto technologijos pastaruju metu itin sparčiai keičia kalbos vartojimo, kūrimo ir apdorojimo formas. Įvairūs įrankiai, pavyzdžiui, automatinio vertimo sistemos, tekstų generavimo platformos ar kalbos atpažinimo programos, kasdien ir plačiai naudojami tiek profesinėje, tiek asmeninėje aplinkoje. Šios technologijos daugiausia kuriamos remiantis didelės apimties kalbiniais duomenimis, taigi ir jų veikimo kokybė tiesiogiai priklauso nuo konkrečios kalbos išteklių, kuriais buvo apmokytos, gausos ir prieinamos. Šis tikslumas ypač svarbus generatyvinio dirbtinio intelekto sistemoms (angl. *Generative Artificial Intelligence (Generative AI)*), kurios ne tik apdoroja pateiktą informaciją, bet ir kuria naują turinį – tekstus, kitus kalbinius vienetus, verčia. Tokie modeliai, gebėdami atpažinti skirtingų kalbų struktūras, gali generuoti nuoseklų ir taisyklingą tekstą, vis dėlto pateikiamo turinio natūralumas, sklandumas ir patikimumas labai priklauso nuo duomenų kokybės bei apimties.

Šiame kontekste lietuvių kalba, kaip nedidelės kalbinės bendruomenės komunikacijos instrumentas, susiduria su specifiniais iššūkiais. Viešuosiuose šaltiniuose dažnai pabrėžiama, kad mažesnių kalbų integracija į DI sistemas yra lėtesnė ir ne visuomet tolygi, palyginti su globaliai dominuojančiomis kalbomis, tokiomis kaip anglų, ispanų ar kinų. Dėl riboto duomenų kiekio ir sudėtingos lietuvių kalbos struktūros DI įrankiai gali ne visai tiksliai interpretuoti, versti ar generuoti tekstus šia kalba. Taigi, aktualus klausimas: kiek efektyviai lietuvių kalba funkcionuoja šiuolaikinėse kalbos technologijose ir kokios jos ateities perspektyvos skaitmeninėje sferoje.

Šio darbo tikslas – remiantis viešaisiais šaltiniais apžvelgti lietuvių kalbos funkcionavimą dirbtinio intelekto įrankiuose ir įvertinti pagrindinius iššūkius bei galimybes. Analizuojant aktualius dokumentus, institucijų komunikaciją, technologinių sprendimų kūrėjų pateikiamas įžvalgas, siekiama nustatyti, ar lietuvių kalba yra pakankamai integruota į šiuolaikines DI sistemas, taip pat – kokios tendencijos gali lemti jos raidą dirbtiniame intelekto.

2. Europinis kontekstas, teisinis pagrindas

Dirbtinio intelekto plėtra Europos Sąjungos (ES) lygmeniu susijusi su siekiu užtikrinti visų oficialiųjų jos kalbų lygiavertį atstovavimą skaitmeninėje erdvėje. Dėl to itin svarbi tampa DI technologijų reglamentacija, apimanti tiek kalbų prieinamumo užtikrinimą, tiek duomenų kokybės, etikos ir skaidrumo reikalavimus. Šiame kontekste ES politikos DI sektoriuje formuotojai technologinę pažangą siekia suderinti su kalbų įvairove ir lingvistinių teisių apsauga.

Europos Sąjunga atlieka įvairias šios srities reguliavimo ir koordinavimo funkcijas – jos institucijos rengia pamatinius teisės dokumentus ir standartus, reikšmingus nacionalinei politikai, taip pat numato DI vystymosi strategijas. Teisėkūros procese dėl DI ES mastu dalyvauja šios svarbiausios jos institucijos:

Europos Komisija, sudaryta iš 27 komisarų, veikia kaip vykdomosios valdžios organas, rengia įstatymus;

Europos Parlamentas atstovauja ES piliečiams; **Europos Sąjungos Taryba**, sudaryta iš atitinkamos srities ministrų, kartu su Parlamentu svarsto Komisijos pateiktus įstatymų projektus – oficialiai patvirtinus

Analitinių apžvalgų archyvas: <https://lnb.lt/istekliai/kiti-istekliai/analitines-apzvalgos>

tekstą, jis tampa oficialiu ES leidiniu.¹ Be šių institucijų, DI kontekste minimas **Europos Komisijos DI biuras**, veikiantis kaip ES dirbtinio intelekto ekspertizės centras. Jo vaidmuo yra skatinti, kad būtų kuriamas ir diegiamas žmonijai ir ekonomikai naudingas DI. Biuras taip pat remia patikimo DI kūrimą, saugo nuo rizikų, yra atsakingas už DI akto įgyvendinimą.²

ES reglamentavimas, nuostatos, kryptys. Bendrija atlieka svarbią funkciją formuojant bendrą teisinę ir reguliavimo sistemą dėl skaitmeninių technologijų, tarp jų ir DI, plėtros visose valstybėse narėse. Ji yra įstatymų leidėja ir turi teisėkūros galią tvirtinti kitus reglamentus ir direktyvas, kurie yra tiesiogiai taikomi ir padeda užtikrinti vienodą taisyklių laikymąsi bendrojoje DI rinkoje. Be to, ES derina valstybių narių politiką tokiose srityse kaip duomenų apsauga, technologijų vystymas ir inovacijų skatinimas. Šiuo metu galioja ne vienas teisės aktas, svarbus DI sistemų reguliacijai ES mastu, o svarbiausi dokumentai ir iniciatyvos yra šie:

- **ES Dirbtinio intelekto aktas** (Reglamentas (ES) 2024/1689). 2024 m. vasarą priimtame ir tų pačių metų rugpjūčio 1 d. įsigaliojusiame dokumente nurodoma, kad **jo „paskirtis - pagerinti vidaus rinkos veikimą nustatant vienodą teisinę sistemą“**. Šis reglamentavimas visų pirma skirtas DI sistemoms kurti, pateikti rinkai ir naudoti ES paisant jos vertybių, skatinti į žmogų orientuoto ir patikimo dirbtinio intelekto įsisavinimą kartu užtikrinant aukšto lygio sveikatos, saugumo, pagrindinių teisių, įtvirtintų Europos Sąjungos pagrindinių teisių chartijoje, apsaugą, įskaitant demokratijos, teisinės valstybės ir aplinkos apsaugą, saugoti nuo žalingo DI sistemų poveikio ES ir remti inovacijas. Šiuo reglamentu užtikrinamas laisvas DI grindžiamų prekių ir paslaugų tarpvalstybinis judėjimas, taip užkertant kelią valstybėms narėms nustatyti apribojimus DI sistemų kūrimui, teikimui rinkai ir naudojimui, nebent tai aiškiai leidžiama pagal šį reglamentą. Nors dokumente tiesiogiai kalbų integracijos į DI modelius principai nėra numatyti, tai susiję su reikalavimais duomenų aprėpčiai, skaidrumui, siekiant išvengti šališkumo. Todėl ir **kokybės standartai sėkmingiau įgyvendinami esant geresniam duomenų prieinamumui konkrečiomis kalbomis**.³
- **2030 m. Skaitmeninio dešimtmečio politikos programa** yra ES planas, kuriame išskelti šie esminiai tikslai: ugdyti skaitmeninių įgūdžių turinčius gyventojus ir ruošti aukštos kvalifikacijos skaitmeninės srities specialistus; skatinti saugią ir tvarią šios srities infrastruktūrą; vykdyti įmonių skaitmeninę transformaciją; įgyvendinti viešųjų paslaugų skaitmenizaciją. Čia kalbų plėtra DI nėra formuluojama kaip savarankiška politikos dalis, vis dėlto DI infrastruktūros gerinimas neatsiejamas nuo kalbų technologijų raidos.⁴
- **Kalbų technologijos ir daugiakalbystė.** Didelis dėmesys skiriamas dirbtiniu intelektu grindžiamų kalbinių sprendimų, tokių kaip automatinis vertimas, kalbos atpažinimas ir teksto analizė, vystymui bei diegimui viešajame ir privačiajame sektoriuose. Taip pat pabrėžiama kalbinių duomenų,

¹ Institutional Context. (n. d.) Prieiga per internetą: <https://artificialintelligenceact.eu/context/>

² Europos Komisija. (2022 m. gruodžio 14 d.). Europos dirbtinio intelekto biuras. Prieiga per internetą: <https://digital-strategy.ec.europa.eu/en/policies/ai-office>

³ Europos Sąjunga. (2024 m. birželio 13 d.). Europos Parlamento ir Tarybos reglamentas (ES) 2024/1689. Prieiga per internetą: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

⁴ Europos Sąjunga. (2022 m. gruodžio 14 d.). Europos Parlamento ir Tarybos sprendimas (ES) 2022/2481. Prieiga per internetą: <https://eur-lex.europa.eu/legal-content/LT/TXT/?uri=CELEX:32022D2481>

inovacijų ir Europos technologinio savarankiškumo svarba. Europos Sąjunga skatina kalbinių technologijų plėtrą siekdama mažinti kalbos barjerus skaitmeninėje erdvėje ir užtikrinti daugiakalbystę. Organizacija orientuojasi į priemonių kūrimą, tam tikslui panaudodama naujausias DI technologijas, o DI mokomas taikant didžiulius kalbinius išteklius.⁵

- **Skaitmeninės Europos programa (DIGITAL).** ES finansavimo programa, skirta sudaryti sąlygas įmonėms, gyventojams ir viešojo sektoriaus institucijoms plačiau naudotis skaitmeninėmis technologijomis. Pagrindinės rėmimo sritys: superkompiuterija, dirbtinis intelektas, kibernetinis saugumas, pažangūs skaitmeniniai įgūdžiai, o jos neatsiejamos nuo kalbų technologijų DI įrankiuose būklės.⁶
- **Kalbų lygybės (ELE) iniciatyva.** Šio projekto, vykusio 2021 m. sausį–2022 m. birželį, tikslas – sukurti Europos kalbų lygybės programą, apimančią strateginę mokslinių tyrimų, inovacijų ir įgyvendinimo kryptį bei veiksmų planą, skirtą iki 2030 m. užtikrinti visapusišką skaitmeninę kalbų lygybę Europoje. Projektą vystė 52 dalyvių konsorciumas iš visų Europos šalių, mokslo ir tyrimų sektorių. Situacija dėl Europos kalbų skaitmeninėje terpėje pasižymi disbalansu, nes daugiausia joje funkcionuoja anglų kalba, ir apskritai įvairių kalbų situacija itin skiriasi. Buvo numatyti konkretūs žingsniai, kurie, remiantis projekto metu suformuluotomis rekomendacijomis, padėtų šią nelygybę sumažinti.⁷
- **Kalbų technologijų aljansas (ALT-EDIC).** Jam priklauso septyniolika valstybių narių, tarp kurių – ir Lietuva. ALT-EDIC siekia kurti bendrą Europos duomenų infrastruktūrą ir kalbos technologijų paslaugas, stiprinti technologinį konkurencingumą bei palaikyti kultūrinę įvairovę. Kalbinių bei multimodalinių duomenų rinkimas ir konsolidacija sudaro palankias sąlygas pažangiems daugiakalbiams didiesiems kalbos modeliams sukurti.⁸
- **Europos skaitmeninės infrastruktūros konsorciumas (ESIK)** – teisinė struktūra, sukurta pagal 2030 m. Skaitmeninio dešimtmečio politikos programą, kad ES valstybės narės galėtų veiksmingai sutelkti išteklius jungtinei skaitmeninei infrastruktūrai plėtoti. Užtikrinant bendrus standartus, orientuojamasi į daugiašalius projektus, ypač į tuos, kurių valstybės narės pačios įgyvendinti negali.⁹

Formuojant europinę DI politiką, dalyvauja aukščiausios ES institucijos. Jos įstatymuose įtvirtinusios bendras dirbtinio intelekto ir kalbų technologijų plėtros kryptis ir teisinius pagrindus, galiojančius visose Bendrijos narėse. Įvairiose programose ir iniciatyvose apibrėžtos kitos svarbios gairės, kurios turi užtikrinti, kad DI būtų kuriamas ir taikomas atsakingai, laikantis skaidrumo, saugumo bei žmogaus teisių apsaugos principų. Be to, ES remia mokslinius tyrimus, skatina inovacijas, siekia, kad į DI integruojant kalbas būtų atsižvelgiama į kultūrinę ir kalbinę Bendrijos įvairovę.

⁵ Europos Komisija. (2024 m. spalio 28 d.). Kalbų technologijos ir daugiakalbystė. Prieiga per internetą: <https://digital-strategy.ec.europa.eu/en/policies/language-technologies-multilingualism>

⁶ Europos Komisija. (2025 m. rugsėjo 26 d.). Skaitmeninės Europos programa. Prieiga per internetą: <https://digital-strategy.ec.europa.eu/en/activities/digital-programme>

⁷ European Language Equality. (n. d.). Developing an agenda and a roadmap for achieving full digital language equality in Europe by 2030. Prieiga per internetą: <https://european-language-equality.eu/>

⁸ Europos Komisija. (n. d.). Kalbų technologijų aljansas EDIC: Kalbų ir kultūrų įvairovės išsaugojimas Europoje ir technologinės kompetencijos bei lyderystės skatinimas. Prieiga per internetą: https://language-data-space.ec.europa.eu/related-initiatives/alt-edic_en

⁹ Kas yra „Europe Direct“ informacijos centras? (n. d.). Prieiga per internetą: <https://www.alt-edic.eu/lt/what-is-an-edic/>

3. Nacionalinis teisinis reguliavimas, institucijos

Lietuvos teisinė bazė DI srityje iš esmės remiasi ES dokumentuose suformuluotomis nuostatomis. Būdama jos nare, šalis atsižvelgia į organizacijos mastu taikomus reglamentus, juos lokaliai pritaiko ir detalizuoja. DI kūrimo, diegimo ir naudojimo principai glaudžiai siejasi su valstybinės lietuvių kalbos išsaugojimu. Esminė sąlyga – jos plėtra skaitmeninėje terpėje, todėl ypač svarbu užtikrinti ne tik duomenų lietuvių kalba prieinamumą, bet ir konkurencingumą daugiakalbėse DI sistemose. Atsakingų institucijų ir subjektų koordinuota veikla yra būtina siekiant garantuoti nuoseklų DI plėtros ir lietuvių kalbos technologinio vystymosi suderinamumą.

DI politika Lietuvoje iš esmės yra daugiapakopė. Aukščiausios institucijos ją formuoja, o joms pavaldžios įstaigos – atsakingos už įgyvendinimą ar priežiūrą. Aptariant lietuvių kalbos funkcionavimą DI ši valdymo sąranga yra svarbi. Kalbos technologijų plėtra ir lietuvių kalbos išteklių kūrimas, prieinamumas ir apsauga iš dalies priklauso nuo to, kokie instituciniai sprendimai bus priimami ir kokią DI kryptį pasirinks politikos formuotojai ir įgyvendintojai valstybiniame bei privačiame sektoriuose. Apskritai mažesnių kalbų skaitmeninę pažangą ir pritaikymą DI modeliuose lemia ne vien rinkos diktuojamos sąlygos, bet ir viešosios politikos darbotvarkė bei valstybės strateginiai tikslai.

Lietuvos Respublikos ekonomikos ir inovacijų ministerija (EIMIN) DI vystymo sektoriuje yra svarbi strateginių ir politinių iniciatyvų koordinatore. Jos veiklos nuostatuose įtvirtintas šis esminis tikslas: „formuoti valstybės politiką skaitmeninės plėtros srityje ir organizuoti, koordinuoti ir kontroliuoti šios politikos įgyvendinimą“. ¹⁰ Taigi Ministerija yra pagrindinė institucija, nustatanti DI politikos ir ekosistemos plėtros kryptis Lietuvoje. Ji atsakinga už inovacijų skatinimą ir palankių sąlygų DI vystyti kūrimą. Be kita ko, EIMIN **finansuoja lietuvių kalbos išteklių skaitmeninimo ir su jos funkcionavimu DI susijusius projektus**, taip pat kartu su Valstybės skaitmeninių sprendimų agentūra Lietuvoje kuria viešajam sektoriui skirtą Dirbtinio intelekto kompetencijų centrą „GovAI“. ¹¹

Ministerijos parengtos **2026–2035 metų nacionalinės dirbtinio intelekto strateginės gairės** – dokumentas, kuriame išdėstyti greito perėjimo nuo atskirų iniciatyvų prie bendros, nacionaliniu mastu veikiančios dirbtinio intelekto ekosistemos, kuri užtikrintų koordinuotą veiklą, siekiniai. Gairėse formuluojami šie su lietuvių kalba susiję teiginiai¹²:

- **technologinis savarankiškumas traktuojamas kaip lietuvių kalbos, kultūrinės tapatybės, demokratijos atsparumo bei nacionalinio saugumo klausimas.** „Lietuvai kaip mažai valstybei DI plėtra yra svarbi ne tik dėl ekonomikos modernizavimo, bet ir dėl lietuvių kalbos gyvavimo, tapatybės, demokratijos tvirtumo ir nacionalinio saugumo skaitmeninėje erdvėje.“¹³

¹⁰ Nutarimas dėl Lietuvos Respublikos ekonomikos ir inovacijų ministerijos nuostatų patvirtinimo. Lietuvos Respublikos Vyriausybės 1998 m. liepos 23 d. nutarimas Nr. 921. Prieiga per internetą: <https://www.e-tar.lt/portal/lt/legalAct/TAR.3CE5C3AA17FB/asr?csr=8530853368953836970>

¹¹ Lietuvos Respublikos ekonomikos ir inovacijų ministerija. (2026). DI plėtra. Prieiga per internetą: <https://eimin.lrv.lt/lt/skaitmenine-lietuva/1-kryptis-di-pletra/>

¹² Lietuvos Respublikos ekonomikos ir inovacijų ministerija. (2026). 2026–2035 metų nacionalinės dirbtinio intelekto strateginės gairės. Prieiga per internetą: https://eimin.lrv.lt/public/canonical/1776768134/6547/Nacionalines_DI_strategines_gaires.pdf

¹³ Ten pat.

- **Pasitelkus DI, būtų galima sukurti lietuvių kalba veikiančią „TOP 20 gyvenimo įvykių“ agentą**, kuris formuotų iš anksto parengtų prašymų šablonus, sumažintų statuso tikrinimo poreikį ir nukreiptų vartotoją atlikti tolimesnius veiksmus.
- **Siekiama kurti lietuvių kalbai pritaikytus ir specializuotus (angl. *on-edge/on-device*) modelius**, dėl kurių šalis taptų konkurencingesnė tarptautinėje erdvėje. Kita šalies konkurencinės jėgos sąlyga - **lietuvių kalbai pritaikyta DI ekosistema**.

Kaip vykdomosios valdžios institucijos DI infrastruktūros plėtros kontekste svarbų vaidmenį atlieka **Lietuvos Respublikos kultūros ir Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos** - jos formuoja kalbos, švietimo, kultūros ir skaitmeninių kompetencijų politiką. Trumpai tariant, Kultūros ministerija rūpinasi kalbos politika, lietuvių kalbos ištekliais, kultūros paveldu, skaitmenine kultūra. Švietimo ministerijos kompetencijos sritys - DI raštingumas, švietimo turinys, mokymas naudotis technologijomis lietuvių kalba. Abi institucijos kartu daro įtaką tam, kaip visuomenė supranta DI, kiek naudoja kalbos technologijas, ar ugdo skaitmeninius įgūdžius. Su DI plėtra susiję principai ir ministerijų veiklos kryptys įtvirtinti jų strateginiuose, planavimo, veiklos ir kituose dokumentuose.^{14, 15}

Valstybinė lietuvių kalbos komisija (VLKK) - kalbos institucija, kuri formuoja kalbos politiką, įgyvendina Valstybinės kalbos įstatymą, yra atsakinga už kalbos norminimą ir standartizavimą, konsultuoja jos taisyklingumo klausimais. Komisijos vaidmuo DI ir lietuvių kalbos sąveikos kontekste reikšmingas ne tik dėl išvardytųjų aspektų, ji dalyvauja vystant kalbos technologijas bei į DI sistemas diegiant lietuvių kalbą. Itin aktualūs aspektai: terminijos, kalbos išteklių standartizavimas, rekomendacijos dėl lietuvių kalbos vartojimo skaitmeninėje aplinkoje, DI sprendimų suderinamumas su lietuvių kalbos normomis ir visuomenės poreikiais. Aptariant VLKK funkcijas DI srityje, svarbu pabrėžti jos indėlį į kalbos infrastruktūros ir kalbinių išteklių vystymą - Kalbos technologijų pakomisė koordinuoja, teikia ekspertines rekomendacijas ir siūlymus, skatina lietuvių kalbos integravimą į skaitmenines technologijas, bendradarbiauja su DI įrankių kūrėjais.¹⁶

Lietuvių kalbos plėtros skaitmeninėje terpėje ir kalbos technologijų pažangos 2021-2027 metų gairėse, Lietuvos Respublikos Seimo priimtose 2020 m. paskutiniame ketvirtyje, numatyta, kad už jų įgyvendinimo stebėseną atsakinga VLKK. Pirmiausia dokumente akcentuojamas kalbos technologijų lietuvių kalba vystymas - aptariamoms konkrečioms technologinėms priemonėms, skirtoms darbui su kalba, ir ją apdorojantys įrankiai. Pavyzdžiui, tai - kalbos atpažinimo, kalbos sintezės, mašininio vertimo, tekstų analizės ir kt. mechanizmai. Dirbtinis intelektas, naudodamas kalbos išteklius, sukuria naują kalbą, tai yra rašytinį tekstą ar kitokio formato turinį, ir taip sąveikauja su vartotoju, ieškančiu aktualios informacijos.

¹⁴ Lietuvos Respublikos kultūros ministerija. (2026 m. balandžio 2 d.). Strateginiai dokumentai. Prieiga per internetą:

<https://lrkm.lrv.lt/lt/administracine-informacija/planavimo-dokumentai/strateginiai-dokumentai/>

¹⁵ Lietuvos Respublikos Švietimo, mokslo ir sporto ministerija. (2026 m. balandžio 16 d.). Planavimo dokumentai. Prieiga per internetą:

<https://smsm.lrv.lt/lt/administracine-informacija/planavimo-dokumentai/>

¹⁶ Dėl lietuvių kalbos plėtros skaitmeninėje terpėje ir kalbos technologijų pažangos 2021-2027 metų gairių patvirtinimo. Lietuvos Respublikos Seimo 2020 m. spalio 13 d. priimtas nutarimas Nr. XIII-3324. Prieiga per internetą: <https://www.e-tar.lt/portal/lt/legalAct/71152ab00eee11ebb74de75171d26d52>

Remiantis Gairėmis, galima išskirti esminius strateginius valstybės tikslus, susijusius su lietuvių kalbos plėtra informacinėse technologijose ir DI¹⁷:

- **lietuvių kalbos konkurencingumas** skaitmeninėje erdvėje;
- kad **DI pilnavertiškai veiktų lietuvių kalba, svarbūs skaitmeniniai ištekliai**, ir tai traktuojama kaip jos gyvybingumo išsaugojimo sąlyga;
- **kalbos duomenų, modelių, išteklių gausinimas ir kokybė yra esminiai veiksniai pažangiems DI sprendimams.**

DI plėtrai svarbios ir kitos institucijos. Viešoji įstaiga **Inovacijų agentūra** skatina DI modernizaciją, finansuoja dirbtinio intelekto projektus ir teikia pagalbą verslui bei viešajam sektoriui, diegiantiems pažangias technologijas. Atlieka mokslo ir verslo sinergijos funkciją, siekia kad DI sprendimai būtų praktiškai pritaikomi šalies ekonomikai kelti.¹⁸ Valstybės institucija **Ryšių reguliavimo tarnyba** užtikrina, kad DI ir skaitmeninės technologijos būtų diegiamos saugiai, laikantis teisės aktų, pirmiausia – ES DI akto, vartotojų apsaugos ir duomenų saugumo reikalavimų. Ji taip pat prisideda prie skaitmeninės infrastruktūros reguliavimo, kuri yra būtina, kad DI sistemos veiktų sklandžiai.¹⁹ Apibendrinant, abi institucijos – DI ekosistemos plėtros dalyvės: viena skatina augimą ir inovacijas, o kita užtikrina, kad tai būtų vykdoma atsakingai ir taikant kontrolę. Informacinės visuomenės modernizacijos procesuose šalyje dalyvauja ir **Valstybės skaitmeninių sprendimų agentūra**, jos veiklos tikslas – skaitmeninės transformacijos politikos įgyvendinimas.²⁰

Kalbant apie DI vystymą, įskaitant kalbos technologijų plėtrą ir lietuvių kalbos diegimą DI modeliuose, svarbią vietą šioje infrastruktūroje užima **mokslo organizacijos**, bendradarbiaujančios su valdžios institucijomis ir verslo sektoriais. Šiame kontekste pirmiausia minėtinas **Vilniaus universiteto** Matematikos ir informatikos bei Filologijos fakultetų mokslininkų vaisingas bendradarbiavimas. Didelis proveržis padarytas įgyvendinant projektą LIEPA (iš viso 3 etapai, paskutinis baigėsi šių metų balandžio mėnesio pabaigoje). Sukurtas lietuvių šnekos garsynas, išvystytas kalbos sintezatorius ir pan. Projekte taip pat dalyvauja **Vytauto Didžiojo universitetas** ir **Lietuvių kalbos institutas**.²¹ Aukštosios mokyklos ne tik prisideda prie DI plėtros, lietuvių kalbos integracijos jo modeliuose, apskritai jose palaipsniui kinta studijų procesas, turinys, nes pradeda atsirasti kombinuotų, su technologijų lingvistika susijusių studijų programų. Be to, didėjančios DI galimybės ir mokslininkams atveria daugiau perspektyvų produktyviems tyrimams vykdyti. Dar aptariamoje srityje aktyviai veikia **Kauno technologijų** bei **Vilnius Tech universitetai**.²²

¹⁷ Ten pat.

¹⁸ Inovacijų agentūra. (n. d.). Naujienos. Prieiga per internetą: <https://inovacijuagentura.lt/news?lang=lt>

¹⁹ Lietuvos Respublikos ryšių reguliavimo tarnyba. (n. d.). DI informacija. Prieiga per internetą: <https://rrt.lt/veiklos-sritys/skaitmenine-erdve/di-informacija>

²⁰ Valstybės skaitmeninių sprendimų agentūra. (2025 m. liepos 30 d.). Vizija, misija, veiklos tikslas ir vertybės. Prieiga per internetą: <https://vssa.lrv.lt/lt/apie-vssa/vizija-ir-vertybes/>

²¹ Vilniaus universitetas. (n. d.). Kai dirbtinis intelektas prabyla lietuviškai: VU projektas LIEPA, atveriantis galimybes filologams. Prieiga per internetą: <https://mif.vu.lt/lt/kas-vyksta-fakultete/naujienos/fakulteto-naujienos/5077-%C5%A1nekos-technologij%C5%B3-projektui-vadovaujanti-vu-mokslinink%C4%97-%E2%80%9Egeb%C4%97jimai-nepriklauso-nuo-lyties%E2%80%9C-2>

²² Kaip dirbtinio intelekto studijos keičia Lietuvos aukštąjį mokslą. (2025 d. gruodžio 8 d.). Prieiga per internetą: <https://faceit.lt/kaip-dirbtinio-intelektu-studijos-keicia-lietuvos-aukstaji-mokslo/>

Būtina paminėti ir privataus sektoriaus įmones, sėkmingai užpildančias nišą kalbos technologijų ir DI srityje. Tiek individualiai, tiek bendradarbiaudamos su valstybinėmis institucijomis, jos eksperimentuoja, išbando ir generuoja produktus, susijusius su kalbos technologijomis ir lietuvių kalbos pritaikymu DI infrastruktūrai. Įmonė „**Neurotechnology**“ specializuojasi natūralios kalbos apdorojimo srityje, o „**Tilde**“ siūlo plačias su skaitmenine kalba susijusias paslaugas – nuo vertimo iki kalbos apdorojimo sistemų.^{23, 24}

Dirbtinio intelekto plėtra Lietuvoje vykdoma bendromis įvairių institucijų pastangomis. Procese kaip strateginius prioritetus nustatančios organizacijos dalyvauja ministerijos, taip pat kalbos politiką formuojanti Valstybinė lietuvių kalbos komisija, įvairios viešojo sektoriaus įstaigos ir universitetai, privačios įmonės. Šio lauko veikėjai pagal jiems priskirtas kompetencijų sritis rūpinasi ne tik technologine pažanga, bet ir lietuvių kalbos išsaugojimu bei pritaikymu DI sprendimams. Toks koordinuotas darbas – pagrindas, kad inovacijos būtų derinamos su lietuvių kalbos visaverčiu funkcionavimu šiuolaikinėse technologijose.

4. Lietuvių kalbos padėtis DI aplinkoje

Dirbtinis intelektas dažnai susijęs su kalba – ja grindžiamas duomenų apdorojimas, sąveika su vartotoju, informacijos generavimas ir perteikimas. Todėl technologinis kalbos prieinamumas tiesiogiai veikia DI sistemų funkcionalumą ir rezultatų kokybę. Sparčiai vystantis DI modeliams, atsiranda įvairių kalbų gausių išteklių nepakankamo prieinamumo bei kalbų technologinės nelygybės problemų. Šiame skyriuje aptariama dabartinė lietuvių kalbos būklė dirbtinio intelekto sistemose, tendencijos bei galimos tolesnės raidos kryptys.

Skaitmeninės visuomenės transformacijos procesai nulėmė didėjantį technologijų vaidmenį kasdieniame gyvenime ir apdorojant informaciją. Dėl šios priežasties **kalba šiandien yra ne vien tik komunikacijos priemonė - ji pamatinis išteklius, užtikrinantis sklandžias DI funkcijas**. Dėl šių pokyčių atsirado poreikis užtikrinti ir lietuvių kalbos visavertį funkcionavimą skaitmeninėje erdvėje bei jos integraciją į pažangias technologijas.

Dar prieš kelerius metus skaitmeninio virsmo eroje apie kalbų būklę kalbėta jų technologijų kontekste, pavyzdžiui, tai – balso atpažinimo, rašybos ir gramatikos, automatinio vertimo modeliai, pokalbių robotai ir asistentai, paieškos sistemos ir pan. Jau tada atkreiptas dėmesys į mažųjų kalbų nekomercinį pobūdį.

Kalbos, turinčios mažą kalbėtojų skaičių, traktuojamos kaip ekonomiškai nepatrauklios, „neatsiperkančios“, todėl ir į skaitmeninių įrankių jomis plėtojimą investavimas vangesnis. Tokia nekomercine laikoma ir lietuvių kalba, vis dėlto jos funkcionavimas skaitmeninių paslaugų erdvėje

²³ Neurotechnology. (n. d.). Dirbtinis intelektas. Prieiga per internetą: <https://www.neurotechnology.com/artificial-intelligence.html>

²⁴ Tilde. (n. d.). Kalbos DI sprendimai saugiams ir efektyviems darbo procesams. Prieiga per internetą: <https://tilde.ai/lt/#services>

traktuojamas kaip būtina sąlyga kalbos raidai. Kompiuterinės lingvistikos mokslininko Andriaus Utkos tvirtinimu: „... pasmerkta ta kalba, kurios nėra skaitmeninėje erdvėje arba kurios ten yra labai mažai.“²⁵

Naujausiais duomenimis, didžioji dalis viso pasaulio DI modelių gerai veikia tik 100 kalbų, o kitos – šios pažangos užribyje. **Didžiausios investicijos šiame sektoriuje skiriamos anglų kalbai**, o daugelis mažųjų kalbų, ypač tų, kurios kovoja dėl išlikimo, ne tik ignoruojamos, bet ir dėl šališkumo matmens atrenkant duomenis yra diskriminuojamos. Tai turi tiek kultūrinių, tiek ekonominių pasekmių.²⁶ Nors omenyje turimos daugiausia regioninės mažumų kalbos, vis dėlto situacija verčia atkreipti dėmesį ir apskritai į daugelio mažesnių šalių padėtį. Taigi, nors tokios nedidelės kalbos kaip lietuvių (tai – viena iš oficialiųjų ES kalbų) dėl riboto komercinio patrauklumo sulaukia mažiau investuotojų į kalbos technologijas dėmesio, jų aktyvus plėtojimas skaitmeninėje erdvėje yra būtinas siekiant užtikrinti kalbos gyvybingumą ir ateitį.

Nepaisant nuogaštavimų dėl įvairių kalbų padėties, hierarchijos ir net galios pozicijų viena kitos atžvilgiu, lietuvių kalba pastaraisiais metais sparčiai žengia į dirbtinio intelekto technologijas. Kaip aptarta ankstesniuose skyriuose, nuoseklūs kalbų išteklių perkėlimo į skaitmeninę terpę veiksmai reglamentuojami ir ES, ir nacionaliniu lygmenimis. Lietuvių integracijos į DI raida nušviečiama nevienareikšmiškai – **nors džiaugiamasi pavieniais laimėjimais, vis dėlto akcentuojamas išteklių lietuvių kalba trūkumas ir tai, kad lietuvių kalba DI kontekste lieka nuošalyje.**²⁷

Pasaulyje sparčiai vystantis DI technologijoms, vienas esminių uždavinių tampa **didžiųjų kalbos modelių (angl. Large language model)**, tai yra dirbtinio intelekto sistemų, kurios sukurtos „suprasti“ ir generuoti žmogaus kalbą, plėtra. Paprasčiau tariant, tai – programos, kurios mokosi iš milžiniško kiekio tekstų ir remiantis jais gali: rašyti tekstus, atsakyti į klausimus, versti iš įvairių kalbų, pratęsti ir užbaigti mintį, padaryti turinio santrauką, klasifikuoti duomenis ir pan.²⁸

Iliustruojant, kas įvykdyta šioje srityje, galima paminėti kai kuriuos konkrečius atvejus. Pastaraisiais metais įgyvendinama daug svarbių projektų, į kuriuos įsitraukia ir viešojo, ir privataus sektoriaus subjektai. Pavyzdžiui, 2025 metų pabaigoje panaudojant **Bendrąjį lietuvių kalbos tekstyną sukurtas atvirosios prieigos Mažasis lietuvių kalbos vektorizuotas modelis** (LT-MLKM-modernBERT). Jo veikimas patikrintas taikant įvardytųjų esybių atpažinimo užduotį, tekstuose jis geba sėkmingai atpažinti asmenvardžius, vietovardžius, datas ir kitą svarbią informaciją, todėl modelis gali būti naudojamas tekstų anonimizavimui, informacijos paieškai, sentimentų analizei bei kitoms užduotims, o šiai funkcijai pritaikyta

²⁵ Lietuvių kalbos institutas. (2021 m. rugpjūčio 2 d.). Ar lietuvių kalba pajėgi varžytis pasaulinėse informacinių technologijų lenktynėse? Prieiga per internetą: <https://lki.lt/ar-lietuviu-kalba-pajegi-varzytis-pasaulinese-informaciniu-technologiju-lenktynese/>

²⁶ Dare, A. (2026 m. sausio 7 d.). The AI Bias Nobody Talks About: How Training Data Is Erasing Minority Languages. Prieiga per internetą: <https://medium.com/@afolabidare50/the-ai-bias-nobody-talks-about-how-training-data-is-erasing-minority-languages-65c81707c9fa>

²⁷ Maumevičienė, D. (2025 m. gruodžio 9 d.). KTU mokslininkė: dėl DI plėtros kalbos ir vertimo specialistų svarba žymiai išaugo. Prieiga per internetą: https://www.lrt.lt/naujienos/svietimas/45/2605136/ktu-mokslininke-del-di-pletros-kalbos-ir-vertimo-specialistu-svarba-zymiai-isaugo?srsId=AfmBOopVZyqIPyxpGq9NeQ1jp8alD5_jB9H_ycx7mBm-Q-baDUfteSp7

²⁸ What are Large Language Models (LLMs)? (n. d.). Prieiga per internetą: <https://www.salesforce.com/artificial-intelligence/what-are-large-language-models/#what-are-llms>

jo versija (LT-NER-modernBERT) yra viešai prieinama.²⁹ Kalbos technologijų startuolis „Kablelis.lt“ **oficialiai pristatė „Kamilę“ - lietuvių kalbai pritaikytą dirbtinio intelekto modelį**. Kūrėjų atlikti bandymai atskleidė netikėtą rezultatą: rašybos, skyrybos ir gramatikos taisymo srityje šis įrankis pranoksta populiariausius pasaulinius kalbos modelius.³⁰

Vis dėlto, net ir stebint teigiamas tendencijas, šis procesas kelia tam tikrų iššūkių. Naujausios lietuvių kalbos pritaikymo DI sistemose kryptys, reali padėtis ir probleminiai aspektai buvo svarstyti 2026 m. gegužės 5 d. Vilniaus apskrities Adomo Mickevičiaus bibliotekoje vykusioje **konferencijoje „Lietuvių kalba x Dirbtinis intelektas“**. Joje aptarta **išteklių lietuvių kalba DI situacija, lietuvių kalba DI sprendimuose, DI taikymas vertimo darbe ES institucijose ir kiti aktualūs susiję klausimai**, surengti keli praktiniai užsiėmimai. Kalbėta tiek apie naujas galimybes ir teigiamą technologijų įtaką kalbos sklaidai bei prieinamumui, tiek apie sunkumus, susijusius su reguliavimu, duomenų kokybe, etika ir mažųjų kalbų išlikimu sparčios technologinės plėtros kontekste. Remiantis renginyje pranešėjų pateikta informacija, galima išskirti esminius dirbtinio intelekto ir lietuvių kalbos plėtros aspektus³¹:

- **2022 m. datuojami kaip pažangių generatyvinio dirbtinio intelekto įrankių pradžios metai.**
- Tarp įvairių įmonių vyksta vadinamosios **technologinės lenktynės**; kalbos modeliams kurti skiriamos milijardinės sumos.
- **Daugiausia investuojama į didžiąsias kalbas, pirmiausia - anglų, kinų.** Taigi, DI būdingas anglocentrizmas ir sinocentrizmas, tai reiškia, kad **dirbtinis intelektas grindžiamas daugiausia tomis kalbomis kalbančių regionų vertybių sistema.**
- **Privalu gerinti lietuvių kalbos palaikymą DI** ir tam skirti daugiau finansinių resursų.
- Į lietuvių kalbos ir DI 16 projektų (mašininis vertimas, šnekos atpažinimas, šnekos sintezė, tekstynai, garsynai ir kt.) **investuoti 26 mln. Eur iš Ekonomikos gaivinimo ir atsparumo didinimo priemonės lėšų.** Didiesiems gamintojams nesame įdomūs, todėl minėti **atviri duomenys naudingi vietiniam verslui**, kuriančiam įvairias lietuviškai suprantančias ir kalbančias programas; Europos Sąjunga ir toliau investuoja į panašius projektus.
- **Mažoms kalboms, įskaitant lietuvių, būtinas nepriklausomumas**, tai reiškia, svarbi savų kalbos modelių, pasitelkiamų DI, plėtra.
- Reikia vystyti vietinę Bendrijos DI infrastruktūrą - tuo tikslu būtina **ES lygmeniu investuoti į mokslininkų pasirengimą kurti nepriklausomus kalbos modelius, skirtus DI**; siekti, kad būtume konkurencingi, o mažesnės ES kalbos nebūtų užgožtos.
- **Palaipsniui gerėja lietuvių kalbos reprezentavimas**, pavyzdžiui, DI pateikia vis sklandesnius sakinius, kurių struktūra būdinga šios kalbos sistemai, ir turinys nebegeneruojamas vien pagal anglų kalbos sakinių konstrukcijų analogiją.

²⁹ Valstybės skaitmeninių sprendimų agentūra. (2025 m. lapkričio 10 d.). Sukurtas pirmasis lietuvių kalbos dirbtinio intelekto modelis: lietuvių tyrėjų žingsnis į DI ateitį. Prieiga per internetą: <https://vssa.lrv.lt/lt/naujienos/sukurtas-pirmasis-lietuviu-kalbos-dirbtinio-intelektu-modelis-lietuviu-tyreju-zingsnis-i-di-ateiti-yx9/>

³⁰ Kaip apmokėme lietuvių kalbos taisymo modelį „Kamilę“. (2026 m. kovo 19 d.). Prieiga per internetą: <https://kablelis.lt/laiskai/llm-treniravimas>

³¹ Konferencija „Lietuvių kalba X dirbtinis intelektas“: [vaizdo įrašas]. (2026 m. gegužės 5 d.). Prieiga per internetą: https://www.youtube.com/watch?v=XGPH_SZIZyY

- Daug dėmesio **kalbos technologijų srityje skiriama pokalbių robotams**, pavyzdžiui, po 2022 m. tokiu įrankiu laikomas ir virtualaus asistento „ChatGPT“ pokalbių langelis. Be to, pokalbių robotai gali būti adaptuoti įmonėms, juos jau naudoja Valstybinė mokesčių inspekcija, valstybės įmonė „Regitra“ ir kitos įstaigos, verslas.
- Akcentuota, **kad dirbtinis intelektas yra žmogaus sukurtas, ir galutinis sprendimo priėmėjas, atliekant sudėtingas užduotis, visada yra žmogus.**
- Keltas klausimas, ar lietuvių kalba paranki DI integracijai ir ar ilgainiui bus perteikiamos jai būdingos subtilybės – ironija, tarmybės ir kt. **Ko DI išmokysime, tą ir turėsime, vis dėlto labai specifiniai dalykai, tokie kaip kirčiavimas, į DI diegiami bus itin lėtai.**
- **Generatyvinio DI suklestėjimas - individualumo ir savito, unikalaus stiliaus sumenkimo priešastis;** DI prilygintas jėgai, mūsų pasaulį paverčiančia nuobodžiu ir monotonišku.
- Nors **dėl DI plėtros transformavosi vertėjo specialybė**, vystosi mašininis vertimas, neuroninis mašininis vertimas, vis dėlto **ši profesija neišnyks taip greitai, kaip tikėtasi, visų pirma - dėl kokybės priežiūros veiksnio.** Šis darbas – hibridinis procesas, mažėja rankinių užduočių, o vertėjo funkcija susijusi su kūrybiškumu ir vertės užtikrinimu.
- Remiantis informacija apie **ES oficialiųjų ir neoficialiųjų kalbų skaitmeninimą**, natūralu, kad pirmauja anglų ir vokiečių kalbos, **Lietuva sąrašė rikiuojasi 24 pozicijoje, ji itin žema oficialiųjų ES kalbų kontekste** (ankstesnių metų skaičiavimais).
- Šaliai **trūksta specialistų, gebančių dirbti su kalbos duomenimis** – kompiuterinės lingvistikos, tekstynų lingvistikos profesionalų, DI etikos konsultantų kalbos srityje ir kt. Dėl to **universitetams vertėtų peržiūrėti studijų kryptių klasifikatorių.**
- Konferencijoje buvo minima, kad **ES turi 24 oficialiąsias kalbas**, todėl vertėjo darbas jos institucijose itin svarbus – **bet kuris organizacijos pilietis savo nuomonę ir užklausą gali pateikti gimtąja kalba.** Dėl DI įrankių darbas tampa produktyvesnis. Apskritai ES teikia daug DI grindžiamų daugiakalbų paslaugų, skirtų viešajam sektoriui, akademinėi bendruomenei, laisvai samdomiems vertėjams, įmonėms ir pan. Reikia pridurti, kad užklaustos vykdomos mašininio vertimo būdu, todėl nesklandumų pasitaiko. **Paprastai kalbos technologijų vystytojams lietuvių kalba nėra prioritetinga, bet ES vertimo segmente ji yra lygiavertė kitoms oficialiosioms kalboms.**

Nuo 2022 m. generatyvinio dirbtinio intelekto plėtra itin sparti, dėl to didžiosios technologijų bendrovės į kalbos modelių kūrimą investuoja milžiniškas lėšas. Tačiau daugiausia dėmesio skiriama anglų, kinų ir kitoms didesnėms kalboms, todėl mažesnės, tarp jų ir lietuvių kalba, susiduria su tam tikrais konkurenciniais keblumais skaitmeninėje erdvėje. Lietuvoje siekiama stiprinti lietuvių kalbos palaikymą DI sistemose, investuojama į kalbos technologijų projektus, bet akcentuojamas reikalingų specialistų trūkumas ir vis dar per mažas dėmesys lietuvių kalbos duomenų skaitmeninimui. Šiame kontekste itin pabrėžiamas žmogaus, kaip intelektualios ir kūrybiškos būtybės, vaidmuo DI informacijos priežiūros ir vertinimo procese.

5. Apibendrinimas ir akcentai, į ką atkreipti dėmesį ateityje


Dirbtinio intelekto plėtra Europos Sąjungos lygmeniu susijusi su siekiu užtikrinti visų oficialiųjų jos kalbų lygiavertį atstovavimą skaitmeninėje erdvėje. Dėl to itin svarbi tampa DI technologijų reglamentacija, apimanti tiek kalbų prieinamumo užtikrinimą, tiek duomenų kokybės, etikos ir skaidrumo reikalavimus. Šiame kontekste ES politikos DI sektoriuje formuotojai technologinę pažangą siekia suderinti su kalbų įvairove ir lingvistinių teisių apsauga. Formuojant europinę DI politiką, dalyvauja aukščiausios ES institucijos. Jos įstatymuose įtvirtinusios bendras dirbtinio intelekto ir kalbų technologijų plėtros kryptis ir teisinius pagrindus, galiojančius visose ES narėse. Įvairiose programose ir iniciatyvose apibrėžtos kitos svarbios gairės, kurios turi užtikrinti, kad DI būtų kuriamas ir taikomas atsakingai, laikantis skaidrumo, saugumo bei žmogaus teisių apsaugos principų. Be to, Bendrija remia mokslinius tyrimus, skatina inovacijas, siekia, kad į DI integruojant kalbas būtų atsižvelgiama į kultūrinę ir kalbinę ES įvairovę.

Lietuvos teisinė bazė DI srityje iš esmės remiasi ES dokumentuose suformuluotomis nuostatomis. Būdama jos nare, šalis atsižvelgia į organizacijos mastu taikomus reglamentus, juos lokaliai pritaiko ir detalizuoja. DI kūrimo, diegimo ir naudojimo principai glaudžiai siejasi su valstybinės lietuvių kalbos išsaugojimu. Esminė sąlyga – jos plėtra skaitmeninėje terpėje, todėl ypač svarbu užtikrinti ne tik duomenų lietuvių kalba prieinamumą, bet ir konkurencingumą daugiakalbėse DI sistemose. Atsakingų institucijų ir subjektų koordinuota veikla yra būtina siekiant garantuoti nuoseklų DI plėtros ir lietuvių kalbos technologinio vystymosi suderinamumą. Ši sritis Lietuvoje plėtojama bendromis pastangomis. Procese kaip strateginius prioritetus nustatančios organizacijos dalyvauja ministerijos, taip pat kalbos politiką formuojanti Valstybinė lietuvių kalbos komisija, įvairios viešojo sektoriaus įstaigos ir universitetai, privačios įmonės. Šio lauko veikėjai pagal jiems priskirtas kompetencijų sritis rūpinasi ne tik technologine pažanga, bet ir lietuvių kalbos išsaugojimu bei pritaikymu DI sprendimams. Toks koordinuotas darbas – pagrindas, kad inovacijos būtų derinamos su lietuvių kalbos visaverčiu funkcionavimu šiuolaikinėse technologijose.

Dirbtinis intelektas dažnai susijęs su kalba – ja grindžiamas duomenų apdorojimas, sąveika su vartotoju, informacijos generavimas ir perteikimas. Todėl technologinis kalbos prieinamumas tiesiogiai veikia DI sistemų funkcionalumą ir rezultatų kokybę. Sparčiai vystantis DI modeliams, atsiranda įvairių kalbų gausių išteklių nepakankamo prieinamumo bei kalbos technologinės nelygybės problema. Nuo 2022 m. generatyvinio dirbtinio intelekto plėtra itin sparti, dėl to didžiosios technologijų bendrovės į kalbos modelių kūrimą investuoja milžiniškas sumas. Tačiau daugiausia dėmesio skiriama anglų, kinų ir kitoms didesnėms kalboms, todėl mažesnės, tarp jų ir lietuvių kalba, susiduria su tam tikrais konkurenciniais keblumais skaitmeninėje erdvėje. Lietuvoje siekiama stiprinti lietuvių kalbos palaikymą DI sistemose, investuojama į kalbos technologijų projektus, bet akcentuojamas reikalingų specialistų trūkumas ir vis dar per mažas dėmesys lietuvių kalbos duomenų skaitmeninimui. Šiame kontekste itin pabrėžiamas žmogaus, kaip intelektualios ir kūrybiškos būtybės, vaidmuo DI informacijos priežiūros ir vertinimo procese.

Analitinių apžvalgų archyvas: <https://lnb.lt/istekliai/kiti-istekliai/analitines-apzvalgos>

Cituojant arba naudojant šį šaltinį, nuoroda į Nacionalinę biblioteką yra būtina.
Šioje apžvalgoje pateikiamas turinys nėra oficiali Nacionalinės bibliotekos nuomonė.

LNB / IAS / R / 2026-05-28 / Psl. 12 iš 13
© Nacionalinė biblioteka, 2026 /  Слава Україні!

- Svarbiausių ES ir nacionalinių veiksnių dėl DI politikos formavimo, ypač kalbų integravimo į DI, apžvalga atskleidė, kad Bendrijos lygmeniu akcentuojama kalbinė įvairovė, daugiakalbystės skatinimas, vis dėlto daugiausia lėšų skiriama DI įrankiams didžiosiomis kalbomis, pavyzdžiui, anglų ir vokiečių. Nebeįtartinai, šiuo atveju svarus ir privataus kapitalo vaidmuo. Statistiniai duomenys taip pat patvirtina geresnes komerciškai patrauklių kalbų pozicijas, palankesnę ir geriau joms pritaikytą skaitmeninę terpę. Nacionaliniu mastu, natūralu, dėmesys koncentruojamas į valstybinės kalbos pritaikymą dirbtiniame intelekto. Apibendrinus pasakytina, kad tik nuoseklus ES ir nacionalinių iniciatyvų derinimas gali užtikrinti, kad kalbinė įvairovė ir lygiavertiškumas skaitmeninėje erdvėje būtų ne deklaratyvi nuostata, o realiai veikiantis principas.
- Lietuvių kalbos plėtra ir funkcionavimas DI turi ir palankiai vertinamų, ir tobulintinų aspektų. Pirmiausia apčiuopiamas progresas matomas stebint valstybės ir privačių institucijų skaitmeninės perspektyvos prioritetus. Čia įsitraukia aukščiausios šalies institucijos ir mokslo įstaigos, taip pat specializuojasi privataus sektoriaus subjektai. Nuo 2022 m., spartėjant generatyvinio DI plėtrai, lietuvių kalba nuosekliai integruojama į DI paslaugas ir sistemas. Tuo tikslu kuriami kalbiniai ištekliai ir modeliai, būtini užtikrinti sklandų DI veikimą. Vis dėlto dėl minėtųjų veiksnių, pavyzdžiui, mažo lietuvių kalbos konkurencingumo, trūksta didelių kokybiškų duomenų rinkinių. Be to, technologijos pritaikytos didesnėms kalboms, lietuvių kalbos struktūra itin sudėtinga, o mažesnė rinka lemia ir menkesnes investicijas į šią sritį. Taigi, situacija vertintina kaip dvilypė, nes valstybinės kalbos išsaugojimas siejamas su jos integracija į skaitmeninę terpę, ir atitinkami žingsniai jau daromi. Kita vertus, mažesni finansiniai resursai ir technologiniai apribojimai sąlygoja lėtesnę lietuvių kalbos pažangą DI aplinkoje.
- Laipsniška raida šioje srityje neįmanoma be aukšto profilio kompiuterinės lingvistikos specialistų, kurių, kaip pabrėžia procese dalyvaujantys mokslininkai, šiuo metu itin trūksta. Evoluicijuojant informacijos erdvei ir paslaugoms, sparčiau turi kisti ir studijų programos – itin pageidautini kvalifikuoti tarpdisciplininiai specialistai, įgiję tiek lingvistikos, tiek informatikos kompetencijų. Kaip matyti, 2026 m. universitetai jau siūlo kai kurias tokio pobūdžio studijų programas: Kalba ir dirbtinio intelekto valdymas, Daugiakalbystė ir skaitmeninės technologijos, Modernioji lingvistika. Kadangi DI kaita yra itin spėri, būtinas ne tik nuolatinis dėmesys studijų turiniui, bet ir atitinkama reakcija bei korekcijos, keičiantis rinkos poreikiams. Tik sisteminis inovacijų skatinimo ir švietimo politikos krypties formavimas padėtų Lietuvai sumažinti kalbinę atskirtį DI infrastruktūroje.