

TECHNINĖ SPECIFIKACIJA**TURINYS**

1. Esamos būsenos aprašymas.....	8
1.1. Lietuvių kalbos išteklių informacinė sistema (LKIIS)	8
1.2. LKIIS architektūra.....	8
2. Perkamų paslaugų apimtis ir reikalavimai	13
3. Paslaugų suteikimo vieta	13
4. Paslaugų suteikimo laikotarpis	14

1. Esamos būsenos aprašymas

1.1. Lietuvių kalbos išteklių informacinė sistema (LKIIS)

LKIIS buvo sukurta įgyvendinant projektą „IRT sprendimų bei turinio, padedančių išsaugoti lietuvių kalbą viešojoje erdvėje, kūrimas bei galimybių naudotis jais sudarymas“. Įgyvendinto projekto apimtyje sukurtas susistemintų lietuvių kalbos išteklių pagrindas ir parengtas tolimesnei plėtrai. Suskaitmeninta 11 vienakalbių ir dvikalbių žodynų ir 5 kartotekos, sukurtos elektroninės šių išteklių duomenų bazės ir jų valdymo įrankiai, sukurta vieša šių išteklių prieiga visuomenei bei integruotos elektroninės paslaugos (e.mokymai, e. kartotekos, e. žodynai).

Projekto „IRT sprendimų bei turinio, padedančių išsaugoti lietuvių kalbą viešojoje erdvėje, kūrimas bei galimybių naudotis jais sudarymas“ apimtyje sukurta informacinė sistema LKIIS. Vadovaujantis Lietuvių kalbos instituto direktoriaus 2013 m. vasario 13 d. įsakymu Nr. R-46 patvirtintais Lietuvių kalbos išteklių informacinės sistemos nuostatais, LKIIS tikslas – kompiuterizuoti su lietuvių kalbos išteklių, apimančių lietuvių kalbos žodyną ir tautosakos kartotekas, vienakalbius ir daugiakalbius žodynus, lietuvių kalbos šriftą „Palemonas“, išsaugojimu susijusius procesus bei padaryti šiuos lietuvių kalbos išteklius prieinamus elektroninėje erdvėje.

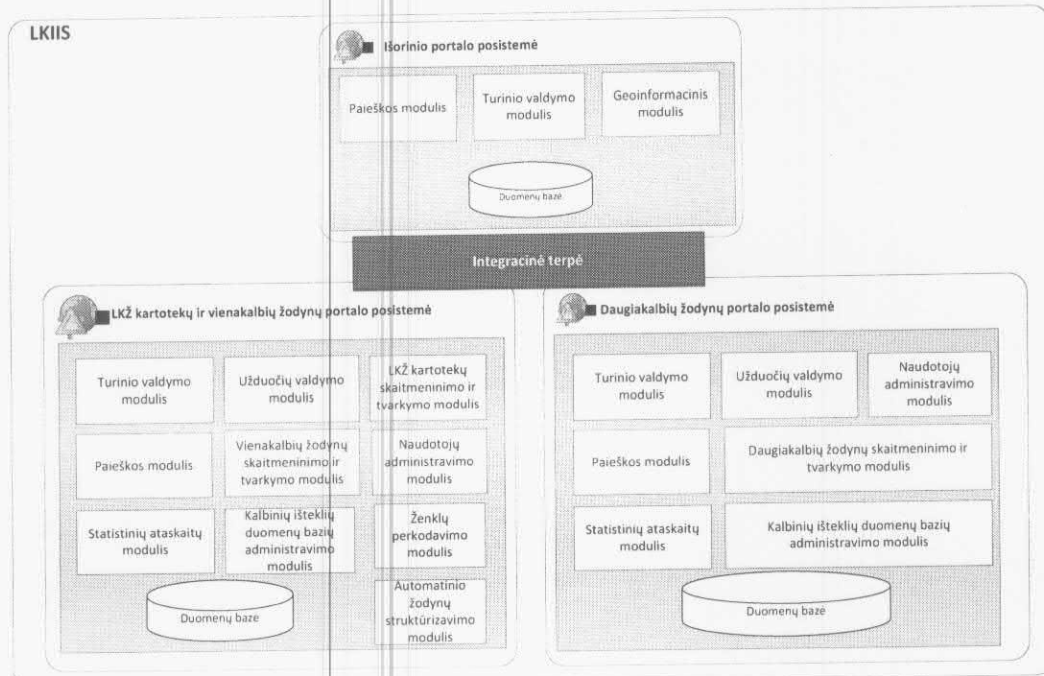
Vadovaujantis Lietuvių kalbos instituto direktoriaus 2013 m. vasario 13 d. įsakymu Nr. R-46 patvirtintais Lietuvių kalbos išteklių informacinės sistemos nuostatais, pagrindinės LKIIS funkcijos yra:

- Lietuvių kalbos didžiojo žodyno kartotekų skaitmeninimas ir publikavimas;
- Vienakalbių ir daugiakalbių žodynų skaitmeninimas ir naudojimosi jais elektroninės paslaugos teikimas;
- Šrifto „Palemonas“ ir edukacinių, lietuvių kalbai skirtų priemonių publikavimas internete;
- Skaitmenintų lietuvių kalbos išteklių susiejimas su metaduomenimis;
- Informacijos paieška;
- Ilgalaikis suskaitmenintų lietuvių kalbos išteklių kaupimas ir saugojimas;
- Informacijos apie suskaitmenintus lietuvių kalbos išteklius teikimas kitoms Lietuvos ir tarptautinėms informacinėms sistemoms.

LKIIS valdytojas yra LKI.

1.2. LKIIS architektūra

LKIIS architektūra realizuota į servišus orientuotos architektūros technologijos (angl. SOA) pagrindu. Tokia architektūra suteikia galimybę sukurti lanksčią sistemą ir orientuotą į paslaugų teikimą su galimybe įdiegti efektyvius paslaugų ir sistemų integravimo sprendimus. Esama LKIIS architektūra pavaizduota paveiksle žemiau.



1 paveikslas. Esama LKIIS architektūra

1 lentelė. LKIIS komponentai

Komponento pavadinimas	Paskirtis	Techninė realizacija
Išorinio portalo posistemė		
Paieškos modulis	Paieškos modulis, užtikrina visose LKIIS duomenų bazėse (lietuvių kalbos didžiojo žodyno kartotekų, vienakalbių ir daugiakalbių žodynų duomenų bazėse) saugomo turinio indeksavimą, paieškos užklausų vykdymą ir rezultatų grąžinimą Paslaugos gavėjams.	Apache Solar paieškos ir indeksavimo programinė įranga.
Turinio valdymo modulis	Portalo turinio valdymo modulis, suteikia galimybę administruoti bei valdyti tiek išorinio, tiek ir vidinių portalų turinį.	Liferay CE 6.1 portalo platforma.
GIS modulis	Geoferencinio modulio pagalba užtikrinama: Kortelės užrašymo geografinės informacijos, nurodančios kurioje vietoje buvo sudaryta kortelė susiejimas su geoinformacinės sistemos duomenimis. Kartotekos įrašų pateikimas interaktyviame žemėlapyje. Kortelės užrašymo geografinės informacijos susiejimas su Lietuvos istorinių vietovardžių duomenų bazės įrašais.	ArcGIS 10
Duomenų bazė	Užtikrinamas LKIIS išteklių saugojimas ir loginis susiejimas tarpusavyje.	Atvirojo kodo reliacinių duomenų bazių valdymo sistema – PostgreSQL.
Integracinė terpė		
Integracinė terpė	Vykdo duomenų mainus tarp įvairiose technologinėse terpėse funkcionuojančių informacinių sistemų komponentų. Integracinėje terpėje yra sukurtos sąsajos su Lietuvių literatūros ir tautosakos institute saugomomis tautosakos kartotekų duomenų bazėmis: Mūsų kartoteka; Pokario partizanų dainų kartoteka; Liaudies tikėjimų kartoteka.	Mule ESB
LKŽ kartotekų ir vienakalbių žodynų portalo posistemė		

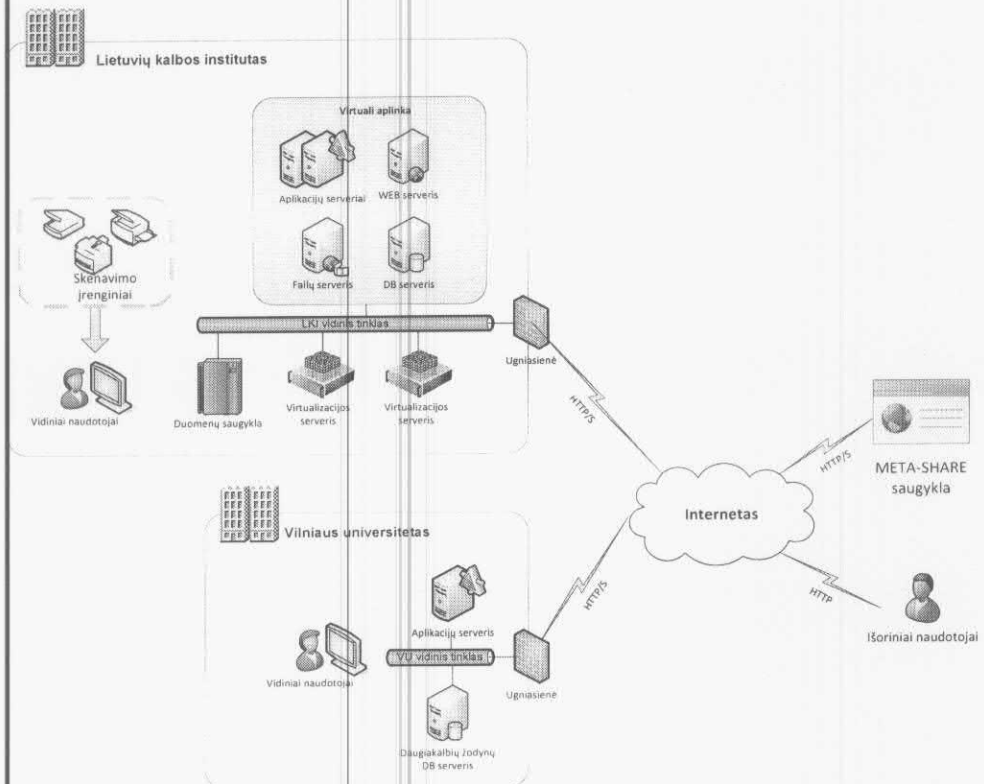
Turinio modulis	valdymo	Modulis, suteikiantis galimybę administruoti bei valdyti portalo turinį.	LKIIS portalų turinio valdymui bus naudojama Liferay platforma.
Paieškos modulis		Modulis, užtikrinantis lietuvių kalbos didžiojo žodyno kartotekų ir vienakalbių žodynų duomenų bazėse saugomų įrašų indeksavimą, paieškos užklausų vykdymą ir rezultatų grąžinimą LKIIS naudotojams.	Paieškos modulis realizuotas panaudojant „Apache Solr“ paieškos indeksavimo variklį.
Statistinių modulis	ataskaitų	Modulis, skirtas įvairių ataskaitų apie suskaitmenintus lietuvių kalbos didžiojo žodyno kartotekų ir vienakalbių žodynų išteklius parengimui.	Ataskaitų modulis realizuotas Birt ataskaitų kūrimo ir generavimo platformos pagrindu (http://eclipse.org/birt/phoenix/intro/).
Naudotojų administravimo modulis		Modulis, skirtas posistemės naudotojų paskyrų sukūrimui ir jų administravimui, naudotojų ir paslaugos gavėjų prieigos teisių posistemėje valdymui.	Naudotojų administravimas, autentifikavimas ir autorizavimas vykdomas Liferay platformos priemonėmis.
Vienakalbių žodynų skaitmeninimo ir tvarkymo modulis		Modulis skirtas vienakalbių žodynų skaitmeninimui, tvarkymui ir išsaugojimui vienakalbių žodynų duomenų bazėje.	Modulis realizuotas kaip Liferay portalo įskiepis (angl. <i>portlet</i>). Optiniam simbolių atpažinimui naudojama ABBYY Fine Reader. Nuskenuoti, atpažinti ir sutvarkyti žodynų tekstai saugomi .doc arba .docx formatu.
Kalbinių duomenų administravimo modulis	išteklių bazių	Modulis, skirtas kurti naujas ir redaguoti/papildyti esamas kalbinių išteklių duomenų bazes	Modulis realizuotas kaip Liferay portalo įskiepis (angl. <i>portlet</i>). Kalbinių išteklių struktūros aprašomos XML schemomis, importavimo ir eksportavimo transformacijomis, įrašo pavadinimo formavimo transformacija ir pan.
Lietuvių didžiojo žodyno (LKŽ) kartotekų skaitmeninimo ir tvarkymo modulis	kalbos	Modulis skirtas lietuvių kalbos didžiojo žodyno kartotekų skaitmeninimui, tvarkymui ir išsaugojimui lietuvių kalbos didžiojo žodyno kartotekų duomenų bazėje.	Modulis realizuotas kaip Liferay portalo įskiepis (angl. <i>portlet</i>). Iš tinklinio katalogo bylos yra automatiškai importuojamos į sistemą ir išsaugomos kartotekos kortelių skenuotų vaizdų sąrašė. Kortelių vaizdai saugomi FreeNAS failų serveryje.
Ženklių perkodavimo modulis		Privataus diapazono ženklų kodų perkodavimas į komponentų sekas.	
Automatinio žodynų struktūravimo modulis		Automatinis žodyno įrašų išgavimas iš .docx formato bylų ir struktūrizavimas, pagal LKIIS apibrėžtą žodyno duomenų modelį. Automatinio struktūrizavimo priemonės parengiamos šiems žodynams: sinonimų, antonimų, palyginimų, frazeologijos, dabartinės lietuvių kalbos žodynas ir sisteminis žodynas. Sustruktūruoti įrašai importuojami į LKIIS.	
Duomenų bazė		Realizuotos vienakalbių išteklių duomenų struktūros.	Atvirojo kodo reliacinių duomenų bazių valdymo sistema – PostgreSQL.
Daugiakalbių žodynų portalo posistemė			
Turinio modulis	valdymo	Modulis, suteikiantis galimybę administruoti bei valdyti daugiakalbių žodynų portalo turinį.	LKIIS portalų turinio valdymui bus naudojama Liferay platforma. Liferay portalo puslapiai konstruojami iš įskiepių (angl. <i>portlets</i>), kuriuos galima traktuoti kaip atskiras aplikacijas, pateikiančias tam tikrą informaciją ar funkcionalumą. Iš esmės naudotojui pateikiamas portalo puslapis yra sudarytas įskiepių rinkinio, kurie sudaro bendrą HTML puslapio vaizdą. Portale diegiami įskiepai turi būti suderinami su JSR-

			286 specifikacija.
Paieškos modulis		Modulis, užtikrinantis daugiakalbių žodynų duomenų bazėje saugomų įrašų indeksavimą, paieškos užklausų vykdymą ir rezultatų grąžinimą LKIIS naudotojams.	Paieškos modulis bus realizuojamas panaudojant „Apache Solr“ paieškos indeksavimo variklį.
Statistinių ataskaitų modulis	ataskaitų	Statistinių ataskaitų modulis, suteikia priemones lengvai ir greitai parengti įvairias ataskaitas apie suskaitmenintus ir LKIIS saugomus lietuvių kalbos išteklius.	Ataskaitų modulis realizuotas Birt ataskaitų kūrimo ir generavimo platformos pagrindu. Birt Ataskaitų formavimo komponentas integruotas LKIIS portaluose, panaudojant LifeRay portalo įskiepi Visioneo.
Užduočių valdymo modulis	valdymo	Užduočių valdymo modulis, skirtas atlikti veiklos procesų modeliavimą, kuris leidžia lengviau pritaikyti LKIIS, atsiradus naujiems poreikiams.	Užduočių valdymo modulis realizuotas naudojant Kaleo Workflow paketą. Užduočių sekų aprašymui naudojami XML dokumentai. Kaleo užduočių valdymo modulis diegiamas kaip Liferay portalo įskiepis.
Naudotojų administravimo modulis		Modulis skirtas naudotojų paskyrų sukūrimui ir jų administravimui, naudotojų ir Paslaugos gavėjų prieigos teisių posistemėje valdymui.	Naudotojų administravimas vykdomas Liferay portalo priemonėmis.
Daugiakalbių žodynų skaitmeninimo ir duomenų tvarkymo modulis	žodynų ir duomenų tvarkymo	Modulis skirtas daugiakalbių žodynų skaitmeninimui, tvarkymui ir išsaugojimui daugiakalbių žodynų duomenų bazėje.	Modulis realizuotas kaip Liferay portalo įskiepis (angl. <i>portlet</i>).
Kalbinių duomenų administravimo modulis	išteklių bazių	Modulis, skirtas kurti naujas ir redaguoti/papildyti esamas kalbinių išteklių duomenų bazes	Modulis realizuotas kaip Liferay portalo įskiepis (angl. <i>portlet</i>). Kalbinių išteklių struktūros aprašomos XML schemomis, importavimo ir eksportavimo transformacijomis, įrašo pavadinimo formavimo transformacija ir pan.
Duomenų bazė		Realizuotos daugiakalbių išteklių duomenų struktūros	Atvirojo kodo reliacinių duomenų bazių valdymo sistema – PostgreSQL.

LKIIS yra geografiškai paskirstyta sistema:

- Lietuvių kalbos didžiojo žodyno kartotekos bei vienakalbiai žodynai skaitmeninami ir saugomi Lietuvių kalbos institute.
- Daugiakalbiai žodynai saugomi Vilniaus universitete.

Paveiksle žemiau pateikiama apibendrinta LKIIS fizinės architektūros schema.



2 paveikslas. LKIIS fizinės architektūros schema

2. Perkamų paslaugų apimtis ir reikalavimai

Šiuo metu Lietuvių kalbos institutas (toliau – LKI, Perkančioji organizacija) įgyvendina projektą „Lietuvių kalbos išteklių informacinės sistemos plėtra (E.Kalba)“, finansuojamą 2014-2020 metų Europos Sąjungos fondų investicijų veiksmų programos 2 prioriteto „Informacinės visuomenės skatinimas“ įgyvendinimo priemonės Nr. 02.3.1-CPVA-V-527 „Lietuvių kalba informaciniuose technologijose“ lėšomis (toliau – Projektas).

Projekto tikslas – didinti lietuvių kalbos išteklių prieinamumą.

Pagrindinis projekto uždavinys – modernizuoti *Lietuvių kalbos išteklių informacinę sistemą* (toliau – LKIIS), integruojant naujus reprezentatyvius lietuvių kalbos išteklius, sukuriant lietuvių kalbos žodžių prasminį tinklą ir naujas e.paslaugas vartotojams.

Šiuo pirkimu siekiama atlikti parengiamuosius darbus prieš vykdant pagrindines LKIIS modernizavimo veiklas.

Paslaugos turi būti teikiamos vadovaujantis galiojančiais Lietuvos Respublikos teisės aktais, reglamentuojančiais informacinių technologijų valdymą.

Toliau lentelėje detalizuojami pagrindiniai reikalavimai parengiamiesiems darbams.

Reikalavimo Nr.	Reikalavimas
1.	Turi būti atliktas duomenų sinchronizavimo tarp vidinio portalo ir išorinio portalo greitaveikos optimizavimas. Vidiniame portale pakeisti duomenys į išorinį portalą turi būti publikuojami ne vėliau kaip po 1 val.
2.	Turi būti atliktas failų saugyklos greitaveikos optimizavimas ir archyvinių duomenų ištrynimas, siekiant užtikrinti stabilų saugyklos veikimą ir galimybę talpinti naujai integruojamų kalbinių išteklių duomenis.
3.	Turi būti atliktas išorinio portalo paieškos indekso optimizavimas. Išoriniame portale paieškos užklausa turi būti aptarnaujama ne ilgiau kaip per 3 sekundes.
4.	Turi būti atliktas dabartinės lietuvių kalbos žodyno duomenų struktūros optimizavimas: a. Kirčiuotų variantų antraštiniuose žodžiuose atskyrimas; b. Gale einančių iliustracijų (kai baigiasi sutrumpinimu) atskyrimas; c. Pažymų perkėlimas po kirčiuotės; d. Žodžių formų atskyrimas.
5.	Turi būti atliktas šių kalbos išteklių išeksportavimas į XML formatą: 1. Vienakalbiai žodynai: a. Dabartinės lietuvių kalbos žodynas b. Sinonimų žodynas c. Antonimų žodynas d. Frazeologijos žodynas 2. Dvikalbiai žodynai: a. Anglų – lietuvių kalbų žodynas
6.	Turi būti atlikta visų LKIIS veikimui naudojamų fizinių ir virtualių serverių inventorizacija. Mažiausiai turi būti identifikuojamas serverio pavadinimas, paskirtis, tinklo informacija, techniniai resursai, esamas apkrovimas.
7.	Vykdydamas parengiamuosius darbus Tiekėjas neturi trikdyti LKIIS veikimo, išskyrus trumpalaikius stabdymus su Perkančiąja organizacija suderintu laiku.

3. Paslaugų suteikimo vieta

Paslaugų teikimo vieta – Lietuvių kalbos institutas, P. Vileišio g. 5, 10308 Vilnius.

4. Paslaugų suteikimo laikotarpis

Paslaugų suteikimo laikotarpis – 60 dienų nuo sutarties pasirašymo dienos.
