



**Vilniaus
universitetas**

Projektavimo mąstymo automatinio vertinimo sistemos modeliavimas

Informatikos inžinerijos krypties doktorantų konferencija
Veiklos ataskaita už 2026 m. kovo 31 d. - 2026 m. spalio 8 d.

Darbo vadovė - dr. Anita JUŠKEVIČIENĖ
Doktorantūros pradžios ir pabaigos metai: 2023-10-01 - 2027-09-30
Studijų metai: tretį

Studijų planas ir jo vykdymo suvestinė

Vilniaus
universitetas

Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2023/2024)	2	2
II (2024/2025)	2	2
III (2025/2026)		
IV (2026/2027)		
Iš viso:	4	4

Studijų planas ir jo vykdymo suvestinė

Vilniaus
universitetas

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse ²		Nacionalinėse ³		Su citav. rodikliu ⁴			Be citav. rodikliu ⁴		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta ⁶	Būklė ⁷	Planas	Įvykdyta ⁶	Būklė ⁷
I (2023/2024)	0		0		0			0		
II (2024/2025)	0	1		4				0		
III (2025/2026)	1	1			1	0	Pateikta			
IV (2026/2027)	1	1			1	0	Pateikta			
Iš viso:	2	3		4	2					

Mokslinių tyrimų ir disertacijos rengimo etapai

Darbo pavadinimas	Atlikimo terminai	Pastabos
2. Mokslinio tyrimo vykdymas:		
2.3. Empirinis tyrimas:		
2.3.1. Projektavimo mąstymo vertinimo sistemos ekspertinis vertinimas.	2026 m. II ketvirtis - 2026 m. III ketvirtis	Vykdoma: gauti 6 užbaigti klausimynai (9 įrašai); numatoma užbaigti 2026 m. IV ketv.
2.3.2. Projektavimo mąstymo vertinimo sistemos eksperimentinis aprobavimas.	2026 m. II ketvirtis - 2027 m. I ketvirtis	Vykdoma: parengtas 362 atsakymų rinkinys, atliktas techninis prototipo išbandymas.

Tyrimo objektas ir tikslas

Tyrimo objektas: projektavimo mąstymo idėjų generavimo etapo automatinio vertinimo sistemos modelis, integruojantis kūrybinio rezultato vertinimą ir proceso bei konteksto požymių analizę.

Darbo tikslas: sukurti ir empiriškai įvertinti projektavimo mąstymo idėjų generavimo etapo automatinio vertinimo sistemos modelį, integruojantį kūrybinio rezultato vertinimą su proceso ir konteksto požymių analize ir sudarantį galimybę pateikti paaiškinamą vertinimo rezultatą.

Uždaviniai

1. Išanalizuoti projektavimo mąstymo vertinimo teorinius požiūrius ir rodiklius, išskiriant idėjų generavimo etapo kūrybinio rezultato, proceso ir konteksto aspektus, reikšmingus automatiniam vertinimui.
2. Išanalizuoti mašininio mokymosi ir natūralios kalbos apdorojimo metodų taikymo galimybes automatiniam projektavimo mąstymo idėjų generavimo etapo vertinimui.
3. **Sukurti projektavimo mąstymo idėjų generavimo etapo automatinio vertinimo sistemos modelį, jungiantį kūrybinio rezultato vertinimą su proceso ir konteksto požymių analize, ir jį patikslinti pagal ekspertinio vertinimo rezultatus.**
4. Eksperimentiškai įvertinti sukurto automatinio vertinimo sistemos modelio veikimą, lyginant jo ir pasirinktų esamų automatinio vertinimo metodų rezultatus su žmogaus vertintojų įverčiais.
5. Remiantis ekspertinio vertinimo ir eksperimentinio tyrimo rezultatais, nustatyti sukurto modelio tobulinimo ir praktinio taikymo kryptis.

Sisteminės literatūros apžvalgos rezultatai

Vilniaus
universitetas

Atlikta sisteminė literatūros apžvalga pagal PRISMA 2020; į kokybinę sintezę įtraukti **79 empiriniai tyrimai**, publikuoti 2019-2025 m.

Nustatyta, kad PM **idėjų generavimo etapas** yra vienas tinkamiausių automatizuotam vertinimui, nes jo kūrybinį rezultatą galima apibrėžti aiškiomis dimensijomis, o jų vertinimui jau taikomi semantinės analizės ir DI metodai.

Idėjų generavimo rezultatui vertinti dažniausiai taikomos keturios divergentinio mąstymo dimensijos: **sklandumas, lankstumas, originalumas ir detalumas**.

Automatinio originalumo vertinimo metodai vystosi nuo semantinio nuotolio rodiklių iki prižiūrimu mokymusi patikslintų didžiųjų kalbos modelių. Naujesni tyrimai rodo, kad **LLM pagrįsti modeliai gali geriau atitikti žmogaus vertintojų įverčius nei vien semantinio nuotolio metodai**, todėl semantinis nuotolis laikytinas vienu originalumo požymiu, o ne viso originalumo konstrukto pakaitalu.

Hibridiniai žmogaus ir DI vertinimo modeliai yra perspektyvūs, tačiau jų empirinis taikymas edukaciniame PM kontekste dar ribotas.

Nustatyta tyrimo spraga ir modelio kūrimo prielaidos

Vilniaus
universitetas

Esami automatiniai metodai daugiausia vertina **galutinį idėjos rezultatą**, o informacija apie idėjos kūrimo procesą naudojama ribotai.

Semantinio nuotolio metodai naudingi kaip **originalumo požymis**, tačiau jų galimybės ribotos: jie neapėmė viso originalumo konstrukto, kitų kūrybinio rezultato dimensijų ir idėjos kūrimo proceso.

LLM pagrįsti metodai gerina originalumo vertinimo atitiktį žmogaus vertinimui, tačiau išlieka paaiškinamumo, šališkumo ir proceso duomenų integravimo klausimai.

Trūksta modelių, sistemiškai jungiančių **kūrybinio rezultato vertinimą su proceso ir konteksto požymių analize**.

Automatiniame vertinime svarbu išlaikyti galimybę žmogui peržiūrėti neaiškius ar nepatikimus atvejus.

Šie rezultatai pagrindė poreikį kurti **integruotą projektavimo mąstymo idėjų generavimo etapo automatinio vertinimo sistemos modelį**.

Kūrybiškumo dimensijos idėjų generavimo etape

Vilniaus
universitetas

Dimensija	Apibrėžimas	Pavyzdys	Technika
Sklandumas (Fluency)	Sugeneruotų tinkamų idėjų skaičius	Daug skirtingų sprendimų pasiūlymų per trumpą laiką	Idėjų vienetų skaičiavimas
Lankstumas (Flexibility)	Skirtingų idėjų kategorijų įvairovė	Idėjos apie funkcijas, procesus, vartotojus, technologijas	Semantinis klasterizavimas / kategorijų įvairovės analizė
Originalumas (Originality)	Idėjų naujumas ir retumas kitų atsakymų atžvilgiu	Netipiški ar nestandartiniai sprendimai	Semantinis atstumas (SemDis) + LLM / modelių grįstas vertinimas
Detalumas (Elaboration)	Idėjos išplėtojimo ir paaiškinimo lygis	Aiškiai aprašyti žingsniai, komponentai, pagrindimas	Išplėtojimo komponentų atpažinimas (NLP); teksto ilgis tik kontrolei

(Torrance, 1990; Runco & Jaeger, 2012; Ludwig et al., 2021).

Automatinio vertinimo sistemos modelio tyrimo kryptys

Sukurti ir empiriškai įvertinti PM idėjų generavimo etapo automatinio vertinimo sistemos modelį, jungiantį kūrybinio rezultato vertinimą su proceso ir konteksto požymių analize.

Metodai

Kiekybiniai: ekspertinio vertinimo duomenų analizė; automatinių ir žmogaus vertintojų įverčių palyginimas; statistinė analizė.

Kokybiniai: mokslinės literatūros ir dokumentų analizė; atvirų ekspertų komentarų teminė analizė.

Dalyviai

Bendrojo ugdymo mokiniai; tikslinės atrankos ekspertai, vertinantys modelį v1.0; žmogaus vertintojai eksperimentiniame etape.

Projektavimo mąstymo idėjų generavimo etapo automatinio vertinimo sistemos architektūrinis modelis

Vilniaus
universitetas

Problema: esami automatinio vertinimo metodai daugiausia orientuoti į galutinį idėjos rezultatą, o idėjos kūrimo proceso ir konteksto duomenys vertinime naudojami ribotai.

Tyrimo sprendimas: kuriamas automatinio vertinimo sistemos modelis, kuris:

- vertina besimokančiojo idėjų generavimo rezultatą pagal kūrybiškumo dimensijas;
- analizuoja idėjos kūrimo proceso ir konteksto požymius;
- integruoja abiejų vertinimo dalių rezultatus;
- įvertina rezultato patikimumą ir numato žmogaus peržiūrą neaiškiais atvejais;
- pateikia paaiškinamą vertinimo rezultatą.

Kūrybinio rezultato vertinimo rodikliai sistemoje

Vertinamas konstruktas - PM idėjų generavimo etapo kūrybinis rezultatas.

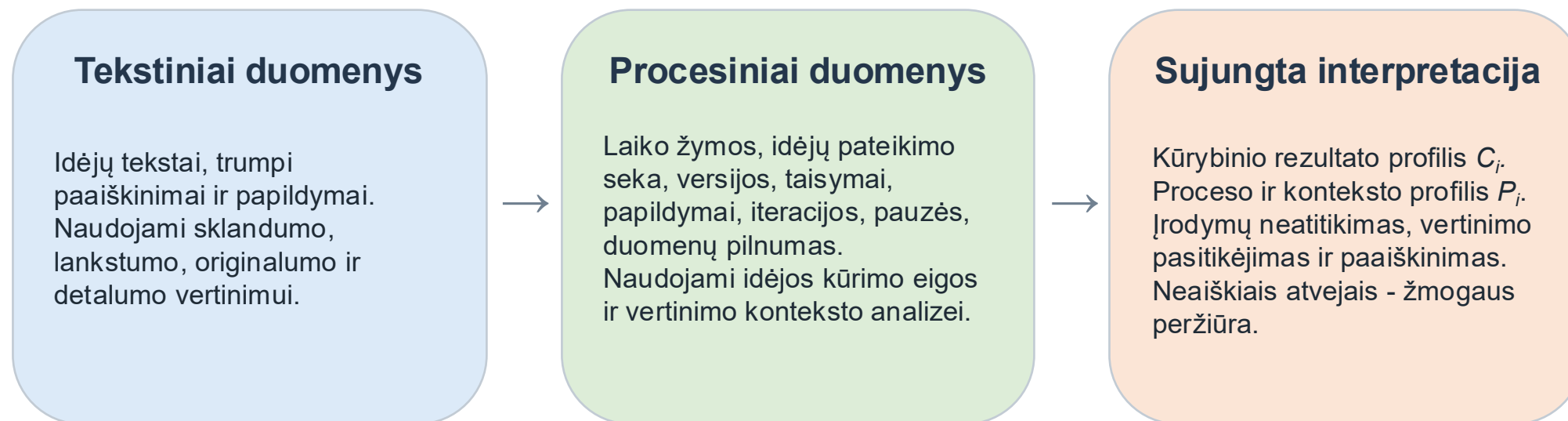
Vilniaus
universitetas

Rodiklis	Apibrėžimas	Žymėjimas / vertinimo būdas	Paaiškinimas
Sklandumas (Fluency)	Pateiktų tinkamų idėjų kiekis	S_{flu}	Siejamas su pateiktų, prasmingų ir užduotį atitinkančių idėjų skaičiumi; atskiriami pasikartojantys atsakymai.
Lankstumas (Flexibility)	Idėjų priklausymas skirtingoms semantinėms kategorijoms ar sprendimo kryptims	S_{flex}	Siejamas su skirtingų kategorijų skaičiumi, semantine įvairove ir perėjimu tarp sprendimo krypčių.
Originalumas (Originality)	Idėjos neįprastumas ir nutolimas nuo įprastų ar dažnai pasitaikančių sprendimų	S_{orig}	Siejamas su semantiniu nutolimu, retumu, modeliu grįstu originalumo įverčiu ir tinkamumu užduoties kontekstui; semantinis nuotolis yra vienas originalumo požymių, bet ne visas originalumas.
Detalumas (Elaboration)	Idėjos išplėtojimo ir paaiškinimo lygis	S_{elab}	Siejamas su veikimo principo, paskirties, naudotojo, situacijos, papildomų savybių ir sąlygų aprašymu.

$$C_i = (S_{flu}, S_{flex}, S_{orig}, S_{elab})$$

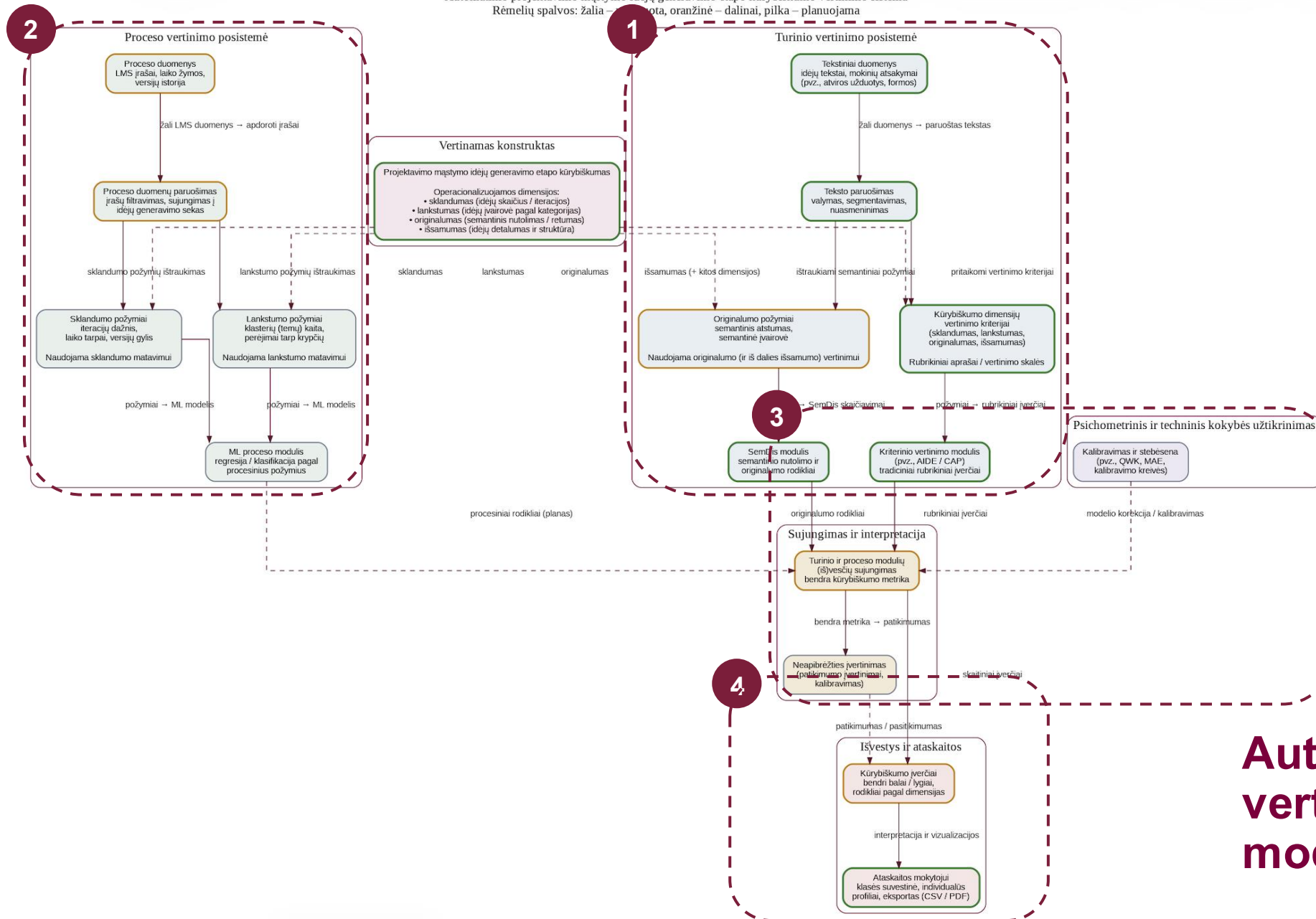
C_i - mokinio i kūrybinio rezultato profilis.

Teksto ilgis nelaikomas detalumo matu; naudojamas tik kaip kontrolinis kintamasis.

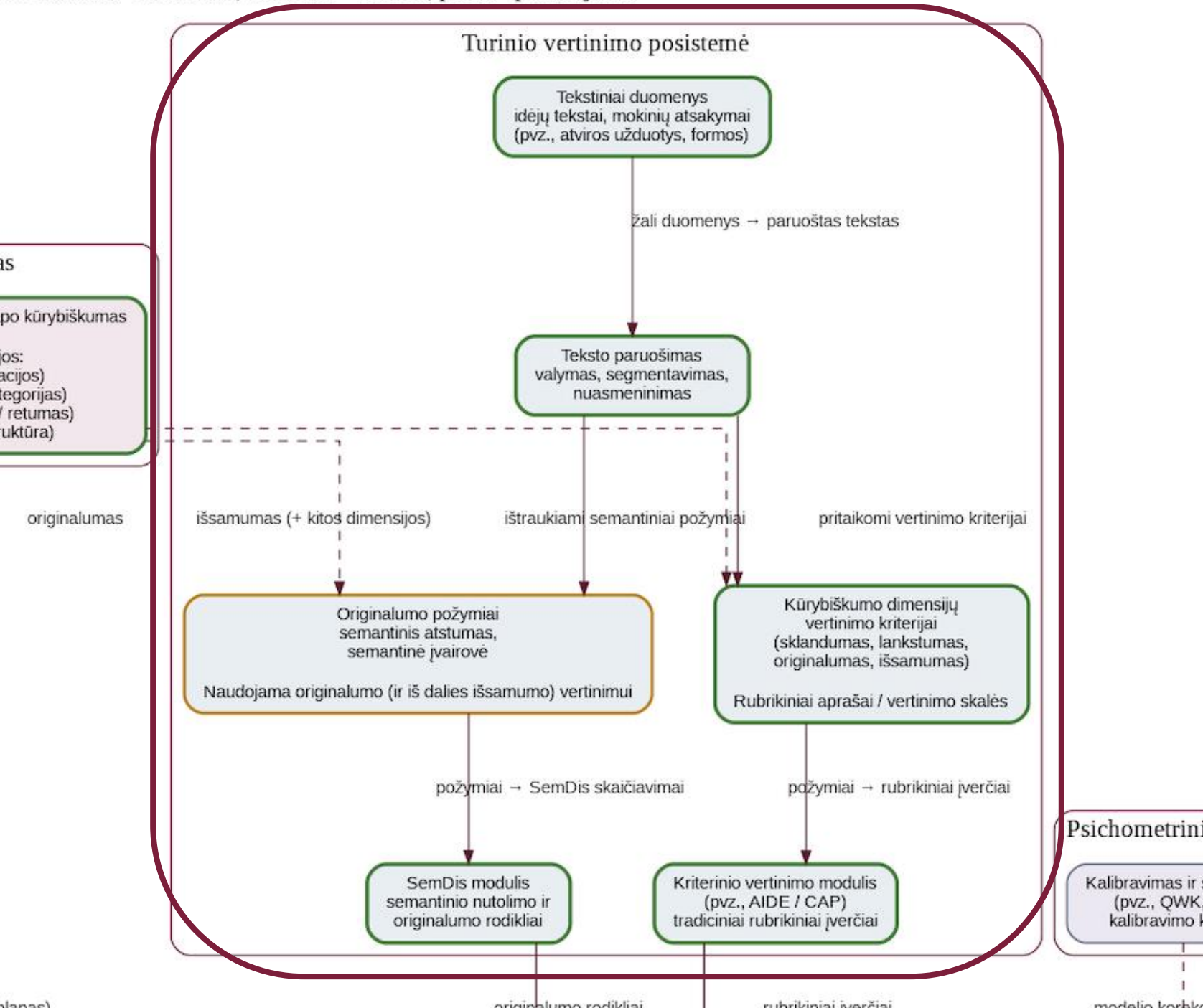


Principas: rezultato ir proceso duomenys nėra paprastai sumuojami - jie interpretuojami kaip skirtingi įrodymų šaltiniai, o neaiškūs atvejai gali būti perduodami žmogaus peržiūrai.

Automatinio projektavimo mąstymo idėjų generavimo etapo kūrybiškumo vertinimo sistema
Rėmelių spalvos: žalia – turinio, oranžinė – daliniai, pilka – planuojama



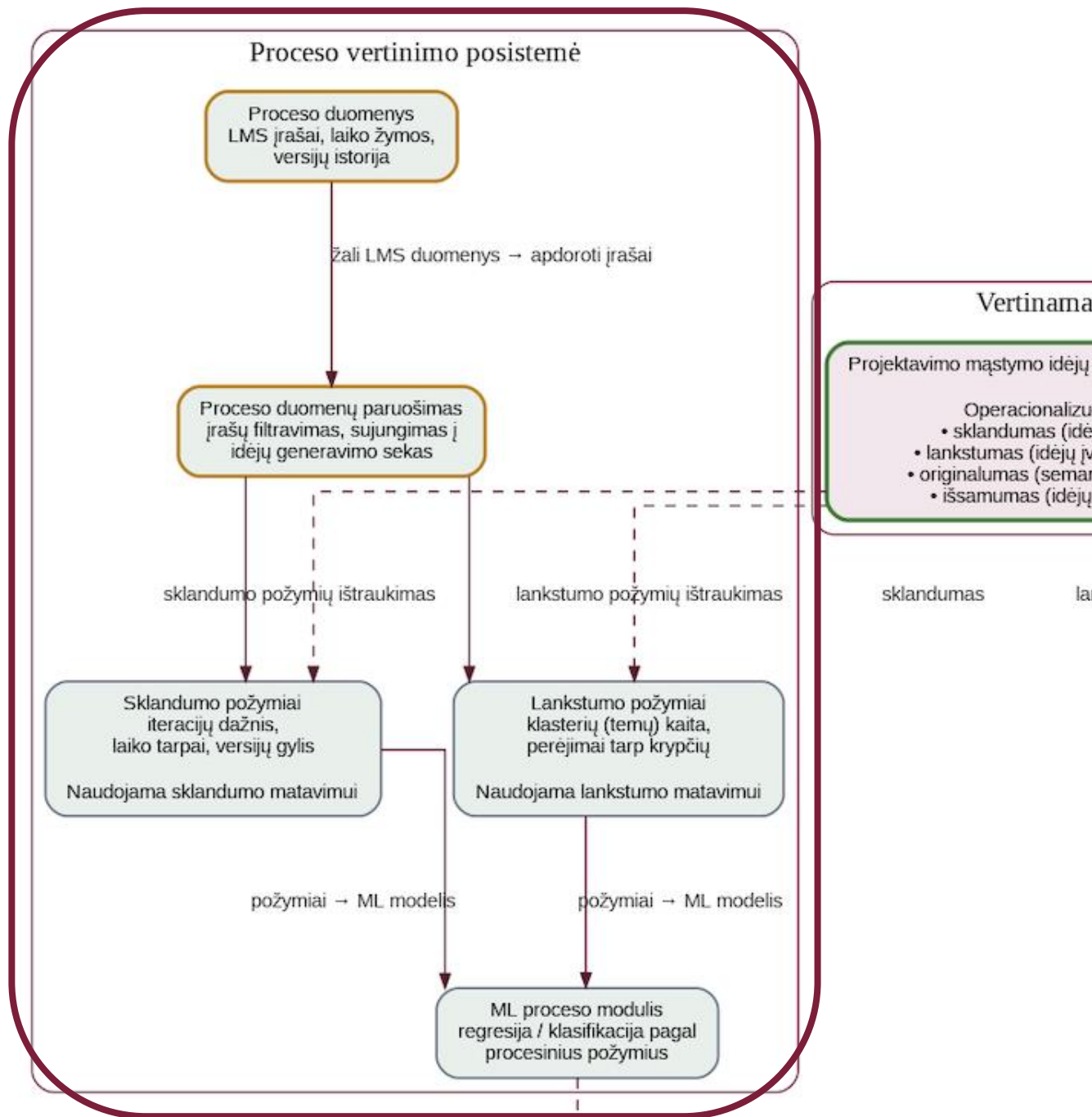
Automatinio
vertinimo sistemos
modelis v1.0



1 / 4 · Modelis v1.0

Turinio vertinimo posistemė

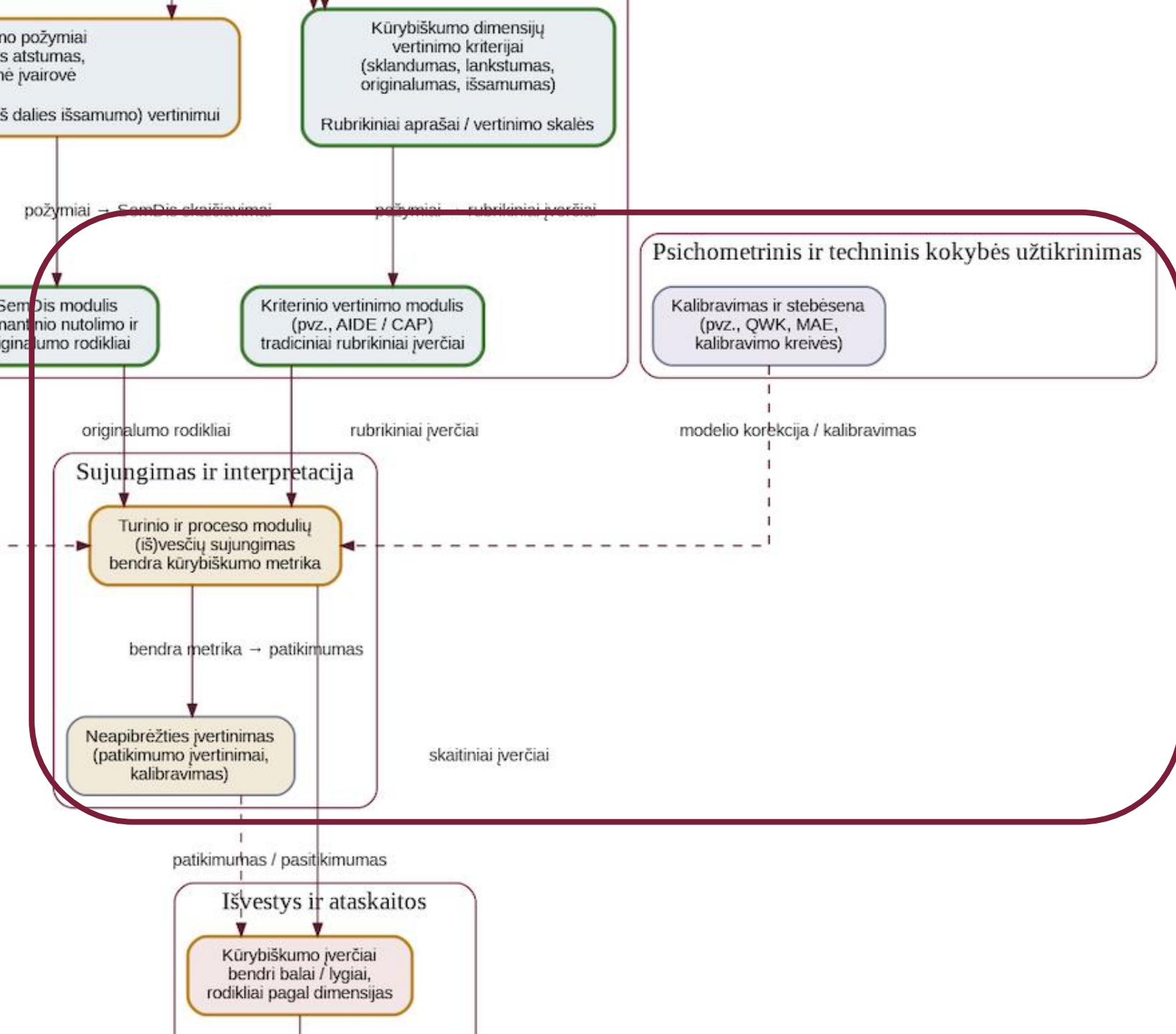
- v1.1 pavadinimas: kūrybinio rezultato vertinimo posistemė.
- Realizuota (žali rėmeliai): tekstų paruošimas, SemDis ir AIDE / CAP moduliai.
- Formuoja kūrybinio rezultato profilį C_i (sklandumas, lankstumas, originalumas, detalumas). Schemoje „išsamumas“ = detalumas.
- Ekspertų signalas (B3): semantinis nuotolis yra tik vienas originalumo požymis.
- v1.1: atskirti semantinį nuotolį, retumą, prasmingumą ir tinkamumą užduočiai; eksperimentiniame etape palyginti semantinio nuotolio ir LLM pagrįstus originalumo vertinimo metodus.



2 / 4 · Modelis v1.0

Proceso vertinimo posistemė

- v1.1 pavadinimas: proceso ir konteksto požymių analizės posistemė.
- Dalinai realizuota (oranžiniai rėmeliai); ML proceso modulis planuojamas.
- v1.0 schemeje proceso požymiai naudoti sklandumui ir lankstumui matuoti.
- v1.1 pakeitimas: proceso požymiai nėra kūrybiškumo balai; jie formuoja profilį P_i ir vertinimo pasitikėjimą.
- Ekspertai: 5 iš 6 siūlo ilgesnes pauzes tikslinti arba neįtraukti.



3 / 4 · Modelis v1.0

Sujungimas ir interpretacija

- v1.0: turinio ir proceso išvestys sujungiamos į bendrą kūrybiškumo metriką.
- v1.1 pakeitimas: rezultatas ir procesas nesumuojami; tikrinama neatitiktis $d_i = D(C_i, P_i)$ ir pasitikėjimas q_i .
- Kai $d_i > \tau$ arba $q_i < \gamma$, atvejis perduodamas žmogaus peržiūrai.
- D forma, τ ir γ kalibruojami eksperimentinio aprobavimo metu.
- Kalibravimui ir stebėsenai numatyti QWK, MAE ir kiti rodikliai.

Neapibrėžties įvertinimas
(patikimumo įvertinimai,
kalibravimas)

skaitiniai įverčiai

patikimumas / pasitikimumas

Išvestys ir ataskaitos

Kūrybiškumo įverčiai
bendri balai / lygiai,
rodikliai pagal dimensijas

interpretacija ir vizualizacijos

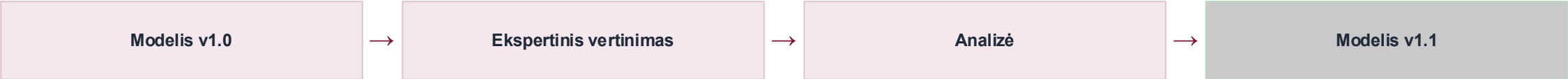
Ataskaitos mokytojui
klasės suvestinė, individualūs
profiliai, eksportas (CSV / PDF)

4 / 4 · Modelis v1.0

Išvestys ir ataskaitos

- Kūrybiškumo įverčiai pagal dimensijas, mokinio ir klasės lygmeniu.
- Ataskaitos mokytojui: suvestinės, individualūs profiliai, eksportas (CSV / PDF).
- Realizuota prototipo lygiu (žr. prototipo skaidres).
- v1.1: išvestyje pateikiamas ir pasitikėjimo lygis bei žmogaus peržiūros žymė.

Modelio v1.0 ekspertinis vertinimas



EKSPERTINIO VERTINIMO METODIKA

Elementas	Taikymas šiame tyrime
Metodas	Vieno etapo struktūruotas ekspertinis vertinimas
Atranka	Tikslinė ekspertų atranka pagal kompetencijos kriterijus
Vertinimo objektas	Automatinio PM idėjų generavimo etapo vertinimo modelis v1.0
Priemonės	M1 modelio aprašas + M2 klausimynas

VERTINIMO DUOMENYS IR ANALIZĖ

Duomenų tipas	Analizė
1-5 skalė + NK / NI	Md, IQR, 4-5 %, NK / NI dažniai; Aiken V galutinėje analizėje
Sprendimų matrica	Palikti / patikslinti / neįtraukti / NK dažniai
Atviri komentarai	Teminė analizė ir v1.0 → v1.1 pakeitimų registras

DUOMENŲ BŪSENA

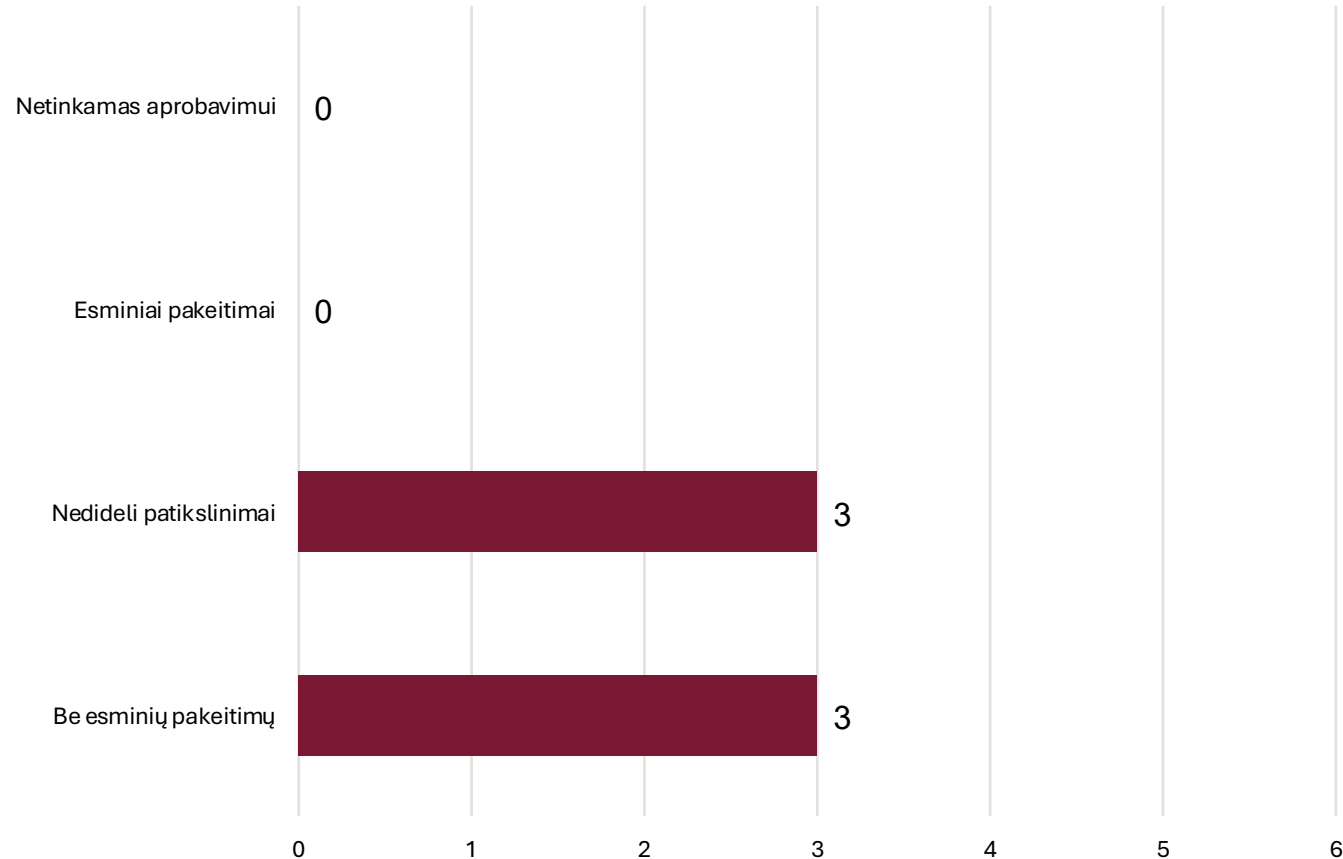
Visi įrašai	Užbaigti	Neužbaigti	Tušti
9	6	2	1

Preliminari kiekybinė analizė skaičiuojama tik iš 6 užbaigtų ekspertų klausimynų. Neužbaigti, bet turinį turintys įrašai neįtraukiami į šias skaidres.

Preliminarus bendras modelio v1.0 įvertinimas

Bendras pasirengimas eksperimentiniam aprobavimui, n = 6

Vilniaus
universitetas



3

be esminių
pakeitimų

3

su nedideliais
patikslinimais

0

esminių pakeitimų

INTERPRETACIJA

Visi 6 užbaigti ekspertiniai vertinimai rodo, kad modelis tinkamas pereiti į eksperimentinį aprobavimą.

Pusė ekspertų siūlo tik nedidelius patikslinimus, todėl kitas žingsnis - ne modelio atmetimas, o v1.0 patikslinimas į v1.1.

Preliminari išvada galutinama tik uždarius duomenų rinkimą.

Teiginių palaikymas pagal modelio dalis

Preliminarus 1-5 skalės teiginių vertinimas pagal iš anksto nustatytą kriterijų

28/29

teiginių atitinka palaikymo kriterijų

Kriterijus: $Md \geq 4$, $IQR \leq 1$ ir bent 75 % galiojančių įverčių yra 4 arba 5.

Dalis	Teiginiai	Atitinka kriterijų
A. Bendras pagrįstumas	5	5 / 5
B. Kūrybinis rezultatas	8	7 / 8
C. Procesas ir kontekstas	5	5 / 5
D. Susiejimas ir neapibrėžtumas	6	6 / 6
E. Paaiškinamumas / patikrinamumas	5	5 / 5
Iš viso	29	28 / 29

VIENTELIS IŠSISKIRIANTIS TEIGINYS

Teiginys	Md	IQR	4-5
B3. Semantinis nejprastumas kaip originalumo požymis	5	1,5	66,7 %

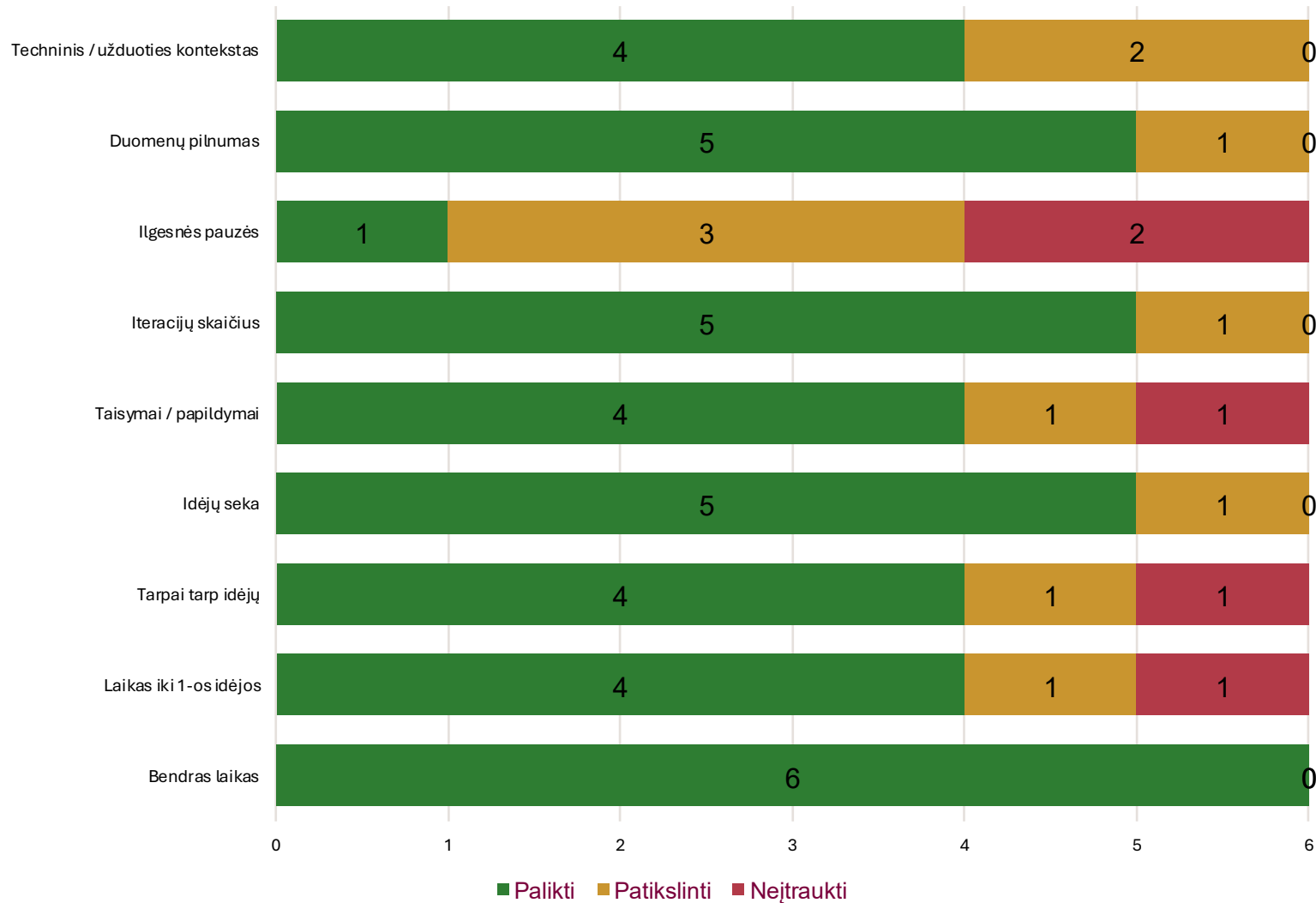
Interpretacija: ekspertai neatmeta originalumo dimensijos. Signalas rodo, kad semantinio nejprastumo vaidmenį reikia aiškiau apibrėžti ir atskirti nuo viso originalumo konstrukto.

Galutinėje analizėje B3 bus interpretuojamas kartu su komentarais apie originalumą, semantinį nuotolį ir užduoties kontekstą.

Proceso ir konteksto požymių sprendimų matrica

Ekspertų sprendimai dėl proceso požymių įtraukimo į modelį, n = 6

Vilniaus universitetas



PAGRINDINIAI SIGNALAI

5/6

ilgesnės pauzės siūlo
tikslinti arba neįtraukti

2/6

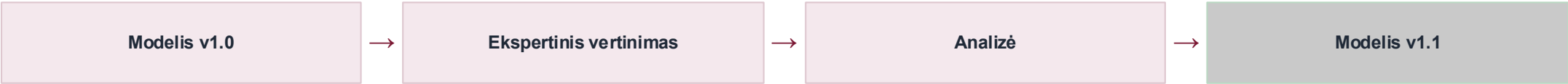
siūlo tikslinti techninius
/ konteksto požymius

Ilgesnės pauzės nėra stabilus kūrybiškumo požymis: jų reikšmė gali priklausyti nuo užduoties, darbo tempo, techninių trikdžių ir refleksijos pobūdžio.

v1.1 kryptis: aiškiau apibrėžti, kada proceso požymis informuoja vertinimą, o kada tik sumažina pasitikėjimą automatine interpretacija.

Preliminarios modelio v1.1 tobulinimo kryptys

Ekspertų signalai verčiami į modelio patikslinimo sprendimus



Ekspertų identifikuota problema	Preliminari interpretacija	v1.1 veiksmas
Ilgesnių pauzių interpretacija	Ryškiausias proceso požymių tikslinimo signalas	Tikslinti arba jungti su kitais laiko požymiais
D(Ci, Pi) vertinimo būdas	Koncepcija palaikoma, bet funkcijos forma dar nekalibruota	Apibrėžti eksperimentiniame etape ir susieti su pasitikėjimo rodikliu qi
Semantinis neįprastumas	Semantinis nuotolis nėra visas originalumo konstruktas	Atskirti semantinį signalą, retumą, prasmingumą ir užduoties tinkamumą ir eksperimentiškai palyginti semantinio nuotolio bei LLM pagrįstus vertinimo metodus.
Kalbiniai ir rašymo gebėjimai	Galimas šališkumas vertinant tekstinį atsakymą, teksto ilgį naudoti tik kaip kontrolinį kintamąjį	Numatyti interpretavimo ribas ir šališkumo kontrolę
Žmogaus peržiūros principas	Reikalingas neaiškiems ar mažo pasitikėjimo atvejams	Išlaikyti žmogaus peržiūros mechanizmą modelyje

Ekspertinio vertinimo vaidmuo modelyje

Konceptualus modelio patikrinimas ir pagrįstas patikslinimas

Vilniaus
universitetas

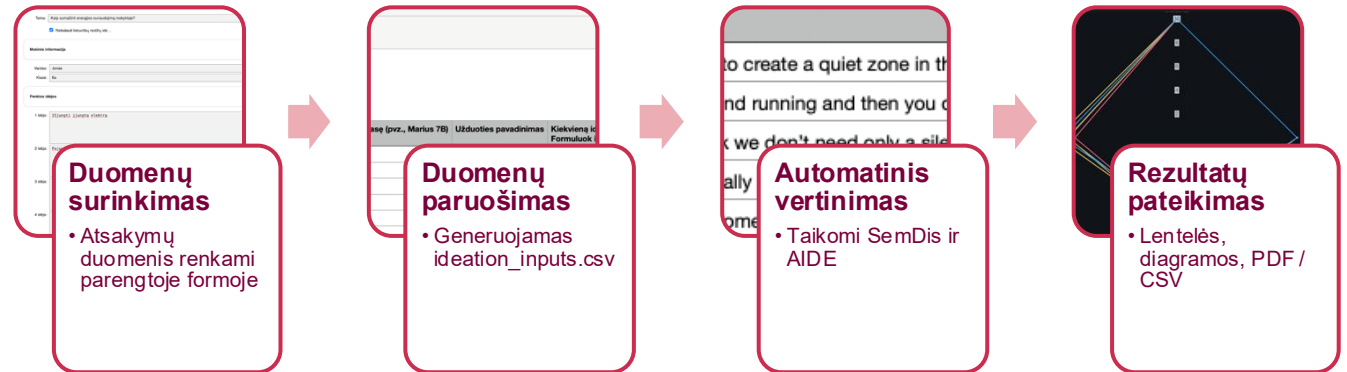
Disertacijos dalis	Ką pateikti
Metodika	Ekspertų atranka, M1/M2 priemonės, 1-5 + NK/NI, sprendimų matricos, analizės taisyklės
Rezultatai	Md, IQR, 4-5 %, NK/NI dažniai; proceso požymių matricos; komentarų temos
Modelio patikslinimas	v1.0 → v1.1 pakeitimų registras: kas keista, kodėl ir kuo remiantis
Ribotumai	Ekspertinis vertinimas neįrodo automatinio vertinimo tikslumo; tai tikrinama eksperimentinio aprobavimo etape

Dabartinė prototipo realizacija

Vilniaus
universitetas

Įgyvendinta

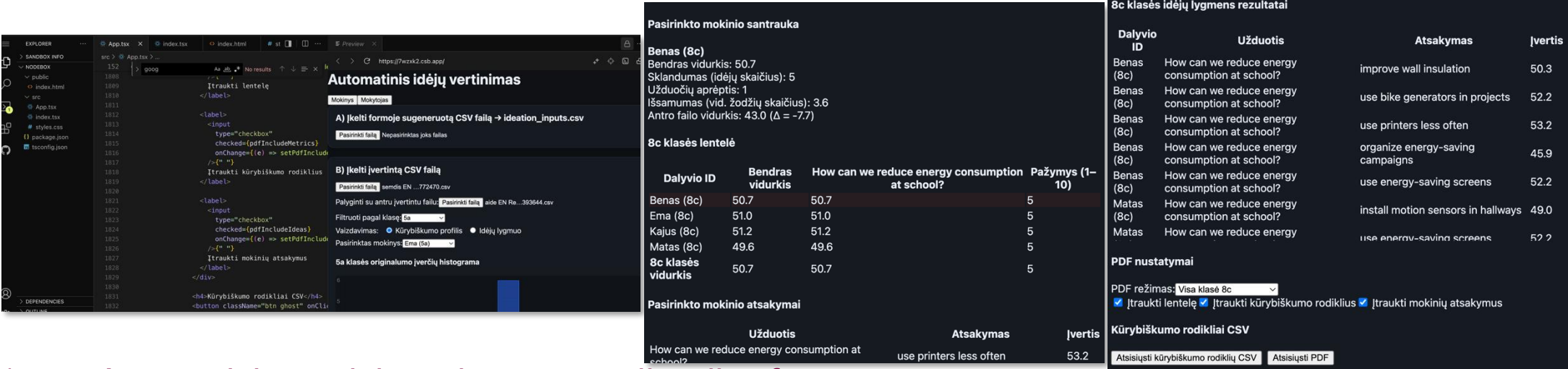
- Automatinis tekstinių duomenų paruošimas.
- SemDis ir AIDE vertinimo modulių taikymas / integravimas.
- Idėjų ir mokinio lygmens įverčių formavimas.
- Rezultatų lentelių ir vizualizacijų generavimas.
- Preliminarus automatinių vertinimo metodų palyginimas.



Dalimai įgyvendinta / toliau realizuojama

- Proceso ir konteksto duomenų surinkimas ir analizė.
- Kūrybinio rezultato ir proceso profilių integravimo logika.
- Pasitikėjimo ir žmogaus peržiūros mechanizmas.
- Alternatyvių originalumo vertinimo metodų, įskaitant LLM pagrįstą vertinimą, palyginimas.

Dabartinė modelio realizacija



1 pav. Automatinio vertinimo sistemos realizacijos fragmentas



2 pav. Mokinio ir klasės vidurkio palyginimas

Dabartinė modelio realizacijos ataskaita

Kkrybiškumo rezultats ataskaita

Rodymo sritis: Tik klas 8c

Eksporto režimas: Visa rodoma grup

Dalyvio ID	Bendras vidurkis	How can we reduce energy consumption at school?
Benas (8c)	50.7	50.7
Ema (8c)	51.0	51.0
Kajus (8c)	51.2	51.2
Matas (8c)	49.6	49.6
8 c klas s vidurkis	50.7	50.7

Kkrybiškumo rodikliai

Dalyvio ID	Sklandumas	Užduo is apr ptis	Originalumas	Išsamumas
Benas (8c)	5	1	50.7	3.6
Ema (8c)	5	1	51.0	5.2
Kajus (8c)	5	1	51.2	5.2
Matas (8c)	5	1	49.6	4.4

Mokinio atsakymai: Benas (8c)

Užduotis	Atsakymas	.vertis
How can we reduce energy consumption at school?	use printers less often	53.2
How can we reduce energy consumption at school?	use bike generators in projects	52.2
How can we reduce energy consumption at school?	use energy-saving screens	52.2
How can we reduce energy consumption at school?	improve wall insulation	50.3
How can we reduce energy consumption at school?	organize energy-saving campaigns	45.9

Mokinio atsakymai: Ema (8c)

Užduotis	Atsakymas	.vertis
How can we reduce energy consumption at school?	turn off lights in empty classrooms	58.6
How can we reduce energy consumption at school?	use thick curtains in winter	52.9
How can we reduce energy consumption at school?	use timers for devices	48.7
How can we reduce energy consumption at school?	shut down computers after lessons	47.5
How can we reduce energy consumption at school?	replace old equipment with efficient devices	47.4

3 pav. Ataskaitos fragmentas

Pirminiai rezultatai

Vertinamas aspektas	Gautas rezultatas	Būsena
Duomenų paruošimas	Mokinių atsakymai automatiškai parengiami vertinimui	Veikia prototipo lygiu
Kūrybinio rezultato vertinimas	Gaunami SemDis ir AIDE pagrindu formuojami įverčiai	Preliminariai išbandyta
Rezultatų agregavimas	Formuojamos idėjos, mokinio ir klasės lygmens suvestinės	Veikia prototipo lygiu
Rezultatų vizualizavimas	Generuojami kūrybiškumo profiliai, histogramos ir lentelės	Veikia prototipo lygiu
Proceso ir konteksto analizė	Realizuota iš dalies	Plėtojama
Integruotas modelio vertinimas	Kūrybinio rezultato ir proceso duomenų susiejimo logika dar tikslinama	Plėtojama
Eksperimentinis tikslumo vertinimas	Dar neatliktas	Numatyta aprobavimo etape

Preliminariam prototipo išbandymui panaudotas 362 mokinių atsakymų duomenų rinkinys. Šiame etape vertintas techninis prototipo veikimas ir duomenų apdorojimo bei vertinimo grandinės realizavimas; modelio vertinimo tikslumas bus nustatomas eksperimentinio aprobavimo etape.

- Sisteminės literatūros analizės pagrindu **patikslinta tyrimo spraga** ir pagrįstas poreikis integruoti kūrybinio rezultato vertinimą su proceso ir konteksto požymių analize.
- Sudarytas PM idėjų generavimo etapo automatinio vertinimo sistemos modelis v1.0, apimantis kūrybinio rezultato vertinimo posistemę ir proceso bei konteksto požymių analizės posistemę.
- **Apibrėžtas kūrybinio rezultato profilis**: sklandumas, lankstumas, originalumas ir detalumas; nustatyti pagrindiniai proceso ir konteksto požymiai.
- **Sukurtas pirminis sistemos prototipas**; preliminariam techniniam išbandymui panaudotas **362 mokinių atsakymų duomenų rinkinys**.
- **Originalumo vertinime** semantinis nuotolis traktuojamas kaip vienas iš požymių; **numatomas alternatyvių modelių grįstų metodų palyginimas**.
- **Parengta modelio v1.0 ekspertinio vertinimo metodika ir instrumentai, pradėtas struktūruotas ekspertinis vertinimas**. Gauti 6 užbaigti ekspertų vertinimai; 28 iš 29 teiginių atitinka palaikymo kriterijų.
- Parengti ir pateikti moksliniai darbai: ICERI2026 straipsnis priimtas, kitas straipsnis pateiktas žurnalui.

2026-10-08 - 2027-03-31 darbo planas

Vilniaus
universitetas

- Modelio v1.0 ekspertinio vertinimo užbaigimas ir v1.1 parengimas.
- Eksperimentinio aprobavimo metodikos užbaigimas.
- Pusiau automatizuoto prototipo plėtra.
- Automatinio ir žmogaus vertinimo palyginimas, įskaitant semantinio nuotolio ir LLM pagrįstų originalumo metodų analizę, bei proceso ir konteksto požymių vertinimas.
- Publikacijų rengimo ir recenzavimo procedūrų tęstinumas.

Literatūros sąrašas

1. **Beaty, R. E., & Johnson, D. R. (2021).** Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, 53(2), 757-780. <https://doi.org/10.3758/s13428-020-01453-w>
2. **Beaty, R. E., Zeitle, D. C., Baker, B. S., & Kenett, Y. N. (2021).** Forward flow and creative thought: Assessing associative cognition and its role in divergent thinking. *Thinking Skills and Creativity*, 41, 100859. <https://doi.org/10.1016/j.tsc.2021.100859>
3. **Guilford, J. P. (1950).** Creativity. *American Psychologist*, 5(9), 444-454. <https://doi.org/10.1037/h0063487>
4. **Ludwig, S., Mayer, C., Hansen, C., Eilers, K., & Brandt, S. A. (2021).** Automated essay scoring using transformer models. *Psych*, 3(4), 897-915. <https://doi.org/10.3390/psych3040056>
5. **Mesbah, S., Arous, I., Yang, J., & Bozzon, A. (2023).** HybridEval: A human-AI collaborative approach for evaluating design ideas at scale. In *Proceedings of the ACM Web Conference 2023* (pp. 3837-3848). <https://doi.org/10.1145/3543507.3583496>
6. **Organisciak, P., Acar, S., Dumas, D., & Berthiaume, K. (2023).** Beyond semantic distance: Automated scoring of divergent thinking greatly improves with large language models. *Thinking Skills and Creativity*, 49, 101356. <https://doi.org/10.1016/j.tsc.2023.101356>
7. **Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021).** The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
8. **Patterson, J. D., Pronchick, J., Panchanadikar, R., Fuge, M., van Hell, J. G., Miller, S. R., Johnson, D. R., & Beaty, R. E. (2025).** CAP: The Creativity Assessment Platform for Online Testing and Automated Scoring. *Behavior Research Methods*, 57(9), 264. <https://doi.org/10.3758/s13428-025-02761-9>
9. **Runco, M. A., & Jaeger, G. J. (2012).** The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96. <https://doi.org/10.1080/10400419.2012.650092>
10. **Torrance, E. P. (1990).** *Torrance Tests of Creative Thinking: Norms-technical manual*. Scholastic Testing Service.



**Vilniaus
universitetas**

Ačiū už dėmesį