



LIETUVOS MOKSLO TARYBOS FINANSUOTO IMPRESS-U PROGRAMOS PROJEKTO

„IRIS-HEP: MOKSLO PROGRAMINĖS ĮRANGOS PLĖTIMAS DALELIŲ FIZIKOS REIKMĖMS UKRAINAI, LENKIJAI IR LIETUVA“

**(Nr. S-IMPRESSU-24-4, reg. data 2024-06-26 d.,
Pakeista Nr. S-IMPRESSU-26-2, reg. data 2026-03-19 d.)**

2026 M. STUDENTŲ STAŽUOČIŲ ATASKAITOS

Projekto vykdytojas: Vilniaus Universitetas

Projekto vadovas: dr. Andrius Juodagalvis

2026 Vilnius

CMS pile-up duomenų kompresija, naudojant giliojo mokymosi VQ-VAE modelį

Stažuotojas P. Balčiūnas, stažuotės konsultantas T. Raila

2026-04-30

Stažuotė „CMS Pile-up duomenų kompresija taikant giliojo mokymosi VQ-VAE modelį (2026-02-02 – 2026-04-18) gavo LMT finansavimą pagal projektą „IRIS-HEP: Mokslo programinės įrangos kūrimo bendradarbiavimo plėtimas į Ukrainą, Lenkiją ir Lietuvą“ (sutarties Nr. S-IMPRESSU-24-4 ir S-IMPRESSU-26-2, vadovas A. Juodagalvis). Europos branduolinių mokslinių tyrimų laboratorijoje CERN atliktos stažuotės metu buvo vykdomi prieš stažuotę numatyti darbai, kurių aprašas pateiktas pridėtame priede.

Pagrindinis stažuotės tikslas buvo pritaikyti giliojo mokymosi vektorių kvantizacijos variacinio autoenkoderio (VQ-VAE) modelį aukštos energijos *pile-up* duomenų kompresijai. Numatytieji stažuotės tikslai buvo įgyvendinti. Svarbiausios stažuotės veiklos buvo šios:

1. Su stažuotės konsultantu sudaryti stažuotės darbų planą. Sulaukti stažuotojų projekto vadovų, P. Elmer ir A. Juodagalvio, pritarimų numatytai veiklai. Užsiregistruoti IRIS-HEP.org puslapyje <https://iris-hep.org/fellows>.
2. Paruošti darbo vietą VQ-VAE modelio duomenų kompresijos tyrimams atlikti.
3. Ištirti CMS detektoriaus *pile-up* duomenų rinkio struktūrą ir savybes.
4. Suprogramuoti CMS duomenų rinkinio filtravimo funkcijas, kurie būtų naudingos tradicinių mašininio mokymosi modelių duomenų kompresijos eksperimentuose.
5. Remiantis moksliniais šaltiniais, atlikti VQ-VAE modelio architektūros ir veikimo analizę.
6. Naudojant Python giliojo mokymosi biblioteką PyTorch ir vektorių kvantizacijos strategijas, parengti VQ-VAE duomenų kompresijos eksperimentus su 2×2 , 4×4 ir 8×8 *pile-up* dalelių susidūrimų pikselių grupėmis.
7. Pritaikyti VQ-VAE modelio hiperparametrų optimizavimo algoritmus.
8. Išanalizuoti VQ-VAE modelio duomenų kompresijos rezultatus, taikant kompresijos kokybės vertinimo metrikas.
9. Pristatyti tarpinius stažuotės projekto rezultatus savaitiniuose susitikimuose konsultantui, aptarti darbo eigą ir gauti grįžtamąjį ryšį.
10. Sudalyvauti CERN organizuojamuose praktiniuose mokymuose, skirtuose tobulinti žinias, susijusias su duomenų analitika, mašininio mokymosi modeliais ir Python programavimu.
11. Atliekamas projektas buvo pristatytas 41-tame „Lithuanians@CERN“ susirinkime, kuris įvyko 2026-03-05. Renginio nuoroda: <https://indico.cern.ch/event/1657849/>

Atliekant stažuotei suformuluotas užduotis, buvo naudota ši programinė įranga: CMSFS, lima, lxplus, root manager, PyTorch, Jupyter Notebooks, VQ-VAE explainer, conda ir uv paketų valdymo įrankis; bei Python programavimo kalba.

Stažuotės metu sukurta programinė įranga yra viešai prieinama adresu <https://github.com/pauliusB20/vq-vae-experiments/tree/main>. Ji bus toliau vystoma bei naudojama kitų tyrėjų visame pasaulyje.

Pagrindinė stažuotės nauda buvo tokia:

1. Patobulinti programavimo, duomenų analitikos ir mašininio mokymosi žinias.
2. Sustiprinti anglų kalbos žinias.
3. Užmegzti naujus profesinius ryšius.
4. Susipažinti su praktiniais aukštos energijos fizikos uždavinių sprendimo metodais kompiuterinėje erdvėje.

5. Susipažinti su CERN Large Hadron Collider greitintuvo veikimu ir jo detektoriais (CMS, LHCb, ATLAS, ALICE).
6. Išgirsti CERN mokslininkų profesinės karjeros istorijas.
7. Sužinoti, koks yra mokslininkų gyvenimas CERN mokslinių tyrimų centre.

Panašias stažuotes rekomenduoju tęsti, nes tai puiki galimybė pagilinti turimas mokslines žinias, pamatyti, kaip realiai dirba mokslininkai, įgyti stipresnių profesinei karjerai reikalingų kompetencijų, gauti grįžtamąjį ryšį apie mokslinius tyrimus ir užmegzti naujų profesinių ryšių.

Stazuotojas

P. Balčiūnas

E-paštas: paulius.balciunas@mif.stud.vu.lt

Charged-Particles Reconstruction at Muon Colliders: 3D Sensor Charge Digitization Model

Stażuotoja Ernesta Idzelytė, stažuotės konsultantas Angira Rastogi
2026-08-15

Stażuotė „Charged-Particles Reconstruction at Muon Colliders: 3D Sensor Charge Digitization Model“ (nuo 2026-07-01– iki 2026-09-23) gavo LMT finansavimą pagal projektą „IRIS-HEP: Mokslo programinės įrangos kūrimo bendradarbiavimo plėtimas į Ukrainą, Lenkiją ir Lietuvą“ (sutarties Nr. S-IMPRESSU-24-4 ir S-IMPRESSU-26-2, vadovas A. Juodagalvis). Europos branduolinių mokslinių tyrimų laboratorijoje CERN atliktos stažuotės metu buvo vykdomi prieš stažuotę numatyti darbai, kurių aprašas pateiktas pridėtame priede.

Pagrindinis stažuotės užduoties tikslas buvo įtraukti 3D silicio jutiklių modelį miuonų greitintuvo programinėje įrangoje. Numatytieji stažuotės tikslai buvo įgyvendinti, pradėti darbai bus tęsiami nuotoliniu būdu. Svarbiausios stažuotės veiklos buvo šios:

1. Su stažuotės konsultantu sudaryti stažuotės darbų planą. Sulaukti stažuotės projekto vadovų, P. Elmer ir A. Juodagalvio, pritarimų numatytiems veiksliams. Užsiregistruoti IRIS-HEP.org puslapyje <https://iris-hep.org/fellows>.
2. Atsisiųsti testuojamos programinės įrangos „mucoll-benchmarks“ (<https://github.com/MuonColliderSoft/mucoll-benchmarks>) pirminį kodą. Jį sukompiliuoti ir paleisti pilną generavimo, simuliacijos, skaitmeninimo ir rekonstrukcijos grandinę.
3. Pagal numatytą planą modifikuoti programinę įrangą, pritaikant ją numatytiems tikslams: modifikuoti „ClusterShapeAnalysis“ paketą, sukurti CMake darbo aplinką, modifikuoti geometrijos aprašus, keičiant vidinio trekerio sluoksnius į 3D jutiklių geometriją, paruošti palyginamuosius grafikus tarp abiejų geometrijų.
4. Duoti patikrinti programinę įrangą mentoriams ir gauti jų atsiliepimus. Koreguoti programinį kodą.
5. Parengti ataskaitą, aprašančią programinės įrangos modifikavimo eigą, patobulinimus ir analizės rezultatus.
6. Atliekamas projektas buvo pristatytas 43-ame „Lithuanians@CERN“ susirinkime, kuris įvyko 2026-08-13. Renginio nuoroda: <https://indico.cern.ch/event/1715834/>.

Atliekant stažuotei suformuluotas užduotis, buvo naudota ši programinė įranga: Key4hep programinė įranga (Gaudi, DD4hep, Geant4, EDM4hep), ROOT, Docker bei programavimo kalbos C++, Python.

Stażuotės metu sukurta programinė įranga patalpinta <https://github.com/ernesta57>.

Pagrindinė stažuotės nauda buvo tokia:

1. Tarptautinė aplinka
2. Bendravimas su kolegomis, mentoriais
3. Ekskursijos laboratorijose
4. Įgytos programavimo žinios

Panašias stažuotes rekomenduoju tęsti, nes gaunama galimybė bendradarbiauti su patyrusiais mokslininkais, įgyti praktinių žinių, pamatyti, kaip realiai įgyvendinamos idėjos programinėje įrangoje. Be to, buvimas Europos branduolinių mokslinių tyrimų laboratorijoje CERN motyvuoja tolimesnei akademiniai veiklai.

Stażuotoja
E-paštas: ernesta.idzelyte@ff.vu.lt

E. Idzelytė

MODIFYING EXISTING TTREES IN UPROOT

Stażuotojas Jokūbas Mačiulis, stažuotės konsultantas Andres Rios-Tascon

2026-08-27

Stażuotė „Modifying existing TTrees in Uproot“ (2026-06-22 – 2026-08-31) gavo LMT finansavimą pagal projektą „IRIS-HEP: Mokslo programinės įrangos kūrimo bendradarbiavimo plėtimas į Ukrainą, Lenkiją ir Lietuvą“ (sutarties Nr. S-IMPRESSU-24-4 ir S-IMPRESSU-26-2, vadovas A. Juodagalvis). Europos branduolinių mokslinių tyrimų laboratorijoje CERN atliktos stažuotės metu buvo vykdomi prieš stažuotę numatyti darbai, kurių aprašas pateiktas pridėtame priede.

Pagrindinis stažuotės užduoties tikslas buvo įgyvendinti galimybę modifikuoti esamus ROOT formato failus Python bibliotekoje uproot — tai yra, pridėti naujus stulpelius ir eilutes į esamus TTrees bei RNTuple failus tiesiogiai, neperrašant viso failo. Numatytieji stažuotės tikslai buvo įgyvendinti. Svarbiausios stažuotės veiklos buvo šios:

1. Su stažuotės konsultantu sudaryti stažuotės darbų planą. Sulaukti stažuotės projekto vadovų, P. Elmer ir A. Juodagalvio, pritarimų numatyti veiklai. Užsiregistruoti IRIS-HEP.org puslapyje <https://iris-hep.org/fellows>.
2. Atsisiųsti atviro kodo Python bibliotekos uproot (<https://github.com/scikit-hep/uproot5>) pirminį kodą. Sukompiliuoti ir įvykdyti pirminius patikrinimus.
3. Pagal numatytą planą sukurti `add_branches()`, `extend()` ir `add_fields()` metodus, leidžiančius modifikuoti esamus TTrees ir RNTuple failus (pridėti eilutes/stulpelius) neperrašant viso failo.
4. Naują programinę įrangą patikrinti naudojant testus ir CERN ROOT programinę įrangą.
5. Atliktus darbus reguliariai aptarti su mentoriumi per savaitinius susitikimus.
6. Pateikti du pull request'us į oficialią uproot biblioteką, kurie šiuo metu peržiūrimi mentoriams.
7. Atliekamas projektas buvo pristatytas 43-tame „Lithuanians@CERN“ susirinkime, kuris įvyko 2026-08-13. Renginio nuoroda: <https://indico.cern.ch/event/1715834/>

Atliekant stažuotei suformuluotas užduotis, buvo naudota ši programinė įranga: GitHub, PyROOT, Linux terminalas, Visual Studio Code; bei programavimo kalba Python.

Stażuotės metu sukurta programinė įranga yra viešai prieinama adresu <https://github.com/Yokubas/uproot5> (asmeninė talpykla) ir <https://github.com/scikit-hep/uproot5> (oficiali talpykla); pateikti pakeitimai: <https://github.com/scikit-hep/uproot5/pull/1690> (TTree) ir <https://github.com/scikit-hep/uproot5/pull/1687> (RNTuple). Ji bus toliau vystoma bei naudojama kitų tyrėjų visame pasaulyje.

Pagrindinė stažuotės nauda buvo tokia:

1. Įgyti praktiniai įgūdžiai kuriant atviro kodo mokslinę programinę įrangą tarptautiniame kontekste
2. Įgytos žinios apie uproot biblioteką bei HEP duomenų formatus (TTrees, RNTuple)
3. Patirtis dirbant su dideliais atviro kodo projektais – kodo peržiūra, testų rašymas, pull request'ų teikimas
4. Tarptautinė aplinka ir bendravimas su mokslininkais iš viso pasaulio
5. Galimybė prisidėti prie programinės įrangos, kuri bus naudojama fizikų visame pasaulyje

Panašias stažuotes rekomenduoju tęsti, nes tai unikali galimybė prisidėti prie realiai naudojamos mokslinės programinės įrangos, įgyti praktinės patirties tarptautinėje aplinkoje ir susipažinti su aukšto lygio moksliniais tyrimais CERN. Stažuotė suteikia galimybę ne tik tobulinti techninius įgūdžius, bet ir užmegzti vertingų profesinių ryšių, pamatyti save tarptautiniame kontekste.

Stazuotojas

Jokūbas Mačiulis

E-paštas: jokubui.maciuliui@gmail.com