



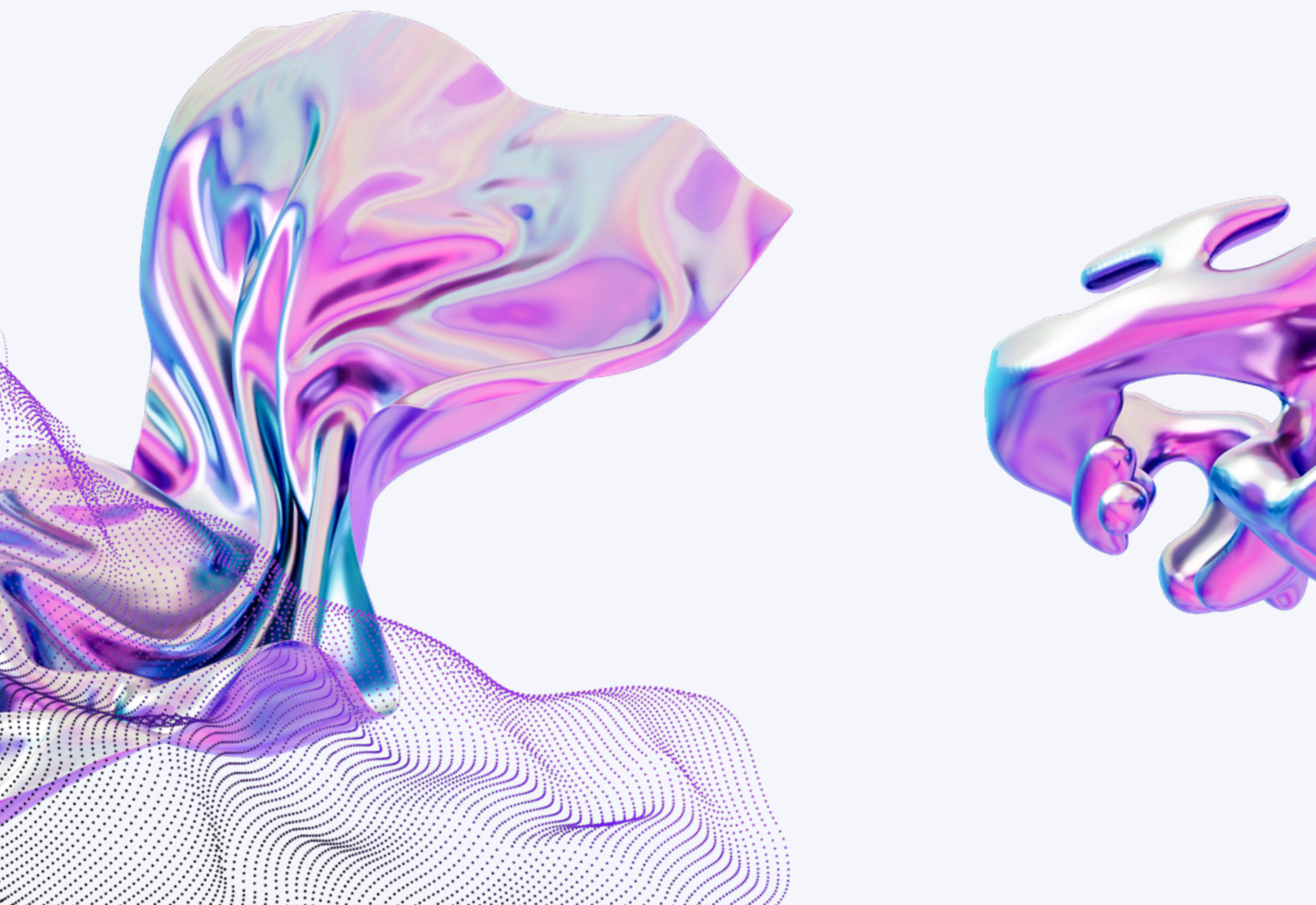
LIETUVOS
DIRBTINIO INTELEKTO
ASOCIACIJA



Medicinos
fakultetas

Konferencijos „Dirbtinio intelekto technologijos medicinoje” tezės

2026





LIETUVOS
DIRBTINIO INTELEKTO
ASOCIACIJA



Medicinos
fakultetas

Konferencijos „Dirbtinio intelekto technologijos medicinoje” tezės

2026 spalio 20 d.

Programinis komitetas:

Doc. dr. Linas Petkevičius

Prof. dr. Arvydas Laurinavičius

Prof. habil. dr. Gintautas Dzemyda

Įžanga

Konferencija „Dirbtinio intelekto technologijos medicinoje“ skirta pristatyti mokslinių darbų tyrimus informatikos ir medicinos srityse taikant dirbtinio intelekto technologijas. Šio renginio tikslas – supažindinti su kitų mokslo ir verslo įstaigų atliekamais tyrimais, naujausiais taikomais metodais ir paskatinti domėtis naujausiais proveržiais dirbtinio intelekto technologijose. Konferencijoje savo pranešimus skaitė tyrėjai iš VUL Santaros Klinikų, Lietuvos sveikatos mokslų universiteto, VilniusTech ir Vilniaus universiteto.

Konferenciją organizuoja Lietuvos mokslų akademija (LMA) kartu su Lietuvos dirbtinio intelekto asociacija (LDIA), bei Vilniaus universiteto, Medicinos fakulteto, Medicinos Mokslų centru. LMA – tai įstaiga, jungianti žymiausius Lietuvos ir savo veikla su Lietuva susijusius mokslininkus. Ji yra nepriklausoma Lietuvos Respublikos Seimo, Vyriausybės ir jai pavaldžių institucijų ekspertė ir patarėja mokslo bei studijų, kultūros, socialinės raidos, ūkio, gamtosaugos, sveikatos apsaugos, technologijų bei kitais klausimais. LDIA - tai verslo įmonės, startuolius, investuotojus, mokslo įstaigas ir profesionalius ekspertus vienijanti asociacija. LDIA rengia viešus renginius, suburia Lietuvos DI ekosistemą į vieną vietą, taip stiprinanti šalies DI potencialą.

Konferencijos tezės - parengtos mokslinių pristatymų santraukos. Tikime, kad dalyvavimas šioje LMA kuruojamoje konferencijoje paskatins tyrėjus ir kitus dirbančius medicinoje toliau tęsti inovatyvią mokslinę veiklą ir dar labiau taikyti naujus DI sprendimus.

Organizatoriai:

Doc. dr. Linas Petkevičius

Prof. habil. dr. Gintautas Dzemyda

Prof. dr. Arvydas Laurinavičius

Turinys

Dirbtinio intelekto taikymo medicininėje diagnostikoje teisiniai iššūkiai ir praktinio įgyvendinimo modelis Lietuvoje	5
Interpretuojamas kalbos modelio asistentas struktūrinių chromosominių persitvarkymų trūkio taškų įrodymams vertinti retųjų ligų diagnostikoje	6
Cancer Case Extraction from Full Electronic Health Records Using Open-Source Small Language Models	8
„Vibe coding“ patologijoje: reali sistema, iššūkiai ir rezultatai	9
From Questionnaires to Computable Evidence: Preparing Biobank Survey Data for Health AI	10
Dviejų apdorojimo strategijų palyginimas struktūrizuojant informaciją iš patologinių aprašų	11
Medicininį mokymo duomenų vertimo į lietuvių kalbą strategijos	12
Triukšmo mažinimas pilno lauko optinės koherentinės tomografijos trimačiuose vaizduose	13
Dirbtinis intelektas medicinos tyrimuose – bioetinio vertinimo patirtys (2020–2026 m.)	14
Nuo klinikinių duomenų iki dirbtinio intelekto: RAPIDS projektas retųjų ligų identifikavimui	15
Diagnostic Reliability of Multi-Label Chest X-Ray AI outside the Training Hospital: Accuracy, Calibration and Fairness across Seven Architectures	16
DI vėžio molekulinį mechanizmų visumos struktūrizavime	17
Sveikimas realiuoju laiku: dirbtinis intelektas ir dėvimieji prietaisai ankstyvai pooperacinių komplikacijų diagnostikai	18
Dirbtinio intelekto taikymas odontologijoje: vyresnio amžiaus pacientų osifikuojančio miozito kalcifikatų detekcija panoraminėse rentgenogramose	19
Dirbtinio intelekto panaudojimas H. pylori infekcijos endoskopinėje diagnostikoje . . .	20
Viso pjūvio morfometrijos ir Banff vertinimo sistemos palyginimas prognozuojant inksto transplantato praradimą	21
Neurologinės baigties po širdies sustojimo prognozavimas taikant kiekybinę EEG analizę ir mašininio mokymosi metodus	22
„Shortcut learning“ melanomos klasifikavimo užduotyje taikant dirbtinį intelektą . . .	23
Dirbtinio intelekto metodų taikymas žmogaus fiziologinių signalų analizei	24
Klasifikavimo modeliai medicinai	25

Dirbtinio intelekto taikymo medicininėje diagnostikoje teisiniai iššūkiai ir praktinio įgyvendinimo modelis Lietuvoje

doc. dr. Neringa Gaubienė, Miglė Bernatonytė

Vilniaus universitetas, Teisės fakultetas, Vilnius, Lietuva

Dirbtinio intelekto (DI) sistemų integravimas į medicininę diagnostiką atveria reikšmingas galimybes didinti diagnostikos tikslumą, efektyvumą ir sveikatos priežiūros paslaugų prieinamumą, tačiau kartu sukuria kompleksines teisines ir organizacines rizikas. Europos Sąjungos Dirbtinio intelekto akto, medicinos priemonių teisinio reguliavimo ir asmens duomenų apsaugos reikalavimų sąveika lemia, kad DI sistemų diegimas medicininėje diagnostikoje negali būti vertinamas vien kaip technologinis sprendimas – jo teisėtumas ir saugus naudojimas priklauso nuo nuoseklaus rizikų valdymo, žmogaus priežiūros, duomenų valdymo, dokumentavimo ir atsakomybės paskirstymo mechanizmų.

Pranešime pristatomi 2026 m. vasarą studentų mokslinės praktikos metu Miglės Bernatonytės ir projekto vadovės dr. Neringos Gaubienės atlikto tyrimo rezultatai. Tyrimo tikslas – **identifikuoti pagrindinius teisinius ir reguliacinius reikalavimus, taikomus medicininėje diagnostikoje naudojamoms didelės rizikos DI sistemoms, įvertinti jų praktinio įgyvendinimo sveikatos priežiūros įstaigose probleminius aspektus ir suformuoti Lietuvos kontekstui pritaiktą DI sistemų diegimo, naudojimo ir stebėsenos modelį.**

Tyrime taikyti sisteminis ir lingvistinis bei teleologinis teisės aiškinimo metodai, analizuojant Europos Sąjungos Dirbtinio intelekto akto, medicinos priemonių (MDR), in vitro diagnostikos medicinos priemonių (IVDR) ir Bendrojo duomenų apsaugos reglamento (BDAR) nuostatų tarpusavio sąveiką. Teisiniams, techniniams, organizaciniais ir klinikiniais reikalavimams identifikuoti bei susisteminti taikyta teminė kokybinė turinio analizė. Parengto praktinio modelio ir rekomendacijų pagrįstumas bei pritaikomumas įvertinti ekspertinio vertinimo būdu.

Tyrimas parodė, kad vien formalus atitikties atskiriems reguliavimo reikalavimams užtikrinimas nėra pakankamas saugiam DI naudojimui medicininėje diagnostikoje. Sveikatos priežiūros įstaigoms būtinas integruotas valdymo modelis, apimantis visą DI sistemos gyvavimo ciklą – nuo sistemos teisinio statuso ir rizikos kategorijos nustatymo bei tiekėjo ir atitikties dokumentų patikros iki institucinio pasirengimo, sistemos parengimo klinikiniam naudojimui, saugaus naudojimo ir tęstinės veikimo stebėsenos. Tyrimo pagrindu suformuotas **šešių etapų DI sistemos gyvavimo ciklo modelis**, kuriame sistemiškai integruojami rizikos valdymo, žmogaus priežiūros, duomenų apsaugos, atsakomybių paskirstymo, dokumentavimo, veikimo stebėsenos ir incidentų valdymo reikalavimai.

Siūlomas modelis sudaro prielaidas ES teisinius reikalavimus transformuoti į konkrečius organizacinius ir procedūrinius veiksmus ir gali būti naudojamas kaip metodologinis pagrindas atsakingam DI sistemų diegimui bei naudojimui Lietuvos sveikatos priežiūros įstaigose.

Raktažodžiai: dirbtinis intelektas; medicininė diagnostika; didelės rizikos DI sistemos; ES Dirbtinio intelekto aktas; medicinos priemonės; žmogaus priežiūra.

Finansavimas. Tyrimas atliktas vykdant Lietuvos mokslo tarybos (LMTLT) finansuojamą Studentų vasaros praktikos projektą „Dirbtinio intelekto taikymo medicininuose tyrimuose metodologijos parengimas: reguliaciniai iššūkiai ir praktinio įgyvendinimo modelis Lietuvoje“, sutarties Nr. P-SV4-26-207.

Interpretuojamas kalbos modelio asistentas struktūrinių chromosominių persitvarkymų trūkio taškų įrodymams vertinti retųjų ligų diagnostikoje

Vytautas Rimas¹, Eglė Preikšaitienė², Živilė Maldžienė², Erinija Pranckevičienė²

¹ Vilniaus universitetas, Medicinos fakultetas

² Vilniaus universitetas, Medicinos fakultetas, Biomedicinos mokslų institutas, Žmogaus ir medicininės genetikos katedra

Struktūriniai chromosominiai persitvarkymai, ypač subalansuotos translokacijos, yra viena sunkiausių retųjų ligų diagnostikos užduočių: jie nekeičia DNR kopijų skaičiaus, o tai, ar jie sukelia ligą, priklauso nuo tikslios chromosomų trūkio vietos. Kalbos modeliai galėtų padėti interpretuoti sekoskaitos duomenis, tačiau linę teigti faktus, kurių duomenys nepatvirtina. Tikslas – sukurti asistentą, kurio kiekvieną teiginį galima atsekti iki jį pagrindžiančio matavimo, ir nustatyti, kas riboja jo išvadas.

Metodai. Asistentą sudaro kalbos modelis ir įrankiai – nedidelės programos, kurias modelis paleidžia genomo/egzomo sekoskaitos duomenims palyginti su referentiniu genomu (BAM failui) ir kurios grąžina trūkio taškų pozicijų genome parametrus. Duomenis modelis mato tik per šiuos įrankius, prijungtus per modelio kontekstinį protokolą (angl. MCP). MCP padaro modelio užklausas labai specifines sprendžiamam uždaviniui. Pavyzdžiui, užklausus apie geną specifinėje pozicijoje, bus atrinkti tik iš anksto paruošti įrankiai darbui su genais. Kontekstinio protokolo įrankių rinkinyje keturi įrankiai matuoja ir atskirai įvertina požymius, kad chromosoma tam tikroje vietoje trūko ir susijungė su kita: skaitinių poras, kurių galai sulyginti skirtingose chromosomose, skaitinius, nukirptus ties trūkio tašku, skaitinius, padalytus tarp dviejų sujungtų chromosomų, ir sekos padengimą atspindintį kopijų skaičių. Kiti įrankiai tikrina duomenų kokybę, nustato geną trūkio vietoje, tikrina partnerinį trūkio tašką ir sukuria integratyvios genomo naršyklės (angl. Integrative Genomics Viewer, IGV) vaizdus. Pats asistentas genomo neskenuoja: standartinė programa DELLY peržiūri visą BAM rinkmeną ir sudaro kandidatinių trūkio taškų sąrašą, kurį asistentas filtruoja. Testavimui pasirinkti vieši NA12878 viso genomo duomenys (BAM rinkmena). Į viešus NA12878 viso genomo duomenis žinomose vietose buvo įterpta dvylika subalansuotų translokacijų tarp 20 ir 21 chromosomų. Naujai įterptas jungtis modeliuojančios trumpos DNR sekos suformuotos ART įrankiu (angl. Amplification of Read Tools). Tirti keturi kalbos modeliai: debesijos Claude Sonnet 5 ir Claude Opus 5 bei atvirieji Qwen2.5 7B ir Qwen3.5 4B, veikiantys per Ollama įprastame kompiuteryje su 8 GB vaizdo plokšte, todėl duomenys išlieka lokaliaje atmintyje ir yra apsaugoti. Modelių atsakymai apie trūkio taškus buvo testuojami daug kartų.

Rezultatai. Sukurta vartotojo sąsaja kaip žiniatiklio programa, leidžianti pasiekti įrankius ir komunikuoti su DI asistentu. Joje naudotojas įkelia DELLY programą nustatytą trūkio taškų sąrašą, mato, kiek kandidatų pašalina kiekvienas filtras, atveria kandidatą arba įveda genomo regiono kandidato pozicijos koordinates. Tame pačiame lange galima užduoti klausimą kalbos modeliui. Tikrinant sistemą iš 24 sukurtų jungčių rastos 16: visos 8 unikalioje sekoje, visos 8 šalia pasikartojančių sekų. Blogai įterptos jungtys į blogai kartografuojamas sritis buvo neidentifikuotos, nes tose srityse negalima trumpų sekų (angl. reads) vienareikšmiškai priskirti vietai genome. DELLY programa neįtraukė tokių sričių į kandidatų sąrašą. Trūkio taškų kandidatai yra ranguojami pagal įrodymų stiprumą. Įrodymus sudaro trumpos sekos, kurių poroje yra labai nutolę viena nuo kitos genome, ir taip pat fragmentuotos trumpų sekų poros. Kuo daugiau tokių požymių, tuo stipresni įrodymai apie trūkio tašką.

Išvados. Sukurtas sistemos prototipas nustatyti trūkio taškų kandidatams. Pademonstruotas

sistemos gebėjimas teisingai identifikuoti iki 100% sumodeliuotų trūkio taškų gerai kartografuojamose genomo srityse. Naujumas: chromosomų struktūrinių persitvarkymų variantų bei subalansuotų translokacijų trūkio taškuose asistentas kiekvieną įrodymą pateikia su visais jį grąžinusio įrankio parametrais. Asistentas yra skaidrus, nes pats parodo naudotojui savo ribas. Kalbos modelis nedaro haliucinacijų todėl, kad analizės rezultatus jis gali paaiškinti tik naudodamas informaciją, gautą iš kontekstinio protokolo įrankio. Ši interpretuojama DI sistema pateikia ne tik išvadas, bet ir faktus, reikalingus jai patikrinti. Ribotumai: rezultatai gauti su sumodeliuotomis translokacijomis viešo genomo chromosomose. Kandidatų paieška ne visada identifikuoja trūkius blogai kartografuojamose genomo srityse. Šiuo metu įrankių slenksčiai nustatyti pagal literatūrą ir turi būti kalibruojami pagal analizuojamus duomenis. Tolesni darbai: sistema bus testuojama su COLO829 viešais prieinamais duomenimis, kuriuose trūkio taškai yra žinomi. Sistema bus tikrinama su klinikiniais genomais, sprendimus priimant klinikiniam genetikui.

Raktažodžiai: struktūriniai chromosomų persitvarkymai, kalbos modeliai, sekoskaita.

Literatūros šaltiniai

1. Model Context Protocol. Specification. 2026. <https://modelcontextprotocol.io/specification>
2. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–26.
3. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–2079.
4. Rausch T, Zichner T, Schlattl A, et al. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28(18):i333–i339.
5. Byrska-Bishop M, Evani US, Zhao X, et al. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. Cell. 2022;185(18):3426–3440.
6. Huang W, Li L, Myers JR, Marth GT. ART: a next-generation sequencing read simulator. Bioinformatics. 2012;28(4):593–594.
7. Anthropic. System Card: Claude Sonnet 5, System Card: Claude Opus 5. 2026. <https://anthropic.com/claude-sonnet-5-system-card>, <https://anthropic.com/claude-opus-5-system-card>
8. Qwen Team. Qwen2.5 Technical Report. arXiv:2412.15115. 2024.
9. Qwen Team. Qwen3.5: Towards Native Multimodal Agents. 2026. <https://qwen.ai/blog?id=qwen3.5>
10. Ollama. <https://github.com/ollama/ollama>
11. Espejo Valle-Inclan J, Besselink N, de Bruijn E, et al. A multi-platform reference for somatic structural variation detection. Cell Genomics. 2022;2(6):100139.

Cancer Case Extraction from Full Electronic Health Records Using Open-Source Small Language Models

Justas Mileika, Tadas Žvirblis

Cancer Registry, Laboratory of Cancer Epidemiology, National Cancer Institute, Vilnius, Lithuania

Language models have been widely adopted across domains such as mathematics, software engineering and scientific writing, where they now handle tasks that once required specialist effort. This work explores whether they can bring similar gains to cancer registration. Electronic health records (EHRs) capture a patient’s entire medical history, including diagnoses, staging and pathology findings, but much of this information is buried in free-text prose spread across many forms, which a model must read through to locate the properties that define each cancer case. Because medical data is highly sensitive and must be handled with care, we deliberately use a small open-source model that runs locally on modest hardware, so patient data never leaves the registry.

We treat the patient, rather than the individual form, as the unit of extraction: all of a patient’s EHR forms are pooled and processed together. A first prompt runs over the whole form set and returns every registrable ICD-10 code found for that patient. Since one patient can carry several registrable codes, every subsequent step is resolved per code rather than per patient. For each code, separate prompts extract the diagnosis date, clinical and pathological TNM stage, grade and morphology, and each code produces its own cancer case form, so a patient with several tumours yields several forms. Separating the extractions in this way allows each variable to have its own task-specific prompt and decomposes the complex task of cancer case extraction into smaller, more manageable subtasks. All model outputs are constrained to JSON and validated against a fixed schema. All steps are run with Qwen3.6-35B-A3B, an open-source mixture-of-experts model.

We evaluated the pipeline on 2,000 human-labelled patients, each with at most seven EHR forms, which represents the easier end of the form-count distribution. Every variable except the ICD-10 code was evaluated conditioned on the ground-truth code, so that only that variable’s extraction is measured. The model identified the ICD-10 diagnosis code with 87.1% accuracy. TNM staging was extracted most reliably: pathological T, N and M reached 95.0%, 90.6% and 92.8%, and clinical T, N and M reached 90.2%, 85.9% and 89.1%. Grade (78.7%), diagnosis date (77.8%) and morphology (71.0%) proved more challenging.

These results show that a small, locally deployed open-source model can extract most cancer registry variables from full-record, largely unstructured EHR data with useful accuracy, without patient data leaving the institution. Variables that are usually stated explicitly, such as TNM stage, are extracted most reliably, while those that often require interpretation or reconciling several mentions across forms, such as diagnosis date, grade and morphology, lag behind. Future work will extend evaluation to patients with more than seven forms and target the weaker variables through prompt refinement and few-shot examples. As more labelled data is collected, models could also be trained and specialised on each of these smaller subtasks. We also plan to benchmark other open-source models and to integrate the pipeline into registrars’ workflow as a human-in-the-loop assistant.

Funding. This project has received funding from the Research Council of Lithuania (LMTLT), agreement No No S-ITP-26-11.

„Vibe coding“ patologijoje: reali sistema, iššūkiai ir rezultatai

Marius Kazlauskas

Valstybinis patologijos centras; Vilniaus universiteto Medicinos fakultetas (VU MF) Transliacinių sveikatos tyrimų institutas (TSTI)

Pats neparašiau nė vienos kodo eilutės – visą kodą parašė DI. Tačiau suvaldyti daugiau nei 40 tūkstančių kodo eilučių ir per mėnesį vienam sukurti sistemą, kuriai keliama aukšti funkcionalumo reikalavimai jautrioje medicinos srityje, buvo didelis iššūkis.

Gyvai parodysiu sukurtoje sistemoje paruoštas interaktyvias sinoptines patologijos formas, parengtas pagal tarptautinius standartus (pvz., Banff, ICCR). Pademonstruosiu, kaip supaprastiname formų atnaujinimą pasirodžius naujoms standartų versijoms ir kokią formų atvaizdavimo lankstumą pavyko pasiekti.

Pranešime papasakosiu visą kūrimo eigą: kokius sprendimus teko priimti, kur klydau, ką taisiau, kokius įrankius naudojavau ir kokių atsisakiau. Pasidalysiu, kas dirbant su DI programavimo agentu iš tiesų svarbu, kaip išlaikiau kontrolę, kur vis dar matau silpnų vietų, kokie iššūkiai buvo ir kokie dar laukia.

Nors kodo pats nerašiau, tačiau 12 metų programavimo patirtis pravertė – papasakosiu, kur ir kodėl.

From Questionnaires to Computable Evidence: Preparing Biobank Survey Data for Health AI

Hugas Jasinskas

Vilnius University, Lithuanian Population and Rare Diseases Biobank

Epidemiological questionnaires remain a cornerstone of biobank and cohort research. They capture exposures, behaviors and lifestyle factors that biological, clinical and molecular data cannot. Yet as AI becomes central to biomedical research, accuracy in the traditional epidemiological sense is no longer enough. Data used as model inputs must also be temporally consistent, semantically harmonized and suitable for automated processing. Gaps introduced at the stage of data collection propagate silently through machine-learning pipelines and can distort results and produce misleading conclusions.

This presentation, based on the monograph *From Questionnaires to Computable Evidence: Standardizing Epidemiological Data Pathways for Multimodal AI in Health* [Iveta Petraityte, Hugas Tauras Jasinskas, Giedre Kvedaraviciene], presents a data engineering framework for transforming questionnaire responses into AI-ready data assets. We first describe how raw responses become standardized, machine-readable variables and computable phenotypes. We also address the main sources of feature noise: recall and social desirability bias, and non-random missingness, which large programs such as All of Us have shown to be systematically linked to participant characteristics.

The central focus is temporal integrity. We show why a clearly defined baseline (t0) is essential for separating predictors from outcomes, and how experience from the UK, China Kadoorie, Estonian and Qatar biobanks supports collecting questionnaire data, measurements and biological samples within a single, timestamped assessment. We introduce feature volatility and phase alignment: the acceptable gap between questionnaire and sampling depends on how quickly a feature changes and on the biological integration window of the corresponding biomarker, from days for cotinine to months for HbA1c. We then discuss temporal, target and preprocessing leakage, which can substantially inflate model performance, and show how versioned baseline data freezes, database locks, preregistered analysis plans and TRIPOD+AI reporting protect against them.

Finally, we show how semantic standardization through the OMOP Common Data Model and automated quality-control frameworks supports FAIR, reproducible analysis, and how federated infrastructures such as HealthData@EU enable cross-border data discovery and reuse while preserving local governance. Interoperability, however, cannot compensate for poorly defined baselines or misaligned data modalities.

Attendees will leave with a practical checklist for building multimodal health AI on questionnaire-derived data: define t0, align phases, freeze before analysis, standardize, then federate.

Keywords: biobanking, epidemiological questionnaires, target leakage, data standardization, OMOP CDM, multimodal health AI

Dviejų apdorojimo strategijų palyginimas struktūrizuojant informaciją iš patologinių aprašų

Tomas Griguola^{1,2}, Dalius Matuzevičius², Tadas Žvirblis¹, Linas Petkevičius³

¹ Vėžio registras, Vėžio epidemiologijos laboratorija, Nacionalinis vėžio institutas ² Vilniaus Gedimino technikos universitetas ³ Vilniaus universitetas

Nacionalinis vėžio registras didelę dalį duomenų gauna nestruktūrizuota forma: patologijos aprašai rašomi laisvu tekstu, o registruojamas įrašas reikalauja tikslų laukų – TLK-10 kodo, lokalizacijos, morfologijos kodo, diferenciacijos laipsnio, pTNM ir stadijos. Šiuos laukus suveda registro darbuotojai. Darbas nėra vien perrašymas: tą patį pacientą dažnai aprašo keli skirtingu metu parengti pranešimai, todėl reikia nuspręsti, kurie iš jų priklauso tam pačiam vėžio atvejui, kuri reikšmė registruojama pagal registravimo tvarką ir kaip iš jos gaunama galutinė stadija.

Didieji kalbos modeliai dalį šio darbo leidžia automatizuoti, tačiau viena modelio užklausa pateikia atsakymą be pagrindimo: lieka neaišku, kuris pranešimas ir kuris sprendimas lėmė galutinę reikšmę. Ankstesni šios užduoties tyrimai daugiausia lygino skirtingų kalbos modelių našumą klasifikavimo uždaviniuose, čia lyginami ne modeliai, o sistemos sandaros variantai, ir keliamas klausimas, ką suteikia uždavinio išskaidymas.

Lyginamos dvi konfigūracijos. Bazinėje visas paciento pranešimų rinkinys apdorojamas viena kalbos modelio užklausa, kuria remiantis modelis pats turi nustatyti vėžio atvejų skaičių ir visas galutines reikšmes. Daugiapakopėje vaidmenys atskirti pagal principą, kad kalbos modelis naudojamas tik ten, kur reikia interpretuoti laisvą tekstą, o registro logika, kuri yra fiksuota, įgyvendinama determinuotomis taisyklėmis.

Grandinę sudaro nuoseklūs žingsniai. Kalbos modelis iš kiekvieno pranešimo išgauna lokalizaciją, morfologijos aprašymą, TLK-10 kodą, diferenciacijos laipsnį ir TNM reikšmes. Pranešimai rikiuojami chronologiškai ir grupuojami į vėžio atvejus pagal TLK-10, morfologijos ir lokalizacijos sutapimą. Modelio sprendimo prašoma tik tada, kai požymiai sutampa iš dalies. Atvejui taikomos registro taisyklės: registravimo langas, pažangiausios TNM reikšmės atranka ir stadijos nustatymas pagal TNM klasifikacijos lenteles. Kiekvienas įrašas saugo ir reikšmės kilmę, pritaikytą taisyklę ir įspėjimus, todėl atskirą sprendimą galima peržiūrėti nekartojant viso apdorojimo.

Naudojamas MedGemma-27B modelis, veikiantis lokaliai NVIDIA DGX Spark infrastruktūroje, todėl pacientų duomenys neišeina už įstaigos ribų. Abi konfigūracijos apdoroja tą patį duomenų poaibį, o prognozuoti vėžio atvejai su etaloniniais siejami pagal pranešimų identifikatorių sutapimą. Pranešime pristatoma sistemos sandara, vertinimo metodika ir abiejų konfigūracijų palyginimas.

Finansavimas. Finansavimą skyrė Lietuvos mokslo taryba (LMTLT), sutarties Nr. S-ITP-26-11.

Medicininį mokymo duomenų vertimo į lietuvių kalbą strategijos

Augustė Raišytė^{1,2}, Dalius Matuzevičius², Tadas Žvirblis³, Linas Petkevičius¹

¹ *Vilniaus universitetas;* ² *Vilniaus Gedimino technikos universitetas;* ³ *Vėžio registras, Vėžio epidemiologijos laboratorija, Nacionalinis vėžio institutas*

Didžiųjų kalbos modelių (angl. Large Language Models, LLM) pritaikymą medicinos srityje lietuvių kalba riboja kokybiškų, specializuotų mokymo duomenų trūkumas. Vienas iš būdų šiai problemai spręsti – esamų anglų kalba parengtų medicininių duomenų rinkinių vertimas į lietuvių kalbą. Tačiau medicininių tekstų vertimui keliama itin aukšti semantinio tikslumo ir terminologinio nuoseklumo reikalavimai, nes vertimo klaidos gali iškreipti medicininę informaciją ir neigiamai paveikti mokomų modelių kokybę.

Šio tyrimo tikslas – ištirti ir palyginti medicininių mokymo duomenų vertimo į lietuvių kalbą metodus bei nustatyti strategijas, leidžiančias sumažinti vertimo klaidų skaičių. Tyrime analizuojami skirtingais didžiaisiais kalbos modeliais ir specializuotais mašininio vertimo modeliais atlikti vertimai, vertinant jų semantinį tikslumą, medicininės terminijos atitiktį ir informacijos išsaugojimą. Dėmesys skiriamas specializuotų medicininių žodynų ir tekstynų integravimui į vertimo procesą, juos naudojant tiek kaip papildomą kontekstą kalbos modeliams, tiek kaip priemonę automatiškai aptikti ir taisyti terminologines bei semantines klaidas. Taip pat tiriamos skirtingos automatinio vertimo kokybės vertinimo ir klaidų aptikimo strategijos, analizuojamas jų veiksmingumas ir pritaikomumas didelės apimties duomenų rinkiniams.

Remiantis tyrimo rezultatais, siekiama parengti praktines medicininių mokymo duomenų rinkinių vertimui į lietuvių kalbą rekomendacijas, apimančias vertimo modelių pasirinkimą, terminologinių išteklių panaudojimą bei automatizuotą vertimų kokybės kontrolę.

Finansavimas. Finansavimą skyrė Lietuvos mokslo taryba (LMTLT), sutarties Nr. S-ITP-26-11.

Triukšmo mažinimas pilno lauko optinės koherentinės tomografijos trimačiuose vaizduose

Radoslav Titko¹, Egidijus Auksorius², Dalius Matuzevičius¹

¹ *Vilniaus Gedimino technikos universitetas (VILNIUS TECH), Elektronikos fakultetas, Elektroninių sistemų katedra, Vilnius, Lietuva* ² *Valstybinis mokslinių tyrimų institutas Fizinių ir technologijos mokslų centras (FTMC), Vilnius, Lietuva*

Pilno lauko optinė koherentinė tomografija (angl. full-field optical coherence tomography) leidžia sparčiai ir neinvaziškai gauti ląstelinės skiriamosios gebos audinių tūrius, tačiau jos vaizdus degraduoja multiplikatyvus ir erdviškai koreliuotas speklinis triukšmas (angl. speckle noise) bei šiai technologijai būdinga kryžminė sklaida (angl. crosstalk). Darbe apžvelgiami giliuoju mokymusi grįsti OKT speklių slopinimo metodai ir analizuojamas jų tinkamumas trimatiems FF-OCT duomenims. Nustatyta, kad dauguma prižiūrimumų, savipriežiūros ir difuzijos modelių sukurti dvimačiams OKT vaizdams ir remiasi nepriklausomo triukšmo prielaida, kurio speklinis triukšmas neatitinka, o nedaugelis žinomų giliojo mokymosi taikymų FF-OCT neišnaudoja trimatės duomenų struktūros.

Eksperimentai atlikti su trimačiais FF-OCT duomenimis, gautais naujai kuriamą vaizdinimo sistema. Kadangi etaloninių be triukšmo FF-OCT vaizdų praktiškai neįmanoma gauti, lyginami neprižiūrimumo giliojo mokymosi metodai, kuriuose neuroninis tinklas mokomas vien iš triukšmingų vaizdų. Noise2Noise metodas išnaudoja tai, kad FF-OCT tame pačiame gylyje vidurkinimui užfiksuoja kelis kadrus kurie rodo tą patį audinį, bet triukšmas juose bent iš dalies skiriasi, todėl vienas kadras paduodamas tinklui, o kitas naudojamas kaip tikslas. Noise2Void metodui užtenka vieno kadro, dalis jo taškų paslepama, o tinklas mokosi atspėti jų vertes iš aplinkinių taškų. Abiem atvejais tinklas išmoksta atkurti audinio struktūrą, bet ne atsitiktinį triukšmą. Rezultatai lyginami su Gauso ir medianos filtrais pagal etaloninio vaizdo nereikalaujančius signalo ir triukšmo (SNR) bei kontrasto ir triukšmo (CNR) santykius, vertinant jų kitimą didėjant gyliui.

Dirbtinis intelektas medicinos tyrimuose – bioetinio vertinimo patirtys (2020–2026 m.)

Lina Zabulienė, Bronius Buckus, Odeta Geleželytė, Ingrida Lisauskienė, Laura Paškevičienė, Sigita Petraitienė, Igoris Šatkauskas, Renata Klevienė

Vilniaus regioninis biomedicininų tyrimų etikos komitetas, Vilniaus universitetas

Įvadas. Dirbtinio intelekto (DI) sistemos medicinoje kuriamos, mokomos ir validuojamos naudojant paciento sveikatos duomenis, medicininius vaizdus ir biologinius ėminius. Todėl biomedicininuose tyrimuose, kuriuose šie duomenys naudojami DI sistemoms kurti ar vertinti, svarbus bioetinis vertinimas. Tarptautinė praktika ir komiteto ekspertinė patirtis rodo, kad tyrėjams ir DI sistemų kūrėjams ne visada aišku, kokius etinius, duomenų apsaugos ir tyrimo pagrįstumo aspektus būtina aptarti tyrimo dokumentuose.

Tikslas. Išanalizuoti Vilniaus regioninio Biomedicininų tyrimų etikos komiteto (toliau – Komitetas) pastabas dėl tyrimų, kuriuose taikomos DI sistemos ir skaitmeninės technologijos, dokumentų ir nustatyti pasikartojančias problemas ir jų pobūdžio pokyčius 2020–2026 m.

Metodai. Atlikta 19 komiteto pirminio vertinimo raštų, parengtų 2020–2026 m., kokybinė turinio analizė. Analizuotas pastabų pobūdis, jos sugrupuotos į temines kategorijas ir nagrinėtos trimis laikotarpiais (2020–2021, 2022–2024, 2025–2026 m.). Konkretūs tyrimai, tyrėjai, užsakovai ir tyrimo centrai neidentifikuojami.

Rezultatai. Dauguma nagrinėtų tyrimų buvo skirti diagnostikai ar prognozei, naudojant medicininius vaizdus ir elektroninius sveikatos įrašus; mažesnė dalis – gydymui, paciento stebėsenai ir edukacijai. Tarp naujesnių nagrinėtų kryptių išryškėjo DI taikymas gydytojų darbo procesams. Dažniausiai Komitetas teikė pirmines pastabas dėl asmens duomenų tvarkymo (18 iš 19 raštų), sutartinių santykių su partneriais (12), tyrėjų kvalifikacijos ir vaidmenų (12) ir dokumentų tarpusavio neatitikčių (12). Nuo 2022 m. kartojo pseudonimizuotų ir anonimizuotų duomenų painiojimas (7 raštuose). Pastabos dėl galimo interesų konflikto nustatytos tik 2020–2021 m. laikotarpio raštuose. 2025–2026 m. komitetas dažniau klausė apie DI sistemos gyvavimo ciklą – DI sistemos naudojimą pasibaigus tyrimui, atitiktį medicinos priemonių reikalavimams, šališkumo vertinimą ir galimybę DI mokymui naudoti sintetinius ar anonimizuotus duomenis.

Išvados. Nagrinėtuose Komiteto vertinimo raštuose išryškėjo pasikartojantys klausimai dėl asmens duomenų tvarkymo, partnerių atsakomybių, dokumentų nuoseklumo, DI sistemos pagrįstumo ir tolesnio naudojimo. Laikotarpių palyginimas leidžia įžvelgti dėmesio plėtrą nuo atskirų dokumentų ir duomenų tvarkymo klausimų iki DI sistemos gyvavimo ciklo vertinimo, tačiau šiuos pastebėjimus reikėtų vertinti kaip įžvalgas dėl nedidelės analizuotų raštų imties. Rengiant DI tyrimo protokolą rekomenduojama aiškiai aprašyti duomenų srautus ir partnerių atsakomybes, atskirti pseudonimizavimą ir anonimizavimą, pagrįsti pacientų duomenų naudojimo būtinybę ir iš anksto numatyti DI sistemos vertinimo metrikas ir tolesnio naudojimo sąlygas.

Raktažodžiai: *dirbtinis intelektas, biomedicininų tyrimų etika, asmens duomenų apsauga, pseudonimizavimas, informuotas asmens sutikimas.*

Nuo klinikinių duomenų iki dirbtinio intelekto: RAPIDS projektas retųjų ligų identifikavimui

Agnė Kerpauskienė

Visceralinės medicinos transliacinių tyrimų centras, Klinikinės medicinos institutas, Medicinos fakultetas, Vilniaus universitetas; VUL Santaros klinikos

Vilniaus universiteto tematinės mokslo grupės **RAPIDS (Rare Disease Phenotype Informed Data-driven Screening, vadovauja prof. A. Laurinavičius)** tikslas – kurti duomenų mokslu ir DI pagrįstus metodus, leidžiančius automatiškai atpažinti retosioms ligoms būdingus fenotipinius požymius ir identifikuoti potencialiai nediagnozuotus pacientus.

Pirmuoju klinikiniu modeliu pasirinktas Alporto sindromas, pasižymintis aiškiu genetiniu pagrindu, heterogenišku fenotipu ir turima genetiškai patvirtintų pacientų kohorta. Atlikto pilotinio tyrimo metu sukurta medicininių terminų atpažinimo sistema, pritaikyti natūralios kalbos apdorojimo (NLP) ir pacientų klasterizavimo metodai. Gauti rezultatai patvirtino šių metodų potencialą iš nestruktūruotų klinikinių tekstų išgauti reikšmingą fenotipinę informaciją.

Tolesni RAPIDS tyrimų etapai (2026–2029 m.) apima skirtingų sveikatos duomenų šaltinių integraciją, fenotipinių duomenų standartizavimą, NLP ir mašininio mokymosi modelių kūrimą bei klinikinį validavimą, naudojant VUL Santaros klinikų ir Valstybės duomenų agentūros duomenis. Vėlesniuose etapuose numatoma sukurtus algoritmus pritaikyti kitoms retosioms ligoms ir plėtoti prognostinius modelius.

Ilgalaikis projekto siekis – sukurti plečiamą, DI pagrįstą retųjų ligų fenotipinės atrankos sistemą, kuri sudarytų prielaidas ankstyvesnei diagnostikai, efektyvesniam sveikatos duomenų panaudojimui ir personalizuotos medicinos plėtrai.

Diagnostic Reliability of Multi-Label Chest X-Ray AI outside the Training Hospital: Accuracy, Calibration and Fairness across Seven Architectures

Paulius Bundza, Justas Trinkūnas

Vilnius Gediminas Technical University, Faculty of Fundamental Sciences, Department of Information Systems, Vilnius, Lithuania

Chest radiographs remain the first-line imaging examination in acute care, yet interpretation is labour-intensive under radiologist shortages. Multi-label deep-learning models can flag several thoracic findings from a single image, but published diagnostic accuracy is almost always measured on the training hospital. In a new institution, scanners, protocols, patient mix and labelling conventions induce domain shift. Uncalibrated probabilities and unequal subgroup performance further limit safe use as a diagnostic aid.

Seven TorchXRayVision classifiers were compared: six DenseNet-121 networks (224×224) trained on NIH ChestX-ray14, CheXpert, MIMIC-CXR (CheXpert or NegBio labels), PadChest, or a multi-source mix, and one ResNet-50 (512×512) trained on the multi-source mix. In-distribution evaluation used NIH ChestX-ray14 (112,120 frontal images; a 1,000-image subset for ResNet-50). External evaluation used a Brazilian BRAX v1.1.0 frontal cohort (18,424 images) absent from every training set. Comparison used seven shared labels: atelectasis, cardiomegaly, consolidation, edema, pleural effusion, pneumonia and pneumothorax. Models were scored with AUROC, average precision, Brier score and expected calibration error (ECE), with 500-iteration patient-level cluster bootstrap 95% confidence intervals. Fairness was assessed by sex, age band, radiographic view (AP versus PA) and equipment manufacturer. The same classifiers are served by a deployable prototype (MCADS; <https://mcads.cloud>) with Grad-CAM explanations, out-of-distribution quality gating and 14-structure anatomical segmentation.

On BRAX, macro AUROC ranged from 0.652 to 0.741. Multi-source or richly annotated training separated the top three models (PadChest DenseNet 0.741, multi-source DenseNet 0.738, ResNet-50 0.733) from four single-source North-American models (0.652–0.677). Overlapping models lost 3.4–8.7 percentage points of macro AUROC relative to NIH. Discrimination and calibration were largely independent: ResNet-50 achieved the best ECE (0.028) while ranking only third in AUROC, whereas a MIMIC CheXpert-labeler DenseNet reached ECE 0.392. Subgroup gaps were clinically large. AP versus PA AUROC differed by up to 21.0 percentage points (NIH-trained DenseNet fell to chance-level 0.499 on AP views). For the multi-source DenseNet, paediatric (ages 0–4) and geriatric (85+) bands fell to 0.563 and 0.420. Sex differences were modest (at most +0.032 for males).

In-distribution leaderboards systematically overstate diagnostic robustness. Before clinical use, chest X-ray AI should undergo external validation, probability calibration and explicit warnings for portable AP radiographs and extreme age, rather than being selected by a single AUROC. Limitations include NLP-derived labels, extreme rarity of some BRAX findings (edema $n = 9$) and a single external hospital.

DI vėžio molekulinį mechanizmų visumos struktūrizavime

Kęstutis K. Urba

Tarptautinis projektas „K-ras iniciatyva“

Vėžio aprašo biologiniame lygmenyje uždavinys yra išspręstas per pastaruosius du dešimtmečius daugiausia amerikiečių tyrėjų Douglas Hanahan ir Robert A. Weinberg pastangomis. Jų straipsniai, 2000, 2011, 2022, 2026-ais metais, išryškinę 14-a pagrindinių vėžio savybių ir požymių, tarp kurių svarbiausias laikytini a) spartesnė proliferacija b) imuninio atsako slopinimas c) naviko aplinkos destrukcija d) naviko kapiliarizacija, limfaangiogenezė ir metastazavimas, jau susilaukė keliasdešimties tūkstančių citavimų ir tapo pripažinta vėžio biologijos klasika. Vienok, itin reikšmingo klinikai – efektyvių vaistinės terapijos protokolų kūrimui tumorigenezės identifikavimo uždavinio gvildenimas reikalauja žinomų piktybinės ląstelės ir jų visumos aplinkos molekulinį mechanizmų apjungimo su biologiniu aprašu schematizavimo - struktūrizavimo, kaip sisteminės metodologijos, priešpastatytos gajai fragmentinei, išraiška. Žmogiškose galiose yra bent jau tumorigenezės procesų šerdies schematizavimas: gautas grafinis atvaizdas, kurį sėkmingai papildė ir redagavo DI, o galutiniame variante išryškėjo 15-a lemiančių vėžio raidą grandžių, kurių superpozicijos brėžia daugybę galimų scenarijų su tuo pačiu rezultatu – išryškėjimu biologiniame apraše. Kadangi su tumorigeneze ir jos raida siejama 750 žmogaus genų iš daugiau nei 20-ties tūkstančių, tai pilna tumorigenezę nusakanti schematizuota struktūra, kurioje įvairiais scenarijais vyksta audinio piktybėjimo procesai, yra už žmogiškojo kognityvumo galimybių ribų, tikintis modelinės pagalbos iš DI, kuris turėtų remtis ir pvz. PubMed straipsnių bazėje daugiau nei 5.5 mln. esančių publikacijų, skirtų vėžiui, rezultatais.

Sveikimas realiuoju laiku: dirbtinis intelektas ir dėvimieji prietaisai ankstyvai pooperacinių komplikacijų diagnostikai

Marijus Ambrazevičius, Tomas Poškus

Vilniaus universiteto Medicinos fakulteto Klinikinės medicinos instituto Gastroenterologijos, nefrourologijos ir chirurgijos klinika

Įvadas.

Pooperacinės komplikacijos tebėra viena pagrindinių nepalankių abdominalinės chirurgijos išiečių priežasčių. Įprastai paciento būklė po operacijos stebima periodiškai matuojant gyvybinius rodiklius, kliniškai vertinant pacientą ir remiantis subjektyviais sveikimo rodikliais. Taikant tokį modelį, subtilūs fiziologiniai pokyčiai, atsirandantys dar prieš komplikacijai pasireiškiant kliniškai, gali būti pastebėti per vėlai. Dėvimieji išmanieji prietaisai leidžia nuolat ir neinvasiškai registruoti paciento fiziologinius ir funkcinės būklės rodiklius, o dirbtinis intelektas (DI) – paversti šiuos daugiamačius, laike kintančius duomenis individualizuotais skaitmeniniais biomarkeriais.

Atlikta naujausios literatūros apžvalga, daugiausia remiantis 2025–2026 m. PubMed duomenų bazėje indeksuotais tyrimais. Analizuoti tyrimai, kuriuose dėvimaisiais prietaisais registruotas žingsnių skaičius, fizinis aktyvumas, širdies susitraukimų dažnis, širdies ritmo variabilumas, miegas, deguonies saturacija ir kiti fiziologiniai rodikliai. Vertinta, kaip šie rodikliai susiję su pooperacinėmis komplikacijomis, gydymo trukme, pakartotinėmis hospitalizacijomis ir funkcinio atsistatymo dinamika, taip pat ar DI ir mašininio mokymosi algoritmais galima prognozuoti nepalankias baigtis.

Rezultatai.

Naujausi duomenys rodo, kad dėvimųjų prietaisų vertė pooperacinėje medicinoje vis dažniau siejama ne su pavieniais rodikliais, o su individualios sveikimo trajektorijos analize. Didelės apimties tyrime, kuriame analizuoti 1965 chirurginių pacientų duomenys, didesnis fizinis aktyvumas po operacijos buvo susijęs su trumpesne hospitalizacija ir mažesne komplikacijų bei pakartotinės hospitalizacijos rizika per 30 ir 90 dienų.

Ypač daug žadanti DI taikymo kryptis – kompleksinė fiziologinių signalų analizė. 2025 m. tyrime vartotojams skirtais dėvimaisiais prietaisais registruoti širdies ritmo ir aktyvumo duomenys buvo analizuojami mašininio mokymosi algoritmais. Taip pooperacines komplikacijas pavyko prognozuoti iki 3 dienų anksčiau, nei jos buvo diagnozuotos kliniškai; modelio jautrumas siekė 91 %, specifiškumas – 74 %.

Išvados.

Dėvimieji prietaisai kartu su DI leidžia pereiti nuo periodinės, reaktyvios stebėsenos prie nuolatinio ir individualizuoto paciento sveikimo trajektorijos vertinimo. Vis dėlto trūksta didelių prospektyvių daugiacentrių tyrimų, įrodančių, kad DI pagrįstos perspėjimo sistemos mažina komplikacijų dažnį ar mirštamumą. Praktinį taikymą taip pat riboja trūkstami duomenys, signalų kokybė, klaidingi perspėjimai ir integracija į klinikinę darbo eigą.

Dirbtinio intelekto taikymas odontologijoje: vyresnio amžiaus pacientų osifikuojančio miozito kalcifikatų detekcija panoraminėse rentgenogramose

Miglė Baranauskaitė, prof. dr. (HP) Alina Pūrienė

Vilniaus universiteto Medicinos fakulteto Odontologijos institutas

Tikslas. Sukurti bei įvertinti mašininio mokymo modelį, skirtą osifikuojančio miozito kalcifikatų atpažinimui panoraminėse rentgenogramose, ir atlikti šių radinių paplitimo bei pasiskirstymo pagal amžių, lytį ir lokalizaciją analizę.

Metodika. Retrospektyviajam tyrimui, kuriam gautas Vilniaus regioninio biomedicininio tyrimų etikos komiteto leidimas, panaudotos 6000 nuasmenintų vyresnių nei 45 metų pacientų panoraminių rentgenogramų (moterų ir vyrų santykis – 1:1). Vaizdų analizę atliko apmokyta tyrėja; jos vertinimo patikimumas patvirtintas pakartotinai išanalizavus 15 į pagrindinę imtį neįtrauktų rentgenogramų. Osifikuojančio miozito kalcifikatai identifikuoti naudojant programinę įrangą „MicroDicom“, o jų anotavimas atliktas platformoje „MakeSense AI“, suskirstant darinius į vienpusius (kairėje arba dešinėje pusėje) ir abipusius. Automatinei detekcijai sukurtas dirbtinio intelekto modelis, naudojant programavimo kalbą „Python 3.8“. Modelio veiksmingumas validuotas su nepriklausomu testiniu rinkiniu, taikant F1 balo ir *IoU* (angl. *Intersection over Union*) metrikas. Statistinė duomenų analizė atlikta paketu „IBM SPSS v.26“, taikant aprašomąją statistiką, Chi-kvadrato (χ^2), Mann–Whitney U testus bei koreliacijos analizę (skirtumai laikyti statistiškai reikšmingais, kai $p < 0,05$).

Rezultatai. Išanalizavus 6000 panoraminių rentgenogramų, osifikuojančio miozito kalcifikatai nustatyti 45 tiriamiesiems (0,8 %): 18 vyrų ir 27 moterims. Vyrų grupėje vienpusiai radiniai nustatyti 16 pacientų (88,9 %), abipusiai – 2 (11,1 %); moterų grupėje šie rodikliai atitinkamai siekė 25 (92,6 %) ir 2 (7,4 %). Vertinant kalcifikatų pasiskirstymą pagal lytį, statistiškai reikšmingų skirtumų nerasta ($p = 0,178$). Iš viso identifikuoti 49 kalcifikatai: dešinėje pusėje $n = 23$, kairėje $n = 18$, o abipusė lokalizacija nustatyta 4 pacientams. Dešinės pusės dominavimas buvo statistiškai reikšmingas ($\chi^2 = 31,364$; $p < 0,001$). Pacientų amžius neturėjo statistiškai reikšmingos įtakos vienpusių ar abipusių kalcifikatų pasireiškimui ($p > 0,05$). Sukurtas DI modelis, kurio F1 metrikos vertė viršijo 0,80, pademonstravo gerą tikslumo ir jautrumo balansą, būtiną patikimai kalcifikatų detekcijai.

Išvados. Sukurtas mašininio mokymo modelis geba efektyviai atpažinti osifikuojančio miozito kalcifikatus vyresnio amžiaus pacientų panoraminėse rentgenogramose. Nustatyta statistiškai reikšminga kalcifikatų pasiskirstymo asimetrija su vyraujančia lokalizacija dešinėje pusėje. DI taikymas leidžia optimizuoti poliligitumu pasižyminčios senėjančios visuomenės radiologinę praktiką, užtikrinant ankstyvą minkštųjų audinių patologijų diferencinę diagnostiką.

Raktažodžiai. Osifikuojantis miozitas; kalcifikatai; panoraminė rentgenograma; dirbtinis intelektas; senėjanti visuomenė.

Dirbtinio intelekto panaudojimas *H. pylori* infekcijos endoskopinėje diagnostikoje

Geistė Tubutytė, Jonas Povilavičius, dr. Arūnas Strumila

Vilniaus universiteto Santaros klinikos Vaikų ligoninė, Vaikų chirurgijos, ortopedijos ir traumatologijos centras

Įvadas. *Helicobacter pylori* infekcijos diagnostikoje endoskopinis skrandžio gleivinės vertinimas gali suteikti svarbios informacijos apie infekcijos tikimybę. Kyoto klasifikacija leidžia sistemingai vertinti endoskopinius gleivinės požymius, tačiau jų atpažinimas priklauso nuo endoskopo patirties ir gali būti subjektyvus. Dirbtinio intelekto (DI) technologijos atveria galimybę automatizuoti endoskopinių vaizdų analizę ir standartizuoti diagnostinių požymių vertinimą.

Tikslas. Sukurti DI pagrįstą sistemą, gebančią iš endoskopinio skrandžio vaizdo automatiškai atpažinti *H. pylori* infekcijai būdingus gleivinės požymius ir įvertinti infekcijos tikimybę.

Metodai. Sukurtas DI pagrįstas algoritmas, integruotas į išmaniojo telefono programėlę. Sistema analizuoja įkeltą endoskopinį skrandžio vaizdą ir automatiškai identifikuoja *H. pylori* infekcijai būdingus gleivinės požymius, įskaitant atrofiją, žarninę metaplaziją, nodulinį gastritą ir kitus endoskopinius pokyčius. Atpažintų požymių visuma naudojama *H. pylori* infekcijos tikimybei įvertinti. Algoritmo kūrimo ir klinikinio pagrindo vertinimui panaudoti 91 vaiko endoskopinių tyrimų duomenys. Šioje kohortoje endoskopinių požymių vertė vertinta pagal Kyoto klasifikaciją ir lyginta su *H. pylori* infekcijos nustatymo rezultatais.

Rezultatai. 91 vaiko kohortoje Kyoto klasifikacija parodė gerą gebėjimą diferencijuoti pacientus pagal *H. pylori* infekcijos tikimybę (ROC AUC – 0,83). Remiantis šiais rezultatais sukurta DI sistema, kuri iš endoskopinio vaizdo savarankiškai atpažįsta vertinamus gleivinės požymius ir, integruodama jų visumą, pateikia *H. pylori* infekcijos tikimybės įvertį. Automatizuotas vaizdinių požymių vertinimas leidžia sumažinti subjektyvaus žmogaus vertinimo įtaką ir sudaro prielaidas standartizuotai endoskopinių vaizdų interpretacijai.

Išvados. Sukurta DI pagrįsta sistema suteikia galimybę automatizuoti endoskopinių skrandžio vaizdų analizę ir įvertinti *H. pylori* infekcijos tikimybę. Šis sprendimas gali būti perspektyvi skaitmeninė priemonė endoskopinei diagnostikai, o tolesnė validacija leis įvertinti jo diagnostinį tikslumą ir klinikinio pritaikymo galimybes.

Viso pjūvio morfometrijos ir Banff vertinimo sistemos palyginimas prognozuojant inksto transplantato praradimą

Renaldas Augulis, Allan Rasmusson, Akshat Singh, Kuang-Yu Jen, Arvydas Laurinavičius

Valstybinis patologijos centras; Vilniaus universitetas; University of California, Davis

Įvadas. Banff klasifikacija yra tarptautinis inksto transplantato biopsijų vertinimo standartas, tačiau jos balai yra pusiau kiekybiniai ir gali skirtis priklausomai nuo vertintojo. Giliojo mokymosi algoritmai leidžia segmentuoti ir išmatuoti audinio struktūras visame inksto biopтата skaitmeniniame pjūvyje. Tyrimo tikslas buvo nustatyti, ar tokie morfometriniai matavimai papildė Banff vertinimą prognozuojant transplantato praradimo laiką.

Metodai. Retrospektyviame vieno centro tyrime išanalizavome 342 pacientų pirmąsias transplantato biopsijas. Tyrimo baigtis buvo transplantato praradimas, apibrėžtas kaip grįžimas į nuolatinę dializę arba pakartotinė transplantacija. Patologijos duomenys (Banff balai ir C4d), klinikinė informacija ir morfometriniai rodikliai buvo sujungti su pacientų ir transplantatų baigčių duomenimis Valstybės duomenų agentūros (VDA) analitinėje aplinkoje. Picrosirius red dažytuose pjūviuose apskaičiavome segmentuotų kanalėlių ir intersticiumo struktūrų morfometrinius rodiklius bei faktorius, išskirtus iš vietinių matavimų šešiakampių gardelėje. Šiuos rodiklius ir faktorius palyginome su Banff balais ir įtraukėme į modelius kartu su iš Banff balų ir C4d sudarytais atmetimo fenotipais. Į visus modelius taip pat įtraukėme paciento amžių, lytį ir laiką nuo transplantacijos.

Rezultatai. Stebėjimo trukmės mediana buvo 4,4 metų; transplantatą prarado 109 pacientai. Pirminėje kryžminėje patikroje morfometrijos modelio C indeksas (CI) siekė 0,820, Banff balų modelio 0,771, o fenotipų modelio 0,800. Morfometriją sujungus su fenotipais, CI padidėjo iki 0,823. Šis jungtinis modelis pagerino fenotipų modelio skiriamąją gebą (CI prieaugis 0,023, 95 % pasikliautinis intervalas 0,001–0,046). Šešiakampių gardelių faktoriai pridėtinės vertės prognozuojant transplantato praradimą nesuteikė. Nekeistus modelius pritaikius 95 tų pačių pacientų pakartotinėms biopsijoms, histologija pagrįsti modeliai prognozavo transplantato praradimą. Be to, rizikos įverčio padidėjimas tarp pirmosios ir antrosios biopsijos prognozavo transplantato praradimą nepriklausomai nuo pirmosios biopsijos rizikos įverčio. Išorinėje kohortoje (69 pacientai, 18 transplantato praradimų) jungtinis modelis išlaikė skiriamąją gebą ($C = 0,843$, 95 % pasikliautinis intervalas 0,75–0,92).

Išvados. Viso pjūvio morfometrija suteikia interpretuojamą, nuo vertintojo nepriklausančią prognostinę informaciją ir gali papildyti Banff vertinimą. Prieš taikant klinikinėje praktikoje būtinas patvirtinimas daugiacentriuose tyrimuose.

Neurologinės baigties po širdies sustojimo prognozavimas taikant kiekybinę EEG analizę ir mašininio mokymosi metodus

Martynas Mikalauskas

Elektronikos fakultetas, elektroninių sistemų katedra Vilniaus Gedimino technikos universitetas

Neurologinės baigties prognozavimas po širdies sustojimo išlieka sudėtingu klinikiniu uždaviniu, ypač pacientams, kurie po gaivinimo išlieka komos būsenos. Ilgalaikė elektroencefalografijos (EEG) stebėseną leidžia vertinti smegenų elektrinio aktyvumo pokyčius, tačiau didelė įrašų apimtis, nevienoda signalų kokybė ir kintanti jų prognostinė vertė apsunkina sistemingą duomenų analizę. Todėl aktualūs automatizuoti metodai, leidžiantys atrinkti kokybiškus EEG segmentus, išskirti kiekybinius požymius ir įvertinti jų tinkamumą neurologinei baigčiai prognozuoti skirtingais stebėsenos laikotarpiais.

Tyrimo tikslas – sukurti ir įvertinti automatizuotą EEG signalų apdorojimo ir mašininio mokymosi metodiką neurologinei baigčiai po širdies sustojimo prognozuoti, atsižvelgiant į EEG signalų kokybę, kiekybinius požymius ir skirtingus stebėsenos laikotarpius.

Tyrimas atliktas naudojant viešai prieinamą I-CARE duomenų rinkinį. Sukurta automatizuota EEG signalų apdorojimo seka, apimanti filtravimą, kokybės vertinimą, tinkamų 10-ties sekundžių trukmės epochų atranką, transformavimą į bipolinį montazą ir kiekybinių požymių išgavimą. Analizuotos spektrinės, statistinės, amplitudės dinamikos ir signalo kompleksiskumo charakteristikos. Neurologinei baigčiai prognozuoti taikyti penki mašininio mokymosi algoritmai ir dviejų lygių sanglaudos ansamblinis modelis. Naudojant EEG duomenis, sukaupus per pirmąsias 24, 48 ir 72 valandas po širdies sustojimo, visi metodai optimizuoti ir vertinti taikant įdėtinę kryžminę validaciją.

Taikant automatizuotą EEG epochų atranką, tinkami signalų segmentai nustatyti 76,7 % analizuotų pacientų stebėsenos valandų. Sanglaudos ansamblinio modelio diskriminacinė geba skirtingais stebėsenos laikotarpiais išliko panaši (AUROC 0,789–0,801), tačiau didžiausias jautrumas esant 95 % specifiškumui (FPR = 0,05) pasiektas 48 val. horizonte (0,512), palyginti su 24 val. (0,412) ir 72 val. (0,378). Tuo tarpu 72 val. horizonte nustatytas didžiausias subalansuotas tikslumas (0,745) ir specifiškumas (0,820).

Gauti rezultatai rodo, kad ilgesnė EEG stebėseną savaime neužtikrina geresnės neurologinių baigčių prognozavimo gebos, o palankiausias stebėsenos laikotarpis gali priklausyti nuo pasirinkto vertinimo kriterijaus. Nors bendroji modelio diskriminacinė geba visais vertintais laikotarpiais išliko panaši, didžiausias jautrumas esant aukštam specifiškumui nustatytas 48-tos val. vertinimuose. Tai rodo, kad panaši bendroji modelio diskriminacinė geba nebūtinai reiškia vienodą jautrumą esant aukštam specifiškumui. Vis dėlto skirtinga pacientų imčių sudėtis ir išorinės validacijos stoka riboja rezultatų apibendrinamumą. Tolesniuose tyrimuose tikslinga vertinti EEG požymių laikinę dinamiką, modelių patikimumą nepriklausomose pacientų imtyse ir jų pritaikomumą kompleksinei neurologinei prognozei.

„Shortcut learning“ melanomos klasifikavimo užduotyje taikant dirbtinį intelektą

Adriana Voišvilo

Vilniaus universitetas

Dirbtinio intelekto modeliai vis dažniau taikomi medicininių vaizdų analizei, tačiau geri klasifikavimo rezultatai nebūtinai reiškia, kad modelis priima sprendimus remdamasis kliniškai reikšmingais požymiais. Viena iš galimų problemų yra „*shortcut learning*“, kuri pasireiškia, kai modelis išmoksta pasinaudoti požymiais, kurie gali būti tiesiogiai nesusiję su sprendžiamą užduotimi. Pranešimo metu bus aptariamos mokslinėje literatūroje nagrinėjamos „*shortcut learning*“ priežastys ir galimos pasekmės dirbtinio intelekto patikimumui medicinoje.

Taip pat bus pristatytas tyrimas, kuriame naudojamas giluminio mokymosi konvoliucinis neuroninis tinklas „*EfficientNet-B3*“, skirtas odos darinių vaizdams klasifikuoti į dvi klases: melanomą ir ne melanomą. Tyrimui naudojamas ISIC 2020 duomenų rinkinys, kuriame pateikiami odos darinių vaizdai ir papildoma informacija apie pacientus ir pažeidimus. Tyrimo tikslas – įvertinti, ar modelio vaizdų reprezentacijose užkoduota informacija apie paciento lytį, amžiaus grupę ir odos darinio anatominę vietą, ir išanalizuoti, ar melanomos klasifikavimo rezultatai skiriasi tarp šių požymių grupių.

Modelis mokomas atpažinti melanomą tik iš odos darinių nuotraukų, nenaudojant papildomos informacijos apie pacientą. Vėliau tikrinama, ar iš modelio išminktų vaizdų reprezentacijų galima nustatyti pacientų lytį, amžiaus grupę ir odos darinio anatominę vietą. Tam apmokomi klasifikatoriai, o jų rezultatai vertinami naudojant testavimo duomenis. Taip pat lyginami melanomos klasifikavimo rezultatai skirtingose pacientų grupėse. Pagal gautus rezultatus galima įvertinti, ar modelio reprezentacijose yra informacijos apie demografinius ir anatominius požymius.

Dirbtinio intelekto metodų taikymas žmogaus fiziologinių signalų analizei

Anastasija Grubinskienė

Vilniaus Gedimino technikos universitetas, Elektroninių sistemų katedra

Žmogaus fiziologiniai signalai, tokie kaip širdies ritmas ir jo kintamumas, odos elektrinis laidumas, temperatūra ir judesys, vis dažniau registruojami įterptinėse ir nešiojamose sistemose. Kasdienėje aplinkoje šie matavimai yra individualūs ir triukšmingi, todėl vien tradicinių slenksčių dažnai nepakanka. Pranešime nagrinėjamas dirbtinio intelekto metodų taikymas tokių signalų analizei.

Tyrimas siejamas su disertacijos tema „Žmogaus biomedicininį signalų įterptinėse sistemose metodų analizė“. Lyginami skirtingi analizės būdai: signalo paruošimas, požymių išskyrimas iš laiko eilučių, asmeninės ramybės bazės sudarymas ir mokymosi modeliai.

Keliamas uždavinys ne tik nustatyti fiziologinės būklės pokytį, bet įvertinti jį objektyviai, remiantis matuojamais signalais, o ne vien subjektyviu aprašymu. Ypatingas dėmesys skiriamas galimybei prognozuoti artėjančius įvykius, pavyzdžiui, skausmą, pagal ramybės signalus, užregistruotus prieš įvykio pradžią.

Pranešime aptariama, kokie dirbtinio intelekto metodai tinkami objektyviam fiziologinės būklės vertinimui ir įvykių prognozei.

Klasifikavimo modeliai medicinai

doc. dr. Linas Petkevičius

Vilniaus universitetas, Matematikos ir informatikos fakultetas, Informatikos institutas, Dirbtinio intelekto metodų laboratorija, Vilnius, Lietuva

Naujos kartos sprendimų modeliai, tokie kaip Jev, Laya, bei kalbos modeliai MedGemma ar Qwen, gali būti pritaikyti į sprendimų modelius, keičia klasifikavimo užduotis medicinoje. Vietoj laisvo teksto generavimo su kalbos modeliais šie modeliai skaito laisvą ar klinikinį tekstą (pokalbių, laboratorinius rezultatus ar elektroninius sveikatos įrašus), įvertina iš anksto apibrėžtas klases (diagnozes, skubumo lygį, rekomendacijas ar simptomų klasifikaciją) ir grąžina pasirinkimą kartu su kalibruotomis tikimybėmis. Dėl to jie tampa greitesni vykdyti, lengviau integruojami į sistemas ir geriau pritaikyti konkrečioms medicininėms klasifikavimo užduotims, o naudojant vietinį mokymą galima pritaikyti modelius prie įstaigos duomenų neprarandant konfidencialumo. Pranešimo metu bus pristatomi šių modelių statistikos adaptuojant kalbos modelius ir Lietuviškiems tekstams.

Raktažodžiai: dirbtinis intelektas; medicininė diagnostika; didelės rizikos DI sistemos; ES Dirbtinio intelekto aktas; medicinos priemonės; žmogaus priežiūra.

Finansavimas. Finansavimą skyrė Lietuvos mokslo taryba (LMTLT), sutarties Nr. S-ITP-26-11.