



LIETUVOS RESPUBLIKOS ŠVIETIMO IR MOKSLO MINISTRAS

ĮSAKYMAS

**DĖL MOKSLO IR STUDIJŲ INSTITUCIJOS DELEGAVIMO ATSTOVAUTI LIETUVAI
BENDROSIOS KALBOS IŠTEKLIŲ IR TECHNOLOGIJŲ INFRASTRUKTŪROS
EUROPOS MOKSLINIŲ TYRIMŲ INFRASTRUKTŪROS KONSORCIUME IR
LIETUVOS NARYSTĖS BENDROSIOS KALBOS IŠTEKLIŲ IR TECHNOLOGIJŲ
INFRASTRUKTŪROS EUROPOS MOKSLINIŲ TYRIMŲ INFRASTRUKTŪROS
KONSORCIUME PLANO PATVIRTINIMO**

2015 m. birželio 16 d. Nr. V-612
Vilnius

Įgyvendindama Lietuvos Respublikos Vyriausybės 2014 m. rugsėjo 3 d. nutarimo Nr. 875 „Dėl valstybės turto investavimo“ 2.2 papunktį ir Dalyvavimo tarptautinėse mokslinių tyrimų infrastruktūrose tvarkos aprašo, patvirtinto Lietuvos Respublikos švietimo ir mokslo ministro 2012 m. liepos 2 d. įsakymu Nr. V-1068 „Dėl Dalyvavimo tarptautinėse mokslinių tyrimų infrastruktūrose tvarkos aprašo patvirtinimo“, 10 ir 14 punktus:

1. D e l e g u o j u Vytauto Didžiojo universitetą atstovauti Lietuvai tarptautinėje mokslinių tyrimų infrastruktūroje – Bendrosios kalbos išteklių ir technologijų infrastruktūros Europos mokslinių tyrimų infrastruktūros konsorciume.

2. T v i r t i n u Lietuvos narystės Bendrosios kalbos išteklių ir technologijų infrastruktūros Europos mokslinių tyrimų infrastruktūros konsorciume planą (pridedama).

Švietimo ir mokslo ministrė

Audronė Pitrėnienė

LIETUVOS NARYSTĖS BENDROSIOS KALBOS IŠTEKLIŲ IR TECHNOLOGIJŲ INFRASTRUKTŪROS EUROPOS MOKSLINIŲ TYRIMŲ INFRASTRUKTŪROS KONSORCIUME PLANAS

I SKYRIUS BENDROSIOS NUOSTATOS

1. Lietuvos narystės Bendrosios kalbos išteklių ir technologijų infrastruktūros Europos mokslinių tyrimų infrastruktūros konsorciume (toliau – CLARIN ERIC) planas (toliau – Planas) yra skirtas išlaidų padengimo planui sudaryti ir finansavimo šaltiniams nurodyti, narystės trukmei CLARIN ERIC ir pasiektiniams rezultatams įvardyti, taip pat – prieigos prie CLARIN ERIC suteikimui apibrėžti.

2. Nacionalinis konsorciumas – Vytauto Didžiojo universiteto, Kauno technologijos universiteto ir Vilniaus universiteto konsorciumas, įgyvendinantis Planą.

3. Narystės CLARIN ERIC trukmė – Lietuvos Respublikos Vyriausybės 2014 m. rugsėjo 3 d. nutarimo Nr. 875 „Dėl valstybės turto investavimo“ 1 punktu patvirtinta narystė iki 2016 metų gruodžio 31 d. Pasibaigus šiam laikotarpiui narystė, Lietuvos Respublikos Vyriausybės sprendimu, gali būti pratęsta.

4. Naryste CLARIN ERIC siekiama užtikrinti Lietuvos humanitarinių ir socialinių mokslų plėtrą, sukuriant tyrėjams prieigą prie bendros Europinės platformos, apimančios kalbų pagrindu sukurtus skaitmeninius teksto, garso, vaizdo išteklius ir jų analizei būtinus įrankius.

5. Potencialūs CLARIN ERIC naudotojai Lietuvoje – mokslininkai ir kiti tyrėjai ar jų grupės, taip pat – mokslo ir studijų institucijos, vykdančios mokslinius tyrimus atitinkamoje srityje.

II SKYRIUS NARYSTĖS ĮGYVENDINIMAS

6. Su Lietuvos naryste CLARIN ERIC susijusių išlaidų padengimo planas ir finansavimo šaltiniai:

Eil. Nr.	Išlaidų pavadinimas	2015 metai (Eur)	2016 metai (Eur)	2015-2016 metų laikotarpiui (Eur)	Finansavimo šaltiniai	Įgyvendinanti institucija ar jų konsorciumas
1.	Darbo užmokestis	90528	90528	181056	Valstybės biudžeto lėšos	Lietuvos mokslo taryba
		7340	7340	14680	Nacionalinio konsorciumo narių lėšos	Nacionalinis konsorciumas
2.	Socialinio draudimo ir kitos teisės aktais numatytos įmokos	28177	28177	56354	Valstybės biudžeto lėšos	Lietuvos mokslo taryba
		2285	2285	4570	Nacionalinio konsorciumo narių lėšos	Nacionalinis konsorciumas
3.	Išlaidos paslaugoms	55653	0	55653	Valstybės biudžeto lėšos	Lietuvos mokslo taryba
4.	Komandiruočių išlaidos	6170	6170	12340	Valstybės biudžeto lėšos	Lietuvos mokslo taryba

5.	Išlaidos rezultatų viešinimui	2702	2702	5404	Valstybės biudžeto lėšos	Lietuvos mokslo taryba
6.	Kompiuterinė/serverinė technika, infrastruktūra, eksploatacija, priežiūra	11984	11984	23968	Nacionalinio konsorciumo narių lėšos	Nacionalinis konsorciumas
7.	Narystės CLARIN ERIC mokestis	12522	12773	25295	Valstybės biudžeto lėšos	Lietuvos mokslo taryba
8.	Iš viso:	217361	161959	379320		
9.	Iš viso (valstybės biudžeto lėšos):	195752	140350	336102		
10.	Iš viso (nacionalinio konsorciumo narių lėšos):	21609	21609	43218		

7. Per 2015-2016 metų laikotarpį pasiektini rezultatai:

Duomenų centro, teikiančio paslaugas CLARIN ERIC infrastruktūrai, įkūrimas:
 Vytauto Didžiojo universitetas (toliau – VDU) turi įkurti nacionalinį CLARIN ERIC centrą (toliau – CLARIN–LT). VDU rūpinsis CLARIN–LT sukūrimu, patalpomis, elektra, atsarginių kopijų užtikrinimu ir serverių priežiūra, kuriuose 2015-2016 metais bus įkurtas duomenų repozitoriumas ir internetinis portalas, užtikrinantis prieigą prie duomenų, įrankių ir paslaugų. Techninė Centro infrastruktūra turi atitikti CLARIN ERIC B tipo centrų reikalavimus;
 Kauno technologijos universitetas suteiks papildomą kompiuterinę/serverinę techniką CLARIN–LT.

Žinių infrastruktūros, koordinuojamos tarptautiniame lygmenyje, sukūrimas:
 VDU turi įkurti Lietuvių kalbos technologijų konsultacijų centrą (angl. *HelpDesk*). Šis centras suteiks informaciją ir konsultuos įvairiais kalbos technologijų klausimais.
 Kalbos: lietuvių ir anglų.

Kauno technologijos universitetas (toliau – KTU) turi įkurti Semantinių technologijų taikymo konsultacijų centrą. Konsultacijų centras teiks informaciją ir konsultuos šiais klausimais:
 ontologijų kūrimas ir taikymas semantiniame tekstų anotavime ir paieškoje;
 semantinis tekstų anotavimas;
 semantinių technologijų taikymas paieškoje;
 veiklos žodynų ir taisyklių specifikuojamas ribota natūralia kalba taikant semantinio verslo žodyno ir verslo taisyklių (angl. *Semantic Of Business Vocabulary and Business Rules*) standartą;
 veiklos žodynų ir taisyklių taikymas.
 Kalbos: lietuvių ir anglų.

Teisinė veikla (atsakinga institucija: Vilniaus universitetas (toliau – VU):
 rinkti, apibendrinti, paruošti publikavimui bei kaupti vienoje vietoje dokumentus, susijusius su teisiniais CLARIN–LT klausimais;
 tvarkyti licencijų šablonų rinkinį, kuris yra palaikomas CLARIN ERIC infrastruktūroje, ir pritaikyti jį ne tik CLARIN–LT, bet ir viso CLARIN ERIC reikmėms;
 sukurti ir įgyvendinti tinkamas procedūras, kurios galėtų būti pritaikomos vertinant ir apdatuojant naujas teises rekomendacijas;
 glaudžiai konsultuotis su CLARIN ERIC Techninių centrų komitetu tam, kad teisinių klausimų sprendimas ir jų pritaikymas, kuriant CLARIN B tipo centrą, atitiktų CLARIN ERIC reikalavimus; užtikrinti teisinių klausimų harmonizavimą tarp CLARIN ERIC ir CLARIN–LT.

Egzistuojančių išteklių tolimesnis vystymas CLARIN–LT:

Rašytinės kalbos ištekliai:

Internetinių naujienų tekstynas (Bendrasis internetinis tekstynas (toliau – BIT) bus papildytas 200 mln. žodžių:

Už tekstyno automatinio rinkimo priemonės konfigūravimą, palaikymą ir kaupiamo turinio kokybės priežiūrą atsakingas VDU;

Už tekstyno kaupimo aplinką (infrastruktūrą) ir repozitoriumo funkcionavimo užtikrinimą atsakingas KTU.

Dabartinės lietuvių kalbos tekstynas (toliau – DLKT) bus papildytas 5 mln. žodžių, taip pat bus pagerinta prieigos sąsaja prie tekstyno (atsakinga institucija: VDU).

Sakytinės kalbos ištekliai:

10 val. lietuvių kalbos garsynas:

3 val. apimties lietuvių kalbos dviejų dialektų garsynas (atsakinga institucija: KTU);

7 val. apimties bendras lietuvių kalbos garsynas (atsakinga institucija: VDU).

Papildomai sukauptas 5 val. garsynas:

Lietuvių kalbos garsyno plėtra 2 val. apimties 3-iu dialektu (atsakinga institucija: KTU);

Bendro lietuvių kalbos garsyno plėtra 3 val. apimties (atsakinga institucija: VDU).

Įrankiai:

Lietuvių kalbos tekstų kamienų analizės įrankis ir internetinės paslauga (atsakinga institucija: VU).

Paslaugos:

Pagerinta prieiga tyrėjams prie BIT tekstyno:

Tekstyno prieigos funkcionalumo gerinimas (atsakinga institucija: VDU);

Tekstyno prieigos infrastruktūros gerinimas (atsakinga institucija: KTU).

Pagerinta prieiga tyrėjams prie Dabartinės lietuvių kalbos tekstyno (DLKT) tekstyno (atsakinga institucija: VDU);

Pagerinta morfologinės analizės duomenų bazė ir morfologinė internetinė paslauga (atsakinga institucija: VU);

Pagerinta lietuvių kalbos tekstų anotavimo paslauga (atsakinga institucija: VDU);

Pagerinta lietuvių kalbos valdymo frazių atpažinimo paslauga (atsakinga institucija: VDU);

Pagerinta lietuvių kalbos valdymo balsų internetinė paslauga (asmenų vardų ir vietovardžių atpažinimas, atsakinga institucija: KTU).

Leksiniai ištekliai:

Lietuvių kalbos leksinės duomenų bazės (angl. *Wordnet*) papildymas 20 000 lietuvių kalbos sinonimų rinkinių su grafine sąsaja ir atsisiunčiamais pavyzdžiais (atsakinga institucija: VDU);

Lietuvių kalbos morfologinės analizės ir gramatikos duomenų bazės vystymas su fleksinėmis taisyklėmis ir žodynais *Hunspell* formate (atsakinga institucija: VU).

Naujų išteklių, įrankių ir paslaugų sukūrimas CLARIN–LT:

CLARIN–LT oficialios svetainės sukūrimas (atsakinga institucija: VDU);

Lietuvių kalbos Sintaksinių medžių banko sukūrimas (ne mažiau nei 300 sintaksinių medžių) (atsakinga institucija: VDU).

Naujų išteklių, įrankių ir paslaugų sukūrimas CLARIN-LT:

Rašytinės kalbos ištekliai:

Interneto lietuviškų viešųjų socialinių tekstų tekstynas (atsakinga institucija: VDU);

Lietuvių kalbos Sintaksinių medžių banko tolimesnis kūrimas (ne mažiau nei 2000 sintaksinių medžių, atsakinga institucija: VDU).

Paslaugų įgyvendinimas CLARIN ERIC infrastruktūros reikmėms:

Federacinis tapatybių valdymas (atsakinga institucija: VDU KLC),
 Komponentų metaduomenų infrastruktūros (angl. *CMDI*) tvarkymas (atsakinga institucija: VDU KLC),
 Pastovios identifikavimo sistemos (angl. *Persistent Identifier*) įdiegimas (atsakinga institucija: VDU KLC, KTU ISK),
 CLARIN-LT autentifikavimo sistemos (angl. *IdPs*) įdiegimas (atsakinga institucija: VDU KLC),
 Duomenų kokybės sertifikatas (angl. *Seal of Approval*) (atsakinga institucija: VDU KLC).

8. Prieigos prie CLARIN ERIC ir CLARIN–LT suteikimo Lietuvos mokslo ir studijų institucijoms ir kitiems subjektams tvarka:

Prieiga prie CLARIN ERIC ir CLARIN–LT infrastruktūros Lietuvos mokslo ir studijų institucijoms ir kitiems subjektams suteikiama, laikantis CLARIN ERIC Statuto 7 skyriuje išdėstytais atviros prieigos, prisijungimo ir autorių teisių apsaugos politikos nuostatais ir principais. Lietuvos mokslo ir studijų institucijos ir kiti subjektai (pvz., muziejai, bibliotekos) turi atvirą prieigą prie CLARIN ERIC ir CLARIN-LT duomenų, įrankių ir paslaugų. Atvira prieiga suteikiama visiems šių institucijų darbuotojams, mokslininkams ir kitiems tyrėjams, studentams. Minėtoms institucijoms atvira prieiga suteikiama, laikantis CLARIN ERIC nustatytų institucijos autentifikavimo principų. Už tinkamą ir savalaikį šių Lietuvos institucijų autentifikavimo klausimų sutvarkymą atsako Nacionalinis konsorciumas. Pastarasis tai atlieka nuolat apklaudamas minėtas institucijas ir informaciją apie atvirą prieigą skelbdamas konsorciumo svetainėje. Bendrą konsorciumo narių politiką duomenų saugos klausimais ir autorių teisių apsaugos klausimais atspindi konsorciumo narių 2015 m. kovo 31 d. Jungtinės veiklos sutartis Nr. 7-19/BS-230000-809 ir principai, fiksuoti konsorciumo narių teisės aktuose. Nacionalinio konsorciumo narių informacinių sistemų duomenų sauga užtikrinama vadovaujantis Valstybės informacinių išteklių valdymo įstatymu ir Bendraisiais reikalavimais organizacinėms ir techninėms duomenų saugumo priemonėms (toliau – Bendrieji reikalavimai), patvirtintais Valstybinės duomenų apsaugos inspekcijos direktoriaus 2008 m. lapkričio 12 d. įsakymu Nr. 1T-71 (1.12) „Dėl Bendrųjų reikalavimų organizacinėms ir techninėms duomenų saugumo priemonėms patvirtinimo“.

Nacionalinio konsorciumo narių informacinių sistemų valdytojas įgyvendina kontrolės priemones, skirtas apsaugoti informacinės sistemos duomenis, informaciją ir jų kopijas nuo atsitiktinio ar neteisėto sunaikinimo, pakeitimo, sugadinimo, atskleidimo, neteisėto pasisavinimo, paskelbimo, pateikimo ar kitokio panaudojimo, taip pat nuo bet kokio kito neteisėto tvarkymo.

Už informacijos saugą pagal kompetenciją atsako Nacionalinio konsorciumo narių informacinių sistemų valdytojas ir tvarkytojas. Nacionalinio konsorciumo narių informacinių sistemų tvarkytojai privalo saugos nuostatuose ir kituose saugos dokumentuose nustatyta tvarka užtikrinti reikiamas administracines, technines ir organizacines duomenų ir informacijos saugos priemones ir tokių priemonių laikymąsi.

9. Institucija, atstovaujanti Lietuvai (rekvizitai):

Pavadinimas	Vytauto Didžiojo universitetas
Akronimas	VDU
Teisinė forma	Viešoji įstaiga
Adresas	K. Donelaičio g. 58, LT-44029 Kaunas
Interneto svetainė	http://www.vdu.lt
El. paštas	info@vdu.lt
Telefonas	(8 37) 222 739
Kodas	111950396
Asmuo kontaktui: dr. Andrius Utka, Vytauto Didžiojo universiteto docentas, Kompiuterinės lingvistikos centro vadovas, tel.: +370 37 327841, el.p. a.utka@hmf.vdu.lt .	