

VU Matematikos ir informatikos fakulteto / VU Faculty of Mathematics and Informatics
Informatikos magistro studijų programos / Informatics Masters Study Programme
Mokslo tiriamųjų darbų temos 2023 - 2024 m. / List of Research Work topics for 2023 - 2024

Nr.	Dėstytojas / Supervisor	Tema LT / Topic in Lithuanian	Tema En / Topic in English	Pastabos / About
1	Adomas Birštunas	Skaičiavimai BDI logikai	Calculi for BDI logic	
2	Adomas Birštunas	Skaičiavimai modalinių logikų fragmentams	Calculi for fragments of modal logics	
3	Aistis Raudys	Vertybinių popierių prekybos sistemų „peroptimizavimo“ efekto tyrimas ir prognozavimas naudojant mašininio mokymosi metodus	Investigation and prediction of the effect of overoptimization of securities trading systems using machine learning methods	VP prekybos sistemos suoptimizuotos naudojant istorinius duomenis realybėje dažnai nuvilia. Reikėtų ištirti kokie faktoriai ir kiek lemia prekybinių sistemų peroptimizavimą. Tai leistų kurti sistemas kurios ir realybėje prekiauja pelningai.
4	Aistis Raudys	Algoritminis menas, kaip algoritmai naudojami meno dirbiniams kurti	Algorithmic art as algorithms used to create works of art	
5	Aistis Raudys	XGBoost medžių ir kitų sprendimo medžių tyrimas	Investigation of XGBoost trees and other decision trees	
6	Aistis Raudys	Alternatyvių duomenų analizė finansų rinkose	Analysis of alternative data in financial markets	
7	Aistis Raudys	Akių sekimo technologijų panaudojimas išmaniuose įrenginiuose	Use of eye tracking technologies in smart devices	
8	Aistis Raudys	Dirbtinis intelektas IoT įrenginiuose, neuroniniai tinklai mikrokontroleriuose ir mikrokompiuteriuose	Artificial intelligence in IoT devices, neural network for microcontrollers	Naudojant Arduino ar pan. MC realizuoti DNT ir naudojant sensorius atpažinti aplinka ar objektus, pritaikyti IoT. Tensorflow Light. Ištirti kaip geriausia realizuoti neuroninius tinklus MC ir naudojant sensorius atpažinti aplinka ar objektus pritaikyti IoT.
9	Aistis Raudys	Pasikartojimo suradimo algoritmai laiko eilutėse	Finding Repeated Structure in Time Series: Algorithms and Applications	https://www.cs.unm.edu/~mueen/Tutorial/SDM2015Tutorial2.pdf
10	Aistis Raudys	Multikriterinio optimizavimo metodų tyrimas	Research of multicriteria optimization methods	
11	Aistis Raudys	Balso atpažinimas mikrokompiuteriuose	Voice recognition in microcomputers	
12	Aistis Raudys	Balso generavimas mikrokompiuteriuose	Voice generation in microcomputers	
13	Aistis Raudys	Giliųjų neuroninių tinklų priešiškas apmokymas garso atpažinime – kaip apgauti balso atpažinimo sistemas	Hostile training in deep neural networks in voice recognition - how to trick voice recognition systems	
14	Aistis Raudys	Multikriterinis pastiprintas mokymas	Multicriteria reinforcement learning	
15	Aistis Raudys	3D/PointCloud ir Kinect v2 / Intel® RealSense vaizdų atpažinime	3D / PointCloud and Kinect v2 / Intel® RealSense image recognition	
16	Aistis Raudys	Žmonių srauto prognozavimas viešame transporte naudojant DI metodus bei autobusų tvarkaraščio ir dydžio planavimas	People traffic forecasting in public transport using DI methods and bus schedule and size planning	
17	Aistis Raudys	Lietuviško teksto generavimo metodai naudojant generatyvinius neuroninius tinklus.	Lithuanian text generation methods using generative neural networks.	
18	Aistis Raudys	Pastiprinto mokymo (ang. reinforcement learning) ir neuroninių tinklų taikymas vertybinių popierių prekyboje	Application of reinforcement learning and neural networks in securities trading	
19	Aistis Raudys	Interaktyvių reklamų kūrimas naudojant mašininio matymo neuroninius tinklus (OpenCV ir pan.)	Creating interactive advertisements using machine vision neural networks (OpenCV, etc.)	

20	Aistis Raudys	Artimiausių kaimynų metodo tobulinimas optimizuojant atstumo funkciją	Improving the nearest neighbors method by optimizing the distance function	
21	Aistis Raudys	Kelių kalbėtojų atskyrimas įrašant keliais mikrofonais (garso šaltinio trianguliacija). Garso šaltinių atskyrimas iš įrašo - kalbėtojo atskyrimas naudojant DNN ir kitus metodus	Separation of multiple speakers by recording with multiple microphones (triangulation of an audio source).	
22	Aistis Raudys	Realaus laiko aktyvus triukšmo panaikinimas (ang. noise canceling) garso įrašuose ir/arba naudojant kelis mikrofonus naudojant neuroninius tinklus.	Separation of audio sources from a recording - Separation of a speaker using DNN and other methods	
23	Aistis Raudys	Vaizdus generuojantys neuroniniai tinklai	Real-time active noise canceling in audio recordings and / or using multiple microphones using neural networks.	
24	Aistis Raudys	Kalbos kokybės rekonstravimas/gerinimas naudojant dirbtinio intelekto metodus (pvz: 8khz upsamplingas į 44khz, telefoninių pokalbių kokybės pagerinimas).	Reconstruction / improvement of speech quality using artificial intelligence methods (eg: 8khz upsampling to 44khz, improvement of telephone call quality).	
25	Aistis Raudys	Kalbos/raktažodžių atpažinimas mikrokompiuteriuose ir mikrokontroleriuose (kaip kad ok google, alexa).	Language / keyword recognition on microcomputers and microcontrollers (like ok google, alexa).	
26	Aistis Raudys	Kalbą generuojantys neuroniniai tinklai lietuvių kalbai, DeepVoice, WaveNet ir kt. (Neural networks for text to speech application)	Language generating neural networks for Lithuanian, DeepVoice, WaveNet and others. (Neural networks for text to speech application)	Pritaikyti gilius neuroninius tinklus lietuvių kalbos balso generavimui.
27	Aistis Raudys	Kalbą atpažįstantys neuroniniai tinklai lietuvių kalbai, DeepSpeech ir kt.	Speech-recognizing neural networks for the Lithuanian language, DeepSpeech and others.	Pritaikyti gilius neuroninius tinklus lietuvių kalbos balso atpažinimui.
28	Aistis Raudys	FPGA panaudojimas gilaus mokymo/finansų optimizavimo/garso atpažinimo uždaviniuose	Use of FPGA in deep learning / financial optimization / sound recognition tasks	
29	Aistis Raudys	Dirbtinis intelektas robotikoje: optimalus variklio valdymas naudojant neuroninius tinklus.	Artificial intelligence in robotics: optimal motor control using neural networks.	
30	Aistis Raudys	K-kaimynų klasifikavimo algoritmo realizavimas ir tobulinamas GPU/CUDA aplinkoje naudojant PyTorch	Implementation and improvement of K-neighbour classification algorithm in GPU / CUDA environment using PyTorch	
31	Aistis Raudys	Lietuvių kalbos rašybos taisymas naudojant gilaus mokymo metodus.	Correction of Lithuanian spelling using deep teaching methods.	
32	Aistis Raudys	Skraidančių mirko objektų sekimas	Tracking flying micro objects for destruction.	Naudojant OpenCV optiniame sraute surasti skraidantį objektą ir jį sunaikinti.
33	Aistis Raudys	Neuroninių tinklų panaudojimas roboto balansavimui	Development of neural network based self-balancing robot	Naudojant giroskopus sukurti balansuojantį robotą panaudojant neuroninius tinklus ir palyginant su tradiciniais metodais. Šiam tikslui panaudoti coppeliarobotics.com, vėliau realybėje
34	Aistis Raudys	Dirbtinio intelekto metodų panaudojimas autonomiam besimokančiam robotui	Autonomous Self-Learning Robot	Panaudoti roboto sensorių duomenis mokymuisi. Pradžioje virtualioje aplinkoje coppeliarobotics.com, vėliau realybėje.
35	Aistis Raudys	Neuroninių tinklų panaudojimas roboto rankos judėjimo kelio optimizavimui	Use of neural networks to optimize the path of robot arm movement	Palyginti neuroninių tinklų ir kitus metodus roboto rankos kelio optimizavimui. Pradžioje coppeliarobotics.com aplinkoje vėliau realybėje.
36	Aistis Raudys	Optimizavimo algoritmų pritaikymas prekybinėms sistemoms kriterijų pakeičiant pelnu arba Sharpe santykiu	Application of optimization algorithms for trading systems, replacing the criterion with a profit or Sharpe ratio	Kainų matrica pritaikyta simuliuoti profit arba sharpe kriterijų
37	Algirdas Lančinskas	Optimizavimo algoritmų lyginamoji analizė	Benchmarking of optimization algorithms	
38	Algirdas Lančinskas	Jūrinių konteinerių krovos optimizavimas	Optimization of marine container loading	
39	Algirdas Lančinskas	Paslaugas teikiančių objektų vietų parinkimas	Optimization of facility locations	

40	Audronė Lupeikienė	Struktūrinio sudėtingumo mažinimas kuriant didelis sistemas	Reduction of structural complexity in designing very big systems	Example of very big system - system for an entrprise in Industry 5.0.
41	Darius Plikynas	Propagandos aptikimas ir klasifikavimas socialinės medijos pranešimuose naudojant gilios mokymo bei semantinės analizės	Detection and classification of propaganda in social media news channels, using deep learning and semantic analysis methods	
42	Gintaras Skersys	Turbo kodų tyrimas	Study of Turbo Codes	
43	Gintautas Dzemyda	Geometrinio Daugiamatčių skalių metodo, skirto daugiamatčiams duomenims vizualizuoti, tyrimas	Investigation of the Geometric multidimensional scaling for multidimensional data visualization	Eksperimentinis metodo tyrimas, siekiant kuo geresnės duomenų vizualizavimo kokybės
44	Gintautas Dzemyda	Kepenų radiologinių vaizdų analizė siekiant nustatyti vėžio pažeistą vietą	Analysis of radiological images of the liver to identify the site of cancer	Vaizdų analizės uždavinys
45	Gintautas Dzemyda	Duomenų klasifikatorių efektyvumo priklausomybės nuo besikeičiančios mokymo imties tyrimas	A study of the dependence of the efficiency of data classifiers on a changing training sample	Reikės patyrinėti kelis klasikinius duomenų klasifikatorius
46	Gintautas Dzemyda	Geometrinio daugiamatčių skalių metodo daugiaekstremiškumo tyrimas	A study of multiextremality of the geometric multidimensional scaling	Eksperimentinis metodo tyrimas, siekiant įvertinti optimizuojamos funkcijos daugiaekstremiškumą.
47	Igor Katin	Building a big data platform for business	Building a big data platform for business	Platformos didelių duomenų apdorojimui, tendencijų paieškai, produktų ar konkurentų analizei kūrimas. Development of a platform for processing big data, finding trends, analyzing products or competitors.
48	Igor Katin	Duomenų saugojimas blockchain	Data storage in the blockchain	Dažnai versle reikia užtikrinti, kad tam tikri duomenys būtų nuosekliai saugomi, jų negalima būtų keisti, klastoti ar pažeisti. Dėl to vienas iš galimų sprendimų juo saugoti naudojant blockchain technologiją.
49	Igoris Belovas	Tolygiųjų skaičių laipsnių savybių b -tainėje skaičiavimo sistemoje tyrimas	Repdigits base b powers properties study	Tolygieji skaičiai – tai natūralieji skaičiai, sudaryti iš vienodų skaitmenų, t. y., jei b yra skaičiavimo sistemos pagrindas, tai tolygūs skaičius yra išreiškiamas formule $a(b^n - 1) / (b - 1)$, kur $0 < a < b$ yra pasikartojantis skaitmuo, n yra pasikartojimų skaičius. Tolygiųjų skaičių sveikųjų laipsnių formulių b -tainėje skaičiavimo sistemoje tyrimas pareikalaus ir kompiuterinių eksperimentų, ir analizinių išvedimų.
50	Igoris Belovas	Rymano dzeta funkcijos skaičiavimo algoritmų CPU ir GPU lygiagrečio lyginamoji analizė	Comparative analysis of CPU and GPU parallelizations of the Riemann zeta function calculation algorithms	Rymano dzeta funkcija yra vienas iš svarbiausių analizinės skaičių teorijos įrankių. Su Rymano dzeta funkcijos reikšmių pasiskirstymo savybėmis yra glaudžiai susijusi Rymano hipotezė – viena iš aktualiausių šiuolaikinės matematikos neišspręstų problemų. Darbo tikslas – susipažinti su kai kuriais šiuolaikiniais dzeta funkcijos reikšmių skaičiavimo algoritmais, realizuoti juos programiškai ir atlikti realizacijų CPU ir GPU lygiagrečio efektyvumo palyginimą.

51	Igoris Belovas	Dokumentų duomenų parengimas natūralios kalbos apdorojimo uždaviniams	Preparing document data for natural language processing tasks	Įmonių veikloje versle svarbi informacija fiksuojama dokumentuose (kad atitiktų įstatymų reikalavimus ir liktų galimybė atsekti istoriją). Kai veikloje dalyvauja kelios įmonės, jos tais dokumentais apsieičia (sutartys, sąskaitos). Jau kuris laikas apsieitimui nebėra naudojami popieriniai dokumentai, bet jais apsieičiama elektroniniais būdais (dažniausiai el. paštu ir per savitarnos svetaines). Gavus dokumentą (pdf, Word ar Excel formato), reikia iš jo ištraukti tekstą, o po to iš teksto – dokumente esančią svarbią informaciją (įmonių duomenis, datas, sumas). Informacijos ištraukimui naudojami natūralios kalbos apdorojimo metodai, tokie kaip įvardintų esybių atpažinimas (named entity recognition), tačiau jiems reikalingas teksto formatas, o ne dokumentas specializuotų programų formatu. Taigi, duomenų parengimas yra vienas svarbiausių etapų dokumentuose esančios informacijos apdorojime, įgalinantis atlikti tolesnius žingsnius. Darbo tikslas – surasti, įvertinti ir atrinkti automatizuotus įrankius ar bibliotekas, ištraukiančius tekstą ir pagalbinę informaciją apie tekstą (teksto blokus, jų išdėstymą) tam, kad vėlesniame etape būtų galima atlikti natūralios kalbos apdorojimo uždavotus. Pasiūlytas sprendimas turi mokėti dirbti ne tik su angliškais, bet ir lietuviškais tektais.
52	Irmantas Radavičius	Adaptyvūs rikiavimo algoritmai	Adaptive sorting algorithms	Tema rezervuota studentui DZ
53	Jolita Bernatavičienė	Vaizdo rezoliucijos didinimas naudojant gilaus mokymosi algoritmus	Image resolution enhancement using deep learning algorithms	Tiriant medicininius vaizdus dažnai susiduriama su skirtingos kokybės vaizdais, kurie skiriasi tiek raiška, tiek dydžiu. Tyrimo tikslas - ištirti ir palyginti giliojo mokymo metodus taikomus vaizdų rezoliucijos didinimui ir įvertinti jų efektyvumą sprendžiant klasifikavimo uždavinius.
54	Julius Žilinskas	Pusiau teigiamai apibrėžtas optimizavimas	Semidefinite programming	
55	Julius Žilinskas	Kalnų kelionių technikos trasų įveikimo modeliavimas ir optimizavimas	Imitation and optimization of accomplishment of mountaineering routes	
56	Karolis Petrauskas	TLA+ specifikacijų išskyrimas iš programinio kodo BEAM virtualiai mašinai.	Extracting TLA+ specifications out of a program for a BEAM virtual machine.	Formal languages are widely used to design and verify distributed systems and algorithms. One of such languages is TLA+ (https://lampion.azurewebsites.net/ta/book-02-08-08.pdf). Specifications are designed at a high level using TLA+, thus there is a gap between an implementation and the spec. Erlang/OTP is a programming language designed for building distributed systems and is used to build online game servers, banking applications, systems integrations and other. This language is compiled to several intermediate formats, that can be used to extract program structure. The aim of this topic is to develop a translation from a BEAM based language (Erlang/Elixir/LFE or other) to a TLA+ specification. The extracted specification then could be used to prove a refinement mapping between the abstract specification and the extracted one thus narrowing the gap between the spec and the implementation. One of the attempts to formalize the semantics of the intermediate format can be found here < https://dl.acm.org/doi/10.1145/3406085.3409008 >. See also references in the related work section.

57	Karolis Petrauskas	Formalių specifikacijų taikymas projektuojant paskirstytas sistemas.	Applying formal specifications to design distributed systems.	A student should develop a formal specification in the TLA+ language (https://lampport.azurewebsites.net/tla/book-02-08-08.pdf) for a chosen distributed algorithm (e.g. Kademlia) or a system (e.g. Kafka replication, Riak KV, RabbitMQ distribution, LevelDB). Another specification language can be chosen, if it will be more appropriate for the specific task. The specification will be used to investigate properties of the selected system or the algorithm. If any flaws are identified, solutions should be provided and validated formally. In the case of notable results, a student is encouraged to present them in a scientific conference.
58	Linas Laibinis	Paskirstytų, gedimams atsparių, dinamiškai konfiguruojamų programų sistemų formalus modeliavimas ir verifikavimas	Formal modelling and verification of distributed, fault tolerant, dynamically reconfigurable software-based systems	Darbo metu yra sukuriama pasirinktos paskirstytos programinės sistemos formalus modelis (remiantis jos aprašymu ar reikalavimų sąrašu). Sukurtas modelis yra analizuojamas ir verifikuojamas, naudojantis automatinio įrodymo ar modelių patikrinimo įrankiais. Atsparumas gedimams ir galimybė dinamiškai konfiguruoti tokias sistemas priklausomai nuo besikeičiančios aplinkos yra esminės savybės, į kurias fokusuojamasi verifikavimo metu.
59	Linas Laibinis	Sistemų prototipų kūrimas ir skaitinis įvertinimas naudojant diskretinių įvykių simuliacijos metodus	System prototyping and quantitative assessment by discrete event simulation techniques	Diskretinių įvykių simuliacijos aplinkos leidžia sukurti būsimos paskirstytos sistemos prototipą, aprašyti joje vykstančius įvykius, komunikacijas tarp komponentų, galimus gedimus ir modeliuoti laiko pauzes. Darbo metu sukurtos sistemos prototipo simuliacijos duotą galimybę skaitiškai įvertinti sistemos darbo charakteristikas (efektyvumą, patikimumą ir t.t.) bei tuo pačiu palyginti skirtingas tokių sistemų konfigūracijas.
60	Linas Laibinis	Paskirstytų sistemų modeliavimas ir verifikavimas remiantis statistinio modelių patikrinimo metodais	Modelling and verification of distributed systems using statistical model checking methods	Darbo metu sukurtas paskirstytos sistemos modelis (prototipas) yra analizuojamas statistinio modelių patikrinimo metodais (pvz., naudojantis automatizuota aplinka Uppaal). Analizės rezultatas yra statistinis ar tikimybinis suformuluotų sistemos savybių (efektyvumo, patikimumo ir t.t.) ar jos skaitinių apribojimų įvertinimas bei grafinis pateikimas.
61	Linas Laibinis	Paskirstytų programinių sistemų kūrimas ir skaitinis įvertinimas naudojantis komunikavimo šablonais	Design and quantitative evaluation of distributed software-based systems using communication patterns	Darbo metu yra skaitiškai įvertinami paskirstytų sistemų kūrime naudojami komunikavimo šablonai ir jų kompozicijos. Identifikuojami sistemos ar šablonų parametrai, tiesiogiai įtakojantys sistemos darbą, ir jų optimalumas yra analizuojamas naudojantis statistinio ar įprasto modelių patikrinimo metodais.
62	Linas Laibinis	Programų sistemų ir algoritmų verifikavimas automatizuotose teoremų įrodymo aplinkose	Verification of software-based systems and algorithms in automated theorem proving environments	Darbe apžvelgiamos plačiai naudojamos automatizuotos teoremų įrodymo ir sistemų verifikavimo aplinkos. Konkretios programų sistemos ar algoritmo verifikavimas yra atliekamas pasirinktoje automatizuoto teoremų įrodymo aplinkoje.
63	Linas Litvinas	Dirbtiniai neuroniniai tinklai laiko eilutėms	Artificial neural networks for time series	
64	Linas Litvinas	Biojutiklio daugiakriterinis optimizavimas	Multiobjective optimization of biosensor	
65	Linas Petkevičius	Edvinis ir laikinis aptikimas palydovinių vaizdų stebėjime	Spatial-temporal detection in satellite imaging monitoring	
66	Mindaugas Bloznelis	Epidemijos plitimas tinkle	Epidemic spread in networks	Epidemijos plitimo ir kontrolės (įvairių stabdymo priemonių taikymas) modeliavimas dideliuose socialiniuose tinkluose.
67	Mindaugas Bloznelis	Tinklo bendruomenių paieška	Network community detection.	Tinklo bendruomenių atskleidimas. Algoritmų veikimas realiuose tinkluose, algoritmų analizė sintetiniuose tinkluose. (Spektrinis algoritmas, atsitiktinio klaidžiojimo algoritmas)
68	Olga Kurasova	Dirbtinių neuroninių tinklų modelių paaiškinamumo ir interpretuojamumo tyrimai	Explainability and interpretability of artificial neural network models	

69	Olga Kurasova	Pusiau prižiūrimo mokymosi metodai medicininių vaizdų analizės uždaviniams spręsti	Semi-supervised learning methods for medical image analysis	
70	Olga Kurasova	Automatinio mašininio mokymosi ir neuroninio tinklo architektūros paieškos tyrimai	Research on automated machine learning (AutoML) and neural network architecture search (NAS)	
71	Pijus Kasparaitis	Neuroninių tinklų spartinimas	Acceleration of neural networks	
72	Pijus Kasparaitis	Lietuvių kalbos sintezė naudojant neuroninius tinklus	Text-to-speech synthesis of Lithuanian based on neural networks	
73	Rimantas Vaicekuskas	Pamaininio darbo tvarkaraščio optimizavimo algoritmų lygiagretus įgyvendinimas	Parallel implementation of shift scheduling optimization algorithms	

74	Rimantas Vaicekuskas	Pamaininio darbo tvarkaraščio optimizavimo algoritmai	Shift sheduling optimization algorithms	
75	Saulius Gražulis	Gardelės parametrų išgavimas iš žmogui skirtų tekstų ir COD apžvalginės DB kūrimas	Extracting Unit cell parameters of crystals from human-readable scientific papers	Deja, kol kas ne visuose straipsniuose paskelbtos kristalų struktūros yra atvirai aprašytos viešai prieinamais duomenimis. Daugeliui tokių „paslėptų“ struktūrų straipsniuose nurodomi pagrindiniai kristalo parametrai – gardelės konstantos, simetrijos grupė, molekulės cheminė formulė ir pavadinimas. Tai, nors ir nepilni, bet vis vien labai vertingi duomenys, nes leidžia preliminariai identifikuoti kristalinę medžiagą. Kursiniame darbe bus pasiūlyta surinkti visų publikuotų kristalų struktūrų bibliografijas ir pagal atvirai prieinamus straipsnių tekstus arba santraukas (abstracts), naudojant paprastas Perl reguliarias išraiškas (o vėliau gal būt ir sudėtingesnę lingvistinę analizę) automatiškai išgauti kiekviename straipsnyje publikuotų medžiagų kristalų parametrus.
76	Saulius Gražulis	Informacijos išgavimas iš mokslinių straipsnių	Extracting scientific facts from research papers	Pastaruoju metu publikuojamų mokslinių straipsnių skaičius auga eksponentiškai [1]. Tikėtina, kad šiuo metu straipsnių paskelbiama tiek daug, jog net ekspertai negali fiziškai perskaityti visų straipsnių savo darbų tematika, jau nekalbant apie gretimų tyrimų krypčių straipsnius. Mokslo žinios tampa išsklaidytos po daugelį žmogui sunkiai aprėpiamų šaltinių, jas vis sunkiau susisteminti, o tarpdisciplininiai tyrimai – vienas svarbių inovacijos šaltinių – dėl šių priežasčių neatskleidžia viso savo potencialo. Turėtų būti įmanoma sukurti apmokytus DNT ir/arba kitas kompiuterines sistemas, gebančias iš žmogui skirtu tekstu (mokslinio straipsnio, patento, monografijos, disertacijos, preprinto) išgauti faktinę mokslinę informaciją, pvz. medžiagos formulę, kristalo ir molekulės struktūros aprašymą, molekulės chemines ir fizines savybes, tų savybių epistemologinį statusą (matavimo rezultatas, teorinis skaičiavimas), ir pateikti šią informaciją formalizuotu, automatiniame apdorojimui tinkamu pavidalu, pavyzdžiui, formaliai specifiikuotų CIF, CML a SDF failų pavidalu, reliacinių duomenų lentelių pavidalu, RDF failų pavidalu naudojant kurią nors žiniatinklio ontologiją ar pan. Darbe bus pasiūlyta, pasitelkiant įvairias teksto analizės priemones (teksto segmentatorius, simbolių atpažinimo programas, tokias kaip 'tesseract', dirbtinius neuroninius tinklus, mašinių mokymo programas) išgauti iš mokslinių straipsnių vertingus faktus ir apjungti juos į sistemingą duomenų bazę greitai ir išsamiai paieškai.

77	Saulius Gražulis	Naršyklės įskiepis kristalografinėi informacijai rinkti	Browser plugin for collection of scientific information	Šiuolaikiniame mokslo pasaulyje labai didelė informacijos dalis yra išbarstyta po daugybę teksto puslapių (mokslinių straipsnių, disertacijų, monografijų), ir jų sisteminga paieška kompiuterinėmis priemonėmis yra labai apsunkinta. Deja, automatinė tekstų analizė ne visada sugeba teisingai išgauti reikiamus duomenis, ir tenka įdėti daug žmogaus rankų darbo, įvedant informaciją į kompiuterizuotas duomenų bazines. Siūloma sukurti įrankį, kuris palengvintų tyrėjams informacijos surinkimą ir kaupimą vienoje gerai struktūruotoje duomenų bazėje. Siūlomas darbas būtų skirtas įskiepiai, kuris padeda rinkti kristalografinę ir cheminę informaciją, sukurti. Tai leistų šimtams tyrėjų vienu metu visame pasaulyje rinkti informaciją, dėti ją į bendrą duomenų bazę ir po to bendrai naudotis surinktais duomenimis. Toks modelis jau neblogai pasiteisino renkant citavimo ir bibliografinius duomenis; naudinga būtų pritaikyti jį kitoms mokslo duomenų rūšims.
78	Saulius Gražulis	Mažų molekulių, randomų COD DB, susiejimas su PDB	Linking the Crystallography Open Database (COD) with the Protein Data Bank (PDB)	COD duomenų bazėje sukaupta virš 450000 įrašų apie mažų molekulių kristalų struktūras. Apie 2/3 jų yra organinės molekulės. Dalis tų molekulių yra ligandai, kofaktoriai ar vaistai, galintys jungtis su baltymų molekulėmis, o šių baltymų struktūras gali būti randamos PDB duomenų bazėje. Darbo metu bus pasiūlyta surasti tas molekules COD duomenų bazėje, kurios yra chemiškai tokios pačios, kaip ir PDB archyve rasti ligandai, ir susieti COD su PDB, naudojant PDB papildomos informacijos pateikimo JSON schemą.
79	Saulius Gražulis	Saugumo auditas COD ir panašiose interneto svetainėse	Security audit in the Crystallography Open Database and related Web sites.	Žiniatinklis (World Wide Web) buvo sukurtas tam, kad įgalintų mokslininkus greitai ir patogiai keistis tyrimų rezultatais bei kita informacija. Naujos žiniatinklio priemonės leidžia ne tik parsisiųsti informaciją iš nutolusios svetainės, bet ir interaktyviai pateikti informaciją šioms svetainėms arba paleisti procesus nutolusiuose serveriuose. Deja, šios galimybės atveria eilę saugumo spragų – serveriai gali būti panaudojami ne taip ir ne tam, kam jie buvo sukurti, ir tai gali atsitikti be svetainės kūrėjo žinios. Šiame darbe, naudojant įvairius svetainių audito įrankius (e.g. Kali Linux), bus pasiūlyta patikrinti įvairių mokslo duomenims skirtų interneto svetainių (COD, RestfulDB, Web Scriptlets) saugumą; kitaip tariant, pabandyti „nulauzti“;) programinę įrangą, kurios pagrindu veikia šios svetainės, ir pasiūlyti, kaip galima nuo išnagrinėtų grėsmių apsisaugoti.
80	Saulius Gražulis	Teisingumo įrodymų palaikymas daugiaparaigminėje programavimo kalboje Starta	Integrating correctness proofs into multi-paradigm programming language Starta.	Šiuolaikiniai aukšto lygio programavimo kalbų kompiliatoriai stebimai padidino programuotojų darbo efektyvumą, palyginus su autokodais ar assemblerio kalbomis. Deja, nuo to laiko nauja "magiška kulka", leidžianti pasiekti dar didesnių darbo našumų ar aukštesnės programų kokybės, taip ir nėra atrasta. Vienas iš stebėtinų dabartinių kompiliatorių trūkumų – visiškai atotrūkis tarp teorinių darbų, skirtų programų teisingumo įrodymams, ir praktiniam programavimui naudojamų programavimo kalbų bei kompiliatorių. Darbo eigoje bus siūloma realizuoti automatinę sistemą teiginiams apie programą įrodyti, integruotą į aukšto lygio programavimo kalbos kompiliatorių. Pirmame etape planuojama naudoti specialiai tam tikslui sukurtą programavimo kalbą ir kompiliatorių, kuriuos, reikalui esant, galima keisti, siekiant supaprastinti įrodymų išvedimą; tačiau sistema turėtų būti pakankamai moduliarizuota, kad įrodymų posistemę galima būtų perkelti į kitų kalbų (Java, C, C++, C#, Perl, Python, etc.) kompiliatorius.

81	Saulius Gražulis	Mažų molekulių kristalų kontaktų paviršiai	Contact surfaces in small molecule crystals	Siūloma peržvelgti visas COD organinių kristalų struktūras, visų pirma tas, kuriuose yra vaistinių medžiagų molekules ar j jas panašios molekules. Surasti šių molekulių kontaktus su *savo pačių* kristalais, aprašyti šių kontaktų paviršius. Surasti tas molekules, kurių kompleksai su baltymais patalpinti PDB archyve. Palyginti mažos molekules kristalo ir baltymo kontaktinius paviršius; nustatyti, ar pagal šių paviršių panašumą galima prognozuoti susirišimą su baltymu.
82	Saulius Gražulis	Paskirstytos, patikimos ir atsparios trikdžiams bei padalinimams COD duomenų bazės kūrimas.	Creating reliable and fault-tolerant server system for Crystallography Open Database	Šiuo metu COD duomenų bazė, organizuota kaip centrinis (angl. "master") serveris, kurio duomenis patikimumo dėlei replikuoja visa eilė antrinių (angl. "mirror") kompiuterių. Tokia sistema, deja, neužtikrina nenutrūkstamo sistemos darbo, centriniam serveriui išėjus iš rikiuotės ar nutrūkus ryšiu tarp centrinio serverio ir Interneto. Darbo metu bus siūloma realizuoti paskirstytą, lygių serverių mainais (angl. "peer-to-peer") paremtą sistemą, atsparią sistemos padalinimui, užtikrinančią duomenų neprieštarumą (consistency) ir minimizuojančią sistemos prastovas. Nors CAP teorema teigia, kad visų trijų tikslų (neprieštarumo, prieinamumo ir atsparumo padalinimams) vienu metu pasiekti neįmanoma, bus bandoma surasti inžinerinius sprendimus, leidžiančius minimizuoti prastovas (prastovas, duomenų praradimą ir pan.), atsiradusius dėl to, kad COD bus realizuota kaip paskirstyta duomenų bazė. Darbo metu reikės išnagrinėti įvairius galimus sistemos variantus, įvertinant įvairius kompromisus (prieinamumas/neprieštarumas, prieinamumas/patikimumas ir pan.).
83	Saulius Gražulis	Didelės apimties duomenų archyvavimas paskirstytoje duomenų saugykloje	Archiving high volume scientific data in a distributed peer-to-peer repository	Naujausias IUCr (Tarptautinės kristalografų sąjungos, angl. International Union of Crystallography) rekomendacijos siūlo archyvuoti visus pradinius duomenis, panaudotus struktūros nustatymui, įskaitant difrakcijos (išsklaidytų Rentgeno spindulių) vaizdus, užregistruotus monokristalinių difraktometrų. Šios rekomendacijos įgyvendinimas kelia naujus iššūkius -- bus reikalingas gerokai didesnis pastovios atminties (diskų, juostų) kiekis, negu naudotas iki šiol, ir duomenys turi būti prieinami bent jau ateinančius dešimtmečius, t.y. pergyventi kelias kompiuterinės įrangos kartas. Visa tai susiję su papildomomis sąnaudomis ir duomenų laikymo kaštais. Vienas iš galimų šių problemų sprendimo būdų -- panaudoti paskirstytą, daugelio institucijų ir/arba individų palaikomą duomenų archyvavimo sistemą, turinčią pakankamą duomenų perteklumą, užtikrinantį patikimą sistemos darbą ilgą laiką. Tokios sistemos prototipas buvo išbandytas 2018 m. studentų bakalauro darbų metu. Darbo metu bus siūloma realizuoti veikiančią, mokslininkams tinkamą sistemą Tahoe-LAFS ir/arba IPFS pagrindu.

84	Saulius Gražulis	Vidutinių trimačių simetrijos grupių apskaičiavimas iš keturmačių moduluotų struktūrų simetrijos operatorių.	Computing average space groups from 4D and higher-dimensional space group operators	Pastaruoju metu daugėja informacijos apie medžiagos būvį, kuris, nors ir turi daugumą kristalo savybių (pvz., sklaido Rentgeno spindulius siaurais koncertuotais atspindžiais), nėra tikras kristalas, nes negali turėti periodinės gardelės, suderinamos su stebima objekto ar sklaidymo vaizdo simetrija, tokią kaip penkto laipsnio simetrijos ašis. Tai – kvazikristalai (http://en.wikipedia.org/wiki/Quasicrystals) ir (nebendramatės) moduluotos struktūros. Šioms struktūroms aprašyti kuriamas matematinis aparatas, panaudojantis simetrijos grupių teoriją. Pasirodo, kad neperiodines trimates struktūras galima aprašyti kaip periodinių struktūrų daugiamatėse erdvėse pjūvius. Pavyzdžiui, kai kurias moduluotas struktūras galima nagrinėti kaip periodinių 4-mačių gardelių pjūvius. Perėjimas į aukštesnių matavimų erdves leidžia panaudoti jau žinomą erdvių simetrijos grupių mat. aparatą, ir kompaktiškai aprašyti neperiodines struktūras. Darbo metu bus siūloma sukurti programinę įrangą, kuri tikrintų keturmačių simetrijos grupių aprašymus, pagal šiuos aprašymus sukurtų vidutinius nmoduluotos trimatės simetrijos grupės aprašus, ir integruoti šiuos algoritmus į duomenų bazę COD, kad būtų galima efektyviai kaupti ir tvarkyti neperiodinių medžiagos pavyzdžių aprašymus.
85	Saulius Gražulis	BOINC serverio ir klientų parengimas statistiniams skaičiavimams ir jų pritraukymas COD duomenų bazės analizei	Creating BOINC clients and setting up BOINC server for large volume data analysis in the COD	Statistiniai skaičiavimai, paremti Bajeso statistikos principais, duoda universalią ir koherentišką skaičiavimo metodiką, bet reikalauja itin daug skaičiavimo resursų. Vienas iš būdų tokius resursus surinkti – panaudoti masinį paralelizmą „savanorių skaičiuotojų“ (angl. "volunteer computing") pateiktuose kompiuteriuose. Šiuo principu yra paremta Berklio universiteto BOINC sistema. Darbo metu bus siūloma: a) paleisti BOINC sistemos serverį; b) parašyti paprasčiausius BOINC klientus; c) parašyti klientus, skirtus COD atstumų ir jungčių parametrų tikimybių pasiskirstymų pasiskirstymų radimui ir atnaujinimui, naudojant Bajeso statistikos metodus, ir skaičiavimų organizavimas.
86	Saulius Gražulis	C kalba parašytų programų transliavimas į Java JVM baitų kodą.	Translating C into Java bytecode	Mūsų turimas CIF sintaksinis analizatorius parašytas C programavimo kalba. C kodas yra efektyvus, perkeliamas į daugelį OS ir skaičiavimo platformų, leidžia sukurti sąsajas (angl. bindings) daugeliui programavimo kalbų: Perl'ui, Pitonui, C++ ir kitoms. Java kalba parašytose programose būtų galima naudoti šią biblioteką JNI sąsajos pagalba, tačiau toks sprendimas Java „pasaulyje“ turi savų trūkumų: kai kurios platformos, pvz. Android/Dalvik arba HTML5/JS, nepalaiko JNI sąsajų, arba šių sąsajų naudojimas per daug sudėtingas, o Java taikomosioms programoms naudojamas klases tenka perkompiliuoti kiekvienai procesoriaus architektūrai vietoj to, kad panaudoti Java baitų kodą ir, jei reikia, turimus JIT optimizuojančius kompiliatorius. Idėja yra surasti transliatorių, kuris transliuotų C kalbos programas į Java virtualios mašinos kodą. Tyrimo eigoje reikės nustatyti ir aprašyti, kaip suderinti C standartą atitinkančias vykdymo aplinkas ir Java virtualias mašinas. Jei reikės, bus modifikuojamas egzistuojantis C kompiliatorius (tcc, lcc arba gcc), nutaikant jį JVM kodo generavimui. Tikslas yra sukompiliuoti turimą CIF sintaksinio analizatoriaus tekstą į gryną JVM baitų kodą.
87	Saulius Gražulis	Taisyklėmis paremtos ekspertinės sistemos sukūrimas COD esančių atomų cheminiams tipams nustatyti	Rule-based expert system for deriving chemical properties from crystal structures	Siūloma paversti molekulių aprašymus Prolog duomenų baze ir užkoduoti chemines žinias taip, kad būtų galima automatiškai gauti išvadas apie molekulių ir jose sujungtų atomų savybes (hibridizaciją, geometriją, dalinius ir formalius krūvius, reakcingumą).
88	Saulius Gražulis	dRel programavimo kalbos realizavimas	Implementing dRel programming language	Siūloma įgyvendinti dRel programavimo kalbos interpretatorių CIF failuose įterptų funkcijų interpretavimui.

89	Saulius Gražulis	Molekulinių mazgų ir sankabų analizė COD duomenų bazėje	Finding molecular knots in the Crystallography Open Database	Molekulių kovalentiniai ryšiai gali būti traktuojami kaip tvirti strypeliai, jungiantys atomus (briaunos, jungiančios grafo viršūnes). Jei tokiame molekuliniame grafe yra ciklai, jie gali sudaryti mazgus arba sankabas (pavyzdžiui, sukabintus žiedus, Boromėjaus žiedus ar sudėtingesnes topologines struktūras). Kol kas nėra efektyvių programų, kurios leistų aptikti ir suklasifikuoti tokias struktūras, aptiktas molekuliniuose medžiagos kristaluose. Darbo metu bus siūloma sukurti programas, generuojančias 3D grafo aprašymą pagal kristalografines informacijos CIF failus su kristalų struktūromis, pritaikyti šias programas Crystallography Open Database (COD, https://www.crystallography.net/) duomenų bazei, sugeneruoti iš gauto molekulinio grafo mazgo aprašymą, ir pagal mazgo aprašymą pritaikyti įvairius mazgo invariantus skaičiuojančius algoritmus. Darbo pabaigoje sukursime aptiktų mazgų duomenų bazę.
90	Saulius Gražulis	Kvantinės mechanikos metodų pritaikymas kristalografinių duomenų validavimui	Validating experimental crystallographic data using first-principles quantum mechanics	Darbe bus siūloma pritaikyti kelias atviro kodo kvantinės mechanikos programas (Abinit, Quantum Espresso) atviros kristalografines duomenų bazės Crystallography Open Database (COD, https://www.crystallography.net/) duomenims patikrinti ir galimoms eksperimento ar duomenų tvarkymo klaidoms aptikti.
91	Saulius Gražulis	DNT ir mašinių mokymo pritaikymas kristalų savybėms prognozuoti	Applying ANN and machine learning for crystal property prediction	Nauji Dirbtinių neuroninių tinklų (DNT) ir mašinių mokymo algoritmai leidžia aptikti dėsningumus ir atpažinti bruožus didelėse duomenų masyvuose, kurie seniau buvo neprieinami išsamiai analizei. Atviroje kristalografinėje duomenų bazėje Crystallography Open Database (COD, https://www.crystallography.net/) yra sukaupia virš 450 tūkst įrašų apie kristalų struktūras, o susieti straipsniai talpina informaciją apie šių kristalų savybes. Darbe bus siūloma panaudoti COD DB duomenų imtį DNT ar mašinių mokymo sistemai apmokyti, siekiant prognozuoti įvairias kristalo savybes (pvz. kristalo elementarus narvelio tūrį, lydymosi temperatūrą ir pan.). Gauti tinklai gali būti naudojami COD ir kitų publikuotų duomenų validavimui, naujų kristalų savybių nustatymui.
92	Saulius Gražulis	Interaktyvios COD recenzavimo svetainės kūrimas	Interactive collaboration platform for the Crystallography Open Database	Interaktyvios programų archyvų svetainės, tokios kaip GitLab, GitGub ar BitBucket, gerai užsirekomendavo programinės įrangos kūrimo procese. Darbe bus siūloma pritaikyti analogiškus programų kūrimo įrankius mokslo duomenims tvarkyti. Atvira kristalografinė duomenų bazė Crystallography Open Database (COD, https://www.crystallography.net/) sėkmingai naudoja Subversijos versijų kontrolės sistemą duomenims versijuoti ir kaupti. Natūralu panaudoti sistemą, analogišką aukščiau minėtoms priemonėms, pvz. Redmine, kuri leistų kristalografams aptarinėti ir taisyti kristalų struktūrų duomenis, panašiai kaip programuotojai aptarinėja ir taiso programas. Darbo metu reikės sukurti COD ir Redmine svetainių sąsają ir išbandyti kristalografinių duomenų valdymo srautą realaus gyvenimo sąlygomis.
93	Saulius Gražulis	COD P1 narvelių skaičiavimas	Computing chemical structures in P1 cell for the Crystallography Open Database	Kristalografiniai duomenų failai pateikia minimalų parametrų rinkinį, būtina kristalo struktūrai atstatyti naudojant kristalo simetrijos grupės operatorius. Toks aprašymas dažnai nepateikia visos chemikų tikslams reikalingos informacijos; pavyzdžiui, kristale molekulė gali būti aprašyta, nurodant tik jos dalį, o likusi dalis turi būti suskaičiuota, pritaikant informaciją apie simetriją. Toks atvaizdavimo būdas chemikams yra nepatogus ir sukelia sunkumų tolimesnėje duomenų analizėje. Darbe bus siūloma pagaminti iš COD duomenų bazę su pilna informacija apie kristalų elementarius narvelius, t.y. sugeneruojant visus atomus, kurių reikia, norint aprašyti kristalo struktūrą tik elementarių transliacijų pagalba (taip vadinamus P1 narvelius). Šioje duomenų bazėje galima atlikti tolimesnę analizę, pavyzdžiui optinių izomerų paiešką.

94	Saulius Gražulis	Srautinio CIF parserio (sintaksinio analizatoriaus) sukūrimas	Creating a streaming CIF parser	Dabar paplitę Crystallography Interchange File (CIF) failų formato sintaksiniai analizatoriai veikia, naudodami DOM modelį, t.y. visas failas perskaitomas į atmintį ir tada apdorojamas. Toks metodas visiškai netinka dideliems konkatenuotiems CIF srautams skaityti, pvz. visiems PDB arba COD duomenų bazių įrašams Unix konvejerijoje apdoroti. Siūloma sukurti jau esamo cod-tools sintaksinio analizatoriaus pagrindu (C/Bison) srautinį analizatorių, t.y. tokį analizatorių, kuris perskaitytų ir grąžintų failo informaciją po vieną įrašą, ir leistų kreiptis į save daug kartų, prašysiant sintaksinę analizę nuo tos failo vietos, kurioje buvo sustojęs.
95	Saulius Gražulis	Kokybiško CIF parserio (sintaksinio analizatoriaus) sukūrimas Java ir C# (.NET) platformoms	High quality CIF parser for Java and .NET platforms	CIF (Crystallography Interchange Format) yra paplitęs struktūruotas duomenų formatas kristalografinių duomenų archyvavimui ir mainams. Yra parašyti kokybiški sintaksiniai šio formato analizatoriai C, Perl ir Python kalbomis. Deja, Java ir .NET aplinkoms šie analizatoriai nėra labai tinkami, nes minėtos platformos sukelia daug keblumų, paleidžiant joms svetimomis kalbomis parašytas bibliotekas. Pvz. Java JNI sąsaja sunkiai perkeliama į mobilias platformas ir dėl to programuotojai labiau vertina „gryna Java“ („pure Java“) parašytus modulius. Darbe bus siūloma perkelti CIF (CIF 1.1 ir CIF 2.0) standarto sintaksinius analizatorius į grynos Java ir C# terpes.
96	Saulius Grigaitis	Blokų grandinių technologijų "įrodymo turtu" protokolai	Blockchain Proof-of-Stake Protocols	Ištirti "įrodymo turtu" (angl. Proof of Stake) protokolus, fokusuojantis į naujausius pasiekimus Ethereum 2.0 "įrodymo turtu" protokole. Pasiūlyti patobulinimus ir juos eksperimentiškai ištirti. (Research Proof-of-Stake protocols focusing on the latest achievements in Ethereum 2.0 Proof-of-Stake protocol. Propose protocol improvements and conduct experiments.)
97	Saulius Grigaitis	Žinių neatskleidžiantys kriptografiniai metodai (PlonK ir kt.)	Zero Knowledge Proofs (PlonK etc.)	Ištirti naujausius žinių neatskleidžiančių kriptografinių metodų pasiekimus (pvz. PlonK)(Research the latest attachments in Zero Knowledge Proofs (PlonK etc.)).
98	Viačeslav Pozdniakov	Funkcinės programavimo kalbos su priklausomais tipais kompiliavimas į Go programavimo kalbą	Golang as a compilation target for functional programming languages with dependent types	
99	Viktor Medvedev	Ne vaizdinių duomenų transformavimo į vaizdus būdai, skirti konvoliucinio neuroninio tinklo architektūrai	Techniques for transforming non-image data into images for use in convolutional neural network architecture	Nagrinėjami metodai, leidžiantys nevaizdinius duomenis transformuoti į vaizdus. Taigi, konvoliucinio neuroninio tinklo (CNN) galimybės gali būti panaudotos ir duomenims, nesusijusiems su vaizdais.
100	Virginijus Marcinkevičius	Garso šaltinio lokalizavimo metodų tyrimas	Investigation of sound source direction detection methods	Tikslas palyginti giliais neuroniniais tinklais grįstų algoritimų garso šaltinio lokalizavimo galimybes, palyginus su klasikiniais metodais ir pasiūlyti algoritimų modifikacijas siekiant pagerinti algoritmų tikslumą.
101	Virginijus Marcinkevičius	Roboto manipulatoriaus valdymo algoritmų palyginimas	Comparison of control algorithms for robot manipulators	Darbe bus naudojamas TurtleBot robotas su manipulatoriumi. Tikslas ištirti kelis metodus skirtus išmokyti robotą paimti įvairius objektus ir juos perdėti į kitą vietą.